Report from Academic Salon, 30th November – 1st December, 2023

High-Performance and Low Latency Networks and Systems

Editors

Prof. Dr.-Ing. Georg Carle

Technical University of Munich, School of Computation, Information and Technology, Chair of Network Architectures and Services

Abstract

This report documents the program and the outcomes of the "Academic Salon on High-Performance and Low Latency Networks and Systems", a workshop held on November 30th and December 1st 2023. Participants of this academic salon discussed topics in the domains of High-Performance Packet Processing, Testbeds for Network Experiments, Time-Sensitive Networking, and others. The two-day workshop was partitioned into four sessions. This report presents summaries of discussions following the various impulse talks given during the individual sessions. The report closely follows the structure of the academic salon itself.

Seminar 30th November – 1st December, 2023,
URL: https://net.in.tum.de/talks/workshops/academic_salon_23.html
Edited in cooperation with Stefan Lachnit, Marcel Kempf, Max Helm, and Johannes Späth

Academic Salon
Chair of Network Architectures and Services, Technical University of Munich, Germany

1 Introduction

Georg Carle (Technical University of Munich– Munich, Germany)

The *Academic Salon on High-Performance and Low Latency Networks and Systems* is an event that aims to bring together scientists from academia and industry working in the areas of High Speed and Low-Latency Communication, Programmable Network Components, and their application in the industrial domain. Of particular interest were topics concerning network-related innovations, in particular, techniques that utilize innovative processing and I/O technologies, like P4 or RDMA, methodologies and infrastructure to evaluate network devices, and applications of low latency communication and modeling in the industrial domain.

The goal of this workshop series is to address some of the key challenges in the areas of high-performance and low-latency network systems in relation to:

- Network-related innovations, in particular, architecture innovations such as acceleration techniques for packet processing in Network Interface Cards and Switches
- Innovative processing and I/O technologies, such as P4-based protocol processing and remote DMA technologies
- Design, specification, verification, implementation, measurement, testing, and analysis of advanced network components
- Methods, tools, and research infrastructure for performance assessment
- Applications and use-cases, in particular of the industrial domain.

This academic salon continues the success of the two previous installments, building on the concept and experiences of the previous years [Car+21][Car+22]. As before, a major goal of the academic salon is not only the exchange of raw content but also to bring the community closer together. Recordings and slides of this year's academic salon are available at https://net.in.tum.de/events/academic_salon_23/academic_salon_23.html.

2 Contents

3 Innovations for High-Performance Packet Processing

The first session of this academic salon comprised three talks addressing different challenges and new approaches in the field of high-performance packet processing. Following a brief but warm welcome and introduction by the host and session chair, Georg Carle, the invited speakers and their impulse talks sparked discussions on their topics. The session was concluded with a panel discussion.

3.1 High-speed stateful packet processing

Tom Barbette (UCLouvain – Louvain-la-Neuve, Belgium)

The session was opened with a talk by Tom Barbette entitled "High-speed stateful packet processing". He proposes a method to reach a throughput of up to 1 Tbit/s for stateful packet processing using programmable P4 switches and commodity servers. Related work by the speaker includes [Sca+23]. The following questions and comments arose in the discussion after the talk:

QUESTION: Is it possible to modify the payload after the processing by the switch?

ANSWER: We target use cases where the payload doesn't need to be modified. This is a limitation of the Intel Tofino switch, which can only operate on the packet headers. The existing implementation could be extended to perform final actions on the payload (like packet encryption) by a separate host after the packet is processed on the switch.

QUESTION: If you were not confined by the technical limitations of the Tofino switch, could similar results also be achieved with a bus technology like CXL?

ANSWER: We are moving towards disaggregated servers, so memory might be accessible by the switch. The first step towards disaggregation is CXL Memory, which makes it possible to share memory over the PCIe backplane. One approach could be to make spare memory capacity directly accessible to the switch using CXL.

QUESTION: The presented approach assumes that no rules are executed directly on the switch. Typically, only a small fraction of flows is responsible for most of the traffic. Would it be possible to offload some frequently used rules to the switch?

ANSWER: State is maintained only on the network functions running on the servers. It would be possible to offload some simple processing directly to the switch. These kinds of offloads are also used with other implementations of network functions like eBPF. An issue of offloading some parts of a network function and distributing the implementation over multiple devices is the increased complexity.

QUESTION: Are you able to combine the header and the payload without using recirculation?

ANSWER: Yes, although our approach could be considered "outside recirculation". The Tofino switch sends header data to a server for processing. The packets and their headers are rejoined after the headers are sent back to the switch.

3.2 P4TG: 1Tb/s Traffic Generation for Ethernet/IP Networks

Steffen Lindner (Eberhard Karls University of Tübingen – Tübingen, Germany)

Steffen Lindner's talk entitled "P4TG: 1Tb/s Traffic Generation for Ethernet/IP Networks" introduced the implementation of a high-performance flexible packet generator based on an Intel Tofino P4 programmable switch. Using this hardware-based approach, it is possible to generate traffic exceeding 100G. The talk is related to recently published work by the speaker [LHM23]. During the discussion following the talk, the following topics were discussed:

QUESTION: Is it necessary to recompile the P4 code when reconfiguring the traffic generator?
ANSWER: No, the P4 program is static. All configuration can be performed by changing the content of tables, allowing for fast modification of the generator parameters.

