Program

15th Cloud Control Workshop

Sandhamn, Sweden
June 11–13, 2019

Practical information

Hotel contact:
Sandhamn Seglarhotell AB
Box 124, 130 39 Sandhamn, Sweden
e-mail: reception@sandhamn.com
Telephone: +46-8-574 504 00

Hotel reception opening hours:
The reception has 24 hour service.

Check-in and -out
Most rooms will be available for check-in on arrival, but definite details will be provided on site. Check-out until 10.00.

Payment
Rooms, meals, and some drinks are covered by the workshop registration fee. For any additional costs put on your room, like telephone, additional drinks, etc, please ensure to pay at check-out.

Luggage room
A luggage room is made available on arrival.

Relax facilities and other leisure activities
The hotel relax facilities are open 7.00 to 20.00.
There are nice 3 km, 5 km, and 8 km trails for walking or running. Maps are available at the reception.
There is also a map for a trail associated to the TV series The Sandhamn Murders.
Canoes and bicycles are available for rental just next to the hotel, but from a separate provider.

Bar
Seglarbaren is open until 24:00

Breakfast
Breakfast is served from 7.00.

Coffee breaks
Coffee during coffee breaks is served in Seglarsalen.

Meals
All meals are served in the main dining room.

Internet
WiFi throughout the premises. Password is Almabar1

Weather
As Sandhamn is located rather far out in the archipelago, expect it to be slightly chillier than, e.g., in the main Stockholm area.

In case everything else fails
Erik's phone number is +46 70 315 3928. Anticipate limited ability to answer around bus departure on June 11.

Tuesday June 11th

8:45 Gathering at bus outside Arlanda airport, Terminal 4
9:00 Bus departure
Ferry departure from Vaxholm slightly before 10.00. If joining by the ferry, please be there no later than 9.45.
Coffee is served on the ferry

11:30 **Lunch**

Session Chair: Karl-Erik Årzén, Lund University, Sweden

12:30 *Workshop Introduction*
Erik Elmroth, Umeå University, Sweden

12:45 **Keynote: *Measuring and Optimizing Tail Latency***
Kathryn S McKinley, Google Cloud, USA

13:30 *Towards Specialized Edge Computing: Challenges, Benefits, and Applications*
Prashant Shenoy, University of Massachusetts, USA

13:50 *Are AI-enabled Platforms the Future of Cloud Operations?*
Odej Kao, TU Berlin, Germany

14:10

Discussion 1 (Utkiken)	Discussion 2 (Sandö)	Discussion 3 (Alma)	Discussion 4 (Kungen)	Discussion 5 (Klubben)
Datacenters without Borders, Mazin Yousif, DXC Technology, USA	*Application of serverless to 5G networks,* David Breitgand, IBM Haifa, Israel	*Multi-tenancy in distributed edge infrastructures,* Thijs Metsch, Intel, Ireland	*On common abstractions between application and resource management systems,* Christian Pérez, INRIA, France	*Distributed Resource Control across the Cloud and Edge,* Ming Zhao, Arizona State University, USA

15:30 **Coffee**

Session Chair: Joe Butler, Intel Labs, Ireland

16.00 *Modeling and Analysis of Performance Interaction and Tradeoff in the Cloud*
Y.C. Tay, National University of Singapore, Singapore

16.20 *Scheduling Parallel Programs under Power Constraints: Algorithmic Advances and Practical Concerns*
Henry Hoffman, University of Chicago, USA

16.40 ***Break***

17:00 ***Poster Reception with Refreshments*** (Outdoor pool area)

1. *Make the Most out of Last Level Cache in Intel Processors.* Amir Roozbeh, Ericsson Research/KTH, Sweden
2. *Decentralized Collective Learning for Self-managed Sharing Economies.* Evangelos Pournaras, ETH Zurich, Zurich, Switzerland
3. *Methodological Principles for Reproducible Performance Evaluation in Cloud Computing.* Nikolas Herbst, University of Würzburg, Germany, Alexandru Iosup, VU Amsterdam, the Netherlands, and Ahmed Ali-Eldin, UMASS
4. *Safe and Efficient Remote Application Code Execution on Disaggregated NVM Storage with eBPF.* Kornilios Kourtis, IBM Research, Zurich, Animesh Trivedi, VU Amsterdam, and Nikolas Ioannou, IBM Research, Zurich
5. *Statistical Inference for the Self-Adaptive Cloud.* Johan Ruuskanen & Anton Cervin, Lund University
6. *Chameleon: A Hybrid, Proactive Auto-Scaling Mechanism on a Level-Playing Field.* Nikolas Herbst, University of Würzburg, Germany & Ahmed Ali-Eldin, UMASS
7. *Telescope: An Automatic Feature Extraction and Transformation Approach for Seasonal Time Series/Workload Forecasting on a Level-Playing Field.* Nikolas Herbst, University of Würzburg, Germany
8. *MicroRCA: Root cause analysis system for Container-based Micro-service Environment.* Li Wu, Elastisys & TU Berlin
9. *Network Slicing in Massive MIMO Radio Access Network.* Haorui Peng, Lund University, Sweden
10. *ASA - Adaptive Scheduling Algorithm for Scientific Workflows.* Abel Souza, Umeå University
11. *Equivalent G/G/1 Modeling for Server Systems with Redundant Requests.* Tommi Nylander, Lund University
12. *Toward Elasticity Control for Edge Data Centers.* Chanh Nguyen, Umeå University
13. *Cost Optimization for Containerized Applications in the Cloud.* Mulugeta Ayalew Tamiru, Elastisys and University of Rennes 1
14. *Democratizing Data Analytics: Crowd-sourcing Decentralized Collective Measurements.* Evangelos Pournaras, ETH Zurich
15. *HoloScale -- Combining horizontal and vertical scaling.* Victor Millnert, Lund University
16. *Fruit Fly-Based Simulated Annealing Optimization Scheme for Resource Scheduling in Mobile Edge Clouds.* Danlami Gabi, Monica Vitali, Umeå University, Sweden
17. *Serverless VNF Orchestration in 5G Networks: lessons from 5G-MEDIA.* David Breitgand, Avi Weit, Stamatia Rizou, David Griffin, Ugur Acar, Gino Carrozzo, Nikolaos Zioulis, Pasquale Andriani, Francesco Iadanza, IBM Research – Haifa, Israel
18. *EdgeDroid: An Experimental Approach to Benchmarking Human-in-the-Loop Applications.* Manuel Olguín Muñoz, KTH Royal Institute of Technology
19. *KnowledgeNet: Disaggregated and Distributed Training and Serving of Deep Neural Networks.* Saman Biookaghazadeh, Yitao Chen, Kaiqi Zhao, Ming Zhao, Arizona State University

19:00 **Dinner**

Wednesday June 12th

7:00	**Breakfast**				
	Session Chair: Andrew Chien University of Chicago, USA				
8:15	*Announcements*				
8:20	*The Cloud Is Not Enough: Enabling the Edge with Cryptographically Hardened Data Capsules* John Kubiatowicz, UC Berkeley, USA				
8:40	*Managing Resources in End-to-End Machine Learning Pipelines* Jim Dowling, Royal Institute of Technology, Sweden				
9:00	*Edge Computing: Challenges and Opportunities* Eyal De Lara, University of Toronto, Canada				
9:20	*Intelligent automation at scale for Cloud Services* Mazin Yousif, DXC Technology, USA				
9:40	*Fractal: Automated Application Scaling* Richard Mortier, Cambridge University, UK				
10:00	**Coffee**				
	Session Chair: Maria Kihl, Lund University, Sweden				
10:30	*Automatic Diversification of SaaS and PaaS* Benoit Baudry, Royal Institute of Technology, Sweden				
10:50	*Who Will Write Those Industry Production Papers?* Indranil Gupta, University of Illinois, Urbana-Champaign, USA				
11:10	Discussion 6 (Utkiken) *From the Edge to the Cloud and back again,* John Kubiatowicz, UC Berkeley, USA	Discussion 7 (Sandö) *Feedback loops: what for, where, how, how many ... ?,* Eric Rutten, INRIA, France	Discussion 8 (Alma) *Serverless Platforms:Resource Management and Benchmarking,* Animesh Trivedi and Alexandru Iosup, VU Amsterdam & Nikolas Herbst, University of Würzburg, Germany	Discussion 9 (Kungen) *Can We Build a Cloud in Which Users Own Their Data?,* Anton Burtsev, University of California, Irvine, USA	Discussion 10 (Klubben) *Beyond the Edge: A New Market Paradigm for Mobile Resource Providers,* Danlami Gabi & Monica Vitali, Umeå University, Sweden
12:30	**Lunch**				
	Session Chair: Jose A Fortes, University of Florida, USA				
13:45	*Google Flex* Steve Webster, Google, USA				
14:05	*Burstable Instances: Models, Equilibria, and Revenue* Carlee Joe-Wong, Carnegie Mellon University, USA				
14.25	*Where blockchains fail (and why cloud computing is of no help)* Maarten van Steen, University of Twente, The Netherlands				
14.45	Discussion 11 (Utkiken) *Experimentation on cloud,* Lucas Nussbaum, Université de Lorraine/Inria, France & Maarten van Steen, University of Twente, The Netherlands	Discussion 12 (Sandö) *Systems for Machine Learning,* Jim Dowling & Seif Haridi, KTH, Sweden	Discussion 13 (Alma) *Challenges of a high throughput scheduler in a real world environment,* Gonzalo Rodrigo, Apple, USA	Discussion14(Kungen) *Service meshes for cloud control,* Cristian Klein & Johan Tordsson, Umeå University and Elastisys, Sweden	Discussion 15 (Klubben) *Distributed data in edge,* Mina Sedaghat & Ola Angelsmark, Ericsson, Sweden
16:00	**Coffee & Group Photo**				
16:30	Social outdoor activities				
19:00	**Dinner**				

