

# Debessay (Debish) Fesehaye

+217 819 2880

debish10@gmail.com

<https://www.linkedin.com/pub/debish-fesehaye/58/a17/26> or <http://debish.weebly.com>

## EDUCATION

---

### **Ph.D. in Computer Science:**

*University of Illinois-Urbana Champaign (UIUC), IL, USA,*

*Completion date:* August 2013.

*Thesis Title:* Efficient Cross-layer Routing and Congestion Control Architectures for Distributed Systems. *Thesis Advisor:* Professor Klara Nahrstedt

### **Master of Science in Computer Science:**

*University of Stellenbosch, SA, Cum Laude (G.P.A. = 4.0/4.0)*

*Thesis Title:* Analytic Models of TCP/IP Performance. *Thesis Advisor:* Professor Anthony E Krzesinski

### **B.Sc. Hnrs Degree in Operational Analysis:**

(A combination of Applied Mathematics, Computer Science and Statistics)

*University of Stellenbosch, SA, Cum Laude.*

## RECENT WORK EXPERIENCE

---

### **VMware, Inc.: January 2016 - Present**

*Job:* Sr. Member of Technical Staff (Palo Alto, CA, USA) position doing Research and Development (R&D) and implementation work.

*Tasks:* Design Algorithms for Small & Large Scale Systems, Coding/Implementation, Building Prototypes, Experimenting, Discussing with Various Teams, Filing Patents, Writing Papers.

*Job Main Keywords:* System Anomaly Detection, System Root Cause Analysis, Data Analytics, Machine Learning, Group Clustering, Metric Ranking, Distributed Systems.

*Languages and Tools:* Java, Python, Spark Mlib, Weka, scikit-learn.

### **Neustar Inc.: June 2013 - December 2015**

*Job:* Sr. (Data) Scientist (UIUC Research Park, June 2013 - Oct 2014) and Lead Software Engineer (San Francisco, Nov 2014 - Present) positions doing Research and Development (R&D) and implementation work.

*Tasks:* Design Algorithms for Small & Large Scale Systems, Coding/Implementation, Building Prototypes, Experimenting, Discussing with Product Teams, Filing Patents, Writing Papers, Open-Sourcing.

*Job Main Keywords:* Big Data Security Analytics (IP Intelligence/Reputation, Big Data Mining, Machine Learning, Malicious Website/IP Detection, String Similarity Metrics and Algorithms, Big Data Visualization, UltraDNS data, Registry data), AdAdvisor Efficiency (Device Identity/Device Fingerprinting), Communications Networks (Content Routing, CDN-SDN, Cloud Networking).

*Languages and Tools:* C++/C, Java, Hive/SQL, Apache Spark, Spark Mlib, Weka, Mahout, Hadoop, Map Reduce, Hadoop Pipes, Apache Http Server, MySQL, MangoDB, PHP, Bash & TCL/oTCL Scripting.

### **University of Illinois at Urbana Champaign: August 2007 - June 2013**

*Roles:* Teaching and Research Assistant, Collaboration with Boeing R&D

### **Huawei USA R&D: Spring 2013**

*Role:* Fulltime Engineering Intern

### **Qualcomm R&D: Summers of 2009, 2010, 2011**

*Role:* Fulltime Engineering Intern

## **PROJECTS**

---

### **CURRENT PROJECTS**

#### **Performance Analytics**

Performance Analytics of Large Scale Systems:

*Role:* Researching, Designing and Implementing Algorithms for Performance Anomaly Detection and Root Cause Analysis of Large Scale Distributed Systems in Offline and Realtime Modes.

*Status:* Designed and implemented multiple root cause metric identification, anomaly detection and group (physical/virtual device, time series groups, etc) clustering algorithms. Currently in the process of integrating algorithms into VMware vSAN performance analysis and other VMware products. (January 2016 - current).

*Software/tools:* Java, Python, Spark MLib, Weka, SciPy, scikit-learn.

### **PAST (Less Active) PROJECTS**

#### **Security Analytics (Big Data) Projects**

IP Intelligence Products (IP Geolocation, IP Reputation):

*Role:* Researching, Designing and Implementing IP Reputation (IPR) Scoring Mechanisms using existing and custom machine learning algorithms on multiple Big Data sources.

