



УНИВЕРЗИТЕТ
СВ. КИРИЛ И МЕТОДИЈ



ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ НАУКИ
И КОМПЈУТЕРСКО ИНЖЕНЕРСТВО

Пресметување со високи перформанси – во Европа, регионот и кај нас

Доц. Д-р Анастас Мишев
Факултет за информатички науки и компјутерско
инженерство
Универзитет Св. Кирил и Методиј
Скопје

Photo and copyright: Thilo Bubek

Агенда

- Пресметување со високи перформанси
- Историски преглед
- Модели на пресметување
- Топ 500
- HPC@EU
- HPC@SEE
- HPC@MK
- Визија и иднина

Што е HPC?

- Не постои строга дефиниција
 - Пресметување на компјутери со високи перформанси
- Решавање на проблеми со огромни процесорски и мемориски побарувања
 - Пример: проблем што ПЦ го решава за 8 часа, Cray C90 го решава за 0,002 секунди
 - Но чини 26.000.000 \$
- Брзината е пресудна
- Клучен збор : паралелизмот

Што опфаќа НРС?

○ Хардвер

- Процесорска и системска архитектура
- Поврзувачка мрежа

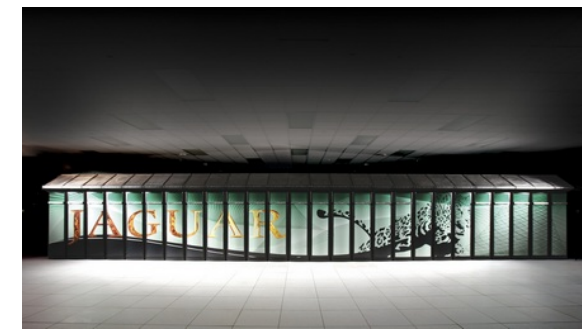
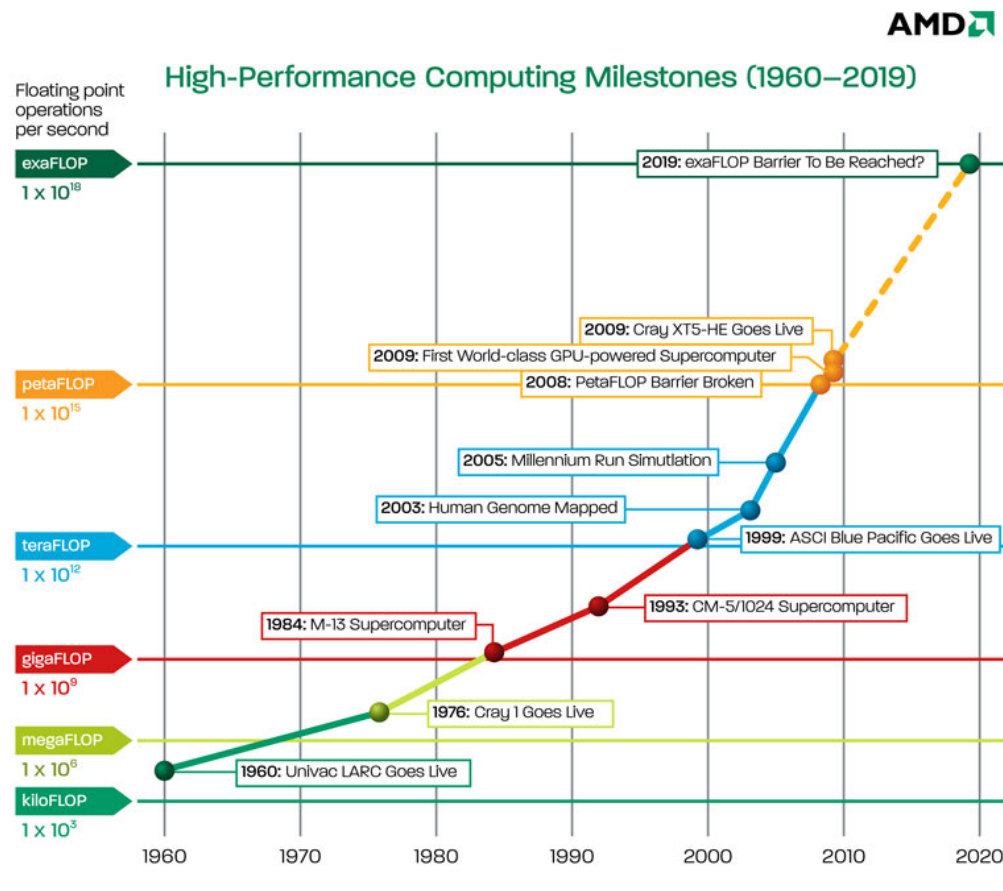
○ Софтвер

- Оперативни системи
- Програмски јазици
- Библиотеки за поддршка
- Паралелни алгоритми

Историјата се повторува

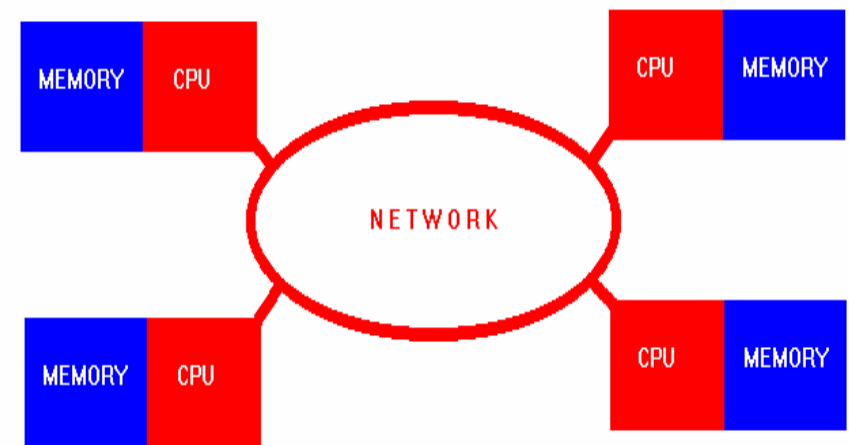
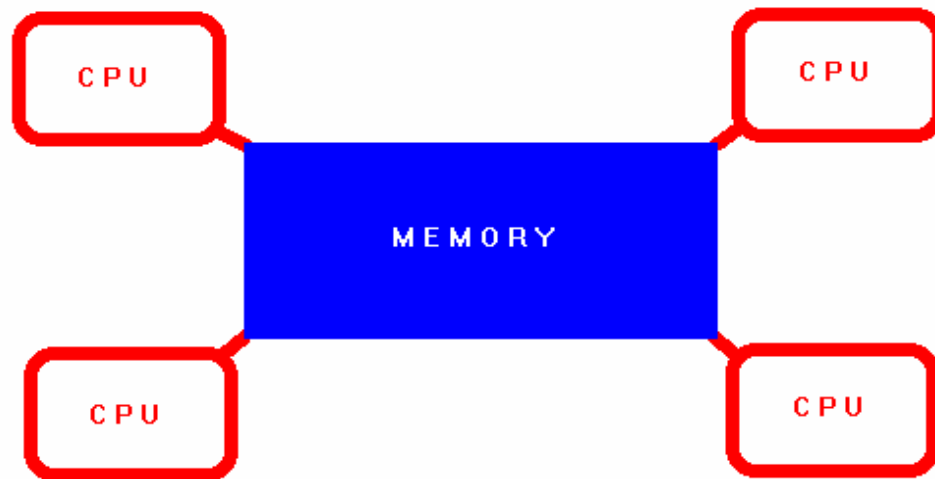
- Се' се врти во круг
- Пред 30-тина години, доминираа “џиновите”
 - Централни компјутери наречени mainframes
- Следи РС револуцијата
- Денес повторно се враќаме на почетната идеа
 - Пресметувачките ресурси се централизираат
 - Складирачките ресурси се централизираат
 - Се' е во “облакот”
 - Некои нови “џинови”