Thursday June 13th

7:00	**Breakfast**
	Session Chair: Azimeh Sefidcon, Ericsson, Sweden
8:15	*Announcements*
8:20	*SplitServe: combining lambdas and VMs for effective overload management of stateful, multi-stage, parallel processing frameworks* Bhuvan Urgaonkar, The Pennsylvania State University, USA
8:40	*The emergence of hybrid clouds* Eran Raichstein, IBM Research, Israel
9:00	*Network-aware energy-efficient virtual machine management in distributed Cloud infrastructures with on-site photovoltaic production* Anne-Cécile Orgerie, CNRS - IRISA, France
9:20	*Dynamic cloud storage management* Khuzaima Daudjee, University of Waterloo, Canada
9:40	*Multi-Period Portfolio Optimization for "Cloud" Scheduling* Ahmed Ali-Eldin, Umeå University/UMASS, Sweden/USA
10:00	**Coffee**
	Session Chair: Guillaume Pierre, Université de Rennes 1, France
10.30	*Design of a "Zero-carbon Cloud" Scientific Computing Facility: Economics, Operation, and Environmental Impacts* Andrew A. Chien, University of Chicago, USA
10:50	*Towards Software-defined Monitoring* Patrick Eugster, University of Lugano, Switzerland

11:10	Discussion 16 (Utkiken)	Discussion 17 (Sandö)	Discussion 18 (Alma)	Discussion19(Kungen)	Discussion 20 (Klubben)
	Reducing the Environmental Impact of Cloud Computing, Andrew A. Chien, University of Chicago, USA	*Challenges with mobile edge,* Björn Skubic & Morgan Lindqvist, Ericsson, Sweden	*Anomaly detection in edge/fog computing,* Odej Kao, TU Berlin, Germany & Lilly Wu, Elastisys, Sweden	*Breaking the limits of Kubernetes cluster federation for 5G telco clouds,* Lars Larsson, Umeå University, Sweden	*Title* Name, affiliation, country

12:30	*Closing* Erik Elmroth, Umeå University, Sweden
12.45	**Lunch**

13:45 **Ferry departure from Sandhamn**

Buses depart from Vaxholm approximately 15:15

16:30 **Bus arrival at Arlanda airport, Terminal 4**

Presentations, Posters, and Discussion Sessions in order of Appearance

13th Cloud Control Workshop

June 13 – 15, 2018

Tuesday, June 11th

12:30 *Workshop Introduction*
Erik Elmroth, Umeå University, Sweden

14:45 **Keynote:** ***Measuring and Optimizing Tail Latency***
Kathryn S McKinley, Google Cloud, USA

Interactive web services require careful engineering to optimize tail latencies (99th percentile or higher) or users abandon the service. Optimizing these systems is challenging because tail latency is hard to measure and because non-deterministic root causes span networking, software, and hardware. We describe Shim, a new tool and application profiling methodology, that reads out hardware and software signals at 1000 cycle granularities without instrumenting applications at 1% overheads, which is orders-of-magnitude more frequent than interrupt based sampling. We also describe Google's kernel trace scheduling framework. These tools divide tail latency into long requests and non-deterministic events, inspiring different optimizations and challenges.

After technical questions, I will show how myths, such as the sole genius, are counter productive to science and innovation. In contrast, social science shows diverse teams achieve better outcomes in controlled and uncontrolled studies. I will put these results in the context of my personal research experiences. I challenge you to build diverse research teams and mentor the next generation, so that your team will produce more innovative work

13:30 *Towards Specialized Edge Computing: Challenges, Benefits, and Applications*
Prashant Shenoy, University of Massachusetts, USA

In this talk, I will argue that the era of general-purpose computing is rapidly evolving into one of special-purpose computing due to technological advances that allow for inexpensive hardware devices and accelerators to optimize specific classes of application workloads. Edge computing has not been immune to these trends, and it is now feasible to specialize edge deployments for workloads such as machine learning analytics, AI on the edge, speech, and augmented reality using low-cost specialized hardware. I will discuss the implications of these technology trends of the future mobile and IOT-based edge applications and present new challenges that will need to be addressed to fully exploit these trends.

13:50 *Are AI-enabled Platforms the Future of Cloud Operations?*
Odej Kao, TU Berlin, Germany

The softwareization and virtualization trend, the increasing number of IoT applications with dynamically linked devices and the embedding in real-world (smart) environments drive the creation of large multilayered systems. Consequently, the complexity of the systems is steadily growing up to a level, where it is impossible for human operators to oversee and holistically manage the systems without additional support and automation. Uninterrupted services with guaranteed latencies, response times, and other QoS parameters, are however mandatory prerequisite for many of the data-driven and autonomous applications deployed these days. Therefore, losing control is not a feasible option for any system or infrastructure, as downtimes and malfunctioning components may have a crucial impact beyond financial and image damage. On the other hand, the software-defined technology on all layers of the infrastructure stack opens new possibilities to control not only the server landscape, but also the connected frontend devices and the communication paths. This optimization potential can be used to increase the dependability and the reliability of the overall system. The next piece in the puzzle aims at decreasing the reaction time, in case an urgent activity of a system administrator – for example due to performance problems (e.g. tuning), to component/system failures (e.g. outages, degraded performance), or due to security incidents – is necessary. All these examples describe situations, where the system operates outside of the normal (expected or pre-defined) parameters. Thus, the system exposes an anomaly that must be registered, recognized, and remediated, before it leads to a component or a system failure. Much effort was spent in recent years on developing and deploying advanced mechanisms for monitoring the systems, aggregating and analyzing the data, detecting the root cause of the anomaly, and finally remediating it, often referred to as artificial intelligence for IT operations (AIOps). These AIOps methods and the reported preliminary results promise a new era in the IT system

administration with obvious benefits in performance, availability, and stability. But they come with the cost and risks due to the increased danger of losing control over the system. The case of an AI-supported administration failing while attempting the recovery of anomalous system may even create a more complex challenge for the human administrator to stabilize the system, as the AI-supported operations need to be recognized and rolled back prior to solving the original problem. The key question thus refers to AI-Governance and to finding an acceptable balance of autonomous decision-making and operation for the sake of performance increase vs. robustness of the system in terms that a human operator understands the current situation and is always capable of controlling the system.

14:10 Discussion 1 (Utkiken)

Datacenters without Borders, Mazin Yousif, DXC Technology, USA

The complexity around applications in an enterprise is growing exponentially making it difficult to establish full visibility all the apps an enterprise owns. Such complexity arises from a long transitional period where an enterprise may have applications running on-premise; another set of applications running on various public cloud providers such as Amazon AWS and Microsoft Azure; yet another set of applications running in private cloud providers' datacenters; possibly co-loco and another set in Hybrid clouds…. Further, it is expected that applications may move across various cloud and on-premise deployments. You get the "complexity" picture and the "Datacenter without Borders" concept. So the key question is how can a CIO have a full handle at their applications landscape. This is what we will be discussing here.

14:10 Discussion 2 (Sandö)

Application of serverless to 5G networks, David Breitgand, IBM Haifa, Israel

Serverless programming takes software development and operation by a storm. The two big promises of the new paradigm are: (1) lower barriers for software engineers and (2) improved cost-efficiency for irregular workloads that might require instantaneous elasticity.

5G networks are evolving rapidly as a new ubiquitous cellular network infrastructure that seamlessly converges with traditional communication networks. Although universal adoption of the 5G networks is not expected before 2025, this year (2019) we start witnessing first commercial 5G pilots by some major telcos with practically every telco company ranking 5G very high in its priorities. An inherent feature of the 5G network architecture is softwarization with VNFs running in the data centers embedded with the infrastructure and organized hierarchically.

In the last few years, as virtualization advanced, telcos underwent a fundamental transformation replacing many PNFs by VNFs. The VNF model of today is based heavily on Virtual Machines (VM) with the control plane technologies, such as OpenStack to manage them. The VM model underpins the ETSI MANO standards and guides the inner workings of the VNFI orchestrators, such as OSM. However, another revolution is coming, the cloud native one. This new transformation is powered by the container technology, new container-oriented control planes, such Kubernetes, and architectural principles collectively termed "cloud native" (although this name does not necessarily imply using cloud). In the brave new world envisioned by the proponents of the cloud-native telco transformation, VNFs might be containers running on Kubernetes.

