PARALLEL DATA LABORATORY
CARNEGIE MELLON UNIVERSITY

PDL Packet Fall Update

NEWSLETTER ON PDL ACTIVITIES AND EVENTS • FALL 2009
http://www.pdl.cmu.edu/

PDL CONSORTIUM MEMBERS

American Power Conversion
Data Domain
EMC
Facebook
Google
Hewlett-Packard Labs
Hitachi
IBM
Intel
LSI
Microsoft Research
NEC Laboratories
NetApp
Oracle
Seagate Technology
Sun Microsystems
Symantec
VMware

CONTENTS

THE PDL PACKET

EDITOR
Joan Digney

CONTACTS
Greg Ganger
PDL Director

Bill Courtright
PDL Executive Director

Karen Lindenfelser
PDL Business Administrator

The Parallel Data Laboratory
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891

TEL 412-268-6716

FAX 412-268-3010

http://www.pdl.cmu.edu/Publications/

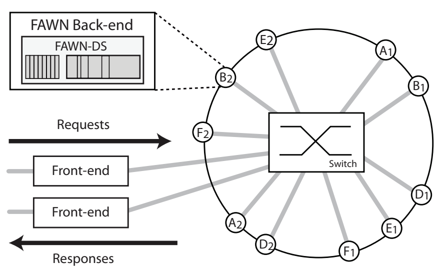
SELECTED RECENT PUBLICATIONS

FAWN: A Fast Array of Wimpy Nodes

Andersen, Franklin, Kaminsky, Phanishayee, Tan & Vasudevan

Proc. 22nd ACM Symposium on Operating Systems Principles (SOSP 2009), Big Sky, MT. October 2009. Best Paper Award.

This paper presents a new cluster architecture for low-power data-intensive computing. FAWN couples low-power embedded CPUs to small amounts of local flash storage, and balances computation and I/O capabilities to enable efficient, massively parallel access to data. The key contributions of this paper are the principles of the FAWN architecture and the design and implementation of FAWN-KV—a consistent, replicated, highly available, and high-performance key-value storage system built on a FAWN prototype. Our design centers around purely log-structured datastores that provide the basis for high performance on flash storage, as well as for replication and consistency obtained using chain replication on a consistent hashing ring. Our evaluation demonstrates that FAWN clusters can handle roughly 350 key-value queries per Joule of energy—two orders of magnitude more than a disk-based system.



FAWN-KV Architecture.

...And eat it too: High read performance in write-optimized HPC I/O middleware file formats

Polte, Lofstead, Bent & Gibson

4th Petascale Data Storage Workshop. Held in conjunction with SC09. November 15, 2009. Portland, Oregon.

As HPC applications run at larger scales, in order to maintain an acceptable ratio of time spent performing useful computation work to time spent performing I/O, write bandwidth to the underlying storage system must increase proportionally to the increase in the computation size. Unfortunately, popular scientific self-describing file formats such as netCDF and HDF5 are designed with a focus on portability and flexibility rather than write bandwidth. To provide sufficient write bandwidth to continue to support the demands of scientific applications, the HPC community has developed a number of I/O middleware layers. A few examples include the Adaptable IO System (ADIOS), a library developed at Oak Ridge National Laboratory and the Parallel Log-structured Filesystem (PLFS), a stackable FUSE filesystem developed at Los Alamos National Laboratory. While both of these middleware layers dramatically improve write performance by writing in a write-optimized log-structured file format, the obvious concern with any write optimized file format

continued on page 3

October 2009

Adrian Perrig Wins Award for Innovative Cybersecurity Research

Adrian Perrig was awarded a Security 7 Award from Information Security magazine for innovative cybersecurity research in academia. Perrig, technical director of Carnegie Mellon CyLab, a professor in the departments of Electrical and Computer Engineering and Engineering and Public Policy, and the School of Computer Science, will be recognized in the magazine's October issue. The magazine's editor, Michael S. Mimoso, said the awards recognize the achievements of security practitioners and researchers in a variety of industries, including education.

--CMU 8.5x11 News, Oct 15, 2009

October 2009

Honorable Mention for Leskovec's Dissertation

Congratulations to Jure Leskovec who has been awarded an honorable mention in the SCS Dissertation of the year competition for his thesis on "Dynamics of Large Networks." We wish him luck as his dissertation is also being submitted as one of Carnegie Mellon's entries for the ACM doctoral dissertation award.

