Research Statement

Ioannis Papapanagiotou

My *current and future plans* include investigating problems in the areas of Distributed Systems/Cloud Computing, High Performance Deduplication Systems and Internet of Things. My research is based on a three-step approach: (a) identify a unique engineering problem in the aforementioned domains; (b) develop the appropriate theoretical framework; (c) evaluate through simulation, emulation or prototyping. In what follows, I will cover my ongoing projects, my vision and my research accomplishments.

1 Current Projects & Vision

1.1 Cloud Computing

I have been working on several layers of the Cloud from designing Infrastructure as a Service (IaaS) layer systems to building distributed systems that can be offered as a Platform as a Service (PaaS). My work in this domain combines both extensive research as well as cross-organization practical experience from developing, maintaining and offering a PaaS itself[1].

In the IaaS domain, my work involves designed archictures with guaranteed performance Service Level Objectives (SLOs). The SLOs play a crucial role in cloud services as they are based on provisioning resources on demand, otherwise known as cloud elasticity. The goal of this project is to investigate central orchestration techniques such that the performance of both the compute and the storage services are optimized.

In our first work [30], we designed an SLO-aware resource scheduling for cloud storage systems in which we achieved a 20% improvement in performance violations, with fewer storage nodes, and higher throughput per backend system, compared to the default scheduler in OpenStack Cinder project. We extended this work taking into account multiple volume requests per second and multiple backend systems [31]. We showed that the proposed scheduler can guarantee performance SLOs even in a dynamic workload, increase the I/O throughput of the volume that have been already provisioned in the backend systems, scale to a higher arrival rate for the volume requests, and finally minimize the number of active hosts (or else the energy consumption). The volume placement process is based on an APX-hard multi-dimensional Vector Bin Packing (VBP_d) algorithm. In order to reduce the complexity we propose a heuristic named Modified Vector Best Fit Decreasing (MVBFD). Finally, we extended this work towards a black-box approach. More specifically, we proposed a scheduler based on machine learning that can dynamically adapt based on the workload, and can efficiently provide a schedulign decision with zero knowledge of the underlying hardware [25]. Our scheduler designs for block storage systems is based on the principles of the OpenStack's Cinder scheduler. Part of this work is sponsored through a NetApp Faculty Fellowship.

We also introduced Serifos, an autonomous performance modeling and load balancing system designed for Solid State Disks (SSDs)-based Cloud storage systems [32]. Serfios takes into account the characteristics of the SSD storage units and constructs hardware dependent workload consolidation models. Thus Serifos is able to predict the latency caused by workload interference and the average latency of concurrent workloads. Furthermore, Serifos leverages an I/O load balancing algorithm to dynamically balance the volumes across the cluster. Experimental results indicate that Serifos consolidation model is able to maintain the mean prediction error around 10% for heterogeneous hardware. As a result of Serifos load balancing, we found that the variance and the maximum average latency are reduced by 82% and 52%, respectively. The supported Service Level Objectives (SLOs) on the testbed improve 43% on average latency, 32% on the maximum read and 63% on the maximum write latency.

My future vision is to extend our work towards enabling (i) batch-scheduling for multi-tenant architectures, and (ii) optimized cloud orchestrator with SLO guarantees. The optimal allocation of the virtual

[1]https://github.com/Netflix/dynomite

resources will be decided through an optimization orchestrator, i.e. maximizing a utility function $u_i = f_i(N_i)$ that will adapt multi-service allocation for each application α_i. N_i will be the virtual resource allocation vector of application α_i. Moreover, we are looking in several "verticals" of the Cloud computing such as healthcare [5].

