Common NFVI Telco Taskforce

Reference Model Chapter 3: Infrastructure Abstraction and Modeling

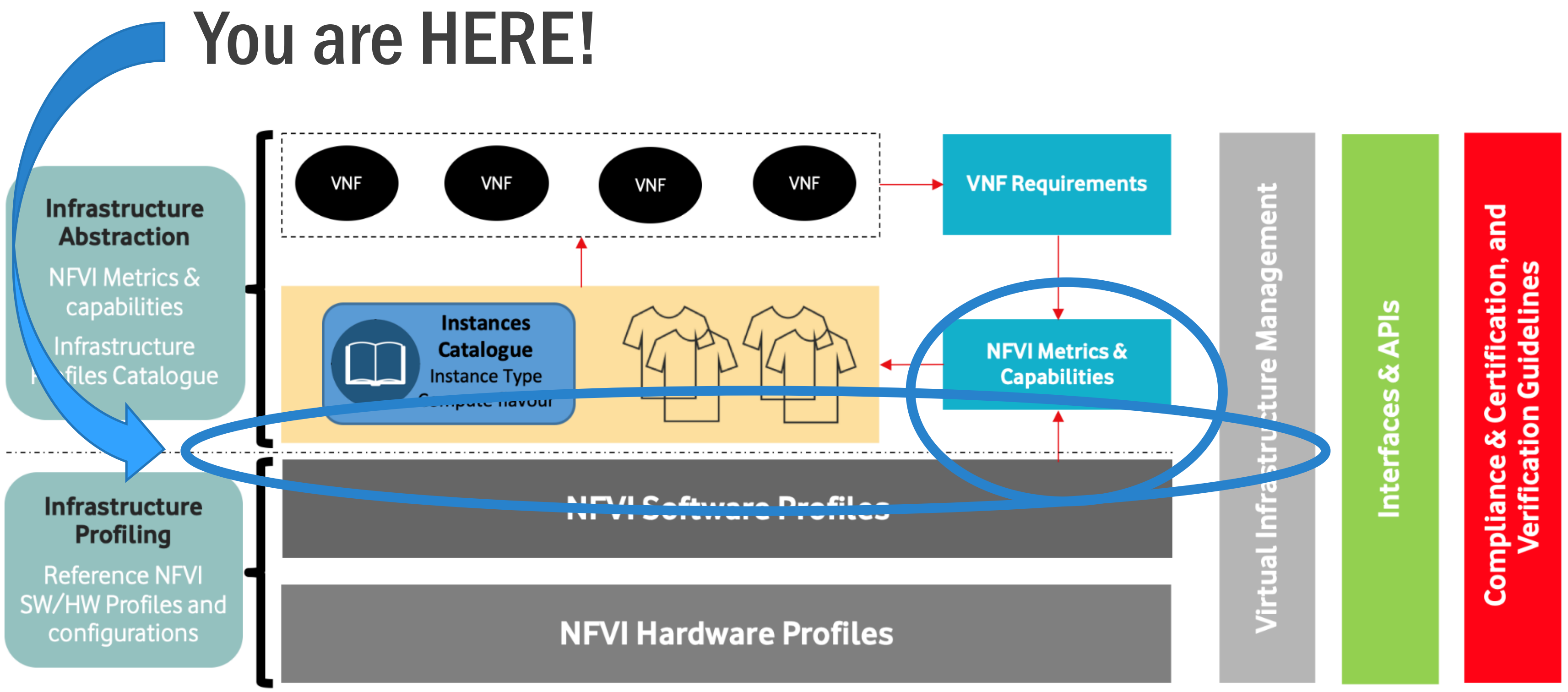
Sabine Demel, Deutsche Telekom
Mark Shostak, AT&T