October 2009

Best Paper Award from SOSP'09 in Big Sky, Montana!

Huge congratulations to Amar Phanishayee, Jason Franklin, Lawrence Tan, Vijay Vasudevan and Dave Andersen on their best paper award at the 22nd ACM Symposium on Operating Systems Principles (SOSP '09). Their paper "FAWN: A Fast Array of Wimpy Nodes" presents a new cluster architecture for low-power data-intensive computing.

August 2009

Nikos Hardavellas Appointed to June & Donald Brewer Chair of EE/CS at Northwestern

Congratulations to Nikos, June and Donald Brewer Assistant Professor of Electrical Engineering and Computer Science at Northwestern University. He has been appointed to the endowed chair for a two-year period from September 1, 2009 to August 31, 2011. Along with the title and honor, Prof. Hardavellas will receive a discretionary fund for each of the two years. This chair is awarded to Northwestern University's very best young faculty in the McCormick School of Engineering.

August 2009

Cranor Receives NSF Funding for Interdisciplinary Doctoral Program in Privacy & Security

Associate Professor Lorrie Cranor and her colleagues received a five-year, $3 million grant from the National Science Foundation (NSF) to establish a Ph.D. program in usable privacy and security. "Carnegie Mellon's CyLab Usable Privacy and Security (CUPS) Doctoral Training Program will offer Ph.D. students a new cross-disciplinary training experience that helps them produce solutions to ongoing tensions between security, privacy and usability," said Cranor, associate professor in the Institute for Software Research, the Department of Engineering and Public Policy and Carnegie Mellon CyLab. She noted that students will be actively involved in Carnegie Mellon's broad usable privacy and security research, which spans three major approaches: finding ways to build systems that "just work" without involving humans in security-critical functions; finding ways of making secure systems intuitive and easy to use; and finding ways to effectively teach humans how to perform security-critical tasks.

--CMU 8.5x11 News, August 27, 2009

August 2009

Priya follows up the YinzCam with iBurgh

Pittsburgh is the first U.S. city with its own iPhone app. iBurgh, developed by Priya Narasimhan and her research group, allows users to take a picture of civic problems such as potholes, graffiti or other hazards and directly upload them, accompanied by a GPS location, to city council and other municipal administration authorities for review. Previous to the iBurgh app, Priya and her group launched the YinzCam, another mobile phone app which allows hockey fans to view replays and alternate action angles at Pittsburgh Penguins hockey games on their phones or other handheld WiFi devices.

To top off all Priya's good news, YinzCam made Network World's top 10 list of sports innovations to love! (http://www.networkworld.com/slideshows/2009/081809-sports-technologies.html -- slide 10).

continued on page 8

continued from page 1

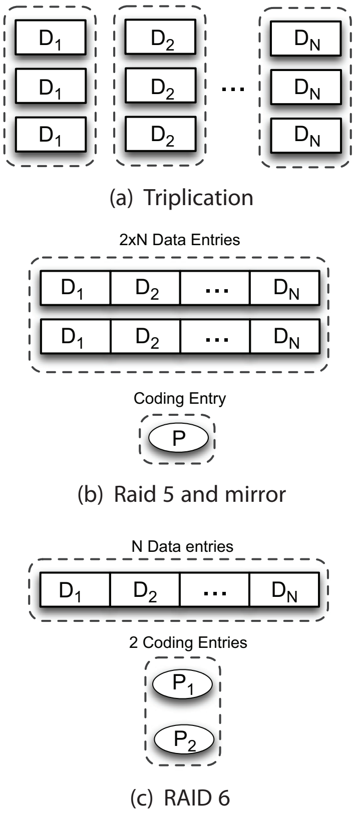
would be a corresponding penalty on reads. Simulation results require efficient read-back for visualization and analytics and while most checkpoint files are never used, the efficiency of restart is still important in the face of inevitable failures. The utility of write speed improving middleware would be greatly diminished without acceptable read performance. In this paper we examine the read performance on large parallel machines and compare these to reading data either natively or to other popular file formats. We compare the reading performance in two different scenarios: 1) Reading back restarts from the same number of processors which wrote the data and 2) Reading back restart data from a different number of processors which wrote the data and demonstrate that the log-structured file formats actually improve read-back performance due to the particular writing and read back patterns of HPC checkpoints and applications.