1.2 High Performance Network Deduplication

The increasing demand for faster networks in the High Performance Computing era is outpacing the ability of the carriers to upgrade their Tier-1 infrastructure. The importance of building this needed infrastructure, such as 100Gbps networks, has been identified by DOE [7], while key research labs, such as Lawrence Berkeley National Laboratory, are currently looking into pathological problems of networks such as the Energy Service Network (ESnet). Recent work highlighting the efficiency problem has shown that utilization can be as low as 5% for files of size of 1MB, and as large as 50% for files of size of 16MB. A potential solution can be the implementation of Data Redundancy Elimination (DRE) systems, which can identify and remove duplicate chunks of data [28]. However, such systems have only been deployed in an enterprise low-bandwidth setting. In my work [11], I have shown that DRE systems cannot currently handle traffic at the Gbps scale, because of bottlenecks that occur due to CPU and persistent storage space capacity constraints [6]. At high-bandwidth links, DRE systems may also also impose an extra delay due to TCP retransmissions. I have also demonstrated that Web proxy caches are inefficient in some scenarios due to the Zipf Law property of Web objects [10]. Nonetheless, I believe that a DRE system redesign can fundamentally increase our ability to solve current network limitations. In our recent work [4], we proposed a Hardware Accelarated Redundancy Elimination in Network Systems (HARENS) that can significantly improve the performance of redundancy elimination by leveraging General Purpose Graphics Processing Unit (GPGPU) optimizations, as well as other optimizations such as the use of a hierarchical multi-threaded pipeline, and memory efficient techniques. We also open sourced our work[2] as generic library for the research community to evaluate and use. Some of this work was supported by established the first NVIDIA CUDA Research Center at Purdue University[3]. *My research goal is to develop the foundation and advance the knowledge in the area of high performance computations for in line network processing.*

Hybrid Byte Level Caching: Towards this effort we first designed a hybrid level cache. Our trace driven analysis showed that some applications are more prone to packet-level deduplication, whereas others cannot benefit from deduplication (traffic through VPN). Hence, I designed and patented an apparatus that uses an intelligent scheduler to decide the optimal deduplication strategy [1, 3]. The proposed implementation showed more than twice as much savings compared to standalone proxy caches, and higher savings compared to byte caches. It also required a third of the memory to store the hashes, and provided a system speedup equal to two [11]. Several additional optimizations were also identified and filed numerous patent disclosures in this area with IBM. One of them was a distributed implementation of a byte caching systems [2]

Our current work focuses on redesigning the DRE systems and deploy them at a university network using Commercial Off the Self (COTS) hardware. I take the following principles into account in my in-line network processing design: (a) encoding performance - the speed of calculating the encoded or redundant data; (b) decoding performance - the speed of recalculating data or coding information following one or more packet losses; (c) easy of implementation - the complexity of the technique; (d) cost of implementation - licensing issues, as many coding techniques are patented; (e) interoperability - across multiple vendors or backward compatibility with current implementations.

1.3 Internet of Things

The third topic of my current research is around the Internet of Things (IoT). I am focused on two directions. The first is the secure communication of the Vehicular IoT. The car receives messages through local sensors, vehicular networks (inter-vehicle communication), and other connected entities. In this situation, the communication has to be secure, fast, and reliable. For instance, consider a group of autonomous vehicles driving on a high-way at high-speeds. Once a vehicle brakes suddenly, this should be broadcasted to other vehicles to avoid collision. It is critical to ensure the authentication of such messages to prevent an

[2] https://github.com/ipapapa/HARENS
[3] https://www.purdue.edu/newsroom/releases/2015/Q1/purdue-university-named-nvidia-cuda-research-center.html

attacker from injecting/manipulating messages. If the delay introduced by the authentication operations is comparable to the transmission delay, then a car may not be able to brake in time. In this recent project, we develop Hardware-Accelerated authentication schemes that enables practical realization of cryptographic signatures for vehicular networks. Our system is based on Multi-Processor Systems on Chip (MpSoC) and we are able to provide significantly faster calculation of signatures with the same level of security. This work is in collaboration with Purdue University's Center for Education and Research in Information Assurance and Security (CERIAS) and Oregon State University. Our most recent results can be found in [26] and in [27].

The second part of this work focuses on the implementation of the IoT in Smart Buildings. We are currently looking into technologies like Geofencing. Geofencing is a Zone Based Location Based Service (LBS) which defines a virtual fence around a certain Point of Interest. In order to be able to provide any type of geofence based service in a smart building the accurate information of the microlocation must be known. However, there has been little innovation in the microlocation area. Hence, we have looked into techniques such as particle filtering, and extended Kalman filters [33] and have open sourced an end-to-end Cloud based Micro-Location framework based on Appel's iBeacons[4]. A detailed survey of this area can be found in [35] and in the IEEE IoT Journal [34].

2 Past Projects

2.1 Network Data Analytics

This work was performed while I was working in IBM TJ Watson Research Center. We devised and patented an Operating System (OS) fingerprinting methodology that identifies each device type such as iOS, Android, BlackBerry, Windows, Mac OS X and Linux. More specifically, this methodology uses association rule mining to determine a potential correlation and trustworthiness of several protocol headers. I prototyped and deployed this idea in two wireless campuses, namely NC State University and IBM TJ Watson Research Center. The former included mostly student smartphone devices, and the latter corporate laptops. My technique was able to accurately identify 98% of the devices in both networks [15], a figure that higher $10-20\%$ from other state-of-the art OS fingerprinting techniques.

