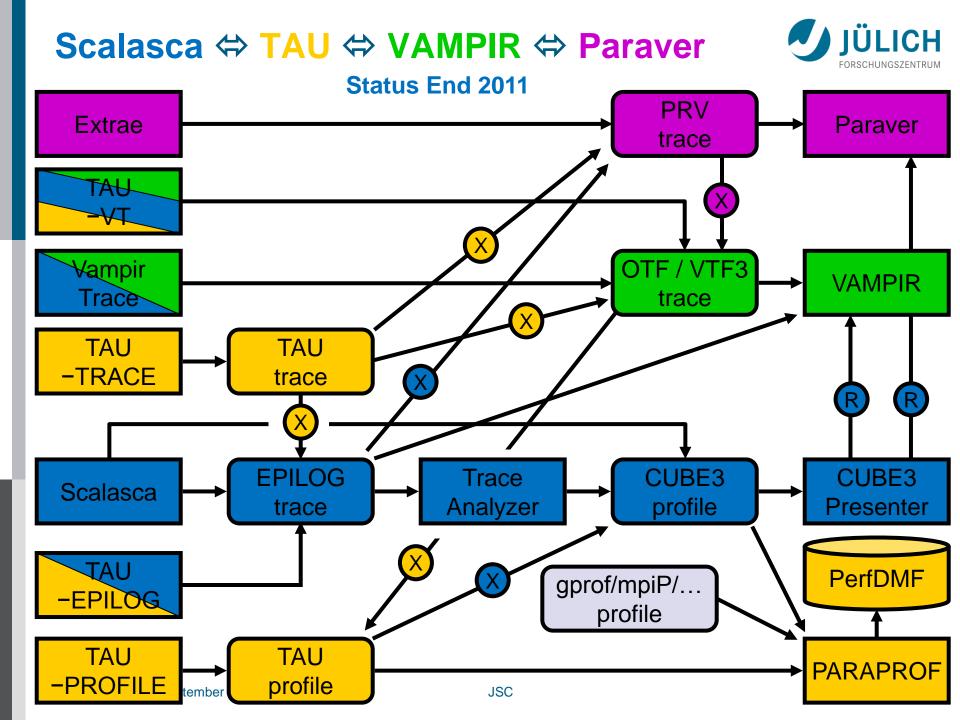JÜLICH
FORSCHUNGSZENTRUM

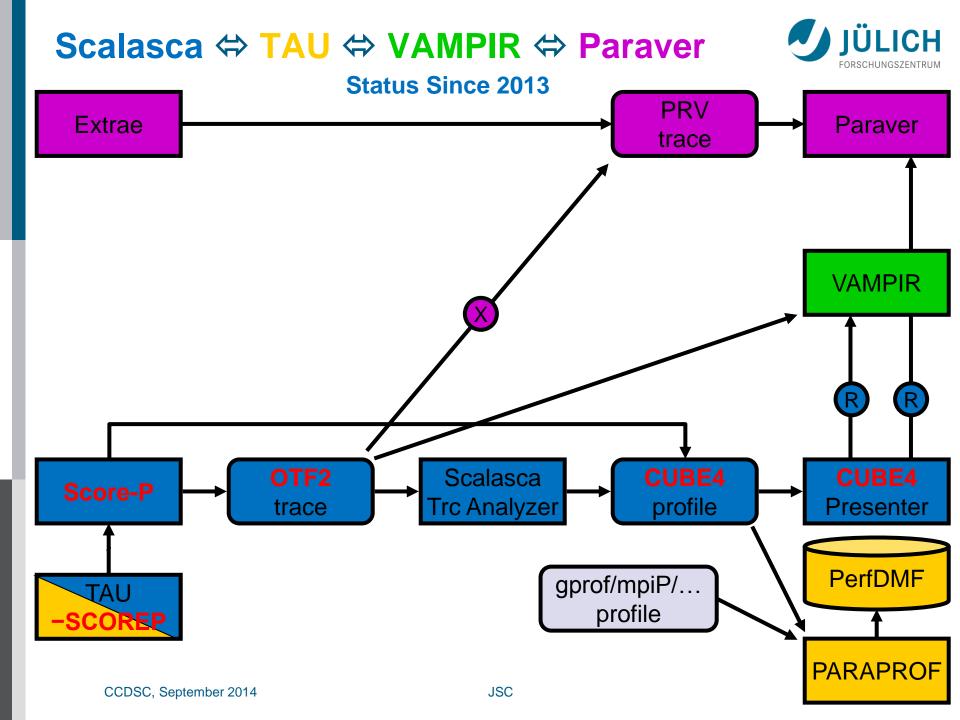CCDSC 2014

The Score-P Tool Ecosystem

Bernd Mohr

Jülich Supercomputing Centre

Integration

- Need integrated tool (environment)
  for all levels of parallelization
  - Inter-node (MPI, PGAS, SHMEM)
  - Intra-node (OpenMP, multi-threading, multi-tasking)
  - Accelerators (CUDA, OpenCL)

- Integration with **performance modeling and prediction**

- No tool fits all requirements
  - **Interoperability of tools**
  - Integration via open interfaces

Scalasca ⇔ TAU ⇔ VAMPIR ⇔ Paraver

Status End 2011

Scalasca ⇔ **TAU** ⇔ **VAMPIR** ⇔ **Paraver**

Status Since 2013

Score-P Functionality

- Provide typical functionality for HPC performance tools
- Instrumentation (various methods)
  - Multi-process paradigms (MPI, SHMEM)
  - Thread-parallel paradigms (OpenMP, POSIX threads)