DiskReduce: RAID for Data-Intensive Scalable Computing

Fan, Tantisiriroj, Xiao & Gibson

4th Petascale Data Storage Workshop. Held in conjunction with SC09. November 15, 2009. Portland, Oregon.

Data-intensive file systems, developed for Internet services and popular in cloud computing, provide high reliability and availability by replicating data, typically three copies of everything. While high performance computing, which has comparable scale, and smaller scale enterprise storage systems get similar tolerance for multiple failures from lower overhead erasure encoding, or RAID, organizations. DiskReduce is a modification of the Hadoop file system (HDFS) enabling asynchronous compression of initially triplicated data down to RAID-class redundancy overheads, and asynchronous repair of lost data. In addition to increasing a cluster's storage capacity as seen by its users by up to a factor of three, DiskReduce can delay encoding long enough to deliver the performance benefits of multiple data copies. DiskReduce also gathers data into encoding groups that is likely to be deleted closely enough in time to avoid re-encoding data as it is deleted



Codewords providing protection against double node failures.

Co-scheduling of Disk Head Time in Cluster-based Storage

Wachs & Ganger

28th International Symposium On Reliable Distributed Systems September 27-30, 2009. Niagara Falls, New York, U.S.A.

Disk timeslicing is a promising technique for storage performance insulation. To work with cluster-based storage, however, timeslices associated with striped data must be co-scheduled on the corresponding servers. This paper describes algorithms for determining global timeslice schedules and mechanisms for coordinating the independent server activities. Experiments with a prototype show that, combined, they can provide performance insulation for workloads sharing a storage cluster—each workload realizes a configured minimum efficiency within its timeslices regardless of the activities of the other workloads.

Improving OLTP Scalability using Speculative Lock Inheritance

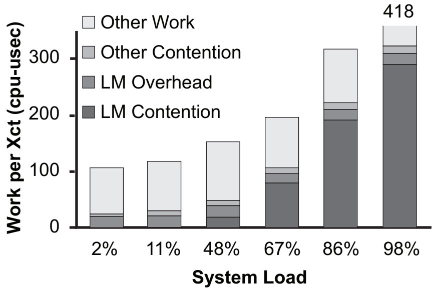
Johnson, Pandis & Ailamaki

35th International Conference on Very Large Data Bases (VLDB2009), Lyon, France, August 2009.

Transaction processing workloads provide ample request level concurrency which highly parallel architectures can exploit. However, the resulting heavy utilization of core database services also causes resource contention within the database engine itself and limits scalability. Meanwhile, many database workloads consist of short transactions which access only a few database records each, often with stringent response time requirements. Performance of these short transactions is determined largely by the amount of overhead the database engine imposes for services such as logging, locking, and transaction management.

This paper highlights the negative scalability impact of database locking, an effect which is especially severe for

continued on page 4



Lock manager overhead as system load increases.

continued from page 3

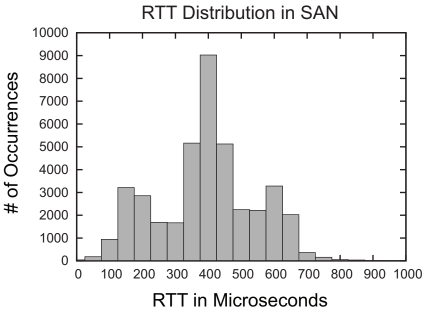
short transactions running on highly concurrent multicore hardware. We propose and evaluate Speculative Lock Inheritance, a technique where hot database locks pass directly from transaction to transaction, bypassing the lock manager bottleneck. We implement SLI in the Shore-MT storage manager and show that lock inheritance fundamentally improves scalability by decoupling the number of simultaneous requests for popular locks from the number of threads in the system, eliminating contention within the lock manager even as core counts continue to increase. We achieve this effect with only minor changes to the lock manager and without changes to consistency or other application-visible effects.

Safe and Effective Fine-grained TCP Retransmissions for Datacenter Communication

Vasudevan, Phanishayee, Shah, Krevat, Andersen, Ganger, Gibson & Mueller

SIGCOMM'09, August 17–21, 2009, Barcelona, Spain.

