

Riza O. Suminto

Address Cloudera
5470 Great America Pkwy, Santa Clara, CA 95054

Web <https://www.rizasuminto.com>
Email riza.on@gmail.com

Education

2013 - 2019 **University of Chicago**, Chicago, IL, USA
Ph.D. in Computer Sciences
Thesis: Mitigating Cascading Performance Failures and Outages in Cloud Systems
Advisor: Prof. Haryadi S. Gunawi

2007 - 2011 **Universitas Gadjah Mada**, Yogyakarta, Indonesia
Bachelor in Computer Science, graduated with cum laude and 3.68 GPA

Research Interests

Areas Distributed Systems, Cloud Computing, and Operating Systems.
Focuses Reliability and Performance.

Work Experiences

May 2022 - Present **Apache Impala PMC Member and Committer**, The Apache Software Foundation
Project Management Committee of [Apache Impala](#).

Jan 2020 - Present **Senior Software Engineer**, Cloudera, Palo Alto, CA, USA
Working on research and development of [Apache Impala](#).

Jan 2014 - Dec 2019 **Research Assistant**, University of Chicago, Chicago, IL, USA
Working on “*improving performance reliability of distributed systems and cloud systems*” topic with [Prof. Haryadi S. Gunawi](#) in [UCARE](#) Group.

May 2018 - Aug 2018 **Performance Engineer Intern**, Cloudera, Palo Alto, CA, USA
Working on “*automatic microbenchmark of public cloud storage*”, such as Amazon Simple Storage Service (S3) and Azure Data Lake Storage (ADLS).

May 2016 - Aug 2016 **Research Intern**, Samsung Research America, Mountain View, CA, USA
Worked on “*robust distributed IOT system*” project, with [Masoud Saeida Ardekani](#), and [Rayman Preet Singh](#).

Sept 2013 - Dec 2013 **Teaching Assistant**, University of Chicago, Chicago, IL, USA
TA in Operating Systems course.

2010 - 2011 **Client Developer Specialist**, Microsoft Innovation Center (MIC) Yogyakarta
Responsible for dealing with product innovation associated with desktop applications, analysis, and optimization.

Publications

Jeffrey F. Lukman, Huan Ke, Cesar A. Stuardo, [Riza O. Suminto](#), Daniar H. Kurniawan, Dikaimin Simon, Satria Priambada, Chen Tian, Feng Ye, Tanakorn Leesatapornwongsa, Aarti Gupta, Shan Lu, Haryadi S. Gunawi. **FlyMC: Highly Scalable Testing for Complex Interleavings in Distributed Systems**. In *Proceedings of the 14th EuroSys Conference (EuroSys)*, March 2019.

Cesar A. Stuardo, Tanakorn Leesatapornwongsa, Riza O. Suminto, Huan Ke, Jeffrey F. Lukman, Wei-Chiu Chuang, Shan Lu, Haryadi S. Gunawi. **ScaleCheck: A Single-Machine Approach for Discovering Scalability Bugs in Large Distributed Systems.** In *Proceedings of the 17th USENIX Conference on File and Storage Technologies (FAST)*, February 2019.

Praditya Raudi, Dody Dharma, Afwarman Manaf, [Riza O. Suminto](#), Haryadi S. Gunawi **Detection of Nodes with Degraded Network Interface Card at Startup Mapreduce Phase using Path-Based Approach.** In *Proceedings of 2018 International Conference on Electrical Engineering and Computer Science (ICEECS)*, November 2018.

Haryadi S. Gunawi, [Riza O. Suminto](#), Russell Sears, Casey Golliher, Swaminathan Sundararaman, Xing Lin, Tim Emami, Weiguang Sheng, Nematollah Bidokhti, Caitie McCaffrey, Deepthi Srinivasan, Biswaranjan Panda, Gary Grider, Parks M. Fields, Kevin Harms, Robert B. Ross, Andree Jacobson, Robert Ricci, Kirk Webb, Peter Alvaro, H. Birali Runesha, Mingzhe Hao, Huaicheng Li **Fail-Slow at Scale: Evidence of Hardware Performance Faults in Large Production Systems.** In *ACM Transactions on Storage (TOS), Volume 14, Issue #3*, November 2018.