  - Accelerator-based paradigms (CUDA)
  - And their combination
- Flexible measurement without re-compilation:
  - Basic and advanced profile generation
  - Event trace recording
  - Online access to profiling data
- Highly scalable I/O functionality

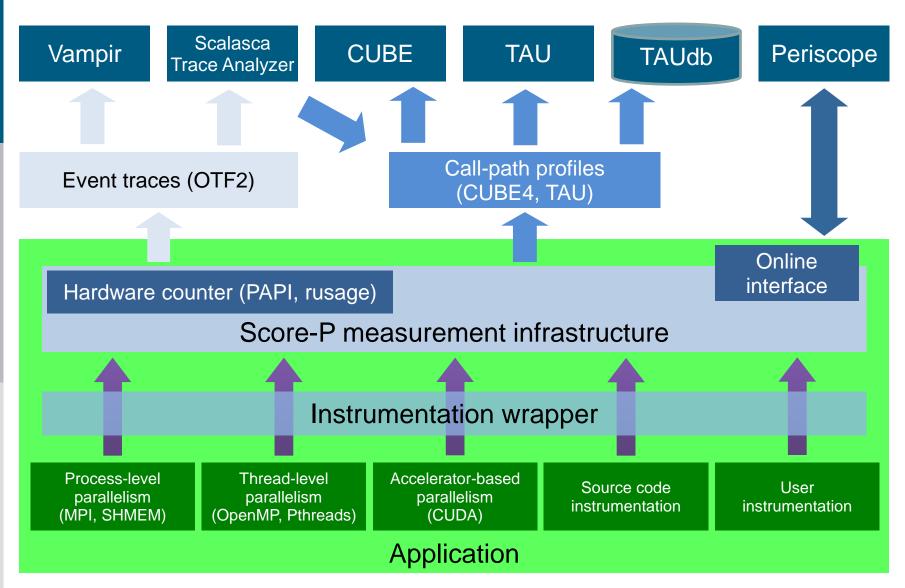- Support all fundamental concepts of partner's tools

Non-functional Requirements

- **Portability:** support all major HPC platforms
  - IBM Blue Gene, Cray X*, Fujitsu K/FX10
  - x86, x86_64, PPC, Sparc, ARM clusters (Linux, AIX, Solaris)
- **Scalability**
  - Petascale, supporting platforms with more than 100K cores
- **Low measurement overhead**
  - Typically less than 5%
- **Robustness and QA**
  - Nightly Builds, Continuous Integration Testing Framework
- Easy and uniform installation through **EasyBuild**
- Open Source: New BSD License

