Operational Control of Internal Transport

ERIM Ph.D. series Research in Management 1
Also appeared as: TRAIL Thesis Series nr. T2000/5, The Netherlands TRAIL Research School

ISBN 90-5892-004-6

Operational Control
of Internal Transport

Besturingsystemen voor intern transport

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr. ir. J.H. van Bemmel
en volgens het besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 28 september 2000 om 16.00 uur

door

Jaap Robert van der Meer

geboren te Harderwijk

PROMOTIECOMMISSIE

Promotoren:	Prof.dr. M.B.M. de Koster
	Prof.dr.ir. R. Dekker
Overige leden:	Prof.dr. S.L. van de Velde
	Prof.dr.ir. J.J.M. Evers
	Prof.dr. A. van Harten

Preface

The trouble with writing a dissertation is writing the dissertation. The trouble in the years of education, research and teaching preceding writing the dissertation seems to diminish when you are ready for this final step. At least in my situation.

This dissertation considers the control of guided vehicles in vehicle-based internal transport systems found in facilities such as warehouses, production plants, distribution centers and transshipment terminals. It is a combination of theoretical and practical research, keeping practical applications in mind. It has been the result of four years of research and hard work, with contributions of different people whom I would like to thank. I regret that I can only mention a few by name here.

It all started at the department of Econometrics at the Erasmus University Rotterdam where Rommert Dekker supervised me with my Master's thesis. While writing my thesis, he asked me if I would like to work on some consultancy type projects. During these projects, I met many Ph.D. candidates who did similar work, or so it seemed. Partly based on the positive experiences I had working on the projects and socializing with the Ph.D. candidates, I decided that a Ph.D. research project would be something for me too. Encouraged by Raymond Plasmeijer, Arjan Berkelaar and Marcel Kleijn (just to name a few), I started my research project supervised by René de Koster as first promoter and Rommert Dekker as second promoter.

René has actually been both my first promoter and daily supervisor. He has endured me for the last four years and is still as patient and understanding as before. His knowledge and practical experiences were of great value for the research and myself. He has helped me to write this dissertation in such a way that others can also understand what has been done, and he gave me plenty of freedom to carry out my own activities.

I thank Kees Jan Roodbergen for his good company during the last four years and especially during the first year when we shared an office. With a similar academic background and field of research we could talk freely, easily understand each other and have irrelevant (but amusing) discussions which left others clueless.

Discussions that often left me clueless took place in the coffee-corner during lunchtime with other department members. I would like to thank them since those discussions were surprisingly helpful to free my mind.
I thank Murthy Halemane for bringing lunchtime to my attention whenever I was too preoccupied in my work. Many times he would call me with “it’s time for lunch, bring your bucket with homemade macaroni”.
Part of my mental relaxation was to exert myself physically in the gym. There I met many friends who have enriched my life in many different ways. I thank Marcel Teeuw for lifting some of the heavy weights off my shoulder.
Finally, I would like to acknowledge all my (former) colleagues within the group of Logistic Management and the department of Econometrics for their support and for providing a pleasant working atmosphere.

Robert van der Meer
Rotterdam, August 2000

Table of contents

Chapter 1

Introduction

The Netherlands has an excellent geographical location in terms of accessibility by air, waterways, and roads stretching far across Europe. The geographical advantage is the basis for high quality logistic businesses and competitive Dutch enterprises, measuring thousands of national and international warehouses, distribution centers, production plants, transshipment terminals etc. Common logistic activities for all these facilities are storage, transshipment and physical distribution of materials, products, etc. The associated performance objectives generally are to fulfill customer needs at low costs with a high service rate, a high degree of flexibility and short delivery times. The availability, reliability and quality of products (and services) therefore have become the key to competitive strength of facilities. These are the accomplishments of the firm's primary process, which consists of systems and activities that add value to the products and services provided to customers. Flows of products and materials are central to the primary process of many firms and as a consequence, efficient and effective material flow management has drawn much attention. The complex and work intensive issues of flow management are used in combination with material handling systems. Material handling systems are complicated and integrated combinations of material, machines and people. Since labor rates were relatively low in the early stages of the industrial development, manpower was used freely. Efficiency in space utilization, material handling and (vehicle-) control systems was given little consideration. However, rapid development of technology in handling equipment and increasing cost of labor and material, compelled management to take appropriate decisions concerning design and operations of material handling environments.

In the design stage of facilities like manufacturing systems, warehouses, production plants, transshipment terminals, a number of questions have to be answered concerning the facility systems design, layout design and the handling system design.
The facility systems for manufacturing type facilities include the structure of the facility and technical systems for lighting, cooling, heating, compressed air, ventilation, water, etc. The layout consists of all equipment, machinery, production and personnel areas within the building. The handling system consists of the materials, personnel, information and equipment required to satisfy the facility interactions. There exists a strong relationship between these design functions. It is difficult to consider one without considering the other.

The following sections place and describe the setting, motivation and objectives of this dissertation in more detail. The analysis of this dissertation is focussed on control systems for vehicles in vehicle-based internal transportation environments. In order to obtain a better idea of this concept, we will first give some examples of typical vehicle-based internal transportation environments and briefly discuss a few associated aspects.
Next, the subjects associated with facility design are discussed in more detail. We focus on material handling systems and the types of material transport equipment. Since material handling and plant layout are tied together, we will also discuss some issues concerned with layout design. Furthermore, we describe what types of systems are found in vehicle-based transportation environments and how the vehicles can be controlled. We conclude with the motivation and outline of the dissertation.

1.1 Examples of vehicle-based internal transport

Facilities such as warehouses, distribution centers, production plants and transshipment terminals all have vehicle-based material transport systems in common. These systems take care of the internal transport of materials. So the vehicle transport systems are the link for materials between different locations of the transport environment. The success of a material handling system is therefore to a large degree dependent on the efficiency of the vehicle system. Thus the ability to design, implement and control these systems is important. This observation is the main motivation for the analysis in this dissertation.

The next example is used to demonstrate and explain the process of internal transport in more detail. This makes the concept of internal transport more recognizable and easier to place in a different context.

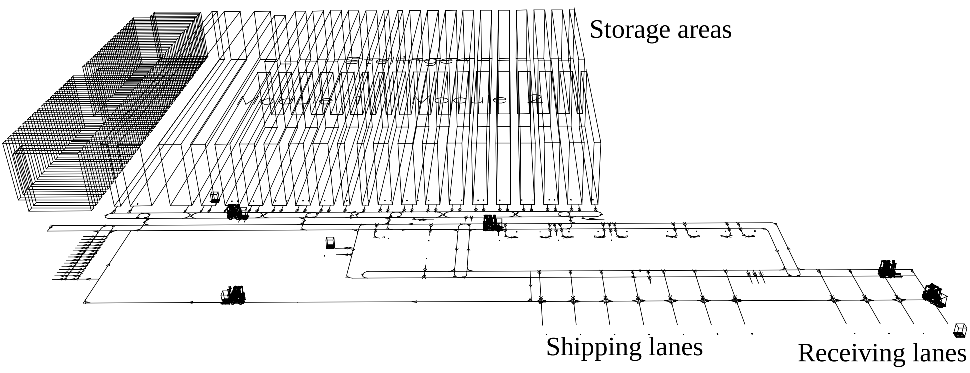

Figure 1. Example of a warehouse. The structure at the top represents the storage area. The network at the bottom represents the vehicle path layout.

Figure 1 shows a front view of an example warehouse. The structure at the top represents the storage areas where (pallet) loads are temporarily stored. The wire network at the lower part of the figure represents the (guide) paths on which vehicles travel. Order requests for (stored) loads are received from customers or clients such as other facilities, retail or department stores. These orders are received on-line by telephone, fax, email or otherwise during the day, and will be considered for shipment the same day if they are received before a certain cut-off time, otherwise the requests are shipped the next day. The requests are clustered partly manually and partly using a decision support system in orders with the same due times, the same (country of) destination, priority, the same carrier trucks, etc. The warehouse management system (see Section 1.4 for more details) then allocates the orders to locations within the storage areas, and order picking routes are determined. Order pickers have to travel (this can be done using automated systems, manually on foot or using orderpicking trucks) through the aisles between the racks in the storage area to visit the locations to collect the orders (outbound loads). These routes are determined with objective criteria such as: minimize the work for the picker or collect materials from the same production run. The sequence of the orders picked depends on the number of order lines, the combination of collecting and returning loads, whether the order involves full pallet picks or combinations with sub pallet picks, etc. Since the pick times and travel times of pickers are stochastic (due to the size or shape of the loads, full or sub-pallet picks, acceleration/deceleration effects, aisle change times, additional requests received during picking, failure of equipment, absence of stock, etc.), the drop off instants of the loads from the storage area by the pickers at the outbound locations are also stochastic. These outbound locations are the pick up locations where vehicles traveling on the guide paths retrieve the loads and transport them to their final destination.
The storage areas are replenished with inbound loads, which are delivered by trucks during the day. Although the approximate arrival times and contents of the trucks are known, the exact arrival time and contents are not known, which makes scheduling the transportation of the loads beforehand impossible. Furthermore, determining combinations of delivering loads and retrieving loads at a particular location beforehand will also be impossible since the exact arrival times of vehicles or release times of loads is not known beforehand.
The loads are moved between other stations in the network as well. For example, to stations for checking-in the load for validation of the contents, for applying a packing list, for labeling the pallet with an identification tag, for repalletization, for problem pallets, etc. This vehicle movement process is called internal transport. The efficiency of the warehouse activities is also dependent on the efficiency of the internal transportation activities.

In automated systems, the load transfer locations can be programmed into a vehicle-control system in advance. Such a control system can be a central controller or computer, which assigns transport tasks (loads) to vehicles. The process of selecting and assigning transport tasks to vehicles is also called dispatching. The vehicles are in general dispatched on-line, i.e. based on real-time information, since the uncertainty of the load release and delivery times makes vehicle dispatching beforehand (scheduling) impossible. Monitoring vehicle positions and traffic control can also be performed by the central computer, or through local controllers of which each controls a section of the path or network.

Internal transport can be found in many other types of facilities as well. The transportation of materials in flexible manufacturing systems (FMS), the transport of luggage at airplane terminals and containers at container transshipment terminals can also be considered as internal transport. Although the transport activities of the last two examples take place outdoors, the vehicles stay in a well-defined and limited operating area.
In this dissertation, internal transport includes all vehicle-based transport systems that share a closed network, limited to and within physical boundaries, like the building of the warehouse or the perimeter (fence) of the transshipment terminal etc.

The focus of this dissertation is the operational control of vehicle-based internal transport systems. After the design of the facility and the choice of the storage and handling equipment, the operating strategies determine the efficiency or performance of the facility. The operation of the storage and retrieval process, and the operational control of internal transport influence the efficiency of material handling operations as a whole.
The next section will first describe in detail the activities, functions and equipment associated with material handling.

1.2 Material handling

There are many activities that occur as part of the transshipment and storage process in facilities like warehouses, distribution centers, production plants and transshipment terminals. The following list includes the general activities, or tasks, found in most of these facilities:

- Receiving materials
- Transportation of materials from receiving to storage areas
- Storage of materials
- Collecting materials (order-picking)
- Internal transport between different areas within the facility
- Value addition to materials or product customization (Value added logistics)
- Shipment of materials

Receiving the materials consists of unloading, identifying and checking the quantity and quality of the material and preparing the material for storage. Preparing the material for storage is optional. This can include repacking the material if it is received in bulk, or restacking the material on a special pallet, such as a slave pallet or in-house pallet.
The transportation of materials consists of moving the materials on product carriers (internal transport) to areas where the product carriers with materials can be stored.

Storage is the physical containment of the materials within storage areas until it is needed again. The type, size, shape and volume of the products or materials to be stored determine the characteristics of the storage area. The position of the stored material for example, often depends on the demand-characteristics of the product. It is preferable to place products with a high demand rate close to entrances and exits of the storage area. Since the storage and retrieval frequency of these products is relatively high, placing these products at the front end of the storage area can save a lot of transportation time.
When needed, the material is collected from the storage area. This process, (also referred to as order picking), can be triggered when an order for certain products is placed by a client or by another process-area. It is one of the basic services within a facility and also determines to a high degree the design of the storage area of the facility.
An optional step after collecting the materials from the storage area is value-added logistics (VAL). VAL is the customization of the product, which may include activities such as, pricing, packaging and labeling the product.
The shipping stage may include a number of tasks. The products may be checked for completeness, sorted for a certain destination or client, packaged, and loaded for shipment.

The handling of materials from the first until the last of the activities mentioned above is referred to as *material handling*. However, there is no unique definition of material handling. For example, Kulwiec (1985) defines material handling as:

"*Materials handling is a system or combination of methods, facilities, labor, and equipment for moving, packaging and storing of materials to meet specific objectives*"

In more recent years, the concept of material handling has become broader. Tompkins et al. (1996) add concepts such as volume, time, space and costs and use the following definition:

"*Material handling means providing the right amount of the right material, in the right condition, at the right place, at the right time, in the right position, in the right sequence, and for the right cost, by using the right method(s)*"

The definition of Tompkins et al. (1996) above is very broad, and is similar to the definitions of logistics and marketing. Specific for material handling are the fundamentals material (what has to be moved and in what quantity), the move (something or someone that executes the move) and the method of handling materials. An organized and systematic method for the analysis of issues concerned with material handling is Systematic Handling Analysis (SHA), see Muther (1973). To perform SHA, certain information is needed of certain factors.
First, information is needed concerning the product (or material) to be moved, the quantity and the product carrier. A directly concerned issue is the unit-load principle. A unit-load can be defined as the unit to be moved or handled at one time (including the product carrier). In some cases the unit-load is one product item (a bottle or screw); in other situations the unit-load is one pallet with several cartons. Bulk-load is another category of

material. Bulk material handling is characterized by continuous-flow operations, involving materials with similar characteristics to those of fluids. Examples include oil, gas, coal, salt, dry powders, and grain.
Next is the information needed on the route and the sequence the product is to be physically handled, and the services necessary to support the activities involved.
Lastly, information is needed concerning the time horizon involved in which (capital) investments are made, movements occur, and performance is measured.
These factors also influence the material handling costs. For example, the cost per unit transported reduces as the quantity transported increases, and increases as the length of the route increases. It has been estimated that between 20 and 50% of the total operating expenses within manufacturing can be attributed to material handling (Tompkins et al., 1996). Furthermore, material handling is estimated to represent between 15 to 70% of the total cost of a product. It is therefore the first place to be considered for cost reduction. Simply handling less is not the answer, since material handling is a very valuable activity. But reduced inventories, improved safety (less damage) and improved (smarter) material control can reduce the costs.

The technology associated with material handling has changed dramatically during the last two decades, mostly due to the introduction of computers and automation. Today, the significance and role of material handling are understood better. Productivity and flexibility are the primary goals of today’s automation technology that can be achieved in highly integrated material handling environments. Therefore, a carefully designed and efficiently managed material handling system is crucial for achieving this integration.
Modern systems can be classified as mechanized (also called conventional) and automated. In general, labor constitutes a high percentage of the overall costs in mechanized systems. In contrast, automated systems attempt to minimize the labor element as much as possible by making capital investments in equipment. In addition, automated material handling (AMH) is expected to operate faster, more accurately and more reliably (in terms of safety and less damaged material) than a mechanized system. However, AMH poses more serious and challenging operational control problems, which in turn, increase by the level of automation.

Figure 2 shows a classification of material handling equipment and the various equipment types used; see also De Koster (1995a). It should be noted that this overview focuses on material transport equipment and is therefore not exhaustive. Furthermore, the growth in technology results in new types of equipment being rapidly developed. Other types of equipment within material handling systems include storage and retrieval equipment and automatic identification and communication equipment.
Storage and retrieval equipment is used within the storage area to store and retrieve loads. Automatic identification and communication equipment includes radio frequency data terminals, bar code readers and communication devices used with all other equipment to coordinate and automate information handling requirements (see Section 1.4).

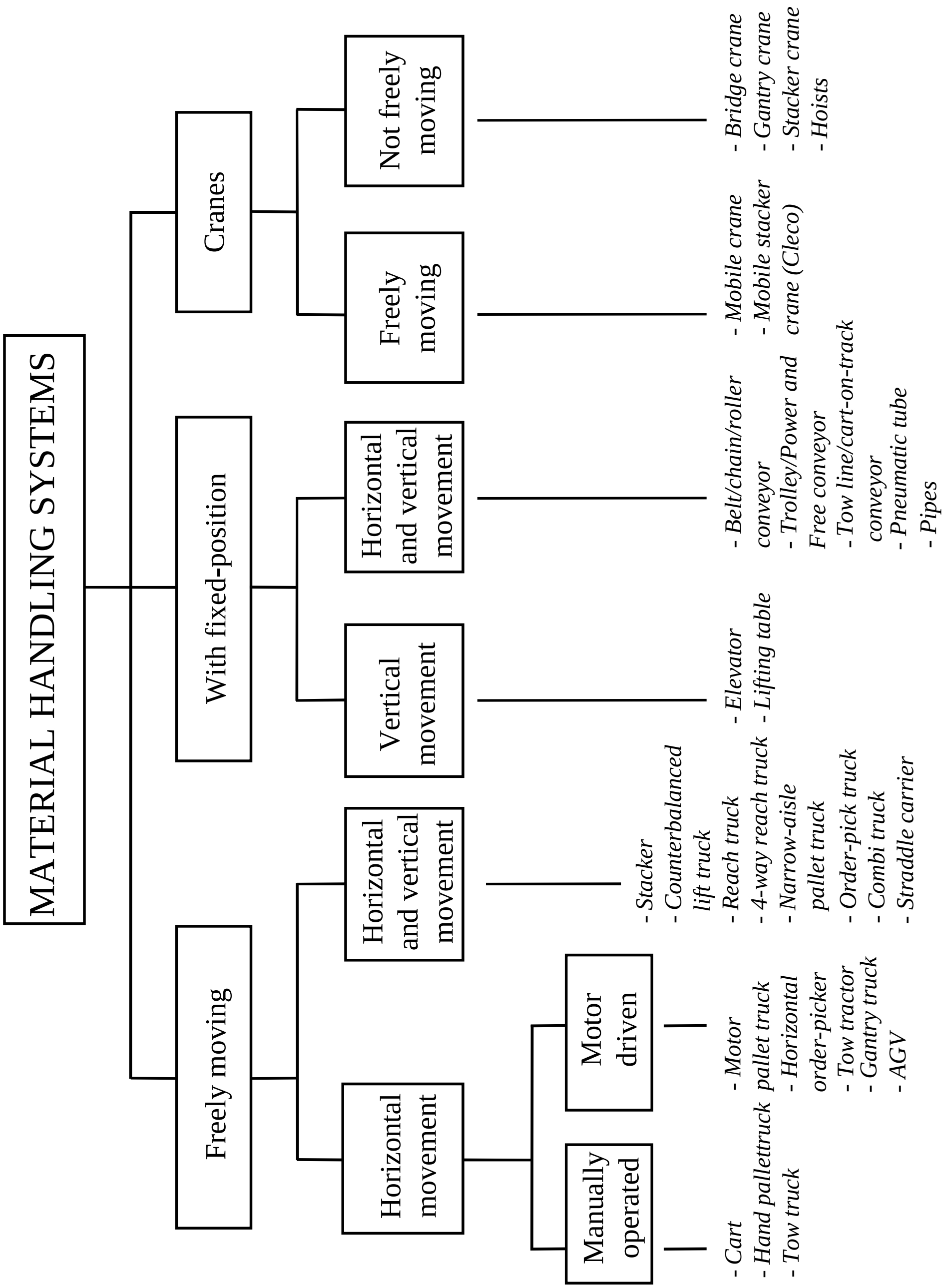

Figure 2. Overview of material handling systems

The next section focuses on the different types of material transport equipment commonly used. We briefly describe the function of some basic equipment types and explain when to use a certain type of equipment.

1.2.1 Material transport equipment

The selection of the proper material handling system is one of the requirements a facility designer has to make. The system planner should identify the alternative equipment for the task and perform a feasibility and economic analysis to select the best type of equipment.
The most commonly used internal transport equipment types are cranes, conveyors and industrial-vehicles, see Figure 2.
Cranes are generally used for horizontal and vertical movement of material from one point to another in the same general area. They can be used for a variety of product-types with respect to weight, volume and material. Typical cranes used for internal transport are bridge cranes, gantry cranes and stacker cranes.
Conveyors are a continuous means of transport, often used when large volumes of conveyable goods have to be transported over rather short distances. Conveyors generally have a fixed position and are therefore static with little flexibility.

Compared to conveyors, industrial-vehicles are discrete (i.e. not continuous), have higher flexibility in routing and in the material that can be transported. Industrial-vehicles are generally used in environments with low intensity material flow with relatively long transport distances. Two categories of industrial-vehicles are defined: manually operated and motor driven.
The manually operated industrial-vehicles are used for short distances and include handcarts, hand trucks and pallet jacks. These are popular vehicles due to their simplicity and low price.
Motor driven vehicles allow the operator to ride the vehicle to, from and between locations. They are used for transporting materials over longer distances and typically offer additional weight and storage height capacity compared to manually operated vehicles. The following sections discuss some motor driven vehicles listed in Figure 2 in more detail.
The motor pallet truck can be seen as a motorized version of the pallet jack (hand pallet truck), used when the travel distances are too long for walking. The motorized version of the hand truck is the tow tractor or tractor trailer. It can pull a train of connected trailers to transport several loads simultaneously.
The “workhorse” of materials handling is the counterbalanced lift truck. These trucks can be used to perform loading, unloading, transportation and stacking tasks and can be adopted to handle a wide range of product-types. The vehicles can be diesel, gas or battery powered. Figure 3 shows an example of two typical counterbalanced lifttrucks. In this case forklift trucks (FLTs). Such vehicles, can achieve lift weight capacities of up to 5000 kg, and transport the loads on the forks in front of the vehicle. The load handling capacity is

determined by the weight of the counter balance over the rear wheels. Since these trucks usually turn within a storage aisle to retrieve a pallet load, the aisle width required to operate is wider than required for some other lift truck alternatives, (see De Koster, 1995b).

Figure 3. Example of typical counterbalanced lifttrucks ***Photo: courtesy of Still BV***

The relatively low equipment costs and flexibility are the main advantage of lift trucks. Figure 4 shows the numbers of (gas and battery powered) forklift trucks and other transport equipment (including: motor pallet trucks, stackers and reach trucks) sold over the last years in the Netherlands, (see Stad, 2000). It shows a general increase in investments made for material handling equipment used for internal transport.

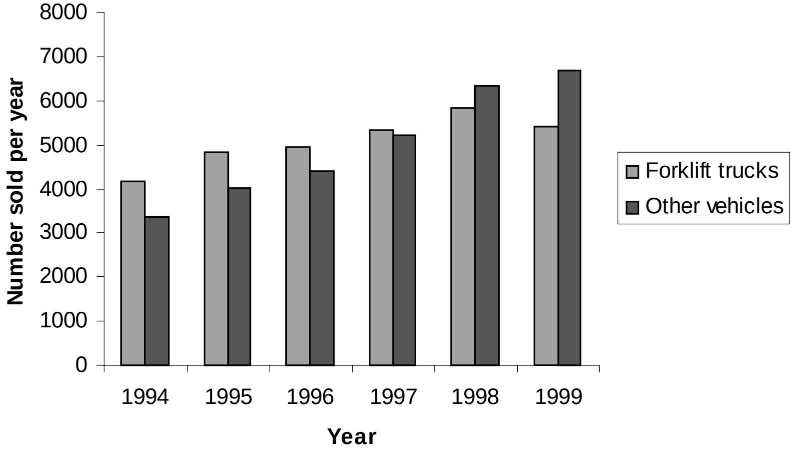

Figure 4. Sales-numbers of industrial vehicles

The cost of a vehicle varies according to the type and the expected performance of the truck. Other factors that influence the cost and choice for a certain vehicle include:

- The handling tasks to be performed by the vehicle
- The weight, shape and size of the material to be transported
- The product carrier on/in which the material has to be transported
- Available maneuvering space (width of the aisles, height of the building)

A non-counterbalanced vehicle is the straddle carrier. This vehicle carries the load within the wheelbase of the truck underneath the driver. The vehicle straddles the load, picks it up and transports it to the desired location. Because the load must be contained within the straddles, the vehicle is not suited for extremely wide loads since the width of the working aisles must also increase. Straddle carriers are primarily used outdoors for very heavy, long and bulky loads and can even be used to handle sea containers. They are therefore also used at container transshipment terminals where they can pick up containers high enough from the ground such that the loaded vehicle can travel ‘over’ one or several stacked containers if necessary. This will be discussed in more detail in Chapter 5.

Automated guided vehicles (AGVs) are distinguished from other industrial-vehicles by the elimination of human intervention from guiding the movement of the vehicle. An AGV system is a set of cooperating driverless (unmanned) vehicles, which transport goods and materials navigated on vehicle guide paths via control mechanisms. Such paths can be wires embedded in the floor with an alternating current to induce a magnetic field, which is detected by antennae attached at the bottom of the vehicles. Guide paths can also be based on a reflector-based system with markings on the floor by reflecting tape, paint or chemicals. Other systems include: laser sensors, sonar, inertial guidance systems where the position is determined based on odometry and systems using a host controller or computer based on beacon systems with vehicle-mounted receivers and stationary beam transmitters.

The latter systems are based on more recent technology that allows the vehicles to operate without physical guide paths; i.e. free-ranging automated guided vehicles.
The vehicles are capable of responding to changing transport patterns and can be integrated into fully automated control systems. AGV systems (AGVSs) can be found in many types of job shop environments and automated warehouses (where the AGVS is made to combine horizontal and vertical load movements). In more complex systems where there are a large number of vehicles transporting loads, a high level of control is necessary for efficient routing, dispatching and collision avoidance of the vehicles. A common method for collision avoidance is zone blocking, in which the path is partitioned into zones, and a vehicle may never enter a zone already occupied by another vehicle.
Depending on the design of the AGV, each vehicle can carry one or several loads in the form of pallet-loads, tote trays, etc. In some cases, the AGV can automatically load or unload the materials at any of several stations along the path.
There is a multitude of AGV types. The towing vehicles were the first types of AGVs introduced. They became very popular since they can pull a magnitude of trailer types. The AGV fork truck has the ability to service loads where the height of the transfer varies at stop locations (i.e. at floor level, stands and in some cases racks), whereas the AGV pallet trucks are designed to only pick up and drop off loads at floor level. Other AGV types include unit-load vehicles, light-load vehicles and assembly-line vehicles. We refer to Müller (1983) for an extended discussion for reasons for implementation, the implementation and experiences with AGVs.

There are large similarities between the control of AGVs and the control of manned vehicles equipped with vehicle-mounted wireless truck-terminals. Vehicle-mounted terminals receive signals transmitted over a radio channel. These radio frequency, in short RF-terminals, make it possible for manned vehicles to be controlled similarly to AGVs. In that case, the driver of the vehicle reads the instructions from the display of the vehicle-mounted terminal. The instructions are usually a set of letter and number combinations that correspond to unique locations and loads within the facility. These inform the driver to go that specific location to retrieve a specific load and transport it to the instructed destination. The route that has to be taken is usually left to the driver. This gives the driver (and the system) a high degree of flexibility in terms of path selection in case of disruptions, obstructions by other vehicles, congestion, crossing behavior at intersections, bi-directional traffic flows, etc. Furthermore, no physical wired guide paths are necessary. The vehicles operate like free-ranging vehicles and travel on so-called virtual flow paths.
Since automated and RF-guided vehicles can be controlled similarly, we will generally refer with ‘Guided Vehicles’ (GVs) to both types of vehicles in this dissertation.

1.3 Layout design

It was indicated in the previous section that the choice of material handling equipment depends on the layout of the facility. For example, a wide straddle carrier cannot travel through small aisles and the height of a truck must be smaller than doorways and corridors between certain areas. Vice versa, one cannot move any material at all if there is no space in which to move them. Furthermore, if the space of the facility is efficiently used, the facility can be relatively smaller, which leads to less personnel and handling equipment being needed and reduced costs. This indicates that plant layout and material handling are tied together. This section briefly discusses some issues concerned with layout design.

The procedures to aid facility planners in designing layouts can be classified in two main categories: construction type and improvement type.
Improvement type layout methods generate layout alternatives based on an existing layout. The need for improvement of the facility layout can arise under a variety of circumstances. For example, changes in the processing sequence for products, changes in production quantities (introduction of new products or elimination of products) or replacement of equipment can all trigger the need for changes in the facility layout.
Construction procedures basically involve developing a completely new layout. The construction type determines the possibilities and flexibility left for the improvement type layout methods. For example, it is almost impossible to change the structure of the facility once it is built. If layout changes are frequently required, it is desirable to plan for change and to develop a flexible layout. This flexibility can be achieved with a modular design of the facility to cope with changes in the volume or types of products, without disrupting other parts of the facility.
The size and type of building are also very important for the choice of handling equipment. Facilities with open-air activities (for example outside storage, loading or unloading activities) need weather resistant handling equipment. In general, the costs of a building increase as the building covers more ground. Since land is expensive, storage buildings are usually short and relatively high. Conventional warehouse buildings, with less than 7 meters overhead clearance, are designed to use vehicles like counterbalanced lift trucks to take care of internal transport activities. Conventional buildings up to 13 meters are often equipped with high-bay stacking or order picking trucks. Higher storage buildings like high-bay warehouses need cranes to store materials. Some other possible aspects that determine the structure of the building are:

- the weight-capacity of the floors
- separating areas with fire-walls
- the location of the receiving and shipping lanes with respect to exit roads
- the location of the management offices, etc.

Common criteria and objectives for the layout within a facility include minimizing costs, distances, delays and congestion by efficient utilization of space, equipment and personnel. Furthermore, the layout should be expandable and have an efficient flow of material and information.
Muther (1973) developed the Systematic Layout Planning (SLP) procedure. This step-by-step process eventually leads to a generally accepted facility layout. It is generally used for designing new facilities. The SLP approach consists of four phases. In the first phase the

location of the facility is determined. In the second phase, the general layout is determined based on an analysis of the relationship between activities and a material flow analysis. The space required for each activity and the size of different departments of the facility is also determined in this phase. In the third phase, a number of detailed layout alternatives with the general positions of material, locations and departments are developed and evaluated. In the fourth stage, the preferred alternative is recommended and the different aspects of the facility are realized and the equipment is installed.
We have argued that the design of the facility layout and the material handling system are greatly dependent on each other. It is difficult to determine which to design first, the material handling system or the facility layout. Since material handling and layout are tied together, it is best to design both simultaneously. One method is to design a number of alternative handling systems, to design the appropriate layout for each and choose the preferred alternative.

1.4 Facility control systems

Automated control of material requires the awareness of the location, amount, origin, destination and schedule of materials. These are the functions of automatic information, identification, communication and control technologies. It is important that the right choice of information, communication and control systems is made in the design stage of a facility, since it is difficult to replace and adapt equipment and software once the facility is in operation. These systems can be categorized in different levels, see Jacobs et al. (1995) which are represented by Figure 5, see also De Koster and Van den Broek-Serlé (1999).

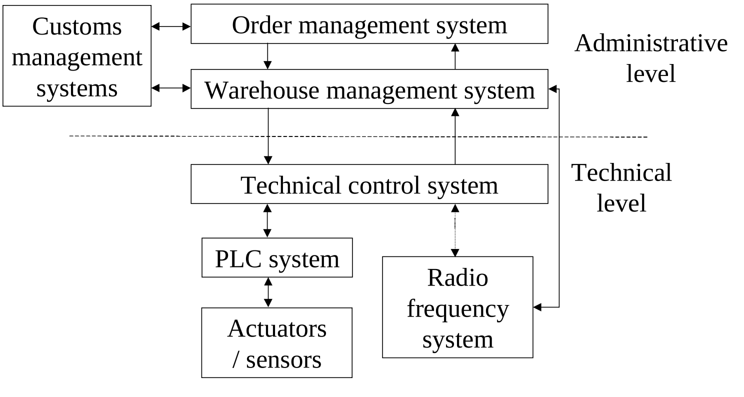

Figure 5. Hierarchical structure of company information systems

At the administrative level, central information systems such as ERP systems, order management systems and warehouse management systems (WMSs) are used to manage financial issues (invoices, wages), purchase orders, checking customer credit, sales, inventories, etc. These systems usually concern issues with a long-term horizon. At the technical level, technical control systems are responsible for timing and locations of order assignments, for different storage and order-picking strategies and route planning (management of product-location and movements). These processes are carried out by systems at the administrative level in cases where facilities have one integrated WMS for all tasks. The issues concerned at this level are more dynamic compared to the administrative level.

Another level, the operational level, can be distinguished for many facilities. This level is often (partly) integrated with the technical level. At the operational level, control systems are used for controlling and positioning equipment (such as AGV or automatic cranes), material flow control, and communication between different handling systems (data-transmission). The information at this level is short-lived and decisions and communications are based on real-time events. The speed and complexity of the issues at this level often demand a powerful system, which communicates or is integrated with the systems of the technical level.
At the operational level, wireless communication between handling systems is based on infrared or radio frequency signals. The advantages of both techniques are fast and flawless communication, which can be used for real-time data-transmissions. Radio-frequency identification uses transponders (tags) in the floor or attached to objects (loads, pallet racks, etc.). The transponders and receivers communicate using radio signals. Infrared systems are generally used by (laser) scanning devices in combination with bar codes (license plates) or labels. Other identification and communication systems include voice recognition, optical character recognition, vision systems and magnetic stripes.

Standard-software packages can be used for the standard procedures most often found at the administrative-level, see De Koster and Van den Broek-Serlé (1999). Using standard-software saves development costs and time, and is often based on proven technology, i.e. other facilities have used the software and (most) start-up errors have been eliminated.
The procedures at the technical and operational-level are less often standard and standard software will be insufficient; especially, if the facility operations are unique or relatively more complex. In this case standardized-software should be customized (tailor-made) or customized-software should be developed.
Similar to data-transmissions that are used within the facilities for communication between equipment, EDI (electronic data interchange) can be used between facilities to exchange standard messages electronically. EDI-messages generally concern purchase orders, invoices, advance shipment notices, order responses and product master data. A major advantage is that information (order-lists, etc.) can be received electronically and directly linked to or interfaced with information systems which can automatically process the order, before the material is physically collected. This time can be used to collect the material from the storage facility to reduce space, costs and waiting time of trucks eventually collecting the material.
As mentioned above, many facilities use a separate information system for material-flow control. These systems are designed to utilize personnel and equipment (vehicles)

efficiently, and are often equipped to interface with radio frequency or infrared communication systems. The next section discusses control systems at the operational-level for vehicle-based internal transport environments in more detail.

1.4.1 Vehicle-based control systems

A fundamental problem of the vehicle control problem is to determine which vehicle should transport which load and when. One can compare this with the dispatching strategies of taxi companies. In that case, customers call in specifying that they want to be picked up at a certain location and then dropped off at some other location. Both the taxi company and vehicle dispatching problem have in common the situation that serving any demand results in the need to visit some other point. The desired times to be picked up and dropped off may be specified, or can be assumed to be as soon as possible. Vehicles that are dispatched to the desired locations serve demands or requests. Typical objectives are to minimize the time that customers must wait for a vehicle or to minimize the number of vehicles needed to serve all customers at some level of service. Similarly, the performance of a vehicle-control system can be measured in several dimensions, such as, flow times, delays, conformance to due times, required throughput, waiting time of goods, idle time of transport vehicles, etc. In internal transport environments, vehicle-based control systems are concerned with the issues of dispatching the vehicles. The efficiency of a vehicle control system is sensitive to operational design parameters, which include: vehicle path layout, track capacity (uni- or bi-directional), track control, the number of trip exchanges between load transfer points, location of the transfer points, the number of vehicles needed, design of the vehicle (single or multiple-load capacity), system reliability, and the logic of the vehicle-control system.

The realization and installation of vehicle-control systems is complex. Some installations face start-up problems directly related to the vehicle-control system, others with the materials such as the tags, RF-equipment, computers, vehicles, etc, (especially if different parts of equipment come from different vendors). Although such problems are usually resolved in a reasonable amount of time, some systems are continuously plagued with problems. As a result, the vehicle-control systems and hardware/software associated with it, not only define the critical path in the system installations, but also are becoming increasingly complicated and more expensive to develop. Vendors and suppliers responding to the demand for vehicle control systems are developing increasingly sophisticated systems. Many of these intelligent control systems use a central computer application, which continuously monitors all the vehicles and the status of the system. In some cases, these systems are tailor-made (customized) with dispatching rules which best suit the objectives and needs of the environment. For example, when certain vehicles have to take care of all inbound jobs, or loads must always travel in pairs, or idle vehicles must park at a specific place.

In general, dispatching rules can make use of two types of operating decisions. The first determines which load should be matched to a vehicle when the vehicle is ready for the next task (vehicle-initiated dispatching). The second determines which vehicle is selected when loads initiate transportation requests (workcenter or load-initiated dispatching). In most cases, vehicle type and status are the only factors that dictate whether a certain vehicle is a possible candidate for a task assignment or not. There are two approaches to task allocation with respect to vehicle status. One approach is to allocate a vehicle to a transport order as soon as the request is received. The second is to defer the task assignment until a vehicle has completed its current task. The first approach can also be used for pre-planning where every vehicle can be allocated a sequentially ordered list of jobs or a set of outstanding transport orders simultaneously. In more complex situations it might be possible to add jobs to the job-list of the vehicle. The advantage of pre-planning is that all vehicles compete to satisfy a given objective, such as to minimize vehicle empty travel time or minimize total travel distance. However, there is an implicit assumption that the control system keeps track of all tasks and is able to monitor the vehicle task lists in advance.

Two situations are possible with the second approach when the central controller receives a move request. The first is that one or more vehicles are immediately available. In that case, a vehicle can be selected by the controller based on an assignment rule. In the second situation, all vehicles are busy and a move request is put in a queue of transport orders. When a vehicle completes the task, the controller selects a transport order from the queue and assigns this one to the vehicle. When two or more locations have outstanding requests, the transport order is again selected based on an assignment rule.

In most cases the decisions have to be made based on real-time events (see also Section 1.1). Due to the high degree of stochasticity/randomness within each transport environment, the vehicles are dispatched on-line. This is why on-line control rules are commonly used in distribution and production environments since load-vehicle scheduling in advance is near to impossible. A company might know beforehand that a car, truck, train or ship will arrive that day to bring or pick up products, but it is not known exactly when the truck or ship will arrive, which products are involved or in what sequence they are loaded. Other reasons, such as failure of equipment and avoiding deadlocks, have also led to the use of on-line vehicle dispatching.

1.5 Motivation and outline of the dissertation

It is the aim of the dissertation to structure the decisions that have to be taken regarding vehicle-based internal transportation, and provide results (a framework) by which decision makers can decide which vehicle control system fills the need and guarantees the desired performance best. The decisions can be subdivided into three broad areas:

- Strategic decisions

- Tactical decisions
- Operational decisions

At the level of the strategic decisions, the first step is finding the relevant long-term constraints and performance criteria, such as: estimating material flows, and the type of systems needed to handle the flows. A further step is, given an estimate of the material flow, to develop a well structured layout for the stations to be visited and the vehicle path network, to estimate the number of vehicles needed to handle the material flow and decide which operating system is needed to control the vehicles; i.e. the tactical decisions. Operational decisions involve all day-to-day operational and scheduling decisions. This means decisions about which vehicle is matched to a load (or vice versa) and when.
There is a high interaction between the strategic, tactical and operational decisions. For example, the number of vehicles is dependent on the control of the vehicles, and the control of the vehicles is dependent on the restrictions, the performance criteria and the layout. An integrated approach to these levels of decision seems impossible and one often uses a nested approach, where first the strategic decisions are made based on rough tactical and operational ideas, followed by fine-tuning the tactical and operational decisions. Although the focus of this dissertation is on the operational level, other issues concerned (layout, vehicle design, etc.), will also be studied.

It is in general difficult to address all issues concerning the operational decisions. However, in order to justify investments for vehicle control systems it is necessary to investigate possible reductions in the number of vehicles needed (and drivers if FLTs are used), and to indicate the impact on response times and throughput times. If there are too many vehicles, then the capital investment is too high, and there is also a greater probability of congestion. The latter will lead to high waiting times for loads to be picked up. On the other hand, if there are not enough vehicles, load waiting times will be high too, and due times will not be met.

The objective of this dissertation is to gain more insight into issues concerning internal transport and relative performance of common dispatching rules used for internal transport systems using guided vehicles. The analysis should be placed in a setting where management using or considers using a guided vehicle system for internal transport wants to know the effect and performance of often used and well known vehicle dispatching rules to choose the best rules for their environment.
Next to some theoretical models, three internal transportation environments of three different companies are discussed in detail in this dissertation. The theoretical models are based on integer programming and heuristic algorithms. However, internal transport environments in practice are too complex for theoretical models. To validate and investigate some ideas and results of the theoretical models for practical situations, the three companies are modeled and studied using detailed simulation studies with real company data. The work and studies at those companies have resulted in case projects in which some vehicle dispatching rules have actually been implemented. In order to relate and make our results accessible to industry, we analyze generally vehicle dispatching rules

that are easy to understand by practitioners and easy to implement in logistics software packages. Although this implies that the overall best or optimal dispatching rule associated with a given material handling system and internal transport environment may not be found, our analysis is more realistic from a practical point of view.

We conclude this introductory chapter with an outline of the remainder of this dissertation. In Chapter 2, an extended overview is provided concerning issues that deal with (automated) vehicle-based internal transport. The overview is supported by relevant discussions by other researchers that have shared their results in the literature. This chapter, partly based on Van der Meer (1999), is meant to explain the issues involved in vehicle-based internal transport and to help define the context of the problems involved with internal transport in more detail.

In Chapter 3, the research discusses mathematical modeling of vehicle-based internal transport systems. Based on Van der Meer and De Koster (1999a), two basic warehouse layouts are modeled and the internal transport is formulated as a pick-up and delivery problem with time windows. When all move requests at a facility are assumed to be known in advance, an efficient schedule can be made off-line to move all requests while meeting certain performance criteria. The idea is to use the performance of off-line control as a benchmark for on-line control. Results show that considerable performance gains can be realized if all information is known beforehand and if off-line control is used instead of on-line control. We investigate how many extra on-line controlled vehicles are needed to approximate off-line performance. Furthermore, the effects of increasing the on-line controlled fleet capacity by using dual-load vehicles are also investigated.

The availability of exact prior information is not very likely in practice, since last minute updates and unexpected failure of equipment create a stochastic environment. Scheduling vehicles or loads a complete day in advance is therefore near to impossible. In fact, the longer the planning horizon, the less reliable the information will be.

In Chapter 4, the model of Chapter 3 is extended. Off-line control is compared with on-line control rules using pre-arrival load information. The load release times are in that case given a few moments in advance, i.e. before the load has actually arrived at the location where it can be picked up. Although only a small portion of information is given beforehand, the extra time, which otherwise would have been vehicle idle time, is now used by vehicles to travel to the next released load. With the use of on-line control with pre-arrival information, load waiting times reduce considerably. It will also be shown that load waiting times can reduce considerably when a right dwell point strategy for idle vehicles is used, i.e. by parking idle vehicles near or at locations where the next transport is likely to be requested. Furthermore, we investigate the effects of assigning the closest moving vehicle to a load. The distance to the vehicle is calculated by calculating the distance which the vehicle still has to travel to finish its current assignment plus the distance the vehicle then has to travel to the requesting load. We also look at the effect on load waiting times when off-line controlled vehicles encounter small perturbations in the actual release times of loads. In this case, the vehicles arrive a little later or earlier than the expected release time of the load, while on-line controlled vehicles are dispatched using

updated real-time information. This chapter is based on Van der Meer and De Koster (1999b).
Actual vehicle dispatching rules are studied for three practical cases in Chapter 5. Based on De Koster and Van der Meer (1998), Van der Meer and De Koster (1998, 1999c, 2000), a variety of dispatching rules are analyzed under numerous conditions (batch-release of loads, multiple-load vehicles, using pre-arrival information, varying fleet size, etc.). It is demonstrated that there seems to be a certain ranking for vehicle dispatching rules and that the ranking of some rules is unaffected by the various conditions.
The main results, conclusions and subjects for further research are discussed in the final chapter of this dissertation, Chapter 6. At the end of the final chapter, some guidelines are provided for selecting control systems in practice.

Chapter 2

Literature review

In this chapter the literature on the control of guided vehicles for vehicle-based internal transport systems is reviewed. Most of the relevant literature on vehicle-based internal transport systems discusses issues concerning automated guided vehicle systems (AGVSs) since automated systems pose research areas which are increasingly being explored by many researchers. Basically, the relevant issues at facilities using internal transport (such as: warehouses, job shops, manufacturing plants, terminals, etc.) can be divided into several main categories, following the strategic, tactical and operational decision areas discussed in Chapter 1. The review starts at the level of the strategic decisions with an introduction on vehicle-based internal transport systems followed by a discussion of performance criteria.

Several sections discuss issues concerned at the tactical level. Section 2.3 on design and guide path layout discusses the layout for the stations to be visited and the vehicle path network. Section 2.4 discusses literature about estimating the number of vehicles needed to handle the material flow, followed by a discussion on the design of the vehicle (uni-load or multi-load vehicles) in Section 2.5.

Operational decisions involve decisions about which control rules are needed to control the vehicles. There are two basic control approaches for vehicle-based transport systems, off-line and on-line (or real-time) control. Both types of control are discussed in detail in Sections 2.6 and 2.7.

There is a high interaction between all issues. For instance, the type of vehicle chosen influences the number of vehicles needed, and the type of vehicle depends on the design of the facility and guide path layout. Furthermore, all the above mentioned issues, including vehicle positioning strategies and traffic control, influence the performance of the vehicle transport system. The latter two are also discussed in more detail in Sections 2.8 and 2.9.

The remainder of this chapter discusses the subjects mentioned above in more detail, supported with an overview of relevant research found in the refereed literature.

2.1 Vehicle-based internal transport systems

A key element of various operations in a facility is the integration of material handling systems. Without the integration of material handling systems, many handling systems in a facility will remain as collections or islands of automation. As stated in the previous chapter, guided vehicles are one of the most important material handling equipment that can provide the link between receiving, storage, handling and shipping areas and the associated handling equipment used at those areas. Therefore, the control of the guided vehicles is important. Examples of systems which control guided vehicles are those where:

– vehicles drive in a pre-defined loop (decentralized control),
– vehicles claim loads waiting for transportation (centralized control),
– loads claim vehicles for transportation (centralized control).

In each case, the design of the vehicle path layout, the number of vehicles required and the vehicle control system plays an important role in the operation of internal transport.
At present, there is still no systematic way to select a unique vehicle path layout for an AGVS. Specification of vehicle path layout is often done by experience and common sense procedures rather than through any exact solution methodology. After specifying the vehicle path layout, the problem of vehicle requirements (the number of vehicles necessary) is addressed. Because of the magnitude and the complexity of the layout problem, the vehicle requirement problem is generally not addressed at the same time as the layout requirement problem. Thus, for a given material flow volume (the number of loads to be transported from one location to another) and vehicle path layout, the number of vehicles required to service the handling needs of the facility depends on the vehicle dispatching rules that are used by the vehicle control system and the performance criteria which have to be met. The vehicle dispatching problem (task assignment) is concerned with how load pick-up and delivery orders are assigned to vehicles, and plays a major role in the performance of internal transport.

More than 20 years ago, a summary of over 100 job shop scheduling rules in a classification scheme for simple priority rules, combination of simple priority rules, weighted priority indexes, heuristic scheduling rules and other rules was presented by Panwalker and Iskander (1977). In their paper they simply list rules and make an attempt to explain the general idea behind them and do not draw any conclusions on which rules perform better than others. Blackstone et al. (1982) and Montazeri and Van Wassenhove (1990) provide a similar survey of scheduling rules for manufacturing job shop operations. Initially, the rules mentioned in those papers are not intended for vehicle control. However, due to the similarities between assigning jobs to machines and assigning vehicles to transportation tasks, some ideas can be used to derive vehicle dispatching rules. The first major published works on AGVS can be traced back to the early 1980's, starting with papers of Maxwell and Muckstadt (1982) and Egbelu and Tanchoco (1984). Koff (1987) illustrates the major functions of AGV systems and describes how these functions are executed. The purpose is to provide a better understanding of AGV systems, and to assist in the planning of systems. Ganesharajah and Sriskandarajah (1995) provide a more recent survey of research about scheduling and dispatching AGVs. Although this survey discusses several topics concerning AGV studies, it does not provide an exhaustive overview and many main contributions are omitted.

2.2 Performance criteria

Once the decision has been made to develop a material handling system, mathematical or simulation modeling can be used to investigate the impact of several factors, like layout design, the number of vehicles needed or dispatching rules, on the performance of the system. But not all systems use the same criteria to measure the performance of the system. For example, in a system where the objective is to maximize vehicle utilization, the performance is said to improve if the vehicles are fully utilized. In other systems the objective can be to minimize the load response time or to minimize the (empty) vehicle travel time. In this section some performance criteria commonly found in literature and practice are discussed. Several performance criteria, however, are mentioned but not discussed extensively. We start by pointing out the most commonly used objectives that define the performance criteria for a certain situation:

- Maximize vehicle utilization
- Balance vehicle utilization
- Minimize the maximum makespan
- Balance machine or workcenter workload
- Minimize queue lengths
- Minimize vehicle travel time
- Minimize vehicle empty travel time
- Minimize average load response time
- Minimize average load throughput time
- Conformance to due times
- Minimize the number of vehicles

Utilizing vehicles fully means that vehicle breakdowns or malfunctions should be kept to a minimum. Erickson (1987) proposes electronic diagnostics for AGV safety and maintainability. Krishnamurthy et al. (1993) developed a heuristic for routing AGVs in a bi-directional network with the objective to minimize the makespan (i.e. the total time needed until the last request has been served). In most cases, the objective to minimize the makespan is used when all demand assignments are known in advance and vehicle routes are calculated beforehand. To avoid that certain vehicles have relatively much more tasks than others do, the makespan is minimized. Jawahar et al. (1998) minimizes the makespan while attempting to link the AGV schedule with the production schedule. In Jaikumar and

Solomon (1990), the objective is to maximize machine utilization. Furthermore, for a given level of service, they minimize the vehicle travel time. Kaspi and Tanchoco (1990) optimize a similar performance criterion. In their model, the total transportation distance (i.e. flow volume times distance) is minimized.

