departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

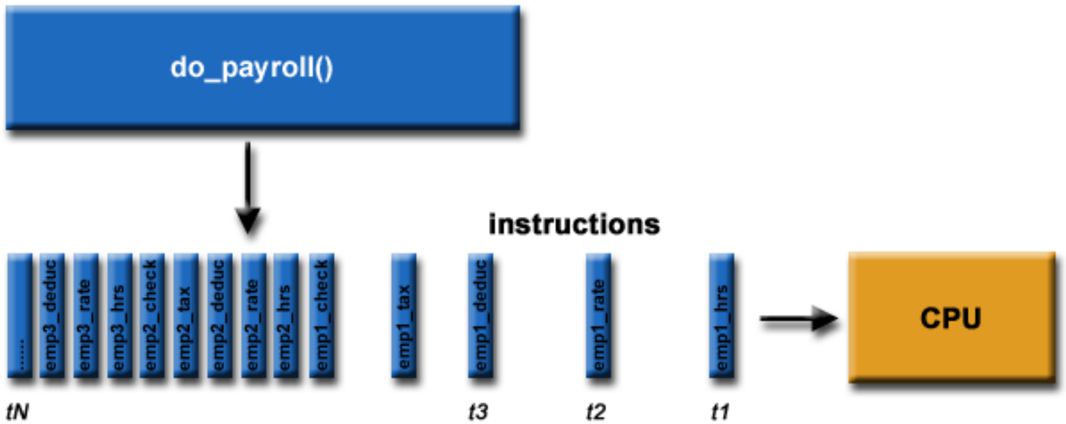
Concurrency and Parallelism
(Concorrência e Paralelismo – CP 11158)

Lecture 1
— Introduction to Parallel Computing —

Slides based in:
https://computing.llnl.gov/tutorials/parallel_comp/

What is Parallel Computing?

- Traditionally, software has been written for serial computation:
  - To be run on a single computer having a single Central Processing Unit (CPU)
  - A problem is broken into a discrete series of instructions
  - Instructions are executed one after another (sequentially)
    - Only one instruction may execute at any moment in time



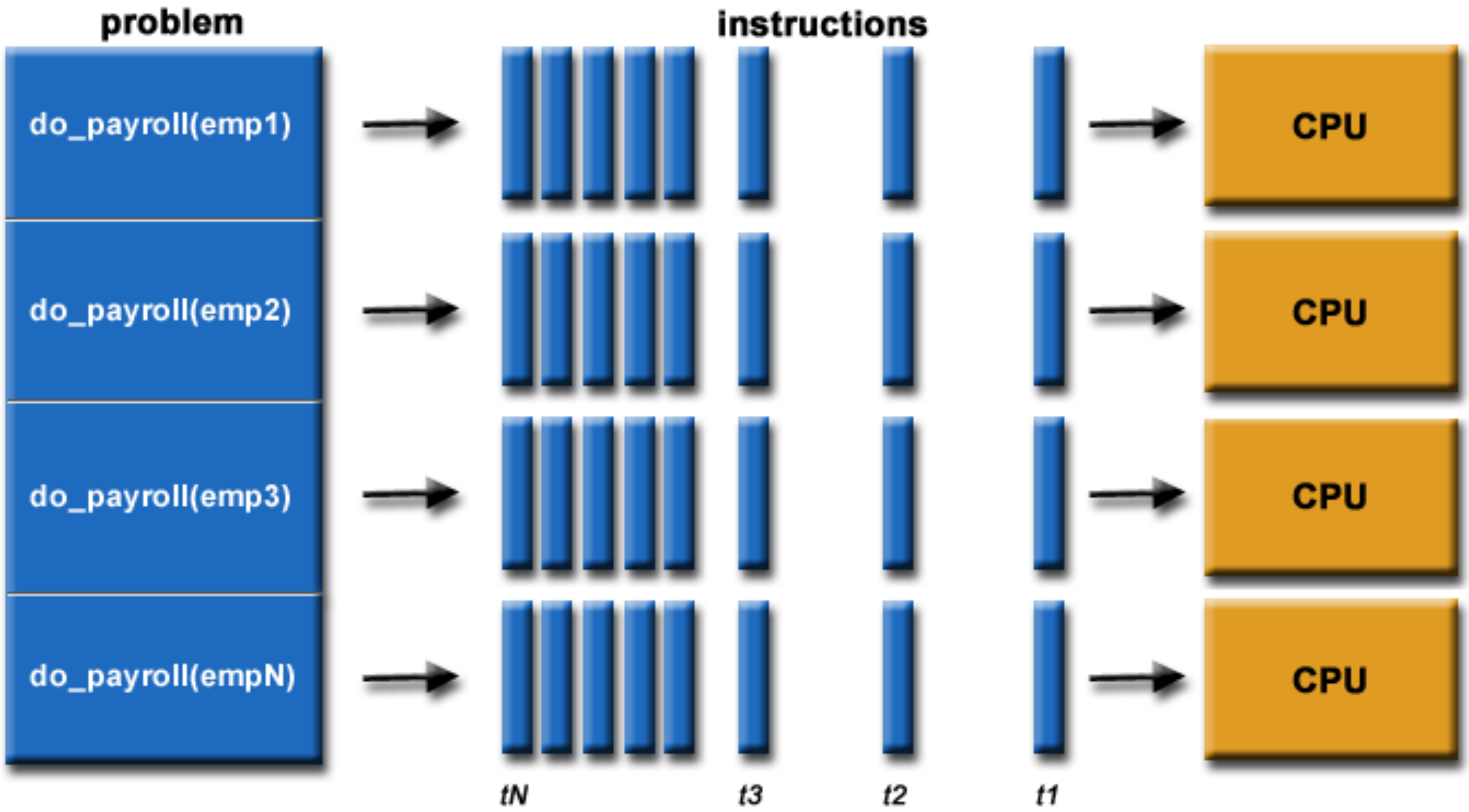
departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

What is Parallel Computing?