Score-P Architecture

Score-P Partners

- Forschungszentrum Jülich, Germany
- German Research School for Simulation Sciences, Aachen, Germany
- Gesellschaft für numerische Simulation mbH Braunschweig, Germany
- RWTH Aachen, Germany
- Technische Universität Dresden, Germany
- Technische Universität München, Germany
- University of Oregon, Eugene, USA

Past Funded Integration Projects

- **SILC (01/2009 to 12/2011)**
  - Unified measurement system (Score-P) for Vampir, Scalasca, Periscope
- **PRIMA (08/2009 to 10/2013)**
  - Integration of TAU and Scalasca
- **LMAC (08/2011 to 07/2013)**
  - Evolution of Score-P
  - Analysis of performance dynamics
- **H4H (10/2010 to 09/2013)**
  - Hybrid programming for heterogeneous platforms
- **HOPSA (02/2011 to 01/2013)**
  - Integration of system and application monitoring

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

ITEA 2
INFORMATION TECHNOLOGY FOR EUROPEAN ADVANCEMENT

MINISTRY OF EDUCATION AND SCIENCE
OF THE RUSSIAN FEDERATION

SEVENTH FRAMEWORK
PROGRAMME

Current Funded Integration Projects

- **Score-E (10/2013 to 09/2016)**
  - Analysis and Optimization of Energy Consumption

- **PRIMA-X (11/2014 to 10/2017)**
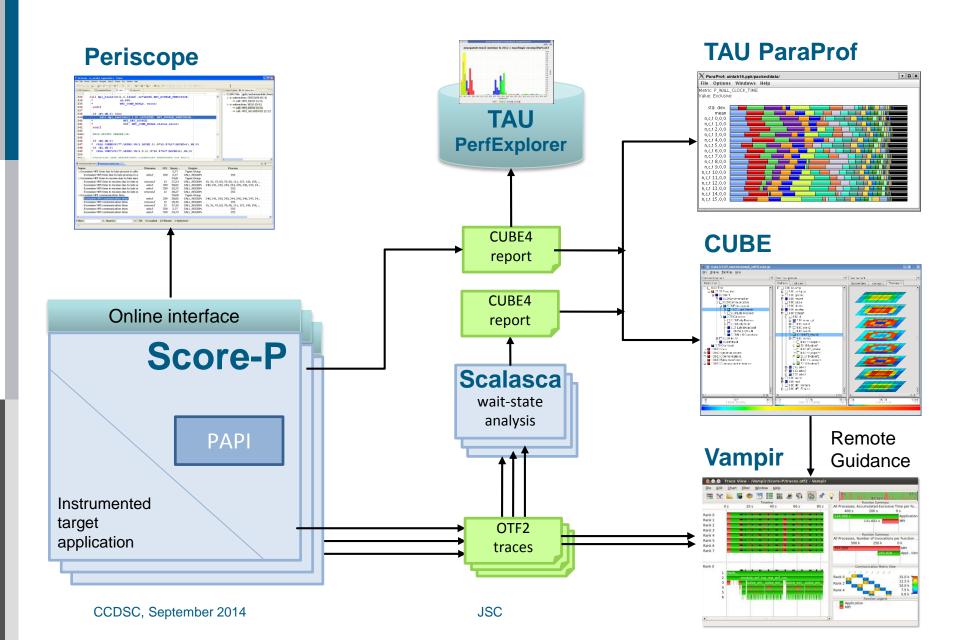 - Extreme scale monitoring and analysis

