

Report from Dagstuhl Seminar 19101

Analysis, Design, and Control of Predictable Interconnected Systems

Edited by

Kunal Agrawal¹, Enrico Bini², and Giovanni Stea³

1 Washington University – St. Louis, US, kunal@seas.wustl.edu

2 University of Turin, IT, enrico.bini@unito.it

3 University of Pisa, IT, giovanni.stea@unipi.it

Abstract

We call “Interconnected Systems” any collection of systems distributed over a metric space whose behavior is influenced by its neighborhood. Examples of interconnected systems exist at very different scales: different cores over the same silicon, different sub-systems in vehicles, communicating nodes over either a physical (e.g., optical) network, or – more recently – virtualized network. Examples also exist in contexts which are not related to computing or communication. Smart Grids (of energy production, distribution, and consumption) and Intelligent Transportation Systems are just two notable examples. The common characteristic among all these examples is the presence of a spatially distributed demand of resources (energy, computing, communication bandwidth, etc.) which needs to be matched with a spatially distributed supply. Often times demands and availability of resources of different types (e.g., computing and link bandwidth in virtualized network environments) need to be matched simultaneously.

Time predictability is a key requirement for above systems. Despite this, the strong market pressure has often led to “quick and dirty” best-effort solutions, which make it extremely challenging to predict the behavior of such systems.

Research communities have developed formal theories for predictability which are specialized to each application domain or type of resource (e.g., schedulability analysis for real-time systems or network calculus for communication systems). However, the emerging application domains (virtualized networks, cyber-physical systems, etc.) clearly require a unified, holistic approach. By leveraging the expertise, vision and interactions of scientists that have addressed predictability in different areas, the proposed seminar aims at constructing a common ground for the theory supporting the analysis, the design, and the control of predictable interconnected systems.

Seminar March 3–8, 2019 – <http://www.dagstuhl.de/19101>

2012 ACM Subject Classification Networks → Cloud computing, General and reference → Performance, Computer systems organization → Real-time systems, Software and its engineering → Scheduling

Keywords and phrases distributed resource management, network calculus, real-time systems

Digital Object Identifier 10.4230/DagRep.9.3.1



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 3.0 Unported license

Analysis, Design, and Control of Predictable Interconnected Systems, *Dagstuhl Reports*, Vol. 9, Issue 3, pp. 1–15
Editors: Kunal Agrawal, Enrico Bini, and Giovanni Stea



DAGSTUHL
REPORTS


Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Kunal Agrawal

Enrico Bini

Giovanni Stea

License  Creative Commons BY 3.0 Unported license
© Kunal Agrawal, Enrico Bini, and Giovanni Stea

Computing capacity surrounds our environment more and more. The exploitation of this diffused infrastructure poses a number of challenges. A notable one is the guarantee of predictable performance. Many modern computing platforms require high degrees of predictability in their timing characteristics – for instance, in avionics and automotive applications, it is essential that the tasks complete in a timely manner in order to take appropriate actions in response to a developing situation. With increasing complexity and heterogeneity in functionality required by time-sensitive applications, there is an increasing need for designing distributed interconnected platforms that respond to time-varying requests in a highly predictable way.

The world of networking is undergoing an analogous transformation. There is a major shift towards smarter and more autonomous networks, the so-called self-driving networks. The goal is to mimic the success that cloud-computing techniques and concepts had on the transformation of the IT infrastructure. The latter made it possible to allow physical resources, such as computing and storage nodes, to be shared among users through the use of virtual resources (Network Function Virtualization). Virtualization of network functions offers an efficient way to meet dynamic user demands in a cost-effective manner. However, the guarantee of predictability in NFV is, to date, an open problem.

With increasing complexity and heterogeneity in functionality required by timing-sensitive applications, there is need for designing distributed and interconnected platforms that provide high predictability. On the one hand, with the advent of Cyber-physical systems and Industry 4.0, real-time systems are becoming more and more networked systems. On the other hand, networking scenarios where a-priori quantifiable guarantees on the worst-case traversal time are required as a prerequisite for a correct application-level computation are becoming commonplace. Therefore, it is necessary to develop methods to ensure compliance with end-to-end deadlines in distributed systems where *both* computation and communication are involved.

