PRACE
Partnership for
Advanced Computing in Europe

Lennart Johnsson

Royal Institute of Technology

Stockholm, Sweden

Department of Computer Science
University of Houston

PRACE Vision and Mission

- **Vision**: Enable and support European global leadership in public and private research and development.

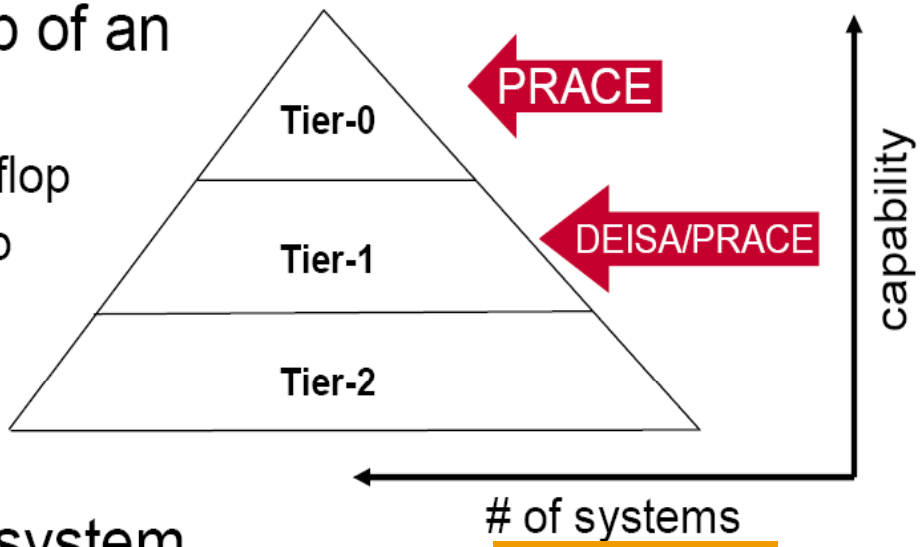
- **Mission**: Contribute to the advancement of European competitiveness in industry and research through the provisioning of world leading persistent High-End Computing infrastructure

PRACE - History

- **HPC for Europe Project** (HPCEUR) 2004 – June 2006
  - Focus on Scientific Case
- **HPC in Europe Task Force** (HET) Fall 2006
  - Identify current European HPC status in industry and academia
  - Develop an HPC Roadmap for Europe
  - Recommend processes for implementing the roadmap
- **European Strategy Forum for Research Infrastructures** (ESFRI, founded 2002) 2006 Roadmap
  - HPC Service for Computing and Data Treatment (1 of 35 pan-European Research Infrastructures)
- **PRACE Initiative formed 2007**
  - 14 European countries signed MoU April 16, 2007
  - 20 European countries by end of 2009
- **PRACE Research Infrastructure** (RI) legal entity (PRACE AISBL) formed April 23, 2010

The ESFRI Vision for a European HPC service

- European HPC-facilities at the top of an HPC provisioning pyramid
  - Tier-0: 3-6 European Centres for Petaflop
  - Tier-0: ? European Centres for Exaflop
  - Tier-1: National Centres
  - Tier-2: Regional/University Centres



- Creation of a European HPC ecosystem involving all stakeholders
  - HPC service providers on all tiers
  - Grid Infrastructures
  - Scientific and industrial user communities
  - The European HPC hard- and software industry

PRACE AISBL

- PRACE AISBL (Association International Sans But Lucratif) is a Belgian legal entity seated in Brussels formed April 23 2010 for providing a persistent pan-European Research Infrastructure for High-End Computing and associated services. Member countries currently are
  - Austria
  - Bulgaria
  - Cyprus
  - Czech Republic
  - Finland
  - France
  - Germany
  - Greece
  - Ireland
  - Italy
  - Netherlands
  - Norway
  - Poland
  - Portugal
  - Serbia
  - Spain
  - Sweden
  - Switzerland
  - Turkey
  - United Kingdom

Interest to join by Belgium, Hungary and Israel)

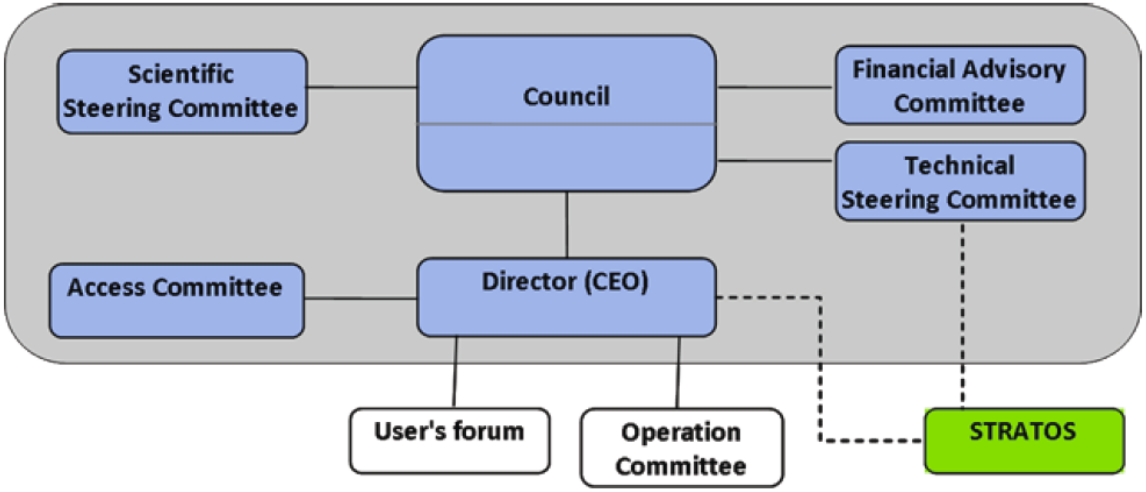
Commitments to PRACE AISBL

- Hosting Partners:
  (Germany, France, Italy, Spain)
  - Binding commitments to contribute 100 M€ over 5 years in terms of Tier-0 cycles
  - Contribution measured by TCO
- All partners:
  - Binding commitment to share PRACE AISBL Head-Quarters costs equally
- EU Commission (expected)
  - 70 M€ in FP7

