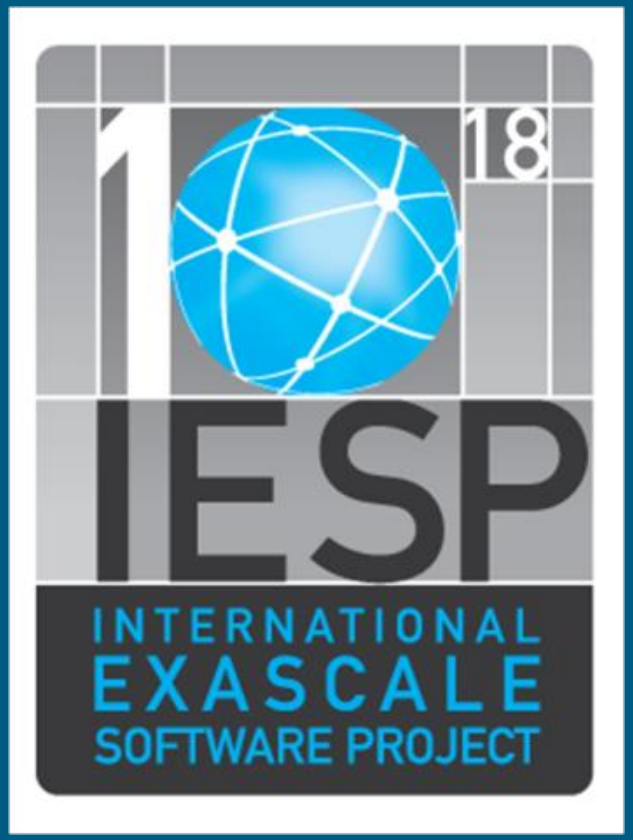
JÜLICH
FORSCHUNGSZENTRUM



Main Results of the Previous Meetings and Overview of Whitepapers

Bernd Mohr
Jülich Supercomputing Centre

Main (Re)source for this Summary

- IESP website (especially Documents section)
  - http://www.exascale.org/iesp/IESP:Documents
- Contents
  - IESP Exploratory Meeting SC08, Austin
    - Presentations
  - Meeting 1, Santa Fe
    - Presentations, breakout summaries
    - 15 whitepapers
  - Meeting 2, Paris
    - Presentations, breakout summaries
    - 15 whitepapers
  - Relevant background material
    - 11 reports (some quite extensive)

Santa Fe, April 6-8, 2009

- 67 attendees
- Agenda
  - Various talks about status of HPC landscape in Japan, US, Europe
  - Panel “Software Barriers for HPC”
  - Breakout 1: Technical roadmap discussion
  - Breakout 2: Collaboration model and funding

Santa Fe, High Level Summary

- 1 day of talks
- 0.5 day of 2-way parallel breakouts
- 14 out of 15 whitepaper written by people from one institution

- Meeting mainly to
  - Understand what IESP is about
  - Understand status of HPC landscape in other continents
  - Position ourselves
  - Get organized and started

- Technical roadmap breakout group discussed
  more process than content
- But identified important topics/persons to write position papers

Paris, June 27-29, 2009

- 73 attendees
- Agenda
  - Reports from last meeting
  - Breakout 1: Technical challenges (software)
  - Breakout 2: Computational challenges (applications)
  - Breakout 3: Management challenges (resource providers and industry partners)
  - Breakout 4: National and international funding

Paris, High Level Summary

- 0.5 day of talks
- 1.5 days of 4-way parallel breakouts
- 10 out of 15 whitepapers multi-institutional cooperations

- All breakout groups
  - identified key structural elements of coordinated software roadmap
  - prepared detailed reports about work and discussions
- Extensive 35 page report
- Provides basis for creating draft roadmap in Japan meeting

Paris, Software Group

- Identified, discussed, and assessed **important research topics**

- For each topic created **1-page summary**
  (⇨ “Summary working group slides, complete”)
  - Situation
  - State of the art
  - Needed actions
  - Roadmap + research

- Partial **categorization**
  [P] Problematic already at petascale
  [M] Currently manageable, problematic later at some point
  [ok] Need research but appear manageable now

Paris, Software Group

- **Inter-node research topics (breakout group 1B)**
  - I/O **[P]**
  - Resiliency (RAS, HA, Fault prevention, detection, recovery) **[P]**
  - Eliminate bottlenecks to strong scaling **[P]**
  - Parallel programming models (distributed memory / PGAS) **[P]**
  - Parallel debugging **[M]**
  - Performance monitoring, feedback, parallel autotuning **[M]**
  - Power and facilities **[M]**
  - Resource provisioning (workflow scheduling) **[M]**
  - Interaction with external resources (clouds, archiving, real-time data streams) **[M]**
  - System integration and management **[ok]**

