


PC GRID use cases and requirements

Hiroyuki Takashima, Ph.D
OGF22, Enterprise Grids Requirements RG
Cambridge, 25 February 2008



Agenda

Enterprise Grids Requirements RG

- Overview of Novartis computing
- Scenarios
- Approaches
- Case studies and gap analysis
- Key technical requirements

Overview of Novartis GRID computing Enterprise Grids Requirements RG

Novartis Institutes for Biomedical Research (NIBR) site map

The map shows the following sites: Emeryville (USA), Cambridge (MA, USA - Headquarters), East Hanover (NJ, USA), Basel (Switzerland), Horsham (UK), Vienna (Austria), and Tsukuba (Japan).

Basel, Switzerland | Cambridge, MA (Headquarters) | East Hanover, NJ | Emeryville, CA | Horsham, UK | Tsukuba, Japan | Vienna, Austria

Overview of Novartis GRID computing Enterprise Grids Requirements RG

Users and applications: Group A, Group B, Group C, Group D, Group E

Job submission protocol

HPC technologies integration/sharing layer

Cluster existing HPC infrastructure: SUN HPC server, TimeLogic HPC server, SGI HPC server, SGI HPC server, SGI HPC server, COMPAQ HPC Cluster

Large shared Linux Clusters

Shared PC grid

Shared Multi-processor servers

External Collaborations

4 | PC GRID use cases and requirements | H. Takahashi | 2

Overview of Novartis GRID computing *Enterprise Grids Requirements RG*

The 22nd Open Grid Forum - OGF22
February 25-28, 2008
Cambridge, MA USA

Wednesday, February 27

2:00 pm - 3:30 pm

Pharma, Biotech and Life Sciences Workshop (Ballroom A/B/C) (90 mins)