Haryadi S. Gunawi, [Riza O. Suminto](#), Russell Sears, Casey Golliher, Swaminathan Sundararaman, Xing Lin, Tim Emami, Weiguang Sheng, Nematollah Bidokhti, Caitie McCaffrey, Gary Grider, Parks M. Fields, Kevin Harms, Robert B. Ross, Andree Jacobson, Robert Ricci, Kirk Webb, Peter Alvaro, H. Birali Runesha, Mingzhe Hao, Huaicheng Li. **Fail-Slow at Scale: Evidence of Hardware Performance Faults in Large Production Systems.** In *Proceedings of the 16th USENIX Conference on File and Storage Technologies (FAST)*, February 2018.

Masoud Saeida Ardekani, Rayman Preet Singh, Nitin Agrawal, Douglas B. Terry, [Riza O. Suminto](#). **Rivulet: Fault-Tolerant Platform for Smart-Home Applications.** In *Proceedings of the 18th International Middleware Conference (Middleware)*, December 2017.

Mingzhe Hao, Huaicheng Li, Michael Hao Tong, Chrisma Pakha, [Riza O. Suminto](#), Cesar A. Stuardo, Andrew A. Chien, Haryadi S. Gunawi. **MittOS: Operating System Supports for Millisecond Tail Tolerance in Data- Parallel Storage.** In *Proceedings of the 26th ACM Symposium on Operating Systems Principles (SOSP)*, October 2017.

[Riza O. Suminto](#), Cesar Stuardo, Alexandra Clark, Huan Ke, Tanakorn Leesatapornwongsa, Bo Fu, Daniar H. Kurniawan, Vincentius Martin, Uma Maheswara Rao G., and Haryadi S. Gunawi. **PBSE: A Robust Path-Based Speculative Execution for Degraded-Network Tail Tolerance in Data-Parallel Frameworks.** In *Proceedings of the 8th ACM Symposium on Cloud Computing (SoCC)*, September 2017.

Haryadi Gunawi, Huan Ke, Tanakorn Leesatapornwongsa, Jeffrey Lukman, Cesar Stuardo, and [Riza O. Suminto](#). **Scalability Bugs: When 100-Node Testing is Not Enough.** In *Proceedings of the 16th Workshop on Hot Topics in Operating Systems (HotOS)*, May 2017.

Haryadi S. Gunawi, Mingzhe Hao, [Riza O. Suminto](#), Agung Laksono, Anang D. Satria, Jeffry Adityatama, Kurnia J. Eliazar. **Why Does the Cloud Stop Computing? Lessons from Hundreds of Service Outages.** In *Proceedings of the 7th ACM Symposium on Cloud Computing (SoCC)*, October 2016.

Haryadi S. Gunawi., Thanh Do, Agung Laksono, Mingzhe Hao, Tanakorn Leesatapornwongsa, Jeffrey F. Lukman, and [Riza O. Suminto](#). **What Bugs Live in the Cloud?: A Study of Issues in Scalable Distributed Systems.** In *USENIX;login: Magazine, Vol. 40, No. 4.*, 2015

[Riza O. Suminto](#), Agung Laksono and Anang D. Satria, Thanh Do, Haryadi S. Gunawi. **Towards Pre- Deployment Detection of Performance Failures in Cloud Distributed Systems.** In *Proceedings of the 7th USENIX Conference on Hot Topics in Cloud Computing (HotCloud)*, July 2015.

