JÜLICH
FORSCHUNGSZENTRUM

Trans-Atlantic Performance Tool Integration

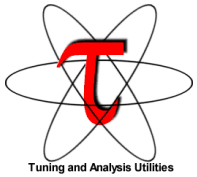
2014-06-22 | Markus Geimer
Jülich Supercomputing Centre
m.geimer@fz-juelich.de

Challenges of increasing parallelism

- Optimization of applications more difficult
  - More CPUs / multi-core / heterogeneity
  - Every doubling of scale reveals new bottlenecks
- Also new demands on performance tools
  - ***Scalable*** to keep up with HPC systems and applications
  - ***Efficient*** to meet performance expectations
  - ***Effective to use*** so that programmer productivity is maximized
  - ***Interoperable*** to perform different analyses on a single measurement data set

HPC performance tool landscape

- **TAU**
  - University of Oregon, US
  - http://tau.uoregon.edu



- **Scalasca**
  - JSC/GRS, Germany
  - http://www.scalasca.org

- **HPCToolkit**
  - Rice University, US
  - http://hpctoolkit.org

- **Open|SpeedShop**
  - Krell Institute, US
  - http://www.openspeedshop.org

- **Extrae/Paraver**
  - BSC, Spain
  - http://www.bsc.es/paraver

- **Periscope**
  - TU Munich, Germany
  - http://www.lrr.in.tum.de/~periscop/

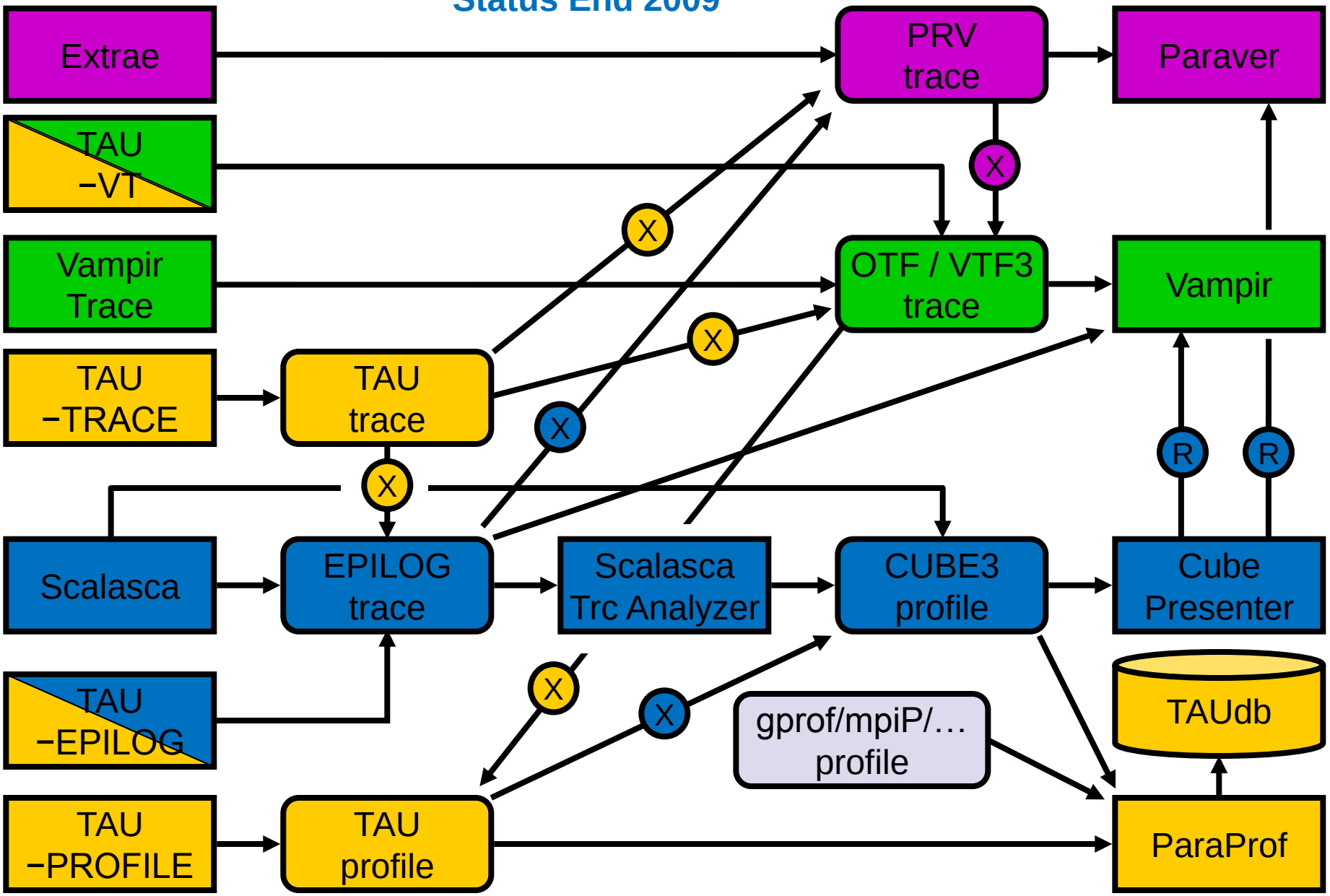
- **VampirTrace/Vampir**
  - TU Dresden, Germany
  - http://www.vampir.eu