In most real world applications (i.e. practical applications such as those described in Chapter 5), the average service time of vehicles to load transportation requests is the most popular criterion to evaluate. The service time consists of response time (i.e. load waiting time), the loaded travel time and transfer time (i.e. load pickup and delivery time). The latter two cannot be influenced much since loaded vehicles travel directly to the load delivery location and load pick-up and set-down operations are mechanized which usually implies fixed handling times or handling times with very little variation. This leaves minimizing the expected load waiting time as a reasonable criterion for defining the system performance. It is desirable to keep the average load waiting time as small as possible in order to provide quick service to waiting trucks at the shipping lanes, release new space for small output buffers, quickly transport perishable products to cooled areas, service other handling equipment, etc. In some environments the release time of a load is directly related to the due time of the final product. If some of the parts stay excessive time in a buffer, they may delay the entire production process.

Bozer et al. (1994) indirectly minimize the expected load waiting time by minimizing the deadhead-traveling time. Deadhead-traveling time is the (empty) travel time from the point at which the vehicle is idle until the point the load is picked up. Egbelu (1987) also measures performance by minimizing the empty travel time. Malmborg (1991) provides models for tightened upper and lower bounds under different dispatching rules to minimize empty travel time. A closely related performance measurement is throughput capacity. Srinivasan et al. (1994), and Bozer and Srinivasan (1991) define the performance of their system by maximizing the throughput capacity. The disadvantage when maximizing the throughput capacity is that some workstations or vehicles can be over-utilized (have too much work to work properly). Over-utilized machines or workstations are subject to failure and can become potential bottleneck areas. If there is no slack time in the system, one failure can have large effects. Similarly, a broken down vehicle in a system where all vehicles are heavily utilized can cause a significant loss in transportation capacity resulting in higher vehicle utilization amongst the remaining vehicles, overflowing buffers, over-utilized workstations, etc.

One method to avoid over-utilization is workload balancing. The vehicle dispatching rule in Kim et al. (1999) is based on the idea of workload balancing; it tries to balance the workload between machines and vehicles as well as the workload among the machines. Other rules encountered minimize the maximum queue length, as studied by Egbelu and Tanchoco (1984) or minimize the average queue length, (Hodgson et al., 1987).

Most of the mentioned performance criteria are based on rules for push production systems. Those are rules were the source has a demand that should be served and that machines or other equipment could be starved for work. Pull-type dispatching rules that supply machines for work is more in line with just-in-time (JIT) environments. JIT systems are known to have the ability to adapt to changes in demand while maintaining greatly

reduced work-in-process (WIP) and short lead times. Gunasekaran and Lyu (1997) provide a discussion of the JIT philosophy in material handling systems. Integrated scheduling of material handling and manufacturing activities for JIT production is considered by Anwar and Nagi (1998). The objective of the integrated problem is to minimize the cumulative lead time of the overall production schedule and to reduce material handling costs. Nakano and Ohno (1999) present an AGV pull-type dispatching rule. The average utilization of the machines is taken for performance evaluation of the AGV model.
Of all mentioned performance criteria, minimizing the expected load waiting time is most often encountered. Minimizing the expected waiting time usually means that empty travel time is low, queue lengths remain small and throughput is high. Other performance measurements encountered are conformance to due times and handling the required throughput with a minimum number of vehicles needed. Due time conformance is used to load trucks or ships on time which need to transport the materials elsewhere or tune in on cycle times of cranes or other material handling systems. Using the minimum number of vehicles needed to handle the required throughput is actually an objective for all internal transport environments. The discussion on that subject is resumed in detail in Section 2.4.

2.3 Design and guide path layout

The main issues concerning the design stage include choosing the appropriate material handling equipment and determining the layout. The design of internal transport systems is greatly dependent on the allocation of floor space, layout of storage zones, and the arrangement of handling stations. A review presented by Ashayeri and Gelders (1985) discusses several types of solution procedures for warehouse design optimization. They conclude that the most practical approach is to combine analytical and simulation methods. Their contribution, however, is somewhat outdated and more recent contributions based on heuristic procedures have shown promising results. A three-phase heuristic procedure for warehouse layout to increase floor space utilization and decrease material handling is provided by Larson et al. (1997). Their procedure, however, is limited to rectangular shaped warehouses. If the warehouse is not rectangular, it should be partitioned into the minimum number of rectangular sections. Other contributions with discussions on facility planning and material handling can be found in Matson and White (1982), Mahadevan and Narendran (1990, 1994), Rajagopalan and Heragu (1997), Tompkins et al. (1996) and Kochhar and Heragu (1999).
The vehicle guide path is usually represented such that aisle intersections and pickup and delivery locations can be considered as nodes on a graph connected by a set of arcs. The arcs describe the vehicle flow or guide path that vehicles can travel on when moving from node to node. Directed arcs between two nodes indicate the direction of vehicle flow. Costs can be assigned to the arcs that represent the distance between the two end points of the segment or the time required by a vehicle to cross the arc. In Figure 6, an example is

given for a uni-directional (left), a bi-directional (middle) and a multiple-lane (right) guide path. Other network designs generally are combinations of these three network designs.

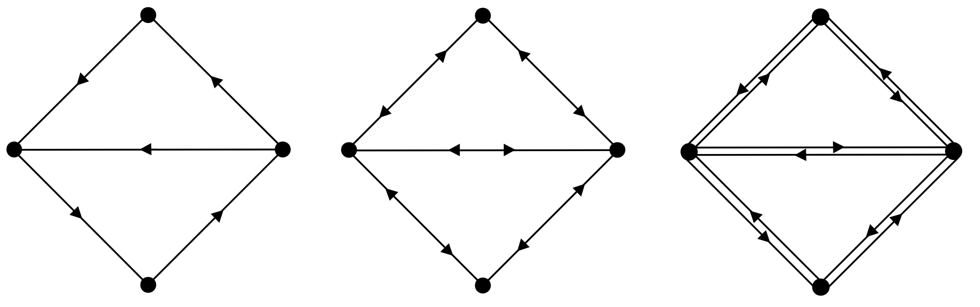

Figure 6. Example of a simple uni-directional, bi-directional and multiple-lane guide path respectively

One of the justifications of uni-directional guide paths is their simplicity in design and control. It seems justified to use uni-directional paths in large systems that employ many vehicles. Whereas the use of uni-directional flow in simple systems, which use relatively few vehicles, may not be appropriate. On the other hand, using bi-directional paths in complex systems with many vehicles can significantly increase delays due to congestion. Generally, bi-directional systems are intuitively more attractive in terms of shorter travel distances. In some simple systems, the gain in productivity that results from the use of bi-directional flow can easily compensate for the added cost in acquiring better control software or more complex vehicles. Vehicles and vehicle control for bi-directional guide paths are more expensive and complex since extra sensors and controllers for collision detection are needed (in the case of automated guided vehicles). Such vehicles are also called bi-directional vehicles, since they are usually physically symmetrical and indifferent in forward and backward travel.

The multiple-lane model is similar to the uni-directional model where nodes are connected with two uni-directional paths, both in opposite directions. The multiple-lane model is intended to overcome the deficiencies of the uni-directional model and use the advantages of bi-directional systems; however, the multiple-lane model needs more floor space in terms of vehicle paths.

The mixed model is an attempt to combine the advantages of the three models mentioned above.

Gaskins and Tanchoco (1987) and Kaspi and Tanchoco (1990) present a zero-one integer programming model for assisting in the choice for the direction of each arc on the graph and the location of pickup and delivery points. Egbelu and Tanchoco (1986), and Kim and Tanchoco (1993) compare the effect of shop throughput for bi-directional with uni-directional traffic flow guide paths. Using simulation, they show that bi-directional guide paths can lead to increased productivity. Sinriech (1995) provides a literature review on the AGV flowpath design problem. This review mainly concentrates on uni-directional flow network and travel directions aspects. Furthermore, various parameters involved in material handling flow systems are discussed, but details of the work done in individual papers are not discussed. Hsieh and Sha (1997) present a heuristic algorithm for the design of AGV facilities solvable in polynomial time. To increase productivity even further, virtual flow paths for free-ranging vehicles can be introduced. Gaskins et al. (1989) address the problem of defining such flow paths for free-ranging vehicles. The shortest loop design problems form another interesting area for investigation. In these problems, the shortest loop covering at least one edge of each cell of a block layout is determined. Asef-Vaziri et al. (2000) address such problems and propose several simplifications in order to reduce the size of the problem.

2.4 Estimating the number of vehicles needed

Vehicle-based material handling systems involve high expenses. These expenses can be split in operating costs (energy, maintenance costs etc.) and investment or capital costs. The capital costs include software, i.e. the costs of the controller performing the scheduling, routing and dispatching of the vehicles and hardware costs. Hardware costs include vehicles, controller links, guide path equipment, etc. The costs of vehicles and vehicle components especially can be very high. This means that an over-estimation of vehicles adds to unnecessary costs. In terms of economic analysis, the number of vehicles is usually minimized at a certain level of load throughput of the system. Although the actual problem is stochastic, due to breakdowns, random transport requests, vehicle blocking etc., the problem is usually solved assuming a deterministic or steady state situation. In some studies, the number of vehicles is considered as given. This can be a situation in need of re-optimization, or where a higher authority sets new assumptions.
Many factors affect the number of vehicles required to handle the throughput in a system adequately; the following list includes the main factors:

1. The system layout
2. The number of loads to be transported between transfer points per time unit
3. System reliability
4. Type of vehicle
5. Speed of travel
6. The vehicle dispatching strategy

Shelton and Jones (1987) identify more detailed attributes that can be considered in the selection of the type of AGVs for a certain system identified by the user. The presented procedure serves as a valuable decision aid, but should not be the sole basis of comparisons between AGVs. In practice, the reliability and operating speed of vehicles can be obtained for from the manufacturers (points 3 and 5). Furthermore, the layout related problems (points 1 and 2) are usually addressed before the vehicle requirement problem.

This narrows the number of factors concerned with estimating the number of vehicles needed down, but does not yet solve the problem and therefore uncovers a potential research area.

As mentioned, the number of vehicles needed to support the material handling requirements is usually determined after the layout of the vehicle guide paths has been defined. The estimation approaches are usually analytical, or simulation-based. In Maxwell and Muckstadt (1982), a methodological framework for calculating the required number of vehicles is presented. Due to the complexity of the dynamic behavior in vehicle systems, the system cannot be captured in a model involving only mathematical techniques. Newton (1985) uses a simulation methodology for determining the number of AGVs needed encompassing various problem scenarios. Tanchoco et al. (1987) compare an analytical technique with a simulation-based technique. It turns out that, when the two approaches are used jointly, the number of simulation runs required to converge to a solution is potentially reduced. Egbelu (1987) presents the use of non-simulation approaches in estimating vehicle requirements in an automated guided vehicle based transport system. However, many simplifying assumptions are made which means the methods can only be used as rough estimates. Some calculation examples based on methods of Maxwell and Muckstadt (1982) are given in Askin and Standridge (1993). Other studies determining the number of vehicles required are provided by Mahadevan and Narendran (1993) and Rajotia et al. (1998). In Gobal and Kasilingam (1991), and Kasilingam and Gobal (1996), the number of vehicles required is based on minimizing the *cost* of vehicle idle time and the waiting time *costs* of parts at machines. Sinriech and Tanchoco (1992a) use throughput performance and costs, enhanced by management decision tables, to determine the number of AGVs needed. The literature mentioned above that deals with the calculation of the number of vehicles needed is concerned with single-load (or uni-load) vehicles. These are vehicles that have the capacity of transporting just one load at a time. One might wonder whether using one vehicle with multiple-load capacity is just as good as using several vehicles with uni-load capacity.

2.5 Multiple-load vehicle capacity

An alternative to increasing the vehicle fleet capacity to improve vehicle availability for material transport is to introduce vehicles with multiple load-carrying capacity. A multiple or multi-load vehicle can pick up additional loads while transporting a previously assigned task. The use of multi-load vehicles can therefore reduce the amount of empty trip time of vehicles; also the total distance traveled is likely to reduce. However, different and often more complex considerations are involved when dispatching partially loaded vehicles compared to empty vehicles. Assigning new tasks to partially loaded vehicles may cause delivery delays for loads already loaded due to preemption of previously assigned tasks. A further complication is the selection of the next destination of the vehicle because positions on the vehicle can be assigned to loads with distinct locations. Hodgson et al. (1987)

provide a heuristic rule for a single uni-load and a dual-load vehicle traveling on a rather simple network. Their rule is based on a Markov decision process attempting to minimize the empty travel time of vehicle, and is dynamic in the sense that the destination of the empty vehicle is reevaluated each time it passes a station. The station the vehicle passes becomes more desirable if it is close to the current location of the vehicle, or if a load from a certain station has to be delivered to a station with another load waiting. The model has several shortcomings, only one vehicle is used in each experiment and the vehicle network is rather simple. Furthermore, the benefits of the rule decrease as the number of jobs per time unit increases.
Thonemann and Brandeau (1996) and Sinriech and Palni (1998) also study a manufacturing system arranged around a single loop, serviced by a single multiple-load carrier. They show that, as vehicle capacity increases, the first-encountered-first-served control rule performs reasonably well compared to optimal schedules. In another paper, Thonemann and Brandeau (1997) extended the model to a zoned AGVS with multiple vehicles with multiple-load capacity. The vehicles are controlled by a simple “go-when-filled” dispatch rule where workcenters demand raw material from a central storage depot. Özden (1988) presents a simulation study of multiple-load AGVs in a flexible manufacturing system. In the simulation study it is observed that the throughput from the FMS during a constant period of time, behaves very much in a concave fashion as a function of design factors like: the traffic pattern, the number of AGVs, queue capacity and the number of pallets. Nayyar and Khator (1993) use a network layout to compare the performance of multi-load vehicles with uni-load vehicles. Occeña and Yokota (1991) model a multiple-load AGVS in a just-in-time (JIT) environment. With computer simulations they evaluate the effects of increasing the carrying capacity of a single AGV on the ability of the system to meet high throughput requirements. Bilge and Tanchoco (1997) discuss several issues related to multi-load AGV systems and demonstrate the potential benefits of using them. They also incorporate network congestion and guide path layout design issues. In Duinkerken et al. (1996) a simulation model for inter terminal transport is presented which compares the performance between multi-load vehicles and uni-load vehicles. They show that even with extreme high numbers of vehicles, the performance of the multi-load vehicles remain clearly poorer than the uni-load vehicles. This is due to the batch-type work method associated with the multi-load vehicles. Loads can be kept back on a multi-load vehicle such that other loads that become available later can also be loaded. Other results are peaks in handling the batches of loads at the destination and in the number of vehicles waiting to be unloaded and the loss of time may lead to a decrease in performance.

2.6 Off-line vehicle control systems

Two main approaches are used for assigning pickup tasks to vehicles; one is dynamic and the other is static. In the dynamic approach, only a certain number of pick up assignments are known and used to construct the vehicle routes. When new assignments arise, the vehicle routes are re-evaluated. This means that claims of vehicles or loads can be released and requests can be reallocated. The dynamic approach is only effective if the calculations of the new assignments can be made quickly and if the information about the positions, traveling times, handling times of current assignments etc. of the vehicles is accurate. The static approach, discussed further in this section, is the case where *all* pickup assignments are made to vehicles at the beginning of the period or (daily) shift. In this case, the complete vehicle routes can be calculated and constructed off-line, i.e. some time before the vehicles carry them out. However, this assignment approach is more amenable to operations by which the supposed times of order requests are known exactly in advance.

The off-line version of the GV management problem can be modeled using the ideas of vehicle routing and scheduling problems with time windows (VRPTW). In this type of scheduling problems, a fleet of heterogeneous vehicles housed at one depot must satisfy transportation requests. A transportation request consists of picking up a certain number of customers (or loads) at a predetermined pick-up location (origin) during a departure time interval and transporting them to a predetermined delivery location (depot). These departure time windows are based on desired pick-up time requests of the customers. Although heuristics for these type of problems have been found to be effective in solving problems of practical size, optimal approaches have lagged far behind. Kolen et al. (1987) extend the shortest *q*-path relaxation algorithm of Christofides et al. (1981) and is referred to (see Solomon and Desrosiers, 1988, and Dumas et al., 1991) as the only application of the *exact* method to solve such time constrained vehicle routing problems to optimality. The largest problems solved to optimality involved four vehicles servicing 14 customers and three vehicles servicing 15 customers.

The pick up and delivery problem with time windows (PDPTW) is a generalization of the VRPTW since the destinations in the VRPTW are all the common depot and the destinations in the PDPTW can all be at a different location. This also implies that the destination should be reached during the arrival time interval, which is based on the desired delivery time requests of the customers.

The analogy with material handling environments is clear. Instead of picking up customers at a certain location from a certain point in time onwards, materials or loads have to be picked up from certain stations after they have been released and request transportation. Furthermore, the destination of the load can be at any other location of the vehicle guide path.

The PDPTW usually consists of minimizing several objectives, such as: minimization of the number of vehicles or minimization of the total distance or travel time of the vehicles. In the original pick-up and delivery problems, people are transported, and another objective encountered is minimizing the inconvenience created by pick-ups or deliveries performed sooner or later than desired by the customers. This later context of the PDPTW is called dial-a-ride problems (DARP). The analogy of minimizing the inconvenience can be seen as minimizing the expected response or waiting time of loads.

The general multiple-vehicle pick-up and delivery problem with time windows (*m*-PDPTW) involves the use of *m* vehicles. Surveys of this area are provided by Solomon and Desrosiers (1988), Desrochers et al. (1988), and Savelsbergh and Sol (1995). An extensive discussion on the PDPTW is given by Dumas et al. (1991). They use a column generation scheme with a constrained shortest path subproblem to solve cases of the PDPTW with high demands at each customer, it is not designed to solve large scale DARPs. Within the PDPTW context, the unit-load vehicle capacity problem can be defined as a multiple traveling salesman problem with time windows (m-TSPTW) with m uncapacitated vehicles. An optimal algorithm for the 1-TSPTW is provided by Dumas et al. (1995). Using a new elimination scheme taking advantage of the time window constraints, they significantly reduce the state space and the number of state transactions, which greatly enhances the performance of a relatively well established dynamic programming approach. In the studies described previously, a fleet of homogeneous vehicles is assumed. The emphasis of Leung et al. (1987) is the assignment of AGVs with different speed and load capacity, but without load pick up and delivery time windows. Using a mixed integer programming formulation the vehicles assignments are made while minimizing the total vehicle travel time. The experiments are restricted to two vehicles.

The previously mentioned vehicle routing and scheduling problems are actually classic problems in operations research and are known to be $\mathcal{NP}$-hard problems, and when each request also specifies a time window in which the request must be served, it is even $\mathcal{NP}$-complete to decide whether a feasible route for a vehicle exists. This implies that it is very unlikely that there exists an algorithm that will find the optimal solution in computation time that is polynomial in the ‘size’ of the problem. In other words, it cannot be guaranteed that instances of $\mathcal{NP}$-hard problems are solvable to optimality within reasonable time. What constitutes as reasonable time may be highly dependent on the environment in which the algorithm will be used; that is, it depends on whether the algorithm needs to solve the logistics problem for example, in real-time.

The alternative to using optimization algorithms is to use heuristic (or approximation) algorithms. Desrosiers et al. (1988) use Lagrangian relaxation methods for solving the minimum fleet size problem for traveling salesman problems with time windows. For the multi-vehicle DARPTW, many heuristics have been designed to solve large scale problems. Most of these approximation algorithms are based on parallel insertion procedures, as discussed in Jaw et al. (1986).

Blair and Vasques (1987) present a heuristic algorithm for routing AGVs into tours with the objective of minimizing the maximum tour length. In their model, minimizing the length of a tour is the same as minimizing the time of a tour, and thus to reduce the largest delay encountered by any material move transaction. Such heuristic solutions can be found faster compared to exact solutions, which makes the use of heuristic procedures more attractive for dynamic environments.

2.7 On-line vehicle control systems

The uncertain and ever-changing nature of internal transport in warehouses, production plants, job shops, transshipment terminals etc., makes it virtually impossible to plan moves ahead of time as described in the previous section. Rarely will any such environment satisfy the assumption of perfect predictability of order requests. The unpredictability arises from the stochastic nature of activities in material handling environments (see also Section 1.1). The stochastic nature can also be the result of hardware/software failure, unpredictability of cycle times of material handling equipment, congestion, late arrivals of material, unpredictability of vehicle driving times due to congestion or acceleration and deceleration effects, etc.
These situations make off-line vehicle control unpractical. Instead, dynamic dispatching rules are needed to service real-time transport requests. Such control systems based on real-time decision making are also called on-line systems.
Basically the relevant literature of on-line vehicle control systems can be divided into two major categories: decentralized control and centralized control. The next sections discuss these topics in more detail.

2.7.1 Decentralized control systems

Traditionally, vehicle systems have been implemented and analyzed assuming that every vehicle is allowed to visit any pick up/delivery location in the system. For automated control, one of the simplest implementations is one in which all vehicles circulate in the same direction in a closed loop, as shown in Figure 7.

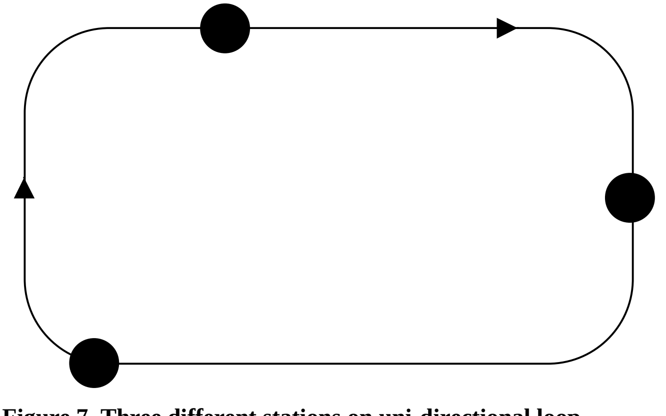

Figure 7. Three different stations on uni-directional loop

In case of a single loop, the vehicles do not communicate with a central computer or controller, but operate using only local information and on-board controlling devises. Since there is no involvement of a central computer or controller, these systems are said to be decentralized.
Decentralized control can be further subdivided into control with loops, control with multiple non-overlapping loops (tandem system), and control with multiple partially overlapping loops. A tandem system is composed of a set of loops. The transport of a load from one loop to another can be done via a load transfer point (see Figure 8), an intermediate station with a buffer at the intersection of the loops, or both vehicles must meet simultaneously for the transfer. Usually, only one vehicle services all stations in a loop.

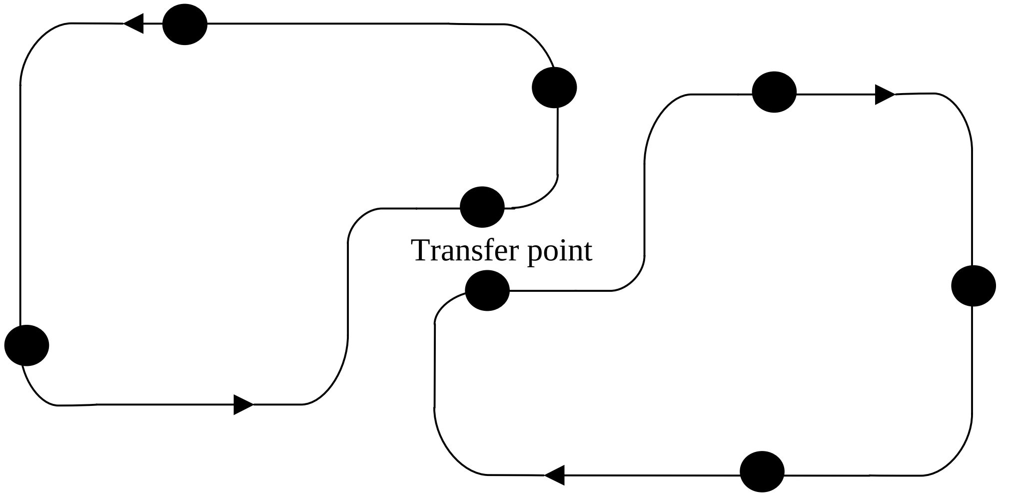

Figure 8. A tandem system composed of two loops and a load transfer point

The advantages of tandem layouts are the elimination of vehicle blocking due to congestion and the ease of traffic control. They also require a less complicated control system that can be duplicated for each loop. This can save costs in control system development and implementation. Furthermore, most of these systems can be analyzed using analytical models with standard queuing theory, or with simulation. The disadvantages of loop systems are the need to transfer loads when transporting material across loops. Also more floor space and perhaps more guide path as well as more pickup and delivery locations to interface adjacent loops will be required. Furthermore, the number of vehicles must increase as the number of loops increase, there is a low tolerance for vehicle breakdowns and the workload must be balanced to avoid over-utilizing a vehicle in a loop creating a bottle-neck loop.

A decentralized heuristic to control AGVs in a simple loop is studied by Bartholdi and Platzman (1989). In their study, a single AGV, which can carry up to three loads, is traveling a simple uni-directional loop and transports loads according to the first-encountered-first-served (FEFS) rule. With the FEFS rule, the AGV circulates a loop continuously. Whenever the vehicle has space available, it picks up the first load it encounters, which will then be delivered whenever the destination is reached. Other studies that address the performance of single-loop systems are carried out by Tanchoco and Sinriech (1992), Sinriech and Tanchoco (1992b, 1993) and Sinriech et al. (1996).

Bozer and Srinivasan (1989, 1991) introduce a conceptually simple and intuitive approach where the system is decomposed into non-overlapping, single-vehicle loops operating in tandem. They also develop an analytical model to study the throughput performance of a single vehicle loop. The model can also be used to measure the impact of using a bi-directional vehicle, reconfiguring the guide path, adding new stations and changing the flow values. Another paper by Bozer and Srinivasan (1992) discusses a partition scheme to select a set of loops for the tandem configuration. Ross et al. (1996) provides another comparative study. They compare the performance of a tandem AGV system with that of conventional AGV track systems where vehicles are allowed to visit any point on the network. The outcome is that the tandem configuration performed as effectively as the conventional control system. Similarly, Bischak and Stevens (1991) provide another evaluation of tandem configurations. Using simulation they show that, because of trips requiring delivery across loops, the tandem system has a higher expected travel time per load and thus a greater average time in the system of loads than with the conventional control system.

Srinivasan et al. (1994) present a general-purpose analytical model to compute the approximate throughput capacity of a material handling system used in a manufacturing setting. For given flow data, the model can be used to rapidly determine the throughput capacity of a wide range of handling and layout alternatives.

Johnson and Brandeau (1993) considered the problem of designing a multi-vehicle AGV system as an addition to an existing non-automated material handling system. The pool of vehicles is modeled as an $M/G/c$ queuing system and the design model is formulated as a binary program. They illustrate their model with an example of an actual design problem, and present computational experience for other example design problems. Also using queuing theory to model an AGVS as a closed queuing network, Wysk et al. (1987) use a spread sheet analysis for evaluating AGV systems.

More general than loop or tandem layouts are the vehicle systems on network layouts; these will be discussed in the next section.

2.7.2 Centralized control systems

The second control type is centralized control. In centralized control systems, a central controller or computer keeps track of all movements regarding internal transport. Such control systems are also called knowledge-based systems because they use a database with information about where loads are to be picked up and/or delivered, about vehicle (last) positions and status, and assigns loads to vehicles (or vice versa) according to specified logistic rules. In general, all tasks related to the management of the vehicles are carried out by means of such a knowledge-based central controller. These tasks include:

- Maintaining a database on status (and possibly the location) of each vehicle
- Checking the input and output queues of workstations, buffers or machine centers
- Receiving transport orders for loads
- Prioritizing and keeping track of outstanding transport orders
- Assigning transport orders to specific vehicles

The centralized controller communicates with local controllers at the queues and controllers on board the vehicles to perform these tasks. The vehicle-task assignment made by the central controller can be invoked on events such as:

- Completion of a delivery task of a vehicle
- Completion of a pickup task of a vehicle
- A transport request by a load

The first two are referred to as vehicle-initiated dispatching rules, the last one is referred to as load or workstation-initiated dispatching rules. The basic idea of vehicle-initiated task assignment rules is prioritization of outstanding move requests based on some parameter such as the distance from the vehicle, the queue size of the workstation with the move request or the elapsed time since the move request was transmitted. Dispatching rules for both task assignment types have been proposed by various researchers. Variations of the nearest-workstation-first (NWF) rule are the most commonly cited rule. This rule has also been referred to as shortest-travel-time-first (STTF) by Egbelu and Tanchoco (1984) and the vehicle-looks-for-work (VLFW) rule by Newton (1985). Under this rule, a vehicle is sent to the closest load with a transport request. The closeness of a load can be defined in terms of travel time or distance. In case the closeness is defined as travel distance, this rule is also referred to as the shortest-travel-distance-first (STDF) rule. Although the advantage in terms of minimizing empty travel time is obvious, there is also a disadvantage. The rule is sensitive to the layout of load locations in the facilities. Since vehicles are available for reassignment when they are released from a previous assignment, the release points generally correspond to delivery locations. If the pickup point of some load turns out not to be the nearest to any vehicle, according to the NWF rule, such a load may never qualify to be transported by a vehicle. Since new deliveries could continue to take place, the output queue of the affected load-location will grow to its maximum capacity.
Using dispatching rules that take queue sizes into account can decrease the probability of overflowing output queues. In Egbelu and Tanchoco (1984) and Sabuncuoglu (1998), several queue size rules were introduced, such as the maximum-outgoing-queue-size (MOQS) rule and the minimum-remaining-outgoing-queue-space (MROQS) rule. With the MOQS rule the decision is to dispatch a vehicle to the workcenter with the largest number of loads waiting to be picked up in the outgoing queue of that workcenter. Dispatching decisions under MROQS is based on the remaining space that is available at the outgoing queues. The vehicles are dispatched to the workcenter with the minimum remaining space in the outgoing queue. The basis of this rule is to reduce the possibility of workcenter blocking.

Characterizations of automated guided vehicle dispatching rules can be found in Egbelu and Tanchoco (1984). In their paper, some heuristic rules for dispatching AGVs in a job shop environment are presented.
Taghaboni and Tanchoco (1988) describe an intelligent controller for a fleet of free-ranging AGVs. Their vehicle controller will perform dispatching, routing and scheduling tasks for the vehicles, and is capable of detecting and preventing collisions before it occurs. A comprehensive review and discussion of the procedures proposed in the literature of research on vehicle management systems is provided by Co and Tanchoco (1991). In the same year, King and Wilson (1991) present a review of AGV systems design and scheduling. In that paper they give a review of the literature relevant to the system design, routing and scheduling, and justification and implementation of AGV systems. An approximate analytical model to estimate the expected waiting times for move requests that occur in single-vehicle trip-based handling systems is presented in a paper by Bozer et al. (1994). They assume that the empty vehicle is dispatched according to the modified first-come-first-served (MOD FCFS) rule, which is comparable in performance to the shortest-travel-time-first (STTF) rule, which they also introduce in their paper. The MOD FCFS rule is a modification of the traditional first-come-first-served rule in which vehicles are assigned to pickup loads sequentially in chronological order as requests for vehicles are received. When a request for a vehicle is placed and the request cannot be immediately satisfied, the time of the request is saved. The saved request and time are used for future assignment decisions. Mantel and Landeweerd (1995) give a discussion of operational control with a centralized vehicle control system. They also mention a classification of vehicle control with time windows. In a case study they try to improve the lead time performance using different dispatching rules.
Kodali (1997) describes a knowledge-based system for selecting an AGV and a workcenter from a set of workcenters simultaneously requesting the service for transport of a part. Hwang and Kim (1998) use so-called bidding functions to dispatch vehicles in an AGVS. Wang and Hafeez (1994) used Petri net models to compare the performance of tandem and conventional systems. Petri nets are a formal graphical modeling tool well suited for the description of systems which exhibit synchronization for shared resources (like locations or transfer points). See also Zeng et al. (1991) for a set of formal definitions of Petri nets modeling for AGV systems. Similarly, Yim and Linn (1993) used Petri net simulation models to analyze the effect of different AGV push and pull dispatching rules. Klein and Kim (1996) compare several single-attribute dispatching rules with multi-attribute rules of which some can choose the next transport task considering multiple criteria based on fuzzy logic. They showed by simulation that multi-attribute rules that partially include the NWF rule in their decision making process can outperform single-attribute rules such as the FCFS and MOQS rule.

Since the requirements of vehicle control systems are becoming more demanding due to the increasing complexity of the environment, the need of AGV systems with increased flexibility and reliability has increased as well. Some argue that it has become more important to be able to dynamically change the AGV job queue and path. Control systems should be smarter and be able to improvise if necessary, especially in the presence of

interruptions. Interruptions in AGV systems are not uncommon. A few situations that lead to interruptions are: AGV malfunction, delays caused by objects on the AGV paths or manual intervention. Narasimhan et al. (1999) use simulation to analyze re-routing AGVs that encounter interruptions. A route database is used to obtain quickly previously generated tasks when a vehicle is interrupted.
Hao and Lai (1996) use a self-organizing neural network to transport the requests in a non-conflicting manner and in the shortest time. Bostel and Sagar (1996) present a neural-network-based method for dynamic control systems for AGVs. The results of their simulations indicate that the system provides increased flexibility and allows a vehicle to deal with the situation in which an AGV breaks down. Another typical approach to solving a dynamic problem is to treat the problem as a series of static models solved on a rolling horizon basis (see Rachamadugu et al., 1986). In this way, exact algorithms can be used. The solution can be updated at regular intervals or whenever a change in the status of the system occurs. However, such an approach can be computationally impractical in real-time operations and is not discussed much in AGV literature.

2.8 Vehicle positioning strategies

Vehicle idleness occurs when a vehicle has completed a task but there is no immediate pickup task to reassign to the vehicle. At that time, the vehicle controller can instruct the vehicle to park at a certain location to wait for further transport orders. Those parking locations are also called the home locations, depots or dwell points of the vehicles when they become idle. Unless a material handling environment is overloaded, the occurrence of vehicle idleness is inevitable. It is desirable to reduce the empty vehicle travel time from its present location to a request at a workstation to deliver a load to another workstation, since this travel time also imposes load waiting time. One of the control decisions in the operation of vehicle systems is to determine the home or parking locations of idle vehicles. The objectives concerned include:

(a) minimization of maximum vehicle empty travel time from the parking position to the load pickup point,
(b) minimization of average vehicle empty travel time, and
(c) balancing distribution of idle vehicles in the network.

The easiest alternative in a system with random transport patterns is to park the vehicle close to the station where it last unloads. However, doing so may cause congestion around that station. The key to effective vehicle parking is to distribute the vehicles so that they are strategically located for future demands. One can think of positioning taxi's or emergency vehicles on strategic places in a city such that they can respond quickly.

The following rules used for positioning idle vehicles have been discussed most in the literature:

1. Central zone positioning rules
2. Circulatory loop positioning rule
3. Point of release positioning rule

With a central zone positioning rule, certain parking areas in the vehicle network have been designated for buffering idle vehicles. Regardless of the position where the vehicles become idle, they are routed to parking areas to await reassignment. These areas can be close to stations with a high probability of a load transport request (hence reducing expected load waiting times), or at battery- recharge or fuel stations.
With a circulatory loop positioning strategy, one or more cruising loop for idle vehicles are defined. When a vehicle becomes idle, it travels to one of the loops until a transport order is received.
With a point of release positioning rule, a vehicle remains at the point of the last release until it is reassigned. One disadvantage of this rule is that idle vehicles can block the path to surrounding stations for other vehicles.

Literature for dwell point positioning has received considerable attention concerning automated storage and retrieval systems (AS/RS). In these systems, the dwell point is the position where the storage and retrieval machine resides when the system is idle. The dwell point positioning problem for AS/RS is usually studied for situations where one storage and retrieval machine serves one aisle of the AS/RS. Van den Berg (1996) discusses such systems in detail.
Most of the literature that discusses dwell-point strategies for guided vehicle-based systems involves studies of selecting a home location of a vehicle in a single loop. Chang and Egbelu (1996) discuss such problems. They calculate the best position to minimize the expected response time for a single vehicle in a uni-directional and a bi-directional loop. Kim and Kim (1997) consider a uni-directional guide path with exponentially distributed interarrival times between the orders and a single vehicle. Because of its simplicity, the system can be modeled as a discrete-time stationary Markov chain.
Literature on the use of multiple vehicles in a loop is rare due of traffic control problems involved in such systems. Egbelu (1993) however, used several vehicles in a loop and determined how to best preposition them to minimize the empty vehicle travel time. The loop is actually made of two loops laid side-by-side with opposite flow direction to create a system similar to a bi-directional system. Kim (1995) proposes a dynamic positioning strategy where a new positioning location is assigned whenever a vehicle becomes idle. Both cases of uni-directional and bi-directional guide paths are considered. Gademann and Van de Velde (2000) consider the problem of positioning *m* AGVs in a loop layout with *n* stations. They provide an overview of time complexities for uni-directional and bi-directional flow systems and show that criteria like maximum response time and average response time can be minimized in polynomial time for any number of vehicles. Dynamic programming algorithms to find the optimal dwell point locations in a single loop system

for the objectives of minimizing the maximum response time and minimizing the mean response time have been studied by Lee and Ventura (1999a) and Ventura and Lee (1999). The same algorithms can also be applied to the (single vehicle) tandem loop and the tandem loop multiple vehicle (TLMV) layouts by considering transfer points to be additional pick-up and delivery points, see Lee and Ventura 1999b and Ventura and Lee (2000). In these studies, once the destination of an idle vehicle is determined, its travel may not be interrupted, see also Lu and Gerchak (1998). Chapter 4 of this dissertation will also study dwell point strategies and introduce a form of interruption such that not only idle vehicles can be assigned a request but that vehicles *traveling* to a dwell point can also be (pre-)assigned to a request.
In practice, the company may define the parking location because vehicles may or can only park in certain areas. These areas can be defined for safety reasons, to avoid congestion, to allow a change of drivers, to recharge the vehicle's battery, etc. Such practical issues are often overlooked or omitted in theoretical models of vehicle-based internal transport systems. The next section describes more issues which are often omitted in modeling a system since they can make the theoretical models too complex.

2.9 Traffic control and blocking prevention

To use vehicles effectively means that they should not be blocked or collide. Sensing elements on the vehicles can keep the vehicles from running too close to each other to prevent collisions. Some systems, where the distance between parallel paths is wide enough to let vehicles pass each other, still need special traffic control to guide vehicles in curves. In curves loads on the vehicles can swing out or vehicles can swerve such that passing vehicles traveling parallel in curves may collide.
In an AGVS, a deadlock (also called gridlock or systemlock) occurs when two or more vehicles are blocking each other's paths such that none of the tasks can continue any longer (stalemate). A deadlock necessitates manual intervention to move vehicles and clear the blockage causing significant loss in system performance. A simple example of an aisle-type deadlock occurs when two vehicles moving in opposite directions block each other's path, as illustrated in Figure 9.

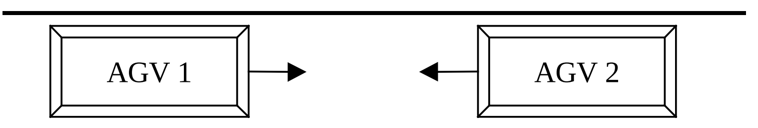

Figure 9. Example deadlock in an aisle

To avoid such problems in practice, automated vehicle systems are usually equipped with traffic control systems.

Systems that divide the total path into sections or zones to properly control the traffic are very popular, see Malmborg (1990). The traffic controller prevents two vehicles from ever being in the same zone. Vehicles travel from one zone to another and a vehicle is stopped if it attempts to enter a zone that already contains a vehicle. Zones, sections, pick up locations etc., can be recognized using labels. Labels can be magnets or sensors embedded in the floor. A label marks each section. When a vehicle crosses a label, it can be instructed to stop, slow down or perform some other task. Similarly to zones, Faraji and Bata (1994), and Heragu and Gupta (1994) form cells to eliminate vehicle interference and system locking. Taghaboni and Tanchoco (1988) incorporated a subroutine in the routing procedure to check if more than one vehicle can pass an intersection simultaneously. Krishnamurthy et al. (1993) developed a column generation based heuristic for conflict-free routing of AGVs in a bi-directional network. Lee and Ventura (2000) use a colored Petri net model in which deadlocks are represented by circuits in order to provide conflict-free routings in automated handling systems. Control laws for deadlock free operation are addressed in Wu (1999) in which necessary and sufficient conditions are modeled using a Petri net model. The control laws are easy to implement and can be embedded into a real-time scheduler. Lindeijer and Evers (1999) introduce the concept of agile high-performance traffic-control systems. Their traffic-control system, called TRACES, can handle high traffic densities on any scale. Semaphores are introduced to indicate freely available capacity such that the vehicle capacity of tracks or intersections can be protected.

2.10 Concluding remarks

In this chapter the relevant control issues at facilities using vehicle-based internal transport were discussed, supported with related research found in the refereed literature. Literature on on-line vehicle scheduling with the objective of minimizing load waiting times is encountered most, but not abundantly. Most literature is on very small-scale (loop) models with little relation with real-life situations. The theoretical models that do exist are usually simplifications of practical situations or simple loop models meant as a theoretical exercise. Practical cases are usually too large for theoretical models and are discussed in simplified simulation studies, thereby losing general applicability. Studies of large case studies are usually made and reported exclusively for the firm concerned and rarely reported in the literature. Furthermore, literature comparing different case studies seeking commonalties for general applicability has not been encountered. This is one of the objectives of this dissertation.

The objective in the next chapters is to relate theoretical vehicle routing problems with real-time internal transport. We will see the difference between optimal and practical control of internal transport, and why optimal control is impossible in practice. Alternatives are provided to approximate optimal vehicle control using real-time dispatching rules and a sensitivity analysis of the optimal vehicle routes is given. Chapter 5, following the chapters with the theoretical studies, provides detailed simulation studies of three large-scale practical case studies concerning operational control of internal transport, and relates some of the theoretical results to practice. These studies are meant to help close the gap in knowledge concerning the general performance of (common) dispatching rules for internal transport in practice. Chapter 5 also contains a brief extension of the literature review. The literature review of Chapter 5 concerns the transshipment operations at container terminals, focussing on the relevant control issues concerning vehicle-based transport of containers.

Chapter 3

Off-line versus on-line vehicle control systems

Vehicles are usually dispatched on-line in computer controlled vehicle-based internal transport systems, since information on load release times, origins and destinations is available only at the last moment. Based on the available information, real-time decisions are made to match loads and vehicles accordingly to serve the load transportation requests. In the theoretical case of off-line control, all load origins, destinations, release instants and transportation times are known in advance. In this case, exact algorithms (such as mixed integer programming) can be used to calculate vehicle routes in such a way that a certain objective function is optimized. Heuristic rules can also be used to quickly find good and sometimes optimal solutions for off-line vehicle control.

As mentioned before, off-line control is possible if all move requests at a facility were known in advance. This means that all necessary data must be available in time and ready to be used for computer algorithms. With the growing use of electronic data interchange (EDI), more information can be made known in advance. However, the data is usually not exact and complete enough as far as timing and locations of loads are concerned to allow off-line control. Scheduling vehicles or loads a complete day in advance is therefore near to impossible. In fact, the longer the planning horizon, the less reliable it will be.
Since it is our objective to compare off-line and on-line control performance, we assume that all information needed for off-line control can be obtained and vehicles can be scheduled optimally. Vehicles are controlled off-line by formulating the situation as a multi-vehicle pick-up and delivery problem with time windows, which is solved using mixed integer programming (MIP) for the situation of single-load vehicles (see Section 2.6 for a more extensive discussion on these type of problems). This exact method using exact information can be seen as an off-line control rule and is compared to the performance of two on-line dispatching rules using the same data made available at the moment of load-release. In the literature overview many different performance objectives were encountered. The most common objective used is minimizing the expected load waiting times, which is also a common performance objective used in practice since it is directly related to the due time of the load. When vehicles with multiple-load capacity are used with the objective of minimizing load waiting times, the load transportation times can still increase. The average load transportation times can increase since a load can remain on the vehicle while other loads are simultaneously being picked up and dropped off by the same vehicle. Therefore, when multi-load vehicles are used, the effects on the actual time

needed to pick up the load from the load origin and deliver the load to its destination is not clear. So in this chapter, the performance of a vehicle control rule is defined as the average load throughput time, and the performance is said to increase as the average load throughput time (the average load waiting time plus transportation time) decreases.
The idea is to use off-line control for single-load vehicles as a benchmark for on-line performance. The following will be investigated:

- the value of having all information needed for off-line control
- the value of additional vehicles for on-line control
- the value of additional load-capacity of vehicles for on-line control.

Solving vehicle schedules to optimality requires (MIP) algorithms that are complex, time consuming and difficult to integrate in vehicle control software. To be able to deal with the latter difficulties and our suspicions that the performance difference between off-line and on-line might be quite large in any case, a heuristic rule will also be used.
Solomon (1987) describes a variety of common heuristic rules used for vehicle routing type problems. These include Savings, Nearest-Neighbor, Insertion and Sweep type heuristics. Based on a comparison study on the performance of the heuristic rules, Solomon (1987) recommends the use of Insertion type heuristics. In this chapter we will also describe an Insertion type heuristic to be used for off-line vehicle control systems and for small problem sizes we compare the results with the optimal results. The Insertion heuristic will then also be used to study larger problem instances.

The vehicles will be dispatched in two different layout environments to investigate the effects of different topologies for different types of dispatching rules. It will also be shown that the performance difference between on-line dispatching and off-line control depends mainly on the load throughput and the spread of load-release instances. In low throughput environments, vehicles can become idle and park when dispatched on-line. When controlled off-line, this idle time is used to travel to the next load transportation assignment, hence, compared to on-line dispatching, reducing the average load waiting times and possibly the average throughput times.
In the remainder of this chapter, the problem is formulated and modeled in more detail. First, the different types of vehicle control rules are explained. Then the different layouts in which the vehicles will travel will be defined. Next, the different types of load generation instances and load throughput levels will be introduced. The results will show the effects of different combinations of layouts, dispatching rules, throughput levels etc. Furthermore, fleet size variations are introduced to increase the performance of on-line dispatching in an attempt to approximate off-line vehicle control and study the effects under different circumstances.

3.1 Problem formulation

In essence, the problem involves the satisfaction of a set of load transportation requests (jobs) by a fleet of vehicles housed at a depot. Loads can be pallets, crates, containers etc. A transportation request consists of picking up a certain number of loads at predetermined pick-up locations during departure time intervals and transporting them to predetermined delivery locations. The departure time intervals, or time windows, are based on desired pick-up time requests specified by the load release system. We consider the general case where loads are released to a vehicle-based transportation system at a certain time i.e. the release time, and need to be transported to a certain destination. The objective is to minimize the average time between the release time of the load and the drop-off time of the load at its destination. Loads have to be picked up after their release time (start of time window) by one of the vehicles and brought to their destination in such a way that the average load throughput time is minimal. The performance of a rule is defined as the sum of the load waiting times and load travel time for this chapter.

In the off-line case, where all transport jobs, including release times, are known in advance, the problem can be modeled as a multi-vehicle pick-up and delivery problem with time windows (*m*-PDPTW) where the objective is to minimize the load throughput time. The following sections describe the off-line control rules used. The vehicle routes can be optimized with mixed integer programming algorithms. The *m*-PDPTW is $\mathcal{NP}$-hard and the algorithms to solve this problem optimally are very time and memory consuming. We therefore also describe a heuristic to solve larger problem instances.

3.1.1 The multi-vehicle Pick-up and Delivery Problem with Time Windows

In the general *m*-PDPTW model the vehicles must pick up a load at the load origin between the start and end of the pick-up time window and deliver the load at its destination between the start and end of the destination time window. However, the time window formulation for the study in this chapter is different. The release time of the load defines the start of the pick-up time window, so the loads can be picked up any time after that and must be delivered directly. Furthermore, to keep the problem computationally tractable, only uni-load vehicles are used for our MIP algorithm. This is also referred to as a full-truck load problem, see Savelsbergh and Sol (1995).

A full-truck load pick-up and delivery problem can be formulated as a Traveling Salesman Problem (TSP) by representing a transportation job as a single job-node (instead of an origin and destination location-node) in which the travel time from job i to job j (t'_{ij}) equals the travel time from the origin of job i to the destination of i ($t_{i^+i^-}$) plus the travel time from the destination of i to the origin of j ($t_{i^- j^+}$). So $t'_{ij} = t_{i^+ i^-} + t_{i^- j^+}$. This origin to origin formulation would cause problems for the time window constraints for destination locations. However, in the formulation of this chapter, there are only start time constraints at the origins. The PDPTW in this special case can therefore be formulated as a TSP with time windows (TSPTW) with the objective to minimize the load throughput time, which is the sum of all load waiting times plus load travel times.

With multi-load vehicles the load travel times would not be unique since a load's travel time can increase when another load is picked up and dropped off first by the loaded vehicle. In the case of uni-load vehicles, the load transportation times are constant and the objective reduces to the minimization of the load waiting time (we will add the loaded trip times at the end). This in turn is also referred to a Traveling Repairman Problem with time windows (TRPTW) (see also Ball et al., 1995). The formulation in the next section has no restriction for the end-time of the time window; this is the main difference with other TRPTW formulations found in literature and is discussed in more detail in the next section.

3.1.2 The Traveling Repairman Problem with Time Windows

We give the formulation for the TRPTW involving a single depot (which is represented by a node where the vehicles start from to serve their first job, and return to after completing their last job) and a homogeneous fleet of vehicles for the models studied in this chapter. The notation used is listed in Table 1, the mathematical formulation is listed in Formulation 1.