Dagstuhl seminar “Analysis, design and control of predictable interconnected systems” gathered more than 40 researchers from four continents, with a good balance of seniority and gender. These researchers came from different communities, including Network Calculus, Real-time Systems, Control Theory, Performance Evaluation and Data-flow, which have been discussing similar problems in the recent past, sometimes also using similar methodologies, but different notations and hypotheses, and focusing on different characteristics. The aim of the seminar was to foster cross-fertilization and inter-community networking, using new, contemporary problems which are interesting for and discussed within more than one community. The seminar hosted 26 talks. Some of these were “background-levelling” talks given by well-recognized senior experts in the respective communities, with the aim of crossing the lexicon and notation gap between communities. Other talks, building on the shared background provided by the former, discussed interesting novel problems and promising application areas. The open, schedule-as-you-go format for the talks and the time left open for free networking activities fostered an open environment. The general conclusion, which can be gathered by the (mostly enthusiastic) comments in the final survey, is that the involved communities have more in common with each other than the attendees thought, and cross-fertilization is necessary to tackle new problems in a holistic approach.

2 Table of Contents

Executive Summary

<i>Kunal Agrawal, Enrico Bini, and Giovanni Stea</i>	2
----------------------------------------------------------------	---

Overview of Talks


Network reservations for dynamic RT systems <i>Luis Almeida</i>	5
Real-Time Scheduling <i>Sanjoy K. Baruah</i>	5
Hierarchical Scheduling, supply functions and friends <i>Enrico Bini</i>	6
Performance in networks with cyclic dependencies <i>Anne Bouillard</i>	6
Analysis of MPPA Network on Chip – Is analysis of FIFO network closed in theory and open in practice? <i>Marc Boyer</i>	6
Performance Analysis based on Arrivals and Delays <i>Peter Buchholz</i>	7
On Resilient Distributed Control of Dynamical Flow Networks <i>Giacomo Como</i>	7
Challenges for Assured Autonomy in Safety-Critical CPS <i>Arvind Easwaran</i>	7
Industrial Cloud & 5G <i>Johan Eker</i>	8
Delay-Constrained Routing Problems <i>Laura Galli</i>	8
DeepTMA: Predicting Effective Contention Models for Network Calculus using Graph Neural Networks <i>Fabien Geyer</i>	8
Information Theory as a Guide to Resource Allocation <i>Sathish Gopalakrishnan</i>	9
Queuing Analysis of Wireless Systems – No Waste of Time! <i>James Gross</i>	9
An Introduction to Stochastic Network Calculus <i>Yuming Jiang</i>	9
Real-Time Calculus – A Tutorial <i>Pratyush Kumar</i>	10
What can arrival curves do for helping online algorithms to meet deadlines? <i>Kai Lampka</i>	10
An Introduction to Network Calculus <i>Jean-Yves Le Boudec</i>	10

Actor-Oriented Models of Computation for Predictable Interconnected Systems <i>Edward A. Lee</i>	11
A Fluid-Flow Interpretation of SCED Scheduling <i>Jörg Liebeherr</i>	11
Single queue equivalence for redundant requests and cloning <i>Martina Maggio</i>	11
Buffer-aware worst-case Timing analysis of wormhole NoCs using Network Calculus <i>Ahlem Mifdaoui</i>	12
Topics on providing end-to-end deadline guarantees in cloud robotics network <i>Victor Millnert</i>	12
Network Performance Tomography. A Revisit of an “Old” Problem <i>Kui Wu</i>	13
Multiple supply estimations in compositional real-time systems with mixed-criticality scheduling <i>Kecheng Yang</i>	13
Network calculus in non-congestion network <i>Jiayi Zhang</i>	14
Feedback Control with Stability Guarantees over Wireless Multi-hop Networks <i>Marco Zimmerling</i>	14
Participants	15

