

Liebe ITG-Mitglieder,

aus den Berichten zu Fachtagungen ist besonders der Start einer neuen Fachtagungsreihe zu „Internet of Things/Industrie 4.0“ hervorzuheben, die in ihrer ersten Tagung einen vielversprechenden interdisziplinären Ansatz verfolgte.

In mehreren Beiträgen widmet sich diese ITG-news den neuen Möglichkeiten von Edge Cloud Services. In dem Funktionen rein softwaremäßig abgebildet werden, können Dienste sehr schnell bereitgestellt und gestartet werden. So können zum einen die im Netz insgesamt verfügbaren Computing-Ressourcen optimal ausgelastet werden und zum anderen Funktionen gezielt zur Erfüllung kritischer Dienstanforderungen (z.B. Latenzzeiten) direkt auf Basis von SDN auf Netzwerkgeräten ausgeführt werden.

Souveränität im Forschungsbetrieb stellt andere Anforderungen als Souveränität in der Produktion oder im Betrieb.

Mit der Ausarbeitung eines Positionspapiers zur „technologischen Souveränität“ möchte die ITG sich in die aktuelle Debatte, wie „souverän“ wir in der Entwicklung und Nutzung zukünftiger Technologien sein wollen, einbringen. Aus Sicht der ITG wäre es wünschenswert, für unseren Wirtschaftsstandort relevante Technologiefelder systematisch identifizieren und vor allem aber auch bewerten zu können. Das, was unter „souveränem Handeln“ verstanden wird, unterscheidet sich deutlich, je nachdem, welche Rolle jemand in der Wertschöpfungskette einnimmt. Souveränität im Forschungsbetrieb stellt andere Anforderungen als Souveränität in der Produktion oder im Betrieb. Das Positionspapier entwickelt darauf aufbauend ein Konzept



zur Systematisierung. Als Koordinator dieses Positionspapiers nehme ich die Gelegenheit dieses Editorials wahr, Sie in Abstimmung mit dem Vorsitzenden, Herrn Prof. Schotten, über diese Initiative zu informieren.

**DR.-ING. KLAUS ILLGNER-FEHNS**

ITG-Vorstandsmitglied

### VDE|ITG-POSITIONSPAPIER

## Zur technologischen Souveränität

Die Informationstechnische Gesellschaft im VDE (ITG) hat ein neues Positionspapier „Zur technologischen Souveränität: Ein Beitrag zur Systematisierung“ erarbeitet. Ziel des Positionspapiers ist, u.a. einen Vorschlag zur Systematisierung des Begriffs „Souveränität“ zu machen, der sich

aus den Anforderungen entlang der Wertschöpfungskette und relevanten Technologiefeldern ableitet.

Weitere Informationen dazu finden Sie auf // [www.vde.com/de/itg](http://www.vde.com/de/itg)

## Inhalt

<b>Meldungen</b>	<b>02</b>
<b>Thema</b>	
Edge Cloud Services	<b>04</b>
<b>Personalia</b>	<b>12</b>
<b>Termine</b>	<b>12</b>

## 1. FACHTAGUNG „INTERNET OF THINGS/INDUSTRIE 4.0“

# Aktuelle Technologien und Anwendungen des Machine Learnings

Vom 27. bis 28. September 2018 fand in Dresden die 1. Fachtagung „Internet of Things/Industrie 4.0“ statt.

Die 1. Fachtagung „Internet of Things/Industrie 4.0“ wurde auf Initiative der Mitglieder des gleichnamigen VDE-ITG-Fachausschusses KT6 ins Leben gerufen und unter Leitung von Dr. Jörg Benze vom VDE-Bezirksverein Dresden in den Räumen der T-Systems Multimedia Solutions GmbH veranstaltet. Die Fachtagung will eine interdisziplinäre Plattform für Industrie und Beratungsunternehmen, Hochschulen, Forschungseinrichtungen und die öffentliche Verwaltung zur Diskussion aktueller Fragestellungen sein. Neben den technischen Aspekten werden auch die durch den digitalen Wandel zu erwartenden juristischen und gesellschaftlichen Veränderungen betrachtet. Der Know-how-Austausch zwischen den verschiedenen Branchenspezialisten aus den Ingenieurwissenschaften, aus Ökonomie, Recht und öffentlicher Verwaltung soll weiter gefördert werden. Zwölf Referenten diskutierten mit den Teilnehmern in insgesamt zwei Keynotes und vier Sessions lebhaft folgende Schwerpunkte:

- Geschäftsmodelle (Ökosysteme) für IoT und Industrie 4.0,
- IoT-Anwendungen, juristische und gesellschaftliche Aspekte,
- Kommunikation und Security als Rückgrat für IoT und Industrie 4.0,
- Industrie 4.0 in Theorie und Praxis.



Eine gesamtheitliche Betrachtung konnte durch die anwesenden hochkarätigen Referenten verschiedenster Fachrichtungen den Teilnehmern dargestellt werden. Es sind technologische Entwicklungen, erfolgreiche neue Geschäftsmodelle und anhand exemplarischer Praxisbeispiele erfolgreich gegangene Realisierungswege vorgestellt worden. Eine Paneldiskussion am Ende jedes Tages gab den Teilnehmern reichlich Gelegenheit,

mit den Referenten ausführlich zu diskutieren.

Mit 35 Teilnehmern war diese erste Fachtagung gut besucht und wurde mit einem gemeinsamen Abendessen mit Blick auf das Dresdner Altstadt-Panorama von der Canaletto-Terrasse des Hotels Bellevue abgerundet.

### DR.-ING. JÖRG BENZE

T-Systems Multimedia Solutions GmbH



Alle Fotos: VDE-Bezirksverein Dresden e.V.

Die 1. Fachtagung „Internet of Things/Industrie 4.0“ fand in Dresden in den Räumen der T-Systems Multimedia Solutions GmbH statt; das Abendessen auf der Terrasse des Hotels Bellevue bot einen Panoramablick auf die Dresdner Altstadt

## Tradition trifft Innovation

Vom 11. bis 14. Februar 2019 fand die „International ITG Conference on Systems, Communications and Coding (SCC 2019)“ an der Universität Rostock statt. Im 600. Jubiläumsjahr der Universität trafen sich Wissenschaftler aus 15 Nationen, um über aktuelle Trends der Kommunikationstechnik zu diskutieren.

Anlässlich des 600-jährigen Bestehens der Universität Rostock wurde die diesjährige SCC vom 11. bis 14. Februar 2019 in Rostock ausgetragen. Über 100 Wissenschaftler aus 15 Nationen tauschten in der festlichen Aula des Hauptgebäudes wissenschaftliche Ergebnisse aus. Auch die SCC selbst feierte ein kleines Jubiläum. Sie startete 1994, also vor 25 Jahren, als deutsche ITG-Fachtagung „Codierung für Quelle, Kanal und Übertragung“. Aufgrund der rasanten Entwicklung dieses Gebiets wurde sie im Jahr 2000 zur „International Conference on Source and Channel Coding“. 2013 verbreiterte sich das Themenspektrum und es erfolgte die Umbenennung auf den heutigen Namen. Mittlerweile hat sich die „International ITG Conference on Systems, Communications and Coding“ in der Fachwelt fest etabliert und repräsentiert eine wichtige in Deutschland beheimatete Konferenz mit internationaler Ausstrahlung. Sie wird vom ITG-Fachausschuss KT 1 „Informations- und Systemtheorie“ organisiert und findet alle zwei Jahre statt.