Index sets	
N	set of nodes {0, ..., n + 1} for the vehicle network, indexed by i and j
P	set of nodes {1, ..., n} other than the depot nodes
V	set of vehicles {1, ..., $\|V\|$} to be routed where $\|V\|$ is the number of vehicles, indexed by v
Parameters	
n	number of load transportation jobs $\|P\|$, associate to job i a node i
r_i	release time of the load at node i, (which defines the start of the pick-up time window)
t_{ij}	travel distance/time from i to j for each distinct i, j in N (that is from the origin of load i to the origin of load j)
Variables	
x^v_{ij}	binary flow variables which equal 1 if vehicle v travels from node i to node j and zero otherwise, $v \in V$, $i, j \in N$
D_i	time at which service at node i begins, $i \in P$
D^v_0	time vehicle v leaves the start depot (node 0), $v \in V$
D^v_{n+1}	time vehicle v returns to the end depot (node n + 1), $v \in V$

Table 1. Notation for the TRPTW

As vehicles travel on the network transporting loads form one location to another, some locations are visited more than once. However, a (job-)node is associated to each transportation job in order to assign a unique service or departure-time to each job. Therefore different nodes may refer to the same physical location at which a transport request was placed. Since each vehicle starts and ends its route at the depot, the depot would be associated with several service-times. However, a variable can only be associated with one value. Therefore extra dummy service-time variables ($D^v{}_0$, $D^v{}_{n+1}$ were $v \in V$) are introduced for the depots which all refer to the same physical depot location.

$$Min \sum_{i \in P} \left(D_i - r_i \right) \quad (1)$$

$$Subject \quad to$$

$$\sum_{v \in V} \sum_{j \in N} x_{ij}^v = 1 \quad \forall i \in P \quad (2)$$

$$\sum_{j \in N} x_{ij}^v - \sum_{j \in N} x_{ji}^v = 0 \quad \forall i \in P, \forall v \in V \quad (3)$$

$$\sum_{j \in P} x_{0j}^v = 1 \quad \forall v \in V \quad (4)$$

$$\sum_{i \in P} x_{in+1}^v = 1 \quad \forall v \in V \quad (5)$$

$$x_{ij}^v = 1 \Rightarrow D_i + t_{ij} \leq D_j \quad \forall i, j \in P, \forall v \in V \quad (6)$$

$$x_{0j}^v = 1 \Rightarrow D_0^v + t_{0j} \leq D_j \quad \forall j \in P, \forall v \in V \quad (7)$$

$$x_{in+1}^v = 1 \Rightarrow D_i + t_{in+1} \leq D_{n+1}^v \quad \forall i \in P, \forall v \in V \quad (8)$$

$$D_i \geq r_i \quad \forall i \in P \quad (9)$$

$$D_0^v = 0 \quad \forall v \in V \quad (10)$$

$$D_{n+1}^v \geq 0 \quad \forall v \in V \quad (11)$$

$$\sum_{v \in V} \sum_{j \in N} x_{n+1j}^v = 0 \quad (12)$$

$$\sum_{v \in V} \sum_{i \in N} x_{i0}^v = 0 \quad (13)$$

$$x_{ij}^v \quad binary \quad \forall i, j \in N, \forall v \in V \quad (14)$$

Formulation 1. The mathematical formulation of the TRPTW

We seek to minimize the sum of the load waiting time (see equation (1)), i.e. the sum of differences between the departure time D_i of a vehicle at node i, and the release time/earliest possible pick-up time r_i of the load at that node. The corresponding objective implicitly minimizes the *average* load waiting time as well. And when the loaded trip times are added, it also minimizes the average load throughput time. If a vehicle arrives at

a node before the load is released, the vehicle must wait. Constraints (2)-(5) and (12)-(13) form a multi-commodity flow formulation, in which constraint (2) ensures that all nodes are visited exactly once. Constraint (3) in turn ensures that a vehicle arriving at a node will also leave that node. Furthermore, vehicles must leave the starting node (constraint (4)), and constraint (13) makes sure that no vehicle can return to the starting node. Constraints (5) and (12) make sure vehicles arrive at the end node and never leave from the end depot respectively.
Next, constraints (6)-(8) describe the compatibility requirements between routes and schedules, while constraints (9)-(11) are the time window constraints. Constraint (9) defines the start of the pick-up time window, since vehicles can come any time after the release time; there is no constraint for the end time.

Constraints (6)-(8) in Formulation 1 are not linear, but can be rewritten in an equivalent linear form using a large constant *M*:

$$D_i + t_{ij} - D_j \le M\left(1 - x_{ij}^v\right) \quad \forall i, j \in P, \forall v \in V \qquad (6')$$

$$D_0^v + t_{0j} - D_j \le M\left(1 - x_{0j}^v\right) \quad \forall j \in P, \forall v \in V \qquad (7')$$

$$D_i + t_{in+1} - D_{n+1}^v \le M\left(1 - x_{in+1}^v\right) \quad \forall i \in P, \forall v \in V \qquad (8')$$

Constraints (6')-(8') also impose increasing times at the nodes of the route. Thus, eliminating possible cycles. These constraints are in fact a generalization of the subtour elimination constraints proposed by Miller et al. (1960).
Since a node has a pick-up time window with a start time only, a vehicle can arrive any time after that. This will result in a large number of possible routes. By introducing an end for the pick-up time window like constraint (9') below, some routes are eliminated and thereby the speed of finding the optimum is increased.

$$D_i \le r_i + C \qquad \forall i \in P \qquad (9')$$

In this case, a constant *C* is used to form a time window of length *C* in which a vehicle should pick up the load at node *i*. However, setting the pick-up time window too narrow will lead to a suboptimal solution (possibly even an infeasible one if all feasible routes are eliminated, *C* should then be increased). The (suboptimal) value of the objective function of this 'previous run solution' can be used to set a new end time for the pick-up time window. Adding the 'previous run solution' to all load release times will create new end times for the pick-up time window (see constraint (9'') next page). This will lead to the optimal value when the MIP is run again, since the optimal value will always be smaller (or equal) than the time windows created with a 'suboptimal' answer.

More formally, since

$$\left[\sum_{i\in P}(D_i - r_i)\right]^{optimal\ solution} \le \left[\sum_{i\in P}(D_i - r_i)\right]^{previous\ run}$$

it follows that all individual waiting times of the optimal solution are smaller than the sum of the waiting times, and

$$(D_i - r_i)^{optimal\ solution} \le \left[\sum_{i\in P}(D_i - r_i)\right]^{optimal\ solution} \le \left[\sum_{i\in P}(D_i - r_i)\right]^{previous\ run} \qquad \forall i \in P$$

which leads to

$$D_i \le r_i + \left[\sum_{i\in P}(D_i - r_i)\right]^{previous\ run} \qquad \forall i \in P \qquad (9'')$$

This means that for the optimal value, the MIP should be run again with constraint (9‘) replaced with (9‘‘). However, the second run is only necessary when the solution of the first run is larger than *C*. Note that the solution of a heuristic algorithm (such as Insertion) can also be used to estimate *C*. With such an (over)estimate of *C*, only one run is necessary to obtain the optimal value. However, the running time can be relatively high since the bound can be rather weak.

3.1.3 The Insertion rule

When using CPLEX to solve the MIP model, memory problems (over 125 MB of RAM was available) and long running times (on an IBM/RS6000 model 370) were soon encountered for relatively small problems (see also Section 3.6.1). To decrease the running time, increase the problem size and increase the chance of practical implementation, we also analyzed the results with an Insertion heuristic. Insertion heuristics have been studied for a variety of vehicle-routing problems (see Solomon, 1987), dial-a-ride problems (see Jaw et al., 1986) and traveling-salesman problems (see Gendreau et al., 1992). Insertion heuristics have shown very promising results in these studies. For off-line vehicle-control based on the traveling repairman problem of Section 3.1.2, we will describe an Insertion type heuristic and compare the results with the optimal results for small problem sizes. The Insertion heuristic will then also be used for larger problem instances.

The pseudocode of the Insertion algorithm used to construct the vehicle routes off-line is presented in Algorithm 1. After sorting the jobs in increasing order in terms of the release time, the position with the cheapest insertion cost of the job is calculated for each job. This is the minimal extra waiting time needed to add job i to a vehicle route v. Since the number of candidate positions is at most n, and the number of jobs considered to one of the positions is at most n, an algorithm that enumerates all jobs for all candidate positions (the Insertion heuristic) will have a time complexity of $O(n^2)$. The algorithm is actually carried out twice in case a job can be inserted in different vehicle routes with the same costs (ties). The first time the data about the possible insertion position is not updated in case of a tie, (so the job is assigned to the first route encountered with that solution). The second time the data about the possible insertion position is also updated when a tie is encountered, (so the job is assigned to the last route encountered with that solution).

Perform jobs to job-nodes transformation;
Construct the node-list by sorting tasks on increasing release times;
for v := 1 **to** $|V|$ **do**
Initialize vehicle route v with depot $D^v{}_0$;
for i := 1 **to** n **do {**
    for v := 1 **to** $|V|$ **do {**
        for j := 1 **to** (nr. nodes of vehicle route v) + 1 **do {**
        Temporarily insert node i at position j of route v;
        Recalculate the sum of differences between departure and release times (waiting times) for all inserted nodes so far; **}}**
    Insert node i at position j of vehicle route v for which the total sum of waiting times was minimal; **}**
Report total waiting time;

Algorithm 1. Simplified algorithm of the Insertion heuristic in pseudocode

The next situation of four locations (including the depot), three jobs and two vehicles is an example of the Insertion procedure. The travel times between the four locations for this example are shown in Table 2. Table 3 gives the jobs to job-nodes transformation (origin-to-origin) and Table 4 the corresponding node to node travel times. For example, the travel time from node 2 to node 3 in Table 4, is the travel time from location 1 to location 3 (the third job row in Table 3), which equals 20 (see Table 2) plus the travel time from (destination) location 3 to (origin) location 1.

Location	Depot	1	2	3
Depot	0	10	10	20
1	10	0	10	20
2	10	10	0	10
3	20	20	10	0

Table 2. Travel times between locations

Load release time (r_i)	Job	Node (i)
0	Depart depot	0
9	From location 3 to location 1	1
15	From location 1 to location 3	2
19	From location 1 to location 2	3

Table 3. Jobs to job-nodes transformation

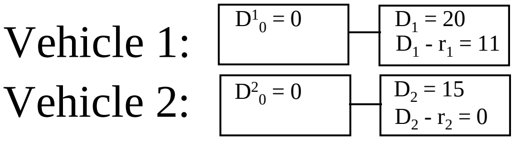

Figure 10. Load to vehicle assignments after two insertions

Node	0	1	2	3
0	-	20	10	10
1	-	-	20	20
2	-	20	-	40
3	-	20	20	-

Table 4. Travel times between nodes

Figure 10 and Figure 11 represent how the vehicle routes are constructed using the Insertion heuristic with the notation of the TRPTW for the departure and release times, the release times of Table 3 and the node to node travel times of Table 4. After two Insertion steps, each vehicle has one job and the sum of the load waiting times is as small as possible (11), as shown in Figure 10. In the next step, the third job is inserted in the most favorable position of the route for one of the vehicles, giving the four alternatives shown in Figure 11.

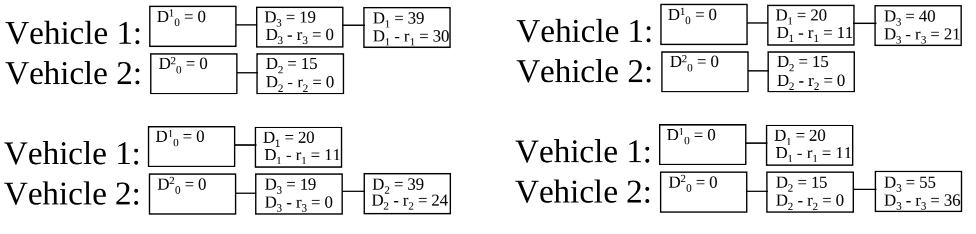

Figure 11. Load to vehicle assignment possibilities when inserting job 3

In this case, the first alternative leads to the smallest total waiting time (30), and job 3 is inserted at the beginning of the route (after leaving the depot) of vehicle 1.
In the case of more jobs and vehicles, the algorithm proceeds in a similar fashion checking all possibilities until all jobs are assigned to a vehicle.

Although this algorithm will not guarantee an optimal route, we can still use the value of off-line control systems when using a simple heuristic by demonstrating that the solutions are sufficiently close to the optimum. These solutions can still be further improved by using more advanced heuristics such as those using column generation techniques (see Dumas et al, 1991). The latter are beyond the scope of this study and will not be discussed.

3.1.4 On-line dispatching rules

Although off-line vehicle control minimizes the fleet size and load throughput times, it is impossible to obtain all the data necessary in practice. The exact time a load is released is usually not known in advance. We therefore also describe two on-line control rules by which real-time decisions of control are triggered at the release time of the loads (the moment they can be transported). Both rules use the same data used with off-line control. In this case, however, the information of the jobs is made available to the vehicles at the release time of the loads.

The dispatching rules have been carefully chosen using previous studies found in literature. From the literature overview in the previous chapter, we know that there are vehicle-initiated dispatching rules and workcenter or load-initiated dispatching rules. Furthermore, most rules are either distance-based, like shortest-travel-distance-first (STDF) or time-based, like first-come-first-served (FCFS). An example of a load-initiated time-based rule would be longest-idle-vehicle-first. Using this rule, the vehicle that has been idle for the longest time in the system is matched to the load placing a transport request. The relative advantage of this rule is unclear; except that perhaps the vehicle utilization is more balanced among all vehicles. A load-initiative rule like nearest-vehicle-first (NVF) makes more sense. In that case, the closest idle vehicle in the system is matched to the load placing a transport request. Intuitively, this will minimize the vehicle empty travel time to the load and thereby the load waiting time. Minimizing the load waiting time is important since the objective is to minimize the average load throughput time; the sum of the load waiting time plus the load travel time.
Since the first rule is a load-initiative distance-based rule, the second rule will be a vehicle-initiative time-based rule. The following sections will describe both rules in more detail.

Nearest-Vehicle-First

Under this rule, the load or workcenter has the dispatching initiative. When a load is released at a workcenter, the workcenter places a move request. The shortest distance along the traveling paths to every available (idle and motionless) vehicle is then calculated. The idle vehicle, whose travel distance to the load is the shortest, will be awoken to be dispatched. On the other hand, when a vehicle becomes idle, it searches for the closest waiting load in the system, i.e., at that point the dispatching initiative is at the vehicle and the rule used is similar to shortest-travel-distance-first (STDF). If there are no vehicle requests for loads in the system, the (empty) vehicles will park at their current locations and become idle until a request becomes available (the point of release positioning rule described in Chapter 2).

Modified First-Come-First-Served

The FCFS rule is a vehicle-initiated dispatching rule. A vehicle delivering a load at the input queue of a station first inspects the output queue of that station. The vehicle is then assigned to the oldest request (longest waiting load) at that station if one or more loads is found. However, if the output queue of that station is empty, the vehicle serves the oldest request in the entire system. If there are no move requests in the system at all, the vehicle will park at that location and becomes idle until a move request becomes available.

Multi-load vehicle dispatching

Multi-load vehicle dispatching is based on the concept of closest task. Therefore, a multi-load vehicle picks up as many loads as it can carry from its current location before moving away. When the vehicle moves, it either delivers one of its loads or picks up another load if it has remaining capacity. The vehicle only looks for additional loads to pick up that are closer in distance than the closest destination of its onboard loads. If the vehicle goes to deliver a load, it always goes to the closest among the destinations of its onboard loads.

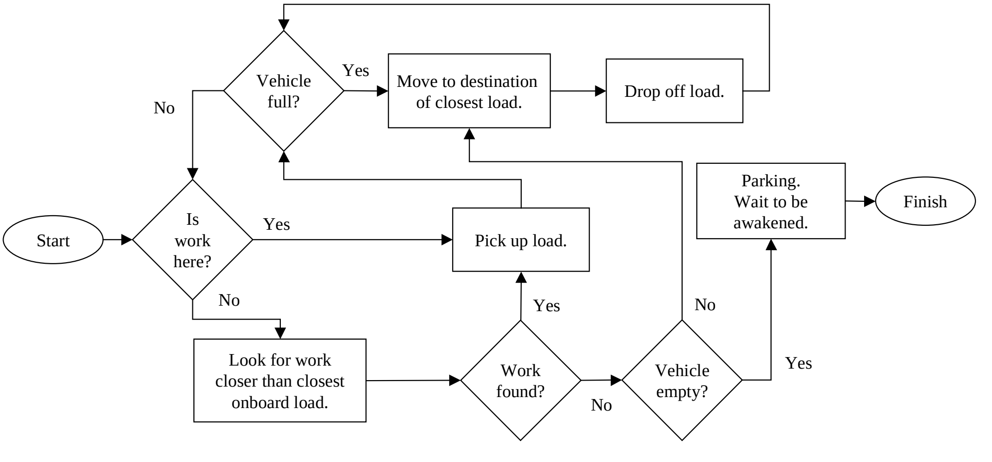

Figure 12. Vehicle dispatching behavior

The concept of closest task for multi-vehicle dispatching applies to the previously described NVF and FCFS dispatching rules. The flowchart of Figure 12 shows the decisions made during vehicle-initiated dispatching. When a vehicle drops off a load, the vehicle continues by checking for (additional) work. When vehicles are parked when a load is released in the system, the (idle) vehicles are awoken which then check for (additional) work. Figure 13 shows how vehicle behavior is affected with load-initiated rules. If a load is released and no (idle) vehicle is found with remaining capacity, the dispatching initiative is passed to the vehicles.

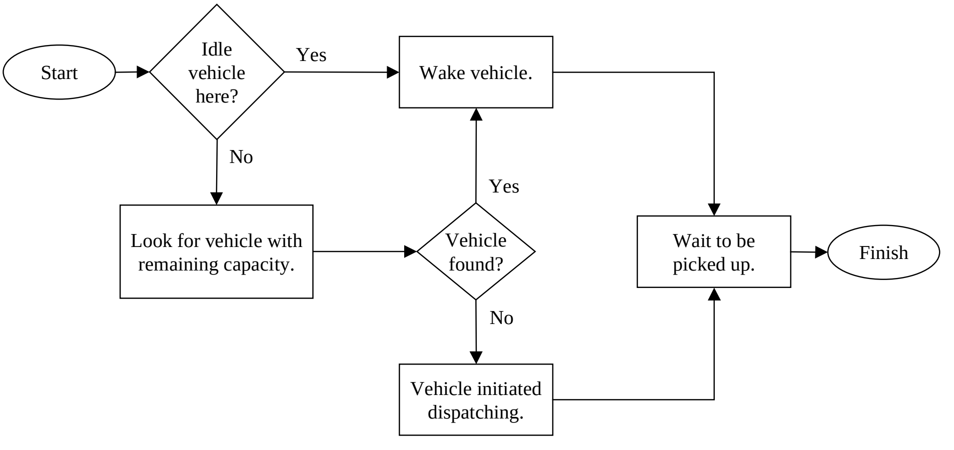

Figure 13. How loads affect vehicle behavior

So the performance of the NVF and modified FCFS rules is mainly characterized by the dispatching rule triggered by the first onboard load when multi-load vehicles are used. In this study the capacity of the vehicles is at most two; i.e. dual-load vehicles.

3.2 The U-layout and I-layout environments

Figure 14 gives a representation of an I-layout and U-layout transportation environment respectively; two common warehouse layouts found in practice. The dashed lines represent the contours of the building. The solid lines represent the network on which the guided vehicles travel. The vehicles are stored in the vehicle depot and also start and end their daily tasks at the depot. The other nodes on the vehicle path represent different locations (the origins and destinations of loads) which the vehicles visit to serve transportation requests. The numbers beside the paths represent the distance units between the nodes when that path is followed. These numbers can also be seen as time units since the vehicles travel with constant unit speed, i.e. one distance unit per time unit.

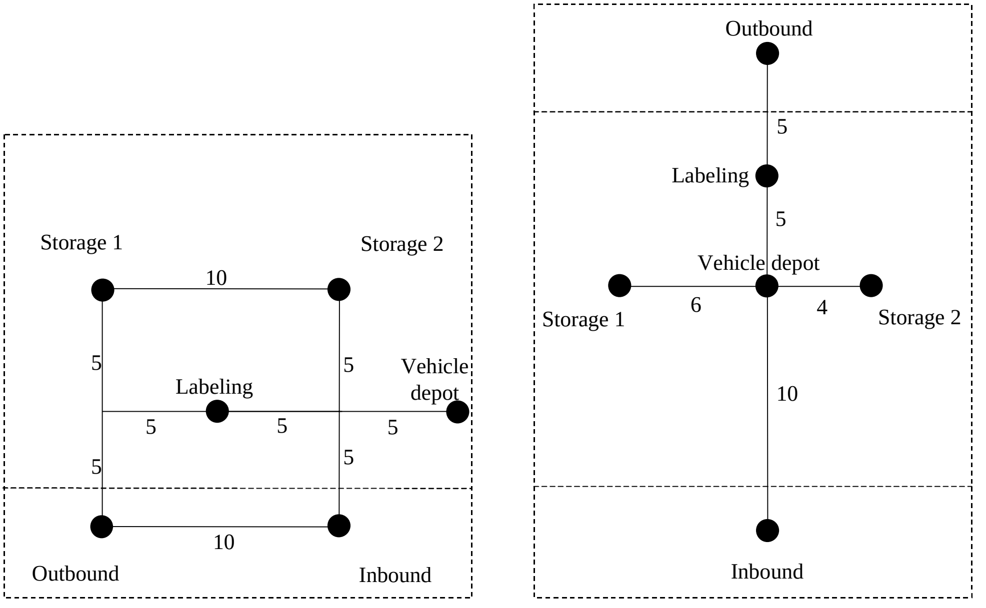

Figure 14. Representation of the U-layout (left) and I-layout (right) warehouses

The design of the facility is mainly dependent on the nature of activities being performed inside the facility and the access to outside transportation facilities, see Tompkins et al. (1996). If both receiving and shipping occur simultaneously, then close supervision is required to ensure that received goods and goods to be shipped are not mixed.

If storage is one of the main functions of the warehouse, then both the receiving (Inbound) and shipping (Outbound) lanes are usually at one side of the building. The result is a so-called U-layout warehouse with a rectangular shape (see Figure 14). In this way it is possible to partly utilize the same docks, personnel and handling equipment for shipping and receiving operations. The storage modules are at the other side of the building and the stations with for example, added value logistics (VAL), in this case a labeling station, in the middle of the warehouse.

The I-layout warehouse (see Figure 14) is an example warehouse commonly used in situations where transshipment is the most important process and storage is less important. Loads are received at one end and leave at the other end. Hence, the receiving (Inbound) and shipping (Outbound) lanes are at opposite ends of the warehouse, and all other stations more or less in the middle.

An extra advantage of the U-layout is the greater possibility for double-plays (combining inbound trips with outbound trips) since the Outbound and Inbound areas are relatively closer to each other. This means that vehicles may be better utilized since empty travel times decrease (and load waiting times possibly decrease).

In this case, the advantage of the I-layout is the smaller transport distances for stored material from Storage 1 and 2 towards the Outbound lanes. The disadvantage is the greater distance between the Inbound and Outbound areas, which slightly increases the average distance between any location to any other location. Observe that the I-layout is less symmetric in distances than the U-layout. This has been done on purpose, in order to investigate whether symmetry has an effect on the performance of certain dispatching rules. Intuitively, one can imagine that a distance-based dispatching rule works better if there are differences in the travel distances, like in a non-symmetrical environment. The vehicle paths for both warehouses are bi-directional and vehicles may pass each other if necessary. The pick up and set down times of loads are negligible and idle vehicles park at their current location.

In the case of the example warehouses, Inbound loads arrive at the Inbound area and are transported to Storage 1 or Storage 2. At Storage 1 and Storage 2, loads that need to be transported are sent to the Labeling area. From the Labeling area Outbound loads are transported to the Outbound area. In both U-layout and I-layout situations, the average inbound travel time is the same. For the U-layout, the travel time is either 10 or 20 units; this means 15 units on average. For the I-layout, the travel time is either 16 or 14 units; this also means 15 units on average. The Outbound loads first go through the Labeling area. In the U-layout this means that those loads always travel 20 units. As mentioned before, the Outbound loads of the I-layout have a travel time advantage. In this case the Outbound loads travel 15 units on average.
So there are three classes with a total of 5 job types:

Class 1: Inbound
1) Inbound to Storage 1 (travel time: 20 time units for U-layout and 16 for I-layout);
2) Inbound to Storage 2 (travel time: 10 time units for U-layout and 14 for I-layout);

Class 2: Labeling
3) Storage 1 to Labeling (travel time: 10 time units for U-layout and 11 for I-layout);
4) Storage 2 to Labeling (travel time: 10 time units for U-layout and 9 for I-layout);

Class 3: Outbound
5) Labeling to Outbound (travel time: 10 time units for U-layout and 5 for I-layout).

In general, two vehicles are used for the transportation jobs. However, in case of on-line dispatching, more vehicles are needed to transport all loads in the given time period. The jobs are generated such that three different daily shifts are constructed with different throughput characteristics, as described in the following sections.

3.3 Random shifts

To keep the MIP problem computationally tractable, the number of jobs could not exceed 12 (see also Section 3.6.1). Since there are three classes, the idea is to generate four jobs of each class on average. Using a uniform distribution, job types are generated at random (where the Outbound job type is weighed twice). It is then easy to see that the (daily) shift of 12 jobs on the U-layout has a total loaded trip time of 140 time units (see Table 5) on average, or 70 time units per vehicle on average. When the empty trip time is estimated at 80% of the loaded trip time, or 56 time units per vehicle, the total trip time will be 126 time units.

	U-layout	I-layout
Job type	Loaded travel time units	Loaded travel time units
Inbound	$2\times20 + 2\times10$	$2\times16 + 2\times14$
Labeling	4×10	$2\times11 + 2\times9$
Outbound	4×10	4×5
Total	140	120

Table 5. Average total loaded travel time units per layout

Although similar calculations for the I-layout result in an average total trip time of 108 time units per vehicle (see Table 5), the same data (transport jobs and load release times) generated with the calculations of the U-layout is used for both layouts. Jobs for both layouts are therefore generated between 0 to 126 time units (the daily shift) from a uniform distribution. These jobs are then assigned to the vehicles according to the dispatching rule used. Observe that one vehicle can have more job assignments than another. The average load throughput times are calculated over a total of 10 different generated shifts.

3.4 Structured shifts

In the case of Random shifts, the 12 jobs are uniformly generated over a period of 126 time units. With Structured shifts, one-third of the jobs are uniformly generated over the first 40% of the total shift time and consist (only) of Inbound jobs. The last 40% of the shift consist of Outbound jobs (also one-third of the total amount of jobs), and the middle 40% (so there is an overlap of 10% on each side) consists Labeling jobs, see also Figure 15.

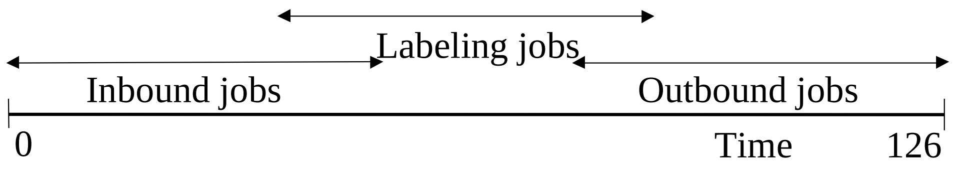

Figure 15. Structured dispersion of jobs over the shift

The Structured shifts comes from the idea that in many warehouse and manufacturing situations; inbound jobs precede processing jobs, which in turn precede outbound jobs during a day.

3.5 High throughput shifts

In order to investigate the dependency of the performance gap between on-line dispatching and off-line control on throughput and vehicle utilization, we increased the throughput level. In many environments, peaks of high workloads can be observed. When the jobs are generated in a shorter time period, the probability of on-line dispatched vehicles *waiting* for a transport assignment is reduced, similar to peak behavior. So by reducing vehicle idle times, the extra load waiting times will be reduced and we can expect that the performance of on-line dispatching and off-line control will be closer together. This is the main idea of High throughput shifts.
The jobs are uniformly generated in a structured shift as shown in Figure 15. The number of jobs, however, is increased to 60 (20 Inbound, 20 Labeling and 20 Outbound). In this case, the jobs are generated over a period equal to the average loaded trip time of the U-layout (so 10*20 + 50*10 = 700 time units) plus an extra 20% to account for (some) of the empty trip time; in total 420 time units per vehicle. Again, although similar calculations for the I-layout result in an average total trip time of 360 time units per vehicle, the data generated with the calculations of the U-layout is used for both layouts. The Insertion heuristic is used as the off-line control rule since exact off-line control appeared to be intractable for this situation.

3.6 Results

The discussion of the results will start by presenting the performance gap, i.e. the differences in expected load throughput times, between off-line and on-line controlled uni-load guided vehicles. It will also be shown to what extent the on-line controlled vehicle fleet has to increase to approximate off-line performance. Next, we present the effects of using dual-load vehicles with on-line control and compare the results with increasing the fleet size with uni-load and dual-load vehicles.

3.6.1 Varying the number of vehicles

Table 6 gives an overview of the average throughput times (see ‘Average’) and standard deviation (see ‘St. dev.’) of the throughput times for 10 runs for both off-line control and on-line dispatching for both layout types in the Random shifts situation. For the off-line rules, ‘Optimal’ refers to the optimal solution from the TRPTW, i.e. the minimum load throughput times possible, when two guided vehicles (GVs) are used. In case of 12 jobs, the computation times for the Optimal solution solving the TRPTW with CPLEX, varied

from two minutes to several hours on an IBM/RS6000 class computer, and in some instances up to 120 Mb of memory was required for the branch and bound tree. The throughput times in the 'Insertion' column represent the throughput times obtained when two vehicles are routed with the Insertion heuristic (less than one second computation time). This leads to the optimal solution in several instances; overall it deviates about 4% and 5% from the optimal value (see 'Deviation' in Table 6) for the U-layout and I-layout respectively. When on-line dispatching rules are used, the loads have to wait about twice as long (see the waiting time results in the next chapter) to be transported with the same number of vehicles and the average throughput time is more than 50% higher in the I-layout case. Notice that the average throughput times are smaller in the I-layout environment. This is due to the travel time advantage for Outbound loads.
To bring the average throughput times with on-line dispatching within 10% of the optimal solution, the fleet size had to be doubled to four vehicles.

	Off-line control		On-line control: NVF			On-line control: FCFS		
	Optimal	Insertion	2 GVs	3 GVs	4 GVs	2 GVs	3 GVs	4 GVs
U-layout								
Average	235.2	244.1	328.2	260.7	245.1	321.5	261.4	248.0
St. dev.	20.7	22.9	41.5	24.4	22.9	47.5	26.4	18.3
Deviation	-	4 %	40 %	11 %	4 %	37 %	11 %	5 %
I-layout								
Average	195.0	204.6	300.7	235.2	210.3	304.4	230.5	219.6
St. dev	44.6	45.1	44.3	33.7	27.0	38.2	45.5	26.9
Deviation	-	5 %	54 %	21 %	8 %	56 %	18 %	13 %

Table 6. Average load throughput times with Random shifts (10 runs consisting of 12 jobs)

Even when the fleet size is quadrupled (not in table), the optimal off-line rule outperforms the on-line rules on average. This is due to the fact that off-line controlled vehicles can use idle time to move closer to the next task, hence reducing load waiting time. In the Random shifts case, the vehicle idle time with on-line control appears to be about 15% (not shown in the table), this means that there is some slack in the system which could be used to route the vehicles to the next assignment and reduce the load waiting times. This will be demonstrated in the next chapter.

	Off-line control		On-line control: NVF			On-line control: FCFS		
	Optimal	Insertion	2 GVs	3 GVs	4 GVs	2 GVs	3 GVs	4 GVs
U-layout								
Average	215.6	218.5	274.8	221.8	211.0	277.0	214.6	193.0
St. dev.	12.8	12.8	20.3	19.6	15.6	19.2	16.5	11.0
Deviation	-	1 %	27 %	3 %	-2 %	28 %	0 %	-10 %
I-layout								
Average	198.4	201.1	261.7	197.9	173.1	262.8	197.8	175.3
St. dev.	23.6	25.2	19.9	17.4	10.8	23.6	16.3	13.7
Deviation	-	1 %	32 %	0 %	-13 %	32 %	0 %	-12 %

Table 7. Average load throughput times with Structured shifts (10 runs consisting of 12 jobs)

Next, the case of Structured shifts is studied. Because of the design of the warehouse and the overlapping periods within the shifts, a combination of dropping off an Inbound load and picking up a load for Labeling or the combination Labeling and Outbound loads, leads to smaller waiting times and thereby smaller load throughput times with the uni-load vehicles. This can be seen by comparing the values of Table 6 with those of Table 7.
For two vehicles, the average load throughput times with online dispatching are about 30% higher than the optimum. The difference between Insertion and the Optimal value is about 1% and the fleet only needs 50% extra vehicles instead of twice as many to approximate Off-line performance within 10%. We can conclude that Structured shifts leads to a better performance for both off-line and on-line control than when the jobs are Random in a shift. Although the differences in the average load throughput times between NVF and FCFS are small, it seems that NVF is a little more favorable in the I-layout environment and FCFS in the U-layout environment (in both cases two out of three times on average, see Table 7). Since the U-layout is rather symmetrical in travel times, the dispatching decisions with a time-based rule turn out to be more favorable. In the less symmetrical I-layout, the dispatching decisions can be made based on different travel distances and the distance-based rule turns out to be more favorable.

In the next experiment, the transport request intensity is increased. This means that there will be less idle time for the vehicles, which will reduce the performance gap between off-line control and on-line dispatching. A total of 60 loads are generated in a time frame that has a length of 1.2 times the load transport time. This will be done in a similar manner, as was the case for 12 loads. The extra 20% is added to account for (some) empty trip time.
We continued the study without calculating the Optimal performance with TRPTW since this led to high running times and computer memory problems. Since the Insertion heuristic leads to very satisfactory results (see the previous part of this section) in a very simple and quick way, we will continue to use Insertion for the off-line control.

	Off-line control: Insertion		On-line control: NVF		On-line control: FCFS	
	2 GVs	3 GVs	2 GVs	3 GVs	2 GVs	3 GVs
U-layout						
Average	6422.8	2006.3	6907.3	2355.1	6805.0	2417.8
St. dev.	629.0	361.3	700.0	382.4	582.0	412.1
Deviation	-	-	8 %	17 %	6 %	21 %
I-layout						
Average	6073.2	2095.2	6518.7	2592.3	7265.9	2586.6
St. dev.	730.0	847.8	710.2	834.1	757.3	791.3
Deviation	-	-	7 %	24 %	20 %	23 %

Table 8. Average load throughput times with High throughput shifts (10 runs with 60 jobs)

The results in Table 8 show that the deviation in load throughput times between on-line dispatching and off-line control is smaller when extra waiting time is eliminated by removing the slack in the system. For the 2-vehicle situation, the load throughput times with on-line dispatching are about 8% and 7% higher, and 17% and 24% for the 3-vehicle situation for the U-layout and I-layout respectively. Considering that the load throughput times with Insertion were up to 5% above the optimum (see the results for Random shifts), we expect that the deviation with the optimal performance is still reasonable.

3.6.2 Varying the capacity of vehicles

The alternative to increasing the number of vehicles is to increase the vehicle capacity. In general, two uni-load vehicles are more expensive than one dual-load vehicle, while the number of loads that can be transported simultaneously is the same.
With doubling the fleet size as in the previous section, we risk congestion, etc. Instead, we would like to see the effects of doubling the vehicle capacity, although the control of dual-load vehicles is more complex than the control of uni-load vehicles as explained earlier. The dual-load vehicles are used with NVF and FCFS dispatching only and are also compared with the performance (average load throughput times) of off-line controlled (Insertion) uni-load vehicles. From the results in Table 9, we see that increasing the vehicle capacity leads to a limited increase in performance and diminishes as the fleet size increases. The performance of two dual-load vehicles, (when four loads can be transported simultaneously), is worse than the performance of three uni-load vehicles. In fact, four uni-load vehicles (see Table 6) outperform three dual-load vehicles (see Table 9).

	Off-line control		On-line control: NVF				On-line control: FCFS			
	Insertion: 2 GVs	Insertion: 3 GVs	2 GVs cap. 1	2 GVs cap. 2	3 GVs cap. 1	3 GVs cap. 2	2 GVs cap. 1	2 GVs cap. 2	3 GVs cap. 1	3 GVs cap. 2
U-layout										
Average	244.1	176.1	328.2	300.7	260.7	255.0	321.5	298.2	261.4	256.8
St. dev.	22.9	13.9	41.5	38.6	24.4	24.3	47.5	36.5	26.4	29.2
Deviation	-	-	34 %	23 %	48 %	45 %	32 %	22 %	48 %	46 %
I-layout										
Average	204.6	141.1	300.7	272.1	235.2	233.7	304.4	278.6	230.5	240.0
St. dev.	45.1	15.4	44.3	43.7	33.7	30.2	38.2	43.4	45.5	33.1
Deviation	-	-	47 %	33 %	67 %	66 %	49 %	36 %	63 %	70 %

Table 9. Average load throughput times for Random shifts (12 jobs)

In Table 9 we can also see the negative effects of dual-load vehicles on the load throughput time. Although the load waiting time can reduce when dual-load vehicles are used, the load transportation time can increase. Load transportation times can increase since loads do not have to be delivered immediately after being picked up. Certain loads can remain on the vehicle while other loads are serviced with the remaining vehicle capacity. The result is that the sum of the load transportation time and load waiting time (i.e. defined as the load throughput time), can then also increase. This phenomenon can be seen in the I-layout environment when three vehicles are dispatched with the FCFS rule. When the vehicle capacity increases, the average load throughput time increases from 230.5 time units to 240 time units.

Since transportation jobs are more structured in the Structured shifts, the opportunity for combining transportation jobs with dual-load vehicles increases. It can be seen in Table 10 that the deviations are more favorable compared to Random shifts in Table 9. The load throughput time can still increase as can be seen in the I-layout environment when three vehicles are dispatched with the FCFS rule, but to a less extent compared to Random shifts.

	Off-line control		On-line control: NVF				On-line control: FCFS			
	Insertion: 2 GVs	Insertion: 3 GVs	2 GVs cap. 1	2 GVs cap. 2	3 GVs cap. 1	3 GVs cap. 2	2 GVs cap. 1	2 GVs cap. 2	3 GVs cap. 1	3 GVs cap. 2
U-layout										
Average	218.5	164.1	274.8	256.9	221.8	219.2	277.0	256.9	214.6	210.3
St. dev.	12.8	10.8	20.3	21.9	19.6	19.9	19.2	21.9	16.5	13.1
Deviation	-	-	26 %	18 %	35 %	34 %	27 %	18 %	31 %	28 %
I-layout										
Average	201.1	150.1	261.7	245.1	197.9	197.1	262.8	244.1	197.6	197.7
St. dev.	25.2	8.9	19.9	14.4	17.4	18.5	23.6	14.2	16.7	18.5
Deviation	-	-	30 %	22 %	32 %	31 %	31 %	21 %	32 %	32 %

Table 10. Average load throughput times for Structured shifts (12 jobs)

We also see that increasing the vehicle capacity leads to a smaller decrease in average throughput times compared to Random shifts, and the decrease in throughput times diminishes as the fleet size increases. In fact, using dual-load vehicles with an increased fleet size does not lead to a significant increase in performance.

	Off-line control		On-line control: NVF				On-line control: FCFS			
	Insertion: 2 GVs	Insertion: 3 GVs	2 GVs cap. 1	2 GVs cap. 2	3 GVs cap. 1	3 GVs cap. 2	2 GVs cap. 1	2 GVs cap. 2	3 GVs cap. 1	3 GVs cap. 2
U-layout										
Average	6422.8	2006.3	6907.3	2012.6	2355.1	1477.5	6805.0	2052.4	2417.8	1470.8
St. dev.	689.0	361.3	700.0	196.4	382.4	96.1	582.0	237.0	412.1	103.2
Deviation	-	-	8 %	-69 %	17 %	-26 %	6 %	-68 %	21 %	-27 %
I-layout										
Average	6073.2	2095.2	6518.7	1998.8	2592.3	1328.3	7265.9	2033.5	2586.6	1334.0
St. dev.	730.0	847.8	710.2	278.5	834.1	88.5	757.3	368.4	791.3	109.8
Deviation	-	-	7 %	-67 %	24 %	-37 %	20 %	-67 %	23 %	-36 %

Table 11. Average load throughput times for High throughput shifts (60 jobs)

From Table 11 it is clear that using two dual-load vehicles leads to smaller average load throughput times than three uni-load vehicles in the High throughput case (this was the reverse for Random and Structured shifts). Apparently, adding vehicle capacity in an environment with high vehicle utilization has a greater impact on the performance than adding capacity in environments with low vehicle utilization. In this case, the performance still significantly improves when three uni-load vehicles are replaced with three dual-load vehicles. Although the differences in average load throughput times for on-line control are rather small, the phenomenon that FCFS generally leads to smaller average load throughput times in a symmetrical layout compared to a less symmetrical layout, and NVF generally leads to smaller average load throughput times in a less symmetrical layout compared to a symmetrical layout seems to occur for the dual-load vehicles case as well.

3.7 Concluding remarks

In this chapter we compared the average load throughput times of several off-line control and on-line dispatching rules for guided vehicles used for internal transport. Using off-line control means that all information on load release times, origins and destinations has to be known in advance. This is not a real situation found in practice due to the stochastic nature of internal transportation environments. However, for theoretical purposes we assumed that all information is available when off-line control rules are used. The performance is defined as the average load throughput time; the time needed to serve a transport request from the moment a load is (physically) released to the system and ready for transport until it is dropped off at its destination, (i.e., the load waiting time plus the load transportation time).
Our goal was to study the performance gap (difference in average load throughput times) between on-line control and off-line dispatching and to investigate how this gap is affected when the fleet size is increased and when the on-line vehicle capacity is increased -in combination with increasing the fleet size, for different dispatching rules in different layout environments.

- The results show that for different studied layouts and shifts, considerable gains on performance (reductions in average load throughput times) are possible with off-line algorithms (exact and heuristics) if the system is relatively quiet, i.e. dispatch requests are spread out evenly (low throughput) and vehicles have relatively high idle times (in this case about 15% or more). This is due to reductions in load waiting time by already traveling to a load before it has been physically released. Therefore the load can be picked up relatively sooner, which leads to a reduction in average load waiting times and in most cases the average load throughput times. In low throughput systems, we see that the fleet size has to increase by 50% or more to obtain similar results to the Optimal routing. However, in systems with high throughput, and therefore a smaller opportunity to reduce load waiting time, the performance of on-line control is already satisfactory (in our case differences of 6-20%).

- Table 12 summarizes the results for adding extra vehicle capacity when two or three vehicles are used. (The value between brackets represents the performance of *three* off-line controlled uni-load vehicles relative to *two* off-line controlled uni-load vehicles). It is clear that heavily utilized GVs benefit most from adding vehicle capacity. The benefits decrease as the fleet size increases.

Control form: vehicle types	Performance deviation by heavily utilized GVs (High throughput shifts)	Performance deviation by GVs with idle time (Structured shifts)
U-layout	NVF / FCFS	NVF / FCFS
Off-line: 2 uni-load GVs	-	-
On-line: 2 uni-load GVs	8 % / 6 %	26 % / 27 %
On-line: 2 dual-load GVs	-69 % / -68 %	18 % / 18 %
Off-line: 3 uni-load GVs	(-69 %)	(-25 %)
On-line: 3 uni-load GVs	17 % / 21 %	35 % / 31 %
On-line: 3 dual-load GVs	-26 % / -27 %	34 % / 28 %
I-layout	NVF / FCFS	NVF / FCFS
Off-line: 2 uni-load GVs	-	-
On-line: 2 uni-load GVs	7 % / 20 %	30 % / 31 %
On-line: 2 dual-load GVs	-67 % / -67 %	22 % / 21 %
Off-line: 3 uni-load GVs	(-65 %)	(-25 %)
On-line: 3 uni-load GVs	24 % / 23 %	32 % / 32 %
On-line: 3 dual-load GVs	-37 % / -36 %	31 % / 32 %

Table 12. Average load throughput time deviations between off-line control (Insertion) and on-line dispatching for changes in fleet capacity

In Table 13, we can also see the effects on on-line performance when the fleet capacity increases, compared with two GVs controlled off-line. In the case of high throughput environments, the difference between off-line control and on-line dispatching performance is in the standard situation already almost negligible (8% or less, as can be seen the second column of Table 13). In fact, adding vehicles to the fleet or adding capacity to the vehicles improves the performance beyond the off-line (standard) situation.

Situation	Performance deviation by heavily utilized GVs (High throughput shifts)	Performance deviation by GVs with idle time (Structured shifts)
U-layout	NVF / FCFS	NVF / FCFS
Standard (2 GVs)	8 % / 6 %	26 % / 27 %
50 % extra GVs	-63 % / -62 %	2 % / -2 %
Dual-load GVs	-69 % / -68 %	18 % / 18 %
50 % extra + dual-load GVs	-77 % / -77 %	0.3 % / -4 %
I-layout	NVF / FCFS	NVF / FCFS
Standard (2 GVs)	7 % / 20 %	30 % / 31 %
50 % extra GVs	-57 % / -57 %	-2 % / -2 %
Dual-load GVs	-67 % / -67 %	22 % / 21 %
50 % extra + dual-load GVs	-78 % / -78 %	-2 % / -2 %

Table 13. Average load throughput time deviations for several situations of on-line dispatching relative to off-line control (Insertion) with two uni-load vehicles

- Increasing the fleet size by 50% for the Structured shifts with vehicle idle time (relatively low throughput environment), leads to a performance deviation of 2% for NVF and -2% for FCFS from the off-line heuristic. However, this is still better than doubling the vehicle capacity, which leads to a performance deviation of 18%. The effect of increasing the number of vehicles in a low throughput environment is so dominant that combining the effects of extra vehicles plus extra vehicle capacity does not lead to a combined increase in performance. Table 13 reveals which steps could be taken to close the performance gap between off-line control and on-line dispatching for a certain environment.

- Careful study also reveals that the NVF rule seems to perform more often more favorable in a non-symmetric layout environment and FCFS more favorable in a symmetric environment. This seems logical since decisions based on symmetric distances with the NVF rule are similar to random load-to-vehicle assignments. Furthermore, the standard deviations of the average load throughput times are in general higher for on-line control compared to off-line control and decreases as the fleet capacity increases. This is due to the phenomenon that when the fleet capacity is relatively greater during peak periods of load releases, the maximum load waiting times decrease.

In the next chapter we will investigate the effects of using some of the vehicle idle time in on-line vehicle dispatching situations, to travel to the next transport request just before the load is physically released. Intuitively, traveling to the load before it is released should reduce the load waiting time. However, the load release data should be correct. If the actual release time deviates form the expected release time, the vehicle-to-load allocation can become unfavorable. The latter will also be studied in the next chapter.

Chapter 4

Approximating off-line vehicle control with on-line dispatching rules

In this chapter, we proceed with the assumption that some load information is available a short moment before the actual physical release of the loads with on-line dispatching. Usually, the vehicles are dispatched on-line, because information on load origins, release times and destinations is only available at the last moment. In some instances, such load information can be given a short moment in advance (this can be a stochastic time period), i.e. a conveyor or a crane carrying a load that will be dropped off for pick-up in a few moments or a pallet wrapping station that will be ready with the load in a few seconds. If load pre-arrival information is available (information about the load before the load has arrived at the location where it can be picked up), the load can already send out a release signal in order to claim transport. The extra time between this virtual release time and the actual release time can be used to schedule the vehicles more favorably compared to ordinary on-line control rules without the use of pre-arrival information (Section 4.2.1).
When vehicles are instructed to park near a location that is most likely to place the next transportation request, the response time to pick up the loads can be reduced. We therefore also study in Section 4.2.3 the performance effects when a vehicle dwell point strategy is used which instructs the vehicles to park near the location of the next expected transport request. Furthermore, we also assume in Section 4.2.4 that the locations of moving vehicles are known at any time by the central controller and that this information can be used to pre-assign idle and moving vehicles to loads requesting transport. In this case the controller calculates the distance that vehicles must travel to reach the requesting pick up. The vehicle with the shortest distance to travel gets the assignment, taking into account delivery of any on-board load or retrieval and delivery of a previously assigned load to the vehicle.

This chapter is an extension of the previous chapter with the restriction that only uni-load vehicles are considered. This means, since travel times are fixed, that the throughput times can only be influenced by load waiting times. The performance objective in this chapter is therefore to minimize the average load waiting times. In this chapter, the average load waiting times of optimal vehicle schedules are compared with the average load waiting times obtained with the on-line NVF and FCFS dispatching rules and a few variations of these rules, which will be described later. The idea is still to use off-line control

performance as a benchmark for on-line performance and to evaluate the value of some extra information with respect to load release times (pre-arrival information), using dwell point strategies and pre-assigning (moving) vehicles to loads.
Since release times are not known in advance in reality, we also investigate the impact of slight disturbances in the release times in Section 4.2.6. To do this, we deviate the load release times for some shifts, while the number of jobs and job types remain the same. In that case, the load release times are perturbed and the loads can be released relatively earlier or later than before the perturbation. The new shifts are then served with the original 'old' route when off-line control is used. It is possible that vehicles arrive relatively too early to pick up the load and have to wait. With on-line dispatching, the situation remains real-time and the vehicles can seize the opportunity to pick up the job that was released relatively sooner. The deviation in performance between on-line dispatching and off-line control should therefore decrease. In fact, it is even possible that the performance of off-line control becomes worse than the performance with on-line control since the off-line controlled vehicles cannot seize possible opportunities to change their route when deviations are encountered.

Using the same simplified U-layout and I-layout warehouse models of the previous chapter, we will show that, although the performance gap between off-line and on-line control can be considerable, using pre-arrival information, dwell points and assigning moving vehicles to loads can reduce the average load waiting times, (but only to a certain extent). Furthermore, the performance gap also decreases as the actual release times deviate slightly from those that were used to calculate the vehicle routes off-line.
The following section briefly describes the models and control rules used for this chapter. The results will then be discussed in detail, followed by concluding remarks. We refer to the previous chapter for a detailed description of the off-line control rules, the description of the layouts and how the loads are generated.

4.1 Model and control rules

Although off-line vehicle control can minimize the average load waiting times, the exact time a load is released (necessary in warehousing and manufacturing practice) is usually not known far in advance. In this chapter we use the same on-line control rules described in the previous chapter, by which decisions of control are triggered at the release time of the loads (the moment they can be transported) or a few moments earlier in case pre-arrival information is available. The first rule is nearest-vehicle-first (NVF), a load-initiated distance-based rule, the second rule is modified first-come-first-served (FCFS), a vehicle-initiated time-based rule. We will also study several new dispatching rules that make use of the NVF and FCFS rules. In the following sections we will describe the model and control rules in more detail.