3 Overview of Talks

3.1 Network reservations for dynamic RT systems


Luis Almeida (University of Porto, PT)

License  Creative Commons BY 3.0 Unported license
© Luis Almeida

Recent growing frameworks such as the IoT, IIoT, Cloud/Fog/Edge computing, CPS, etc, bring the networking platforms on which they rely to the spotlight, as first class citizens of an increasingly software-dependent landscape. As a result, networks play an increasingly central role in supporting the needed system-wide properties, particularly, scalability, adaptivity, timeliness, safety and security. This trend has been generating many hard challenges and boosted an unprecedented engagement from the research community. Specifically in our case, we have been working to provide openness and adaptivity together with timeliness guarantees. This combination seems fundamental to support inherently dynamic applications in a resource efficient way, covering not only the cases of systems of systems, systems with variable number of users, components or resources but also systems that undergo frequent live maintenance and even reconfiguration during their lifetime. Examples range from autonomous vehicles to collaborative robotics, remote interactions, fog/edge computing, flexible manufacturing, etc. We postulate that combining openness and adaptivity with guaranteed timeliness can only be achieved with an adequate communication abstraction supported on adequate protocols. To this end, we have been proposing channel reservation-based communication as a means to provide scalable and open latency-constrained communication and thus enable a more efficient system design. This presentation will provide a brief tour of our recent work on flexible and composable approaches to real-time communication. We will end highlighting some open challenges towards adequate networking infrastructures for effective dynamic real-time systems.

3.2 Real-Time Scheduling


Sanjoy K. Baruah (Washington University, US)

License  Creative Commons BY 3.0 Unported license
© Sanjoy K. Baruah

Real-time scheduling theory studies problems associated with the objective of achieving efficient implementations of systems that are subject to timing constraints. As do other disciplines, real-time scheduling theory has a (fortunately, rather small) common core of knowledge – assumptions, terminology, basic results. I will provide a high-level introduction to some of this common core of real-time scheduling theory.

3.3 Hierarchical Scheduling, supply functions and friends

Enrico Bini (University of Turin, IT)

License  Creative Commons BY 3.0 Unported license
© Enrico Bini

Main reference Enrico Bini, Marko Bertogna, Sanjoy K. Baruah: “Virtual Multiprocessor Platforms: Specification and Use”, in Proc. of the 30th IEEE Real-Time Systems Symposium, RTSS 2009, Washington, DC, USA, 1-4 December 2009, pp. 437–446, IEEE Computer Society, 2009.

URL <https://doi.org/10.1109/RTSS.2009.35>

The term hierarchical scheduling denotes a spectrum of theories developed starting from the 90s to ensure the isolation of real-time applications. In this presentation, I will present the main results of hierarchical scheduling, emphasizing the connections with network calculus (namely the similarity between EDF demand bound function and arrival curves, as well as the similarity between supply bound functions and service curves). Finally, the extension to multiprocessor scheduling will be presented showing the difficulties of a straightforward extension of the single processor concepts.

3.4 Performance in networks with cyclic dependencies

Anne Bouillard (Nokia Bell Labs – Nozay, FR)

License  Creative Commons BY 3.0 Unported license
© Anne Bouillard

This talk discusses the behaviour of cluster of servers when requests are cloned. In cloud computing cloning a request means generating duplicates of the request and sending these duplicates to a given number of servers – the cloning factor. The response to the request is generated by the first server that terminates the associated computation. Cloning has been extensively studied, but the theoretical results so far have been limited to specific arrival rates and service rates distributions. In this talk, I will relax these assumptions and present our ongoing research on a G/G/1 queuing model that is equivalent to a cloning cluster.