QUESTION: Is it possible to modify the payload of packets either dynamically or statically?
ANSWER: The payload is random at the initial configuration but does not change during the runtime of the generator. You could rewrite the start of the payload as far as it can be parsed. This is limited by the Intel Tofino hardware restrictions.

QUESTION: Is it possible to adapt the implementation to randomly modify data in the packet payload like for generating random GTP tunnels?
ANSWER: As long as the respective payload fields are at the portion of the packet that is parsable by the Tofino switch, this could be implemented in a similar way to the existing implementation for random IP headers.

QUESTION: Do all packets include the P4TG header?
ANSWER: The P4TG header is inserted after the UDP header and is required for features like packet loss detection. If the tested application is not UDP-sensitive, it will not see this header. If you use monitor mode, the P4TG header is not added.

QUESTION: Is the tool easy to install and configure?
ANSWER: The controller can be easily installed using Docker-compose and Makefiles. The tool is already being used by several others.

3.3 Never Miss Twice – Add-on-Miss Table Updates in Software Data Planes

Manuel Simon (Technical University of Munich – Munich, Germany)

The final talk of the first session was held by Manuel Simon. He presents the implementation and evaluation of integrating add-on-miss data plane table updates to T4P4S. This approach allows for data plane table updates with a low impact on the introduced latency and achievable throughput, enabling new applications for P4. Related work by the speaker includes [SGC23] and [Sim+21]. The following questions were raised during the discussion following the talk.

QUESTION: In your measurements: Did you consider the round-trip latency of packets or only the latency for the table lookup or insertion?
ANSWER: We measured the round-trip latency, including the reception of the packet at the DuT, the processing in the P4 pipeline, and the transmission of the packet.

QUESTION: Can you estimate the latency, which is introduced by the lookup or update itself?

ANSWER: By considering the difference between cases where there is only a table lookup and cases with an additional insertion, we can estimate the latency introduced by the table update. We also investigated how the entire processing costs can be broken down into individual sub-components (e.g. packet processing by DPDK, different P4 mechanisms) on different P4 targets.

QUESTION: What happens to packets that are received during the delay introduced by the table update?

ANSWER: The software target used for the measurements follows a run-to-completion model to optimize the cache usage. In this case, packets are queued and processed sequentially. Some hardware targets use pipelining to process multiple packets simultaneously.

QUESTION: What is the throughput achievable with this approach?

ANSWER: For the approach implemented in T4P4S we achieved 1.2 Gbit/s per core for every packet triggering an update and 2.2 Gbit/s for a more realistic scenario. This can be scaled up with multiple cores using RSS.

Panel Discussion: What is the future of high-performance packet processing architectures?

The panel discussion was opened by an open question outlining the scope of this discussion: We see various approaches to implement packet processing and offloading at the host level: smart NICs with P4 capabilities, or more conventional software-based tools like DPDK and eBPF. Is it clear what the future architecture for high-performance packet processing systems will be?

First, it was commented that ongoing scaling of semiconductor density allows for the implementation of additional offloading mechanisms and accelerators like smart NICs, including a P4 programmable pipeline and programmable processing cores. The best model, which is best suited, for leveraging these accelerator implementations is still unclear.

Addressing the talk by Tom Barbette, the question was raised if software based packet processing systems, like presented in the talk, are able to achieve zero packet loss in addition to achieving high data rates. In use cases where packet loss is not acceptable, ASIC-based solutions might be required.

Then, the discussion returned to the question of the best architecture and software interface for current and future accelerator and offloading hardware. Low-level access to hardware accelerators, while efficient, might require expertise in programming the specific hardware architecture and, therefore, also be more error-prone. Relying on a more straightforward language, like P4, which is common to many diverse architectures and letting a compiler generate and optimize hardware-specific implementations, decreases the effort to implement solutions using hardware accelerators. This moves the hardware-specific implementation and optimization efforts to the compiler development. It was also noted that the optimal architecture and implementation is strongly dependent on the targeted use case.

Next, the idea was raised to let cloud providers implement accelerators following the approach of Functions As A Service, letting the operator deal with implementation details and hardware. In the following discussion, the problem was raised that allowing a cloud provider to handle the highly complex implementations, especially for security-related network functions, might not be acceptable for all users who prefer simpler and easier approaches like P4, which are manageable by the end user.

As a remark from the audience, the difference between accelerator use cases for end-host NICs and network switches was noted. For switches, high-speed packet processing in datacenters will be implemented using a pipeline architecture, while more sophisticated functions will be run on a multi-core run-to-completion architecture. The open questions mostly center around architectures for NIC offloading. It was commented that current smart NICs achieve high performance when offloading specific functions like encryption in purpose-built accelerators on the NIC. The next generation of smart NICs will have the capacity to process all packets at high data rates internally, while the currently available solutions are limited by slow or limited processing of the processor integrated into the NIC. The best architecture for implementing and utilizing these architectures is still an open research question.

