computer engineering graduate seminars

A place for cool research, thought-provoking questions, and free food!

Subscribe to the mailing list ececompseminar-list@ecn.purdue.edu for announcements.

Time & Place
Wed 12.30-1.30pm (may change on some weeks) | EE 317

Participate!

1. Choose up to three open slots below for your talk. Remember to leave your email there.
2. After your time is confirmed, directly edit the following schedule to add title, abstract, & bio
   1. Please do so at least one week before the talk
3. Email the talk info to Mary Ann.

Speaker: please edit the corresponding row to reflect your talk

Spring 2019

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker /(Prof)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 9</td>
<td>Yanjun Wang (Qiu)</td>
<td>A Decidable Logic for Tree Data-Structures with Measurements</td>
</tr>
<tr>
<td>Jan 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar 6</td>
<td>Yilong Zhang</td>
<td>Farewell to Servers: Hardware, Software, and Network Approaches towards Datacenter Resource Disaggregation</td>
</tr>
<tr>
<td>Mar 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar 20</td>
<td>Tiantu Xu</td>
<td></td>
</tr>
<tr>
<td>Mar 27</td>
<td>Felix Lin</td>
<td>Supercharge analytics over large IoT data: an OS approach</td>
</tr>
<tr>
<td>Apr 3</td>
<td>Sanjay Rao</td>
<td></td>
</tr>
<tr>
<td>Apr 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Talks

March 6, 2019: Farewell to Servers: Hardware, Software, and Network Approaches towards Datacenter Resource Disaggregation

Abstract:

Datacenters have been using the "monolithic" server model for decades, where each server has a motherboard that hosts a set of hardware devices like CPU and DRAM. This monolithic architecture is easy to deploy but cannot offer efficient resource packing, hardware elasticity, failure isolation, or good support for heterogeneity. I believe that recent and future hardware and software application trends call for a rethinking of the long-standing server-centric model. Our solution is to "disaggregate" monolithic servers into distributed, network-attached hardware components that can each manage its own resource and can fail independently.

In this talk, I will talk about my lab's efforts in building the next-generation disaggregated datacenters at system software, hardware, and network levels. Specifically, I will focus on two systems: LegoOS, a new distributed operating system designed and built from scratch for managing disaggregated resource devices (OSDI’18 Best Paper), and LegoFPGA, an efficient, flexible, and easy-to-use FPGA-based hardware solution designed for heterogeneous, programmable, and disaggregated datacenters. LegoFPGA is an ongoing work and the LegoFPGA team would welcome and appreciate feedback and suggestions!

Bio:

Yiying Zhang is an assistant professor in the School of Electrical and Computer Engineering at Purdue University. Her research interests span operating systems, distributed systems, computer architecture, and datacenter networking. She also works on the intersection of systems and programming language, security, and AI/ML. Yiying received her Ph.D. from the Department of Computer Sciences at the University of Wisconsin-Madison under the supervision of Andrea and Remzi Arpaci-Dusseau and worked as a postdoctoral scholar at the University of California, San Diego before joining Purdue.

Mar 27: Supercharge analytics over large IoT data: an OS approach

Abstract:

This talk overviews our inquiry in systems software for taming large IoT data. I will describe two systems that target two crucial, complementary analytics paradigms: a data engine for processing large telemetry streams (hot data); a data store for querying large archival videos (cold data). The two systems exploit emerging hardware (e.g. 3D-stacked DRAM) as well as emerging workloads (e.g. neural networks). Both systems advance the state-of-the-art performance by orders of magnitude.

Our experiences highlight the significance of designing OSes for specific scenarios, where the OSes play key roles: mapping AI to new hardware, dynamically configuring AI, and trading off among competing objectives.

Bio:

Felix Xiaozhu Lin is an assistant professor of Purdue ECE. He received PhD in CS from Rice and BS/MS from Tsinghua. At Purdue, he leads the Xroads systems exploration lab (XSEL) to accelerate and secure important computing scenarios. He and his students measure and build systems software with diverse techniques, including novel OS structures, kernel subsystem design, binary translation, and user-level runtimes. See http://xsel.rocks for more information.

He is a recipient of a Google Faculty Award, an NSF CRII award, and the best paper award from ASPLOS’14.

2018 Fall schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker (Prof)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/26</td>
<td>Yizhou Shan (Y. Zhang)</td>
<td>LegoOS: A Disseminated, Distributed OS for Hardware Resource Disaggregation</td>
</tr>
</tbody>
</table>
Talks

December 5, 2018: D2P: Automatically Creating Distributed Dynamic Programming Codes

Abstract:
Dynamic Programming (DP) algorithms are common targets for parallelization, and, as these algorithms are applied to larger inputs, distributed implementations become necessary. However, creating distributed-memory solutions involves the challenges of task creation, program and data partitioning, communication optimization, and task scheduling. In this paper we present D2P, an end-to-end system for automatically transforming a specification of any recursive DP algorithm into distributed-memory implementation of the algorithm. When given a pseudo-code of a recursive DP algorithm, D2P automatically generates the corresponding MPI-based implementation. Our evaluation of the generated distributed implementations shows that they are efficient and scalable. Moreover, D2P-generated implementations are faster than implementations generated by recent general distributed DP frameworks, and are competitive with (and often faster than) hand-written implementations.