Note: GDP spread among PRACE partners a factor of ~200

Governance of the Association

- Modeled after successful examples of existing RIs
- Council as main decision making body
- Director with strong managing mandate
- Scientific Steering Committee and Access Committee to give scientific advice and to steer the Peer Review process
- Further committees will be instantiated by the Council as needed



PRACE Board of Directors (interim)

- **France**
  - Jean-Philipe Nominé (Financial Work, Dissemination)
- **Germany**
  - Thomas Eickermann (HQ-local Organizational Tasks, Legal Work)
- **Italy**
  - Sergio Bernardi (Legal Work, Peer Review, Business Plan)
- **Spain**
  - Sergi Girona, Chair (Business Plan, Peer Review)
- **Non-hosting partners (16)**
  - Lennart Johnsson (Dissemination, Financial Work)

Scientific Steering Committee

Scope and basic rules defined in the Statutes of the AISBL

- The SSC is responsible for giving opinions on all matters of ascientific and technical nature
  - Maximum of 21 members
  - Members appointed by Council based on a list of candidates prepared by the SSC
- Two year term (renewable twice)
- Propose the members of the Access Committee
- Resolutions by simple majority

Scientific Steering Committee

(Approved by PRACE Council 10/5/2010)

- Richard Kenway (chair) (UK,particle physics)

- Baldasano(Spain,environment)
- Binder (Germany, statistical physics)
- Carloni (Italy,biological physics)
- Ciccotti (Italy, statistical physics)
- Frenkel (Netherlands, molecular simulations)
- Joussaume (France, environment)
- Moore (Switzerland, astrophysics)
- Muenster (Germany, particle physics)
- Nieminen (Finland, materials)
- Orozco (Spain, life sciences)
- Ottaviani (France, plasma physics)
- Parrinello (Switzerland, chemistry)
- Pironneau (France, mathematics)
- Poinsot (France, engineering)
- Portegies Zwart (Netherlands, astrophysics)
- Quarteroni (Italy, engineering)
- Ruud (Norway, chemistry)
- Schroeder (Germany, engineering)
- Silva (Portugal, plasma physics)
- Valencia (Spain, bioinformatics)

Access Committee

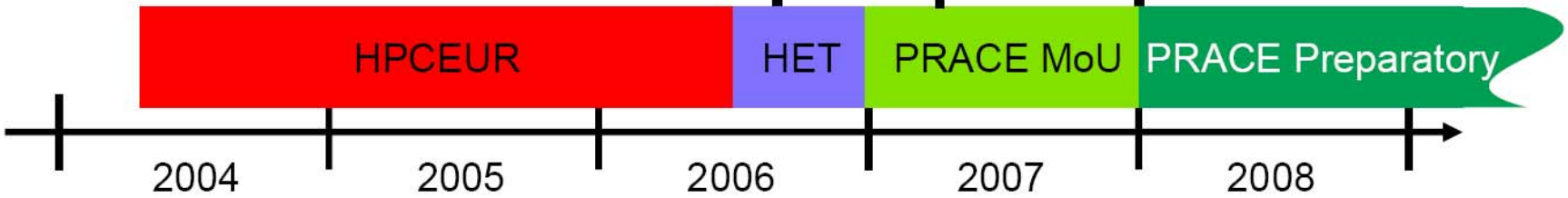
Responsible for giving opinions on the scientific use of Tier-0 Infrastructure, and providing recommendations on the allocation of PRACE RI resources based on the Peer Review process

- Proposed by the SSC based on their personal experience in the areas of science
  - Appointed by the Council
  - Minimum of 5 members
  - Two years term (renewable once)
  - Half of the members shall be replaced every year
  - The Access Committee shall define its internal working rules

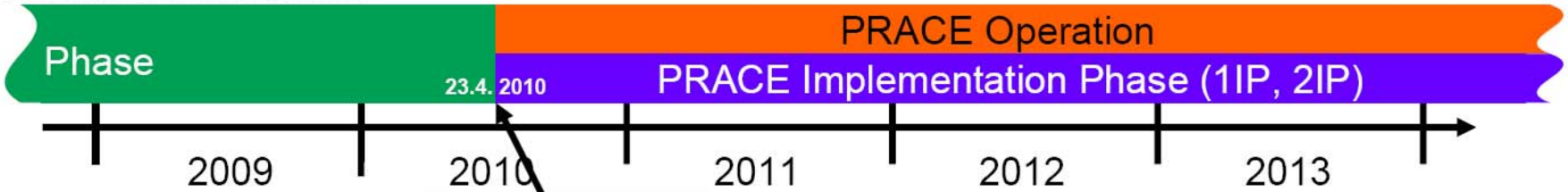
PRACE EU projects

- PRACE Preparatory Phase project
  - 1/1/2008 – 6/30/2010, 20M€, 50% from EU Commission
- PRACE 1st Implementation Phase project
  - 7/1/2010 – 6/30/2012, 20M€ from the EU Commission plus 8.5 M€ from partners
- PRACE 2nd Implementation Phase project (proposal in preparation)
  - 20M€ from the EU Commission plus ~20M€ from partners

PRACE History and first steps



EU-Grant: INFSO-RI-211528, 10 Mio. €



PRACE (AISBL), a legal entity seated in Brussels

PRACE RI Systems

- 1st PRACE System
  - BG/P by Gauss Center for Supercomputing at Juelich
    - 294912 CPU cores, 144 TB memory
    - 1 PFlop/s peak performance
    - 825.5 TFlop/s Linpack
    - 600 I/O nodes (10GigE) > 60 GB/s I/O
    - 2.2 MW power consumption
    - 35% for PRACE
- 2nd PRACE System (2010/2011) - GENCI
  - Intel based system by Bull
    - ~90,000 cores
    - 1.6 PFlop/s peak performance
- 3rd PRACE System 2011 HLRS
- 4th PRACE System 2012 LRZ
- 5th and 6th Systems 2012/2013 – Cineca and BSC