23 July 2019

THE LINUX FOUNDATION

GSMA

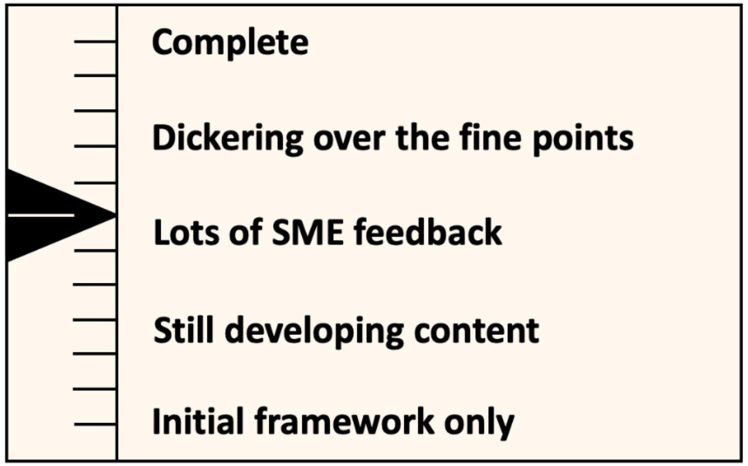
Infra Abstraction and Modeling



Modeling: Outline and Maturity

Table of Contents

Bogo-Meter rating



Infra Abstraction Purpose

The abstraction model presented in this chapter specifies a common set of virtual infrastructure resources which NFVI will need to provide to be able to host most of the typical VNF/CNF workloads required by the operator community.

The lack of a common understanding of which resources and corresponding capabilities a suitable NFVI should provide may lead to several issues which could negatively impact the time and the cost for on-boarding and maintaining these solutions on top of a virtualised infrastructure. For Example:

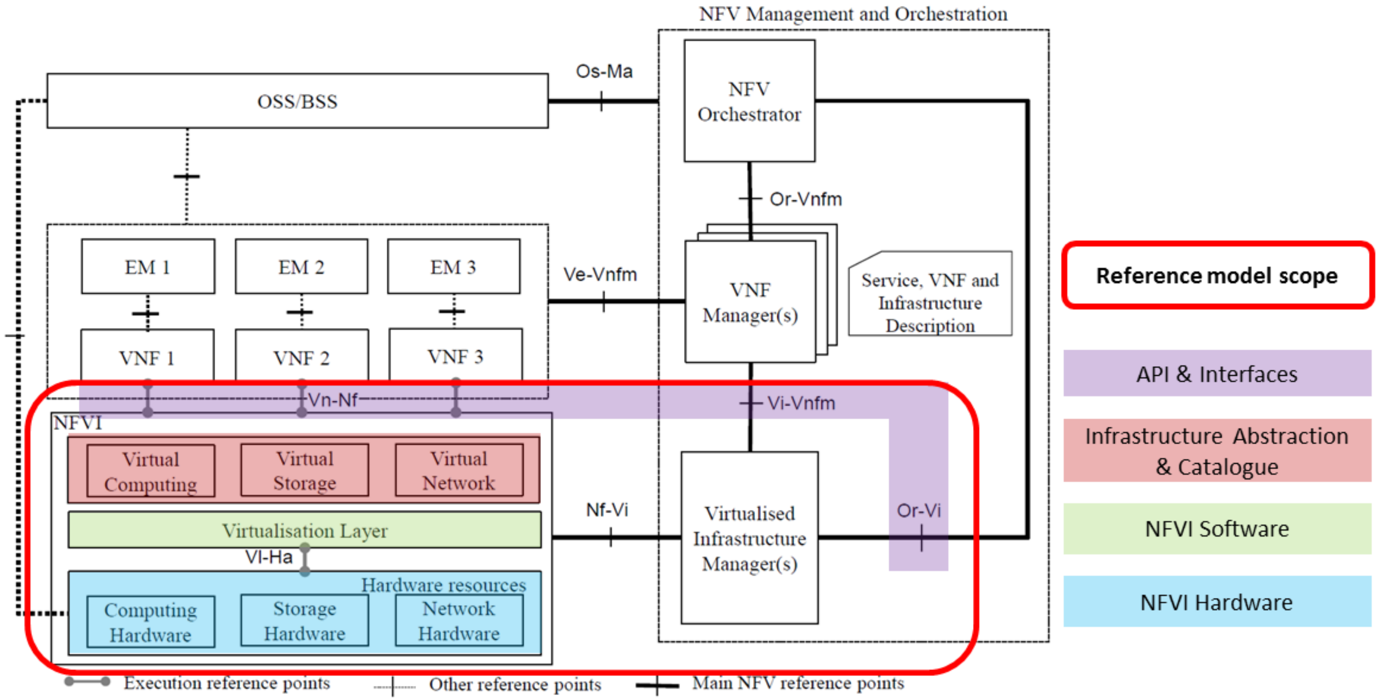
Supporting any kind of workload specific requirements (e.g. regarding network acceleration or API access) might result in having to establish different silo of NFVIs for each workload type.

