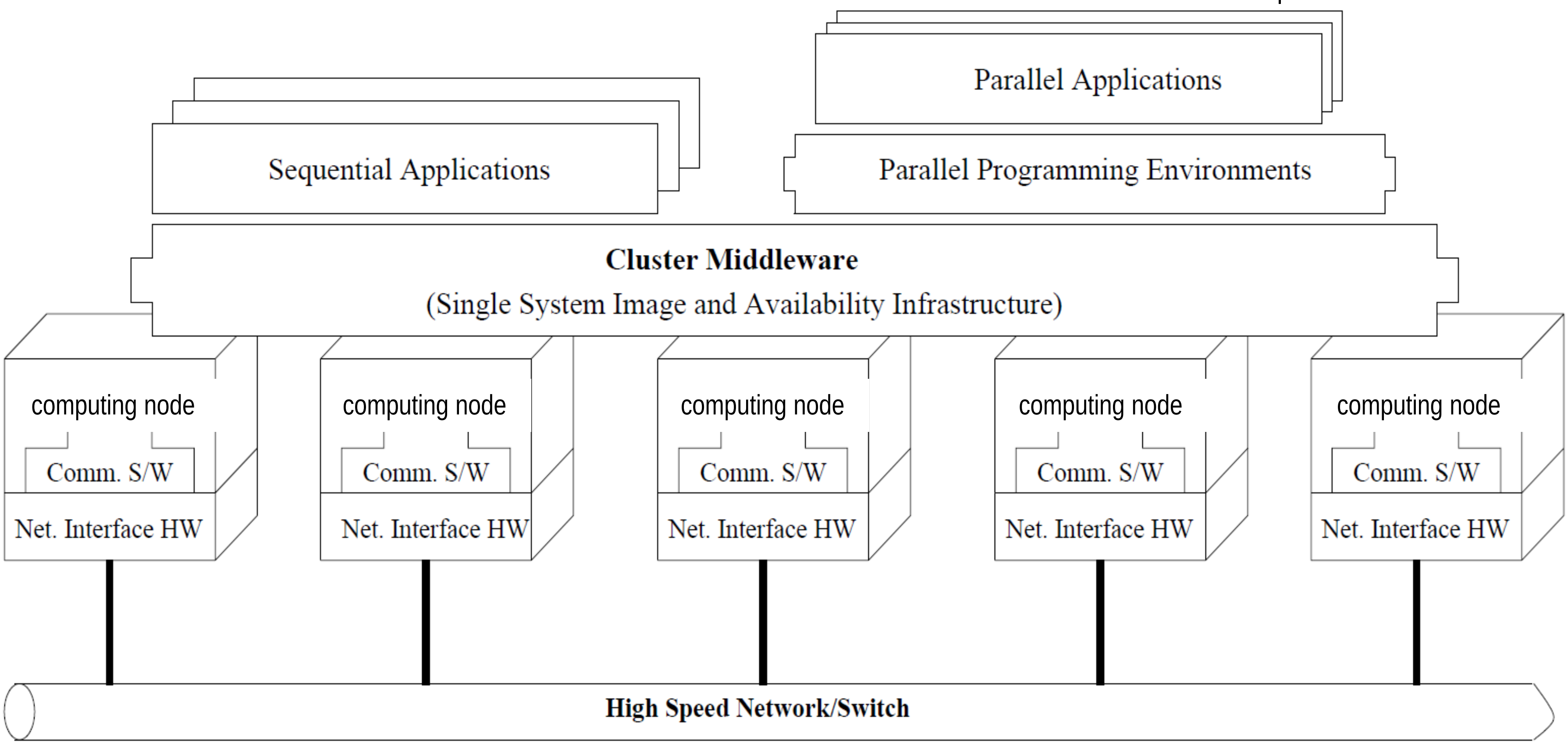
U. de Valparaíso, Escuela Ingeniería Industrial

U. de Chile, Centro de Modelamiento Matemático



Introduction to Parallel Computing

Gonzalo Hernandez

Parallel Programming Models: Agenda

1) HPC in Chile: NLHPC

2) Why Use Parallel Computing ?

3) Some Computing Intensive Applications

4) Cluster Computing (Very Brief Introduction)

5) Cluster Computing Trends

6) Some Old Supercomputers

7) Supercomputers Evolution

CMM
Center for Mathematical Modeling
Advanced Math for Business

National Laboratory for High Performance Computing (NLHPC)

First National Call for Major Scientific and Technological Equipment for User Facility Centers
2010

usach

Universidad de Talca

Reuna Ciencia y Educación en Red

fcfm

MISSION

To consolidate a national facility for HPC by offering top quality services and advanced training to answer the national demand for scientific computing, developing links between research groups, the industry and the public sector

VISION

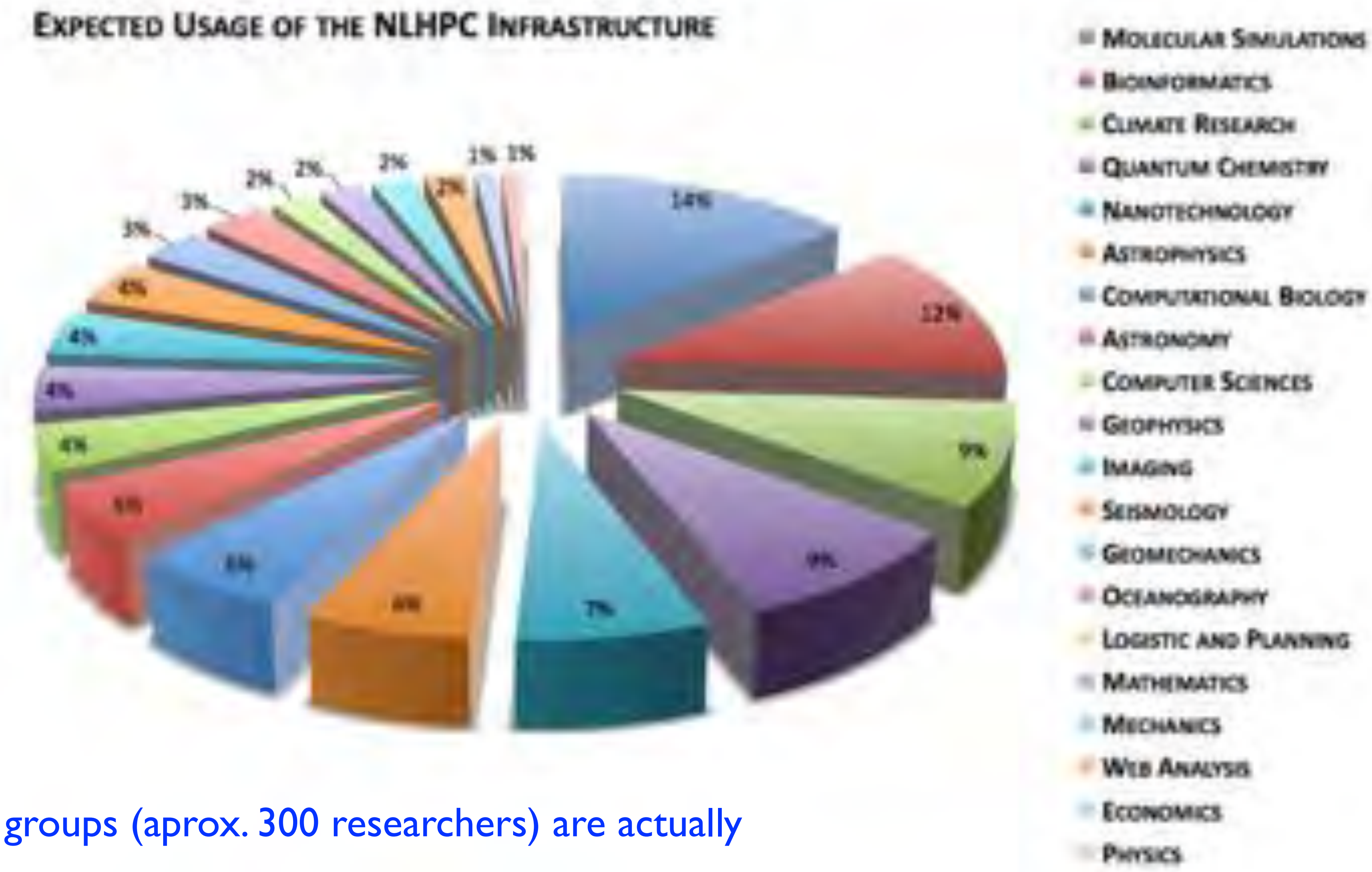
We envision the NLHPC as a competitive facility of world-class quality for research services in HPC

Objectives

Critical IT partner to large-scale and data intensive research projects based on HPC

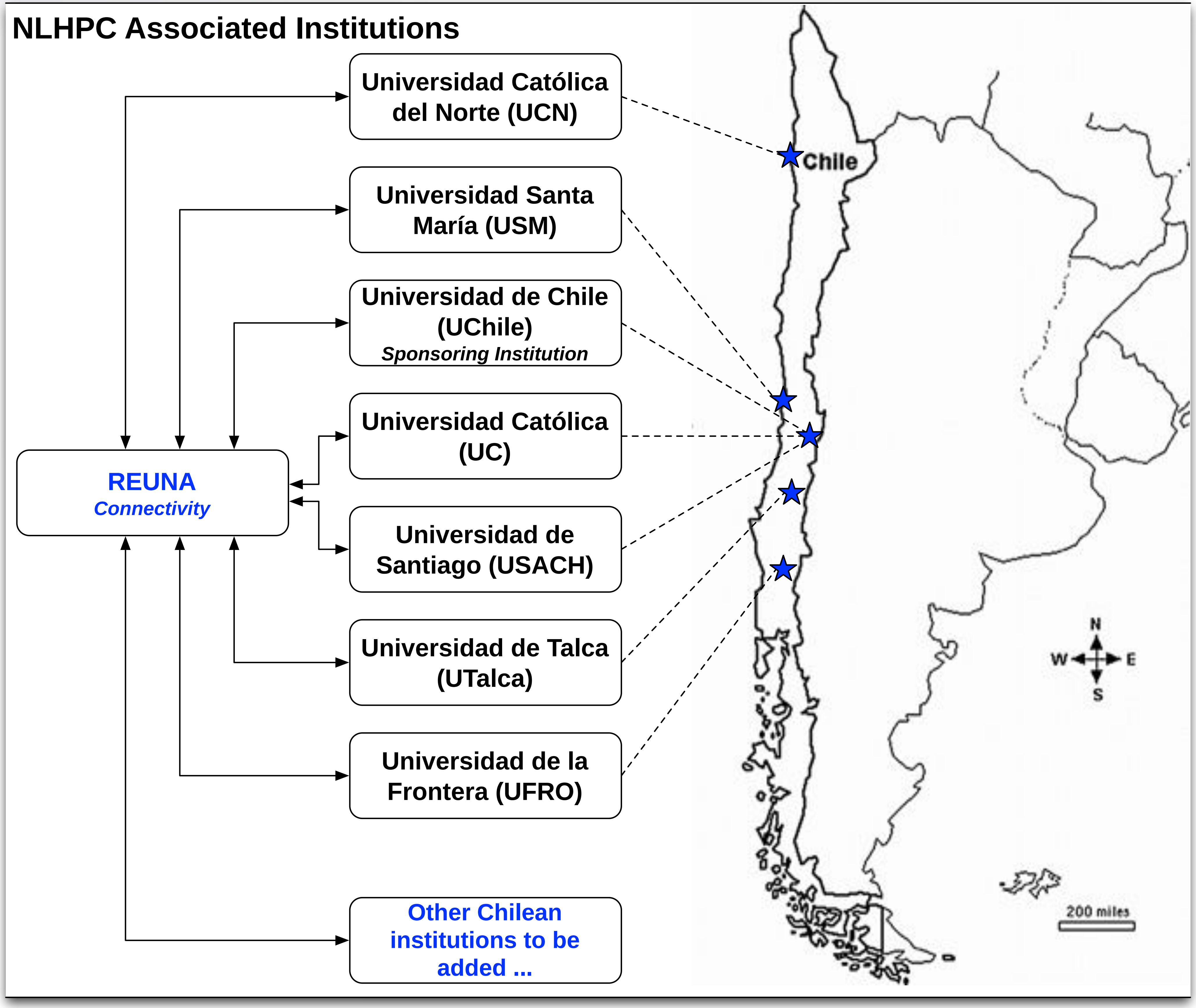
1. To develop and support powerful and reliable computing and network resources enabling national researchers to have access to HPC capacity to solve computing- and data-intensive scientific problems.
2. To stimulate and participate in the creation of a national high-capacity network for the transmission of research data.
3. To use and explore innovative architectures and techniques to accelerate scientific computing.
4. To trigger new industrial initiatives in HPC.
5. To provide insights into advanced modeling needs through the NLHPC's scientific network.
6. To help educate the next generation of scientists and engineers in HPC.
7. To increase social awareness of the role of HPC in contemporary society and technological development.

Estimated Scientific Demand for HPC: A Key Issue for for improving Chilean competitiveness

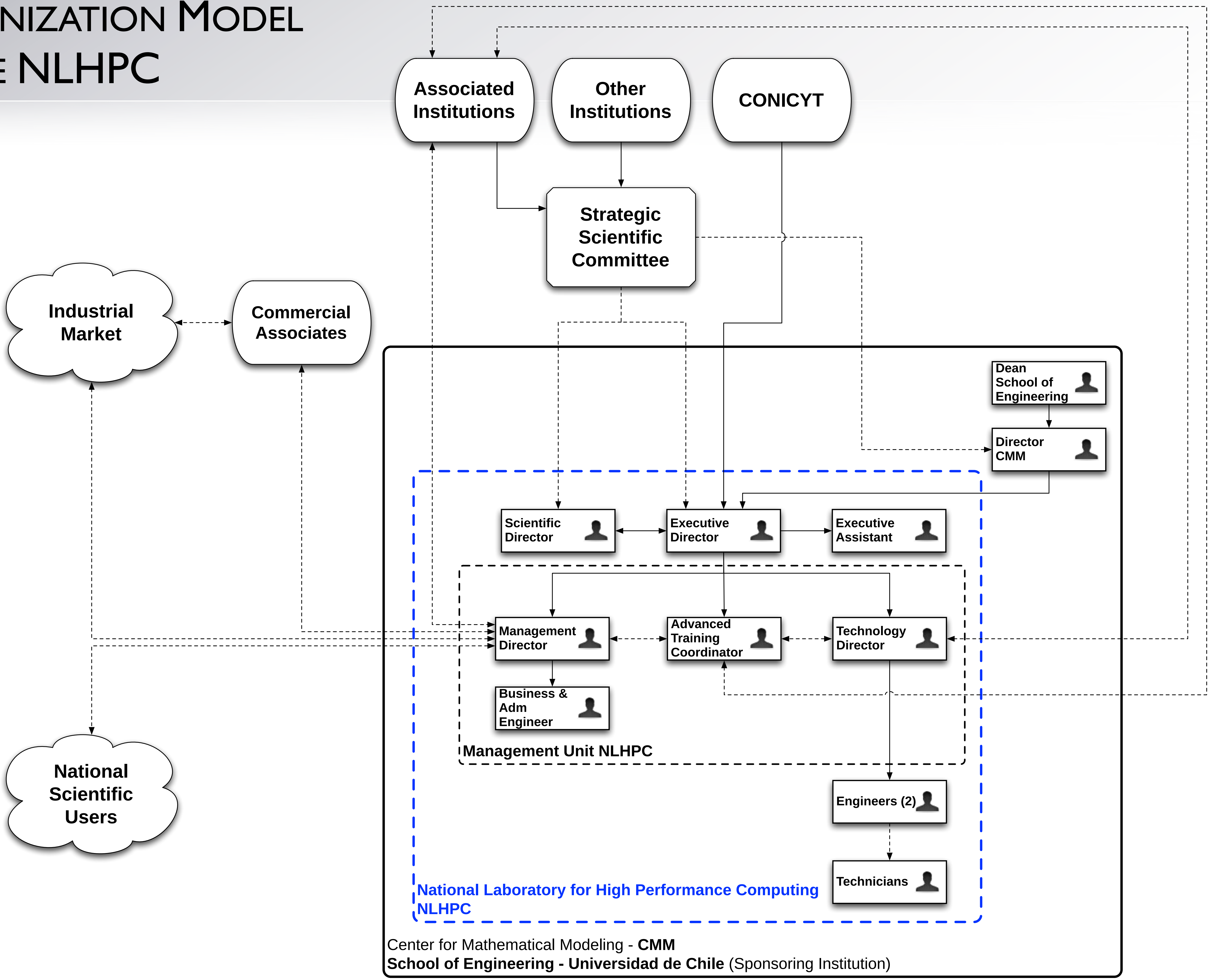


- Near to 50 Chilean research groups (aprox. 300 researchers) are actually demanding HPC resources
- The expansion of Chilean research in data-driven sciences such as Omics, Bioinformatics and Astronomy, requires a national facility for HPC services
- Scientific contributions in areas such Molecular Simulations, Quantum chemistry and Astrophysics are highly dependent on the availability of HPC capacities
- Astronomy will become a highly demanding research field for HPC: LSST and AURA will produce more than 40TB of raw data per night

ASSOCIATIVE MODEL: AN OPEN MODEL



ORGANIZATION MODEL OF THE NLHPC



The NLHPC Strategic Scientific Committee (SSC)

Eduardo Vera
CMM - UChile
Advanced Networks
NLHPC Executive Director

Alejandro Jofré
CMM - UChile
Mathematics
Committee's Scientific Director

Servet Martinez
CMM - UChile
National Prize (Mathematics)

José Maza
FCFM - UChile
National Prize (Astronomy)

Pablo Valenzuela
Fund. Ciencia para la Vida
National Prize (Biotechnology)

Marco Serón
UCN
Computer Sciences

Mauricio Marín
USACH
Computer Sciences

Luis Salinas
USM
Computer Sciences

Alvaro Soto
UC
Robotics

Nélson Padilla
UC
Astronomy

Chris Smith
AURA
Astronomy

Danilo Gonzalez
UTalca
Molecular Simulations - Nanobiotechnology

Andrés Ávila
UFRO
Mathematics

José Palacios
REUNA
Advanced Networks

Myrna Araneda
Weather Agency - Chilean Air Force
Meteorology

The Executive Director will be the ultimate responsible for the management of the NLHPC:

- Leads the NLHPC Team
- Reports to the SSC
- Responds to CONICYT

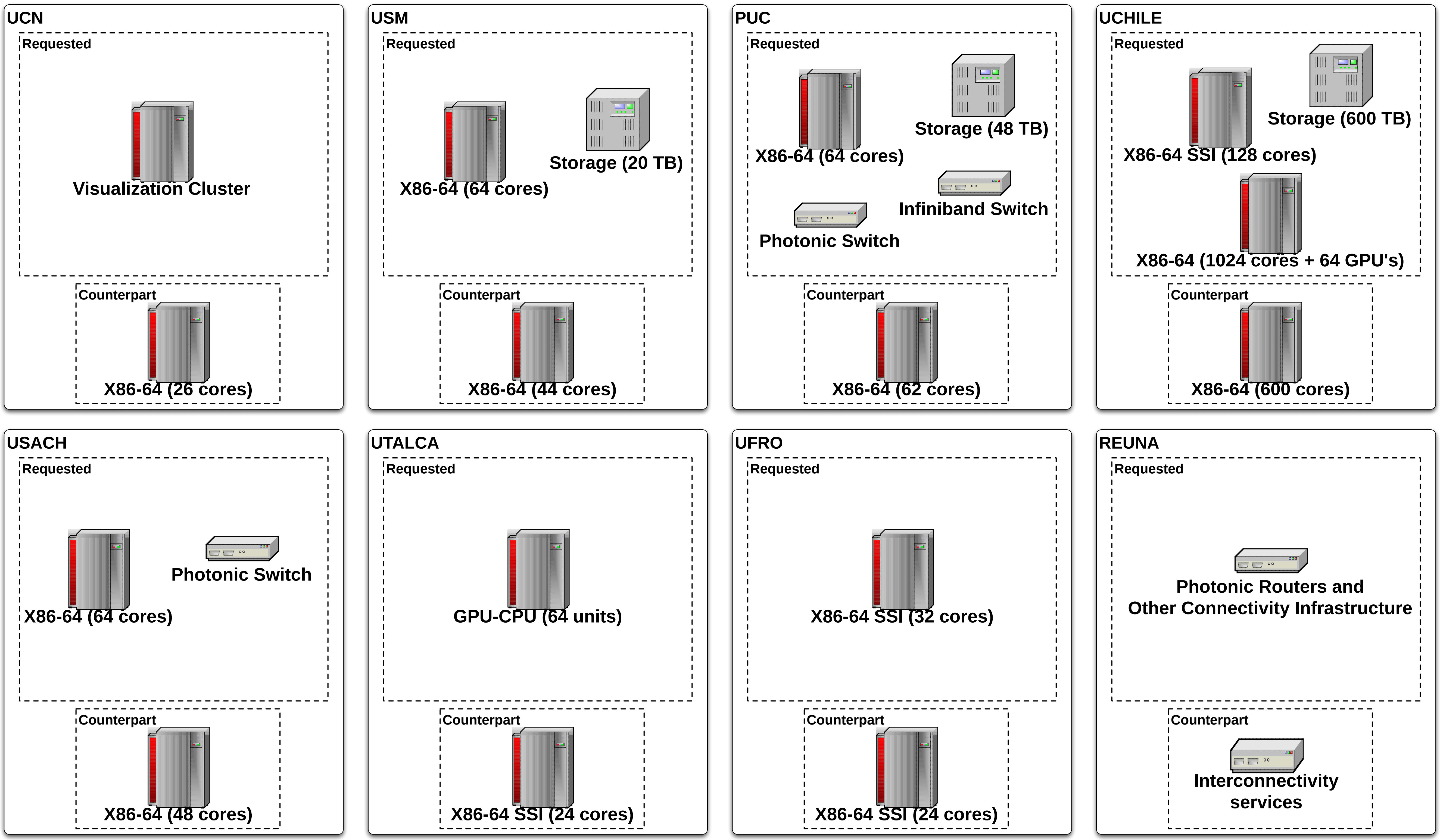
The SSC will meet at least twice per year to:

- Advice on strategic planning of the NLPHC
- Advice on the priorities for the NLHPC's infrastructure usage
- Analyze performance's accountability

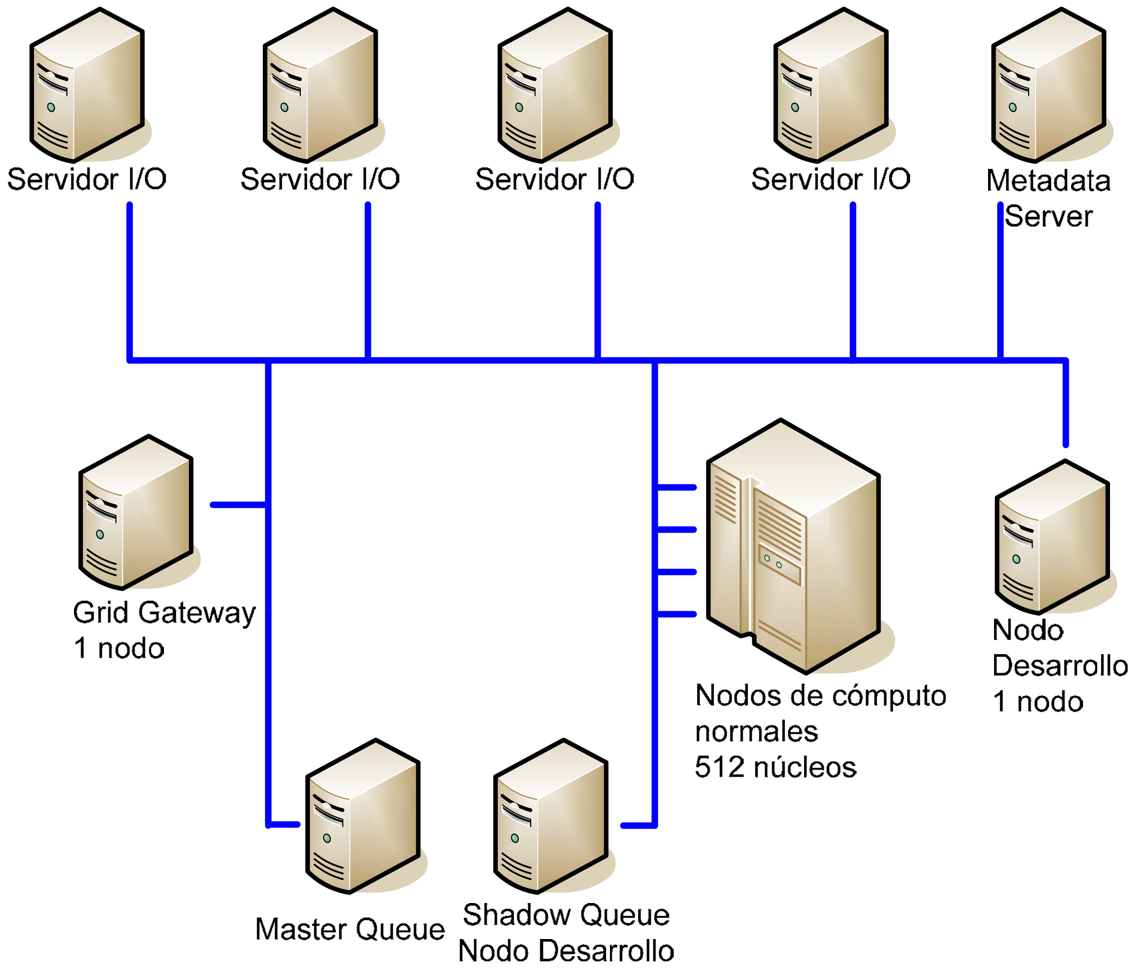
NLHPC INTERNATIONAL NETWORK: GLOBAL OUTLOOK

- Access to advanced training
- Technological support network
- Collaboration agreements and joint research projects
- Participation on international HPC networks and projects: Gellato Federation, PRAGMA, GISELA
- CMM's technology partners: IBM, HP, Microsoft and NTT

The NLHPC Supercomputing Infrastructure



NLHPC: Levque (Relámpago)

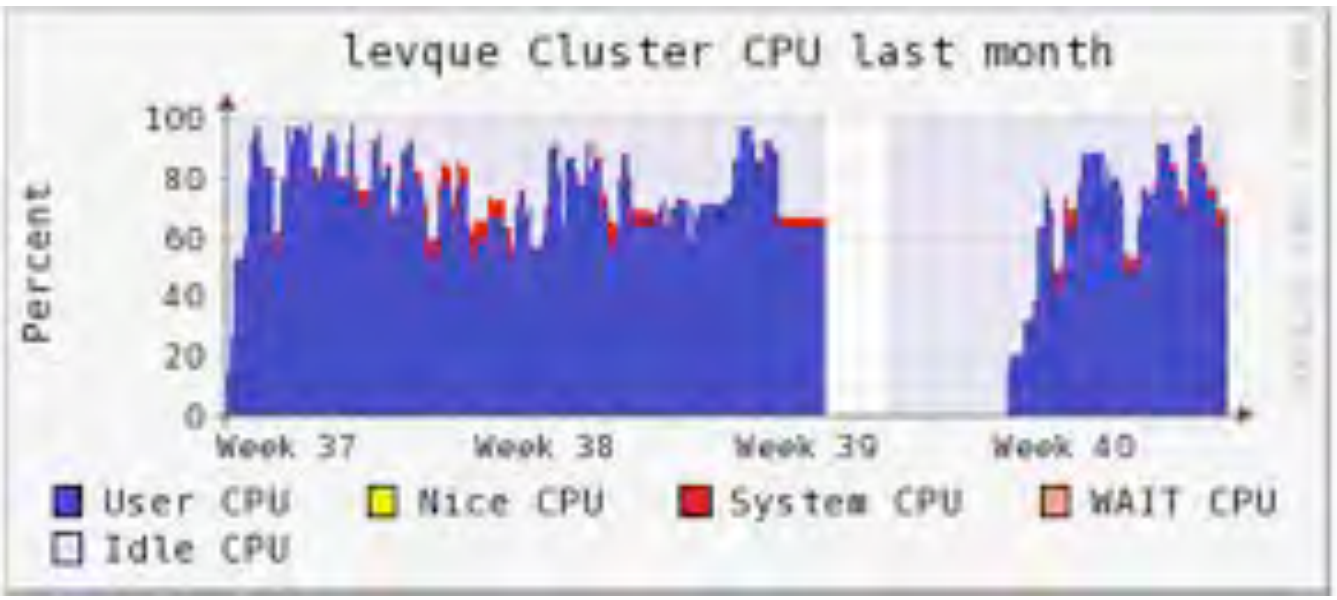


Red Infiniband para I/O y MPI

Current CMM Technology: An Advanced Computing Cluster Open to the Chilean Scientific Community

IBM iDataplex machine (Levque)

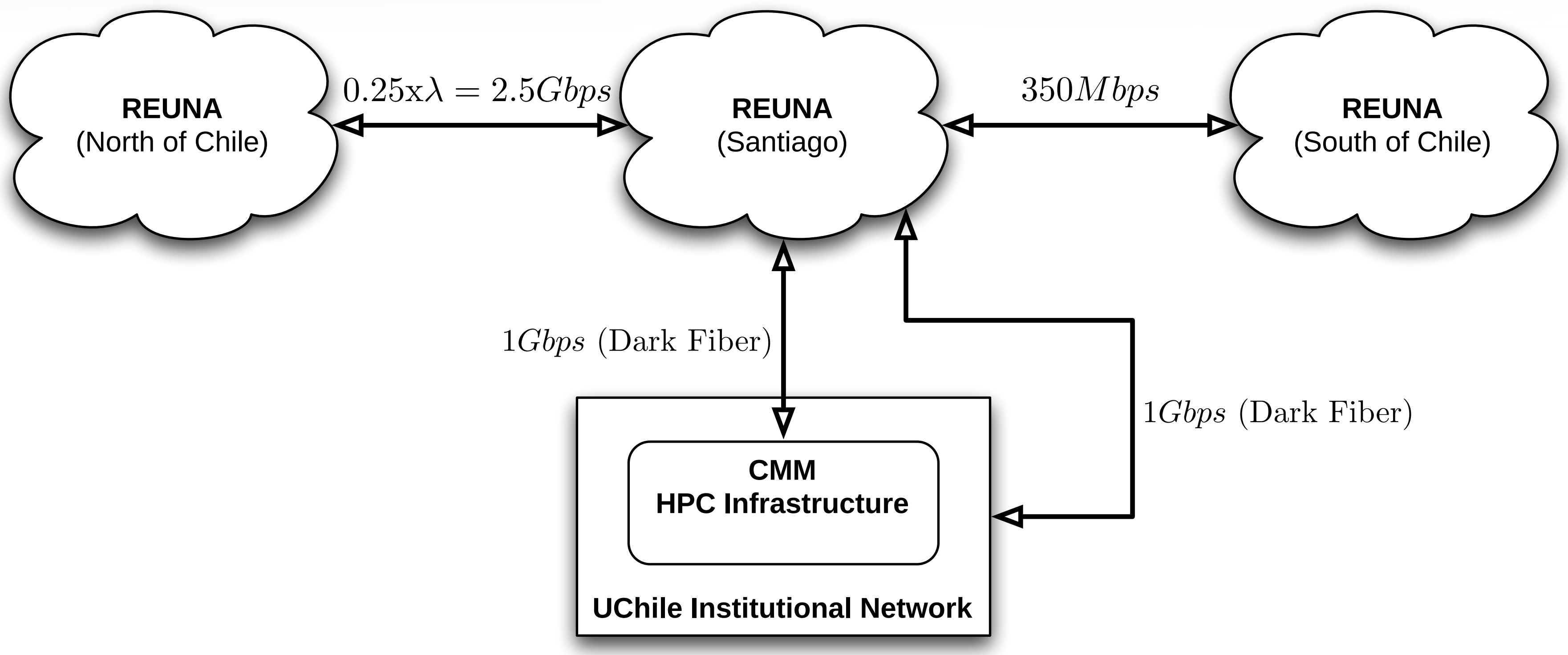
- 66 computing nodes
- 600 cores, Intel Xeon X5550 (nehalem)
- 3 GB per core
- Two IB ports 40 GB/s
- 100% non-blocking Infiniband interconnection Switch Qlogic 12800-180
- Nodes for queuing, metadata storage, applications testing, development, grid access and webservices access



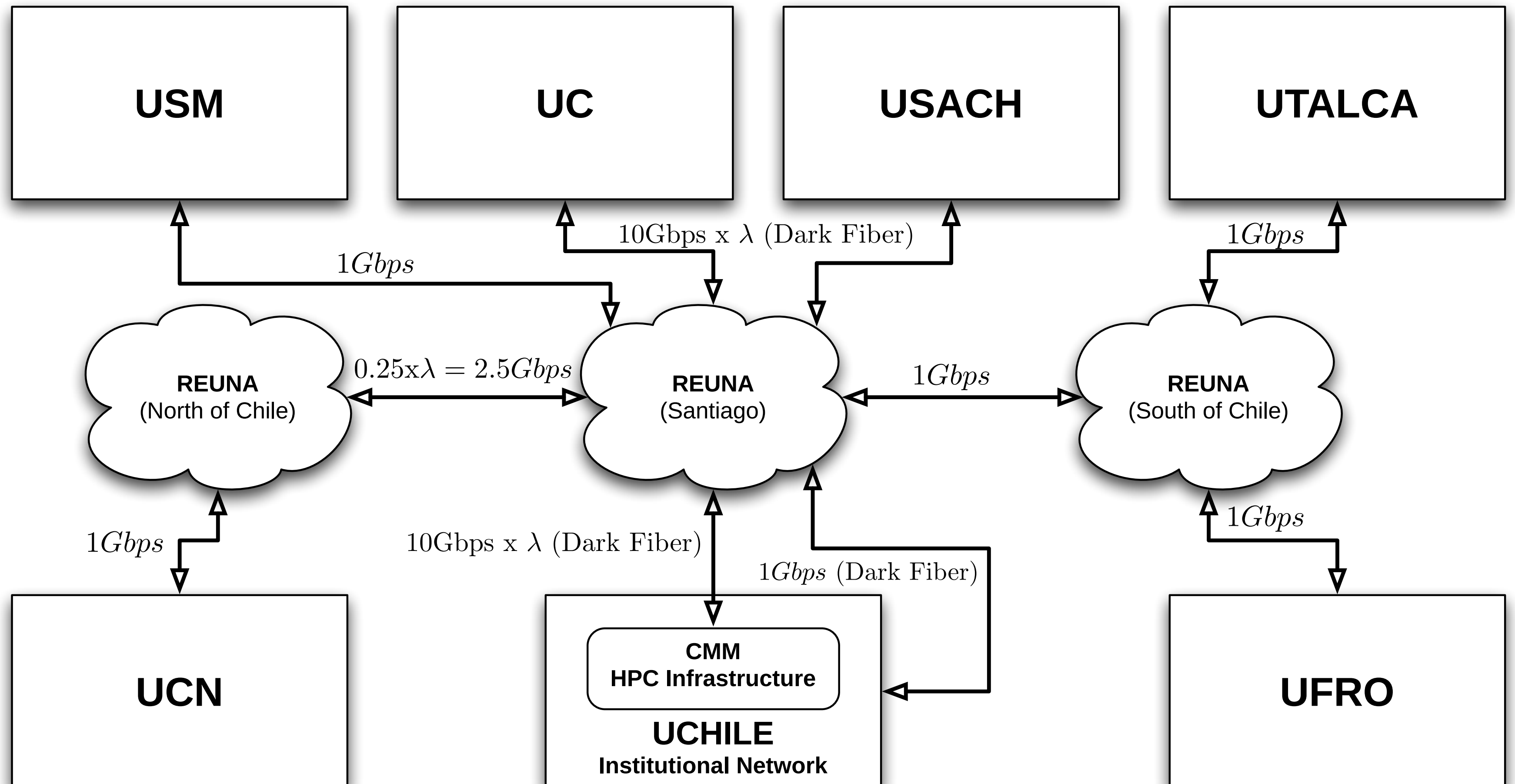
- Data gathered during last month with only 20% of potential users
- 75% of average CPU load