Coming back to a previous question, it was added that relying on cloud providers to deal with offloading hardware may also be problematic or not possible because of issues in the context of privacy regulations.

After a question from the audience, the discussion focused on the trade-off between flexibility and usability of the provided hardware. Also, implementing accelerators for specific hardware might be a waste of silicon if these features remain unused in most applications. Additionally, it is noted that implementations should be flexible enough to be able to deal with evolving protocols and applications, to prevent accelerator hardware from becoming obsolete. Here, processing on DPUs in combination with languages like P4 could be applied.

Finally, the discussion ended with exploring the benefits of languages like P4. One of its benefits is the lower burden of entrance allowing accelerating research and development of new applications. P4 programmable switches, like the Intel Tofino, are a sweet spot between powerful hardware, ease of programmability, and flexibility, making it relevant even with changing protocols and applications. Additionally, the use of P4 also allows users to reuse building blocks on diverse types of accelerator hardware leading to platform-independent network function implementations.

4 Tools and Research Infrastructures for Network Experiments

In the second session, chaired by Georg Carle, two talks covering approaches, testbeds, and tools for reproducible network experiments were presented. After these talks, there was a panel discussion.

4.1 Rethink testbeds with blueprints

Damien Saucez (INRIA – Sophia-Antipolis, France)

The first talk by Damien Saucez, entitled "Rethink testbeds with blueprints", presents a new approach to design testbeds called "blueprints" and gives a detailed example of a 5G network blueprint. During the discussion after the talk, the following points were raised:

QUESTION: The presented advice and approaches to develop experiments, will realistically not be followed in practice, due to time constraints. Can we automatically detect changes in the test environment (e.g., using artificial intelligence)?

ANSWER: That's what will be presented in the next talk by Sebastian Gallenmüller. After investing the time to properly define an experiment you will save time afterwards. Some tools do these things automatically (e.g., Ansible). We try to use industry-standard tools for our testbeds.

QUESTION: In France, INRIA operates a nationwide IoT-LAB testbed. Is there any relation to this?

ANSWER: SLICES-RI is a research project in Europe to build a common research instrument, e.g., for IoT research. Normally, this is built from the ground up. In practice, we integrate existing testbeds (like IoT-LAB) into the new infrastructure.

QUESTION: How did you solve the issue with dependencies? What are recommendations from your side?

ANSWER: Initially, we took the wrong approach: Patching code to work with new versions. What we do now is to log information on the versions used. Using this information you are able to go back and look at changelogs and fix problems if something breaks. Some people use transparent proxies to fix versions of utilized tools. This is a problem if the researcher actually wants a newer version. Therefore, being completely transparent doesn't work.

QUESTION: With the introduction of LLMs and AI: How far are we from analyzing and running experiments without human interaction (zero-touch management)?

ANSWER: I think no human interaction will not be possible. For finding the root cause of issues (e.g., version of the used tools, analyzing logs), this might be feasible in the next decade. The fix of the problem might be harder to automate. We are far from making an automated deployment based on a general experiment description.

4.2 The SLICES/pos Framework: A Methodology and Toolchain for Reproducible Network Experiments

Sebastian Gallenmüller (Technical University of Munich – Munich, Germany)

The second talk of the session had the title "The SLICES/pos Framework: A Methodology and Toolchain for Reproducible Network Experiments". Sebastian Gallenmüller presented the pos framework for creating reproducible network experiments. Additionally, he shared his experience from the artifact evaluation of CoNEXT 2023. Relevant related work by the speaker includes [Gal+21]. The following questions were asked after the talk:

QUESTION: Is pos available to deploy on other testbeds?
ANSWER: Not yet. We are trying to create a version that works without infrastructure specific to our deployment. We are currently in the process of dividing the orchestration part and the experiment workflow part and porting the experiment workflow to other testbeds, like Clouldlab and Chameleon. We will then release this more generic component.

COMMENT: In artifact evaluation, the setup (configuring machines, wiring, ...) is the most problematic part compared to the evaluation of results.
ANSWER: I agree. In the SLICES initiative, we try to have a federated testbed. We are trying to make the pos controller portable to other testbeds.

QUESTION: Regarding your experience in the CoNEXT artifact evaluation. How difficult was it to find reviewers? Are you aware of any mechanism to give credit to the reviewers, such as a best reviewer award?
ANSWER: First, for the nomination, we handed out a call and got approximately 40 replies. We chose 36 reviewers, which was enough for the whole evaluation. We are handing out community contribution awards for the authors but not the reviewers. Currently, we do not have any kind of incentive for the reviewers. Half of the reviewers were experienced users (already did an artifact evaluation at least once), so there seems to be enough incentive to participate.
COMMENTS: The community drives towards replicability. Therefore, papers having badges, will probably have more impact. The best process to gain the experience to perform replicable experiments and get the respective badges, is to take part in artifact evaluation.