Dispatching with pre-arrival information

When pre-arrival information of the load is available, the release signal for transport is given a few moments before the load is physically ready for transport. This virtual load release signal varies from 0 to 20 time units (the maximum travel time between two locations in both the U- and I-layout) before the physical release time.
In each case, two uni-load vehicles are routed on the same U-layout and I-layout environment and with the same NVF and modified FCFS dispatching rules described in Chapter 3.
Initially, for vehicle dispatching with pre-arrival information, the jobs are generated in three different ways:

1. Random shifts: 12 jobs are uniformly generated over a shift 126 time units. The total length of the shift is 126 time units.
2. Structured shifts: 12 jobs are generated over a shift of 1.8 times the loaded travel time. The Inbound jobs are uniformly generated in the beginning of the shift; the Labeling jobs are uniformly and independently generated in the middle of the shift and the Outbound jobs are uniformly and independently generated at the end of the shift.
3. High throughput shifts: 60 jobs are generated the same way as structured shifts over a shift of 1.2 times the loaded travel time. The total length of the shift is 420 time units.

When the pre-arrival time increases, it becomes more likely that the vehicles arrive before the physical release of the loads. In that case, the vehicles have to wait until the loads are ready for transport. This vehicle waiting time could be used to serve other requests and result in a more favorable vehicle to job allocation. To investigate this, we also study in Section 4.2.2 a case where jobs can be suspended when the vehicle waiting time for an expected load at a particular location is higher than the average vehicle travel time to another pick up location.

Dispatching with vehicle dwell points

In the case that vehicles are dispatched using an alternative dwell point strategy, the vehicles are sent to park at the station where the next transport request is most likely to be released. The next request is most obvious in case that the shifts are structured. In that case, a '*Structured dwell point strategy*' is used where idle vehicles are sent to park at Receiving the beginning of the shifts, sent at random to Storage 1 or Storage 2 (unless they become idle at one of these locations) in the middle of the shifts and sent to Labeling at the end of the shifts. Because the Inbound and Labeling jobs, and the Labeling and Outbound jobs partly overlap (see Figure 15), vehicles which become idle during these overlap times are sent to one of the locations (Receiving, Storage 1 or Storage 2 for the first overlap and Storage 1, Storage 2 or Labeling for the second overlap) for those jobs at random (unless they are already at one of those locations).

We will also use a dwell point strategy when vehicles are dispatched with Random shifts in order to investigate the effects of sending vehicles to park when the shifts are not structured. With Random shifts, the likelihood for the next transportation requests is the same for each pick up location. We therefore determine a dwell point location (a separate one for both layouts) on the vehicle network that has the smallest average travel time to any pick up location. More formally, we select the dwell point which minimizes the average (empty) travel distance/time from the dwell point to the pick up points, where the travel times to the pick up points are weighed with the number of expected pick up jobs at those pick up points. Table 14 shows the average travel times from the best dwell points to the pick up locations. This means that with Random shifts, Labeling is used as the dwell point for the U-layout, and the Depot is used as the dwell point for the I-layout. Any other dwell point results in a higher average travel time, ranging from $8\frac{1}{3}$ (for Receiving and Storage 2 on the U-layout) to $13\frac{1}{3}$ (for Shipping on the I-layout) time units, which may increase the average response time of the vehicle to the load which in turn increases the average waiting time.

	U-layout	I-layout
Pick up location : weight	Travel time from Labeling	Travel time from Depot
Receiving : 4	10	10
Storage 1 : 2	10	6
Storage 2 : 2	10	4
Labeling : 4	0	5
Average travel time	$6\frac{2}{3}$	$6\frac{2}{3}$

Table 14. Average travel times from dwell points to pick up points for the U-layout and I-layout with Random shifts

If a vehicle moving to a dwell point must reach that point before becoming eligible to pick up a load, it may pass by waiting loads and thereby possibly decreasing the performance of the system. Therefore, the closest parked (idle) and ‘going to park’ vehicles can be matched to loads. This is done in order to prevent loads being matched to vehicles that are parked or which have just dropped off a load at a distant location when other empty vehicles traveling to dwell points are closer. The distances from the loads to the vehicles are defined as the distances that the parked or going to park vehicles still have to travel before they can reach the loads placing the requests. When no parked or going to park vehicle is available, the load will eventually be matched to a vehicle using the standard NVF or FCFS rule (described in the previous chapter). The flowchart of Figure 16 shows how these vehicle-to-load assignments are made when the dwell point strategies are used. Since a dwell point strategy can only work when vehicles need to park, we cannot study a dwell point strategy for the High throughput shifts.

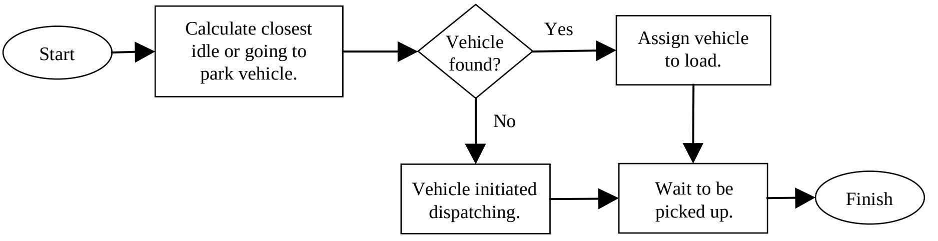

Figure 16. How loads affect vehicle behavior with dwell point strategies

Pre-assigning moving vehicles to loads

When loads can be assigned to the closest parked or any moving vehicles, the distances from the loads to the vehicles are defined as the distances which the moving or parked vehicles still have to travel before they can reach the loads placing the requests. The calculations for the vehicle-to-load assignments are triggered as soon as a load is released to the transport system, (see Figure 17).

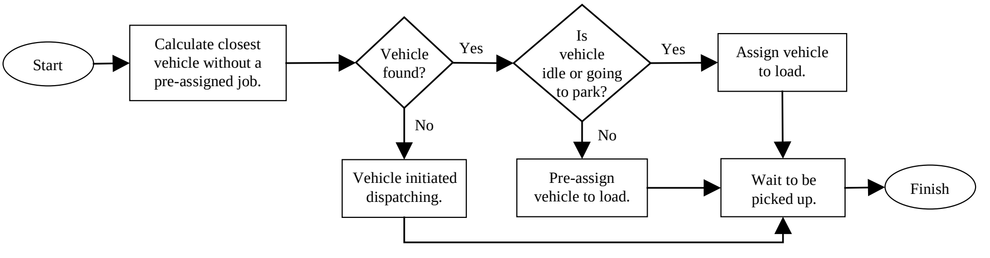

Figure 17. How loads affect vehicle behavior with the pre-assigning moving vehicles to loads strategy

Vehicles already delivering a load can be pre-assigned one extra pick up request, but must deliver their current load before traveling to the new pick up point. Similarly, vehicles moving to retrieve a load can be pre-assigned a second request, which is carried out after the currently assigned load has been retrieved and delivered. The vehicle's travel distance is therefore calculated based on its status. When the vehicle is 'parked' or 'going to park', the distance to the load is defined as the distance from the vehicle's current location to the load's pick up point. When the vehicle is 'retrieving' a load, the distance to the load is defined as the distance from the vehicle's current location to the first load's pick up point plus the distance to the first load's destination plus the distance to the new load's pick up point. When the vehicle is 'delivering' a load, the distance to the load is defined as the distance to the vehicle's on-board load's destination plus the distance to the new load's pick up point. The vehicles can have no more than two loads assigned to it, and once a vehicle is assigned to a load, it cannot be reassigned. When no vehicle is pre-assigned to a load entering the system and requesting to be picked up (for example when all vehicles are assigned two jobs), the load will eventually be matched to a vehicle using the (standard) NVF or FCFS rule.

Dispatching with perturbed release times

Since we also investigate the case where the exact data on release times needed for off-line control is not available, we take one of the instances of the Random shifts (Run 1), and create 10 more shifts by deviating the release times uniformly with 5% of the shifts length. (In this case, the release times can decrease or increase up to 5% of 126 time units, the length of the shift, which means that the loads can be released up to 6 time units earlier or later.) The 10 new shifts are then served with the original ‘old’ route when off-line Insertion control is used. It is then possible that the loads are released a little later and that vehicles arrive relatively too soon for the pick up and have to wait. It could be more favorable for the vehicle to pick up another job instead, but it has a pre-fixed route. It is also possible that the loads are released a little sooner and the vehicle will arrive relatively too late. This means extra load waiting time. With on-line dispatching, it is possible that the vehicle seizes the opportunity to pick up the job that was released relatively sooner.
Similarly, one of the Structured shifts undergoes perturbation. As shown in Figure 15, the transport requests of Inbound and Labeling, and Labeling and Outbound jobs, overlap 10% in the original Structured shifts. With the 5% perturbation on the load release times, the overlap can increase to 20%.
Just like the Structured shifts, one of the High throughput shifts undergoes perturbation. This also means that the overlap of Inbound and Labeling, and Labeling and Outbound jobs in the perturbed High throughput shifts can increase to 20%. Since there is little vehicle idle time in this situation, a vehicle traveling an old route with off-line control can never make up for the time lost when it had to wait for a load that was released relatively later. We expect the load throughput times with off-line control to increase to a relatively higher level than in the cases with some system slack, like the Random shifts and Structured shifts.

4.2 Results

The discussion of the results will start by presenting the performance gap in terms of expected load waiting times between off-line and on-line controlled guided vehicles using pre-arrival information. It will be shown that dispatching vehicles using pre-arrival information can reduce the average load waiting times considerably. However, using too much pre-arrival information can dispatch the vehicles unfavorably, since vehicles arrive relatively too early and wait for the load to be ready for physical transport. This vehicle waiting time could be used to serve another job instead such that the waiting time decreases. This is investigated in Section 4.2.2 where such transportation tasks are suspended. In Section 4.2.3, we present the results using dwell point strategies, i.e. when vehicles are instructed to park at the location where the following task is likely to occur, followed by the results when the closest moving or parked vehicle can be (pre-)assigned to loads. This section is followed by some results that are obtained when the vehicles are dispatched using dwell point strategies combined with pre-arrival information. Finally, we present the effects of routing vehicles on shifts with perturbed release times. In the on-line situations, the vehicles are still dispatched as if the release times appear in real-time, so no different behavior is expected. With off-line control, the vehicles are actually traveling an ‘old’ pre-fixed route. In general, changes encountered between the expected and actual load release times while executing a pre-fixed route will result in a loss of performance.

4.2.1 Dispatching with pre-arrival information

Table 15 gives an overview of the average total load waiting times and the standard deviation of the waiting times for 10 runs for both types of control (on-line and off-line) and the corresponding layout used. For the off-line rules, ‘Optimal’ refers to the average optimal solutions from the TRPTW, i.e. the average minimum load waiting times when two guided vehicles are used. The average waiting times in the ‘Insertion’ column represent the average load waiting times obtained when two vehicles are routed with the Insertion heuristic. This leads to the optimal solution in several instances, overall it deviates about 9% and 12% from the optimal value (see ‘Deviation’ in Table 15) for the U-layout and I-layout respectively. When on-line control rules are used without pre-arrival information (summarized in the rows ‘NVF: 0’ and ‘FCFS: 0’), the ‘Deviation’ from the optimum waiting time varies from 90% to 138%. This means that the loads have to wait about twice as long to be picked up with on-line control compared to off-line control.

	U-layout			I-layout		
Rule	Average (10 runs)	Standard deviation	Deviation from optimum	Average (10 runs)	Standard deviation	Deviation from optimum
Optimal	96.2	12.9	-	79.1	35.6	-
Insertion	105.1	13.5	9 %	88.7	36.5	12 %
NVF: 0	189.2	37.0	97 %	184.4	38.9	133 %
NVF: 5	153.1	27.6	59 %	145.0	42.5	83 %
NVF: 10	126.7	27.0	32 %	118.2	48.5	49 %
NVF: 15	115.3	18.5	20 %	121.5	45.9	54 %
NVF: 20	138.2	42.2	44 %	123.7	48.9	56 %
FCFS: 0	182.5	28.1	90 %	188.5	33.5	138 %
FCFS: 5	147.4	24.3	53 %	145.1	38.4	83 %
FCFS: 10	129.4	30.7	35 %	122.3	44.6	55 %
FCFS: 15	125.2	33.6	30 %	107.5	49.2	36 %
FCFS: 20	128.9	36.3	34 %	102.8	49.4	30 %

Table 15. Average load waiting time for several control rules with Random shifts. The pre-arrival time for on-line control is expressed in time units

With on-line control, the vehicles park at their current location if there are no transport request available. The parked vehicles are idle until a new transport job is available. There is about 15% vehicle idle time in the case without pre-arrival information (not in table). Some of this time can be used when pre-arrival information is available to travel to the next transportation job, hence, possibly reducing load waiting times. The 'NVF: 15' row, shows that a pre-arrival time of 15 time units will decrease the load waiting time deviation to 20% from the optimum in the U-layout environment.

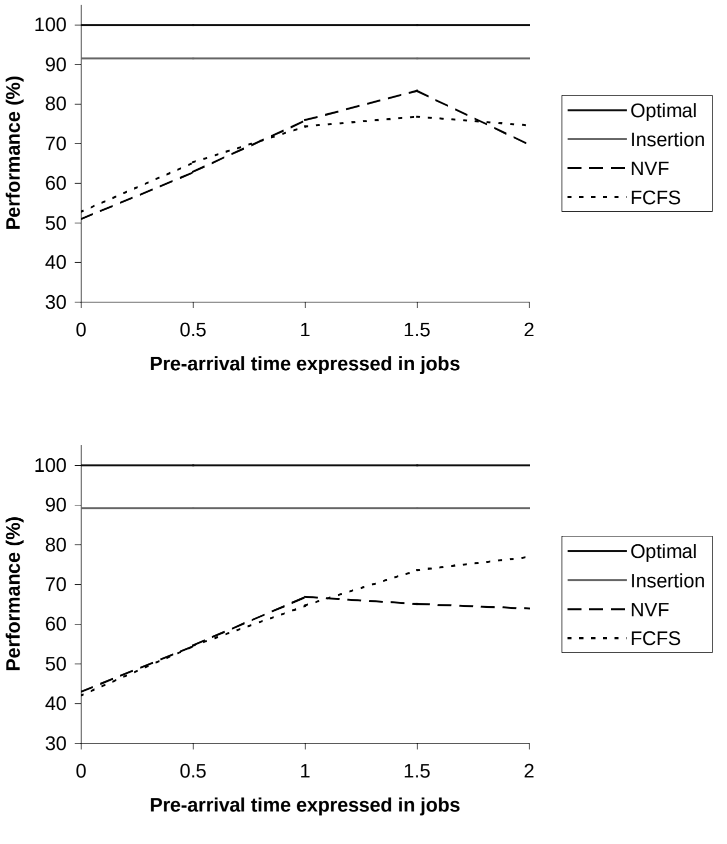

Figure 18. Performance of control rules for the U-layout (top) and I-layout (bottom) as a function of pre-arrival time expressed in the number of jobs with Random shifts

In Figure 18, the optimal value for the load waiting time is represented with a line at 100% performance. The performance of the Insertion heuristic, which deviates about 9% (see Table 15) from the optimum in the U-layout with NVF, is represented with a line at about 92% performance (the ratio of the average load waiting times: 96.2/105.1). The *x*-axis is represented in the average number of jobs of which pre-arrival information is available. A pre-arrival period of 5 time units is comparable to looking 0.5 ($\doteq \frac{12}{126} * 5$) loads ahead on average. Similarly, a pre-arrival period of 20 time units is comparable to looking two loads ahead on average. Using this scale, it can be shown how much pre-arrival information is needed in terms of the number of loads to decrease the average load waiting times of on-line control. We see that the performance of the on-line control rules increase as the pre-arrival time increases. For example, the average waiting times with NVF on the U-layout decrease from 189.2 to 115.3 time units, i.e. by 39% (not in table), when 15 time units of pre-arrival information is available. At a certain point, however, the vehicles are scheduled unfavorably and the performance decreases again. Unfavorable scheduling occurs if partial information is given too far in advance. Vehicles are then allocated to loads too early and the allocation should be reconsidered when new information is made available (see Section 4.2.2).

The Structured shifts represent a more realistic generation order of the load transportation jobs. Because of the design of the layouts and the overlapping periods within the shifts, a combination of dropping off an Inbound load and picking up a load for Labeling or the combination Labeling and Outbound loads, leads to smaller waiting times especially for the U-layout. This can be seen by comparing the values of Table 16 with those of Table 15. The average waiting times for the U-layout reduce with almost 20% for off-line control and 30% for both layouts using on-line control without pre-arrival information. For the I-layout, the load waiting time with ‘FCFS: 0’ and ‘NVF: 0’ is about 80% higher than the optimum, this was more than 130% with Random shifts.

The difference between Insertion and the Optimal value is reduced to 4% or less, and a pre-arrival period of 10 time units is necessary (about one load on average) to get within 26% of the Off-line performance. This means that less pre-arrival information is needed to reduce the average load waiting times with Structured shifts compared to Random shifts.

	U-layout			I-layout		
Rule	Average (10 runs)	Standard deviation	Deviation from optimum	Average (10 runs)	Standard deviation	Deviation from optimum
Optimal	80.6	10.9	-	78.8	24.5	-
Insertion	83.5	12.2	4 %	81.5	26.1	3 %
NVF: 0	139.8	17.4	73 %	142.1	20.4	80 %
NVF: 5	108.6	13.5	35 %	109.2	26.7	39 %
NVF: 10	87.3	13.7	8 %	98.9	27.8	26 %
NVF: 15	93.8	22.3	16 %	108.8	23.3	38 %
NVF: 20	119.1	25.7	48 %	121.4	35.5	54 %
FCFS: 0	142.0	16.8	76 %	143.2	24.1	82 %
FCFS: 5	111.5	14.2	38 %	106.7	23.3	35 %
FCFS: 10	93.0	16.6	15 %	96.1	23.8	22 %
FCFS: 15	100.9	27.3	25 %	96.4	25.1	22 %
FCFS: 20	122.2	25.0	52 %	110.8	24.7	41 %

Table 16. Average load waiting time for several control rules for Structured shifts. The pre-arrival time for on-line control is expressed in time units

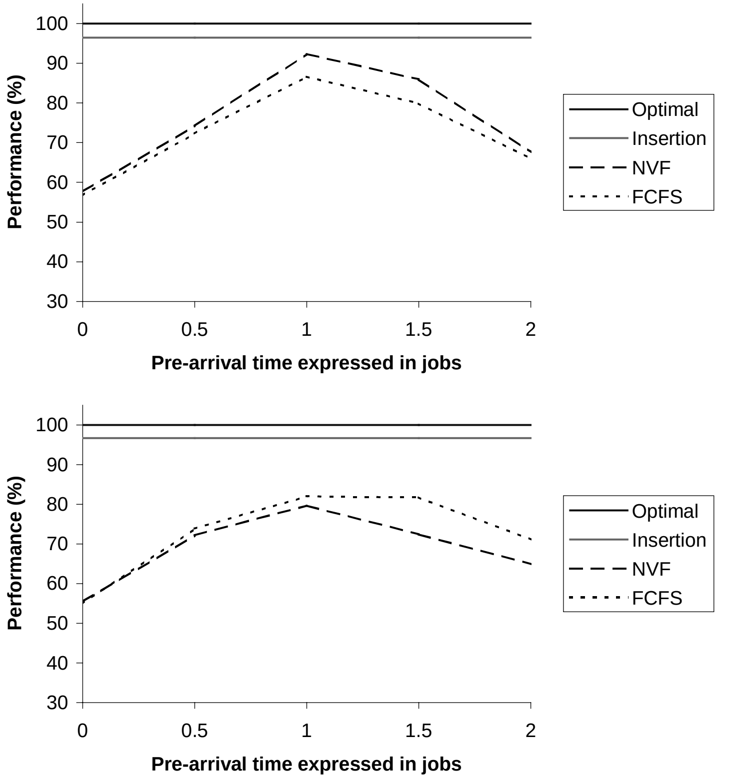

Figure 19. Performance of control rules for the U-layout (top) and I-layout (bottom) as a function of pre-arrival time expressed in the number of jobs with Structured shifts

Figure 19 shows a similar trend as Figure 18. It is clear that pre-arrival information can reduce the load waiting time and thus increase the performance of the control rule. In each case the average waiting times are most favorable when 10 time units of pre-arrival information is available, i.e. when information of about one load is available in advance. In those cases the average load waiting times decrease from 30% (i.e. from 142.1 to 98.9 time units) obtained with NVF on the I-layout, to 39% for NVF on the U-layout. Using more information than one load in advance reduces the positive effect on the performance for the on-line dispatching rules.

For the case of High throughput shifts (where there is no vehicle idle time), the number of jobs is increased to 60. The optimal performance with TRPTW could not be calculated since this led to high running times and computer memory problems. Since the Insertion heuristic leads to very satisfactory results (see the previous sections) in a very simple and quick way (less than one second calculation time), we will continue to use Insertion for the off-line control.

The results in Table 17 show that the deviation in load waiting times between on-line and off-line control is smaller when extra empty vehicle travel time is reduced by eliminating the slack in the system. Load waiting time deviations in High throughput shifts without pre-arrival information are about the same as the deviations with pre-arrival information in Structured shifts. So in busy internal transport systems (with high throughput), where vehicles have little to no idle time, the on-line control rules already work satisfactory. In this case, pre-arrival information has little effect on the performance of on-line control.

	U-layout			I-layout		
Rule	Average (10 runs)	Standard deviation	Deviation from Insertion	Average (10 runs)	Standard deviation	Deviation from Insertion
Insertion	5722.8	608	-	5114.3	353.7	-
NVF: 0	6207.3	680.5	8 %	5559.8	277.9	9 %
NVF: 5	6075.9	720.3	6 %	5462.0	313.0	7 %
NVF: 10	6090.0	771.0	6 %	5610.8	302.7	10 %
NVF: 15	6199.5	701.4	8 %	6146.0	447.2	20 %
NVF: 20	6448.4	658.6	13 %	6845.8	447.2	33 %
FCFS: 0	6105.0	561.6	7 %	6307.0	275.8	23 %
FCFS: 5	5945.2	600.6	4 %	6231.1	313.0	22 %
FCFS: 10	5883.4	614.4	3 %	6241.5	307.7	22 %
FCFS: 15	5961.0	597.4	4 %	6362.4	268.7	24 %
FCFS: 20	6082.7	571.0	6 %	6483.7	482.0	27 %

Table 17. Average load waiting times for several control rules for High throughput shifts. The pre-arrival time for on-line control is expressed in time units

Figure 20 shows similar results for the High Throughput cases compared to the cases described above, pre-arrival information of about one load yields maximum performance.

Note that the performance with the NVF rule without pre-arrival information (NVF: 0) in the I-layout environment, seems to be more favorable compared to the modified FCFS (FCFS: 0) in the I-layout environment. This can be observed for the Random shifts and Structured shifts cases as well.

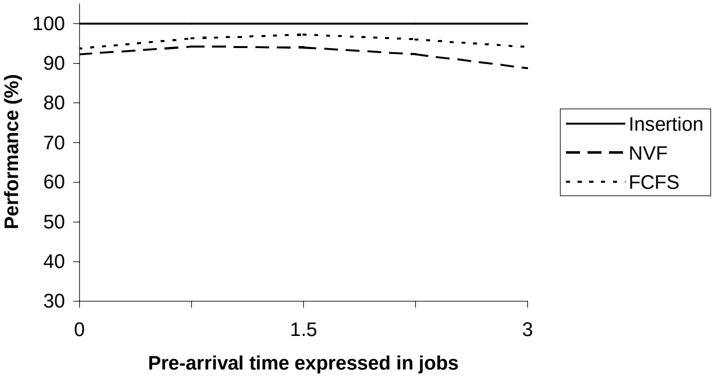

Figure 20. Performance of control rules for the U-layout (top) and I-layout (bottom) as a function of pre-arrival time expressed in the number of jobs with High throughput shifts

4.2.2 Dispatching experiments with suspended jobs

In the previous section, vehicles dropping off loads at an input queue of a station first check the output queue of that station. This happens for both the NVF and the modified FCFS rule; for the NVF rule the load claims the closest vehicle, which is at that moment at the same station, and the vehicles dispatched with the modified FCFS rule always inspect the output queue of the last drop off location first before checking any other queue. As a result, when pre-arrival information is given far in advance about a load which will be made available near or at the current parking or delivery location of the vehicle, vehicles will be matched to that load and wait until the load is physically ready for transport. The idea is that the vehicles could have used this vehicle waiting time more efficiently by suspending this load-to-vehicle match and serve another job at another location instead.

In this section we present the results of a dispatching rule variation where the vehicle suspends a task at its current parking or delivery location when the expected waiting time of the vehicle is sufficiently long to serve another request first. The suspended job can be served by the same or another vehicle, as soon as the expected vehicle waiting time is smaller than the average travel time to another pick up location.
The suspended job rule is introduced to investigate if the increase in average waiting times when too much pre-arrival information is used (see the previous section) can be reduced. The experiment is carried out for the dispatching rule, layout and shift type combination of the previous section which has shown the largest increase in average load waiting times when too much pre-arrival information is used. The largest increase in average load waiting times occurred with the NVF dispatching rule using High throughput shifts on the I-layout environment. As mentioned above, jobs are suspended when the expected vehicle waiting time (the time vehicles have to wait at the load release location until the load can actually be transported) is larger than the expected travel time to another pick up point. Using the distance/time units for the I-layout (see Section 3.2), the average travel time from each pick up location to any other pick up location is equal to $9\frac{3}{8}$ time units. Table 18 and Figure 21 show the results when jobs can be suspended.

Situation	Using pre-arrival information		Using pre-arrival information with suspended claims	
Rule	Average (10 runs)	Deviation from Insertion	Average (10 runs)	Deviation from Insertion
Insertion	5114.3	-	5114.3	-
NVF: 0	5559.8	9 %	5559.8	9 %
NVF: 5	5462.0	7 %	5462.0	7 %
NVF: 10	5610.8	10 %	5610.8	10 %
NVF: 15	6146.0	20 %	5946.2	16 %
NVF: 20	6845.8	33 %	6364.6	24 %

Table 18. Pre-arrival information versus pre-arrival information with suspended claims for high throughput shifts on the I-layout environment using NVF

Since the time limit used to suspend jobs is almost 10 time units ($9\frac{3}{8}$ to be exact), no actual changes in performance were expected to occur when pre-arrival information is available up to 10 time units beforehand. When more than 10 time units of pre-arrival information is available, jobs can be suspended and changes in the load-to-vehicle allocation occur. The results show that relative improvements are possible but that suspending jobs this way is not enough to eventually prevent some decrease in performance (increase in average load waiting times). It also shows that even when the vehicle-to-load assignments can be suspended, on-line dispatching with (too much) pre-arrival information is still greedy and inefficient in the sense that some assignments are irrevocable. On the other hand, the slight improvements also indicate that a more complex rule, which can (re-)schedule all loads to vehicles based on available information within the pre-arrival time, can increase the performance even more and possibly prevent any decrease in performance. Algorithms for such complex rules can be categorized as on-line optimization algorithms, which are generally too complex and time consuming to be used in (real-time) practical environments.

Figure 21. Performance of control rules for the I-layout as a function of pre-arrival time expressed in the number of jobs with High throughput shifts

4.2.3 Dispatching with vehicle dwell points

Another way to reduce average load waiting times is to park idle vehicles at or near locations where the next transportation request is likely to originate (similar to central zone positioning, described in Section 2.8). However, if a vehicle moving to a parking location must reach the parking point before becoming eligible to pick up a load, it may pass by waiting loads and thereby wasting vehicle capacity and possibly decreasing the performance of the system. In this section we investigate the dwell point strategies (discussed in Section 4.1) by which the closest parked or going to park vehicle is considered to be matched to a load requesting transport.

Table 19 and Table 20 show the average load waiting times in the standard situation without using dwell point strategies and when the dwell point strategies are used for Random shifts and Structured shifts respectively. The most likely location for the next transportation task is only known for the structured shifts, and the Structured dwell point strategy is used (the dwell point location changes during the shift). For the Random shifts, the vehicles are sent to park at the location from which the average travel time to the next pick up location is minimal. In the case of the U-layout the dwell point is at the Labeling station, and in the case of the I-layout the dwell point is at the Depot (see Table 14).
The results show that considerable gains can be made when dwell point strategies are used. The results are most favorable for the Structured shifts where the vehicles are parked at the

most likely point of the next transport request (the Structured dwell point strategy). For Random shifts the results are most favorable on the I-layout. This is due to the various transport distances that are in some cases longer than the distances on the U-layout although the average transport distance from the dwell points are the same (see Table 14). When distances between locations are far, sending vehicles to a dwell point will reduce the reaction or response time of the vehicle to a load request, hence reducing vehicle waiting time relatively more.

	Off-line control		On-line control: NVF		On-line control: FCFS	
	Optimal	Insertion	Standard	Dwell points	Standard	Dwell points
U-layout						
Average	96.2	105.1	189.2	170.6	182.5	169.6
St. dev.	12.9	13.5	37.0	36.1	28.1	36.5
Deviation	-	9 %	97 %	77 %	90 %	76 %
I-layout						
Average	79.1	88.7	184.4	150.5	188.5	149.8
St. dev	35.6	36.5	38.9	38.8	33.5	41.6
Deviation	-	12 %	133 %	90 %	138 %	89 %

Table 19. Average load waiting times with Random shifts for the standard situation and when the dwell point strategies for Random shifts are used

The waiting time reduction using the dwell point strategy with the FCFS on the I-layout with Structured shifts is most favorable. In this case the waiting time deviation becomes almost half the deviation for the (standard) situation without the dwell point strategy (see Table 20).

	Off-line control		On-line control: NVF		On-line control: FCFS	
	Optimal	Insertion	Standard	Dwell points	Standard	Dwell points
U-layout						
Average	80.6	83.5	139.8	116.4	142.0	116.4
St. dev.	10.9	12.2	17.4	21.0	16.8	21.0
Deviation	-	4 %	73 %	44 %	76 %	44 %
I-layout						
Average	78.8	81.5	142.1	114.2	143.2	112.0
St. dev	24.5	26.1	20.4	33.5	24.1	29.0
Deviation	-	3 %	80 %	45 %	82 %	42 %

Table 20. Average load waiting times with Structured shifts for the standard situation and when the Structured dwell point strategy is used

Intuitively, the best parking strategy for random load generations would be parking the vehicles randomly, this would be similar to the point of release strategy for Random shifts. Intuitively it would also be better not to park the vehicles at locations without pick up requests such as the Shipping station or the depot. The results for the I-layout in Table 19 are counter examples for both intuitions. With the dwell point strategies studied in this section it is possible that both vehicles are sent to park to the same dwell point. Intuitively, further reductions in average load waiting times could be possible when vehicles are balanced over several dwell points on the network (see also Ventura and Lee, 2000). In that case, vehicles can be sent to alternative dwell points when certain dwell points are already occupied by a vehicle. Such strategies might be useful for environments with numerous vehicles. The probability that both vehicles are idle at the same dwell point is rather small for the environments studied in this section, and any possible effects of unbalanced idle vehicles are therefore negligible.

	Off-line control		On-line control: NVF		On-line control: FCFS	
	Optimal	Insertion	Standard	Dwell points	Standard	Dwell points
U-layout						
Average	96.2	105.1	189.2	172.9	182.5	172.7
St. dev.	12.9	13.5	37.0	29.4	28.1	25.9
Deviation	-	9 %	97 %	80 %	90 %	80 %
I-layout						
Average	79.1	88.7	184.4	174.9	188.5	171.9
St. dev	35.6	36.5	38.9	45.7	33.5	43.8
Deviation	-	12 %	133 %	121 %	138 %	117 %

Table 21. Average load waiting times with Random shifts for the standard situation and when the Structured dwell point strategy is used

Table 21 shows the results of an experiment when the *Structured dwell point strategy* is used for *Random shifts*. This situation could occur when the Structured dwell point strategy is implemented but when the actual release locations of loads are random. In this case, the vehicles are instructed to park at a pick up point which varies during the shift, but which is not necessarily the pick up point closest to or is likely to place the next transport requests. The results show that the point of release dwell point strategy can still be outperformed when both parked and going to park vehicles can be matched to a load, even when it is possible that the vehicles are instructed to move away from the location with the next transport request. Note, however, that the average waiting time reductions for the Random shifts are less favorable when the Structured dwell point strategy is used, compared to the waiting times for the Random shifts with dwell point strategies especially designed for the Random shifts (Table 19).

4.2.4 Pre-assigning moving vehicles to loads

The average load waiting times reductions for Random shifts using the Structured dwell point strategy (see the previous section) partly illustrates that matching the closest idle or going to park vehicle to a load can reduce average load waiting times. In this section, not only idle or going to park vehicles, but also vehicles retrieving or delivering a load are considered to be matched to a load by calculating the distance that those vehicles still have to travel to reach the requesting pick up. The (moving) vehicle closest to the load will be assigned to the pick up task. In case the vehicle is retrieving or delivering a load, no other tasks should be pre-assigned to the vehicle (see Figure 17). The advantage of pre-assigning a moving vehicle to a load, is that the load will be matched to the vehicle that can pick up the load soonest (since there are no acceleration/deceleration or congestion effects), given the current situation of the system, and possibly creating favorable double-plays. A disadvantage of pre-assigning (moving) vehicles to loads is that certain favorable pick up combinations can be lost since vehicles that have been pre-assigned to loads cannot claim other loads that have become available while the vehicles are handling their current tasks (claims are irrevocable). Instead, vehicles are only considered for assignment to a newly generated load (or vice versa) when their current assigned tasks or pre-assigned tasks are completed or if they are idle or going to park.

	Off-line control		On-line control: NVF		On-line control: FCFS	
	Optimal	Insertion	Standard	Pre-assignment	Standard	Pre-assignment
U-layout						
Average	96.2	105.1	189.2	169.7	182.5	170.6
St. dev.	12.9	13.5	37.0	42.5	28.1	39.1
Deviation	-	9 %	97 %	76 %	90 %	77 %
I-layout						
Average	79.1	88.7	184.4	179.6	188.5	184.2
St. dev	35.6	36.5	38.9	36.6	33.5	44.9
Deviation	-	12 %	133 %	127 %	138 %	133 %

Table 22. Average load waiting times by pre-assignment of (moving) vehicles to loads on Random shifts

The results in Table 22 show the advantage of pre-assigning moving vehicles to loads for Random shifts. The situation of NVF on the U-layout is most favorable where the deviation from the optimum reduces from 97% to 76%.

	Off-line control		On-line control: NVF		On-line control: FCFS	
	Optimal	Insertion	Standard	Pre-assignment	Standard	Pre-assignment
U-layout						
Average	80.6	83.5	139.8	139.8	142.0	143.6
St. dev.	10.9	12.2	17.4	17.4	16.8	19.8
Deviation	-	4 %	73 %	73 %	76 %	78 %
I-layout						
Average	78.8	81.5	142.1	141.6	143.2	143.9
St. dev	24.5	26.1	20.4	23.5	24.1	29.9
Deviation	-	3 %	80 %	80 %	82 %	83 %

Table 23. Average load waiting times by pre-assignment of (moving) vehicles to loads on Structured shifts

Table 23 shows that the waiting times do not necessarily benefit from pre-assigning moving vehicles to loads. In fact, the performance slightly decreases if the FCFS dispatching rule is used. In this case the pre-assignment strategy on structured shifts unfavorably changed some pick up combinations (created by the overlap of job types) and certain double-play situations were missed because some load-to-vehicle allocations ware made too soon. A similar situations occurred in Sections 4.2.1 and 4.2.2 when vehicle-to-load assignments were made too soon when relatively too much pre-arrival information was used. In Section 4.2.2 it was shown that the irrevocable pre-assignments are greedy and inefficient and that improvements can be realized when vehicle-to-load assignments can be suspended. Based on the ideas of suspended jobs, similar methods can be used when pre-assigning moving vehicles to loads to improve the performance. For example, by reconsidering all pre-assigned jobs every time a new job is released or when a load is dropped off and new calculations have to be made.

4.2.5 Combining pre-arrival information and dwell point strategies

We have seen in the previous sections that certain strategies can increase the performance of a system (reduce average load waiting times) but can also decrease the performance. For example, the use of pre-arrival information reduces waiting times but only marginal for high throughput shifts and using too much pre-arrival information increases the average load waiting times for the on-line rules studied here. This in turn can be slightly restrained by suspending jobs. Furthermore, pre-assigning moving vehicles to loads can be favorable for Random shifts but not so much for Structured shifts. And the average load waiting times obtained with the use of dwell point strategies are favorable for any situation but are most favorable for Structured shifts on the I-layout.

In this section we investigate two combinations of the two most favorable strategies: dwell point strategies, which reduces the average waiting times in each case, and using pre-arrival information, which reduces the waiting times when limited information is given beforehand. The layout, dispatching rule, pre-arrival time and shift type combination is determined as follows:

1. The performance gains of using the dwell point strategy relative to the standard situation are sorted in ascending order, so the layout, dispatching rule and shift type combination with the best performance gain is ranked first and the poorest combination is ranked last (in this case 8^{th}).
2. Similarly, the performance gains of using the most favorable amount of pre-arrival information relative to the standard situation are sorted in ascending order and the combinations are ranked.
3. The layout, dispatching rule, pre-arrival time and shift type combinations with the lowest (best) total rank and highest (poorest) total rank calculated by adding the individual ranks of 1) and 2) will be investigated for the dwell point and pre-arrival information combination.

In this case the best total rank obtained is three, which is calculated as follows: the FCFS rule with Random shifts on the I-layout that scores rank 2 for the dwell point strategy, and scores rank 1 with 20 units of pre-arrival time. The poorest total rank is 15; calculated as follows: the FCFS rule with Random shifts on the U-layout scores rank 8 for the dwell point strategy, and scores rank 7 with 15 units of pre-arrival time.

	Off-line control		On-line control: FCFS			
	Optimal	Insertion	Standard	Dwell points	Pre-arrival	Combination
U-layout						
Average	96.2	105.1	182.5	169.6	125.2	115.5
St. dev	12.9	13.5	28.1	36.5	33.6	44.6
Deviation	-	9 %	90 %	76 %	30 %	20 %
I-layout						
Average	79.1	88.7	188.5	149.8	102.8	102.8
St. dev	35.6	36.5	33.5	41.6	49.4	49.4
Deviation	-	12 %	138 %	89 %	30 %	30 %

Table 24. Average load waiting times combining pre-arrival information and the dwell point strategy with Random shifts and the FCFS dispatching rule

Table 24 shows that the average load waiting times can decrease in the situation in which the least favorable (rank 8) dwell point strategy is combined with pre-arrival information. In this case the individual result were poor enough to have room for improvement, which is utilized by the combination. The results also show that combining the two most favorable strategies may not necessarily result in a gain in performance. The waiting time reductions with pre-arrival information are dominant in the poorest ranked combination (the I-layout combination with the two most favorable performance gains) and the average waiting times with the combined strategies remains the same (with respect to the results with pre-arrival information).

The average load waiting times even increased slightly in one experiment (not shown in table) with the best ranking combination with the FCFS rule, I-layout and *Random shifts* when 20 units of pre-arrival information and the *Structured dwell point strategy*. Although the average load waiting times decreased for when strategies were used individually (see Table 15 for the strategy using pre-arrival information and Table 21 for the results of the Structured dwell point strategy on Random shifts).

Based on the different results mention above it is difficult to conclude when strategies should be combined. Especially since it is not known beforehand what the rank of the individual strategy was. However, when the results obtained when the ‘wrong’ dwell point strategy is combined with pre-arrival information are omitted, it can be concluded that pre-arrival information and dwell point strategies can be combined to obtain more favorable average load waiting times.

4.2.6 Dispatching with perturbed release times

To study the effects on average load waiting times when off-line routed vehicles encounter small changes (of 5%) in the release time, perturbed load release times were introduced. In this way, off-line controlled vehicles are actually routed in a stochastic environment. Intuitively, on-line vehicle dispatching should be more stable (or robust) since the vehicles still encounter the release times in real-time. With off-line control, however, once an allocation is made it is irrevocable, i.e. not reconsidered later (since this would imply on-line control). The results with the on-line dispatching rules presented in this section are obtained from the standard situation, i.e. without pre-arrival time, suspended jobs, dwell point strategies or moving vehicle-to-load pre-assignment.

Run 1 without Perturbation				10 perturbations of Run 1		
Layout	U-layout					
Rule	Insertion	NVF	FCFS	$Insertion_{98}$	NVF	FCFS
Average	98	210	172	105.2	185.0	190.6
Deviation	-	114 %	76 %	-	76 %	81 %
Ratio	-	-	-	1.39	1.29	1.25
Layout	I-layout					
Rule	Insertion	NVF	FCFS	$Insertion_{44}$	NVF	FCFS
Average	44	173	173	52.0	197.9	188.1
Deviation	-	293 %	293 %	-	281 %	262 %
Ratio	-	-	-	1.78	1.47	1.35

Table 25. Load waiting times with perturbation on Random shifts

In Table 25 we use the ‘Ratio’ of the highest and lowest waiting times (worst and best performance value) of 10 runs to obtain an impression of the robustness in the average load waiting times. The ratio for off-line control is in both cases higher than the ratio of on-line controlled vehicles. This indicates that routing vehicles on an ‘old’ schedule is less robust than using on-line control to seize the opportunity to service jobs in a more favorable sequence. This can also be seen by the general reduction of the deviation from off-line control. Notice that the average load waiting times increase when the vehicles are routed with the ‘old’ off-line schedule denoted by ‘$Insertion_{98}$’ and ‘$Insertion_{44}$’ for the U-layout and I-layout respectively. This was expected since the time lost by unfavorable scheduling is not compensated. For on-line control the change in average waiting times seem to behave unpredictably. This is due to the fact that vehicles are dispatched on relatively new or different shifts with the same jobs shuffled in time. However, since vehicles can seize the opportunity to serve the transport requests relatively more favorably, the performance deviation with off-line control decreases.

Table 26 shows the results for perturbed release times with Structured shifts. As in the case of to Random shifts, perturbed release times result in a decrease in performance for off-line control and therefore a relative increase in performance for on-line control (from 59% to 35% deviations from off-line control for the U-layout). This is due to the increase in average load waiting times when transport requests are served with outdated vehicle routes. Notice that in this case the Ratio (robustness) obtained with off-line control on the I-layout is relatively small and is in fact not (or too small to be) outperformed by the Ratio’s obtained with on-line control.

Run 1 without Perturbation				10 perturbations of Run 1		
Layout	U-layout					
Rule	Insertion	NVF	FCFS	Insertion_{83}	NVF	FCFS
Average	83	132	132	93.2	126.3	126.3
Deviation	-	59 %	59 %	-	35 %	35 %
Ratio	-	-	-	1.43	1.25	1.25
Layout	I-layout					
Rule	Insertion	NVF	FCFS	Insertion_{131}	NVF	FCFS
Average	131	147	147	138.9	147.2	149.1
Deviation	-	12 %	12 %	-	6 %	7 %
Ratio	-	-	-	1.18	1.28	1.28

Table 26. Waiting times with perturbation on Structured shifts

In the standard High throughput situation, the performance with on-line dispatched vehicles deviates slightly from the performance with off-line controlled vehicles. We already deduced that this deviation is much smaller compared to the previous two shift compositions, because of the lack of vehicle idle time. The performance of off-line control deteriorates so much when the non-idle vehicles are guided with a previously calculated off-line route, that on-line dispatched vehicles can outperform off-line controlled vehicles. This can be seen by the negative ‘Deviation’ values in Table 27. Apparently, the on-line dispatched vehicles seize the opportunity to transport the loads in a different sequence when the load release times change. The off-line controlled vehicles will try to complete the original route and will arrive relatively too late when loads are released sooner (instead of doing another job) or have to wait when loads are released relatively later.

Run 1 without Perturbation				10 perturbations of Run 1		
Layout	U-layout					
Rule	Insertion	NVF	FCFS	Insertion_{5186}	NVF	FCFS
Average	5186	5738	5508	6269.4	5564.7	5374.7
Deviation	-	11 %	6 %	-	-11 %	-14 %
Ratio	-	-	-	1.23	1.08	1.13
Layout	I-layout					
Rule	Insertion	NVF	FCFS	Insertion_{5723}	NVF	FCFS
Average	5723	6025.2	6738	6045.9	6000.8	6761.4
Deviation	-	5 %	18 %	-	-0.7 %	12 %
Ratio	-	-	-	1.06	1.09	1.07

Table 27. Waiting times with perturbation on High throughput shifts

The deviation Ratio for off-line control in the I-layout is about the same for off-line control and on-line control. However, the Ratio for off-line control in the U-layout environment is higher than the Ratio for on-line control. This supports the idea that the performance with on-line control in a stochastic environment is more stable than the performance of off-line control in a stochastic environment when the load-to-vehicle allocations are irrevocable.

4.3 Concluding remarks

Using off-line control means that all information on load release times, origins and destinations has to be known in advance. This is not a situation found in internal transportation situations in practice due to the stochastic nature of such environments (see Sections 1.1 and 2.7). However, often some information can be made available before the load has physically arrived or has been released at the pick up location, i.e. load pre-arrival information. In this chapter, this pre-arrival information on load release times is used to reduce the average load waiting times when vehicles are controlled using on-line dispatching rules. The resulting average load waiting times define the performance of the vehicle control rules, which increases as the average load waiting times decrease. A performance gain (i.e. decrease in average load waiting time) can be realized by already traveling to a load before it is physically released, therefore the load is picked up relatively sooner. However, if the information of too many loads is made available beforehand, the load to vehicle allocation can become unfavorable, which in turn increases the average load waiting times and reduces the performance (see Table 28). The average load waiting times obtained for on-line control using pre-arrival information are compared to the theoretical situation where all information is known beforehand, for particular internal transportation environments under various circumstances. Table 28 gives an overview of the performance gap in average load waiting times between on-line and off-line control for different control rules, layouts and throughput levels.

Control rule	NVF					FCFS					Insertion
Loads known ahead	0	0.5	1	1.5	2	0	0.5	1	1.5	2	All
U-layout											
Structured shifts	60%	77%	96%	89%	70%	59%	75%	90%	83%	68%	100%
High throughput shifts	92%	94%*	94%*	94%	93%*	93%	95%*	97%*	97%	96%*	100%
I-layout											
Structured shifts	57%	75%	82%	75%	67%	57%	76%	85%	85%	74%	100%
High throughput shifts	92%	93%*	93%*	91%	86%*	81%	82%*	82%*	82%	81%*	100%

* Interpolated value

Table 28. Average performance of the on-line control rules with pre-arrival information relative to the performance of the off-line Insertion heuristic (100%)

Since the average load waiting times of the High throughput shifts are calculated with the Insertion heuristic, all performance calculations of Table 28 are given relative to the off-line performance of the Insertion heuristic. For example, the average load waiting times are smallest with the off-line Insertion rule, and therefore defines the best performance possible of 100%. The average performance of the NVF control rule in the U-layout with Structured shifts without pre-arrival information (denoted as 0 ‘Loads known ahead’) is the waiting time obtained with Insertion divided by the waiting time obtained with NVF, i.e. (see Table 16) 83.5/139.8 ($\doteq$ 0.6); denoted by 60%. Note that for low throughput environments (Structured shifts), the performance without pre-arrival information is in

general about 60% of the performance that could be obtained with off-line control. However, using some pre-arrival information of the loads, will increase the performance of the on-line control rules. The pre-arrival information of only one load on average increases on-line performance such that it approximates off-line performance.
For High throughput environments, the average load waiting times with on-line control already approximates off-line performance well without the use of pre-arrival information. Using pre-arrival information increases performance slightly and using pre-arrival information of too many loads can even reduce the performance to a level worse than the performance without pre-arrival information. This effect can be (slightly) compensated as is shown in an experiment where transport request are suspended when the expected vehicle waiting time for a load is greater than the average traveling time to another pick up point (Section 4.2.2).

Table 29 shows the relative increase in performance (decrease in average load waiting times) when the dwell point strategies are used and when moving vehicles can be pre-assigned to a load, compared to the standard situation (without pre-arrival information, suspended jobs, the dwell point strategy, or pre-assigning moving vehicles to loads). For example, the 10% decrease in average load waiting times noted in Table 29 for NVF with the dwell point strategy on Random shifts is calculated (using the values obtained from Table 19) by: (189.2-170.6)/189.2*100%. As expected, the average load waiting times with Structured shifts benefit most from the dwell point strategy since the vehicles can be sent to park near the next transport request with more accuracy.
The FCFS rule with the dwell point strategy also starts out as a load-initiative rule since the load entering the system first checks for the closest idle or going to park vehicle. Furthermore, the FCFS rule has distance-based elements due to the ‘closest’ criterion. As a result, the performance increase using the FCFS rule is relatively higher compared to the NVF rule, which was already a load-initiative distance-based rule.

Strategy	Dwell point		Pre-assign (moving) vehicles	
Control rule	NVF	FCFS	NVF	FCFS
U-layout				
Random shifts	10 %	7 %	10 %	7 %
Structured shifts	17 %	18 %	0 %	-1 %
I-layout				
Random shifts	18 %	21 %	3 %	2 %
Structured shifts	20 %	22 %	0 %	0 %

Table 29. Decrease in average load waiting times using different strategies relative to the standard situations

Table 29 also shows that pre-assigning moving vehicles to loads can lead to a performance increase, but may in some cases also lead to a decrease in performance (increase or negative decrease in load waiting times). This decrease is caused by unfavorable allocations when pre-assignments of vehicles to loads cannot be undone when potential more favorable combinations are possible when new loads enter the system. This is one of the drawbacks of the pre-assignment strategy. Another drawback is the practical value of this strategy since the information needed about the exact location of moving vehicles is not known in practice (due to acceleration/deceleration, congestion varying speeds, etc.). We will discuss this in more detail in the next chapter.