- Is the simultaneous use of multiple compute resources to solve a computational problem:
  - To be run using multiple processors
  - 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 processors
  - An overall control/coordination mechanism is employed

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

What is Parallel Computing?

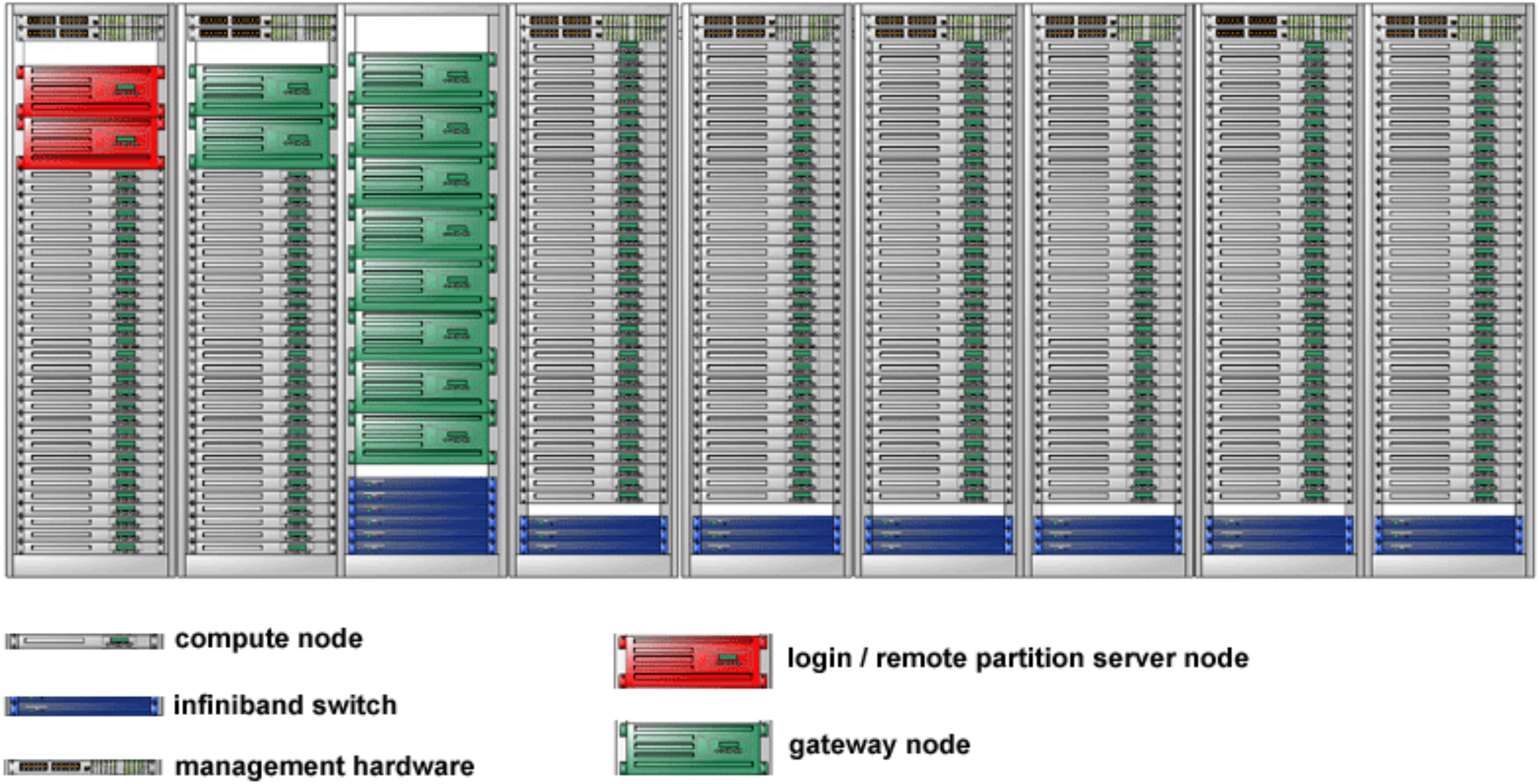


What is Parallel Computing?

- The computational problem should be able to:
  - Be broken apart into discrete pieces of work that can be solved simultaneously
  - Execute multiple program instructions at any moment in time
  - Be solved in less time with multiple compute resources than with a single compute resource
- The compute resources might be:
  - A single computer with multiple processors
  - An arbitrary number of computers connected by a network
  - A combination of both

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

What is Parallel Computing?