This paper presents a practical solution to a problem facing high-fan-in, high-bandwidth synchronized TCP workloads in datacenter Ethernets|the TCP incast problem. In these networks, receivers can experience a drastic reduction in application throughput when simultaneously requesting data from many servers using TCP. Inbound data overfills small switch buffers, leading to TCP timeouts lasting hundreds of milliseconds. For many datacenter workloads that have a barrier synchronization requirement (e.g., filesystem reads and parallel data-intensive queries), throughput is reduced by up to 90%. For latency-sensitive applications, TCP timeouts in the datacenter impose delays of hundreds of milliseconds in networks with round-trip-times in microseconds. Our practical solution uses high-resolution timers to enable microsecond-granularity TCP timeouts. We demonstrate that this technique is effective in avoiding TCP incast collapse in simulation and in real-world experiments. We show that eliminating the minimum retransmission timeout bound is safe for all environments, including the wide-area.



During an incast experiment on a cluster RTTs increase by 4 times the baseline RTT (100μs)on average with spikes as high as 800μs. This produces *RTO* values in the range of 1-3ms, resulting in an RTO_{min} of 1ms being as effective as 200μs in today's networks.

On the Inapproximability of M/G/K: Why Two Moments of Job Size Distribution are Not Enough

Gupta, Harchol-Balter, Dai & Zwart

Queueing Systems: Theory and Applications, August 2009.

The M/G/K queueing system is one of the oldest models for multiserver systems and has been the topic of performance papers for almost half a century. However, even now, only coarse approximations exist for its mean waiting time. All the closed-form (nonnumerical) approximations in the literature are based on (at most) the first two moments of the job size distribution. In this paper we prove that no approximation based on only the first two moments can be accurate for all job size distributions, and we provide a lower bound on the inapproximability ratio, which we refer to as "the gap." This is the first such result in the literature to address "the gap." The proof technique behind this result is novel as well and combines mean value analysis, sample path techniques, scheduling, regenerative arguments, and asymptotic estimates. Finally, our work provides insight into the effect of higher moments of the job size distribution on the mean waiting time.

The Perspective Home Storage team at the PDL Spring Industry Visit Day. From L to R, Zoheb Shivani, Nitin Gupta, Brandon Salmon and Michelle Mazurek.

RTG: A Recursive Realistic Graph Generator using Random Typing

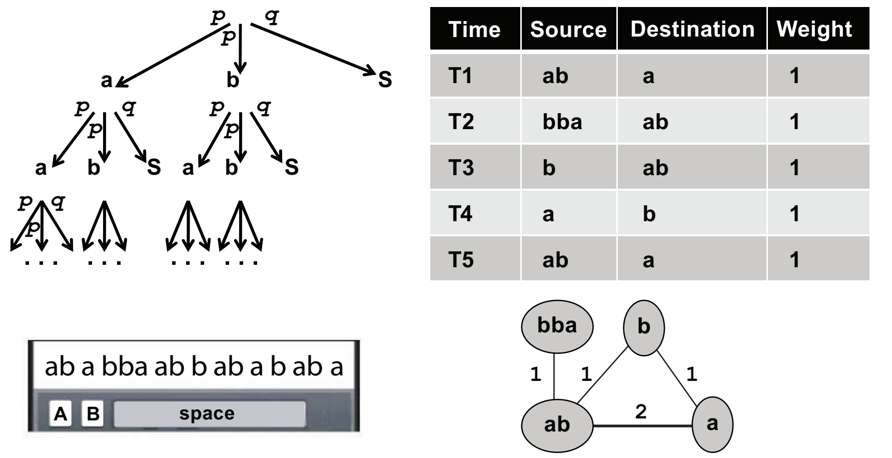
Akoglu & Faloutsos

ECML PKDD, Bled, Slovenia, Sept. 2009. Best Knowledge Discovery Paper award.