NLHPC Advanced Network Technology: Strengthening Academic Networks across the country

Current situation



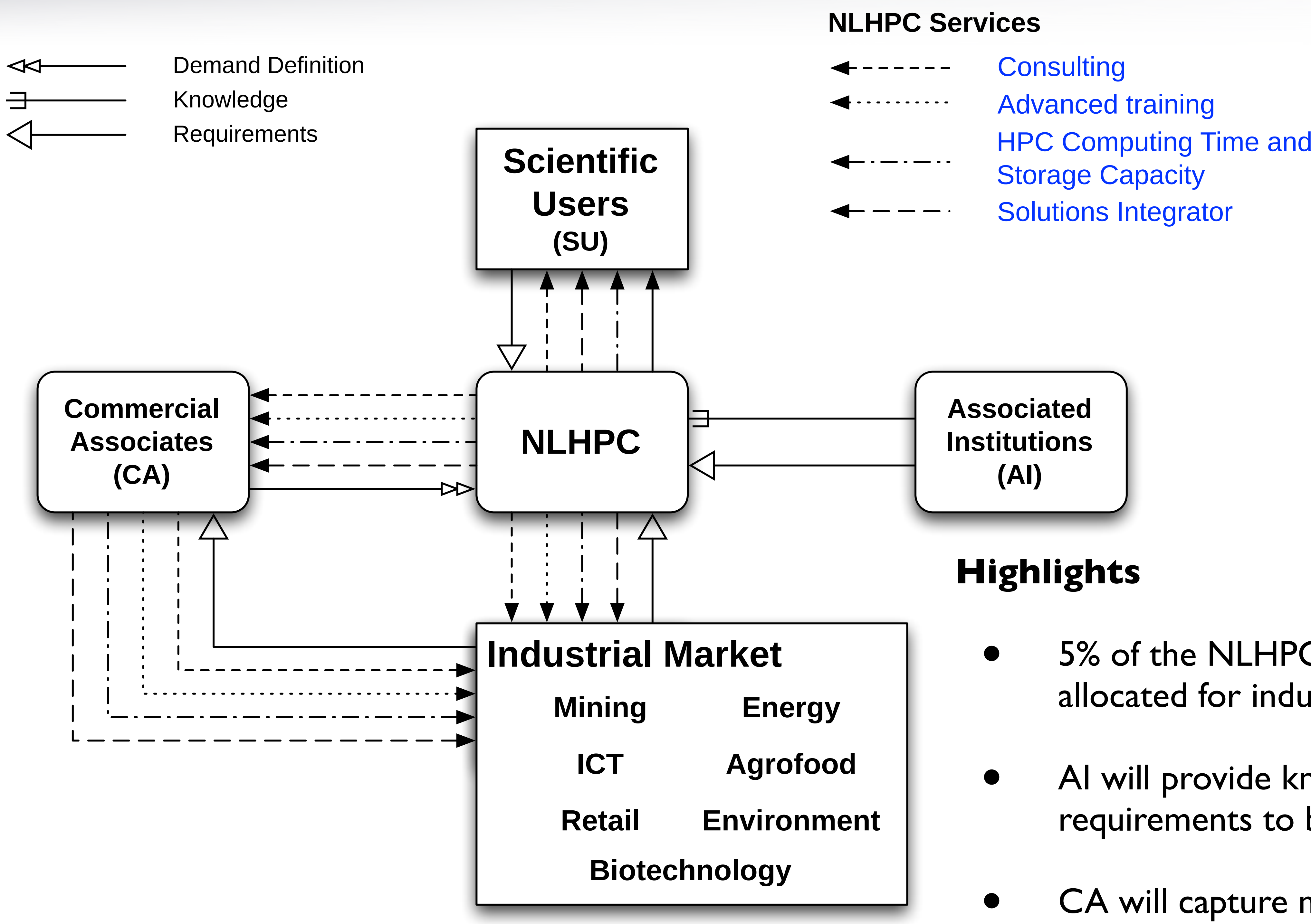
NLHPC situation



Additional impact of the network enhancement:

- e-learning
- Video-conferencing
- Accessing e-infrastructures
- Distributing knowledge

Sustainability Model: Supporting Marginal Costs



Highlights

- 5% of the NLHPC infrastructure will be allocated for industrial applications
- AI will provide knowledge and requirements to be satisfied by the NLHPC
- CA will capture market requirements to be satisfied by the NLHPC
- NLHPC can join the AI, as Solutions Integrator, to satisfy complex industrial requirements

Industrial Liason: Open Innovation Program in HPC

The CMM Approach to Industrial Liason

- Strong ties with the National and International industry
- Yearly budget of the CMM US$7,000,000
- 50% of the CMM's budget comes from Industrial applied projects

Open Innovation Program in HPC

- Annual memberships with no fees
- Deep discussions of HPC impacts on Industry
- Business-to-business collaboration
- Research agreements with identifiable deliverables
- Resource-use agreements for HPC
- Co-location of personnel

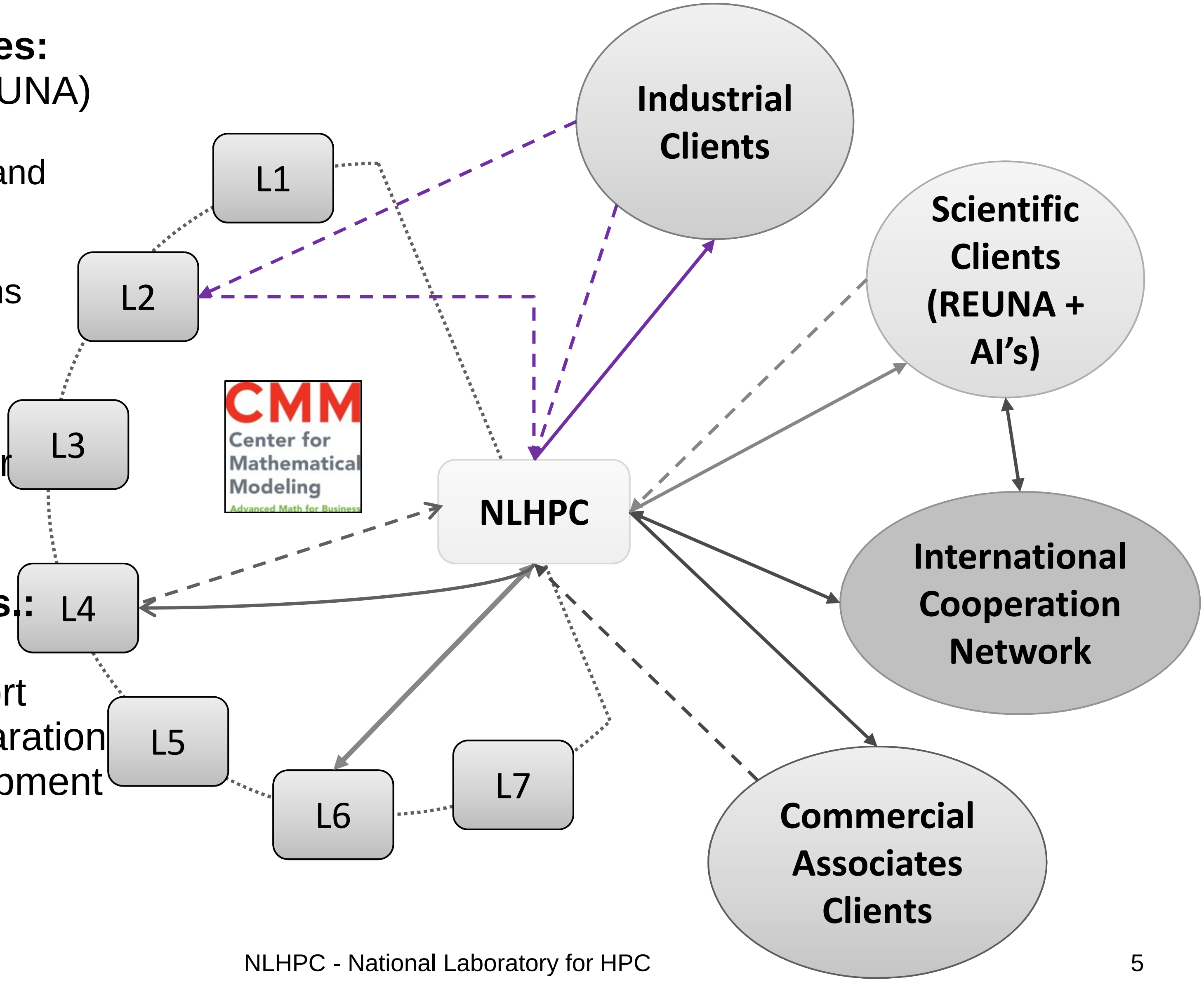
NLHPC Training & Partnership: 2011 - 2015

Training Activities:

- E-learning (REUNA)
  Videoconference
  Training on demand
  Digital library
  Online courses
  Virtual classrooms
  Virtual Labs
- Seminars
- Short Courses
- Semce Modular Courses

Partnership Acts.:

- Meetings
- Training Support
- Proposal Preparation
- Project Development



Training Proposal for NLHPC

- Goals:
  - Bring up to date the requirements for education and training of the Chilean scientific community in the area of HPC.
  - The development of a culture of efficient use of the techniques and methods of HPC as a tool to study real system models found in science, engineering or industry.
  - The education of under- and postgraduate students in the area of HPC by means of courses in engineering careers, diplomas and postgraduate programs.
  - The realization of training activities which enable continuous learning of classical techniques and new HPC technologies (Hardware systems, Programming models, Languages, Libraries)

NLHPC: Organization and Participation in Conferences

Organization and Participation in Conferences and Workshop

- Organization of five CLGrid workshops
  - Short training courses including laboratory sessions
  - Plenary conferences.
  - Presentations of scientific articles and thesis projects.
  - Talks by companies related to HPC.
- HPC 2010 & 2009 Latam: Latin American HPC Symposium, Mar del Plata & Buenos Aires, Argentina.
- CLCAR 2009: Latin American Conference on HPC, Venezuela.
- XIII WSDP: Workshop on Parallel and Distributed Systems, November 9 – 14, Santiago, Chile.

NLHPC: Participation in PASI School 2011

- PASI School (Pan-American Advanced Studies Institute): “Scientific computing in the Americas: the challenge of massive parallelism”.
- http://www.bu.edu/pasi/
- Dates: January 03 - 15 enero, 2011.
- Place: Santa María University, Valparaíso, Chile.
- International Committee: Lorena Barba (Boston University), Matthew Knepley (University of Chicago), Luis Miguel de la Cruz (UNAM, Mexico).
- Local Committee:

  Luis Salinas, Javier Cañas & Raquel Pezoa (U. Santa María),

  Gonzalo Hernández (CMM, U. of Chile)

NLHPC: HPC School Chillán, December 06 – 10, 2010

- Initiative of Universidad del Bio-Bio and HPC Laboratory CMM.
- Dates: January 03 - 15 enero, 2011.
- Place: Universidad del Bio-Bio, Chillan, Chile.
- Organizing Committee: Natanael Guerrero (U. Bio-Bio)

  Gonzalo Hernández (CMM, U. of Chile)

  Courses with hands on sessions:

  – Cluster Computing
  – MPI and OpenMP Libraries
  – Performance Evaluation od Parallel Aplications
  – Climate Modeling

NLHPC: Training Workshop September 09, 2010

- 15:00-16:00 hrs.: Enabling Computational and Data Intensive Research at Brown University through Shared Resources: Challenges, Benefits, Experiences, and Prospects, Jan Hesthaven, Applied Mathematics, Brown University.
- 16:30-17:00 hrs.: Takeshi Asahi, Variational Methods for Digital Image Segmentation, Denoising and Classification, CMM U. Chile.
- 17:00-17:30 hrs.: Finding Optimal Sequences of Extraction in Mine Planning, Jorge Amaya, CMM U. Chile.
- 17:30-18:00 CCV-Brown: Computational Science using GPGPU's, Jan Hesthaven, Applied Mathematics, Brown University.

NLHPC: Training Workshop August 03, 2010

- 15:00 hrs. - 15:45 hrs.: Cluster Levque Architecture, Claudio Baeza, HPC Laboratory CMM.
- 15:45 hrs. - 16:30 hrs.: Parallel Programming Models, Gonzalo Hernandez, HPC Laboratory CMM.
- 17:00 hrs. - 18:00 hrs.: MPI Library, Gonzalo Hernandez, HPC Laboratory CMM.
- 18:00 hrs. - 18:45 hrs.: ScaLAPACK Library, Oscar Peredo, Doctoral Program University of Calaluña & Barcelona Supercomputing Center.

NLHPC: GPU Day January 06, 2010

- 10:00-13:00 hrs.: CPU/GPU Tutorial, Felipe Cruz, PhD (C), Dep. Mathematics, University of Bristol.
- 15:00 hrs.: TESLA Supercomputing, Arturo Allel, NVIDIA.
- 15:30 hrs.: INTEL Vision for HPC based in CPU/GPU, Agustin March, Lab. Manager at INTEL Software Argentina.
- 16:30 hrs.: Toward GPU-accelerated mesh free flow simulation, Lorena Barba, Mechanical Engineering Dep., Boston University.
- 17:30 - 18:30 hrs.: 42 TFlops Hierarchical N-body Simulations on GPUs with Applications in both Astrophysics and Turbulence, Rio Yokota, Department of Mathematics, University of Bristol.

EDUCATING THE NEXT GENERATION OF SCIENTISTS AND ENGINEERS IN HPC

- To consolidate undergraduate courses in HPC
- To offer advanced courses in HPC for Doctoral training
- To organize periodic seminars on HPC
- To organize the NLHPC Annual Symposium
- To organize Industrial Workshops to promote the adoption of HPC by the industry

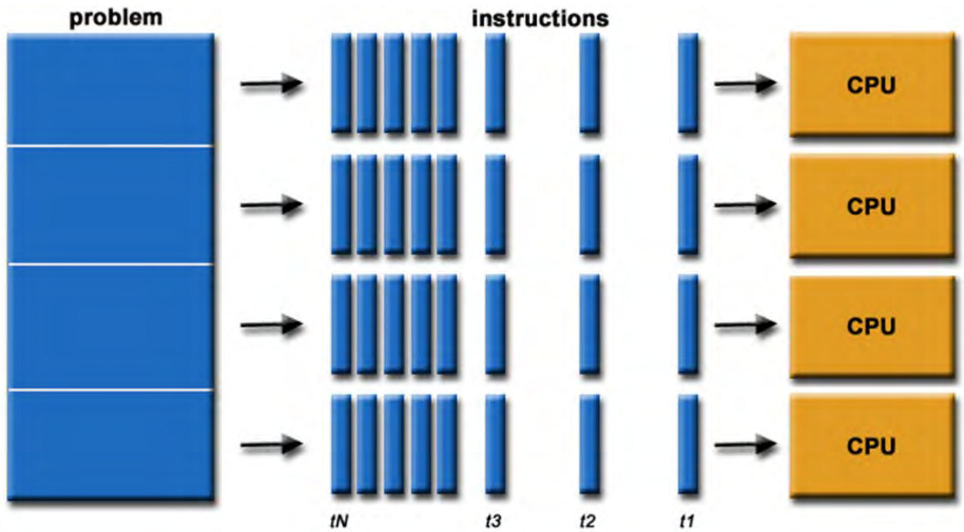
- Some training events on HPC:
  - GPU Day: University of Bristol, Boston University, Nvidia and Intel Argentina (01-2010)
  - Supercomputing techniques in Astrophysics: PUC (04 - 2010)
  - Workshop on HPC by Brown Supercomputer Center and CMM (09-2010)
  - Five CLGrid Workshops: Santiago (2), Antofagasta, Temuco, Valparaiso (2006 - 2008)
  - PASI School for Scientific Computing in the Americas: the challenge of massive parallelism: USM and CMM (01-2011)
  - HPC School Chillán: UBB and CMM (12 - 2010)

Outreach: Increasing the social awareness of the role of HPC in contemporary society and technological development

- Open Innovation program in HPC
- High-school guided visits
- HPC installation at MIM (Interactive Science Museum)
- Explora Projects
- HPC Cup
- General public lectures on the HPC impact in; Mining Industry, Astronomy, Biotechnology, Climate Modeling, Early Alerts in Critical Events

Why Use Parallel Computing ?

