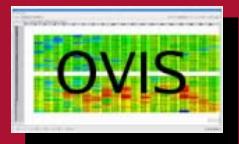




OVIS: A Tool for Intelligent, Real-time Monitoring of Computational Clusters

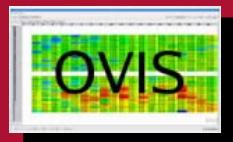
*Matthew Wong, Jim Brandt, Ann Gentile,
José Ortega, Philippe Pébay, David Thompson
(with Youssef Marzouk)*



*Sandia National Laboratories
Livermore, CA*



Sandia
National
Laboratories



Abstract

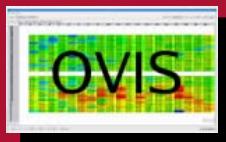
- OVIS is an open-source software tool for intelligent, real-time monitoring of computational clusters
- Visualization of deterministic information of cluster nodes:
 - CPU temperature
 - fan speed
 - memory error rate
 - etc.
- Built-in statistical tools for cluster analysis and prediction of future cluster health problems



Traditional RAS Tools

Ganglia, Supermon and commonly-supplied vendor RAS systems

- Nodes within clusters treated in singleton
- Manufacturer determined extreme limits define thresholds used for failure detection/avoidance



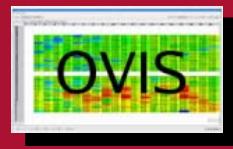


Statistical approach

- Enables earlier detection of abnormal behavior
- Can enable controlled failure avoidance
- Includes statistical and correlation tools

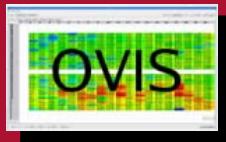
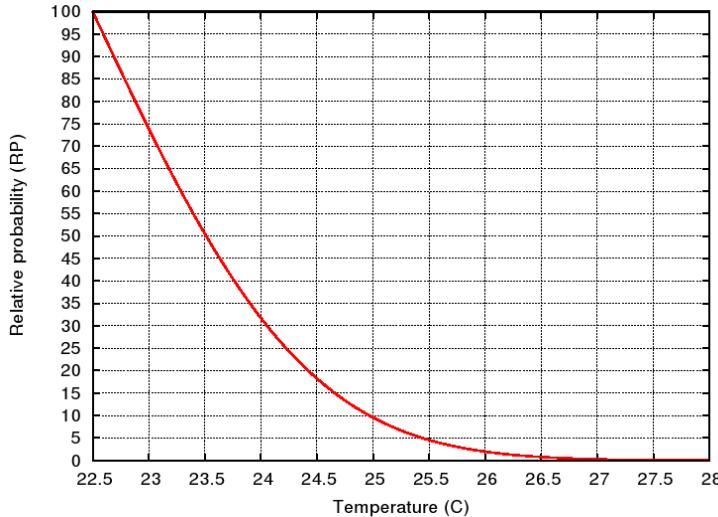
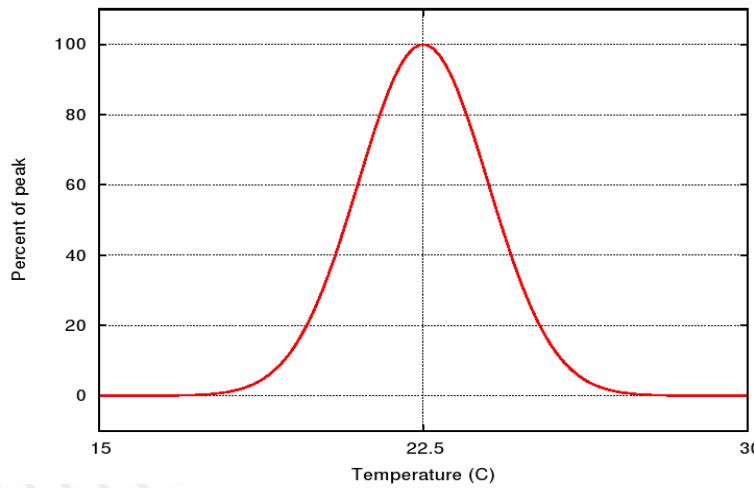
Spatial organization

- Data is displayed in a meaningful realistic geographic layout
- Immediate visual feedback aids in environmental understanding and configuration
- Color map facilitates intuitive visualization of state





Statistical Approach: An example of using a PDF to determine a measure of the relative probability of a value occurring



Spatial Approach

File Edit View Tools Time Help

System: Shasta Variable: CPU1 Speed: Time: Apr 7, 2005 11:30:33 AM Go

Plane 1 Plane 2 Plane 3 Plane 4 Plane 5 Plane 6 Plane 7 Plane 8 Plane 9 Plane 10

Rack 1	Rack 2	Rack 3		Rack 5			Rack 10	Rack 9	Rack 8
	cn48 42.00	cn80 41.00		cn112 44.00			cn198		
	cn47 40.00	cn79 43.00		cn111 43.00			cn197		
	cn46 38.00	cn78 43.00		cn110 42.00			cn196		cn138
	cn45 40.00	cn77 41.00		cn109 45.00			cn195		cn137
	cn44 36.00	cn76 40.00		cn108 43.00			cn194		cn136
	cn43 36.00	cn75 43.00		cn107 40.00			cn193		cn135
	cn42 39.00	cn74 42.00		cn106 38.00			cn192		cn134
	cn41 38.00	cn73 42.00		cn105 42.00			cn191		cn133
	cn40 41.00	cn72 43.00		cn104 41.00			cn190	cn188	cn132
	cn39 36.00	cn71 38.00		cn103 40.00			cn189	cn187	cn131
cn16 28.00	cn88 40.00	cn70 39.00		cn102 39.00			cn188	cn186	cn130
cn15 29.00	cn87 34.00	cn69 40.00		cn101 41.00			cn187	cn185	cn129
cn14 29.00	cn86 35.00	cn68 41.00		cn100 43.00			cn186	cn184	cn128
cn13 29.00	cn85 38.00	cn67 37.00		cn99 39.00			cn185	cn183	cn127
cn12 28.00	cn84 35.00	cn66 39.00		cn98 37.00			cn184	cn182	cn126
cn11 27.00	cn83 38.00	cn65 41.00		cn97 43.00			cn183	cn181	cn125
cn10 27.00							cn182	cn180	cn124
cn9 28.00							cn181	cn179	cn123
cn8 29.00							cn180	cn178	cn122
cn7 28.00							cn179	cn177	cn121
	cn6 26.00	cn64 36.00		cn96 34.00			cn178	cn176	cn120
	cn5 26.00	cn63 36.00	cn63 39.00	cn95 40.00			cn175	cn174	cn119
cn4 28.00	cn62 36.00	cn62 37.00		cn94 38.00			cn174	cn173	cn118
cn4 27.00	cn61 36.00	cn61 40.00		cn93 38.00			cn173	cn172	cn117
cn3 26.00	cn60 36.00	cn60 36.00		cn92 39.00			cn172	cn171	cn116
cn3 26.00	cn59 36.00	cn59 37.00		cn91 36.00			cn171	cn170	cn115
cn2 26.00	cn58 36.00	cn58 37.00		cn90 37.00			cn170	cn169	cn114
	cn57 36.00	cn57 38.00		cn89 39.00			cn169	cn168	cn113
	cn56 36.00	cn56 39.00		cn88 37.00			cn168	cn167	cn112
	cn55 36.00	cn55 35.00		cn87 40.00			cn167	cn166	cn111
	cn54 36.00	cn54 38.00		cn86 31.00			cn166	cn165	cn110
	cn53 36.00	cn53 35.00		cn85 40.00			cn165	cn164	cn109
	cn52 31.00	cn52 33.00		cn84 33.00			cn164	cn163	cn108
	cn51 30.00	cn51 40.00		cn83 39.00			cn163	cn162	cn107
	cn50 35.00	cn50 38.00		cn82 36.00			cn162	cn161	cn106
	cn49 36.00	cn49 39.00		cn81 39.00			cn161	cn160	cn105

From: Apr 7, 2005 11:28:30 AM To: Sliding End Apr 7, 2005 11:33:36 AM Sampling Period (sec): 1

Statistics Tool

The figure shows the Ovis software interface. At the top, a menu bar includes File, Edit, View, Tools, Time, and Help. Below the menu is a toolbar with buttons for System (Shasta), Variable (CPU1), Speed (clock icon), Time (set to Apr 7, 2005 11:30:33 AM), and Go. A navigation bar at the top right lists Plane 1 through Plane 10. The main area features a 10x10 grid of colored cells representing data for Racks 1 through 10 and Planes 1 through 10. The color scale ranges from blue (low values) to yellow (high values). A statistics tool window is open on the right, titled 'Ovis - Statistics Tool'. It contains fields for Node, Rack, and System, and date/time pickers for From (Apr 7, 2005 11:28:30 AM), At (Apr 7, 2005 11:28:33 AM), and To (Apr 7, 2005 11:33:36 AM). The 'Variable' is set to CPU1, and the 'Results File Prefix' is OvisStatistics-. Below these fields, the window displays basic statistics for variable CPU1: sample size (15120), minimum (20), maximum (51), mean (35.99), median (38), mode 1 (41), mode 2 (43), unbiased variance (60.46), sample variance (60.45), sample skewness (-0.386), and sample kurtosis (1.881). A legend on the right side of the heatmap provides numerical values for the colors: Node (cn8 23.00), CPU1: 39, CPU2: 40, FAN1: 12, FAN2: 12, dRX: 8918, and dTX: 24872. At the bottom, there are buttons for Clear, Calculate, Save, and Close, along with date/time pickers for From (Apr 7, 2005 11:28:30 AM) and To (Apr 7, 2005 11:30:33 AM), and a 'Sliding End' checkbox.