In this discussion session, we will consider the following questions:

- What a cloud native VNF orchestration should look like? Is MANO adequate or even needed in the future?
- Are there advantages in running third party software inside the providers' 5G data centers alongside their core VNFs?
- If the answer to the previous question is yes, then should this third-party software be packaged, deployed, and orchestrated as ETSI MANO compatible VNFs or should these be serverless functions to match the public cloud vendors offerings and allow low entry barriers for developers?
- Are there natural use cases for FaaS related to VNF orchestration?
- What gaps exist in the current serverless frameworks that should be closed before they can serve as an infrastructure for VNFs?
- Can Kubernetes be our next generation VNF orchestrator?
- What are the research challenges in this new technological cycle combining VNFs, FaaS, and 5G networks?

We will briefly introduce our experience and insights accrued in 5G-MEDIA H2020 project (http://www.5gmedia.eu/), which pioneered application of serverless to VNF orchestration for media-intensive applications and then proceed to the discussion points above.

14:10 Discussion 3 (Alma)

Multi-tenancy in distributed edge infrastructures, Thijs Metsch, Intel, Ireland

During this discussion we want to explore the challenges, use cases and approaches to deal with multi-tenant distributed compute environment at the edge. Current deployments of edge-infrastructures show limit amount of compute capacity being available. The need to maximize ROI for the infrastructure owner, while customers of edge services might expose

dynamic usage patterns demand that – similarly to trends previously seen in cloud computing and NFV – the edge infrastructure is shared among multiple tenants. The distributed nature and dynamic usage of edge resources furthermore brings unseen challenges – e.g. the movement of data belonging to a tenant to the appropriate infrastructure resources on-demand. In general multi-tenancy brings a set of challenges – for example but not limited to: security & trust, fair-sharing of resources, storage and sharing of data, …. For those and others we want to look at use-cases and - if existing - approaches to tackle those.

14:10 Discussion 4 (Kungen)

On common abstractions between application and resource management systems, Christian Pérez, INRIA, France

RMS provide ressources based on some abstractions: for example, it can VMs for Cloud, a job for HPC/batch systems, or just processes for classical OS. While the job description slowly evolves from rigid job (fixed number of ressources), to some forms of moldability and malleability (usually a variable number of ressources), they are rather limited to express needs for more advanced applications. In the cloud world, several languages, such as TOSCA, have been proposed to describe the structure of the applications. They enable to manage ressources at a finer grain but they are still limited to some application structures. This session is about the relationships between applications and RMS with respect to the description of the structure of an application and the usage that can be done to efficiently manage the ressource (user and system concerns, static and dynamic) and in particular its impact on scheduling issues.

14:10 Discussion 5 (Klubben)

Distributed Resource Control across the Cloud and Edge, Ming Zhao, Arizona State University, USA

Smart devices (smart cameras, phones, cars, etc.) are becoming pervasive in our society. Many such devices nowadays have not only sensing abilities and network connectivity but also non-trivial computational power and storage capacity, which allow them to provide computing and data services at the edge of the network with unique capabilities such as situation awareness and real-time responsiveness. Simultaneously, traditional cloud computing datacenters also continue to scale up, providing the capacities for intensive computations and massive data storage cost-effectively. These two computing paradigms can be potentially combined to take advantage of their complementary strengths and support many important applications, but how to optimally control the resources distributed across edge and cloud is a challenging problem to developing an effective edge+cloud computing solution.

16.00 *Modeling and Analysis of Performance Interaction and Tradeoff in the Cloud*

Y.C. Tay, National University of Singapore, Singapore

Multiple forces (e.g. data contention, resource contention and precedence constraints) and performance tradeoffs (e.g. energy vs latency) affect job execution in a datacenter. This talk points out some techniques (e.g. model decomposition and bottleneck analysis) that can facilitate the analysis and control of job performance.

16:20 *Scheduling Parallel Programs under Power Constraints: Algorithmic Advances and Practical Concerns*

Henry Hoffman, University of Chicago, USA

Scheduling parallel programs for minimal execution time has been a widely studied problem for decades. Researchers have looked at the problem of scheduling dependent programs (expressed as directed acyclic graphs, or DAGs) and the problem of scheduling independent applications given a shared power budget. While each of these problems is NP-hard, they have simple, greedy approximations that are just a small constant factor from optimal. Interestingly, however, applying greedy algorithms to the combined problem of scheduling a DAG with a power constraint can differ from optimal by a factor of O(N). We recently (in collaboration with Demirci and Kim) proposed a divide and conquer algorithm that brings the approximation factor down to O(log N). In this talk we briefly give an overview of greedy approaches (upon which all implemented schedulers are based), highlight the divide and conquer algorithm, and talk about the practical challenges that would need to be overcome to implement the divide and conquer algorithm on modern distributed systems.

17:00 ***Poster Reception with Refreshments*** (Outdoor pool area)

Poster 1: Make the Most out of Last Level Cache in Intel Processors.

Amir Roozbeh, Ericsson Research/KTH, Sweden

In modern (Intel) processors, Last Level Cache (LLC) is divided into multiple slices and an undocumented hashing algorithm (aka Complex Addressing) maps different parts of memory address space among these slices to increase the effective memory bandwidth. After a careful study of Intel's Complex Addressing, we introduce a slice-aware memory management scheme, wherein frequently used data can be accessed faster via the LLC. Using our proposed scheme, we show that a key-value store can potentially improve its average performance ~12.2% and ~11.4% for 100% & 95% GET workloads, respectively. Furthermore, we propose CacheDirector, a network I/O solution which extends Direct Data I/O (DDIO) and places the packet's header in the slice of the LLC that is closest to the relevant processing core. We implemented CacheDirector as an extension to DPDK and evaluated our proposed solution for latency-critical

applications in Network Function Virtualization (NFV) systems. Evaluation results show that CacheDirector makes packet processing faster by reducing tail latencies (90-99th percentiles) by up to 119 µs (~21.5%) for optimized NFV service chains that are running at 100 Gbps. Finally, we analyze the effectiveness of slice-aware memory management to realize cache isolation.

Poster 2: Decentralized Collective Learning for Self-managed Sharing Economies.
Evangelos Pournaras, ETH Zurich, Zurich, Switzerland

EPOS, the Economic Planning and Optimized Selections (http://epos-net.org), is a fully decentralized networked system designed for participatory multi-objective optimization forming a public good and supporting sharing economies. It performs collective decision-making among agents that autonomously generate a set of options from which they make a choice. Each agent is a human actor, a piece of software or a hybrid system of both that locally generates in a self-determined way a set of plan that define how some resources are allocated. For example, a plan may define the energy demand of a residential appliance in a future horizon or the availability of bicycles in the bicycle stations of a city. A set of several plans per agent represents alternative options, equivalent or not for the agent. EPOS is capable of steering such highly complex systems of combinatorial complexity to desirable outcomes by structuring agent interactions in dynamic self-organized tree topologies and performing bottom-up collective decision-making using fitness functions designed to solve particular problems. for instance, preventing blackouts in smart grids by load-shifting or load-adjustment. Given its decentralization, scalability, local autonomy and collective decision-making, it can promote participation, fairness and sustainability in the sharing economies and application domains of energy, transportation, voting, Smart Cities and others.

Poster 3: Methodological Principles for Reproducible Performance Evaluation in Cloud Computing.
Nikolas Herbst, University of Würzburg, Germany, Alexandru Iosup, VU Amsterdam, the Netherlands, and Ahmed Ali-Eldin, UMASS

The rapid adoption and the diversification of cloud computing technology exacerbate the importance of a sound experimental methodology for this domain. This work investigates how to measure and report performance in the cloud, and how well the cloud research community is already doing it. We propose a set of eight important methodological principles that combine best-practices from nearby fields with concepts applicable only to clouds, and with new ideas about the time-accuracy trade-off. We show how these principles are applicable using a auto-scaling as a practical use-case experiment. To this end, we analyze the ability of the newly released SPEC Cloud IaaS 2018 benchmark to follow the principles, and showcase real-world experimental studies in common cloud environments that meet the principles. Last, we report on a systematic literature review including top conferences and journals in the field, from 2012 to 2017, analyzing if the practice of reporting cloud performance measurements follows the proposed eight principles. Worryingly, this systematic survey and the subsequent two-round human reviews, reveal that few of the published studies follow the eight experimental principles. We conclude that, although these important principles are simple and basic, the cloud community is yet to adopt them broadly to deliver sound measurement of cloud environments. (further contributors from SPEC Research Cloud: Alessandro Vittorio Papadopoulos, Laurens Versluis, André Bauer, Jóakim von Kistowski, Cristina Abad, J. Nelson Amaral, and Petr Tuma - PDF)

Poster 4: Safe and Efficient Remote Application Code Execution on Disaggregated NVM Storage with eBPF.
Kornilios Kourtis, IBM Research, Zurich, Animesh Trivedi, VU Amsterdam, and Nikolas Ioannou, IBM Research, Zurich

With rapid improvements in NVM storage devices, the performance bottleneck is gradually shifting to the network, thus giving rise to the notion of "data movement wall". To reduce the amount of data movement over the network, researchers have proposed near-data computing by shipping operations and compute-extensions closer to storage devices. However, running arbitrary, user-provided extensions in a shared, disaggregated storage environment presents multiple challenges regarding safety, isolation, and performance. Instead of approaching this problem from scratch, in this work we make a case for leveraging the Linux kernel eBPF framework to program disaggregated NVM storage devices. eBPF offers a safe, verifiable, and high-performance way of executing untrusted, user-defined code in a shared runtime. In this work, we describe our initial experiences towards building a first prototype that supports remote operations on storage using eBPF, discuss the limitations of our approach, and directions for addressing them. The work is in initial stages where we seek feedback regarding our approach, potential applications, and counter proposals.

