Data-centric-infrastructure-as-a-service PoC Reference

Classification: APPROVED REFERENCE DOCUMENT

Confidentiality: PUBLIC

Version 1.0

July 21, 2022

[DCIaaS PoC Reference]

Legal

Contents

1. Introduction 5

1.1. Context of this document within IOWN GF 5

1.2. Structure of this document 6

2. Scope: focus on services, not concrete applications 7

3. Example reference cases 8

3.1. Area management with security monitoring 8

3.1.1. Overall structure of the application 8

3.1.2. Key technological gaps to address with initial DCIaaS PoC activities 8

3.1.3. Key features to consider as background for initial DCIaaS PoC activities 8

3.1.4. Advanced DCIaaS features for composing LSNs from resource pools 9

3.2. VRAN deployment 9

3.2.1. Overall structure of the use-case 10

3.2.2. Key technological gaps to address with initial DCIaaS PoC activities 10

3.2.3. Key features to consider as background for initial DCIaaS PoC activities 10

3.2.4. Advanced DCIaaS features for creating VRAN nodes from resource pools 11

3.3. Data hub disaggregated deployment 11

3.3.1. Overall structure of the use-case 11

3.3.2. Key technological gaps to address with initial DCIaaS PoC activities 12

3.3.3. Key features to consider as background for initial DCIaaS PoC activities 12

3.3.4. Advanced DCIaaS features for managing migration and storage node configuration. 13

3.4. Common requirements: QoS-aware communication stacks, data pipelines, and networks 14

4. Desired features for initial PoCs 15

4.1. Composing heterogeneous computing systems by enabling and releasing of LSNs 15

4.1.1. Description of the desirable feature 15

4.1.2. Example of a possible PoC demonstrating the feature 16

4.2. QoS-aware network supporting links with extreme QoS between components of different LSNs . 17

4.2.1. Description of the desirable feature 17

4.2.2. Example of a possible PoC demonstrating the feature 18

4.3. QoS-aware data pipelines between application level endpoints 18

4.3.1. Description of the desirable feature 18

4.3.2. Example of a possible PoC demonstrating the feature 19

5. Expected benchmarks 20
5.1. Energy efficiency 20
5.1.1. Background of the benchmark 20
5.1.2. KPIs to evaluate 20
5.2. Network service quality 20
5.2.1. Background of the benchmark 20
5.2.2. KPIs to evaluate 21
5.3. Applicability of the QoS-aware network for time synchronization 21
5.3.1. Background of the benchmark 21
5.3.2. KPIs to evaluate 22
6. Summary 23
Annex A. Outlook: future DCIaaS PoC roadmap 24
A.1. Second step PoCs 24
A.2. Third step PoCs 25
References 26
History 27

List of Figures

Figure 3-1: Suggested key features of the PoC in the context of the IOWN GF area management use-case... 9
Figure 3-2: Suggested key features of the PoC in context of VRAN use-cases. 11
Figure 3-3: Suggested key features of the PoC in context of the IOWN GF data hub technologies. 13
Figure 4-1: Example of heterogeneous systems composition PoC structure. 17
Figure 4-2: Example of QoS-aware network PoC structure. 18
Figure 4-3: Example of QoS-aware data pipeline PoC structure. 19

List of Tables

Table A-1: Envisioned timeline of PoC steps. 24

1. Introduction

This section explains and outlines the context and structure of this document.

1.1. Context of this document within IOWN GF

In 2021, IOWN Global Forum (IOWN GF) identified a number of future use-cases with growing market demand that are complex to realize with current technology. Examples of these IOWN GF use-cases include:

- 5G/6G virtual radio access networks (VRAN)/open radio access networks (Open RAN) network function virtualization (NFV),
- area management with artificial intelligence (AI) security monitoring, and
- AI-integrated communication applications such as interactive music performance events with massive and bi-directional streaming.

In turn, market demand for these use-cases causes demand for computing infrastructures that can support these use-cases. Such infrastructures are expected to:

- allow applications of the use-cases as mentioned above to utilize heterogeneous computing resources to cost- and energy-efficiently support the extreme amounts of computation of these use-cases,
- support assembling of computing environments dynamically from disaggregated hardware component pools (called logical service nodes (LSNs) below),
- provide their services on-demand whenever the customer (requestor) needs the services,
- allow platform service providers to become tenants on top of such a base computing infrastructure, and thereby enable such providers to augment LSNs with their own specific additional functionality before giving their customers access to such LSNs,
- host such LSNs in geographically dispersed data centers,
- provide communication mechanisms as well as control and management mechanisms that can satisfy the extreme quality-of-service (QoS) requirements of the use-cases as mentioned above,
- provide the basis to realize computing with deterministic QoS for latency-sensitive and real-time applications that also make use of the communication mechanisms mentioned above, and
- isolate tenants of the infrastructure from each other both performance- and security-wise.

IOWN GF has defined the IOWN GF data-centric-infrastructure (DCI) [DCIFA] and the IOWN GF all-photonics-network (APN) [APNFA] to fulfill these expectations together: APN provides a foundation for dynamically creating optical paths with deterministic quality; DCI builds on top of APN to provide its services. As the next step, the viability of DCI and APN as solutions to realize such massive heterogeneous computation and extremely high QoS communication for future applications needs to be demonstrated.

Data-centric-infrastructure-as-a-service (DCIaaS, cf. [DCIFA](Sec. 4)) is an open architecture that is capable of providing a generic and heterogeneous computing and communication infrastructure. This infrastructure can be consumed by multiple tenants and their applications. Therefore, DCIaaS is a fundamental and central element, connecting multiple IOWN GF technologies into a single multi-tenant infrastructure. Consequently, DCIaaS also becomes a base technology for other IOWN GF technologies, such as the IOWN data hub (IDH) data management framework, and the application of IOWN GF technologies for mobile networks (IMN).

Any application or use-case wanting to benefit from any IOWN GF technology regarding computing or communication will need to use DCIaaS; accordingly, the viability of DCIaaS should be demonstrated as early as possible.

Therefore, the purpose of this proof-of-concept (PoC) Reference is to encourage IOWN GF members to conduct PoC activities and provide feedback to IOWN GF by preparing PoC Reports. Using these reports, IOWN GF plans to

- use the results described in the PoC Reports to refine the IOWN GF DCI architecture,
- foster the market for IOWN GF ecosystem products and services by showcasing to potential DCI-as-a-service providers and application creators that IOWN GF DCI and APN technologies will solve future customer business needs, and
- raise wider interest of hardware/system software vendors to create hardware for IOWN GF technologies in general and IOWN GF DCI in particular by demonstrating that business-ready products are within reach.