Die SCC 2019 begann am Montag, dem 11. Februar, mit zwei Tutorials zu den Themen „The Information Bottleneck Method in Communications“ und „Communications and Control“, bevor am Abend eine Welcome Reception die Konferenz eröffnete. Die drei folgenden Tage gliederten sich in 15 Sessions, die elf eingeladene, 25 reguläre Vorträge sowie 28 Poster-Beiträge umfassten. Insbesondere die langen Poster-Sessions gaben vielfältige Diskussionsmöglichkeiten und wurden von den Teilnehmern intensiv genutzt. Die SCC ist daher auch für junge Wissenschaftler mit innovativen Ideen eine gefragte Plattform. Das Programm wurde am Mittwoch mit einem Bankett im historischen Lokschuppen abgerundet.

Mit der SCC wird versucht, eine Brücke zwischen Grundlagenforschung und ihren Anwendungen in der Kommunikationstechnik zu schlagen und Universitäten wie auch Industrie gleichermaßen anzusprechen. Wichtige Schlüsselthemen wurden durch elf eingeladene Beiträge international renommierter Wissenschaftler präsen-

tiert und gaben den Rahmen des technischen Programms vor. In der optischen Übertragungstechnik wurden beispielsweise Möglichkeiten zur Optimierung der Signalraumkonstellationen und die nichtlineare Fourier-Transformation erörtert. Weitere Beiträge thematisierten die Probleme der Post-Quantum-Kryptographie und stellten Verfahren für den inhaltlich anonymisierten Zugriff auf Datenbanken vor, welche der Datenbank keine Informationen über die tatsächlich gesuchten Inhalte geben. Ferner wurde der Einsatz von maschinellem Lernen in der Kommunikationstechnik thematisiert. Neben Diskussionen über das Potenzial neuronaler Netze wurden auch neue Clustering-Verfahren zur effizienten Implementierung von Algorithmen vorgestellt. Etliche weitere Beiträge beschäftigten sich mit speziellen Fragen der Kanalcodierung in Kommunikationssystemen.

Insbesondere die fünfte Mobilfunkgeneration gilt als wichtige Basistechnologie für die automatisierte Industrie 4.0, das autonome Fahren und das Internet der Dinge (Internet of Things, IoT). Aus diesem Grund wurde eine Panel-Diskussion zum Thema 5G mit Vertretern der Industrie von Prof. Stephan ten Brink organisiert. Sie kam bei den Teilnehmern sehr gut an und offenbarte viele offene Fragestellungen, die als interessante Forschungsthemen in den nächsten Jahren zu klären sind. Die Einbeziehung des Publikums zeigte ein durchaus heterogenes Meinungsbild über zukünftige Trends in diesem Bereich.

Weitere Informationen und Details zum Programm sind unter <https://scc2019.net> zu finden. Die nächste SCC wird 2021 in Ulm stattfinden und von Prof. Robert Fischer organisiert werden.

### PROF. DR.-ING. VOLKER KÜHN

Institut für Nachrichtentechnik  
Fakultät für Informatik und Elektrotechnik  
Universität Rostock



Vortrag in der festlichen Aula der Universität Rostock

# The Serverless Computing Paradigm

*Serverless computing offers a viable option for a vast range of Edge Cloud services.*

Communication networks need to be both adaptive and scalable to support cloud services that extend to the edge of the network. The last few years have seen an explosive growth of software-defined networking (SDN) and network function virtualization (NFV) to address this need. Both technologies enable services to be decoupled from hardware, so that software functionality is no longer constrained by the underlying hardware and can evolve independently. SDN and NFV aim to virtualize the entire network. Services are implemented in software and deployed across a distributed cloud computing infrastructure that extends to the network edge. Achieving this goal requires a careful consideration of the computing paradigm used by these virtualized services to successfully transition to software-based networks with the needed agility, robustness and scalability.

## The evolution of a new cloud computing paradigm

Several virtualization technologies have been developed in the cloud domain, such as Virtual Machines (VMs), containers and Unikernels. Each technology has influenced the design, development and operation of applications.

**Cloud computing** uses virtual machines (VMs) to run customers' operating systems (OS) in isolation. A hypervisor running on the actual physical machine manages the sharing of the hardware resources among VMs, allowing multiple VMs to be run on the same host. Legacy applications running on common servers can be quickly onboarded, which simplifies the adoption of the cloud computing paradigm. Application developers need to estimate the size and the number of virtual machines required

to meet service demand. Since VMs need to boot the guest OS when they start and typically comprise several application modules in a single instance, they tend to have long start-up latencies. The setup cost of an instance amortizes with operation time, making VMs well suited for long running services but also sluggish. Hence, cloud applications are often overprovisioned to make sure they can meet sudden demand spikes. Scaling VMs requires continuous monitoring and often complicated pattern prediction to get the scaling actions right.

**Cloud-native computing**, or the microservice architectural style, is a paradigm that regards the application as a set of services similar to service-oriented architecture (SOA). The use of containers allows developers to package services with their runtime environment. Unlike VMs, containers do not simulate hardware but have the components share an underlying OS kernel of the host. Essentially, the services run as processes, with an additional layer of protection for resource isolation. The memory footprint can be much smaller than that of a VM, which enables containers to start an order of magnitude faster. Since containers have overlapping fault domains, Unikernels [1] have more recently gained attention from the community

to leverage hardware virtualization mechanisms. Unikernels are tiny runtime environments where a target application is statically linked to a minimalistic library operating system. This allows the application to run directly on the virtual hardware abstraction of a hypervisor. Unikernels have similar isolation properties as VMs, while being lighter weight because of the reduction in the memory footprint and system initialization. The cloud-native service paradigm still requires the application developer to estimate re-

source requirements and to monitor instances continuously to scale the number of service instances accordingly.

**Serverless computing** has attracted significant interest from both industry and academia as a new paradigm since the launch of Amazon Lambda in 2014. With Serverless computing, application developers no longer need to manage servers (hence, the name), guess the resource demand or scale the application – these tasks are delegated to the Serverless platform. One popular realization of Serverless computing paradigm is the Function-as-a-Service (FaaS) model, where the applications are written as individual functions that can be separately triggered by various types of events (e.g. external web requests, internal database triggers). Upon each request, the platform allocates an ephemeral runtime environment for executing the functions. Finer granularity enables the platform to schedule each execution, making resource allocations responsive to load variations and applications inherently scalable. As resources are released after execution, the Serverless platform can avoid idle instances and operate more resource-efficiently.

## Advantages and challenges of Serverless computing

The new paradigm promises less orchestration overhead and more cost reductions. Managing and provisioning server resources for applications currently presents a significant overhead and responsibility for application developers. Serverless computing alleviates the problem of right-sizing allocations by shifting the responsibility to the platform, which allows developers to focus on their application logic, while the platform is enabled

to better leverage its infrastructure and jointly optimize for application performance and resource efficiency.