QUESTION: Is there a feedback loop back to the author who attempted to get an artifact evaluation badge but failed?
ANSWER: Yes, reviewers can communicate with the authors, allowing them to fix problems. Not all of the authors and reviewers were responsive, so this can take some time. Getting the available badges is quite easy. The functional badge and reproducible badge are significantly more difficult.

QUESTION: Is there a time threshold to spend on getting the artifacts working?
ANSWER: For us, this was the conference date.

Panel Discussion: Which Research Infrastructure and Tools Foster Best Scientific Collaborations?

The discussion started with a short introduction followed by an open question by the chair of this session, Georg Carle, to guide the discussion: *Which research infrastructure and tools*

foster the best scientific collaboration?.

The initial comment was that experiment infrastructure needs to be useful for researchers, therefore being built by researchers, targeting the needs of this community. Additionally, it was noted that a portable experiment methodology is a key factor that is still missing in current testbeds. Being able to move experiments between diverse testbeds, without requiring the rewriting of experiment scripts could be achieved by having a well-defined structure, to which experiments adhere.

Further, it was added that not only the testbed but also the experiment itself needs to be constructed in a portable way. Based on this comment, the discussion turned to teaching measurement methodology to students. It was noted that currently the scientific method is not properly taught in university classes. As a note from the audience, it was commented that at FAU, there is a variety of courses for PhD students, but there is no course yet on artifact preparation. Maybe, offering such courses or workshops would help the community to get into such topics.

Damien Sauchez noted that at his research team at INRIA, even for internships doing mainly mathematical modeling, all work is organized as code sprints according to an agile development methodology. Additionally, managing everything in a Gitlab pipeline ensures that all results are documented automatically. When finishing the work all the required information is there.

Following a comment from the audience, the panel moved to a discussion about ideas of how to incentivize creating reproducible experiment and measurement artifacts, keeping results available in the long-term, and covering the cost of performing artifact evaluation. One possible incentive that was mentioned was the ease of creating demos if the experiment is already developed in a reproducible way.

For the incentive of taking part as a reviewer, it was noted that motivation is regional and individual. Most people taking part in the artifact evaluation process are part of a single community, mostly associated with only a few professors. One proposed incentive to take part in the reviewing process is that the reviewer gets access to the paper and corresponding artifacts earlier than others.

Regarding the long-term availability of artifacts, the SLICES project became part of the ESFRI roadmap, a European program with the goal of creating long-term available scientific infrastructures. This will help to fund SLICES for roughly 20 years.

Next, a comment from the audience raised the problem that sometimes the fact that important parts of artifacts are missing is only noticed after the conference. Additionally, the idea was given that publishing a minimized usable version (e.g. using emulators such as Mininet) would make it easier to test results generally, even if performance results would not be comparable to measurements on a real testbed.

Regarding these points, an example for this issue was given: There was a paper at SIGCOMM, where the authors only were able to publish a subset of the measurement data. This does not fit with the current definition of reproducibility (accessible, documented, complete). The decision of how to handle these properties is subjective and dependent on the reviewer. There are open discussions on how these definitions should be interpreted.

Regarding the idea of awarding the artifact evaluation badge after the conference, the problem was raised that there would be less motivation to work on artifacts after the paper was already published. Having only the artifacts available badge before the conference serves as an additional incentive to not only publish artifacts but also create reproducible and documented measurement scripts later, as only having one badge might raise suspicion.

As a comment from the audience, the issue was raised that some measurements (like

wireless experiments) are unpredictable, making reproducibility difficult. Also, having long-running experiments makes repeating measurements time-consuming. Instead, one could create a system that testifies that the measurement results were not tampered with. In the following discussion, ideas were exchanged on how such a system could be implemented: Matching result binaries with digitally signed logs and timestamps created by the testbed was one idea. It was commented that if one follows a methodology similar to pos, experiment results are likely to be correct. pos could certify that experiment metadata is correct and that a specific code was running on the host. The execution and results of this software can not be certified. Additionally, it was added that if a user is determined to cheat and tamper with result data, it is very difficult to detect.

5 Networking in industrial scenarios

The first session on the second day was opened by a short introduction from the host and session chair, Georg Carle. This session comprised four talks covering methods and evaluations for networking in industrial scenarios. The addresses topics included approaches, architectures, and implementations for TSN. To conclude the session, the speakers took part in a panel discussion.

5.1 Dynamic Reservation of Ultra-Reliable Real-Time Streams in Time-Sensitive Networking

Lisa Maile (Friedrich-Alexander University – Erlangen-Nürnberg, Germany)

The session is opened by a talk entitled "Dynamic Reservation of Ultra-Reliable Real-Time Streams in Time-Sensitive Networking". Lisa Maile presents an overview of her research on dynamic flow reservation. During the discussion after the talk, the following topics were addressed:

QUESTION: Does the system work with existing standard TSN signaling or are extensions required?

