Workshop on Managing Systems via Log Analysis and Machine Learning Techniques (SLAML '10)

Sponsored by USENIX, the Advanced Computing Systems Association

http://www.usenix.org/slaml10

October 3, 2010   **Vancouver, BC, Canada**

SLAML '10 will be held immediately preceding the 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI '10), which will take place October 4–6, 2010.

SLAML '10 combines the Workshop on the Analysis of System Logs (WASL) and the Workshop on Tackling Computer Systems Problems with Machine Learning Techniques (SysML). We welcome contributions related to either of these important and related topics.

> **NEW! We are now soliciting both 8-page full papers and 3-page position papers.** See the revised schedule and Submission Instructions below for details.

Important Dates

Full paper submissions due: *Sunday, July 11, 2010, 11:59 p.m. PDT*
Notification of acceptance: *Friday, August 20, 2010*
Final papers due: *Thursday, September 16, 2010*

Workshop Organizers

Program Co-Chairs

Greg Bronevetsky, *Lawrence Livermore National Laboratory*
Kathryn Mohror, *Lawrence Livermore National Laboratory*
Alice Zheng, *Microsoft Research*

Program Committee

Trishul Chilimbi, *Microsoft Research*
Anton Chuvakin, *Security Warrior Consulting*
Ira Cohen, *HP Labs*
Archana Ganapathi, *University of California, Berkeley*
Moises Goldszmidt, *Microsoft Research*
Vijay K. Gurbani, *Bell Laboratories, Alcatel-Lucent*
Daniel V. Klein, *LoneWolf Systems*
Ethan Miller, *University of California, Santa Cruz*
Priya Narasimhan, *Carnegie Mellon University*
Adam Oliner, *Stanford University*
Marcus J. Ranum, *Tenable Security*
Mark Sandler, *Google*
Jon Stearley, *Sandia National Laboratory*
Lin Tan, *University of Waterloo*
Eno Thereska, *Microsoft Research*
Shobha Venkataraman, *AT&T Research*
Wei Xu, *University of California, Berkeley*

Overview

Modern large-scale systems are challenging to manage. Fortunately, as these systems generate massive amounts of performance and diagnostic data, there is an opportunity to make system administration and development simpler via automated techniques to extract actionable information from the data. This workshop addresses this problem in two thrusts: (i) the analysis of raw system data logs and (ii) the application of machine learning to systems problems. We expect the large overlap in these topics to promote a rich interchange of ideas between the areas. **NEW!** We are soliciting both 8-page full papers and 3-page position papers. See the Submission Instructions below for details.

Log Analysis: It is well known that raw system logs are an abundant source of information for the analysis and diagnosis of system problems

and prediction of future system events. However, a lack of organization and semantic consistency between system data from various software and hardware vendors means that most of this information content is wasted. Current approaches to extracting information from the raw system data capture only a fraction of the information available and do not scale to the large systems common in business and supercomputing environments. It is thus a significant research challenge to determine how to better process and combine information from these data sources.

Machine Learning: The large scale of available data requires automated and machine-assisted analysis. Statistical machine learning techniques have recently shown great promise in meeting the challenges of scale and complexity in datacenter-scale and Internet-scale computing systems. However, applying these techniques to real systems scenarios requires careful analysis and engineering of the techniques to fit them to specific scenarios; there is sometimes also the opportunity to develop new algorithms specific to systems scenarios. This workshop thrust thus also presents a substantial research area: the exploration of new approaches to using machine learning to help us understand, measure, and diagnose complex systems.

Topics

Topics include but are not limited to:

- Reports on publicly available sources of sample system logs
- Prediction of malfunction or misuse based on system data
- Statistical analysis of system logs
- Applications of Natural-Language Processing (NLP) to system data
- Techniques for system log analysis, comparison, standardization, compression, anonymization, and visualization
- Applications of log analysis to system administration problems
- Use of machine learning techniques to address reliability, performance, power management, security, fault diagnosis, scheduling, or manageability issues
- Challenges of scale in applying machine learning to large systems
- Integration of machine learning into real-world systems and processes
- Evaluating the quality of learned models, including assessing the confidence/reliability of models and comparisons between different methods

Submission Instructions

Interested speakers should submit their full papers or position papers by July 11, 2010, via the Web submission form on the SLAML '10 Call for Papers Web site, http://www.usenix.org/slaml10/cfp. All papers will be subject to peer review under conference standards. Authors may choose to submit a paper anonymously or with author names visible to reviewers. Experience reports and papers on work in progress are welcome as long as there is a clear contribution. Paper submissions may be accepted as papers for the regular workshop program or as posters for the poster session. Submissions must be in PDF format. Full papers must be no longer than eight 8.5" x 11" pages and position papers must be no longer than three pages, including figures and tables, but not including references. Papers must be formatted in two columns, using 10 point type on 12 point (single-spaced) leading, with the text block being no more than 6.5" wide by 9" deep.

All papers will be available online to registered attendees before the workshop. If your accepted paper should not be published prior to the event, please notify production@usenix.org. The papers will be available online to everyone beginning on the first day of the workshop, October 3, 2010.

Papers accompanied by nondisclosure agreement forms will not be considered. Accepted submissions will be treated as confidential prior to publication on the USENIX SLAML '10 Web site; rejected submissions will be permanently treated as confidential.

Simultaneous submission of the same work to multiple venues, submission of previously published work, or plagiarism constitutes dishonesty or fraud. USENIX, like other scientific and technical conferences and journals, prohibits these practices and may take action against authors who have committed them. See the USENIX Conference Submissions Policy at http://www.usenix.org/submissionspolicy for details.

Questions? Contact your program co-chairs, slaml10chairs@usenix.org, or the USENIX office, submissionspolicy@usenix.org.