# **CURRICULUM VITAE - YANG HU**

#### **CONTACT INFORMATION**

Ph.D. candidate Phone: (352) 281-3364 huyang.ece@ufl.edu URL: http://plaza.ufl.edu/huyang.ece

### **Research Interests**

#327 Benton Hall University of Florida Gainesville, FL 32611

Architecture of server systems, networking infrastructures, and data centers.
Virtualized platform for networking, computing, and storage; Cloud and IoT platforms.
System design for availability, security, sustainability; Power, renewable energy-powered system.
Heterogeneous system design for <b>big data</b> , <u>network security</u> , and <u>deep learning</u> .

# EDUCATION

University of Florida Ph.D. in Electrical and Computer Engineering	Aug. 2011 – Aug. 2017 (expected)
Advisor: Tao Li	
Tsinghua University, China	Sep. 2008 – Jun. 2011
M.S. in Microelectronic Engineering Advisor: Shouyi Yin	
Tioniin University China	Son 2002 Jul 2007
B.S. in Electronic and Information Engineering	3ep. 2005 – Jul. 2007
AWARDS AND HONORS	

Best Paper Nomination of HPCA 2017, ranked No.1 in 231 submissions	2017
Travel Grant for 23rd IEEE Symp. on High Performance Computer Architecture (HPCA)	2017
Travel Grant for 49th Int'l Symp. on Micro Architecture (MICRO)	2016
Travel Grant for 42th Int'l Symp. on Computer Architecture (ISCA)	2015
Best of IEEE Computer Architecture Letters Award	2015
Travel Grant for 46th Int'l Symp. on Micro Architecture (MICRO)	2013
Travel Grant for 39th Int'l Symp. on Computer Architecture (ISCA)	2012
UF Alumni Fellowship, the highest graduate student awards available at UF	2011~2015

# ACADEMIC TALKS

Towards "Full Containerization" in Container Based Network Function Virtualization	ASPLOS, 2017
Towards Efficient Server Architecture for Virtualized Network Function Deployment	MICRO, 2016
HOPE: Enabling Efficient Service Orchestration in Software-Defined Data Centers	<i>ICS</i> , 2016
Towards Sustainable in-Situ Server Systems in the Big Data Era	<i>ISCA</i> , 2015

#### **INVITED TALKS**

Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the University of Houston, Mar, 2017 Communication Path Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the **Communication Path** University of Central Florida, Mar, 2017 Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the University of Louisiana at Lafavette, Mar, 2017 **Communication Path** Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the San Francisco State University, Mar, 2017 Communication Path Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the **Communication Path** Utah State University, Mar, 2017 Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the **Communication Path** The University of Texas at Dallas, Mar, 2017 Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the University of Pittsburgh, Mar, 2017 **Communication Path** Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the **Communication Path** Florida State University, Feb, 2017 Towards Efficient Architecture Design in IoT Big Data Era: Integrating Intelligence into the **Communication Path** University of North Texas, Feb, 2017 Towards Efficient Server Architecture for Virtualized Network Function Deployment Capital Normal University, China, 2016 Towards Emerging Data Center Architecture in IoT Era

Tsinghua University, China, 2016

#### FUNDING RAISING

My research work has led to proposals that have been awarded following grants. I also contributed to the write-ups of these proposals.

NSF, CCF-1527535, Optimizing Consolidation Efficiency of Emerging Virtualized Cloud Applications on<br/>Contemporary Server Architecture (Highly Competitive, \$460k)2015

NSF, CNS-1423090, Leveraging Distributed Generation and Adaptive Energy Storage Management forEfficient and Scale-out Renewable Powered Data Center Design (Highly Competitive, \$450k)2014

NSF, CNS-1117261, Enabling Renewable Energy Powered Sustainable High Performance Computer Architectures and System (Highly Competitive, \$200k) 2011

#### **PROFESSIONAL SERVICES**

Program Committee Member:	The 4th Workshop on Efficient Data Center Systems (EDCS), 2016
	The 15th IEEE International Conference on Pervasive, Intelligence and
	Computing (PI-Com), 2017
Invited Journal Reviewer:	IEEE Transactions on Parallel and Distributed Systems (TPDS), IEEE
	Computer Architecture Letters (CAL)