- **[Score-P]**
  - JSC/GRS, TUD, TUM, RWTH, UO
  - http://www.score-p.org

Scalasca ⇔ TAU ⇔ Vampir ⇔ Paraver

JÜLICH FORSCHUNGSZENTRUM

Status End 2009



Integration projects

- SILC (2009-2011)
  - Develop a common, next-gen measurement infrastructure for Scalasca, Vampir and Periscope incl. open data formats
  - Partners: JSC, GRS, TUD, TUM, RWTH, GNS

- PRIMA (2009-2013)
  - Reengineering core components of TAU and Scalasca
  - Tighter integration and improved interoperability
  - Partners: UOregon, JSC

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Joining forces

- Partners of both consortia quickly agreed to collaborate
  - Recognized overlapping goals & benefits from sharing resources, ideas, and designs
  - Key: Early communication

- Funding agencies were open-minded, too
  - BMBF approved UOregon as associated partner in SILC
  - DOE did not object either

The Score-P idea

- Start a community effort for a common infrastructure
  - Score-P instrumentation and measurement system
  - Common data formats OTF2 and CUBE4
  - Scalability target: Petascale systems
- Developer perspective:
  - Save manpower by sharing development resources
  - Invest in new analysis functionality and scalability
  - Save efforts for maintenance, testing, porting, support, training
- User perspective:
  - Single learning curve
  - Single installation, fewer version updates
  - Interoperability and data exchange

Organizing the collaboration

- Phone conferences
  - Monthly status telecons
  - Many developer telecons scheduled on demand
- Joint project meetings
  - 2x per year, 2-3 days, incl. developer discussions
  - Typically attached to a conference or workshop
- Multiple research visits
  - JSC ⇨ UOregon, up to 3 months
  - Between partners in Germany, typically 1 week

Essential: Common infrastructure

Mailman	• Mailing lists
subversion	• Version control • Commit hooks
trac	• Integrated SCM & Project Management
bitten	• Continuous integration
Review Board	• Code review
Uncrustify	• Consistent code formatting
GNU autotools (patched)	• Build- and packaging system

Review Board

Essential: Common procedures

- Documented in project wiki
  - E.g., coding guidelines, naming conventions, branching & merging procedure, …
- Example: Designs drafted in wiki, announced via mailing list
  - Allows everyone to review and comment
  - Continuously updated to address comments
    - Trac wiki keeps history of revisions and changes
  - Implement in feature branch once consensus is reached
  - Code review before reintegration into trunk

Continuity: Governance model

- Tedious to define, but necessary
  - Loosely based on Qt governance model
- Specifies roles & responsibilities
  - User, contributor, maintainer, consortium partner, …
- Defines procedures
  - How to contribute?
  - How are roles assigned/revoked?
  - How to become a partner of the consortium?
  - How to resolve conflicts?
  - Voting rules

