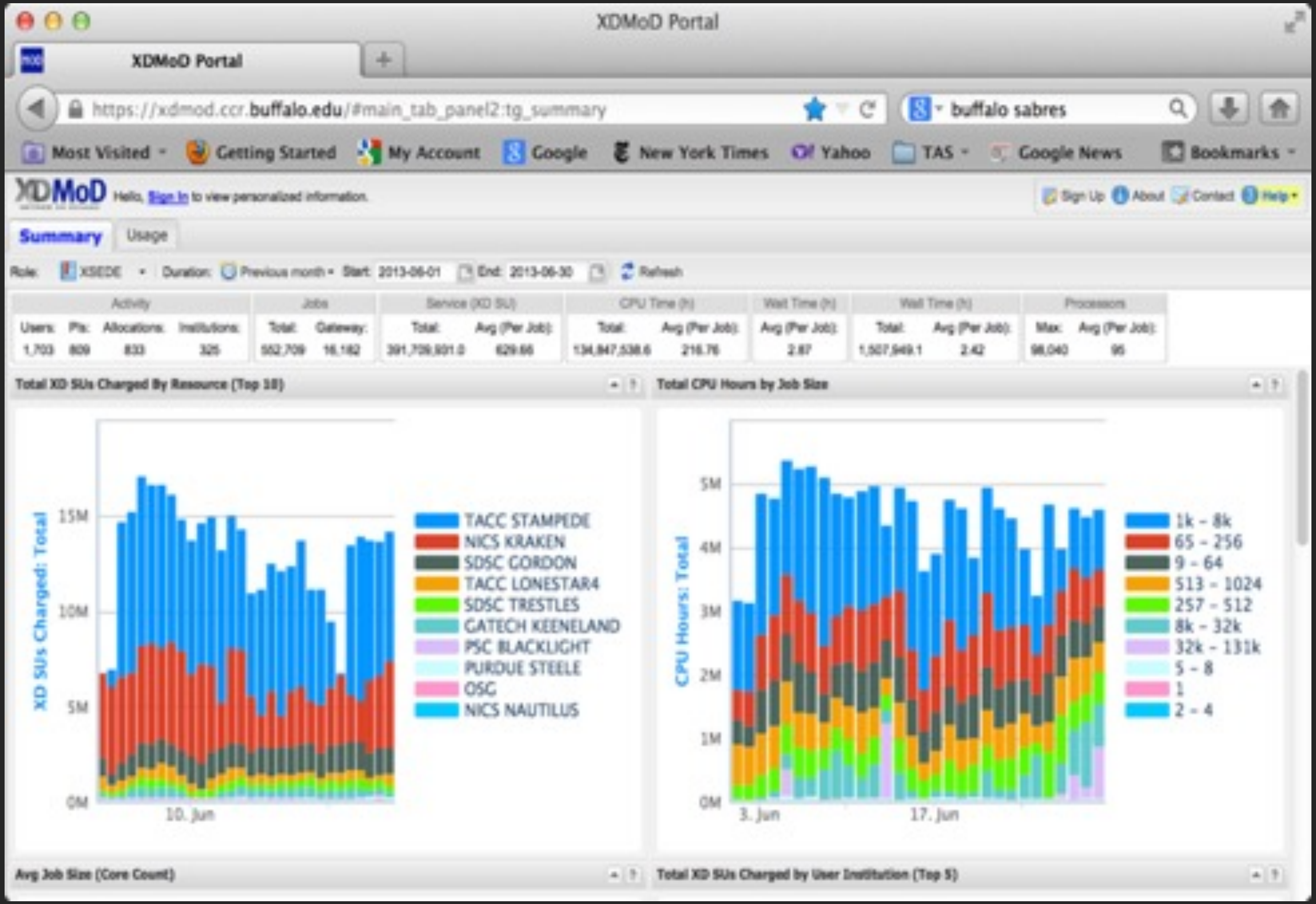
# MEASURING PERFORMANCE AND USAGE – EVOLUTION OF THE MEASURING AND MONITORING OF NSF SUPERCOMPUTING

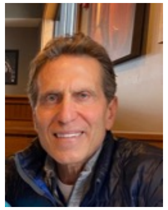

ABANI K PATRA

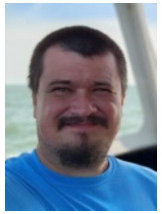
STERN PROFESSOR OF MATHEMATICS AND DIRECTOR, DATA INTENSIVE STUDIES CENTER

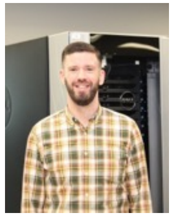
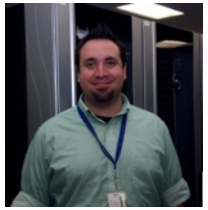
TUFTS UNIVERSITY, MEDFORD, MA 02155

ABANI.PATRA@TUFTS.EDU

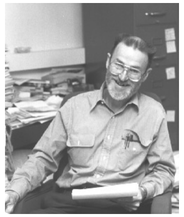


“The Buffalo Bulls”

T. Furlani R. de Leon, J. White, N. Simakov, C. Saeli, A. Weeden, G. Dean, R. Rathsam. S. Gallo, …




“The Texas Longhorns”

And many others over the years!