PRACE RI Access

- Access strictly by PRACE peer review
- Free-of-charge for European scientific communities
- Three types of access
  - Preparatory access (Technical Review only)
    - Scalability
    - Code development
    - Code development with PRACE support
  - Project access – for a specific project, grant period ~1 year (Technical and Scientific Review)
  - Program access – resources managed by a community (Technical and Scientific Review)
- Early access call opened May 10 and closed June 10, 2010
- Start of provision: 1.8.2010
- 1st regular call opened June 15, closed August 15, 2010
- 2nd regular will open November 1, 2010, close January 11, 2011
- Further calls every 6 months
- 1st Preparatory Access call will open November 1, 2010

PRACE RI Review procedures for Project and Program proposals

- Technical peer review (system and code suitability) by hosting centre representatives
- Scientific peer review by 3 external reviewers
- Applicants have the right to comment on the reviewers remarks; these remarks are sent together with the reviewers comments to the Access Committee (Prioritization Panel in lieu of the AC)
- Formally the Council ratifies the prioritization list
- The Director informs (in writing) the applicants of the computing grants, on behalf of the Council
- Applicants have the right to appeal the decision of the Council

PRACE RI Access

Early Access Call

- Belgium 1
- Bulgaria 2
- Denmark 1
- Finland 2
- France 6
- Germany 10
- Germany/UK 1
- Greece 2
- Ireland 2
- Italy 4
- Netherlands 2
- Portugal 4
- Spain 13
- Sweden 1
- Switzerland 3
- UK 12

68 proposals from 15 countries
Scientific Review of 39 proposals
10 proposals granted access
Request 5x available resources

- Astrophysics - 7
- Chemistry and Materials - 11
- Earth Sciences and Environment - 4
- Engineering and Energy - 15
- Fundamental Physics - 17
- Mathematics and Computing - 3
- Medicine and Life Sciences - 8

PRACE RI Access

1st Regular Access Call

- Belgium 1
- Bulgaria 1
- Cyprus 2
- Denmark 2
- Finland 1
- France 6
- Germany 7
- Greece 2
- Hungary 1
- Ireland 3
- Italy 5
- Netherlands 1
- Poland 1
- Portugal 2
- Spain 10
- Switzerland 1
- UK 13

- Astrophysics - 4
- Chemistry and Materials - 8
- Earth Sciences and Environment - 1
- Engineering and Energy - 6
- Fundamental Physics - 10
- Mathematics and Computing - 1
- Medicine and Life Sciences - 3

59 proposals from 17 countries
Evaluation in progress

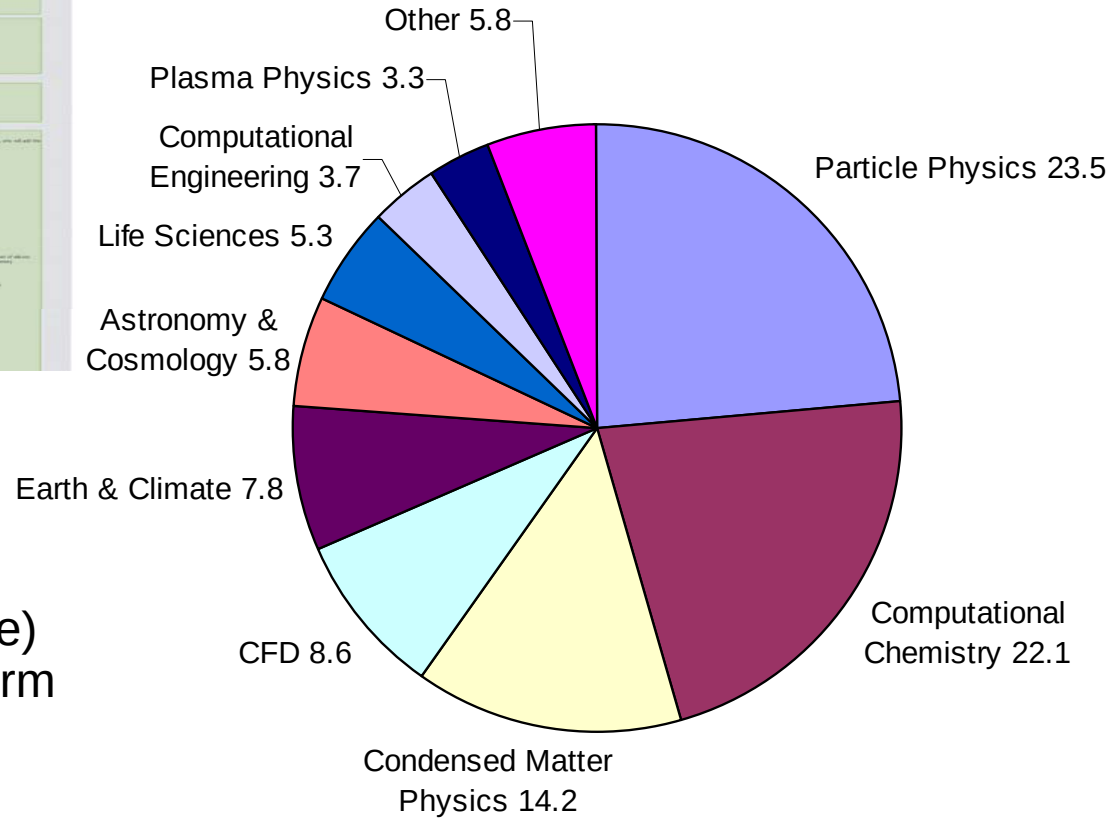
PRACE Preparatory Phase Project Categorization of Applications (2008/9)

- Survey of 24 PRACE partner systems (>10 TF) and 69 applications (100+ responses)

- Linpack Equivalent Flops (LEFs) defined as

(percentage of system utilization (time) relative to available time) x (platform Linpack Rmax)

used for "normalization"



Language use by PRACE surveyed application codes

Language	No. of applications
Fortran90	50
C90	22
Fortran77	15
C++	10
C99	7
Python	3
Perl	2
Mathematica	1

About 50% use more than one base language

16 out of the 69 application codes combine Fortran with C or C++

PRACE EU Deliverable D6.1

PRACE Benchmark Suite

- 12 core applications
  - NAMD, VASP, QCD, CPMD, GADGET, Code Saturne, TORB, ECHAM5, NEMO, CP2K, GROMACS, N3D
- 8 additional applications:
  - AVBP, HELIUM, TRIPOLI_4, PEPC, GPAW, ALYA, SIESTA, BSIT
- Synthetic benchmarks for architecture evaluation
- Integrated into JuBE (Juelich Benchmark Environment)