The Real World is Massively Parallel

- In the natural world, many complex, interrelated events are happening at the same time, yet within a temporal sequence.
- Compared to serial computing, parallel computing is much better suited for modeling, simulating and understanding complex, real world phenomena.
- For example, imagine modeling these serially:

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

The Real World is Massively Parallel

Galaxy Formation

Planetary Movments

Climate Change

Rush Hour Traffic

Plate Tectonics

Weather

Auto Assembly

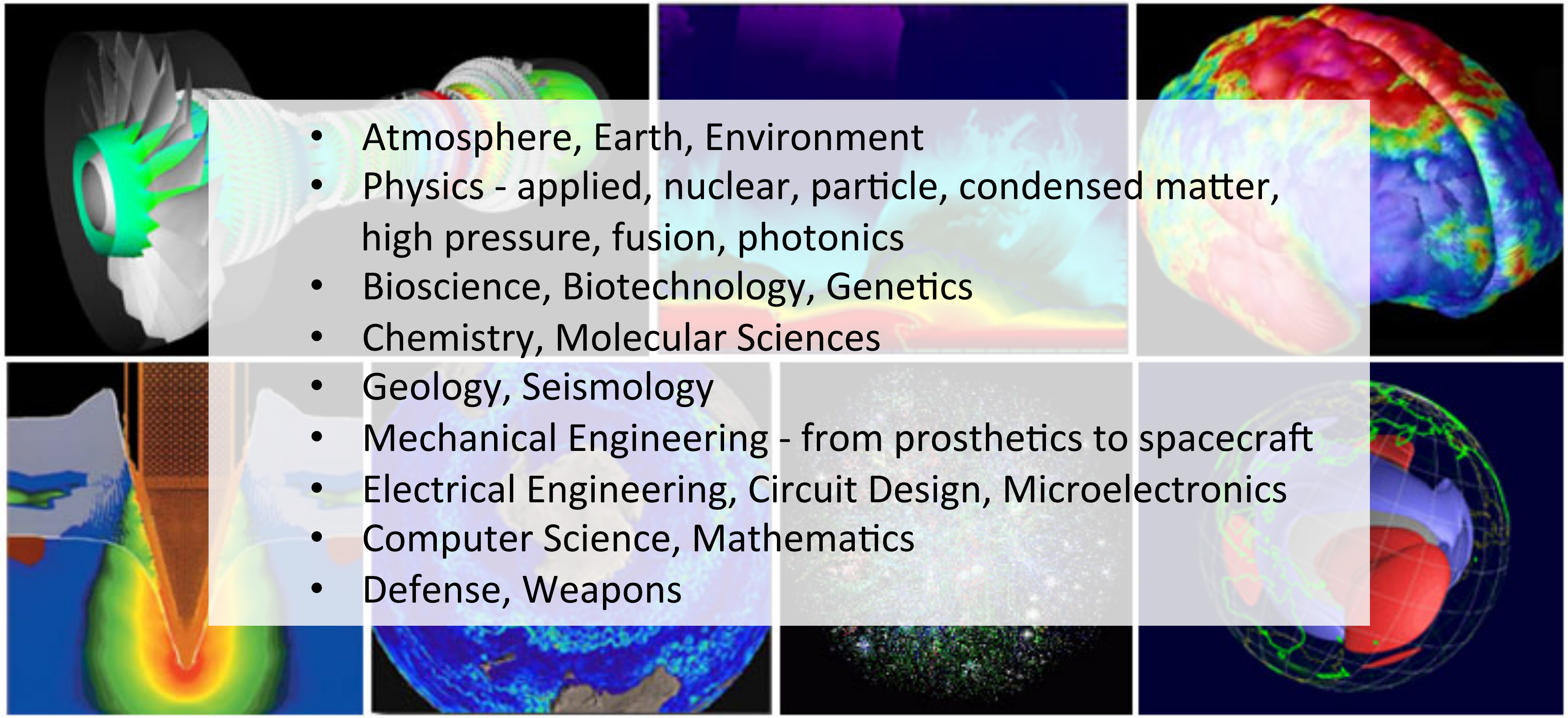
Jet Construction

Drive-thru Lunch

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Uses for Parallel Computing

- Modeling difficult problems in many areas of science and engineering



- Atmosphere, Earth, Environment
- Physics - applied, nuclear, particle, condensed matter, high pressure, fusion, photonics
- Bioscience, Biotechnology, Genetics
- Chemistry, Molecular Sciences
- Geology, Seismology
- Mechanical Engineering - from prosthetics to spacecraft
- Electrical Engineering, Circuit Design, Microelectronics
- Computer Science, Mathematics
- Defense, Weapons

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Uses for Parallel Computing

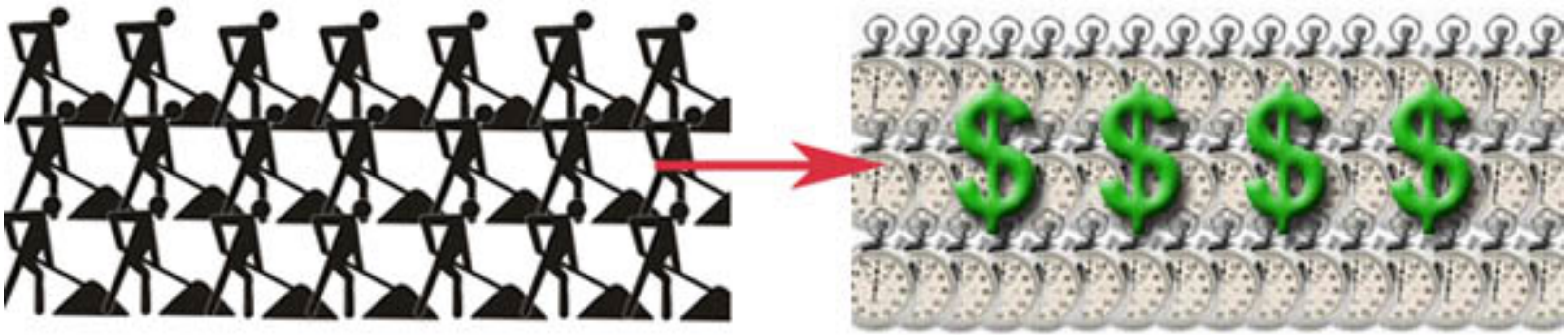
- Industrial and Commercial

- Databases, data mining
- Oil exploration
- Web search engines, web based business services
- Medical imaging and diagnosis
- Pharmaceutical design
- Financial and economic modeling
- Management of national and multi-national corporations
- Advanced graphics and virtual reality, particularly in the entertainment industry
- Networked video and multi-media technologies
- Collaborative work environments

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Why Use Parallel Computing?

