

Research Statement

Changbin Liu

<http://www.cis.upenn.edu/~changbl/>

1 Introduction

My research centers around dealing with the challenges of managing diverse resources (*i.e.*, compute, storage and network) in distributed systems. My grand vision is to make the building and management of large scale distributed systems safe, reliable, and automated with minimum human effort. In the context of cloud computing, this involves building a robust system to orchestrate cloud resource provisioning, configuration, management and decommissioning. This orchestration is challenging due to the rapid growth of data center sizes, high failure rate of commodity hardware, and the increasing sophistication of cloud services.

To address these challenges, my dissertation work proposes a transactional approach to build and manage cloud resources with data-centric techniques and declarative languages to automatically optimize resource orchestrations. Specifically, it makes following two contributions: (1) A highly available transactional resource orchestration framework for building Infrastructure-as-a-Service (IaaS) cloud. Transactional orchestration and semantics provide a clean abstraction which enables cloud operators to focus on developing high level cloud services without worrying about the complexities of accessing and managing underlying volatile distributed resources; and (2) An automated resource orchestration platform that allows cloud operators to declaratively specify optimization goals and constraints based on provider operational objectives and customer service level agreements (SLAs). Based on these specifications, orchestration commands are automatically generated to optimize resource configurations and allocations within the cloud.

My approach to research involves a combination of analysis, language design, implementation, deployments and empirical measurements on testbeds such as ShadowNet and ORBIT. I actively involve researchers from diverse backgrounds, and take an inter-disciplinary approach to solving problems. As a graduate student, I have written ten first-author publications with five industry lab researchers, and have successfully lead and managed six masters students and three undergraduate students. My work has been published at premier conferences in a variety of research domains, such as networking (ICNP [4], TON [6], COMSNETS [5], HotNets [17]), systems (USENIX ATC [11]), database (CIDR [12], VLDB [13]), and cloud (SOCC [9]). I have strong system building abilities, and enjoy building real and sophisticated systems. All of my research projects have been (or in the process) open-source released to the public, as well as live-demonstrated at SIGCOMM [15], NSDI [10] and COMSNETS [5]. It is worth mentioning that my work on transactional cloud orchestration was highlighted in recent AT&T's official announcement [3, 1] to be integrated into the popular OpenStack platform.

2 Dissertation Work

My dissertation work focuses on *transactional automated* cloud resource orchestration, which is divided into two parts of work: orchestrating cloud resources in a transactional style, and further automating resource provisioning and configuration to satisfy providers operational objectives and customer SLAs.

Transactional cloud orchestration. TROPIC¹ [12, 11, 10] is a highly available, transactional resource orchestration platform for building IaaS cloud infrastructures. TROPIC's orchestration procedures that manipulate physical resources are transactional, automatically guaranteeing atomicity, consistency, isolation and durability (ACID) properties for cloud operations. Transactional semantics provide a clean abstraction for cloud operators to focus on implementing high level cloud services without worrying about the complexities of accessing and managing underlying volatile resources. My experience in building cloud services on top of TROPIC demonstrates its usability in handling errors, enforcing service and engineering rules, and eliminating race conditions. The extensive evaluation of the TROPIC prototype deployed on the ShadowNet testbed shows its capability to support production-scale workload with high degrees of concurrency and provide high availability with low overhead, while always ensuring the transactional semantics of cloud operations.

Automated cloud orchestration. Building on top of the transactional orchestration layer provided by TROPIC, I have further designed and implemented an automated resource orchestration platform called