In practice, the exact data on release times become more unreliable as the time to the actual release of the load increases. Using perturbed load release times, we investigated the effects of routing vehicles with outdated information. As the given release times become unreliable, the load waiting times with off-line control start to deviate since vehicles are routed on an ‘old’ schedule. This deviation is measured with the Ratio of the lowest and highest performance and the reduction of the average load waiting time deviations between off-line and on-line controlled vehicles. For the on-line controlled vehicles, the perturbation makes little difference since the loads are still released with (updated) real-time information, which is stochastic in perspective of the vehicles. But for the off-line controlled vehicles, the ‘new’ release-schedule does not match the ‘old’ schedule of the pre-calculated route. The vehicles can therefore arrive too early at a location and have to wait for the loads to be physically released and leave another load waiting to be picked up at that time, or at a later time since the vehicle has been delayed. On the other hand, vehicles can arrive as scheduled at a location (or at a later moment because of a previous delay) while the load has actually been released earlier than expected. This can also cause extra load waiting time. From Table 30 we can see that when release times are perturbed, the differences in average load waiting times between on-line dispatching and off-line control reduces. In fact when the vehicles are fully utilized, on-line dispatching rules can outperform off-line control. This is represented by the negative deviation values in Table 30.

Situation	Performance deviation by heavily utilized GVs (High throughput shifts)	Performance deviation by GVs with idle time (Structured shifts)
U-layout	NVF / FCFS	NVF / FCFS
Standard	8 % / 7 %	67 % / 70 %
Release time perturbation	-11 % / -14 %	35 % / 35 %
I-layout	NVF / FCFS	NVF / FCFS
Standard	9 % / 23 %	74 % / 76 %
Release time perturbation	-0.7 % / 12 %	6 % / 7 %

Table 30. Average load waiting time deviations with on-line dispatching relative to off-line (Insertion) control

Results in Section 4.3 have shown that off-line control is more sensitive to unexpected deviations in release times than on-line control. The results are increasing load waiting times (which in turn reduces the performance gap) and a general increase in the Ratio of the highest and lowest average waiting times, which indicates unstability (see for example Table 25). On-line control is more robust since it can seize the opportunity to service jobs in a more favorable sequence. This is logical, unexpected deviation or unreliable information creates a stochastic environment, unsuitable for off-line control. Furthermore,

the NVF rule in the standard situation (without pre-arrival information, etc.) seems to perform slightly more favorable in a non-symmetric layout environment. A similar observation was made in the previous chapter.
Since it is impossible to have all load information perfectly in advance anyway, off-line control will not be studied for the internal transport in cases in practice described in the next chapter.

In the next chapter, we will investigate the effects of different dispatching rules for three practical case studies. Several ideas from this chapter and the previous chapter are studied for the practical situations. The three cases have been studied and analyzed using detailed simulation models. We will look at the effects of using simple dispatching rules versus more complex dispatching rules, increasing the number of vehicles versus increasing the vehicle capacity and using limited load pre-arrival information. Furthermore, we will attempt to classify the different dispatching rules according to their performance, and investigate similarities in the classifications of certain dispatching rules for different environments.

Chapter 5

Control of vehicle-based transport systems in practice

In Chapter 2 we discussed push and pull-type dispatching rules and discussed that on-line vehicle dispatching rules can make use of two types of operating decisions. The first determines which load should be matched to a vehicle when the vehicle is ready for the next task (vehicle-initiated dispatching). The second determines which vehicle is selected when loads initiate transportation requests (workcenter or load-initiated dispatching). We have also discussed literature (see Section 2.2) which study control rules used to balance workloads, minimize travel distances, minimize queue lengths, minimize average load waiting times, etc. On-line control rules, such as first-come-first-served (FCFS), nearest-vehicle-first (NVF) and nearest-workstation-first (NWF) are relatively simple and commonly used. Such rules perform reasonably well in general, but behave different for various performance criteria. For example, NWF may minimize empty travel time, but can ignore remote areas; and FCFS assigns a vehicle to the oldest transportation request in the system, but has no special regard for the vehicle empty travel time.
From the results of the previous chapters we can see that the criteria used to measure the performance of internal transport systems are not only influenced by the type of dispatching rule used, but by many other factors as well. Based on the results of the previous chapters, the literature and intuition, a number of factors of a transportation environment that can influence the performance of a vehicle transport system are listed in Table 31.

Dispatching rule (vehicle-initiated, load-initiated, time-based, distance-based, etc.)
Empty/loaded behavior decision rules (dwell point strategies, pre- or re-assign vehicles)
Vehicles (number, uni- or multi-load, speed, pick up and set down times, etc.)
Path layout (remote areas, various distances, equal distances, dwell point locations)
Transport time/distance (dispersion, long, short)
Load throughput (high, low)
Work-load intensity (peak periods, evenly spread)
Track control (congestion, zone control, collision avoidance)
Degree of load release time and location predictability (pre-arrival time)

Table 31. Examples of factors that influence the performance of a vehicle transport system

Although the list is not complete, it provides us with an indication of the many possible factors that can influence any type of performance measurement. Mathematical or simulation modeling of a vehicle transport system can be used to investigate the impact of several factors of the environment on the performance of the system. In this chapter, we will study the performance of different centralized GV control systems in three different environments. Each environment is characterized by different factors that can influence the performance of the vehicle-control system. We will describe the different factors when we describe the models of the three companies later in this chapter and reflect on the differences and similarities in Section 5.5, i.e. after all individual details of the companies have been described. The three companies, which are studied using detailed simulation models with real company data, are:

1. A European distribution center (EDC) for computer components,
2. A production plant for packaging glass and,
3. A transshipment terminal for sea containers.

The EDC is relatively small. The 5 GVs have to move inbound pallets from the receiving lanes to the storage areas. Also during the day, outbound pallets from storage are moved to the shipping lanes. In total, about 600 pallets have to be moved per day. These pallets have to be moved as quickly as possible to serve trucks at the receiving and shipping lanes and keep queue lengths limited. The main objective is therefore to minimize the average pallet waiting time, i.e. the difference between the release time of the pallet until a vehicle picks up the pallet. This can be achieved with a large number of GVs. However, operating GVs is expensive (vehicle plus driver costs) and the objective is therefore to meet the required daily throughput with a minimum number of GVs and minimize the average pallet waiting time.

In the production plant, 11 GVs move pallets coming from the production area on three conveyors to 8 different main storage areas 24 hours a day, 365 days per year. These inbound pallets are released for transport when the glass has cooled down far enough. The pallets should be picked up from the conveyors and stored as quickly as possible since there is limited buffering capacity at the conveyors. Also during weekdays, pallets are moved in batches to be loaded on trucks that park outside at the front of the main storage areas. These trucks that come to pick up a batch of pallets should also be served as soon as possible. In total about 1600 pallets have to be moved per day.

At the transshipment terminal, about 3000 containers have to be unloaded per Jumbo (i.e. very large) container vessels (JCVs) by 6 quay cranes. About 25 ALVs then transport the containers to the main storage area called the stack yard. At the stack yard the stacking cranes transport the containers within the stack. Only when the quay cranes have finished unloading their part of the vessel, they can start loading the vessel with other containers, and the process is reversed. In total about 3000 containers have to be loaded per JCV. The containers have to be moved as quickly as possible since the (quay and stacking) cranes have limited buffer space to set down containers, where the ALVs can pick up and deliver containers. Although the performance of the cranes is not the topic in this study, the cranes

operate in such a way that the average lead-time (time needed to unload and load the vessel) of the vessels is within contract agreements (24 hours).
More detailed characteristics and factors that influence the performance of each environment will be described in much more detail in Sections 5.2, 5.3 and 5.4.

To take care of internal transport in practice, a centralized computer system, (such as a warehouse management system) assigns GVs to loads (or vice versa). These systems keep track of inventory and the movements of loads and vehicles. In many cases, these systems are tailor-made according to the objectives and needs of the environment (see also Section 1.4). For example, when certain vehicles have to take care of all inbound jobs, or loads must always travel in pairs, or idle vehicles must park at a specific place. In each case, the dispatching decisions have to be made on-line, based on real-time events, due to the high degree of stochasticity within each transport environment. For example, a company might know beforehand that a truck will arrive that day to bring or pick up products, but it is not known exactly when the truck arrives or what products are involved (see also Sections 1.1 and 2.7). For the case of transshipment terminals, *much* information of the containers to be unloaded and loaded off and on board vessels is known in advance, but the fact that not all information is known and that the crane and vehicle operation times are very stochastic makes scheduling vehicles beforehand almost impossible. Other reasons such as failure of equipment and avoiding deadlocks have also resulted in the use of on-line vehicle dispatching.

In all three cases, on-line dispatching rules are used to control the vehicles, and the relevant performance criteria studied include meeting the required throughput with a minimum number of vehicles (and vehicle utilization), and minimizing the average load waiting time. We will also investigate whether the same dispatching rules behave similar in different environments or if the relative performances of the dispatching rules depend on the environment. Furthermore, we will investigate the performance of the current dispatching rules used by the different companies and compare them to several standard (common) rules described in the literature. Finally, we will investigate possible performance gains when load pre-arrival information is available and whether this changes the ranking of the dispatching rules. Dispatching vehicles using load pre-arrival information (if available in advance) can lead to considerable reductions in average load waiting times as shown in Chapter 4. In the three cases (and generally in practice), pre-arrival information about load positions and release times could be made available in the case that a load will be dropped off soon by a conveyor, crane, truck, labeling or work station. In this case, the load can send out a signal a few moments before the actual (physical) release time. This extra time can be used to dispatch the vehicles more favorably in terms of load waiting time. On the other hand, in case a vehicle arrives just before the load is actually ready for transport, the vehicle must wait. This vehicle waiting time could have been used to transport another load first, especially if the vehicle waiting time is quite high, i.e. if the pre-arrival time is too high (see Sections 4.2.1 and 4.2.2).
The actual vehicle control systems used at the studied companies only know the distances between different workstations and parking locations (this is generally true for vehicle

control systems in practice). Usually the distances of moving vehicles at any other point on the vehicle network is not known exactly, or not known at all. The lack of exact information is due to the fact that vehicles do not report their position while traveling. Although an estimate of the location could be made using the last point in time that the vehicle passed a station and using the speed of the vehicle. However, this would be a rough estimate since the traveling speed of the vehicle is usually not exactly known, especially taking into account effects such as acceleration, deceleration, congestion, collision control, interference, failure of equipment, etc. Such information is impossible to obtain especially when the vehicles are manned, and directed via vehicle-mounted RF-terminals. With manned vehicles, the speed, acceleration, etc, and even the route the driver takes vary too much to be of any use for the central controller. Furthermore, these same criteria also influence the time still needed to travel the (estimated) distance to the request. Since it is too complex and not realistic to assume that all distance information is available while dispatching vehicles on-line for internal transport in practice, we will not use the dispatching rules based on pre-assigning moving vehicles to loads as described for the theoretical models in Chapter 4.

The studies of the three companies are based on detailed simulation models using data from the company information and control systems and expert judgements. The cases are modeled in the AutoMod™ simulation package (see AutoMod™), which is specialized in (animated) simulation of material handling environments. Simulation is necessary because of the complexity of the models. Due to the stochastic nature of the systems and the many decisions that are real-time based, formulating an analytical model is too complex. Using the simulation models it is possible to implement specific details which would have to be simplified or omitted when the situations are formulated with analytical models. Furthermore, using the animation simplifies the verification of the models and can make possible bottlenecks more apparent.

In the next section we describe the vehicle dispatching rules that will be used in all three case studies to dispatch the vehicles. The following sections describe the case studies, introduce specific vehicle dispatching rules for those companies and present the results found for each case. Since the scale and terminology of the material handling activities at container terminals are different compared to the activities of warehouses and DCs discussed so far, we will also describe the handling activities found at container transshipment terminals and review some studies found in the literature about container terminals in Section 5.4. The general conclusions and similarities between the case specific results will be discussed in detail in the concluding remarks.

5.1 Common dispatching rules for all cases

For the studies in this chapter, we made a selection of three common dispatching rules described in literature (see Chapter 2) which could also be implemented at all three companies using their current vehicle dispatching systems. Two of the rules, the nearest-

vehicle-first and modified first-come-first-served rule were motivated and discussed in the previous chapters and are repeated for the sake of completeness.
The results of the previous chapters indicate that the distance-based dispatching rule outperforms the time-based dispatching rule when there is a diversity of transportation distances (i.e. when the transportation distances between different locations on the network are different). The vehicle traveling networks for the cases studied in this chapter are large and complex, with a diversity of transportation distances. We therefore expect that the distance-based rule will outperform the time-based rule. To further investigate the difference between vehicle-initiated and workcenter-initiated dispatching rules, we will also use a distance-based vehicle-initiated dispatching rule called nearest-workstation-first (NWF).

Nearest-Workstation-First

Under this rule, the vehicle has the dispatching initiative. Using this rule, a vehicle dropping off a load becomes idle and requests the central controller to search for the closest available load (with respect to the position of the idle vehicle) to be transported. The closeness is measured in terms of travel distance. However, a facility layout may contain a few remote stations. These remote stations are not near to other vehicle release points and may therefore never qualify to receive a vehicle dispatch. This illustrates the major drawback of this rule; it is sensitive to the layout of the facilities. It should be made clear that the closest vehicle in distance is not necessarily the closest in travel time. This phenomenon is due to acceleration and deceleration effects, speed restrictions on some paths or variable vehicle speed or a congested travel network. If there are no move requests in the system when the vehicle is looking for work, the vehicle will park at the nearest parking location and becomes idle until a move request becomes available (the central zone positioning rule described in Chapter 2).

Nearest-Vehicle-First

Under this rule, the load or workcenter has the dispatching initiative. When a load is released at a workcenter, the workcenter places a move request. The shortest distance along the traveling paths to every available (idle and motionless) vehicle is then calculated. The idle vehicle, whose travel distance to the load is the shortest, will be awoken to be dispatched. On the other hand, when a vehicle becomes idle without receiving a request, it searches for the closest waiting load in the system, i.e., at that point the dispatching initiative is at the vehicle and the rule used is NWF. If there are no vehicle requests for loads in the system, the (empty) vehicles will park at the nearest parking location and become idle until a request becomes available.

Modified First-Come-First-Served

The modified FCFS rule is a vehicle-initiated dispatching rule. A vehicle delivering a load at an input queue of a station first inspects the output queue of that station. The vehicle is then assigned to the oldest request (longest waiting load) at that station if one or more loads is found. However, if the output queue of that station is empty, the vehicle serves the oldest request in the entire system. If there are no move requests in the system at all, the vehicle will park at the nearest parking location and becomes idle until a move request becomes available.

5.2 Case study of a European Distribution Center

The first case concerns the transportation of pallet loads at the European distribution center (EDC) of a multinational wholesaler in computer hardware and software. This wholesaler distributes computer products to different retail stores in Europe and anticipates how much to purchase and store to be able to comply with the demand of the retailers. Because computer products change quickly over time, it is necessary to keep inventory levels low and the storage times as short as possible. A large part of the incoming products are packed in cartons, stacked per product on pallets. A central warehouse management system (WMS) keeps track of inventory and the position of stored products. The EDC can be divided into several areas (see Figure 22) with a total GV operating area of 40 by 140 meters.

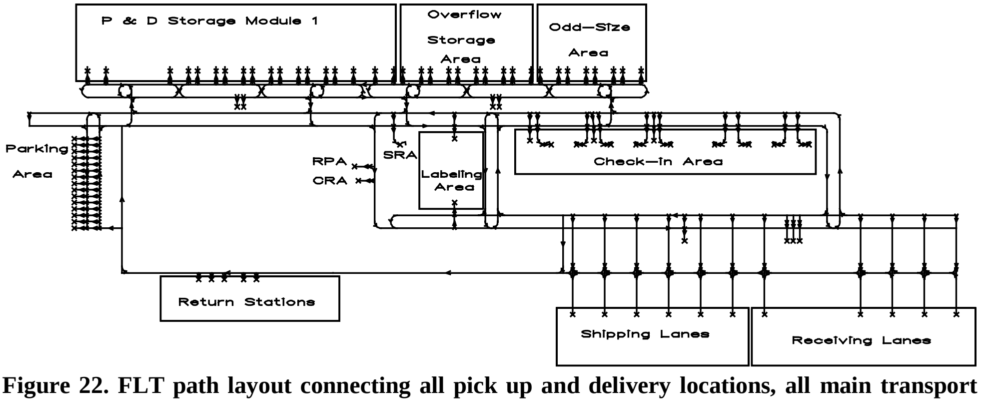

Figure 22. FLT path layout connecting all pick up and delivery locations, all main transport tracks are uni-directional

Each weekday, trucks arrive at the *Receiving Lanes* of the DC where the pallets (loads) are unloaded. In total there are 5 Receiving Lanes. If the cartons on the pallets contain returned or broken products they are manually transported to one of the 5 *Return stations*. The pallets are also manually transported to one of twelve *Check-in Area* stations if the content of the cartons is unclear. At each of the previously mentioned stations, the pallets are labeled with a so-called license plate number (bar code). This license plate number contains information about the content of the cartons and the location the pallet should be brought to. At the moment the license plate is placed on the pallet, the pallet is entered into the WMS.

The other material transport between the different stations of the EDC is taken care of by 5 forklift trucks (FLTs) with vehicle-mounted wireless truck-terminals. If the cartons on the pallet are odd-shaped, or if the pallet is one of many with the same product, it will be transported to the *Odd-Size* or *Overflow Storage Area* respectively. The Odd-Size Storage Area and the Overflow Storage Area have 10 and 8 pick-up & drop (P&D) locations respectively. Otherwise the pallets go to one of the 18 P&D locations of *P&D Storage Module 1*. Within the storage modules, pallets are stored and orders are picked using high-bay orderpicking trucks. From Storage Module 1, pallets can be transported to the Repalletization Area (*RPA*), the Shelf Replenishment Area (*SRA*), the Central Return Area (*CRA*), the *Shipping Lanes* or to the *Labeling Area* (see the material flow diagram in Figure 23). The Labeling Area has one delivery station and one pick up station. RPA, CRA and SRA have one station each, and there are 6 shipping lanes in total (see Figure 22).

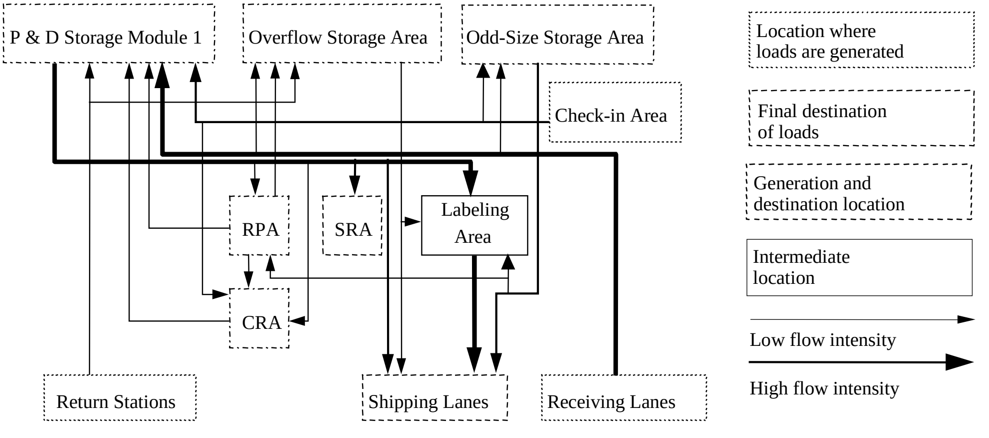

Figure 23. Material flow between all locations of the EDC

From RPA, pallets move to Storage Module 1 or to CRA. At SRA the cartons of the pallets are placed on a conveyor belt, and will be transported to the shelf area where products are handpicked.
Pallets at CRA always move to Storage Module 1. At the Labeling Area, pallets receive customer stickers and packing lists. The Shipping Lanes are the final stations where trucks arrive at dock doors to transport products to retail stores.
The data of the load release times have been measured for a period of 6 weeks. An example data-file from the vehicle-control system of the EDC is presented in Appendix 2. The example printout is an extraction of a list of clock-times at which loads are released to be transported by a certain vehicle at the load-origin and transported to the load-

destination. These files are used to generate the origin-destination matrix shown in Table 32, which represents the average throughput per day from one location to another. The vehicles always use the path with the shortest travel distance when traveling to a pick up or delivery location. Loads are generated at the Return Stations, RPA, CRA, Storage Areas, Check-in Area and Receiving Lanes (see Figure 23). All other locations are end (destination) locations or intermediate locations.

From / To		1	2	3	4	5	6	7	8	9	10	11	**Total**
1	Labeling Area	0	0	159	0	0	0	0	0	0	0	0	159
2	Check-in Area	0	0	0	0	0	0	22	0	0	0	0	22
3	Shipping Lanes	0	0	0	0	0	0	0	0	0	0	0	0
4	Receiving Lanes	0	0	0	0	0	0	109	2	2	0	0	113
5	SRA	0	0	0	0	0	0	0	0	0	0	0	0
6	RPA	0	0	0	0	0	0	9	0	0	0	1	10
7	P&D Storage Module 1	144	0	31	0	17	5	2	0	0	0	0	199
8	Overflow Storage Area	4	0	12	0	0	0	0	0	0	0	0	16
9	Odd-Size Area	11	0	40	0	0	1	0	0	0	0	0	52
10	Return Stations	0	0	0	0	0	0	6	0	0	0	0	6
11	CRA	0	0	0	0	0	0	4	0	0	0	0	4
Total		159	0	242	0	17	6	152	2	2	0	1	**581**

Table 32. Total throughput in pallets per day (obtained from a 6 week period)

The distribution of the interarrival times of the release times between loads have been tested using a χ^2-test The distributions investigated were a uniform, normal, and exponential distribution. It appears that the interarrival times between load releases can be modeled properly using an exponential distribution where the times between releases of a certain loads (interarrival times) for transport depend on the time of day. The interarrival times of loads are independently exponentially generated, where each day is in turn divided into four periods. Period 1: from the start of the day until the coffee break, period 2: from the coffee break until lunch, period 3: from lunch until the tea break, and period 4: from the tea break until the end of the working day. These periods are introduced to realistically represent the variation in the interarrival rates over the day. For example, in period 4 there are more loads transported to the shipping lanes than in period 1. Each type of transport from a certain area is independently generated at its own rate where the interarrival times of the load release times follow a Poisson process.
For each dispatching scenario a new *model-run* is defined and executed. All the parameters are kept the same for each run (unless specified otherwise). These include: the material flow (see Figure 23 and Table 32), the number and locations of loads generated in the system, load generation instants, the speed of the vehicles, vehicle capacity, the paths via which the vehicles may travel (see Figure 22), the pick up and set down time of the load by the vehicle, the number of simulated days and the number of working hours per day. Loads are independently generated until 10 days are expired. The results presented in Section 5.2.2 are the averages obtained over the complete simulation period, these results can be

considered reliable due to the fairly large number of observations (5800 transport movements). Table 33 gives a summary of some other values of the simulation model.

GV speed	2 m/s
Acceleration/deceleration	0.5 m/s^2
Pick up time of a load	15 s
Set down time of a load	15 s
Vehicle capacity (unit-load)	1 load (pallet)
Number of working hours per day	7.5 hours
Simulation period	10 days

Table 33. The parameters used for each scenario

The general aim is to find a control system for the vehicles such that loads are transported on the track layout to and from the correct location with a given flow intensity (see Figure 23), while keeping the number of vehicles (and drivers) as low as possible and keeping load waiting times as short as possible. Short load waiting times, or response times are important to realize due times for the trucks waiting at the Shipping Lanes, keep queues at stations as short as possible, etc. To find the control system capable of this task, the case has been implemented with the common dispatching rules described in Section 5.2 and the case specific vehicle dispatching rules described in the next section.

5.2.1 Case specific dispatching rules for the European Distribution Center

The first described case specific transportation control system of vehicles at the EDC can be classified as *decentralized*, the last as *pre-arrival* and the others as *centralized*. A typical decentralized way of control is a control system where vehicles drive in fixed loops and perform the first transportation task they encounter. The loop configuration that we discuss below has partially overlapping loops, and will be compared with configurations where vehicles are not fixed to a dedicated loop. Initially the loop configuration was to be implemented at the EDC and is therefore subject of investigation. The actual control system implemented using location priority-lists (work-list dispatching) will also be investigated and can be classified as centralized control.

At this company the centralized controller (a WMS) considers all available transportation tasks simultaneously. The central computer uses global information to keep track of real-time activities and assigns idle vehicles to loads (or vice versa) on-line accordingly.

Pre-arrival control is an extension of centralized control in which (some) information on future transport requests is available and can be used to dispatch the vehicles. In Chapter 4, vehicles were scheduled with off-line control and with on-line dispatching rules when load pre-arrival information is available. In the case of the EDC we will also study the performance gain on average load waiting times when load pre-arrival information is available for on-line dispatching.

The idle vehicle positioning or dwell point strategy has been defined by the company as follows: A vehicle should immediately pick up the next available load after dropping off a load, if no load is available the vehicle (driver) is instructed to park at the nearest free parking location (similar to the central zone positioning rule). In Section 5.2.3 we also briefly discuss a hypothetical situation where the vehicles are allowed to park at their last load drop off location (the point of release positioning rule). Although that case is hypothetical (since idle vehicles may block the path for other vehicles, vehicles must recharge at the parking locations, etc.), it can help to indicate the changes in average load waiting times for a new layout where vehicles can park anywhere.
The following sections describe the control systems individually in more detail.

Decentralized control

(a) First-Encountered-First-Served (FEFS)
Under this rule, all the pick-up and delivery locations in the warehouse are divided in two main uni-directional loops (see Figure 24). Loop 1 (bold printed) contains the Return Stations, RPA, CRA and Labeling Area. Loop 2 consists of all stations except the Return Stations. Thus there are two partially overlapping loops for the vehicles. The idea of using the overlapping loop construction is that no use of intermediate or interface stations has to be made, which could normally increase the travel time of the load. Furthermore, this loop configuration has a small partial overlap. Changing the loop configuration or adding more loops increases the partial overlap such that stations which are already covered by a vehicle will be covered by additional vehicles which implies more vehicles and vehicle empty trip time and a reduction in system performance.
Each GV is assigned to a fixed loop, with one GV in loop 1, and 6 vehicles in loop 2; this is done in order to handle the entire throughput acceptably (see the results in Section 5.2.2). The vehicles of one loop are not allowed to pick up pallets in the other loop, i.e. pallets can only be *delivered* at locations of the same or the other loop. A vehicle which delivers a pallet in the other loop immediately returns to the nearest point of its own loop. The vehicles are in this case always in motion, driving in their own loop checking for work at the stations they pass. If there is no work at a particular station, the GV travels to the next station in the loop. If there is work at that location, the load is picked up and transported to its destination (which could be in the other loop). In other words, the vehicles perform the first transportation task they encounter. Hence, the first-encountered-first-served dispatching rule.

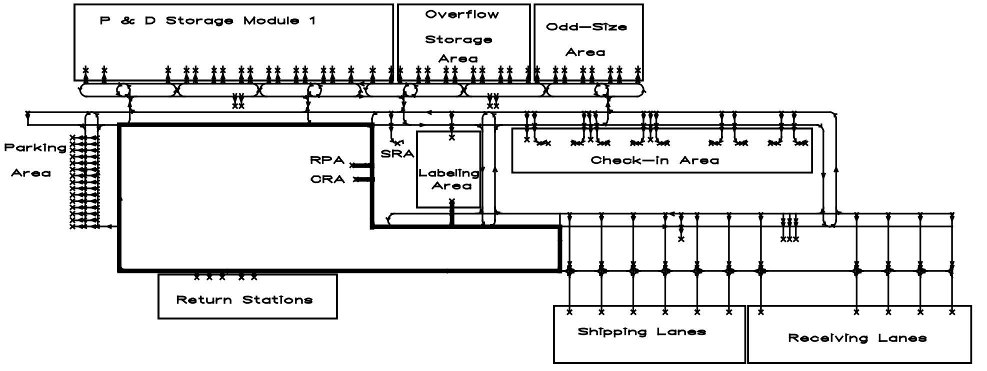

Figure 24. Path layout connecting all pick up and delivery locations. The bold printed paths belong to Loop 1, the other paths to Loop 2. All main transport tracks are uni-directional

The FEFS rule is easy to understand and implement, and is therefore a common way of vehicle-control in warehouse environments. However, it can be very inefficient. If more stations are added, or if the vehicle path layout is changed, then the design of the loops changes also. This means that a new vehicle-to-loop assignment has to be made to balance the performance in the new loops.

To make more efficient use of the vehicle, we introduce case specific vehicle-control systems that use a centralized controller.

Centralized control

(b) Work-List Dispatching (WLD)

This dispatching rule is actually used at the EDC. Using this centralized control system, the vehicles have the dispatching initiative and are in constant communication with a central computer. The central computer keeps track of the pallets and the so-called work-lists (WLs). With WLD it is possible to give priorities to certain locations where loads are to be picked up. Each delivery or drop-off location has a work-list (see Figure 25 for an example). The central controller is triggered to search the work-list of a location when a vehicle visits that location (to park or drop off a load). The WLs contain locations or areas that have to be searched in sequence for loads to be picked up, after the drop-off location has been inspected. If there are no more locations to check on the list, and still no work has been found, the GV is instructed to park at the nearest parking place, and waits until it is called for again.

There are many work-lists at the EDC, a unique one for every drop-off location. For example, at the labeling area, the first search location on the work-list is *labeling area* then *P&D module 1* then the *return stations* etc., at the end of the list *all* remaining stations are checked for possible work (see Figure 25).

Drop-off Area	Shipping Lanes	SRA	RPA	P&D Storage Module 1
Search Areas (in sequence)	Odd-Size Area Overflow Area Check-in stations Receiving Lanes P&D Module 1 Labeling Area ALL	RPA CRA P&D Module 1 Return Stations Receiving Lanes Labeling Area ALL	RPA CRA Return Stations Receiving Lanes P&D Module 1 Labeling Area ALL	P&D Module 1 RPA CRA Return Stations Overflow Area Receiving Lanes Labeling Area ALL

Drop-off Area	Overflow Area	Odd-Size Area	CRA	Labeling Area
Search Area (in sequence)	Overflow Area Odd-Size Area P&D Module 1 RPA CRA Return Stations ALL	Odd-Size Area Overflow Area P&D Module 1 RPA CRA Return Stations ALL	CRA RPA P&D Module 1 Return Stations Receiving Lanes Labeling Area ALL	Labeling Area P&D Module 1 Return Stations Receiving Lanes ALL

Figure 25. Work-lists for all delivery locations for the control system with work-lists

The work-lists are constructed in such a way, that in most cases, the locations around the current position of the GV are checked first for work. Furthermore, the route the vehicles should follow next is consistent (in most cases) with the uni-directional flow of the paths. This reduces the probability of circulating around without a load, to pick up a load that has been made available just ‘behind’ the current location of the idle vehicle.

Although these work-lists are constructed in such a way that the WMS searches for work in neighboring locations, they might not give the best results. Because the Return Stations, RPA and CRA do not appear in every work-list, or appear on the top of the WLs, it is expected that the load waiting times (or pallet response times) will be rather high for these areas. To decrease these pallet response (or waiting) times, the work-lists are updated as described in (d).

(c) Load-List Dispatching (LLD)

A load-list is a list of locations at a load pick-up point. This list is searched to wake up an idle vehicle when a load places a transport request. The newly awoken vehicle triggers the control system to scan the WL of the current vehicle location to find a load transportation request, (i.e. WLD). Since the vehicle scans the work-list, it may find a higher priority load than the load that woke it.

Using this rule the first dispatching initiative lies with the load; however, the vehicle will determine the move request. This rule has been studied to investigate the differences between vehicle-initiated and load-initiated dispatching using priority-lists. If there are no vehicle requests in the system, the (empty) vehicle will park at the nearest parking location and become idle until a request becomes available.

(d) Updated-Work-List Dispatching (UWLD)

This control system is the same as WLD described in (b) above. The difference is that more priority is given to stations where relatively little happens. Due to the structure of the current work-lists, the pallet response times may be high at the CRA, RPA and Return Stations. With WLD, these stations have lower priority since relatively few transport requests occur (see Table 32). The result is that the more busy areas are always checked first for work. Because many pallets need to be moved there, it is probable that the GVs are instructed to leave immediately, and instructions to go to the Return Areas are therefore rare. This may result in high pallet response times at the Return Areas. To investigate this, the work lists are updated (see Figure 26) with those areas placed on top of every work-list (i.e. they now have the most priority).

Drop-off Area	Shipping Lanes	SRA	RPA	P&D Storage Module 1
Search Areas (in sequence)	RPA CRA Return Stations Odd-Size Area Overflow Area Check-in stations Receiving Lanes P&D Module 1 Labeling Area ALL	RPA CRA Return Stations P&D Module 1 Receiving Lanes Labeling Area ALL	RPA CRA Return Stations Receiving Lanes P&D Module 1 Labeling Area ALL	RPA CRA Return Stations P&D Module 1 Overflow Area Receiving Lanes Labeling Area ALL

Drop-off Area	Overflow Area	Odd-Size Area	CRA	Labeling Area
Search Area (in sequence)	RPA CRA Return Stations Overflow Area Odd-Size Area P&D Module 1 ALL	RPA CRA Return Stations Odd-Size Area Overflow Area P&D Module 1 ALL	RPA CRA Return Stations P&D Module 1 Receiving Lanes Labeling Area ALL	Labeling Area P&D Module 1 Return Stations Receiving Lanes ALL

Figure 26. Work-lists for all delivery locations for the control system with updated work-lists

(e) Single-Work-List Flow-intensity-based Dispatching (SWLFD)

To update and maintain all the work-lists is time consuming in practice, especially if stations are added and the number of work-lists increases. To ease the workload of maintaining all the work-lists, and to keep the advantage of a centralized computer control, a new control system with only one work-list in the total system is studied.

The first difference between this control system and the previously described WLs control systems is that there is only one work-list altogether. However, use is still made of a centralized computer for the vehicle control system. Every drop-off location has the same work-list (see Figure 27).

All Stations
CRA
Return Stations
RPA
Overflow Area
Check-in Stations
Odd-Size Area
Receiving Lanes
Labeling Area
P&D Module 1
ALL

Figure 27. The work-list for all delivery locations for single work-list flow intensity based dispatching

The second main difference with the other WLD rules is that the location with the lowest daily outgoing flow intensity (see Table 32) is placed on top of the work-list. The locations are added to the list in ascending order of flow intensity. The idea is that locations with little work will now not be neglected so soon. Because relatively little happens at these locations, they are often skipped quickly and the central computer ends up checking the busy areas anyway. This is a very simple control system, since there is only one list, which is based only on the flow intensity per location or area. It is also rather easy to implement and easily maintained.

Pre-arrival control

(f) Dispatching with Pre-arrival Information (DPI)

This rule uses the common dispatching rules previously described in Section 5.1 and the dispatching rules (a)-(c) described above. The difference is that the load gives a signal x time units in advance of its actual release time. The time between the actual release, and the virtual release x time units before, can be interpreted as a forecast time. This gives the vehicle the opportunity to travel to the load before the load is physically ready for transport. The results of Chapter 4 indicate that the use of pre-arrival information can reduce average load waiting times. However, an increase in load waiting time is also possible.

In the case of the EDC, the cranes in the storage areas for example, can trigger this pre-arrival information of loads about 15 seconds in advance. In the studies with pre-arrival control it is assumed that 5, 10 or 15 seconds of load pre-arrival information is available in advance. Considering that about 581 (see Table 32) loads have to be moved per day (7.5 hours), 15 seconds pre-arrival information is similar to looking $\left(\frac{581 * 15}{7.5 * 3600}\right) = 0.32$ loads ahead on average.

5.2.2 Results for the EDC

The parameters are kept the same in each dispatching scenario in order to make a fair comparison to rank the dispatching rules accordingly (see Section 5.2.3). The only variables in the model are the (logic of the) vehicle dispatching rule, the number of vehicles and the pre-arrival time.
The performance criteria that are most important in the distribution center, and which have been used to evaluate the different dispatching rules are the following:

- The number of vehicles needed to handle the required throughput
- Load waiting times
- Vehicle idle time (or percentage of utilization)
- Maximum number of loads waiting at any time
- Complexity of dispatching rule

Some of these criteria might be contradictory. For example, an extra vehicle can lead to a reduction of load waiting times. A complex dispatching rule could need relatively few vehicles, but use them with full utilization (which is unfavorable in view of long load waiting times, inflexibility and sensitivity to failures of equipment). Therefore, we need to rank the performance criteria. Since investments and operating costs are the most important for the EDC, we will try to keep the number of vehicles to handle the required throughput as low as possible. In first instance, however, to compare the performance between different dispatching rules properly, the number of vehicles used with each dispatching rule is the same. Therefore, the load waiting time is the main performance criterion to determine the rankings. Waiting times should be small, so that delivery schedules are met and queues or buffers do not overflow. The criteria vehicle utilization, maximum number of loads waiting and rule-complexity are used to break ties.

Results Decentral Control System Using FEFS Dispatching

The decentralized First-Encountered-First-Served dispatching rule is the simplest rule described. Using only local information the vehicles continuously move from station to station checking if there is any work. Because the vehicles are always in motion, the average vehicle utilization is 100%, which may lead to problems as soon as a vehicle breaks down. Another shortcoming of only using local information is that no pre-arrival time is available. In total 7 vehicles, one in loop 1 (the bold printed paths in Figure 24) and 6 in loop 2 (the other paths of Figure 24), are needed to handle the entire throughput with an acceptable average total load waiting time. This acceptable level has been set by the management of the EDC at about three minutes. When only 6 vehicles in total are used, the average total waiting time increases to more than four minutes.

Area	Mean
Labeling Area	122
Return Areas	148
P & D Modules	195
Receiving Lanes	116
Check-in Stations	169
Inbound	138
Outbound	166
Total	158
Average Utilization	100 %
Max. nr. of loads waiting	14

Table 34. Average load waiting times per area in seconds, with 7 vehicles and FEFS dispatching.

Table 34 gives a detailed overview of the load waiting times when 7 vehicles are used. In this table, the load waiting times at the P&D Storage Modules, Overflow Area and Odd-Size Area, are grouped under 'P&D Modules'. The pallet response times for the Return Stations, RPA and CRA are grouped together as 'Return Areas'. All average waiting times of the loads leaving the EDC from the P&D storage modules and the labeling area are grouped under 'Outbound'. The waiting times of all other load movements are grouped under 'Inbound'. The maximum number of loads that are waiting for transport at a certain time is 14 (see Table 34).

Results Using Modified FCFS Dispatching

The modified FCFS rule, without the use of pre-arrival information, only needs 6 vehicles to realize an average load waiting time of 116 seconds with a vehicle utilization rate of 69%. To compare the waiting times using the same number of vehicles as the other centralized dispatching rules, the model was recalculated with 5 vehicles, since the other dispatching rules described in the next sections only need 5 vehicles to obtain waiting times less than the acceptable level of about three minutes. Using only 5 vehicles and no pre-arrival information ($x = 0$) the average total load waiting time increases to 194 seconds (see Table 35). The maximum average vehicle utilization of 80% is the maximum utilization measured when the vehicles are dispatched with or without pre-arrival information and represents the percentage of time needed to retrieve and deliver loads. When the pre-arrival time is set to 5 seconds (which corresponds to looking about 0.1 jobs ahead on average), the average load waiting time decreased by 10 seconds. Using a virtual release time 5 seconds before the actual release changed the allocation of vehicles in such a way that the mean waiting time decreased more than proportional. The reverse is also possible which can be seen in Table 35 when the pre-arrival time is set to 10 seconds, i.e. looking 0.2 jobs ahead. In that case the vehicles are allocated unfavorably and the mean waiting time increases relative to $x = 5$ with three seconds.

Pre-arrival time (*x*)	Mean waiting time
0	194
5	184
10	187
15	183
Max. average utilization	80 %
Max. nr. of loads waiting	21

Table 35. Average total load waiting times in seconds with modified FCFS

The maximum number of loads waiting at one time when vehicles are dispatched with or without pre-arrival information is 21. This is 50% more compared to the previous rule. It should be noted, however, that using the FEFS rule 7 GVs were necessary instead of the 5 used now. This is an improvement of 28.6% for the number of vehicles needed and illustrates the more efficient use of information of the central systems compared to the decentral control systems.

Results Using Work-List Dispatching

The WLD rule without pre-arrival time uses 5 vehicles to realize an average load waiting time of 169 seconds with a vehicle utilization rate of 80% (see Table 36). When pre-arrival time is available, the average load waiting time monotonically reduces with three seconds (when x = 5 seconds) to 13 seconds (when x = 15 seconds).

Pre-arrival time (*x*)	Mean waiting time
0	169
5	166
10	160
15	156
Max. average utilization	80 %
Max. nr. of loads waiting	20

Table 36. Average total load waiting times in seconds with WLD

The average load waiting times per area in Table 37, compared to the waiting times with FEFS in Table 34, show that dedicating a vehicle to loop 1 can reduce the average load waiting times of the Return Areas by about 100 seconds. Furthermore, the waiting times at the receiving lanes were considerably less with FEFS as well. However, the WLD rule uses about 29% less vehicles compared to the FEFS rule, while the average total load waiting times remain comparable.
Compared to the modified FCFS rule, the WLD rule has almost same maximum number of waiting loads (20), and is more efficient with respect to the load waiting time. The waiting times using WLD decreases by about 30 seconds.

Area	Mean (x = 0)
Labeling Area	112
Return Areas	249
P & D Modules	127
Receiving Lanes	350
Check-in Stations	116
Inbound	280
Outbound	122
Total	169

Table 37. Average load waiting times per area in seconds with WLD

Results Using Updated-Work-List Dispatching (UWLD)

Although the average total load waiting time for the work-lists control system (without pre-arrival information) is less than the waiting time of the decentralized FEFS system, the average load waiting time at the Return Areas has increased. This was expected, because the positions of the Return Areas are not very high on the work-lists, if they are on the lists at all (see Figure 25). The result is that other locations are given more priority and the locations in the Return Areas are often neglected.

The waiting times of the loads when the updated work-lists are used are presented in Table 38. As can be seen in the last row, the average load waiting time is comparable with the previous control system (see Table 37) with a maximum of 20 loads waiting to be picked up at a certain point in time. As expected, however, the response time at the Return Areas has decreased considerably, from 249 seconds to 102 seconds. Giving more priority to the Return Areas in the work-lists, has a similar effect for the waiting times of the Return Areas as dedicating a vehicle to loop 1 which includes the stations of those areas.

Area	Mean
Labeling Area	115
Return Areas	102
P & D Modules	128
Receiving Lanes	340
Check-in Stations	117
Inbound	265
Outbound	123
Total	166
Average Utilization	80 %
Max. nr. Of loads waiting	20

Table 38. Average load waiting times per area in seconds with UWLD

Another result of giving the Return Areas more priority is a slight increase in the load response time of some other areas. However, this is only a few seconds and the average total load waiting times remain comparable to the waiting times with WLD. The vehicle utilization is about 80% as well.

Results Using Single Flow-intensity-based Work-List Dispatching

This control system makes use of only one work-list (see Figure 27), to which the control system refers every time a vehicles becomes idle (i.e. after delivering a load). In total there are still only 5 vehicles necessary. The maximum number of loads waiting to be picked up at a certain point in time is only 12, the lowest result compared to all other dispatching rules. The results of SFWLD are presented in Table 39. As can be seen, the average load waiting times of the individual areas, except for the waiting time at the P&D modules, are comparable to, or less than, that of the previous control systems (see Table 34, Table 37 and Table 38). However, the waiting time at P&D Modules has increased. This is no surprise because these stations are given the least priority. This is due to the fact that most of the material flow takes place here, and so they are at the bottom of the flow intensity based work-list. This results in a slight increase of the average total waiting time of about 10 seconds.

Area	Mean
Labeling Area	106
Return Areas	107
P & D Modules	256
Receiving Lanes	122
Check-in Stations	101
Inbound	133
Outbound	201
Total	180
Average Utilization	80.5 %
Max. nr. of loads waiting	12

Table 39. Average load waiting times per area in seconds with SFWLD

Due to the simplicity of this control rule and the reasonable results for the average waiting times, especially since only 5 vehicles are needed, it may be said that this control system scores well, and is comparable to the results obtained with the WLD rule.

Results Using Load-List Dispatching

The LLD rule without pre-arrival information ($x = 0$) uses 5 vehicles to get an average total load waiting time of 166 seconds (see Table 40). The results of this rule are about the same as the WLD rule; the maximum vehicle utilization rate calculated for all pre-arrival times is 79% and a maximum of 20 loads are simultaneously waiting to be picked up. In this respect it is difficult to draw conclusions whether the vehicle-initiated WLD rule or the load-initiated LLD rule leads to a better performance. However, as the pre-arrival time increases, the LLD rule tends to lead to more favorable average load waiting times than WLD. This can be seen by comparing Table 36 with Table 40.

Pre-arrival time (*x*)	Mean waiting time
0	166
5	161
10	155
15	149
Max. average utilization	79 %
Max. nr. of loads waiting	20

Table 40. Average total load waiting times in seconds with LLD

Results Using Nearest-Workstation-First Dispatching

Dispatching 5 vehicles with the NWF rule without pre-arrival information results in an average total load waiting time of 134 seconds, with a maximum vehicle utilization rate of 78% (see Table 41). Although this is a moderately simple dispatching rule, its performance with respect to the number of vehicles needed, average load waiting time and complexity is the best until now. Even without the use of pre-arrival information, the load waiting times are about 30 seconds lower then when WLD is used. Compared to modified FCFS the average total load waiting time decreases even by one minute.

Pre-arrival time (*x*)	Mean waiting time
0	134
5	129
10	126
15	121
Max. average utilization	78 %
Max. nr. of loads waiting	14

Table 41. Average total load waiting times in seconds with NWF dispatching

Another improvement is shown in the maximum number of loads waiting for transport at a certain time instant. This maximum is 12 loads when the pre-arrival time is 0 or 5 seconds (not in the Table) and a maximum of 14 loads when the pre-arrival time is 10 or 15 seconds.

Results Using Nearest-Vehicle-First Dispatching

The last distance-based control system is the load-initiative nearest-vehicle-first (NVF) dispatching rule. When only four vehicles are used, the average vehicle utilization is about 90% and the average load waiting times are about 5 minutes. When 5 vehicles are used to handle the entire throughput, the average load waiting times decrease below the acceptation level of about three minutes since dispatching without pre-arrival information results in an average total load waiting time of 129 seconds (see Table 42). Just like the modified FCFS rule, an increase in the pre-arrival information time does not mean a decrease in the average load waiting time. When the pre-arrival time is 5 seconds ($x = 5$), the decrease in total waiting time is also 5 seconds. However, when x is set to 10 seconds, the vehicle allocation is changed in such a way that it is less favorable than when x was set to 5 seconds. The average load waiting times decrease again when the pre-arrival time is 15 seconds.

Pre-arrival time (x)	Mean waiting time
0	129
5	124
10	126
15	123

Table 42. Average total load waiting times in seconds with NVF dispatching

The results are similar to NWF where the vehicle had the initiative, although in this case the maximum number of loads waiting for transport is 15. Table 43 shows that there are no areas with exceptionally high peaks in average load waiting times. The balanced waiting times show a similar pattern as when vehicles are dedicated to a loop with the FEFS dispatching rule. Note however, that with the FEFS rule, 7 vehicles were used, resulting in an average waiting time of 158 seconds.

Area	Mean ($x = 0$)
Labeling Area	92
Return Areas	145
P & D Modules	140
Receiving Lanes	153
Check-in Stations	111
Inbound	145
Outbound	122
Total	129
Average utilization	76 %
Max. nr. of loads waiting	15

Table 43. Average load waiting times per area in seconds with NVF dispatching

5.2.3 Ranking dispatching rules for the distribution center

In the previous section we investigated the performance of a number of different vehicle dispatching rules for a particular European distribution center. These rules include one decentralized dispatching rule with two partially overlapping loops based on first-encountered-first-served (FEFS) dispatching, and several centralized knowledge-based

dispatching rules. With the performance results we classify (or rank) these rules to see which rule can be considered best given the circumstances. The main performance criteria we look at include: the number of vehicles to adequately handle the throughput, the average time a load has to wait for transport (load waiting time), the utilization rate of the vehicles, the maximum number of loads waiting at a certain time, and the complexity of the dispatching rule.
Moreover, we will look at the rank of the rules when using pre-arrival information for the loads. From Chapter 4 we know that using pre-arrival information makes it possible to improve the performance even further. However, the pre-arrival time should not be too long in a stochastic environment or the information may become unreliable.

Table 44 gives a summary of the results obtained. Dashed lines are drawn where the difference between the average load waiting times of two successive rules is greater than 5% or when the number of vehicles needed changes. The results show that less GVs are needed with centralized control in order to obtain comparable load waiting times with the decentralized control system. In total, the conventional decentralized system needs 40% more GVs. In view of complexity it should rank amongst the best, however this rule has a few disadvantages. First of all, it makes use of local information only. The vehicles drive around in loops until they actually bump into work. This has as effect that the vehicle utilization is 100%. In case a vehicle breaks down in loop 1, then the buffers of the workstations in that loop will overflow. A vehicle break down in loop 2 will result in more work for the remaining vehicles and the average load waiting times will increase. A reduction of the number of vehicles to 6 (one in loop 1 and five in loop 2), increases the load waiting time to 258 seconds. Although this rule is the simplest of all, it will be ranked lowest due to the large number of vehicles needed and high vehicle utilization rate. Decentralized vehicle control is thus outperformed by centralized control.
Within the class of centralized control systems, improvements in performance can still be realized, this is represented by the ranking. The centralized control rule ranked lowest is modified FCFS (see Mod. FCFS in Table 44), even though it is in complexity the simplest centralized rule and has about the same vehicle utilization and maximum number of waiting loads as the work-lists and load-lists based rules. However, the average load waiting time is about 30 seconds higher on average. If 6 vehicles were used, the load waiting time would reduce to 116 seconds. This would be better than NVF, but for this performance improvement an extra vehicle is needed.

Table 44 shows a remarkable result with modified FCFS. Using 5 seconds of pre-arrival information changes the allocation of vehicles in such a way that the mean waiting time decreases more than proportionally. The reverse is also possible. When the pre-arrival time is changed to 10 seconds, the mean waiting time increases by three seconds with respect to the waiting time when 5 seconds of pre-arrival information is available, (a similar effect is shown by NVF).