- Parallel computing is the simultaneous use of multiple computing resources (clusters) to solve a problem:
  - The applications will run using multiple CPUs
  - A problem is broken into discrete parts that can be solved concurrently
  - Each part is further broken down to a series of instructions
  - Instructions from each part execute simultaneously on different CPUs



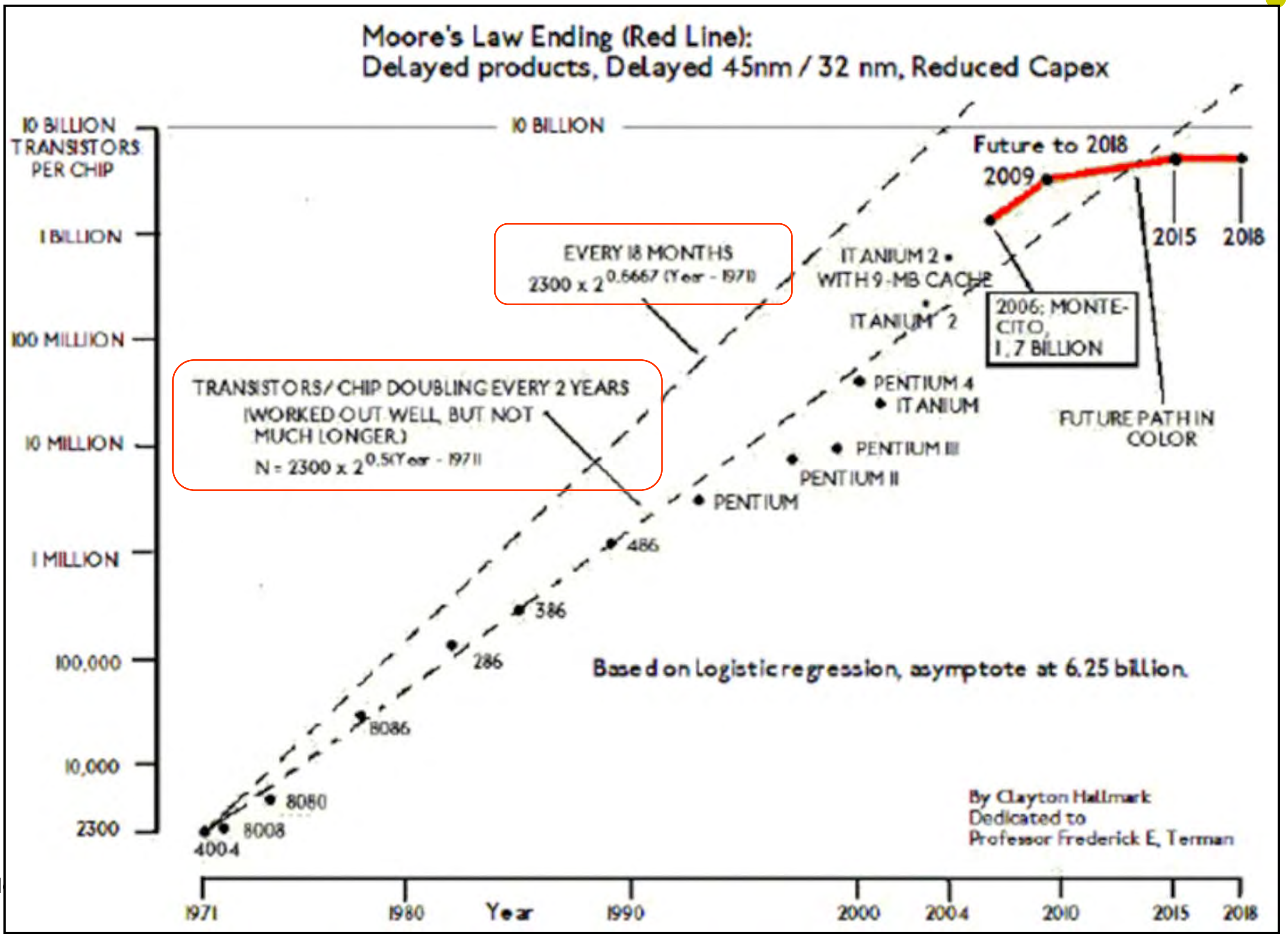
Why Use Parallel Computing ?

- Decrease computing time.
- Increase precision of computations.
- Solve larger and more complex problems by implementing concurrency (simultaneous computations).
- Taking advantage of non-local resources: Using available computing resources on a wide area network or Internet (Grid Computing).
- Cost savings: Using multiple computing resources instead of paying for time on a supercomputer.
- Overcoming memory constraints: Using the memories of multiple computers for large and complex problems.

Why Use Parallel Computing ?

- Limits to serial computing: Both physical and practical reasons constraint building faster serial computers:
  - Transmission speeds: The speed of a serial computer depends on how fast data can move through hardware.
  - Absolute limits are the speed of light 30 [cm/ns] and the transmission limit of copper wire 9 [cm/ns].
  - Increasing speeds require increasing proximity of processing elements.
  - Restrictions to miniaturization: Processor technology is allowing an increasing number of transistors to be placed on a chip. However, even with molecular or atomic-level components, a limit will be reached on how small components can be.
  - Economic limitations: It is increasingly expensive to make a single processor faster. Using a larger number of moderately fast commodity processors to achieve the same performance is less expensive.

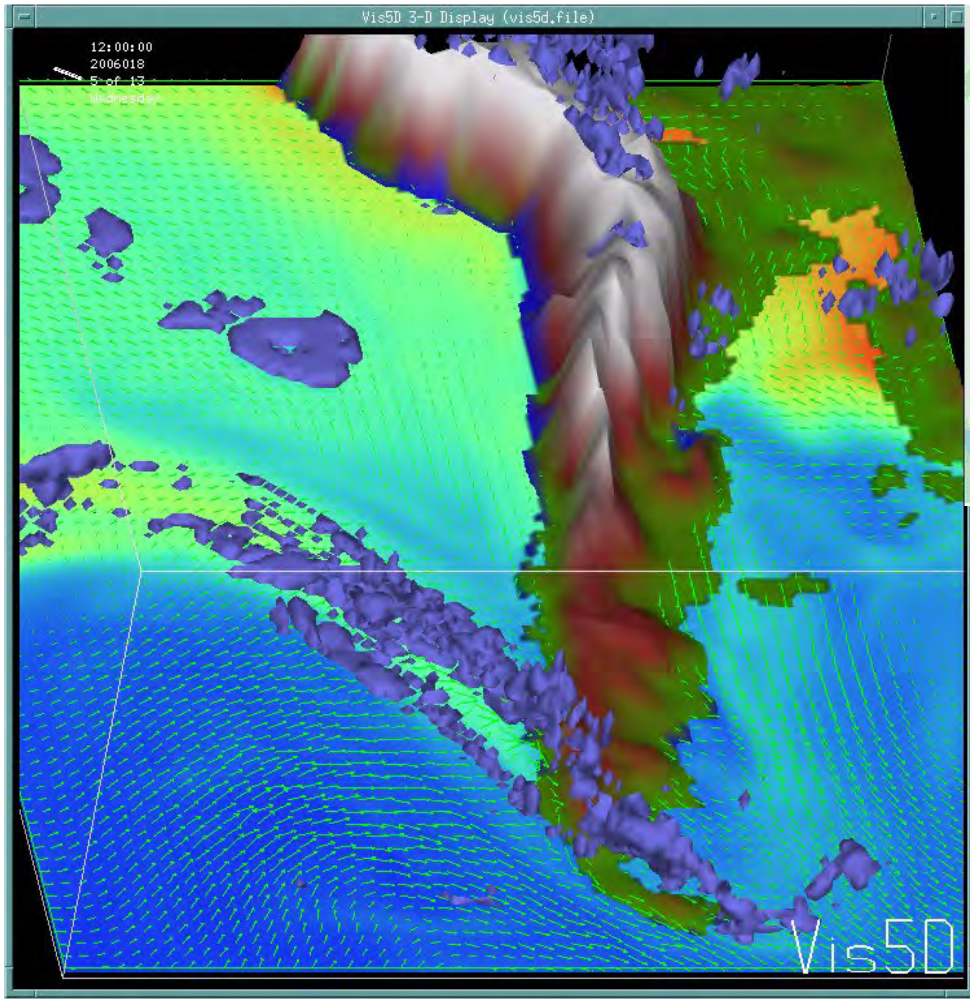
Why Use Parallel Computing ?

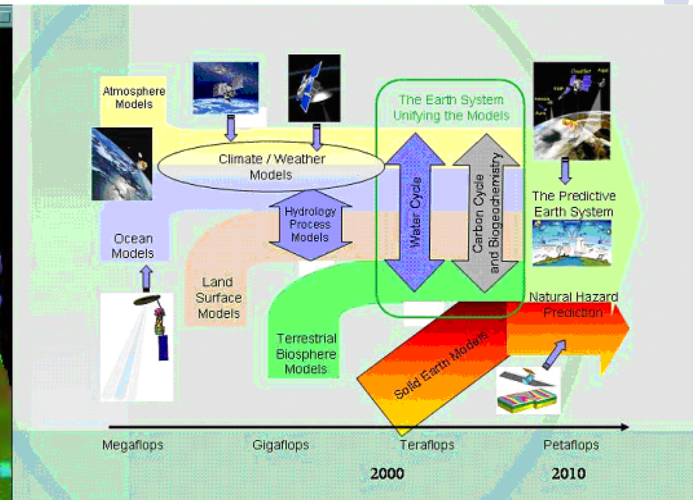


Why Use Parallel Computing ?

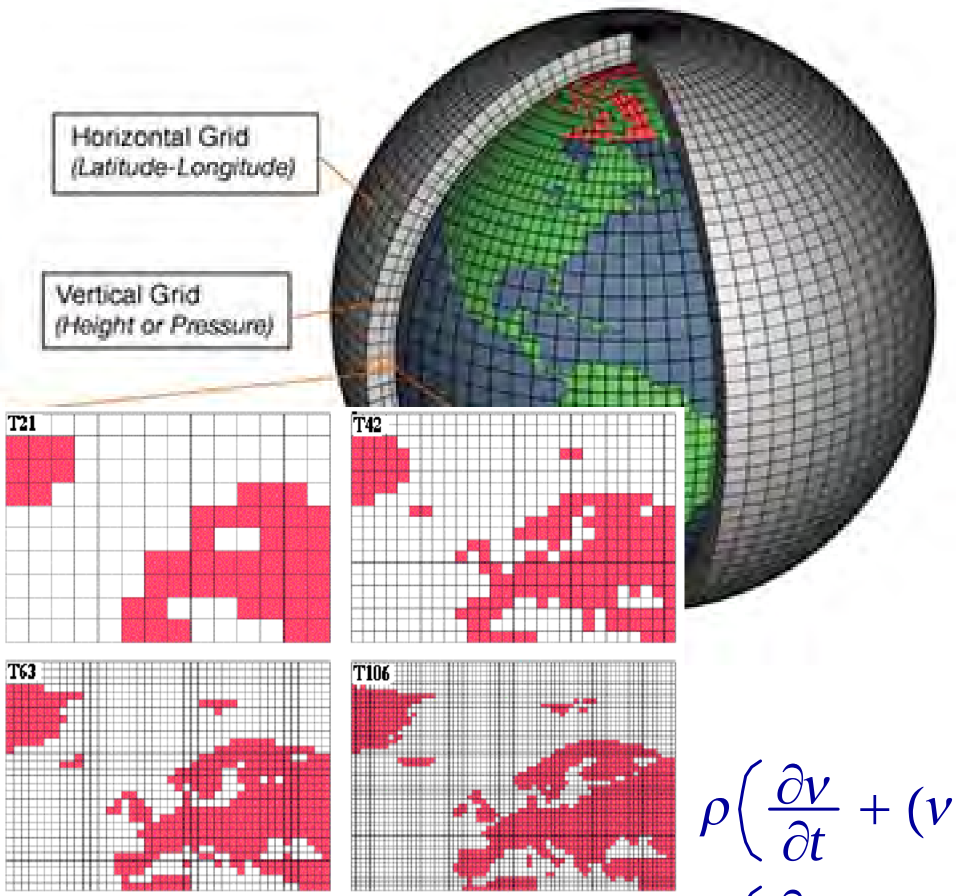
- During the past 10 years, the trends show that parallelism is the present and future of computing:
  - Faster networks with larger bandwidth and lower latency
  - Multi-processor computer architectures
  - Parallel and distributed systems
  - Different parallel programming models
  - Software for High performance computing: Compilers, Math and Simulation Libraries, Performance Evaluation.
  - Cheaper components

Some Computing Intensive Applications: Climate Modeling





Some Computing Intensive Applications: Climate Modeling



$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho v) = 0$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0$$

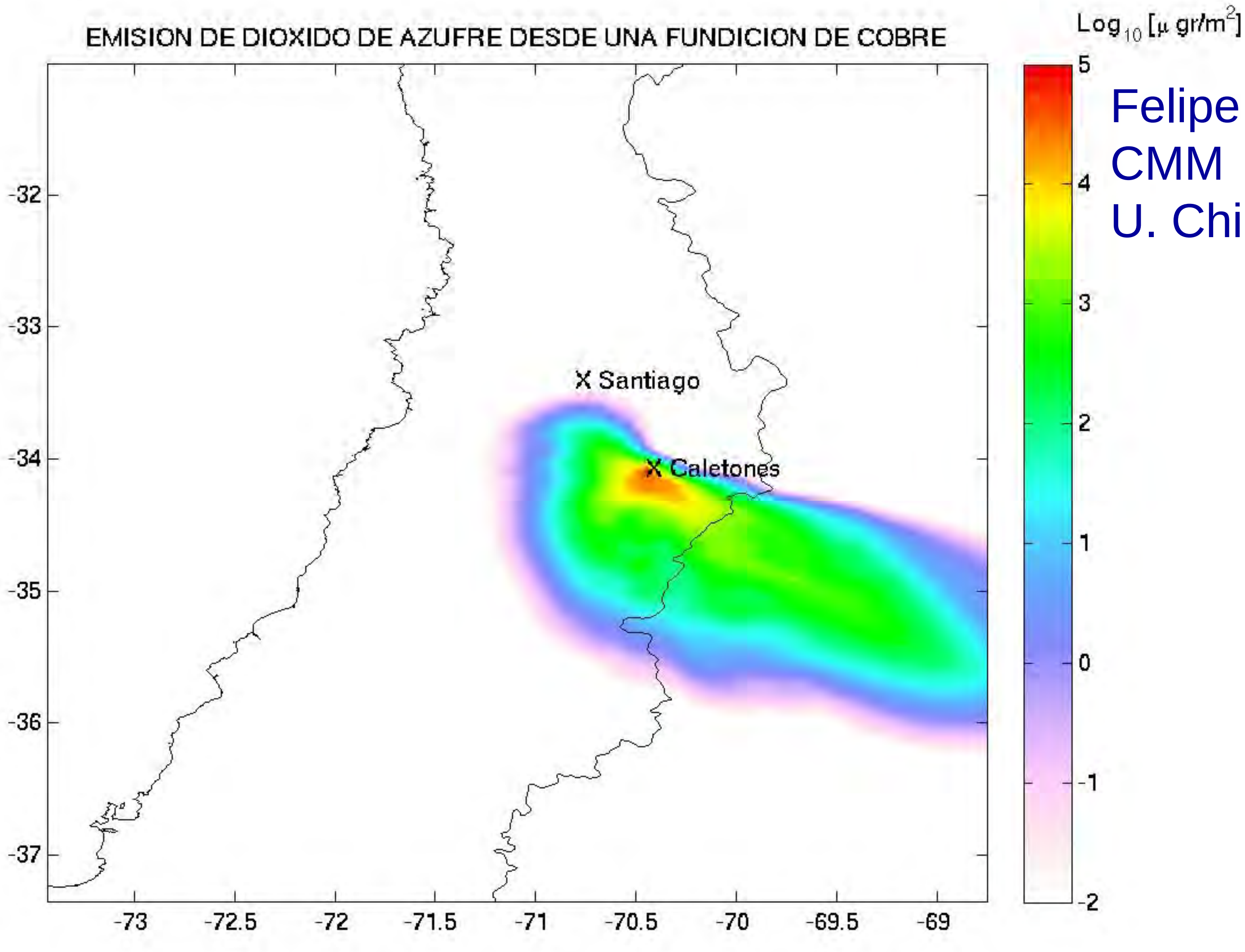
$$\frac{\partial \theta}{\partial t} + (v \bullet \nabla)\theta = \tau$$

$$\frac{\partial h}{\partial t} = -(v \bullet \nabla)h + E - P$$

$$\rho\left(\frac{\partial v}{\partial t} + (v \bullet \nabla)v\right) = F_g + F_p + F_f + F_C$$

$$\rho\left(\frac{\partial v}{\partial t} + (v \bullet \nabla)v\right) = -\rho g\hat{k} - \nabla p - 2\rho\Omega \times v + F_f$$

Some Computing Intensive Applications: Climate Modeling



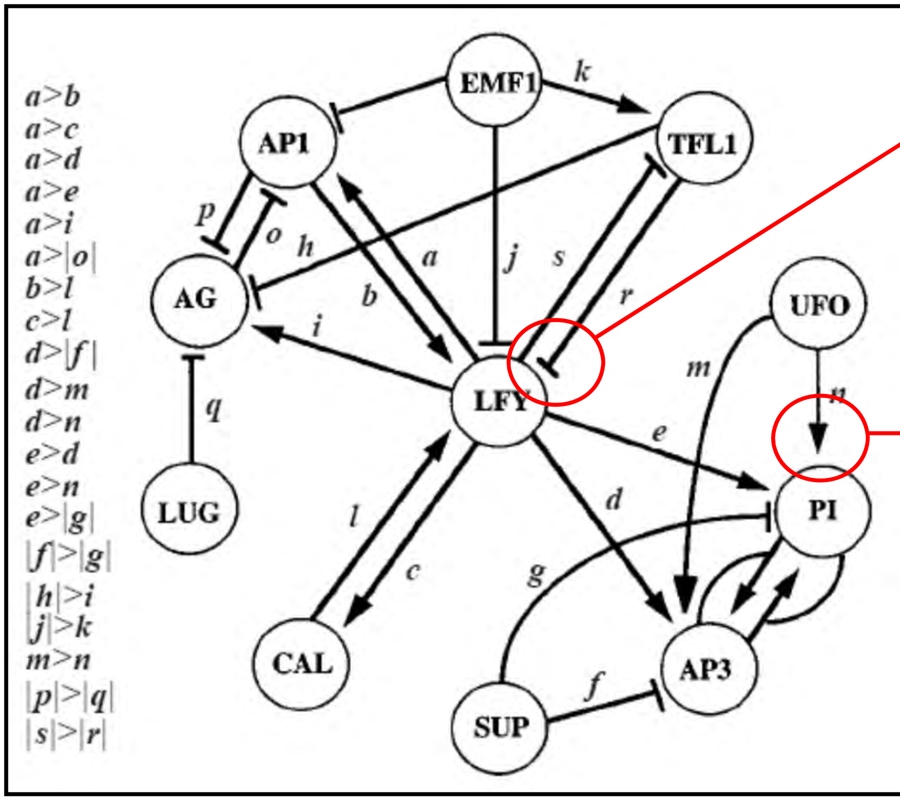
Felipe Macias
CMM
U. Chile

Some Computing Intensive Applications: Morphex: Boolean Networks in Biology

- The main objective of the project is to obtain via a modeling and simulation approach a better understanding of the structure and dynamics of metabolic pathways of gene regulatory networks involved in the morphogenesis of animals and plants.
- One of the plants is the Arabidopsis Thaliana.
- The genetic regulation network of the flower morphogenesis will be studied.
- Genes and proteins interactions will be modeled as a boolean network.