¹Stands for Transactional Resource Orchestration Platform In the Cloud

COPE² [9]. COPE is a distributed platform that allows cloud providers to perform declarative automated cloud resource orchestration. In COPE, cloud providers specify system-wide constraints and goals based on provider operational objectives and customer SLAs using Colog, a declarative policy language geared towards specifying distributed constraint optimizations. Colog typically results in orders of magnitude less code compared to imperative implementations. COPE integrates the RapidNet [2] declarative networking engine with an off-the-shelf constraint solver, and uses novel query processing strategies for processing Colog programs, by combining the use of bottom-up distributed Datalog evaluation with top-down goal-oriented constraint solving. COPE takes policy specifications and cloud system states as input and then automatically generates orchestration commands to optimize resource allocations such that provider operational objectives and customer SLAs can be better met. COPE is evaluated using realistic scenarios and workloads derived from production cloud services that orchestrate compute, storage, and network resources within and across geographically distributed data centers for load balancing and consolidation. In addition, I have designed and developed Cologne [13, 8], a general declarative distributed constraint optimization engine for executing Colog programs. Interestingly, this platform is also usable for a wide range of optimization scenarios in distributed systems. See Section 4 for more details on another use case of Cologne for managing and optimizing channel selection and routing in wireless networks [5, 6, 4, 14, 7].

3 Future Work

I briefly outline my future plans, providing both a short-term plan on extending my dissertation, and a longer-term agenda.

3.1 Extensions to Dissertation

Moving forward, I am in the process of releasing TROPIC and COPE to the open-source community for use by other researchers and practitioners to quickly build reliable cloud and further automate cloud services. I have also started integrating their capabilities into the popular OpenStack platform. This integration of transactional task management is highlighted in recent AT&T's official announcement [1, 3] of collaboration with OpenStack to launch to a developer friendly cloud.

While TROPIC is designed within a data center, I recognize the emerging trend of services spanning geographically distributed data centers. Therefore I intend to investigate distributed orchestrations coordinated by multiple TROPIC instances. The open research issues include distributed transactions, data replication and partitioning, and their interactions with manageability.

COPE uses Colog declarative policy language to specify distributed constraint optimizations given providers operational objectives and customers SLAs. I intend to formally demonstrate that Colog rules map correctly to the underlying mathematical formulations. Towards this goal, I am leveraging formal analysis techniques that I have developed in the FVN project (See Section 4) to prove the correctness of Colog programs. As an interesting follow-up work, I intend to explore the use of high-level declarative languages for formally reasoning about the convergence time and optimality of distributed systems configurations.

3.2 Looking Beyond

Look beyond my dissertation topic, my long-term goal is to enable automated building and management of large scale distributed systems with safety and robustness. My dissertation work is the first step towards this vision, and I would like to leverage TROPIC and COPE framework to further explore following areas.

Cloud ecosystem. Building upon TROPIC and COPE, I hope to incorporate into the integrated framework recent models on resource provisioning and deployment analysis in the cloud. I plan to build and evaluate novel and more sophisticated cloud services in the framework to further explore new opportunities, which include cloud storage provisioning and distribution, virtualized desktop, database consolidation, and network management tasks incurred by cloud orchestration. The eventual goal is to build a cloud ecosystem which provides a variety of services ranging from infrastructure, platform, software and web. One interesting use case within this ecosystem is in mobile computing. It is expected that as mobile devices become more pervasive, heavy computing tasks will be mostly shifted into the backend in the cloud. Driven by this emerging trend, I plan to explore the use of my dissertation work to deploy mobile computing services more effectively. Some research questions here include reducing data access latency and increasing data

²Stands for Cloud Orchestration Policy Engine

bandwidth between mobile devices and the cloud, providing reliable and highly available services, considering geographical factors which may affect performance.

Distributed optimization. COPE is a flexible distributed constraint optimization platform, which not only applies to the cloud, but also to a variety of vastly different domains. I believe that such a platform has tremendous practical value in facilitating extensible distributed systems optimizations. In fact my prior experience demonstrates COPE’s approach in managing and optimizing adaptive wireless channel selection and routing. As future work, I plan to apply this approach to explore additional use cases in a wide range of emerging domains that involve distributed constraint optimization, including decentralized data analysis and model fitting, resource allocations in other distributed settings, network design and optimizations, etc.