- **RAPID (04/2014 to 03/2015)**
  - Enhanced support for node-level programming models
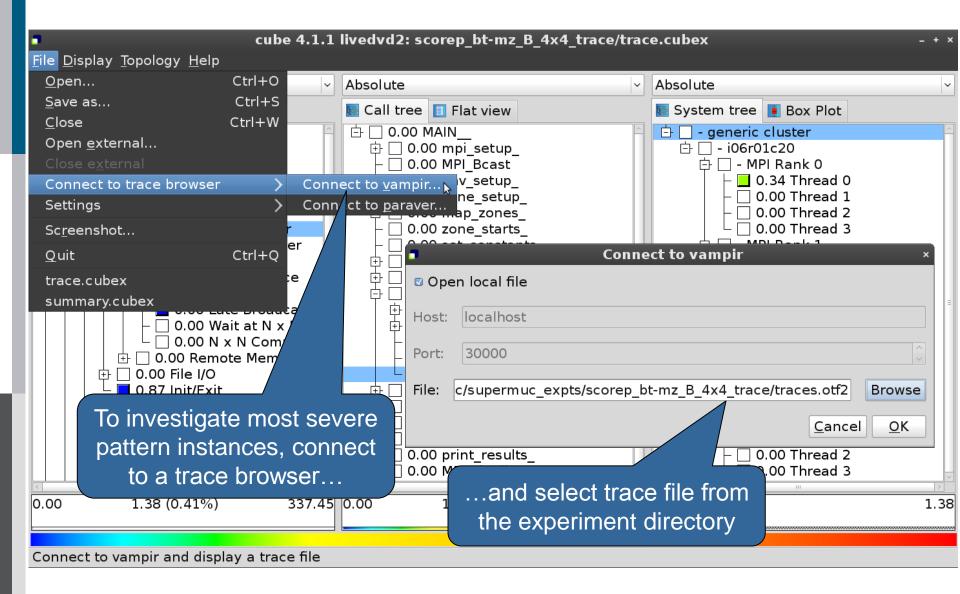   - POSIX, ACE, Qt threads
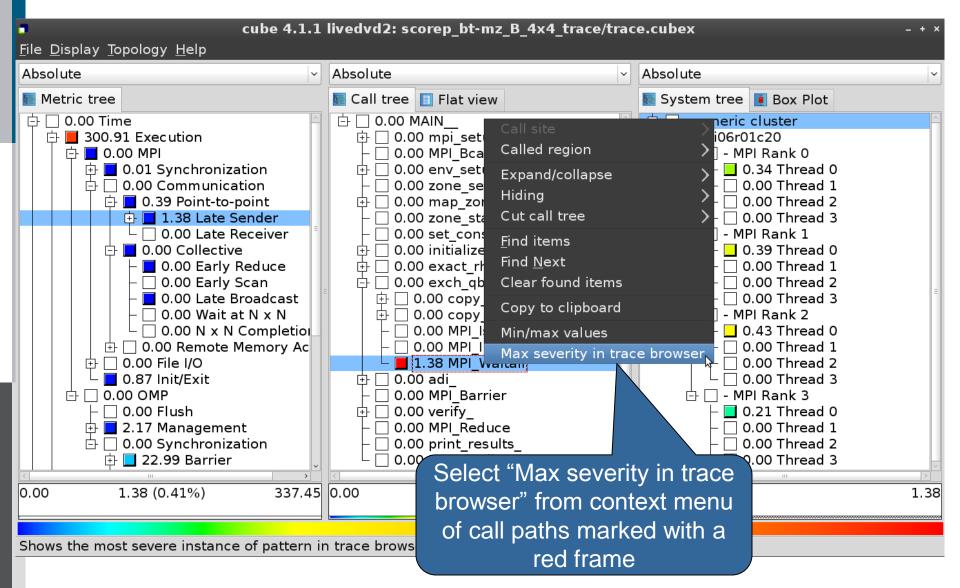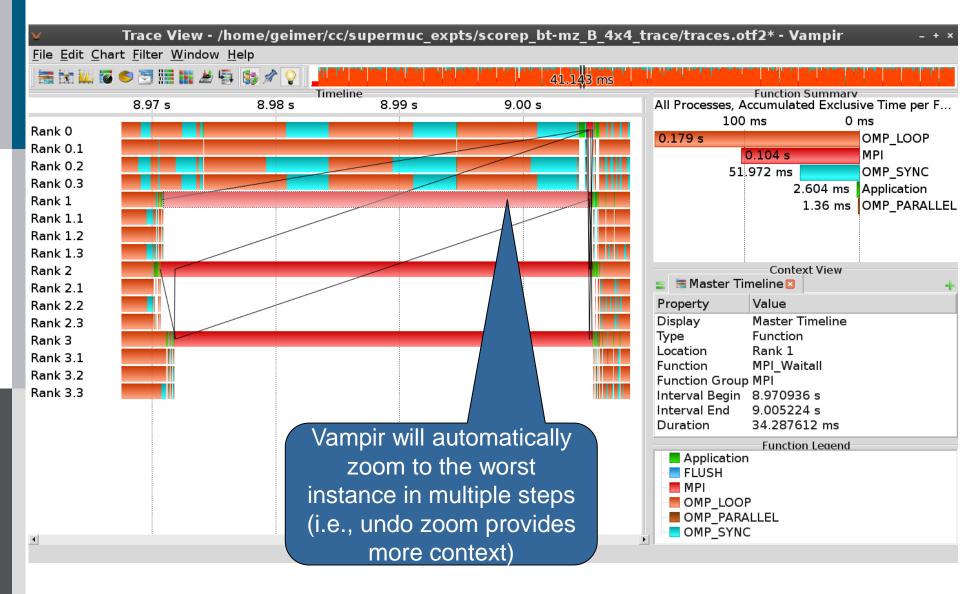   - MTAPI, TBB