Some Computing Intensive Applications: Morphex: Boolean Networks in Biology

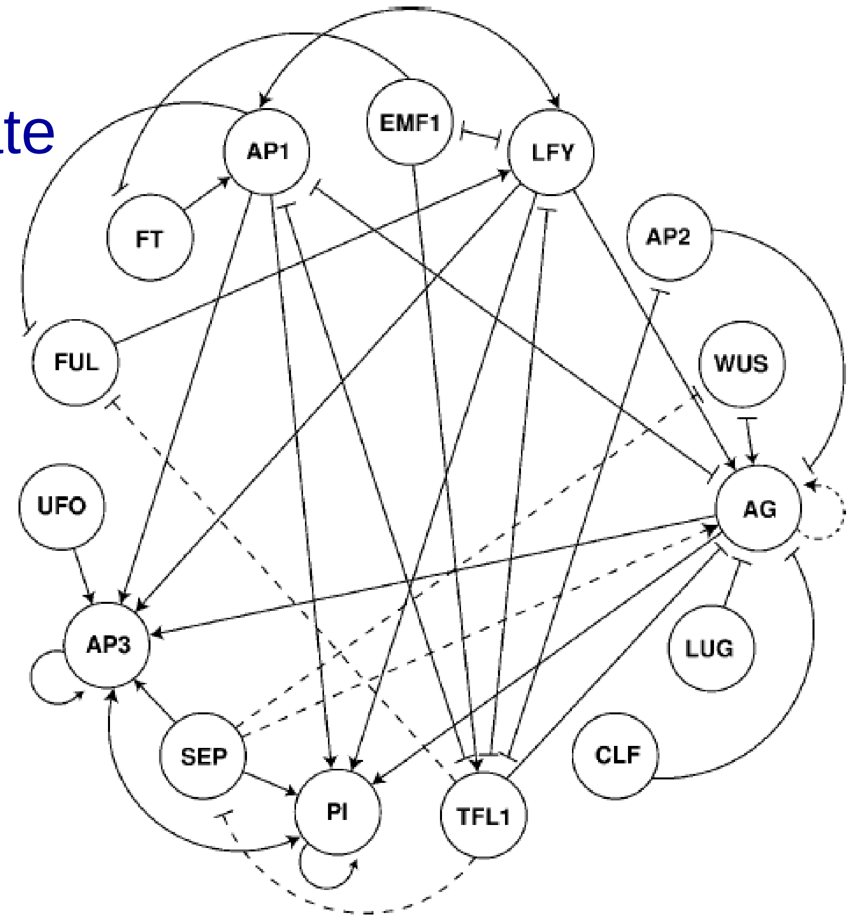
- First model:



Dynamics of the Genetic Regulatory Network for Arabidopsis Thaliana Flower Morphogenesis, L. Mendoza, E. Alvarez-Bulla, J. T. Biology, V. 193, pp. 307–319, 1998.

Some Computing Intensive Applications: Morphex: Boolean Networks in Biology

- Gene network architecture for the Arabidopsis floral organ fate determination:



A Gene Regulatory Network Model for Cell-Fate Determination during Arabidopsis thaliana Flower Development, The Plant Cell, Vol. 16, 2923–2939, 2004.

Some Computing Intensive Applications: Morphex: Boolean Networks in Biology

- The objective of this talk is to present the first version of the library: Boolean Networks Numerical Dynamics, that was designed to study computationally the dynamics of boolean functions:
  - Define a boolean function (threshold function):

$$f(x) = \left(f_i(x) \right)^n_{i=1} \Rightarrow f_i(x) = \begin{cases} 1 & \text{if } \sum_{j=1}^{n} w_{ij} x_j - \theta_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

  - Define different deterministic update schema:

$$\left.\begin{array}{llll} 1 \quad 1 \quad 1 \quad 1 & \longleftarrow & \text{Synchronous} \\ 1 \quad 2 \quad 3 \quad 4 & \longleftarrow & \text{Asynchronous} \\ 1 \quad 1 \quad 2 \quad 2 & \longleftarrow & \text{Two blocks: 1,2 ; 3,4} \\ 1 \quad 2 \quad 2 \quad 3 & \longleftarrow & \text{Three blocks: 1 ; 2,3 ; 4} \end{array}\right\} \text{Block Iteration Modes}$$

Some Computing Intensive Applications: Morphex: Boolean Networks in Biology

- Numerical Dynamics of Boolean Networks:
  - For a specific block iteration mode compute all the attractors over the hypercube
  - For all block iteration modes compute all the attractors
  - Visualizes the dynamics for an initial condition
- This library was programmed in Matlab® and utilizes some commands of the Random Boolean Network Toolbox:

  Christian Schwarzer

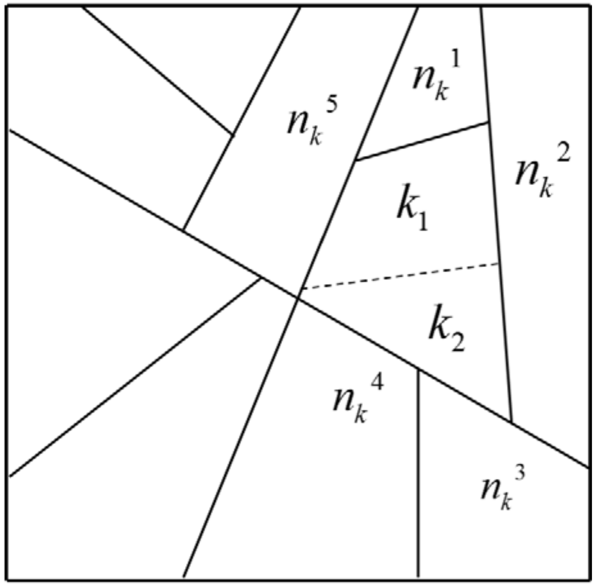
  Logic System Laboratory

  Swiss Federal Institute of Technology in Lausanne

Some Computing Intensive Applications: Fragmentation Models Neighborhood Interaction

Model assumptions:

1) Continuous material and point flaws.
2) Fracture forces generated by neighborhood interaction (not random!).
3) Multiple fragmentation:
   - Point flaws
   - Cut planes orthogonal to larger forces.
   - Self-similar process.
4) Mass conservation.
5) Stopping condition: minimal fragment size, stopping probability.

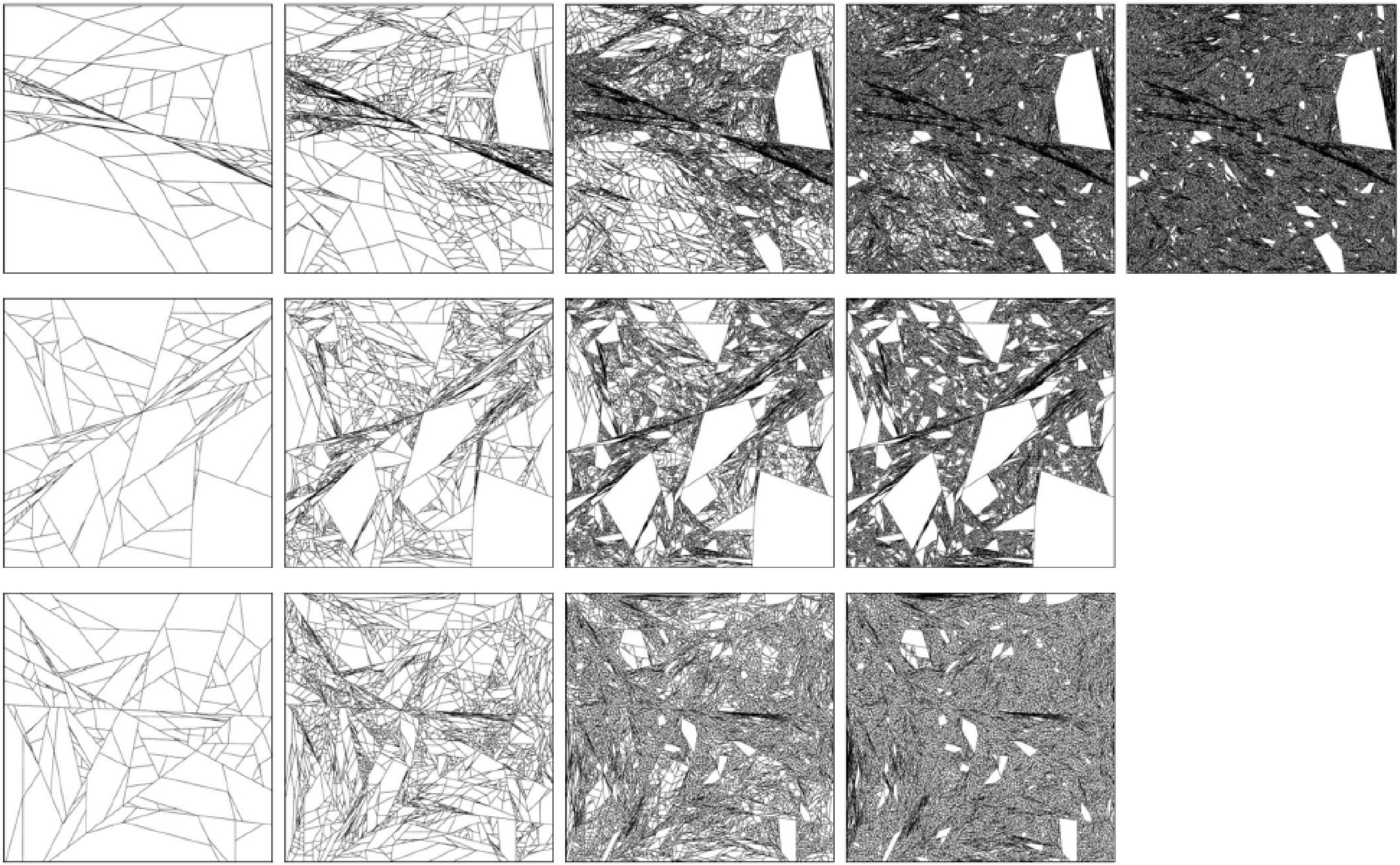


$$f_{i_k,k} \propto \int_{\partial B_{i_k,k}} \underline{\underline{\sigma}}\,\vec{n} = lenght(\partial B_{i_k,k})$$

Some Computing Intensive Applications: Fragmentation Models Neighborhood Interaction

Stage 0.	Initializations:	Define the value of the parameters (q, m_{fs}, p)
		Generate q random point flaws.
Stage 1.	Stopping Condition:	If (list of active fragments is empty) $\Rightarrow$ STOP
		For each fragment f such that $status(f) = active$
		If $(area(f) < m_{fs}$ or $random\_number \leq p)$
		$status(f) = not\ active$
		Else Goto Stage 2.
Stage 2.	Fragmentation Process:	Determine the neighbors of f: all the fragments that have a part of a boundary in common with f.
		Apply the fragmentation rule defined in assumption A3.
		Define for the new fragments g: $status(g) = active$
		Store the vertices of the new cuts in $lines\_mfs\_p$.
		Goto Stage 1.

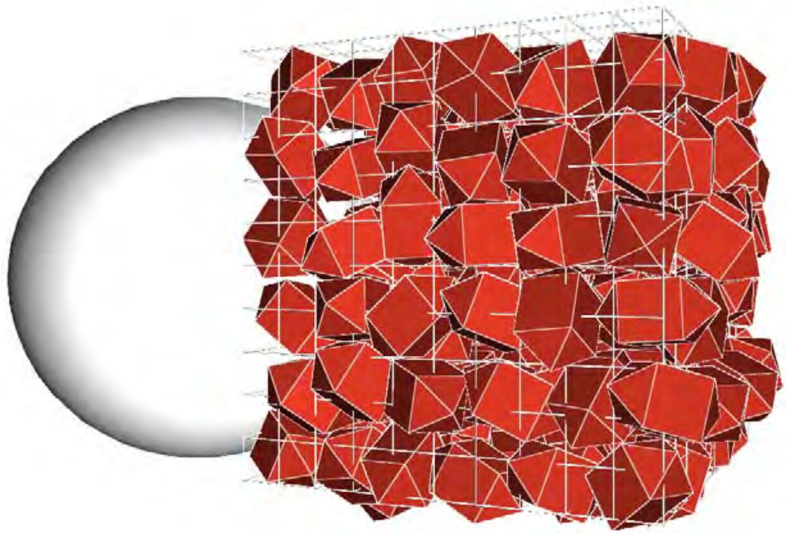
Some Computing Intensive Applications: Fragmentation Models Neighborhood Interaction

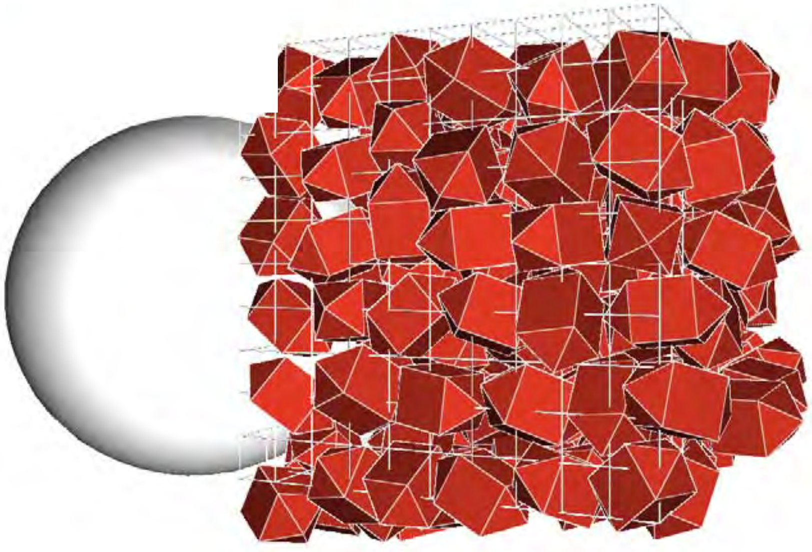


Some Computing Intensive Applications: Fragmentation Models Neighborhood Interaction

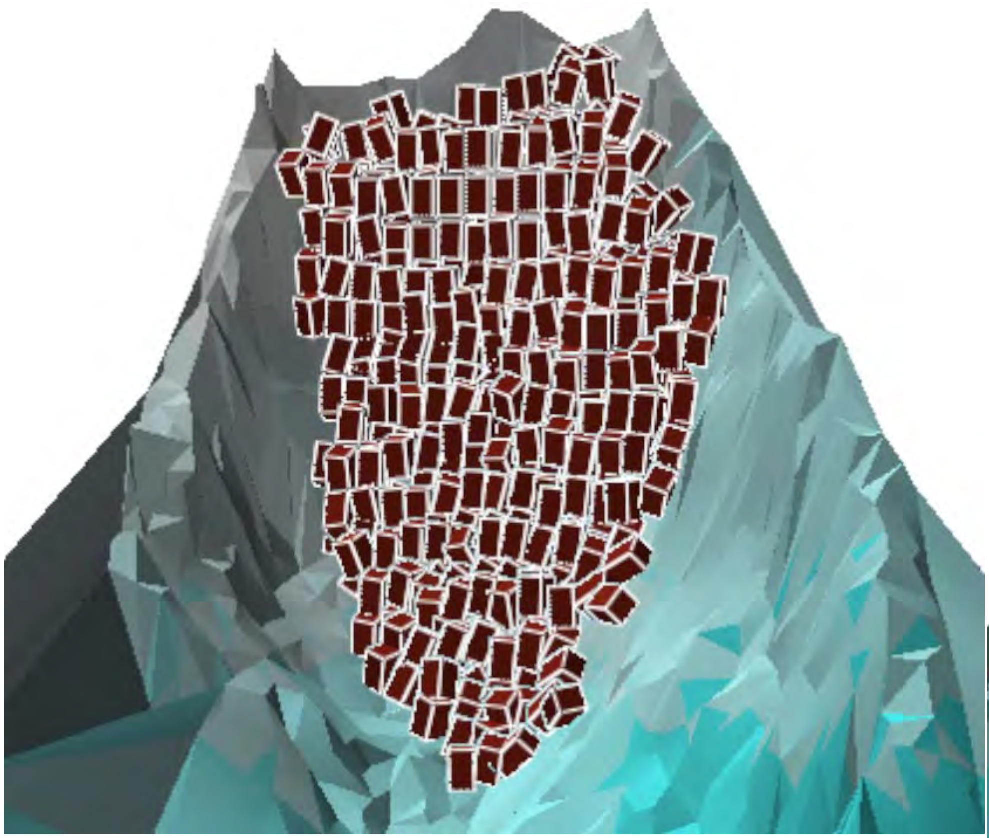
Some Computing Intensive Applications: Fragmentation Models Neighborhood Interaction

Some Computing Intensive Applications: Numerical Simulation Granular Flows (CCTVal+IFFSTAR)