Correlation Tool

Ovis

File Edit View Tools Time Help

System: Shasta Variable: CPU1 Speed: Time: Apr 7, 2005 11:30:33 AM Go

Plane 1 Plane 2 Plane 3 Plane 4 Plane 5 Plane 6 Plane 7 Plane 8 Plane 9 Plane 10

Rack 1	Rack 2	Rack 3	Rack 4	Rack 5
cn18 42.00	cn80 41.00		cn112 44.00	
cn17 40.00	cn79 43.00		cn111 43.00	
cn16 39.00	cn78 43.00		cn130 42.00	
cn15 40.00	cn77 41.00		cn109 45.00	
cn44 39.00	cn76 40.00		cn108 43.00	
cn43 36.00	cn75 43.00		cn107 40.00	
cn42 39.00	cn74 42.00		cn106 38.00	
cn41 38.00	cn73 42.00		cn105 42.00	
cn40 41.00	cn72 43.00		cn104 41.00	
cn39 39.00	cn71 38.00		cn103 40.00	
cn16 28.00	cn70 39.00		cn102 39.00	
cn15 26.00	cn77 34.00	cn69 40.00	cn101 41.00	
cn14 29.00	cn76 35.00	cn68 41.00	cn100 43.00	
cn13 29.00	cn75 38.00	cn67 37.00	cn69 39.00	
cn12 26.00	cn74 35.00	cn66 39.00	cn68 37.00	
cn11 27.00	cn73 38.00	cn65 41.00	cn67 43.00	
cn10 27.00				
cn9 28.00				
cn8 29.00				
cn7 28.00				
	cn62 23.00	cn64 36.00	cn66 34.00	
cn65 25.00	cn61 23.00	cn63 29.00	cn65 40.00	
cn5 28.00	cn60 23.00	cn62 37.00	cn64 38.00	
cn4 27.00	cn59 23.00	cn61 40.00	cn63 38.00	
cn3 24.00	cn58 22.00	cn60 36.00	cn62 39.00	
cn2 25.00	cn57 21.00	cn59 37.00	cn61 36.00	
cn1 34.00	cn56 29.00	cn58 37.00	cn60 37.00	
	cn25 25.00	cn57 36.00	cn69 39.00	
	cn54 23.00	cn56 39.00	cn68 37.00	
	cn23 24.00	cn55 35.00	cn67 40.00	
	cn22 22.00	cn54 38.00	cn66 31.00	
	cn21 22.00	cn53 35.00	cn65 40.00	
	cn20 21.00	cn52 33.00	cn64 33.00	
	cn19 23.00	cn51 40.00	cn63 39.00	
	cn18 26.00	cn50 38.00	cn62 38.00	
	cn17 26.00	cn49 39.00	cn61 39.00	

From: Apr 7, 2005 11:28:30 AM To: Sliding End Apr 7, 2005 11:33:36

Ovis - Correlation Tool

Node: Rack: System

From: Apr 7, 2005 11:28:30 AM To: Apr 7, 2005 11:33:36 AM

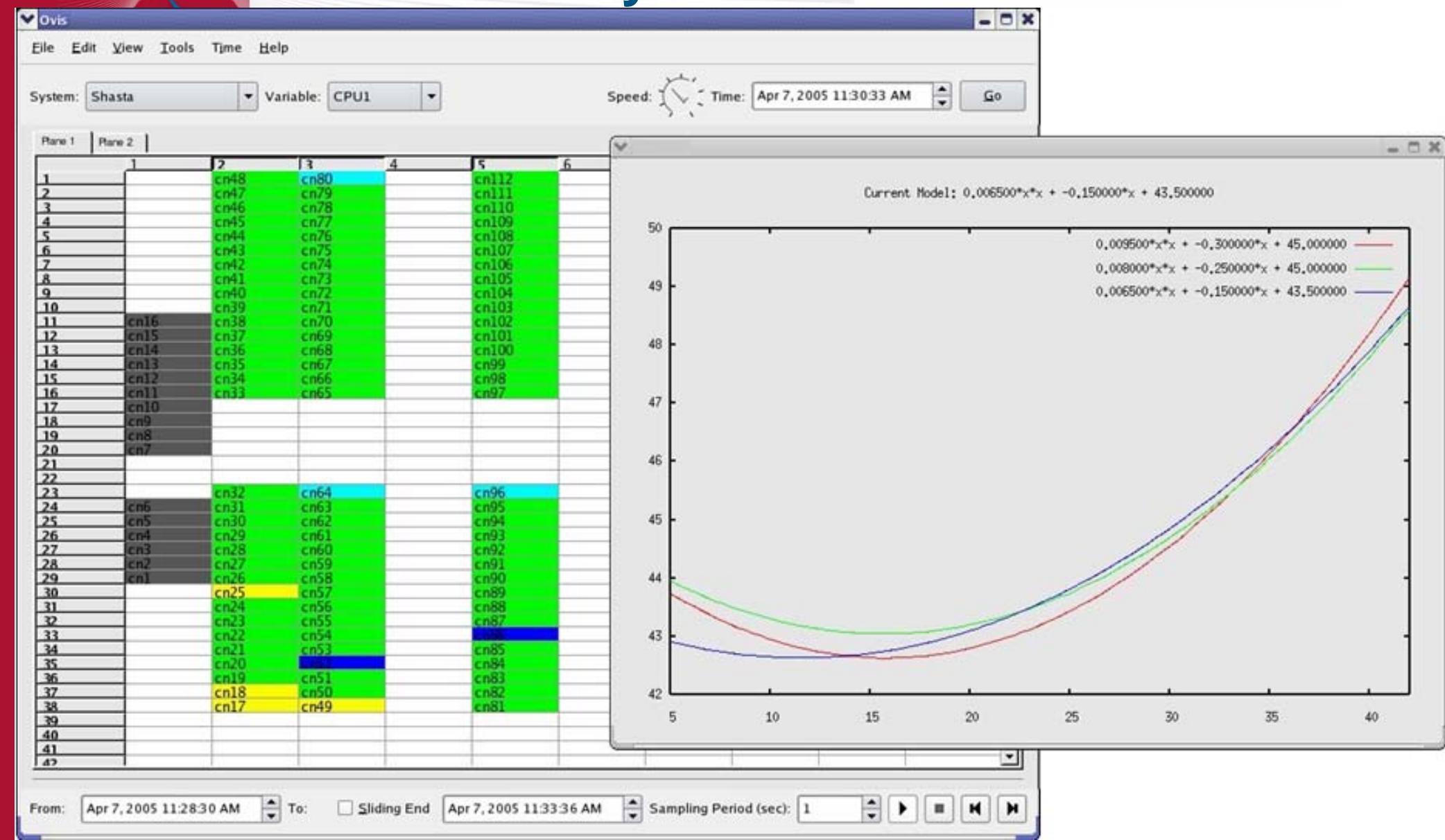
Variables: CPU1 CPU2

Results File Prefix: OvisCorrelation-

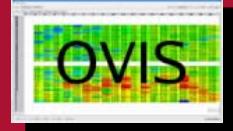
Linear correlation between variables CPU1 and CPU2:
sample size: 15120
mean of variable 1: 35.99
mean of variable 2: 38.2
sample variance of variable 1: 60.45
sample variance of variable 2: 77.24
covariance: 65.67
variable 2 on 1 regression slope: 1.086
variable 2 on 1 regression intersect: -0.8932
variable 1 on 2 regression slope: 0.8502
variable 1 on 2 regression intersect: 3.509
linear correlation coefficient: 0.961

