Computer Science Research at CC-IN2P3

F. Suter

IN2P3 Computing Center

IHEP Delegation – June 10, 2014

A Research Activity in CS, Why?

- (Re)Establish links with researchers in Computer Science
  - Kinda broken since the Data-Grid project
  - Common topics and issues
    - But different concerns, expectations, ways of working, ...
    - Research is not production and *vice-versa*
  - at local, national, and international levels
- Interact with other teams of the CC IN2P3
  - Give pointers to projects, teams, people
- Bring a different point of view on CC's activities
  - Coming from HPC
  - and as a researcher with different concerns, expectations, ...
- Initiate or participate to answers to Calls for Proposals
- Promote CC's activities in the CS community (resources, devels, ...)

An Activity to Develop

- No extra cost for the CC
  - Fully autonomous for 6 years, thanks to funded projects
  - Even contribute to the global budget and manpower
- Funded research projects are a good way to
  - Get extra manpower (at least)
  - Increase visibility
  - Foster collaborations
    - Within and outside the IN2P3 Computing Center
    - With research and industry
  - **Condition:** be coherent (or overlap) with regular activities
- Be proactive to face forthcoming challenges
  - Lead, not follow!
  - Requires upstream studies even on not production ready technologies

SimGrid: a Versatile Simulator of Distributed Apps

- Scientific approach of Large-Scale Distributed Systems simulation
- Propose ready to use tools enforcing methodological best practices

Scientific Instrument

- Validated, Scalable, Usable; Modular; Portable
- Grounded +100 papers; 100 members on simgrid-user@; Open Source
- Simulates real programs (using specific API), not models; C, Java, Lua

Scientific Object (and lab)

- Allows comparison of network and middleware performance models
- Experimental (but on par with SotA) Model Checker; Soon an emulator

Scientific Project for 15 years

- Collaboration between Inria / CNRS / University of Hawai'i

Supported by several funded ANR projects

The SONGS Project

- Big ANR project
  - 1.8M€, 4 years, 7 sites
  - 9 permanent FTE (21 people) + 8 funded FTE (16 positions)
- **S**imulation **O**f **N**ext **G**eneration **S**ystems
  - Better understanding of Data-Grids, P2P, HPC, and Cloud systems
  - Based on **SimGrid**
  - Declarations of interest by IBM and CERN
- At CC IN2P3
  - 50% of my own research time, DataGrid WP leader
    - Very active in the HPC work package too
    - also in charge of communication and dissemination
  - Involves the storage team
    - Engineer shared 60/40 between research and storage
    - Try a new way of intern collaboration

The Dimensioning Issue

The problem

- More and more sciences rely on computing infrastructures to produce scientific results
  - Large experiments, data deluge, fourth paradigm, ...
- They need more CPUs, more storage, and faster networks

$\Rightarrow$ Data centers have to increase their capacities

Common practices

- Rely on expertise of system administrators and/or users
  - $\Rightarrow$ Empirical decisions
- A true dimensioning story, here at CC-IN2P3
  - What should be the next upgrade of the parallel cluster?
    - More cores or a high performance network?
  - Users were asked for their preference $\Rightarrow$ Lack of objectivity