ANSWER: For the centralized approach only the algorithms on the controller are not part of the standard. All the communication with the network is already present and well-defined (priorities, idle slopes). For the decentralized protocol, we showed that information, that is already distributed, is sufficient. Additionally, we showed that communicating information about the shaping of traffic to neighboring nodes can increase the throughput by about 20%. This would require a new signaling process.

QUESTION: Did you consider using more parallel queues to achieve better granularity in the arbitration and decrease the delay?

ANSWER: In our publication, we tested for up to 4 queues. As a user of our framework, it is possible to define the maximum number of queues. The heuristic then determines how many queues are actually needed.

QUESTION: What kind of sharpers did you evaluate?

ANSWER: Currently, only credit-based sharpers are implemented. Other sharpers like Deficit Round Robin or Strict Priority have similar characteristics.

QUESTION: Regarding the path selection in the online optimization: Will new flows impact the flows configured during the offline optimization?

ANSWER: The static flows are configured with the delay budgets, so there is a minimum reservation that will always be present. Therefore they don't change if new flows will be added. One idea, also considered in related work, is to use separate queues for static traffic and dynamic traffic. This is still an open research question.

QUESTION: Your framework is very generally applicable. Is there an application domain where it is most likely to be used?

ANSWER: It is applicable in applications where dynamic traffic occurs. Maybe in 5G Backends or Industrial Communication, where devices connect from different locations and require reconfiguring of the flows.

5.2 P4-PSFP: P4-Based Per-Stream Filtering and Policing for Time-Sensitive Networking

Steffen Lindner (Eberhard Karls University of Tübingen – Tübingen, Germany)

In the talk "P4-PSFP: P4-Based Per-Stream Filtering and Policing for Time-Sensitive Networking" by Steffen Lindner, the implementation of a mechanism to enforce sending patterns in TSN networks based on a programmable P4 switch was presented. The talk refers to previously published related work [ILM23]. The discussion after the talk covered the following topics:

QUESTION: Given the number of streams that are supported by the relatively powerful Tofino ASIC: Is this a good usage of hardware resources, or would a specialized ASIC for PSFP be more energy and silicon space-efficient?

ANSWER: The first question is if PSFP is needed in a 100G network. In application areas like factory automation with data rates in the range from $100\,\mathrm{Mbit\,s^{-1}}$ to $1\,\mathrm{Gbit\,s^{-1}}$ are common. With these lower data rates, it is definitely possible to rescale parts of the ASIC. Additionally, the presented flow limitation refers to ternary table entries. It is possible to match multiple streams with a single table entry, increasing the number of supported flows.

QUESTION: Could the number of supported flows easily be exceeded at data rates of $100\,\mathrm{Gbit\,s^{-1}}$? As shown in the presentation the number of supported flows may even be lower when performing more complex matches.

ANSWER: The presented limitations represent the amount of supported independent stream gates. With a large number of flows it would also be possible to bundle multiple streams using the ternary functions and performing the identification on a per-class basis. A hybrid form where only some flows are identified exactly, and others are grouped into classes could also be implemented.

5.3 OpenCNC: A software-defined control plane for automating TSN network configuration

Hamza Chahed (Karlstad University – Karlstad, Sweden)

The third talk with the title "OpenCNC: A software-defined control plane for automating TSN network configuration" is presented by Hamza Chahed. The talk gives an overview of OpenCNC, which is a tool that implements a control plan to automatically manage a TSN network. The following questions were raised during the discussion after the talk:

QUESTION: You presented three different approaches for TAS optimization (ILP, AI, and Metaheuristics). Are you implementing all three of them, or are you comparing them against existing solutions?

ANSWER: All three approaches were implemented for this work. A future task is to compare these implementations to other existing work.

QUESTION: What was the motivation or advantage to create a new ILP formulation instead of using existing implementations?

ANSWER: There are three different types of existing formulations: ILP with continuous variables and with discrete variables. A new approach relies on the ordering of frames and inferring exact sending times. For the last approach, which shows promising initial results, no existing implementations exist.

5.4 Containing Low Tail-Latencies in Packet Processing

Max Helm (Technical University of Munich – Munich, Germany)

The final talk of the third session was presented by Max Helm. In this talk titled "Containing Low Tail-Latencies in Packet Processing", he presents two approaches to model tail latencies. This talk is based on two publications [Wie+23][HC23]. After the talk, the following questions were discussed:

QUESTION: Regarding the network calculus part: How did you determine the curves?
ANSWER: The arrival curves are based on basic assumptions about the flow rate and rustiness. For the service curves, we assumed typical processing delays determined from previous measurements.

QUESTION: Did you consider different ways to determine the service curves from the measurements?
ANSWER: We did not consider this in this work. The used parameters are approximations.