Paris, Software Group

- **Intra-node research topics (breakout group 1A)**
  - 1-billion way concurrency and load balancing **[P]**
  - Locality and distributed data **[P]**
  - Sustainability and interoperability **[M]**
  - Operating systems **[M]**
  - Algorithms + Software Libraries **[M]**
  - Performance **[M]**
  - Fault tolerance **[P]**
- **Topics for future discussion**
  - Characterization of a node
  - Intra-node and inter-node interplay
  - Evolution / revolution
  - Connection to application drivers

Paris, Applications Group

- **Selected representative application domains**
  - Weather, climate, earth sciences
  - Astrophysics, HEP, and plasma physics
  - Materials science, chemistry, nanoscience
  - Life sciences
  - Engineering
  - Finance & optimization
- **Identified software challenges** as seen from application drivers
  - Still need to be compared and integrated with issues from software group
- **Collected list of experts** for further insight, analysis and validation of roadmap issues
- More detailed report by Jean-Yves (tomorrow)

Paris, Providers/Vendors Group

- Discussed
  - Lessons to be learn from other large-scale projects
  - Implications of HW and SW on HPC market
- strong indications that **vertical design approach is needed**

- **Focused on potential mechanisms and models**
  - that could facilitate **collaboration** between
    - exascale community
    - industry partners
    - government
  - in a mutually productive and sustainable way

Paris, Providers/Vendors Group

- Identified collaboration models
  - **Funded investigation**
    - Research exploration resulting in insight and knowledge
  - **Full defined purchase**
    - "Standard" product acquisition
  - **Design a development (no product)**
    - Creating and demonstrating prototypes
  - **Co-design and co-development**
    - Like purchase but shared risk (buyer + vendor)
  - **Base plus value add**
    - Produces baseline product + incentives to add more value
- Funding profiles greatly effect type + sustainability of models

Paris, Funding Group

- Group discussed
  - **Existing funding programs and cooperation mechanisms** to support exascale research
    - Cross-atlantic bilteral funding feasable and not uncommon
  - **Motivations and conditions for cooperation** between funding agencies
  - **Actions for near future**

- Recognize that the combined urgency and scope of problem necessitate a coordinated, global response
  - Requires strategy / methodology for exascale SW integration
  - **IESP roadmap plays key role** to allow agencies to plan program and define initiatives

Paris, Funding Group

- Agreed that international collaboration in exascale
  - should involve joint road mapping
  - periodic workshops
  - coordination of investments to avoid excess duplication
  - standardization
  - as well as joint projects in specific cases
  - Open source policies necessary for effective reuse

- Identified the **need for structured and stable framework for international collaboration**
  - currently often bilateral, driven by individuals
  - Multilateral at institution / agency / political level could help

Whitepapers Summary

1. Programming models and languages
   - **1-billion way concurrency and load balancing**
     - Enough threads for millions of core to hide latency + LB
   - **Locality and distributed data**
     - Need standard + scalable way to express scope and locality
   - **Eliminate bottlenecks to strong (and weak) scaling**
     - Linear scaling not enough for exascale
     - Limited memory will not always allow weak scaling
   - **Sustainability and interoperability**
     - Large existing code base
     - New models must provide incremental transition path
   - **Programming model support for fault tolerance**

Whitepapers Summary

❶ Programming models and languages

- The Biggest Need: A New Model of Computation **[SF]**
  Thomas Sterling (LSU)
- Slouching Towards Exascale **[SF]**
  Rusty Lusk (ANL)
- BSC Vision Towards Exascale **[SF]**
  Mateo Valero (BSC)
- Models of Computation - Enabling Exascale **[Pa]**
  Thomas Sterling (LSU)
- Programmability Issues **[Pa]**
  Vivek Sarkar (Rice U.), Jesus Labarta (UPC), Mitsuhisa Sato (U. of Tsukuba), Barbara Chapman (U. of Houston)
- Programming Models at Exascale: adaptive runtime systems, incomplete simple languages, and interoperability **[Pa]**
  Sanjay Kale (UIUC)

Whitepapers Summary

❷ System software

- **Operating systems**
  - Need to handle heterogeneous HW and unconventional memory structures, leave scheduling to runtime systems
- **Virtualization, dynamic provisioning, partitioning**
  - Provide fault isolation
- **System integration and management**
  - More systematic approaches needed

- The Case for A Hierarchal System Model for Linux Clusters **[SF]**
  Mark Seager, Brent Gorda (LLNL)
- Resource Management **[Pa]**
  Barney McCabe (ORNL), Hugo Falter (ParTec)
- Operating Systems for Exascale **[Ts]**
- John Shalf (LBL), Thomas Sterling (LSU)

Whitepapers Summary

❸ Tools for parallel programming

- **Performance analysis, monitoring, and workload analysis**
  - Light-weight, scalable monitoring facilities needed
  - More modeling, on-line, automation, clustering, …
  - More integration with compilers and runtime systems
- **Parallel debugging**
  - Current approaches to don't scale, new ones needed