НРС низ времето

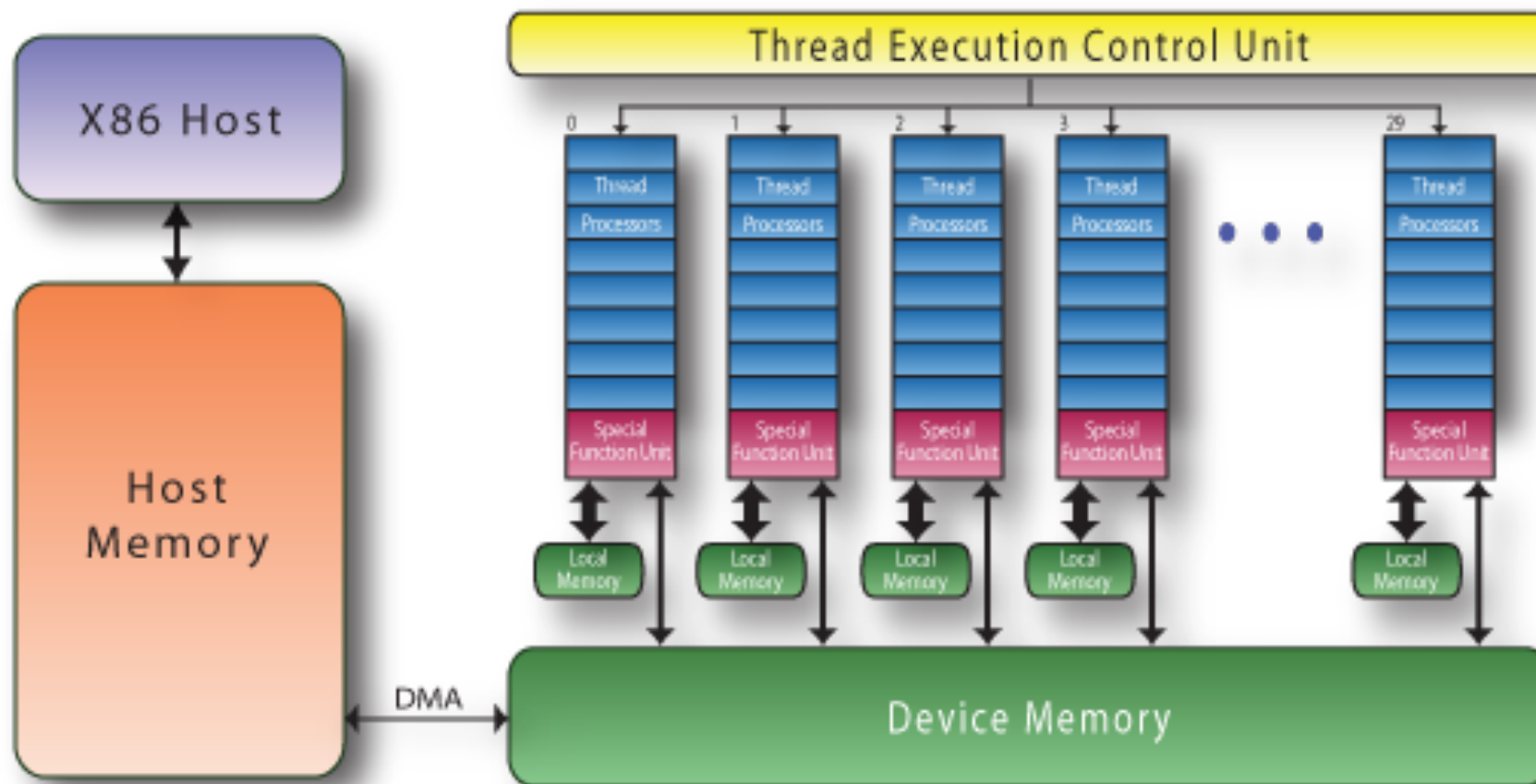


©2010 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, combinations thereof, are trademarks of Advanced Micro Devices, Inc. All other trademarks are the property of their respective owners.

Модели на паралелно пресметување



Нови трендови - GPGPU



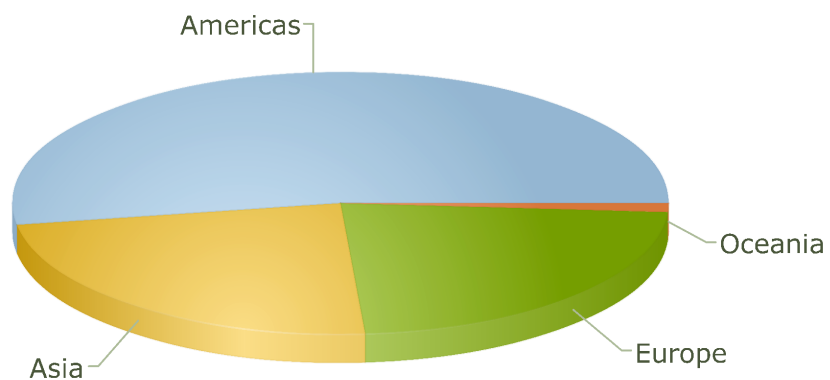
Мерење на силите – top500.org

- Централна локација за споредба
- Рангирање според постигнати резултати на конкретен тест
- 2 листи годишно
 - Ноември
 - Јуни