Differential DHCP lease time: I investigated the impact that new types of wireless devices, such as smartphones, have on DHCP. I used two one-month long traces, collected at the gateway of NC State University and IBM TJ Watson Research Center, and compared side-by-side their DHCP usage patterns. I showed that DHCP implementations vary among device types, and have an effect on DHCP lease durations. To improve network address utilization, without introducing any protocol changes, I proposed a new leasing strategy which takes into account the device types. This strategy, compared to current approaches, provides a sixfold improvement in the address utilization without considerably increasing the DHCP overhead. My work was presented in ACM IMC 2012 [15] and granted a patent [9, 8].

Web Traffic: In this work, we focused on the areas where the traffic between mobile and laptop is different. More specifically, we looked into video delivery as it was the bulk of the network traffic requested by the end-user. Using the traces collected in the aforementioned locations, we studied the effectiveness of the device caches, as well as the effectiveness of the proxy caches. We showed that smartphone caches are not effective at all. Hence, we proposed caching polices for smartphone devices in which only a 10MB cache would provide the maximum achievable savings. In addition, we showed that there would be an considerable improvement in the Service Provider's bandwidth if proxy caches would be able to handle partially downloaded objects. My work was presented in ACM SIGMETRICS 2012 [16].

2.2 Functionality Distribution in Aggregation Networks

This work was performed in collaboration with Cisco Systems. We used Operations Research principles in the area of network design & dimensioning problems. This work was motivated by the lack of established principles on the design of next generation aggregation architectures, which have a tree structure. Specifically, most of the design problems address the issue of deploying new infrastructure. However, the cost to build

[4] https://github.com/ipapapa/IoT-MicroLocation

out new infrastructure may be prohibitively expensive. In addition, vendors have introduced multipurpose edge "systems" (backplanes) that incorporate different functionalities in modular "sub-systems" (line cards). Finally, Internet Service Providers (ISPs) have to deal with the explosion of video traffic in different formats, e.g. multicast, peer-to-peer, mobile etc. Hence, it was of vital importance for the ISPs to re-engineer their infrastructure to meet these challenges. Moreover, Network Design Problems (NDPs) have to be revisited to include those multipurpose edge systems, instead of single-functionality devices (such as L2 switches or L3 routers).

Considering the above challenges, we developed a network design model that goes beyond the well investigated *location* and *dimensioning* problems. Our modeling approach is applied to both designing an aggregation architecture, as well as upgrading the current infrastructure. We proposed models based on edge "systems", rather than network elements. Edge systems may support different types of intelligence, either as part of the backplane or as part of the modular "sub-systems", i.e. high-end line cards. Finally, I evaluated the model with two close-to-real-life scenarios and multiple traffic profiles, based on data provided by EU ISPs. My results yield that ISPs will benefit from distributing the functionalities closer to the subscriber. Furthermore, sensitivity analysis suggested that ISPs should invest on systems that support more line cards and interfaces, rather than on devices with higher bandwidth. Our work was presented in IEEE CCNC 2010 [12] and was published in the IEEE Transactions on Network and Service Management [13].

2.3 Performance Analysis of Wireless Networks

Finally, I worked with analytical decomposition of wireless protocols. Specifically, I studied the Quality of Service (QoS) in WiFi networks (IEEE 802.11e). The motivation for this study was twofold; first, some IEEE 802.11e properties were not properly captured by prior analyses, and second, the current standards do not define principles for strict QoS [20, 29]. I used three common analytical approaches for wireless networks, namely Discrete Time Markov Chains, Closed form Queueing networks, and analysis using elementary conditional probabilities. I proposed several enhancements that included (a) QoS class differentiation, (b) Block-Acknowledgments, as defined in the IEEE 802.11e standard, and (c) erroneous channel conditions. I demonstrated that Block-ACKs may improve the QoS under erroneous channel conditions [17, 18]. I also proposed a methodology to incorporate non-markovian traffic to the analytical approaches. In addition, my colleagues and I devised a scheme that combines TDMA for the VoIP class, and CSMA/CA for the rest of the QoS classes [23]. Our analytical results indicated that the saturation throughput was higher and the end-to-end delay was lower. Moreover, we showed that with the proper Admission Control, the aforementioned scheme may provide QoS guarantees [24]. We verified all these results through Discrete Event Simulations. We also extended the simulation analysis based on a Response Surface Methodology (RSM) in order to determine the effect of parameters from multiple layers of the network architecture in the performance of the wireless client [14]. During my collaboration with the VTT Research Institute, I also applied the same analytical approaches on the energy consumption of WiMAX networks (IEEE 802.16e) [22, 21, 19].