We propose a new, recursive model to generate realistic graphs, evolving over time. Our model has the following properties: it is (a) flexible, capable of generating the cross product of weighted/unweighted, directed/ undirected, uni/bipartite graphs; (b) realistic, giving graphs that obey eleven static and dynamic laws that real graphs follow (we formally prove that for several of the (power) laws and we estimate their exponents as a function of the model parameters); (c) parsimonious, requiring only four parameters. (d) fast, being linear on the number of edges; (e) simple, intuitively leading to the generation of macroscopic patterns. We empirically show that our model mimics two real-world graphs very well: Blognet (unipartite, undirected, unweighted) with 27K nodes and 125K

continued on page 5

continued from page 5



Time	Source	Destination	Weight
T1	ab	a	1
T2	bba	ab	1
T3	b	ab	1
T4	a	b	1
T5	ab	a	1

Illustration of the RTG-IE. Upper left: how words are (recursively) generated on a keyboard with two equiprobable keys, 'a' and 'b', and a space bar; lower left:a keyboard is used to randomly type words, separated by the space character; upper right: how words are organized in pairs to create source and destination nodes in the graph over time; lower right: the output graph; each node label corresponds to a unique word, while labels on edges denote weights.

edges; and Committee-to-Candidate campaign donations (bipartite, directed, weighted) with 23K nodes and 880K edges. We also show how to handle time so that edge/weight additions are bursty and self-similar.

No Downtime for Data Conversions: Rethinking Hot Upgrades

Dumitraş & Narasimhan

Carnegie Mellon University Parallel Data Lab Technical Report CMU-PDL-09-106. July 2009.

Unavailability in enterprise systems is usually the result of planned events, such as upgrades, rather than failures. Major system upgrades entail complex data conversions that are difficult to perform on the fly, in the face of live workloads. Minimizing the downtime imposed by such conversions is a time-intensive and error-prone manual process. We present Imago, a system that aims to simplify the upgrade process, and we show that it can eliminate all the causes of planned downtime recorded during the upgrade history of one of the ten most popular websites. Building on the lessons learned from past research on live upgrades in middleware systems, Imago trades off a need for additional storage resources for the ability to perform end-to-end, enterprise upgrades online, with minimal application-specific knowledge.

Optimal Power Allocation in Server Farms

Gandhi, Harchol-Balter, Das & Lefurgy

Proceedings of ACM SIGMETRICS 2009 Conference on Measurement and Modeling of Computer Systems. Seattle, WA, June 2009.

Server farms today consume more than 1.5% of the total electricity in the U.S. at a cost of nearly $4.5 billion. Given the rising cost of energy, many industries are now seeking solutions for how to best make use of their available power. An important question which arises in this context is how to distribute available power among servers in a server farm so as to get maximum performance.

By giving more power to a server, one can get higher server frequency (speed). Hence it is commonly believed that, for a given power budget, performance can be maximized by operating servers at their highest power levels. However, it is also conceivable that one might prefer to run servers at their lowest power levels, which allows more servers to be turned on for a given power budget. To fully understand the effect of power allocation on performance in a server farm with a fixed power budget, we introduce a queueing theoretic model, which allows us to predict the optimal power allocation in a variety of scenarios. Results are verified via extensive experiments on an IBM BladeCenter. We find that the optimal power allocation varies for different scenarios. In particular, it is not always optimal to run servers at their maximum power levels. There are scenarios where it might be optimal to run servers at their lowest power levels or at some intermediate power levels. Our analysis shows that the optimal power allocation is non-obvious and depends on many factors such as the power-to-frequency relationship in the processors, the arrival rate of jobs, the maximum server frequency, the lowest attainable server frequency and the server farm configuration. Furthermore, our theoretical model allows us to explore more general settings than we can implement, including arbitrarily large server farms and different power-to-frequency curves. Importantly, we show that the optimal power allocation can significantly improve server farm performance, by a factor of typically 1.4 and as much as a factor of 5 in some cases.

Perspective: Semantic Data Management for the Home

Salmon, Schlosser, L. Cranor & Ganger

;LOGIN: Vol. 34, No. 5

Distributed storage is coming home. An increasing number of home and personal electronic devices create, use, and display digitized forms of music, images, and videos, as well as more conventional files (e.g. financial records and contact lists). In-home

continued on page 6

continued from page 5

networks enable these devices to communicate, and a variety of device-specific and datatype-specific tools are emerging. The transition to digital homes gives exciting new capabilities to users, but it also makes them responsible for administration tasks which in other settings are usually handled by dedicated professionals.