Dispatching rule	Initiative	Number of vehicles	Waiting time in sec. ($x = 0$ / $x = 5$ / $x = 10$ / $x = 15$)	Vehicle utilization	Max. number of loads waiting
NVF	Load	5	129 / 124 / 126 / 123	76 %	15
NWF	Vehicle	5	134 / 129 / 126 / 121	78 %	14
LLD	Load	5	166 / 161 / 155 / 149	79 %	20
WLD	Vehicle	5	169 / 166 / 160 / 156	80 %	20
Mod. FCFS	Vehicle	5	194 / 184 / 187 / 183	80 %	21
FEFS	-	7	158	100 %	14

Table 44. Summary of results, the ranking of the various dispatch rules

The next two rules, LLD and WLD are very similar. Both make use of priority lists. However, with LLD the workcenter scans the load lists to wake a vehicle which then scans its work-list, and with WLD the vehicle has the initiative and scans the work-lists to claim a load. It is difficult to say which of the two is best. Although both are practically the same for all criteria, LLD has slightly more favorable waiting times. In any case both are classified below NVF and NWF, which means that the performance can still be increased.
Both dispatching rules on top are considered best for the EDC, although the average load waiting times of NWF are slightly higher than NVF rule when the pre-arrival information is (x =) 0 or 5 seconds. Notice that these two rules are practically the same, except that NWF is a vehicle-initiated rule and NVF is a load or workcenter-initiated rule.

In Section 5.2.1 we mentioned a hypothetical dwell point strategy where the vehicles are allowed to park at their last load drop off location. In the situation where vehicles moving to a parking location must reach the parking point before becoming eligible to pick up a load, it may pass by waiting loads that are released during the time that the vehicle needs to reach its parking destination. When vehicles can park at their last release point, they become idle and are available immediately for the next transport request. This strategy would be similar to assigning loads to moving vehicles, except that the problems with the position, distance or time to moving vehicles do not have to be considered. Although the case is hypothetical in practice since idle vehicles may block the path for other vehicles, vehicles must recharge at the parking locations, vehicles must park at certain locations for safety reasons, change drivers, etc., it can help to indicate the changes in average load waiting times for a new layout where vehicles could park anywhere or using dispatching rules that can consider moving vehicles. The result of this strategy with the NVF rule indicates that the average load waiting time can decrease by about 10%.
In conclusion, it can be said that the centralized systems outperform the decentralized system and the systems using any kind of pre-arrival information can outperform the standard centralized dispatching rules. Within the class of centralized rules there is also a ranking of dispatching rules, although the performance when dispatching vehicles using work-lists is almost insensitive to the structure of the lists. But a simple rule like modified FCFS that looks for the first load in time can be outperformed by a moderately simple rule like NWF that looks for the closest load in distance. However, the results of Chapter 3 and 4 have shown that distance-based rules would not perform as well when all travel distances between locations are the same, but this is not the case of the EDC. Furthermore, with

respect to load waiting time, load-initiative distance-based rules perform slightly better than vehicle-initiative distance-based rules. This is due to the fact with vehicle-initiated rules, the vehicles may not be able to claim a load when it is too far away while a closer alternative load is available.

5.2.4 Varying fleet size and batch arrivals of loads

In the previous sections, we modeled the EDC to look at how different vehicle dispatching rules can be classified, and how the performance of internal transport is affected if pre-arrival information of the release time of loads is available. The model will now be extended.
When a truck with loads arrives at the receiving area of a warehouse, the loads are released for transport in batches, rather than one at a time. The release of loads in batches at the receiving area is more realistic than loads being released one by one. In practice, when a truck arrives to deliver pallets with loads, the data of the pallets are entered into the WMS in small groups (batches) and the release of the loads to the transport system follows in a similar fashion. When loads are released for pick-up in batches (often two or three at once), and uni-load vehicles are used, one or more loads will be left behind, which will *increase* the average load waiting times. On the other hand, using multi-load vehicles, transportation jobs can be combined which will *decrease* the average load waiting times. However, by combining load transports, the average load *transportation* time will increase since some loads remain relatively longer on the vehicle (while being delivered). We will therefore also look at the effects of the load throughput time, i.e. the load waiting time plus the load transportation time.
Furthermore, we will combine batch releases of loads and the use of multi-load vehicles and classify the dispatching rules accordingly.

The model of the EDC has been adapted to generate batch arrivals of loads at the receiving lanes and check-in area, and the use of multi-load vehicles. We would like to see which dispatching rule gives the best performance and if this is consistent with earlier findings.

Generating batch arrivals of loads

The release times will still be generated following a Poisson process. However, when more than one load is generated simultaneously, the period between load generations increases in proportion. So if one load is generated every t time units, then n loads are generated every $n*t$ time units. Loads are generated and released in this way at the receiving lanes and check-in area. At these locations the batch size is one, two or three loads. For a model-run the batch size is either one, two or three, i.e. *all* load-generations for loads released for transport at those locations for that model-run are in batch sizes of one, two or three loads respectively.

Dispatching rules using multiple-load vehicle capacity

An alternative to increasing the number of vehicles needed to handle the transport requests is to use vehicles with multiple-load capacity (see Chapter 3). In practice, multiple-load vehicle systems are not very common. The capacity rarely exceeds two loads. In Chapter 2 and 3 we discussed some advantages of using multi-load vehicles, such as:

- combining transportation jobs to reduce load waiting times
- fewer vehicles are necessary to handle the required throughput, which can
- improve traffic efficiency
- lower average load waiting times

Some of the disadvantages of using multi-load vehicles included:

- vehicles are more expensive
- more maneuvering space for the (larger) vehicles is needed, which increases (storage) costs
- vehicle dispatching rules are more complex
- higher average load transportation times

Multi-load vehicle dispatching is based on the concept of *closest task* as described in Section 3.1.4. Since multi-load vehicle dispatching rules are based on the concept of closest task, the behavior of a multiple-loaded vehicle and the assignment of partially loaded vehicles to additional loads (or vice versa) is similar for all dispatching rules.
When vehicles are dispatched according to the decentralized FEFS rule, an extra instance concerning the vehicle capacity is evaluated. With the decentralized control rule, the layout of the EDC is divided in two (overlapping) loops (see Figure 24). So next to the scenarios where all vehicles have capacity 1 or 2 (or even 3), an additional scenario is evaluated where the vehicle of loop 1 (the smallest loop), has capacity 1 and the other 6 vehicles in loop 2 have capacity 2.

5.2.5 Results

As in Section 5.2.2, the parameters are kept the same in each dispatching scenario (model-run) in order to make a fair (rank) comparison between the dispatching rules (see Section 5.2.6). The only differences are the arrivals of loads in batches of 1, 2 or 3 and the vehicle capacity of 1 or 2 (or 3) loads depending on the scenario.
We have seen that the rank of the dispatching rules is highly dependent on the number of vehicles needed and the load waiting times. In the case of multi-load dispatching, the load

waiting times can reduce while the load transportation times can increase. We will therefore substitute the performance criterion ‘average load waiting times’ with ‘average load throughput times’. So the performance criteria we look at in this case to evaluate the robustness of the different dispatching rules are the following:

- The number of vehicles needed to handle the required throughput
- Average load throughput time (= average load waiting time + transportation time)
- Vehicle idle time (or percentage of utilization)
- Complexity of dispatching rule

The utilization of a (multi-load or uni-load) vehicle is calculated by adding the percentage of time used for delivering and retrieving loads. Another way of calculating this is by taking the percentage of idle time (i.e. the percentage of time used for going to the parking location and parking) from the percentage of total time available (100%). This means that multi-load vehicles with only one load are also considered as fully utilized during that trip.

Table 45 explains how the results are tabulated for the different dispatching rules in the following sections. For example, when loads are released in batches of two and the vehicle capacity is one, one should look at the cell ‘Batch 2/Capacity 1’. In total there are four statistics in this cell.
The first statistic in the cell is the average total load throughput time in seconds. The second statistic of the first row of the cell is the average load waiting time. This statistic is also reported to study the effects of multi-load vehicles on the average load waiting times and the proportional changes in load transportation times.
The second row of the cell states the number of vehicles necessary to handle the required load throughput. The last statistic in the cell represents the percentage of vehicle utilization. This is usually between 65 and 85% except for the FEFS rule where the vehicles never park and have a utilization of 100%.

Batch \ Capacity	1	
2	Average load throughput time Number of vehicles	Average load waiting time Vehicle utilization (%)

Table 45. Explanation of result tables

Results Using FEFS Dispatching

The decentralized first-encountered-first-served dispatching rule is the simplest of all. Because this control system makes use of loops, it is possible to make a distinction between the vehicles by giving them different vehicle capacity. An extra column has been added to Table 46 where loop 1 has one uni-load vehicle and loop 2 has 6 dual-load vehicles. Because the vehicles are always in motion, the utilization is 100%. As expected, the average load waiting times decrease as the vehicle capacity increases, and increases as the batch size increases. Waiting times increase with batch releases because one or more loads have to wait longer when only one load is picked up if uni-load vehicles are used. Increasing the load capacity of the vehicle can compensate this effect. The vehicle can then carry more loads at a time and the average load waiting time decreases, (which reduces the throughput time as well). When batch releases and the use of multi-load vehicles are combined (the diagonal movement from cell 1:1 to cell 2:2), the average throughput time also decreases. This can be explained intuitively by the fact that the arrival of loads changes but the number of loads stays the same and the number of transportation units increases. Thus the system behaves as if the *unit-load* changes from one pallet to two pallets.

Notice, that the difference between the average load throughput times and average waiting times for uni-load vehicles is about (269 - 158 =) 111 seconds. This difference is the average load transportation time, which increases (to about 130 seconds), as the capacity of the vehicles increase.

Batch \ Capacity	1		2		1 for Loop 1 2 for Loop 2	
1	269 1+6	158 100	251 1+6	127 100	251 1+6	130 100
2	279 1+6	168 100	259 1+6	131 100	260 1+6	135 100
3	287 1+6	177 100	260 1+6	132 100	259 1+6	134 100

Table 46. Results First-Encountered-First-Served dispatching

When only loop 2 is provided with dual-load vehicles, the performance is almost the same as when all vehicles have dual-load capacity. This means that a single uni-load vehicle in loop 1 is sufficient to handle the entire throughput in that loop, which would therefore be the cheaper option for using multi-load vehicles. The fact that a single uni-load vehicle in loop 1 is sufficient is not really a surprise, since there are no batch arrivals generated for the stations in that loop. Therefore, the relative benefit of using a dual-load vehicle to combine transportation requests in that relative low throughput area is negligible.

Results Using Modified FCFS Dispatching

Although the positive effects of adding vehicle capacity and the negative effects of batch releases of loads show the same trends as with FEFS dispatching on the average load waiting times and throughput times, the relative difference is greater. Adding capacity (when 5 vehicles are used) leads to a reduction in average load throughput time of 21% and more, while increasing the batch size from one to three leads to an increase of nearly 13% for the case of uni-load vehicles.

Batch \ Capacity	1		1		1		2	
1	305 5	194 80	228 6	116 75.7	209 7	97 59.1	242 5	106 75.7
2	317 5	205 82.2	242 6	130 70.0	214 7	102 60.4	251 5	112 76.1
3	344 5	234 73.4	246 6	134 63.7	221 7	109 55.2	254 5	117 68.7

Table 47. Results modified First-Come-First-Served dispatching

The increase in throughput times, when increasing the batch size, is smaller for dual-load vehicles than for uni-load vehicles. Furthermore, the results in Table 47 show a decrease in the relative reduction in average load throughput and waiting times when the number of vehicles increases. Also, the vehicle utilization decreases as the fleet size increases. This can be explained by the fact that transportation jobs are combined when dual-load vehicles are used and vehicles are left with relatively more idle time. Similarly, when the number of vehicles increases, the same amount of work is balanced over relatively more vehicles, thus reducing the average vehicle utilization. In order to obtain a comparable throughput time between situations of dual-load vehicles and uni-load vehicles, it appears that for all batch sizes nearly 6 uni-load vehicles are necessary to yield the same performance as 5 dual-load vehicles. So one could say here, the performance of one dual-load vehicle is about the same as 1.2 uni-load vehicles.

Results Using Work-List Dispatching

The results with WL dispatching (see Table 48) are almost identical with modified FCFS dispatching when the capacity is two (and 5 vehicles are used). The mean reason is that the multi-load vehicles are dispatched similarly, i.e. based on the concept of closest task. When the vehicle capacity is one, the results with WLD are more favorable except that the relative increase of load throughput time is larger when the batch size changes from one to three. The location based WLD rule therefore outperforms the time based modified FCFS rule. This is consistent with earlier mentioned results.

Batch \ Capacity	1		2	
1	281 5	169 79.7	242 5	107 75.7
2	290 5	179 81.7	246 5	112 75.2
3	322 5	213 72.4	251 5	116 67.8

Table 48. Results Work-List dispatching

Results Using Load-List Dispatching

The results of LL dispatching are practically identical to those of WL dispatching when 5 vehicles are used (see Table 49). This is consistent with the results of in Section 5.2.2. The performances of both rules are very similar and when multi-load vehicles are used, both rules have about the same level of complexity. So LLD dispatching is preferred, due to the fact that the average load waiting times are slightly more favorable. The results show further that the number of uni-load vehicles has to increase to 6 to obtain a similar average waiting (and throughput) time as 5 dual-load vehicles.

Batch \ Capacity	1		1		1		2	
1	278 5	166 78.5	215 6	103 65.9	196 7	84 55.6	236 5	102 73.3
2	289 5	177 81.4	226 6	114 68.2	204 7	92 57.8	242 5	108 74.0
3	313 5	203 72.2	235 6	123 72.1	211 7	99 53.2	248 5	113 67.7

Table 49. Results Load-List dispatching

Results Using Nearest-Workstation-First Dispatching

The NWF results with dual-load vehicles in Table 50 are comparable to the results of the other dispatching rules. However, when the vehicle capacity is one and the number of vehicles is 5, the load throughput times show a noteworthy reduction. Even the worst result (for load throughput time) when the batch size is three and the vehicle capacity is one is better than the best result of LLD dispatching (see Table 49). The distance-based NWF rule therefore outperforms the previous location-priority based rules.

Batch \ Capacity	1		2	
1	246 5	134 76.7	234 5	101 74.3
2	257 5	146 79.4	241 5	106 74.4
3	275 5	165 71.5	245 5	111 67.0

Table 50. Results Nearest-Workstation-First dispatching

Results Using Nearest-Vehicle-First Dispatching

The results of the NVF rule are slightly better than the results obtained with NWF dispatching. Table 51 shows that the average load waiting time indeed decreases as the capacity of the vehicle increases. However, the advantage decreases as the vehicle capacity increases, in this case even triple-load vehicles are investigated. The results also show that the load transportation time, i.e. the difference between the throughput time and the waiting time, increases as the vehicle capacity increases. When uni-load vehicles are used, the vehicle transportation time is about 111 seconds. When dual-load vehicles are used the transportation time increases by 18.9% to (229 - 97 =) 132 seconds. However, the average load throughput times decreases, this makes clear that the reduction in load waiting time outweighs the increase in transportation time, (which leads to the reduction in the average load throughput time).
The results of dual-load and triple-load vehicles are rather similar with respect to the average throughput time and vehicle utilization. The average load waiting time with triple-load vehicles are relatively more favorable. However, the throughput time increases as the vehicle capacity increases from two to three. This is due to the increase in load transportation time to about (232 - 92 =) 140 seconds. So the relative increase in average load transportation times outweighs the relative decrease in average load waiting times. It is therefore not favorable in this case to use multi-load vehicles with capacity 3 at the EDC.

Batch \ Capacity	1		1		2		3	
1	241	129	208	96	229	97	232	92
	5	77.1	6	63.9	5	72.8	5	72.0
2	258	147	217	105	236	102	237	96
	5	76.8	6	67.1	5	73.2	5	73.2
3	271	161	225	113	241	107	242	101
	5	70.3	6	70.9	5	67.3	5	66.2

Table 51. Results Nearest-Vehicle-First dispatching

5.2.6 Ranking dispatching rules for the EDC with varied fleet sizes

In the previous section we looked into the results of several dispatching rules when loads are released in different batch sizes and vehicles have multi-load capacity. Table 52 gives a summary of the results obtained. Dashed lines are drawn where the difference between the average load throughput times obtained with uni-load vehicles (Cap. = 1) of two successive rules is greater than 5% or when the number of vehicles needed changes. The dashed lines are extended for dual-load vehicles (Cap. = 2) although the relative differences in the average throughput times are less noticeable. The rank of the (multi-load) dispatching rules with batch arrivals of loads with respect to the average load throughput time remains the same as their (uni-load) counterpart without batch arrivals described Section 5.2.3.

The decentralized FEFS rule still requires more vehicles relative to any of the centralized rules even when multi-load vehicles are used. This is due to the poor use of information of the location and status of vehicles and loads, and the fact that vehicles are dedicated to a loop.

Within the group of centralized dispatching rules, three subgroups can be defined, e.g.: time-based dispatching, priority-based dispatching and distance-based dispatching. Time-based dispatching, represented by modified FCFS, performs the least well of the centralized rules. However, when multi-load vehicles are used, it becomes comparable with the other rules. This is actually no surprise because the dispatching rules of multi-load vehicles are more or less similar; based on the concept of closest task.

Dispatching rule	Initiative	Number of vehicles	Load throughput time in sec. (Cap. = 1, Batch = 1 / 2 / 3)	Load throughput time in sec. (Cap. = 2, Batch = 1 / 2 / 3)
NVF	Load	5	241 / 258 / 271	229 / 236 / 241
NWF	Vehicle	5	246 / 257 / 275	234 / 241 / 245
LLD	Load	5	278 / 289 / 313	236 / 242 / 248
WLD	Vehicle	5	281 / 290 / 322	242 / 246 / 251
Mod. FCFS	Vehicle	5	305 / 317 / 344	242 / 251 / 254
FEFS	-	7	269 / 279 / 287	251 / 259 / 260

Table 52. Summary of results, the ranking of the various dispatching rules, with respect to average load throughput times and the number of vehicles needed

The location-based rules are subdivided in a vehicle-initiative rule, represented by WLD, and a load-initiative rule, represented by LLD. Although there is little difference, the results with LLD are slightly more favorable than WL dispatching and are therefore ranked higher. The difference between the performance of modified FCFS and WLD (more than 20 seconds waiting times) is similar to the difference between LLD and NWF.

The distance-based rules are also subdivided in a vehicle-initiative rule, represented by NWF and a load-initiative rule, represented by NVF dispatching. Again there is little difference, but NVF dispatching has more favorable results and is ranked higher than NWF dispatching.

For a fixed batch size, increasing the capacity of the vehicle leads to a reduction of the average throughput time (see Table 52). The magnitude of the reduction is stronger for larger batch sizes. This is intuitive, as the opportunity for combining loads with multi-load vehicles increases for larger batch sizes. However, in Table 51 we have seen that increasing the capacity of the vehicles to three works contra-productively (the average load throughput times increases again).

In conclusion, a more realistic model taking into account batch release of loads (in batch size 2 or 3) at the receiving lanes and check-in area, increases the average load throughput time by 15% for uni-load vehicles and about 5% for multi-load vehicles. Using dual-load vehicles, the average load waiting times can decrease by about 35% (see the detailed results in Section 5.2.5). Furthermore, about 20% more uni-load vehicles are needed to yield approximately the same results as with dual-load vehicles. This means that the costs of a dual-load vehicle should be less than 20% higher than those of an uni-load vehicle in order to be cost effective when the objective is to reduce the average load throughput times without increasing the number of vehicles. This percentage is situation dependent. In Chapters 3 and 4 we have seen that the performance is sensitive to the utilization of the vehicles. The higher the vehicle-utilization, the more effect multi-load vehicles will have and the sooner the costs for multi-load vehicles can be justified.

5.3 Case study of a glass production plant

The second case concerns the transportation of pallet loads at a production plant of packaging glass. At the production area of the glass plant, trucks arrive with cullet (recycled glass), sand, soda ash and lime stone. These raw materials are proportionately mixed into batches and melted in a furnace at a temperature of about 1580° C. The melted mass is then cut into drops which are automatically molded into the desired packaging shapes (bottles, jars, etc.). The glassware is then cooled in a carefully controlled process and rigorously inspected to ensure that the dimension and strength meet specification.
The glassware is stored after production at the site until the clients (other companies that fill the glassware) collect the products for their own use. About 400 different glassware products, varying from jars to bottles, are produced. With three glass melting ovens and 9 production lines, 9 different glassware products are produced simultaneously, 24 hours per day, 365 days per year. After production and inspection, the glassware is automatically stacked on pallets, which are then wrapped in plastic foil and moved by three pallet-conveyors to the 'landing zone', located in one of the storage areas, see Figure 28. Each conveyor has a buffer capacity of about 6 pallets.
The GVs (dual-load RF-guided FLTs) then transport two pallets at a time, which arrive at the *conveyors* in pairs, to one of the 8 storage areas (denoted by S1 through S8 in Figure 28). The 8 main storage areas have a total storage space of 55000 square meters where the pallets can be stacked three pallets high in block stack. The total GV operating area, represented by Figure 28, is 315 by 540 meters. Some pallets that arrive at the 'landing zone' are transported back to the beginning of the conveyor to be re-foiled with plastic. On very rare occasions (at the start of a new production-run), some pallets are transported from the 'landing zone' to the *Crush* area. The glassware of these 'start of the run pallets' may contain small defects and are destroyed (crushed) as a precaution.
In some instances, the glassware is transported from the storage areas to the value added logistics (*VAL*) area. At the VAL area, the glassware is wrapped with a customized sleeve (a label on a jar or a bottle). After customization in the VAL area, the glassware is transported by the conveyors back to the 'landing zone'.
The stored glassware can only be collected during weekdays. The collectors (customers) park their trucks in front of the appropriate storage area where the truck is loaded by the GVs.

The pallets are always moved in pairs and a transportation request is always for two loads. This means that the *unit-load* at the glass production plant can be considered as one dual-pallet.

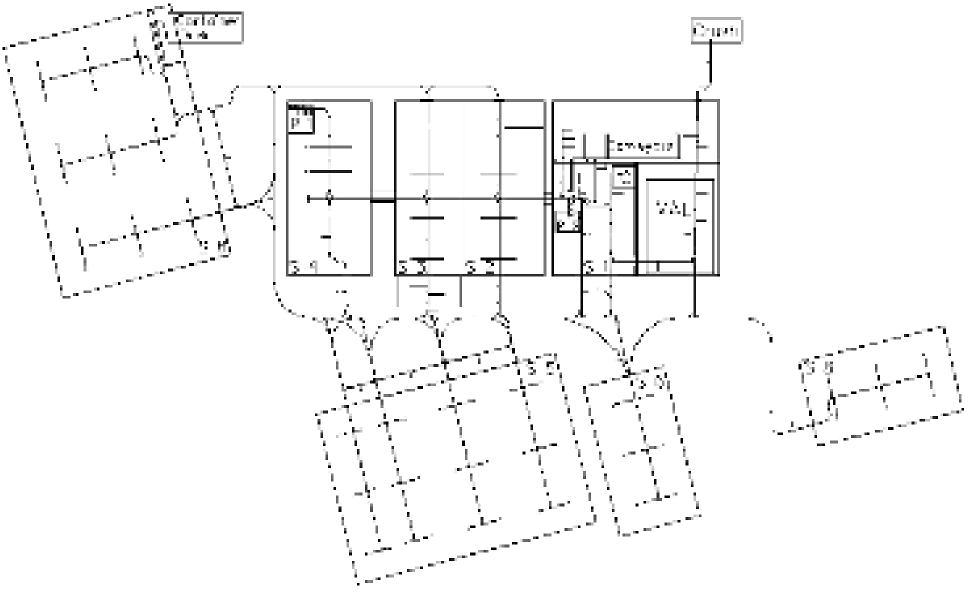

Figure 28. GV path layout connecting all pick up and delivery locations

Figure 29 shows the material flow concerning the transportation tasks for the vehicles at the glass production plant. Between 1200 and 1400 (1300 on average) production pallets, 200-250 VAL pallets and 60-70 ‘extra foil pallets’ arrive at the ‘landing zone’ by conveyor per day. These *Inbound* pallets are stored by product type in stows of 90-120 pallets. About four pallets on average have to be transported to the Crush area per day. Within the storage areas, about 200-250 pallets per day are transported in batches of 10 pallets to the VAL area and 200 pallets are relocated in batches of 2-60 pallets within a storage area for storage space optimization. The VAL and relocation moves will be referred to as the *Internal* moves.

On average, 1820 *Outbound* pallets have to be moved per day in batches of 28 pallets to 65 trucks which arrive just outside the storage areas between 6.00 am and 10.00 p.m., except in the weekends. In 20% of the cases, the trucks must visit two storage areas to be completely loaded. On average, 10% of all outbound pallets from S8 leave via the *container dock* (see Figure 28) instead of the main door of S8. This is because 10% of the trucks that arrive there can only be loaded from the back. Furthermore, there are peak arrivals of trucks during the day, since more trucks arrive in the morning and late afternoon compared to the early afternoon and the evening. Note that the in Figure 29 mentioned pallet moves from the Conveyors include the 225 pallets received from VAL and the 65 ‘extra foil pallets’ (which means that there are 1300 *production* pallet moves on average), and that 1820 outbound moves to the trucks per weekday also means (1820*5/7 =) 1300 moves on average per day.

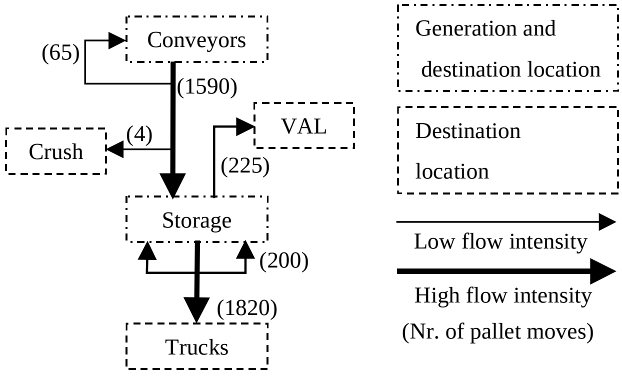

Figure 29. Average weekday material flow between all locations of the glass production plant

In general, 11 guided vehicles are used 24 hours a day, 365 days per year. The vehicles are free to move anywhere on the paths of the defined operating area (see Figure 28) and can pass each other if necessary. However, there is room for only one vehicle at a time at the pick-up and drop locations of the conveyors, trucks and stows in the storage buildings.
Similar as the EDC, the idle vehicle positioning strategy has been defined by the company. When the GVs temporarily have no transportation task, they will park at the closest free parking place. This can be at *P1* (near the vehicle-depot and coffee-corner) or *P2* (near the pallet-conveyor and cafeteria).

The layout of the production plant (see Figure 28) and other relevant specifications of the environment and FLTs (see Table 53) have been modeled in the AutoMod™ simulation software package. The data on load release times, origins and destinations come directly from the database of the WMS of the company and expert judgements. Other parameters such as vehicle speed, pick up times etc. come from (more than 100) careful measurements made at the production plant. A distinction has been made for the pick-up and set down times of loads which are stacked at different heights, (see Table 53).

Speed of loaded GVs on straight paths	2.5 m/s
Speed of loaded GVs in curves	2 m/s
Speed of empty GVs on straight paths	3.5 m/s
Speed of empty GVs in curves	3 m/s
Acceleration/deceleration	1 m/s^2
Pick up time of a load (Height dependent)	13, 19, 28 s
Set down time of a load (Height dependent)	14, 19, 28 s
Vehicle capacity	1 dual-pallet
Simulation period	28 days
Number of working hours per day	24 hours

Table 53. The parameters used for each scenario for the glass production plant

Since pallets are always moved in pairs, the unit-load is one dual-pallet and the glass production plant can be modeled as a uni-load environment (with half the number of pallets to be transported). The length of the simulation is 28 days (such that four weekends are included in the results) in which 45065 dual-load pallets (which means about 90130 single-loads) are independently generated. The rather constant characteristics of the production lines has as result that the interarrival times between the release times of the loads arriving by conveyors at the ‘landing zone’ can be modeled by a uniform distribution. All interarrival times are independently generated. Similarly, trucks collecting Outbound loads are generated with uniform interarrival times since truck arrivals are balanced over the day as much as possible. However, four periods are introduced to realistically represent the variation in the interarrival rates of trucks (Outbound loads) during the day. Table 53 gives a summary of some other values of the model.

5.3.1 Case specific dispatching rules for the glass production plant

Next to the common dispatching rules of Section 5.1, we propose some case specific dispatching rules used at the production plant. Similar to the specific rules of the EDC, the specific rules of the glass production plant can be classified as *centralized* and the last as *pre-arrival*. The dispatching rule currently used at the production plant is discussed first. Next, a similar rule is described to investigate the effects of removing certain restrictions of the dispatching rule currently used.
The following sections describe the control systems individually in more detail.

Centralized control

(a) Dedicated-Dispatching (DD)
The dispatching rule actually used at the production plant is dedicated-dispatching. Of the 11 GVs, 5 vehicles are dedicated to the Inbound jobs, two vehicles are dedicated to all Internal jobs (the relocation moves for storage space optimization and the pallet moves to the VAL area) and the remaining four vehicles are dedicated to all Outbound moves. Since there are no Outbound jobs at night and in the weekends, the remaining four ‘Outbound vehicles’ are free to do any other task and the vehicle dispatching behavior follows Figure 30. In all cases, all idle vehicles searching for a task will first claim the load (of the type which the vehicle is dedicated to) closest to a vehicle within 100 meters (like NWF). The idea is that vehicles will have less empty travel time. If there is no task closer than 100 meters the vehicle will claim the load that has been waiting longest in the entire system (like FCFS). If there is still no transportation job the vehicles will park at the closest parking location (see for the general claim behavior of the vehicles).

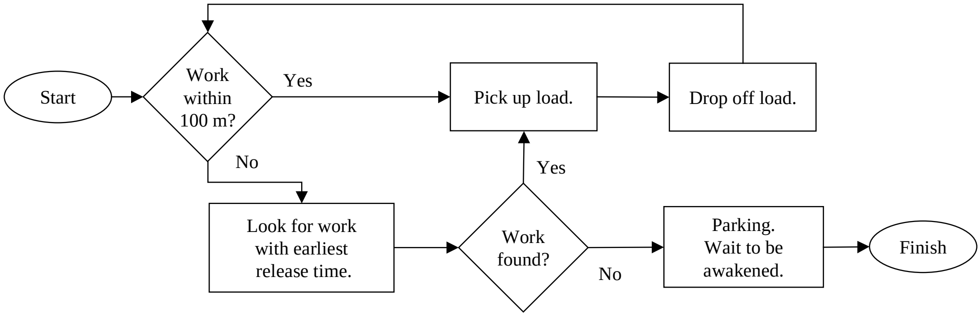

Figure 30. Dispatching behavior with C100FCFS

(b) Closer than 100 meters, First-Come-First-Served Dispatching (C100FCFS)
The C100FCFS dispatching rule is similar to the DD rule except that the 11 vehicles are not dedicated to a task, but free to claim any available load in the system. So no extra check about the load type is needed to match vehicles to loads. The flowchart of Figure 30 shows the dispatching decisions made during C100FCFS. This is still a special rule since it is a hybrid rule of distance (claim the load closer than 100 meters) and time (or else claim the longest waiting load in the system).

Pre-arrival control

(c) Dispatching with Pre-arrival Information (DPI)
This rule uses the common dispatching rules previously described in Section 5.1 and the dispatching rules (a) and (b) described above. The difference is that the load gives a signal (x =) 30, 60, 90 or 120 seconds in advance, before its physical release time. This pre-arrival information can, for example, already be triggered as soon as Inbound loads are placed on the conveyors. Outbound loads can already be released when the trucks arrive at the gate. There is no pre-arrival information for loads that are moved to VAL or for storage space optimization (Internal loads).
Considering that about (7*1525+5*1820)/2 Inbound and Outbound (unit-)loads have to be moved per week, about 1412.5 Inbound and Outbound unit-loads are moved per day on average. The operational time is 24 hours per day, so 60 seconds pre-arrival information is similar to looking $\left(\frac{1412.5*60}{24*3600}\right) = 0.98$ Inbound and Outbound loads ahead on average.

5.3.2 Results for the production plant

The model-parameters are kept the same for each executed run with the previously described vehicle dispatching rules. The only variables in the model are the (logic of the) vehicle dispatching rule, the number of vehicles and the pre-arrival time. This ensures that a fair comparison can be made to rank the dispatching rules according to certain performance criteria. The performance criteria that are most important in the glass production plant, and which have been used to evaluate the different dispatching rules are the following:

- The number of vehicles needed to handle the required throughput
- Average load waiting times
- Vehicle idle time (or percentage of utilization)

Since investments and operating costs are considered very important at the production plant, we will try to keep the number of vehicles to handle the required throughput as low as possible. The current number of vehicles used defines the benchmark for the number of vehicles needed and the average load waiting times. In first instance, however, to compare the performance between different dispatching rules properly, the number of vehicles used with each dispatching rule is the same. In this case 11 vehicles. Therefore, the load waiting time is the main performance criterion to determine the ranking. The criterion vehicle utilization is used to break ties.

Results Using Dedicated-Dispatching

The Dedicated Dispatching rule is similar to the decentralized First-Encountered-First-Served dispatching rule of the European distribution center. Both rules are relatively easy to understand and have a dedicated element. With FEFS at the EDC, vehicles are dedicated to perform transportation tasks for a particular loop (loop 1 or loop 2). With DD at the production plant, vehicles are dedicated to perform transportation tasks for a particular load-type (Inbound, Outbound or Internal). When vehicles were dedicated to a loop in the EDC, they could not ‘help out’ the vehicles in the other loop at times with high workloads. This in turn increased the average load waiting times at the EDC, unless the number of vehicles increases. We expect a similar effect at the production plant. For example, Outbound loads are generated in batches of 28 pallets on average, the result is a relatively high number of simultaneous transport requests which implies a temporarily peak in the work load for the Outbound dedicated vehicles. If some of the vehicles dedicated to other loads were available, they could ‘help out’ the Outbound dedicated vehicles, and as a result decrease the average load waiting times of Outbound loads.

Load-type	Waiting time when $x = 0$	Waiting time when $x = 30$	Waiting time when $x = 60$	Waiting time when $x = 90$	Waiting time when $x = 120$
Inbound	35	11	3	0	0
Outbound	373	373	328	340	304
Internal	482	461	478	465	481
Total	227	212	194	195	184
Utilization	31.3 %	32.1 %	33.8 %	36.2 %	39.1 %

Table 54. Average load waiting times per load-type with DD at various pre-arrival times (x) in seconds

Table 54 shows the average load waiting times when 11 vehicles are dispatched with the DD rule for the specific load-types. The average total load waiting times are represented in the row named ‘Total’. The last row shows the average vehicle utilization realized to obtain the waiting time results. The average waiting times of the Inbound load decreases as the load pre-arrival time increases. In fact, the average waiting times are 0 when pre-arrival information is available (x =) 90 seconds (or more) beforehand. This means that vehicles arrive before the load is physically released and the vehicles have to wait temporarily until the load is available. This vehicle waiting time should not be confused with vehicle idle time. The vehicles are still in the ‘load-retrieving’ status while waiting for the release of the load. The result is that part of the vehicle idle time is transferred to the vehicle utilization time. This phenomenon can be seen in Table 54 by the increasing average vehicle utilization as more pre-arrival time is made available.

Results Using Closer than 100 meters, FCFS Dispatching

With C100FCFS, the 11 vehicles are dispatched similar to DD except that none of the vehicles are dedicated to a certain load-type. This means that any (idle) vehicle can be assigned to a certain transportation task. This ‘relaxes’ the dedication constraint of DD as if vehicles can ‘help out’ other vehicles during peak workloads.

Load-type	Waiting time when x = 0	Waiting time when x = 30	Waiting time when x = 60	Waiting time when x = 90	Waiting time when x = 120
Inbound	91	61	50	47	36
Outbound	289	248	227	222	200
Internal	316	299	308	330	304
Total	200	166	154	153	135
Utilization	44.1 %	43.5 %	45.0 %	47.3 %	49.4 %

Table 55. Average load waiting times per load-type with C100FCFS at various pre-arrival times (x) in seconds

The results can be observed in Table 55. The average total waiting times with C100FCFS for a fixed pre-arrival time decrease when compared to DD, but the average waiting times of the Inbound loads increase. This suggests that the formally Inbound dedicated vehicles are used to ‘help out’ the vehicles dedicated to Outbound and Internal jobs. Hence, increasing the waiting times of Inbound loads and reducing the average waiting times of Outbound and Internal loads.

The average total load waiting times decrease as the pre-arrival time increases. This is mainly due to the reductions in the waiting times of the Inbound and Outbound loads when the pre-arrival time increases. The waiting times of the Internal loads are not influenced as much. This can be explained by the fact that pre-arrival information is only available for Inbound and Outbound load-types.

Results Using Modified First-Come-First-Served Dispatching

With the modified FCFS rule, the GVs are dispatched similar to the C100FCFS rule, except that idle vehicles are in first instance matched to the longest waiting load in the entire system and not in first instance matched to a load closer than 100 meters. In fact, modified FCFS dispatching can be considered as ‘Closer than 0 meters, FCFS’ dispatching. This can also be seen as eliminating the distance-element (or extending the time-element) of the C100FCFS dispatching rule.
Table 56 shows the average load waiting times with modified FCFS dispatching. The average total waiting times with modified FCFS are comparable to the results with C100FCFS dispatching when no pre-arrival information is available. The performance with C100FCFS becomes more favorable than modified FCFS when pre-arrival information becomes available. The load waiting times with modified FCFS even increase (slightly) when pre-arrival information is available 120 seconds beforehand. So it is not favorable to extend the time-element of the C100FCFS rule.

Load-type	Waiting time when x = 0	Waiting time when x = 30	Waiting time when x = 60	Waiting time when x = 90	Waiting time when x = 120
Inbound	98	76	63	53	55
Outbound	283	262	233	213	213
Internal	302	301	315	312	325
Total	198	179	164	149	153
Utilization	43.5 %	44.2 %	45.3 %	46.8 %	49.9 %

Table 56. Average load waiting times per load-type with modified FCFS at various pre-arrival times (*x*) in seconds

Results Using Nearest-Workstation-First Dispatching

When the initial distance of 100 meters of the C100FCFS rule is extended to the largest transport distance between two locations, then (idle) vehicles can be matched to the closest load in the entire system. This is similar to eliminating the time-element (or extending the distance-element) of the C100FCFS dispatching rule. The result is exactly how vehicles and loads are matched with NWF dispatching.
Table 57 shows the results with the NWF rule. These average load waiting times are smaller that the waiting times obtained with modified FCFS. So extending the distance element of the C100FCFS rule is more favorable than extending the time-element. This is consistent with earlier results in which the performances with distance-based rules are more favorable than time-based rules in environments with a variety of transport distances

Load-type	Waiting time when x = 0	Waiting time when x = 30	Waiting time when x = 60	Waiting time when x = 90	Waiting time when x = 120
Inbound	68	46	30	22	17
Outbound	274	245	213	195	191
Internal	305	311	300	297	288
Total	180	160	138	126	120
Utilization	42.3 %	43.5 %	44.6 %	46.5 %	48.9 %

Table 57. Average load waiting times per load-type with NWF at various pre-arrival times (*x*) in seconds

The average total load waiting times with NWF reduces with one minute when two minutes of pre-arrival time is available. The largest contribution to this total reduction is the relatively high reduction (83 seconds) of the Outbound load waiting times.

Results Using Nearest-Vehicle-First Dispatching

The most favorable results are obtained with the load-initiated NVF dispatching rule. Table 58 and Figure 31 show a steady decrease of the Inbound, Outbound and Average total load waiting times.

Load-type	Waiting time when x = 0	Waiting time when x = 30	Waiting time when x = 60	Waiting time when x = 90	Waiting time when x = 120
Inbound	62	42	28	23	17
Outbound	270	244	213	202	188
Internal	303	301	294	308	301
Total	175	155	135	131	120
Utilization	42.0 %	42.8 %	44.4 %	46.5 %	48.6 %

Table 58. Average load waiting times per load-type with NVF at various pre-arrival times (*x*) in seconds

The average waiting times of the Internal moves are more or less unaffected when pre-arrival information is available. As stated before, this is mostly due to the fact that pre-arrival information is only available about Inbound and Outbound moves.

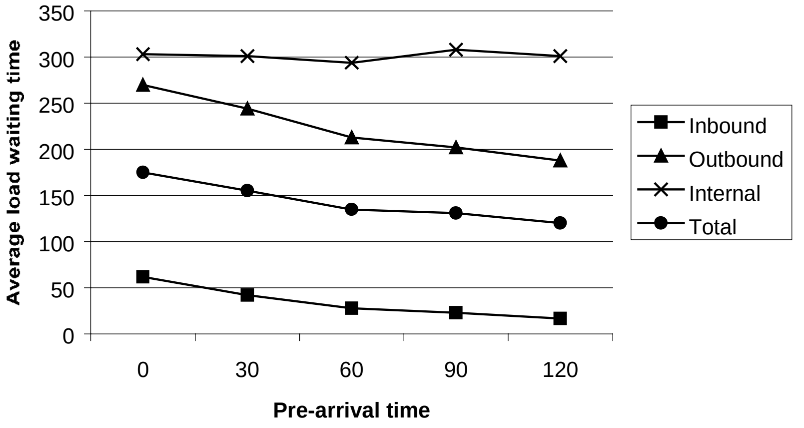

Figure 31. Average waiting times per load-type with NVF dispatching at various pre-arrival times

5.3.3 Ranking dispatching rules for the production plant

Table 59 gives a summary of the results obtained for the glass production plant. The best dispatching rules of the EDC (see previous case) are also the best rules studied in the production environment. In this case, it is still difficult to say which of the two rules is best, although the average total load waiting times are slightly in favor of NVF. In any case, both distance-based rules outperform the other three studied dispatching rules.

C100FCFS is a hybrid version of modified FCFS. Although both rules perform rather similar without the use of pre-arrival information (x = 0), it appears that with pre-arrival information the simple addition of claiming a load closer than 100 meters before claiming the oldest load in the system has a greater impact on the performance than with modified FCFS.

Dispatching rule	Number of vehicles	Average load waiting time in sec. (x = 0/x = 30/x = 60/x = 90/x = 120)	Vehicle utilization (%)
NVF	11	175 / 155 / 135 / 131 / 120	42 – 49
NWF	11	180 / 160 / 138 / 126 / 120	42 – 49
C100FCFS	11	200 / 166 / 154 / 153 / 135	44 – 49
Mod. FCFS	11	198 / 179 / 164 / 149 /153	44 – 50
DD	11	227 / 212 / 194 / 195 / 184	31 – 39

Table 59. Summary of results, the ranking of the various dispatch rules

The current rule used at the production plant is Dedicated-Dispatching (DD). This rule (see the last row of Table 59) is similar to C100FCFS and uses the same number of vehicles as the other rules. However, the vehicles are dedicated to a certain load-type (or task). The load waiting times are relatively much higher compared to the other rules. It is surprising that, in the case of C100FCFS (with x = 30), the absolute decrease in load waiting time is the same as the pre-arrival time, since the pre-arrival time is only available to Inbound and Outbound tasks. The DD is clearly outperformed and can easily be improved by releasing the vehicle dedication constraints. This way, the dispatching rule changes to C100FCFS and can still be improved by increasing the initial distance of 100 meters. If this initial distance is increased to the longest length between two locations of the vehicle paths of the production plant, the rule changes to NWF.

If the average load waiting times of DD are satisfactory, than the DC has to consider using NVF with 8 vehicles. Using NVF with 8 vehicles results in an average total load waiting time of 225 seconds, which is comparable to the average load waiting time of 227 seconds with DD. This means that three vehicles (plus 15 drivers due to round the clock operations) can be saved which will lead to considerable cost-savings.
Vice versa, 17 vehicles (i.e. 6 extra vehicles, two for each load-type, plus 30 drivers) have to be used with DD for an average total load waiting time of 173 seconds, which is comparable to the average load waiting time of 175 seconds with NVF when 11 vehicles are used.

5.4 Description of marine container transshipment terminals

Overseas trade is inevitable in our global economy since more than two thirds of the world is covered with water. The predominant method of transporting bulk cargo in overseas trade has become containerized cargo. In the early stages of containerized transport, containers were transported over sea by regular vessels, and loading and unloading containers on or off vessels were handled with traditional cranes. The increasing popularity of containerized transport motivated carriers and stevedores to modernize their operations. Vessels were specially built and equipped with holds and cells for the transportation of containers. Special cranes with spreaders to hoist containers were designed with which containers could be handled more accurate and faster. These developments have led to the design and realization of modern mega-scale (automated) container terminals where containers are handled efficiently.
The role operations research plays, and an overview of its application at transshipment terminals is discussed in Dekker et al. (1995). They state that too much focus on technology can lead to underperformance. An extended literature overview about transshipment of containers at container terminals is provided by Vis and De Koster (2000).
In the remainder of this section we briefly describe the operations involved at container transshipment terminals. It should be noted that the following discussion is not intended to be an exhaustive overview of container transshipment terminals, but merely as an extended

introduction to familiarize the reader with vehicle-based internal-transport concepts and activities for the last case discussed in this dissertation.

The core business of container terminals is mega-scale loading and unloading of ships, barges, trucks, and trains, and attending to the containers. The containers are thus transshipped form one mode of transportation to another, i.e. intermodal transport of containers. The most common container-lengths are 20, 40 and 45 foot. A standard container, measuring 20 by 8 by 8 foot is referred to as a TEU (twenty-foot-equivalent-unit). The modern container terminals are equipped with sophisticated container handling equipment especially designed to handle these types of containers. The extreme peaks in workloads associated with large container vessels demand a high degree of automation of the container handling equipment. The advantage is continuous operation 24 hours a day, 365 days per year, in all kinds of weather conditions. In Section 1.3 we mentioned that facilities with open air activities need weather resistant handling equipment. This is especially true for container terminals since all activities are open air activities. A container terminal is in this sense an open air-warehouse, see Celen and Leijn (1996).

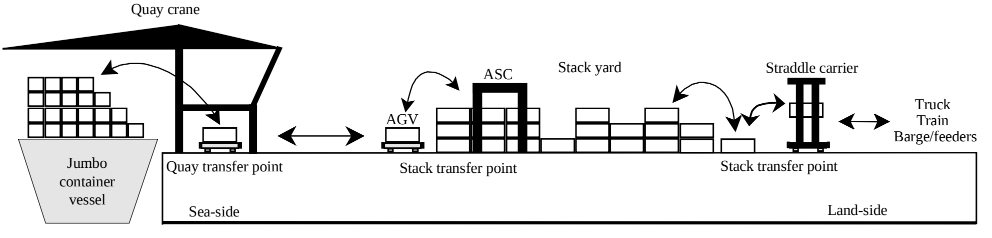

Figure 32. Overview of container moves at a typical container transshipment terminal

Figure 32 shows a typical overview of the activities at a marine container transshipment terminal. The transfer of containers to and from the vessels should be carried out rapidly and efficiently in order to reduce docking time (and costs) of ships at ports. Daganzo (1989), Peterkofsky and Daganzo (1990) and Kim and Bae (1999) study minimizing the delay of ships at ports. When a ship arrives at the terminal, quay cranes take care of the unloading and loading of containers off or on the ship. Unloading the vessel is relatively simple. Containers are picked-up one at a time by the quay cranes and dropped off at the quay. The complexity of the loading activities involves the retrieval of the right containers of clients and to match them to the right place (hold, bay or cell) of the ship. The sequence in which containers arrive at the quay cranes to be loaded onto the ships is therefore very important and should be managed carefully. Arranging the cargo on board a vessel is called stowage planning or preplanning. Shields (1984) provides a computer-aided preplanning system designed to aid in this planning process.
When the vessels are unloaded or loaded, containers are transported from the quay to the stack or vice versa (see Figure 32). The most common transport vehicles used for transporting single containers include: straddle carriers, forklift trucks, reach stackers, multi-trailer systems, AGVs and Automated lifting vehicles (see Section 1.2.1 for a discussion of these vehicles).

Figure 33. Foreground: AGV loaded with a 40 foot container; Background: yard crane *Photo: courtesy of ECT*

A typical example of an AGV found at container terminals is shown in Figure 33. The AGVs are loaded and unloaded by handling equipment such as cranes. The control of AGVs at container terminals is crucial to the system performance due to the large number of vehicles used. Evers and Koppers (1996) present a modeling technique of traffic-control imposed by a hierarchical system of so-called semaphores. The semaphore controls the admission of approaching vehicles individually for a specified part of the vehicle track. By admitting multiple vehicles at the same time, the technique can increase the vehicle capacity of an area (as opposed to zone control where only one vehicle is allowed in an area at a time) and is very useful for facilities with a large number of vehicles.

Automatic lifting vehicles (ALVs) have the ability to lift a container without additional handling equipment, similar to a straddle carrier. Ballis and Abacoumkin (1996) provide a computer simulation model of a container terminal equipped with straddle carriers that can be used to investigate alternative terminal configurations. Kim and Kim (1999b) discuss how to route straddle carriers during loading operations of containers. Their objective is to minimize the total travel time of the straddle carriers.

A method for planning inter-terminal-transport using multi-trailer systems (MTSs) is presented by Kurstjens et al. (1996). An MTS is a truck pulling 5 trailers (so a multi-load vehicle with a total length of about 75 meters) which can be loaded with up to ten TEU. In a simulation model, they show how planning can reduce the total number of empty trips and the waiting time for multi-trailer systems at the terminal. In Duinkerken et al. (1996), three types of vehicles are investigated for inter-terminal-transport: AGVs, ALVs and MTSs. They conclude that a great deal of effort into the control of MTSs is needed to attain an acceptable performance. Both MTSs and AGVs are held up in queues of full vehicles waiting to be serviced by other handling equipment (cranes). However, ALVs are not dependent on other handling equipment, and the number of ALVs needed is half the number of AGVs needed to obtain a comparable service rate.

Similar to storing and retrieving goods in storage areas of warehouses and DCs, containers are stored in stack yards at container transshipment terminals. The stack yard (see Figure 32) is situated between the transfer area of containers between the container yard and the container vessel (sea-side) and the transfer area of containers between the container yard and the road/rail/inland water system (land-side). Containers are stored and retrieved in and from the stack by stacking cranes, like automated-storage and retrieval systems of warehouses. The transport vehicles pick-up or drop off containers at the transfer points at the sea-side and land-side ends of the stack. The same transfer points are used by the stacking cranes to pick-up containers to be stored, or drop off retrieved containers from the stack. The stacking cranes can be automated; (automated stacking cranes, ASCs).