PRACE Preparatory Phase project
(ended 6/30/2010)

PRACE prototypes: A selection of systems and components with existing, near-existing or emerging platform technologies

Objectives:

- Evaluate before buying Peta-scale systems
- Assessment of technology and architectures
- Share experience between partners
- Prepare benchmarks
- Foresee technology evolutions
- Foster collaborations between providers and users

PRACE Preparatory Phase project (ended 6/30/2010)

- (Near) existing technologies for 2009-2010
  - Full-featured systems, either:
    - % of large existing production systems (*scaling*)
    - or extensions of existing production systems (*technology updates*)
    - or dedicated (smaller) systems (*new technology*)
- Emerging technologies for 2011 and beyond
  - Mostly components/subsystems
  - Strong focus on density, efficiency, energy efficiency, programming methodology and productivity
  - Mostly attached processors (accelerators)

PRACE Prototypes

(Near) existing technologies for 2009-2010

IBM BlueGene/P (FZJ)
01-2008

IBM Power6 (SARA)
07-2008

Cray XT5 (CSC)
11-2008

IBM Cell/Power (BSC)
12-2008

NEC SX9
vector part (HLRS) 02-2009
x86 part 04-2009

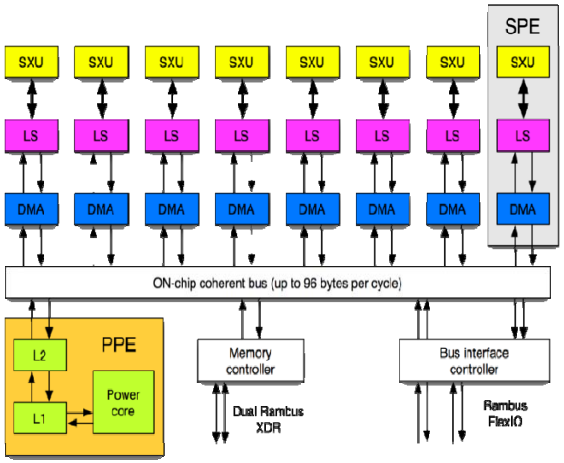
Bull Nehalem/INCA (CEA)
06-2009

PRACE Prototypes

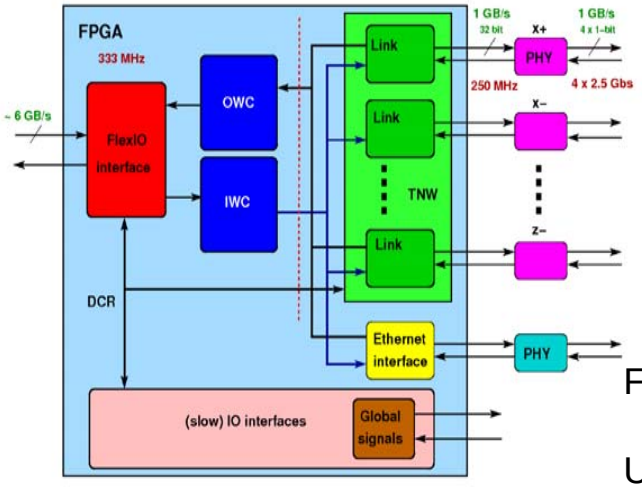
Emerging technologies

Prototypes	Installation Site	Targeted Components
SNIC-KTH	KTH, Sweden	Energy efficient computing
eQPACE	JSC, Germany	Interconnects, energy efficiency and density
LRZ-CINES 1	CINES, France BAdW-LRZ, Germany	Intel Nehalem-EP, ClearSpeed and QDR Infiniband
LRZ-CINES 2	BAdW-LRZ, Germany	Intel Nehalem-EX, Numalink5, Intel Larrabee
Maxwell FPGA	EPCC, UK	FPGA, energy efficiency and programing
ClearSpeed	NCF, Netherlands	ClearSpeed
Accelerator efficiency	PSNC, Poland, SFTC, UK	Power consumption, porting of applications
XC4-IO	CINECA, Italy	I/O and File System perf/, SSD for metadata,
PGAS Compiler	CSCS, Switzerland	PGAS programming model
Hybrid Technology	CEA, France	GPGPU, HMPP
RapidMind	BAdW-LRZ, Germany	Programming models for hybrid systems
PGAS Programming	CSC, Finland	Performance of UPC and CAF
Parallel GPU	CSC, Finland	Parallelizing CUDA, porting CUDA to OpenCL

eQPACE (extended QCD PArallel computing on Cell)

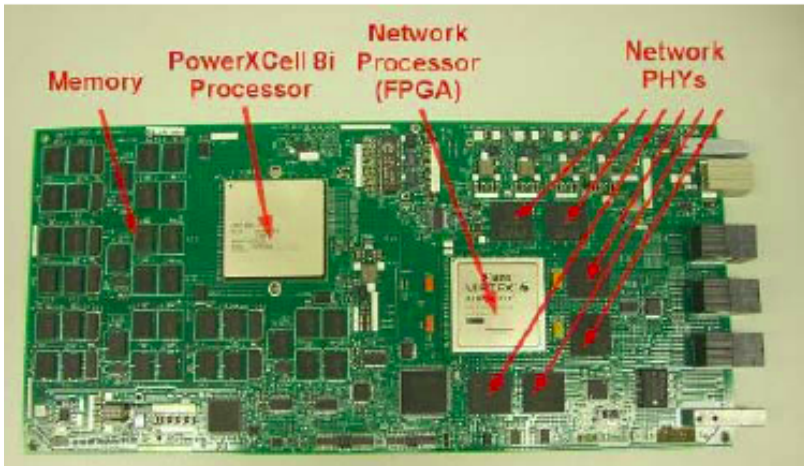


Cell processor PowerXCell 8i



eQPACE FPGA network processor (extension of QPACE)

ForschungZentrum Juelich
Univ of Regensburg
Univ of Wuppertal



eQPACE board

eQPACE with frontend at JSC

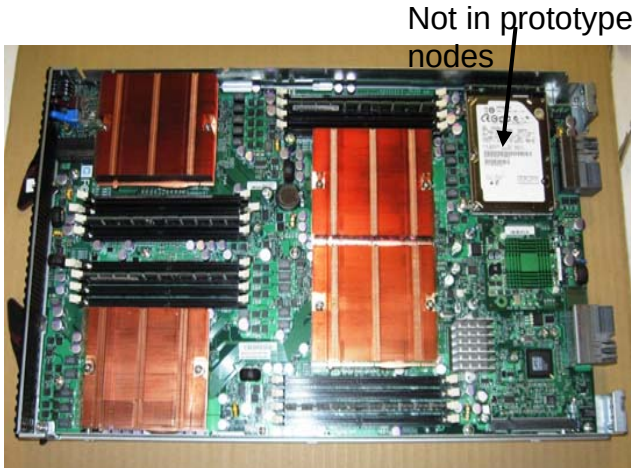
eQPACE Greenest of Green ….