In the cloud resource rental model, resources are provisioned even when they remain unused. With Serverless computing, applications only occupy resources when they are being executed by the platform, thus being very resource efficient. In commercial scenarios, customers are charged based on how many resources are utilized by their applications and there is no cost resulting from idle resources.

Despite these advantages, there are considerations to make for applications that employ Serverless computing. The primary concern is the start-up latency per invocation to handle a request. Since the platform allocates an ephemeral runtime to execute a function, the time spent for preparing this environment could be in the 10s to 100s of milliseconds, or even longer, depending on the virtualization technology used and whether the function invocation happens to incur a “warm” or a “cold” start. For short-lived functions, these start-up latencies impose a significant overhead. Furthermore, if a sequence of multiple functions is executed to handle a request, these start-up latencies accumulate and incur a significant delay to the application response time.

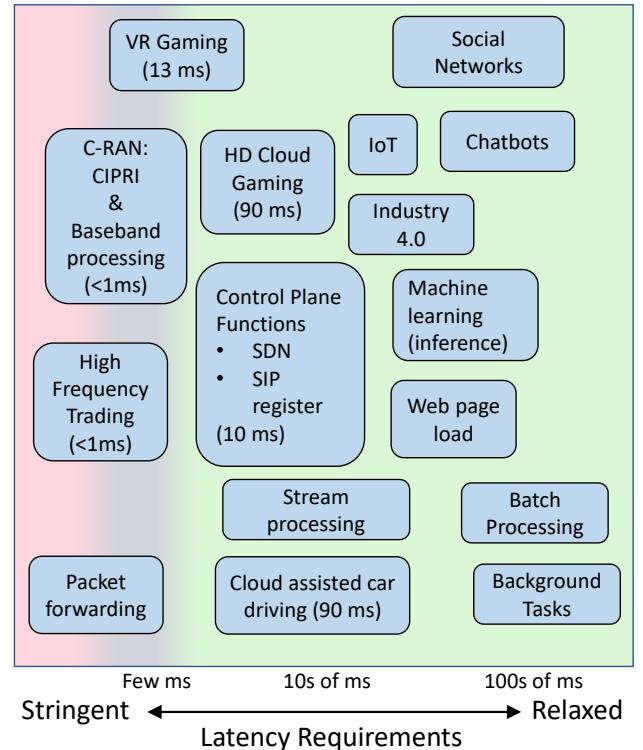
The applications well-suited for Serverless computing depend on their latency requirements, which may range from sub-milliseconds for applications such as packet forwarding and high-frequency trading, to latencies beyond the 100-millisecond range for applications such as batch or asynchronous event processing. Applications that do not have very strict latency requirements, in the order of a few 10s of milliseconds or more, can potentially be implemented in a Serverless manner. These include stream processing, SDN control plane functions, background tasks, IoT-related tasks, etc. On the other hand, applications with very strict latency budget, such as online gaming and augmented reality, may not work well on existing Serverless platforms. Recent approaches [2, 3] for high-performance Serverless computing can push the latency boundary and broaden its use and help more developers

benefit from it. The SAND platform [2] uses application-level sandboxing and a hierarchical message bus, which outperforms the state-of-the-art serverless platforms significantly. For Example, in a commonly-used image processing application, SAND achieves a 43 per cent speedup compared to Apache OpenWhisk.

Another important aspect to be considered is the workload characteristics. Because of the event-driven execution nature, Serverless computing is especially well-suited for applications that have variable workloads. The platform utilizes resources in fine increments to match the exact demand for the function. In the case the demand for a function experiences a sudden increase, the platform automatically creates new instances to handle the increased demand. In doing so, the Serverless paradigm is well suited to handle load spikes, where traditional approaches would over-provision to meet peak demand. The more dynamic the workload of a task, the higher the benefits from adopting the Serverless computing paradigm.

The ephemerality of the execution environment poses a challenge to applications with large contexts. Between executions, the application state needs to be persisted outside the runtime environment. This entails two options to preserve the state: either the state is carried within the message passing between function executions or the state is persisted in a storage system, such that each invocation can fetch its context, operate on it, and then store it at the end of the execution. Many Serverless applications operate in conjunction with a storage service using this principle, which are often part of the Cloud service ecosystem.

The principles to decompose applications for fine-granular development, to pass messages for distributed execution and to persist state for failover are not new design choices for cloud



applications. But whether the application can still operate at a reasonable performance depends largely on the platform’s design. Runtime environment start-up latencies, application state access or an otherwise stable workload define the granularity of the application decomposition at which it can operate economically in the Serverless paradigm. A component with stringent latency requirements may not allow for further decomposition, e.g. real-time video encoder, in which functions need to pass a context with high-resolution frames to the next function, is such an example. The ability to reduce overheads becomes important for the Serverless platform to push the boundaries to broaden its applicability.

## Distributed services in transition to a Serverless paradigm

**Software-Defined Networking (SDN)** transforms network management by separating the network’s control plane from the routers and switches that forward traffic. With such a separation, network elements become simple forwarding devices, and the control logic is written as a distributed service yet logically centralized, which simplifies policy enforcement, network

configuration and service evolution [4]. Packet handling has latency requirements that cannot be met by today's Serverless computing technologies. On the other hand, the SDN controller is a prime candidate. It exercises direct control over switches via a well-defined interface. The event-based control loop allows the controller to manage routing, switching, fire-walling, or other network functions (e.g. load balancing, traffic shaping). Controllers also serve a northbound interface to carry out network application requests, typically formulated as intents that describe the communication properties between multiple network endpoints. As the SDN controller serves events from both southbound and northbound APIs, the rate at which these events arrive can vary significantly. Hence, considerable effort has been dedicated to engineering controllers as highly-concurrent, multi-threaded systems to scale up to enterprise-class networks. These implementations aim at increased parallelism by minimizing synchronization needed between the modules via various mechanisms. Furthermore, SDN controllers are designed for distributed deployment to make them scalable and failure resilient. In order not to risk operability, pluggable modules need to be isolated from one another. These requirements make SDN controllers a great fit for Serverless computing, given that the platform meets its latency requirements.

**Traffic Anomaly Detection** as proposed by Kostas et al. [5] suggests a modular design for distributed denial-of-service (DDoS) mitigation that exploits the network programmability of SDN within the context of NFV. The computation steps required for anomaly detection are decoupled from the data plane on the edge router and divided into a few computation tasks that are executed in a chain, each with increasing computational requirements. When an attack is detected, an on-demand network function instructs the edge router to forward all traffic destined to the victim to another switch, which then filters malicious traffic while preserving benign flows. More specifically, the first step (i.e., statistics collector) harvests statistics from the packet samples, and periodically

exports them to the second component for further investigation. The second stage performs coarse-grained, entropy-based anomaly detection on these statistics, and invokes the third step only if an anomaly is detected. The third component (i.e., heavy-weight anomaly detection) uses a computationally heavy bidirectional count sketch algorithm to verify the attack, and then identifies the victim (i.e., victim identification). Finally, the traffic manipulation function is triggered to update the routing table of the edge router to redirect traffic away from the victim. These computation tasks required for anomaly detection are sporadic and event-driven. Such a workload could be encountered using the Serverless paradigm, which would make these instantiations resource-efficient and flexible.