Clear Calculate Save Close

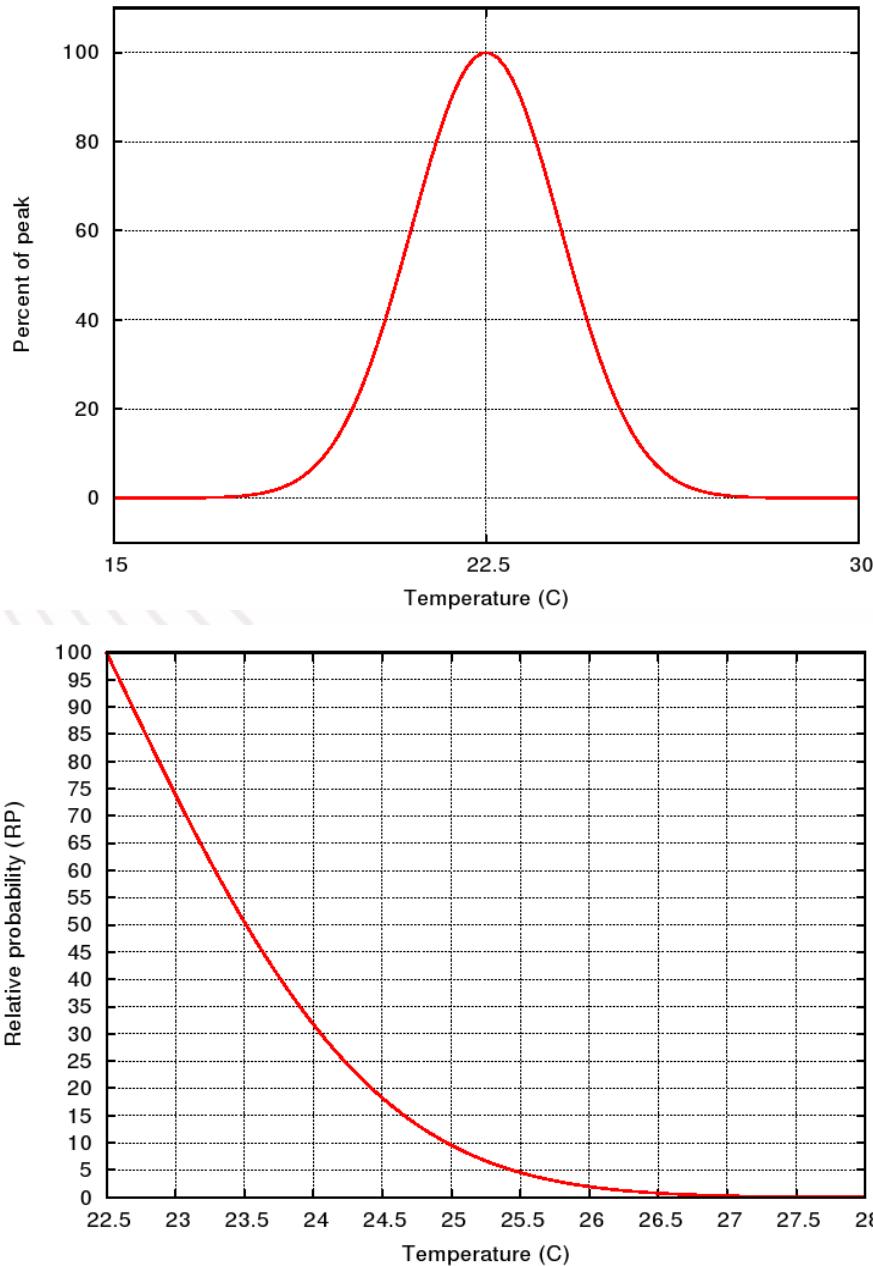
Environmental Modelling Using Bayesian Inference



Sandia
National
Laboratories



Statistical Approach: Plot of PDF of Idle Rack 3 at height 10

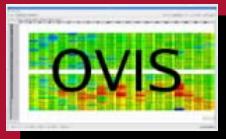


Sandia
National
Laboratories



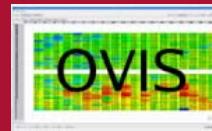
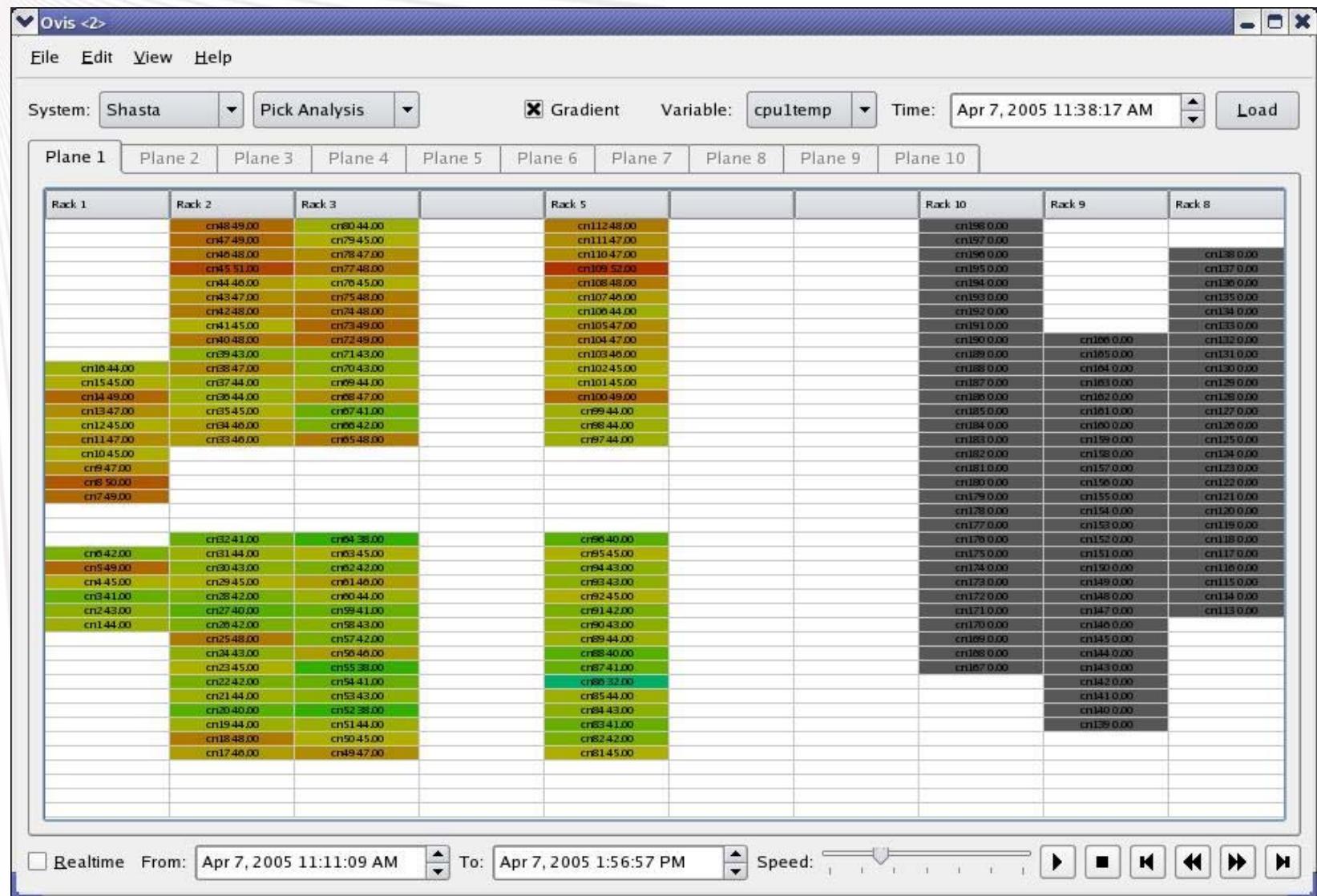
OVIS – Other Features

- User-customizable color schema with automatic gradient shading
- Visualization of real-time data, or playback of historical data
- Easily adaptable to new systems
 - XML cluster description
 - Various readers to fetch collected data (e.g. Ganglia)





Demo

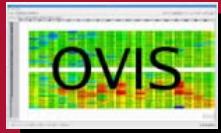


**Sandia
National
Laboratories**



OVIS 2

- 3-D navigable cluster representation providing advanced visualization with drill-down functionality to component level
- Parallel architecture for real-time statistical characterization, modeling and analysis of large (10's of thousands of nodes) clusters
- Further data analysis tools including automated troubleshooting functionality





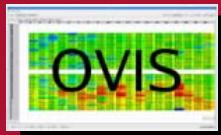
Contacts and Downloads

Contacts

ovis-help@sandia.gov

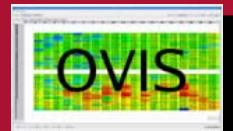
For OVIS downloads and more information:

<http://ovis.ca.sandia.gov>



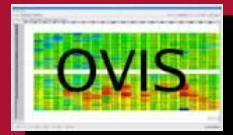


Last Slide





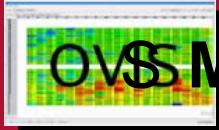
Extras





Problems and Solutions

- P** CPU utilization of different nodes could be different at any point in time → legitimately different CPU temperatures
- S** Model steady state temperature distribution(s) with CPU utilization dependence

- P** All nodes not co-located in space → non-homogeneous environment
-  **OVSS** Model temperature distribution(s) with geographic location (e.g. distance off floor) dependence



Problems and Solutions

P Room temperature not fixed (but does have an acceptable range) → same CPU utilization rate yields different temperatures at different times (though change is slow)

S Use new data to continuously update model parameters and confidence in the model

P If sampling rate too slow --> Cannot look for stability



Sovis Model upper and lower bounds



Sandia
National
Laboratories

Sandia ICC Shasta System Grid > Compute Partition >

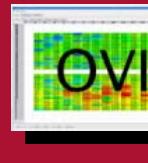
Compute Partition cluster - Physical View | Columns

Verbosity level (Lower is more compact):
 3 2 1

Total CPUs: 396
 Total Memory: 387.7 GB

Total Disk: 6685.0 GB
 Most Full Disk: cn64 (38.3% Used)

