DMITRII USTIUGOV

About

Dmitrii Ustiugov is a PhD student at the University of Edinburgh, co-advised by Prof. Boris Grot (University of Edinburgh) and Prof. Edouard Bugnion (EPFL). Dmitrii's research interests span across Computer Architecture and Computer Systems with a focus on serverless and cloud architectures.

Contact info

10 Crichton Street, Edinburgh, UK, EH8 9AB
Informatics Forum, office 2.05

Website: https://homepages.inf.ed.ac.uk/s1373190
E-mail: dmitrii.ustiugov@ed.ac.uk

Research interests

Broad: Cloud computing, virtual memory, datacenter and rack-scale systems, emerging memory systems, security.
Current focus: Serverless clouds, including benchmarking, virtualization, and high-speed communication.

Education

Ph.D., Computer Science (09.2014 – ongoing) at **EPFL** (09.2014-08.2019) and the **University of Edinburgh** (from 09.2019).
Advisors: **Prof. Boris Grot** (University of Edinburgh) and **Prof. Edouard Bugnion** (EPFL).

- Thesis: "Data-centric serverless cloud architecture".

M.Sc.&B.Sc. (with Honors), **Applied Math and Physics** (09.2008-08.2014) at **MIPT** (Moscow Institute of Physics and Technology)
Advisor: **Dr. Alexander Butuzov** (Intel/MIPT).

- B.Sc. thesis: "CPU performance analysis using critical path methods".
- M.Sc. thesis: "CPU power consumption analysis based on cycle-accurate microarchitecture simulation".

Awards and Fellowships

- Distinguished Artifact Award at ASPLOS'21.
- The winner of 1-minute research pitch competition at the JOBS workshop co-located with MICRO'20.
- Arm Center of Excellence fellowship at the University of Edinburgh, 2019.
- EPFL PhD Fellowship, 2014.

Current Research Projects

- **vHive**: Open-source framework & ecosystem for serverless experimentation, in collaboration with ETH, AWS, Microsoft Research, Arm, Huawei, and the vHive open-source community (users & contributors at 14+ universities and 5 companies). https://github.com/ease-lab/vhive
- **STeLLAR**: Tail latency analyzer framework for commercial serverless clouds. https://github.com/ease-lab/STeLLAR

Past projects

- Architectural support for address translation in virtualized clouds, https://github.com/ease-lab/PTEMagnet
- Bankrupt covert communication channel, http://github.com/ease-lab/bankrupt
- VISA: Vertically integrated server architecture, https://parsa.epfl.ch/visa
- Scale-out NUMA, https://parsa.epfl.ch/sonuma/sonuma.html
- CloudSuite: A benchmark suite for cloud services, https://cloudsuite.ch
- Google Perfkit Benchmarker, http://www.perfkitbenchmarker.org/
- QFlex: Quick and flexible computer architecture simulation, https://qflex.epfl.ch

Refereed Conference Publications

1. **D. Ustiugov**, S. Jesalpura, M.B. Alper, M. Baczun, R. Feyzkhanov, E. Bugnion, B. Grot, and M. Kogias, "Expedited Data Transfers for Serverless Clouds". *Under submission to ASPLOS'22*.
2. **D. Ustiugov**, T. Amariucai, and B. Grot "Analyzing Tail Latency in Serverless Clouds with STeLLAR". *Conditionally accepted (with shepherding) to IISWC'21*.

3. **D. Ustiugov**, P. Petrov, M. Kogias, E. Bugnion, and B. Grot, "Benchmarking, Analysis, and Optimization of Serverless Function Snapshots". *ASPLOS'21*. **Distinguished Artifact Award.**
4. A. Margaritov, **D. Ustiugov**, and B. Grot, "PTEMagnet: Fine-Grained Physical Memory Reservation for Faster Page Walks in Public Clouds". *ASPLOS'21*.
5. A. Margaritov, **D. Ustiugov**, E. Bugnion, and B. Grot, "Prefetched Address Translation". *MICRO'19*.
6. **D. Ustiugov**, A. Daglis, J. Picorel, M. Sutherland, E. Bugnion, B. Falsafi and D. Pnevmatikatos, "Design Guidelines for High-Performance SCM Hierarchies", *MEMSYS'18.*
7. M. Drumond, A. Daglis, N. Mirzadeh, **D. Ustiugov**, J. Picorel, B. Falsafi, B. Grot, and D. Pnevmatikatos, "The Mondrian Data Engine", *ISCA'17.*
8. A. Daglis, **D. Ustiugov**, S. Novakovic, E. Bugnion, B. Falsafi, and B. Grot, "SABRes: Atomic Object Reads for In-Memory Rack-Scale Computing", *MICRO'16*.