**Media Processing** has been targeted by the 5G Media project [6] to use Serverless VNFs for carrying out short tasks running in parallel with the container-based VNFs. One such short task is "repeat clip creation" for online spectators that are watching a live game being played between two players in a virtual world. A spectator may wish to replay a certain portion of the game from a specific camera angle. This task requires collecting the necessary buffered game play data and rendering it from one or more camera angles desired by a specific spectator. The task can be instantiated anytime during the game play. Although it may not be as time-sensitive as the game play itself, multiple instances of this task may need to be created quickly to handle a sudden increase of spectators wishing to replay a particular portion of the game (e.g. a goal event in a soccer game). These workload properties make it a good fit for Serverless computing.

### Serverless is here to stay

Services that are tied to the data-plane such as media processing requires addressable, long-running instances that current Serverless platforms can't provide. But separating the control functions and deploying them on a Serverless platform allows the application to efficiently orchestrate while consuming resources purely on-demand. The

Serverless computing paradigm is conceptually a great fit for implementing SDN controllers, and practically Serverless computing also offers the scalability and resource isolation needed by any distributed Edge Cloud service. Latency-critical functions are currently not suited for scaling by each request, but research efforts show a promising improvement on the latency overheads of Serverless platforms. Serverless computing is here to stay and offers a viable option for a vast range of Edge Cloud services to reap the benefits of commonly automated platform operations.

### References

- [1] Madhavapeddy, A.; Scott,D. J.: Unikernels: the rise of the virtual library operating system. Communications of the ACM, 2014
- [2] Akkus, I. E.; Chen, R.; Rimac, I.; Stein, M.; Satzke, K.; Beck, A.; Aditya, P.; Hilt, V.: SAND: Towards High-Performance Serverless Computing. In: 2018 USENIX Annual Technical Conference (USENIX ATC 18), 2018
- [3] Oakes, E.; Yang, L.; Zhou, D.; Houck, K.; Harter, T.; Arpacı-Dusseau, A. C.; Arpacı-Dusseau, R. H.: SOCK: Rapid Task Provisioning with Serverless-Optimized Containers. In: 2018 USENIX Annual Technical Conference (USENIX ATC 18), 2018
- [4] Kim, H.; Feamster, N.: Improving network management with software defined networking. IEEE Communications Magazine 51 (2013), pp. 114–119
- [5] Kostas,G.; George, A.; Vasilis, M.: A scalable anomaly detection and mitigation architecture for legacy networks via an OpenFlow middlebox. Security and Communication Networks 9, pp. 1958–1970
- [6] Programmable edge-to-cloud virtualization fabric for the 5G Media industry
- [7] Shin,S.; Song,Y.; Lee,T.; Lee, S.; Chung, J.; Porras, P.; Yegneswaran, V.; Noh, J.; Kang, B. B.: Rosemary: A Robust, Secure, and High-performance Network Operating System. In: 2014 ACM SIGSAC Conference on Computer and Communications Security (ACM CCS 14), 2014

**M.SC. MANUEL STEIN**  
**DR. PAARIJAAT ADI-TYA**  
**DR. ISTEMI EKIN AKKUS**  
**DR. RUICHUAN CHEN**  
**DR. IVICA RIMAC**  
**DR. VOLKER HILT**

Nokia Solutions and Networks GmbH & Co. KG,  
Stuttgart

# Flexible Netzwerkarchitekturen für das Internet of Production

Dem Internet of Production liegt die Vision zugrunde, die großen Mengen an Daten, die in der Produktions-technik vorhanden sind, einfach zugänglich, interpretierbar und vernetzbar zu machen, sodass daraus wertvolles Wissen abgeleitet werden kann. Hierbei können Edge Cloud Services einen wichtigen Beitrag leisten, um die zu kommunizierenden Datenmengen beherrschbar und die echtzeitfähige Steuerung von Produktionsprozessen möglich zu machen.

Konzepte wie das Industrielle Internet der Dinge und Industrie 4.0 haben die Erfassung und Verarbeitung von Produktionsdaten zur Automatisierung von Produktionsprozessen sowie die Entwicklung und den Einsatz autonomer Kontrollmechanismen zur Verbesserung der Systemsicherheit und Zuverlässigkeit zum Ziel [1]. Die bisher vorgeschlagenen Konzepte ziehen jedoch hauptsächlich auf spezifische lokal begrenzte Verbesserungen in isolierten Bereichen. Im Gegenzug verfolgt das Internet of Production [2] das Ziel, eine domänen- und unternehmensübergreifende Zusammenarbeit zu ermöglichen, indem Daten aus Produktion, Entwicklung und Nutzung interessierten Parteien in Echtzeit, einer angemessenen Granularität und potenziell in globalem Maßstab zur Verfügung gestellt werden. Dazu soll die Isolierung von Daten, die von Produktionsmaschinen erzeugt werden, überwunden sowie Maschinen und Produktionsstätten auf der ganzen Welt in die Lage versetzt werden, Informationen über eine einzige globale Infrastruktur miteinander auszutauschen.

## Kommunikation im Internet of Production

Die präzise und funktional sichere Steuerung von Produktionsmaschinen erfordert eine höchst zuverlässige und latenzarme Kommunikation zwischen den beteiligten Geräten (Sensoren, Steuerungsgeräten, Aktoren). Zum Erlernen und stetigen Verbessern digitaler Modelle von Produkten und deren Produktionsprozessen wird jedoch die Datenmenge, welche nicht nur zwischen verschiedenen Teilen

einer Produktionslinie, sondern auch zwischen Linien und den jeweiligen Steuerungseinheiten und weiteren übergeordneten Systemen ausgetauscht wird, in Bezug auf Frequenz, Anzahl und Größe stark ansteigen müssen. In diesem Spannungsfeld werden traditionelle Fabriksteuerungsarchitekturen und -netzwerke jedoch an ihre Skalierbarkeitsgrenzen stoßen, wie in Bild 1 dargestellt: Während hohe Datenmengen und Modellkomplexitäten eine zentralisierte Verarbeitung auf (Netzwerken von) speicherprogrammierbaren Steuerungen erschweren und daher an Cloud Services im Internet ausgelagert werden müssen, ist eine Kontrolle der Maschinerie über das Internet aufgrund der Unzuverlässigkeit der Verbindung (Paketverluste, Jitter) nicht möglich. Um das Internet of Production verwirklichen zu können, müssen daher neue, flexiblere Architekturen, die eine horizontale Skalierung und damit eine Verteilung der Verarbeitungslast im Netzwerk selbst ermöglichen, entwickelt werden.

## Vorteile von Edge Services

Um sowohl den Latenzanforderungen als auch den stark steigenden Datenmengen Herr zu werden, erlauben es Edge Cloud Services, Datenverarbeitung und Kontrolle wesentlich

näher an die Produktionslinie zu bringen, als dies bei traditionellen cloud-basierten Diensten möglich ist.

Dabei lassen sich, wie in Bild 2 dargestellt, zwei Arten von Edge Cloud Services unterscheiden:

- In-Network Edge Services, welche leichtgewichtige Aufgaben, wie die Datenvorverarbeitung und kleinere Steuerungsroutinen, direkt in dem Kommunikationsnetzwerk durchführen können, sowie
- klassische Edge Services für speicher- und rechenintensivere Aufgaben.

