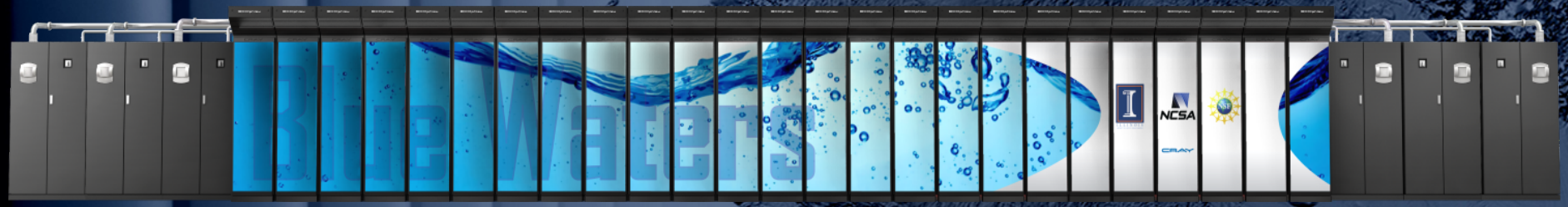
BLUE WATERS-GEO

MAPPING AND MODELING THE WORLD

Filesystem User Experience Monitoring

Mike Showerman



ILLINOIS
NCSA | National Center for Supercomputing Applications

NSF

CRAY

Best Practice: Filesystem User Experience Metrics

What is the Best Practice?

Measure the bahaviors of your filesystem components

Why is it needed?

It is difficult to evaluate the behavior of pasts events on the filesystem and impossible to reproduce

Who does it impact and when?

It enables support staff to identify issues before it becomes a user complaint

Why is this a Best Practice?

Enables us to be proactive in filesystem and user behavior influenced health

Challenges

- Parallel Filesystems in use can be chaotic
- Difficult to understand if the system is operating normally
- Behaviors are often the result of multiple concurrent events
  - Shared resource with contention from multiple sources
  - Striped file performance can be impacted by a single large unstriped access for example
- In the past, best indicators of non-obvious filesystem issue were sometimes user complaints
  - We strive to be proactive, complete

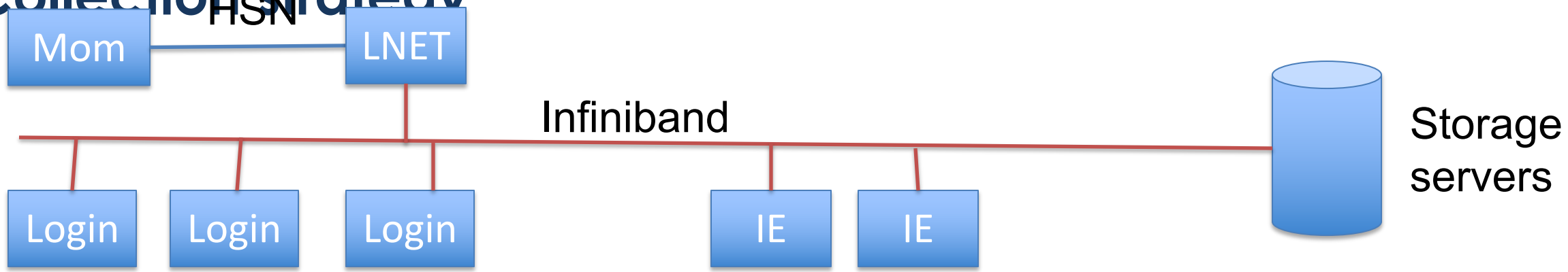
What is UEM

- UEM's goal is to be an "outside envelope" test that measures the filesystem interaction that users are actually experiencing and record for posterity
- Old practices have gaps that UEM encompasses:
  - monitor logs, monitor system and network loads, monitor I/O activity level, interactively sample ("works for me")
- UEM measures the latency to perform filesystem operations on each storage server component
  - 1 create (not used)
  - 2 write
  - 3 rmdir (not used)
  - 4 end (not used)
  - 5 single file create
  - 6 single file remove
- These active probes insert data into our database ISC with 1 minute intervals
- Automated tools as well as realtime interfaces

How the data is collected

- Measure only what is worth the cost
  - Don't measure every mounted client- just representative ones
    - Different hardware, software, networks, usage type, etc
  - Pre-create files on each storage server to avoid unnecessary MDS interaction, which is itself measured in "single file create"
  - Small actual i/o… we use 4k writes/reads
  - Targeting a specific file targets a specific server
- Use volkseti (lightweight text parser) to run the collector
  - Volkseti downloads a configuration file each time it is run
  - Configuration file is dynamically generated
    - Allows us to include the UEM test when desired
    - When we have a large pool and want a single collection, the first node to run volkseti in that class since the collection interval passes has a UEM run in its config. *No unique sampler node requirement!*
    - For Mom and Import/Export only one node will run the test each minute
    - Login nodes all collect at every interval due to high potential variation from user influence
- What about performance impact of probing?
  - Interference tests under various load conditions showed no measurable difference.