Conference Reviewer: ISCA, HPCA, MICRO, ASPLOS, ICS, IPDPS, ICDCS, ICPP, DAC

Workshop Reviewer: ASBD, CoolDC

#### **SELECTED PUBLICATIONS**

# Journals:

- TPDSChao Li, Yang Hu, Juncheng Gu, Jingling Yuan, and Tao Li. "Oasis: Scaling Out Datacenter Sustainably<br/>and Economically". IEEE Transactions on Parallel and Distributed Systems (TPDS), 2016, IEEE
- CAL Longjun Liu, Chao Li, Hongbin Sun, Yang Hu, Jingmin Xin, Nanning Zheng, and Tao Li. "<u>Leveraging</u> <u>Heterogeneous Power for Improving Datacenter Efficiency and Resiliency</u>". IEEE Computer Architecture Letters (CAL), 2015 (Best of Computer Architecture Letters)
- JSC Longjun Liu, Hongbin Sun, Chao Li, **Yang Hu**, Jingmin Xin, Nanning Zheng, and Tao Li. "RE-UPS: An Adaptive Distributed Energy Storage System for Dynamically Managing Solar Energy in Green Datacenters", The Journal of Supercomputing (JSC), Volume 72, Issue 1, pp. 295-316, 2016. Springer
- JCST Chao Li, Rui Wang, **Yang Hu**, Ruijin Zhou, Ming Liu, Longjun Liu, Jingling Yuan, Tao Li, and Depei Qian. "Towards Automated Provisioning and Emergency Handling in Renewable Energy Powered Datacenters", Journal of Computer Science and Technology (JCST), Vol. 29, No. 4, pp. 618-630, 2014. Springer
- IEICEShouyi Yin, Yang Hu, Zhen Zhang, and Shaojun Wei. "Hybrid Wired/Wireless On-Chip Network<br/>Design for Application-Specific SoC", IEICE transactions on electronics 95 (4), 495-505, 2012

#### **Conferences:**

- SC '17 Siyang Li, Jiwu Shu, Youyou Lu, **Yang Hu**, and Tao Li. "LocoFS: A Loosely-Coupled Metadata Service for Distributed File System". The International Conference for High Performance Computing, Networking, Storage and Analysis (**SC** 2017), To appear.
- ISPASS '17 Huixiang Chen, Yang Hu, Meng Wang, Mingcong Song, and Tao Li. "Characterizing GaaS Workloads Under Virtualized GPU Environment". ISPASS 2017
- ASPLOS '17 Yang Hu and Tao Li. "<u>Towards "Full Containerization</u>" in Container Based Network Function <u>Virtualization</u>". ASPLOS 2017. (Acceptance Rate: 53/306 = 17%)
  - HPCA '17 Mingcong Song, Yang Hu, Huixiang Chen, and Tao Li. "<u>Towards Pervasive and User Satisfactory CNN</u> across GPU Microarchitectures". HPCA 2017. (Best Paper Nomination, ranked No.1 in 231 submissions)
- MICRO '16 Yang Hu, Mingcong Song, and Tao Li. "<u>Towards Efficient Server Architecture for Virtualized Network</u> <u>Function Deployment: Implications and Implementations</u>". Proc. of the 49th International Symposium on Microarchitecture (**MICRO**), 2016. (Acceptance Rate: 61/288 = 21.2%)
  - PACT '16 Yang Hu\*, Mingcong Song\*, and Tao Li. "<u>Bridging the Semantic Gaps of GPU Acceleration for Scale-out CNN-based Big Data Processing: Think Big, See Small</u>". Proc. of the 25th International Conference on Parallel Architectures and Compilation Techniques (**PACT**), 2016 (Authors with equal contribution, acceptance Rate: 31/139 = 22.3%)
    - ICS '16 Yang Hu, Chao Li, Longjun Liu, and Tao Li. "<u>HOPE: Enabling Efficient Service Orchestration in</u> <u>Software-Defined Data Centers</u>". Proc. the 30th International Conference on Supercomputing (ICS), Jun. 2016 (Acceptance Rate: 43/178 = 24.2%)
    - ICS '16 Longjun Liu, Chao Li, Hongbin Sun, Yang Hu, Tao Li, Jingmin Xin and Nanning Zheng.. "Towards an

<u>Adaptive Multi-Power-Source Datacenter</u>". Proc. the 30th International Conference on Supercomputing (**ICS**), Jun. 2016 (Acceptance Rate: 43/178 = 24.2%)

- ISCA '15 Yang Hu\*, Chao Li\*, Longjun Liu\*, Juncheng Gu, Mingcong Song, Jingling Yuan, Xiaoyao Liang, and Tao Li. "<u>Towards Sustainable in-Situ Server Systems in the Big Data Era</u>". Proc. the 42nd ACM/IEEE International Symposium on Computer Architecture (ISCA), Jun. 2015 (Authors with equal contribution) (Acceptance Rate: 58/305 = 19%)
- ISCA '15 Longjun Liu\*, Chao Li\*, Hongbin Sun, Yang Hu, Juncheng Gu, Tao Li, Jingmin Xin, and Nanning Zheng. "HEB: Deploying and Managing Hybrid Energy Buffers for Improving Datacenter Efficiency and Economy". Proc. the 42nd ACM/IEEE International Symposium on Computer Architecture (ISCA), Jun. 2015 (Authors with equal contribution) (Acceptance Rate: 58/305 = 19%)
- DSN '15 Longjun Liu, Chao Li, Hongbin Sun, **Yang Hu**, Juncheng Gu, and Tao Li. "<u>BAAT: Towards Dynamically</u> <u>Managing Battery Aging in Green Datacenters</u>". Proceedings of the 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (**DSN**), Jun. 2015 (Acceptance Rate: 50/227 = 22%)
- MICRO '13 Chao Li, Yang Hu, Ruijin Zhou, Ming Liu, Longjun Liu, and Tao Li. "<u>Enabling Datacenter Servers to</u> <u>Scale Out Economically and Sustainably</u>". Proc. of the 46th Intl. Symposium on Microarchitecture MICRO 2013. (Acceptance Rate: 39/269 = 14.5%)
- APCCAS '10 **Yang Hu**, Shouyi Yin, and Shaojun Wei. "<u>Mixed-level modeling for network on chip infrastructure in</u> <u>SoC design</u>", IEEE Asia Pacific Conference on Circuits and Systems (**APCCAS**), 2010