## In-Network Edge Services

Um In-Network Edge Services zu realisieren, ermöglicht Software-defined Networking, mit Open-Flow [3] und P4 [4] als prominenten Protokollen und Programmiersprachen, Verarbeitungsregeln mit niedriger Latenz direkt auf Netzwerkgeräte, wie Router und Switches, zu übertragen, um z.B. eine Vorfilterung von Sensordaten oder Weiterleitungsrichtlinien effizient

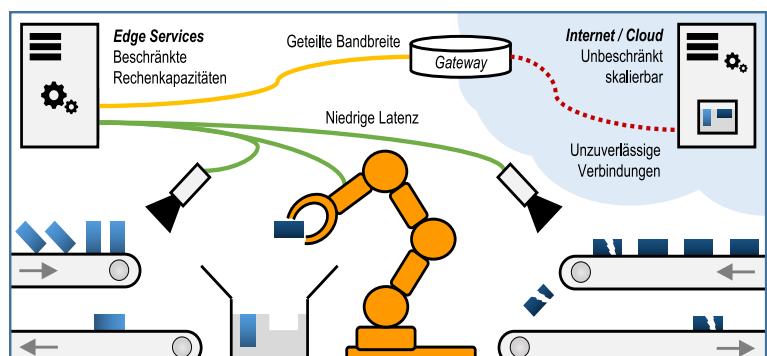


Bild 1: In künftigen Produktionsszenarien werden niedrige Latenz- und hohe Zuverlässigkeitserfordernisse der Maschinensteuerung und -überwachung mit den quasi unbegrenzten, aber nicht zuverlässig erreichbaren Rechen- und Speicherressourcen der Cloud konkurrieren

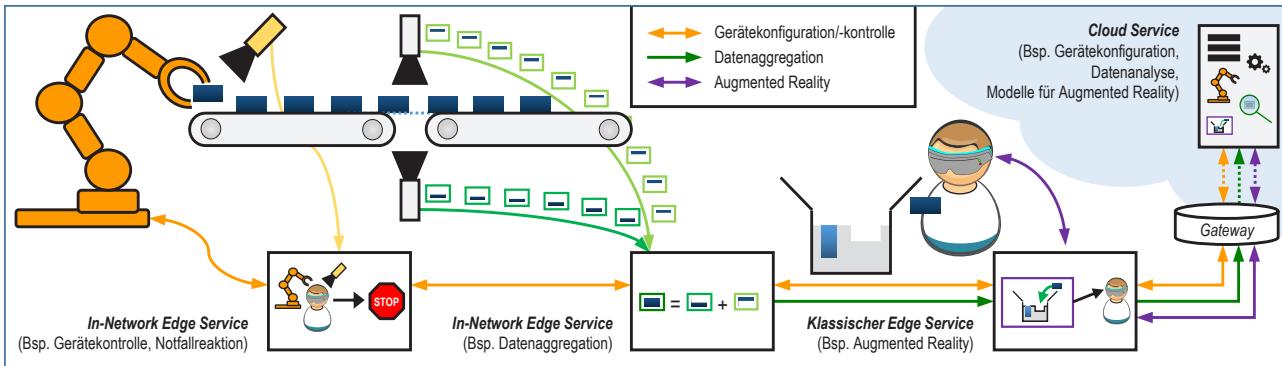


Bild 2: Edge Services in der Produktion treten in zwei Varianten auf: In-Network Services auf Routern, Switches und einfachen Endgeräten für einfache und sicherheitskritische Aufgaben (links, mittig) sowie Services mit hohen Rechenkapazitätsanforderungen auf Edge Clustern (Mitte rechts); entfernte Cloud Services führen Datenanalysen durch und konfigurieren die Edge Services

umzusetzen. Netzwerkgeräte, die ursprünglich nicht für die Ausführung komplexer Algorithmen entwickelt wurden, sind jedoch in ihrer Rechtleistung begrenzt, was die Regeln, die derzeit auf ihnen ausgeführt werden können, stark beschränkt. Erste Ansätze, Software-defined Networking mit ausdrucksstärkeren Sprachen wie eBPF zu kombinieren, zeigen jedoch großes Potenzial zur Ausführung von rudimentären Kontrollalgorithmen direkt auf Netzwerkgeräten [5]. Somit bieten In-Network Edge Services eine effiziente Möglichkeit, um leichtgewichtige Kontroll-, Sicherheits- und weitere Funktionalitäten effizient und so nah wie möglich an der Produktionslinie zu realisieren.

## Klassische Edge Services

Für Fälle, in denen die Fähigkeiten von Netzwerkgeräten nicht ausreichen, können klassische Edge Services, d.h. das Auslagern von Funktionalitäten an leistungsfähigere Maschinen in der Nähe der kontrollierten Prozesse, eingesetzt werden [6]. Während Edge-Geräte zwar über viel Speicherplatz und Rechenkapazitäten verfügen, sind sie jedoch oft mit Standard-Betriebssystemen ausgestattet, deren komplexe Netzwerkstacks schlecht für die Verarbeitung großer Datenströme und Niedriglatenzanforderungen geeignet sind. Lösungen, die darauf abzielen, entweder die Stacks zu umgehen [7] oder die Anwendungslogik teilweise in den Betriebssystemkern auszulagern [8], können eingesetzt werden, um diese Verzögerungen zu verringern.

## Herausforderungen

Edge Cloud Services können einen wichtigen Beitrag leisten, um die im Rahmen des Internet of Production auftretenden Datenmengen beherrschbar zu machen und Latenzanforderungen zu erfüllen. Eine wesentliche Herausforderung dieses neuen Ansatzes zur Verarbeitung von Daten und Steuerung von Prozessen im Produktionsumfeld liegt im Zusammenspiel der zwei verschiedenen Arten von Edge Cloud Services, die zum Einsatz kommen können.

Da klassische Edge Services weiter von den kontrollierten Prozessen entfernt sind als In-Network Edge Services, können klassische Edge Services höchstwahrscheinlich nicht für Prozesse wie die Steuerung von Produktionsmaschinen verwendet werden, die häufig Latenzen unter 1 ms erfordern [9]. Daher ist eine große offene Forschungsfrage, wie ein präzises Zusammenspiel zwischen ressourcen-starken klassischen Edge Services mit mittlerer Latenz und ressourcen-schwachen In-Network Edge Services mit niedriger Latenz realisiert werden kann.

Darüber hinaus ergeben sich Fragen in Hinblick auf IT-Sicherheit und Geschäftsgeheimnisse, wenn zum einen Kontrolllogik eines Drittanbieters in das eigene, gesicherte Produktionsnetzwerk eingespeist wird und zum anderen Edge Services über Firmengrenzen hinweg vertrauliche Daten austauschen müssen.