Poster 5: Statistical Inference for the Self-Adaptive Cloud.
Johan Ruuskanen & Anton Cervin, Lund University

Successful self-adaptive resource provisioning in the cloud relies on accurate tracking and prediction of workload variations and timely detection of changes in the infrastructure. However, the estimation problems in this context becomes challenging due to the massive number of measurements, heterogeneous behavior of tasks and time changing resource requirements. On this poster we will present some early and ongoing work in tackling these issues by event-based particle filtering for server state estimation and Gaussian processes for resource usage prediction.

Poster 6: Chameleon: A Hybrid, Proactive Auto-Scaling Mechanism on a Level-Playing Field.
Nikolas Herbst, University of Würzburg, Germany & Ahmed Ali-Eldin, UMASS

Auto-scalers for clouds promise stable service quality at low costs when facing changing workload intensity. The major public cloud providers provide trigger-based auto-scalers based on thresholds. However, trigger-based auto-scaling has reaction times in the order of minutes. Novel auto-scalers from literature try to overcome the limitations of reactive mechanisms by employing proactive prediction methods. However, the adoption of proactive auto-scalers in production is still very low due to the high risk of relying on a single proactive method.
This work tackles the challenge of reducing this risk by proposing a new hybrid auto-scaling mechanism, called Chameleon, combining multiple different proactive methods coupled with a reactive fallback mechanism. Chameleon employs on-demand, automated time series-based forecasting methods to predict the arriving load intensity in combination with run-time service demand estimation to calculate the required resource consumption per work unit without the need for application instrumentation.
We benchmark Chameleon against five different state-of-the-art proactive and reactive auto-scalers one in three different private and public cloud environments. We generate five different representative workloads each taken from different real-world system traces. Overall, Chameleon achieves the best scaling behavior based on user and elasticity performance metrics, analyzing the results from 400 hours aggregated experiment time.
(IEEE Transactions on Parallel and Distributed Systems, 30(4):800 -- 813, April 2019, IEEE. [DOI])

Poster 7: Telescope: An Automatic Feature Extraction and Transformation Approach for Seasonal Time Series/Workload Forecasting on a Level-Playing Field.
Nikolas Herbst, University of Würzburg, Germany

One central problem of machine learning is the inherent limitation to predict only what has been learned --- stationarity. Any time series property that eludes stationarity poses a challenge for the proper model building. Furthermore, existing forecasting methods lack in reliable forecast accuracy and time-to-result if not applied in their sweet-spots.
In this ongoing work, we propose a fully automated machine learning-based forecasting approach. Our Telescope approach extracts and transforms features from a seasonal input time series and uses them to generate a tailored forecast model.
In a broad competition including latest hybrid forecasters (e.g., M4 competition winner in 2018 from Uber and Prophet from Facebook), established statistical and machine learning-based methods, our approach shows the best forecasting accuracy coupled with a lower and reliable time-to-result.
Furthermore, we showcase that Telescope can significantly improve auto-scaling accuracy and timeliness as it provides more accurate and timely workload predictions to auto-scaler decision engines compared to, e.g., ARIMA, FFT- or ANN-based forecasters.
R package/sources available at http://descartes.tools/telescope)

Poster 8: MicroRCA: Root cause analysis system for Container-based Micro-service Environment.
Li Wu, Elastisys & TU Berlin

Cloud applications have recently shifted from monolithic architecture to micro-service architecture rapidly. However, the complexity of dependency and the dynamic environment pose critical challenges to identify the root causes. To tackle these challenges, we propose MicroRCA, a novel root cause analysis system dedicated for container-based micro-service environment, to localize the anomalous micro-services and identify the root causes in the system. Without instrumenting the source code of micro-services, MicroRCA constructs a multi-layer propagation graph(MPG) and combines the anomaly detection results to pinpoint the possible cause components in the system.

Poster 9: Network Slicing in Massive MIMO Radio Access Network.
Haorui Peng, Lund University, Sweden

Network slicing is one of the efficient solutions to the critical requirements that 5G is addressing. By the definition of 3GPP, a network slice is a collection of network functions and resources that are formed to meet certain network characteristics required by a specific service instance. It generally comes with Software-Defined Network (SDN) and Network Function Virtualisation (NFV) to support its functionality and has been mostly developed in the Core Network (CN). However when considering the Radio Access Network (RAN), the network slicing problem is more challenging due to the limited spectrum and the interferences of radio channels. This presented work will focus on the approaches of wireless visualisation and radio resource allocation problems to enable network slicing in a single-cell RAN, which is based on Massive MIMO, one of the key physical layer candidate techniques for 5G.

Poster 10: ASA - Adaptive Scheduling Algorithm for Scientific Workflows.
Abel Souza, Umeå University

In constrained environments such as High Performance Computing (HPC) infrastructures, resources may not be readily

available for use, making job management challenging since constraints intensify with limited resources. Scientific workflows are used to model, describe, and steer the process of how a scientific campaign develops. New demands regarding the amounts of data being generated from such large scale experiments, planning and allocating resource requests for processing all information on time gets even more complex. In one hand, allocating the max capacity expected for a scientific workflow guarantees execution on time, at the cost of spare and idle resources, which impacts total resource utilization. On the other hand, allocating resources according to on-time and workflow stage demands optimize resource utilization, but negatively impacts the total workflow runtime.
In this paper, we propose ASA: an Adaptive Scheduling Algorithm for scientific workflows. Packaged as a system's library, ASA aims at estimating queue waiting times enabling scientific workflows to optimize job submission planning. We study a novel statistical method for pro-actively placing job requests to minimize waiting times, idle resources, and to reduce workflow's total runtime. This amounts to both learning (the waiting times), and acting on what is learnt thus far, and amounts hence to a realization of the exploration-exploitation trade-off. Experiments with real workflows in real supercomputers show that ASA achieves a middle ground between the two aforementioned ways for resource allocation: faster total runtime, with near optimal resource utilization.

Poster 11: Equivalent G/G/1 Modeling for Server Systems with Redundant Requests.
Tommi Nylander, Lund University

We present a model that allows us to equivalently represent a system of servers with cloned requests, as a single server. The model is very general, and we show that no assumptions on either inter-arrival or service time distributions are required, allowing for, e.g., both heterogeneity and dependencies. Additionally, we show that the model holds for any queuing discipline as well. The key requirement that enables us to use the single server G/G/1 model is that the request clones have to receive synchronized service. We show examples of server systems that fulfill this requirement, and use our G/G/1 model to co-design traditional load-balancing algorithms together with cloning strategies, providing well performing, provably stable designs. Finally, we use our model to evaluate request cloning in heterogeneous server systems.

Poster 12: Toward Elasticity Control for Edge Data Centers.
Chanh Nguyen, Umeå University

Elasticity is one of fundamental properties that MECs must inherently hold in order to become a mature computing platform hosting software applications. However, unlike the current cloud platform, MECs need to cope with more challenges from the densed distribution, the heterogeneity and limitation of resource capacity in Edge Data Centers (EDCs), and the end-user mobility. In this poster, we present our proposed auto scaling technique to help MECs overcome the aforementioned challenges to automatically scale its resource in a proactive manner. The technique utilizes the location information of EDCs and the correlation of workload changes of those EDCs in the proximity to first estimate request arrival rate at EDCs, then the two level controllers (i.e., a local controller, and a neighbor-cluster controller) decide the right amount of EDC's resource to be scaled up/down which ensure persistently the Quality of Service (QoS), while maximizing resource utilization. The experimental result based on the simulated workload(using the real mobility traces) shows that our proposed proactive auto-scaling helps MECs to improving the efficiency of resource provisioning (i.e., reducing the rejected rate, increasing the resource utilization) while maintaining the system stability (preventing the resource oscillation).

Poster 13: Cost Optimization for Containerized Applications in the Cloud.
Mulugeta Ayalew Tamiru, Elastisys and University of Rennes 1

Modern applications are designed and deployed in the form of microservices, which are modular self-contained components. In the last few years, containers have gained popularity as a natural way to package and deploy microservices. At the same time, container orchestration platforms -- for example Kubernetes -- have been proposed for automatically managing these containerized microservices across their entire lifecycles. However, default configurations of the major container orchestration platforms are not optimized for performance, cost or availability. Moreover, meeting these requirements is complex and cannot be met manually with reasonable effort in a reasonable amount of time. The dynamic nature of applications, workloads, user behaviour and mobility make the problem even worse. Therefore, automatic ways of meeting application requirements are required. In this work we propose to address the problem of cost optimization for containerized applications by optimizing application placement (scheduling), re-scheduling, dynamic resource allocation (autoscaling) and dynamically selecting the right priced and right sized VMs from the infrastructure pool. We evaluate the performance of our proposed solution using real-world workloads.