Panel Discussion: Overcoming limitations of current approaches of quality and latency guarantees

The panel was opened with the following question outlining the scope of the following panel: *With our understanding of how we can ensure quality and latency guarantees, what are the limitations that we try to overcome?* First, it was remarked that all the existing tools, like Network Calculus and ILP optimization, work in theory but are too optimistic to work in reality due to effects like processing jitter or other network elements changing the traffic shaping. The question was posed, how to implement algorithms and implementations that allow meeting latency bounds, even without strict definitions of flows, while achieving high utilization, allowing the application of these methods for realistic traffic.

It was further added that latency bounds are violated mainly due to jitter. Variations in the sending time of a talker leading to a gap between theoretical predictions and practical results, were posed as an example. From this example, the need for more robust scheduling algorithms was derived.

Additionally, it was noted that current research is very specialized. For example, some research on TSN mostly ignores the effects of delays on the end devices and only considers delays introduced by the network. Advanced could be made by considering the complete path of a signal through a system. In relation to this comment, it was noted that the bounds determined by Network Calculus typically correlate with the maximum latency but are never true bounds. Therefore, statistical models might be better for some applications in practice, when applying them to standard networking hardware.

Next, a question from the audience guided the conversation towards certification. Addressing the talk by Max Helm, it was asked how the described system could be certified. It was noted that the goal of the described approach is to speed up extensive hardware measurements or simulations if results need to be within some error margin. It was further commented that, depending on the damage caused by a small number of late packets in a specific time interval, this approach could also be useful for certification. The applicability depends on the robustness of the overall system.

Additionally, in regard to the talk about OpenCNC, it was asked if it is possible to request information about processing delays or other information from the devices. This would be needed to perform robust scheduling. In the current state of OpenCNC, this information is not yet available.

Next, the discussion turned to the severity of delay introduced by required processing steps like tracking stream IDs. It was commented that in some scenarios, experiments showed that a single delayed or dropped frame could crash the complete network. This leads to the requirement to ensure that every single packet arrives correctly or to design the overall system to be robust against these packets. Additionally, the tradeoff between network utilization and robustness was emphasized.

A final remark from the audience stated that using network calculus bounds and accounting for very rate events requires a large amount of additional resources, which will remain unused in normal operation. The question was posed if there are any methods to avoid this overprovisioning of the network. Some ways to improve the network calculus bounds could be to model the service curve more precisely. Also, it would be interesting to combine or compare network calculus with other worst-case analysis approaches. Maybe different analysis approaches achieve better results, depending on the specific network layout and properties.

6 Real-Time Networking approaches and tools

The fourth session was chaired by Sebastian Gallenmüller. It comprised four talks showing new approaches and tools to evaluate latency and performance in TSN and networks. Afterward, the speakers joined a panel discussion covering the prediction of latency properties and network monitoring.

6.1 Performance Evaluation and Configuration in Time-Sensitive Networking

David Raunecker (University of Würzburg – Würzburg, Germany)

The first talk of the fourth session by David Raunecker entitled "Performance Evaluation and Configuration in Time-Sensitive Networking" gave an overview of the research done by their research group. Relevant publications regarding this topic include [Ste+23] and [Gri+20]. The discussion and questions following the talk touched on these topics:

QUESTION: Do you intend to implement RAP or even LRP for OMNNeT in INET? Or is the presented work just a high-level implementation of the logic of RAP?
ANSWER: This may be a future step. Currently, we focus on a high-level implementation, comparing the decentralized approach of our implementation to the centralized chameleon method developed at TUM. Concrete implementations are currently not planned.

QUESTION: Do you plan to implement frame duplication?
ANSWER: We do not currently have plans for frame replication. It could be incorporated into the current implementation, as it is built in a way that allows for extending the simulation, but there are no plans yet.

QUESTION: Is the simulator available as open-source code?
ANSWER: We did not publish the code yet, but plan to do so in the future.

QUESTION: Do you plan to realize this system on real hardware? What hardware would this implementation target?
ANSWER: Our current research covers high-level algorithms and approaches. In the short-term I am not aware if there is a plan to realize these systems on real hardware.

6.2 Containing Low Tail-Latencies in Packet Processing Using Lightweight Virtualization

Florian Wiedner (Technical University of Munich – Munich, Germany)

Next was a talk by Florian Wiedner, with the title "Containing Low Tail-Latencies in Packet Processing Using Lightweight Virtualization". He presented his research on achieving low-latency networking in Linux containers. The talk referred to the previously published paper [Wie+23]. After the talk, the following topics were discussed:

QUESTION: Are the measurement scripts available as open-source software?

ANSWER: We published the scripts and measurement data, which can be used to adapt the shown experiments to other testbeds. The exact measurement scripts rely on the pos orchestration service, which is currently not published.

QUESTION: The shown measurements use DPDK. Do you plan to use other technologies like XDP or AF-Packets?

ANSWER: We did not do that yet. We plan to do that in the future. XDP and AF-Packet provide a lot more configuration options and failure possibilities, making their analysis challenging.

QUESTION: Are you considering to investigate virtual TSN switches?

ANSWER: Yes, we are planning to do this as a next step.