## Referenzen

- [1] Glebke, R.; Henze, M.; Wehrle, K.; Niemitz, P.; Trauth, D.; Mattfeld, P.; Bergs, T.: A Case for Integrated Data Processing in Large-Scale Cyber-Physical Systems. In: HICSS, 2019
- [2] Internet of Production (2019). [www.iop.rwth-aachen.de](http://www.iop.rwth-aachen.de)
- [3] McKeown, N.; Anderson, T.; Balakrishnan, H.; Parulkar, G.; Peterson, L.; Rexford, J.; Shenker, S.; Turner, J.: Open-Flow: Enabling Innovation in Campus Networks. In: ACM SIGCOMM CCR. 38 (2008)
- [4] Bosshart, P.; Daly, D.; Gibb, G.; Izzard, M.; McKeown, N.; Rexford, J.; Schlesinger, C.; Talayco, D.; Vahdat, A.; Varghese, G.; Walker, D.: P4: Programming Protocol-Independent Packet Processors. In: ACM SIGCOMM CCR. 44 (2014)
- [5] J. Rüth, R. Glebke, K. Wehrle, V. Causevic, S. Hirche, Towards In-Network Industrial Feedback Control, in: ACM SIGCOMM Net-Compute, 2018.
- [6] Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L.: Edge Computing: Vision and Challenges. Internet of Things Journal 3 (2016)
- [7] Rizzo, L.: Netmap: A Novel Framework for Fast Packet I/O. In: USENIX ATC, 2012
- [8] Schmidt, F.; Hohlfeld, O.; Glebke, R.; Wehrle, K.; Santa: Faster Packet Delivery for Commonly Wished Replies. In: ACM SIGCOMM Poster, 2015
- [9] Frotzscher, A.; Wetzker, U.; Bauer, M.; Rentschler, M.; Beyer, M.; Elspass, S.; Klessig, H.: Requirements and current solutions of wireless communication in industrial automation. In: IEEE ICC, 2014

**DR. MARTIN HENZE  
RENÉ GLEBKE  
PROF. DR.-ING. KLAUS WEHRLE**

Lehrstuhl für Informatik 4 (Kommunikation und verteilte Systeme), RWTH Aachen

# Multi-access Edge Computing – a key enabler for latency critical applications

*Multi-access Edge Computing can help to accelerate time to market and ensure industry adoption.*

MEC (Multi-access Edge Computing) is an application enablement framework that enables innovation co-creation with the ecosystems to provide value-added services from the edge of the network and open new business opportunities with enterprises and verticals, especially in ultra-low latency use cases. The edge of the network is an environment which is characterized by proximity, ultra-low latency and high bandwidth. Furthermore, MEC can provide access to real-time network and context information that can be used by applications to create value. MEC can be considered as the “cloud” for real-time and personalized services, providing an unparalleled experience with reduced latency in addition to highly efficient network operation and service delivery.

Applications and services can be accelerated by increased responsiveness and maximized interactivity from the edge. Robust low-latency can be ensured for critical communication; popular and locally-relevant content can be delivered exactly where users and objects connect. Real-time analytics can be performed at the point of capture and corrective actions, based on observed data, can be taken locally.

Access to a real-time network, context information and service analytics can be used to optimize network and service operation and proactively maintain customer experience. Moreover, context information and location awareness can be used to create highly personalized services which can be tailored to individual needs and preferences.

MEC is a powerful catalyst for innovation across traditional domain borders. It supports the digital transformation and mobilization of the enterprise industries. It is a key enabler for IoT and mission critical, vertical solutions.

As depicted in **Figure 1**, MEC can be brought to all locations: far edge, enterprise edge, access, aggregation hub, distributed data center, centralized data center. It can be deployed at hotspots to support zone applications, such as special services in stadiums, exhibitions, malls, enterprise campuses. It can be deployed at metro-aggregation sites and baseband hotels to support city-wide applications, such as video analytics used across the city, or IoT applications used to monitor a smart-city infrastructure. MEC can also be deployed in a distributed data

center (e.g. cloud RAN), or in specific deployment patterns, as well as in a centralized data center to support network-wide applications, such as Car-to-Car and Car-to-Infrastructure along roads, etc.

The ETSI ISG (Industry Specification Group) MEC was founded in September 2014 by six ETSI members representing the different players in the value-chain: Vodafone and NTT Docomo as tier-1 operators, Nokia and Huawei as tier-1 vendors, Intel as a technology provider and IBM as an application provider. The phase 1 specifications are published [1] and include the API (Application Programming Interfaces) framework for the delivery of services which may be consumed or offered by (locally hosted or remote) authorized applications. The framework enables registration, announcement, discovery and notification of services as well as authentication and authorization of applications providing and consuming services.

The specifications include a set of API principles and guidance for developing and documenting mobile edge service APIs inside or outside of ETSI with the objective of ensuring a

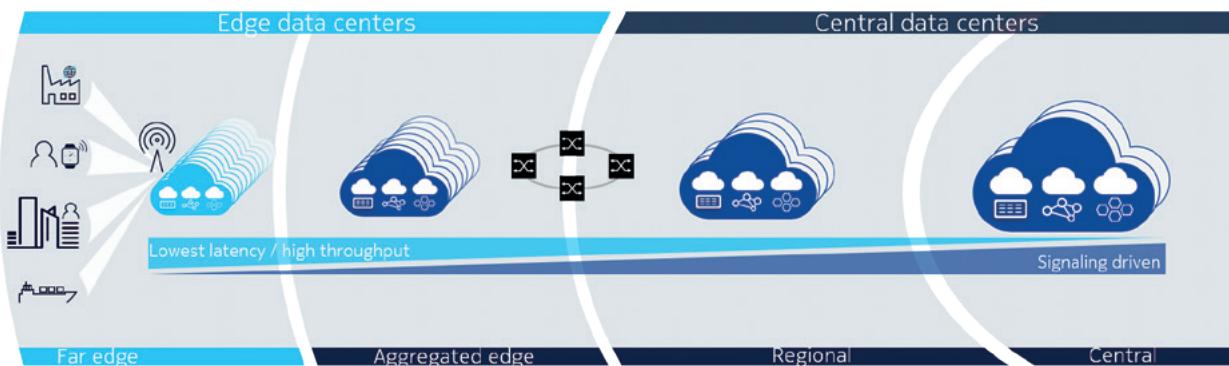


Figure 1: MEC deployment options

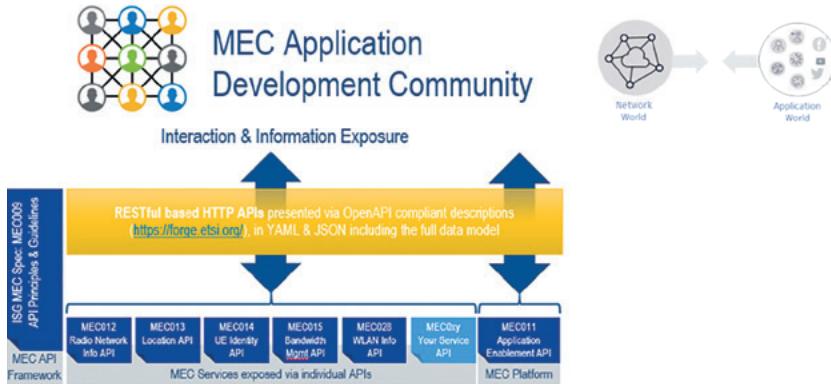


Figure 2: MEC APIs

consistent set of APIs that can be used to empower the developers to create extra value with minimal effort and cost. APIs provide programmatic (read and write) access to abstracted information or actions. ETSI has defined a set of specific service-related APIs for information exposure (see Figure 2), including Radio Network information, Fixed Network information, WLAN (Wireless LAN) Network Information, Location information, UE (User Equipment) Identity, V2X (Vehicle to Vehicle or Vehicle to Infrastructure) information, etc. MEC has used the OpenAPI framework (a.k.a swagger specification) to define and create RESTful APIs. The OpenAPI compliant MEC API description file inherently includes the MEC data structure definitions; it is machine readable and can be used for content validation. The MEC APIs are available for accessibility and validation in electronic format at the ETSI Forge site [2]. The open API descriptors file can be downloaded and used for implementation. Stubs for both service client and server can be created automatically, making it easy to understand what is delivered via the API. Feedback can be provided towards the repository and enhancements can be suggested.