- On the Importance of End-to-end Application Performance Monitoring and Workload Analysis at the Exascalec **[Pa]**
  David Skinner (LBNL), Alok Choudary (Northwestern U.)
- Performance at Exascale **[Pa]**
  Bernd Mohr (JSC), Matthias S. Müller, Wolfgang E. Nagel (ZIH Dresden)
- **MISSING: Parallel debugging**

Whitepapers Summary

❹ Algorithms and software libraries

- Need global synchronization-free (-minimal) algorithms
- Need fault-insensitive and error tolerant algorithms
- New approaches for pre- (e.g. meshing) and post-processing (e.g. visualization)

- XXL Simulation for XXIst Century Power Systems Operation **[SF]**
  J.Y. Berthou, J.F. Hamelin, Etienne de Rocquigny (EDF R&D)
- Application Analysis and Porting in the PRACE Project **[SF]**
  Peter Michielse (NCF)
- PDE-based applications and solvers at extreme scale **[SF]**
  David Keyes (Columbia U. + SciDAC TOPS)
- Climate Change Research at Exascale **[Ts]**
  Giovanni Aloisio, Italo Epicoco (U. of Salento), Silvia Mocavero (CMCC), Mark Taylor (SNL)

Whitepapers Summary

❺ Power and facilities

- Need more power-efficient devices, packaging, facilities
- Power-aware scheduling, …

- **MISSING**

❻ Resiliency and fault tolerance

- Exascale: will have to cope with continuous stream of failures
- Time to checkpoint longer than MTBF
- Vertical and transparent mechanisms needed

- Towards Exascale Resilience **[Pa]**
  Franck Cappello (INRIA), Al Geist (ORNL), Bill Gropp (UIUC), Sanjay Kale (UIUC), Bill Kramer (UIUC), Marc Snir (UIUC)
- Resilience and Fault Tolerance **[Pa]** **[PPT]**
  Al Geist (ORNL), Franck Cappello (INRIA)

Whitepapers Summary

❼ Systems and software integration

- New robust and more stable way of installing and maintaining HPC software stack needed
- Automated testing, change management, …

- A Collaboration and Commercialization Model for Exascale Software Research **[SF]**
  Mark Seager, Brent Gorda (LLNL)
- An Exascale Approach to Software and Hardware Design **[Pa]**
  William Kramer (NCSA), David Skinner(LBNL)
- Consistent Application Performance at Exascale **[Pa]**
  William Kramer(NCSA), David Skinner(LBNL)

Whitepapers Summary

❽ I/O

- File + directory abstractions scalable enough?
- New caching techniques, new tehnologies (e.g. SSD)

- Towards Exascale File I/0 **[Pa]** **[PPT]**
  Yutaka Ishikawa (U. of Tokyo)

❾ Interactions with external resources: clouds, archiving, real-time data streams

- Projection: 100s of exabytes of data generated by 2020

- Towards exascale distributed data management **[Pa]**
  Giovanni Aloisio (CMCC)

Whitepapers Summary

⑩ Meta / Process-related Papers

- Co-design
- Organisation

- The Application Perspective - Seeking Productivity and Performance **[SF]**
  David Barkai (Intel)
- A Proposal for a Capability Centers Consortium **[SF]**
  Bill Gropp, Mark Snir, (NCSA + UIUC)
- NSF IESP Whitepaper **[SF]**
  Abani Patra, Rob Pennington, Ed Seidel (NSF)
- Developing a high performance computing/numerical analysis roadmap **[SF]**
  Ann Trefethen, Nick Higham, Ian Duff, Peter Coveney, (UK)
- Co-design of Architectures and Algorithms **[Pa]**
  Al Geist (ORNL), Sudip Dosanjh (SNL)
- Early application development/tuning and application characterization/segmentation **[Pa]**
  Sanjay Kale (UIUC)

Whitepapers Summary

Multi-topic papers

- Major Computer Science Challenges at Exascale **[Pa]** ❶ ❷ ❹ ❻ ❼
  Al Geist (ORNL), Robert Lucas (ISI)
- Software Challenges of Extreme Scale Computing **[SF]** ❶ ❻ ❿
  Michael Heroux (SNL)
- Musings on the Path Toward Exascale **[SF]** ❻ ❿
  Robert Lucas (ISI/USC)
- Software and Exascale Computing **[SF]** ❶ ❸ ❻ ❾
  Bill Camp (Intel)
- IESP Challenges: Application Development, Translation and Execution Environments ❶ ❷ ❸
  Barbara Chapman (U. of Houston), Richard Graham (ORNL), Barney Maccabe (ORNL), Oscar Hernandez(U. of Houston), Bernd Mohr(JSC), Wolfgang E. Nagel(TUD-ZIH)

What's next?

- Write IESP roadmap!

- SC 2009, Portland, OR
  - **BOF "International Exascale Software Program"**
    Wed, Nov 18, 17:30 – 19:00
  - **Panel "The Road to Exascale: Hardware and Software Challenges"**
    Fri, Nov 20, 10:30 – 12:00