# PATENTS

Chinese Patent, 201510553330, Pending, 2015 Chinese Patent, 201510556261, Pending, 2015 Chinese Patent, 201510558202, Pending, 2015

# TEACHING

EEL 6935 Billion-Transistor Computer Architecture, Spring 2015, University of Florida EEL 6935 Billion-Transistor Computer Architecture, Spring 2014, University of Florida

# Interested Courses to Teach:

- Undergraduate level: computer architectures, computer networks, signals and systems, digital signal processing, communication systems, algorithms, and computer programming.
- Graduate level: advanced computer architectures, computer networks, cloud computing, big data analysis, IoT/cyber-physical systems, emerging network security.

# PRIMARY RESEARCH PROJECTS AND EXPERIENCE

- <u>Optimizing the NFV Deployment on Heterogeneous Server Platform</u> Goal: Explore the challenges of deploying network functions on heterogeneous COTS server and propose an adaptive runtime framework that is tailored for network packet processing applications.
- <u>Optimizing the Networking Performance on Containerized NFV Platform</u> Goal: In this project we propose NetContainer, a framework that achieves a fine-grained hardware resource management for containerized NFV platform.
- <u>Optimizing the Virtual Switch Performance on NUMA based NFV Platform</u> Goal: Explore the performance bottlenecks of NUMA based COTS server in modern NFV deployment. Optimize the thread scheduling of Open vSwitch.

- <u>Orchestrate the Management Workloads in Software-Defined Cloud Data Center</u> Goal: Design a cloud management system to coordinate the power consumption of management workloads in modern software-defined data center running big data workloads.
- <u>In-Situ Big Data Processing Server System for IoT Applications</u> Goal: Design server clusters for IoT big data processing in the field.

### LIST OF REFERENCES

Tao Li (advisor) Professor, Electrical and Computer Engineering Department, University of Florida Program Director, CCF/CISE, National Science Foundation send.Professor.93E4882A49@interfolio.com (352) 392-9510

### Yuan Xie

Professor, Department of Electrical and Computer Engineering, University of California at Santa Barbara IEEE Fellow send.Professor.E49E57DC03@interfolio.com (805) 893-5563

#### Sean Meyn

Professor and Robert C. Pittman Eminent Scholar Chair, Electrical and Computer Engineering Department, University of Florida IEEE Fellow send.Meyn.5A91D4D5A2@interfolio.com (352) 392-8934

#### APPENDIX

\*(Credited to Prof. Tor Aamodt's CV): "For those not familiar with the field of computer architecture, publication in top international conferences is the *preferred* means of disseminating research results in this area. Papers in proceedings of top conferences, often with acceptance rates below 20%, are more important than journals. See Chapter 4 in the 1994 National Academy of Sciences report, Academic Careers for Experimental Computer Scientists and Engineers http://books.nap.edu/html/acesc/ and the 1999 Best Practices Memo: Evaluating Computer Scientists and Engineers For Promotion and Tenure http://www.cra.org/uploads/documents/resources/bpmemos/tenure\_review.pdf. The top computer architecture conferences are ISCA, MICRO, HPCA, ASPLOS and PACT. The review process at these conferences is very rigorous: Program committee (PC) members are internationally recognized experts in the field. Each paper typically receives four or more double-blind reviews (3 or more from PC members) providing detailed feedback. Authors submit responses to questions raised by reviewers prior to the PC meeting, which is attended in-person by PC members. Accepted papers are revised to reflect

reviewer feedback before publication. High quality papers deemed to be on the borderline of accept/reject often undergo an additional round of review by a PC member after revision and before final acceptance for publication (a process known as "shepherding"). Added together, the five conferences above publish only 150 to 200 papers per year. Due to their impact, these conferences are very well attended. Given their importance, papers published in the proceedings of these conferences are read and cited by active researchers in the area even if those researchers did not attend the conference."