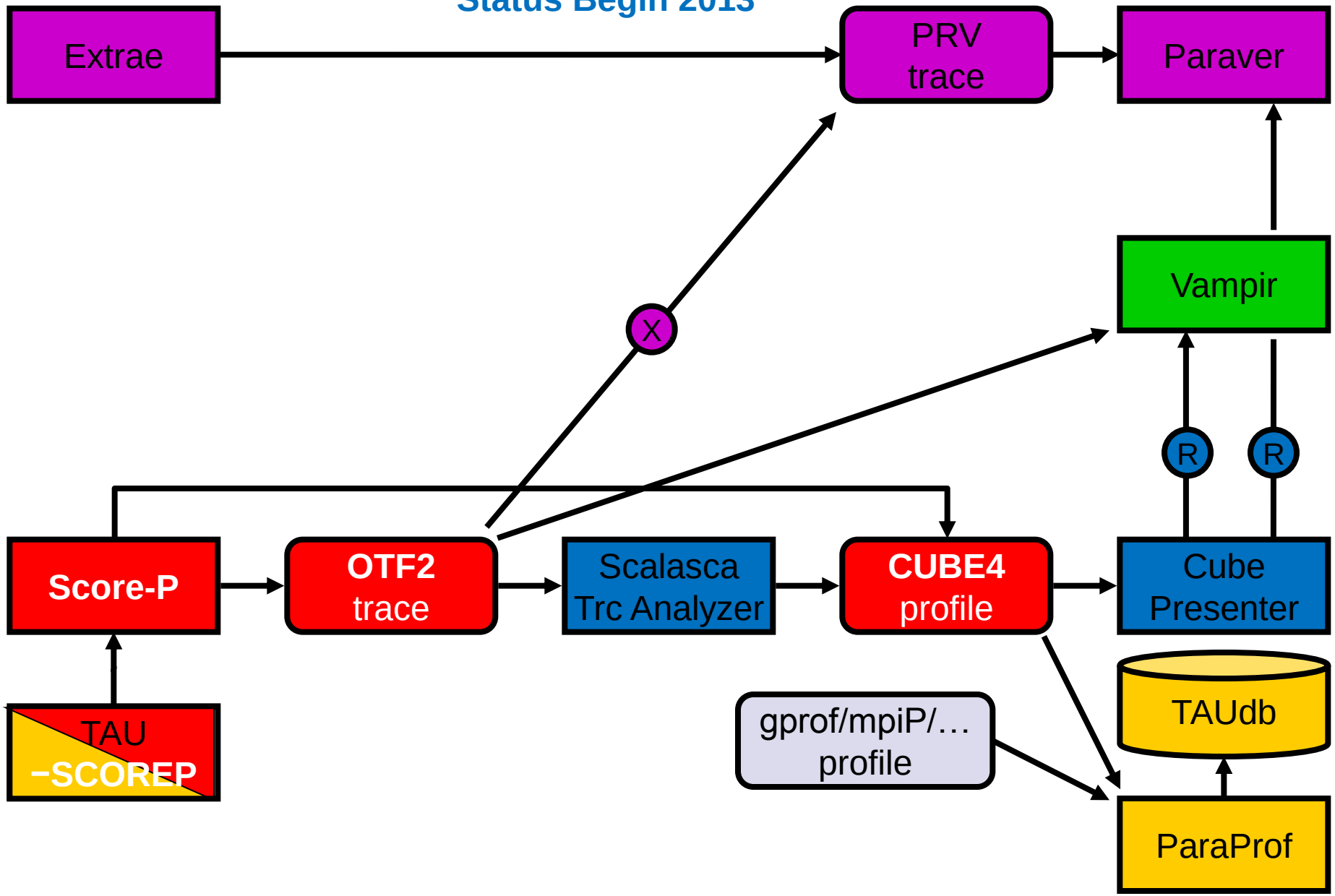
Score-P features and status

- Open source (BSD 3-clause license)
- Fairly portable
  - IBM Blue Gene, Cray XT/XE/XK/XC, Fujitsu FX10 & K computer, SGI Altix, IBM SP & blade clusters, Linux clusters
- Supports common parallel programming paradigms & languages
  - Fortran, C, C++
  - MPI 2.2, OpenMP 3.0, POSIX threads, SHMEM, CUDA, and combinations
- Generation of call-path profiles and event traces using common file formats
- Current release: 1.3rc1 (June 2014)

http://www.score-p.org

JÜLICH FORSCHUNGSZENTRUM

Status Begin 2013



Follow-up projects

- LMAC (2011-2014, BMBF)
  - Extensions for performance dynamics
  - SILC consortium + UOregon (associated)
- Score-E (2013-2016, BMBF)
  - Extensions for optimizing energy consumption
  - SILC consortium + UOregon (associated)
    + 2 further associated partners
- PRIMA-X (2013-2016, DOE)
  - Explore extensions towards Exascale
  - UOregon + GRS/JSC