The Score-P Tool Ecosystem

Periscope

TAU PerfExplorer

TAU ParaProf

CUBE

Vampir

Scalasca wait-state analysis

Online interface

Score-P

PAPI

Instrumented target application

CUBE4 report

CUBE4 report

OTF2 traces

Remote Guidance

Connect to Vampir trace browser

To investigate most severe pattern instances, connect to a trace browser…

…and select trace file from the experiment directory

Show most severe pattern instances

Select "Max severity in trace browser" from context menu of call paths marked with a red frame

Investigate most severe instance in Vampir

CCDSC, September 2014   JSC   14

Integration of Measurement and Modelling
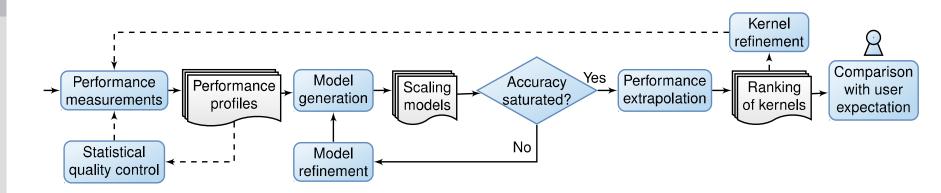
- Example: DFG SPPEXA Catwalk Project

Input

- Set of performance measurements (profiles) on different processor counts $\{p_1, \ldots, p_{max}\}$ w/ weak scaling
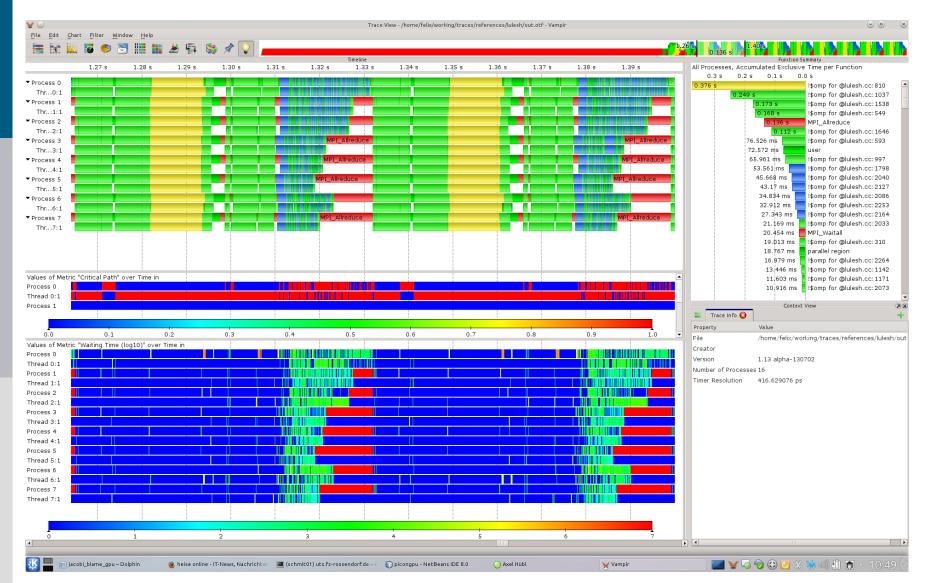- Individual measurement broken down by program region (call path)

