
RESEARCH INTEREST	My research interests include parallel programming, parallel computing, distributed system and networking.	
--------------------------	--	--

EDUCATION	PhD student , Computer Science Cornell University , Ithaca, NY GPA: 4.0	<i>Aug 2007</i> - <i>current</i>
	Master of Engineering , Computer Science, Distinguished Graduates Award Tsinghua University , Beijing, China	<i>Jul 2007</i>
	Bachelor of Science , Mathematics and Physics, Academic Talent Program Tsinghua University , Beijing, China	<i>Jul 2005</i>

SKILLS	<i>Projects in:</i> C/C++, Python, CUDA, Intel IXP network processor C/ASM, Maple <i>Familiar with:</i> HTML, XML, Perl, Java, Matlab <i>Other Tools:</i> TUN, Iptables, SIPp, NS2, OPNET, Gnuplot, LaTeX	
---------------	---	--

WORK EXPERIENCE	System software intern , NVIDIA, Santa Clara, CA CUDA-Enhanced Compression for Remote Graphics <i>Best Intern Presentation in System Software (Resman) Department</i>	<i>May 2008 –</i> <i>Aug 2008</i>
	<p>Motive: As GPU becoming the new computing powerhouse, GPU virtualization can allocate dedicated GPU resources to multiple clients for concurrent, high quality display. For remote clients, dozens of frames will be delivered over the network in a second, easily generating astronomical traffic volume (several Gbps). Compression is thus a must for remote graphics. However, CPU-based compression mechanisms in today's remote desktop fall short of the requirements of many applications. Further, algorithms chosen should adapt to varying tradeoffs between image sharpness and display smoothness.</p> <p>Goal: Design and implement a GPU-based software compression module that converts screenshots into a compact stream ready for network transmission, and meeting the following additional requirements: 1. Real-time processing; 2. Multi-client support; 3. Minimized CPU utilization and PCI-E traffic; 4. Adaptive compression algorithm and ratio.</p> <p>Contribution: Adapted a hardware-oriented, serialized algorithm to a massively parallel CUDA implementation on existing GPU. It meets requirements 1-3 and supports lossless compression. Algorithms are specially tuned to exploit the parallelism and memory hierarchy of NVIDIA GPUs. Novelty of the software solution includes redesigned program and data flow and hardware-specific optimizations. As of performance, the compressing module spent <4ms to process a frame of HD resolution (1920x1200) using GT200 (compared to >130ms on an Intel E2180 CPU). For the first time, a single graphic card can handle HD frame compression for several to a dozen users at 30+fps.</p>	
	Research Intern , IBM China Research Lab, Beijing Evaluation and Optimization of SIP-based Applications on IBM Java Virtual Machine	<i>Aug 2006 -</i> <i>Jan 2007</i>
	<p>Motive: Despite their security advantage and development convenience, performance of servers running SIP-based applications on top of Java virtual machine (JVM) are often questioned, mainly due to the impact of GC. The basic GC policy in IBM JVM can be summarized as a stop-the-world mark-sweep-and-compact strategy. Consequently, GC introduces periodical slow response intervals and message loss.</p> <p>Goal: 1. To streamline servers specifically for SIP-based applications by tuning a number of JVM policies, such as generational GC, concurrent marking and JVM parameters. 2. Calculate performance degradation curves of an optimized server configuration assuming certain user behavior, and point out the inflection</p>	

point of performance.

Contribution: We have deployed a series of experiments on our test bed with various JVM settings. So far, we are able to decide optimal policy configurations under different traffic patterns. A deterministic Petri Net model has also been developed to plot performance degradation curves and to predict the inflection point.

Graduate research assistant, Tsinghua University, Beijing *Mar 2006 - Jul 2007*
Enhance Unstructured Peer-to-Peer System in Privacy and Security, a joint program with Microsoft Research Asia

Motive: While P2P systems have an edge in scalability, they are often questioned about their fairness, privacy and security. On the other hand, efficiency, fairness and survivability are interwoven in many P2P file sharing systems such as BitTorrent. It is important to realize the interdependency of performance indices, and introduce enhancement into one or more aspects of such applications afterwards.

Goal: 1. To point out the how different aspects, such as fairness, survivability, efficiency and etc., are related to each other. 2. To put forward an improved design to enhance efficiency and privacy in BitTorrent like file sharing systems. 3. Provide measurement tools to collect end-to-end performance metrics at low communication cost for a global view of P2P systems.

Contribution: Our preliminary research on takes a modeling approach on the dynamic behavior of BitTorrent like networks. The model reveals how fairness and survivability of a P2P file sharing system are related to the more traditional issues like scalability and efficiency. After realizing our first goal, an enhanced interconnecting mechanism has been designed and implemented to improve efficiency, security and privacy, and this mechanism could be applied to general P2P file-sharing systems.

Graduate research assistant, Tsinghua University, Beijing *Jul 2005 - Aug 2006*
Two projects on Network Processors of Intel IXA University Program, released a software package, **NP**CryptBench

Motive: The functionality and performance of many network security applications rely on an efficient, flexible implementation of cryptographic algorithms on network processors (NPs), especially on data plane (fast path). Necessity of benchmarking cryptographic algorithms on NPs rises in the face of rigid hardware solutions and unclear performance of software-based solutions.

Goal: 1. Implementing popular, representative cryptographic algorithms on network processors, with our focus on data-plane. 2. Benchmarking NPs to locate performance bottlenecks. 3. To alleviate the bottlenecks with generic or model-specific optimizations.