- Save time and/or money
  - In theory, throwing more resources at a task will shorten its time to completion, with potential cost savings. Parallel computers can be built from cheap, commodity components.



Why Use Parallel Computing?

- Solve larger problems
  - Many problems are so large and/or complex that it is impractical or impossible to solve them on a single computer. For example:
    - "Grand Challenge" (en.wikipedia.org/wiki/Grand_Challenge) problems requiring PetaFLOPS and PetaBytes of computing resources.
    - Web search engines/databases processing millions of transactions per second.

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

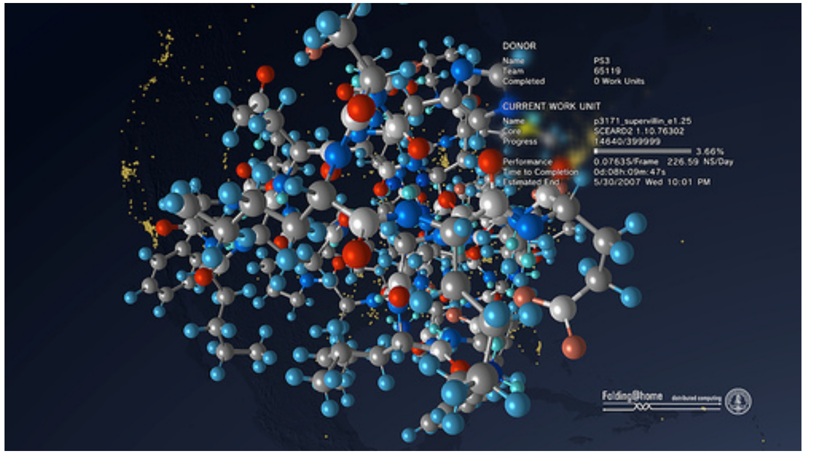
Why Use Parallel Computing?

- Provide concurrency
  - A single compute resource can only do one thing at a time. Multiple computing resources can be doing many things simultaneously. For example, the Access Grid (www.accessgrid.org) provides a global collaboration network where people from around the world can meet and conduct work "virtually".

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Why Use Parallel Computing?

- Use of non-local resources
  - Using compute resources on a wide area network, or even the Internet when local compute resources are scarce. For example:
    - SETI@home (setiathome.berkeley.edu) over 1.3 million users, 3.4 million computers in nearly every country in the world. Source: www.boincsynergy.com/stats/ (June, 2013).
    - Folding@home (folding.stanford.edu) uses over 320,000 computers globally (June, 2013)



departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

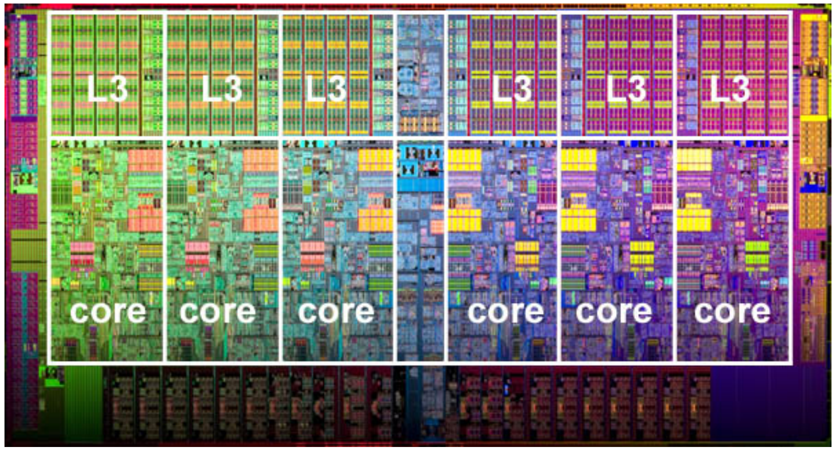
Why Use Parallel Computing?

- Limits to serial computing
  - Both physical and practical reasons pose significant constraints to simply building ever faster serial computers:
    - Transmission speeds
      - » the speed of a serial computer is directly dependent upon how fast data can move through hardware. Absolute limits are the speed of light (30 cm/nanosecond) and the transmission limit of copper wire (9 cm/nanosecond). Increasing speeds necessitate increasing proximity of processing elements.
    - Limits 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.

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Why Use Parallel Computing?