1.2. Structure of this document

The remainder of this document is structured as follows: Directly following, the scope of the PoC activities envisioned is briefly visited in Section 2, emphasizing the focus on DCI platform services. Next, short overviews of the example reference use-cases – area management, VRAN, and data warehouses – are introduced in Section 3, to provide the reader with knowledge about context of the PoC technologies. After that, suggestions for initial PoC features are given in Section 4, and matching benchmarks to evaluate these features are outlined in Section 5. Finally, the document concludes with a summary in Section 6.

Additionally, a roadmap with ideas for additional future DCIaaS PoC activities is presented in Annex A.

Furthermore, the reader is assumed to have basic familiarity with the IOWN GF APN and DCI functional architectures; for more information, readers are advised to consult the respective technical reports [APNFA, DCIFA].

2. Scope: focus on services, not concrete applications

This PoC Reference outlines possible activities to demonstrate the viability of the computing and communication infrastructure proposed by IOWN GF DCI and APN. Therefore, PoCs should be focused on demonstrating particular services a DCI platform has to offer to its tenants. Accordingly, PoC Reports should report the maximum achievable service levels that can be realized, so that application creators get an impression of what final systems will be capable of. In turn, PoCs do not need to present complete or even partial use-case applications – this would be the scope of use-case reference implementation model (RIM) PoCs.

As DCIaaS should provide many features, IOWN GF proposes a multi-step approach as outlined in Annex A. The remainder of this document focuses on the initial step. In this first step, DCIaaS PoCs should prioritize only the most essential features and services. At the current stage, IOWN GF suggests to focus only on those features that are required to conduct the key benchmarks.

3. Example reference cases

This section briefly visits three IOWN GF use-cases to provide the reader with context regarding how DCIaaS services offered to tenants are expected to be used in actual applications. DCIaaS PoCs are not expected to implement these applications.

In the following, the reference cases of area management with security monitoring, VRAN, and data warehouses are introduced in Section 3.1, Section 3.2, and Section 3.3, respectively. Finally, Section 3.4 identifies QoS-aware communication mechanisms as a requirement common to all these showcased reference cases.

3.1. Area management with security monitoring

The purpose of this application is to provide security monitoring for a given area. This use-case and its reference implementation model are extensively described in [AMRIM].

3.1.1. Overall structure of the application

The area management with security monitoring application is structured as follows: A large number of security cameras record footage of the monitored area. The video stream of each camera is sent to a local aggregation node. There, the data streams are bundled and possibly prepared for mid-range transmission to a regional data center. At the following ingestion stage, the video streams are analyzed for events of interest. An example of such events is a person collapsing on the street due to health issues or criminal activities. The detected events are sent to a data hub. From there, an intelligence application uses these events to trigger external actions, such as alerting operators of anomalies. The communication and computation load of the system is expected to vary according to the time of day due to the human activity cycle. These load cycles of the system are expected to be well-predictable.

3.1.2. Key technological gaps to address with initial DCIaaS PoC activities

A number of technological gaps exist that hamper the realization of this use-case with today’s technology. A discussion of current technological gaps together with a discussion of how DCI features can fill these gaps is provided in [AMRIM] (Sec. 4.2, Sec. 5.1): among other issues, means are required to

- reduce the processing and energy overhead of data transfers and
- cope with load fluctuations without needing to resort to hardware over-provisioning.

The following section proposes features to implement in initial PoCs to demonstrate how these gaps can be overcome.

3.1.3. Key features to consider as background for initial DCIaaS PoC activities

Two key services to be provided by DCIaaS to tenants that in turn offer services on application level include

- point-to-point streaming at extremely high bandwidths on tenant user application level and
- on-demand provisioning of logical service nodes (logical servers) (cf. [DCIFA] (Sec. 1.3)) with heterogeneous computing resources such as IPUs/DPUs and GPGPUs executing applications that participate in such point-to-point transfers.

PoC conductors should be particularly aware of these two DCIaaS features as background for their work, as these features are deemed to be key to realize this and other future use-cases. These two features are illustrated in Figure 3-1 below.

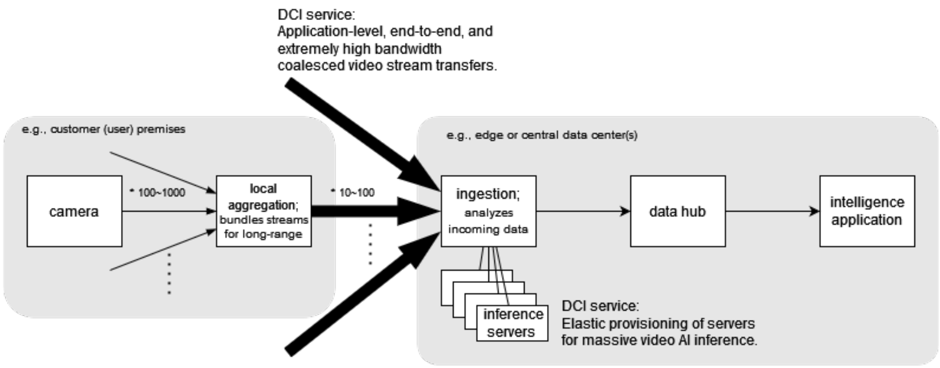

Figure 3-1: Suggested key features of the PoC in the context of the IOWN GF area management use-case.

3.1.4. Advanced DCIaaS features for composing LSNs from resource pools

As hinted above, logical service nodes (servers) used to implement the individual functions of the application on top could have hardware compositions as follows:

- The main tasks of the ingestion stage are accepting high-bandwidth video streams, video decoding, and image analysis by artificial neural networks. Therefore, logical servers created from sets of central processing units (CPUs), network interface cards (NICs), and general-purpose graphics processing units (GPGPUs), or sets of data processing units (DPUs) and GPGPUs seem to be appropriate for these tasks.
- The main tasks of the data hub are related to distributed data processing, e.g., data replication and ensuring the integrity and confidentiality of data. Therefore, logical servers containing components centered around specialized NICs, DPUs, or infrastructure processing units (IPUs), as well as non-volatile memory seem to be appropriate for these tasks.

To increase hardware utilization, IOWN GF further envisions that these various heterogeneous hardware components may be provided by hardware pools. For the area management reference case, the ingestion node as shown in Figure 3-1 is one of the most computationally intensive parts of the application and may benefit from accelerator pools, as is outlined in [AMRIM](Sec. 5.3.2). Furthermore, logical servers may be composed from these resource pools. Therefore, succeeding PoC steps could investigate candidate approaches to create such pools. Such approaches may include:

- System bus extension: This is a technique by which the system bus is extended beyond a single motherboard or server. This allows application software to interact with hardware outside the server as if this hardware was locally installed in the server.
- Software stack interpositioning: This technique works by intercepting application programming interface (API) calls between layers of a software stack such as an artificial intelligence machine learning framework, and replaying these calls in a remote environment.