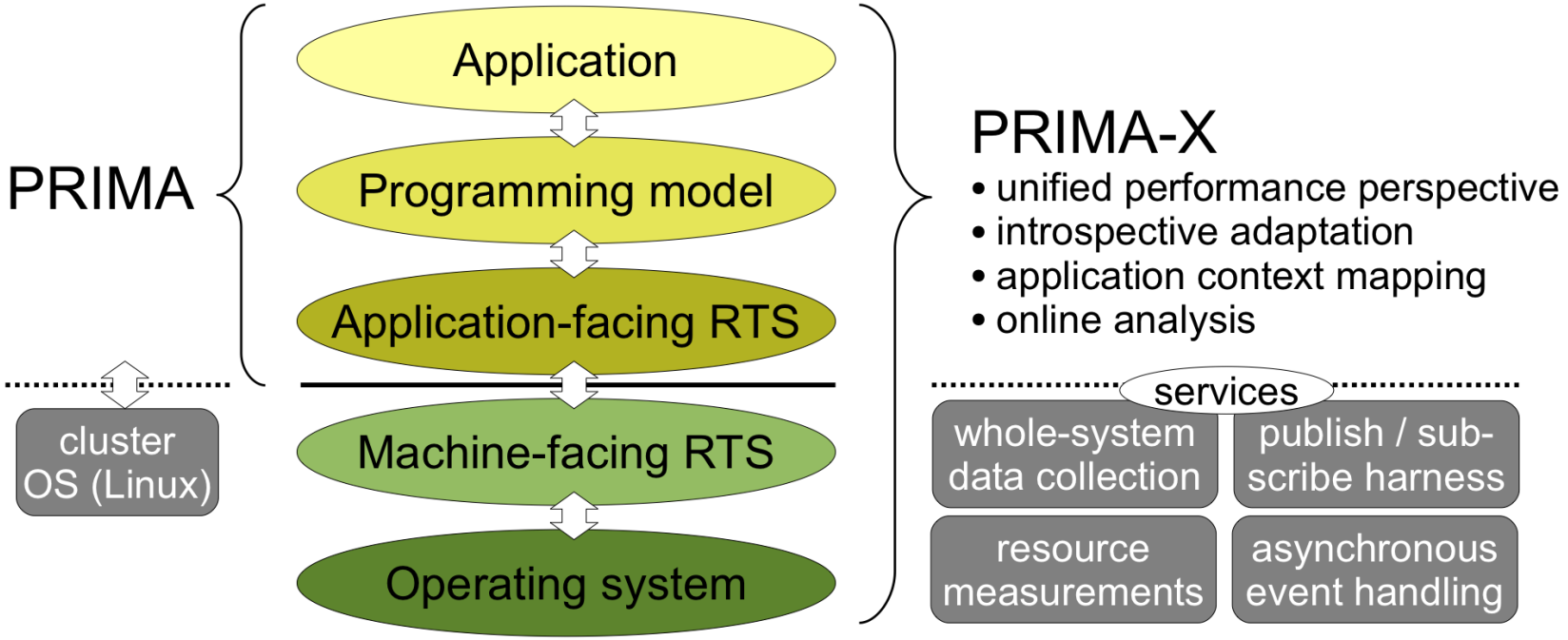
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

PRIMA-X

- *Performance Retargeting of Instrumentation, Measurement, and Analysis Technologies for Exascale Computing: the PRIMA-X Project*
  - DOE ASCR project (renewal of the successful PRIMA project)
  - 1 September 2013 – 31 May 2016
- Goal: Develop a performance analysis system for exascale applications targeting 3 key challenges:
  - Scalability
  - Dynamism
  - Heterogeneity
- Focus is on *application-facing* research and technology for exascale performance measurement and analysis

PRIMA-X Relation to Exascale SW Stack

- PRIMA-X extends PRIMA's 1st-person performance observation with 3rd-person resource measures
- High-level perspective by associating performance information across the exascale SW stack



PRIMA-X – Objectives and Results

- Objectives
  - Extreme scalability
  - Resilient performance metrics
  - Introspective adaption
  - Holistic perspective
  - Sustainable design
- Year 1 results
  - Prototype of a scalable observation system (SOS)
  - Hybrid parallel profiling
    - Integrated event-based sampling and direct measurement
  - In-situ analysis of performance data
  - Support for energy profiling
  - Integrated runtime from Score-P and Score-E projects

Other tool-related int'l collaboration actions

- Virtual Institute – High-Productivity Supercomputing
  - 9 EU partners + UOregon + UTK ICL + LLNL
  - Multi-day "bring your own code" tuning workshops (16 so far)
- Co-organized international workshops
  - Dagstuhl (2002, 2005, 2007, 2010, 2014)
  - Also: Participation in CScADS (2007-2011)
- Tool tutorials with UO, TUD, ICL, TUM (SC'09, SC'10, SC'13, ISC'14)
- Tool tutorials with LLNL (SC'11, ISC'12, SC'12, ISC'13, SC'13)
- Summer internships with LLNL (5 interns so far)
- Tool development (Scalasca, Score-P) for K computer
  - Close collaboration with Fujitsu + RIKEN

Summary / Conclusion

- Nowadays, a *single* tools group will struggle to keep up…
  - Collaboration is indispensable!
- However:
  - One needs to be willing to compromise
  - Requires to give up “not invented here” attitude
- Essential:
  - Common infrastructure
  - Clearly defined procedures
- Cross-national or coordinated funding helps a lot!
  - To get collaboration/community projects started
  - To sustain development with a common focus

Thank you!

Talk A. Knüpfer: “Score-P & Friends: Scalable & Versatile Parallel Performance Analysis with Periscope, Scalasca, TAU & Vampir”
Wednesday, 2:15 pm – 2:45 pm, Hall 5

Visit the GCS/JSC booth #940