# OUTLINE

- History and Motivation
- A little story
- “Audit” and the Creation story of XDMOD
- Workload Analysis for Decision Making
- Advanced analysis –Anomaly detection, …
  - What about the end user
- ACCESS today

# SUPERCOMPUTERS AND MEASUREMENTS

- Computers measure everything!
  - User and Usage data
  - Performance data
  - System Logs
- Shared resource → Quotas and Allocations
- Data and analysis
  → code performance,
- Grants and Annual Reports

*Table A18.2. FY06 Total NUs Available and Delivered*

| | NUs | % |
|---|---|---|
| NUs Available | 2.23 billion | |
| NUs Requested | 2.96 billion | 132.7 |
| NUs Allocated | 1.917 billion | 86.0 |
| NUs Delivered | 1.281 billion | 57.4 |
| ***Usage by Category (% of NUs Delivered)*** | | |
| Academic (allocated) | 1.257 billion | 98.1 |
| Academic (discretionary) | 13.5 million | 1.1 |
| TeraGrid Staff | 10.2 million | 0.8 |

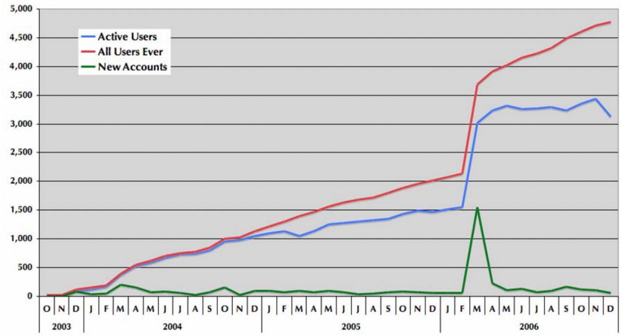



*Figure A18.3. TeraGrid Users and New Accounts*

Plus a dozen " science nuggets"

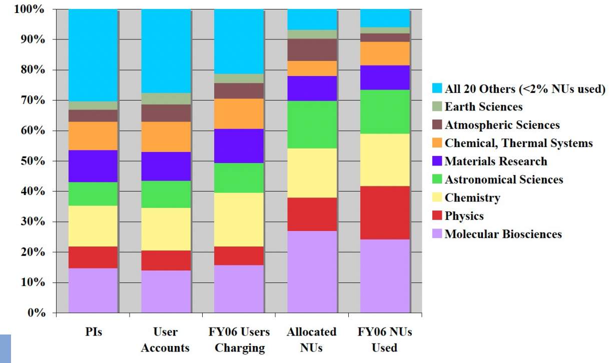



*Figure A17.1. TeraGrid Use by Discipline, FY06*

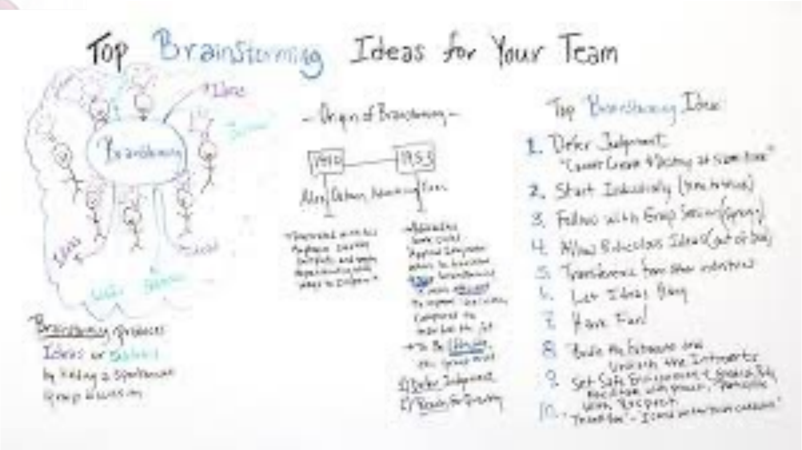
- Simple aggregations
- Hard to answer complex queries about usage, performance, ROI and plan

*Data from TG annual reports.*

1985 NSF Supercomputer Center → NPACI→TERAGRID → XSEDE → ACCESS  2025+

# A LITTLE STORY






- What do users actually do on these machines?
- Who does it?
- How well are they doing it?
- “Industrial Production” → Quality control?
- If you measure it they will make it better!

# QUALITY – DEMING …

**1.** Create constancy of purpose toward improvement of product and service …

**3.** Cease dependence on inspection to achieve quality. Eliminate the need for inspection on a mass basis by building quality into the product in the first place.

**10.** Eliminate slogans, exhortations, and targets for the work force asking for zero defects and new levels of productivity.

**11b**. Eliminate management by objective. Eliminate management by numbers, numerical goals. Substitute leadership.

**13.** Institute a vigorous program of education and self-improvement.

- Annual report and aggregate based management was not optimal
- Data collection needed to be made systematic and designed to elicit information on machine, use and policy
- Process audit based on ALL available data was crucial
- Need dedicated group with focus on measurement and monitoring