Poster 14: Democratizing Data Analytics: Crowd-sourcing Decentralized Collective Measurements.
Evangelos Pournaras, ETH Zurich

This poster introduces a Technology Readiness Level (TRL) 6 live demonstrator for socially responsible real-time data analytics: crowd-sourced decentralized collective measurements by preserving privacy. It provides a proof of concept for the democratization of big data running by citizens, for citizens. The demonstrator connects DIAS, the Dynamic

Intelligent Aggregation Service, and GDELT, the Global Database of Events, Language, and Tone, to monitor planetary activity in different countries. DIAS remains operational for more than 4 months, ex- changing millions of messages to self-adapt to data updates from GDELT. It accurately estimates the total news events providing high responsiveness to mobile devices. The demonstrator comes with a community Slack Bot for status updates of system health.

Poster 15: HoloScale -- Combining horizontal and vertical scaling.
Victor Millnert, Lund University

Holo originates from the Greek word “olos” meaning “whole, entire, complete” as well as “safe and sound”. This is typically not the case when scaling cloud resources. Normally, one has to choose between one of two options: a) vertical scaling, or b) horizontal scaling. In vertical scaling it is possible to quickly and continuously scale existing processing capacity by adding more resources to a machine. The downside, however, is that the operating range is very limited. The other option, horizontal scaling, is where one adds processing capacity by adding more instances which are identical. The benefit of this is that one can scale the processing capacity over a very large range. The downside, however, is that it is usually very slow and with a very coarse control.
In HoloScale we combine both horizontal and vertical scaling in a provably safe way so that we achieve both rapid, continuous scaling (by leveraging vertical scaling) over a large spectrum (by leveraging horizontal scaling). We create a model for this system and solve the control problem by borrowing ideas and concepts from classical mid-ranging control. Finally, we evaluate our proposed solution using real traffic, and show that the system is capable of quickly scaling the computation resources as well as to react tox sudden disturbances, such as virtual machines going down.

Poster 16: Fruit Fly-Based Simulated Annealing Optimization Scheme for Resource Scheduling in Mobile Edge Clouds.
Danlami Gabi, Monica Vitali, Umeå University, Sweden

In Mobile Edge Clouds (MECs), achieving sustainable profit advantage, cost reduction and resource utilization is always a bottleneck when trying to meet the computing needs of resource hungry applications. More concern is to how edge and cloud resource providers can coexist to improve resource allocation in MECs, since each provider has different service quality and might set a different cost for it resource usage. Recent research shows metaheuristic techniques are used to allocate resources to large scale applications in MECs. However, some challenges attributed to the metaheuristic techniques includes entrapment at the local optimal caused by premature convergence and imbalance between the local and global search. These may affect resource allocation in MECs if continually implemented. To address these concerns and ensure efficient resource allocation in MECs, we proposed a Fruit Fly-based Simulated Annealing Optimization Scheme (FSAOS) as an ideal solution. The FSAOS is a combination of Fruit Fly Optimization Algorithm (FOA) and Simulated Annealing (SA). The SA is incorporated into the local search procedure of the FOA to balance between the global and local search and to overcome its premature convergence. Implementation of the FSAOS is carried out on EdgeCloudSim Simulator tool. Simulation results show that the FSAOS can return minimum makespan, execution costs and better resource utilization compared to conventional FOA and PSO schemes. To further unveil how efficient the FSAOS, a statistical analysis based on 95% confidence interval is carried out. Numerical results show FSAOS has outperformed the benchmarked schemes by achieving minimum confidential level. This is an indication that the proposed FSAOS can provide efficient resource allocation in MECs while meeting customers aspirations as well as that of the resource providers

Poster 17: Serverless VNF Orchestration in 5G Networks: lessons from 5G-MEDIA. David Breitgand, Avi Weit, Stamatia Rizou, David Griffin, Ugur Acar, Gino Carrozzo, Nikolaos Zioulis, Pasquale Andriani, Francesco Iadanza, IBM Research – Haifa, Israel

5G networks open up new opportunities for telcos to compete on providing cloud services. While current standards and solutions in NFV are oriented towards Virtual Machines and IaaS technologies, such as OpenStack, a container revolution is looming, and we expect it to take traction with the advance of 5G networks within the next two-three years. Kubernetes becomes a de-facto standard for container orchestration management and we expect it to be adopted by the telcos for the 5G network DCs that are organic parts of the converged network infrastructure. The future of NFV is in adopting the cloud native approach. Arguably, one of the more influential cloud native technologies that enjoys exponential growth is serverless. In this poster we highlight our experience with applying serverless to VNF orchestration for media intensive applications in 5G networks. We examine the pros and cons of this approach, analyze new requirements and illustrate our results by demanding use cases, such as tele-immersive gaming. We argue that serverless can be an extremely useful tool for VNF orchestration if applied properly and outline future research directions.

Poster 18: EdgeDroid: An Experimental Approach to Benchmarking Human-in-the-Loop Applications.
Manuel Olguín Muñoz, KTH Royal Institute of Technology

Many emerging mobile applications, including augmented reality (AR) and wearable cognitive assistance, aim to provide seamless user interaction. However, the complexity of benchmarking these human-in-the-loop applications limits reproducibility and makes performance evaluation difficult. In this poster, we present EdgeDroid, a benchmarking

suite designed to reproducibly evaluate these applications.
Our core idea rests on recording traces of user interaction, which are then replayed at benchmarking time in a controlled fashion based on an underlying model of human behavior. This allows for an automated system that greatly simplifies benchmarking large scale scenarios and stress testing the application. Our results show the benefits of EdgeDroid as a tool for both system designers and application developers

Poster 19: KnowledgeNet: Disaggregated and Distributed Training and Serving of Deep Neural Networks.
Saman Biookaghazadeh, Yitao Chen, Kaiqi Zhao, Ming Zhao, Arizona State University

Deep Neural Networks (DNNs) have a significant impact on numerous applications, such as video processing, virtual/augmented reality, and text processing. The everchanging environment forces the DNN models to evolve, accordingly. Also, the transition from the cloud-only to edgecloud paradigm has made the deployment and training of these models challenging. Addressing these challenges requires new methods and systems for continuous training and distribution of these models in a heterogeneous environment. In this paper, we propose KnowledgeNet (KN), which is a new architectural technique for a simple disaggregation and distribution of the neural networks for both training and serving. Using KN, DNNs can be partitioned into multiple small blocks and be deployed on a distributed set of computational nodes. Also, KN utilizes the knowledge transfer technique to provide small scale models with high accuracy in edge scenarios with limited resources. Preliminary results show that our new method can ensure a state-of-the-art accuracy for a DNN model while being disaggregated among multiple workers. Also, by using knowledge transfer technique, we can compress the model by 62% for deployment, while maintaining the same accuracy.

Wednesday June 12th

8:20 *The Cloud Is Not Enough: Enabling the Edge with Cryptographically Hardened Data Capsules*
John Kubiatowicz, UC Berkeley, USA

Exciting new applications rely on information from widely disparate sources, combining processing and storage in the cloud with embedded sensors, actuators, and processing at the edge (or "Fog"). Unfortunately, the edge of the network consists of islands of trusted hardware interspersed with a vast sea of untrusted components. In fact, privacy, authenticity, and security weaknesses seem to emerge almost daily as application writers "roll their own" infrastructure from scratch. What is needed is something completely different. In this talk, I will describe a standardization effort to refactor the network around cryptographically-hardened bundles of data, called DataCapsules. DataCapsules are the cyberspace equivalent of ubiquitous shipping containers: uniquely named, secured bundles of information transported over a data-centric "narrow-waist" infrastructure called the Global Data Plane (GDP). The GDP partitions the network into Trust Domains (TDs) to allow clients to reason about the trustworthiness of hardware. When combined with trusted computing enclaves, the GDP enables applications to place their persistent information into DataCapsules and dynamically partition their functionality between the cloud and network edge. Applications benefit from new classes of performance optimizations, yielding better QoS, lower latency, and enhanced privacy without risking information integrity or obscuring its provenance.

8:40 *Managing Resources in End-to-End Machine Learning Pipelines*
Jim Dowling, Royal Institute of Technology, Sweden

Pipelines have become the defacto abstraction used to productionalize machine learning (ML) models, as they enable the automated processing of raw data (on typically large numbers of CPU-based containers) into a format that can be used by containers with supported hardware accelerators (GPUs, TPUs) to train models that are then deployed onto elastic infrastructure for serving models over the network to client applications. In this talk, we discuss the challenges of managing resources for such ML pipelines, and how we support this in the Hops platform, using YARN, and the challenges ahead in moving to Kubernetes.