References

[1] R.D. Callaway and I. Papapanagiotou. Dynamic caching module selection for optimized data deduplication, September 18 2014. US Patent App. 13/800,289.

[2] R.D. Callaway, I. PAPAPANAGIOTOU, and A.F. Rodriguez. Distributed data deduplication in enterprise networks, June 9 2016. US Patent App. 14/559,495.

[3] Robert D Callaway and Ioannis Papapanagiotou. Scheduler training for multi-module byte caching, September 4 2014. US Patent App. 14/477,093.

[4] K. Diao, I. Papapanagiotou, and T. J. Hacker. HARENS: Hardware accelerated redundancy elimination in network systems. In *IEEE International Conference on Cloud Computing Technology and Science (CLOUDCOM)*, Dec 2016.

[5] Rahul Ghosh, Ioannis Papapanagiotou, and Keerthana Boloor. A survey on research initiatives for healthcare clouds. *Cloud Computing Applications for Quality Health Care Delivery*, pages 1–18, 2014.

[6] M. Martynov. Challenges for high-speed protocol-independent redundancy eliminating systems. In *International Conference on Computer Communications and Networks (ICCCN)*. IEEE, 2009.

[7] Parag Mhashilkar, Wenji Wu, Hyun Woo Kim, Gabriele Garzoglio, Dave Dykstra, Marko Slyz, and Phil DeMar. Research activities at fermilab for big data movement. In *SC*, 2013.

[8] Erich N Nahum, Ioannis Papapanagiotou, and Vasileios Pappas. Differential dynamic host configuration protocol lease allocation, October 26 2012. US Patent App. 13/661,349.

[9] Erich N Nahum, Ioannis Papapanagiotou, and Vasileios Pappas. Differential dynamic host configuration protocol lease allocation that classifies the client/operating system, September 9 2014. US Patent 8,832,007.

[10] I. Papapanagiotou, E. M. Nahum, and V. Pappas. Smartphones vs. laptops: comparing web browsing behavior and the implications for caching. In *SIGMETRICS*, volume 40, pages 423–424. ACM, 2012.

[11] Ioannis Papapanagiotou, Robert D. Callaway, and Michael Devetsikiotis. Chunk and object level deduplication for web. optimization: A hybrid approach. In *International Communications Conference (ICC), 2012 IEEE*, Jun. 2012.

[12] Ioannis Papapanagiotou and Michael Devetsikiotis. Aggregation network design methodologies for triple play services. In *Consumer Communications and Networking Conference (CCNC), 2010 7th IEEE*, pages 1 –5, jan. 2010.

[13] Ioannis Papapanagiotou, Matthias Falkner, and Michael Devetsikiotis. Optimal functionality placement for multiplay service provider architectures. *IEEE Transactions on Network and Service Management*, 9(3), september 2012.

[14] Ioannis Papapanagiotou, Fabrizio Granelli, Dzmitry Kliazovich, and Michael Devetsikiotis. A metamodeling approach for cross-layer optimization: A framework and application to voice over wifi. *Simulation Modeling Practice and Theory*, 19(9):2117–2129, 2011.

[15] Ioannis Papapanagiotou, Erich M. Nahum, and Vasileios Pappas. Configuring DHCP leases in the smartphone era. In *International Measurement Conference*, IMC '12. ACM, Nov. 2012.

[16] Ioannis Papapanagiotou, Erich M. Nahum, and Vasileios Pappas. Smartphones vs. laptops: comparing web browsing behavior and the implications for caching. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '12, pages 423–424, New York, NY, USA, 2012. ACM.

[17] Ioannis Papapanagiotou, Georgios S. Paschos, and Michael Devetsikiotis. A comparison performance analysis of QoS WLANs: approaches with enhanced features. *Adv. MultiMedia*, 2007(1):1:1–1:13, April 2007.

[18] Ioannis Papapanagiotou, Georgios S. Paschos, Stavros A. Kotsopoulos, and Michael Devetsikiotis. Extension and comparison of QoS-enabled Wi-Fi models in the presence of errors. In *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, pages 2530 –2535, nov. 2007.