3.5 Analysis of MPPA Network on Chip – Is analysis of FIFO network closed in theory and open in practice?


Marc Boyer (ONERA – Toulouse, FR)

License  Creative Commons BY 3.0 Unported license
© Marc Boyer

The MPPA is a many-core with 256 cores, grouped into 16 clusters, connected by a Network on Chip (NoC). This NoC applies Round-Robin arbitration at output ports. With the notion of residual service, this NoC can be reduced into a FIFO (feed-forward) network. Several algorithms and tools can compute an end-to-end delay for each flow (SFA, TFA, deborah, translation into LP problem). They have been compared on this case study, and the unexpected result is that the old TFA algorithm gives the best results, due to the fact that the problem involves link shaping, and considered curves are neither concave nor convex. It opens research on how to enhance the other approaches on such case study.

3.6 Performance Analysis based on Arrivals and Delays


Peter Buchholz (TU Dortmund, DE)

License  Creative Commons BY 3.0 Unported license
© Peter Buchholz

General performance analysis is based on the availability of arrival and service processes and computes then measures like throughputs, delays, populations. On an abstract level service processes are virtual and bounds on arrivals and delays are specified. In the talk we briefly show how this approach can be integrated in network calculus.

3.7 On Resilient Distributed Control of Dynamical Flow Networks

Giacomo Como (Polytechnic University of Torino, IT)

License  Creative Commons BY 3.0 Unported license
© Giacomo Como

Resilience has become a key aspect in the design of contemporary infrastructure networks. This comes as a result of ever-increasing loads, limited physical capacity, and fast-growing levels of interconnectedness and complexity due to the recent technological advancements. The problem has motivated a considerable amount of research within the last few years, particularly focused on the dynamical aspects of network flows, complementing more classical static network flow optimization approaches. In this tutorial, we discuss some notions, techniques, and results concerning the stability, robustness, throughput optimality, and performance analysis of deterministic models of dynamical flow networks. In particular, we focus on distributed architectures for routing, scheduling, and flow control and emphasise the role of structural properties -such as monotonicity, contraction, and separable Lyapunov functions- that enable tractability and scalability.

3.8 Challenges for Assured Autonomy in Safety-Critical CPS


Arvind Easwaran (Nanyang TU – Singapore, SG)

License  Creative Commons BY 3.0 Unported license
© Arvind Easwaran

Modern cyber-physical systems (CPS) such as autonomous vehicles are expected to perform complex functions under highly dynamic operating environments. Increasingly, to cope with the complexity and scale, these functions are being designed (semi) autonomously, including the use of Machine Learning (ML) techniques in some cases. For example, perception systems in autonomous vehicles have been designed using ML techniques such as deep neural networks. A key challenge in the design of such autonomous systems is how to achieve high safety assurance. In this talk, I provide an overview of the key technical challenges for safety assurance of ML based systems, shortcomings of existing approaches in the ML community with respect to those challenges, and some potential research directions to address them.

3.9 Industrial Cloud & 5G


Johan Eker (Lund University / Ericsson Research, SE)

License  Creative Commons BY 3.0 Unported license
© Johan Eker

New technologies such as 5G, cloud & edge computing have the potential to disrupt automation as we know it today, by redefining automation and communication architectures. In this talk we present some opportunities and challenges when digitalizing traditional industries and particularly focus on dynamical aspects such as application mobility and re-orchestration.

3.10 Delay-Constrained Routing Problems

Laura Galli (University of Pisa, IT)

License  Creative Commons BY 3.0 Unported license
© Laura Galli


Main reference Antonio Frangioni, Laura Galli, Giovanni Stea: “QoS routing with worst-case delay constraints: Models, algorithms and performance analysis”, *Computer Communications*, Vol. 103, pp. 104–115, 2017.

URL <https://doi.org/10.1016/j.comcom.2016.09.006>

Delay-Constrained Routing (DCR) problems require to route a new flow in a computer network subject to worst-case end-to-end delay guarantees. The delay of a packet flow has three components, one of which is the “queueing delay”, that depends on the scheduling algorithm implemented by the routers of the network. When flows are not independent of each other, i.e., admitting a new flow changes the delay of the existing ones, admission control policies are necessary to ensure that existing flows do not become latency-unfeasible. It has been recently shown that admission control runs contrary to the usual objective function employed in these models, i.e., minimization of the reserved rates, significantly worsening network performance. In this work we investigate the phenomenon and propose a heuristic way to overcome the problem.