Steve Litster Ph.D, Novartis Institutes for BioMedical Research

Main topic: Storage virtualization

5 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22

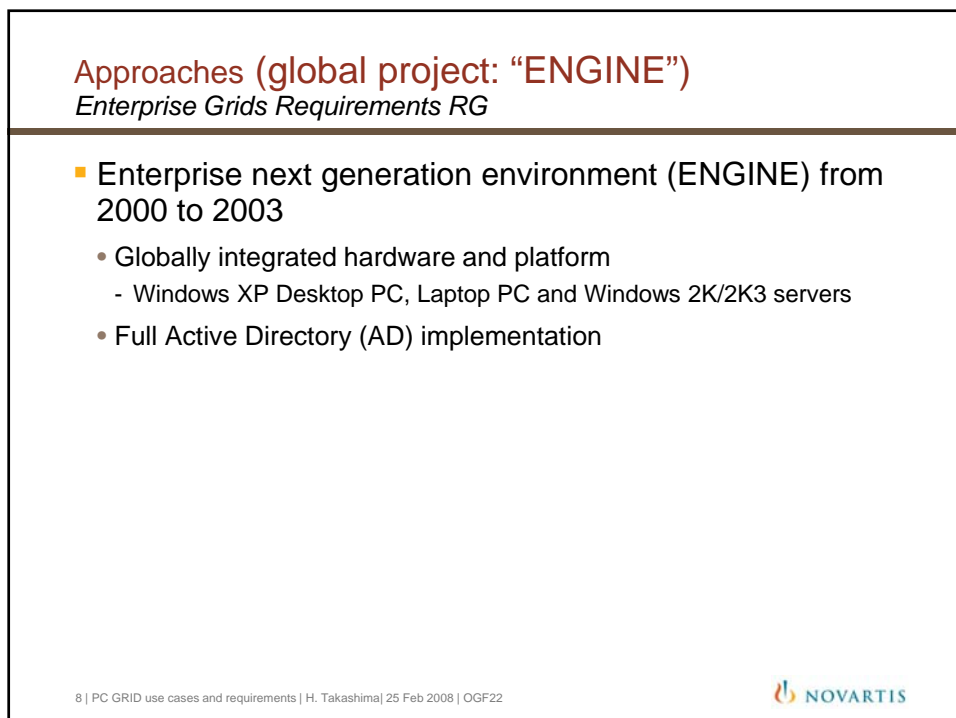
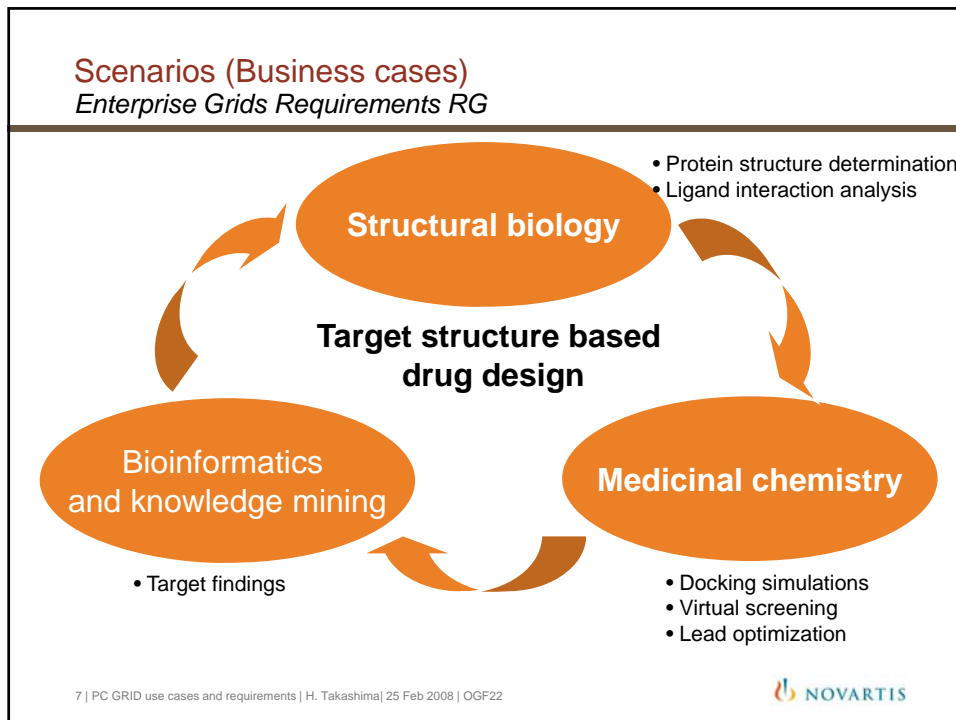