- Limits to serial computing
  - Both physical and practical reasons pose significant constraints to simply building ever faster serial computers:
    - 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 (or better) performance is less expensive.
    - Current computer architectures are increasingly relying upon hardware level parallelism to improve performance:
      - » Multiple execution units
      - » Pipelined instructions
      - » Multi-core



The Future

- During the past 20+ years, the trends indicated by ever faster networks, distributed systems, and multi-processor computer architectures (even at the desktop level) clearly show that parallelism is the future of computing.
- In this same time period, there has been a greater than 1000x increase in supercomputer performance, with no end currently in sight.
- The race is already on for Exascale Computing! (10^{18} FLOPS)

departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

The Future

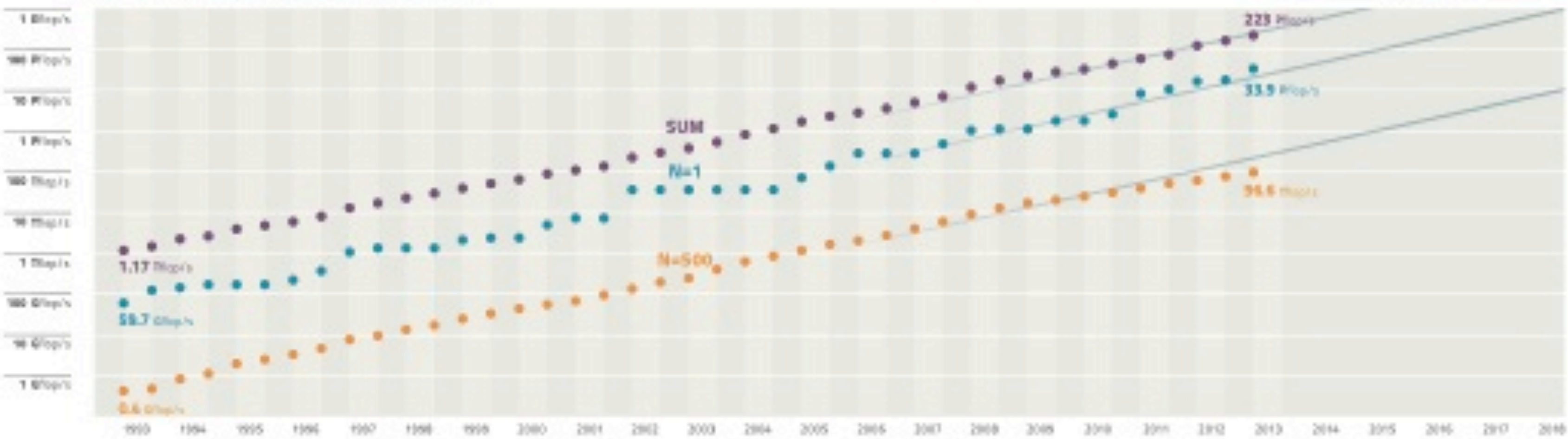
TOP500 JUNE 2013

PRESENTED BY UNIVERSITY OF MANNHEIM · ICL INNOVATIVE COMPUTING LABORATORY · BERKELEY LAB Lawrence Berkeley National Laboratory