*Status:* Newer versions of IPR scoring product released and continuously upgraded with more data sources and customized scoring algorithms (November 2014 - December 2015).

*Software/tools:* Java, Hive, Hive UDF, Hadoop MapReduce, MongoDB, Spark, Spark MLib, Weka, Mahout.

Big Data Mining and Machine Learning:

*Malicious Domain/URL Detection Using Big DNS and Domain Registry Datasets*

*Efficient String Similarity Metric Algorithms and Their Applications to Big Data Mining*

*Fast Machine Learning Algorithm For Multi-Attribute Big Data Analysis*

*Status:* Experimented with various machine learning algorithms with Decision Trees and Random Forests giving the best classification accuracy. Designed various string similarity metrics and implemented (using C++) efficient algorithms to calculate them (Spring 2013 - current).

Live Visualization of Malicious Website/Domain Activities:

The goal of this project is to *observe the malicious on-line activities in real-time using hourly Neustar UltraDNS datasets.*

*Status:* Tool designed, implemented and is operational in collaboration with another Neustar colleague and a CS/UIUC MSc student intern (Summer 2014 ).

#### **AdAdvisor Efficiency Projects**

Networked Device Identity (NDI):

*Enhancing AdAdvisor (targeted advertising) Services Using Privacy Aware NDI/Device Fingerprint-*

*ing Instead of Only Cookies*

*Status:* System designed and analyzed using Neustar real world big Hadoop Cluster datasets. Initial prototype built in collaboration with a CS/UIUC undergraduate student intern (Summer 2014).

## **Communication Networks (Cloud, SDN, CDN, Wireless, etc) and Other Projects**

Cloud Datacenter Architectures & SDNs:

*SCDA: SLA-aware Cloud Datacenter Architecture for Efficient Content Storage and Retrieval*

*Status:* System designed, implemented in the NS2 network simulator (C++ and oTCL based) and analyzed using extensive trace-based experiments.

Private/Personal Clouds with Incentived, Prioritized and Efficient Content Routing:

*CIM: Efficient and Scalable Content Index Manager*

*Hincent: Quick Content Distribution With Priorities and High Incentives*

*Status:* System designed, implemented in the NS2 network simulator (C++ and oTCL based) and analyzed using extensive trace-based experiments. Initial prototype built in collaboration with a CS/UIUC MSc student intern (Summer 2014).

Optimizing Private CDNs using SDNs:

*Optimizing Private Networks to Achieve Realtime SLA and QoS Guarantees at Scale*

*Status:* System designed and analyzed. Initial prototype built in collaboration with a CS/UIUC PhD student intern (Summer 2014).

XDI Personal Cloud Content Index:

The goal of this project is to *build a personal cloud content index application using the XDI personal cloud language.*

*Status:* Application designed and initial prototype built in collaboration with other Neustar colleagues and a CS/UIUC undergrad student intern (Summer 2014 ).

Fast Content Transfer and Congestion Control Protocol:

*QCP: Finishing Flows Faster with A Quick congestion Control Protocol (QCP),*

Illinois (UIUC) Department of Computer Science, Illinois.

Cross-Layer Routing and Congestion Control Architectures:

*Efficient Cross-layer Routing and Congestion Control Architectures for Distributed Systems,*

Illinois (UIUC) Department of Computer Science, Illinois.

DoS and DDoS Mitigation Schemes:

*A Packet Accounting System to Limit the Effects of DoS & DDoS,*

Illinois (UIUC) Department of Computer Science, Illinois, USA.

Named Data Networking (NDN) → Content Centric Networking (CCN):

*Design and Analysis of Scalable NDN-Based Conferencing Architecture,*

Huawei US R&D, (Interim Project), San Jose, CA, USA, March-June 2013.

Cloud-Cloudlet Architectures:

*Analysis and Design of Cloud and Cloudlet Architectures and Services,*

Boeing Research Assistantship (RA), Illinois (UIUC) Department of Computer Science, Illinois. Spring 2012-Spring 2013.