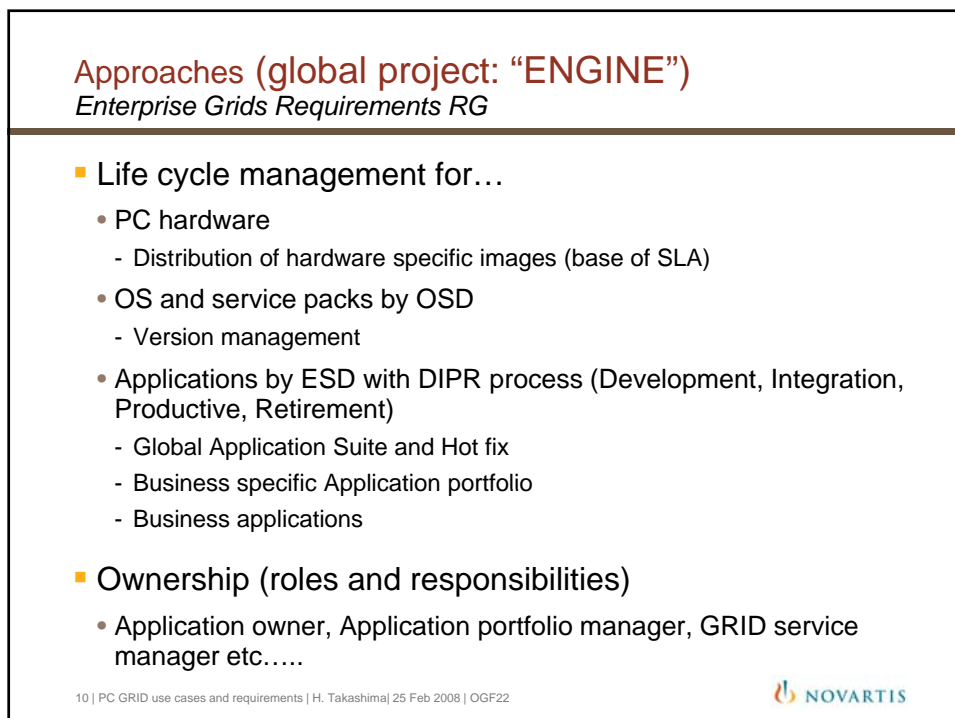
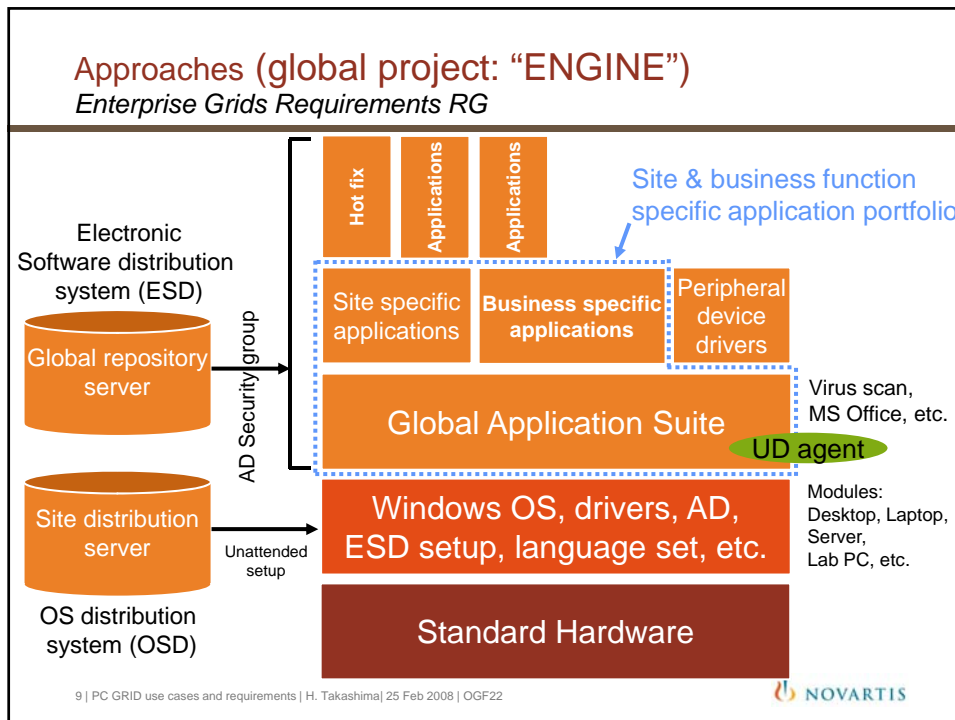
Scenarios (Business cases) *Enterprise Grids Requirements RG*