[19] Ioannis Papapanagiotou, Dimitris Toumpakaris, Jungwon Lee, and Michael Devetsikiotis. A survey on next generation mobile wimax networks: objectives, features and technical challenges. *IEEE Communications Surveys Tutorials*, 11(4):3 –18, quarter 2009.

[20] Ioannis Papapanagiotou, John S. Vardakas, Georgios S. Paschos, Michael D. Logothetis, and Stavros A. Kotsopoulos. Performance evaluation of IEEE 802.11e based on ON-OFF traffic model. In *Proceedings of the 3rd international conference on Mobile multimedia communications*, MobiMedia '07, pages 17:1–17:6, ICST, Brussels, Belgium, Belgium, 2007.

[21] Georgios S. Paschos, Ioannis Papapanagiotou, Christos G. Argyropoulos, and Stavros A. Kotsopoulos. A heuristic strategy for ieee 802.16 wimax scheduler for quality of service. In *Proceedings of 45th Congress FITCE*, 2006.

[22] Georgios S. Paschos, Ioannis Papapanagiotou, and Thomas M. Bonert. *Power Saving Mechanisms in IEEE 802.16e Wireless Networks*. Auerbach Publications, CRC Press, 2009.

[23] Georgios S. Paschos, Ioannis Papapanagiotou, Stavros A. Kotsopoulos, and George K. Karagiannidis. A new mac protocol with pseudo-TDMA behavior for supporting quality of service in 802.11 wireless LANs. *EURASIP J. Wirel. Commun. Netw.*, 2006(3):8:1–8:9, June 2006.

[24] Georgios S. Paschos, Ioannis Papapanagiotou, Efstathios D. Vagenas, and Stavros A. Kotsopoulos. Performance analysis of admission control for 802.11 wlan networks in ad hoc mode. 2006.

[25] B. Ravandi, I. Papapanagiotou, and B. Yang. A black-box self-learning scheduler for cloud block storage systems. In *IEEE International Conference on Cloud Computing (CLOUD)*, July 2016.

[26] A. Singla, A. Mudgerikar, I. Papapanagiotou, and A. A. Yavuz. Haa: Hardware-accelerated authentication for internet of things in mission critical vehicular networks. In *Military Communications Conference, MILCOM 2015 - 2015 IEEE*, pages 1298–1304, Oct 2015.

[27] Ankush Singla, Anand Mudgerikar, Ioannis Papapanagiotou, and Atilla Yavuz. Fast and scalable authentication for vehicular internet of things. In *Proceedings of the 16th Annual Information Security Symposium*, CERIAS '15, pages 22:1–22:1, West Lafayette, IN, 2015. CERIAS - Purdue University.

[28] N.T. Spring and D. Wetherall. A protocol-independent technique for eliminating redundant network traffic. In *SIGCOMM*. ACM, 2000.

[29] John S. Vardakas, Ioannis Papapanagiotou, Michael D. Logothetis, and Stavros A. Kotsopoulos. On the end-to-end delay analysis of the IEEE 802.11 distributed coordination function. In *Internet Monitoring and Protection, 2007. ICIMP 2007. Second International Conference on*, page 16, july 2007.

[30] Zhihao Yao, Ioannis Papapanagiotou, and Robert D. Callaway. Sla-aware resource scheduling for cloud storage. In *Proceedings of IEEE CLOUDNET*, 2014.

[31] Zhihao Yao, Ioannis Papapanagiotou, and Robert D. Callaway. Multi-Dimensional Scheduling in Cloud Storage Systems. In *International Communications Conference (ICC), 2015 IEEE*, 2015.

[32] Zhihao Yao, Ioannis Papapanagiotou, and Rean Griffith. Serifos: Workload consolidation and load balancing for ssd based cloud storage systems. *arXiv preprint arXiv:1512.06432*, 2015.

[33] F. Zafari and I. Papapanagiotou. Enhancing ibeacon based micro-location with particle filtering. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7, Dec 2015.

[34] F. Zafari, I. Papapanagiotou, and K. Christidis. Microlocation for internet-of-things-equipped smart buildings. *IEEE Internet of Things Journal*, 3(1):96–112, Feb 2016.

[35] Faheem Zafari, Ioannis Papapanagiotou, and Konstantinos Christidis. Micro-location for internet of things equipped smart buildings. *CoRR*, abs/1501.01539, 2015.