

Monitoring plans @ DKRZ

Julian M. Kunkel

kunkel@dkrz.de

German Climate Computing Center (DKRZ), Research Group/Scientific Computing (WR)

SC 2015



Outline

- 1 Observations
- 2 Design
- 3 Design
- 4 Summary & Discussion

Motivation & Background

- With our new supercomputer we need a new monitoring system
 - Our previous monitoring solution was specific for Power6/AIX/Loadleveler
 - Roadmap: We must deploy (early) in 2016
- Existing solutions are inflexible/lack important information
- Previously, we also implemented capturing of statistics for e.g. SIOX, PIOvis

Observations

- With every system generation we have new monitoring tools
- Modifications to a tool are needed to cover (most) requirements
- Monitoring is mainly restricted to the system perspective
- For diagnostics, users often use their own app. monitoring
- In procurements, each vendor offers another solution with pro/cons
- Apparently other data centers also implement/extent their (own) tool
 - LDMS/OVIS (Sandia, Open Grid computing)
 - Total Recall (Gregor Mendel Institute, lambda: relient.systems)
 - TACC Stats (TACC)
 - ... [20+]

Design

Goals

- Flexible statistics monitoring for system and application level
- Portable: useful for our next supercomputer

Supported perspectives

- Data center/admin perspective
 - Foster understanding of hardware/software factors and diagnostics
- User perspective
 - Provide a first idea about application performance
 - Helps diagnosing system issues if the job does not perform
- Developers perspective
 - Support performance regression testing
 - Allow to provide arbitrary application metrics
- Project/Consortium leader
 - Show how the system has been used
 - Provide hints about app performance

Covered Information

System side

- Relevant hardware: Storage, tape, compute, etc.
- CPU counters (e.g. via periodic sampling)

Application side

- Job information
 - Resource manager e.g. Slurm
- Support for application phases
- User/application-specific metrics!
 - Could be performance metrics e.g. MLUPS
 - Could be a model diagnostics statistics e.g. avg. temperature
 - Apps instrumented via Darshan or SIOX could supply new metrics

All information should be available in customizable views

Some Requirements

Functional

- Ingesting of application information (resource manager)
- Lightweight API for application phases
- Statistics should be able to be read by tools to enable prescriptive actions
- Configurable sampling frequency (at least upon job start)
- Avoid interaction with application monitoring/tracing tools
- Export of relevant data views for long-term inspection
 - Allow third-party tools to digest data, e.g. via. R

Non-functional

- Low overhead
- Efficient storage of data
- User-friendly management of metrics and views
- Independent to other HPC tools (e.g. resource manager)

Discussion: Conventions/Standards and More

- Assume a vendor could be told to support the cluster monitoring convention “X”
- Any tool(s) that implement it could support the same API
 - Information of all relevant hardware components
 - Correlation to application jobs
 - Plus the features to easily plug in your own metrics

Design Status

- Only a few thoughts, because we are looking forward to collaborations
 - Joint discussions, benefit to everyone
 - 50% more work but portable!
- We will collaborate with everybody that is interested
 - Integrate with OpenHPC by the Linux foundation?
 - Yesterday I learned a SIG is about to be established to tackle this topic
 - BoF: “Monitoring Large-Scale HPC Systems: Data Analytics and Insights”
 - I kept most slides for the sake of discussion

Design decisions

- Tool should work on (many) data centers
- Lightweight C/Fortran/Python/RESTful API to add metrics
- A flexible solution such as Graphite/InfluxDB as backend
- Additional instrumentation built on top of Python

Summary & Discussion

- DKRZ needs a holistic tool for system & application monitoring
- We don't want to implement AGAIN something for only us
 - We have discussed potential collaboration with some centers
 - We will try to contribute to standards/conventions
 - A standard helps preserving ANY investment
- Application metrics could be feed by existing tools such as Darshan & SIOX
- Statistics should be able to be read by tools to enable prescriptive actions
- Everything is on discussion
- Our project mailing list: <http://goo.gl/iTjzJN>