Collection strategy



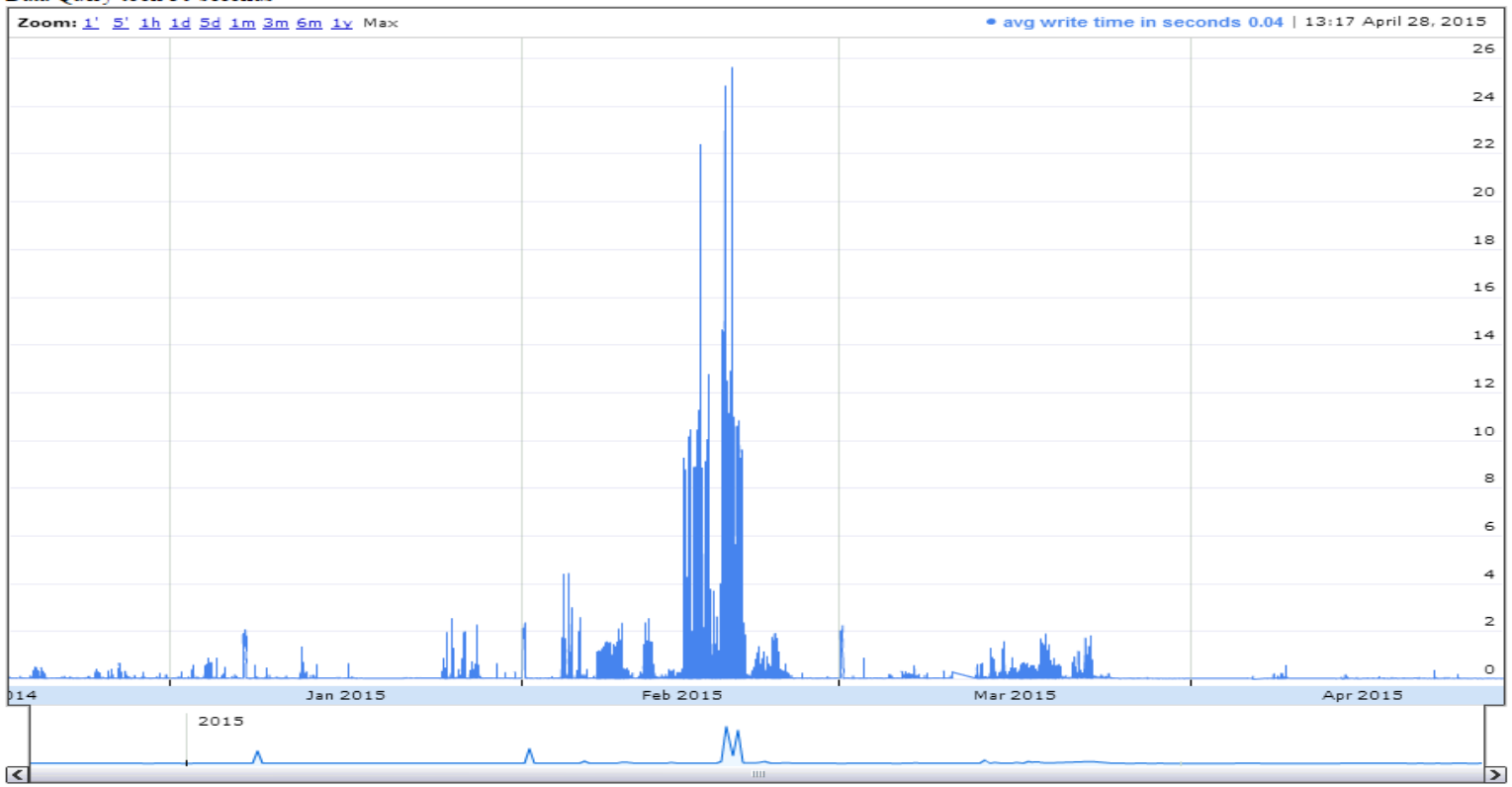
- Data is collected from 3 perspectives
  - Mom nodes to use the path through LNET routers over the high speed network then through Infiniband
  - Import/Export nodes to test direct Infiniband attachment and different hardware/software
  - Every login node to show the impact of multi-user activity
    - RPC exhaustion and such
- Differing results of UEM by perspective can rapidly reduce a problem to a probable source