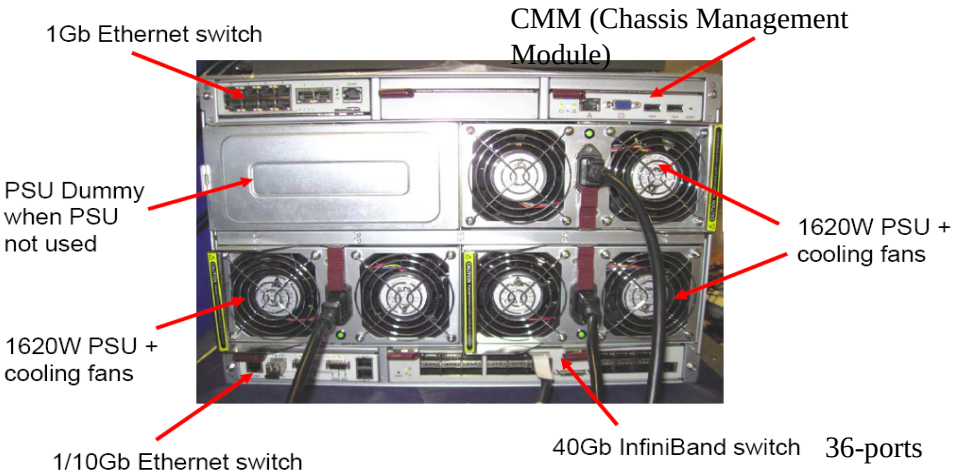
Green500 Rank	MFLOPS/W	Computer June 2010 List	Power (kW)
1	773.38	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.74
1	773.38	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.74
1	773.38	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.74
4	492.64	Nebulae	2580.00
5	458.33	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , IB	276.00
5	458.33	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , IB	138.00
7	444.94	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire IB	2345.50
8	431.88	Mole-8.5 Cluster Xeon L5520 2.26 Ghz, nVidia Tesla, IB	480.00
9	418.47	iDataPlex, Xeon X56xx 6C 2.8 GHz, IB	72.00
10	397.56	iDataPlex, Xeon X56xx 6C 2.66 GHz, IB	72.00

PRACE SNIC/KTH Prototype

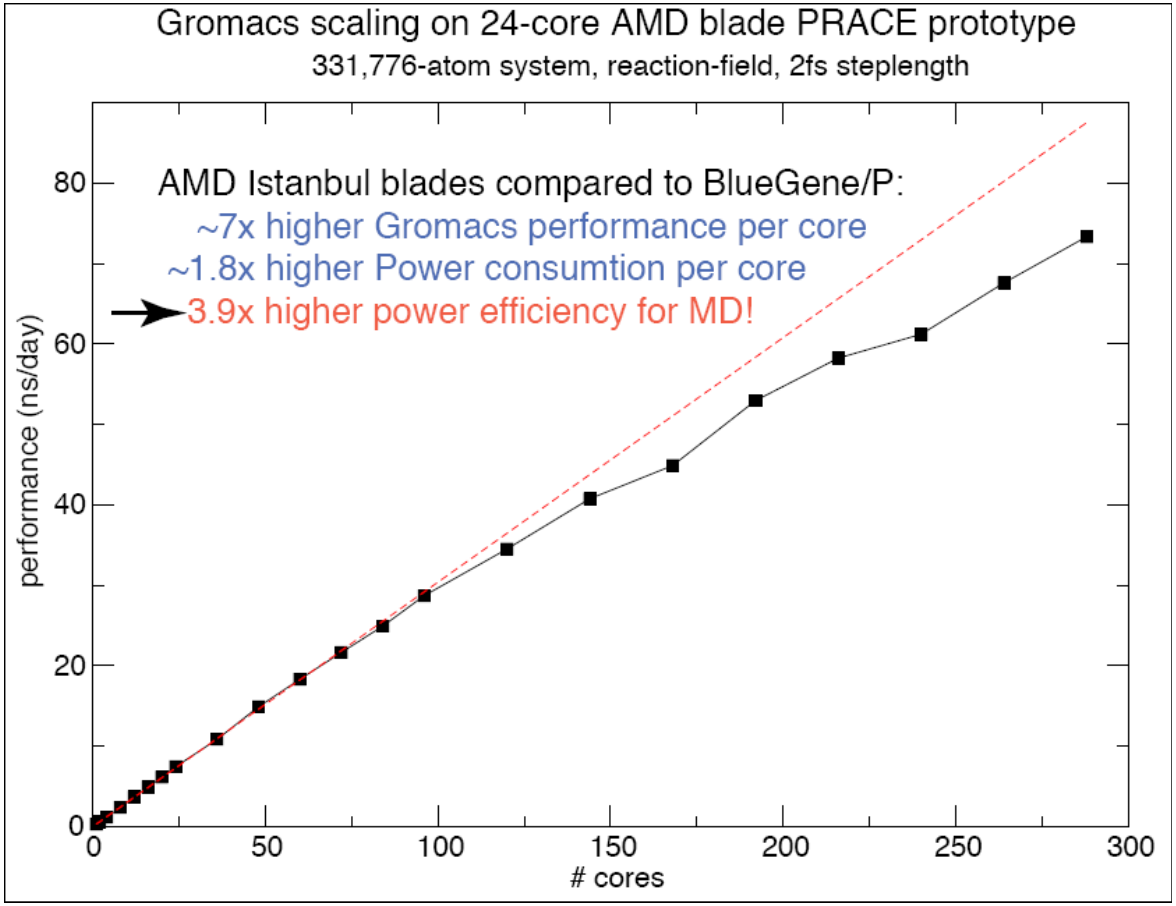
Target: BG/P efficiency using commodity components only an no acceleration