Dimensioning Through Simulation

Why

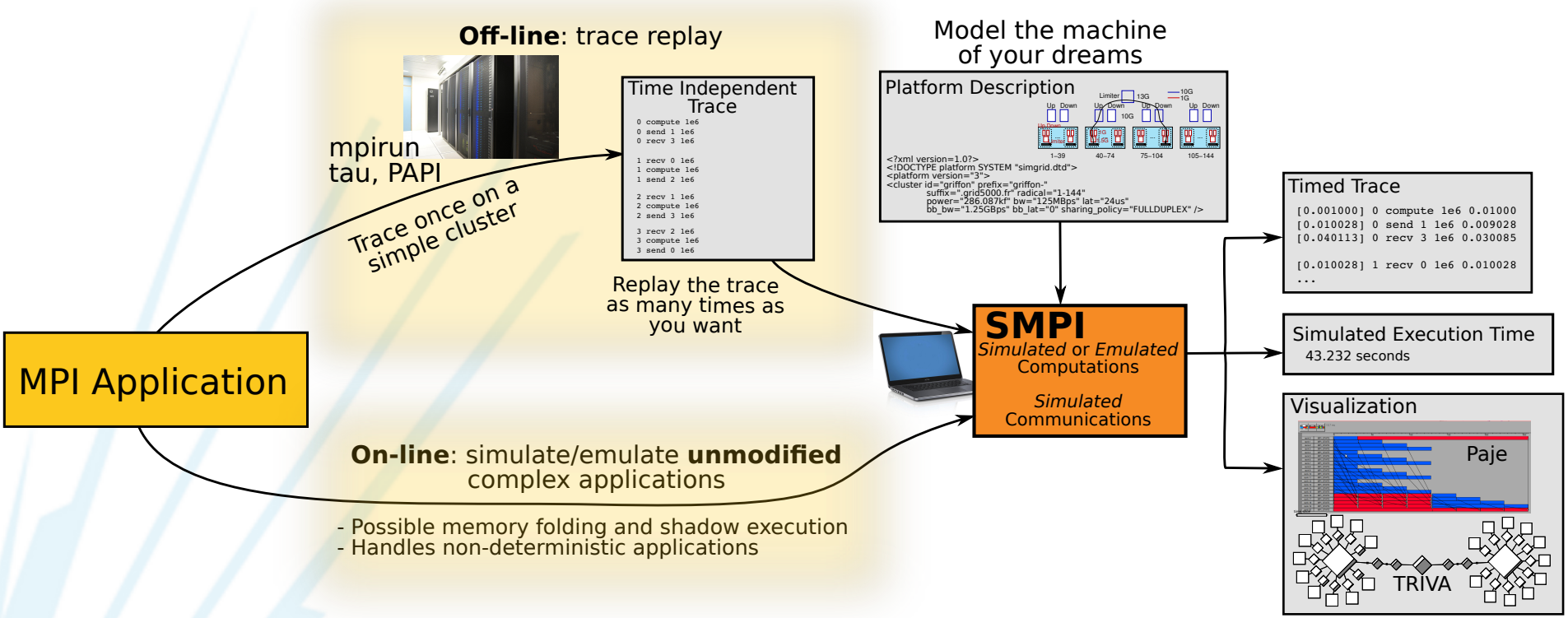
- User and administrator expertise is not enough and too subjective
- A bad desicion may have a high cost
  - Cannot wait for implementation to know the good/bad outcome

$\Rightarrow$ Need for objective indicators by exploring various "what-if" scenarios

How

- Simulation has many advantages
  - Less simplistic than theoretical models (which are useful too)
  - More reproducible than production systems (+ not disruptive)
  - Not as tedious, time/labor consuming than experimental platforms
- Focus on MPI applications
- Two complementary approaches in a single framework
  - Off-line: replay an execution trace
  - On-line: execute the application with some simulated parts

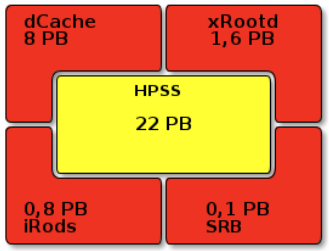
SMPI in a nutshell



Simulation of a Hierarchical Mass Storage System

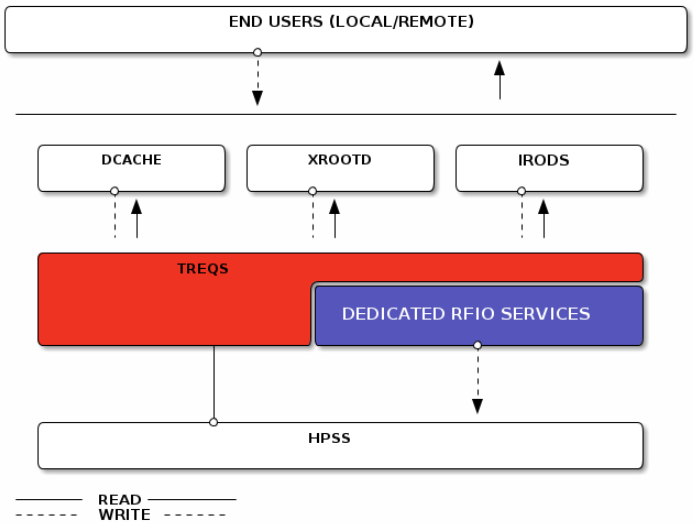
Activities

- Definition, design and implementation of SimGrid's storage API
- Preliminary versions of storage simulators