Teddy Mantoro, Jelita Asian, [Riza Oktavian](#), Media Anugerah Ayu. **Optimal translation of English to Bahasa Indonesia using statistical machine translation system.** In *Proceedings of the 5th International Conference on Information and Communication Technology for the Muslim World (ICT4M)*, March 2013.

Talks

Sept 2017 **PBSE: A Robust Path-Based Speculative Execution for Degraded-Network Tail Tolerance in Data-Parallel Frameworks.** (SoCC '17)

Sept 2017 **Scalability Bugs: When 100-Node Testing is Not Enough.**
(Chameleon User Meeting 2017)

Jul 2015 **Towards Pre-Deployment Detection of Performance Failures in Cloud Distributed Systems.**

Posters

Riza O. Suminto, Cesar Stuardo, Alexandra Clark, Huan Ke, Bo Fu, Tanakorn Leesatapornwongsa, Vincentius Martin, Daniar Kurniawan, Haryadi S. Gunawi. **PBSE: Path-Based Speculative Execution for Robust Tail Tolerance in Data-Parallel Systems**. In *Poster Session of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, November 2016.

Tanakorn Leesatapornwongsa, Cesar Stuardo, Huan Ke, Jeffrey F. Lukman, Riza O. Suminto, Daniar H. Kurniawan, Haryadi S. Gunawi. **SCk: Scale-Checking and Debugging Scalability Bugs on One Machine**. In *Poster Session of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, November 2016.

Riza O. Suminto, Agung Laksono and Anang D. Satria, Thanh Do, Haryadi S. Gunawi. **Limplock Static Analysis: Preventing Slow Performance Lock-up in Cloud Systems**. In *Poster Session of the 11th USENIX Symposium on Operating System Design and Implementation (OSDI)*, October 2014.

Awards and Honors

2018	FAST '18 Student Grant , USENIX
2017	Middleware '17 Student Grant , ACM
2017	SOSP '17 Student Grant , USENIX
2017	SoCC '17 Student Scholarship , ACM
2015	ATC '15 Student Grant , USENIX
2015	UU Fellowship , University of Chicago
2014	OSDI '14 Grants for Women/Diversity Grant , USENIX
2011 & 2010	3rd place , Imagine Cup Indonesia
2007	Bronze Medalist , 19th International Olympiad in Informatics
2006	Gold Medalist , 5th National Science Olympiad in Informatics, Indonesia

Projects

Please click on the titles for the reports or more information

Research Projects

2017 - 2019	FlyMC: Highly Scalable Testing for Complex Interleavings in Cloud Systems (EuroSys '19) FlyMC is scalable testing approach for datacenter/cloud computing systems with unique ability to overcome path/state-space explosion problem in testing workload with complex interleavings of messages and faults. I contribute in this project by building and instrumenting FlyMC to capture race problem in Hadoop MapReduce framework.
2017 - 2019	ScaleCheck: A Single-Machine Approach for Discovering Scalability Bugs in Large Distributed Systems (FAST '19) ScaleCheck is an approach for discovering scalability bugs and for democratizing large-scale testing. ScaleCheck employs a program analysis technique, for finding potential causes of scalability bugs, and series of colocation techniques, for testing implementation code at real scales but doing so on just a commodity PC. I contribute to the project by instrumenting ScaleCheck to capture various scalability bugs in Hadoop HDFS system.