Since system bus extension operates on system bus level, and software stack interpositioning operates on user or driver software level, the techniques are considered to be orthogonal to each other and may even be usable in combination.

3.2. VRAN deployment

VRAN is a recently emerging trend to virtualize the components of radio access networks (RANs) and executes these components on generic hardware augmented with accelerator cards as required. This allows operators to deploy RANs in shorter time and further to elastically scale their RANs up and down depending on the load situation. As a result, initial cost, initial time-to-deploy, and energy consumption are all reduced.

3.2.1. Overall structure of the use-case

Simplifying, a typical 5G network system comprises the following two main networks:

- Radio access network (RAN): consisting of the radio unit (RU), distributed unit (DU) and the central unit (CU), provided by the radio access network operator. These functions convert analog radio signals to domain-specific data packets.
- Core network (CN): The core network including the user plane function (UPF) that connects the RAN to the internet, provided by mobile core operator.

Furthermore, emerging applications such as AI inferencing of user data and low-latency high-precision mapping services that are provided by the enterprise service operators and that offer value to end-users must be executed in close physical proximity to the RAN infrastructure due to latency and security requirements.

This means that mobile applications involve three kinds of workloads: radio access network, mobile core network, and user applications.

In the past, most components executing these workloads were placed in close proximity to the radio antennas. However, there is a trend to disaggregate these components. Except for the RU, which needs to be close to its antenna, in the near future, it is expected that the components of the workloads need to be placed across mobile edge sites, mobile core sites, or in regular cloud data centers. This will be necessary to fulfill the service level requirements of future user applications. In particular, user applications with stringent service requirements will have RAN, CN, and mobile network tenant user application workloads to execute in the same location. Furthermore, the workload of all components is expected to vary significantly over time.

3.2.2. Key technological gaps to address with initial DCIaaS PoC activities

A number of technological gaps exist that hamper the realization of this use-case with today's technology. Among others, DCI attempts to fill the following gaps of current technologies:

- For efficiency, workloads should be accelerated on accelerators specialized for each workload, and due to workload fluctuations, over time different numbers of servers with different hardware configurations are required; however, in current data centers, only limited types of servers are available.
- Accelerators need to handle extremely high data volumes, and accelerators should further be able to terminate connections with incoming data, without having to rely on CPUs for traffic processing. In addition, the network software stacks should abstract the communication mechanisms between accelerators, such as Ethernet network, system bus, or host memory. However, while direct transfer mechanisms between accelerators are available today, these typically are specialized solutions with only limited support for different communication mechanisms and limited interoperability between different types or even classes of computing devices.

The following section proposes features to implement in initial PoCs to demonstrate how these gaps can be overcome.

3.2.3. Key features to consider as background for initial DCIaaS PoC activities

Considering the above, two key services to be provided by DCIaaS to its DCIaaS tenants that in turn offer services on application level include

- extremely high quality-of-service (QoS) point-to-point connection support between DCI tenant user applications and
- on-demand provisioning of heterogeneous types of servers executing applications that participate in such point-to-point transfers and accordingly reconfiguration of the communication mechanism in-between.

PoC conductors should be particularly aware of these two DCIaaS features as background for their work, as these features are key to realize this future use-case. These features are illustrated in Figure 3-2 below.

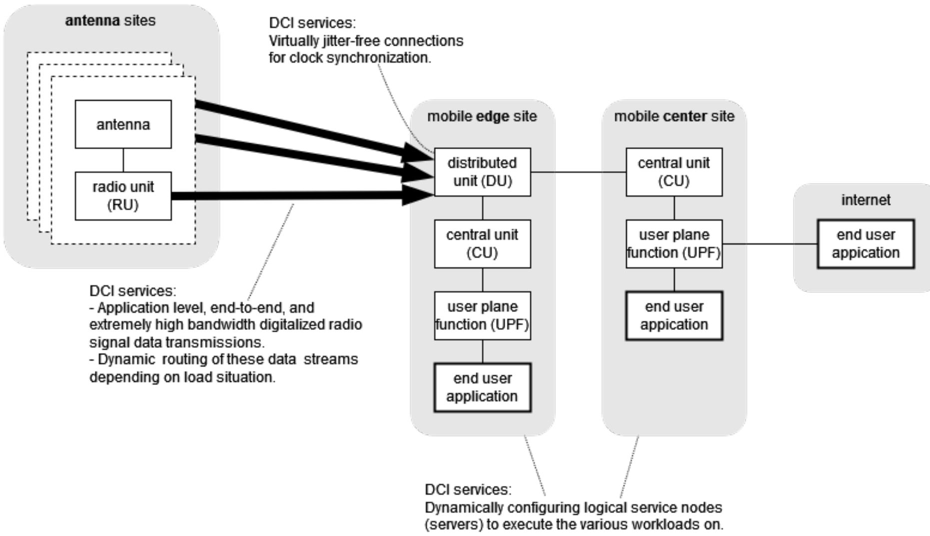

Figure 3-2: Suggested key features of the PoC in context of VRAN use-cases.

Furthermore, different workloads will require different accelerator hardware for energy-efficient execution. For example,

- 5G CU and UPF workloads could be realized using CPUs, specialized NICs, and/or DPUs/IPUs,
- 5G DU workloads could be realized using combinations of CPUs, specialized NICs, low-density parity-check (LDPC) accelerators, and/or DPU/IPU accelerator cards, and
- tenant workloads such as AI inference could be realized using combinations of CPUs and GPGPUs.

3.2.4. Advanced DCIaaS features for creating VRAN nodes from resource pools

Future PoC phases need to demonstrate that logical service nodes (LSNs) can be composed on demand from disaggregated resource pools holding the various components as introduced above.

Additionally, LSNs need to be isolated from each other both performance- and security-wise. The main reason is that applications of multiple different organizations with separate service level agreements are expected to execute on the same hardware pools (DCI physical nodes).

3.3. Data hub disaggregated deployment

This reference case represents the deployment of large-scale databases in the form of IOWN GF data hubs [DHFA]. Currently, typical performance-sensitive database systems implement processing and storage on the same physical servers. Database components can be disaggregated using DCI technologies. This will provide better scalability and elasticity, higher reliability, simplified maintenance, and lower operating cost.

3.3.1. Overall structure of the use-case

Typically, large-scale databases are organized in a three-tiered functional architecture:

1. Front-end functions accept client requests from the outside.
2. Service functions perform the actual processing as requested by the front-end functions.