- **XSEDE → TAS**

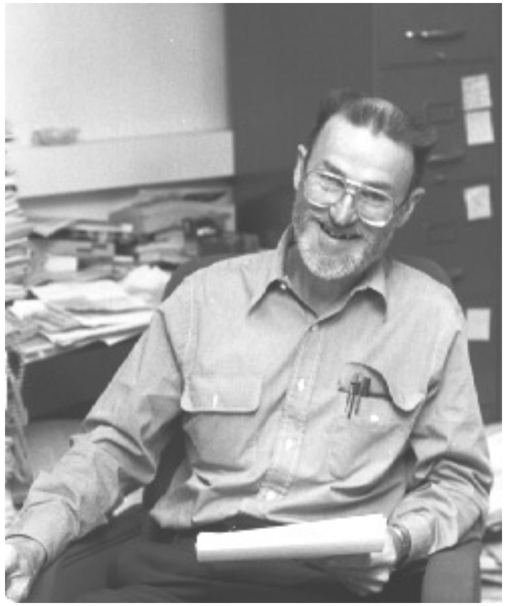
*It is wrong to suppose that if you can't measure it, you can't manage it – a costly myth." – Deming in New Economincs*

# STAKEHOLDERS

| | |
|---|---|
| 01 | PI and End User |
| 02 | Computational Scientist and Support Specialist |
| 03 | Systems Administrator |
| 04 | CI Center Director |
| 05 | NSF Managers and Senior Leadership |
| 06 | External Reviewers |

## WHAT DO THEY NEED/WANT?

- **PI and End User**
  - management, resource selection, application tuning, improved throughput
- **Systems Administrator**
  - System diagnostic and performance optimization, application tuning
- **Computational Scientist and Support Specialist**
  - Tool to facilitate work with end-users to improve job performance and throughput
- **CI Center Director**
  - Comprehensive resource management and planning tool
  - Return on Investment Metrics
- **External Reviewers**
  - Tool for data driven review for verification of best practices and project goals
- **NSF Senior Leadership**
  - Measure the effectiveness of supported programs
  - Inform deployment of future systems to fulfill unmet need



# PUTTING IT TOGETHER

- What are our users doing?  "My boss needs to know"
  - UBMOD → XDMOD
- Supercomputer Performance?
  - TACCSTATS
- How well are they doing it?

# A Workload Analysis of NSF's Innovative HPC Resources Using XDMoD

Nikolay A. Simakov, Joseph P. White, Robert L. DeLeon, Steven M. Gallo, Matthew D. Jones, Jeffrey T. Palmer, Benjamin Plessinger, and Thomas R. Furlani

Center for Computational Research, University at Buffalo, Buffalo NY USA 14203

January 16, 2018

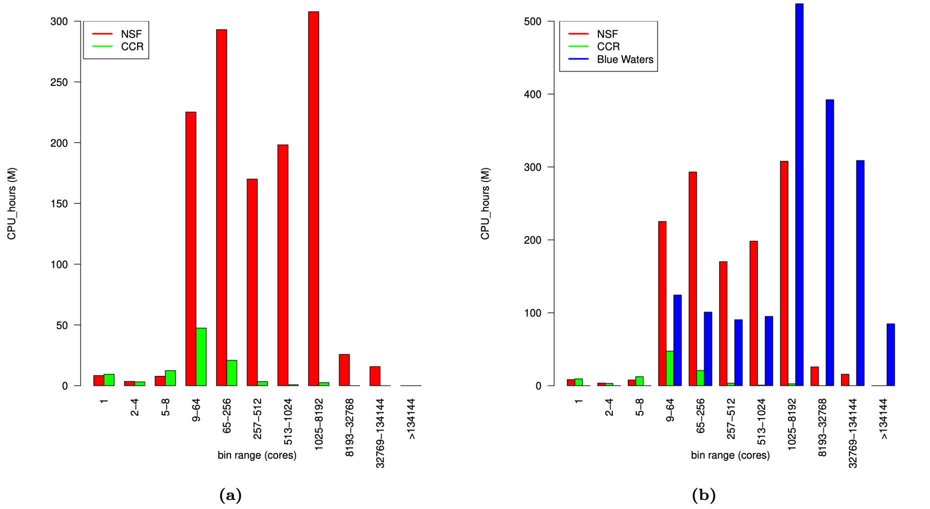



**Figure 1** – Comparison of CPU hours versus job size for a typical academic HPC center (CCR) versus the NSF Innovative HPC resources (a) and Blue Waters (b) in 2016.