Rack 1	Rack 2	Rack 3	Rack 5
cn16 0.20 cpu: 2.99G (2) mem: 1.96G	cn48 0.02 cpu: 2.99G (2) mem: 1.96G	cn80 0.03 cpu: 2.99G (2) mem: 1.96G	cn112 2.00 cpu: 2.99G (2) mem: 1.96G
cn15 0.00 cpu: 2.99G (2) mem: 1.96G	cn47 2.00 cpu: 2.99G (2) mem: 1.96G	cn79 0.02 cpu: 2.99G (2) mem: 1.96G	cn111 2.00 cpu: 2.99G (2) mem: 1.96G
cn14 0.00 cpu: 2.99G (2) mem: 1.96G	cn46 1.99 cpu: 2.99G (2) mem: 1.96G	cn78 0.01 cpu: 2.99G (2) mem: 1.96G	cn110 2.00 cpu: 2.99G (2) mem: 1.96G
cn13 2.00 cpu: 2.99G (2) mem: 1.96G	cn45 1.99 cpu: 2.99G (2) mem: 1.96G	cn77 0.02 cpu: 2.99G (2) mem: 1.96G	cn109 2.00 cpu: 2.99G (2) mem: 1.96G
cn12 2.09 cpu: 2.99G (2) mem: 1.96G	cn44 2.01 cpu: 2.99G (2) mem: 1.96G	cn76 0.00 cpu: 2.99G (2) mem: 1.96G	cn108 1.00 cpu: 2.99G (2) mem: 1.96G
cn11 2.00 cpu: 2.99G (2) mem: 1.96G	cn43 1.99 cpu: 2.99G (2) mem: 1.96G	cn75 0.01 cpu: 2.99G (2) mem: 1.96G	cn107 2.00 cpu: 2.99G (2) mem: 1.96G
cn10 2.00 cpu: 2.99G (2) mem: 1.96G	cn42 1.99 cpu: 2.99G (2) mem: 1.96G	cn74 0.00 cpu: 2.99G (2) mem: 1.96G	cn106 2.00 cpu: 2.99G (2) mem: 1.96G
cn9 2.00 cpu: 2.99G (2) mem: 1.96G	cn41 1.99 cpu: 2.99G (2) mem: 1.96G	cn73 0.00 cpu: 2.99G (2) mem: 1.96G	cn105 2.00 cpu: 2.99G (2) mem: 1.96G
cn8 2.00 cpu: 2.99G (2) mem: 1.96G	cn40 1.99 cpu: 2.99G (2) mem: 1.96G	cn72 0.00 cpu: 2.99G (2) mem: 1.96G	cn104 2.00 cpu: 2.99G (2) mem: 1.96G
cn7 2.01 cpu: 2.99G (2) mem: 1.96G	cn39 1.99 cpu: 2.99G (2) mem: 1.96G	cn71 0.00 cpu: 2.99G (2) mem: 1.96G	cn103 0.00 cpu: 2.99G (2) mem: 1.96G
cn6 2.00 cpu: 2.99G (2) mem: 1.96G	cn38 0.00 cpu: 2.99G (2) mem: 1.96G	cn70 0.00 cpu: 2.99G (2) mem: 1.96G	cn102 2.00 cpu: 2.99G (2) mem: 1.96G
cn5 2.00 cpu: 2.99G (2) mem: 1.96G	cn37 2.00 cpu: 2.99G (2) mem: 1.96G	cn69 0.00 cpu: 2.99G (2) mem: 1.96G	cn101 2.00 cpu: 2.99G (2) mem: 1.96G
cn4 2.00 cpu: 2.99G (2) mem: 1.96G	cn36 2.01 cpu: 2.99G (2) mem: 1.96G	cn68 0.02 cpu: 2.99G (2) mem: 1.96G	cn100 2.00 cpu: 2.99G (2) mem: 1.96G





Node View for Thu, 13 Apr 2006
14:03:15 -0700

[Get Fresh Data](#)

[Host View](#)

[Sandia ICC Shasta System Grid](#) > [Compute Partition](#) > **cn68**

cn68 Info

cn68

Load: 1.94 0.44 1.29

1m 5m 15m

Location: Rack 3, Rank 25, Plane 0.

Last heartbeat received 9 seconds ago.

Uptime 16 days, 0:20

CPU Utilization: 97.4 3.1 0.1

user sys idle

Hardware

CPUs: 2 x 2.99 Ghz

Memory (RAM): 1.96 GB

Local Disk: Using 4.994 of 33.772 GB

Most Full Disk Partition: 21.6% used.

Software

OS: Linux 2.4.21-27.0.2.ELSFsmp (x86)

Booted: March 28, 2006, 12:42 pm

Uptime: 16 days, 0:20

Swap: Using 11.5 of 2000.2 MB swap.



cn68 Overview



This host is up and running.

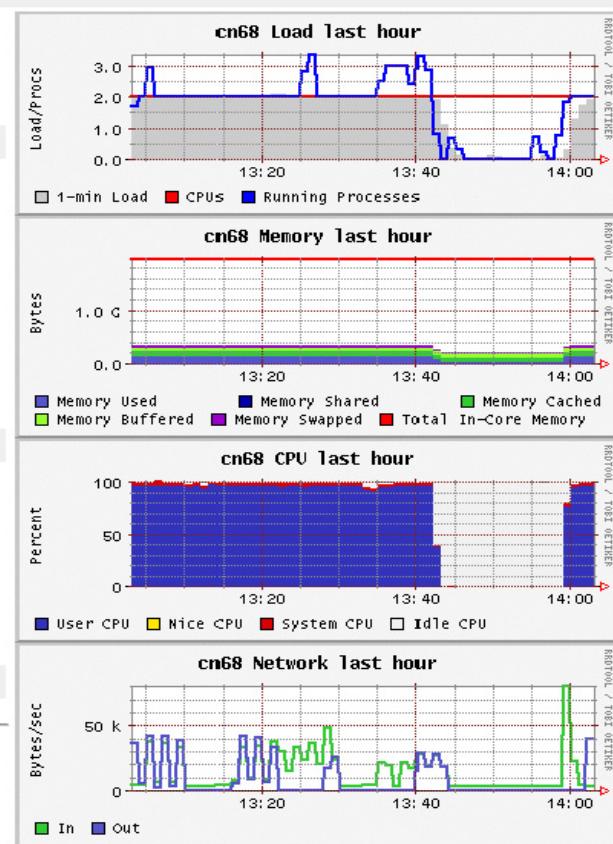
Time and String Metrics

boottime Tue, 28 Mar 2006 12:42:10 -0800
gexec OFF
machine_type x86
os_name Linux
os_release 2.4.21-27.0.2.ELSFSSmp
sys_clock Mon, 3 Apr 2006 09:22:54 -0700
uptime 16 days, 0:20

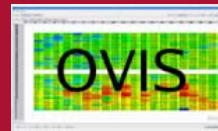
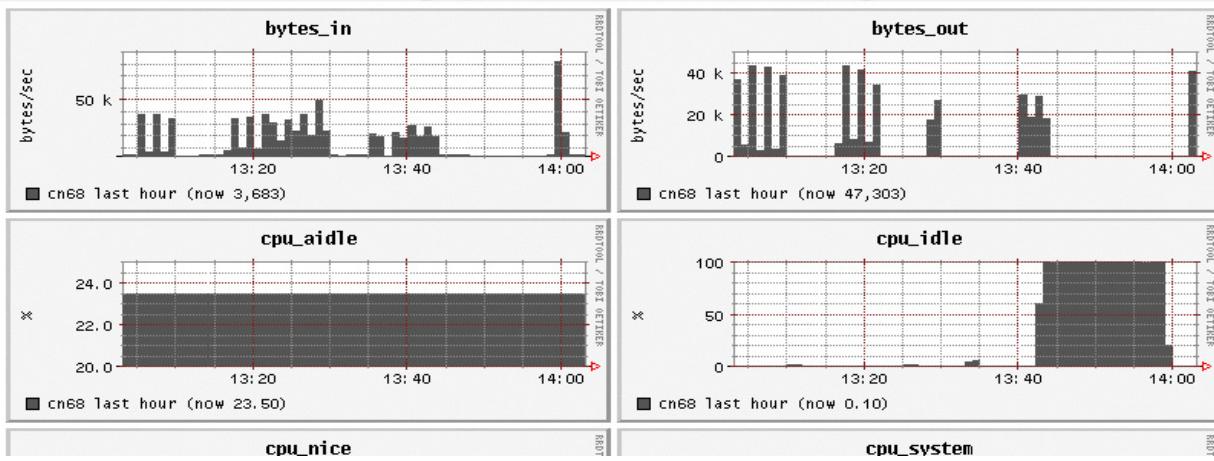
Constant Metrics

cpu_num 2
cpu_speed 3065 MHz
mem_total 2053420 KB
mtu 1500 B
swap_total 2048248 KB

Gmetrics



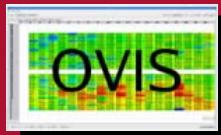
cn68 graphs last hour sorted descending





Why simple statistics aren't enough

Need something about why we are using Bayesian inference.



Example of Bayesian Modeling:

Dependence of Temperature on

Height

- Bayesian learning allows us to incorporate expert knowledge in the model. e.g., it has been noted that on Shasta, the temperature T baseline varies with height h in the cluster.
- However, each individual node should behave similarly under similar conditions, within some variability range due to the manufacturing process.
- Thus, we model T as a Gaussian r.v. with mean $Q(h)$ (Q is a quadratic \Rightarrow 3 degrees of freedom) and variance s . We can then infer the PDF of the 4 unknown parameters based on the data at hand using Bayes' formula.
- But in fact, we are mostly interested in the *most likely* parameters to characterize the model, using $P(X|D,M) \propto P(D|X,M) \times P(X|M)$

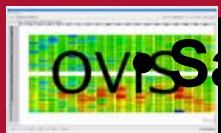


Model training is automatically done until some user-defined maximum likelihood convergence ratio is reached



Problems

- CPU utilization of different nodes could be different at any point in time → legitimately different CPU temperatures
- All nodes not co-located in space → non-homogeneous environment
- Room temperature not fixed (but does have an acceptable range) → same CPU utilization rate yields different temperatures at different times (though change is slow)



Sampling rate too slow --> Cannot look for stability -->
can only compare with upper and lower bounds

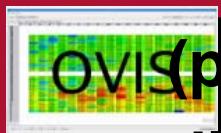


Sandia
National
Laboratories



Solutions

- Model steady state temperature distribution(s) with CPU utilization dependence
- Model temperature distribution(s) with geographic location (e.g. distance off floor) dependence
- Use new data to continuously update model parameters and confidence in the model
- Model upper and lower bounds and call anything in between good though must detect shift in model (perhaps if many are falling out of bounds in same direction).