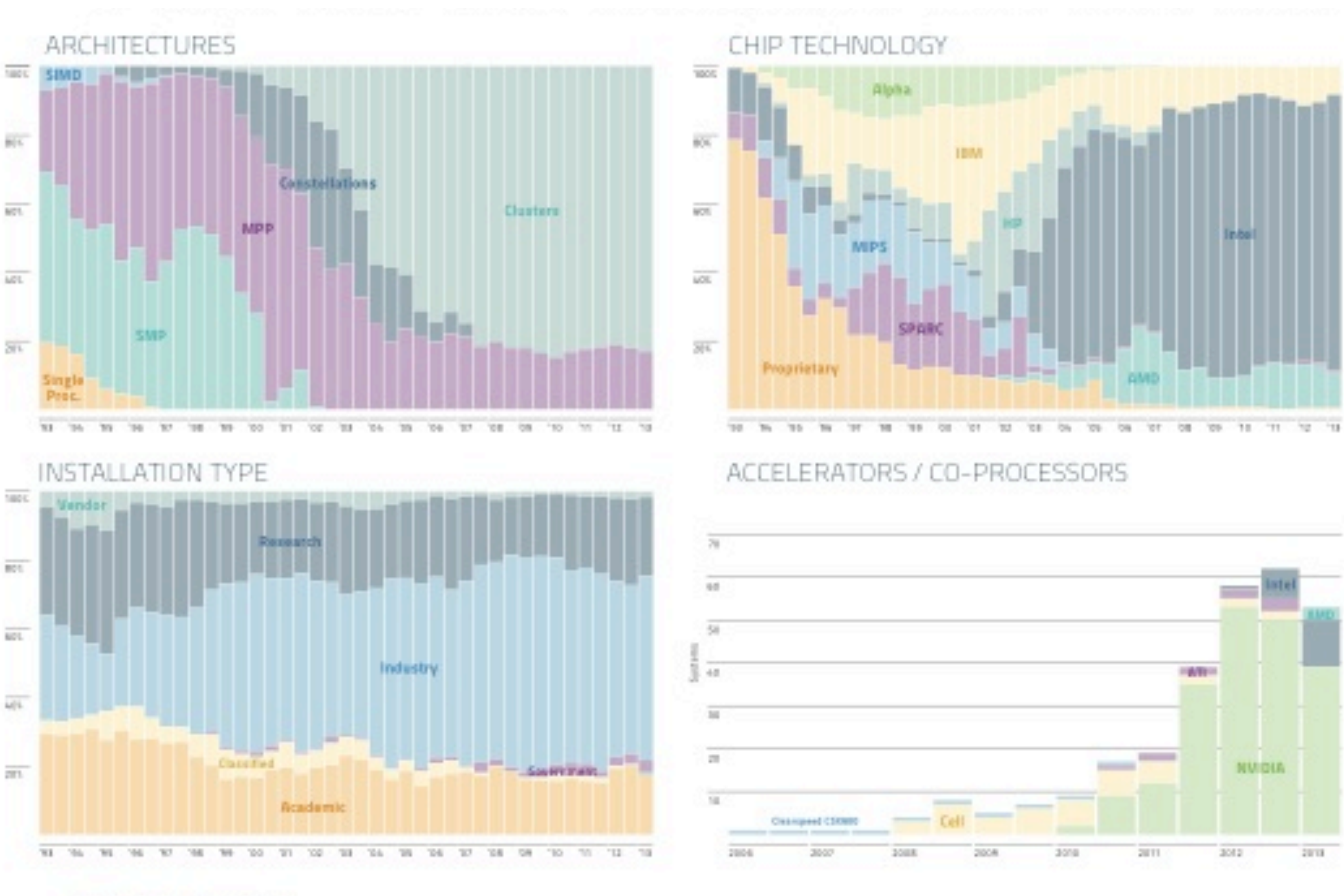
FIND OUT MORE AT www.top500.org

	NAME	SPECS	SITE	COUNTRY	CORES	RMAX PFLOP/S	POWER MW
1	Tianhe-2 (Milkyway-2)	NUDT, Intel Ivy Bridge (12C, 2.2 GHz) & Xeon Phi (57C, 1.1 GHz), Custom interconnect	NUDT	China	3,120,000	33.9	17.8
2	Titan	Cray XK7, Opteron 6274 (16C, 2.2 GHz) + Nvidia Kepler (14C, .732 GHz), Custom interconnect	DOE/SC/ORNL	USA	560,640	17.6	8.3
3	Sequoia	IBM BlueGene/Q, Power BQC (16C, 1.60 GHz), Custom interconnect	DOE/NNSA/LLNL	USA	1,572,864	17.2	7.9
4	K computer	Fujitsu SPARC64 VIIIfx (8C, 2.0GHz), Custom interconnect	RIKEN AICS	Japan	705,024	10.5	12.7
5	Mira	IBM BlueGene/Q, Power BQC (16C, 1.60 GHz), Custom interconnect	DOE/SC/ANL	USA	786,432	8.16	3.95