- Medicinal chemistry
- Structural biology
- Bioinformatics and knowledge mining

6 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22







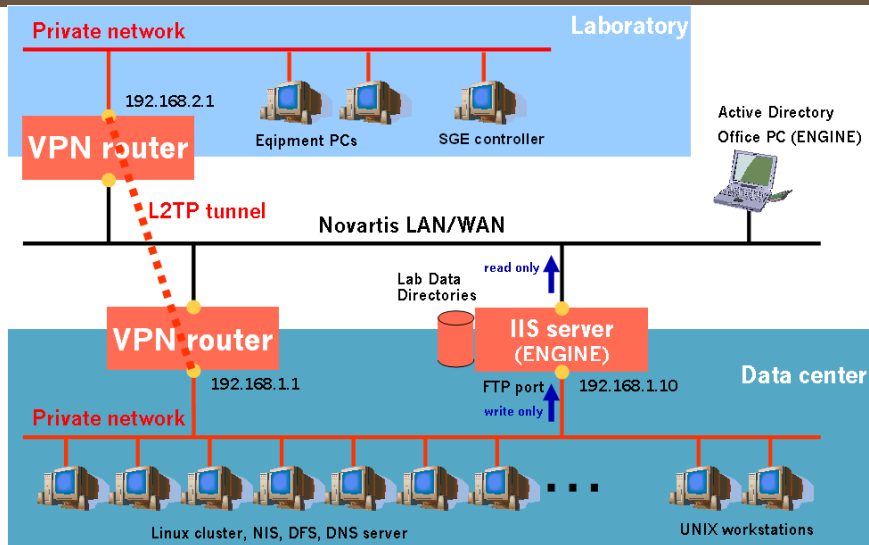
Approaches (Local build in environments) Enterprise Grids Requirements RG

- Linux clusters
 - Site oriented (Basel, Cambridge etc.)
 - Test cluster in Tsukuba
 - 20 Linux PC controlled by Sun Grid Engine (SGE)
 - Logically separated private segments by VPN router

11 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22



Approaches (Local build in environments of Tsukuba) Enterprise Grids Requirements RG



12 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22



Case studies and gap analysis

Enterprise Grids Requirements RG

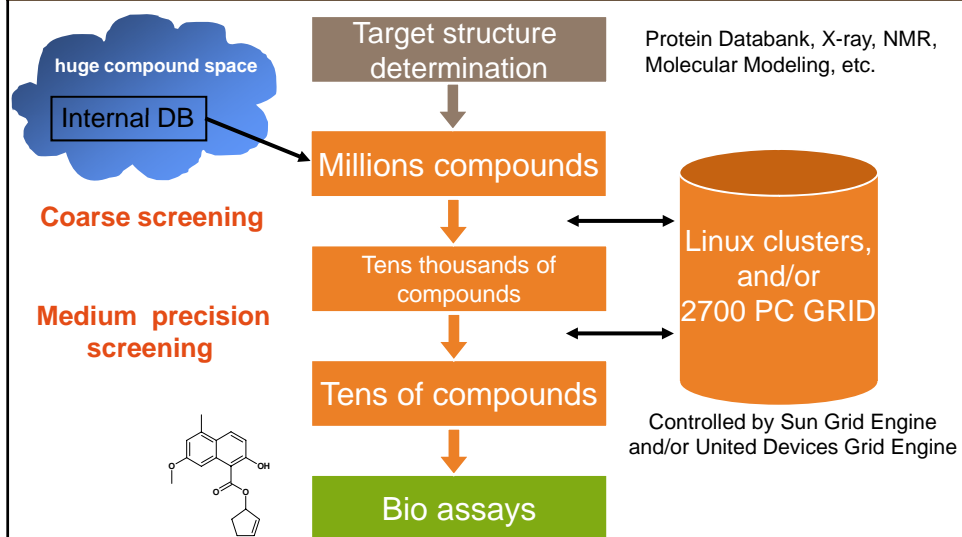
- High-Throughput Docking for *in silico* screening
- Protein structure determination
- Vector/Parallel/Distributed jobs bench marks

13 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22



Case Study 1: High-Throughput Docking for *in silico* screening

Enterprise Grids Requirements RG



14 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22



Case Study 1: High-Throughput Docking for *in silico* screening Enterprise Grids Requirements RG

Discovery of a Potent and Selective Protein Kinase CK2 Inhibitor by High-Throughput Docking
Vangrevelinghe, E., Zimmermann, K., Schoepfer, J., Portmann, R., Fabbro, D., and Furet, P.
(2003) *J. Med. Chem.* **46**, 2656-2662.

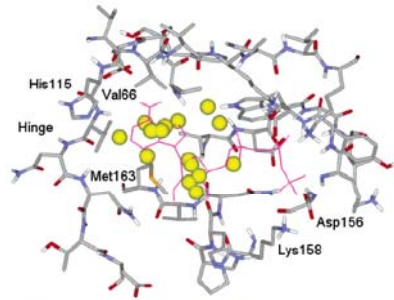


Figure 2. Overlay of the docking site points used (yellow spheres) and the ATP analogue extracted from the crystallographic template structure 1DAW (magenta) inside the active site of the homology model of human CK2a.