9:00 *Edge Computing: Challenges and Opportunities*
Eyal De Lara, University of Toronto, Canada

Edge computing expands the traditional cloud architecture with additional datacenter layers that provide computation and storage closer to the end user or device. For example, a wide-area cloud datacenter which serves a large country can be augmented by a hierarchy of datacenters that provide coverage at the city, neighborhood, and building level. By leveraging geo-replication, edge computing facilitates next generation mobile and IoT applications that require low latency, or produce large volumes of data that can overwhelm the network. This talk will discuss challenges and opportunities that the advent of edge computing raises for the research community.

9:20 *Intelligent automation at scale for Cloud Services*
Mazin Yousif, DXC Technology, USA

Intelligence and automation are trending to become integral components of every computing architecture, deployment, operations, and support; making the surface area where they can be used very large. However, looking at business processes from an end-to-end perspective, there is always a weak link component that can be a human or a physical machine that slows the process and limits scalability. Here is where Intelligent automation, which bring together AI, RPA (Robotic Process Automation), and BPO (Business Process optimization) together, along with lean and agile processing considerably improve performance and scalability. This talk will discuss intelligent automation and how to scale it for large enterprises.

9:40 *Fractal: Automated Application Scaling*
Richard Mortier, Cambridge University, UK

Cloud applications are typically deployed as virtual machines, containers, and even unikernels. Resources are allocated to these deployment units, whether manually, or automatically according to static configuration or a more complex dynamic orchestration system. However, even with the rise of the devops movement and orchestration facilities such as Kubernetes, there is a tendency to separate development from deployment: the application developer does not specify how the application should be deployed.
I will briefly present an exploration of a more extreme point on the devops spectrum: Fractal. Building on previous work developing the MirageOS unikernel framework, Fractal provides an API and extensions to the Jitsu toolstack that allows a unikernel developer to embed orchestration logic inside their application. The result is a unikernel that can self-scale, adjusting the number of live replicas of itself in response to demand based on metrics and resource requirements specific to that particular application.

10:30 *Automatic Diversification of SaaS and PaaS*
Benoit Baudry, Royal Institute of Technology, Sweden

Continuous deployment, micro service architectures, live monitoring are key features that support live experiments on cloud applications. These experiments can consist in testing new the relevance or performance of new features, e.g., canary releases or A/B testing. Some other types of experiments serve in assessing the reliability of the application under certain perturbation.
This talk focuses on the latter type of experiments: proactively perturb and evolve the microservices in order to assess and improve the resilience of the application. I will introduce some recent work that leverage automatic code transformations and container technology to run these experiments based on the following concepts: chaos, randomness and diversity.

10:50 *Who Will Write Those Industry Production Papers?*
Indranil Gupta, University of Illinois, Urbana-Champaign, USA

Hundreds to thousands of production systems [are running/ran/could've run/tried to run] inside the system stacks for cluster scheduling and cloud management. A very small fraction of these systems are published or even known outside the company. Can academic researchers outside the company help?

11:10 Discussion 6 (Utkiken)
From the Edge to the Cloud and back again, John Kubiatowicz, UC Berkeley, USA

Over the last decade and a half, the Cloud has evolved into a mature platform for sophisticated, elastic delivery of services. During this time, the Cloud has experienced a convergence of sorts -- as best practices emerged with respect to mechanisms, computational models, service offerings, etc. Further, the Cloud has become an extremely successful architectural paradigm (and served as fodder for many iterations of this workshop); today, new applications of the Cloud arise almost daily. However, in retrospect, the Cloud has never offered a complete solution for all users and all applications. By centralizing resources, Cloud platforms manage to avoid (or forestall?) sticky questions of security, privacy, and data ownership. One might argue that these complex issues were never *solved* by the Cloud, merely reduced *in perceived risk* by the small number of large service providers. In addition, the centralization of the Cloud complicates latency, bandwidth, and QoS-driven concerns for applications at the edge of the network.

Fast-forward to the explosion of interest in the Internet of Things (IoT) and Edge Computing. This interest is driven by exciting new cyberphysical applications such as smart manufacturing, self-driving vehicles, face-recognition-driven security, and home automation. These applications are characterized by their melding of new and unique resources at the Edge with computation and storage in the Cloud. Although many attempts have been made to develop middleware and computational paradigms that support such melded Edge environments, none of these have yet reached anything close to universal applicability. In addition, Edge environments are notoriously underpowered, unstable, and insecure.

Thus to the charter of this discussion session: These new applications highlight deficiencies in the Cloud model and call into question the conventional wisdom of how to approach intelligent infrastructure. We hope to explore what is needed for a truly functional melding of the Cloud and Edge paradigms. How can we get the best of both environments -- the centralization and stability of the Cloud in cooperation with the flash and variety of the Edge? What new mechanisms do we need at the Edge? Do we need to re-architect the Cloud or Network? Ideally, we will put our preconceived notions at the door and approach this will a greenfield point of view.

11:10 Discussion 7 (Sandö)
Feedback loops: what for, where, how, how many ... ?, Eric Rutten, INRIA, France

The Cloud Control workshop, from its inception, even in its logo, feature the feedback loop centrally to its topic; in recent editions other topics took more space, like Fog and Edge computing, where flexibility is even more present and multi-faceted than in Cloud Data Centers, because of e.g. heterogeneity and mobility. An interesting discussion would be to make a synthesis of what is happening in these topics, from the points of view of infrastructures and systems (mechanisms, metrics), software architecture (layers, hierarchies), control theory (multi-criteria, adaptive control), AI/ML (advantages, limits and dangers) e.a.

11:10 Discussion 8 (Alma)
Serverless Platforms:Resource Management and Benchmarking, Animesh Trivedi and Alexandru Iosup, VU Amsterdam & Nikolas Herbst, University of Würzburg, Germany

Microservices, containers, and serverless computing are part of a trend toward applications composed of many small, self-contained, and automatically managed components. Core to serverless computing, Function-as-a-Service (FaaS) platforms employ state-of-the-art container technology and microservices-based architectures to enable users to orchestrate complex applications comprised of relatively simple components, without the need for systems expertise. Victim of its own hype, and partially due to proprietary technology, currently the community has a limited overview of FaaS platforms.

We define a serverless cloud function as a small, stateless, on-demand service with a single functional responsibility. The function implements specific business logic, depending on the goal of the application. We identify as main characteristics of cloud functions:
(C1) Cloud functions are short-lived: each function takes in typically small input and produces output after a typically short amount of time, which means they can be easily operated automatically.
(C2) A cloud function is devoid of operational logic: any operational concerns are delegated to the platform (operational, cloud-managed) layer, which allows the cloud function to be platform-agnostic.

In this proposed discussion session, we would like to dig into an in depth discussion about the new challenges of resource management beyond the ones in classical cloud data centers. Some examples for new challenges are:
(I) Package-aware scheduling of FaaS functions for optimised execution container reuse and smoother cold starts,
(II) Placement strategies that leverage serverless workflow definitions and data locality information,
(III) Orchestration of serverless workflows and function compositions, across one or several cloud providers.
(IV) Memory, storage, and network engines that facilitate extending serverless operations from the current stateless to more data-intensive execution, without significant performance penalties.
(V) Credible and representative instruments for experimental performance analysis, and metrics and tools for benchmarking.

Regarding point (V), so far, experimental research on serverless computing platforms focusses on cold start measurements of public provider's serverless offerings. Pivotal building blocks for a representative serverless platform benchmarks are missing. Thus, in the second half of the discussion session, we would like to focus the discussion on topics such as: the definition of quality attributes of serverless platform management as intuitive metrics, reference applications and workflows derived from common use cases that can be executed agnostic of the supported language runtimes, a measurement methodology that enables reproducible results in a volatile environment.

More topics related to the resource management and benchmarking of serverless platforms are also welcome. (State them at the start of the session, so they get a fair share of attention.)

11:10 Discussion 9 (Kungen)

Can We Build a Cloud in Which Users Own Their Data?, Anton Burtsev, University of California, Irvine, USA

Today, public and private clouds are de facto platforms for a variety of workflows ranging from video surveillance to large-scale data analytics. A huge fraction of data involved in cloud workflows is privacy and security sensitive. Furthermore, many cloud computations are inherently data-centric and collaborative, i.e., involve data from multiple participants. Oftentimes, participants need to combine their respective (potentially sensitive) data to perform joint cloud-based computation, e.g., several hospitals might want to correlate a large genomic dataset with the dataset of electronic medical records. For example, in large-scale Genome-Wide Association Studies (GWAS) both privacy and security are of paramount concern, given unparalleled sensitivity of genomic data. Naturally, individuals who agree to participate in a GWAS need strong guarantees that their genomes will not be leaked or used beyond a specific study. However, commodity clouds provide very little mechanisms to ensure that sensitive datasets cannot be exfiltrated via an attack on one of the participants or through a malicious cloud provider.

