The 20 Tera Flop Erasmus Computing Grid (ECG)

Tobias A. Knoch [1, 2] and Luc de Zeeuw [3]

with

Rob de Graaf [1, 3], Remko Kleinjan [1, 3], Redmar van Leeuwarden [1, 3], Sander van Vliet [1, 3], Bert Eussen [1, 4], Michael J. Moorhouse [1, 5]

J. J. van der Moolen [3], A. de Klein [4], A. Osterhaus [5], F. G. Grosveld [6]

[1] The GLOBE-Consortium of B. Eussen, M. Moorhouse and Tobias A. Knoch,
[2] Biophysical Genomics in the [5] Department of Cell Biology, [4] Department of Clinical Genetics, [5] Department of Bioinformatics
Erasmus Medical Center, Dr. Molewaterplein 50, NL - 3015 GE Rotterdam, The Netherlands

[3] RIVIO, Hogeschool Rotterdam, Museumpark 40, NL - 3015 CX Rotterdam, The Netherlands

http://www.erasmusmc.nl or T. A.Knoch@erasmusmc.nl

Introduction

Hogeschool Rotterdam
RIVIO Campus

Erasmus Medical Center
Faculty Tower

To meet the enormous computational needs of live-science scientific research as well as clinical diagnostics and treatment at the

Hogeschool Rotterdam and the Erasmus Medical Center

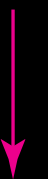
both are currently setting up in an interdisciplinary approach one of the largest desktop computing grids in the world – The Erasmus Computing Grid. The Hogeschool Rotterdam has a capacity of ~4000 desktop computers and the Erasmus Medical Center has a capacity of ~9000 desktop computers, i. e. ~ 20 Tera flops. These computers are located in different networks and even building throughout the city of Rotterdam with a ~8000 desktop computers located within 500m of the Museums Park campus area and connected by optical fibers. The operating systems range from Windows 2000 and Windows XP to Linux and Mac OS driven systems, and are especially diverse at the faculty building of the Erasmus Medical Center. Different usage scenarios and management strategies apply as well in both institutions. Nevertheless, the computers are available for computations 99 % of the time!

Results

The Erasmus Computing Grid is based on the Condor middleware. An own developed management system allows easy installation and management of all the computers in the system. The system was set-up in a way that security and privacy issues, which are especially important for the Erasmus Medical Center are always guarantied. Currently, the Erasmus Computing Grid has reached a capacity of 3 Tera flops. 4 user groups are either in production or testing of their algorithms. Around 15 different user groups are heading towards using the grid in the near future. Expansions are planned.

Goal

current
3 Tera Flops
2006



20 Tera Flops

2007

The 20 Tera flop Erasmus Computing Grid (ECG)

Knoch, T. A. & de Zeeuw, L.

D-Grid All-Hands-Meeting. Universitäts Klinikum, Göttingen, Germany,
23rd - 25th March, 2009.

Abstract

The Set-Up of the 20 Teraflop Erasmus Computing Grid: To meet the enormous computational needs of live-science research as well as clinical diagnostics and treatment the Hogeschool Rotterdam and the Erasmus Medical Center are currently setting up one of the largest desktop computing grids in the world – The Erasmus Computing Grid. Currently 3 Tera flops are operational and in early production, installation up to the today available maximum capacity of 20 Tera flops in both institutions is planned and partly underway. Thus the Erasmus Computing Grid transforms the existing and sustained huge computer capacity available into usable form via a reliable and secure installing and management system, so that the academic and industrial opportunities depending on such huge computing capacities can be realized for the benefit of society.