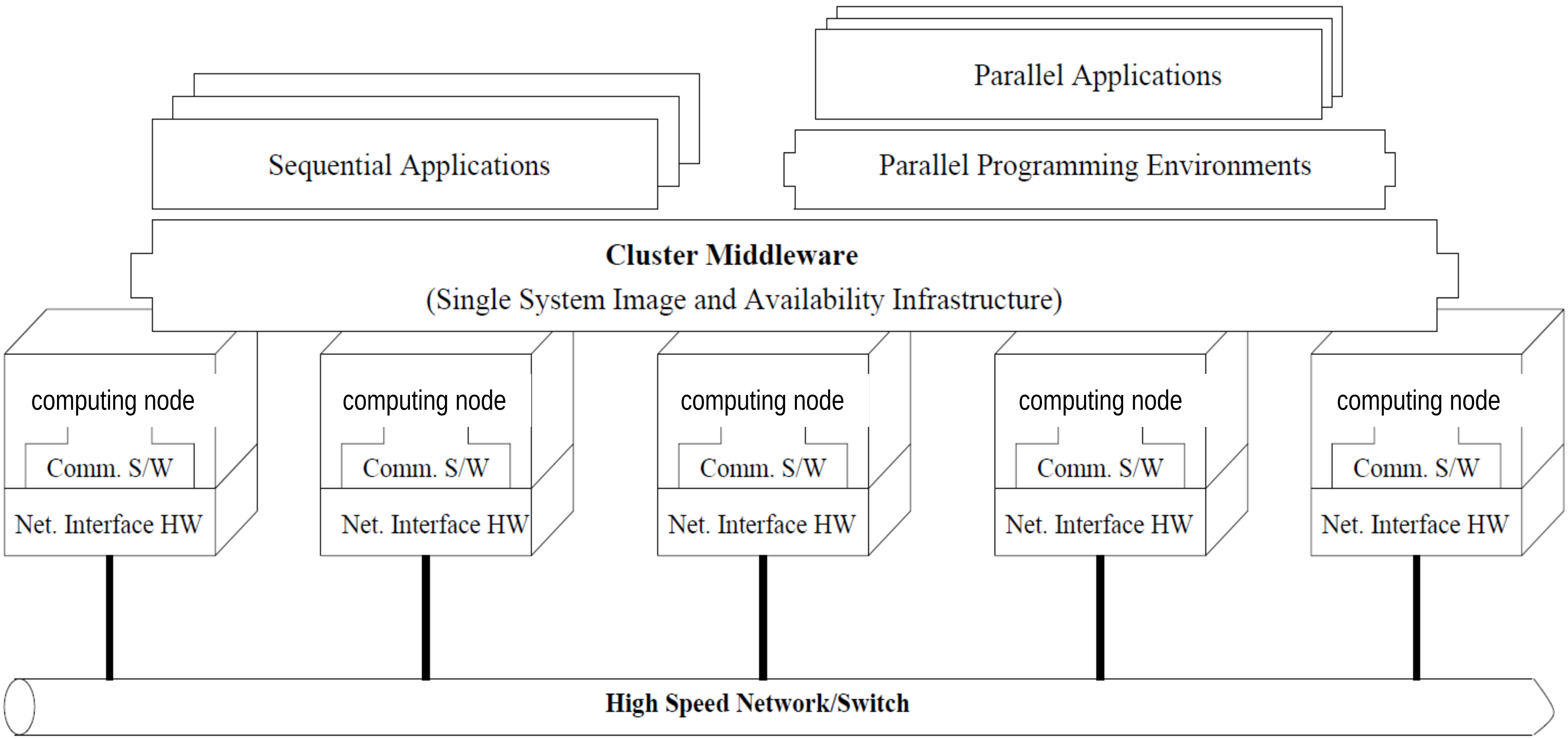




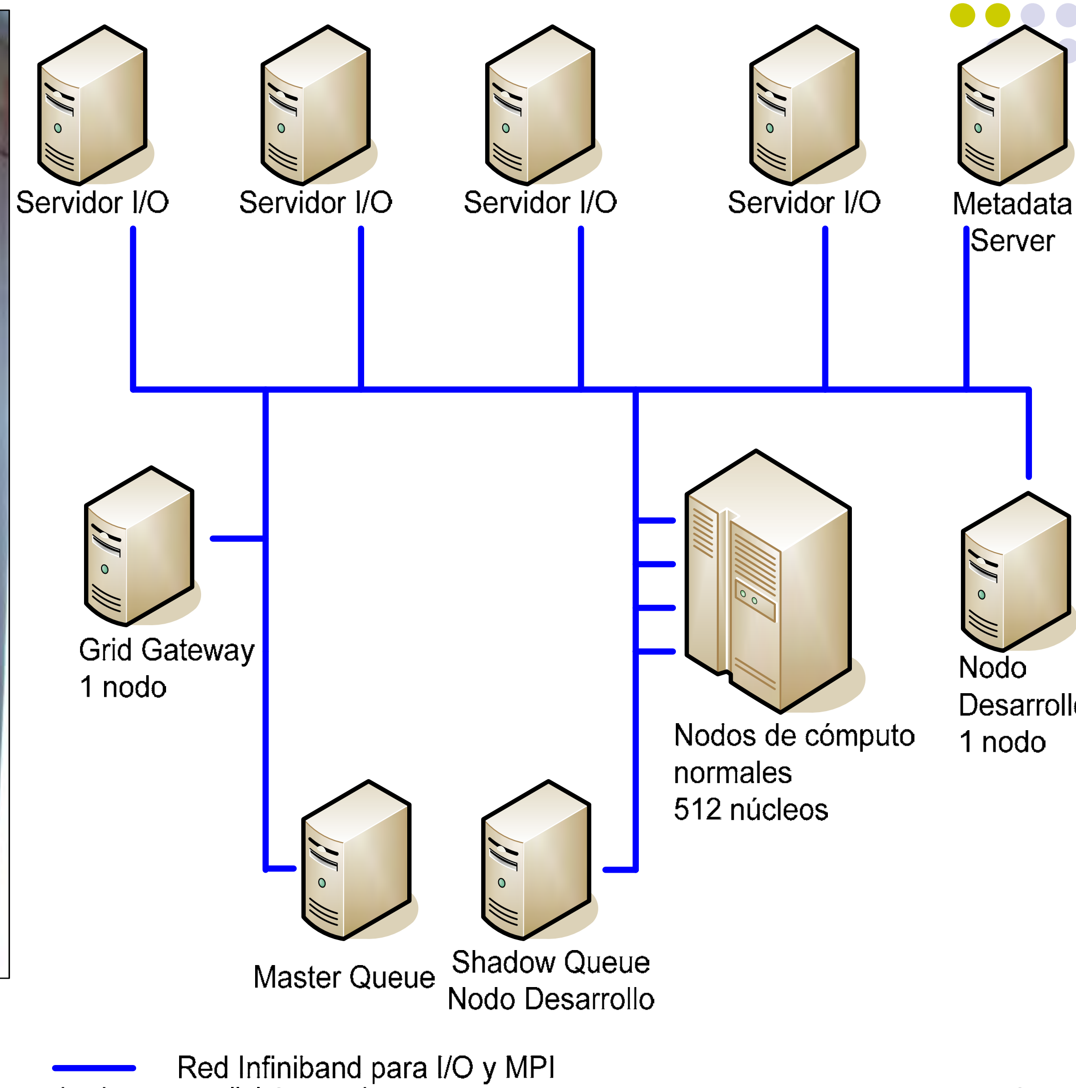
Some Computing Intensive Applications: Numerical Simulation Granular Flows (CCTVal+IFFSTAR)



Cluster Computing

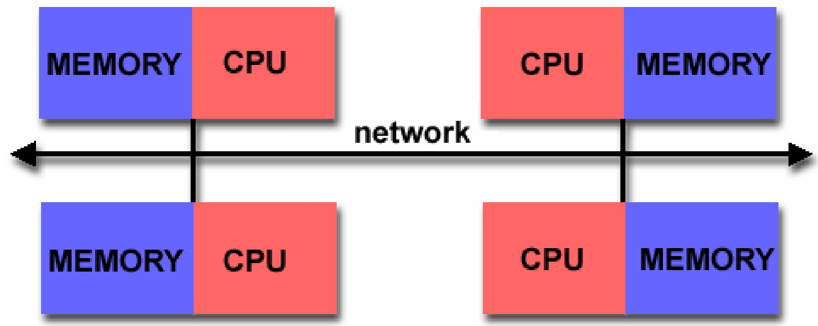


Cluster CMM: Levque



Cluster Computing: Distributed Memory

- Distributed memory systems require a communication network to connect inter-processor memory.



- Processors have their own local memory. Memory addresses in one processor do not map to another processor, so there is no concept of global address space across all processors.
- Each processor has its own local memory. Changes in its local memory have no effect on the memory of other processors.
- When a processor needs access to data in another processor, it is usually the task of the programmer to explicitly define how and when data is communicated and synchronized.

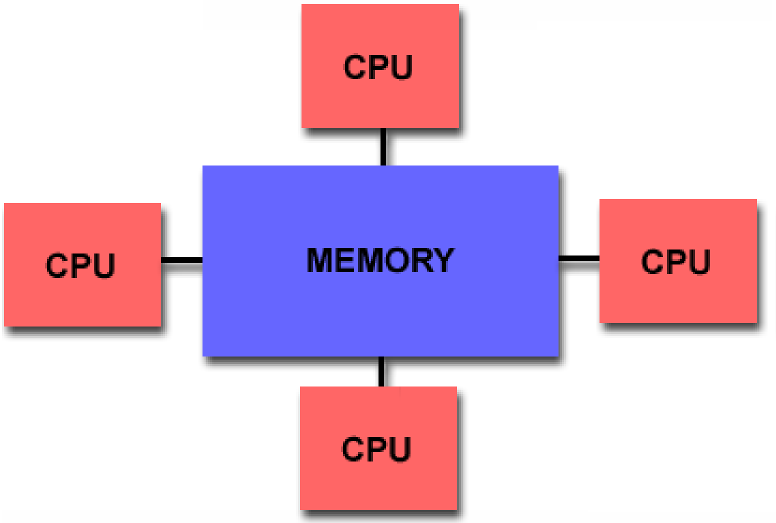
Cluster Computing: Distributed Memory

- The network "fabric" used for data transfer varies widely.

		Latencia [µs]	Bandwidth Teorico [Gbs]	Bandwidth Medido [Gbs]
10 Gb Ethernet		21.0	8.5	1.8 - 4.1
Myrinet GM	spi	7.0	1.98 - 3.9	0.78 - 1.93
	dpi	6.5	3.92 - 7.0	1.90 - 3.66
Dolphin SCI		1.4 - 14.4 (4 - 4096 bytes)	2.34	2.21 - 2.55
Quadrics		1.0 – 5.0	4.13 – 8.32	2.73 - 7.03
Infiniband		6.0 - 8.0	**2.0 – 20.0**	1.5 – 9.0
Infiniband **QDR**		**1.0 – 1.3**	**30.0 – 40.0**	**± 20.0**

Cluster Computing: Shared Memory

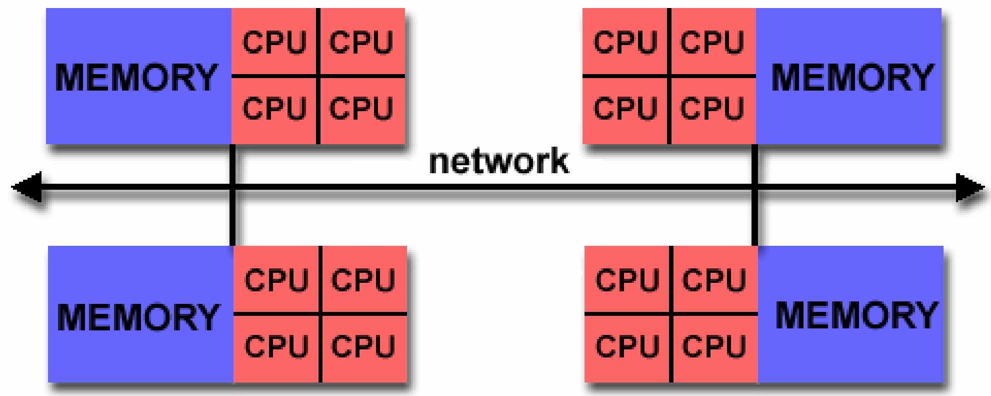
- Shared memory parallel computers vary widely, but generally have in common the ability for all processors to access all memory as global address space.
- Multiple processors can operate independently but share the same memory resources.



- Changes in a memory location generated by one processor are visible to all other processors.
- Shared memory machines can be divided into two main classes based upon memory access times: UMA, NUMA.

Cluster Computing: Hybrid Memory

- The largest and fastest computers in the world today employ both shared and distributed memory architectures.



- The shared memory component is usually a cache coherent SMP machine. Processors on a given SMP can address that machine's memory as global.
- SMPs know only about their own memory - not the memory on another SMP. Therefore, network communications are required to move data from one SMP to another.
- Current trends seem to indicate that this type of memory architecture will continue to prevail and increase at the high end of computing for the foreseeable future.

Cluster Computing: Hybrid Memory

- Advantages:
  - Memory is scalable with number of processors. Increase the number of processors and the size of memory increases proportionately.
  - Each processor can rapidly access its own memory without interference and without the overhead incurred with trying to maintain cache coherency.
  - Cost effectiveness: can use commodity, off-the-shelf processors and networking.
- Disadvantages:
  - The programmer is responsible for many of the details associated with data communication between processors.
  - It may be difficult to map existing data structures, based on global memory, to this memory organization.
  - Non-uniform memory access (NUMA) times.

Cluster Computing: Trends

- Traditionally, parallel computing has been considered to be "the high end of computing" and has been motivated by numerical simulations of complex systems and "Grand Challenge Problems" such as:
  - Weather and climate modeling
  - Bioinformatics
  - Geology: Seismic activity simulation, oil exploration
  - Engineering: transportation and telecommunication networks, mining operations, manufacturing processes, pattern recognition and image processing, mechanical devices, electronic circuits, etc.
- These problems are studied mainly in Universities and Research Centers.

Cluster Computing: Trends

- Today's commercial applications are providing an equal or greater driving force in the development of faster computers. These applications require the processing of large amounts of data in sophisticated ways. For example:
  - Parallel databases
  - Web based business services (business intelligence)
  - Oil exploration
  - Data mining
  - Computer-aided diagnosis in medicine
  - Advanced graphics and virtual reality (entertainment industry)
  - Networked video and multi-media technologies
  - Collaborative work environments

Cluster Computing: Trends

Segments	Count	Share %	Rmax Sum(Tflops)	Processor Sum
Industry	312	62.40	8986	1547276
Research	91	18.20	12480	2155447
Academic	79	15.80	5571	826584
Government	9	1.80	579	74744
Classified	5	1.00	140	22844
Vendor	4	0.80	222	37732
Totals	500	100	27978	4664627

Cluster Computing: Trends

Application Area	Count	Share %	Rmax Sum (Teraflops)	Processor Sum
Not Specified	117	23.40	9761	1510523
Research	78	15.60	85457	1576744
Finance	47	9.40	1268	247716
Information Service	34	6.80	994	161524
Geophysics	31	6.20	899	138888
Logistic Services	30	6.00	835	126068
Semiconductor	22	4.40	578	92416
Information Processing Service	18	3.60	630	126280
Defense	16	3.20	874	109380
Software	16	3.20	433	83192
Telecommunications	11	2.20	278	51784
Service	10	2.00	242	41540
Database	9	1.80	227	49336
Aerospace	7	1.40	369	38864
Energy	7	1.40	236	33712
WWW	7	1.40	228	48896
Weather and Climate Research	6	1.20	386	30016

Cluster Computing: Trends

- The sum of the performance of the world's 500 most powerful supercomputers has increased exponentially from 1.12 to 27951.30 teraflops between 1993 – 2009. If this tendency continues in 2019 the total performance will reach approximately 10^7 teraflops.
- Since 1993, the number of supercomputers dedicated to industrial areas of application has experienced sustained growth.
- Since 2004, the architecture of cluster with multi-core processors has been the most frequently architecture utilized in the group of top 500 supercomputers.

Cluster Computing: Trends

Continents	Count	Share %	Rmax Sum (Teraflops)	Processor Sum
Americas	287	57.40	16920488	2883843
Europe	152	30.40	7468476	1242944
Asia	51	10.20	3307969	490432
Oceania	9	1.80	255128	44784
Africa	1	0.20	25440	2624
Totals	500	100	27977501	4664627

Countries	Count	Share %
United States	277	55.40
United Kingdom	45	9.00
Germany	27	5.40
France	26	5.20
China	21	4.20
Japan	16	3.20
Canada	9	1.80
Austria	8	1.60
New Zealand	8	1.60
Russia	8	1.60

Cluster Computing: Flynt Taxonomy

- Architecture:
  - SIMD
  - MIMD
- Memory:
  - Distributed
  - Shared,
  - Hybrid
  - CPU/GPU
- Programming:
  - Data Parallel
  - Control Flow:

    SPMD, MPMP

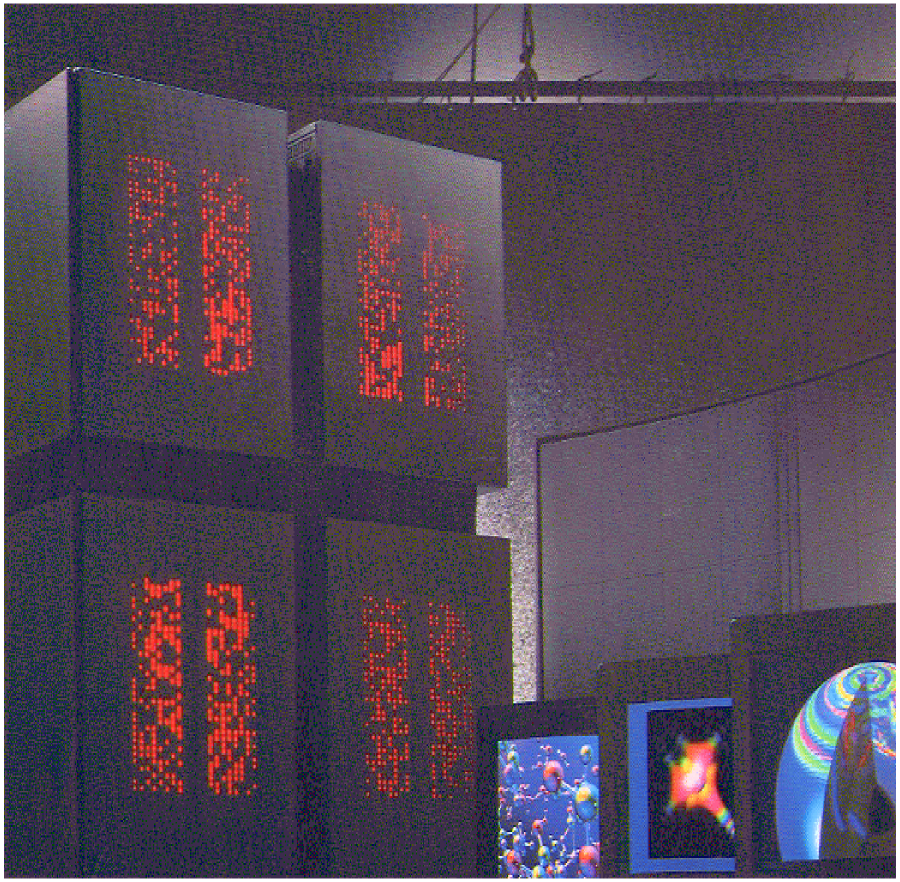
Some Old Supercomputers

Connection Machine CM 2 (1992)