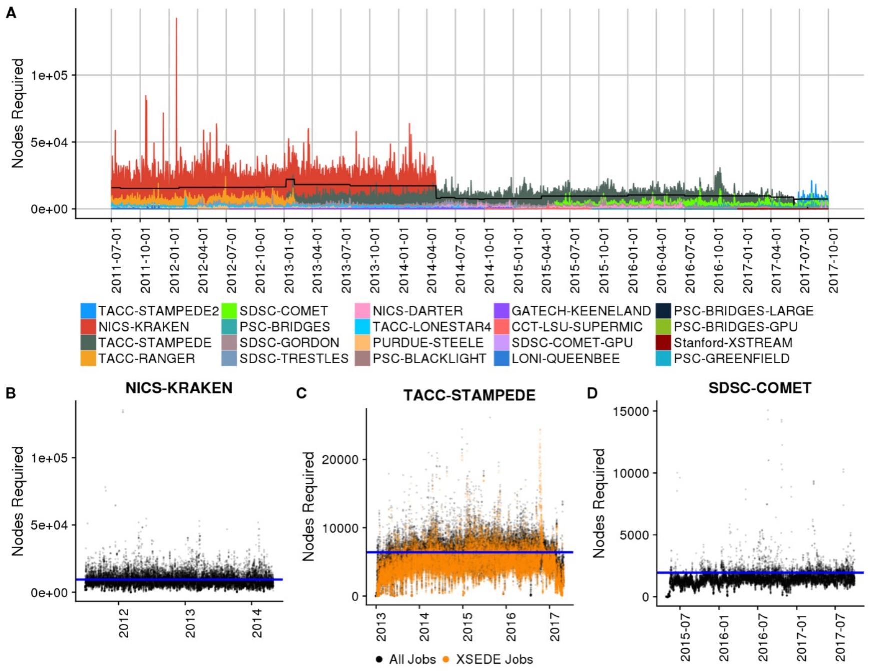



**Figure 73** – **A** - Number of nodes that would be required to run all queued jobs immediately over time. Solid black line represents the available nodes. Nodes required for: `NICS KRAKEN` (**B**), `TACC STAMPEDE` (**C**) and `SDSC COMET` (**D**). Blue line shows total number of nodes for the system. Black dots show required nodes for all jobs and orange for only XSEDE jobs, note that `NICS KRAKEN` and `SDSC COMET` are 100% allocated to XSEDE jobs

Providing a system with no wait time is not realistic. However, we can also frame our analysis in terms of what size would each of the HPC systems need to be to immediately run a given percentage of the queued jobs. This is presented in Table 12 in which we show the number of additional nodes and cores needed for 95% or 99% percent of the queued jobs on each resource to run immediately. At the 95% level, many of the systems would require about a 10 - 20% increase in size. Of course, this analysis assumes that the allocations are not increased to take advantage of the increased resource size.

# FROM THE DATA – ML BASED EVENTS AND ANOMALIES

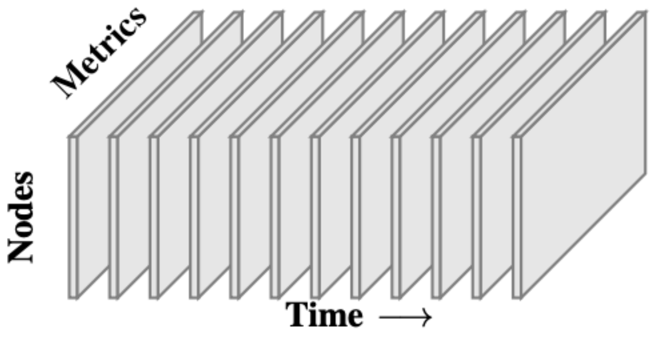



Fig. 1. Conceptual representation of resource usage data as a 3-way tensor consisting of compute **nodes**, usage **metrics** and **time** dimensions as lateral slices. Each dimension is referred to as a *mode* of the tensor.

Low rank Tensor (Tucker and CP) decomposition based modeling of data to detect anomalies. Validation using system log data.

Cluster 2017, Sorunkunlu, Chandola, Patra

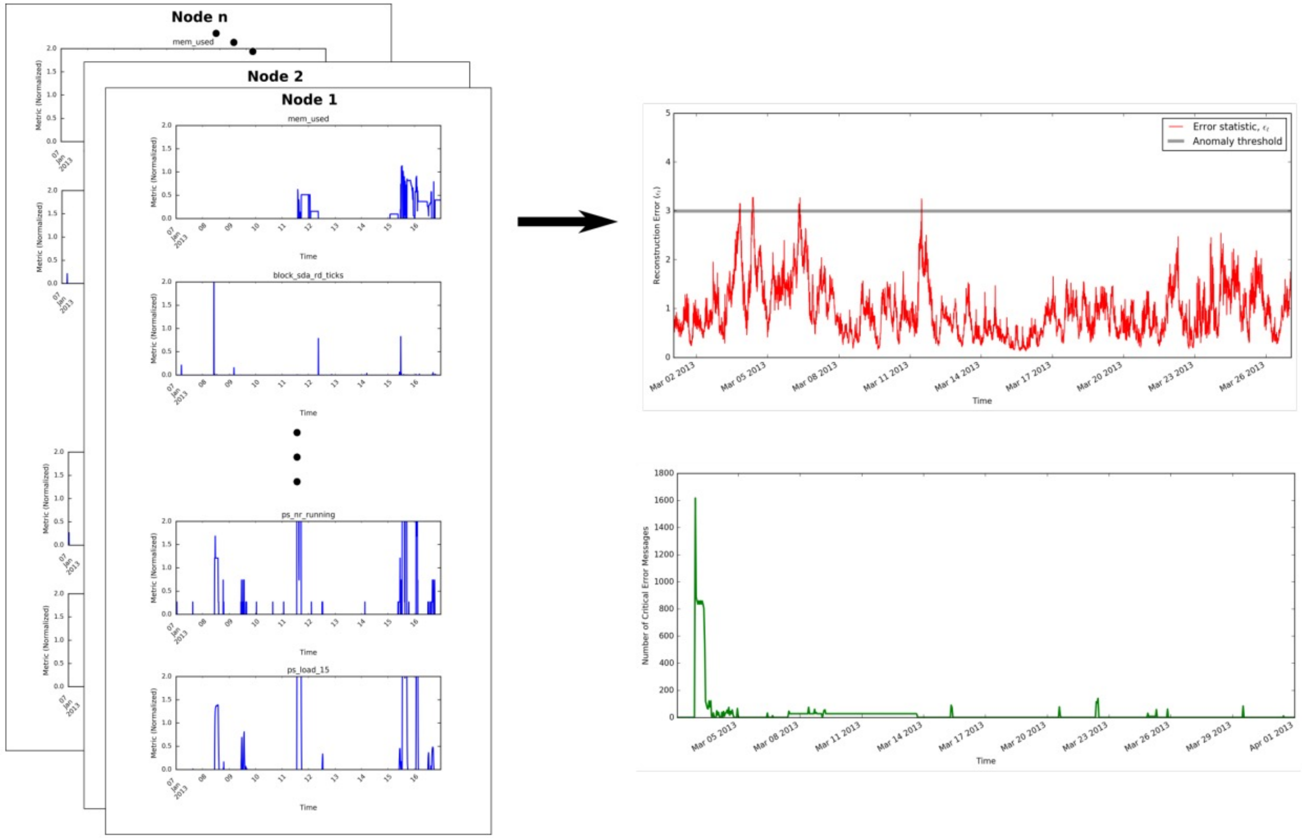
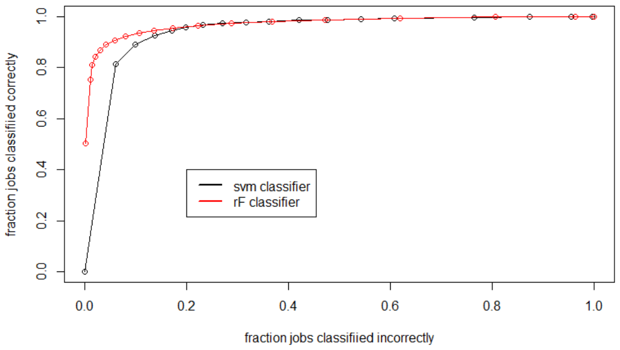



