Chen LIANG

E-mail: cliang@cs.duke.edu

Educational Background

08/2013-present **Duke University** PhD candidate in Computer Science
07/2011-05/2013 **Brown University** Master of Computer Science
09/2007-07/2011 **South China University of Technology** Bachelor of Computer Science and Technology

Publications

Alvin AuYoung; Yadi Ma; Sujata Banerjee; Jeongkeun Lee; Puneet Sharma; Yoshio Turner; **Chen Liang**; Jeff Mogul Democratic Resolution of Resource Conflicts Between SDN Control Programs. *Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies. ACM*, December 2014.

Andrew D. Ferguson, Arjun Guha, **Chen Liang**, Rodrigo Fonseca, and Shriram Krishnamurthi. Participatory Networking: An API for Application Control of SDNs. In *Proc. ACM SIGCOMM 2013,* August 2013

Andrew D. Ferguson, Arjun Guha, **Chen Liang**, Rodrigo Fonseca, and Shriram Krishnamurthi. Hierarchical Policies for Software Defined Networks. *Proc. Workshop on Hot Topics in Software Defined Networks (Hot-SDN),* August 2012

Weiwei Lin, **Chen Liang**, James Z. Wang, Rajkumar Buyya. Bandwidth-aware divisible task scheduling for cloud computing *Software: Practice and Experience*

Weiwei Lin, James Z.Wang, **Chen Liang** and Deyu Qi, A Threshold-based Dynamic Resource Allocation Scheme for Cloud Computing, *Procedia Engineering Volume 23, 2011*

Professional Skills

- Strength and Interest: Distributed Systems, Software Defined Network, Datacenter Network, Big Data
- Programming Language: C/C++, Java, Python, Haskell
- Platform: Hadoop, Floodlight, Beacon, Flowvisor, POX, ZooKeeper

Research Project

Sep. 2014 to Jan. 2015 **Performance diagnosis in larger operational network**

The goal of this project is to understand, correlate and analyze network events and their impact on performance. The data sources we use in this project include router syslog, config files and SNMP data. A main challenge in this project lies in how to characterize the data, how to identify the hidden internal causal relationships. The purpose is to be able to answer questions such, when and how often do such anomalies happen, and what are the events and devices that cause the network behaving abnormally. In doing this, so far it has been involving several statistical techniques from clustering, anomaly detection and causality analysis

May. 2014 to Aug. 2014 **Network resource conflict resolution with Athens**

(This is the project during internship at HP lab)
In this project, we extended Athens, such that the framework takes into account network level resource utilization conflict. More specifically, before this project, the functionality of Athens is mainly focusing on datacenter resources utilization such as host utilization, fault tolerance guarantee, while this particular part of project aimed at extending the methodology into network resources such as switch memory and switch load balancing. Doing this requires resolving new challenges including internal network representation of Athens, modification to the voting mechanism of Athens, etc.

Oct. 2013 to Apr. 2014 On Demand Controller in Software Defined Network Framework
In this project, we aimed to address the controller scalability issue in SDN by dynamically changing controller placement. More specifically, we applied a linear programming model to compute an optimal scheme that takes into account network statistics such as bandwidth, topology and latency. After obtaining the optimal scheme, we use FlowVisor to gather required switch statistics and apply the optimal scheme at run-time.

Sep. 2012 to Dec. 2013 Instrumenting Hadoop with Software Defined Network Framework
Participatory Network(PANE) is a SDN framework built on top of software defined network protocol openflow. And in this project, we took advantage of PANE to achieve fairness and improve performance of Hadoop jobs by making network resource reservations with PANE. We instrumented a deadline-based model into Hadoop such that bandwidth is allocated to satisfy certain deadline requirement.

Mar. 2012 to May. 2012 Benchmarking PANE with ZooKeeper
We designed and developed a benchmark to test the performance of ZooKeeper with and without PANE which makes bandwidth reservation for its flows. We justified the effectiveness of PANE's ability to protect its traffic from being interfered by background traffics. As part of the benchmark tool, a PANE client library in Java is written and made public to support this and further benchmark.

Oct. 2011 to Dec. 2011 End to End Tracing in Hadoop Using X-trace
Research Purpose: Use tracing framework X-trace to characterize Hadoop work flow and analyze its bottleneck
We instrumented Hadoop using an end-to-end tracing framework X-trace to capture all its traffic in Map/Reduce execution. Our implementation is able to find the dependency relationship between all the map and reduce tasks and time stamp of all the events. By searching for the critical path of dependency from the start of job to its end, we are able to get the knowledge of the bottleneck in its work flow.

Jan.2010 to July. 2011 Bandwidth Scheduling in Datacenter Environment
We built a model that represent applications' bandwidth demand on a cluster based on bandwidth capacity and VM placement. The goal of the model is to find the a allocation scheme that maximizes the number of applications with bandwidth demand satisfied. We implemented our model and verified our algorithm in a cloud computing simulator Cloudsim.

Dec. 2010 to Feb. 2011 Generics Cloud Storage Model Based on HDFS
We explored some core features that is missing in HDFS and expanded HDFS to provide better storage service in terms of replication locality and utilization of physical resources. We instrumented Hadoop to support the replication locality based on the topology and load of datanode. My part in this work was that instead of having a randomly chosen node to store replication, the namenode evaluates all nodes based on network topology and datanode capability. By choosing the optimal datanode for every replication, we improve the overall performance and availability of HDFS.

Nov. 2009 to Nov. 2010 Scheduling Algorithm for Cloud Computing under Load Fluctuation

In this project, we aimed to explore an algorithm that attempt achieves better overall resource utilization under frequent load fluctuation in cloud. We built a model based on short term predictions. Our algorithm measure current load of VMs and once their resource consumption reaches certain thresholds, the algorithm re-compute and adjust resource allocations.

Awards and Recognitions

- 07/2009 First Class Scholarship of Academic Year 2008~2009, awarded by SCUT
- 07/2009 Three-merit Scholar of Academic Year 2008~2009, awarded by SCUT