- Architecture: MESH
- Performance: 30 GFlops
- Applications:

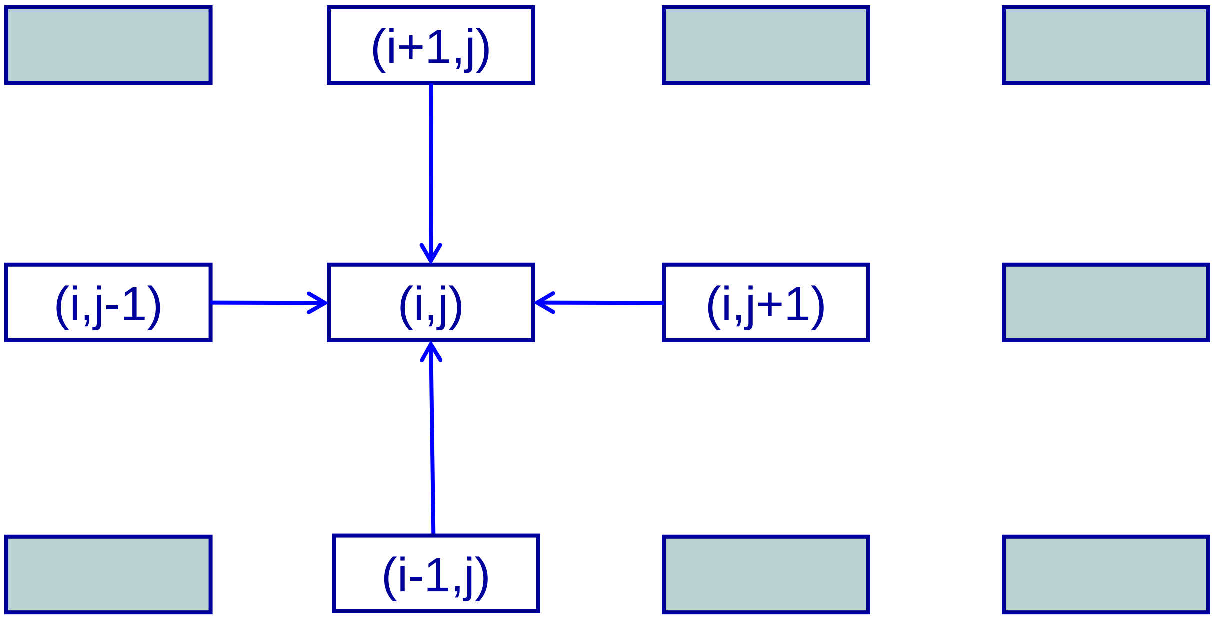
  Image Processing

  Cellular Automata



Some Old Supercomputers

- Connection Machine CM 2 (1992)



Some Old Supercomputers

- Connection Machine CM 2 (1992): NER Program

```
criteria = 0;
while ( criteria == 0 )
        with (lattice)
                u0 = from_grid_dim(&v0,v0,1,1);
                u1 = from_grid_dim(&v0,v0,1,-1);
                u2 = from_grid_dim(&v0,v0,0,1);
                u3 = from_grid_dim(&v0,v0,0,-1);
                M = MAX(MAX(u0,u1),MAX(u2,u3)); M = MAX(M,v0);
                m = MIN(MIN(u0,u1),MIN(u2,u3)); m = MIN(m,v0);
                condition = (M - v0) <= (v0 - m);
                where (condition) v1 = M;
                else v1 = m;
                if ( HAMMING(v0,v1) == 0 ) criteria = 1;
                else v0 = v1;
```

Some Old Supercomputers

Connection Machine CM 5 (1995)

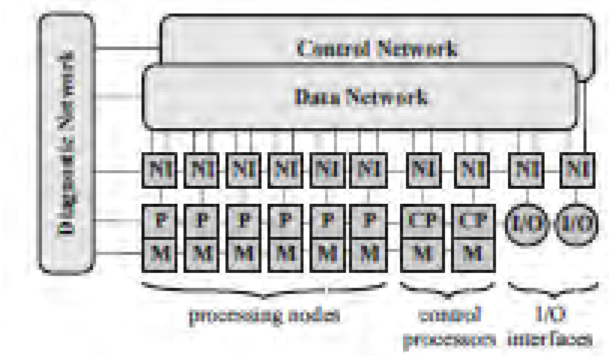
- Architecture: Fat Tree
- Performance: 80 Gflops
- Applications:

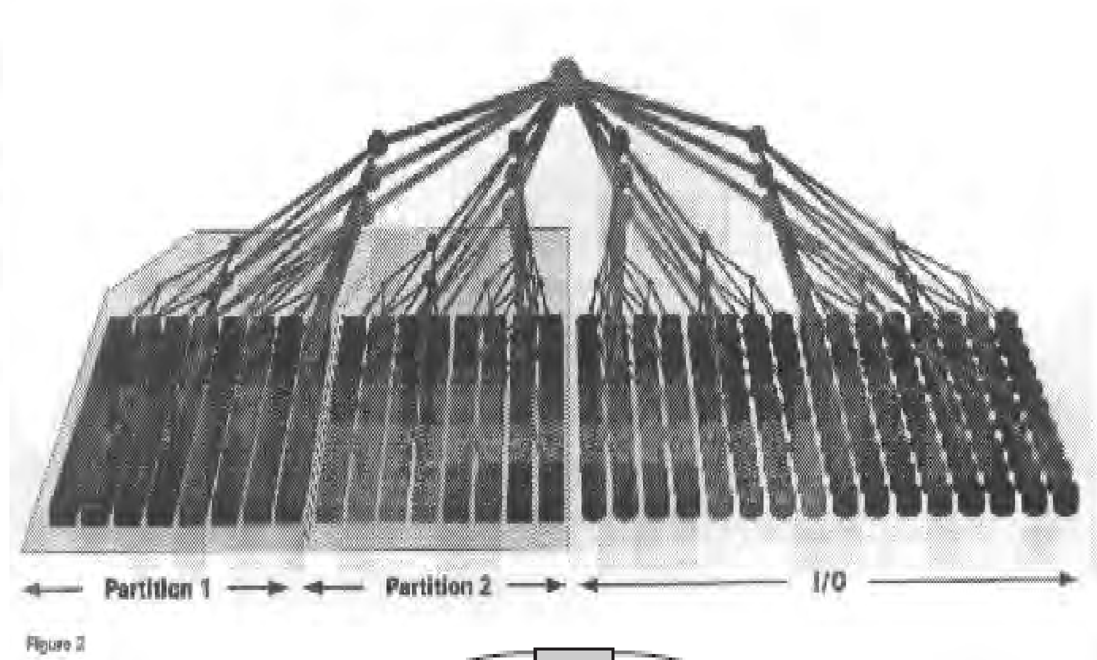
  Climate Modeling

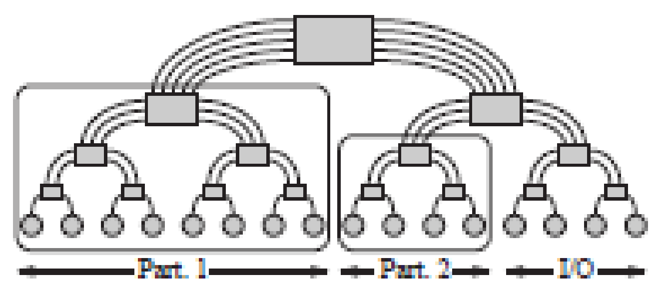
  Physics Models

Some Old Supercomputers

Connection Machine CM 5 (1995):







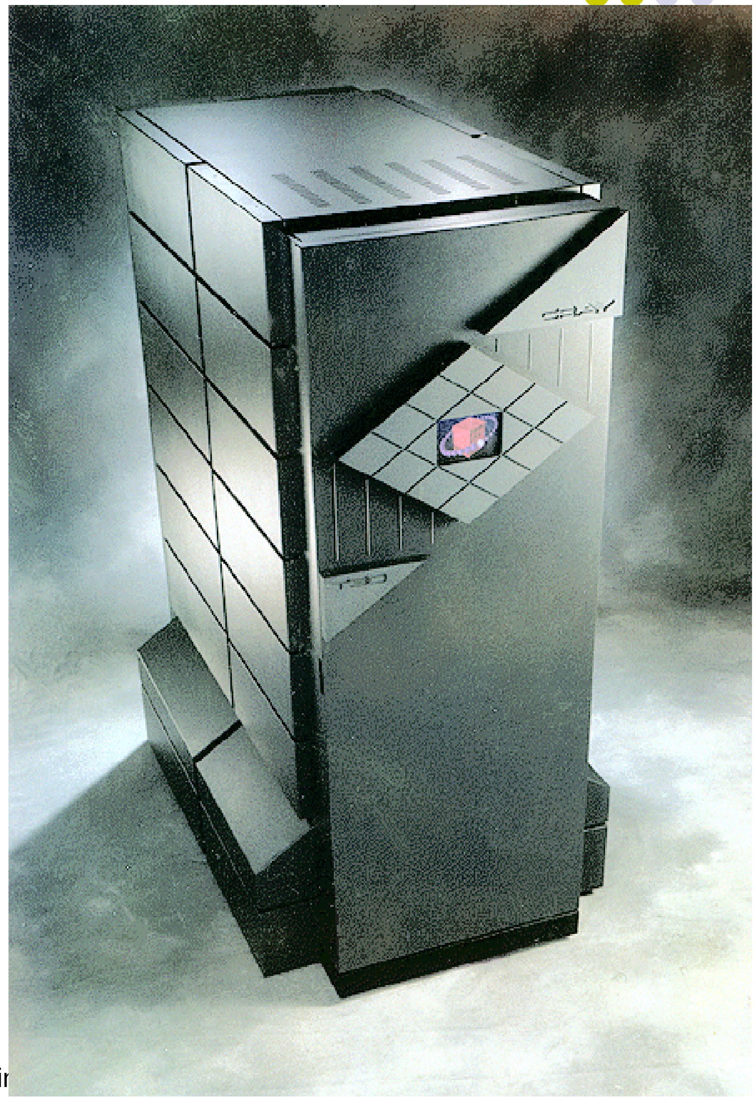
Some Old Supercomputers

Cray T3D (1998):

- Architecture: 3d Torus
- Performance: 160 GFlops
- Applications:

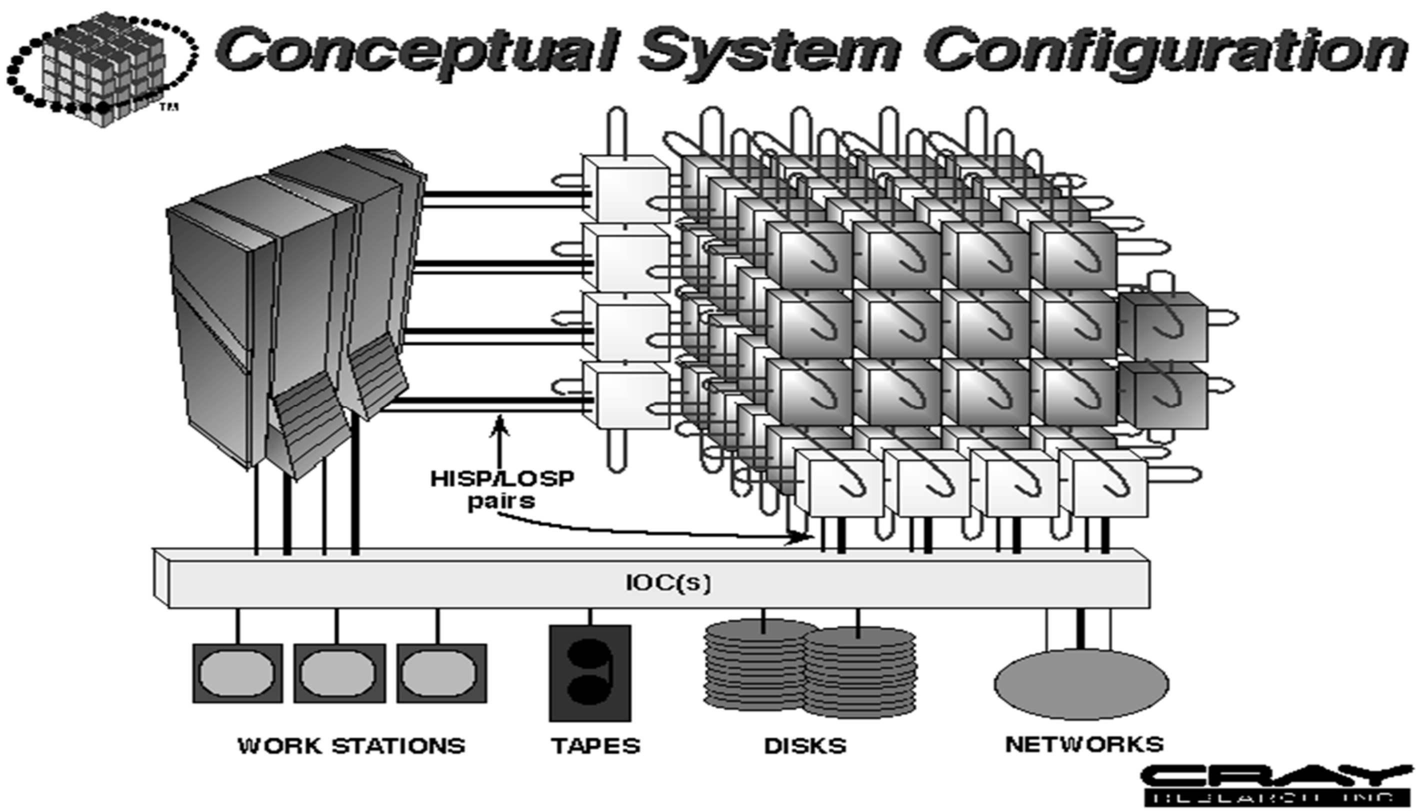
  Climate Modeling

  Physics Models



Some Old Supercomputers

- Cray T3D (1998):



Supercomputers Evolution

- 1992 CM2: 30 GFlops, 1995 CM 5: 100 GFlops
- 1997 Cray T3D: 160 GFlops, 1999 Cray T3E: 1 Tflops
- 2004-2 IBM Blue Gene: 70 TFlops
- 2005 IBM BlueGene/L (DOE/NNSA/LLNL): 137 Tflops,

  PowerPC 440 MHz, 131072 Procesadores, 32768 GB Ram
- 2006 IBM BlueGene/L (DOE/NNSA/LLNL): 280 Tflops,

  PowerPC 440 MHz, 131072 Procesadores, 32768 GB Ram
- 2007 IBM Blue Gene/L (DOE/NNSA/LLNL): 478 Tflops,

  PowerPC 700 MHz, 212992 Cores, 73728 GB RAM

Supercomputers Evolution

IBM Blue Gene/L 2007:

- (DOE/NNSA/LLNL)
- 106.496 Nodos
- 212.992 Cores
- 73728 GB RAM
- Storage 1.89 Peta Bytes
- Procesador PowerPC
- 440 700 MHz 2.8 Gflops
- Peak performance 478 Tflops

Top 500 2008

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband / 2008 IBM	129600	1105.00	1456.70	2483.47
2	Oak Ridge National Laboratory United States	Jaguar - Cray XT5 QC 2.3 GHz / 2008 Cray Inc.	150152	1059.00	1381.40	6950.60
3	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0/2.66 GHz / 2008 SGI	51200	487.01	608.83	2090.00
4	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
5	Argonne National Laboratory United States	Blue Gene/P Solution / 2007 IBM	163840	450.30	557.06	1260.00
6	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband / 2008 Sun Microsystems	62976	433.20	579.38	2000.00
7	NERSC/LBNL United States	Franklin - Cray XT4 QuadCore 2.3 GHz / 2008 Cray Inc.	38642	266.30	355.51	1150.00
8	Oak Ridge National Laboratory United States	Jaguar - Cray XT4 QuadCore 2.1 GHz / 2008 Cray Inc.	30976	205.00	260.20	1580.71
9	NNSA/Sandia National Laboratories United States	Red Storm - Sandia/ Cray Red Storm, XT3/4, 2.4/2.2 GHz dual/quad core / 2008 Cray Inc.	38208	204.20	284.00	2506.00
10	Shanghai Supercomputer Center China	Dawning 5000A - Dawning 5000A, QC Opteron 1.9 Ghz, Infiniband, Windows HPC 2008 / 2008 Dawning	30720	180.60	233.47	

Top 500 2009

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	Oak Ridge National Laboratory United States	Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.60
2	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband / 2009 IBM	122400	1042.00	1375.78	2345.50
3	National Institute for Computational Sciences/University of Tennessee United States	Kraken XT5 - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	98928	831.70	1028.85	
4	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution / 2009 IBM	294912	825.50	1002.70	2268.00
5	National SuperComputer Center in Tianjin/NUDT China	Tianhe-1 - NUDT TH-1 Cluster, Xeon E5540/E5450, ATI Radeon HD 4870 2, Infiniband / 2009 NUDT	71680	563.10	1206.19	
6	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0 GHz/Nehalem EP 2.93 Ghz / 2009 SGI	56320	544.30	673.26	2348.00
7	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
8	Argonne National Laboratory United States	Blue Gene/P Solution / 2007 IBM	163840	458.61	557.06	1260.00
9	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband / 2008 Sun Microsystems	62976	433.20	579.38	2000.00
10	Sandia National Laboratories / National Renewable Energy Laboratory United States	Red Sky - Sun Blade x6275, Xeon X55xx 2.93 Ghz, Infiniband / 2009 Sun Microsystems	41616	423.90	487.74	