Corresponding author email contact: TA.Knoch@taknoch.org

Keywords:

Genome, genomics, genome organization, genome architecture, structural sequencing, architectural sequencing, systems genomics, coevolution, holistic genetics, genome mechanics, genome function, genetics, gene regulation, replication, transcription, repair, homologous recombination, simultaneous co-transfection, cell division, mitosis, metaphase, interphase, cell nucleus, nuclear structure, nuclear organization, chromatin density distribution, nuclear morphology, chromosome territories, subchromosomal domains, chromatin loop aggregates, chromatin rosettes, chromatin loops, chromatin fibre, chromatin density, persistence length, spatial distance measurement, histones, H1.0, H2A, H2B, H3, H4, mH2A1.2, DNA sequence, complete sequenced genomes, molecular transport, obstructed diffusion, anomalous diffusion, percolation, long-range correlations, fractal analysis, scaling analysis, exact yard-stick dimension, box-counting dimension, lacunarity dimension, local nuclear dimension, nuclear diffuseness, parallel super computing, grid computing, volunteer computing, Brownian Dynamics, Monte Carlo, fluorescence in situ hybridization, confocal laser scanning microscopy, fluorescence correlation spectroscopy, super resolution microscopy, spatial precision distance microscopy, auto-fluorescent proteins, CFP, GFP, YFP, DsRed, fusion protein, in vivo labelling, information browser, visual data base access, holistic viewing system, integrative data management, extreme visualization, three-dimensional virtual environment, virtual paper tool.

Literature References

Knoch, T. A. Dreidimensionale Organisation von Chromosomen-Domänen in Simulation und Experiment. (Three-dimensional organization of chromosome domains in simulation and experiment.) *Diploma Thesis*, Faculty for Physics and Astronomy, Ruperto-Carola University, Heidelberg, Germany, 1998, and TAK Press, Tobias A. Knoch, Mannheim, Germany, ISBN 3-00-010685-5 and ISBN 978-3-00-010685-9 (soft cover, 2rd ed.), ISBN 3-00-035857-9 and ISBN 978-3-00-0358857-0 (hard cover, 2rd ed.), ISBN 3-00-035858-7, and ISBN 978-3-00-035858-6 (DVD, 2rd ed.), 1998.

Knoch, T. A., Münkel, C. & Langowski, J. Three-dimensional organization of chromosome territories and the human cell nucleus - about the structure of a self replicating nano fabrication site. *Foresight Institute - Article Archive*, Foresight Institute, Palo Alto, *CA, USA,* http://www. foresight.org, 1- 6, 1998.

Knoch, T. A., Münkel, C. & Langowski, J. Three-Dimensional Organization of Chromosome Territories and the Human Interphase Nucleus. *High Performance Scientific Supercomputing*, editor Wilfried Juling, Scientific Supercomputing Center (SSC) Karlsruhe, University of Karlsruhe (TH), 27- 29, 1999.

Knoch, T. A., Münkel, C. & Langowski, J. Three-dimensional organization of chromosome territories in the human interphase nucleus. *High Performance Computing in Science and Engineering 1999*, editors Krause, E. & Jäger, W., High-Performance Computing Center (HLRS) Stuttgart, University of Stuttgart, Springer Berlin-Heidelberg-New York, ISBN 3-540-66504-8, 229-238, 2000.

Bestvater, F., **Knoch, T. A.**, Langowski, J. & Spiess, E. GFP-Walking: Artificial construct conversions caused by simultaneous cotransfection. *BioTechniques* 32(4), 844-854, 2002.

Knoch, T. A. (editor), Backes, M., Baumgärtner, V., Eysel, G., Fehrenbach, H., Göker, M., Hampl, J., Hampl, U., Hartmann, D., Hitzelberger, H., Nambena, J., Rehberg, U., Schmidt, S., Weber, A., & Weidemann, T. Humanökologische Perspectiven Wechsel - Festschrift zu Ehren des 70. Geburtstags von Prof. Dr. Kurt Egger. Human Ecology Working Group, Ruperto-Carola University of Heidelberg, Heidelberg, Germany, 2002.

Knoch, T. A. Approaching the three-dimensional organization of the human genome: structural-, scaling- and dynamic properties in the simulation of interphase chromosomes and cell nuclei, long- range correlations in complete genomes, *in vivo* quantification of the chromatin distribution, construct conversions in simultaneous co-transfections. *Dissertation*, Ruperto-Carola University, Heidelberg, Germany, and TAK†Press, Tobias A. Knoch, Mannheim, Germany, ISBN 3-00-009959-X and ISBN 978-3-00-009959-5 (soft cover, 3rd ed.), ISBN 3-00-009960-3 and ISBN 978-3-00-009960-1 (hard cover, 3rd ed.), ISBN 3-00-035856-9 and ISBN 978-3-00-010685-9 (DVD, 3rd ed.) 2002.