- BG/P Green500 November 2008 357.14MF/W – 371.67MF/W
- Achieved: 343.91 MF/W
  The prototype has more memory/core with the added memory representing about 10% of the power consumption



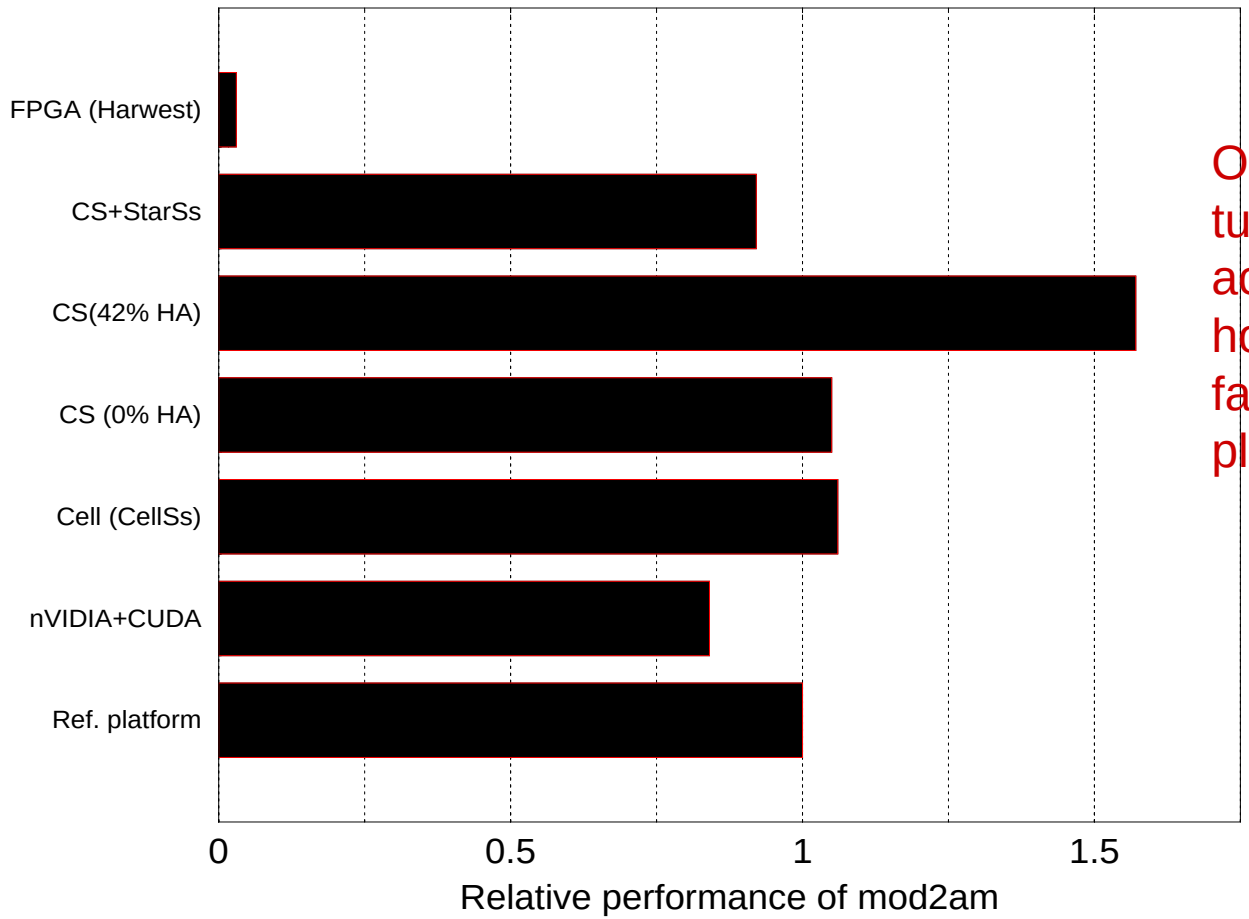


PRACE SNIC/KTH Prototype Gromacs



EuroBen results – mod2am

Performance relative to Reference Platform of mod2am

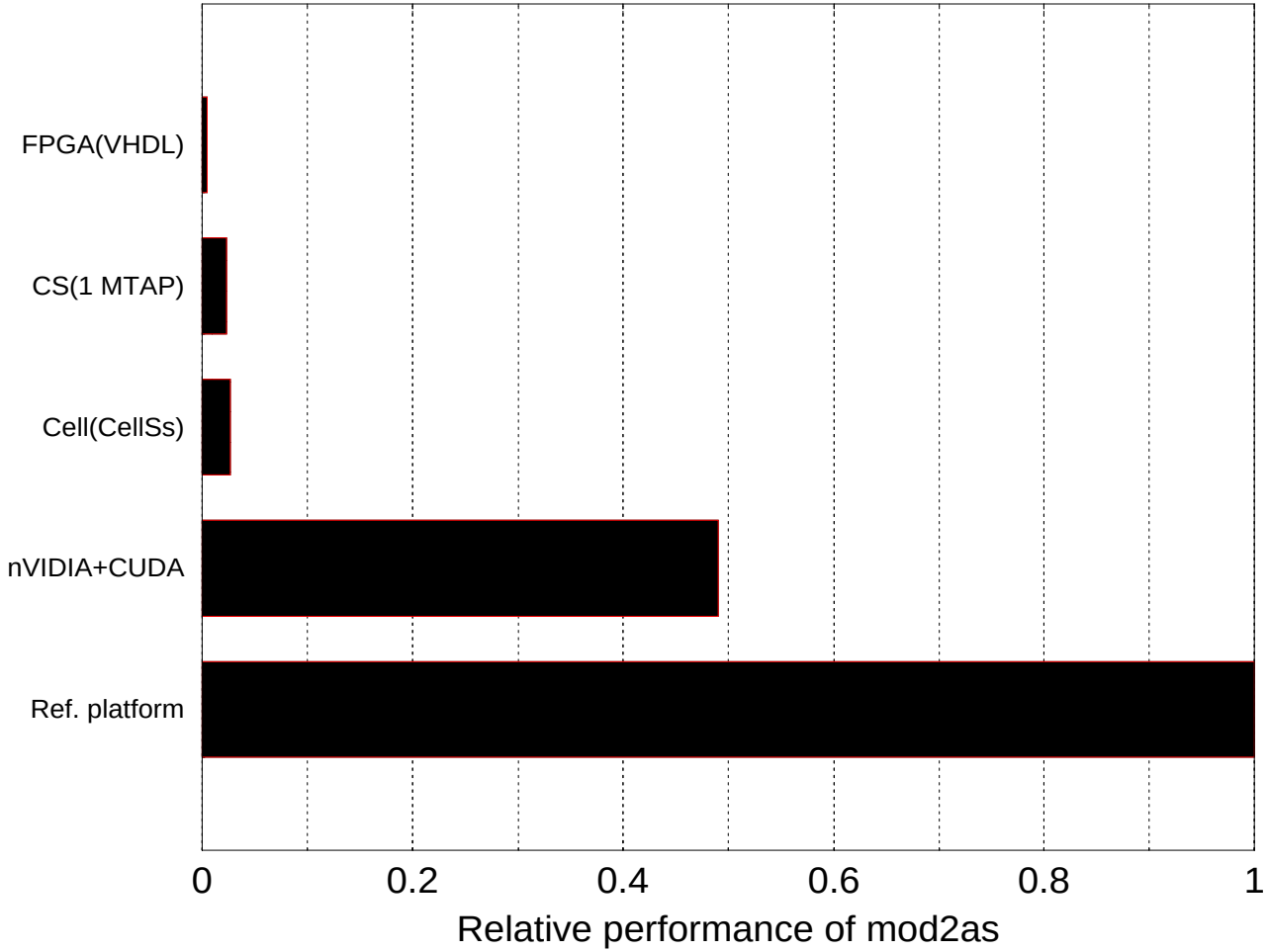


