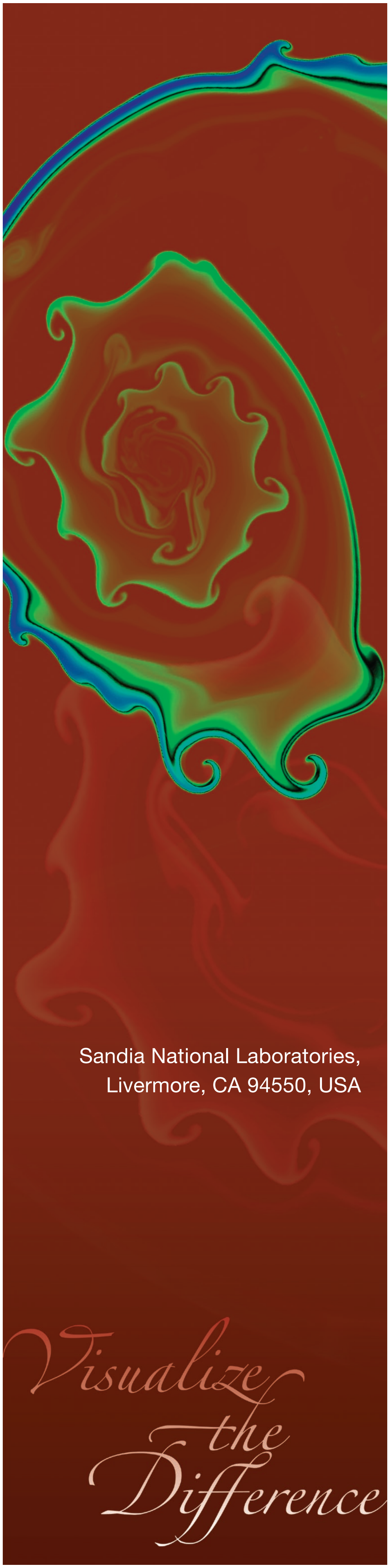
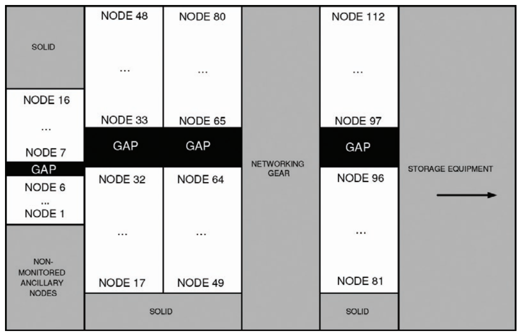
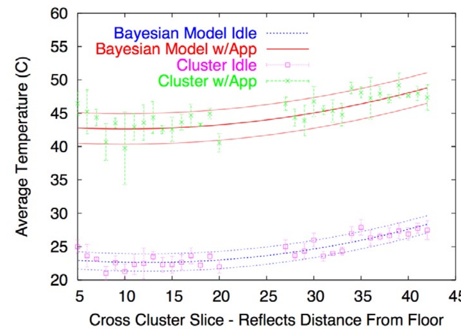


Sandia National Laboratories,
Livermore, CA 94550, USA

Bayesian Inference for Intelligent, Real-time Monitoring of Computational Clusters





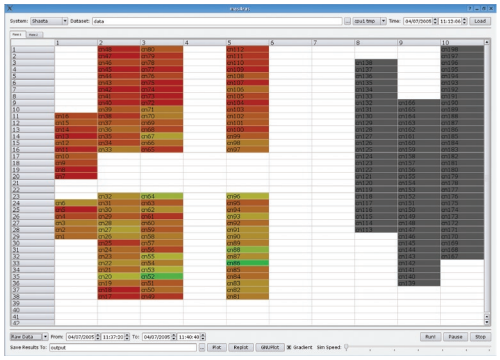
Cluster temperature profile is modeled by $T \sim N(Q(h),\sigma)$, where h and T respectively denote height and temperature and $N(Q(h),\sigma)$ is the normal distribution with mean Q(h) and variance σ. In this case, Q is a quadratic polynomial. The coefficients of Q and the value of σ are identified by Bayesian Inference. Inclusion of σ allows us to determine the likelihood for any T at a given h.

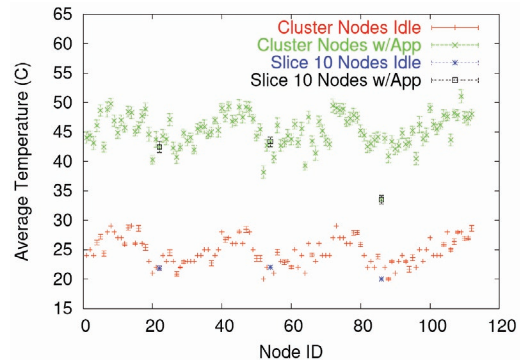
Current monitoring of computational clusters of sizes ranging from tens to tens of thousands of nodes is typically performed in a simple fashion: data are obtained from each node and a predefined rule set is invoked on a per-node basis to any node whose values cross a predetermined threshold. Though this methodology is well suited to smaller clusters, we can use statistical modeling and Bayesian inference to add a great deal of intelligence to the process, as well as determining problems sooner than is possible using static thresholds.

We have developed a tool that detects aberration from normal behavior, where normal is defined as within some administrator-defined likelihood (e.g., 95%), given a model inferred from peers' behaviors. A peer is defined as any node that is under the same workload and has the same environmental conditions. We use Bayesian inference in order to dynamically infer such models that consist of probability distributions of observing sensor values under a variety of conditions (e.g., varying load, idle, ambient temperature, etc.). Individual node values that have small likelihood given the current applicable model are then flagged as aberrant. This can be a much earlier indicator of problems than waiting for the crossing of some threshold that is necessarily set high to preclude too many false alarms.

We provide various built-in models, as well as allowing users to define their own. Since the machine environment is ideally uniform across all nodes, then highly complex models are unlikely to be required. In fact, if such models are required, then it is likely that there are problems in the cluster design and/or the room configuration. Noise terms are included in the models to account for secondary or

ADVANCED SIMULATION & COMPUTING™





Outliers are automatically detected by likelihood. A color-coded graphical display aids in awareness of outliers and patterns. Here, aberrant node 86 is detected by its value having less than 95% likelihood.

unmeasured/unconsidered effects. Overly large noise terms and poor fits to models indicate that the chosen model contributors and the model itself are poorly/inappropriately chosen.

Using the available data (that can also include training data and expert knowledge), we update the probabilistic model in real time as data become available using Bayes Theorem. In addition, a goodness-of-fit estimate is calculated for the models.

Unexpected or unaccountable shifts in the current models indicate global problems or changes in the room. The ability to dynamically and automatically update the model not only allows us to stay current with conditions, but also enables us to seamlessly apply the code to a new cluster with different hardware and environmental conditions. This tool facilitates real-time visualization of the effects of equipment and room configuration changes on the cluster and can be used to tune the cluster or room configurations.

for more information

J. M. Brandt, brandt@ca.sandia.gov,
A. C. Gentile, gentile@ca.sandia.gov
D. J. Hale, djhale@sandia.gov
Y. M. Marzouk, ymarzou@sandia.gov
P. P. Pébay, pppebay@sandia.gov

SAND 2005-6844P
Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

NNSA National Nuclear Security Administration