3. Storage functions provide access to the mass storage backend on which the actual data is stored.

While logically being a clean layered structure, actual implementations are often monolithic in that service (i.e., processing) and storage must be located in close proximity. Typically, processing and storage are located even on the same physical server. This is required due to the performance limitations of the networks inside data centers.

With IOWN GF data-centric-infrastructure (DCI) interconnects, the expectation is that processing and storage of such databases can be disaggregated.

3.3.2. Key technological gaps to address with initial DCIaaS PoC activities

A number of technological gaps exist that hamper the realization of this use-case with today's technology. Among others, DCI attempts to fill the following gaps of current technologies:

- Current data centers provide virtually unlimited any-to-any connections, among which network resources get evenly distributed. However, data warehouses have opposite requirements, i.e., only a bounded number of connections between service and storage functions, but with high and guaranteed service quality.
- Over time, the geographical location of servers where storage functions reside needs to be changed, for example to create new database replicas or to move storage to different locations to improve efficiency. However, again, current networks are not optimized for such extremely long burst-type transmissions, which causes data movement to be impractical in many cases.
- Geographically distributed databases benefit from precisely synchronized time sources to accelerate transaction processing. However, while specialized solutions exist, these are not widely available to end-users in data centers, and other generic current infrastructure technologies do not provide time synchronization functionality with sufficient precision.

The following section proposes features to implement in initial PoCs to demonstrate how these gaps can be overcome.

3.3.3. Key features to consider as background for initial DCIaaS PoC activities

Considering the above, two key services to be provided by DCIaaS to tenants that in turn offer services on application level include

- point-to-point streaming at extremely high bandwidths and with low latency on application level of the data hub framework, and
- on-demand provisioning of multiple kinds of servers that execute software participating in such point-to-point transfers.

These points are illustrated in Figure 3-3: Here, a large number of external database users are accessing a group of front-end nodes. These front-end nodes generate requests to service nodes. Each service node manages in the order of ten storage servers. Storage servers are grouped into replicas with large numbers of servers. Multiple replicas form a replica set. Data is constantly synchronized across peers in a replica set. To increase availability, each replica set may span across reliability domains, which are designed so that major failures do not go beyond a domain. In case of group or node failure, either new nodes are prepared close to the failed nodes, or a complete group of nodes is prepared in a different location, depending on the scale of the failure. Each kind of node, i.e., front-end nodes, service nodes, and storage nodes, all require different hardware configurations.

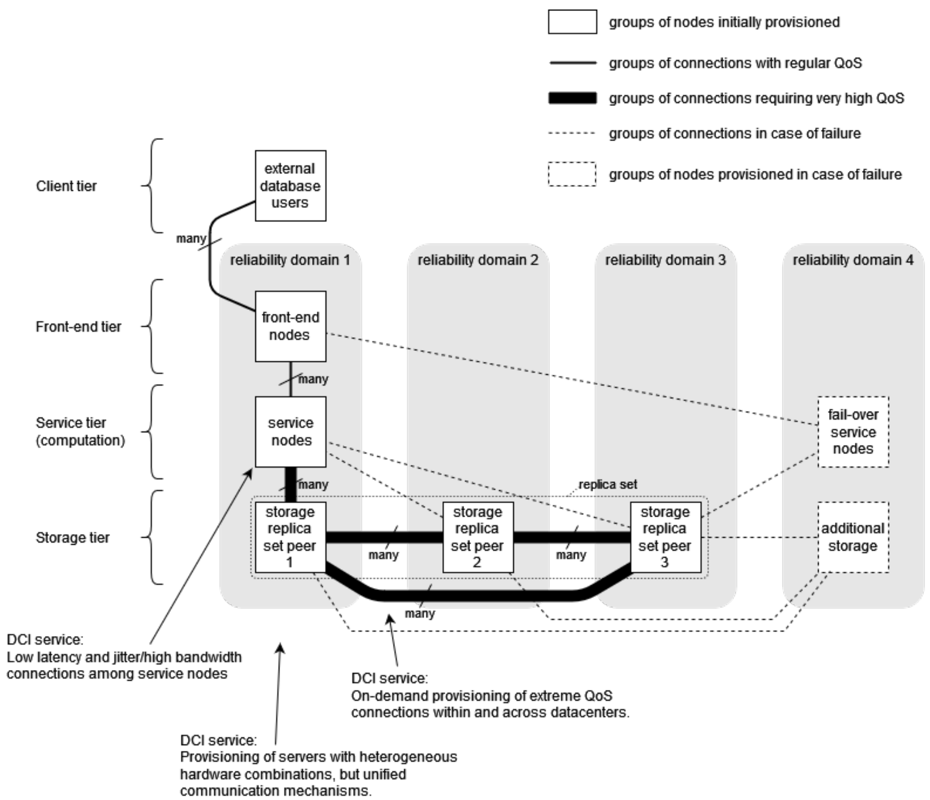

Figure 3-3: Suggested key features of the PoC in context of the IOWN GF data hub technologies.

Accordingly, DCI features realizable by early PoCs are expected to relate to the following functions:

- separating service and storage by providing high-bandwidth connections over longer ranges,
- providing multiple types of servers with heterogeneous configurations in cold or hot stand-by to cope with failures and/or user load fluctuations,
- copying extremely large amounts of data over long distance high-QoS connections to other reliability domains for the creation of new replica sets after failures, and
- providing jitter-free connections among nodes to accelerate transaction processing by supporting extremely precise time synchronization for event time-stamping.

3.3.4. Advanced DCIaaS features for managing migration and storage node configuration

Future PoC phases could demonstrate how storage nodes are created from disaggregated storage element pools tailored to user definitions, resulting in higher storage utilization.

Furthermore, disaggregation might allow to temporarily add accelerator cards from pools directly to storage nodes for specialized tasks requiring tight interaction between storage and accelerator.

Finally, future PoCs could explore how high-QoS connections could be used to streamline live and transparent migration of nodes, and in particular nodes with large amounts of storage.

3.4. Common requirements: QoS-aware communication stacks, data pipelines, and networks

The reference cases introduced above have the common requirement of high-quality point-to-point connections. Due to real-time constraints from the applications, actual implementations will have minimum required quality of service levels that must be maintained for the application to be feasible.

For initial commercialization, service quality close to the maximum achievable by current available transceiver technologies will likely be sufficient (depending on the reference case, e.g., bandwidth of 25G ~ 800G Ethernet transceivers). The reference case implementations are expected to still benefit from even higher service quality, though.

Currently, such transceiver technologies are mainly used to construct the links of back-bone networks, for high-performance computing applications, or for other special applications with direct port-to-port connections such as mobile front-haul links. However, the reference cases require communication with the quality of the service directly provided by such transceivers to be available from application software, between different kinds of computing devices, and over networks within or even among data centers.