Bio:
Nikhil Hegde is a graduate student in the department of electrical and computer engineering advised by Prof. Milind Kulkarni. His current research interests include distributed-memory parallelism of irregular applications.

LegoOS: A Disseminated, Distributed OS for Hardware Resource Disaggregation

Yizhou Shan

Abstract
The monolithic server model where a server is the unit of deployment, operation, and failure is meeting its limits in the face of several recent hardware and application trends. To improve heterogeneity, elasticity, resource utilization, and failure handling in datacenters, we believe that datacenters should break monolithic servers into disaggregated, network-attached hardware components. Despite the promising benefits of hardware resource disaggregation, no existing OSes or software systems can properly manage it.

We propose a new OS model called Splitkernel to manage disaggregated systems. Splitkernel decomposes traditional OS functionalities into loosely-coupled monitors, each of which runs on and manages a hardware component. Using the Splitkernel model, we built Lego, a new OS designed for hardware resource disaggregation. Lego appears to users as a set of distributed servers. Internally, Lego cleanly separates processor, memory, and storage devices both at the hardware level and the OS level. We implemented Lego from scratch and evaluated it by emulating hardware components using commodity servers. Lego improves performance per dollar by 33% to 66% over Linux on popular datacenter applications like TensorFlow, while providing better heterogeneity, elasticity, and failure isolation.

Bio
**2018 Spring schedule**

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/7</td>
<td>Yiyang Chang</td>
<td>Ensuring Network Designs Meet Service Level Objectives under Failures</td>
</tr>
<tr>
<td>3/21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/18</td>
<td>Liwei Guo</td>
<td>Power Sandbox: Power Awareness Redefined</td>
</tr>
<tr>
<td>4/25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Talks**

**Ensuring Network Designs Meet Service Level Objectives under Failures**

*Yiyang Chang*

**Abstract**

Network architects must ensure that networks meet Service Level Objectives (SLOs), i.e., meet a performance metric a desired percentage of time. The recent shift towards tightly provisioned networks has made this challenging given that failures are the norm. In this talk, I will present a principled approach to tackling the problem. It is based on Slice, our new optimization-based algorithm, which analytically classifies potential failure scenarios into those where the network performs acceptably or violates requirements. The approach combines Slice with network modeling and statistical sampling algorithms in novel ways. Our evaluations show that our approach is effective in certifying network designs and achieves orders of magnitude reduction in (and may even eliminate) explicit testing needs, relative to current practice. Further, our approach identifies failure scenarios with unacceptable performance and accommodates a variety of network and failure models.

**Bio**
Yiyang Chang is a Ph.D. student in the School of Electrical and Computer Engineering at Purdue University, advised by Professor Sanjay Rao. He obtained B.S. degree from School of EECS, Peking University. His research interest lies in the area of Computer Networks. He currently focuses on network certification with optimization approaches. Contact him at chang256@purdue.edu

Stop Worrying and Start Embracing Lossy Network: a Datacenter Network Protocol for Approximate Computing

Ke Liu

Abstract

Many datacenter applications such as machine learning and streaming systems do not need the complete set of data to perform their computation. These approximate applications often choose to select a subset of data and then transmit them over a reliable network layer like TCP. With such network-oblivious approaches, there is a huge missing opportunity in improving both the performance and accuracy of datacenter approximate applications. We propose to run approximate applications on a lossy network and to allow packet loss (or even intentionally create more) in a controlled manner. Specifically, we designed a new network protocol called Approximate Transmission Protocol, or ATP, for datacenter approximate applications. Our simulation results show that ATP can improve both the performance and accuracy of approximate applications than network-oblivious solutions.

Bio

Ke Liu is a visiting scholar from Institute of Computing Technology, Chinese Academy of Sciences, advised by Prof. Yiying Zhang from School of Electrical and Computer and Engineering and Prof. Vaneet Aggarwal from School of Industrial Engineering. He received Ph. D. in Department of Information Engineering from The Chinese University of Hong Kong. He is interested in the protocol optimization in computer networks including mobile data network and data center network.

Power Sandbox: Power Awareness Redefined

Liwei Guo

Abstract

Many apps benefit from knowing their power consumption and adapting their behaviors on the fly. To offer apps power knowledge at run time, an OS often meters system power and divides it among apps. Since the impacts of concurrent apps on system power are entangled, this approach not only makes it difficult to reason about power but also results in power side channels, a serious vulnerability. To this end, we introduce a new OS principal called power sandbox, which enables one app to observe the fine-grained power consumption of itself running in its vertical slice of the hardware/software stack. The observed power is insulated from the impacts of other apps. Our contribution is a set of lightweight kernel extensions that simultaneously i) enforce the power sandbox boundaries and ii) confine entailed performance loss to the sandboxed apps. Our experiences on two embedded platforms show that power sandboxes simplify reasoning about power, maintain fairness among apps, and minimize power side channels, thus facilitating construction of power-aware apps.

Bio

Liwei Guo is a second-year Ph.D student at Department of Electrical and Computer Engineering, Purdue University, advised by Prof. Felix Xiao Zhu Lin. He is interested in playing with different kernel subsystems and computer games.