Fig. 2. Workflow for the proposed methodology. Resource usage data from multiple compute nodes (first panel) is combined as a tensor to produce an error metric (second panel). The anomalous spikes are shown to correspond to reported errors extracted from message logs. An example of such messages is shown in green (third panel).

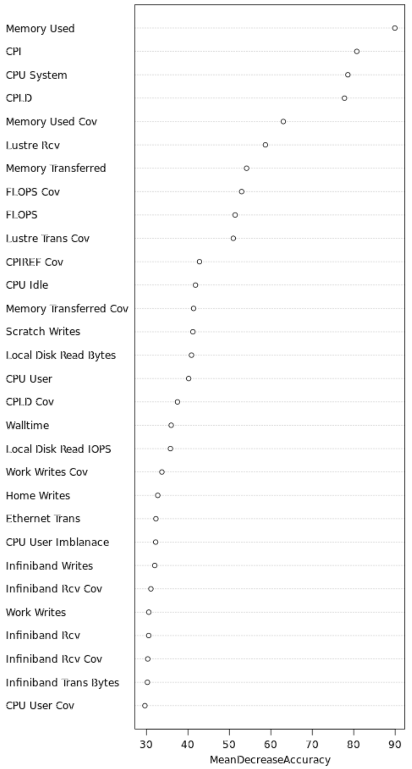
# EARLY USE OF ML

- Machine learning techniques svm and random forest were applied to job accounting and performance data for application classification.
- The results clearly demonstrate that community applications have characteristic signatures which can be exploited for job classification.

| Metric | Description |
|---|---|
| CPU SYSTEM, CPU USER, CPU IDLE | Fraction of CPU time spent in kernel mode, user mode, or idle. |
| CPLD | Clock ticks per L1D cache: Average ratio of clock ticks to L1D cache loads per core. |
| CPI | Clock tics per instruction: Average ratio of clock ticks to instructions per core. |
| FLOP | Total floating point operations. |
| MEMORY USED | Total memory used per node. This value excludes the OS buffer cache and kernel memory. |
| MEMORY TRANSFERRED | Total memory bandwidth in bytes per second. |
| ETHERNET TRANSMIT | Total bytes transmitted over the ethernet device per node. |
| INFINIBAND RECEIVE | Total bytes received over the InfiniBand device per node. |
| HOME WRITE, SCRATCH WRITE | Total bytes per node written to the indicated directory file. |
| LUSTRE RECIEVE | Total data received by the Lustre filesystem driver per node. |
| LOCAL DISK READ IOS, LOCAL DISK READ BYTES | The total local disk reads in bytes per second or operations per second. |
| NODES | Number of nodes on which the job was executed. |




| group_name | number | % mix | % correct |
|---|---|---|---|
| Astrophysics | 8200 | 2.94 | 87.01 |
| benchmark | 1238 | 0.44 | 76.27 |
| CFD | 10405 | 3.74 | 90.63 |
| E&M,photonics | 2922 | 1.05 | 98.39 |
| Lattice QCD | 344 | 0.12 | 89.38 |
| Math | 766 | 0.28 | 73.70 |
| Matlab | 125 | 0.04 | 91.43 |
| MD | 111102 | 39.90 | 98.71 |
| Python | 2041 | 0.73 | 65.67 |
| QC | 7670 | 2.75 | 94.60 |
| QC,ES | 120236 | 43.18 | 98.25 |
| Unknown | 13414 | 4.82 | 87.52 |

