Center for Experimental Research in Computer Systems (CERCS)

A CISE-funded Center

Georgia Institute of Technology, Karsten Schwan, Director, 404.894.2589, schwan@cc.gatech.edu
Georgia Institute of Technology, Calton Pu, 404.385.1106, calton@cc.gatech.edu
Georgia Institute of Technology, Doug Blough, 404.385.1271, doug.blough@ece.gatech.edu
Georgia Institute of Technology, Sudhakar Yalamanchili, 404.894.2940, sudha@ece.gatech.edu
Ohio State University, Jay Ramanathan, 614.565.4187, jaram@cse.ohio-state.edu

Center website: http://www.cercs.gatech.edu/

Attaining Predictably Fast Responses: The "Travelport Flight Shopping Engine"

Researchers at the Center for Experimental Research in Computer Systems (CERCS) have been working with industry to improve the performance of Travelport, a primary computer application (shopping engine) for making travel arrangements (flight shopping). Several prototypes have been developed in the areas of early problem detection, traffic distribution, and multi-core processing. The result of this work was a 35% improvement in the average response time of end user requests for travel options (e.g., possible flights), and a 10% reduction in failures to meet response time requirements for such requests. The economic impact of the breakthrough translated into 20% less hardware purchases by Travelport (the shopping engine provider) and a more competitive position for Travelport's shopping engine in the market place. These technology breakthroughs, therefore, improved an existing product for airline shopping, i.e., Travelport's flight shopping engine. These breakthroughs are relevant to all products that require extensive calculations such as flight shopping, ticketing, hotel reservations, and similar. More importantly, they are relevant to any application for which the amount of processing performed for each request can vary significantly depending on the nature of the requests. For more information contact Karsten Schwan, 404.894.2589, schwan@cc.gatech.edu.

Power-Efficient Data Centers

In the U.S., data center facilities consume approximately 2% of all electricity, with an estimated growth rate of 12% per year. The majority of this datacenter power is either consumed by the computers themselves or by the HVAC units required to keep the computers from overheating. The use of virtual machines has allowed data center managers to conserve power by consolidating data processing onto fewer servers in times low demand. However, that consolidation had not previously been performed while taking into account its impact upon the cooling demands of the data center.

The Georgia Tech CERCS project 'CoolIT', developed in collaboration with the Mechanical Engineering department, allows the synergistic and cooperative management of IT and cooling system resources, adjusting air velocity and the location of active computing simultaneously to minimize power consumption. Initial results attained in CoolIT highlight the interesting trade-offs faced by a coordinated management solution. For instance, at lower cooling air velocities, for a homogeneous set of server systems, an awareness of hot spots in the data center permits the IT management system to operate at close to 100% maximum performance load, whereas without such awareness, there are situations in which only 40% load is achieved. These results illustrate the significant benefits of coordination between computer and facilities power management. For more information, contact Karsten Schwan, 404.894.2589, schwan@cc.gatech.edu.

Insights from High Performance Computing Service Augmentation

Service augmentation is a new concept developed by researchers at the Georgia Tech Center for Experimental Research in Computer Systems (CERCS) to benefit high performance, I/O intensive applications. Receiving the “Best Paper” award at the 2006 Conference on Cluster Computing, service augmentation uses runtime binary code generation to augment complex codes with new functionality, as and when needed by these applications. Originally intended for and applied to enable the online visualization of high performance simulations running on supercomputers, this concept is now showing promise for commercial applications. For retail forecasting, for instance, the concept can be used to efficiently extract data from the complex internal structures used by these codes to make it useful for display to managers and planners. Ongoing work with a startup company is exploring this use of the idea.
For more information, contact Matthew Wolf, 404.385.1278, mwolf@cc.gatech.edu or Karsten Schwan, 404.894.2589, schwan@cc.gatech.edu.

Automated Performance Characterization (APC) During Staging

As the complexity of large-scale enterprise applications increases, providing performance verification through iterative staging becomes an important part of reducing business risks associated with violating sophisticated service-level agreements. Staging is the deployment and validation of the entire application in a controlled testing environment. Currently, performance verification during the staging process is accomplished through either an expensive, cumbersome manual approach or ad hoc automation. Researchers at the Center for Experimental Research in Computer Systems (CERCS) have developed an

automated approach for supporting the monitoring and performance analysis of distributed multi-tiered applications. The process uses code generation and machine learning to automatically determine service level agreement satisfaction and to locate bottlenecks in candidate application deployment scenarios. The intent is to detect bottlenecks and enable the system to self-tune by addressing the detected bottlenecks through redesign; thus using iterative staging until a satisfactory system is designed. Evaluation tools are used to examine and illustrate the effectiveness of the APC monitoring and analysis process in successfully locating performance limitations as part of automated iterative staging. For more information, contact Karsten Schwan, 404.894.2589, schwan@cc.gatech.edu.

Opinion Mining

Opinion Mining (OM) tool set with the following capabilities – search and extract of relevant data from multiple sources, aggregate and pre-process data, analyze data using techniques such as clustering and sentiment detection, summary generation and user group detection. The OM tools have the following features (at a high level):

- Data Extraction: Automatic extraction of unstructured data from multiple data sources for further processing. Cluster Analysis: Process of partitioning the given data into several groups or clusters, so that objects from the same cluster are more similar to each other than objects from different clusters.
- Sentiment Analysis: Determines the attitude of the writer of a given topic, such as finding out whether people have positive opinion or negative opinion.
- User profiling: Grouping users based on common interests, demographics and location.
- Summarization: Creating a summary of data corresponding to each cluster.

With the OM tool set, the enterprise extracts and aggregates reader comments from news blogs and websites. The data is then processed by using the clustering analysis component used to identify prominent themes. Following this sentiment analysis could be applied to extract the underlying sentiment of each cluster. For more information, contact Jay Ramanathan, 614.565.4187, jayram@cse.ohio-state.edu.