Knoch, T. A. Towards a holistic understanding of the human genome by determination and integration of its sequential and three-dimensional organization. *High Performance Computing in Science and Engineering 2003*, editors Krause, E., Jäger, W. & Resch, M., High-Performance Computing Center (HLRS) Stuttgart, University of Stuttgart, Springer Berlin-Heidelberg-New York, ISBN 3- 540-40850-9, 421-440, 2003.

Wachsmuth, M., Weidemann, T., Müller, G., Urs W. Hoffmann-Rohrer, **Knoch, T. A.**, Waldeck, W. & Langowski, J. Analyzing intracellular binding and diffusion with continuous fluorescence photobleaching. *Biophys. J.* 84(5), 3353-3363, 2003.

Weidemann, T., Wachsmuth, M., **Knoch, T. A.**, Müller, G., Waldeck, W. & Langowski, J. Counting nucleosomes in living cells with a combination of fluorescence correlation spectroscopy and confocal imaging. *J. Mol. Biol.* 334(2), 229-240, 2003.

Fejes Tóth, K., **Knoch**, **T. A.**, Wachsmuth, M., Frank-Stöhr, M., Stöhr, M., Bacher, C. P., Müller, G. & Rippe, K. Trichostatin A induced histone acetylation causes decondensation of interphase chromatin. *J. Cell Science* 177, 4277-4287, 2004.

Ermler, S., Krunic, D., **Knoch, T. A.**, Moshir, S., Mai, S., Greulich-Bode, K. M. & Boukamp, P. Cell cycle-dependent 3D distribution of telomeres and telomere repeat-binding factor 2 (TRF2) in HaCaT and HaCaT-myc cells. *Europ. J. Cell Biol.* 83(11-12), 681-690, 2004.

Kost, C., Gama de Oliveira, E., **Knoch, T. A.** & Wirth, R. Spatio-temporal permanence and plasticity of foraging trails in young and mature leaf-cutting ant colonies (*Atta spp.*). *J. Trop. Ecol.* 21(6), 677- 688, 2005.

Winnefeld, M., Grewenig, A., Schnölzer, M., Spring, H., **Knoch, T. A.**, Gan, E. C., Rommelaere, J. & Cziepluch, C. Human SGT interacts with BAG-6/Bat-3/Scythe and cells with reduced levels of either

protein display persistence of few misaligned chromosomes and mitotic arrest. *Exp. Cell Res.* 312, 2500-2514, 2006.

Sax, U., Weisbecker, A., Falkner, J., Viezens, F., Yassene, M., Hartung, M., Bart, J., Krefting, D., **Knoch, T. A.** & Semler, S. Grid-basierte Services für die elektronische Patientenakte der Zukunft. *E- HEALTH-COM - Magazin für Gesundheitstelematik und Telemedizin*, 4(2), 61-63, 2007.

de Zeeuw, L. V., **Knoch, T. A.,** van den Berg, J. & Grosveld, F. G. Erasmus Computing Grid - Het bouwen van een 20 TeraFLOP virtuelle supercomputer. *NIOC proceedings 2007 - het perspective of lange termijn*. editor Frederik, H. NIOC, Amsterdam, The Netherlands, 52-59, 2007.

Rauch, J., **Knoch, T. A.** , Solovei, I., Teller, K. Stein, S., Buiting, K., Horsthemke, B., Langowski, J., Cremer, T., Hausmann, M. & Cremer, C. Lightoptical precision measurements of the Prader- Willi/Angelman Syndrome imprinting locus in human cell nuclei indicate maximum condensation changes in the few hundred nanometer range. *Differentiation* 76(1), 66-82, 2008.