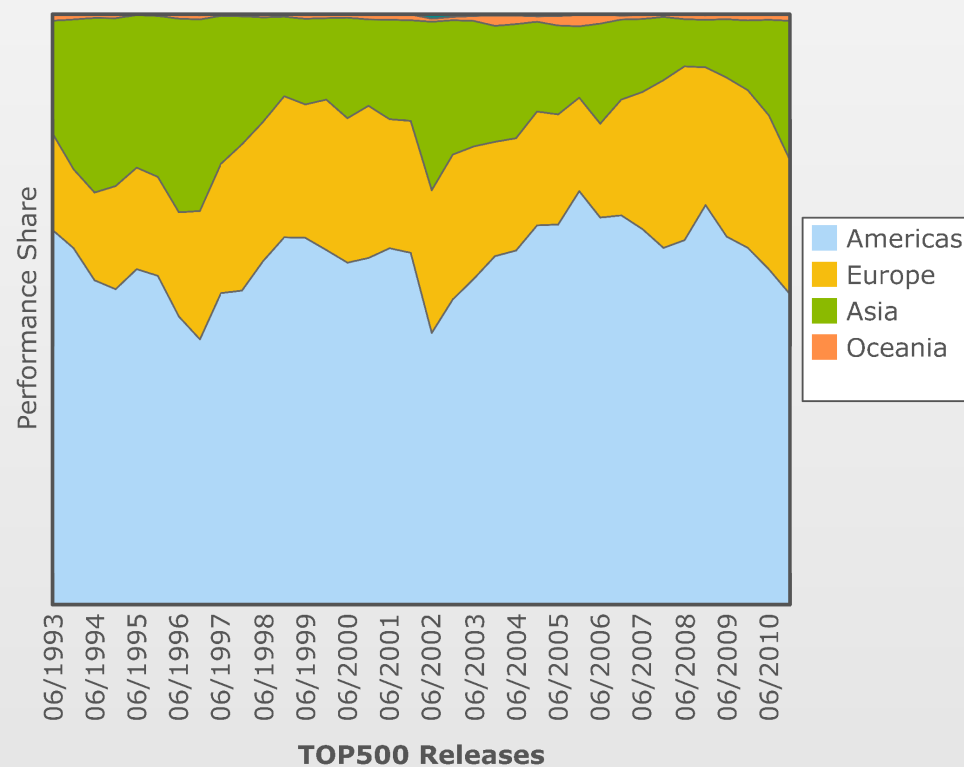
#	Tflops	Име	Систем и процесори	Земја
1	4701.00	<i>Tianhe-1A</i>	NUDT YH Cluster 14,336x6 Xeon + 7168x14 Fermi	Кина
2	2331.00	<i>Jaguar</i>	Cray XT5 224,162 Opteron	САД
3	2984.30	<i>Nebulae</i>	Dawning TC3600 Blade 55,680 Xeon + 64,960 Tesla, InfiniBand	Кина
4	2287.63	<i>TSUBAME 2.0</i>	HP Cluster Platform 3000SL 73,278 Xeon, Fermi	Јапонија
5	1288.63	<i>Hopper</i>	Cray XE6 153,408 Opteron	САД

Однос на силите

Continents / Performance
November 2010



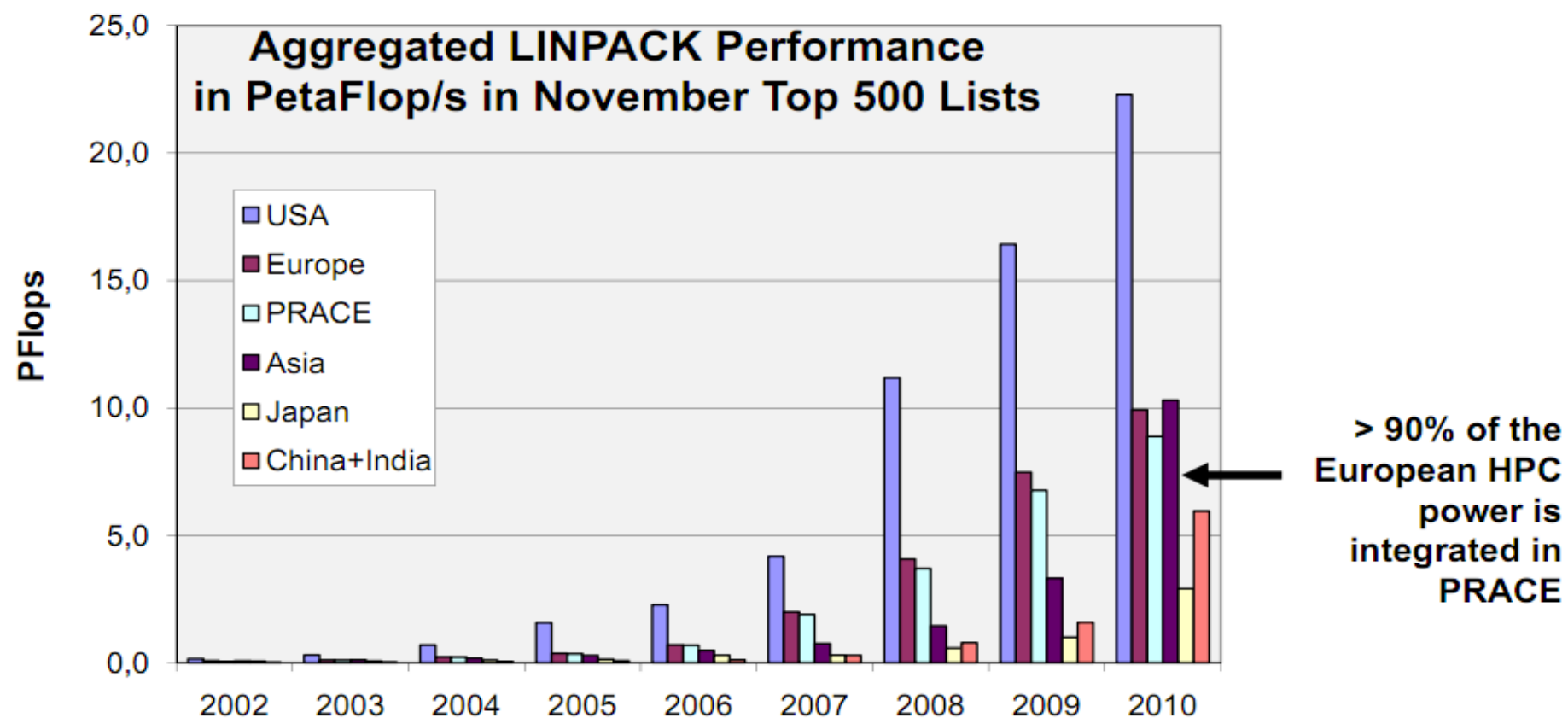
Continents Share Over Time
1993-2010



HRCS@EU

- Стратешка определба на ЕУ
- Проект PRACE
- Прва фаза
 - Траење: 2 години
 - Буџет: 28 милиони ЕУР
- Втора фаза:
 - Траење: 2 години
 - Буџет: 38 милиони ЕУР
- Обединување на европските ресурси
- Финансирање за периодот 2010–2015
 - Повеќе од 400 милиони ЕУР од земјите членки
 - Инвестирање во опрема, но и во човечки потенцијал