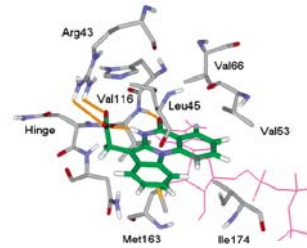


Figure 3. Relative binding modes of compound 4 (green) and the ATP analogue extracted from the crystallographic template structure 1DAW (magenta) in the human CK2a active site. The hydrogen bonds formed with Arg 43 and Val 116 are indicated as orange lines.

15 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22

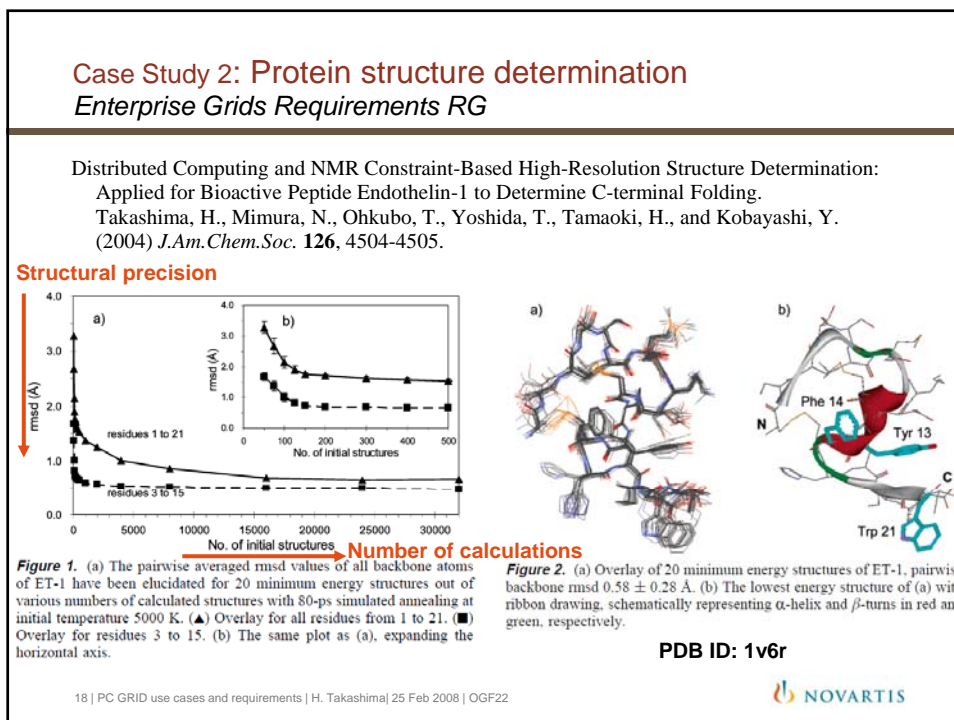
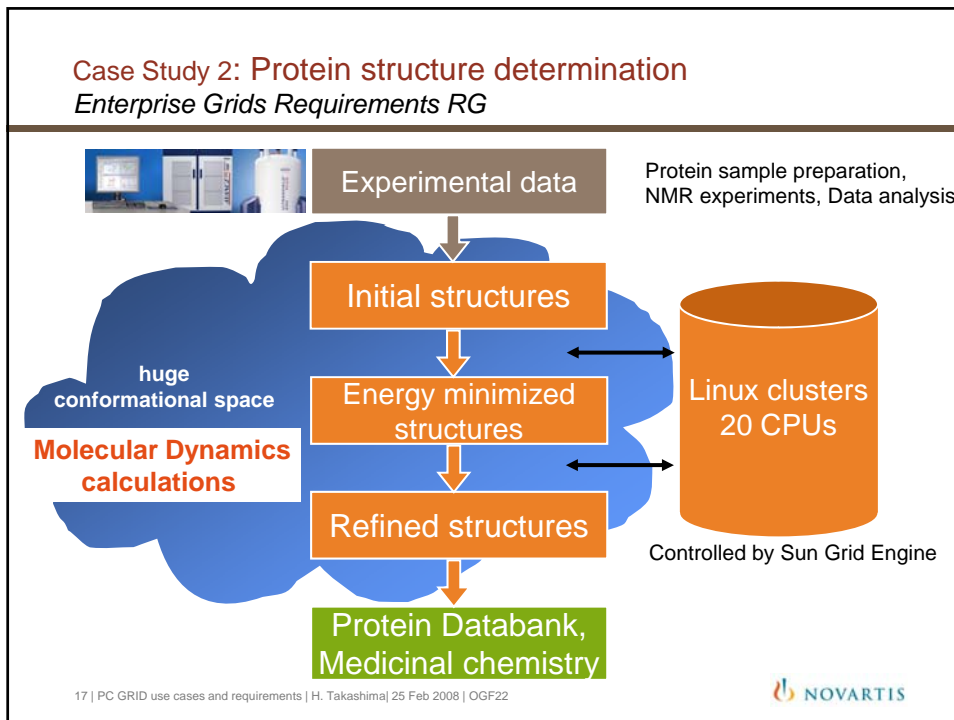