--------------- Previous Talks -------------------

2017 Fall schedule
<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/6</td>
<td>Hongyu Miao</td>
<td>StreamBox: Modern Stream Processing on a Multicore Machine</td>
</tr>
<tr>
<td>9/13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/27</td>
<td>Kangjing Huang</td>
<td>DryadSynth: A Concolic SyGuS Solver</td>
</tr>
<tr>
<td>10/4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/18</td>
<td>Shin-Yeh Tsai</td>
<td>LITE Kernel RDMA Support for Datacenter Applications</td>
</tr>
<tr>
<td>11/1</td>
<td>Bo Fu</td>
<td>Parallel Video Processing Using Embedded Computers</td>
</tr>
<tr>
<td>11/8</td>
<td>Yizhou Shan</td>
<td>Disaggregated Operating System</td>
</tr>
<tr>
<td>11/15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/22</td>
<td>Ashraf Y. Mahgoub</td>
<td>Rafiki: A Middleware for Automated Parameters Tuning</td>
</tr>
<tr>
<td>11/29</td>
<td>Tsung Tai Yeh</td>
<td>Dynamic Warp Collapsing on the GPU</td>
</tr>
<tr>
<td>12/6</td>
<td>Yun Seong Nam</td>
<td>ABRTuner: Improving the Tail Performance of Video Delivery</td>
</tr>
</tbody>
</table>

**Talks**

*Speaker: please append your talk title, abstract, and bio to the following.*

*You may use previous talks as a template.*

**Disaggregated Operating System**

Yizhou Shan

**Abstract**

Recently, there is an emerging interest in datacenter resource disaggregation, an architecture that breaks monolithic servers into independent hardware components connected through network. OSes built for monolithic computers can not handle the distributed nature of disaggregated hardware components. Datacenter distributed systems are built for managing clusters of monolithic computers, not individual hardware components. When traditional OS operations spread across hardware components over the network, these distributed systems fall short. Clearly, we need a new operating system for the disaggregated architecture.

We propose the concept of decomposed operating system for the disaggregated datacenter architecture. The basic idea is simple: When hardware is disaggregated, the operating system should be also.

**Bio**

Yizhou Shan is a second-year Ph.D student at Department of Electrical and Computer Engineering, Purdue University, advised by Prof. Yiying Zhang. His research interests are in distributed systems and NVM.

---

**Parallel Video Processing Using Embedded Computers**
**Bo Fu**

**Abstract**

As frame rates and resolutions of video streams increase, a need for parallel video processing emerges. Most studies offload computation to the cloud, but this is not always possible. For example, solar-powered cameras can be deployed in locations away from power grids. A better choice is to process the data locally on embedded computers without raw video transmission through networks. Parallel computing alleviates the performance bottleneck of a single embedded computer but it degrades analysis accuracy because partitioning video streams breaks the continuity of motion. This paper presents a solution for maintaining accuracy in parallel video processing. A video stream is divided into multiple segments processed on different embedded computers. The segments overlap so that continuous motion can be detected. The system balances workload based on the speed of GPU and CPU to reduce execution time. Experimental results show up to 82.6% improvement in accuracy and 58% reduction in execution time.

**Bio**

Bo Fu is a Ph.D student at Department of Computer Science, Purdue University, advised by Prof. Yung-Hsiang Lu. His research interests are in Systems and Video Processing.

---

**Shin-Yeh Tsai**

**Abstract**

Recently, there is an increasing interest in building datacenter applications with RDMA because of its low-latency, high-throughput, and low-CPU-utilization benefits. However, RDMA is not readily suitable for datacenter applications. It lacks a flexible, high-level abstraction; its performance does not scale; and it does not provide resource sharing or flexible protection. Because of these issues, it is difficult to build RDMA-based applications and to exploit RDMA's performance benefits.

To solve these issues, we built LITE, a Local Indirection Tier for RDMA in the Linux kernel that virtualizes native RDMA into a flexible, high-level, easy-to-use abstraction and allows applications to safely share resources. Despite the widely-held belief that kernel bypassing is essential to RDMA's low-latency performance, we show that using a kernel-level indirection can achieve both flexibility and low-latency, scalable performance at the same time.

**Bio**

Shin-Yeh Tsai is a Ph.D student at Department of Computer Science, Purdue University, advised by Prof. Yiying Zhang. His research interests are in Operating Systems and Networking.

---

**Kangjing Huang**

**Abstract**

The classical formulation of the program-synthesis problem is to find a program that meets a correctness specification given as a logical formula. Recent work on program synthesis and program optimization illustrates many potential benefits of allowing the user to supplement the logical specification with a syntactic template that constrains the space of allowed implementation, which is called Syntax-Guided Synthesis, SyGuS for short. On that basis, we present DRYADSYNTH, a concolic SyGuS solver. The synthesis algorithm combines enumerative search and symbolic search, uses a decision-tree data strcutre in a CEGIS (CounterExample Guided Inductive Synthesis) framework and supports synthesis for CLIA (Conditional Linear Integer Arithmetic) and INV (Invariant Synthesis) problems defined in the SyGuS context. Our solver took part in the SyGuS-COMP2017 competition, in which it achieved fair performance on CLIA problems and achieved outstanding performance on INV problems.

