Session III Discussion Notes

Scribe: Nahur Fonseca, Boston University

Towards Realistic File-System Benchmarks with CodeMRI

Presented by Nitin Agrawal

Discussant: Arif Merchant

This paper presents a preliminary investigation of a new interesting method to build easy to use and representative workloads to evaluate file-systems. The basic principle is to instrument a file-system, with the aid of domain knowledge, to breakdown workloads into micro-functions; then a subset of micro-functions are selected as predictors of a particular type of workload function; and finally a synthetic workload can be reconstructed from the micro-functions.

The following questions were discussed:

1) There is a strong dependence on domain knowledge in order to define the set of micro-functions. How one could create a micro workload without domain knowledge? This is particularly important when one needs to evaluate a new system.

Domain knowledge is important. For instance knowing about the existence of cache and how it interacts with the file-system is important to differentiate between cached-reads and uncached ones.

One approach is to collect information using domain knowledge is to use tools such as *strace*. But it need to be combined with other instrumentations as well.

2) What is the right level of granularity of domain knowledge to include in the benchmarking in order to derive the synthetic models?

It is not clear what the granularity should be. There is no clear cut.

Two important points need to be observed. First, the benchmark has to be generic, and avoid being to adherent to one particular system. In constructing a generic benchmark, many systems need to be profiled and the common patterns need to be extracted. Second, the amount of domain knowledge that is included in the benchmarks has to be high level enough; otherwise, one could use a trace, instead of the benchmark.

3) Workload reconstruction. This part of the method is not ready yet, since this is a work in progress. What are the nice ways that this could be done? What are the challenges?

The reconstruction of workloads is complex because of the interactions of the different components of a file-system, specially in multi-threaded environments.

One limitation of the method of deconstruction of workloads to create the micro-workloads is the reliance of linear models to select the subset of predictors. Multiple threads, caching and applications may create non-linear dependences in the workload functions that are exercised.

One suggestion was to use some type of algebra in the reconstruction of workloads. However one requirement of this approach is that the different parts of the model, in this case, the micro-workloads, be composable; however, in this set up, the dependencies caused by access patterns, high multi threading, caching, may make it hard to apply some type of algebra to solve this problem.

Synchronized Network Emulation: Matching prototypes with complex simulations

Presented by Elias Weingärtner

Discussant: Leana Golubchik

In evaluating a system, one can use simulation, real implementation or emulation, which sits in between. In this paper, a synchronized emulation is proposed. It is a hybrid system where part is being simulated and part is being emulated. The basic problem they solve is to synchronize simulations that are too complex (i.e. need more time to run than what the real system would) with the emulated systems. The basic idea of the solution is to put the emulated systems in virtual hosts, so that the time as perceived by the emulated system can be frozen or slowed down as needed.

Some of the discussion was related to clarification questions:

1) How the emulated systems do not cause time-outs or other timing problems once they are frozen?

There are no stand-alone real systems running during the emulation. A host is either being simulated or emulated in a virtual host. So the clock that the hosts being emulated see is the clock of the virtual host; therefore, once they are frozen, all timers freeze as well; No time-outs or other time synchronization problems will occur.

2) How the virtual host interfere with the evaluation of the system under scrutiny?

This question has to do with the fact that virtualization creates overhead. So the question is: if one is evaluating the performance of a system, how can one discount the performance degradation due to virtualization?

The answer was the following: at the moment, the authors focus is on the correctness of the system, and not with performance. The interference of the virtual host in the performance of the emulated system is subject of future work.

3) How many emulated hosts can co-exist in a virtual host?

This question is related to the former one. The overload of the virtualization increases with the number of virtual hosts in a single machine. How would that affect the results of the emulation?

The answer to that question was that a scalability study needs to be carried out as future work. The experiments in the paper were run with at most two virtual hosts per machine. Since one of the reasons why emulation is needed is to run large scale experiments, this is an important problem to solve.

What are our standards for validation of measurement-based networking research?

Presented by Walter Willinger

Discussant: John C.S. Lui

Walter claimed the community attention to four points to improve the quality of research in network measurements: quality of data, quality of statistics analysis, quality of models, and quality of validation.

1) Quality of data.

Many datasets are available for download to be used in network research. However, the lack of meta-data about this datasets makes it hard to ascertain if this datasets have the right information to support the proposed study.

The datasets need not be of the finest quality in order to extract useful information. The experts who are analyzing these datasets need to bring in the domain knowledge about communication networks and previous results to evaluate what can be learned from the particular dataset at hand.

For instance, it may not be possible to assert that a particular dataset feature presents long-range-dependence properties, but it may be possible to assert that the same feature distribution is highly variable.

2) Availability of data.

Can high quality data be public available?

A few representatives of industry leaders in communication networks are willing to share their datasets, regarded that there is a secure way of preserving the secrecy of private and strategic information.

One idea was proposed to have a data repository where questions could be submitted and only the responses would be sent back.

3) Interplay with other sciences.

It is not only network research that has problems with data collection. Our colleagues from biology and medical sciences have the same problem. Sometimes exacerbated by the rarity of the phenomenon under study.

Crossing boundaries between sciences can be done with the necessary caution of gathering the required domain knowledge. For instance, in the case studies presented in the paper, one could simply check that there is not so advanced network technology to create network elements with high degree of connectivity at the core of the Internet. Such technology just does not exist yet.

4) What could be thought in graduate schools to improve the science of Computer Science and other sciences as well?

The problems illustrated in the paper are common to other areas as well. Can the community come up with ideas to help make the science more rigorous? Even adding something to the curriculum of network research.