How the data is received

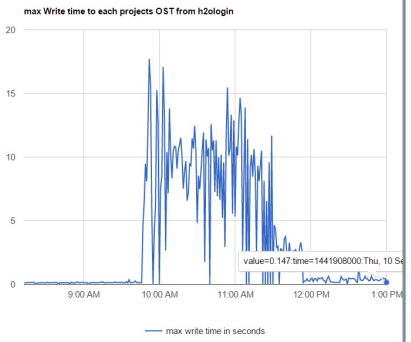
- The output from the UEM tests is directed to a socket on the database server
- Syslog-ng listens for the data and sends it to a parser/db injector
- PHP pages read the database and return results
  - For automated tests, the output is text
  - For interactive, the data is graphical

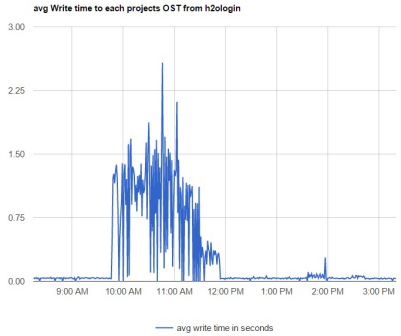
Crucial in finding the root cause of a problem

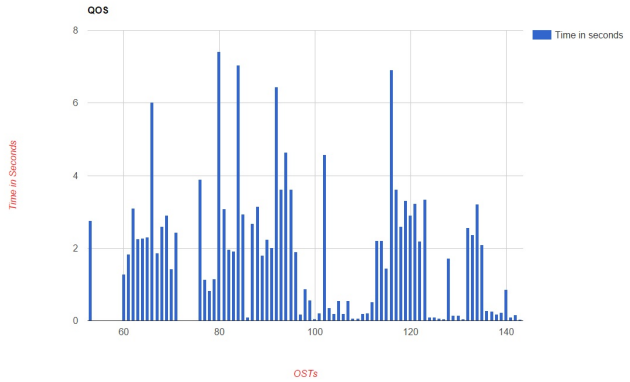
Displaying data for h2ologin
Test Number not set via testnum=, using 2 (write)
Displaying data for home filesystem
No duration Time set via length= , showing data through current time
Start=Mon, 01 Dec 2014 12:14:15 -0600
End =Tue, 28 Apr 2015 13:17:11 -0500
Data Query took 30 seconds

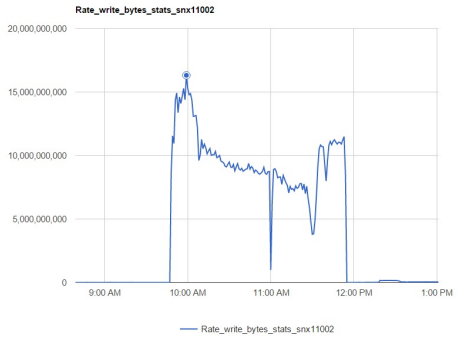


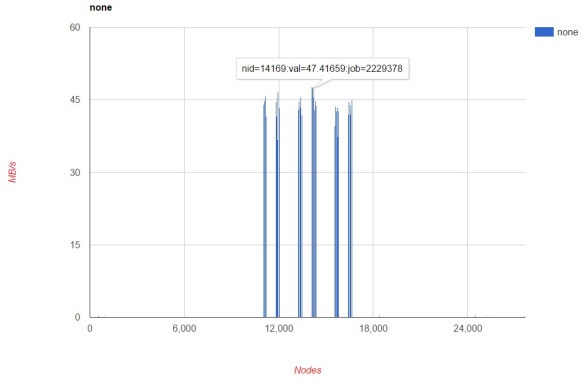
Use Exa

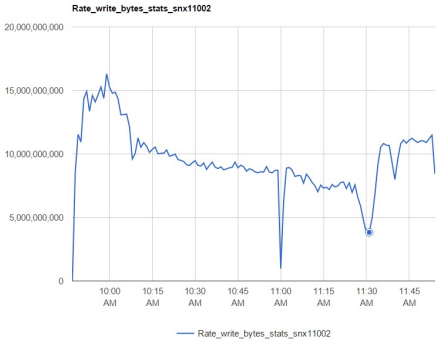












Top line is UEM data and is compared against node monitoring data below

Other Successes

- Slow OST was detected by being a persistent outlier with UEM
  - Had a drive failing that was causing performance issue but not beyond a set error threshold
- Rapidly identifying difference between node local slow downs vs full filesystem slow downs
- Full history for correlation with any reported job slowness
  - Helps rapidly rule-in or rule-out the filesystem early in many cases

What needs to be done

- Adapt and deploy at alternate site
- Evaluate application to alternate filesystems
- Improve or replace volkseti
- Evaluate use of IOping tool to perform measurement
- Develop methods for better integration of active probe data with traditional monitoring data