TCP, Flow Prioritization and Energy-Efficient Resource Utilization:

*Analysis and Design of TCP, Flow Prioritization and Energy-Efficient Resource Utilization Schemes,* Corporate R&D, Qualcomm Inc. (Summer Engineering Intern), May-Sept 2011.

Publish-Subscribe (Message Exchange) Systems:

*Analysis and Design of Co-operative Publish-Subscribe (Message Exchange) Systems*,  
Boeing Research Assistantship (RA), Illinois (UIUC) Department of Computer Science, Illinois, Fall 2010-Spring 2011.

Wireless Interface Management:

*Design and Analysis of Wireless Interface Management and Flow Prioritization Schemes*,  
Corporate R&D, Qualcomm Inc. (Summer Engineering Intern), San Diego, USA, May-August 2010.

3GPP-LTE (4G Wireless Network Technology):

*Analysis and Design of Protocols and Algorithms for 4G Networks (The 3GPP-LTE)*,  
Corporate R&D, Qualcomm Inc. (Summer Engineering Intern), San Diego, USA, June-August 2009.

LiteOS (Sensor Network Operating System):

*Embedded Sensor Network Energy Management System*,  
Adding Energy Management Features to LiteOS  
Illinois (UIUC) Department of Computer Science, Illinois, USA, Spring 2008.

Communication Networks (Wired):

*Analysis and Design of Protocols and Algorithms for Communication Networks*,  
Computer (Network) Engineering Researcher at Ghent University, Faculty of Engineering, Belgium,  
Spring 2006 - Fall 2007.

*Modeling and Analysis of Communication Networks (The Internet)* ,  
Student and Researcher at the University of Stellenbosch, Unit of The Telkom-Siemens Centre of  
Excellence in ATM & Broadband Networks & their Applications, SA, Yr. 2005.

Group Operations Research Projects on:

The *Optimal Weekly Water Release Strategy* for the Kerom Dam in Western Cape,  
*Modeling Human Caused Mortality of the KTP Lion Pride*, Under Different Environmental Condi-  
tions Subject to Birth and Sex Ratio Biases,  
*Simulation Study of Wood Manufacturing Process* at the Cutting Section of Infinite Wood Factory,  
*Modeling the Inventory System* of ClickAbox Cardboard boxes Manufacturing Factory,  
Department of Applied Mathematics, University of Stellenbosch, SA.

## PUBLICATIONS

---

### In Preparation (Ongoing Work):

Debessay Fesehaye and Klara Nahrstedt. *Efficient Server Selection and Quick Content Transfer Architecture for Cloud Datacenter Networks*.

Debessay Fesehaye and Klara Nahrstedt. *Fast Content Routing and Scalable Content Index Management System*.

Debessay Fesehaye. *Weighted String Similarity Metric Algorithms And Their Applications To Data Classification*.

### Published Thesis Works:

Debessay Fesehaye. *Efficient Cross-layer Routing & Congestion Control Architectures for Distributed Systems*, *PhD Dissertation*. Department of Computer Science, University of Illinois at Urbana-Champaign (UIUC), IL, USA. August 2013.  
(Available at <https://www.ideals.illinois.edu/handle/2142/45430>).

Debessay Fesehaye. *Analytic Models of TCP/IP Performance*, MSc Thesis. Department of Computer Science, University of Stellenbosch, SA.  
(Available at <http://scholar.sun.ac.za/handle/10019.1/16608>).

**Published/Accepted Papers (Full Papers):**

*(a) General Distributed Systems and Cloud Computing:*

Debessay Fesehaye and Klara Nahrstedt. *SCDA: SLA-aware Cloud Datacenter Architecture for Efficient Content Storage and Retrieval*. 7th IEEE International Conference on Cloud Computing (IEEE Cloud 2014). Alaska, USA. June 2014. (Longer Version Available as UIUC Technical Report at <https://www.ideals.uiuc.edu/bitstream/handle/2142/36012/latestSCDAversion.pdf>)

Debessay Fesehaye and Jun Wei. *SNC: Scalable Named Data Networking (NDN)-Based Conferencing Architecture*. 11th Annual IEEE Consumer Communications and Networking Conference (CCNC2014)- Special Session on Information Centric Networking, Las Vegas, USA. January 2014.