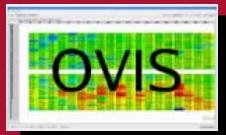
Sandia
National
Laboratories



OVIS's Interface

OVIS can be used either in command-line mode or with a GUI Interface that:

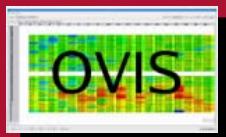
- provides smooth transition between raw data and statistics and derived data and visualization,
- facilitates system administrator analyses for configuration and monitoring of a system,
- supports a variety of drop-in monitoring and analysis modules.





Current functionality

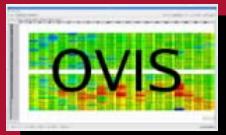
- Provide raw data visualization and archiving
- Provide statistical characterization, Bayesian modeling and analysis – not very fast and needs a homogeneous glob to work on





Posterior

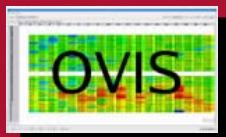
- $P(X|D,M) = P(D|X,M) \times P(X|M) / P(D|M)$
- **$P(X|D,M)$ (posterior) is the probability distribution of model parameters given our data and choice of model.**
- **We chose the particular parameter set which yields the maximum of the PDF as most likely and use it and the model as the basis for comparison with any individual node**





Likelihood

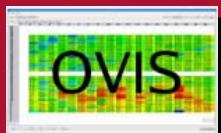
- $P(X|D,M) = P(D|X,M) \times P(X|M) / P(D|M)$
- $P(D|X,M)$ (likelihood) is the probability distribution of actual data over all sets of parameters for the model





Prior

- $P(X|D,M) = P(D|X,M) \times \textcolor{red}{P(X|M)} / P(D|M)$
- **P(X|M) (prior) is the probability distribution of the parameter sets of the model**
 - For initial calculation allows input of expert knowledge
 - Model selection
 - Knowledge of the actual distribution of parameter sets (we use a uniform distribution because of lack of knowledge)
 - Posterior from previous iteration is used on subsequent iterations

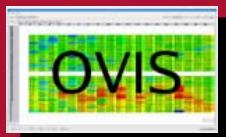


Sandia
National
Laboratories

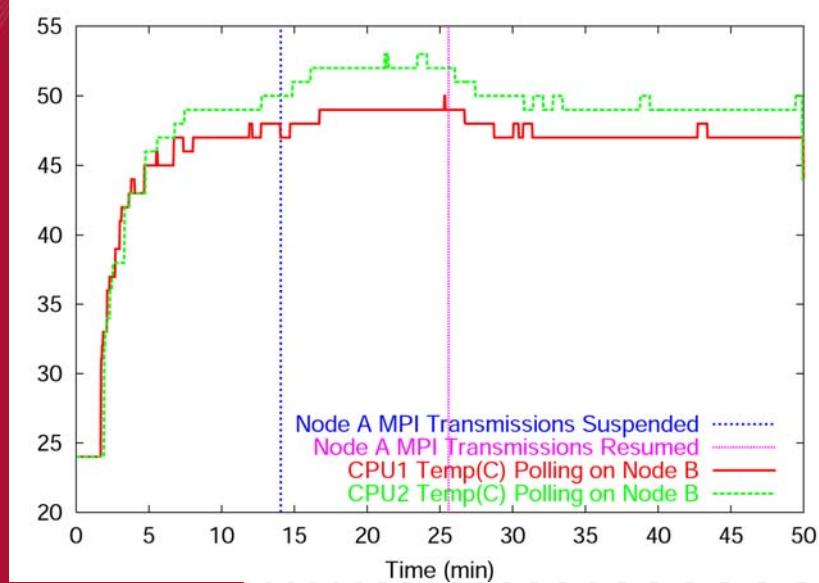


Evidence

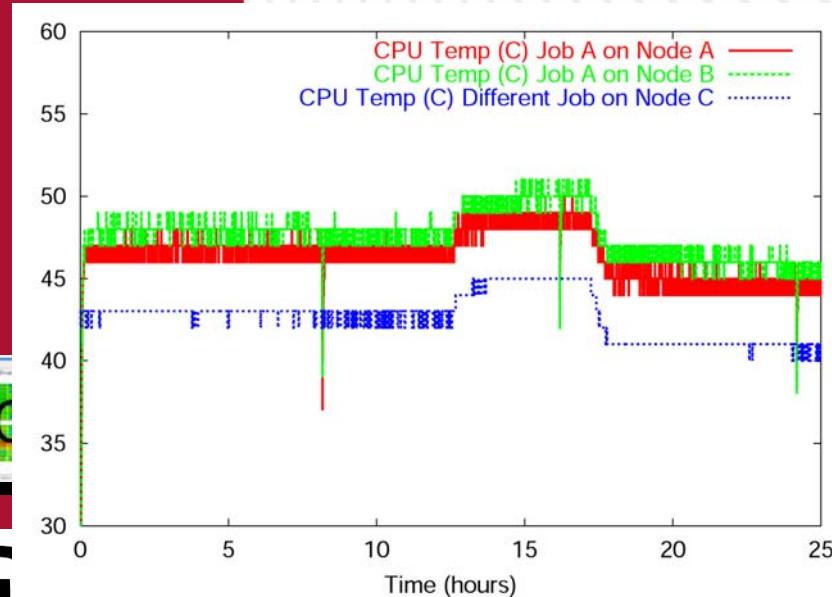
- $P(X|D,M) = P(D|X,M) \times P(X|M) / P(D|M)$
- $P(D|M)$ (evidence) is a normalization term calculated by summing the posteriors before normalization has occurred (total CDF = 1)



Statistically Recognizable Abnormalities



Emulate MPI failure – suspend one node, temperature rises in polling nodes



Room temperature changes – affect nodes in uncorrelated job groups



Demo: Example Analysis Sequence

- General overview of variables:



– Fan Initial

- Faster around gaps

– Altering Airflow - Fan

- Problem solved when plugged racks

– Temperature

- Jobs come in and out (propagation of jobs onto the cluster starting at higher numbers)

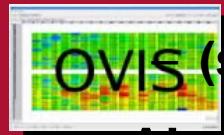
- Natural vertical gradient

- Also hot near floor

- Node 86

- Inference Model (Non-uniform and non-linear environmental effects)

– Fit to Normal distribution with 2nd order polynomial (function of height) as its mean



– OVIS (screen data) converging likelihood, increasing confidence

- Abnormality Detection (Model comparison with the inferred curve)

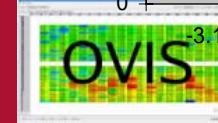
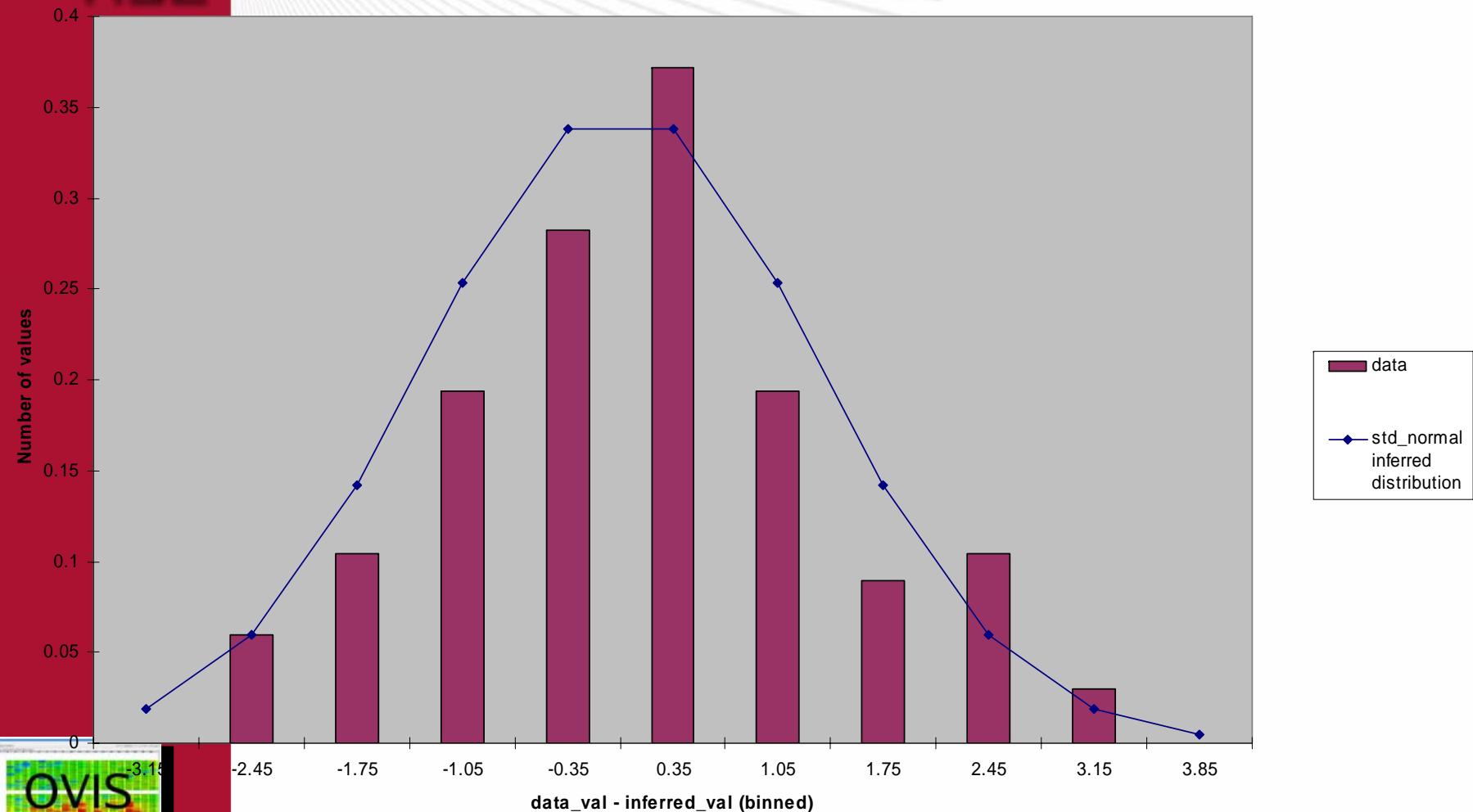