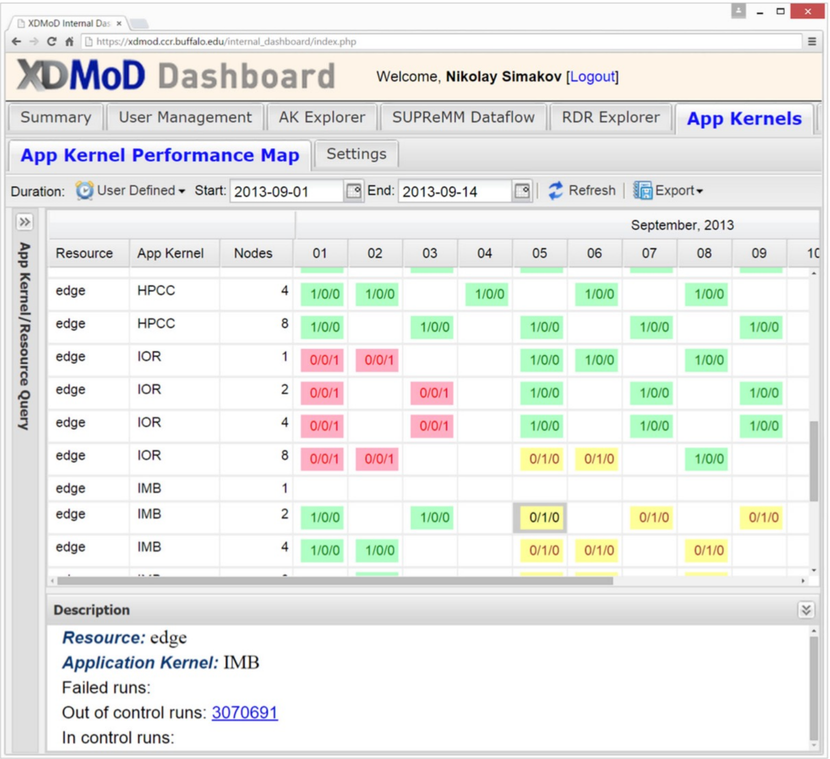


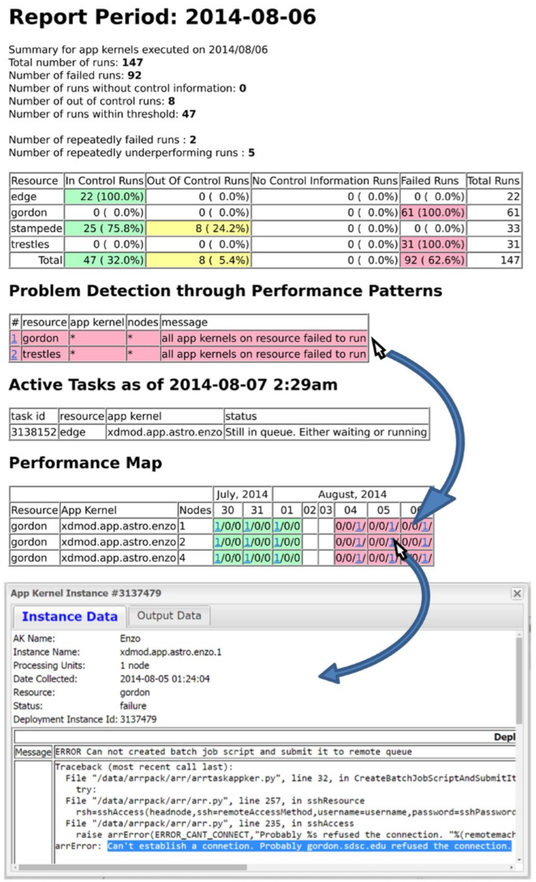

S. M. Gallo *et al.*, "Analysis of XDMoD/SUPReMM Data Using Machine Learning Techniques," *2015 IEEE International Conference on Cluster Computing*, Chicago, IL, USA, 2015, pp. 642-649, doi: 10.1109/CLUSTER.2015.114.

# BUT HOW DO WE GET GOOD DATA – INSTRUMENT!! APPLICATION KERNELS

Tufts UNIVERSITY | Data Intensive Studies Center

- Representative set application surrogates that are well instrumented and light weight!
- These are run periodically and performance stability is examined.
- Kernel performance indicates algorithm possibilities on architectures.

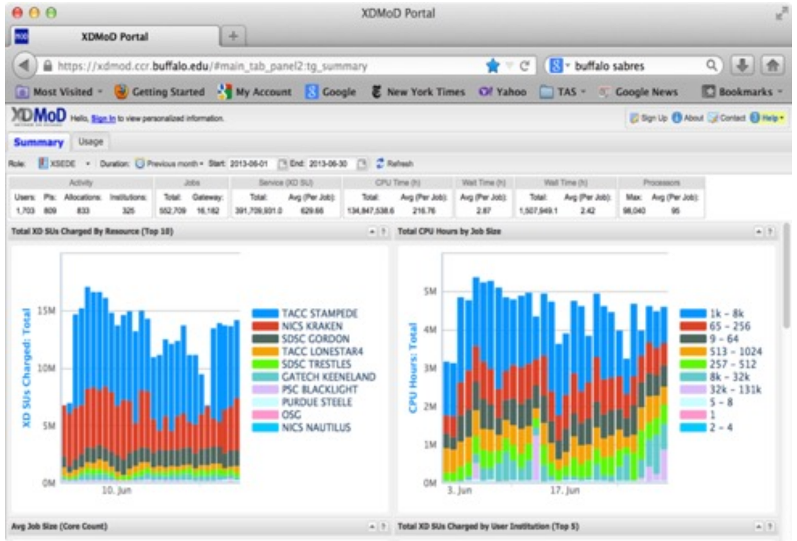



N. A. Simakov, J. P. White, R. L. DeLeon, A. Ghadersohi, T. R. Furlani, M. D. Jones, S. M. Gallo, A. K. Patra, "Application Kernels: HPC Resources Performance Monitoring and Variance Analysis", *Concurrency and Computation: Practice and Experience* (2015).