PRACE unites the European HPC activities to one -voice



PRACE Tier-0 Systems

- 1st Tier-0 System provides cycles since August 1
 - Jugene: BlueGene/P in GCS@Juelich
 - 72 Racks, 1 PFlop/s Peak
 - 35% of capacity provided to PRACE
- 2nd Tier-0 System announced by GENCI on October 5
 - Curie: Bull Cluster with Intel CPUs operated by CEA
 - 1.6 PFlop/s peak in Oct. 2011 (1st step in 2010)
 - Largest fraction of capacity provided to PRACE
- Next Procurements (in alphabetical order)
 - BSC, CINECA, GCS@HLRS, GCS@LRZ
 - Procurement plan based on analysis of user requirements and market



HP@SEE

- Фаќање чекор со ЕУ и светот
- Преку регионални иницијативи и проекти

Пресметување со високи перформанси за истражувачката заедница од југоисточна Европа

High Performance Computing Infrastructure for the South East Europe's Research Communities – HP-SEE



- ❑ **Contract n°:** RI-261499
- ❑ **Project type:** CP & CSA
- ❑ **Call:** INFRA-2010-1.2.3: VRCs
- ❑ **Start date:** 01/09/2010
- ❑ **Duration:** 24 months
- ❑ **Total budget:** 3 885 196 €
- ❑ **Funding from the EC:** 2 100 000 €
- ❑ **Total funded effort, PMs:** 539.5
- ❑ **Web site:** www.hp-see.eu



Partners



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

Participant Organization name	Short Name	Country
Greek Research & Technology Network	GRNET	GR
Institute of Information and Communication Technologies, Bulgarian Academy of Sciences	IICTP-BAS	BG
"Horia Hulubei" National Institute of Research and Development for Physics and Nuclear Engineering	IFIN-HH	RO
Turkish Academic Network & Information Centre	TÜBİTAK-ULAKBİM	TR
National Information Infrastructure Development Institute	NIIF	HU
Institute of Physics Belgrade	IPB	RS
Polytechnic University of Tirana	PuoT	AL
University of Banja Luka	UoBL ETF	BA
SS. Cyril & Methodius University of Skopje	UKIM	MK
University of Montenegro	UOM	ME
Research & Educational Networking Association of Moldova	RENAM	MD
Institute for Informatics & Automation Problems, National Academy of Sciences of Armenia	IIAP-NAS-RA	AM
Georgian Research & Educational Networking Association	GRENA	GE
Azerbaijan Research and Education Association	AZRENA	AZ

SEE HPC Infrastructure - Current and Near Future



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities

- Main current facility
 - IBM Blue Gene/P – **two racks 8192 cores**;
 - Theoretical peak performance: $R_{peak} = 27.85$ Tflops;
 - Energy efficiency: 371.67 MFlops/W: Green top 10
- Smaller HPC installations in **Serbia, Romania, Hungary, Bulgaria**
- **Hungary** recently procured a distributed HPC infrastructure **~30Tflops**
- **Romania** to procure a **BlueGene/P** machine
- **Macedonia** has concrete plans for procurement soon (September 2011)
- Upcoming procurements in **Serbia** and **Greece**

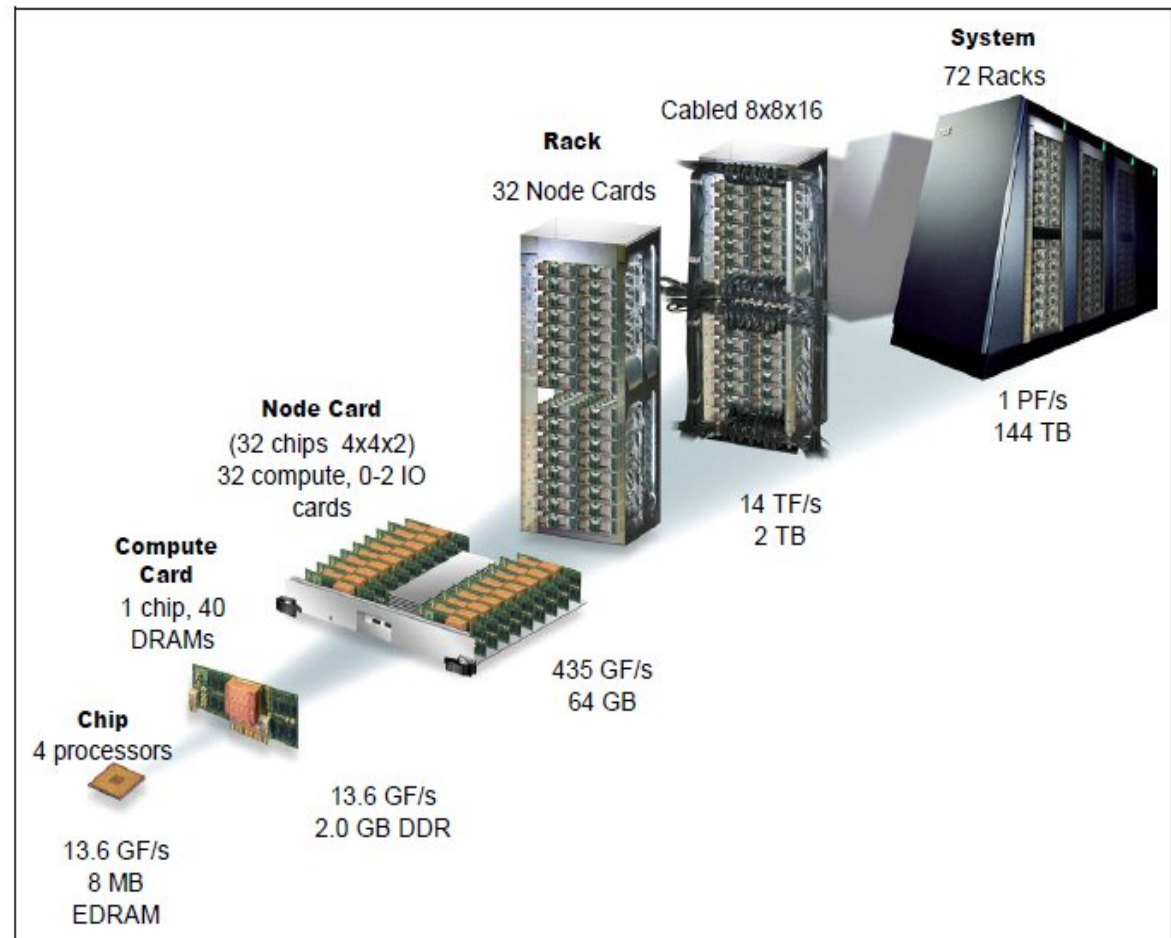


Figure 1-2 Blue Gene/P packaging

НРС@МК

- Подигање на свеста
- Идентификување на целните групи
- Воспоставување на национална ХПЦ иницијатива
- Пристап до регионални ХПЦ ресурси
- Тренинг и дисеминациски настани
- Национална ХПЦ инфраструктура