Case Study 2: Protein structure determination Enterprise Grids Requirements RG

- X-ray
- NMR
- Electron microscopy
- Computer modeling

16 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22





Case Study 2: Protein structure determination

Enterprise Grids Requirements RG

Solution NMR Structure Investigation for Releasing Mechanism of Neocarzinostatin Chromophore from the Holoprotein.

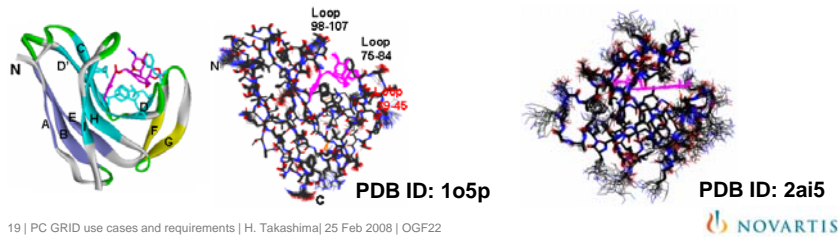
Takashima, H., Yoshida, T., Ishino, T., Hasuda, K., Ohkubo, T., and Kobayashi, Y. (2005) *J. Biol. Chem.* **280**, 11340-11346.

High-Resolution Protein Structure Determination by NMR

Takashima, H. (2006) in *Annual Reports on NMR Spectroscopy* **59**, 235-273. printed by Elsevier Ltd., London

Structure of Cytochrome c552 from a Moderate Thermophilic Bacterium, *Hydrogenophilus thermoluteolus*: Comparative study on the thermostability of cytochrome c.

Nakamura, S., Ichiki, S., Takashima, H., Uchiyama, S., Hasegawa, J., Kobayashi, Y., Sambongi, Y., and Ohkubo, T. (2006) *Biochemistry* **45**, 6115-6123.



Key technical requirements

Enterprise Grids Requirements RG

■ Limitation as is

- Difficult license management
 - Number of CPUs is limited by software licenses
- Scientific application platform
 - Unix/Linux base to Windows base software
- Network performance
 - Especially at remote sites
- Overhead of parallel calculations
- Calculation precisions
 - With or without induced fitting, Free energy calculations, Protein dynamics
Force fields, Solvent and more.....

Key technical requirements

Enterprise Grids Requirements RG

- Requirements to be
 - Precision as high as possible (depending on computer power)
 - Expand available applications
 - Real time simulations
 - **More computer power on demand** (from user's point of view)
 - Software license virtualization

21 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22



Enterprise Grids Requirements RG

Thank you for your collaboration!

22 | PC GRID use cases and requirements | H. Takashima | 25 Feb 2008 | OGF22