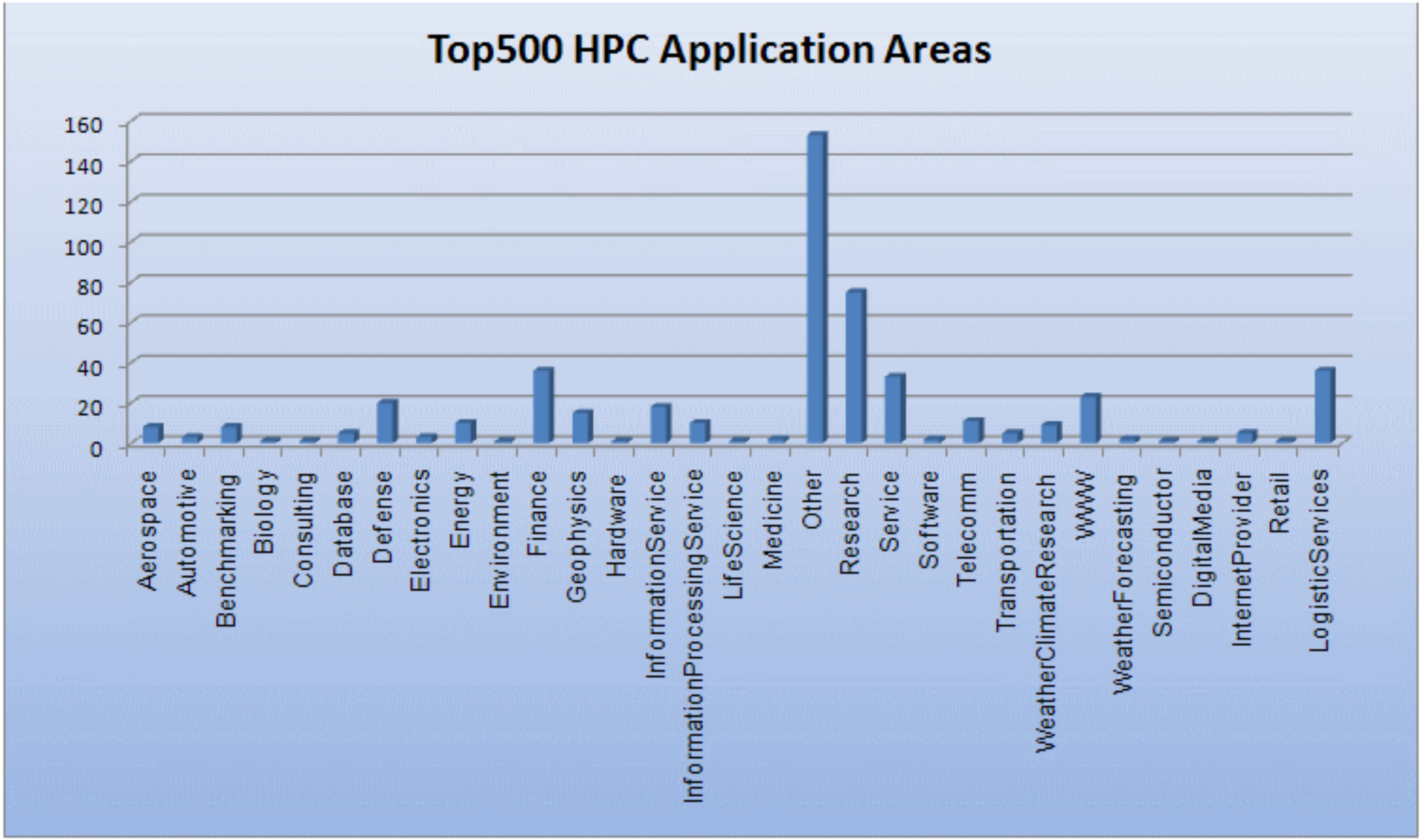
PERFORMANCE DEVELOPMENT



The Future

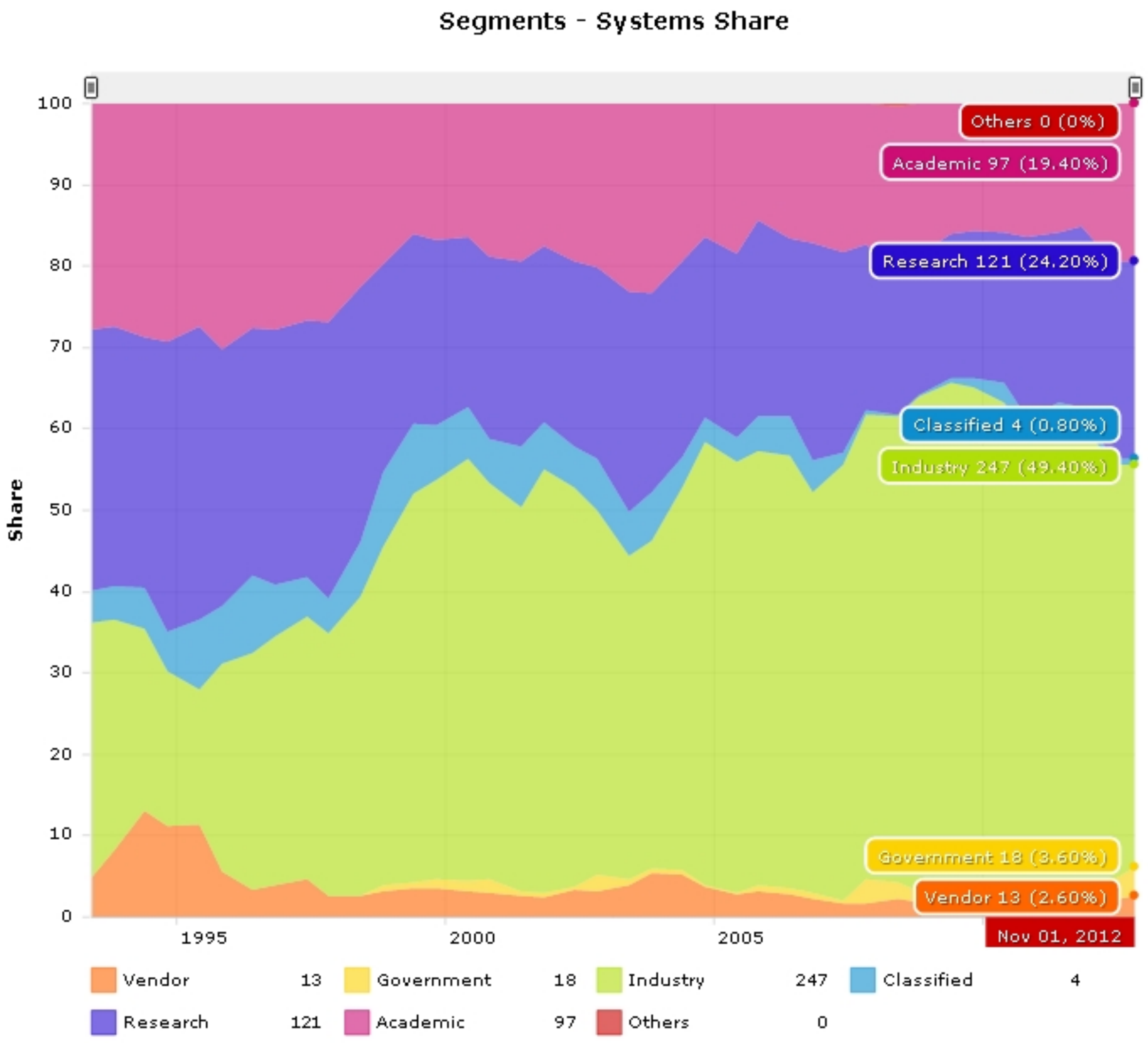

ARCHITECTURES
SIMD
MPP
Constellations
Clusters
SMP
Single Proc.
CHIP TECHNOLOGY
Alpha
IBM
HP
MIPS
Intel
SPARC
Proprietary
AMD
INSTALLATION TYPE
Vendor
Research
Industry
Classified
Academic
ACCELERATORS / CO-PROCESSORS
Intel
AMD
ATI
NVIDIA
Cell
Systems

The Future


Top500 HPC Application Areas
160
140
120
100
80
60
40
20
0
Aerospace
Automotive
Benchmarking
Biology
Consulting
Database
Defense
Electronics
Energy
Environment
Finance
Geophysics
Hardware
InformationService
InformationProcessingService
LifeScience
Medicine
Other
Research
Service
Software
Telecomm
Transportation
WeatherClimateResearch
WWW
WeatherForecasting
Semiconductor
DigitalMedia
InternetProvider
Retail
LogisticServices