Debessay Fesehaye and Klara Nahrstedt. *Hincent: Quick Content Distribution With Priorities and High Incentives*. 10th Annual IEEE Consumer Communications and Networking Conference (CCNC2013), Las Vegas, USA. January 2013.  
(Longer Version Available as UIUC Technical Report at <https://ideals.illinois.edu/bitstream/handle/2142/36035/extendedHincent.pdf>)

*(b) Big Data Mining And Machine Learning:*

Debessay Fesehaye, Lenin Singaravelu, Chien-Chia Chen, Xiaobo Huang, Amitabha Benerjee, Ruijin Zhou, and Rajesh Somasundaran. *Group Clustering Using Inter-Group Dissimilarities*. IEEE the 37th International Conference on Distributed Computing Systems. Atlanta, Georgia, USA. June 2017.

Debessay Fesehaye, Lenin Singaravelu, Amitabha Benerjee, Ruijin Zhou, Xiaobo Huang, Chien-Chia Chen and Rajesh Somasundaran. *Performance Analysis of Large Scale Distributed Systems by Ranking Dominant Features*. VMware Innovation Series: R&D Innovation Offsite (RADIO-2017). San Francisco, CA, USA. May 2017 and The 4th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT 2017), Austin, TX, Dec 2017.

Debessay Fesehaye and Armand Prieditis. *Efficient String Similarity Algorithms and Their Applications to Data Mining & Analysis*. Neustar Tech Summit. San Francisco, USA. June 2014.

*(c) Mobile And Sensor Networks:*

Debessay Fesehaye, Yunlong Gao, Klara Nahrstedt and Guijun Wang. *Impact of Cloudlets on Interactive Mobile Cloud Applications*. 16th IEEE International Enterprise Distributed Object Computing Conference (EDOC 2012). Beijing, China. September 2012.  
(Available at <http://web.engr.illinois.edu/~dkassa2/PID2397309.pdf>).

Debessay Fesehaye, Klara Nahrstedt and Guijun Wang. *Analytical Models of Short-Message Reliability in Mobile Wireless Networks*. 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2011). Miami, FL, USA. October 2011.

Pete Barany, Debessay Fesehaye and Rohit Kapoor. *On the Performance of TCP/IP in 3GPP LTE (4G)*. Qualcomm Technical Conference (QTEC). San Diego, USA. Summer 2010.

Qing Cao, Debessay Fesehaye, Nam Pham, Yusuf Sarwar, and Tarek Abdelzaher. *Virtual Battery: An Energy Reserve Abstraction for Embedded Sensor Networks*. 29th IEEE Real-Time Systems Symposium (RTSS 2008). Barcelona, Spain. November 2008.

(d) *TCP/IP (Internet Traffic) Modeling and Analysis*:

Debessay Fesehaye and Sabine Wittevrongel. *Convergence of the fixed point algorithm of analytical models of reliable Internet protocols (TCP)*. International Conference on Computational Science (ICCS 2006), Springer Lecture Notes in Computer Science, University of Reading, UK. May 2006.

Debessay Fesehaye and Sabine Wittevrongel. *An analytical model of TCP Performance*. 25th IEEE International Performance Computing and Communications Conference (IPCCC 2006). Phoenix, Arizona, USA. April 2006.

Debessay Fesehaye. *Emulating TCP (a Reliable Internet Protocol) using the fixed point algorithm*. 31st IEEE Conference on Local Computer Networks (LCN). Tampa, Florida, USA. November 2006.

Debessay Fesehaye and A. E. Krzesinski. *A queuing network model of TCP performance*. Proceedings of the South African Institute for Computer Scientists and Information Technologists (SAIC-SIT), ACM International Conference Proceedings Series. White River, South Africa.

#### **Patents:**

US Patent Granted (US 9094326 B2): *SYSTEMS AND METHODS FOR COMMUNICATING IN A NETWORK*. By Qualcomm. Co-Inventors: Jay Kumar Sundararajan, Gerardo Giarretta, David William Craig, Julien H. Laganier, Gavin Bernard Horn. November 2011.  
(Available at <http://www.google.com/patents/US9094326>).