**Bio**

Kangjing Huang is a Ph.D student at Department of Electrical and Computer Engineering, Purdue University, advised by Prof. Xiaokang Qiu. His research interests are in Program Languages, Program Synthesis and Verification.
StreamBox: Modern Stream Processing on a Multicore Machine
Hongyu Miao

Abstract
Stream analytics on real-time events has an insatiable demand for throughput and latency. Its performance on a single machine is central to meeting this demand, even in a distributed system. This paper presents a novel stream processing engine called StreamBox that exploits the parallelism and memory hierarchy of modern multicore hardware. StreamBox executes a pipeline of transforms over records that may arrive out-of-order. As records arrive, it groups the records into ordered epochs delineated by watermarks. A watermark guarantees no subsequent record’s event timestamp will precede it.

Our contribution is to produce and manage abundant parallelism by generalizing out-of-order record processing within each epoch to out-of-order epoch processing and by dynamically prioritizing epochs to optimize latency. We introduce a data structure called cascading containers, which dynamically manages concurrency and dependences among epochs in the transform pipeline. StreamBox processes records up to 38 GB/sec (38M Records/sec) with 50 ms latency.

Bio
Hongyu Miao is a PhD student at Department of Electrical and Computer Engineering, Purdue University, advised by Professor Felix Xiaozhu Lin. His research interests are in operating systems and data analytics systems.

Sense-Aid: A framework for enabling network as a service for participatory sensing
Heng Zhang

Abstract
A rapid adoption of smartphones with different types of advanced sensors has led to an increasing trend in the usage of mobile crowdsensing applications, e.g., to create hyperlocal weather maps. However, the high energy consumption of crowdsensing, chiefly due to expensive network communication, has been found to be detrimental to wide-spread adoption. We propose a framework, called SENSE-AID, that can provide energy-efficient mobile crowd-sensing service, coexisting with the cellular network. There are two key innovations in SENSE-AID beyond prior work in mobile crowdsensing (Piggyback Crowdsensing-Sensys13)—the middleware running on the cellular network edge orchestrates multiple devices present in geographical proximity to suppress redundant data collection and communication and it understands the state of each device (radio state, battery state, etc.) to decide which ones should be selected for crowdsensing activities at any point in time. It also provides a simple programming abstraction to help with the development of crowdsensing applications. We show the benefit of SENSE-AID by conducting a user study consisting of 60 students in our campus, compared to a baseline periodic data collection method and Piggyback Crowdsensing. We find that energy saving is 79.9% for a representative case which requires 5 devices to provide barometric values within an area of a circle whose radius is 1 kilometer and requires periodic data collection of 5 minutes. The selection algorithm of SENSE-AID also ensures reasonable fairness in the use of the different devices.

Bio
Heng Zhang is a Ph.D student at Department of Electrical and Computer Engineering, Purdue University, advised by Prof. Saurabh Bagchi. His research interests are in Mobile Sensing and Mobile Computing.

Rafiki: A Middleware for Automated Parameters Tuning
Ashraf Mahgoub

High performance computing (HPC) applications, such as metagenomics and other big data systems, need to store and analyze huge volumes of semi-structured data. Such applications often rely on NoSQL-based datastores, and optimizing these databases is a challenging endeavor, with over 50 configuration parameters in Cassandra alone. As the application executes, database workloads can
change rapidly from read-heavy to write-heavy ones, and a system tuned with a read-optimized configuration becomes suboptimal when the workload becomes write-heavy. In this paper, we present a method and a system for optimizing NoSQL configurations for Cassandra and ScyllaDB when running HPC and metagenomics workloads.

Bio

I am a 2nd year PhD student in the Department of Computer Science, Purdue University. I am advised by Prof. Ananth Grama and Prof. Saurabh Bagchi. My research interests are in reliable and high performance distributed systems.

ABRTuner: Improving the Tail Performance of Video Delivery

Yun Seong Nam

Content providers are interested in providing good video delivery QoE for all users, not just on average. Deployed ABR algorithms do not, in general, span the entire range of network conditions seen in practice (i.e., they have limited dynamic range), so they may perform poorly for some users and/or videos. To improve the tail performance for video delivery, we observe that an ABR algorithm’s limited dynamic range arises from the fact that it often has parameters sensitive to network conditions such as the mean and variance in throughput. Yet, when an ABR algorithm is deployed, it is configured so that it performs well on average. This observation led us to design ABRTuner, which pre-computes, for a given ABR algorithm, the best possible parameter configurations for different network conditions, then dynamically adapts the parameter at run-time for the current network conditions. Using an implementation of ABRTuner in the Dash.js framework, we show that it significantly improves the tail performance of a widely deployed hybrid ABR algorithm and is also better than other recent research proposals in the tail. Unfortunately, even if one were to devise a new ABR approach, as we have done, today’s ecosystem does not permit rapid deployment of these approaches, since ABR implementations are embedded into application frameworks. To address this challenge, we demonstrate that ABRTuner’s control logic can be deployed on the cloud, which can permit fast ABR evolution.

Bio

Yun Seong Nam is a Ph.D. in the School of Electrical and Computer Engineering at Purdue University, advised by Professor Sanjay Rao. His research interest lies in the area of Computer Networks. His current work primarily focuses on building systems that improve the user experience of video delivery.

2017 Spring schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/15</td>
<td>Shuang Zhai</td>
<td>Decelerating Suspend and Resume in Operating Systems</td>
</tr>
<tr>
<td>2/22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/8</td>
<td>Yiyang Chang</td>
<td>Robust Validation of Network Designs Under Uncertain Demands and Failures</td>
</tr>
<tr>
<td>3/15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/29</td>
<td>Kirshanthan Sundararajah</td>
<td>Locality Enhancing Transformations for Nested Recursive Iteration Spaces</td>
</tr>
</tbody>
</table>
Talks

Speaker: please append your talk title, abstract, and bio to the following.