Лабораторија за Грид и ХПЦ

- Проект на владата на РМ: 60 милиони ЕУР за научни лаборатории
- Највисоко оценет предлог од домашните, а особено од странските експерти
- Единствена ИТ лабораторија во првиот повик за набавки
- Потпишан договор со испорачателот
- 01.09.2011 – Официјален почеток на националниот ХПЦ центар, во сопственост на МОН, а управуван од ФИНКИ
- За техничките карактеристики: Горан Кубелка, НЕОКОМ

ФИНКИ - Лабораторија за масивно пресметување, пресметување со високи перформанси и Грид инфраструктура

Goran Kubelka - NEOCOM

Service Executive Manager



2011 Preferred Partner
GOLD



Computing Systems
Specialist

NEOCOM - facts and figures



- Established in May 1990 as shareholder company.
- Complete ICT Solution Provider
- In 2002 Neocom opened own Business Center.
- ISO 9001:2008; 27001:2005 ;14000:2004 certified
- Long term partnership with leading vendors:



Digital since 1990, Compaq since 1993, HP since 2001

Other partners of Neocom:

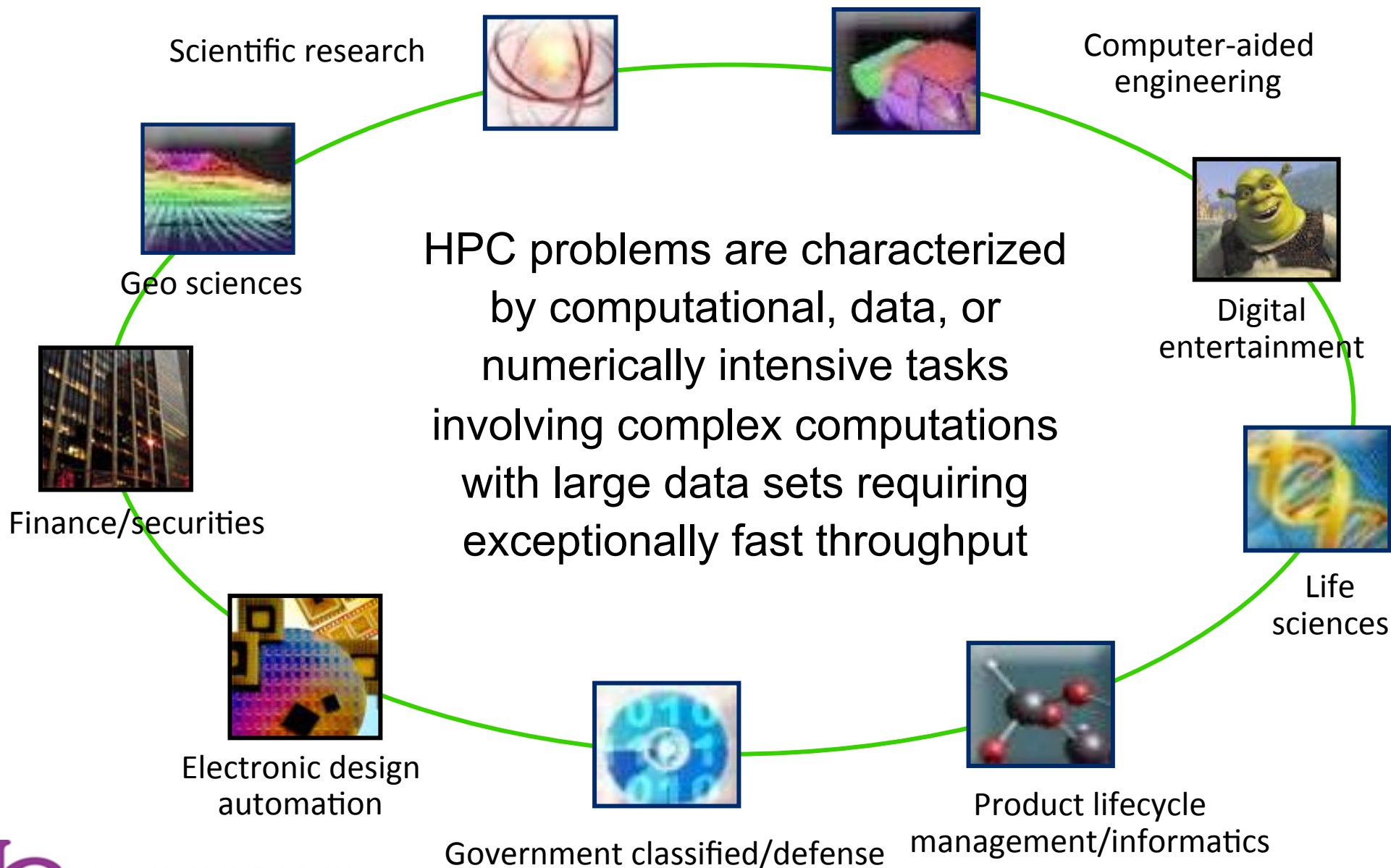
- HP
- Microsoft
- Oracle
- Lenovo
- VMWare
- Computer Associates
- Novell SuSe Linux
- RAD Data Communication
- Cisco
- OCE & KERN

Personnel:

- 50+ Employees
- HP ASE PRO Engineers
(Windows, UNIX, OpenVMS,
Linux, Storages)
- Oracle Certified Professional
- Microsoft Certified Engineers
- CA Certified
- Cisco Certified Engineers
- RAD Certified Engineers