Power Capping Via Forced Idleness

Gandhi, Harchol-Balter, Das, Kephart & Lefurgy

Workshop on Energy-Efficient Design (WEED 09) Austin, Texas, June 2009.

We introduce a novel power capping technique, IdleCap, that achieves higher effective server frequency for a given power constraint than existing techniques. IdleCap works by repeatedly alternating between the highest performance state and a low-power idle state, maintaining a fixed average power budget, while significantly increasing the average processor frequency. In experiments conducted on an IBM BladeCenter HS21 server across three representative workloads, IdleCap reduces the mean response time by up to a factor of 3 when compared to power capping using clock-throttling. Furthermore, we argue how IdleCap applies to next-generation servers using DVFS and advanced idle states.

In Search of an API for Scalable File Systems: Under the table or above it?

Patil, Gibson, Ganger, López, Polte, Tantisiriroj & Xiao

USENIX HotCloud Workshop 2009. June 2009, San Diego CA.

"Big Data" is everywhere – both the IT industry and the scientific computing community are routinely handling terabytes to petabytes of data [24]. This preponderance of data has fueled the development of data-intensive scalable computing (DISC) systems that manage, process and store massive data-sets in a distributed manner. For example, Google and Yahoo have built their respective Internet services stack to distribute processing (MapReduce and Hadoop), to program computation (Sawzall and Pig) and to store the structured output data (Bigtable and HBase). Both these stacks are layered on their respective distributed file systems, GoogleFS [12] and Hadoop distributed FS [15], that are designed "from scratch" to deliver high performance primarily for their anticipated DISC workloads.

However, cluster file systems have been used by the high performance computing (HPC) community at even larger scales for more than a decade. These cluster file systems, including IBM GPFS [28], Panasas PanFS [34], PVFS [26] and Lustre [21], are required to meet the scalability demands of highly parallel I/O access patterns generated by scientific applications that execute simultaneously on tens to hundreds of thousands of nodes. Thus, given the importance of scalable storage to both the DISC and the HPC world, we take a step back and ask ourselves if we are at a point where we can distill the key commonalities of these scalable file systems.

This is not a paper about engineering yet another "right" file system or database, but rather about how do we evolve the most dominant data storage API – the file system interface – to provide the right abstraction for both DISC and HPC applications. What structures should be added to the file system to enable highly scalable and highly concurrent storage? Our goal is not to define the API calls per se, but to identify the file system abstractions that should be exposed to programmers to make their applications more powerful and portable. This paper highlights two such abstractions. First, we show how commodity large-scale file systems can support distributed data processing enabled by the Hadoop/MapReduce style of parallel programming frameworks. And second, we argue for an abstraction that supports indexing and searching based on extensible attributes, by interpreting BigTable [6] as a file system with a filtered directory scan interface.

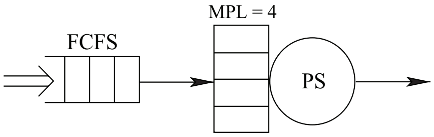
Self-Adaptive Admission Control Policies for Resource-Sharing Systems

Gupta & Harchol-Balter

Proceedings of ACM SIGMETRICS 2009 Conference on Measurement and Modeling of Computer Systems. Seattle, WA, June 2009.

We consider the problem of admission control in resource sharing systems, such as web servers and transaction processing systems, when the job size distribution has high variability, with the aim of minimizing the mean response time. It is well known that in such resource sharing systems, as the number of tasks concurrently sharing the resource is increased, the server throughput initially increases, due to more efficient utilization of resources, but starts falling beyond a certain point, due to resource contention and thrashing. Most admission control mechanisms solve this problem by imposing a fixed upper bound on the number of concurrent transactions allowed into the system, called the Multi-Programming-Limit (MPL), and making the arrivals which find the server full queue up. Almost always, the MPL is chosen to be the point that maximizes server efficiency. In this paper we abstract such resource sharing systems as a Processor Shar-

continued on page 7



A G/G/PS-MPL queue with MPL = 4. Only 4 jobs can simultaneously share the server. The rest must wait outside in FCFS order.