Observation: only the tuned ClearSpeed accelerator with a 42 % host assist is about 50 % faster than the reference platform.

PRACE EU Deliverable D8.3.2

EuroBen results – mod2as

Performance relative to Reference Platform II of mod2as



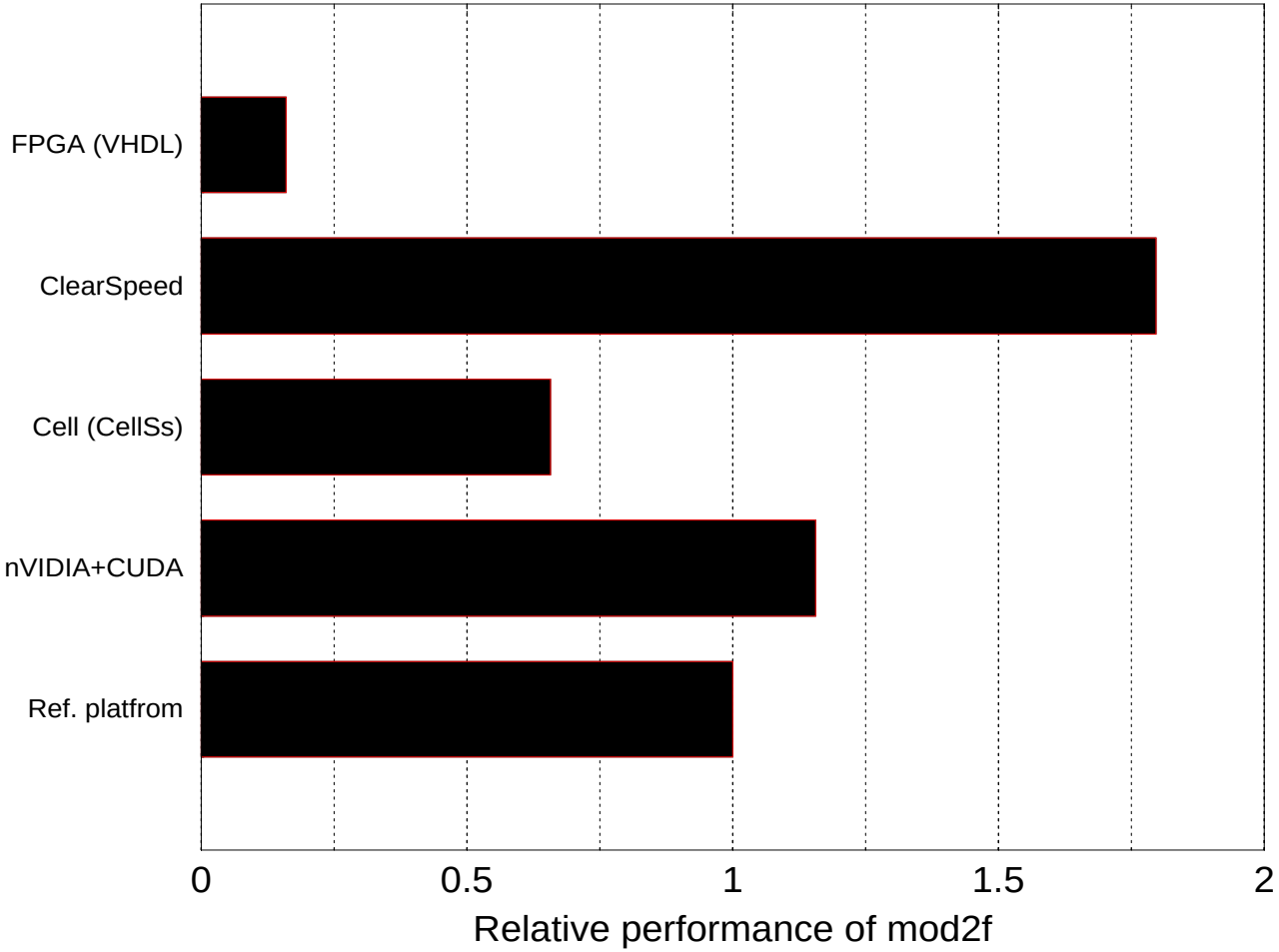
No accelerator is able to attain a decent fraction of the performance of the Reference Platform II (1,392 MFlop/s).