Is it possible to build a cloud in which users own their data, and what architectural and algorithmic mechanisms are involved? The discussion will touch on a broad collection of mechanisms required to implement secure cloud ranging from novel, potentially fully-verified hypervisors aimed to protect the end-hosts, algorithms and hardware mechanisms for secure and oblivious computations, fine-grained access control in the cloud network, and the feasibility of cloud-wide information flow control.

11:10 Discussion 10 (Klubben)

Beyond the Edge: A New Market Paradigm for Mobile Resource Providers, Danlami Gabi & Monica Vitali, Umeå University, Sweden

In MECs, resource allocation is always a bottleneck when meeting the computing needs of resource hungry applications. Recent research shows that both cloud and edge resource providers can coexist to improve the existence of edge computing environments by meeting high computation of resource hungry applications. The coordination of multiple and heterogeneous resource providers is a potential solution but is also a very difficult task due to the fact that each provider has different service quality and might set a different cost on its resources. In this session, we will discuss some of the challenges and points out some open research questions arising from the coordination of both Cloud and MECs resource provider.

13:45 *Google Flex*

Steve Webster, Google, USA

Ensuring that users in a computation cluster can receive strong guarantees about immediate resource availability is surprisingly hard to do without wasting significant amounts of resources. The Google Flex system gives administrators a way to define resource pools that provide a range of strong, statistically-backed guarantees based on user and job behavior, while also reducing human effort by automation tied into the Borg cluster manager, Colossus distributed file system, and many other services. Flex has been widely adopted at Google, and has produced significant resource savings.

14:05 *Burstable Instances: Models, Equilibria, and Revenue*
Carlee Joe-Wong, Carnegie Mellon University, USA

Temporal variation in cloud workloads can hurt both tenants and providers: providers must provision their resources for unexpected peaks in tenant demands, and under traditional on-demand pricing tenants must pay for their workloads' peak demands to ensure that these demands can always be satisfied. To better match the time-varying workloads of tenants and further reduce their costs, leading cloud providers recently introduced a new instance type named burstable instances. In the research community, however, little has been done to understand burstable instances from a theoretical perspective. We present the first framework to model, analyze, and optimize the operation of burstable instances. Specifically, we model the resource provisioning of burstable instances in different service classes, identify key performance metrics, and derive the performance given the resource provisioning decisions. We then characterize the equilibrium behind tenants' responses to the prices offered for different burstable instance service classes, taking into account the impact of tenants' actions on the performance achieved by each service class. In addition, we investigate how a cloud provider can leverage the knowledge of this equilibrium to find the prices that maximize its total revenue. Finally, we validate our framework on real traces and demonstrate its usage to price a public cloud

14:25 *Where blockchains fail (and why cloud computing is of no help)*
Maarten van Steen, University of Twente, The Netherlands

Blockchains have become immensly popular and are high on the list of national and international research and innovation agenda's. This seems to be partly caused by the numerous interesting applications, combined with the promise of full decentralization and high scalability (among others). However, there are some fundamental problems with blockchains, notably when it comes to scalabilty. In this presentation I will focus on these problems and argue that we need to temper expectations concerning blockchains until some of their fundamental issues have been adequately addressed. As computer scientists, we have a special responsibility as the hype around blockchains at points seems to truly unfounded.

14.45 Discussion 11 (Utkiken)
Experimentation on cloud, Lucas Nussbaum, Université de Lorraine/Inria, France & Maarten van Steen, University of Twente, The Netherlands

About 20 years ago, researchers on distributed systems were lacking the facilities to back up their theoretical work with serious experiments that would allow them to measure performance of their solutions across wide-area networks. In 2003, PlanetLab saw the light of day providing the appropriate means for experimentation. Since then, various experimental infrastructures have been set up, including the Distributed ASCI Computer in The Netherlands and Grid5000 in France. With the emergence of the Internet of Things, novel networks, and various forms of edge computing, we are facing a similar situation as 20 years ago: the lack of an appropriate testbed to run large-scale experiments.

In this session, we will be looking at discussing what we actually need as a cloud testbed: what would such a testbed look like, what are the basic requirements, how can experiments be run independently and at the same time, how can a specific configuration for an experiment be realized, etc. There are no answers to these and other questions, so active participation is essential.

14.45 Discussion 12 (Sandö)
Systems for Machine Learning, Jim Dowling & Seif Haridi, KTH, Sweden

Machine learning (ML) introduces both new challenges and opportunities for cloud computing environments. ML applications introduce new challenges that span the application stack from support for hardware acceleration (GPUs/TPUs), resource management, elastic serving of models, and support for distribution when training models and optimizing hyperparameters.
In this discussion group, we will discuss these challenges, within the context of ML pipelines. ML pipelines have recently become the defacto end-to-end abstraction for taking raw data, refining that data, training models with the refined data, and deploying/monitoring the resultant models in production.

14.45 Discussion 13 (Alma)

Challenges of a high throughput scheduler in a real world environment, Gonzalo Rodrigo, Apple, USA

The workload of a single schedule unit can be composed by millions of jobs that arrive at a rate of hundreds to thousands per second. In such scenario, the scheduler has to be able to evaluate the jobs at least at the same rate as they arrive to ensure high utilization and short turnaround around time. Although high throughput rates can be achieved through relatively easy optimizations, real, highly available systems present challenges that slow down the calculations needed to perform scheduling. In this discussion, we will talk about these challenges and existing and possible ways to overcome them. Example of such challenges are:
A highly available scheduler usually requires multiple instances with consistent views of the state of the jobs. Distributed consistency increase the latency of any persisted write to state of a job, and this affects the cost of scheduling calculations.
Workloads are heterogeneous in terms of job requirements vs. resources offering, runtime, size, and importance.
Resources are also heterogeneous: The calculations on ideal placement are becoming increasingly expensive.
Complex job policies (e.g., hierarchical fair-share) are expensive to calculate.
What if preemption is not possible? i.e., a job placement decision cannot be cancelled.
The split brain problem of multi level schedulers.

14.45 Discussion 14 (Kungen)

Service meshes for cloud control, Cristian Klein & Johan Tordsson, Umeå University and Elastisys, Sweden

A current trend in software architectures is the transition from large monolithic applications (single binaries with full functionality) into graphs of (up to) hundreds of loosely coupled microservices. Service meshes are low overhead, configurable layer-7 networks that can be used to interconnect microservices, commonly implemented as highly configurable network proxies deployed in sidecar containers. A service mesh improves observability, as application level metrics, such as throughput and latency, can be monitored without application modification. At the same time, service meshes also provide novel actuators in the form of highly customizable load balancing, as well as circuit breakers. This technology has lately been used for microservice debugging, chaos engineering, as well as to automate canary testing and rollbacks of software updates.

- What are the opportunities to use service meshes in cloud control?
- How would service mesh autonomics complement (or contradict?) traditional resource management approaches such as autoscaling and scheduling (placement of VMs and containers)?

14.45 Discussion 15 (Klubben)

Distributed data in edge, Mina Sedaghat & Ola Angelsmark, Ericsson, Sweden

One idea with the edge is to bring the compute closer to the source of data generation. However, using the edge to only execute stateless code and do simple local calculations, would not be using it to its potential. To fully utilize edge, the applications running and distributed over edge should be able to be stateful and edge platforms should be able to run stateful computations.
For distributed stateful services, where little bits of data are spread everywhere, we face consistency and coordination issues. Without consistency guarantees, applications, devices and users see different versions of data, which can lead to unreliable applications, different executions, data corruption and data loss. — and this is when edge computing becomes difficult.
In this discussion session, we want to tackle the following questions: What are the actual requirements to support stateful applications running on edge? How to handle distributed data? How to store, process and query distributed data? How to manage and coordinate state across a set of edge locations or nodes and synchronize data with consistency guarantees?

Thursday June 13th

8.20 *SplitServe: combining lambdas and VMs for effective overload management of stateful, multi-stage, parallel processing frameworks*
Bhuvan Urgaonkar, The Pennsylvania State University, USA

Since virtual machines (VMs) in the public cloud may take a few to several minutes to launch, autoscaling techniques must navigate a fundamental cost-performance trade-off. They may over-provision VMs and use the extra capacity to hide the latency of bringing on new VMs during a scale-up operation. Alternatively, they may provision tightly (i.e., closer to predicted needs) but suffer from poor performance during scale-up. In this talk, we will describe ongoing work on two systems called BurScale and SplitServe that use recently emergent burstable VMs and serverless lambda functions, respectively, for improving the cost and performance during scale-up operations for diverse workloads including a Web server, memcached, and Apache Spark.