Sax, U., Weisbecker, A., Falkner, J., Viezens, F., Mohammed, Y., Hartung, M., Bart, J., Krefting, D., **Knoch, T. A.** & Semler, S. C. Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die EPA der Zukunft. *Telemedizinführer Deutschland 2008*, editor Jäckel, A. Deutsches Medizinforum, Minerva KG, Darmstadt, ISBN 3-937948-06-6, ISBN-13 9783937948065, 47-51, 2008.

Drägestein, K. A., van Capellen, W. A., van Haren, J. Tsibidis, G. D., Akhmanova, A., **Knoch, T. A.**, Grosveld, F. G. & Galjart, N. Dynamic behavior of GFP-CLIP-170 reveals fast protein turnover on microtubule plus ends. *J. Cell Biol. 180(4)*, 729-737, 2008.

Jhunjhunwala, S., van Zelm, M. C., Peak, M. M., Cutchin, S., Riblet, R., van Dongen, J. J. M., Grosveld, F. G., **Knoch, T. A.**$^{+}$ & Murre, C.$^{+}$ The 3D-structure of the Immunoglobulin Heavy Chain Locus: implications for long-range genomic interactions. *Cell 133(2)*, 265-279, 2008.

Krefting, D., Bart, J., Beronov, K., Dzhimova, O., Falkner, J., Hartung, M., Hoheisel, A., **Knoch, T. A.**, Lingner, T., Mohammed, Y., Peter, K., Rahm, E., Sax, U., Sommerfeld, D., Steinke, T., Tolxdorff, T., Vossberg, M., Viezens, F. & Weisbecker, A. MediGRID - Towards a user friendly secured grid infrastructure. *Future Generation Computer Systems 25(3)*, 326-336, 2008.

Knoch, T. A., Lesnussa, M., Kepper, F. N., Eussen, H. B., & Grosveld, F. G. The GLOBE 3D Genome Platform - Towards a novel system-biological paper tool to integrate the huge complexity of genome organization and function. *Stud. Health. Technol. Inform. 147*, 105-116, 2009.

Knoch, T. A., Baumgärtner, V., de Zeeuw, L. V., Grosveld, F. G., & Egger, K. e-Human Grid Ecology: Understanding and approaching the Inverse Tragedy of the Commons in the e-Grid Society. *Stud. Health. Technol. Inform. 147*, 269-276, 2009.

Dickmann, F., Kaspar, M., Löhnardt, B., **Knoch, T. A.,** & Sax, U. Perspectives of MediGRID. *Stud. Health. Technol. Inform. 147*, 173-182, 2009.

Knoch, T. A., Göcker, M., Lohner, R., Abuseiris, A. & Grosveld, F. G. Fine-structured multi-scaling long-range correlations in completely sequenced genomes - features, origin and classification. *Eur. Biophys. J. 38(6)*, 757-779, 2009.

Dickmann, F., Kaspar, M., Löhnhardt, B., Kepper, N., Viezens, F., Hertel, F., Lesnussa, M., Mohammed, Y., Thiel, A., Steinke, T., Bernarding, J., Krefting, D., **Knoch, T. A.** & Sax, U. Visualization in health-grid environments: a novel service and business approach. *LNCS 5745*, 150-159, 2009.

Dickmann, F., Kaspar, M., Löhnhardt, B., Kepper, N., Viezens, F., Hertel, F., Lesnussa, M., Mohammed, Y., Thiel, A., Steinke, T., Bernarding, J., Krefting, D., **Knoch, T. A.** & Sax, U. Visualization in health-grid environments: a novel service and business approach. *Grid economics and business models - GECON 2009 Proceedings, 6th international workshop, Delft, The Netherlands.* editors Altmann, J., Buyya, R. & Rana, O. F., GECON 2009, LNCS 5745, Springer-Verlag Berlin Heidelberg, ISBN 978-3-642-03863-1, 150-159, 2009.

Estrada, K.*, Abuseiris, A.*, Grosveld, F. G., Uitterlinden, A. G., **Knoch, T. A.**$^{+}$ & Rivadeneira, F.$^{+}$ GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics 25(20), 2750-2752*, 2009.