

SRIA promotion webinar, Feb 3-4, 2021

SRIA: System Architecture

(Chapter 4)

Artur Hecker

Huawei Munich Research Center

Based on Networld2020 SRIAv3.0



**The European Technology Platform for
communications networks and services**



About this document



- Presentation

- from Networld2020 (<https://www.networld2020.eu/>)
- of the System Architecture related research challenges as per its SRIA published 2020:
 - Cf. Chapter 4, System Architecture in
 - SRIA: <https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf>
- at the promotion event held online Feb 3-4, 2021
 - Webinar: <https://www.networld2020.eu/networld2020-sria-webinar-3-4-february-2021/>



Background: Evolution of Networks and Services



- The success of the mobile Internet continues revolutionizing our society
- However, the Internet has changed in at least three aspects
- **Nature of communications:** from servers to services
 - From Client-Server model to a Client to Data Center model (DC: distributed, proprietary)
 - Localized realization (SP PoPs, relative locality, loss of importance of global routing)
- **Nature of services:** from monolithic services to microservices
 - Services as chains of collaborative transactions
 - Repositioning the meaning of a middle box, and, of the networking
- **Nature of provisioning**
 - From bare metal to virtualization
 - From planning to dynamic orchestration of instances
 - Radically changed node/structural dynamics, different topologies
- **Key challenge:** dynamically set up a service environment of determined trustworthiness over microservices of varying trustworthiness levels



Smart Green Intercomputing



- **Evolution**

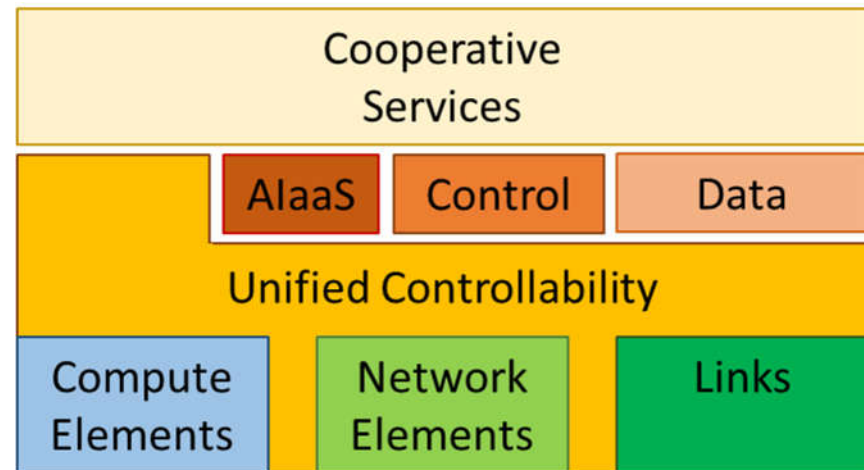
- From inter-working to inter-computing
- From physical boundaries to service boundaries
- From an add-on to a critical infrastructure

- **Future Networking landscape**

- Higher dynamics
- Co-existence of virtual and physical entities
- Physically overlapping policy islands

- **Key challenges**

- Flexible provisioning (any service)
- Unified resource control (execution)
- Elastic execution (increased yield)
- Integrated smartness (decision)



- **Infrastructure dynamically composed of het. resources**

- Features self-organization in the resource control domain
- Exports basic support for decision, control and transport to services

- **Services allocated on the latter, through programmability**

- Max. degrees of freedom
- Max. multiplexing



Virtualised Network Control for Increased Flexibility



- Programmability is key for ICT infrastructure sustainability
 - It requires scalable, pervasive control solutions
 - Yet it is more: programmability creates dynamics that it must be able to cope with
 - Programmability requires “protected modes” to be able to apply and undo changes resiliently in all situations
- Research challenges
 - Autonomically establishing, auto-stabilizing, self-preserving resource control acting as an initial glue beyond the current domain boundaries
 - Must be dynamic, adaptive and elastic in itself
 - Must be able to work with incomplete knowledge



Re-Thinking the Data & Forwarding Planes



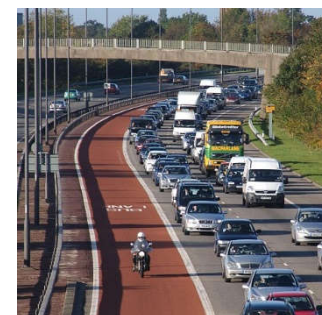
- Design Considerations for an Evolved Data & Forwarding Plane
 - Beyond best effort global host reachability
- Key Research Questions
 - Layering
 - Soft architecting
 - Which actor interplay produces which outcome?
 - Which metadata is required?
- Recommendations for Actions
 - Requires international cooperation
 - Requires large scale testbeds, open source



Efficiency and Resource Management



- Green ICT: how to get more service from resources?
 - Expressed in capacity, nodes, EUR, Watts, ...
- Shared use of a common infrastructure by different tenants requires some form of dynamic resource management
 - Pre-allocation (e.g., hard quotas) does not support post-deployment solutions like programmability
 - Fixed allocations are simple yet exhibit bad yield
 - Require a full *a priori* knowledge
- Three main research challenges:
 - App expression of needs and allocation decision
 - Conflict resolution
 - Runtime decision



AI/ML-based System Evolution



- **Why AI in the System Architecture?**
 - Recent advances in AI/ML enable new applications
 - Increasing complexity, scale and dynamics of ICT
- **Expected: AI as part of basic system services**
 - For service providers: Proliferation of AlaaS in Network Operations
 - For users: AlaaS Proliferation in Service Provisioning
- **Research Challenges**
 - Availability of network-characteristic data sets
 - Methodological training schemes for dynamic environments
 - Distributed or Network-Suitable AI
 - AutoML



Deep Edge, Terminal and IoT Device Integration



- As per SRIA system vision, **the deepest edge is the terminal**
 - In this view, every device is a capacity
 - It should be used to maximize both yield and the degrees of freedom
 - Exceptions are still likely
 - Constrained or unstable devices, owner restrictions
 - May be integrated over a specialized gateway
- **Highly attractive**
 - Numerous, powerful, highest locality, lowest latency
 - Situationally aware *and* integrated in the environment
- **Research Challenges**
 - Privacy and data management
 - Effective management of numerous devices (e.g. IoT, fleets, etc)
 - Architectural Device Integration
 - Common model for the representation of different aspects
 - Delivery models and open APIs
 - Constraint-aware scheduling mechanisms
 - Incentivization and governance schemes



Potential KPIs



- Pervasive resource control
 - Autonomic features
 - Resilience to self-inflicted errors
- Resource efficiency
 - Runtime Service Scheduling efficiency
 - Time required for runtime conflict resolution
 - Networked garbage collection
 - TCO
- Network-suitability of AI platforms
 - Maturity and application
 - Scalability and performance
- Terminal integration
 - Usability / domain boundaries
 - Deployability of complete service architectures on edge-only
- See more details in the published SRIA, Chapter 2