QUESTION: Did you consider if the distribution of latencies changes depending on what you do (e.g., add an encapsulation)?

ANSWER: Currently, we are trying to keep out the effect caused by the application.

QUESTION: You mentioned that there was no difference between VM, bare-metal, and containers. What would you prefer?

ANSWER: There are minimal latency differences. With bare metal, there are fewer possibilities that boundaries will be exceeded. VMs can be used if the amount of available resources allows for it, while containers may be used if there are strict resource limits.

6.3 Simulation and Practice: A Hybrid Experimentation Platform for TSN

Marcin Bosk (Technical University of Munich – Munich, Germany)

In his talk "Simulation and Practice: A Hybrid Experimentation Platform for TSN", Marcin Bosk presented a hybrid experimentation approach allowing for measurements on hardware-based nodes and in a simulation-based environment using the same configuration. The talk is based on a published paper [Bos+23]. In the discussion after the talk, the following topics were discussed:

QUESTION: Do you know what could be the cause of the differences between hardware and simulation?

ANSWER: The processing delay is one of the biggest factors. The simulation assumes processing on the nodes happens instantly. Connections inside of the nodes are modeled as links with infinite bandwidth and zero delay.

COMMENT: I would suggest looking at error budget numbers in standardization, especially in the PHY part.

COMMENT: Hardware sometimes does weird stuff. You could try to measure what the network card actually sends. The NIC may cause additional buffering, leading to bursts in the transmitted traffic.

QUESTION: Why do you use iperf for traffic generation?

ANSWER: We wanted to generate packets in user space because we wanted to use traffic shaping from the Linux kernel. We also wrote our own traffic generator, where we can generate packets as expected by the time-aware shaper. We also considered using MoonGen or

DPDK, but this would require re-implementing some of the traffic shaping implementation.

6.4 Reproducible Experiments of Threshold Cryptography Functions and Trusted Execution Environments

Filip Rezabek (Technical University of Munich – Munich, Germany)

In the final talk of the fourth session, Filip Rezabek presented a talk with the title "Reproducible Experiments of Threshold Cryptography Functions and Trusted Execution Environments", describing the METHODA framework. It extends the existing EnGINE framework to support the assessment of capabilities and performance of TEE solutions in a reproducible and scalable way. The talk related to previously published work [Rez+23]. After the talk, the following points were discussed:

QUESTION: You already showed results for multiple nodes. Did you measure the impact of using a TEE on network performance on a single node?

ANSWER: We have already measured it in two scenarios. When measuring communication between two Kata containers, we observed additional delay on the server side caused by using a bridge instead of directly attaching the SR-IOV interface to provide networking to the container. We are looking into optimizing this problem. The second considered scenario is disk access. In this scenario, we only saw a negligibly small performance impact in the order of fractions of a percent.

Panel Discussion: How well can we predict latency properties of systems in support of time-sensitive applications?

The panel discussion started with an open question: *How well can we predict latency properties of systems in support of time-sensitive applications?*. Regarding this question, it was noted that it is possible to predict the latency with sufficiently accurate models. Verifying the theoretical insights with measurements on real hardware or simulated experiments is also important. It was further noted that with a sufficiently precise latency model, taking into account all hardware effects, it would be possible to accurately predict a system's behavior. To be able to define such a model, it is important to monitor the behavior of real-world systems and adapt models accordingly.

The importance of monitoring rare events, which have a lot of influence on systems but are typically ignored in models and simulations, was also mentioned. It was noted that these events are hard to predict and hard to measure. Additionally, accommodating these rare events costs additional resources. Also, understanding the cause of these events is required, as some effects might also be limited to the measurement setup and do not occur in real systems. This led to the comment that a model can not cover every aspect of a system, and some effects need to be abstracted.

Next, a question from the audience regarding the importance of network monitoring and updating predictions accordingly was posed. It was commented that monitoring the latency guarantees is not useful in systems where even a single packet breaking a latency bound is not acceptable. One way to mitigate this issue could be to introduce additional redundancy, like

parallel paths, requiring additional resources. Monitoring should only complement guaranteed latency bounds and redundant system elements. Additionally, it is essential to make sure that monitoring tools do not influence the tested system while being able to react to events fast enough.

Using monitoring makes it possible to gain insights that are not feasible in an experiment setup, e.g., when running for long measurement durations. With larger network setups, methods that do not influence the measured system, as presented in the talk by Florian Wiedner, are not applicable. This comment led to the question of how a model, initially defined by data gathered using a precise experiment setup, can be updated and adjusted based on less precise monitoring data.

Next, the discussion moved to the question of future hardware improvements and features and their impact on the implementation of TSN solutions. One possible improvement, which was noted, was to increase the variety of available TSN hardware, like switches and NICs, to be able to distinguish between the effects of specific hardware and effects caused by TSN protocols. Also, additional monitoring capabilities and better, more accurate documentation on the network and software level would help to gain a deeper insight. It was further noted that increasing the scalability of NICs, e.g., adding support for TSN functionality inside of virtual functions or elevating limits on the number of virtual interfaces, would allow for precise and simpler experiment setups.