To summarize, the above reference cases demand that

- communication mechanisms provide bandwidth and latency parameters close to what current transceiver technologies provide on direct links,
- minimum service quality of connections, such as for bandwidth, latency, jitter, is requestable by applications, and that
- such communication means are available to the application layer for point-to-point communication.

Below, communication means providing any primitive fulfilling these requirements to user applications are referred to as *“QoS-aware communication stacks”*.

Furthermore, these communication stacks are expected to provide various communication abstractions. The most prominent abstraction, also required by all use-cases above, is expected to be the data pipeline abstraction, with the same requirements as described for the communication stacks.

Such higher-level abstractions will analogously be called *“QoS-aware data pipelines”*.

Finally, for commercially viable implementations, it is further expected that in addition to the ability to provide QoS guarantees in the same range of what transceivers are capable of, the underlying communication mechanism will need to provide switching capabilities among a large number of participants. What kinds of switching are realized is left to implementers.

Such networks will be called *“QoS-aware networks”*.

4. Desired features for initial PoCs

The purpose of IOWN GF technologies is to drastically improve the energy efficiency of computing and communication infrastructure, and at the same time provide means to realize new use-cases with extreme requirements toward communication.

To satisfy the corresponding application requirements, IOWN GF DCIaaS comprises two main aspects. First, DCI provides a composable computing infrastructure supporting the use of various heterogeneous computing resources. Second, DCI provides QoS-aware communication stacks and data pipelines (cf. Section 3.4) supporting extremely high quality-of-service (QoS), such as bounded minimum available bandwidth, bounded maximum latency, and bounded maximum jitter.

Accordingly, for initial PoC activities, IOWN GF proposes that PoCs should aim to

1. demonstrate the viability of the proposed QoS-aware communication stacks as mentioned in Section 3.4 in the context of heterogeneous servers (LSNs) up to application level,
2. demonstrate how QoS-aware networks providing switching or routing capabilities can be implemented, and
3. demonstrate that QoS-aware communication stacks can be used to implement QoS-aware data pipelines that can be established between user application-level endpoints.

Any DCIaaS PoC Report should provide evaluation results of at least one of these three aspects. In particular, initial PoCs targeting only aspect 1 and/or 3 might decide to proceed using an interconnect/network with very limited switching or routing capabilities.

Below, the desirable features for PoC logical service node (LSN) implementations are further discussed in Section 4.1, and the desirable features for PoC communication mechanisms are highlighted in Section 4.2. Finally, the desired features for PoCs of extreme QoS application-level data-pipelines with are presented in Section 4.3.

4.1. Composing heterogeneous computing systems by enabling and releasing of LSNs

The functionality of the envisioned feature is discussed in Section 4.1.1, and following, an example of how a PoC to demonstrate the feature could be structured is outlined in Section 4.1.2.

4.1.1. Description of the desirable feature

Regarding LSNs, an initial PoC should demonstrate two or more different LSN composition patterns that effectively support IOWN GF use-cases in the same application, together with the enabling and release of such LSNs. For example, a minimal implementation could prepare a number of physical servers that have the accelerator and network cards installed as required by the PoC evaluation application as LSN, and then realize enabling and releasing of these LSNs by booting up and shut down physical servers via remote commands. Many more possibilities exist to realize LSNs, e.g., using cloud management frameworks or system bus extension hardware. PoC implementers are free to choose technologies to realize LSNs.

Together with enabling and releasing LSNs, a software stack should be demonstrated that

- provisions and enables use of communication links with extreme traffic parameters (e.g., latency, jitter, bandwidth) as required by IOWN GF use-cases and technologies between these LSNs and
- allows integrating and removing LSNs from the communication network.

PoC Reports should explain how LSNs are realized by the PoC and mention the concrete technologies employed as far as possible.

Furthermore, if for initial PoCs as suggested above each LSN is realized by an individual dedicated physical server, then the server system bus would correspond to the DCI intra-node interconnect, and the network connecting LSNs/servers would correspond to the DCI inter-node interconnect.

4.1.2. Example of a possible PoC demonstrating the feature

An example of such a straight-forward set-up is illustrated in Figure 4-1: three servers equipped with different accelerator types, NICs, and CPUs implement three LSNs. In the following, accelerators, such as smart network interface cards (SmartNICs), data processing units (DPUs), and infrastructure processing units (IPUs), and other extension cards that are equipped with on-board network interfaces are called *network-capable functional cards* (NCFCs). Typical accelerators of today such as graphics processing units (GPUs) or field-programmable gate arrays (FPGAs) that are not equipped with an Ethernet interface, but instead need to rely on other components to provide network connectivity, such as NICs, are called *intra-node functional cards* (INFCs). In this PoC, INFCs may communicate directly or rely on CPUs or other components to utilize network interfaces (illustrated by dotted lines). The DCI FA technical report [DCIFA](Sec. 3.2.1) provides further information about DCI NCFCs and INFCs. The NCFCs and NICs are connected by direct cables or by other means, such as network switches. At the beginning of a PoC demonstration, some servers, e.g., server #1 and server #2 are booted and connections are established (shown by solid lines) between NCFCs of these servers that have different types (e.g., NCFCs of different types might correspond to different types of hardware accelerators). In a following demonstration phase, connections to server #2 could be torn down and server #2 be shut down. In exchange, after that server #3 could be booted up, and connections could be established (shown by dashed lines), again between different types of NCFCs and/or CPUs.

The example in Figure 4-1 displays three different node types and connections using many combinations of functional cards. This large variety is shown primarily for illustrating the many possible options of node types and connection endpoints.

