Software engineering in High Performance Computing

Anastas Mishev
Faculty of Computer Science and Engineering
UKIM

Agenda

- HPC
- HPC Developers
- Problems
- Testing HPC software
- The HP-SEE project
- Further work
- Conclusion

HPC

- Usage of advanced computing systems (aka supercomputers) in solving complex computing problems
- It includes
  - Infrastructure
    - Architectures
    - Interconnection
    - Operating systems
    - Middleware
  - Parallel programming
    - Languages
    - Algorithms

HPC Developers

- Majority of the developers are not computer scientists or engineers
- Their primary goal: computational science
  - Achieving scientific results in their field of research using computers
- Almost all of them lack formal software engineering knowledge

HPC developers from SE perspective

- Teams
  - Individual developers
  - Large, distributed groups
- Code lifetime (development/usage ratio)
  - Short
  - Long
- Users
  - Developers
  - Community users

Some thoughts

- "Their goal is to do science, not execute programs"
- "FLOPS are not measure of science achieved"
- The results are measured not on the program performance, but on the scientific output achieved
- Writing codes that perform efficiently on HPC systems is a means to an end, not an end to itself
- Not always ready to sacrifice to increase program performance, especially if it means maintainability problems.

Technologies

- Skepticism for new technologies
- Key to success: co-existence of the new and old technologies
- Shared resources and problems
  - Batch queues make the debugging hard
  - Remote access even more

SE vs. HPC

- Mismatches
  - OO vs. FORTRAN
  - Frameworks
  - IDE
- How to help
  - Educate
  - Disseminate best practices
  - Reuse

Testing

- Different
- Often try to optimize the methods, not the programs
- Validation
  - Compare to experimental results, where possible
- Some deeper research
- Try to provide more help to the computational scientist
- The environment: HP-SEE community

HP-SEE

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

- **Contract n°:** RI-261499
- **Project type:** CP & CSA
- **Call**: INFRA-2010-1.2.3: VRCs
- **Start date:** 01/09/2010
- **Duration:** 33 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

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

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
nstitute 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-ULAKBIM	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: Rpeak= **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**

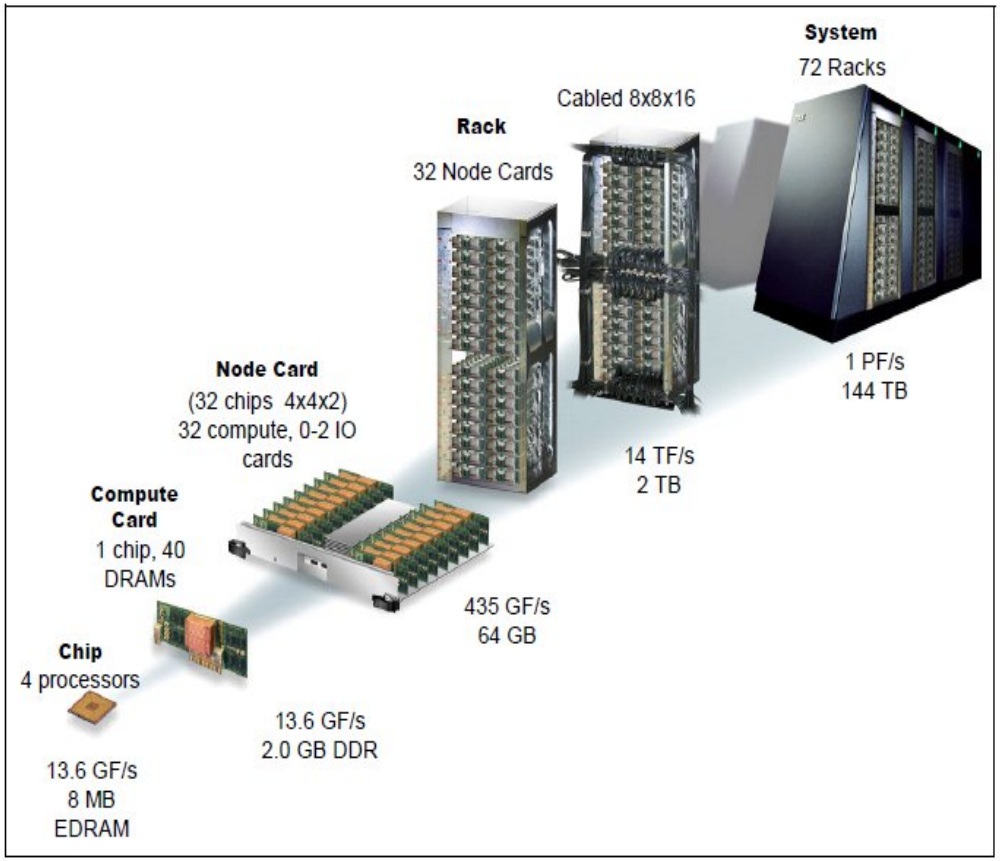

Figure 1-2 Blue Gene/P packaging

HP-SEE Virtual Research Communities

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

- Comp. Physics
  6 countries,
  8 apps.
- Comp. Chemistry
  6 countries,
  7 apps.
- Life Sciences
  5 countries,
  7 apps.

Country	Physics	Chemistry	Life Sciences	TOTAL
Albania	1			1
Armenia			1	1
Bosnia-Herzegovina		1		1
Bulgaria	2	2		4
Georgia			1	1
Greece		1	2	3
Hungary			2	2
Moldova	1			1
Montenegro			1	1
FYR of Macedonia	1	1		2
Romania	2	1		3
Serbia	1	1		2
TOTAL	8	7	7	22

How?

- Surveys
- Dissemination
- Best practices
- But most importantly, education
  - Special courses
  - Courses for computational scientists

Conclusion

- SE and HPC are not always aligned
- Computer scientist need to understand the problems of the computational scientist
- Computational scientists should trust more computer scientists
- Education is critical

Thank you for your attention

Questions