Sandia
National
Laboratories

Node 86

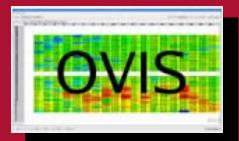
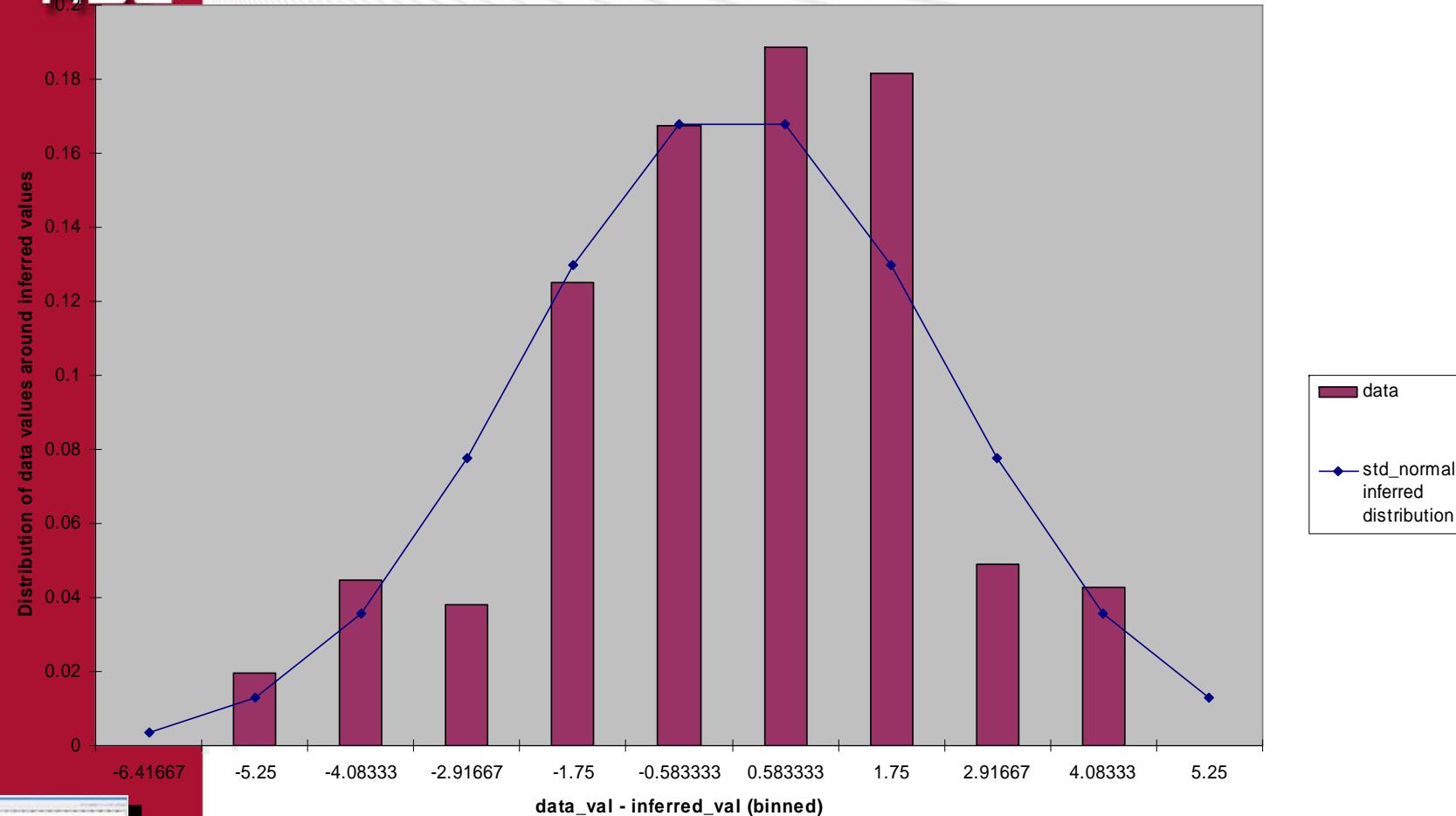


Distribution of data values around inferred values- idle case (1 timestep)