Contribution: A major developer of NPCryptBench, a cryptographic benchmark suite for network processors. By applying the benchmarks on several NP models we obtain first-hand statistics. With these data performance bottlenecks under different workloads and system configurations are successfully located. Summarized benchmark methodology and optimization guidelines from both software and architecture angles.

OTHER PROJECTS

A Highly Parallel Implementation of Yarn-level Knitted Clothing Simulator *Sep 2008 – current*
Advisor: Prof. Doug James

Scaling BGP/MPLS VPN *Sep 2007 – Jan 2008*
Advisor: Prof. Paul Francis

Motive: In recent years, network-based VPN services have gained wide popularity. Consequently, routing and forwarding tables in provider equipment undergo a rapid growth. Observations show that the traditional any-to-any route provision is wasteful under existing traffic patterns. Scalability of network-based VPN services improves if routing and forwarding table size can be reduced significantly.

Goal: 1. Analyze topological and traffic characterizations of existing network-based VPN infrastructure. 2. Design a mechanism to reduce routing/forwarding table size without violating service-level agreements. 3. Implement the mechanism and integrate it into network infrastructure.

Contribution: Comprehensive analyses of collected data have been performed. We study basic aspects such as VPN deployment characteristics, infrastructure topology, traffic patterns and routing prefix utilization. Based on this insight, we find it plausible to introduce a tradeoff between worst-case path length and routing table size. An algorithm is designed to automatically configure route provisioning under performance constraints. We also suggest several simpler strategies to reduce routing table size.

The Analysis and Improvement of Peer-to-Peer File-sharing Systems

May 2006 -
Jul 2007

Best Master Thesis Award, Tsinghua University

Advisor: Prof. Chuang Lin

PUBLICATION

1. **Yao Yue**, Chuang Lin, Zhangxi Tan, “Analyzing the Performance and Fairness of BitTorrent-like Networks Using a General Fluid Model”, **Elsevier Computer Communications**, Volume 29, Issue 18, Nov 2006, pp. 3946-3956
2. **Yao Yue**, Chuang Lin, Zhangxi Tan, “NPCryptBench: A Cryptographic Benchmark for Network Processors”, **ACM SIGARCH Computer Architecture News**, volume 34, issue 1, 2006, pp. 49-56
3. **Yao Yue**, Chuang Lin, Zhangxi Tan, “Analyzing the Performance and Fairness of BitTorrent-like Networks Using a General Fluid Model”, the 49th annual IEEE Global Telecommunications Conference (**GLOBECOM 2006**), San Francisco, California, Nov 2006.
4. Hao Wen, Chuang Lin, Fengyuan Ren, **Yao Yue** and Xiaomeng Huang, “Retransmission or Redundancy: Transmission Reliability in Wireless Sensor Networks”, The Fourth IEEE International Conference on Mobile Ad-hoc and Sensor Systems (**MASS 2007**), Pisa, Italy, Oct 2007
5. Zhen Chen, Chuang Lin, Jia Ni, Dong-Hua Ruan, Bo Zheng, Zhang-Xi Tan, Yi-Xin Jiang, Xue-Hai Peng, An-an Luo, Bing Zhu, **Yao Yue**, Yang Wang, Peter Ungsunan, Feng-Yuan Ren, “AntiWorm NPU-based Parallel Bloom Filters in Giga-Ethernet LAN”, Proceedings of the 2006 IEEE International Conference on Communications (**ICC 2006**), Istanbul, Turkey, Jun 2006
6. Zhen Chen, Chuang Lin, Jia Ni, Dong-Hua Ruan, Bo Zheng, Yi-Xin Jiang, Xue-Hai Peng, Yang Wang, An-an Luo, Bing Zhu, **Yao Yue**, Feng-Yuan Ren, “AntiWorm NPU-based Parallel Bloom Filters for TCP/IP Content Processing in Giga-Ethernet LAN”, Proceedings of The IEEE Conference on Local Computer Networks 30th Anniversary (**LCN 2005**) - Volume 00, pp. 748-755, Sydney, Australia, Nov 2005

SCHOLARSHIP

Olin Fellowship

One year full support for incoming graduate students

2007

Underwriter Scholarship

A travel grant award to attend 2008 Grace Hopper Celebration of Women in Computing, Keystone, CO, sponsored by Yahoo!

2008

Women of Color Scholarship	2006
A travel grant award to attend 2006 Grace Hopper Celebration of Women in Computing, San Diego, sponsored by Google.	
Tsinghua Friends - Siemens A&D Fellowship	2006
For excellence in graduate study, top 1%	
Excellent Student Leadership Award & Scholarship for the Excellency of Social Services, Tsinghua University	2003, 2004
For excellent leadership in student organizations	
Meng Zhaoying Scholarship, Tsinghua University	2002
For excellence in undergraduate study, top 5%	

ACTIVITIES	Teaching Assistant for CS4450/5450: Computer Networks Instructor: Paul Francis Department of Computer Science, Cornell University	<i>Fall 2008</i>
	Grace Hopper Celebration of Women in Computing Attendee Keystone, CO	<i>Oct 2008</i>
	CRA-W Graduate Cohort Attendee Seattle, WA	<i>Mar 2008</i>
	Teaching Assistant for Student Research Training program Department of Computer Science and Technology, Tsinghua University	<i>Fall 2006</i>
	Oral Presentation of Analyzing the Performance and Fairness of BitTorrent-like Networks Using a General Fluid Model, Globecom'06, San Francisco, CA	<i>Nov 2006</i>
	Reviewer (external) for Frontiers on Computer Science in China, IEEE Journals on Selected Areas in Communications and The IEEE/ACM Transactions on Networking	<i>2005-2006</i>
	Vice President of Physics Department Student Union	<i>2002-2003</i>

REFERENCES Available upon request