You may use previous talks as a template.

Decelerating Suspend and Resume in Operating Systems

Shuang Zhai

Abstract

Short-lived tasks have a large impact on mobile computer’s battery life. In executing such tasks, the whole system transitions in and out of the deep sleep mode. This suspend/resume procedure is controlled by the operating system (OS), which consumes a dominating portion of energy. Through characterizing the Linux kernel on a variety of modern system-on-chips (SoCs), we show that the OS suspend/resume mechanism is fundamentally slowed down by various IO devices, which frequently keep CPU waiting.

To minimize energy consumption, we advocate offloading the OS suspend/resume to a miniature processor that waits more efficiently. To this end, we propose a new virtual executor that runs on a miniature core and directly executes the unmodified kernel binary of the main CPU. We construct the virtual executor centering on software-only, cross-ISA binary translation, an approach previously considered prohibitively expensive. Through novel designs and optimizations, we reduce the translation overhead by 5x. The preliminary benchmarks show promising energy efficiency.

A report of this ongoing work will appear at HotMobile ‘17.

Bio

Shuang Zhai is a final-year undergraduate student of Purdue ECE, advised by Professor Felix Xiaozhu Lin. He is interested in operating systems. Contact him at zhais@purdue.edu.

Robust Validation of Network Designs under Uncertain Demands and Failures

Yiyang Chang

Abstract

A key challenge confronting network designers is verifying that their networks can properly cope with a wide range of traffic conditions and failure scenarios, and designing networks to meet this objective. Since many design choices (e.g., topology design, middlebox placement) are made over longer time-scales and cannot be easily adapted, it is important to make these choices in a manner robust to traffic demands and failure scenarios. In this paper, we develop an optimization-theoretic framework to provide guaranteed bounds on link utilization across traffic patterns and failure scenarios of interest, as well as help design networks with guaranteed bounds. The key novelty of our framework is that while some design choices (e.g., middlebox placement) are made in a manner robust to traffic demand, other decisions, especially the actual routing strategy may be optimized for a specific traffic demand. Considering flexible routing strategies is important since oblivious strategies can be unduly conservative, but poses theoretical challenges that we address. We apply our framework to multiple case studies including design of MPLS tunnels, and routing in the presence of middleboxes. Evaluations over real network topologies and traffic data show the promise of the approach.

Bio

Yiyang Chang is a Ph.D. student in the School of Electrical and Computer Engineering at Purdue University, advised by Professor Sanjay Rao. He obtained B.S. degree from School of EECS, Peking University. His research interest lies in the area of Computer Networks. He currently focuses on network synthesis and validation with optimization approaches. Contact him at chang256@purdue.edu.
Locality Transformations for Nested Recursive Iteration Spaces

Kirshanthan Sundararajah

Abstract:

There has been a significant amount of effort invested in designing scheduling transformations such as loop tiling and loop fusion that rearrange the execution of dynamic instances of loop nests to place operations that access the same data close together temporally. In recent years, there has been interest in designing similar transformations that operate on recursive programs, but until now these transformations have only considered simple scenarios: multiple recursions to be fused, or a recursion nested inside a simple loop. This paper develops the first set of scheduling transformations for nested recursions: recursive methods that call other recursive methods. These are the recursive analog to nested loops. We present a transformation called recursion twisting that automatically improves locality at all levels of the memory hierarchy, and show that this transformation can yield substantial performance improvements across several benchmarks that exhibit nested recursion.

Bio:

I have earned my BSc. Eng. (Hons) in 2014 from Department of Electronic and Telecommunication Engineering, University of Moratuwa, Sri Lanka. Currently I am PhD Student, conducting research under the supervision of Professor Milind Kulkarni as part of PLCL (Parallelism, Languages and Compilers Lab). My areas of interest in research are programming languages, compilers, systems and high performance computing. I am particularly interested in enhancing the performance of irregular programs and building a general framework to analyze the optimizations for such programs.

2016 Fall schedule (archived; don't edit)

Speaker: please edit the corresponding row to reflect your talk

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/7</td>
<td>Junghoon Chae</td>
<td>Visual Analytics of Location-Based Social Networks for Decision Support</td>
</tr>
<tr>
<td>9/14</td>
<td>Tara Thomas</td>
<td>Sirius: Neural network based probabilistic assertions for detecting silent data corruption in parallel programs</td>
</tr>
<tr>
<td>9/21</td>
<td>Laith Sakka</td>
<td>Fusing General Recursive Tree Traversals</td>
</tr>
<tr>
<td>9/28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/5</td>
<td>Justus C Ndukaife</td>
<td></td>
</tr>
<tr>
<td>10/12</td>
<td>Anup Mohan</td>
<td>Cloud Resource Management for Big Visual Data Analysis</td>
</tr>
<tr>
<td>10/19</td>
<td>David Ebert</td>
<td>Solving Real World Challenges with Visual Analytics</td>
</tr>
<tr>
<td>10/26</td>
<td>Yiyang Chang</td>
<td>Robust validation of network designs under uncertain demands and failures</td>
</tr>
<tr>
<td>11/2</td>
<td>Shin-Yeh Tsai</td>
<td>FIT: A Flexible Infiniband Tier for Datacenter Software</td>
</tr>
<tr>
<td>11/9</td>
<td>Yizhou Shan</td>
<td>Distributed Shared Persistent Memory</td>
</tr>
<tr>
<td>11/16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/23</td>
<td>Youngsul Koh</td>
<td></td>
</tr>
<tr>
<td>11/30</td>
<td>Ashiwan Sivakumar</td>
<td>Proxy-Assisted Browsing for Low-Latency Web over Cellular Networks</td>
</tr>
</tbody>
</table>
Talks

Speaker: please append your talk title, abstract, and bio to the following.