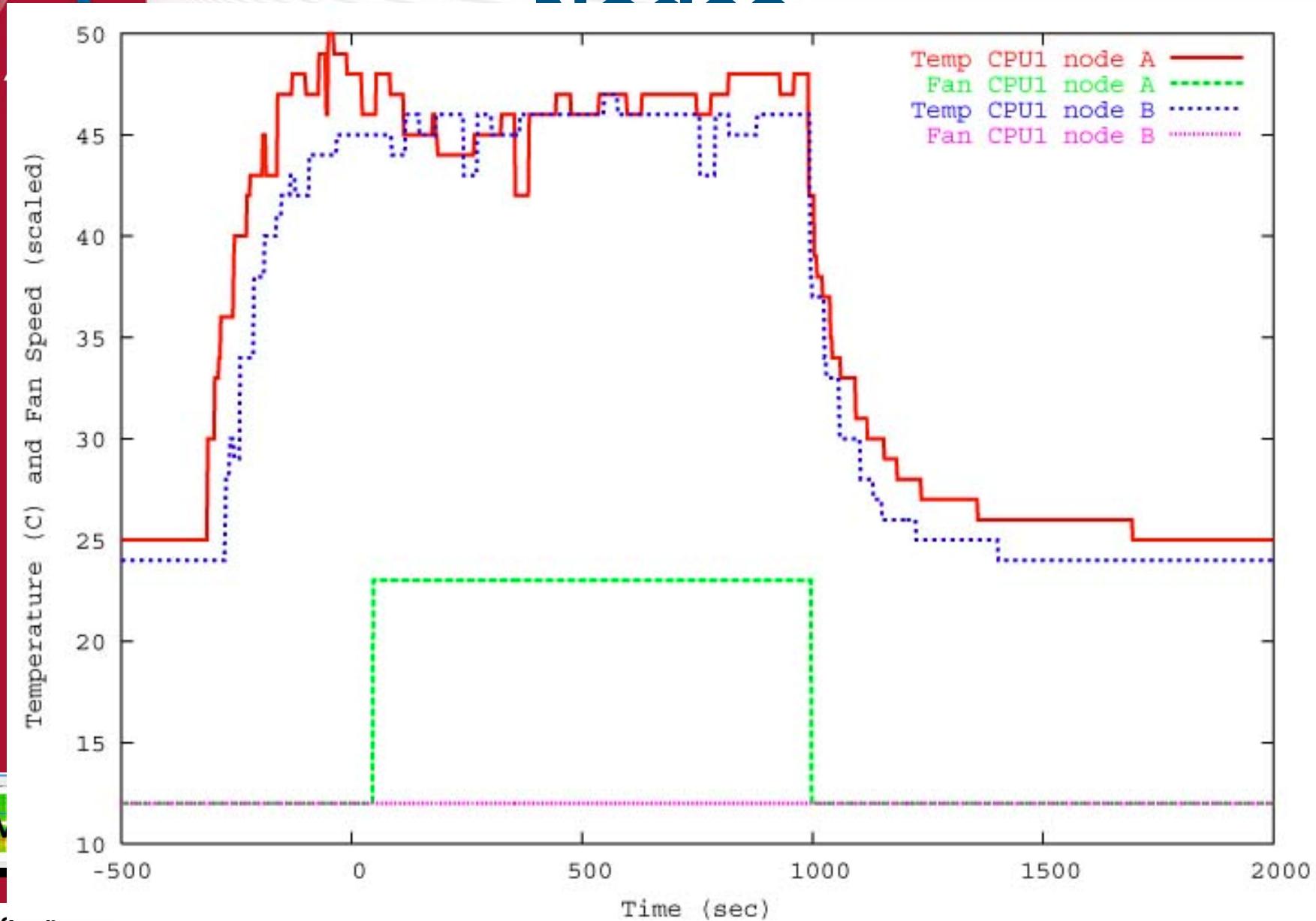




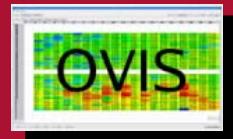
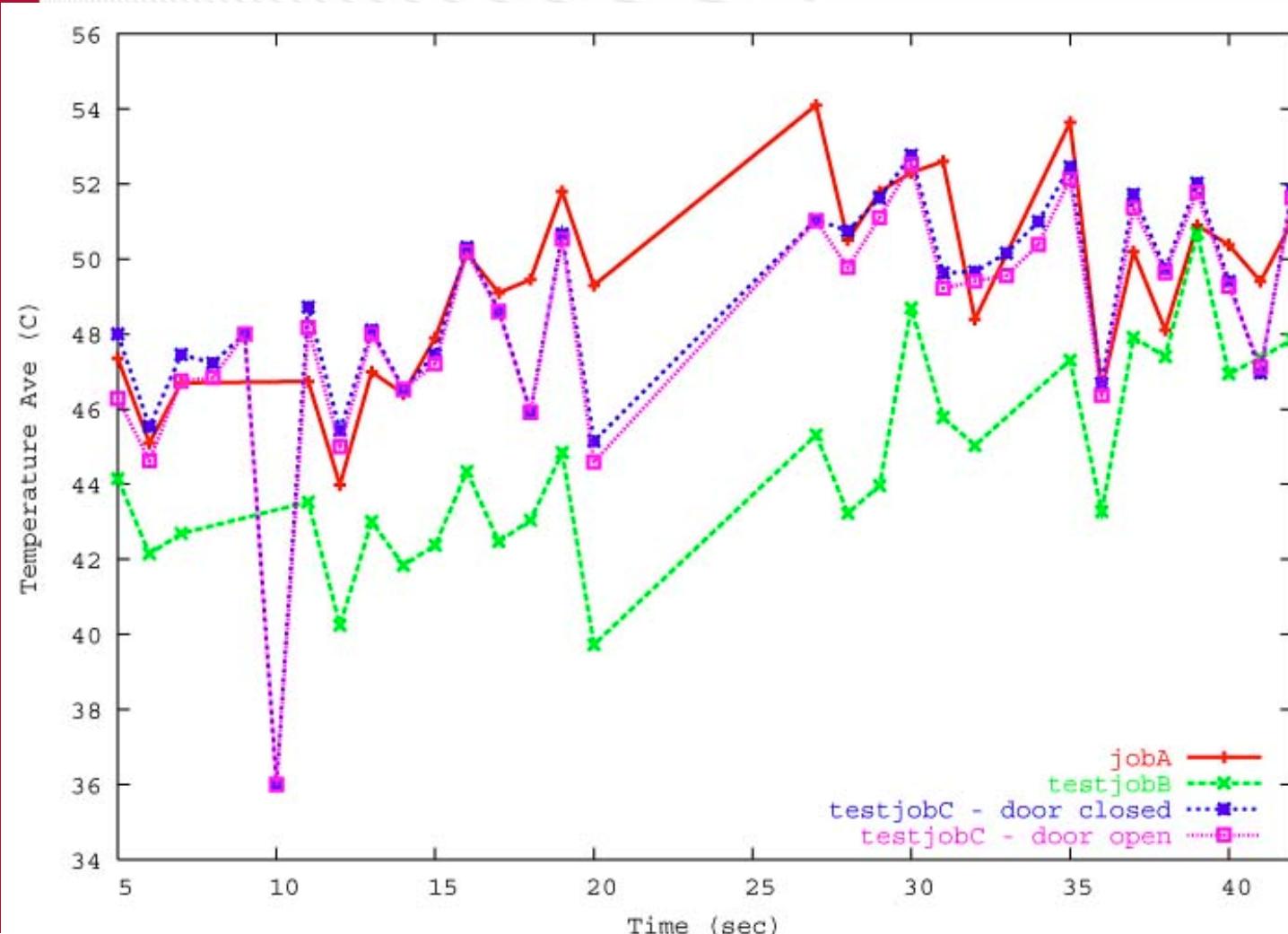
Distribution of data values around inferred values- loaded case (38 timesteps)



Natural Variations in Individual



Characteristics Persist Across Different Job and Environmental Conditions

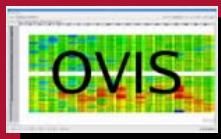


Notes

- Jags are the same in different circs, even if the curves (absolute and relative) arent. Individual variations overwhelm the conditions ?

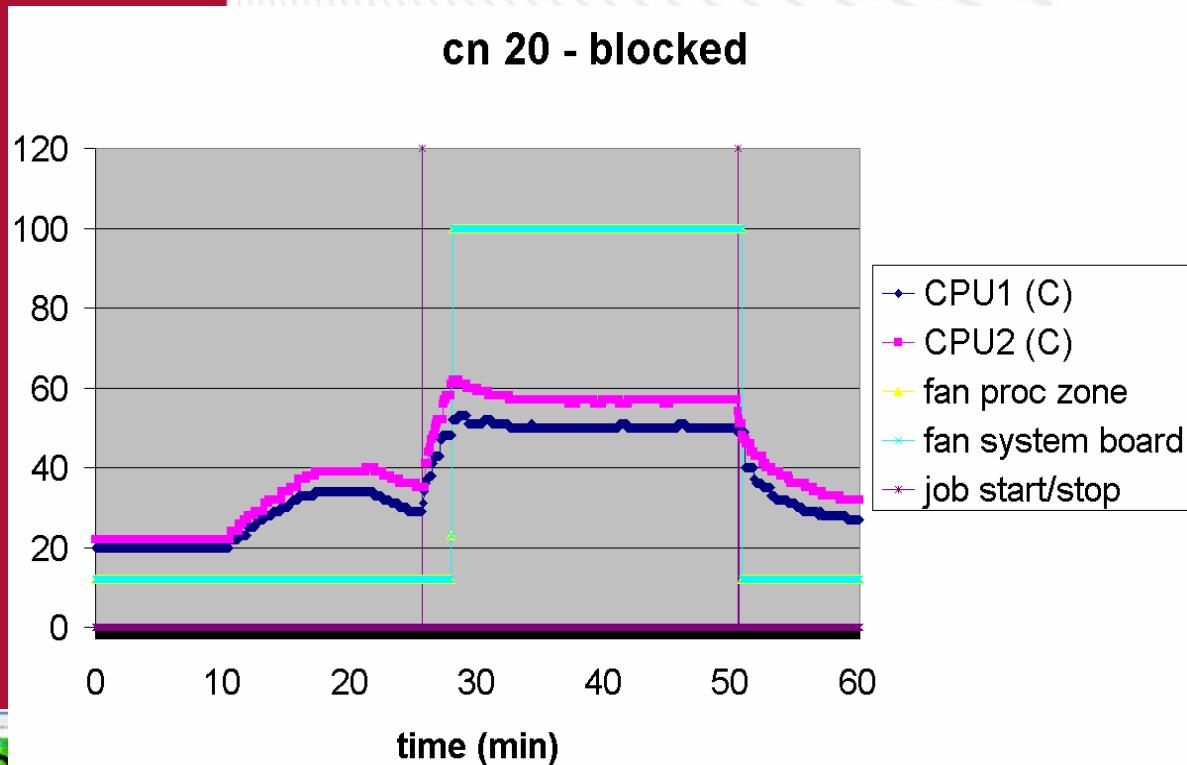
- Research things:

- how does model change as room temp change ? Does that differ from how model changes when nodes run harder (CPU utiliation)
 - Model – not over fitting





Recognizable Abnormalities (cont'd)