To reduce the area required for container storage, the containers are stacked. Although higher stacking of containers will reduce the storage area, it also requires additional handling to retrieve the containers near the ground (below other containers). The relocation of the containers on top of the desired bottom container leads to a reshuffling of the stack. The consequence of this rehandling of the containers is a higher number of unproductive movements of containers within the stack. De Castilho and Daganzo (1993), Kim (1997), and Kim and Kim (1999a) present methods for measuring the expected number of moves required to retrieve containers from stacks. These methods can be used to make a trade-off between handling effort and stack height.
At the land-side of the stack, containers are transferred to or from other modalities like trucks, trains and barges.

Careful attention must be paid to the operation of container terminals. The terminals must have the facilities to provide an adequate level of service to the ships, see Van der Meer et al. (1999). Central to most terminals are the (guided) vehicles that transport containers between the storage yard and the ships. Managing, controlling and operating such a system is very complex. At the operational level the questions are clear: How should the vehicles be dispatched and routed, and what is an effective traffic control mechanism? In the next section (Section 5.4.1), we will investigate several dispatching rules in order to find answers to some of these questions.

5.4.1 Case study of a container transshipment terminal

The case described in this section is inspired and based on the operations and data of ECT, the largest European container transshipment terminal in the port of Rotterdam.
Arriving container vessels have to be unloaded and loaded before they move to the next port. Currently, 5 quay cranes are in operation for the unloading and loading process of the large container vessels. When a container is unloaded from a vessel, the quay cranes places

the container on one of in total 50 automated guided vehicles, which then transports the container to the stack. The stacking cranes will then pick up the containers from the AGVs and store them in the stack yard. When the quay cranes have finished unloading, other containers will be loaded on the vessel and the process is reversed.

Since the introduction of post-Panamax container ships, the number of vessels, the sizes of the vessels and the number of container moves have increased. The capacity of the vessels increases to 10000 TEU or more. The size of these Jumbo container vessels (JCVs) implies that between 5000-7000 (on average 6000) container moves is foreseen to take place before the vessel moves to the next port. The investigated transshipment company wants to serve JCVs in about 24 hours. This simulation case study was set up as a strategic study for a container transshipment company to investigate some scenarios to see whether the objectives could be met in the future. Some of these objectives include reducing the average handling costs of containers by increasing the volume per time unit and servicing the vessels within 24 hours. Since the JCVs will be bigger, it is assumed that 6 quay cranes will be operational for the unloading and loading process at the JCV, and that automatic lifting vehicles (ALVs) will be used to transport the containers from the quay cranes to the stack. When ALVs are used instead of AGVs, the cranes do not have to wait for a vehicle when setting down a container. The containers can then be placed on the ground and are later picked up by a vehicle. Furthermore, a vehicle can drop off a container in front of the cranes and it does not have to wait until the container is picked up from the vehicle (as with AGVs). It is believed that uncoupling the cranes and ALVs with small buffers will increase the performance enough to meet the objectives set. In fact, the number of ALVs needed to handle the required throughput will be about half the number of AGVs currently used (this means that about 25 ALVs will be used).

Lifting vehicles are currently operating in another area of the stack yard. So by using proven technology of the current material handling equipment (quay cranes, stacking cranes, lifting vehicles, etc.), extrapolation of actual data to generate the number of moves required per vessel (see Celen et al., 1997) and a simulation model to combine the data and equipment, we will study the performance of several on-line centralized vehicle dispatching rules in this type of transshipment environment.

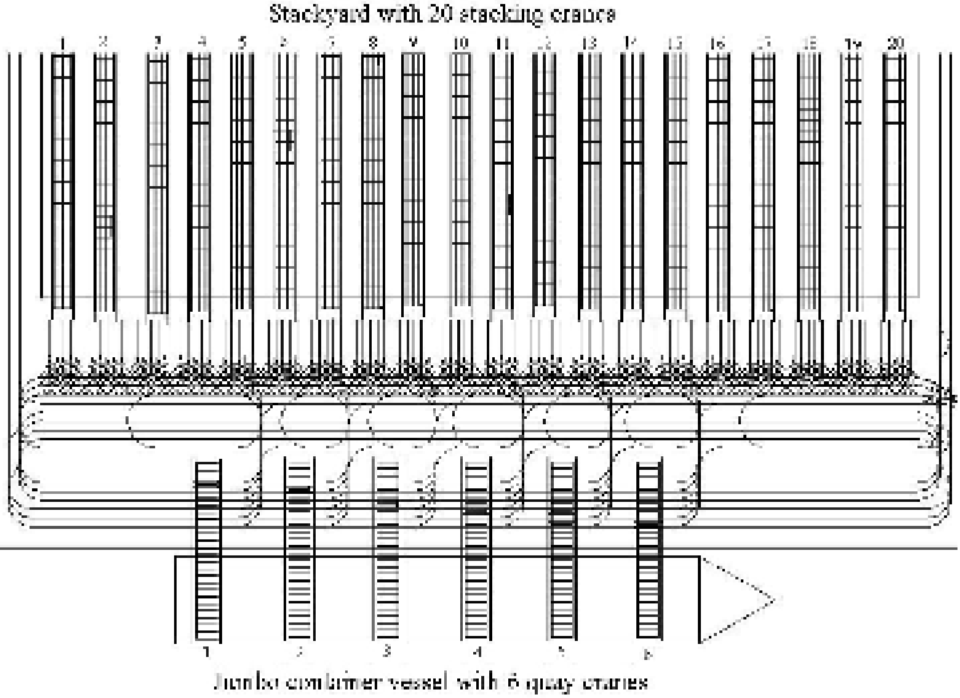

Figure 34. ALV path layout connecting all pick up and delivery locations, all main transport tracks are uni-directional

In the new environment, see Figure 34, there are 6 quay cranes operating the Jumbo vessels. Since the vessel needs about 1 to 1½ hours both to dock and to undock, there are about 21 hours left in which the quay cranes are operational. On average there are 6000 container moves per JCV. This means 500 Inbound containers are unloaded before 500 Outbound containers are loaded per quay crane on average. The Inbound containers are stacked within the stack yard by one of the 20 randomly chosen stacking cranes. These stacking cranes are operational 24 hours a day. There are also land-side tasks at the other end of the stack yard.

Outbound containers are generated randomly from 20 stack lanes operated by one stacking crane each. The containers are then transported by the stacking cranes to the front of the stacking lanes where they are placed on the ground.

An ALV then transports the container to the designated quay crane. There is a capacity of 5 ALVs in front of each stack lane (see Figure 34) where the ALVs can also park when idle. The operational area of the ALV is 120 by 540 meters. The ALVs are bi-directional; i.e. they make no distinction between traveling forwards or backwards. When there are no transport tasks available, the ALVs will park at the closest available parking place in front of the stack lanes. The parked vehicles are balanced over the stacking lanes. This means that only one vehicle can park at a stacking lane at a time unless all stacking lanes already have one parked vehicle. In that case, only two vehicles are allowed to park at one lane at a time unless all lanes already have two parked vehicles, etc. Since the considered model has 20 stacking lanes and 25 ALV, at most two vehicles can be parked at the same stacking lane at any time and only during periods at which all or almost all vehicles are idle (periods in which the vessel docks or undocks).

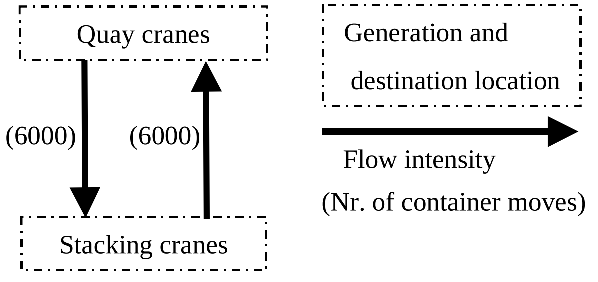

Figure 35. Average material flow between for the ALVs at the transshipment terminal per JCV

Figure 35 shows the material flow that concerns the transportation of containers for the ALVs. Containers are generated in sizes of one TEU and two TEU. There are about 50% containers of one TEU and 50% of two TEU. About 10% of those containers are reefer-containers. These containers can be cooled or heated to preserve perishable products (fruit, meat) and therefore need electricity. To hook them up to electricity they are stored at the back of the stack. The unit-load for the model of the transshipment terminal is one container, since the ALV operations for one TEU, two TEU and reefers containers are the same.

In some cases the ALVs need to drive an extra arc in order to rotate the container in the right direction. Orientation maneuvering is necessary since the doors of the containers must face the back of the vessel. In total, 50% of all moves on average in and out the stack yard require extra orientation maneuvering.

The design of the terminal and other relevant specifications including the ALVs has been modeled in the AutoMod™ simulation software package. The data about the loads vehicles, cranes, etc. come from a data definition study of the company (see Celen et al., 1997) and expert judgements. It is assumed that quay cranes are continuously available to load and unload ships, and load generations are spread evenly over the available operating hours to serve the ship. As mentioned above, the number of container moves varies between 5000 and 7000, which implies between 2500 and 3500 moves for unloading and loading respectively. Each crane is independently assigned to unload between 2500/6 and 3500/6 containers and then load between 2500/6 and 3500/6 containers. Since any number of moves between 2500/6 and 3500/6 is equally likely, the number of moves is drawn form a uniform distribution. The interarrival times of the container generations (on the vessel and in the stack) are drawn from a uniform distribution since load generations are spread evenly over the available operating hours to serve the ship, (i.e. the interarrival times are equal to the estimated crane operation time divided by the average total number of moves to be made). Although the release times of containers to be loaded are uniformly generated within the stack, the release times of these containers when the stacking cranes deliver them at the stack transfer points for the ALVs might be disturbed due to response times of the cranes and possible land-side operations for the stacking cranes, which are also uniformly generated.

In a simulation run, 20 JCVs are served, which means that the performance is measured over 120000 container moves. Table 60 gives a summary of some other values for the model of the transshipment terminal.

Length of vehicle	14 m
Speed of loaded ALVs on straight paths	6 m/s
Speed of loaded ALVs in curves	3 m/s
Speed of empty ALVs on straight paths	7 m/s
Speed of empty ALVs in curves	3 m/s
Acceleration / deceleration (loaded)	0.3 m/s^2 / 0.5 m/s^2
Acceleration / deceleration (empty)	0.5 m/s^2 / 1 m/s^2
Pick up time of a load (constant)	22.2 s
Set down time of a load (constant)	22.2 s
Vehicle capacity	1 load (container)
Simulation period	20 JCVs (about 20 days)

Table 60. The parameters used for the model of the transshipment terminal

5.4.2 Case specific dispatching rules for the container transshipment terminal

Next to the common centralized dispatching rules of Section 5.1, we introduce a new rule using location-priority lists. The specific rules of the terminal are classified as *centralized* or *pre-arrival*. Since the actual situation at the transshipment company involves AGVs, no current dispatching rule for the ALVs is used. The current situation using AGVs involves dedicating a certain number of vehicles to a quay crane. We have seen in the previous cases that dedicating vehicles in any way can lead to underperformance. So we will not explore vehicle dedication options for the new rule in this case. We have also seen that load-initiated dispatching rules generally lead to a more favorable performance than similar rules based on vehicle-initiated dispatching. The loads will therefore have the first dispatching initiative in the new rule. We have also seen that distance-based rules lead to more preferable results than time based rules. The new rule will therefore be mainly distance-based. Some time elements are added to take the advantage to match vehicles to loads successively in terms of release times, when the second dispatching initiative (vehicle-initiated dispatching) is used. Furthermore, location priority-lists are used to balance travel times between quay cranes and stacking cranes.

The following sections describe the control systems individually in more detail.

Centralized control

(a) Hybrid Dispatching

As mentioned, the loads have the first dispatching initiative. In fact, the hybrid-dispatching (HD) rule uses both load-lists and work-lists with time and distance elements.

When loads are dropped by stacking cranes 1-4, the loads first try to wake a vehicle at quay crane 1, and if unsuccessful the closest vehicle in the entire system is awoken for transport. The LLs of stacking cranes 5-8 check quay cranes 2 first before searching for the closest vehicle in the system, etc. The load lists are constructed this way with the idea to balance travel times between quay cranes and stacking cranes. Checking for an idle vehicle at the quay cranes is actually a formality. ALVs at the quay cranes are idle only for a split second when dropping off a load, and are then instructed to park at the stacking lanes or pick-up a load at the stacking lanes.

When quay cranes drop off loads, the LL at that crane is checked. The LL at quay crane 1 checks the locations of stacking cranes 1-4 to wake up the closest vehicle. The LL at quay crane 2 checks stacking cranes 5-8 for the closest vehicle, etc.

Idle or awoken vehicles checking their work-lists at quay crane 1 will try to claim the oldest (longest waiting) load at stacking cranes 1-4. If unsuccessful they check for the closest task in the entire system. If still unsuccessful the ALVs will park at the closest available parking place at the stack lanes. Only one vehicle can park at a stack lane at a time in order to balance the parked vehicles over all stack lanes. Stack lanes can only have multiple parked vehicles when all stack lanes already have a parked vehicle.

Similarly, vehicles checking their WLs at quay crane 2 and 3 will claim the oldest load at stacking crane 5-8 and 8-11 respectively. So there is a small overlap in the WLs of the vehicles. Furthermore, idle vehicles at stacking cranes 1-4 and 5-7 will first try to claim a load at quay crane 6 and 5 respectively, etc. If unsuccessful the vehicles will try to claim the closest load in the entire system; if still unsuccessful they will remain parked.

The Hybrid dispatching rule is thus a combination of a load-initiated distance-based location-priority rule and in second instance a vehicle-initiated time and distance-based location-priority rule.

Pre-arrival control

(b) Dispatching with Pre-arrival Information (DPI)

The DPI rule uses the HD rule (a) and the common dispatching rules of Section 5.1. In this case, a vehicle can be reserved for a pick-up task by the quay cranes or the stacking cranes. The following situations are investigated:

1. Reserve a vehicle as soon as an Inbound container is picked up from the JCV by a quay crane (which means about 30 seconds before it is dropped off)
2. Reserve a vehicle as soon as an Outbound container is picked up in the stack yard by a stacking crane (which means about 70 seconds before it is dropped off)
3. Reserve vehicles both at quay cranes and stacking cranes

5.4.3 Results for the container transshipment terminal

The performance criteria we look at for the transshipment terminal are the same criteria that have been used to evaluate the different dispatching rules for the EDC and the glass production plant, i.e.:

- The number of vehicles needed to handle the required throughput
- Average load waiting times
- Vehicle idle time (or percentage of utilization)

The number of ALVs to handle the required throughput has been set at the estimated number of half the current number of AGVs used (see Duinkerken et al., 1996). This means that the number of ALVs used is fixed at 25. Therefore, the load waiting time is the main performance criterion to determine the rankings. The load waiting times are measured as the time containers are placed on the ground by quay cranes or stacking cranes until picked up by ALVs. So the release times of the containers are the set down times of Inbound containers by quay cranes and set down time of Outbound containers by stacking cranes. This means that waiting times of containers within the stack or within a vessel to be picked up by a crane are not included. In this case we are only concerned with the performance of the ALVs. The criterion vehicle utilization is used to break ties.

Results Using Modified First-Come-First-Served Dispatching

The resulting average waiting times with modified FCFS vehicle dispatching are shown in Table 61. The individual waiting times of the Inbound and Outbound loads are also provided in the results. The average Total waiting times is the average waiting times of the Inbound and Outbound loads. The effects of reserving vehicles for Inbound moves at quay cranes, or Outbound moves at stacking cranes are clearly shown. The results of combining reservations of vehicles as quay cranes and stacking cranes are shown in the last column of Table 61.

Load-type	Waiting time without pre-arrival information	Waiting time with pre-arrival at quay cranes	Waiting time with pre-arrival at stacking cranes	Waiting time with pre-arrival at all cranes
Inbound	123	93	123	93
Outbound	68	68	25	25
Total	95	80	74	59
Utilization	68.5 %	69.0 %	71.8 %	72.1 %

Table 61. Average load waiting times per load-type with modified FCFS at various load pre-arrival instances

Reserving vehicles at the quay cranes alone only reduce the waiting times of the Inbound jobs. In the previous cases of the EDC and the glass production plant, Inbound and

Outbound moves occurred simultaneously and pre-arrival information of Inbound loads also affected the waiting times of the other loads. This can be explained by the fact that the other loads can be picked up sooner if Inbound loads are served quicker. At the transshipment terminal, a crane can only start loading when it has finished unloading the containers. So Inbound and Outbound moves can be seen as successive processes with a very small overlap. This overlap occurs when some quay cranes are loading Outbound loads while some other quay cranes are still unloading Inbound loads. This overlap period is rather small and therefore possible effects are negligible.

Results Using Nearest-Workstation-First Dispatching

Dispatching vehicles at the transshipment terminal with NWF results in practically the same average load waiting times as with dispatching with modified FCFS. Table 62 shows that the average waiting times are at most two seconds more favorable than the results of Table 61. The rather simple dispatching of vehicles from the vessel to the stack and vice versa seems to be insensitive to whether time or distance-based dispatching is used.

Load-type	Waiting time without pre-arrival information	Waiting time with pre-arrival at quay cranes	Waiting time with pre-arrival at stacking cranes	Waiting time with pre-arrival at all cranes
Inbound	123	93	123	93
Outbound	66	66	24	24
Total	94	79	73	58
Utilization	68.4 %	69.1 %	72.7 %	71.6 %

Table 62. Average load waiting times per load-type with NWF at various load pre-arrival instances

Results Using Hybrid Dispatching

The results of the previous two sections show that both the time-based and distance based vehicle-initiated dispatching rules perform comparably. The results with the HR in Table 63 are also comparable with the other two rules. Although the average waiting times for the Outbound loads are slightly more favorable when load-initiated GV dispatching is used with time and distance-based elements.

Load-type	Waiting time without pre-arrival information	Waiting time with pre-arrival at quay cranes	Waiting time with pre-arrival at stacking cranes	Waiting time with pre-arrival at all cranes
Inbound	124	94	125	94
Outbound	60	61	21	22
Total	92	77	73	58
Utilization	67.8 %	68.2 %	71.8 %	71.6 %

Table 63. Average load waiting times per load-type with HD at various load pre-arrival instances

Results Using Nearest-Vehicle-First Dispatching

Finally, when removing the location-priority lists and the time-based elements from the HR, we obtain the nearest-vehicle-first dispatching rule. The results with the NVF rule are shown in Table 64. The differences in waiting times with the HR are in favor of the NVF rule, and are about the same as the differences in waiting times between both vehicle-initiated rules.

Load-type	Waiting time without pre-arrival information	Waiting time with pre-arrival at quay cranes	Waiting time with pre-arrival at stacking cranes	Waiting time with pre-arrival at all cranes
Inbound	123	93	124	93
Outbound	59	59	22	22
Total	91	76	73	58
Utilization	67.4 %	68.0 %	71.7 %	72.7 %

Table 64. Average load waiting times per load-type with NVF at various load pre-arrival instances

The results for the average Inbound load waiting times are comparable for all four dispatching rules. The largest performance differences that can be observed are with the average Outbound load waiting times. The NVF Outbound load waiting times of 59 seconds has reduced about 11% compared to the average Outbound load waiting times of 66 for modified FCFS.

5.4.4 Ranking dispatching rules for the container transshipment terminal

Table 65 gives a summary of the average Total load waiting times (and the vehicle utilization) obtained for the container transshipment terminal. It seems that the average (Total) load waiting times for this environment are rather insensitive to the dispatching rule used, although the ranking (with absolute differences in waiting times) is slightly in favor of the load-initiated driven dispatching rules. The rules behave in a similar way since Inbound and Outbound jobs are separated. This makes combinations of double-play (combining Inbound and Outbound jobs to reduce vehicle empty travel time) almost impossible. In situations where Inbound and Outbound moves occur simultaneously, it is more likely to combine transports with distance based rules such that the relative differences with time based rules are more obvious. Since the relative difference between two successive rules is smaller than 5%, no dashed lines are drawn to separate the performances of the different rules in Table 65.

Dispatching rule	Number of vehicles	Average load waiting time in sec. (x = 0 / PQ / PS / PA)*	Vehicle utilization
NVF	25	91 / 76 / 73 / 58	67 - 73 %
HD	25	92 / 77 / 73 / 58	68 - 72 %
NWF	25	94 / 79 / 73 / 58	68 - 72 %
Mod. FCFS	25	95 / 80 / 74 / 59	69 - 72 %

* x = 0: no pre-arrival information used; PQ: Pre-arrival information used at Quay cranes;
PS: Pre-arrival information used at Stacking cranes; PA: Pre-arrival information used at All cranes.

Table 65. Summary of results, the ranking of the various dispatch rules for the transshipment terminal

The largest relative difference (about 5%) is found between modified FCFS en NVF when no pre-arrival information and pre-arrival information at the quay cranes is used (see the PQ results in Table 65). Using similar arguments used at previous cases, we can say that NVF outperforms modified FCFS. This is consistent with previous results for non-symmetrical environments.

The pre-arrival information at the quay cranes of about 30 seconds is reflected as a decrease by 15 seconds in average load waiting times. This is independent of whether pre-arrival information is used at the stacking cranes or not. The decrease in average load waiting times by 15 seconds when about 30 seconds of pre-arrival information is available, is due to the fact that the average load waiting time is the average of 50% Inbound (quay crane) and 50% Outbound (stacking crane) moves and that the minimum travel time from the stack yard to the quay cranes is larger than 30 seconds.

5.5 Concluding Remarks

It is conceivable that different internal transportation environments characterized by different impact-factors (see Table 31 at the beginning of this chapter) need different vehicle dispatching rules to transport the loads with minimum average load waiting times. In this chapter, we modeled three different environments, and studied the performance of several well known on-line dispatching rules found in literature and some case specific dispatching rules, and ranked the performance of the rules according to:

- the number of guided vehicles (GVs) needed to meet the required throughput
- the average load waiting times
- and vehicle utilization

In all cases, the GVs are dispatched on-line since only real-time information is available. However, the performances of the dispatching rules are also studied when some load pre-arrival information is assumed to be known. The effects of pre-arrival information on the

ranking of the dispatching rules is also studied. In each case, the pre-arrival times are set such that they can indeed be obtained at the cases studied.
The studies of the three companies are based on highly detailed simulation models using real company data. The three environments studied are:

1. A distribution center for computer components with about 600 pallet moves per day (7.5 hours), 5 GVs and a GV operational area of 40 by 140 meters.
2. A production plant for packaging glass with about 1600 pallet moves per day (24 hours), 11 GVs and a GV operational area of 315 by 540 meters.
3. A container transshipment terminal with about 6000 container moves per JCV (about 24 hours), 25 GVs and a GV operational area of 120 by 540 meters.

The three companies have different internal transport systems, described in Sections 5.2, 5.3 and 5.4. Some of the case specific impact-factors and differences that influence the transport systems performance (see also Table 31 at the beginning of this chapter) are mentioned in Table 66. For example, the average number of loads generated per hour (measure of throughput) at the distribution center can be calculated by dividing the average number of pallets to be moved per day by the number of working hours per day, i.e. $581/7.5 \doteq 77$. Since the number of loads generated per hour at the production plant varies for weekdays and weekends, the loads generated per hour have also been specified for these periods (peak variations).

Factors / Company	Loads generated per hour	Load transport time (sec)	Min. vehicle utilization (%)	Nr. of different transport distances (approximately)
Distribution center	77	111	48	800
Production plant	67 (81, 33)	142	24 (29, 12)	2000
Transshipment terminal	250	170	47	240

Table 66. Some case specific impact-factors and differences of the transport systems

The average load transportation time (for uni-load vehicles) at the distribution center of 111 seconds (see also Section 5.2.5) is smallest for the three cases. This transportation time is the average time needed to handle a load and includes load transport, pick up and set down times but excludes the time needed to retrieve the loads. With the previous two statistics, the minimum vehicle utilization can be calculated. This is the minimum percentage of time required by the vehicles to deliver the loads and does not include the percentage of time needed to retrieve the loads. For example, at the distribution center 581 loads with an average transportation time of 111 seconds are transported in 7.5 hours by 5 vehicles. This means that the vehicles use at least $\frac{581}{7.5} * \frac{111}{3600 * 5} \doteq 0.48$ (denoted by 48%) of their available time to deliver the loads. The number of different transport distances gives an indication of the dispersion of the pick up and delivery locations.

The impact-factors in Table 66 indicate the differences between the three practical cases. As mentioned before, we modeled several well-known on-line dispatching rules found in literature and some case specific dispatching rules, and studied the influence of different impact-factors on the performance of the transport systems and the rank of the dispatching rules. The three common dispatching rules found in literature and used in all cases include:

- nearest-vehicle-first (NVF), in which a released load at a certain location or workstation triggers the search for the nearest idle vehicle in the system (load-initiated distance-based rule)
- nearest-workstation-first (NWF), in which an idle vehicle triggers the search for the nearest released load in the system to be transported (vehicle-initiated distance-based rule)
- modified first-come-first-served (Mod. FCFS), in which an idle vehicle triggers the search for the longest load waiting to be transported, (vehicle-initiated time-based rule)

Table 67 shows the general ranking of the common dispatching rules. Although NVF and NWF are rather similar in resulting average load waiting times, NVF is ranked more favorably. Both are in all cases clearly better than Modified FCFS. In most cases the case specific rules ranked somewhere between the three common rules. This happened most often between the distance-based rules (NVF and NWF) and the time-based rule (modified FCFS). Note that this rank holds for all three cases studied and is not influenced by the impact factors, which indicates that the performance of the dispatching rules seem to be insensitive for differences in the transportation environments.

General ranking	Dispatching rule
1	NVF
2	NWF
-	Case specific
3	Mod. FCFS

Table 67. General ranking of dispatching rules for internal transport studied in different environments

The main disadvantage of the vehicle-initiated distance-based NWF rule is that some pick up points of loads turn out not to be nearest to any vehicle during relatively high throughput periods. The load at that pick up point may never qualify to be served by a vehicle in busy systems if other closer loads are always available. Since new deliveries could continue to take place, the output queues of the affected load-locations will grow to their maximum capacity. The distance-based load-initiated rules have the advantage that remote areas will be matched to a vehicle even if the vehicles are at the other end of the vehicle operating area. Time-based rules have a similar advantage. In that case, vehicles or loads do not discriminate jobs with respect to transport distances, but simply serve the next transport request of the systems transport request list, which is sorted according to load release times. The disadvantage of time-based rules like FCFS, is that no special regard is given to minimize vehicle empty travel time, which generally reduces average load waiting times.

Therefore, a load-initiated distance-based dispatching rule like NVF generally leads to satisfactory performance in practice, since vehicle travel distances for practical environments are diverse (the controller can make a unique match), loads can always be matched to vehicles (even if vehicles or loads are in remote areas) and at real-time level, vehicle empty travel time is minimized, hence minimizing load waiting times.

In conclusion, it can be said that vehicle dispatching based on specially designed priority-lists (load-lists or work-lists) or based on dedicating vehicles to certain tasks can be outperformed by a moderately simple rule like NVF, which seems to perform rather well in any type of environment. Furthermore, using realistic pre-arrival information can significantly reduce average load waiting times. We have seen that in the case of the EDC, a pre-arrival period of 15 seconds, or an average of 0.063 loads per vehicle, reduces the waiting times about 5% with NVF and 7% with modified FCFS. About 30 and 70 seconds, or an average of 0.05 and 0.1 load per vehicle of available pre-arrival information at the quay and stacking cranes respectively, reduces the average total waiting times with 57% and 61% with the NVF and modified FCFS rule respectively at the transshipment terminal. Note that the reductions in waiting times are more favorable at the transshipment terminal compared to the EDC. However, the vehicle utilization is lower at the terminal and more pre-arrival information was available. The vehicle utilization at the production plant was even lower than the utilization at the transshipment terminal and relatively more pre-arrival information was available. Using a pre-arrival period at the production plant of 120 seconds, or an average of 0.18 loads per vehicle, reduces the average waiting times (for the Inbound and Outbound loads) about 65% and 77% with the NVF and modified FCFS rule respectively.

The reductions in waiting times when pre-arrival information is available with modified FCFS were in all cases higher than those obtained with NVF. It seems that waiting time reductions are more favorable with the time-based rule. However, the absolute waiting times are more favorable with the distance-based rules. Furthermore, waiting time reductions with the use of pre-arrival information decrease as the vehicle utilization increases. This is consistent with earlier results obtained in Chapter 4.

Chapter 6

Conclusions and further research

In this dissertation we considered the control of guided vehicles of vehicle-based internal transport systems found in many warehouses, distribution centers, production plants, transshipment terminals, etc. Common logistic activities for all these facilities are transshipment, storage, and physical distribution of materials. The material transport in these facilities is generally taken care of by vehicle-based systems that act as the link for materials between different locations within the facilities.

The research started with the observation that the number of automated guided vehicle and mobile-terminal controlled industrial-truck implementations are rapidly increasing in number. A better understanding of the issues involved in the control of internal transport helps managers and designers to increase facility performance.

The performance of a material-flow system can be measured in several dimensions, such as flow times, delays, required throughput, waiting time of goods and idle time of vehicles. The efficiency of a vehicle system is sensitive to operational design parameters, such as vehicle path layout, track capacity, track control (uni- or bi-directional), the number of vehicles needed, design of the vehicle (uni- or multi-load capacity), reliability, and the logic of the vehicle-control system. Due to the high degree of stochasticity/randomness within transport environments, vehicles are dispatched on-line based on real-time events. The objective of this dissertation is to gain more insight into the issues concerned with on-line control of vehicle-based internal transport and the relative performance of common (and less common) dispatching rules.

Next to some theoretical models, the transportation environments of three different companies in practice are discussed. In order to relate and make our results accessible to industry, we have analyzed vehicle dispatching rules that are easy to understand by practitioners and easy to implement in logistics software packages. Although this implies that the overall best or optimal dispatching rule associated with a given internal transport environment may not be found, our analysis is more realistic from a practical point of view.

We start with a brief overview of the results and discuss the most important conclusions in Section 6.1. Section 6.2 discusses subjects for further research, and Section 6.3 concludes the dissertation with some guidelines for selecting control systems in practice.

6.1 Conclusions

The research of this dissertation started with an introduction and an overview of material handling systems, and placed the main subject of this dissertation accordingly. Chapter 2 provides an extensive literature overview of subjects related to vehicle-based internal transport systems. In Chapters 3, 4 and 5, the research focuses on the (mathematical) modeling of vehicle-based internal transport systems. Various issues have emerged in those chapters on topics such as the performance difference between on-line and off-line vehicle control, why on-line systems are more commonly found in stochastic environments, the performance differences between on-line decentralized and centralized systems, the performance effects of time-based and distance-based elements of centralized load-initiated and vehicle-initiated dispatching rules, the effects of using load pre-arrival information, dwell point strategies and pre-assigning (moving) vehicles to loads, and the rank of certain dispatching rules relative to other dispatching rules.
In the following sections we will discuss our achievements and conclusions for this dissertation based on the topics mentioned above.

On-line versus off-line vehicle control

The research in Chapters 3 and 4 focuses on the mathematical modeling of vehicle-based internal transport systems. Two warehouse layouts (U-layout and I-layout) are modeled and the internal transport activities are formulated as a pick-up and delivery problem with time windows. The idea is to use off-line control (with the pick-up and delivery problem modeling) as a benchmark for on-line dispatching rules. Using off-line control means that all information on load release times, origins and destinations has to be known in advance. This is not a real situation found in practice, due to the stochastic nature of internal transportation environments. However, for theoretical purposes we assume that all information is available when off-line control rules are used.

The results show that for the two different layouts studied, considerable reductions in average load waiting times are possible with off-line control (exact and heuristics) if the system is relatively quiet, i.e. relatively few dispatching requests per time unit (low throughput) and vehicles have relatively high idle times (in this case about 15-20%). The lower average waiting times are possible because vehicles can already travel to the loads before they have been physically released. Therefore the loads can be picked up relatively sooner. In low throughput environments, the fleet size with on-line dispatched vehicles has to increase about 50% to obtain similar results compared to off-line control. It is shown (in Table 13) that 50% extra guided vehicles reduces the performance deviation to about -2 or 2% on average.
In the case that vehicles have relatively low idle times, (in this case 0%), there is less opportunity to reduce load waiting times and the performance of on-line control can already be considered satisfactory since the deviations in average load throughput times are about 6 to 20% compared to off-line control. Increasing the fleet size with 50% for the

two environments studied in Chapter 3 outperformed off-line control considerably (about 57 to 63%, as can be seen in Table 13 on page 64).
To close the gap between on-line to off-line control, investments have to be made by either increasing the fleet size, or getting more reliable information on load origins, destinations and release times. However, we will discuss next that more information might still not be satisfactory.

Control with perturbed release times

In Chapter 4 we studied the effects on load waiting times when dispatched vehicles encounter small perturbations (5%) in the actual release times of loads. If the release times are given far in advance they can become unreliable. With unreliable release times, the expected load waiting times with off-line control change, since vehicles are routed on a ‘old’ route. In that case, the vehicles arrive a little later or earlier than the expected release times of the loads. With on-line dispatching, the dispatching perspective remains real-time and vehicles seize the opportunity to pick up a load that was released relatively sooner. Results of routing vehicles with perturbed release times (Table 30 on page 90) show that the deviation in load waiting times between on-line dispatching and off-line control decreases when load release times become unreliable. With heavily utilized vehicles it is even possible that the waiting times of off-line control become worse than the waiting times obtained with on-line dispatching. In other words, on-line control outperforms off-line control in high throughput environments when load release times become unreliable.
We conclude that off-line control is much more sensitive to unexpected deviations in release times than on-line control, which seizes the opportunity to service jobs in a more favorable order. This is logical, unexpected deviations or unreliable information create a stochastic environment, unsuitable for off-line control. On-line control systems are therefore more suitable for practical situations. Those control systems can be further subdivided as discussed in the next sections.

Decentralized versus centralized vehicle control

In Chapter 5 we studied different on-line vehicle control systems used in practice. In the first case study (described in Section 5.2) we investigated a decentralized system with two partially overlapping loops, and several centralized control systems. The results (Section 5.2.2) showed that (29%) fewer vehicles were needed with centralized control in order to obtain similar load waiting times as decentralized control. This is because decentralized control makes inefficient use of information about the location and status of vehicles and loads, and vehicles are dedicated to a loop so that the work cannot be evenly shared (or balanced among the vehicles). Furthermore, when decentralized control is used, the vehicles are constantly in motion, and are therefore more liable to break down, wasting energy and possibly causing unnecessary congestion.

Although the centralized systems win in the number of vehicles needed, they lose in simplicity from the decentralized systems. Decentralized control systems with loops will in general be outperformed by centralized control systems because the latter are more efficient and make better use of available information. So, in view of this and the higher number of vehicles needed, centralized control is ranked more favorable than decentralized control (see also Table 44).
The performance of different centralized control methods is influenced by many factors (see Table 31). Such factors can make a certain dispatching rule more favorable than another. Several factors, which influence the performance of centralized systems, are discussed next.

Load-initiated versus vehicle-initiated dispatching rules

In Chapter 5 we have investigated several load-initiated dispatching rules including nearest-vehicle-first and load-list dispatching, which we compared with several vehicle-initiated dispatching rules including nearest-workstation-first and work-list dispatching. In each case, the average load waiting times obtained with the load-initiated vehicle dispatching rules were slightly more favorable than the waiting times obtained with the vehicle-initiated dispatching rules. Similar results were observed in Chapters 3 and 4. The main disadvantage of vehicle-initiated rules is that when the rules are distance-based, like nearest-workstation-first, certain loads may never qualify to be served by a vehicle in busy systems if other loads are always available. The load-initiated rules have the advantage that all areas will eventually be served by a vehicle, independent of the vehicle utilization rate or the distance of the vehicle to the load.

Time-based versus distance-based dispatching rules

Time-based rules such as modified first-come-first-served have a similar advantage as distance-based rules such as nearest-vehicle-first in which remote areas will be matched to a vehicle even if the vehicles are at the other end of the vehicle operating area. With time-based rules, vehicles or loads do not discriminate jobs with respect to transport distances, but simply serve the next (or oldest) transport request of the systems transport request list, which is sorted according to load release times. The disadvantage of time-based rules is that no special regard is given to minimize vehicle empty travel distance (or time), which generally reduces average load waiting times.
Consequently, the results in Chapters 3, 4 and 5 reveal that distance-based rules generally seem to perform more favorable in layouts where all transport distances are different and time-based rules perform relatively more favorable in environments where all transport distances are similar. This seems logical since decisions based on similar distances with distance-based rules are comparable to inefficient random load-to-vehicle assignments. The distances between locations of environments found in practice are in general not similar but quite different. The results of the practical cases in Chapter 5 show that the

distance-based rules perform better than the time-based rules. In the case of the transshipment terminal with a relatively small number of different transport distances (see also Table 66), the performance obtained with first-come-first-served was quite similar to those obtained with nearest-vehicle-first (see Table 65). This is consistent with the conclusion that distance-based rules perform better in environments with a high dispersion of the pick up and delivery locations where transport distances are different.

Uni-load versus multi-load capacity vehicles

By introducing vehicles with multi-load capacity, more loads can be transported simultaneously. This means that the average load waiting times can decrease because load transports can be combined. But the average load transportation time can increase because certain loads can remain on the vehicle while other loads are served with the remaining vehicle capacity.

When multi-load dispatching is used, the performances of different dispatching rules (such as first-come-first-served and nearest-workstation-first) become comparable (see Section 3.6.2 and Section 5.2.6). This is because dispatching rules of multi-load vehicles in this dissertation are similar; based on the concept of closest task.

When loads are released in batches, and vehicles can only transport one load at a time, one or more loads will be left behind and have to wait (unless two or more vehicles are dispatched to the pick up location simultaneously). This additional waiting time increases the average throughput time of loads (see Table 52 on page 123) as the batch sizes increase. For a fixed batch size, increasing the capacity of the vehicle leads to a reduction of the average throughput time. The magnitude of the reduction is stronger for larger batch sizes. This is due to the fact that the opportunity for combining loads on a multi-load vehicle increases for larger batch sizes.

We have also seen that systems with heavily utilized vehicles benefit most from adding (more vehicles and) load capacity to the vehicles (see Table 13). These benefits decrease as the fleet size increases. Increasing the vehicle capacity too much can also work contra-productively. Since increasing the capacity increases the load transportation times, the average load throughput times can also increase (see Table 51).

Using vehicle dwell point strategies and pre-assigning moving vehicles to loads

We also studied different vehicle dwell point strategies for on-line controlled vehicles in Chapter 4. In the standard case, when vehicles are not assigned to retrieve a new load after delivering a load, they are instructed to park at the location of the last delivery (point of release positioning rule). When the dwell point strategies are used, the vehicles are instructed to park at the location such that the response time to the next transportation task is minimized (central zone positioning rule). When a vehicle moving to a parking location must reach the parking point before becoming eligible to pick up a load, it may pass by waiting loads, wasting vehicle capacity. We therefore investigated the dwell point

strategies such that the closest parked or going to park vehicle can be matched to the requesting loads.
In case the load transport requests appear at random locations, the vehicles are instructed to park at the dwell point location which has the minimum average transport distance to any of the pick up locations. This results in different dwell points for the different layouts. In case the load transport requests appear in a structured fashion, the vehicles are instructed to park at the location at which the next transport request is most likely to be placed. The results (Table 29 on page 89) show that considerable reductions in the average load waiting times can be obtained (up to 22%) for both dispatching rules (first-come-first-served and nearest-workstation-first) and both types of layout when the central zone positioning rule is used instead of the point of release positioning rule. The waiting time reductions are most favorable for the Structured shifts where the most likely location of the next transport request can be calculated more accurately compared to Random shifts.
The idea of matching idle and moving to park vehicles to loads has also been extended to assigning loads to any vehicle independent of its status (idle, going to park, retrieving or delivering,). In that case, the (moving) vehicle that can pick up the requesting load soonest, given the current state of the system, is pre-assigned to the load. The time a load can be picked up is calculated considering the time (or distance) still needed by the vehicles to reach the load. The vehicles can in this case have at most two pre-assigned jobs and will always complete the jobs they are executing first before executing the new assigned loads. This vehicle pre-assignment strategy only marginally (0-10%) increased the performance of the system (see Table 29 on page 89) and can also result in the loss of favorable combinations or double-plays such that the average load waiting times slightly increase. A strategy that is more favorable makes use of some load pre-arrival information, this we will discuss next.

Using load pre-arrival information

In Chapter 4, off-line control is also compared with on-line dispatching rules using load pre-arrival information. Although only a small portion of information is given beforehand for on-line control, the extra time, which otherwise would have been vehicle idle time, can then be used by vehicles to travel to the next released load. Intuitively, the use of on-line control with a pre-arrival period about equal to the average vehicle travel time to retrieve loads, reduces load waiting times considerably. In Chapter 4, this was similar to having information beforehand about 0.6 loads per vehicle on average. The most favorable results were also obtained using this much pre-arrival information. However, if too much information is made available beforehand, the load to vehicle allocation can become unfavorable, which in turn increases the average load waiting times. Experiments also show that the unfavorable allocation can become less unfavorable when jobs are suspended, i.e. vehicles which have to wait for loads at pick up locations are instructed to pick up other loads when the expected vehicle waiting time is higher than the average travel time to other pick up locations.

In practice, the use of pre-arrival information is rather restricted due to the stochasticity of events and handling times in the environments. In general, the exact information about the release of the loads can only be given during the last handling activities of the loads before they are released to the transport system, (i.e. just before a crane or conveyor drops off the load). Using realistic pre-arrival information in practice (see the practical cases in Chapter 5) can also significantly reduce average load waiting times. A pre-arrival period at the distribution center of 15 seconds (or 0.06 loads per vehicle on average) reduces the average waiting times about 5% or more. And in the case of the production plant, a pre-arrival period of 0.18 loads per vehicle on average, reduces load waiting times with 65%. The waiting time reductions when pre-arrival information is available decrease as the vehicle utilization increases. Furthermore, the reductions in waiting times when pre-arrival information is available were more favorable with modified first-come-first-served than those obtained with nearest-vehicle-first. Nevertheless, the absolute waiting times were still more favorable with the nearest-vehicle-first rule.

Ranking vehicle dispatching rules

In the previous chapters, a variety of dispatching rules has been analyzed under numerous conditions (layout variations, dedicating vehicles, batch-release of loads, multiple-load vehicle capacity, varying fleet size, using pre-arrival information, dwell point strategies, etc.). It is conceivable that different internal transportation environments need different vehicle dispatching rules to transport the loads with a minimum number of vehicles and a minimum average load waiting time. Vehicles are dispatched for several theoretical cases and three practical cases with the assumption that only real-time information is available. However, the performances of the dispatching rules have also been calculated when some load pre-arrival information is assumed to be known. It is demonstrated in the discussions of the results for each case in chapter 5 that there seems to be a certain ranking for vehicle dispatching rules and that the ranking of some rules seems insensitive for various conditions.

We have studied different vehicle-initiated, load-initiated, priority-list, time-based and distance-based dispatching rules of which three commonly used dispatching rules have been studied for *each* practical case. The first commonly used rule is the load-initiated distance-based nearest-vehicle-first rule which has the advantage that remote areas will be matched to a vehicle even if the vehicles are at the other end of the vehicle operating area, and at real-time level, minimizes vehicle empty travel time, hence minimizing load waiting times.

With the second common dispatching rule, the vehicle-initiated distance-based nearest-workstation-first rule, it is possible that certain loads in remote areas will not be served, since it is possible that vehicles are matched to other loads in the system that are relatively closer.

The third rule, the vehicle-initiated time-based modified first-come-first-served rule, simply serves the next transport request according to the load release times of all loads in

the system. The advantage is that every request will be served. The disadvantage of this rule is that no special regard is given to minimize vehicle empty travel time.
Table 67 shows that the absolute waiting times for all (practical) cases studied are most favorable with the load-initiated distance-based rule followed by the vehicle-initiated distance-based rule and the vehicle-initiated time-based rule. This ranking is also valid when multi-load capacity vehicles are used and when load pre-arrival information is available.

6.2 Further research

The research discussed in the previous section has focussed on certain topics of vehicle-based internal transport systems and has been restricted by certain assumptions and simplifications. In this section, several topics for further research will be presented.

We have seen that routing vehicles with off-line algorithms can lead to considerable reductions in average load waiting times compared to on-line dispatching rules (see Chapters 3 and 4). This is because off-line algorithms calculate vehicle routes beforehand, based on all load information. However, only some of the information is available a short time in advance, which actually makes dispatching vehicles with some pre-arrival information possible.
The first interesting extension is to combine pre-arrival information with off-line algorithms. Vehicles and loads are matched off-line using all the information available in the pre-arrival information period. The idea is to use a rolling horizon in which the algorithm is called after a certain period (smaller or equal to the pre-arrival time) has expanded or after a certain number of loads has been released within the pre-arrival period. It should be noted that the off-line algorithms, such as dynamic PDPTW, should be able to operate in real-time. Although interests about such systems are rising, the algorithms involved are usually based on simple heuristics to make real-time scheduling possible.
A similar idea of a rolling horizon can be used without pre-arrival information. In this case, when transport requests become available, they are not served but temporarily delayed. After a certain period, or until a certain number of requests have been collected, the off-line algorithm is used to match the vehicles to loads. This type of control is only useful if the reduction in waiting times by using the off-line algorithm outweighs the increase in waiting times due to the delay to collect transport requests.
Another interesting extension is to use pre-availability information about *vehicles*. In this case, vehicles that are about to complete their tasks can also be considered in the vehicle to load assignments. This is similar to pre-assigning loads to moving vehicles discussed in Chapter 4. However, it might also be favorable to use dynamic vehicle to load assignments. In that case, the assignments of vehicles and loads can change if they are not yet executed. Dynamic assignments are especially useful with rolling horizons. The disadvantage is that the locations of moving vehicles should be accurately known at all decision time points in order to make re-assignments correctly.

Another topic for further research is the reduction of the variance of load waiting times. In this case, dispatching rules with combinations of time and distance elements might be interesting (such as multiple-attribute dispatching rules). We have already seen that the dispatching rules with distance elements lead to relatively good results. Time elements can be added to reduce maximum waiting times or give more priority to loads that are waiting relatively longer.

Another method to reduce the average and the variance of load waiting times is the clever positioning of idle vehicles. Vehicle dwell point strategies have been studied for small and simple loop problems, and in this dissertation for simple networks, but literature on this subject for large scale practical problems is deficient. Results in Section 4.2.3 have indicated that dispersing idle vehicles over the vehicle network may reduce the average and the variance of load waiting times. Further research on this topic is needed.

In most cases, vehicles will follow a specific route (the shortest or quickest). Control rules that select paths dynamically may also be an interesting area for research. In such cases, vehicles may select an alternative path in case of congestion, a blocked path or an occupied zone. Control rules that select paths dynamically can be based on artificial intelligence, fuzzy logic or neural networks. However, such control rules can be very complex and confusing which may complicate tracking and tracing of vehicles and loads or restarting the operation in case of a malfunction. Similarly, it might also be interesting to let self learning vehicles choose the next transport task from a list with transportation requests collected by a centralized controller.

Finally, another interesting research topic is to integrate vehicle control with the control and schedules of other handling equipment, such as other vehicles, cranes, conveyors, workstations, etc. Multi-modal transport of loads can be efficient if the control of all handling equipment can be integrated and taken care off by one control system. It should be noted that the best result with an integrated system is not the same as connecting the best results obtained when separate control systems are considered individually. Connecting separate control systems can lead to serious deadlocks. Although the interest and literature about integrated systems is increasing, avoiding deadlocks in vehicle systems and integrated systems still pose other interesting subjects in need of further research.

6.3 Guidelines for designing control systems in practice

In this last section of the dissertation we provide some guidelines for designing and selecting control systems in practice. The guidelines are based on results found in literature (see Chapter 2) and results obtained from studies for this dissertation. It should be noted that these guidelines are not intended to represent an exhaustive decision framework or support system. The guidelines are given for situations generally found in practice in which multiple (single-load) vehicles travel in a complex vehicle network to serve load transport request with the objective to minimize average load waiting times. This does not imply that the overall best or optimal dispatching strategy is given for every arbitrary internal transport environment. Selections of sources on which the guidelines are based are mentioned in the last column of Table 68.

The type of control system is highly dependent on the stochasticity of the environment. As more information about the load to be transported is available, more scheduling can be used to dispatch the vehicles, (see the first two columns of Table 68). In environments in which the exact release times of loads are only known at the actual moment that the load is available for transport, for example in (highly) stochastic environments, off-line scheduling rules cannot be used. We have seen that centralized on-line dispatching with the load-initiated distance-based nearest-vehicle-first rule results in the lowest average load waiting times (in fact also the variance in load waiting times; not discussed in Chapter 5). The best alternative to the load-initiated rule is the vehicle-initiated nearest-workstation-first dispatching rule, which is also a distance-based rule.

Similar results are true when information about the load releases is available a moment before the load is physically ready or available for transport, i.e. with load pre-arrival information. If this information is not exact, dispatching vehicles based using scheduling algorithms (with rolling horizons for example) may be less favorable than dispatching vehicles on-line. This was observed in Section 4.2.6 where routes are calculated off-line and vehicles are dispatched while release times are perturbed.

Decisions based on	Type of dispatching	Rules minimizing average load waiting times	Sources (selection)
Load information available real-time	On-line dispatching rules	- Nearest-vehicle-first - Nearest-workstation-first	- This dissertation (Chapter 5) - Klein and Kim (1996) - Faraji and Batta (1994)
Load information available just before load release	On-line dispatching with load pre-arrival information	- Nearest-vehicle-first - Nearest-workstation-first	- This dissertation (Chapters 4 and 5)
Exact load information available some time before load release	Off-line scheduling based on rolling horizons to update assignments	Scheduling rules based on PDPTW such as insertion type heuristics	- Rachamandugu et al. (1986) - Jaw et al. (1986)
Exact information of all loads available before start of operation	Off-line scheduling rules	- PDPTW - Insertion type heuristics	- This dissertation (Chapters 3 and 4) - Dumas et al. (1991) - Savelsbergh and Sol (1995) - Jaw et al. (1986)

Table 68. Guidelines for selecting control systems

Scheduling vehicles with off-line algorithms using pre-arrival information about loads is not necessarily favorable since pre-arrival information is limited and possibly not exact, and off-line algorithms require some calculation effort and exact information. However, if load information is available well before the load is released then vehicle routes can be calculated using a series of off-line algorithms (static models) solved on a rolling horizon basis (Rachamadugu et al., 1986). Thus, another possibility (besides on-line dispatching with pre-arrival information) would be to use exact models such as those used for pick up and delivery problems with time windows (PDPTW) for static models to determine optimal dispatching decisions with the current status of the system. The solution can be updated at regular intervals or whenever a change in the status of the system occurs. As mentioned (Section 6.2) this methodology can be computationally impractical in real-time operations, even with heuristic algorithms such as insertion (see Jaw et al, 1986), and heuristic dispatching rules are preferred by practitioners (Co and Tanchoco, 1991).