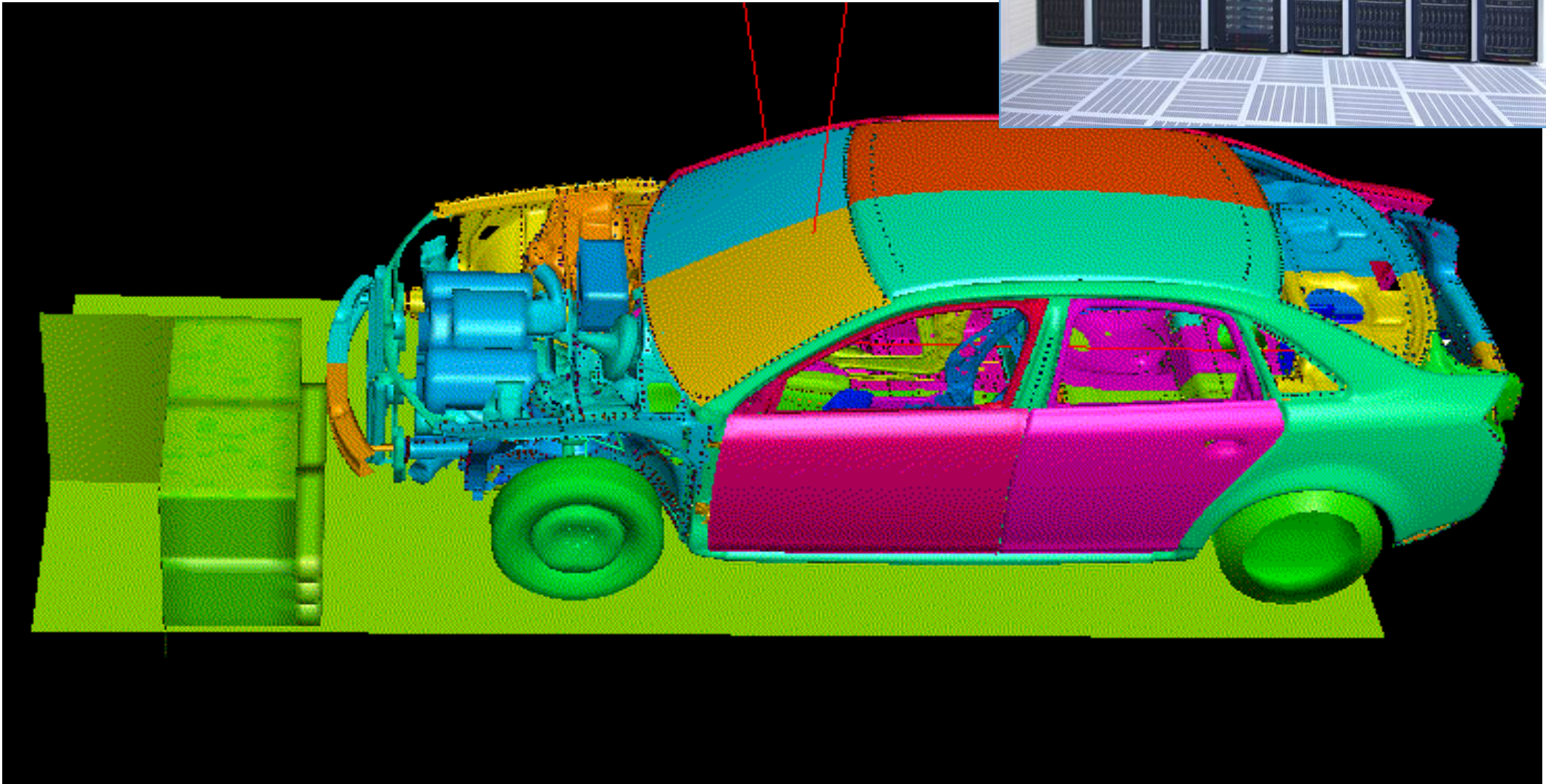
HPC : what and where ?



Automotive Example



600 x BL460 IB Cluster

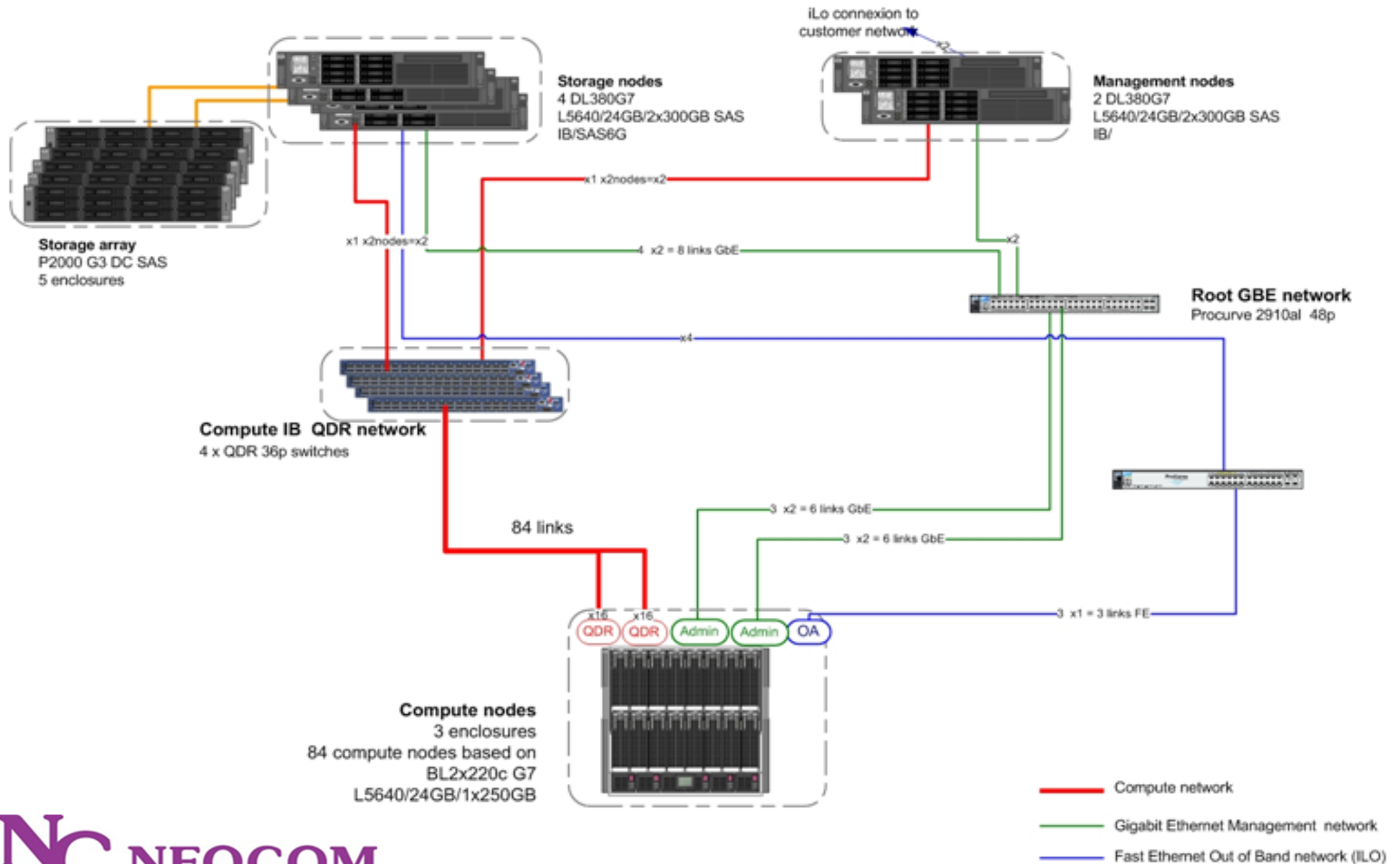


HP CAE Success Stories: www.hp.com/go/cae



FINKI – 84 compute nodes based on BL2x220 G7

Infiniband QDR interconnect



Technical description (1)