continued from page 6

ing (PS) server with state-dependent service rate and a First-Come-First-Served (FCFS) queue, and we analyze the performance of this model from a queueing theoretic perspective. We start by showing that, counter to the common wisdom, the peak efficiency point is not always optimal for minimizing the mean response time. Instead, significant performance gains can be obtained by running the system at less than the peak efficiency. We provide a simple expression for the static MPL that achieves near-optimal mean response time for general distributions. Next we present two traffic-oblivious dynamic admission control policies that adjust the MPL based on the instantaneous queue length while also taking into account the variability of the job size distribution. The structure of our admission control policies is a mixture of fluid control when the number of jobs in the system is high, with a stochastic component when the system is near-empty. We show via simulations that our dynamic policies are much more robust to unknown traffic intensities and burstiness in the arrival process than imposing a static MPL.

Access Control for Home Data Sharing: Attitudes, Needs and Practices

Mazurek, Arsenault, Bresee, Gupta, Ion, Johns, Lee, Liang, Olsen, Salmon, Shay, Vaniea, Bauer, L. Cranor, Ganger & Reiter

Carnegie Mellon University Parallel Data Lab Technical Report CMU-PDL-09-110, October 2009.

As digital content becomes more prevalent in the home, non-technical users are increasingly interested in sharing that content with others and accessing it from multiple devices. Not much is known about how these users think about controlling access to this data. To better understand this, we conducted semi-structured, in-situ interviews with 33 users in 15 households. We found that users create ad-hoc access-control mechanisms that do not always work; that their ideal polices are complex and multi-dimensional; that a priori policy specification is often insufficient; and that people's mental models of access control and security are often misaligned with current systems. We detail these findings and present a set of associated guidelines for designing usable access-control systems for the home environment.

System-Call Based Problem Diagnosis for PVFS

Kasick, Bare, Marinelli, Tan, Gandhi & Narasimhan

Proceedings of the 5th Workshop on Hot Topics in System Dependability (HotDep '09). Lisbon, Portugal. June 2009.

We present a syscall-based approach to automatically diagnose performance problems, server-to-client propagated errors, and server crash/hang problems in PVFS. Our approach compares the statistical and semantic attributes of syscalls across PVFS servers in order to diagnose the culprit server, under these problems, for different file-system benchmarks—dd, PostMark and IOzone—in a PVFS cluster.

Mochi: Visual Log-Analysis Based Tools for Debugging Hadoop

Tan, Pan, Kavulya, Gandhi & Narasimhan

Workshop on Hot Topics in Cloud Computing (HotCloud '09), San Diego, CA, on June 15, 2009.

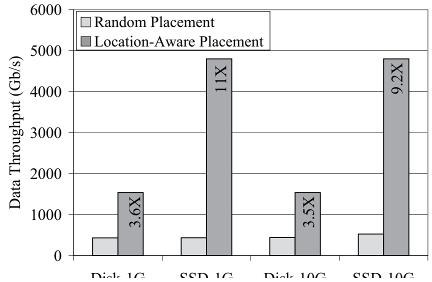
Mochi, a new visual, log-analysis based debugging tool correlates Hadoop's behavior in space, time and volume, and extracts a causal, unified control- and data-flow model of Hadoop across the nodes of a cluster. Mochi's analysis produces visualizations of Hadoop's behavior using which users can reason about and debug performance issues. We provide examples of Mochi's value in revealing a Hadoop job's structure, in optimizing real-world workloads, and in identifying anomalous Hadoop behavior, on the Yahoo! M45 Hadoop cluster.

Tashi: Location-aware Cluster Management

Kozuch, Ryan, Gass, Schlosser, O'Hallaron, Cipar, Krevat, López, Stroucken & Ganger

First Workshop on Automated Control for Datacenters and Clouds (ACDC'09), Barcelona, Spain, June 2009.

Big Data applications, those that require large data corpora either for correctness or for fidelity, are becoming increasingly prevalent. Tashi is a cluster management system designed particularly for enabling cloud computing applications to operate on repositories of Big Data. These applications are extremely scalable but also have very high resource demands. A key technique for making such applications perform well is Location-Awareness. This paper demonstrates that location-aware applications can outperform those that are not location aware by factors of 3-11 and describes two general services developed for Tashi to provide location-awareness independently of the storage system.