US Patent Filed (Application No. US 14/494,761): *METHOD AND APPARATUS AND COMPUTER READABLE MEDIUM FOR COMPUTING STRING SIMILARITY METRIC*. By Neustar. Inventor: Debessay Fesehaye. January 2014.

US Patent Filed (Application No. US 15/877,977): *Group Clustering Using Inter-Group Dissimilarities*. By VMware. Co-Inventors: Lenin Singaravelu, Xiaobo Huang, Amitabha Banerjee, Ruijin Zhou. January 2018.

#### **Open Source Software:**

Open Source Project On String Similarity: *New String Similarity Metrics and Efficient Algorithms to Calculate Them*. By Neustar Inc. Author: Debessay (Debish) Fesehaye. October 2014.  
(Available at <https://github.com/neustar/string-similarity>).

#### **Published Papers (Extended Abstracts):**

Debessay Fesehaye and Klara Nahrstedt. *SCDA: SLA-aware cloud datacenter architecture for efficient content storage and retrieval*. ACM HPDC 2013. NY City, NY, USA. June 2013.

Debessay Fesehaye, Rahul Malik and Klara Nahrstedt. *EDFS: a semi-centralized Efficient Distributed File System*. IFIP/USENIX Middleware 2009. Champaign, IL, USA. December 2009.

Debessay Fesehaye. *Simple Online Cross-layer Traffic Engineering Techniques*. 21st IEEE Annual Computer Communications Workshop. Pittsburgh, PA, USA. February 2007.

Debessay Fesehaye. *A simple, cost-effective and energy-efficient (C3E) location finding algorithm*

for mobile networks. 4th EuroFGI Workshop on New Trends in Modelling, Quantitative Methods and Measurements. Gent, Belgium. May 2007.

Debessay Fesehaye. *A Fast Congestion control Protocol (FCP) for networks (the Internet)*. 4th IEEE International Conference on Information Technology: Research and Education. Tel Aviv, Israel. October 2006. (Longer and Latest Version Available as UIUC Technical Report at <https://ideals.illinois.edu/bitstream/handle/2142/35905/qcpUpdated.pdf>)

Debessay Fesehaye. *An efficient queue management (EQM) technique for networks*. 6th IEEE International Workshop on IP Operations and Management. Springer Lecture Notes in Computer Science. Dublin, Ireland. October 2006.

Debessay Fesehaye. *A Simple Network (Internet) Control Protocol (SNCP)*. Networks: Computation, Communication and Applications (NCCA'06) Workshop. University of Oxford, UK. Sept 2006.

Debessay Fesehaye. *A video QoS analysis using a fixed point algorithm*. 3rd EuroNGI Workshop on New Trends in Modeling, Quantitative Methods and Measurements. Politecnico di Torino, Torino, Italy. June 2006.

### **Technical Reports (With initial versions of research works to be published):**

Debessay Fesehaye, Pengye Xia, P. Brighten Godfrey, Klara Nahrstedt and Steven S. Lumetta. *QCP: Finishing Flows Faster with A Quick congestion Control Protocol (QCP)*. (Available as UIUC Technical Report at <https://ideals.illinois.edu/bitstream/handle/2142/35905/qcpUpdated.pdf>), January 2013. Shorter version published as FCP at ITRE '06 (IEEE conference), Tel-Aviv, Israel.

Debessay Fesehaye, Shameem Ahmed, Thadpong Pongthawornkamol, Klara Nahrstedt and Guijun Wang. *Reliability Trade-off Analysis of Deadline-Sensitive Wireless Messaging Systems*. Department of Computer Science, UIUC, Technical Report. March 2011. (To be Submitted).

Debessay Fesehaye. *Dynamic Mapping of an AS Network into A Smaller Network of Border Routers*. Department of Computer Science, UIUC, Tech Report at <https://www.ideals.illinois.edu/bitstream/handle/2142/17442/debishNetmap.pdf>. Dec 2010.

Debessay Fesehaye and Klara Nahrstedt. *BRTP: Border Routing & Transport Protocol*. Department of Computer Science, UIUC, Technical Report. August 2010.