➤ Calculated performance

- ❑ Rpeak : 9329.28 GFlops
- ❑ Rmax : 8489.65 GFlops

➤ Technical specification

❑ Computing elements

- Blade solution with high density (4 CPUs, 2 boards per half size blade)
- Total number of cores 1008
- Support for hyper-threading, total logical units 2016
- Low voltage CPUs (Xeon L5640)
- 6 cores/CPU, 2 CPU/Motherboard
- QDR Infiniband
- DDR3 RAM, 24GB per Motherboard (2MB / CPU core)

HP ProLiant BL2x220c G7

BL2x220c G7

Processor	Up to 2P, up to 6 cores Intel Xeon 5600 or 5500 Series processors, up to 95w
Memory	6 DDR3 DIMM Sockets per node 96GB maximum per node (with 16GB DIMMs)
Storage	One NHP SFF SATA 9.5mm or SSD per node
Networking	Two 1GbE Ethernet NIC ports per node One Mellanox ConnectX-2 10GbE or QDR IB port per node
Mezzanine Slots	None
Management	HP Integrated Lights-Out 3 (iLO 3)
Form Factor	32 server nodes in 10u
Usage	Dense Compute with 10GbE or IB



Technical description (2)

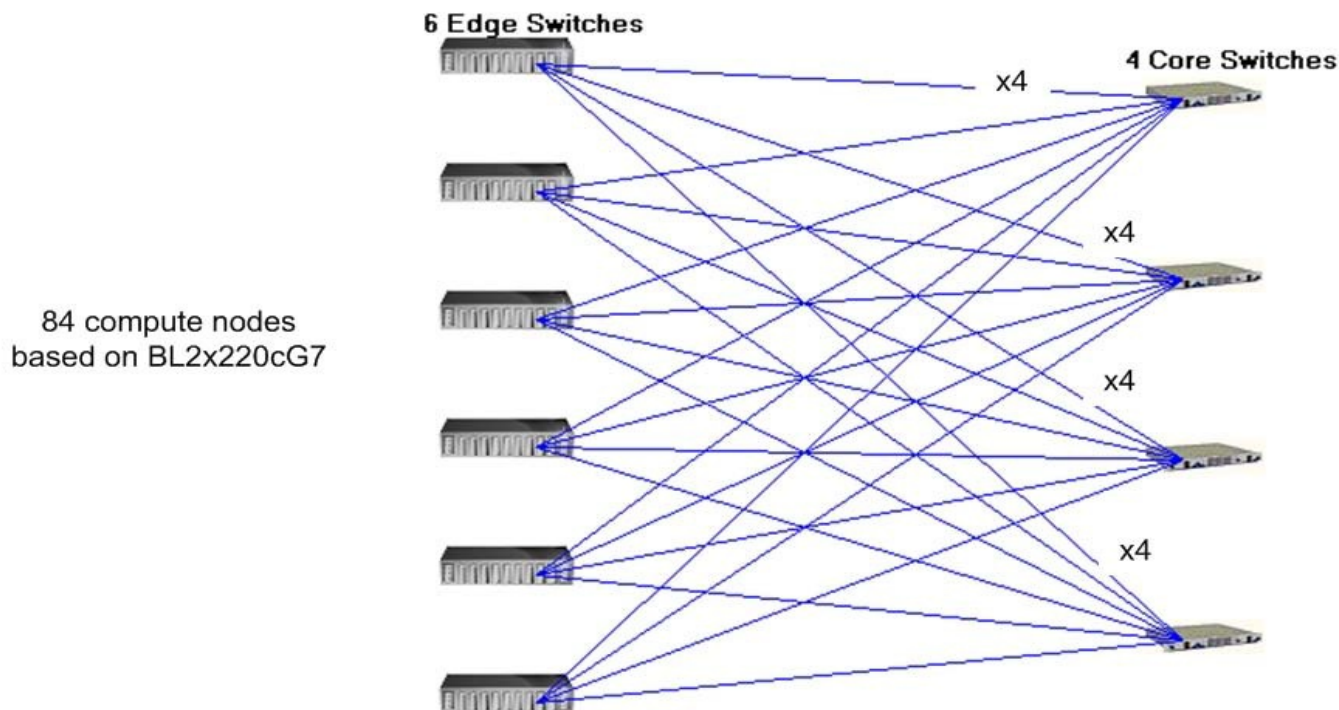
❑ Interconnectivity

- QDR Infiniband (FAT Tree)
- Non-blocking
- 1 connection for each computing element
- 1 connection for each management server

❑ Storage

- Raw capacity ≥ 30 TB
- Hot-swap SAS, 3.0 Gb/s, 10000 rpm
- RAID 0, 1, 5, 6, 10
- 2 controllers, 1GB cache/controller
- 6Gb/s SAS toward the management server