- 2017 **Fail-Slow at Scale: Evidence of Hardware Performance Faults in Large Production Systems** (*FAST '18*)
Fail-slow hardware, is a hardware that is still running and functional but in a degraded mode, slower than its expected performance. We present a study of 101 reports of fail-slow hardware incidents, collected from 12 institutions large-scale cluster and show that all hardware types can exhibit performance faults. I contribute in this project by collecting and classifying anecdotal data from different institutions.
- 2016 - 2017 **MittOS: Operating System Supports for Millisecond Tail Tolerance in Data-Parallel Storage** (*SOSP '17*)
MittOS is an extended Linux kernel with supports for helping applications such as NoSQL storage cut ms-level tail latencies. MittOS principle is that the OS layer should be transparent of resource busyness such that applications can instantly failover to other less-busy replicas without waiting. I make initial implementation to enable application to pass SLA information to kernel IO stack.
- 2016 - 2017 **PBSE: A Robust Path-Based Speculative Execution for Degraded-Network Tail Tolerance in Data-Parallel Frameworks** (*SoCC '17*)
This work reveals fundamental flaws of Speculative Execution (SE) implementations under a unique fault model: performance-degraded network device. To address this fault model, I present *path-based speculative execution (PBSE)*, which contains three key ingredients: path progress (progresses of individual network communications in every stages), path diversity, and path- straggler detection and speculation.
- 2016 **Rivulet: Fault-Tolerant Platform for Smart-Home Applications** (*Middleware '17*)
I join Samsung Research America on summer 2016 as intern. I work on developing Rivulet, a fault-tolerant distributed platform for running smart-home applications. It can tolerate failures typical for a home environment, such as link losses, network partitions, sensor failures, and device crashes. I contribute in this project by building initial implementation of Rivulet delivery service protocol.
- 2015 - 2017 **Scalability Bugs: When 100-Node Testing is Not Enough** (*HotOS '17*)
My colleagues and I present scale-check methodology that empowers developers to check scalability of distributed systems implementations and helps them reproduce and debug scalability bugs economically on one machine. I implement the scale-check methodology to detect scalability bugs on HDFS.
- 2015 - 2016 **Why Do the Clouds Stop? Lessons from Hundreds of High-Profile Outages** (*SoCC '16*)
Our group and I conduct a cloud outage study of most popular applications and providers in cloud system. We collected high-profile outages across 35 popular cloud services published over 6 years from 2009 to 2014 from press releases, credible news provider, and outage report center. We categorize these issues by their root causes, implications, and fix strategies.
- 2014 - 2015 **Towards Pre-Deployment Detection of Performance Failures in Cloud Distributed Systems** (*HotCloud '15*)
This work addresses the following question: *How to detect performance bugs in real distributed systems code and do so prior to deployment and in a fast and complete manner?* I built System Performance Verifier (SPV), a framework that takes real system code and automatically generates the model, environment, and scenarios to permute. SPV parse real system code into formal model, Colored Petri Net (CPN), and verify the model performance using CPN model checker.
- 2012 - 2013 **Optimal translation of English to Bahasa Indonesia using statistical machine translation system** (*ICT4M '13*)
This work searched for the most optimal parameter for statistical machine translation from English to Bahasa Indonesia. I built and cleaned language corpuses of both languages that used in training phase.

Industry Projects

- 2011 **PrimagamaPlus.com** (*MIC Yogyakarta*)
As a systems analyst and software engineer, I built e-learning platform based on ASP.NET, Umbraco CMS, and Silverlight. This e-learning platform serves more than hundred thousand students of Primagama Tutoring Institution.
- 2009 - 2010 **Integrated Hospital Information Systems** (*Happy Land Hospital, Yogyakarta*)
I built web-based integrated hospital information system, using PHP and PostgreSQL. The system handles all of the management of the hospital including patient medical records, pharmacy inventory, finance administration, and others. I worked as a software engineer in this project.

2009 **Pharmacy Warehouse Information Systems** (*Department of Health Yogyakarta*)
As a systems analyst and software engineer, I built information system to assist pharmaceutical warehouse inventory management and distribution to health centers and hospitals in Yogyakarta. Applications developed by a team of OmahTI UGM on behalf of PT. Gamatechno.

Professional Service

2016 **External reviewer** *TCC '16*, IEEE Transactions on Cloud Computing

Technical Skills

Operating Systems	Linux.
Distributed Systems	Apache Impala, Hadoop, HDFS.
Databases	PostgreSQL, MySQL.
Programming	Java, C++, Python, PHP.