Roadmap

1. HPSS / RFIO
   - Clients ↔ Disks ↔ Tapes
2. DCache/XrootD/IRods ↔ HPSS / RFIO
3. TReqs ↔ HPSS / RFIO
   - Implementation of internal scheduler
   - Impact assessment
   - Study scheduling strategies
4. Users ↔ DCache/XrootD ↔ TReqs ↔ HPSS / RFIO
   - Black box vs. Hierarchical storage modeling



Concepts and Models for Storage

File

- Name (= full path) + size [+properties]
- Operations: open, close, seek, read, write, move, cp, rename, and delete
- no UNIX info, no contents, no navigation in tree

Storage Space

- Name + type + capacity + contents + mount point + model
- Operations: get contents and get [total, used, available] size(s)

Models

- linear w/ or w/o latency (for HDD and SDD)
- Matrix-based (for SE on Grids and Clouds)
- Stochatic (for tapes)
- Black box (for disk bay and file system)

"Workflowization" of SNFactory

Collaboration

- Saul Perlmutter's team + ACS Department at LBNL
- CC-IN2P3 + IPNL + Inria in Lyon

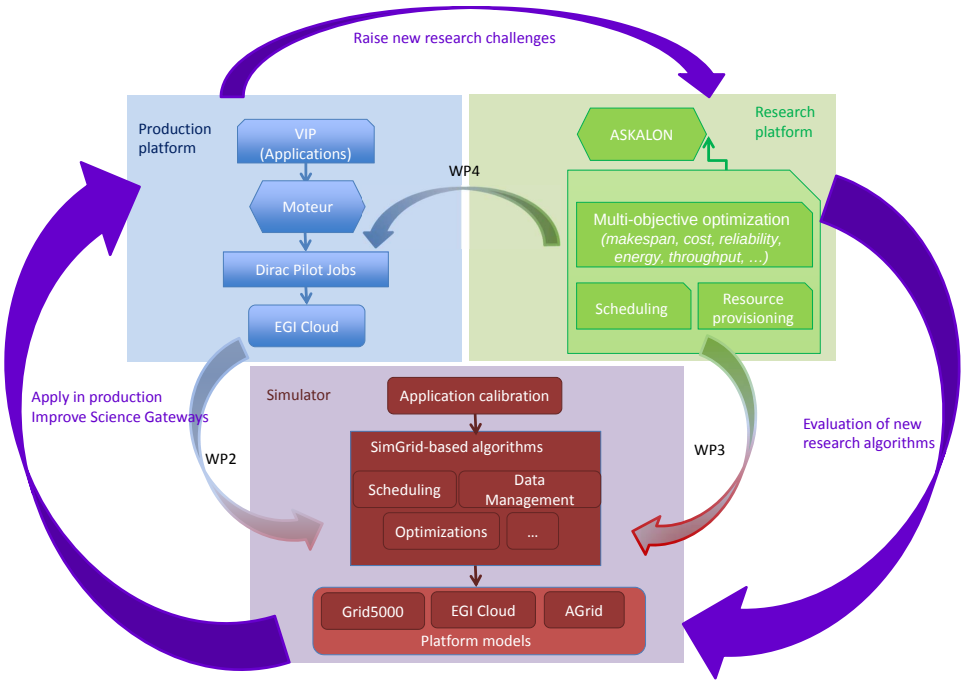
Objectives and method

- Optimize the execution and improve failure reports
  - Focus on the orchestration of the computations
- Method
  - Deconstruct the *plans* into elementary operations
  - Express the workflow in a standard format (DAX files)
  - Run it with a proper workflow engine (Pegasus)
  - Study scheduling strategies

RealSim

- Realistic Simulation of Distributed Computing Infrastructures for Scientific Experiments
- Collaboration with
  - CREATIS: VIP and Dirac/moteur
  - UIBK: Askalon and Scheduling

Closing the loop between research, simulated and production DCIs



- Application of SONGS results and methodology to production environments