In addition, ETSI MEC has specified management and orchestration-related APIs that facilitates the running of applications at the correct location at the right time, based on technical requirements relating to that specific application, for example end-to-end latency, virtualized resources, the MEC services and features that are required to run the application, connectivity requirements, geographic

reach, support for UE (User Equipment) mobility and service continuity, etc. Business parameters, such as cost and business models, can also influence the decision regarding location. If conditions change, the application or its context may be relocated to another MEC platform.

ETSI Phase 2 work extends the applicability of MEC to support application mobility/re-location to ensure service and session continuity, MEC integration with NW slicing, V2X service in multi-operator/multi-domain environment, etc. Note that MEC is covered widely detailed in technical documents of 5GAA as a key building block of V2X services. Advanced automotive applications such as hazard warning, collision warning, cooperative passing, E-tolling, traffic information, infotainment, etc. can all benefit from MEC. MEC was used for the V2X showcased at A9 in Deutsche Telekom's live LTE network that took place already on November 2015 [3].

MEC is a natural element of 5G, helping to satisfy 5G throughput, latency, scalability and automation targets. It offers additional privacy and security and ensures significant cost efficiency. Integrating MEC into 5G architecture creates added value, ensuring efficient network operation, service delivery, and a highly personalized customer experience. 3GPP has specified the key enablers for the integration of MEC in 5G. The Enabling Applications at the edge of the 5G Network [4] and MEC in 5G Networks [5] white papers describe how these enablers can be used to integrate MEC into 5G architecture.

The GSMA association has published white papers related to MEC,

such as business opportunities and monetization options for MEC, stakeholder analysis and feasibility study for MEC, mobile throughout guidance, and implementation guidance for MEC.

It is important to note that the MEC specifications are ready, but the edge cloud ecosystem is still being developed. We are witnessing a rapid expansion of edge-related open source initiatives. It is important to ensure the MEC application platform (API framework and APIs) is leveraged by key edge computing related initiatives to allow the developers to create common architectures and write a single application software module that can run on every edge environment. MEC provides a common API framework that supports all use cases: telco, IT, enterprise, consumers, and is ready for the unexpected; it can be deployed in different cloud environments and allow a single application software module to run on every infrastructure platform. This can help to accelerate time to market and ensure industry adoption.

With its open architecture, open APIs and open ecosystem, MEC provides the openness foundation.

## References

- [1] <https://www.etsi.org/standards-search#page=1&search=&title=1&etsiNumber=1&content=1&version=0&onApproval=1&published=1&historical=1&startDate=1988-01-15&endDate=2019-02-20&harmonized=0&keyword=&TB=826,,835&stdType=&frequency=&mandate=&collection=&sort=2>
- [2] <https://forge.etsi.org>
- [3] [www.prnewswire.com/news-releases/continental-deutsche-telekom-fraunhofer-esk-and-nokia-networks-showcase-first-safety-applications-at-digital-a9-motorway-test-bed-543728312.html](http://www.prnewswire.com/news-releases/continental-deutsche-telekom-fraunhofer-esk-and-nokia-networks-showcase-first-safety-applications-at-digital-a9-motorway-test-bed-543728312.html)
- [4] [www.networks.nokia.com/solutions/multi-access-edge-computing](http://www.networks.nokia.com/solutions/multi-access-edge-computing)
- [5] <https://portal.etsi.org/TBSiteMap/MEC/MECWhitePapers.aspx>

## NURIT SPRECHER

Nokia, Head of Management, Virtualization and Application Enablement Standardization, and Founding Chair of ETSI ISG MEC

# Hamburg Smart Sea Port – 5G-Network-Slicing-Testfeld im realen Betrieb

Im Hamburger Hafen wurde ein 5G-Testfeld eingerichtet, mit dem 5G Network Slicing anhand realer Anwendungsfälle getestet wird.

Im Rahmen des europäischen Horizon-2020-Forschungsprojekts 5G-MoNArch [1] wurde durch die drei Projektpartner Hamburg Port Authority (HPA), Deutsche Telekom und Nokia im Hamburger Hafen ein 5G-Testfeld für Anwendungsfälle des Hafens der Zukunft implementiert. Ziel des Testfelds ist es, Network Slicing, eine Kernkomponente der künftigen 5G-Netze, anhand realer Anwendungsfälle zu testen.

Zu den implementierten Anwendungsfällen gehören als Erstes die Optimierung des Last- und Lieferverkehrs durch intelligente Verkehrsteuerung, für die exemplarisch eine Ampelanlage an die zentrale Verkehrsüberwachung der HPA angebunden wurde. Als Zweites wurde für die kontinuierliche Überwachung der Luft- und Wasserqualität entsprechende Sensorik auf Barkassen der HPA installiert und über das Mobilfunk-Testnetz mit der Betriebszentrale ver-



Bild: Hamburg Port Authority, 2019

Bild 2: Präsentation der Ampelsteuerung in der Verkehrsüberwachung der HPA (v.l.n.r): Jens Meier, CEO der HPA, Markus Breitbach, Deutsche Telekom, und Hendrik Roreger, HPA

bunden. Als dritter Anwendungsfall wird ein mobiles Wartungsteam der HPA mit Virtual-Reality-(VR)-Brillen ausgestattet, um eine Vor-Ort-Unterstützung durch Experten aus der Zentrale zu ermöglichen. Jeder der drei genannten Anwendungsfälle hat

dabei spezifische Anforderungen an die Qualität der Mobilfunkverbindung: Absolute Zuverlässigkeit bei der Überwachung der Ampelanlage, geringster Datenverlust bei gleichzeitiger Mobilität für die Luftgütemessungen und geringe Latenz bei gleichzeitiger hoher Datenrate für die VR-Brillen [2].