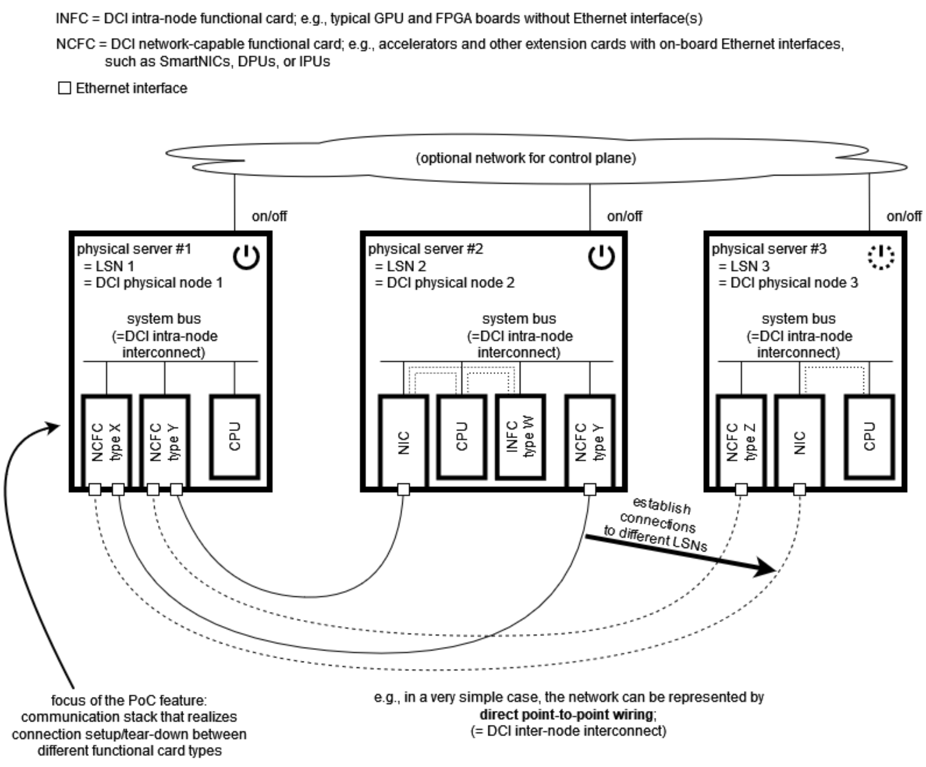

Figure 4-1: Example of heterogeneous systems composition PoC structure.

4.2. QoS-aware network supporting links with extreme QoS between components of different LSNs

The functionality of the envisioned feature is discussed in Section 4.2.1, and following, an example of how a PoC to demonstrate the feature could be structured is outlined in Section 4.2.2.

4.2.1. Description of the desirable feature

Regarding communication mechanisms for IOWN GF use-cases, PoCs should demonstrate interconnect or network technologies that could serve as a basis to implement DCI cluster inter-node interconnects, which connect DCI physical nodes, and/or DCI gateways.

Such communication mechanisms

- should support extreme traffic parameters (e.g., latency, jitter, bandwidth) as required by IOWN GF use-cases and technologies and
- should be QoS-managed.

Possible technologies that might be usable to realize such a network PoC include data center bridging (DCB) networks and optical layer 1 switches, but PoC implementers are encouraged to use the technologies of their choice.

PoC Reports should explain how communication mechanisms are realized and mention the concrete technologies used as far as possible.

4.2.2. Example of a possible PoC demonstrating the feature

An example of a straight-forward PoC structure is illustrated in Figure 4-2: a number of different servers representing different LSNs and DCI physical nodes contain NICs and/or NCFCs (cf. Section 4.1.2). The network interfaces of these cards are connected to the QoS-aware network to be demonstrated as the PoC feature. Test software programs on the NICs, NCFCs, or CPUs perform data transfers, demonstrating extreme traffic parameters and QoS isolation.

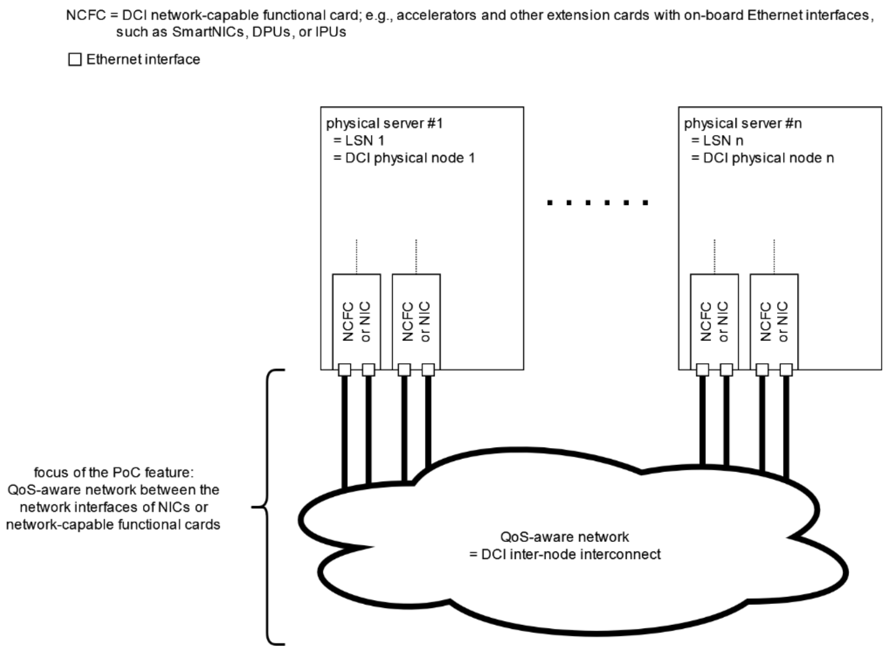

Figure 4-2: Example of QoS-aware network PoC structure.

4.3. QoS-aware data pipelines between application level endpoints

The functionality of the desirable feature is discussed in Section 4.3.1, and following, an example of how a PoC to demonstrate the feature could be structured is outlined in Section 4.3.2.

4.3.1. Description of the desirable feature

As of today, applications can hardly make use of the fastest available communication mechanisms directly. For example, common transfer protocols such as transmission control protocol/internet protocol (TCP/IP) socket pipelines typically incur severe performance penalties over long-range high-bandwidth links, such as 400 Gbps Ethernet. Therefore, PoCs should demonstrate that QoS-aware data pipelines are able to exploit the capabilities of the underlying network as far as possible.

Such software constructs in between application level communication endpoints

- should support extreme traffic parameters (e.g., latency, jitter, bandwidth) as required by IOWN GF use-cases and technologies,
- should support connections between LSN components from different DCI physical nodes,
- should be integrated into a protocol stack that reaches up to application level,
- should provide data pipeline semantics to the application software,
- should be configurable regarding the provided QoS level,

- could further allow connections between heterogeneous components of LSNs, e.g.,
  - between NCFCs (e.g., SmartNICs, DPUs, or IPUs; cf. Section 4.1.2) of different types, or
  - between CPUs and storage elements.

PoC Reports should explain how such pipelines are implemented, considering both hardware and software aspects. The underlying base technologies should be mentioned as far as possible.