The last mentioned situation in which all exact information is available before the vehicles need to be dispatched (see Table 68) is actually not found in practical internal transportation environments. Furthermore, off-line scheduling algorithms often do not consider blocking or congestion in the system. This can change the status of the system similar to situations with perturbed release times (see Section 4.2.6) such that off-line control can be outperformed by relatively simple on-line dispatching rules, such as nearest-workstation-first.

Appendix 1

List of abbreviations

The following abbreviations are used in this dissertation. Text written *italic* is explained elsewhere in the list.

AGV(s)	= Automated Guided Vehicle(s)
AGVS(s)	= Automated Guided Vehicle System(s)
ALV(s)	= Automatic Lifting Vehicle(s)
AMH	= Automated Material Handling
ASC(s)	= Automated Stacking Crane(s)
AS/RS	= Automatic Storage and Retrieval System
C100FCFS	= Closer than 100 m, *FCFS* rule
Cap.	= Capacity
DARP	= Dial and Ride Problem
DC(s)	= Distribution Center(s)
DD	= Dedicated Dispatching
DPI	= Dispatching with Pre-arrival Information
EDC	= European Distribution Center
EDI	= Electronic Data Interchange
ERP	= Enterprise Resource Planning
FCFS	= First-Come-First-Served
FEFS	= First-Encountered-First-Served
FLT	= Forklift Truck
FMS	= Flexible Manufacturing System
GV(s)	= Guided Vehicle(s)
HD	= Hybrid Dispatching
JCV(s)	= Jumbo Container Vessel(s)
JIT	= Just In Time
LL(s)	= Load-List(s)
LLD	= Load-List Dispatching
Max.	= Maximum
MB	= Mega Byte
MHS	= Material Handling System
MIP	= Mixed Integer Programming

Mod. FCFS	= Modified *FCFS*
MOQS	= Maximum-Outgoing-Queue-Size
m-PDPTW	= Multi-vehicle *PDPTW*
MTS(s)	= Multi-Trailer System(s)
m-TSPTW	= Multi *TSP* with Time Windows
NVF	= Nearest-Vehicles-First
NWF	= Nearest-Workstation-First
PDPTW	= Pick up and Delivery Problem with Time Windows
P&D	= Pick-up & Delivery
PLC	= Programmable Logic Controllers
RAM	= Random Access Memory
RF	= Radio Frequency
SHA	= Systematic Handling Analysis
SLP	= Systematic Layout Planning
St. dev.	= Standard Deviation
STDF	= Shortest-Travel-Distance-First
STTF	= Shortest-Travel-Time-First
SWLFD	= Single-Work-List Flow-intensity-based Dispatching
TEU	= Twenty-Foot-Equivalent-Unit
TRACES	= Traffic-Control and Engineering System
TRPTW	= Traveling Repairman Problem with Time Windows
TSP(TW)	= Traveling Salesman Problem (with Time Windows)
VAL	= Value Added Logistics
VLFW	= Vehicle-Looks-For-Work
WL(s)	= Work-List(s)
WLD	= Work-list-Dispatching
WMS	= Warehouse Management System

Appendix 2

Example of a control system data-file

NSPSTP	FLT2	1143907	13-dec	8:06:29	P1- A-02-2-1	D1-04---	D1-04---
NSPSTP	FLT4	1143911	13-dec	8:06:41	P1- F-02-1-1	D1-04---	D1-04---
NFPSTP	FLT2	1143940	13-dec	8:10:57	P1- F-02-2-1	B1-01---	D1-06---
NSPSTP	FLT5	1143941	13-dec	8:11:19	P1- E-02-2-1	D1-04---	D1-04---
NFPSTP	FLT2	1144015	13-dec	8:13:42	B1-01---	D1-02---	D1-02---
NFPSTP	FLT5	1143942	13-dec	8:16:22	P1- F-02-1-1	B1-01---	D1-06---
NFPSTP	FLT5	1144063	13-dec	8:17:04	B1-01---	D1-06---	D1-06---
NSPSTP	FLT2	1144056	13-dec	8:17:58	P1- A-02-2-1	D1-06---	D1-06---
NSPSTP	FLT2	1144064	13-dec	8:23:20	P1- F-02-2-1	B1-01---	D1-02---
LSPRPP	FLT5	1144059	13-dec	8:23:53	P1- E-02-2-1	H1-01---	H1-01---
NFPSTP	FLT2	1144074	13-dec	8:25:18	B1-01---	D1-06---	D1-06---
NFPSTP	FLT5	1144068	13-dec	8:25:20	P1- E-02-1-1	B1-01---	D1-06---
NSPSTP	FLT5	1144078	13-dec	8:27:16	P1- F-02-3-1	D1-09---	D1-09---
NFPSTP	FLT2	1144085	13-dec	8:28:32	P1- F-02-1-1	B1-01---	D1-07---
NSPSTP	FLT2	1144082	13-dec	8:30:33	P1- E-02-3-1	D1-06---	D1-06---
NSPSTP	FLT5	1144094	13-dec	8:30:48	P1- F-02-2-1	D1-07---	D1-07---
NFPSTP	FLT5	1144095	13-dec	8:33:35	P1- E-02-2-1	B1-01---	D1-09---
NFPSTP	FLT5	1144106	13-dec	8:34:22	B1-01---	D1-06---	D1-06---
LFPRPP	FLT2	1144102	13-dec	8:35:30	P1- F-02-3-1	H1-01---	H1-01---
NSPSTP	FLT2	1144093	13-dec	8:37:34	P1- A-02-1-1	B1-01---	D1-02---
NSPSTP	FLT2	1144110	13-dec	8:38:40	B1-01---	D1-02---	D1-02---
NFPSTP	FLT5	1144105	13-dec	8:39:31	P1- F-02-1-1	B1-01---	D1-02---
NSPSTP	FLT1	1144107	13-dec	8:40:54	P1- E-02-1-1	B1-01---	D1-09---
NFPSTP	FLT2	1144116	13-dec	8:41:18	P1- E-02-3-1	B1-01---	D1-07---
NFPSTP	FLT5	1144113	13-dec	8:41:27	B1-01---	D1-07---	D1-07---
NFPSTP	FLT1	1144126	13-dec	8:41:28	B1-01---	D1-09---	D1-09---
NSPSTP	FLT2	1144119	13-dec	8:43:28	P1- F-02-2-1	B1-01---	D1-02---
NFPSTS	FLT1	1143851	13-dec	8:44:05	D1-14---	P1- C-01-1-1	P1- C-20-4-1
NFPSTS	FLT5	1143852	13-dec	8:44:36	D1-14---	P1- C-01-2-1	P1- C-35-4-3
NSPSTP	FLT5	1144149	13-dec	8:45:39	P1- A-02-1-1	B1-01---	D1-09---
NFPSTP	FLT5	1144147	13-dec	8:46:23	B1-01---	D1-02---	D1-02---
NSPSTP	FLT1	1144122	13-dec	8:48:12	P1- A-02-2-1	B1-01---	D1-02---

NFPSTS	FLT5	1143856	13-dec	8:48:31	D1-11---	P1- C-01-1-1	P1- C-27-3-2
NSPSTP	FLT1	1144155	13-dec	8:48:59	B1-01---	D1-09---	D1-09---
NSPSTP	FLT2	1144143	13-dec	8:49:14	B1-01---	D1-02---	D1-02---
NSPSTP	FLT5	1144145	13-dec	8:49:58	P1- C-02-2-1	D1-04---	D1-04---
NFPSTS	FLT2	1143854	13-dec	8:51:40	D1-11---	P1- D-01-1-1	P1- D-16-5-2
NFPSTS	FLT5	1143891	13-dec	8:55:19	D1-11---	P1- F-01-1-1	P1- F-32-3-3
NFPSTP	FLT5	1144157	13-dec	8:56:27	P1- A-02-3-1	B1-01---	D1-07---
NSPSTP	FLT2	1144150	13-dec	8:57:16	P1- E-02-1-1	D1-07---	D1-07---

Explanation:

- The first column represents the type of job or transportation task. Where:
  - HSPSPP is a High-priority Sub Pallet Special Pick
  - LFPRPP is a Low-priority Full Pallet Special Pick
  - LSPRPP is a Low-priority Sub Pallet Replenishment Pick
  - NFPSTP is a Normal-priority Full Pallet Standard Pick
  - NSPSTP is a Normal-priority Sub Pallet Standard Pick
- The second column indicates the (number of the) vehicle which is assigned to the task.
- The third column indicates the task number. Numbers are assigned when the tasks are generated (born), including all tasks by cranes, conveyors etc. Therefore some tasks are not listed in the vehicle data-file activity-list.
- The fourth and fifth column represent the release date (day-month) and time (hour:min:sec) of the load (transportation task for the vehicle).
- The last three columns represent the origin, destination and final destination of the load respectively. Where:
  - The locations starting with B represent the labeling area.
  - The locations starting with C represent the check-in area.
  - The locations starting with D represent the docking lanes at the shipping and receiving area.
  - The locations starting with H1 represent the shelf replenishment area (SRA).
  - The locations starting with K2 represent the repalletization area (RPA).
  - The locations starting with R1 represent the return stations.
  - The locations starting with R2 represent the central return area (CRA).
  - The locations starting with P1-A represent the pick-up and delivery stations of storage module 1. For example, P1-A-02-2-1 is the location at storage module 1, Aisle A, section 02, height 2, and slot 1.
  - The locations starting with P2-A represent the odd-size area.
  - The locations starting with P2-B represent the overflow area.

Note: the destination can be different from the final destination if the load also has to be handled by a crane (to be stored) or if the load has to be picked up later from the labeling station (B1-01---) and transported to one of the docking lanes in the shipping area.

References

Anwar, M.F., and Nagi, R., 1998, Integrated scheduling of material handling and manufacturing activities for just-in-time production of complex assemblies, International Journal of Production Research, Vol. 36, No. 3, pp. 635-681.

Asef-Vaziri, A., Laporte, G., and Sriskandarajah, C., 2000, The block layout shortest loop design problem, IIE Transactions, Vol. 32, pp. 727-734.

Ashayeri, J., and Gelders, L.F., 1985, Warehouse design optimalization, European Journal of Operational Research, Vol. 21, pp. 285-294.

Askin, R.R., and Standridge, C.R., 1993, Modeling and analysis of manufacturing systems, John Wiley & Sons, Inc, pp. 307-323.

AutoMod™ by AutoSimulations™, Inc., 655 Medical Drive, Bountiful, Utah 84010.

Ball, M.O., Magnanti, T.L., Monma, C.L., and Nemhauser, G.L., 1995, Network routing : Handbooks in operations research and management science, Elsevier, Vol. 8, ISSN 0927-0507.

Ballis, A., and Abacoumkin, C., 1996, A container terminal simulation model with animation capabilities, Journal of Advanced Transportation, Vol. 30, No. 1, pp. 37-57.

Bartholdi III, J.J., and Platzman, L.K., 1989, Decentralized control of automated guided vehicles on a simple loop, IIE Transactions, Vol. 21, No. 1, pp. 76-81.

Bilge, Ü., Tanchoco, J.M.A., 1997, AGV systems with multi-load carriers: Basic issues and potential benifits, Journal of Manufacturing Systems, Vol. 16, No. 3, pp. 159-174.

Bischak, D.P., and Stevens, K.B., Jr, 1995, An evaluation of the tandem configuration automated guided vehicle system, Production Planning & Control, Vol. 6, No. 5, pp. 438-444.

Blackstone, J.H., Phillips, D.T., and Hoff, G.L., 1982, A state-of-the-art survey of dispatching rules for manufacturing job shop operations, International Journal of Production Research, Vol. 20, No. 1, pp. 27-45.

Blair, E.L., and Vasques, A., 1987, Optimal routing of driverless vehicles in a flexible material handling system, Material Flow, Vol. 4, pp. 73-83.

Bostel, A.J. and Sagar, V.K., 1996, Dynamic control systems for AGVs, Computing and Engineering Journal, pp. 169-176.

Bozer, Y A., Cho, M., and Srinivasan, M M., 1994, Expected waiting times in single-device trip-based material handling systems, European Journal of Operational Research, Vol. 75, pp. 200-216.

Bozer, Y.A., and Srinivasan, M.M., 1989, Tandem configurations for AGV systems offer simplicity and flexibility, Industrial Engineering, Vol. 2, pp. 23-27.

Bozer, Y.A., and Srinivasan, M.M., 1991, Tandem configurations for automated guided vehicle systems and the analysis of single vehicle loops, IIE Transactions, Vol. 23, No. 1, pp. 72-82.

Bozer, Y.A., and Srinivasan, M.M., 1992, Tandem AGV systems: a partitioning algoritm and performance comparison with conventional AGV systems, European Journal of Operational Research, Vol. 63, pp 173-191.

Celen, H.P, and Leijn, R.A., 1996, Europe combined terminals BV, Delta sea-land terminal, een geautomatiseerd openluchtmagazijn van 60 ha., In: Praktijkboek Magazijnen / Distributiecentra, J.P. Duijker et al. (Eds.), Kluwer Bedrijfswetenschappen, Deventer, pp. 3.14.D-01-18. (In Dutch)

Celen, H.P., Slegtenhorst, R.J.W., Van der Ham, R. TH., Nagel, A., Van den Berg, J., De Vos Burchart, R., Evers, J.J.M., Lindeijer, D.G., Dekker, R., Meersmans, P., De Koster, M.B.M., Van der Meer, J.R., Carlebur, A.F.C., Nooijen, F.J.A.M., 1997, FAMAS, NewCon definitiestudie, CTT Publication 32, ISBN 90-76091-36-6. (In Dutch).

Chang, S., and Egbelu, P.J., 1996, Dynamic relative positioning of AGVs in a loop layout to minimize mean system response time, International Journal of Production Research, Vol. 34, No. 6, pp. 1655-1673.

Christofides, N., Mingozzo, A., and Toth, P., 1981, Exact algorithms for the vehicle routing problem, based on spanning tree and shortest path relaxations, Mathematical Programming, Vol. 20, pp. 255-282.

Co, C.G., and Tanchoco, J.M.A., 1991, A review of research on AGVS vehicle management, Engineering Cost and Production Economics, Vol. 21, pp. 35-42.

Daganzo, C.F., 1989, The crane scheduling problem, Transportation Research B, Vol. 23, No. 3, pp. 159-175.

De Castilho, B., and Daganzo, C.F., 1993, Handling strategies for import containers at marine terminals, Transportation Research B, Vol. 27, No. 2, pp. 151-166.

Dekker, R., Piersma, N., de Bruin, J.J., and Van der Vlist, E., 1995, Operations research in action for mainport Rotterdam, an investigation of routing AGVs at ECT, Kwantitatieve Methoden, Vol. 16, pp. 5-18. (In Dutch)

De Koster, M.B.M., 1995a, Intern transportmaterieel, In: Praktijkboek Magazijnen / Distributiecentra, J.P. Duijker et al. (Eds.), Kluwer Bedrijfswetenschappen, Deventer, Ch. 3.7, pp. 1-4. (In Dutch)

De Koster, M.B.M., 1995b, Rijdend intern transportmaterieel, In: Praktijkboek Magazijnen / Distributiecentra, J.P. Duijker et al. (Eds.), Kluwer Bedrijfswetenschappen, Deventer, pp. 3.7.A-01-26. (In Dutch)

De Koster, M.B.M., and Van der Meer, J.R., 1998, Centralized versus decentralized control of internal transport, a case study, In: Advances in Distribution logistics, B. Fleischmann, J.A.E.E. Van Nunen, M. Grazia Speranza and P. Stähly (Eds.), Springer (Berlin), p. 403-420.

De Koster M.B.M., and Van dern Broek-Serlé, F.N., 1999, WMS in further detail, In: Standard software for warehouse management systems, F. Van den Broek (Ed.), PricewaterhouseCoopers.

Desrochers, J., Lenstra, J.K., Savelsbergh, M.W.P., and Soumis, F., 1988, Vehicle routing with time windows: optimization and approximation, In: Vehicle Routing, B.L. Golden and A.A. Assad (Eds.), Elsevier Science Publishers, pp. 65-84.

Desrosiers, J., Sauvé, M., and Soumis, F., 1988, Lagrangian relaxation methods for solving the minimum fleet size multiple traveling salesman problem with time windows, Management Science, Vol. 34, No. 8, pp. 1005-1022.

Duinkerken, M.B., Ottjes, J.A., Evers, J.J.M., Kurstjens, S.T.G.L., Dekker, R., and Dellaert, N.P., 1996, Simulation studies on inter terminal transport at the Maasvlakte, TRAIL-studies, Nr. 96/3, ISBN 90-5584-020-3.

Dumas, Y., Desrosiers, J., and Soumis, F., 1991, The pickup and delivery problem with time windows, European Journal of Operational Research, Vol. 54, pp. 7-22.

Egbelu, P.J., 1987, The Use of non-simulation approaches in estimationg vehicle requirements in an automated guided vehicle based transport system, Material Flow, Vol. 4, pp. 17-32.

Egbelu, P.J., 1993, Positioning of automated guided vehicles in a loop layout to improve response time, European Journal of Operational Research, Vol. 71, No. 1, pp. 32-44.

Egbelu, P.J., and Tanchoco, J.M.A., 1984, Characterization of automated guided vehicle dispatching rules, International Journal of Production Research, Vol. 22, No. 3, pp. 359-374.

Egbelu, P.J., and Tanchoco, J.M.A., 1986, Potentials for bi-directional guide-path for automated guided vehicle based systems, International Journal of Production Research, Vol. 24, No. 5, pp. 1075-1097.

Erickson, B.E., 1987, Electronic diagnostics for enhanced automatic guided vehicle safety and maintainability, Material Flow, vol. 4, pp. 109-115.

Evers, J.J.M., and Koppers, S.A.J., 1996, Automated guided vehicle traffic control at a container terminal, Transportation Research A, Vol. 30, No. 1, pp. 21-34.

Faraji, M., and Batta, R., 1994, Formaing cells to eliminate vehicle interference and system locking in an AGVS, International Journal of Production Research, Vol. 32, No. 9, pp. 2219-2241.

Gademann, N., and Van de Velde, S., 2000, Positioning automated guided vehicles in a loop layout, To appear in: European Journal of Operational Research, Dec. issue.

Ganesharajah, T., and Sriskandarajah, H., 1995, Survey of scheduling research in AGV-served manufacturing systems, Advances in Instrumentation and Control; proceedings of the international conference and exhibit, Vol. 50, No. 1, pp. 87-94.

Gaskins, R.J., and Tanchoco, J.M.A., 1987, Flow path design for automated guided vehicle systems, International Journal of Production Research, Vol. 25, No. 5, pp. 667-676.

Gaskins, R.J., Tanchoco, J.M.A., and Taghaboni, F., 1989, Virtual flow paths for free-ranging automated guided vehicle systems, International Journal of Production Research, Vol. 27, No. 1, pp. 91-100.

Gendreau, M., Hertz, A., and Laporte, G., 1992, New insertion and postoptimization procedures for the traveling salesman problem, Operations Research, Vol. 40, No. 6, pp. 1086-1094.

Gobal, S.L., and Kasilingam, R.G., 1991, A simulation-model for estimating vehicle requirements in automated guided vehicle systems, Computers and Industrial Engineering, Vol. 21, No. 1-4, pp. 623-627.

Gunasekaran, A., and Lyu, J., 1997, Implementation of just-in-time in a small company; a case study, Production Planning and Control, Vol. 8, No. 4, pp. 406-412.

Hao, G., and Lai, K.K., 1996, Solving the AGV problem via a self-organizing neural network, Journal of Operational Research Society, Vol. 47, pp. 1477-1493.

Heragu, S.S., and Gupta, Y.P., 1994, A heuristic for designing cellular manufacturing facilities, International Journal of Production Research, Vol. 32, No. 1, pp. 125-140

Hodgson, T.J., King, R.E., Monteith, S.K., Schultz, S.R., 1987, Developing control rules for an AGVS using Markov decision processes, Material Flow, Vol. 4, No. 1-2, pp. 85-96.

Hsieh, L.F., and Sha, D.Y., 1997, Heuristic algorithm for the design of facilities layout and AGV routes in tandem AGV systems, International Journal of Industrial Engineering, Vol. 4, No. 1, pp. 52-61.

Hwang, H., and Kim, S.H., 1998, Development of dispatching rules for automated guided vehicle systems, Journal of Manufacturing Systems, Vol. 17, No. 2, pp. 137-143.

Jacobs, F.H.W.M., Van Dort, H.E., and De Koster, M.B.M., 1995, Informatiesystemen binnen magazijnen, In: Praktijkboek Magazijnen / Distributiecentra, J.P. Duijker et al. (Eds.), Kluwer Bedrijfswetenschappen, Deventer, pp. 3.11.A-01-36. (In Dutch)

Jaikumar, R., and Solomon, M.M., 1990, Dynamic scheduling of automated guided vehicles for a certain class of systems, Journal of Manufacturing Systems, Vol. 9, No. 4, pp. 315-323.

Jaw, J., Odini, A., Psaraftis, H., and Wilson, N., 1986, A heuristic algorithm for the multi-vehicle advance-request dail-a-ride problem with time windows, Transportation Research B, Vol. 20, No. 3, pp. 243-246.

Jawahar, N., Aravindan, P., Ponnambalam, S.G., and Suresh, R.K., 1998, AGV schedule integrated with production in flexible manufacturing systems, International Journal of Advanced Manufacturing Technology, Vol. 14, No. 6, pp. 428-440.

Johnson, M.E., and Brandeau, M.L., 1993, An analytical model for design of a multivehicle automated guided vehicle system, Management Science, Vol. 39, No. 12, pp. 1477-1489.

Kasilingam, R.G., and Gobal, S.L., 1996, Vehicle requirements model for automated guided vehicle systems, International Journal of Advanced Manufacturing Technology, Vol. 12, 276-279.

Kaspi, M., and Tanchoco, J.M.A., 1990, Optimal flow path design of unidirectional AGV systems, International Journal of Production Research, Vol. 28, No. 6, pp. 1023-1030.

Kim, K.H., 1995, Positioning of automated guided vehilces in a loop layout to minimize the mean vehicle response time, International Journal of Production Economics, Vol. 39, No. 3, pp. 201-214.

Kim, K.H., 1997, Evaluation of the number of rehandles in container yards, Computers and Industrial Engineering, Vol. 32, No. 4, pp. 701-711.

Kim, K.H., and Bae, J.W., 1999, A dispatching method for automated guided vehicles to minimize delays of containership operations, International Journal of Management Science, Vol. 5, No. 1, pp. 1-25.

Kim, K.H., and Kim, J.Y., 1997, Estimating mean response time and positioning idle vehicles of automated guided vehicle systems in loop layout, Computers and Industrial Engineering, Vol. 33, No. 3-4, pp. 669-672.

Kim, K.H., and Kim, H.B., 1999a, Segregating space allocation models for container inventories in port container terminals, International Journal of Production Economics, Vol. 59, No. 1-3, pp. 415-423.

Kim, K.H., and Kim, K.Y., 1999b, Routing straddle carriers for the loading operation of containers using a beam search algorithm, Computers and Industrial Engineering, Vol. 36, No. 1, pp. 109-136.

Kim, C.W., and Tanchoco, J.M.A., 1993, Operational control of a bidirectional automated guided vehicle system, International Journal of Production Research, Vol. 31, No. 9, pp. 2123-2138.

Kim, C.W., and Tanchoco, J.M.A., and Koo, P.H., 1999, AGV dispatching based on workload balancing, International Journal of Production Research, Vol. 37, No. 17, pp. 4053-4066.

King, R.E., and Wilson C., 1991, A review of automated guided-vehicle sytems design and scheduling, Production Planning & Control, Vol. 2, No. 1, pp. 44-51.

Klein, C.M, and Kim, J., 1996, AGV dispatching, International Journal of Production Research, Vol. 34, No. 1, pp. 95-110.

Kochhar, J.S., and Heragu, S.S., 1999, Facility layout design in a changing environment, International Journal of Production Research, Vol. 37, No. 1, pp. 2429-2426.

Kodali, R., 1997, Knowledge-based systems for selection of an AGV and a workcenter for transport of a part in on-line scheduling of FMS, Production Planning & Control, Vol. 8, No. 2, pp. 114-122.

Koff, G.A., 1987, Automatic guided vehicle systems: Applications, controls and planning, Material Flow, Vol. 4, pp. 3-16.

Kolen, A., Rinnooy Kan, A., and Trienekens, H., 1987, Vehicle routing with time windows, Operations Research, Vol. 35, pp. 266-273.

Krishnamurthy, N.N., Batta, R., and Karwan, M.H., 1993, Developing conflict-free routes for automated guided vehicles, Operations Research, Vol. 41, pp. 1077-1090.

Kulwiec, R. A., Materials handling handbook, John Wiley, New York.

Kurstjens, S.T.G.L., Dekker, R., Dellaert, N.P., Duinkerken, M.B., and Evers, J.J.M., 1996, Planning inter terminal tranport at the Maasvlakte, TRAIL-studies, nr. 96/3, ISBN 90-5584-020-3.

Larson, T.N., March, H., and Kusiak, A., 1997, A heuristic approach to warehouse layout with class-based storage, IIE Transactions, Vol. 29, pp. 337-348.

Lee, C., and Ventura, J.A., 1999a, Optimal idle vehicle positioning in a loop automated guided vehicle system to minimize the mean weighted response time. *Working*

Paper, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA.

Lee, C., and Ventura, J.A., 1999b, Design of tandem loop multiple vehicle layouts for automated guided vehicle systems. *Working Paper*, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA.

Lee, C., and Ventura, J.A., 2000, Deadlock analysis of flexible manufacturing systems using a Petri net model. *Working Paper*, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA.

Leung, L.C., and Khator, S.K., and Kimbler, D.L., 1987, Assignment of AGVs with different vehicle types, Material Flow, Vol. 4, pp. 65-72.

Lindeijer, D.G., and Evers, J.J.M., 1999, TRACES the agile traffic-control and engineering system, In: Proceedings of the 5th TRAIL congress 1999, P.H.L. Bovy (Ed.), Delft, The Netherlands, ISBN 90-407-1996-9.

Lu, X., and Gerchak, Y., 1998, Minimizing the expected response time of an idled server on a line, IIE Transactions, Vol. 30, No. 4, pp. 401-408.

Mahadevan, B., and Narendran, T.T., 1990, Design of an automated guided vehicle-based material handling system for a flexible manufacturing system, International Journal of Production Research, Vol. 28, No. 9, pp. 1611-1622.

Mahadevan, B., and Narendran, T.T., 1993, Estimation of number of AGVs for an FMS: an analytical model, International Journal of Production Research, Vol. 31, No. 7, pp. 1655-1670.

Mahadevan, B., and Narendran, T.T., 1994, A hybrid modelling approach to the design of an AGV-based material handling system, International Journal of Production Research, Vol. 32, No. 9, pp. 2015-2030.

Malmborg, C.J., 1990, A model for the design of zone control automated guided vehicle systems, International Journal of Production Research, Vol. 28, No. 10, pp. 1741-1758.

Malmborg, C.J., 1991, Tightened analytical bounds on the impact of dispatching rules in automated guided vehicle systems, Applications in Mathematical modelling, Vol. 15, pp. 305-311.

Mantel, R.J., and Landeweerd, H.R.A., 1995, Design and operational control of an AGV system, International Journal of Production Economics, Vol. 41, No. 1-3, pp. 257-266.

Matson, J.O., White, J.A., 1982, Operational research and material handling, European Journal of Operational Research, Vol. 11, pp. 309-318.

Maxwel, W.L., and Muckstadt, J.A., 1982, Design of automated guided vehicle systems, IIE Transactions, Vol. 14, No. 2, pp. 114-124.

Miller, C., Tucker, A., and Zemlin, R., 1960, Integer programming formulations and traveling salesman problems, Journal of the ACM, Vol. 7, pp. 326-329.

Montazeri, M., and Van Wassenhove, L.N., 1990, Analysis of scheduling rules for an FMS, International Journal of Production Research, Vol. 28, No. 4, pp. 785-802.

Müller, T., 1983, Automated guided vehicles, IFS (Publications) Ltd, UK, Springer-verlag Berlin, Heidelberg, New York.

Muther, R., 1973, Systematic layout planning 2nd ed., Cahners Books, Boston.

Nakano, M., and Ohno, K., Decomposition algoritm for performance evaluation of AGV systems, Production and Operations Management, Vol. 8, pp. 193-205.

Narasimhan, R., Batta, R., and Karwan, M.H., 1999, Routing automated guided vehicles in the presence of interruptions, International Journal of Production Research, Vol. 37, No. 3, pp. 653-681.

Nayyar, P., and Khator, S., 1993, Operational control of multi-load vehicles in an automated guided vehicle system, Computers and Industrial Engineering, Vol. 25, No. 1-4, pp. 503-506.

Newton, D., 1985, Simulation model helps determine how many automated guided vehicles are needed, Industrial Engineering, Vol. 17, No. 2, pp. 68-78.

Occeña, L.G., and Yokota, T., 1991, Modelling of an automated guided vehicle system (AGVS) in a just-in-time (JIT) environment, International Journal of Production Research, Vol. 29, No. 3, pp. 495-511.

Özden, M., 1988, A simulation study of multiple-load-carrying automated guided vehicles in a flexible manufacturing sytem, International Journal of Production Research, Vol. 26, No. 8, pp. 1353-1366.

Panwalkar, S.S. and Iskander, W., 1977, A survey of scheduling rules, Operations Research, Vol. 25, No. 1, pp. 45-61.

Peterkofsky, R.I., and Daganzo, C.F., 1990, A branch and bound solution method for the crane scheduling problem, Transportation Research B, Vol. 24, No. 3, pp. 159-172.

Rachamadugu, R.V., Raman, N., and Talbot, F.B., 1986, Real time scheduling of an automated manufacturing center, Real Time Optimization in Automated Manufacturing Facilities, NBS, Gaithesburg, Maryland, pp. 293-315.

Rajagopalan, S., and Heragu, S.S., 1997, Advances in discrete material handling system design, Sadhanna-academy Proceedings in Engineering Sciences Vol. 22, No. 2, pp. 281-292.

Rajotia, S., Shanker, K., and Batra, J.L., 1998, Determination of optimal AGV fleet size for an FMS, International Journal of Production Research, Vol. 36, No. 5, pp. 1177-1198.

Ross, E.A., Mahmoodi, F., and Mosier, C.T., 1996, Tandem configuration automated guided vehicle systems: A comparative study, Decision Sciences, Vol. 27, No. 1, pp. 81-102.

Sabuncuoglu, I., 1998, A study of scheduling rules of flexible manufacturing systems: a simulation approach, International Journal of Production Reseach, Vol. 36, No. 2, pp. 527-546.

Savelsbergh, M.W.P., and Sol, M., 1995, The general pickup and delivery problem, Transportation Science, Vol. 29, No. 1, pp. 17-29.

Shelton, D., and Jones, M.S., 1987, A selection method for automated guided vehicles, Material Flow, Vol. 4, pp. 97-107.

Shields, J.J., 1984, Container stowage: a computer aided pre-planning system, Marine Technology, Vol. 21, No. 4, pp. 370-383.

Sinriech, D., and Tanchoco, J.M.A., 1992a, An economic model for determining AGV fleet size, International Journal of Production Research, Vol. 30, No. 6, pp. 1255-1268.

Sinriech, D. and Tanchoco, J.M.A., 1992b, Impact of empty vehicle flow on performance of single-loop AGV systems, International Journal of Production Research, Vol. 30, No. 10, pp. 2237-2252.

Sinriech, D. and Tanchoco, J.M.A., 1993, Solution methods for the mathematical models of single-loop AGV systems, Inernational Journal of Production Research, Vol. 31, pp. 705-725.

Sinriech, D., 1995, Network design models for discrete material flow systems: A literature review, International Journal of Advanced Manufacturing Technology, Vol. 10, pp. 277-291.

Sinreich, D., and Palni, L., 1998, Scheduling pickup and deliveries in a multiple-load discrete carrier environment, In: Progress in Material Handling Research, Graves et al. (Eds.), Ann Arbor, Michigan, pp. 521-550, ISBN 1-882780-03-5.

Sinriech, D., Tanchoco, J.M.A, and Herer, Y.T., 1996, The segmented bidirectional single-loop topology for material flow systems, IIE Transactions, Vol. 28, pp. 40-54.

Solomon, M.M., 1987, Algorithms for the vehicle routing and scheduling problems with time window constraints, Operations Research, Vol. 35, No. 2, pp. 254-265.

Solomon, M.M., and Desrosiers, J., 1988, Time window constrained routing and scheduling problems, Transportation Science, Vol. 22, No. 1, pp. 1-13.

Srinivasan, M.M., Bozer, Y.A., and Cho, M., 1994, Trip-based material handling systems: Throughput capacity analysis, IIE Transactions, Vol. 26, No. 1, pp. 70-89.

Stad, H., 2000, Magazijntrucks niet aan te slepen, In: Logistiekkrant, Vol. 13, No. 1. (In Dutch)

Taghaboni, F., and Tanchoco, 1988, J.M.A., A LISP-based controller for free ranging automated guided vehicle systems, International Journal of Production Research, Vol. 26, No. 2, pp. 173-188.

Tanchoco, J.M.A. and Sinriech, D., 1992, OSL - optimal single-loop guide paths for AGVS, International Journal of Production Research, pp. 665-681.

Tanchoco, J.M.A., Egbelu, P.J., and Taghaboni, F., 1987, Determination of the total number of vehicles in an AGV-based material transport sytem, Material Flow, Vol. 4, pp. 33-51.

Thonemann, U.W., and Brandeau, M.L., 1996, Designing a single-vehicle automated guided vehicle system with multiple load capacity, Transportation Science, Vol. 30, pp. 351-363.

Thonemann, U.W., and Brandeau, M.L., 1997, Designing a zoned automated guided vehicle system with multiple vehicles and multiple load capacity, Operations Research, Vol. 45, No. 6, pp. 857-873.

Tompkins, J.A., White, J.A., Bozer, Y.A., Frazelle, E.H., Tanchoco, J.M.A., and Trevino, J., 1996, Facilities Planning, 2nd ed., John Wiley & Sons, Inc., New York.

Van den Berg, J., 1996, Planning and control of warehousing systems, Ph.D. Thesis, University of Twente.

Van der Meer, J.R., 1999, Operational control of internal transport, In: European Logistics Association Doctorate Workshop, T. Skjøtt-Larsen, France (Ed.), p. 75-89, ISBN 3-9804583-8-5.

Van der Meer, J.R., and De Koster, R., 1998, A classification of control systems for internal transport, In: Progress in Material Handling Research, Graves et al. (Eds.), Ann Arbor, Michigan, p. 633-650, ISBN 1-882780-03-5.

Van der Meer, J.R., and De Koster, R., 1999a, On-line versus off-line vehicle control for internal transport systems, In: Proceedings of the 5th TRAIL Congress 1999, P.H.L. Bovy (Ed.), Delft, The Netherlands, ISBN 90-407-1996-9.

Van der Meer, J.R., and De Koster, R., 1999b, Routing vehicles on-line with forecasted release times to approach off-line performance, In: Proceedings of the 5th TRAIL Congress 1999, P.H.L. Bovy (Ed.), Delft, The Netherlands, ISBN 90-407-1996-9.

Van der Meer, J.R., and De Koster, R., 1999c, Using multiple load vehicles for internal transport with batch arrivals of loads, In: Advances in Distribution Logistics, M. Grazia Speranza and P. Stähly (Eds.), Springer (Berlin), p. 197-214.

Van der Meer, J.R., and De Koster, R., 2000, Centralized internal transport systems analyzed and classified, In: Progress in Material Handling Research, Ann Arbor, Michigan.

Van der Meer, J.R., Vis, I., De Koster, R., Meersmans, P., and Dekker R., 1999 De containerterminal in het FAMAS/NEWCON concept, In: SIMLOG Simulatie en logistiek rond de haven, J. Van Nunen, L. Verspui (Eds.), Eburon, Delft, pp. 89-104, ISBN 90-5166-720-5. (In Dutch)

Ventura, J.A., and Lee, C., 1999, Optimal positioning of automated guided vehicles to minimize the maximum response time in a single loop guide path system. *Working Paper*, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA.

Ventura, J.A., and Lee, C., 2000, A simulation study of idle vehicle positioning policies for automated guided vehicle systems, In: Progress in Material Handling Research, Ann Arbor, Michigan.

Vis, I.F.A. and De Koster, R., 2000, Transshipment of containers at a container terminal: an overview, Management Report, Erasmus University Rotterdam.

Wang, H.P.B., and Hafeez, S.A., 1994, Performance evaluation of tandem and conventional AGV systems using generalized stochastic Petri nets, International Journal of Production Research, Vol. 21, No. 4, pp. 917-932.

Wu, N.Q., 1999, Necessary and sufficient conditions for deadlock-free operation in flexible manufacturing systems using a colored Petri net model, IEEE Transactions on Systems, Man and Cybernetics, Part C-applications and reviews, Vol. 29, No. 2, pp. 192-204.

Wysk, R.A., Egbelu, P.J., Zhou, C., and Ghosh, B.K., 1987, Use of spread sheet analysis for evaluating AGV systems, Material Flow, Vol. 4, pp. 53-64.

Yim, D.S., and Linn, R.J., 1993, Push and pull rules for dispatching automated guided vehicles in a flexible manufacturing system, International Journal of Production Research, Vol. 31, No. 1, pp. 43-57.

Zeng, L., Wang, H.P.B., and Jin, S., 1991, Conflict detection of automated guided vehicles: a Petri net approach, International Journal of Production Research, Vol. 29, No. 5, pp. 865-879.

Samenvatting
(Summary in Dutch)

In dit proefschrift worden besturingsregels bestudeerd voor geleide voertuigen in voertuiggebaseerde interne transportsystemen. Deze interne transportsystemen zijn te vinden in opslag-, overslag- en distributiecentra.
Het onderzoek begint in Hoofdstuk 1 met een introductie en een overzicht van deze transportsystemen en plaatst het onderwerp van het proefschrift in dit kader.
Hoofdstuk 2 geeft een uitgebreid overzicht van de in de literatuur besproken onderwerpen gerelateerd aan voertuiggebaseerde interne transportsystemen. De besproken onderwerpen behandelen onder andere: prestatiecriteria, lay-out ontwerp, schatten van het benodigde aantal voertuigen, voertuigpositioneringstrategieën, en on-line (direct verbonden met een centrale verwerkingseenheid voor onvertraagde besturing) en off-line (niet direct verbonden) voertuigbesturingssystemen.
In Hoofdstuk 3 richt het onderzoek zich op het mathematisch modelleren van voertuiggebaseerde interne transportsystemen. De lay-outs van twee opslag magazijnen (een U-lay-out en een I-lay-out) worden gemodelleerd en de interne transportactiviteiten worden geformuleerd als een soort handelsreizigersprobleem met tijdvensters. De idee is om de prestatie van on-line besturingsregels te vergelijken met off-line besturing. Uit de resultaten blijkt dat voor verschillende bestudeerde lay-outs, behoorlijke reducties in de gemiddelde wachttijd voor ladingen te realiseren zijn voor off-line besturing met relatief weinig transportopdrachten per tijdseenheid (lage doorzet). Wachttijden van ladingen kunnen in zulke gevallen gereduceerd worden, aangezien de voertuigen naar de volgende ladingen kunnen rijden voordat de ladingen werkelijk fysiek aanwezig en vrijgegeven zijn. In omgevingen met een lage doorzet, kan het aantal benodigde voertuigen met on-line besturing 50% hoger zijn dan met off-line besturing om vergelijkbare resultaten te halen in dezelfde situatie. In omgevingen met een hoge doorzet (relatief veel transportopdrachten per tijdseenheid) is de mogelijkheid om wachttijden van ladingen te verkleinen veel beperkter, aangezien de voertuigen toch al voordurend in beweging zijn. Ook blijkt dat omgevingen met zwaar bezette voertuigen het meeste voordeel hebben van een verhoging van de voertuigcapaciteit. Het voordeel neemt af naarmate de capaciteit en het aantal voertuigen toeneemt.

Aangezien exacte voorinformatie over de tijd van vrijgave, oorsprong en bestemming van ladingen in interne transportsystemen in vrijwel geen enkel geval beschikbaar is, is het plannen van voertuigen een dag van tevoren, zoals aangenomen in Hoofdstuk 3, vrijwel onmogelijk. In hoofdstuk 4 wordt het model van Hoofdstuk 3 uitgebreid. Off-line besturing wordt vergeleken met on-line besturingsregels met gedeeltelijke voorinformatie van ladingen. Alhoewel slechts kleine beetjes voorinformatie beschikbaar zijn (enkele seconden), kan deze extra tijd, die anders werkloze tijd voor de voertuigen was, gebruikt worden om naar de volgende opdracht (lading) te rijden. Het gebruik van on-line besturingsregels met een voorinformatie periode die ongeveer gelijk is aan de gemiddelde transporttijd naar de ladingen, kan de gemiddelde wachttijden van de ladingen zodanig reduceren dat deze vergelijkbaar worden met de gemiddelde wachttijden verkregen met off-line besturing. Echter, als er te ver wordt vooruitgekeken met on-line besturing, kan de toewijzing van ladingen aan voertuigen (en vice versa) ongunstig worden, hetgeen de gemiddelde wachttijden zelfs weer kan verhogen. Uit experimenten blijkt dat de verslechtering in gemiddelde wachttijden kunnen worden gereduceerd als opdrachten worden uitgesteld indien de voertuigen langer op de ladingen moeten wachten dan de gemiddelde transporttijd naar andere ophaallocaties.

Vervolgens zijn voor on-line bestuurde voertuigen strategieën onderzocht waarbij voertuigen naar een positie worden gestuurd om te parkeren wanneer er geen opdrachten aanwezig zijn. Bij omgevingen waar de opdrachten op willekeurige momenten op willekeurig plaatsen plaatsvinden wordt de parkeerpositie zodanig gekozen dat de verwachte rijtijd naar de locatie van de eerstvolgende opdracht minimaal is. Bij omgevingen waarbij de opdrachten gestructureerd in de tijd op bepaalde locaties worden vrijgegeven, worden de voertuigen naar de parkeerposities gestuurd in de buurt van de locaties waar de eerstvolgende opdracht het meest waarschijnlijk is. Bovendien kunnen in dit geval niet alleen voertuigen die op een parkeerplaats staan aan een opdracht worden toegewezen, maar ook de dichtstbijzijnde voertuigen die naar een parkeerplaats toe rijden. Resultaten wijzen uit dat met deze parkeerstrategieën voor voertuigen de prestatie van het systeem aanzienlijk kan worden verbeterd ten opzichte van de strategie waarbij voertuigen parkeren op de locatie waar ze hun laatste opdracht hebben afgezet. Bij experimenten waarbij het dichtstbijzijnde (bewegende) voertuig ongeacht de status aan een lading kan worden gekoppeld, daalt de gemiddelde wachttijden van de ladingen in beperkte mate. In enkele gevallen nemen de wachttijden zelfs toe.

Ook zijn de effecten op de wachttijden van ladingen bestudeerd als voertuigen bestuurd worden in omgevingen waar de tijden van vrijgave van ladingen afwijken van de geplande tijden van vrijgave. In de off-line situaties, worden voertuigen via het originele schema op pad gestuurd, welke eigenlijk verouderde informatie bevat. In dat geval, zullen de voertuigen (relatief) vroeger of later bij de ladingen arriveren, hetgeen de gemiddelde wachttijden voor de ladingen ongunstig beïnvloedt. Vanuit het perspectief van on-line besturing verandert er eigenlijk niets. Met on-line besturing kunnen voertuigen van de mogelijkheid gebruik maken om zich aan andere ladingen te koppelen zonder extra wachttijd voor de ladingen (en voertuigen) op te lopen. Mede hierdoor daalt het verschil in wachttijden van ladingen tussen on-line en off-line besturing. In omgevingen met zwaarbezette voertuigen is het zelfs mogelijk dat de wachttijden voor de ladingen met on-line besturing gunstiger zijn dan met off-line besturing.

In praktijksituaties wordt steeds vaker gebruik gemaakt van centrale besturingssystemen om voertuigen aan ladingen te koppelen (of vice versa). Deze systemen bevatten vaak veel maatwerk om te voldoen aan de specifieke wensen en eisen van de klant en de omgeving. Om het effect van verschillende factoren op verschillende besturingsstrategieën in echte transportomgevingen te bestuderen, zijn drie werkelijke voertuiggebaseerde interne transportomgevingen van drie verschillende praktijksituaties in detail gemodelleerd in Hoofdstuk 5. De drie bestudeerde omgevingen zijn:

1. Een Europees distributiecentrum met ongeveer 600 pallet bewegingen per dag, 5 geleide voertuigen en een voertuig-operationeel gebied van 40 bij 140 meter.
2. Een productie omgeving met ongeveer 1600 palletbewegingen per dag, 11 geleide voertuigen en een voertuig-operationeel gebied van 315 bij 540 meter.
3. Een containeroverslagbedrijf met ongeveer 6000 containerbewegingen per schip, 25 automatische geleide hefvoertuigen en een voertuig-operationeel gebied van 120 bij 540 meter.

Voor deze omgevingen zijn verschillende besturingsregels bestudeerd onder verschillende omstandigheden, zoals: voertuigen toewijzen aan bepaalde ladingen of stukken van het terrein, ladingen in groepen vrijgeven, gebruik van multi-capaciteit voertuigen, gebruik van voorinformatie, etc. Het is voor de hand liggend dat verschillende interne transport-omgevingen ook verschillende voertuigbesturingsregels moeten gebruiken om de ladingen te transporteren met zo weinig mogelijk voertuigen en een minimale gemiddelde wachttijd van ladingen. De voertuigen worden voor de drie bestudeerde praktijkomgevingen on-line bestuurd aangezien er aangenomen wordt dat operationele gegevens alleen tijdens de uitvoering (real time) beschikbaar komen. De voertuigen zijn echter ook bestuurd met de aanname dat gedeeltelijke voorinformatie over de tijd van vrijgave, oorsprong en bestemming van ladingen beschikbaar is. Er wordt aangetoond dat er een zekere rangorde bestaat voor de voertuigbesturingsregels en dat deze rangorde vrijwel ongevoelig is voor de verschillende omstandigheden.

Voor de drie praktijkomgevingen zijn voertuig-initiatief, lading-initiatief, tijd-gebaseerde en afstand-gebaseerde besturingsregels bestudeerd.

De afstand-gebaseerde lading-initiatief besturingsregel welke dichtstbijzijnde-voertuig-eerst (NVF) wordt genoemd, heeft als voordeel dat zelfs de meest afgelegen locaties met ladingen worden bediend omdat ladingen zich koppelen aan het dichtstbijzijnde voertuig. Dit leidt er toe dat de rijtijden van lege voertuigen op real time niveau zo klein mogelijk worden gehouden, met als gevolg dat de wachttijden van ladingen zo klein mogelijk worden gehouden.

Met de afstand-gebaseerde voertuig-initiatief regel welke dichtstbijzijnde-lading-eerst (NWF) wordt genoemd, kan het voorkomen dat een lading in een afgelegen gebied niet wordt bediend, aangezien het mogelijk is dat voertuigen zich telkens aan andere dichterbij gelegen ladingen koppelen.

Met de tijd-gebaseerde voertuig-initiatief regel welke oudste-lading-eerst (FCFS) wordt genoemd, wordt een voertuig dat op zoek is naar werk, gekoppeld aan de langst wachtende lading in het systeem. Het voordeel is dat elke lading uiteindelijk wel aan de beurt is om

vervoerd te worden, maar het nadeel is dat er geen aandacht gegeven wordt aan het minimaliseren de onderlinge lege transportafstanden (en daarmee de transporttijden en wachttijden).
Het in de praktijk gebruiken van realistische voorinformatie over ladingen kan de wachttijden van ladingen aanzienlijk reduceren. Een voorinformatieperiode in de bestudeerde productieomgeving van gemiddeld 0,18 ladingen per voertuig heeft als resultaat dat de gemiddelde wachttijden van ladingen ongeveer 65% reduceren. In het geval van het Europees distributiecentrum, heeft een voorinformatie periode van gemiddeld 0,06 ladingen per voertuig als resultaat dat de wachttijden van ladingen met 5% of meer dalen.
De reducties in de wachttijden met behulp van voorinformatie nemen af naarmate de bezettingsgraad van de voertuigen toeneemt. Zwaar bezette voertuigen hebben immers relatief meer te doen en houden minder tijd over om van de voorinformatie gebruik te maken. Opvallend is dat de afname in wachttijden van ladingen wanneer voorinformatie beschikbaar is, gunstiger is met de oudste-lading-eerst (FCFS) regel dan met de dichtstbijzijnde-voertuig-eerst (NVF) regel.
Echter, in alle (praktijk) gevallen is het aantal benodigde voertuigen en de gemiddelde wachttijd van ladingen verkregen met de afstand-gebaseerde lading-initiatief dichtstbijzijnde-voertuig-eerst (NVF) regel het meest gunstig, ook bij het gebruik van multi-capaciteit voertuigen en voorinformatie over ladingen.

Curriculum Vitae

Robert van der Meer was born on December 20th 1970 in Harderwijk, the Netherlands. In 1984, he moved to Mclean Virginia in the USA and finished the Mclean High School in three years. Back in the Netherlands at the age of 16, he continued his education at the International School in Hilversum, which he finished with a Bilingual International Baccalaureate. In 1990 he started his academic life and studied Business-econometrics at Erasmus University Rotterdam. In the first half of 1996 he received his Master's degree after writing a Master's thesis on Multi-component maintenance optimization models. At the same university he took the opportunity to continue his academic life with a Ph.D. project within the group of Logistic Management. His research on the control of vehicle-based internal transport systems has resulted in a number of publications in journals and books and an International Transport & Logistics best-paper award. Next to his Ph.D. activities, he was a member of several committees and the Ph.D.-council of the research school TRAIL. In his spare time he enjoys fitness and playing guitar.