You may use the following as a template.

A holistic approach to lowering latency in geo-distributed multi-tier applications

Shankar Narayanan

Abstract
User perceived end-to-end latency of applications have a huge impact on the revenue for many businesses. Service level agreements (SLAs) on such applications often require bounds on the 90th (and higher) percentile latencies, which must be met while scaling to hundreds of thousands of geographically dispersed users. Improving user perceived performance of such large-scale, multi-tiered applications is challenging. These applications often have complex operating environments, with many user-facing components like web servers and content distribution network (CDN) servers, and many other backend components like application and storage servers. Further, the end user performance often depends on the complex interactions between these components.

The primary focus of my research has been to develop techniques and build systems that help reduce end-to-end application latency. More specifically, it makes the following three contributions. First, it reduces the user-facing latency by priority-aware organization of content within the CDNs. Next, it reduces the backend latency through optimal placement of data at the data store layer taking into account application's data access patterns and judiciously balancing the forces of latency, consistency, and availability. Finally, it improves the performance of the application layer by carefully rerouting requests across different application component replicas. While these solutions can be implemented independently or concurrently at the different layers, each of them explicitly focuses on improving end-to-end application performance.

Bio
Shankar is a Ph.D. candidate at the Department of ECE at Purdue University, working with his advisor Prof. Sanjay Rao. He obtained his Bachelor's Degree in Computer Science with Anna University, one of the premier engineering institutes in India. Broadly, his research interests are in the areas of Computer Networks, Distributed Systems, and Cloud Computing. His work primarily focuses on building systems that improve the performance of geo-distributed, multi-tiered applications.

Visual Analytics of Location-Based Social Networks for Decision Support

Junghoon Chae

Abstract
Recent advances in technology have enabled people to add location information to social networks called Location-Based Social Networks (LBSNs) where people share their communication and whereabouts not only in their daily lives, but also during abnormal situations, such as crisis events. However, since the volume of the data exceeds the boundaries of human analytical capabilities, it is almost impossible to perform a straightforward qualitative analysis of the data. The emerging field of visual analytics has been introduced to tackle such challenges by integrating the approaches from statistical data analysis and human computer interaction into highly interactive visual environments.

Based on the idea of visual analytics, this research contributes the techniques of knowledge discovery in social media data for providing comprehensive situational awareness. We extract valuable hidden information from the huge volume of unstructured social media data and model the extracted information for visualizing meaningful information along with user-centered interactive interfaces. We develop visual analytics techniques and systems for spatial decision support through coupling modeling of spatiotemporal social media data, with scalable and interactive visual environments. These systems allow analysts to detect and examine abnormal events within social media data by integrating automated analytical techniques and visual methods. We provide comprehensive analysis of public behavior response in disaster events through exploring and examining the spatial and temporal distribution of LBSNs. We also propose a trajectory-based visual analytics of LBSNs for anomalous human movement analysis during crises by incorporating a novel classification technique. Finally, we introduce a visual analytics approach for forecasting the overall flow of human crowds.

Bio
Junghoon Chae is a Ph.D. candidate in the School of Electrical and Computer Engineering at Purdue University, working with Prof. David Ebert. His research expertise and interests are, but not limited to, in the areas of visual analytics for large-scale data, spatiotemporal data modeling and visualization, and social media and text data mining. He received his Master of Science degree in Electrical and Computer Engineering from Purdue University in 2011 and Bachelor of Science degree in Computer Engineering and Electrical Engineering (Dual Major) from Kyung Hee University, South Korea in 2008. Contact him at jchae@purdue.edu.
Sirius: Neural network based probabilistic assertions for detecting silent data corruption in parallel programs

Tara Elizabeth Thomas

Abstract

The size and complexity of supercomputing clusters are rapidly increasing to cater to the needs of complex scientific applications. At the same time, the feature size and operating voltage level of the internal components are decreasing. This dual trend makes these machines extremely vulnerable to soft errors or random bit flips. For complex parallel applications, these soft errors can lead to silent data corruption which could lead to large inaccuracies in the final computational results. Hence, it is important to determine the presence and severity of such errors early on, so that proper counter measures can be taken. In this paper, we introduce a tool called Sirius, which can accurately identify silent data corruptions based on the simple insight that there exist spatial and temporal locality within most variables in such programs. Spatial locality means that values of the variable at nodes that are close by in a network sense, are also close numerically. Similarly, temporal locality means that the values change slowly and in a continuous manner with time. Sirius uses neural networks to learn such locality patterns, separately for each critical variable, and produces probabilistic assertions which can be embedded in the code of the parallel program to detect silent data corruptions. We have implemented this technique on parallel benchmark programs - LULESH and CoMD. Our evaluations show that Sirius can detect silent errors in the code with much higher accuracy compared to previously proposed methods. Sirius detected 98% of the silent data corruptions with a false positive rate of less than 0.02 as compared to the false positive rate 0.06 incurred by the state of the art acceleration based prediction (ABP) based technique.