4.3.2. Example of a possible PoC demonstrating the feature

An example of a straight-forward PoC structure is illustrated in Figure 4-3: Two different servers represent two LSNs. On both servers, a benchmark software is executed to evaluate QoS-aware data pipelines. These pipeline constructs are provided to the application via a QoS-aware communication stack. The QoS-aware data pipeline may be implemented by a software library, using lower-level primitives provided by a network driver, or may be provided directly by hardware. The benchmark software may be executed on CPUs, NCFCs, or on one CPU and one NCFC on the other server. In case the benchmark software is executed on a CPU and the server is connected via a NIC, then the server system bus also is an integral part of the hardware implementing the QoS-aware data pipeline. The network interfaces of NICs and NCFCs should feature network interfaces within one order of the fastest available NIC interface products available today, such as 100 Gbps Ethernet or faster. Finally, the two network interfaces between the two servers may be connected directly with a network cable or by other means.

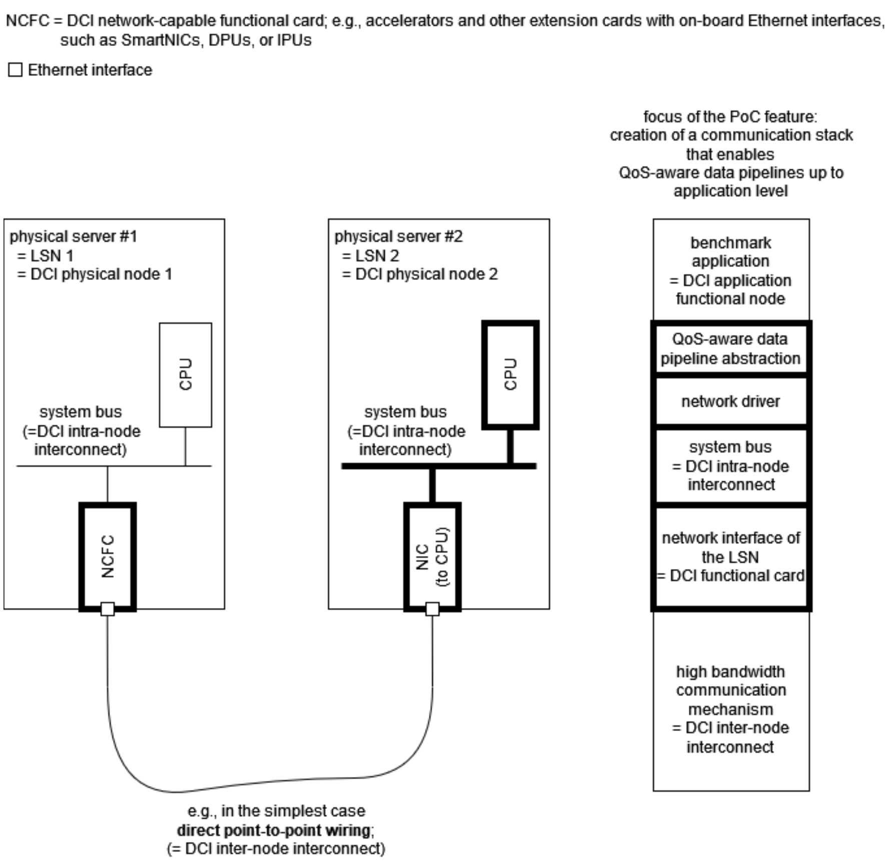

Figure 4-3: Example of QoS-aware data pipeline PoC structure.

5. Expected benchmarks

This section identifies the benchmarks that PoC implementers are expected to conduct. The most important key performance indicators (KPIs) identified by IOWN GF are *energy efficiency, network service quality, and applicability for time synchronization*. Benchmarks for each KPI are outlined in Section 5.1, Section 5.2, and Section 5.3, respectively. For each benchmark, the corresponding section briefly outlines the use-case background of the benchmark and the particular KPIs for evaluation.

5.1. Energy efficiency

This section revisits the background of the proposed energy efficiency benchmark in Section 5.1.1 and then highlights the KPIs to focus on in Section 5.1.2.

5.1.1. Background of the benchmark

One of the primary goals of IOWN GF is to propose technologies that will contribute to reducing the energy consumption of data processing and data transfers. Therefore, PoC conductors are encouraged to report how the technology demonstrated by their PoC will influence the overall system energy efficiency of IOWN GF use-case applications.

5.1.2. KPIs to evaluate

The metric for benchmarks should be selected based on the nature of the application the benchmark pertains to: When continuously processing of data with strict real-time requirements is evaluated, the total amount of Watts consumed by the overall PoC system, to reliably sustain real-time operation should be used as comparison metric. On the other hand, in all other cases when evaluating data processing of discrete jobs or processing without strict real-time requirements, the total amount of Watt-seconds (Joules) consumed by the system (or alternatively, average Watt by seconds) for completing the tasks within a reasonable time frame should be used as comparison metric. Furthermore, the volume and transfer rate of data should be reported. For PoCs that focus mainly on communication, this value should also be used to additionally give normalized power and energy consumption values.

PoC implementers are free to choose a measuring methodology adequate to the conducted benchmarks, and are expected to report about the benchmark methodology.

5.2. Network service quality

This section revisits the background of the proposed network service quality benchmark in Section 5.2.1 and then highlights the KPIs to focus on in Section 5.2.2.

5.2.1. Background of the benchmark

Applications currently under consideration by IOWN GF have in common that these applications all require QoS-aware networks that combing QoS awareness with extremely high technical specifications by today's standards (also cf. Section 3.4). The required QoS is typically so high that with currently available technology, either complex solutions need to be tailored to the specific use-case and user, or the required QoS is so high that the requirements cannot be fulfilled in a reasonable economic manner at all. Therefore, conducted benchmarks should demonstrate that the realized PoC communication mechanism removes some boundaries of today's technologies regarding IOWN GF use-cases. Accordingly, such benchmarks should be independent of and do not take into account throughput, latency, and jitter limitations due to delays in software processing.

5.2.2. KPIs to evaluate

The key service quality properties of communication mechanisms for IOWN GF use-cases are

- minimum communication latency,
- minimum jitter (variation of the latency),
- maximum bandwidth,
- corresponding packet drop rates (alternatively bit error rates), and
- network service availability (calculated per different time quanta)

of end-to-end connections between

- LSN ports (server NIC ports) and
- endpoints in application software;

measurements should be performed between different LSNs that are located on DCI physical nodes and thus communicating via a DCI inter-node interconnect.

PoC Reports should include measurement results of the above KPIs and descriptions of the benchmark setups.

5.3. Applicability of the QoS-aware network for time synchronization

This section revisits the background of the proposed applicability for time synchronization benchmark in Section 5.3.1 and then highlights the KPIs to focus on in Section 5.3.2.