Automated synthesis of distributed optimization programs. As a longer-term agenda, I hope to develop novel techniques to aid distributed system operators to synthesize Colog programs automatically. Currently, COPE and also our wireless configuration use cases require operators to manually specify optimization models including constraints and goals. In large-scale distributed systems, figuring out these goals and constraints may be a non-trivial process. For instance, it was challenging for engineers to pinpoint the root causes of the recent Amazon and Blackberry outages. To automate this process, I plan to explore the use of machine learning and formal analysis techniques (in particular SMT solving) to develop new strategies to automatically *synthesize* goals and constraints, by learning from example configurations and failures in deployed systems. This is an active area of research currently in the programming languages community, as shown by recent work in program sketching, concurrent programming, and auto-completion of spreadsheets. This is an inter-disciplinary research topic that I hope to explore, in collaboration with researchers in machine learning, programming languages, and formal methods.

4 Other Work

Prior to my dissertation research, I worked on a variety of projects. Here I provide a brief summary of each project:

Policy-based Unified Multi-Radio Architecture (PUMA). PUMA [5, 13, 7] is a policy-based extensible platform that combines channel selection and routing within a common declarative framework for multi-radio wireless mesh networks. I lead the design and implementation of PUMA, in collaboration with Raytheon BBN Technologies. In PUMA, users formulate channel selection policies as optimization goals and constraints that are concisely declared using a declarative policy language. PUMA integrates the RapidNet declarative networking engine with a constraint solver to realize a variety of declarative wireless routing and channel selection protocols, and in addition, provides avenues to optimize across route and channel selection policies. I have carried out extensive evaluations of PUMA in simulation and on the ORBIT testbed. Experimental results demonstrate that PUMA can flexibly and efficiently implement a variety of centralized and distributed channel selection protocols that result in significantly higher throughput compared to baseline protocols.

Declarative Adaptive Wireless Networks (DAWN). DAWN [4, 6, 14] is a declarative platform that creates highly adaptive policy-based Mobile Ad hoc Networks (MANETs). Collaborating with Raytheon BBN Technologies under the DARPA Wireless Networks After Next (WNaN) program, I lead the system design and demonstrate that a variety of MANET routing protocols (*i.e.*, proactive, reactive, DTN) can be specified concisely as declarative networks. Moreover, policy-based adaptation can be expressed in the same declarative language to create hybrid adaptive protocols based on various network and traffic conditions. I further propose inter-protocol forwarding techniques that ensure packets are able to seamlessly traverse across clusters of nodes running different protocols selected based on their respective policies. I have developed a full-fledged prototype of DAWN using RapidNet declarative networking system and experimentally validated policy-based adaptive MANETs in dynamic settings using a combination of ns-3 simulations and deployment on the ORBIT testbed. Our evaluation demonstrates that DAWN is a practical approach for deploying MANET protocols, achieving concise specifications, policy-based adaptation, and inter-protocol forwarding at low performance overhead.

RapidNet declarative networking engine. RapidNet [2, 16] is an open-source development toolkit which provides a unified platform for rapid implementation, simulation and experimentation of network protocols. RapidNet utilizes declarative networking to specify network behavior by integrating a declarative networking engine with the emerging network simulator 3 (ns-3) platform. RapidNet has been demonstrated at SIGCOMM [15], and actually deployed on the ORBIT wireless testbed.

Formally Verifiable Networking (FVN). FVN [17, 18] is a novel approach towards unifying the design, specification, implementation, and verification of networking protocols within a logic-based framework. FVN utilizes theorem proving, a well established verification technique where logic-based axioms that automatically capture network semantics are generated, and a user-driven proof process is used to establish network correctness properties. We validate various protocols such as Border Gateway Protocol (BGP) and Hybrid Link State protocol, using both PVS theorem prover and real implementation and experimentation.