Bio

Tara Thomas is a graduate student doing her MS in the Department of Electrical and Computer Engineering, working with Prof. Saurabh Bagchi. She obtained her undergraduate degree from National Institute of Technology, Calicut and worked as an IC Design Engineer at Broadcom Corporation for two years before joining Purdue. Her research broadly focuses on performance improvements in distributed systems, improving reliability of systems and using causality to root cause issues in these systems. Contact her at thom a579@purdue.edu.

Fusing General Recursive Tree Traversals

Laith Sakka

Abstract:

Series of traversals of tree structures arise in numerous contexts: abstract syntax tree traversals in compiler passes, rendering traversals of the DOM in web browsers, kd-tree traversals in computational simulation codes. In each of these settings, a tree is traversed multiple times to compute various values and modify various portions of the tree. While it is relatively easy to write these traversals as separate small updates to the tree, for efficiency reasons, traversals are often manually fused to reduce the number of times that each portion of the tree is traversed: by performing multiple operations on the tree simultaneously, each node of the tree can be visited fewer times, increasing opportunities for optimization and decreasing cache pressure. This fusion process is often done manually, requiring careful understanding of how each of traversals of the tree interact. This paper presents an automatic approach to traversal fusion: tree traversals can be written independently, and then our framework analyzes the dependences between the traversals to determine how they can be fused to reduce the number of visits to each node in the tree. A critical aspect of our framework is that it exploits two opportunities to increase the amount of fusion: i) it automatically integrates code motion, and ii) it supports partial fusion, where portions of one traversal can be fused with another, allowing for a reduction in node visits without requiring that two traversals be fully fused. We implement our framework in Clang, and show across a series of compiler passes written in C++ that our framework can automatically enable a significant reduction in the number of visits to the nodes of an AST, and in traversals with data reuse can substantially improve performance.

Bio:

Laith Sakka is a second year PhD student, Working with professor Milind Kulkarni in PLCL group, his research interests spans through different layers of computer systems, right not his focus is on optimizing recursive traversals .He got his BSc degree in computer engineering from Princess Sumaya University in Jordan. you contact him at lsakka@purdue.edu.

Cloud Resource Management for Big Visual Data Analysis

Anup Mohan

Abstract:

There has been tremendous growth of visual data available on the Internet in recent years. These data may be used for a wide variety of scientific studies about weather, wildlife, traffic, etc. Analyzing and managing such big visual data requires significant computational resources that can be expensive. Transmitting this data over networks also consumes significant energy and time. One type of visual data of particular interest is produced by network cameras providing real-time views. Millions of network cameras
around the world continuously stream data to viewers connected to the Internet. Cloud computing is an ideal choice to meet the resource requirements as it offers resources known as instances with different capabilities and price them according to the usage. There are many options when selecting cloud instances (amounts of memory, number of cores, geographic locations, etc.). The cost of renting a cloud instance varies based on the options. The selection of the cloud instance impacts the accuracy of the analysis being performed. Inefficient provisioning of cloud resources may become costly in pay-per-use cloud computing.

This presentation examines the different factors involved in cloud resource management for analyzing big visual data from globally distributed network cameras. A resource manager is presented that significantly reduces the analysis cost by selecting the types, locations, and number of cloud instances while satisfying the performance and accuracy requirements, and adaptively allocating the cloud resources based on run-time conditions. The trade-offs between quality, accuracy, performance, and cost for big visual data analysis are also discussed.

Bio:

Anup Mohan is a Ph.D. candidate in the Department of Electrical and Computer Engineering at Purdue University. His research interests include large-scale video analysis, cloud computing, and big data analysis. Anup Mohan obtained an M.S. degree from Purdue University in 2013. You can contact him at mohan11@purdue.edu

Robust Validation of Network Designs under Uncertain Demands and Failures

Yiyang Chang

Abstract

A key challenge confronting network designers is verifying that their networks can properly cope with a wide range of traffic conditions and failure scenarios, and designing networks to meet this objective. Since many design choices (e.g., topology design, middlebox placement) are made over longer time-scales and cannot be easily adapted, it is important to make these choices in a manner robust to traffic demands and failure scenarios. In this paper, we develop an optimization-theoretic framework to provide guaranteed bounds on link utilization across traffic patterns and failure scenarios of interest, as well as help design networks with guaranteed bounds. The key novelty of our framework is that while some design choices (e.g., middlebox placement) are made in a manner robust to traffic demand, other decisions, especially the actual routing strategy may be optimized for a specific traffic demand. Considering flexible routing strategies is important since oblivious strategies can be unduly conservative, but poses theoretical challenges that we address. We apply our framework to multiple case studies including design of MPLS tunnels, and routing in the presence of middleboxes. Evaluations over real network topologies and traffic data show the promise of the approach.