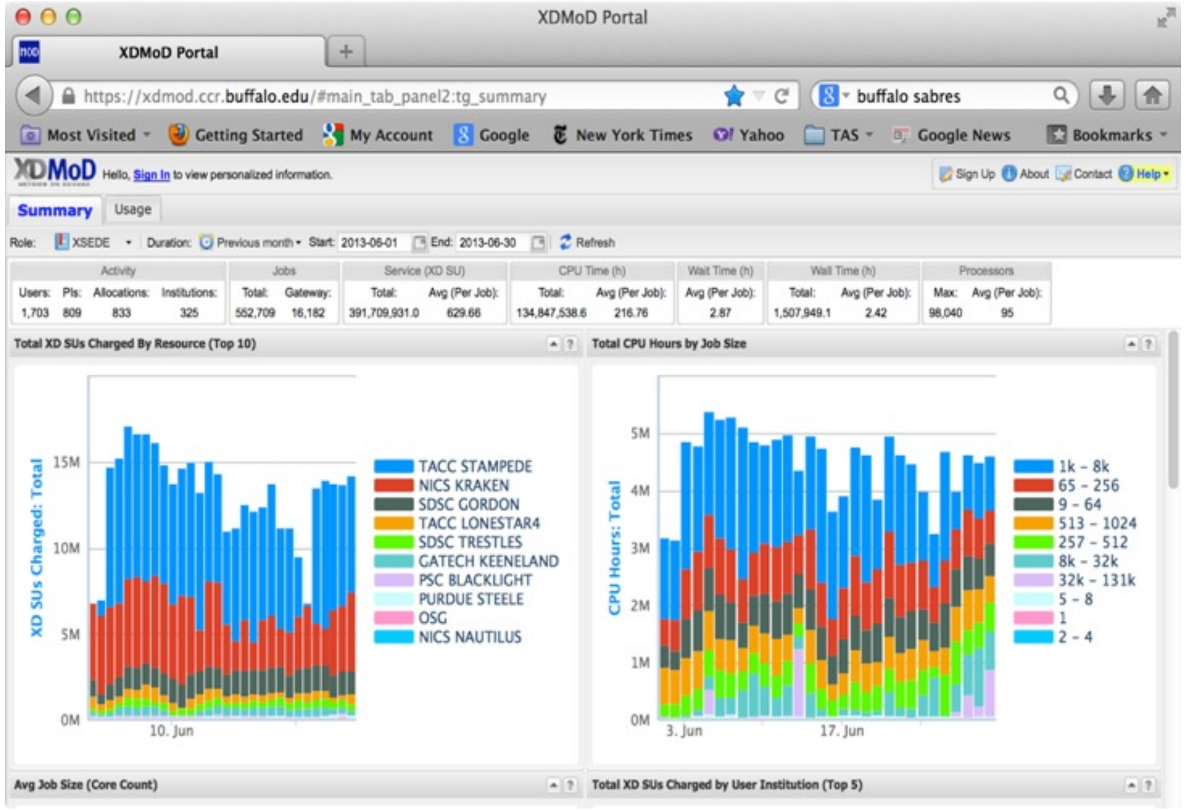
# XDMOD: METRICS ON DEMAND

- **Comprehensive framework for CI system management**
- **Understand and optimize resource utilization and performance**
  - Provide instantaneous and historical information on utilization
  - Measure Quality of Service of CI systems and applications
  - Measure and improve job and system level performance
  - Inform computing system upgrades and procurements
- **XDMoD tool**
  - Analytics Framework for ACCESS
- **Open XDMoD*: Open Source version for CI centers**
  - Used to measure and optimize performance of CI centers
  - 400+ academic, governmet, & commercial installations worldwide
  - [https://open.xdmod.org/](https://open.xdmod.org/)