Refereed Workshop Publications

1. **D. Ustiugov**, P. Petrov, M.R.S. Katebzadeh, and B. Grot, "Bankrupt Covert Channel: Turning Network Predictability into Vulnerability". *Workshop on Offensive Technologies (WOOT) at USENIX Security, 2020*.
2. A. Margaritov, **D. Ustiugov**, E. Bugnion, and B. Grot, "Virtual Address Translation via Learned Page Table Indexes". *In the Workshop on Machine Learning for Systems (MLSys) workshop at NeurIPS, 2018.*

Refereed Journal Publications

1. S. Novakovic, A. Daglis, **D. Ustiugov**, E. Bugnion, B. Falsafi, and B. Grot, "Mitigating Load Imbalance in Distributed Data Serving Through Rack-Scale Memory Pooling", *TOCS, 2019.*
2. M. Drumond, A. Daglis, N. Mirzadeh, **D. Ustiugov**, J. Picorel, B. Falsafi, B. Grot, and D. Pnevmatikatos, "Algorithm/Architecture Co-Design for Near-Memory Processing", *SIGOPS Operating Systems Review, 2018.*

Other Publications

- B. Ziv, G. Haber, L. Rumyantsev, **D. Ustiugov**, "Use Cases for the Critical Path Analyzer Framework". *Software Professionals Conference (SWPC)*, 2013.
- A. Butuzov, O. Shimko, **D. Ustiugov**, "Fast and Easy Ways to Improve Software Development Teamwork Efficiency". *Software Professionals Conference (SWPC)*, 2013.
- K. Garifullin, **D. Ustiugov**, "Performance Simulation and Early Power Modeling of New HW/SW Co-Designed Architecture". *Software Professionals Conference (SWPC)*, 2012.
- N. Kosarev, O. Shimko, **D. Ustiugov**, "Critical path study tool for performance analysis of modern architectures". *Software Professionals Conference (SWPC)*, 2011.
- **D. Ustiugov**, "Survey of the state-of-the-art methods for dynamical power estimation and analysis conducted during early stages of CPU microarchitecture development". *MIPT conf., 2012*.
- **D. Ustiugov**, "CPU power estimation method using software cycle-accurate simulator". *MIPT conf., 2011*.
- **D. Ustiugov**, "Practical mobile banking with multi-factor authentication". *MIPT conf., 2011*.
- **D. Ustiugov**, "High-performance instruction cache for multithreaded architecture". *MIPT conf., 2010*.

Research and Invited Talks

Invited talks and lectures

- "vHive Framework and Ecosystem for Serverless Experimentation", an **invited talk** at Intel (Processor Architecture Research Lab), Aug 2021.
- "Turbocharging Serverless Research with vHive", an **invited talk** at the Workshop on Cloud-Native Future Innovation, Huawei, Jul 2021.
- "Turbocharging Serverless Research with vHive", an **invited talk** at ETH Zurich's Systems Group, May 2021.
- "Cloud Computing: Evolution, Technologies, Future", **invited lecture** at the Operating Systems course at the University of Edinburgh, Mar 2021.
- "Benchmarking, Analysis, and Optimization of Serverless Function Snapshots", an **invited talk** at Amazon Web Services (Amazon Lambda and Firecracker teams), Feb 2021.
- "Prefetched Address Translation", an **invited talk** at the 5th Computing Systems Day at NTUA, Athens, 2020.
- "Towards High-Performance SCM Hierarchies", an **invited talk** at Oracle, 2017.

Research talks

- "Benchmarking, Analysis, and Optimization of Serverless Function Snapshots", ASPLOS'21.