Bio

Yiyang Chang is a Ph.D. student in the School of Electrical and Computer Engineering at Purdue University, advised by Professor Sanjay Rao. He obtained B.S. degree from School of EECS, Peking University. His research interest lies in the area of Computer Networks. He currently focuses on network synthesis and verification with optimization approaches. Contact him at chang256@purdue.edu

FIT: A Flexible Infiniband Tier for Datacenter Software

Shin-Yeh Tsai

Abstract

Recently, there is an increasing interest in building distributed systems with Infiniband (IB) and RDMA in datacenter environments, because of its low-latency performance, RDMA capability, and lossless communication. Despite all these advantages, it is notoriously difficult to use IB and to exploit all its performance benefits. The fundamental cause of this difficulty is the mismatch between IB’s abstraction and the abstraction its datacenter users expect: IB provides low-level, hardware-like primitives while application and system builders intend to use IB to build high-level distributed systems and treat RDMA as a means to access and manage data in memory. We believe that this mismatch can be solved by adding a level of indirection to virtualize Infiniband and to provide its users with a familiar, easy-to-use, flexible abstraction. An immediate question that follows is whether or not adding this level of indirection will strip away the low-latency performance of IB. To answer this question, we built FIT, a Flexible Infiniband Tier, in the Linux kernel. FIT provides a virtualized, flexible, easy-to-use abstraction to both kernel and user-level applications. With this rich abstraction, FIT can easily support memory-based, data I/O based, and network-based applications. Despite the widely-held belief that kernel bypassing is essential to IB’s low-latency performance and that adding another level of indirection will inevitably cause performance overhead, we show that by exploiting, rethinking, and integrating various kernel and IB’s functionalities, we can achieve both flexibility and low-latency performance at the same time. To further demonstrate the benefits of FIT, we built a userlevel, one-sided distributed logging system and a kernellevel distributed shared memory system on top of FIT.

Bio

Shin-Yeh Tsai is a Ph.D. student from the department of Computer Science, advised by Professor Yiying Zhang. His work lies between system and network area. He currently focuses on Infiniband network. Contact him at tsai46@purdue.edu

Distributed Shared Persistent Memory
Yizhou Shan

Abstract

We introduce Distributed Shared Persistent Memory (DSPM), a new framework for using persistent memories in distributed datacenter environments. DSPM provides a new abstraction that allows applications to both perform traditional memory load and store instructions and to name, share, and persist their data. We built Hotpot, a kernel-level DSPM system that provides low-latency, transparent memory accesses, data persistence, data reliability, and high availability. The key ideas of Hotpot are to integrate distributed memory caching and data replication techniques and to exploit application hints. We implemented Hotpot in the Linux kernel and demonstrated its benefits by building a distributed graph engine on Hotpot and porting a NoSql database to Hotpot.

Bio

Yizhou Shan is a first year Ph.D. student in the Department of Electrical and Computer Engineering at Purdue University, advised by Professor Yiying Zhang. His research interests span computer architecture, operating system, and distributed system software. He currently focuses on building next-generation operating system for rack-scale datacenter. Contact him at: shan13@purdue.edu

Proxy-Assisted Browsing for Low-Latency Web over Cellular Networks

Ashiwan Sivakumar

Abstract: With the rapid growth of cellular users, ensuring good user experience is imperative for service providers. However, this is made challenging for network-intensive mobile web applications, due to factors such as high cellular last-mile latencies and resource constraints of both cellular devices and the Radio Access Network (RAN). Further, today’s web download process is ill-suited for cellular networks resulting in high page load times. Despite much recent progress at tackling the challenge with cloud-assistance and new protocols like SPDY, achieving a responsive browsing experience still remains an elusive goal.

In this talk, we first discuss where existing solutions fall short in achieving the above goal. Then we describe our system, PARCEL that efficiently refactors browsing functionality between the mobile device and the cloud based on their respective strengths, and in a manner distinct from traditional browsers and existing cloud-heavy approaches. The proxy executes Web page code (e.g. JavaScript) for a user and proactively pushes objects to the client. Through experiments on live LTE settings we show that PARCEL can achieve 2X latency reduction over traditional browsers. Finally, we will conclude with a brief description of our ongoing work where we tackle the problem of scaling the proxy execution when deployed inside the cellular network edge.

Bio: Ashiwan Sivakumar is a Ph.D. candidate in the School of Electrical and Computer Engineering at Purdue University, advised by Professor Sanjay Rao. He obtained his Bachelor's Degree in Electrical Engineering from Anna University (MIT campus), India. His research interest lies in the area of Computer Networks. His current work primarily focuses on building systems that improve the user experience of mobile Web over cellular networks. Contact him at asivakum@purdue.edu

Previous schedule

2016 Spring

Ensuring Network Designs Meet Service Level Objectives under Failures

Yiyang Chang

Abstract

Network architects must ensure that networks meet Service Level Objectives (SLOs), i.e., meet a performance metric a desired percentage of time. The recent shift towards tightly provisioned networks has made this challenging given that failures are the norm. In this talk, I will present a principled approach to tackling the problem. It is based on Slice, our new optimization-based algorithm, which analytically classifies potential failure scenarios into those where the network performs acceptably or violates requirements. The approach combines Slice with network modeling and statistical sampling algorithms in novel ways. Our evaluations show that our approach is effective in certifying network designs and achieves orders of magnitude reduction in (and may even eliminate) explicit testing needs, relative to current practice. Further, our approach identifies failure scenarios with unacceptable performance and accommodates a variety of network and failure models.

Bio