Debessay Fesehaye and Klara Nahrstedt. *PAS: A Packet Accounting System to Limit the Effects of DoS & DDoS*. Department of Computer Science, UIUC, Technical Report. August 2010.

Debessay Fesehaye, Klara Nahrstedt and Matthew Caesar. *A Network Congestion control Protocol (NCP)*. Department of Computer Science, UIUC, Technical Report. July 2010.

Debessay Fesehaye, Rahul Malik and Klara Nahrstedt. *A Scalable Distributed File System for Cloud Computing*. Department of Computer Science, UIUC, Technical Report. March 2010.

Debessay Fesehaye, Indranil Gupta and Klara Nahrstedt. *A Cross-layer Routing and Congestion Control for Distributed Systems*. Department of Computer Science, UIUC, Technical Report at <https://www.ideals.illinois.edu/handle/2142/10861>. November 2008.

## **TEACHING ASSISTANTSHIP EXPERIENCE**

---

*Data Structures and Programming Principles (CS 225):*

Department of Computer Science, University of Illinois-Urbana Champaign, Fall 2009-Spring 2010.

*Advanced Internetworking (CS 598):*

Department of Computer Science, University of Illinois-Urbana Champaign, Fall 2008.

*Systems Programming (CS 241):*

Department of Computer Science, University of Illinois-Urbana Champaign, Fall 2007-Spring 2008.

*Calculus and Linear Algebra I, II and Analysis and Linear Algebra I, II:*

Department of Mathematics, University of Stellenbosch, Spring and Fall 2005.

## AWARDS

---

### Full Sponsorship

Belgium:

Full sponsorship by *EuroNGI Network of Excellence* to attend the following courses:

Security in IP Networking and Satellite, on board processing, dynamic accesses, at INT Evry, France.

*Performance Engineering and Queueing Network Models (QNM)s* at University of Bradford, UK.

*A summer school in NGI Network Strategy, Design and Dimensioning*, at Santander & Laredo, Cantabria (Spain), Spring-Fall 2006.

*Research Grant* by Ghent University, Belgium, January 2006 to January 2010 ( I declined it in the Fall 2007).

SA (Western Cape):

*Full research sponsorship* by Telkom SA in the University of Stellenbosch Unit of the Telkom-Siemens Centre of Excellence in ATM and Broadband Networks and their Applications.

Israel:

*Full sponsorship* by the Israeli Ministry of Foreign affairs to take a one-month (162hrs) training on educational technology in Jerusalem, Israel.

### Merit bursaries

University of Stellenbosch Merit Bursary (for 2 years).

## COMPUTER PROGRAMMING

---

The *C++/C*, Java, Pascal, Visual Basic, Hive/SQL, Apache Spark, Spark Mlib, Weka, Mahout, Hadoop, Map Reduce, Hadoop Pipes, Apache Http Server, MySQL, MangoDB, PhP, Perl, AWK, Bash *Script* (Linux) programming & TCL/oTCL Scripting, *NS2* (an object-oriented, discrete event driven Network Simulator developed at UC Berkeley written in C++ and OTcl).

## SOME COURSES TAKEN

---

Advanced Distributed Systems, Advanced Operating Systems, Advanced Multimedia Systems, Computer Security II, Distributed Sensor and Cyber-Physical Systems, Advanced Topics in Wireless Networks, Operating Systems Design, Computer Networks, Cryptography and Data Security, Object Databases, Discrete Event Simulation.



Probability Models and Stochastic Simulation I, Probability Models and Stochastic Simulation II, Methods of Operations Research, Mathematical Programming, Graph Theory, Bayesian Statistics (Audited and all assignments completed), Time Series Analysis.

Linear Algebra II, Ordinary Differential Equations, Sampling Survey, Functions of Complex Variables, Introduction to Numerical Methods, Design of Experiments, Statistical Theory of Distributions, Transformation Geometry, Introduction to Probability, Statistical Methods I and II, Theory of Numbers, Modern Algebra, Advanced Calculus I, Differential and Integral Calculus II.

## **REFERENCES**

---

Available on request.