Interconnection topology

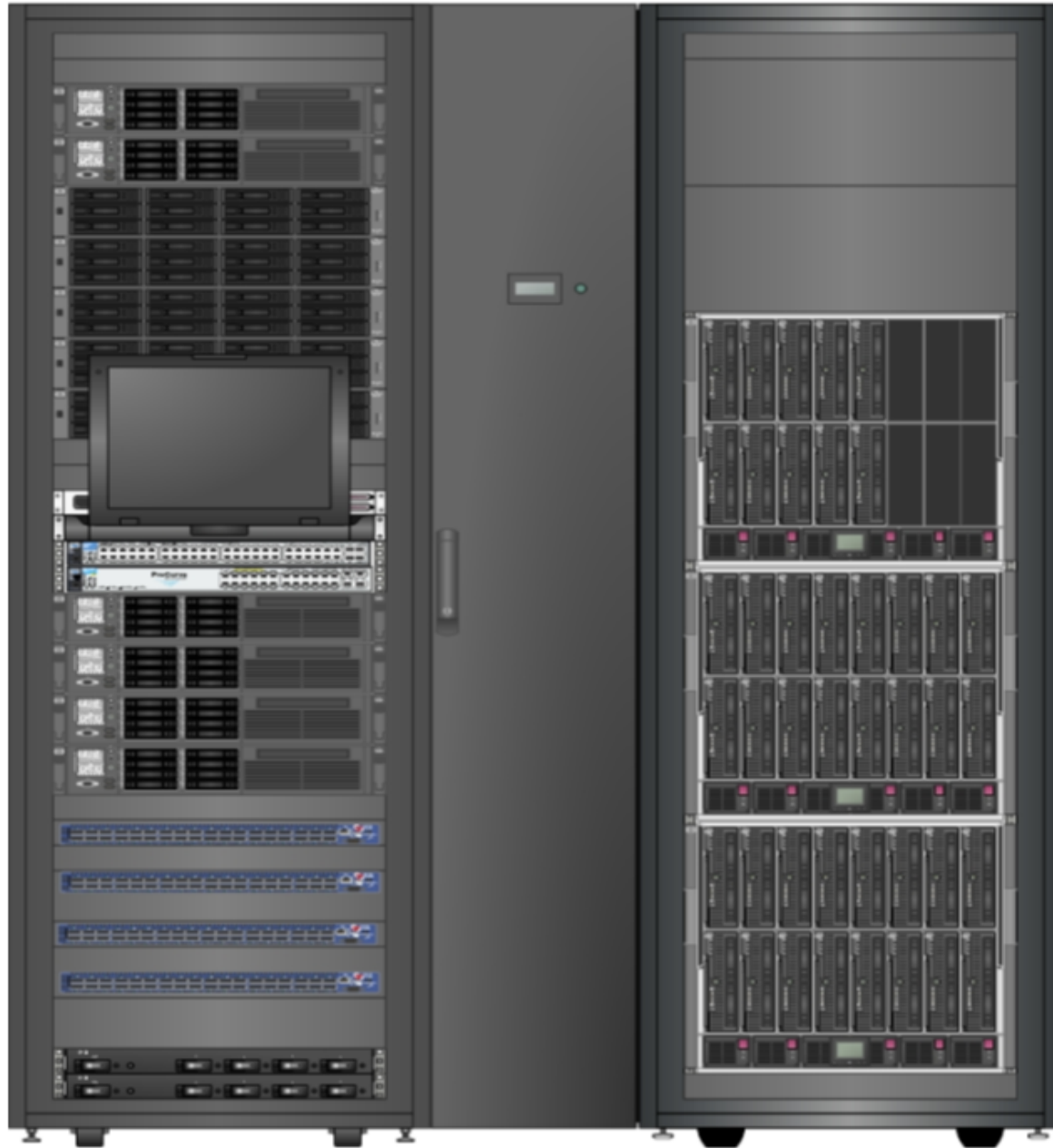


- InfiniBand FAT Tree
- Non- blocking 1:1
- 4xQDR IB/10Gb Flex-10 Adapter
- HP 4X QDR IB Switch BladeSystem c-Class
- Voltaire IB 4X QDR 36P Managed Switch

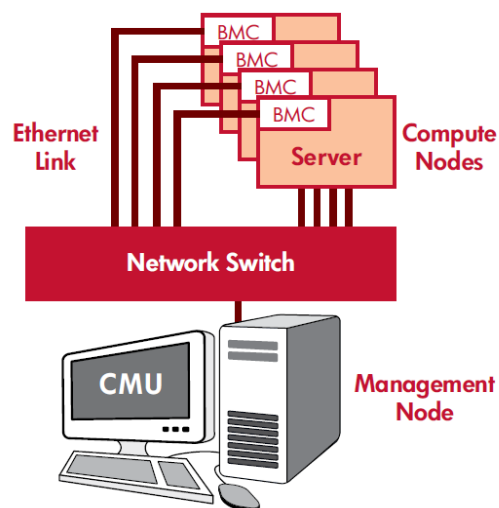
Technical description (3)

- ❑ Management/Storage servers
 - 6 servers
 - Low voltage CPUs (Xeon L5640)
 - 6 cores/CPU, 2 CPU/Motherboard
 - QDR Infiniband
 - DDR3 RAM, 24GB per server
 - SAS communication with the storage
 - 1 GB Ethernet
- ❑ Other
 - Room adoption with Electricity, UPS, Cooling, Racks
 - Power consumption: 30 KW computing and 25 KW for cooling

HPC Front View



HP Cluster Management Utility



- **Initial configuration support**
 - Automatic registration of nodes, with network parameters, in the HP CMU database
- **A robust cloning utility that remotely performs the full compute node installation by system image propagation.**
 - For the first installation or for future updates, the cluster administrator has the capability to propagate a system configuration image to all or a part of the compute nodes in the cluster.
- **A scalable and customizable monitoring tool**
 - At a glance, a system manager can see the complete behaviour of the cluster or part of it without to painfully analyze the performance of the individual compute nodes. Sensors can be easily defined.
 - monitors enable rapid drill down and understanding of cluster performance and system health
 - summarizes the states of all the nodes and a summary per group
 - displays customizable information and alerts on all the nodes of the cluster
- **A graphic user interface for system administration**
 - From HP CMU you can administrate your cluster (eg: halt, boot reboot, power off, broadcast commands from a single keyboard session Interfaces),
- **A command line interface option** for day-to-day system administration, with support for scripting



RACK

HP Modular Cooling System G2 10642 G2 + Expansion Rack



HP Modular Cooling System Management



HP Modular Cooling System



User Name

Password

You have been logged out

[Sign in](#) [Clear](#)

Copyright 2007 Hewlett-Packard, Inc. All rights reserved. HP and the HP logo are registered trademarks of Hewlett-Packard, Inc.

**Modular Cooling System** Cupertino Inazu Proto #2

[Home](#) [Logs](#) [Setup](#) [Help](#)

Menu

- Overview
- Alarms
- Identification

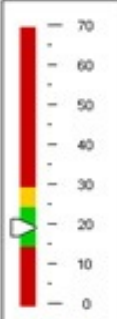
Alarms and Warnings

Updated 11/13/2007, 12:00:54

 No Active Alarms...

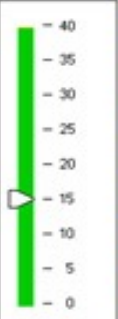
Legend

Overview



19 °C

Server Intake Temp.



15 kW

Cooling

Status	
	Server Intake Temperature 19 °C
	Server Exhaust Temperature 35 °C
	Cooling 15 kW
	Water Inlet Temperature 15 °C
	Water Outlet Temperature 19 °C
	Water Flow 54 l/min
	Water Valve Current 100 % Target 100 %
	Water System Integrity OK 10 Cycles Duration 16 s
	Fan Speed Target Level 2
	Cooling Module 1 Intake 20 °C Exhaust 38 °C 1290 rpm 1275 rpm
	Cooling Module 2 Intake 19 °C Exhaust 36 °C 1290 rpm 1260 rpm
	Cooling Module 3 Intake 18 °C Exhaust 32 °C 1275 rpm 1245 rpm
	AC input 1 OK
	AC input 2 OK
	Magnetic Door Lock Status OK

[Refresh](#)

NC NEOCOM

Визија и иднина

- Технологијата незапирливо напредува
- Паралелизмот станува клучен за подобрување на перформансите
- Вложување во едукација и тренинг за ХПЦ
- Обезбедување одржливост на ХПЦ ресурсите
- Обединување на ресурсите на европско ниво
- Фаќање чекор со современите трендови

Благодариме за вниманието