References

- [1] A better cloud for developers. <http://www.attinnovationspace.com/innovation/story/a7779660>.
- [2] RapidNet Declarative Networking Engine. <http://netdb.cis.upenn.edu/rapidnet/>.
- [3] Transactional Orchestration for Cloud Computing. <http://trac.research.att.com/trac/dmf/>.
- [4] **Changbin Liu**, R. Correa, Xiaozhou Li, P. Basu, B.T. Loo, and Yun Mao. Declarative policy-based adaptive MANET routing. In *17th IEEE International Conference on Network Protocols (ICNP)*, pages 354–363, 2009.
- [5] **Changbin Liu**, Ricardo Correa, Harjot Gill, Tanveer Gill, Xiaozhou Li, Shivkumar Muthukumar, Taher Saeed, Boon Thau Loo, and Prithwish Basu. PUMA: Policy-based Unified Multi-radio Architecture for Agile Mesh Networking (full paper and demonstration). In *The Fourth International Conference on COMMunication Systems and NETWORKS (COMSNETS)*, 2012.
- [6] **Changbin Liu**, Ricardo Correa, Xiaozhou Li, Prithwish Basu, Boon Thau Loo, and Yun Mao. Declarative policy-based adaptive mobile ad hoc networking. *Networking, IEEE/ACM Transactions on*, 2011.
- [7] **Changbin Liu**, Xiaozhou Li, Shivkumar C. Muthukumar, Harjot Gill, Taher Saeed, Boon Thau Loo, and Prithwish Basu. A Policy-based Constraint-solving Platform Towards Extensible Wireless Channel Selection and Routing. In *PRESTO: Proceedings of the ACM workshop on Programmable routers for extensible services of tomorrow*, 2010.
- [8] **Changbin Liu** and Boon Thau Loo. Declarative Constraint Optimization in Distributed Systems. In *Workshop on Languages for Distributed Algorithms (LADA)*, 2012.
- [9] **Changbin Liu**, Boon Thau Loo, and Yun Mao. Declarative automated cloud resource orchestration. In *Proceedings of the 2nd ACM Symposium on Cloud Computing, SOCC '11*, pages 26:1–26:8, 2011.
- [10] **Changbin Liu**, Yun Mao, Xu Chen, Mary F. Fernandez, Boon Thau Loo, and Kobus Van der Merwe. Towards transactional cloud resource orchestration. In *NSDI (poster and demo)*, 2011.
- [11] **Changbin Liu**, Yun Mao, Xu Chen, Mary F. Fernandez, Boon Thau Loo, and Kobus Van der Merwe. TROPIC: Transactional Resource Orchestration Platform In the Cloud. In *2012 USENIX Annual Technical Conference (ATC'12)*, 2012.
- [12] **Changbin Liu**, Yun Mao, Jacobus Van der Merwe, and Mary Fernandez. Cloud Resource Orchestration: A Data-Centric Approach. In *Proceedings of the biennial Conference on Innovative Data Systems Research (CIDR)*, 2011.
- [13] **Changbin Liu**, Lu Ren, Boon Thau Loo, Yun Mao, and Prithwish Basu. A Declarative Distributed Constraint Optimization Platform. In *Accepted at 38th International Conference on Very Large Databases (VLDB'12)*, 2012.
- [14] **Changbin Liu**, Yun Mao, Mihai Oprea, Prithwish Basu, and Boon Thau Loo. A declarative perspective on adaptive manet routing. In *PRESTO: Proceedings of the ACM workshop on Programmable routers for extensible services of tomorrow*, 2008.
- [15] Shivkumar C. Muthukumar, Xiaozhou Li, **Changbin Liu**, Joseph B. Kopena, Mihai Oprea, and Boon Thau Loo. Declarative toolkit for rapid network protocol simulation and experimentation. In *SIGCOMM (demo)*, 2009.
- [16] Shivkumar C. Muthukumar, Xiaozhou Li, **Changbin Liu**, Joseph B. Kopena, Mihai Oprea, Ricardo Correa, Boon Thau Loo, and Prithwish Basu. RapidMesh: declarative toolkit for rapid experimentation of wireless mesh networks. In *4th ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization (WiNTECH)*, 2009.
- [17] Anduo Wang, Limin Jia, **Changbin Liu**, Boon Thau Loo, Oleg Sokolsky, and Prithwish Basu. Formally Verifiable Networking. In *8th Workshop on Hot Topics in Networks (ACM SIGCOMM HotNets-VIII)*, 2009.
- [18] Anduo Wang, Boon Thau Loo, **Changbin Liu**, Oleg Sokolsky, and Prithwish Basu. A Theorem Proving Approach Towards Declarative Networking. In *22nd International Conference on Theorem Proving in Higher Order Logics (TPHOLs) emerging trends proceedings*, 2009.