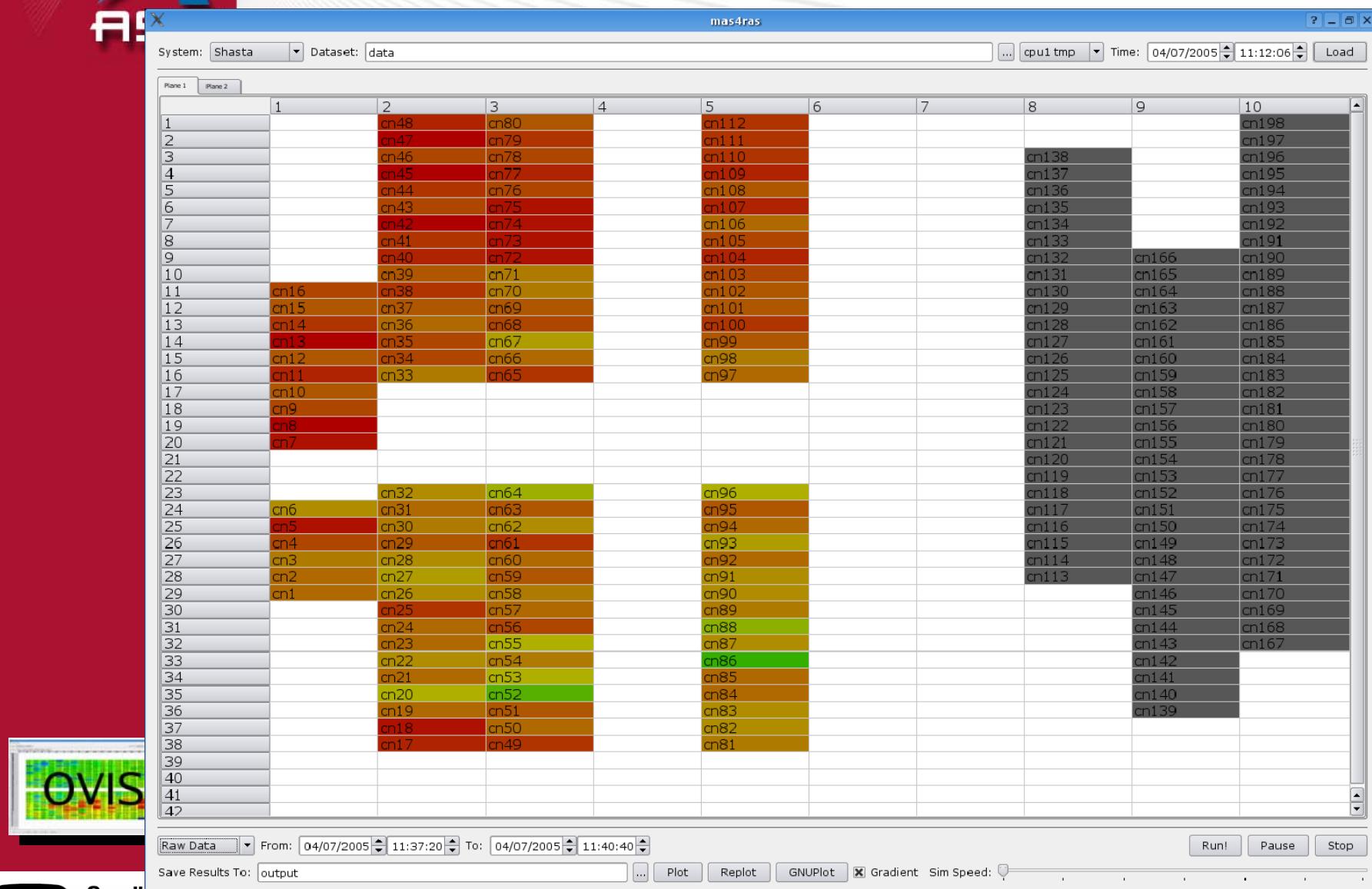


Partially blocked fan
– fan speed
increases in
response to rising
temperatures, while
fans on other nodes
remain at constant
speed



Sandia
National
Laboratories

Non-Uniform, Non-Linear Environmental Effects





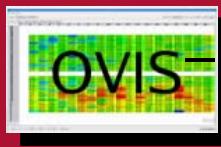
Analysis and Monitoring

- **Analysis research**

- Understanding the system, changing likelihood's, other machine learning issues (clustering, adaptive methods etc.), time dependencies, dynamic issues, plotting capabilities, other variables (e.g., memory error rates).
- Many different jobs – what is a statistically significant set?
- Failure prediction and interactions with resource management

- **Monitoring -- System administrators**

- Initial set up (HP) and effects of changes
- Silently models, monitors and provides descriptive colour maps.



Currently addressing a subset of the problem – thermal issues/airflow/cooling

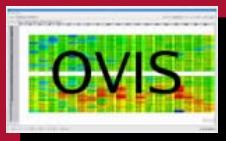


Sandia
National
Laboratories



Needs associated with RAS

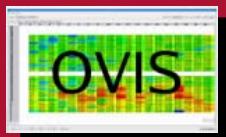
- Advance warning of impending faults
- Fault detection
- Diagnostic help to identify actual problem(s)
- Interface to batch scheduler
- Interface to trigger check-point





Real-Time Requirements for Statistical Approach

- *Data sampling intervals short relative to change in monitored variables*
- *Data processing must keep pace with sampling*
- *Only do comparisons when data is stable*





Probabilistic Characterization Using Bayesian Inference

The keystone of our approach is Bayesian inference.

Reminder: Bayes' Theorem:

$$P(X|D, M) = P(D|X, M) \times P(X|M) / P(D|M)$$

or less formally: posterior = likelihood \times prior / evidence

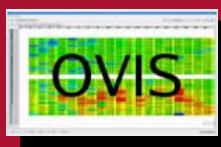
- X is a set of model parameters to be inferred (e.g., polynomial coefficients and variance in the model above);
- D is the data, i.e., measurements of the variables that are present in the model.
- M is the probabilistic model (e.g., temperature is distributed as a Gaussian r.v. whose parameters have a polynomial dependence on OVI⁵ weight in the cluster);





Abnormality Detection Using the Inferred Model

- After model inference has been done (either with training data or "live" data), we have a stochastic model whose parameters optimally fit the data.
 - E.g., for idle rack 3:
 - $T \sim N(0.005 h^2 -0.1 h + 23, 1.5)$
 - Outliers can be defined automatically based on user-defined thresholds
 - $RP(\{h=10, T=23\} | \{0.005, -0.1, 23, 1.5\}, M) \approx 95\%$
 - $RP(\{h=10, T=25\} | \{0.005, -0.1, 23, 1.5\}, M) \approx 25\%$

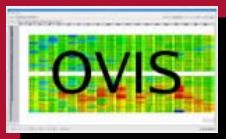


Sandia
National
Laboratories



Approach Summary

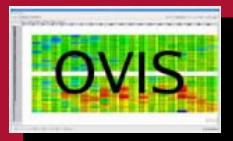
- Use statistical approach for probabilistic modeling using Bayesian inference
- Use these probabilistic characterizations to identify outliers, hot spots, etc.
- Use the constant influx of incoming data to update and improve the existing probabilistic models (machine learning).
- Scope the problem (thermal issues) for technique development then expand (memory errors, fans, voltages, cross-correlations)





How Is This Different and Where Is the Intelligence?

- **Thresholds in terms of probability rather than raw data**
- **Determined by statistical processing on real data**
- **Numerical threshold values are *learned* and dynamic**
 - Will change in response to aging, environmental effects, etc.
- **Environmental modeling**





Ganglia on Shasta

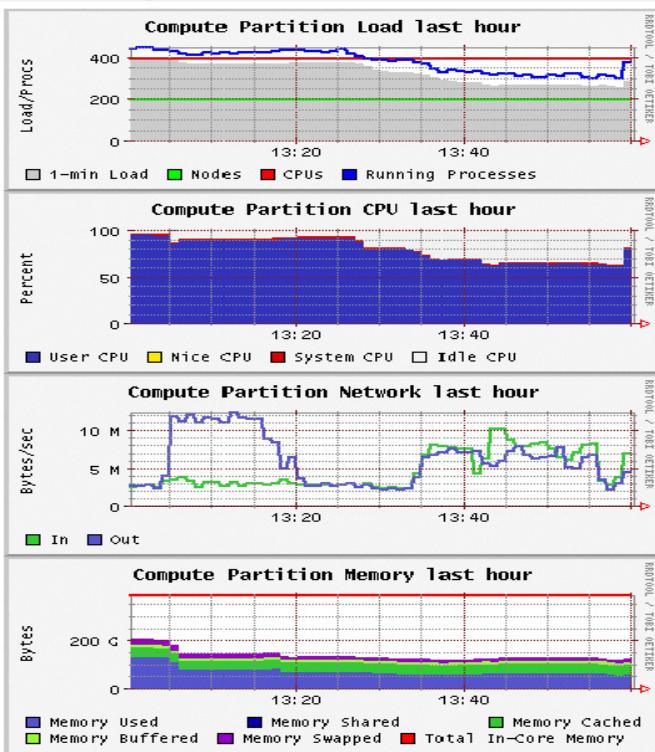
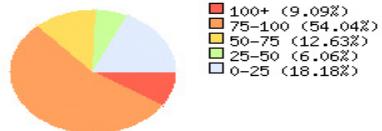
CPUs Total: 396
Hosts up: 198
Hosts down: 0

Avg Load (15, 5, 1m):
74%, 67%, 74%

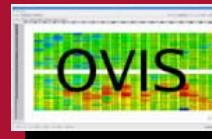
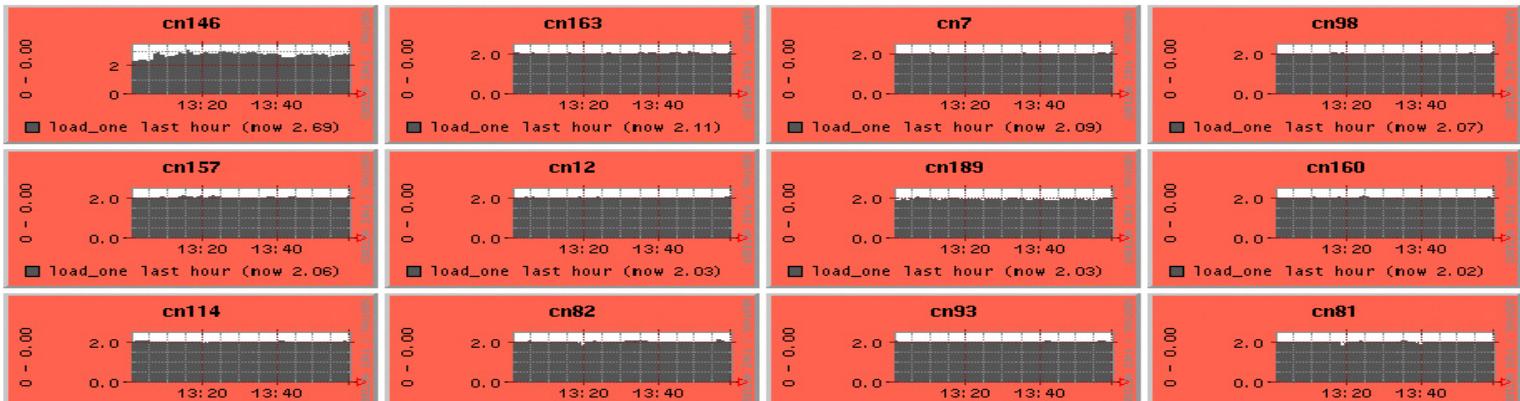
Localtime:
2006-04-13 14:00

Job Queue

Cluster Load Percentages



Show Hosts: yes no | Compute Partition **load_one** last hour sorted **descending** | Columns **4**





Temp vs. Time Plots on Shasta Cluster