Performance comparison of location aware task placement with random task placement. The labels on the data bars show the performance improvement for the Location-Aware Placement relayive to the Random Placement.

continued from page 2

July 2009

Carlos Guestrin Wins Presidential Early Career Award

Carlos Guestrin, the Finmeccanica Assistant Professor of Computer Science and Machine Learning, has won a Presidential Early Career Award for Scientists and Engineers (PECASE), the highest honor bestowed by the U.S. government on scientists and engineers beginning their careers. He was nominated by the Department of Defense, which recognized him last year with the Office of Naval Research's Young Investigator Award.

The PECASE program recognizes 100 scientists and engineers who show exceptional potential for leadership at the frontiers of knowledge. "These extraordinarily gifted young scientists and engineers represent the best in our country," President Obama said. "With their talent, creativity and dedication, I am confident that they will lead their fields in new breakthroughs and discoveries and help us use science and technology to lift up our nation and our world."

For more on the PECASE award and Guestrin's other honors, visit http://www.cmu.edu/news/archive/2009/July/july10_guestrinaward.shtml

--CMU 8.5x11 News, July 16, 2009

June 2009

Greg Ganger Earns Prestigious HP Innovation Research Award

Greg Ganger, a professor of electrical and computer engineering and director of the Parallel Data Lab, is among 60 recipients worldwide who received 2009 HP Innovation Research Awards. The award encourages open collaboration with HP Labs for mutually beneficial, high-impact research.

Ganger, who also received an HP Innovation Lab Award in 2008, will lead a research initiative in collaboration with HP Labs focused on data storage infrastructure issues, based on his winning proposal "Toward Scalable Self-Storage."

Ganger was chosen from a group of nearly 300 applicants from more than 140 universities in 29 countries on a range of topics within the eight high-impact research themes at HP labs - analytics, cloud computing, content transformation, digital commercial print, immersive interaction, information management, intelligent infrastructure and sustainability.

Noah Smith, an assistant professor of language technologies and machine learning at CMU, also received this award.

"This award recognizes the ongoing innovative and cutting-edge work that Carnegie Mellon professors bring to all collaborative research efforts," said Mark S. Kamlet, Carnegie Mellon provost and senior vice president. "We are proud of their accomplishments and the vital impact their research will have for a variety of industry sectors."

--CMU 8.5x11 News, June 17, 2009

June 2009

Jure Leskovek Wins Doctoral Dissertation Award

Jure Leskovec won the prestigious 2009 SIGKDD Doctoral Dissertation Award from the Association of Computing Machinery's Special Interest Group on Knowledge Discovery and Data Mining for his thesis "Dynamics of Large Networks." He was advised by School of Computer Science Professor Christos Faloutsos, who also advised the 2008 runner-up Jimeng Sun. Leskovec will present a short summary of his work at the SIGKDD Conference in Paris on Sunday, June 28. For more: http://www.sigkdd.org/awards_dissertation.php

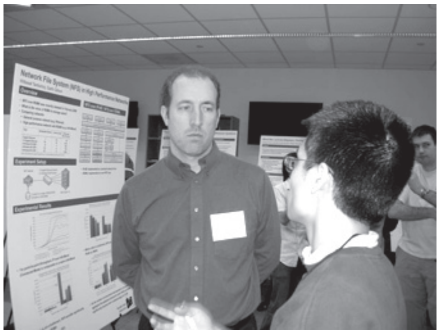
--CMU 8.5x11 News, June 4, 2009

May 2009

Polo Chau Selected as Future Thought Leader by Yahoo!

Yahoo! has named four Ph.D. students in the School of Computer Science among 20 winners of its inaugural Key Scientific Challenges program, which recognizes outstanding graduate-student researchers with the potential to become thought leaders in their fields. Polo Chau (advised by Christos Faloutsos) of the Machine Learning Dept. won recognition in the Search Technologies category. Each recipient receives $5,000 in unrestricted seed funding for their research, exclusive access to certain Yahoo! data sets and the opportunity to collaborate directly with Yahoo! scientists. This summer, they will attend a Yahoo! Graduate Student Summit to present and discuss their work with some of the top minds in academia and industry.

--CMU 8.5x11 News, May 14, 2009



Wittawat Tantisiriroj discusses his poster on Network File Systems in High Performance Networks with Michael Kozuch of Intel.