Output

- List of program regions (kernels) ranked by their predicted execution time at target scale $p_t > p_{max}$
- Or ranked by growth function ($p_t \to \infty$)
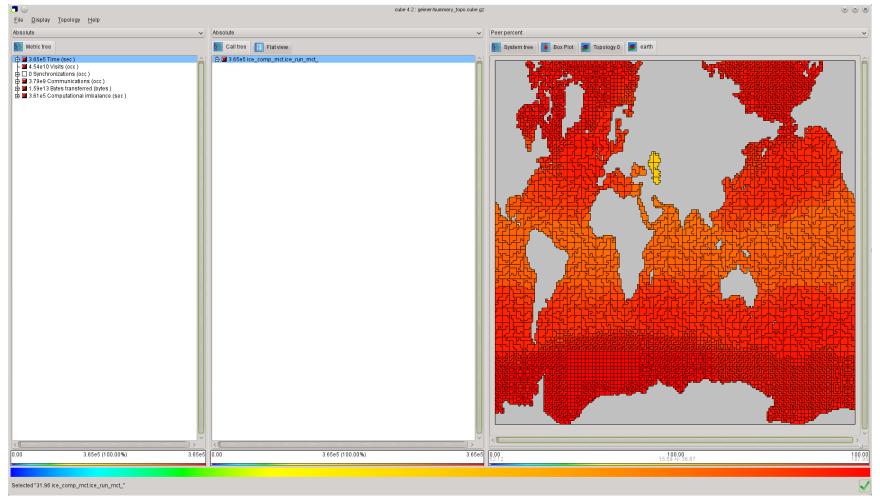
Showcase
Trace Visualization with Vampir

Showcase
Trace Visualization with Vampir

Showcase
Wait-State Analysis with Scalasca

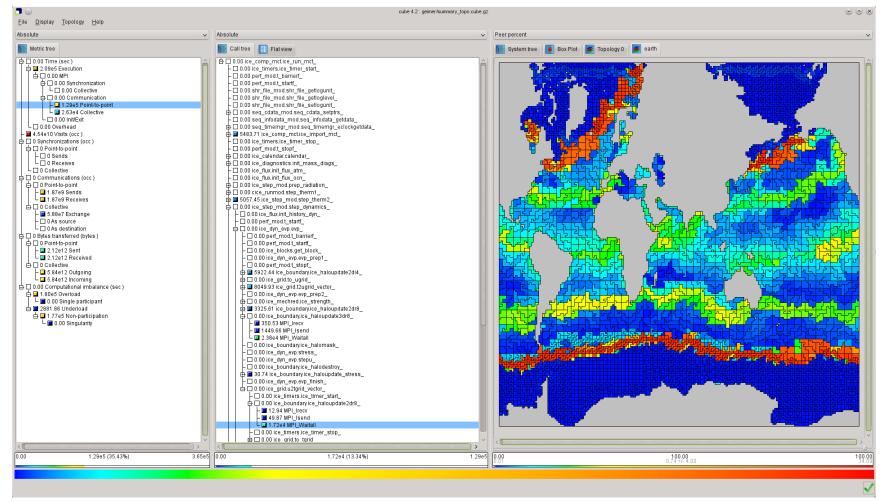
Courtesy John M. Dennis, NCAR
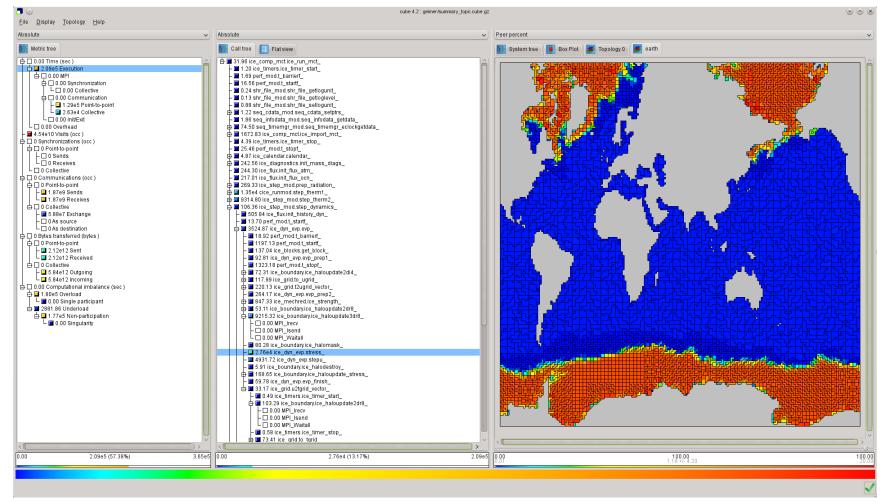
Showcase
Wait-State Analysis with Scalasca