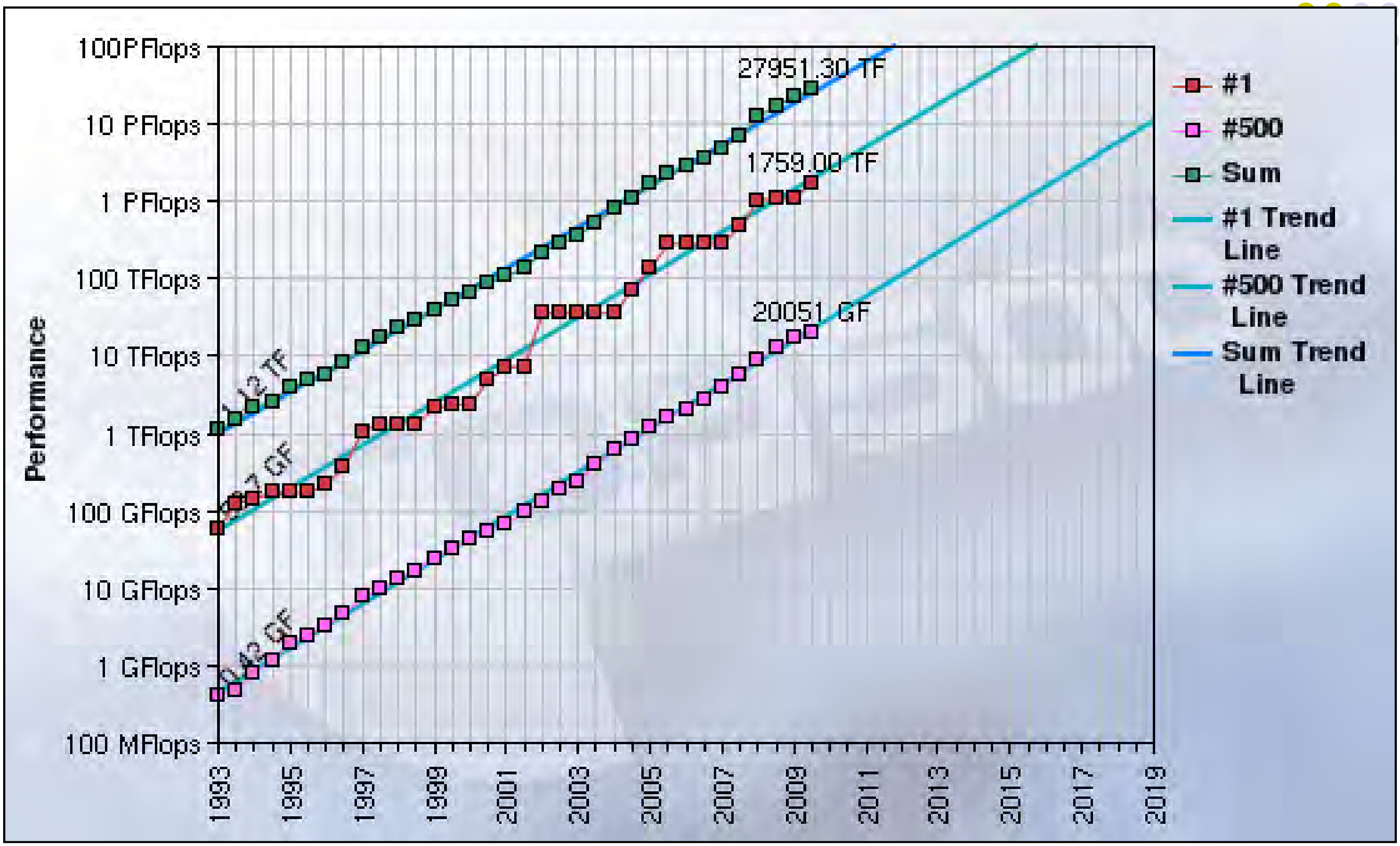
Top 500 2010

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	Oak Ridge National Laboratory United States	Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.60
2	National Supercomputing Centre in Shenzhen (NSCS) China	Nebulae - Dawning TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU / 2010 Dawning	120640	1271.00	2984.30	
3	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband / 2009 IBM	122400	1042.00	1375.78	2345.50
4	National Institute for Computational Sciences/University of Tennessee United States	Kraken XT5 - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	98928	831.70	1028.85	
5	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution / 2009 IBM	294912	825.50	1002.70	2268.00
6	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX/8400EX, Xeon HT QC 3.0/Xeon Westmere 2.93 Ghz, Infiniband / 2010 SGI	81920	772.70	973.29	3096.00
7	National SuperComputer Center in Tianjin/NUDT China	Tianhe-1 - NUDT TH-1 Cluster, Xeon E5540/E5450, ATI Radeon HD 4870 2, Infiniband / 2009 NUDT	71680	563.10	1206.19	
8	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
9	Argonne National Laboratory United States	Intrepid - Blue Gene/P Solution / 2007 IBM	163840	458.61	557.06	1260.00
10	Sandia National Laboratories / National Renewable Energy Laboratory United States	Red Sky - Sun Blade x6275, Xeon X55xx 2.93 Ghz, Infiniband / 2010 Sun Microsystems	42440	433.50	497.40	

Top 500 2011

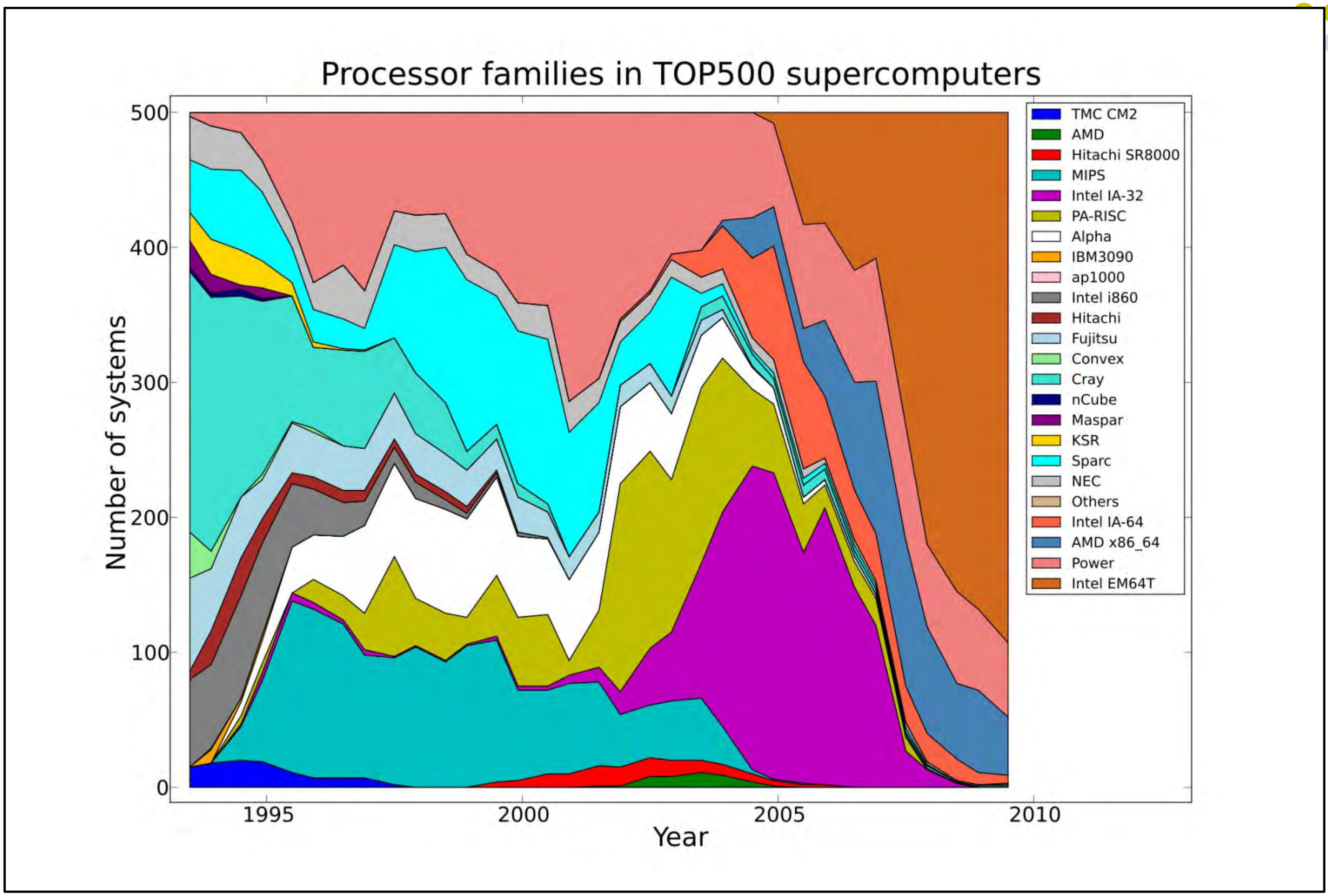
Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect / 2011 Fujitsu	705024	10510.00	11280.38	12659.9
2	National Supercomputing Center in Tianjin China	NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050 / 2010 NUDT	186368	2566.00	4701.00	4040.0
3	DOE/SC/Oak Ridge National Laboratory United States	Cray XT5-HE Opteron 6-core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.0
4	National Supercomputing Centre in Shenzhen (NSCS) China	Dawning TC3600 Blade System, Xeon X5650 6C 2.66GHz, Infiniband QDR, NVIDIA 2050 / 2010 Dawning	120640	1271.00	2984.30	2580.0
5	GSIC Center, Tokyo Institute of Technology Japan	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows / 2010 NEC/HP	73278	1192.00	2287.63	1398.6
6	DOE/NNSA/LANL/SNL United States	Cray XE6, Opteron 6136 8C 2.40GHz, Custom / 2011 Cray Inc.	142272	1110.00	1365.81	3980.0
7	NASA/Ames Research Center/NAS United States	SGI Altix ICE 8200EX/8400EX, Xeon HT QC 3.0/Xeon 5570/5670 2.93 Ghz, Infiniband / 2011 SGI	111104	1088.00	1315.33	4102.0
8	DOE/SC/LBNL/NERSC United States	Cray XE6, Opteron 6172 12C 2.10GHz, Custom / 2010 Cray Inc.	153408	1054.00	1288.63	2910.0
9	Commissariat a l'Energie Atomique (CEA) France	Bull bullx super-node S6010/S6030 / 2010 Bull	138368	1050.00	1254.55	4590.0
10	DOE/NNSA/LANL United States	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband / 2009 IBM	122400	1042.00	1375.78	2345.0

Supercomputers Evolution

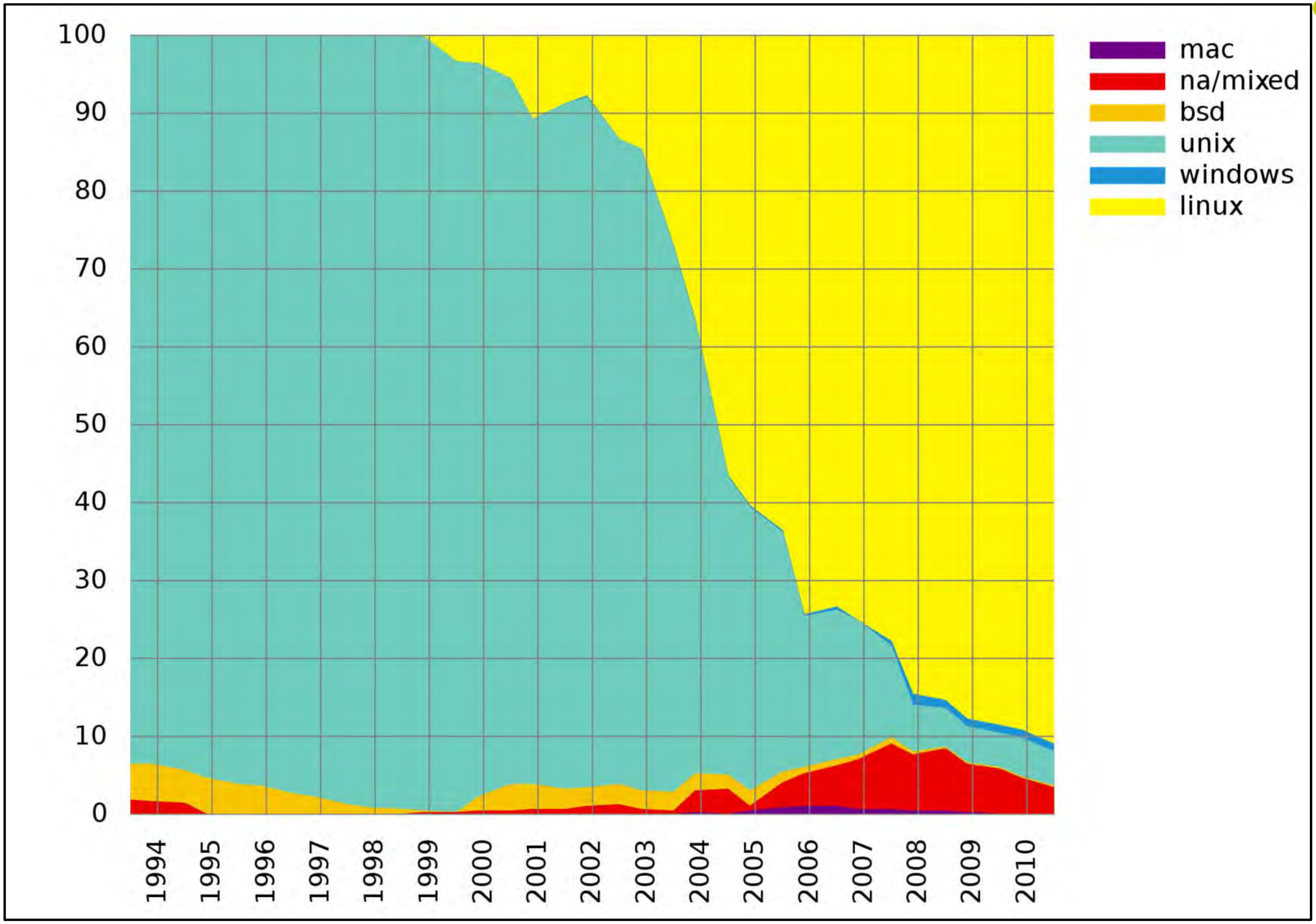


Supercomputers Evolution

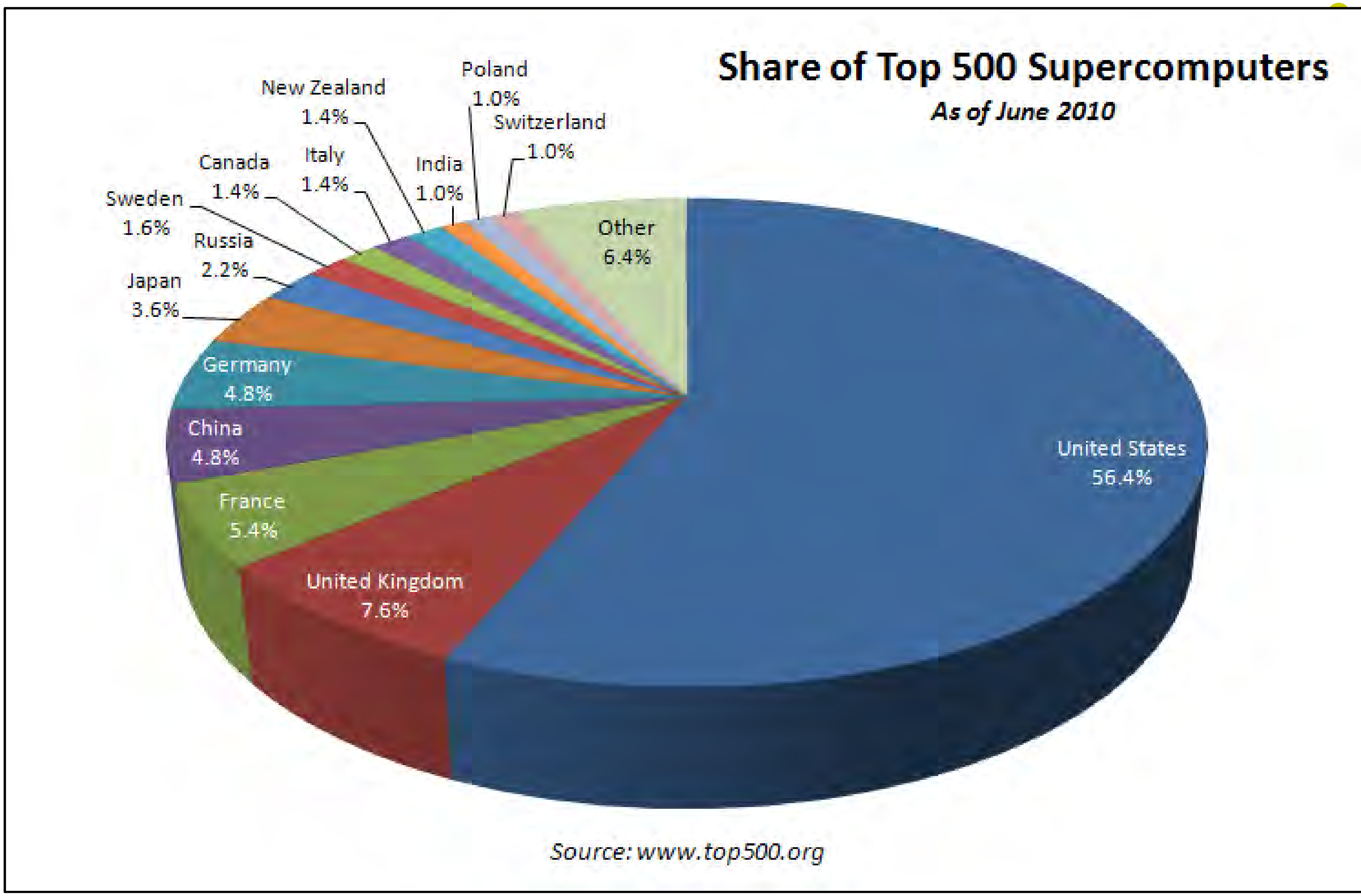
Supercomputers Evolution



Supercomputers Evolution



Supercomputers Evolution



References

1) Parallel Programming with MPI, P. Pacheco, Morgan Kaufmann, 1997.
2) Parallel Programming in C with MPI and OpenMP, M.J. Quinn, McGraw-Hill, 2004.
3) Introduction to Parallel Algorithms, C. Xavier & S. S. Iyengar, Wiley Series on Parallel and Distributed Computing, 1998.
4) www.top500.org