Contributing causes:

- irregular memory references low computation to memory reference ratio (2/3)
- memory bandwidth limitations
- poor support for inner-products for all accelerators evaluated

EuroBen results – mod2f

Performance relative to Reference Platform II of mod2f



As for mod2am only the Clearspeed accelerator achieved a significant improvement over the reference platform

Emerging Technologies Prototypes HPL Summary

Efficiency

Reference platform	91%
Reference platform + Clearspeed	61.8%
Reference platform + GPU	52.5%
SNIC/KTH prototype	79% (preliminary)
eQPACE (Cell)	79.9%

Reference Platform: Dual socket Nehalem 2.53GHz

Emerging Technologies Prototypes HPL Summary

Power Efficiency

Reference platform	240 MF/W
Reference platform + Clearspeed	326 MF/W
Reference platform + GPU	270 MF/W
SNIC/KTH prototype	344 MF/W(prelim.)
eQPACE (Cell)	773 MF/W

Reference Platform: Dual socket Nehalem 2.53GHz

Programmer Productivity
Code size

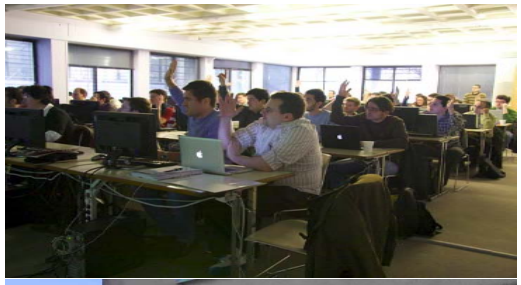
- A few languages stick out positively
  - Chapel ~1/16 LoC
  - RapidMind ~1/4 LoC
- Others require many lines of source code
  - Cn ~2x
  - CUDA ~3x unless CUDA library calls can be used
- No obvious conclusions for CellSs
  - Few for FFT, medium for matrix-matrix multiplication, many for sparse matrix-vector multiplication
- All other languages are comparable
  - MPI+OpenMP often the shortest
  - UPC one of the largest

Programmer Productivity Porting time

- Most ports done within 5 days
- Very short porting time (< 1day) for
  - Chapel, MIP+OpenMP and CAPS HMPP
- Relatively long (5 – 10 days) porting time for
  - CUDA, OpenCL, MPI+CUDA
- Porting time for UPC unnaturally long due to many problems with an immature compiler
- Porting time for CellSs was very long, in part due to programmers attempting to optimize for performance
  - First version running within a few days

Education and Training Highlights

- Petascale training and education needs surveyed Spring 2008
- First PRACE Summer School on Peta-scaling, KTH, Stockholm, August 2008. Platforms: IBM Blue Gene/P (FZJ) (65,536 cores) and Cray XT4 (CSC) (10,816 cores),
- First PRACE Winter School on Scalable Programming Models and Paradigms, GRNET, Athens, February 2009. Platforms IBM Power 6 (3,328 cores) and IBM Cell (1,152 SPE cores)
- Seven Code Porting Workshops in 2009



In total 270 participants in education and training events

PRACE Code Porting Workshops

- GPU and Hybrid system programming using CUDA and CAPS-HMPP, CEA, Paris, April 2009
- Porting and optimization techniques for PRACE applications, CSC, Helsinki, June, 2009
- Porting and optimization techniques for the CRAY XT5, CSCS, Manno, Switzerland, July, 2009
- Porting and optimization techniques for the Clearspeed/Petapath architecture, NCF/SARA, Amsterdam, October, 2009
- Porting and optimization techniques for the NEC/SX-9 (HLRS) and IBM BG/P (FZJ), Cyfronet, Cracow, October, 2009
- Porting and optimization techniques for the IBM Cell (BSC) and GPGPU systems, BSC, Barcelona, October, 2009
- Stream Programming with OpenCL, KTH, Stockholm, December, 2009

PRACE 1st Implementation Phase Project - increased focus on Tier-1 (started 7/1/2010)

- User & Community support through application enabling> 40% of the total effort
  - Support can be requested along with proposals for Preparatory Access to the Tier-0 systems
- Deployment and operation of the Technical Infrastructure
- Collaboration with Communities and other Research Infrastructures
- Development of a model for cross-national Tier-1 access
  - This activity will be extended in the future implementation phase projects
- Cooperation with vendors for future HPC technologies~20% of the total workforce + 5 Million €for prototypes (50% EC-funded)
- Further development of the legal, organizational and financial framework
- Continuation and further extension of the very successful training program started in the Preparatory Phase project

STRATOS (STRAtegic TechnOlogieS)

Leader: Herbert Huber, LRZ, Germany

- An initiative by PRACE to enable vendor collaboration on Evaluation and Development of Next-Generation Technology
- MoU signed by 12 PRACE partner representatives
  - Forschungszentrum Jülich (FZJ), Germany
  - Universität Stuttgart (HLRS), Germany
  - Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften (BADW-LRZ), Germany;
  - Grand Equipement National de Calcul Intensif (GENCI), France;
  - Barcelona Supercomputing Center (BSC), Spain;
  - Netherlands National Computing Facilities Foundation (NCF), the Netherlands;
  - Swedish National Infrastructure for Computing (SNIC), Sweden;
  - CINECA Consorzio Interuniversitario (CINECA), Italy;
  - CSC – IT Center for Science Ltd. (CSC), Finland;
  - Eidgenössiche Technische Hochschule Zürich (ETHZ), Switzerland;
  - Greek Research and Technology Network S.A (GRNET), Greece
  - Poznan Supercomputing and Networking Center (PSNC), Poland
- Industry can join as Associate Members

STRATOS (STRAtegic TechnOlogieS)

- Three Working Groups
  - Technology Watch
    (Lead Jonathan Follows, STFC, United Kingdom)
  - Green IT and HPC Leadership Resources
    (Lennart Johnsson, SNIC/KTH, Sweden)
  - Exa-scale Software Working Group
    (Peter Michelse, NCF, the Netherlands)


PRACE
PARTNERSHIP FOR
ADVANCED COMPUTING
IN EUROPE
IESP October 18, 2010

Thank YOU!