Courtesy John M. Dennis, NCAR

Showcase
Wait-State Analysis with Scalasca

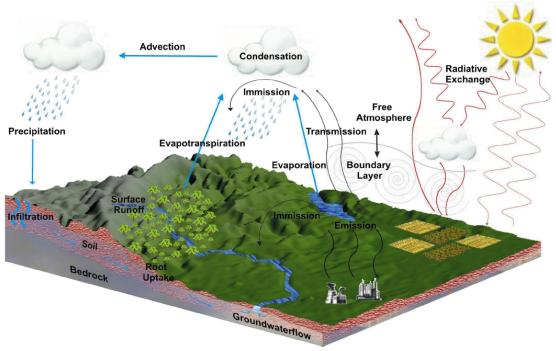Courtesy John M. Dennis, NCAR

Showcase: TerrSysMP

- Scale-consistent highly modular integrated multi-physics sub-surface/surface hydrology-vegetation atmosphere modelling system
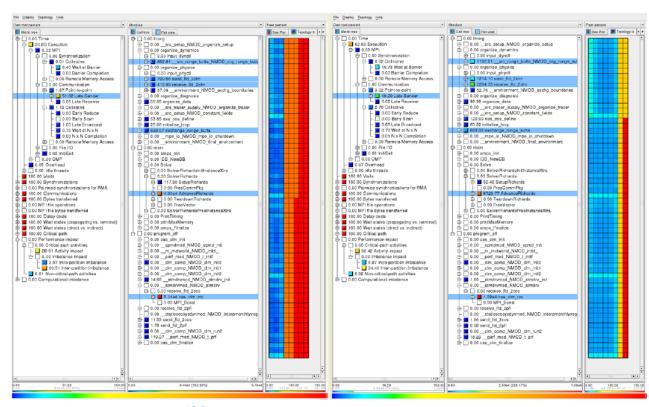

- Fully-coupled MPMD simulation consisting of
  - COSMO (Weather prediction)
  - CLM (Community Land Model)
  - ParFlow (Parallel Watershed Flow)
  - OASIS coupler

Success Story: TerrSysMP

- Identified several sub-components bottlenecks:
  - Inefficient communication patterns
  - Unnecessary/inefficient code blocks
  - Inefficient data structures

- Performance of sub-components improved by factor of 2!

- Scaling improved from 512 to 32768 cores!

Score-P Development

- **New features** of the current Score-P 1.3 relase
  - Instrumentation and monitoring of Pthreads and SHMEM parallel applications
  - Platform support for the K Computer and Fujitsu FX10 systems added

- **Features under development**
  - Sampling as less intrusive alternative to instrumentation
  - Support for the OpenMP tools interface
  - Handling of new MPI 3 functions
  - Support for the Intel MIC (native and offload model)
  - Recording MPI, user, and hardware topologies

Questions?

- Check out
  http://www.score-p.org

- Or contact us at
  support@score-p.org