5.3.1. Background of the benchmark

Time synchronization between servers today is used widely in multiple use-cases. For example, VRAN/ORAN use-cases require radio functions, e.g., RUs and DUs, operate with precise time synchronization and also require synchronous data pipelines. Furthermore, database use-cases may use time synchronization to realize high-performance transaction processing: For example, distributed databases that need to guarantee data consistency (safety) may employ a common time base among servers to remove the need for centralized transaction processing by combining the Paxos consensus algorithm with high-precision time sources. For such database implementations, the transaction processing throughput is bounded by the achievable precision of the time synchronization. Therefore, to improve clock synchronization, current realizations of such databases are typically using specialized hardware such as atomic clocks and/or global positioning system (GPS) signal receivers.

Networks applicable as DCI inter-node interconnect will need to be able to provide communication with bounded latency and jitter. Such communication networks are a natural fit for time synchronization over networks. Therefore, implementers of PoCs that comprise networks as described in Section 3.2 are encouraged to evaluate and report the attainable precision of time synchronization over such networks.

For example, when a PoC implements a high-performance network with low jitter as a feature as described in Section 4.2, a straight-forward experiment setup might consist of two or more microservices running on a DCI and synchronizing their clocks via well-established protocols such as the IEEE 1588 precision time protocol (PTP) over this network.

5.3.2. KPIs to evaluate

The key service quality to evaluate in this benchmark is with what time precision events (or actions) on different physical nodes (servers) can be aligned in time. In this context, such events (or actions) could be

- the reception of a data packet by a SmartNIC or
- the execution of a particular machine instruction of any processing element; such instructions could be used to time-stamp events or actions in application software such as:
    - sending of time-sensitive data or
    - receiving of time-sensitive data.

PoC Reports including time synchronization are expected to report the measured upper bound of the time skew and deviation and how time synchronization is implemented.

If in a real-world situation signals used for such clock synchronization would share a physical link with other data, then the PoC experiments should be performed with different levels of background traffic to simulate such real-world situations.

6. Summary

The sections above give an overview of the DCIaaS PoCs as envisioned by IOWN GF. First, the context of this PoC Reference is introduced in Section 1. Following, the scope of the proposed PoC activities are outlined in Section 2. After that, reference cases to illustrate the background how technologies of the PoCs could be used to overcome current technology gaps are highlighted in Section 3. Then, desired features of initial PoCs are given in Section 4, and finally, expected benchmarks to evaluate implementations of such features are detailed in Section 5.

IOWN GF is looking forward to receive PoC Reports. The experiences in these reports will be reviewed and used to improve and guide the development of the IOWN GF DCI specifications.

Annex A. Outlook: future DCIaaS PoC roadmap

The sections above describe the IOWN GF technology features deemed most important by IOWN GF members: heterogeneous LSN implementations, a QoS-aware network (cf. Section 3.4) in-between, and QoS-aware data pipelines. IOWN GF suggests to prioritize these features for the first PoC step.

In addition to the three features highlighted so far, IOWN GF DCI technologies will provide many more other features. This section lays out in which order these features could be demonstrated in two further PoC steps in the coming years. Of course, PoC implementers are most welcome to pick features from succeeding steps now and demonstrate such features earlier than listed. Table A-1 summarizes the proposed additions in each step and the timeline that the IOWN GF envisions for PoCs to be implemented in the future.

Table A-1: Envisioned timeline of PoC steps.

	STEP 1: THIS POC REFERENCE (2022)	**STEP 2 (2023-2024)**	**STEP 3 (2024-2025)**
Background: **Use-cases and operation scenarios**	• area management • VRAN • data warehouse	• AI-integrated communication • industry management	• deployment of NFV on the same edge computing infrastructure as other use-cases
Features: **LSN composition**	• LSN enabling and releasing only	• dynamic LSN creation/modification • heterogeneous resources pools	• synchronous (time-synchronized) data pipelines
Features: **DCIaaS**	• HW/SW stack for extreme QoS communication	• QoS isolation between applications on one physical node	• multi-tenancy on physical nodes • strong security and QoS isolation between tenants on a given physical node
Features: **Inter-/intra-node interconnect**	• extreme QoS communication	• QoS isolation between different data streams	

In the following, possible use-cases and features of PoC Step 2 are further introduced in Section A.1, and further features for PoC Step 3 are outlined in Section A.2.

A.1. Second step PoCs

The PoCs of the second step are envisioned to be realized between 2023 and 2024.

PoCs are expected to consider requirements from further use-cases, such as:

- combined bandwidth, latency, and massive streaming requirements of AI-integrated communication use-cases and
- extremely strict and short latency requirements of industry management remote robot control use-cases.

Possible new PoC features could include:

- Dynamic composition and modification of LSNs, so that LSNs can be automatically assembled on-the-fly from multiple resource pools providing a variety of different resources to reduce both system capital expenditure (CAPEX) as well as energy consumption.

- QoS isolation between DCIaaS applications executing on the same physical node, which is also a key feature for VRAN use-cases.
- Ability to automatically adapt to varying load conditions based on workload prediction. For example, both area management and IOWN GF technologies for mobile networks (IMN) use-cases are expected to exhibit large differences in their load profile depending on the time of day. A PoC could demonstrate how an orchestrator provisions AI inference servers for area management or DU servers for VRAN proactively and reconfigures the extreme QoS communication mechanisms accordingly.

A.2. Third step PoCs

The PoCs of the third step are envisioned to be realized between 2024 and 2025.

Regarding use-cases, PoCs are expected to showcase the genericity of the IOWN GF architecture. One possible PoC scenario would be to demonstrate that the same IOWN GF infrastructure implementation can be employed for use-cases with different requirements, such as area management and IMN VRAN (NFV).

On the feature side, IOWN GF assumes that by this third step, PoCs could already be implemented that demonstrate functionality that requires strict coordination between multiple functional blocks of DCIaaS systems, such as:

- multi-tenancy, in particular the management of applications of multiple tenants on the same physical node and
- strong security and QoS isolation among tenants and their applications.

References

REFERENCE	DESCRIPTION
[AMRIM]	IOWN Global Forum, "Reference Implementation Model (RIM) for the Area Management Security Use Case," 2022, URL: https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-RIM-for-AM-Use-Case-1.0.pdf
[APNFA]	IOWN Global Forum, "Open All-Photonic Network Functional Architecture", 2022, URL: https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Open-APN-Functional-Architecture-1.0-1.pdf
[DCIFA]	IOWN Global Forum, "Data-Centric Infrastructure Functional Architecture," 2022, URL: https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-DCI-Functional-Architecture-1.0-1.pdf
[DHFA]	IOWN Global Forum, "Data Hub Functional Architecture," 2022, URL: (https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Data-Hub-Functional-Architecture-1.0-1.pdf

History

REVISION	RELEASE DATE	SUMMARY OF CHANGES
1.0	July 21, 2022	Initial Release