Das Mobilfunk-Testnetz im Hafengebiet (**Bild 1**) implementiert Ende-zu-Ende-Network-Slicing, um alle drei Anwendungsfälle mit der geforderten Qualität und voneinander isoliert über eine gemeinsame Infrastruktur betreiben zu können. Mithilfe der eingesetzten Lifecycle-Management-Software können applikationspezifische Network Slices im Testnetz innerhalb von Minuten konfiguriert und eingerichtet werden. Dies konnte im November 2018 bei

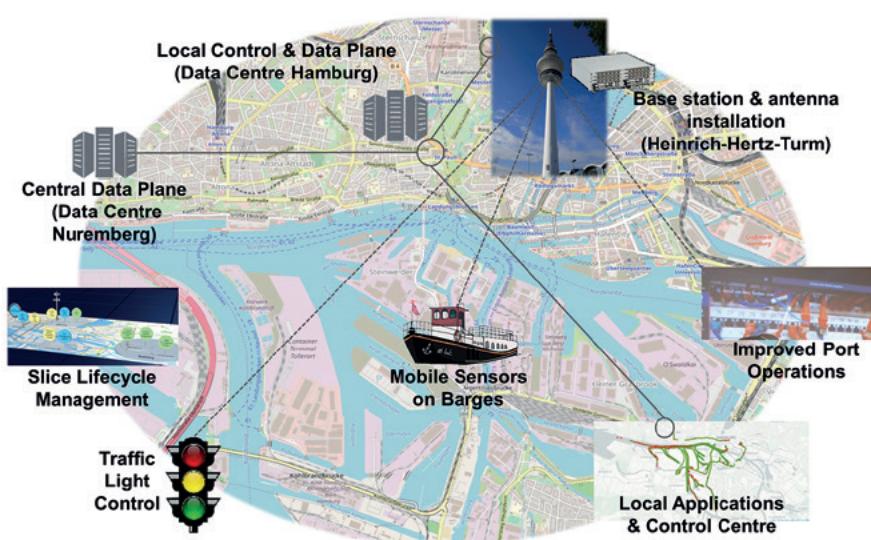


Bild 1: Aufbau und Anwendungsfälle des Testfelds in Hamburg

einem Workshop [3, 4] der Hamburger Hafenbehörde mit zahlreichen Vertretern von am Hafen ansässigen Firmen demonstriert werden (**Bild 2**); zudem wurde gezeigt, wie die Verkehrsströme der einzelnen Network Slices voneinander entkoppelt sind. Dies ist enorm wichtig im industriellen Umfeld, wo zum Beispiel eine Lastzunahme in einem Slice keine Auswirkung auf kritische Dienste in einem anderen Slice haben darf.

Der Aufbau der Testumgebung in Hamburg ist größtenteils abgeschlossen. Die in einer Höhe von 180 Meter installierten Antennen der Basisstation decken den zentralen Bereich des Hafens ab. Durch die starke Überlappung der Zellen und die imple-

mentierte Dual-Connectivity-Lösung kann insbesondere für die Ampelanlage und die mobilen Sensoren auf den Barkassen eine hohe Zuverlässigkeit der Verbindung sichergestellt werden. Das Kernnetz ist auf ein lokales Datenzentrum in Hamburg sowie ein zentrales Datenzentrum in Nürnberg aufgeteilt, wobei die Steuerebene (Control Plane) lokal, die Datenverarbeitung (Data Plane) sowohl lokal als auch zentral durchgeführt werden kann.

## Referenzen

- [1] 5G Mobile Network Architecture for diverse services, use cases and applications. <https://5G-monarch.eu>
- [2] Rost, P.; Breitbach, M.; Roreger, H.; Erman, B.; Mannweiler, C.; Miller, R.; Viering, I.: Custo-

mized Industrial Networks – Network Slicing Trial at Hamburg Seaport. IEEE Wireless Communications Magazine 25 (2018) H. 5

[3] [www.nokia.com/news/releases/2018/11/06/port-of-hamburg-5g-applications-pass-field-test](http://www.nokia.com/news/releases/2018/11/06-port-of-hamburg-5g-applications-pass-field-test)

[4] Port Technology, 7. Nov. 2018. [www.porttechnology.org/news/port\\_of\\_hamburg\\_hails\\_5g\\_success](http://www.porttechnology.org/news/port_of_hamburg_hails_5g_success)

## DIPL.-ING. LARS CHRISTOPH SCHMELZ

### DR.- ING. PETER M. ROST

Nokia Solutions and Networks GmbH & Co. KG,  
München

### DR.-ING. MARKUS BREITBACH

Deutsche Telekom AG, Bonn

### DIPL.-INF. HENDRIK ROREGER

Hamburg Port Authority AöR, Hamburg

# Personalia

## MIT DER VDE-EHRENURKUNDE WURDE PROF. DR. DIRK HEBERLING, RWTH AACHEN UND FHG, AUSGEZEICHNET.

Übergeben wurde die Auszeichnung von Prof. Dr. Joachim Ender (im Bild links), Sprecher des Fachbereiches HAT.



## Veranstaltungen

**Hinweis:** Weitere Veranstaltungen finden Sie auf den Seiten 46 und 47 des VDE dialog.

**08.–09.04.2019, Kaiserslautern**  
**MBMV 2019 – Methoden und Beschreibungssprachen zur Modellierung und Verifikation von Schaltungen und Systemen**  
ITG, GMM, GI  
[// www.vde.com/de/itg/veranstaltungen](http://www.vde.com/de/itg/veranstaltungen)

**08.–09.05.2019, Leipzig**  
**Photonische Netze 2019**  
ITG-FA KT 3, ITG-FG 5.3.3  
[// www.vde.com/de/itg/veranstaltungen](http://www.vde.com/de/itg/veranstaltungen)

**08.–09.05.2019, Clausthal-Zellerfeld**  
**Simulation Science Workshop 2019**  
ITG, TU Clausthal, SWZ Simulation Science Service, Universität Göttingen  
[// www.simscience2019.tu-clausthal.de](http://www.simscience2019.tu-clausthal.de)

**04.–05.06.2019, Magdeburg**  
**Smart SysTech 2019 – European Conference on Smart Objects, Systems and Technologies**  
ITG, AIM-D e.V., Fraunhofer IPMS  
[// www.vde.com/de/itg/veranstaltungen](http://www.vde.com/de/itg/veranstaltungen)

**15.–16.05.2019, Osnabrück**  
**24. ITG-Fachtagung Mobilfunk-kommunikation**  
ITG-FA KT 2  
[// www.vde.com/de/itg/veranstaltungen](http://www.vde.com/de/itg/veranstaltungen)

**24.–25.09.2019, Mannheim**  
**Digitale Welt Rhein-Neckar**  
ITG, GPM, VDE Bezirksverein Kurpfalz  
[// www.vde.com/de/itg/veranstaltungen](http://www.vde.com/de/itg/veranstaltungen)

**07.–09.10.2019, Ostfildern**  
**Quantentechnologie – Chancen für**

## neue Anwendungen in Industrie und Gesellschaft

ITG, TAE, VDE Bezirksverein Württemberg e.V.  
[// www.vde.com/de/itg/veranstaltungen](http://www.vde.com/de/itg/veranstaltungen)

## Impressum

### ITG-news

Herausgeber: Informationstechnische Gesellschaft im VDE, Frankfurt am Main

Redaktion: Dr. Volker Schanz,  
Silvia Buhlmann

Telefon: 069/6308-360/-362  
E-Mail: [itg@vde.com](mailto:itg@vde.com)  
Internet: [www.vde.com/itg](http://www.vde.com/itg)

Konzept und Realisation: HEALTH-CARE-COM GmbH, ein Unternehmen der VDE VERLAG GmbH, Projektleitung: Anne Wolf

Druck: Heenemann GmbH & Co. KG, Berlin