3.11 DeepTMA: Predicting Effective Contention Models for Network Calculus using Graph Neural Networks

Fabien Geyer (TU München, DE)

License  Creative Commons BY 3.0 Unported license
© Fabien Geyer

Network calculus computes end-to-end delay bounds for individual data flows in networks of aggregate schedulers. It searches for the best model bounding resource contention between these flows at each scheduler. Analyzing networks, this leads to complex dependency structures and finding the tightest delay bounds becomes a resource intensive task. The exhaustive search for the best combination of contention models is known as Tandem Matching Analysis (TMA). The challenge TMA overcomes is that a contention model in one location of the network can have huge impact on one in another location. These locations can, however, be many analysis steps apart from each other. TMA can derive delay bounds with high degree of tightness but needs several hours of computations to do so. We avoid the effort of exhaustive search altogether by predicting the best contention models for each location

in the network. For effective predictions, our main contribution in this paper is a novel framework combining graph-based deep learning and Network Calculus (NC) models. The framework learns from NC, predicts best NC models and feeds them back to NC. Deriving a first heuristic from this framework, called DeepTMA, we achieve provably valid bounds that are very competitive with TMA. We observe a maximum relative error below 6%, while execution times remain nearly constant and outperform TMA in moderately sized networks by several orders of magnitude.

3.12 Information Theory as a Guide to Resource Allocation

Sathish Gopalakrishnan (University of British Columbia – Vancouver, CA)

License © Creative Commons BY 3.0 Unported license
© Sathish Gopalakrishnan

In modern embedded systems that process data, we may decide how to allocate resources across competing tasks using the value of information that we may obtain from these tasks. Such an approach differs from the use of utility-loss indices that are more commonly applied when tasks may be standard control tasks.

3.13 Queuing Analysis of Wireless Systems – No Waste of Time!

James Gross (KTH Royal Institute of Technology – Stockholm, SE)

License © Creative Commons BY 3.0 Unported license
© James Gross

The recent industrial interest in ultra-reliable low latency wireless communication systems has triggered a significant interest in evaluating different wireless system designs with respect to their delay performance. We show in this talk that stochastic network calculus offers here a suitable tool for the investigation of meaningful questions such as training length choice, antenna configuration, resource allocation and rate adaptation. We finally also discuss our most recent efforts to characterize transient system performance of wireless networks.

3.14 An Introduction to Stochastic Network Calculus

Yuming Jiang (NTNU – Trondheim, NO)


License © Creative Commons BY 3.0 Unported license
© Yuming Jiang

Stochastic Network Calculus (SNC) is the stochastic extension of Network Calculus (NC), intended for use in finding latency and backlog bounds with targeted violation probability. It belongs to the branch of queueing theory that “deals with methods for finding approximation or bounding behavior for queues [Kleinrock, 1976]”. In the talk, the basic min-plus SNC models, stochastic arrival curve (SAC) and stochastic service curve (SSC), are introduced, followed by a discussion on the need of their variations, based on sample-path analysis. In addition, the inherent dependence between the SAC and SSC models and the consequences on the analysis results are highlighted. Two techniques to exploit independence information

in the arrival and service processes are introduced. At the end of the talk, the development status of max-plus SNC is briefed, followed by a discussion on applying / extending SNC to two fundamental network cases: multi-server and wireless channel.