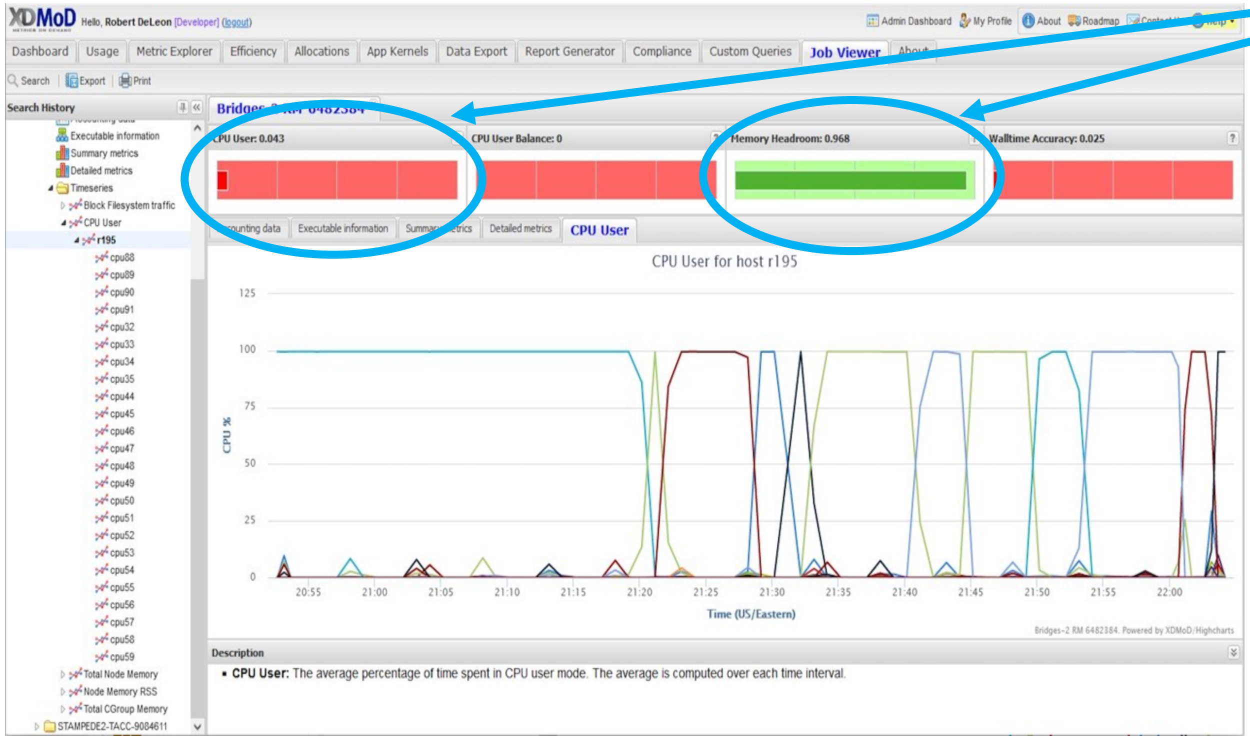
# PRIMARY COMPONENTS

- **Open XDMoD Portal**
  - Monitor system usage
- **Application Kernels**
  - Measure Quality of Service
- **Single Job Viewer Tab**
  - Measure and improve job performance
- **Job Efficiency Tab**
  - System wide view of job efficiency



# SINGLE JOB VIEWER

“Traffic Lights” provide quick overview of performance

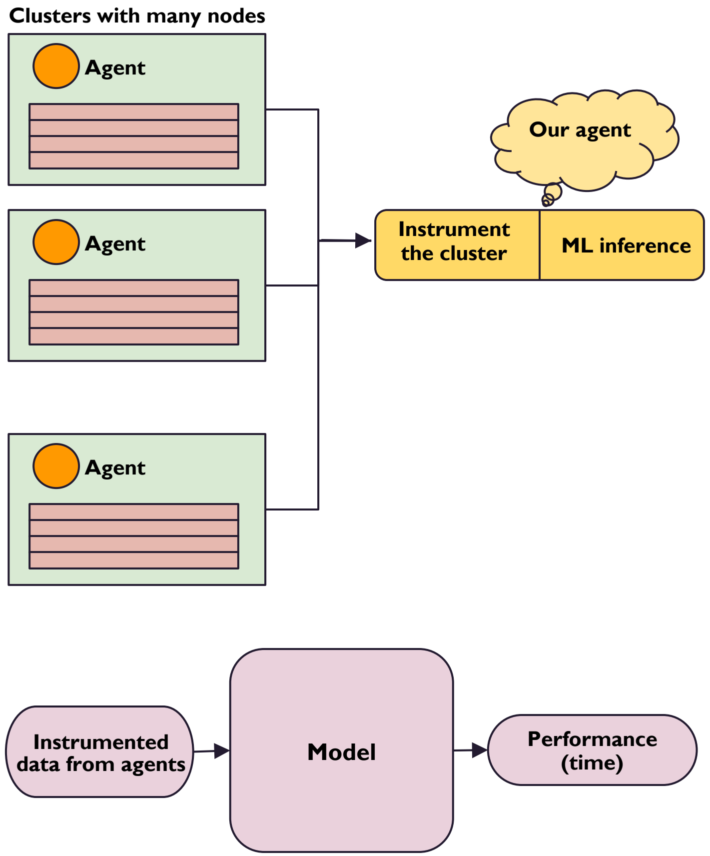



***Bridges-2 RM-6482384 job requests 24 cpus but only uses one at a time switching between the various cpus***

- **Detailed job-level performance data**
- **Run on every cluster node and collects hardware counter data**
  - CPU usage, Memory usage, I/O usage
- **Identify poorly performing jobs (users) and applications**
  - Support personnel/Users can use tools to identify/diagnose problems

Data Analytics framework – talk at 11:00

# THE FUTURE




- Our framework aims to facilitate problem identification by creating a controlled environment for data collection.
- Agents instrument the cluster, including machines and applications (memory, CPU, I/O, cache, network) to manipulate the data for measuring the performance
- Agent says: “this is the problem!”
- Leverage expert knowledge and instrumented data collected by monitoring agents to impose constraints (e.g., limited resources, high network traffic) that can lead to poor performance
- Identify the problem: list all the problems, use multi-level classification
- Build a lightweight model that can perform root cause analysis in real-time
- Experiment with various models, benchmark the performance, select the fastest one

# FINALLY

- Metrics are great – BUT we are in the service of science!

- Instrument, measure, analyze, improve → connect to real goal!
- Automation is essential → ML/AI …
- Improvement of science returns is everyone's goal all the time.