- "Bankrupt Covert Channel: Turning Network Predictability into Vulnerability", WOOT'20 co-located with USENIX Security.
- "Design Guidelines for High-Performance SCM Hierarchies", MEMSYS'18.
- "Hardware Support for Remote Atomic Reads in Rack-Scale Systems", the EuroSys Doctoral Workshop (EuroDW) co-located with EuroSys, 2016.

Community Service and Professional Activity

- EuroSys'21 Shadow TPC member.
- External reviewer for ISCA 2019, ATC 2019, and Computer Architecture Letters (CAL) 2019 and 2020.
- Student ACM member since 2015, student USENIX member since 2020.

Contribution to Grant Proposals Writing

- Proposal funded by **Huawei** ($250,000): Efficient Serverless Applications via Communication-aware Function Composition. PI: Boris Grot, 2021.
- Proposal funded by **Google** ($73,000): Accelerating Address Translation via a Learned Page Table Index. PI: Boris Grot, 2019.
- Proposal funded by **Oracle Labs** ($90,000): Scalable Memory Server Architecture for Disaggregated Memory Systems. PI: Virendra Marathe, Co-PI: Babak Falsafi, 2018.

Professional Experience

Research intern at **Oracle Labs**, Apr – Jun 2017.
- Mentors: Dr. Virendra Marathe and Stephen Byan.
- Analysis of persistent memory applications and disaggregated memory systems.

Senior Lua/Perl developer at **IPONWEB**, Dec 2013 – Aug 2014.
- Manager: Lev Leontiev.
- Developing a high-load big data software platform for real-time bidding.

Software simulation intern (full-time engineer since Jan 2012) at **Intel**, Jun 2010 – Nov 2013.
- Mentors: Dr. Alexander Butuzov, Vladimir Gnatyuk.
- CPU performance/power/energy studies using software simulators and analyzer tools.

Teaching Assistant

- Introduction to Computer Systems, Uni of Edinburgh, Fall 2019, 2020.
- Extreme Computing, Uni of Edinburgh, Fall 2019.
- Mathematical analysis (MAN), EPFL, Spring 2018.
- Computer Architecture I, EPFL, Fall 2016, 2017.
- Probability and Statistics, EPFL, Spring 2016.
- Introduction to Multiprocessor Architecture, EPFL, Fall 2015.

Student Research Projects Supervision

Shyam Jesalpura, intern, Uni of Edinburgh / BITS, Jan-Aug 2021.
- High-speed communication fabric for serverless clouds.

Mert Bora Alper, intern, Uni of Edinburgh, Jun-Aug 2021.
- Benchmarking methodology for serverless clouds.

Michal Baczun, intern, Uni of Edinburgh, Jun-Aug 2021.
- Representative suite of serverless workloads.

Yuchen Niu, BSc, Uni of Edinburgh, Sep 2020 – Apr 2021.
- Implications of multi-tenancy on serverless hosts.

Theodor Amariucai, BSc, Uni of Edinburgh, Sep 2020 – Jul 2021.
- Tail-latency analysis framework for serverless clouds.

Plamen Petrov, BSc, Uni of Edinburgh, Jun 2019 – Nov 2020.
- End-to-end serverless benchmarking framework.
- RDMA networks security and covert communication.

Ivy Wang, BSc, Uni of Edinburgh, Sep 2019 – Apr 2020.
- Software support for contiguous page tables allocation.

Sean Mullan, BSc, Uni of Edinburgh, Sep 2019 – Apr 2020.
- Design space exploration for TLB prefetching.

Lei Yan, MSc, EPFL/RWTH Aachen University, Jan – Aug 2018.
- Design space exploration for user-level cooperative scheduling for latency-critical cloud services.

Siddharth Gupta, PhD candidate, EPFL, Sep – Dec 2017.
- Analysis of persistent memory systems on modern CPUs.

Nikhil Gupta, intern, EPFL/IIT Delhi, May – Aug 2016.
- Building a robust infrastructure for QFlex (Flexus) simulation framework.

Virgile Neu, BSc, EPFL, Jan – May 2016.
- Analyzing CPU front-end efficiency using perf counters.