3.15 Real-Time Calculus – A Tutorial


Pratyush Kumar (Indian Institute of Technology – Madras, IN)

License  Creative Commons BY 3.0 Unported license
© Pratyush Kumar

This tutorial on Real-Time Calculus (RTC) will introduce the foundational concepts of RTC and extends up to advanced applications and results. It will begin with presenting the formal representation of arrival and supply functions, and the greedy processing component which is the fundamental unit of modelling in RTC. We will briefly intuit the different operations under a min-plus algebra and their application to compute bounds on worst-case response time and buffer space usage. We will discuss how to apply these results to compositional analysis of embedded systems. We will then switch gears and see more advanced applications of RTC. First, we will discuss how the abstractions of RTC are also surprisingly applicable for computing thermal bounds on processors. Second, we will discuss the use of RTC for interface-based design. Finally, we will discuss the combination of stateful models with RTC.

3.16 What can arrival curves do for helping online algorithms to meet deadlines?

Kai Lampka (Elektrobit Automotive GmbH – München, DE)

License  Creative Commons BY 3.0 Unported license
© Kai Lampka

Online Workload Monitoring with Arrival Curves and its applications

3.17 An Introduction to Network Calculus

Jean-Yves Le Boudec (EPFL – Lausanne, CH)

License  Creative Commons BY 3.0 Unported license
© Jean-Yves Le Boudec

Network Calculus is a set of theories and tools that was developed for the deterministic analysis of queuing systems arising in communication networks. It is based on min-plus algebra, and sometimes max-plus algebra. It is used to compute latency and backlog bounds, with proofs that can be formally verified. It can also provide fundamental insights into physical properties of delay systems such as “Pay Bursts Only Once “ or “Re-Shaping is For Free”. In this level-setting talk, we review the basic concepts of arrival curve, service curve, their expression with min-plus algebra, and their composition. We also describe shapers (also called minimal regulators) and their properties. We illustrate the use of these concepts on classical per-flow and class-based networks.

3.18 Actor-Oriented Models of Computation for Predictable Interconnected Systems

Edward A. Lee (University of California – Berkeley, US)

License © Creative Commons BY 3.0 Unported license
© Edward A. Lee

Broadly, an actor is a software component with private state that communicates with other actors via message passing. Networks of actors form a distributed queuing system amenable to both network calculus analysis and real-time scheduling. Questions of interest include “how can queues be kept bounded?” and “will the system deadlock?” and “will messages be delivered reliably on time?” In this talk, I will review varying semantic models that have an actor-oriented flavor, particularly with an eye towards the suitability of these models for formal analysis. The models I will review include various flavors of dataflow, Hewitt actors, synchronous/reactive systems, and discrete-event systems. I will discuss properties that make such models suitable or unsuitable for the design of predictable interconnected systems. I will illustrate these properties with several variants of software realizations that are in use today and will outline a model that I believe combines the best strengths of these while avoiding many of the pitfalls.

3.19 A Fluid-Flow Interpretation of SCED Scheduling

Jörg Liebeherr (University of Toronto, CA)

License © Creative Commons BY 3.0 Unported license
© Jörg Liebeherr

We show that a fluid-flow interpretation of Service Curve Earliest Deadline First (SCED) scheduling simplifies deadline derivations for this scheduler. By exploiting the recently reported isomorphism between min-plus and max-plus network calculus and expressing deadlines in a max-plus algebra, deadline computations no longer require explicit pseudo-inverse computations. SCED deadlines are provided for latency-rate as well as a class of piecewise linear service curves.

3.20 Single queue equivalence for redundant requests and cloning


Martina Maggio (Lund University, SE)

License © Creative Commons BY 3.0 Unported license
© Martina Maggio

This talk discusses the behaviour of cluster of servers when requests are cloned. In cloud computing cloning a request means generating duplicates of the request and sending these duplicates to a given number of servers – the cloning factor. The response to the request is generated by the first server that terminates the associated computation. Cloning has been extensively studied, but the theoretical results so far have been limited to specific arrival rates and service rates distributions. In this talk, I will relax these assumptions and present our ongoing research on a G/G/1 queuing model that is equivalent to a cloning cluster.