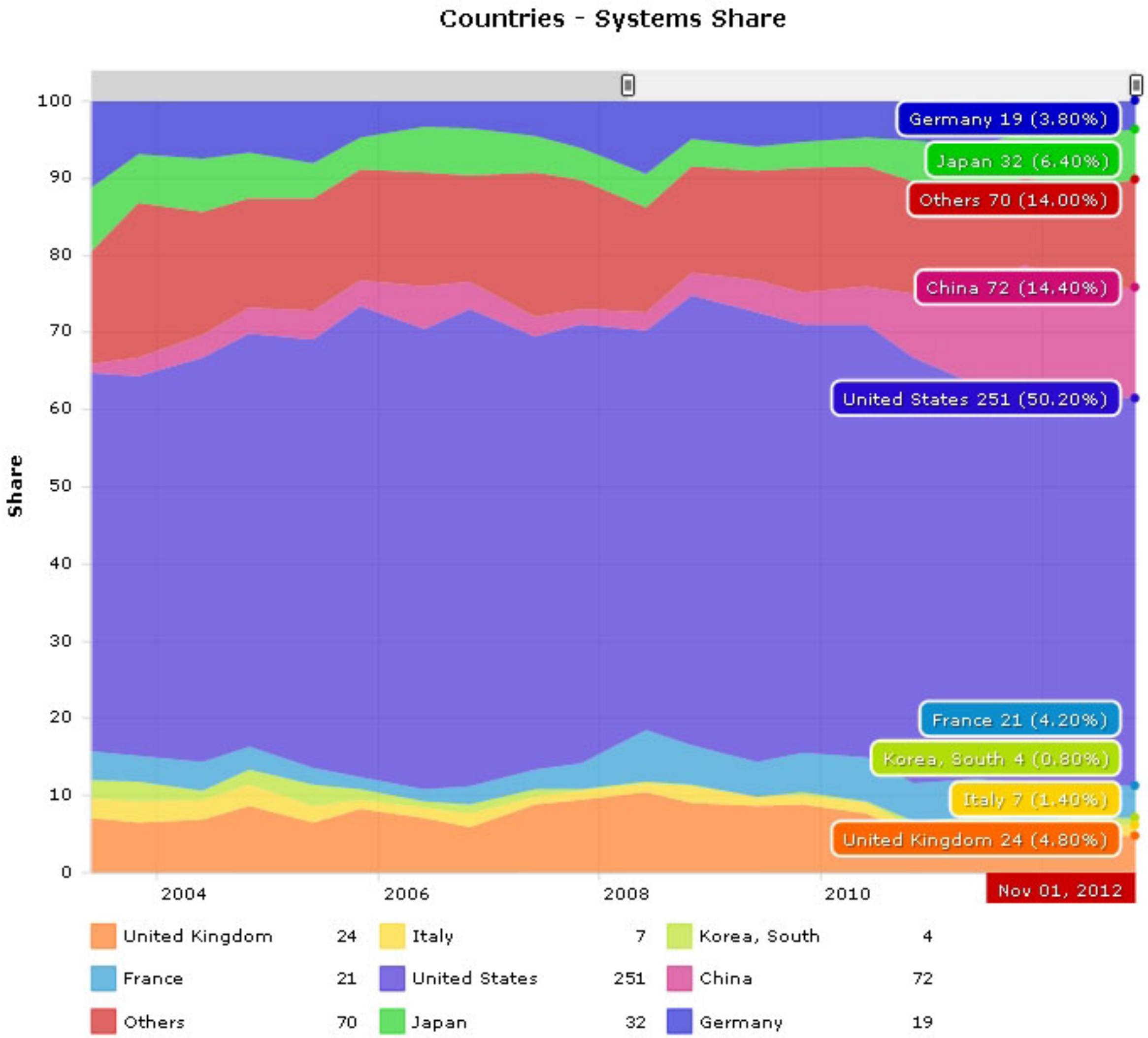
departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

The Future



departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

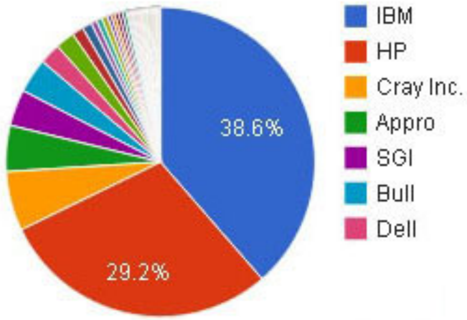
The Future



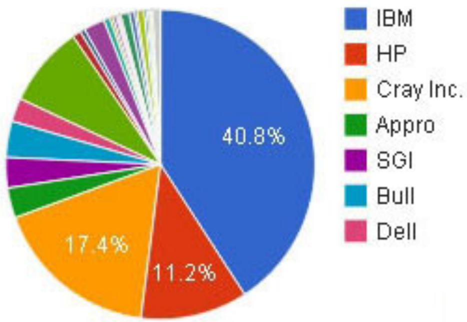
departamento de informática
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

The Future

Vendors System Share



Vendors Performance Share



Vendors	Count	System Share (%)
IBM	193	38.6
HP	146	29.2
Cray Inc.	31	6.2
Appro	24	4.8
SGI	19	3.8
Bull	18	3.6
Dell	11	2.2
Fujitsu	10	2
Oracle	6	1.2
Hitachi	5	1
NUDT	3	0.6
NRCPCET	3	0.6
Megware	3	0.6
RSC Group	2	0.4
Self-made	2	0.4
Dell/Sun/IBM	2	0.4
Supermicro	2	0.4
Dawning	2	0.4
T-Platforms	1	0.2
Adtech	1	0.2
HP/WPRO	1	0.2
NEC	1	0.2
ManyCoreSoft	1	0.2
Atipa	1	0.2
Inspur	1	0.2
NEC/HP	1	0.2
IPE, Nvidia, Tyan	1	0.2
Intel	1	0.2
Clustervision/Supermicro	1	0.2
Itautec	1	0.2
Acer Group	1	0.2
Xenon Systems	1	0.2
Raytheon/Aspen Systems	1	0.2
Eurotech	1	0.2
Lenovo	1	0.2
RSC SKIF	1	0.2

The END