7 Wrap-Up

The academic salon ended with some final remarks and comments by Georg Carle, thanking speakers and participants for their contributions. Additionally, feedback on the event organization and program was discussed to be considered for possible future events.

7.1 Final Remarks

Feedback for the event and its organization was very positive. A follow-up event was suggested by several participants.

References

[Bos+23] Marcin Bosk et al. “Simulation and Practice: A Hybrid Experimentation Platform for TSN.” In: *IFIP Networking Conference, IFIP Networking 2023, Barcelona, Spain, June 12-15, 2023*. IEEE, 2023, pp. 1–9. DOI: 10.23919/IFIPNETWORKING57963.2023.10186364. URL: https://doi.org/10.23919/IFIPNetworking57963.2023.10186364.

[Car+21] Georg Carle et al. *Academic Salon on Time-Sensitive Networking and Deterministic Applications*. Oct. 2021. URL: https://net.in.tum.de/events/academic_salon_21.html.

[Car+22] Georg Carle et al. *Academic Salon on Low-Latency Communication, Programmable Network Components and In-Network Computation*. Sept. 2022. URL: https://net.in.tum.de/events/academic_salon_22.html.

[Gal+21] Sebastian Gallenmüller et al. “The pos framework: a methodology and toolchain for reproducible network experiments.” In: *CoNEXT ’21: The 17th International Conference on emerging Networking EXperiments and Technologies, Virtual Event, Munich, Germany, December 7 - 10, 2021*. Ed. by Georg Carle and Jörg Ott. ACM, 2021, pp. 259–266. DOI: 10.1145/3485983.3494841. URL: https://doi.org/10.1145/3485983.3494841.

[Gri+20] Alexej Grigorjew et al. “Bounded Latency with Bridge-Local Stream Reservation and Strict Priority Queuing.” In: *11th International Conference on Network of the Future, NoF 2020, Bordeaux, France, October 12-14, 2020*. Ed. by Prosper Chemouil et al. IEEE, 2020, pp. 55–63. DOI: 10.1109/NOF50125.2020.9249224. URL: https://doi.org/10.1109/NoF50125.2020.9249224.

[HC23] Max Helm and Georg Carle. “Predicting Latency Quantiles using Network Calculus-assisted GNNs.” In: *Proceedings of the 2nd on Graph Neural Networking Workshop 2023*. Paris, France: Association for Computing Machinery, Dec. 2023, 13â€“18. ISBN: 9798400704482. DOI: 10.1145/3630049.3630173.

[ILM23] Fabian Ihle, Steffen Lindner, and Michael Menth. *P4-PSFP: P4-Based Per-Stream Filtering and Policing for Time-Sensitive Networking*. 2023. arXiv: 2311.07385 [cs.NI].

[LHM23] Steffen Lindner, Marco Häberle, and Michael Menth. “P4TG: 1 Tb/s Traffic Generation for Ethernet/IP Networks.” In: *IEEE Access* 11 (2023), pp. 17525–17535. DOI: 10.1109/ACCESS.2023.3246262. URL: https://doi.org/10.1109/ACCESS.2023.3246262.

[Rez+23] Filip Rezabek et al. “Multilayer Environment and Toolchain for Holistic NetwOrk Design and Analysis.” In: Nov. 2023.

[Sca+23] Mariano Scazzariello et al. “A High-Speed Stateful Packet Processing Approach for Tbps Programmable Switches.” In: *20th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2023, Boston, MA, April 17-19, 2023*. Ed. by Mahesh Balakrishnan and Manya Ghobadi. USENIX Association, 2023, pp. 1237–1255. URL: https://www.usenix.org/conference/nsdi23/presentation/scazzariello.

[SGC23] Manuel Simon, Sebastian Gallenmüller, and Georg Carle. “Never Miss Twice - Add-On-Miss Table Updates in Software Data Planes.” In: *KuVS Fachgespräch - Würzburg Workshop on Modeling, Analysis and Simulation of Next-Generation Communication Networks 2023 (WueWoWAS’23)*. Würzburg, Germany, June 2023, p. 5. DOI: 10.25972/OPUS-32207.

[Sim+21] Manuel Simon et al. “High-Performance Match-Action Table Updates from within Programmable Software Data Planes.” In: *4th P4 Workshop in Europe (EUROP4)*. Virtual Event, Dec. 2021. DOI: 10.1145/3493425.3502759.

[Ste+23] Geißler Stefan et al. “DARTA: Generation of Autocorrelated Random Numbers using Discrete AutoRegression To Anything.” In: *35th International Teletraffic Congress (ITC-35)*. Torino, Italy, Sept. 2023, pp. 1–9.

[Wie+23] Florian Wiedner et al. “Containing Low Tail-Latencies in Packet Processing Using Lightweight Virtualization.” In: *2023 35rd International Teletraffic Congress (ITC-35)*. Turin, Italy, Oct. 2023.