3.21 Buffer-aware worst-case Timing analysis of wormhole NoCs using Network Calculus

Ahlem Mifdaoui (ISAE – Toulouse, FR)

License  Creative Commons BY 3.0 Unported license
© Ahlem Mifdaoui

Conducting worst-case timing analyses for wormhole Networks-on-chip (NoCs) is a fundamental aspect to guarantee real-time requirements, but it is known to be a challenging issue due to complex congestion patterns that can occur. In that respect, we introduce in this paper a new buffer-aware timing analysis of wormhole NoCs based on Network Calculus. Our main idea consists in considering the flows serialization phenomena along the path of a flow of interest (foi), by paying the bursts of interfering flows only at the first convergence point, and refining the interference patterns for the foi accounting for the limited buffer size. The derived delay bounds are analyzed and compared to available results of existing approaches, based on Scheduling Theory as well as Compositional Performance Analysis (CPA). In doing this, we highlight a noticeable enhancement of the delay bounds tightness with our approach, in comparison to the existing ones.

3.22 Topics on providing end-to-end deadline guarantees in cloud robotics network


Victor Millnert (Lund University, SE)

License  Creative Commons BY 3.0 Unported license
© Victor Millnert

In this presentation I will, on a high level, cover some topics explored during my PhD. The focus of these topics has been on how to provide end-to-end deadline guarantees on top of Cloud/IoT/Edge-networks. After starting off by presenting some motivation from a cloud-robotics scenario, I will discuss one way to perform response-time control of a timing-critical cloud service. By then taking a step back, we will see that by properly controlling the response-times of all the nodes in the network, we can provide end-to-end deadline guarantees for the robotic applications using the services of the network. We can also show that this is possible even when the topology of the network is dynamic and changes over time, i.e., when applications join/leave and when cloud services join/ leave. I will conclude my presentation by briefly discussing what I think is an important open problem; how to learn and scale virtual network functions if you know nothing about them. Is it possible to dynamically update simple network calculus models, and still enjoy good guarantees?

3.23 Network Performance Tomography. A Revisit of an “Old” Problem

Kui Wu (University of Victoria, CA)

License  Creative Commons BY 3.0 Unported license
© Kui Wu

Network performance tomography infers performance metrics on internal network links with end-to-end measurements. Existing results in this domain are mainly Boolean-based, i.e., they check whether or not a link is identifiable, and return the exact value on identifiable links. If a link is not identifiable, however, Boolean-based solution gives no performance result for the link. We extended Boolean-based network tomography to bound-based network tomography where the lower and upper bounds are derived for unidentifiable links. We develop efficient algorithms for different objectives, e.g., obtaining the tightest total error bound or minimizing the maximum error bound across all the links. We also present two methods that can significantly reduce the total number of measurement paths required for deriving the tightest bounds. At the end of the talk, some open problems and potential applications of network tomography in domains beyond computer networks will be introduced.

3.24 Multiple supply estimations in compositional real-time systems with mixed-criticality scheduling


Kecheng Yang (Texas State University – San Marcos, US)

License  Creative Commons BY 3.0 Unported license
© Kecheng Yang

In traditional real-time systems on dedicated physical platforms, controlling and analyzing the task workloads, or the demand, have been the focus to ensure the predictability of the system. In such systems, the processing capacity of dedicated processors, or the supply, is simply linear to the length of the time interval and therefore is predictable. However, component-based design and virtualizations have been the trend in developing advanced modern interconnected systems, in which the supply may be provided in a more complicated form than linear and ensuring the predictability of the supply becomes a new challenge for the predictability of the entire system. On the other hand, to achieve a sweet point between predictability and efficiency, mixed-criticality scheduling theory has been proposed and studied, where the workload may be modeled by multiple estimations of different levels of confidence. Thus, the question arises: can we extend this idea to the supply side in component-based, virtualized systems? That is, can we also use multiple estimations to characterize the processor supply so that multiple levels of guarantees can be provided?