8:40 *The emergence of hybrid clouds*
Eran Raichstein, IBM Research, Israel

During last couple of years consumption of remote resources; compute, storage and variety of services became common for home-users as well as for enterprises. This resource consumption model is known today as the 'public' cloud. In contradiction stands the legacy model; 'on-prem' a resource consumption model hat was and still is the basis for all leading enterprises for many years. Now we start to see real erosion in that base model even for workload that were considered 'never move to cloud'. 'public' cloud remote resources are available virtually, consumed globally, and can be accesses from everywhere over the internet. In addition remote resources are available from variety of vendors in various pricing models and with significant functional overlap. With such amount of options most enterprises end-up today choosing a mixture of local and remote resources from multiple vendors in parallel. This phenomena is known as the 'Hybrid cloud' (sometimes also referred to as 'multi cloud') and is a basic challenge for all modern enterprise IT teams. Pick and choose the (a) best-or-breed or the (b) best-fit are common strategies for each such service / function but how those resources interact one with the other? and what are the communication challenges between all those services? (remote and local). How to make sure that those services can scale to the same level? and have commutative SLA that meet the need? Those questions require good planning. In-addition, on going changes to both internal requirements and to the capabilities of the services require decisions that allow flexibility and future agility. In this session we will introduce the 'Hybrid cloud' model and compare with the legacy 'on prem' model. We will present the current state of 'Hybrid cloud's focus on the available services and functionality. We will discuss several of the challenges that the 'Hybrid cloud' model impose and describe the communication challenges between multiple diverged segments composing customer level 'Hybrid cloud's.

9.00 *Network-aware energy-efficient virtual machine management in distributed Cloud infrastructures with on-site photovoltaic production*
Anne-Cécile Orgerie, CNRS - IRISA, France

Distributed Clouds are nowadays an essential component for providing Internet services to always more numerous connected devices. This growth leads the energy consumption of these distributed infrastructures to be a worrying environmental and economic concern. In order to reduce energy costs and carbon footprint, Cloud providers could resort to producing onsite renewable energy, with solar panels for instance. In this talk, I will present our approach NEMESIS: a Network-aware Energy-efficient Management framework for distributEd cloudS Infrastructures with on-Site photovoltaic production. NEMESIS optimizes VM placement and balances VM migration and green energy consumption in Cloud infrastructure embedding geographically distributed data centers with on-site photovoltaic power supply.

9:20 *Dynamic cloud storage management*
Khuzaima Daudjee, University of Waterloo, Canada

Cloud storage systems typically choose between replicating or erasure encoding data. I will present an overview of EC-Store, our erasure-coded storage system that incorporates strategies for dynamic data access and movement based on workload access patterns, and that can achieve better performance while incurring lower storage overhead than replication.

9:40 *Multi-Period Portfolio Optimization for "Cloud" Scheduling*
Ahmed Ali-Eldin, Umeå University/UMASS, Sweden/USA

Portfolio optimization is a technique used in finance and stock markets to optimize the selection of a portfolio of assets from all the possible sets of assets/stocks under risk according to some objective function such as maximize the risk adjusted return. This technique has found numerous applications in Computing science, including for scheduling and algorithm selection. In this talk, I will briefly describe an improved portfolio selection mechanism called Multi-Period

Portfolio optimization which tries to solve the portfolio problem when there are some predicted dynamics such as a price change or a change in risk. I will then talk about how this technique can be used for server scheduling with autoscaling, briefly describing a system we built to run Web applications on transient computing resources. I will also briefly discuss how this technique can find applications in other scheduling problems. This talk is based on our paper to appear in HPDC 2019, which has been shortlisted for the best paper award.

10:30 *Design of a "Zero-carbon Cloud" Scientific Computing Facility: Economics, Operation, and Environmental Impacts*
Andrew A. Chien, University of Chicago, USA

We recently proposed to the US National Science Foundation to construct and deploy a "Zero-carbon cloud" style scientific computing facility. Preparation included economic, environmental, and productivity analysis. Further design included the computing system and operational plan. Key results include:

1) A facility expected to deliver >200M XSEDE Service Units/year for parallel jobs as large as 7,800 cores
2) Projected 90% availability for the largest jobs
3) Projected elimination of 4.8M metric tons of CO2 emissions (308,000 US person-years) over the 5-year project life
4) Economic parity with traditional reliable computing deployment models, after allowing for the varied subsidies US universities provide to those facilities.

We will describe the analysis and system design, its highlights, its limitations, and prospects for further benefits.

10:50 *Towards Software-defined Monitoring*
Patrick Eugster, University of Lugano, Switzerland

Current network monitoring paradigms (e.g., sFlow, IPFIX) are collector-based, with (1) agents executing on network devices which are "lightweight" in that they collect raw samples or simple statistics for given flows, and pipe these straightforwardly to (2) a centralized collector which computes a global picture of the network status. Intuitively, gathering all information in one place may seem to yield the most complete and accurate picture of the network state, and may be simple from an implementation perspective. Some monitoring tasks may benefit from, or even require, a centralized component. However, sending raw samples and statistics from hundreds or thousands of datacenter (DC) network appliances to a single analyzer can congest network and/or collector, and will likely incur high latency for ultimately triggering corresponding reactions on these appliances. While more efficient solutions exist for specific monitoring tasks and/or for highly specific hard- or software platforms, these are not generalizable.

We propose the abstraction of monitoring seeds which highly accurately poll statistics, probe packets, track state, and execute actions locally to network devices; these execute very efficiently in a decentralized manner and interact among each other and with a global analyzer only in specific, well-defined states. To enforce this model we propose a domain-specific programming language. This language simplifies the task of precisely describing monitoring tasks without exact knowledge of network topology and resources, by leveraging the well-known intuitive abstraction of state machine, specialized to capture communication patterns, resource utilization levels, placement constraints, as well as actions such as TCAM modifications. The runtime system re-locates monitoring seeds between devices without disruptions, and globally optimizes placement for co-existing monitoring applications based on a novel optimization algorithm that takes into account network device resources, different costs, and beneficial aggregation factors. In preliminary experiments heavy hitters were detected in a DC of SAP in ~1ms, compared to 4ms for the best existing approach, or 100ms for sFlow.

11:10 Discussion 16 (Utkiken)
Reducing the Environmental Impact of Cloud Computing, Andrew A. Chien, University of Chicago, USA

The growth of use of computing, and particularly cloud computing, continues without any sign of a limit. As such, it is increasingly important to find ways to reduce the environmental impact of cloud computing with concerns ranging from power/carbon footprint, but also construction, materials, computing equipment, and so on.

11:10 Discussion 17 (Sandö)
Challenges with mobile edge, Björn Skubic & Morgan Lindqvist, Ericsson, Sweden

The emergence of 5G and edge cloud breaks some of the assumptions of the traditional "one-size fits all" of the cloud model. 5G enables a wider range of services with requirements beyond best-effort mobile broadband, from mission critical connectivity to massive Internet of Things, where connectivity is service dependent. Edge-based services increase the need for control over where services are hosted, as well as connectivity between end-users and service instances. This creates new complex dependencies between services, connectivity and edge-cloud infrastructure.
- How will this impact the adoption of edge-based services?
- What will be the future model for hosting edge-based services?
- One of the biggest challenges is the mobility of the client. To reach the connectivity requirement, we may need to

change the used service instance mid-session.
- How is the (TCP/UDP/QUIC) session moved?
- Should the client application be aware of the move?
- How will it affect the service delivery model?
- How should edge infrastructure be aware of the intent of the services to allow for sufficient control whilst maintaining simplicity?

11:10 Discussion 18 (Alma)
Anomaly detection in edge/fog computing, Odej Kao, TU Berlin, Germany & Lilly Wu, Elastisys, Sweden

With running application close to the users, edge/fog computing provides a better performance to the edge/fog-based applications than clouds. However the highly distributed, the hierarchy architecture between IoT devices, edge/fog and cloud, and large-scale of application yield an imminent requirement of autonomous operation of the application and edge/fog infrastructures. As most of the edge/fog applications are latency-sensitive, once an abnormal behavior happens, how could we detect it in an efficient way, and localize the real cause from this complex edge/fog environment and recover it with leverage the resources of surrounding edge/fog nodes and cloud.

In this discussion session, we want to discuss the following questions:
1. How to define the QoS of edge/fog computing, and what's the difference from typical cloud.
2. To guarantee the QoS in edge/fog computing, what requirements in anomaly detection need to meet?
3. What's the possible kinds of anomalies in edge/fog computing, any differences between the typical cloud
4. To detect anomalies in edge/fog computing, what challenges we will meet? Any advantages we could take from fog/edge computing for anomaly detection, root cause analysis and remediation
5. Regarding to the highly distribution, hierarchy architecture in IoT devices, fog/edge nodes and cloud, how to distribute the monitoring, anomaly detection, root cause analysis and remediation among them.

11:10 Discussion 19 (Kungen)
Breaking the limits of Kubernetes cluster federation for 5G telco clouds, Lars Larsson, Umeå University, Sweden

To deliver on the promises of 5G, computation and data must dynamically be kept closer to end-users and allow for efficient point-to-point communication across the global infrastructure. Kubernetes is emerging as an open source industry standard for deploying containerized applications. The current proposal for Kubernetes federation will, according to its designers and creators, scale to about a dozen clusters. While sufficient for today's globally deployed multi-region applications, it is woefully insufficient for supporting the 5G telco cloud of tomorrow. What are the limiting factors, where can trade-offs be made, and what new approaches should be considered to support federation of thousands of clusters?

11:10 Discussion 20 (Klubben)
Open parallel slot

12:30 *Closing*
Erik Elmroth, Umeå University, Sweden