3.25 Network calculus in non-congestion network

Jiayi Zhang (Huawei Technology – Beijing, CN)

License  Creative Commons BY 3.0 Unported license
© Jiayi Zhang

Recently, as time sensitive network (TSN) and deterministic network (DetNet) focus on the bounded latency, both standard groups and industry go to network calculus. Can current network calculus results, based on arrival curve, service curve and min-plus algebra fit the requirements in TSN? How to establish model for Internet traffic as arrival curve, switch or router as service curve? How can we reduce the pessimism when the network scale goes large, maybe by using appropriate NC methods that fit certain scenario, eg. IP carrier network, data center networks? Will network measurements help improve latency and buffering estimation, and guide the network tuning? This presentation summarizes questions and observations from network device vendor's perspective in pursuing IP non-congestion network solution.

3.26 Feedback Control with Stability Guarantees over Wireless Multi-hop Networks

Marco Zimmerling (TU Dresden, DE)

License  Creative Commons BY 3.0 Unported license
© Marco Zimmerling

Closing feedback loops fast and over long distances is key to emerging cyber-physical systems; for example, robot motion control and swarm coordination require update intervals of tens of milliseconds. Low-power wireless technology is often preferred for its low cost, small form factor, and flexibility, especially if the devices support multi-hop communication. So far, however, feedback control over wireless multi-hop networks has only been shown for update intervals on the order of seconds without stability analysis. In this talk, I will present a wireless embedded system design that tames imperfections impairing control performance, such as jitter and message loss. Our approach entails avoiding resource interference throughout the system and decoupling higher-level logic from the time-varying network state. As a result, control design and analysis are greatly simplified, allowing us to provide conditions to formally verify closed-loop stability for physical systems with linear time-invariant dynamics. I will also present results from a cyber-physical testbed with 20 wireless nodes and multiple cart-pole systems that demonstrate the capabilities of our approach and confirm our theoretical results.

Participants

- Kunal Agrawal
Washington University –
St. Louis, US
- Luis Almeida
University of Porto, PT
- Sanjoy K. Baruah
Washington University, US
- Enrico Bini
University of Turin, IT
- Steffen Bondorf
NTNU – Trondheim, NO
- Anne Bouillard
Nokia Bell Labs – Nozay, FR
- Marc Boyer
ONERA – Toulouse, FR
- Peter Buchholz
TU Dortmund, DE
- Giacomo Como
Polytechnic University of
Torino, IT
- Arvind Easwaran
Nanyang TU – Singapore, SG
- Pontus Ekberg
Uppsala University, SE
- Johan Eker
Lund University / Ericsson
Research, SE
- Markus Fidler
Leibniz Universität
Hannover, DE
- Laura Galli
University of Pisa, IT
- Fabien Geyer
TU München, DE
- Sathish Gopalakrishnan
University of British Columbia –
Vancouver, CA
- James Gross
KTH Royal Institute of
Technology – Stockholm, SE
- Chadlia Jerad
University of Manouba, TN
- Yuming Jiang
NTNU – Trondheim, NO
- Li Jing
NJIT – Newark, US
- Pratyush Kumar
Indian Institute of Technology –
Madras, IN
- Kai Lampka
Elektrobit Automotive GmbH –
München, DE
- Jean-Yves Le Boudec
EPFL – Lausanne, CH
- Edward A. Lee
University of California –
Berkeley, US
- Jörg Liebeherr
University of Toronto, CA
- Martina Maggio
Lund University, SE
- Ahlem Mifdaoui
ISAE – Toulouse, FR
- Victor Millnert
Lund University, SE
- Geoffrey Nelissen
CISTER Research Center –
Porto, PT
- Paul Nikolaus
TU Kaiserslautern, DE
- Amr Rizk
TU Darmstadt, DE
- Ketan Savla
USC – Los Angeles, US
- Giovanni Stea
University of Pisa, IT
- Niklas Ueter
TU Dortmund, DE
- Tongtong Wang
Huawei Technology – Beijing, CN
- Kui Wu
University of Victoria, CA
- Kecheng Yang
Texas State University –
San Marcos, US
- Jiayi Zhang
Huawei Technology – Beijing, CN
- Marco Zimmerling
TU Dresden, DE