* Footnote text

Model's Scope (NFVI + VIM)

Scope of Capabilities and Metrics modeling includes ETSI NFVI plus VIM

CNTT | Mapping to ETSI NFV Architecture (NFV002-v1.2.1)

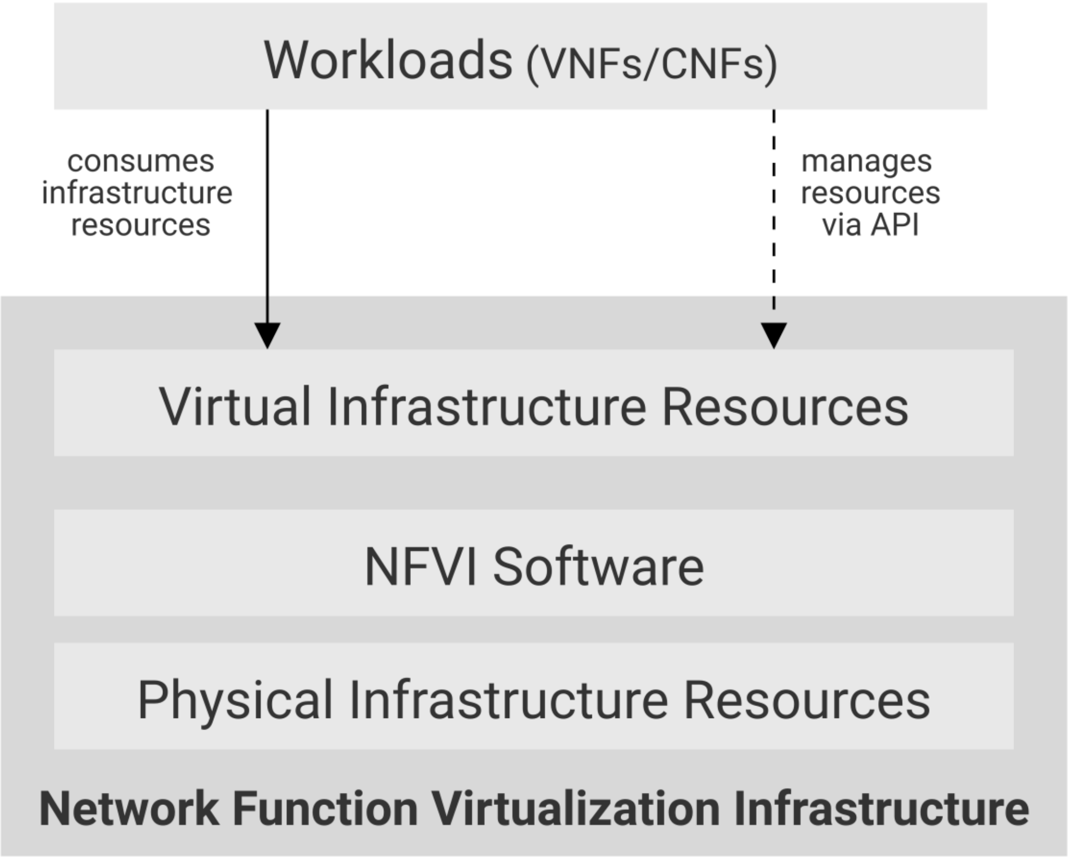


C2 General

THE LINUX FOUNDATION

GSMA

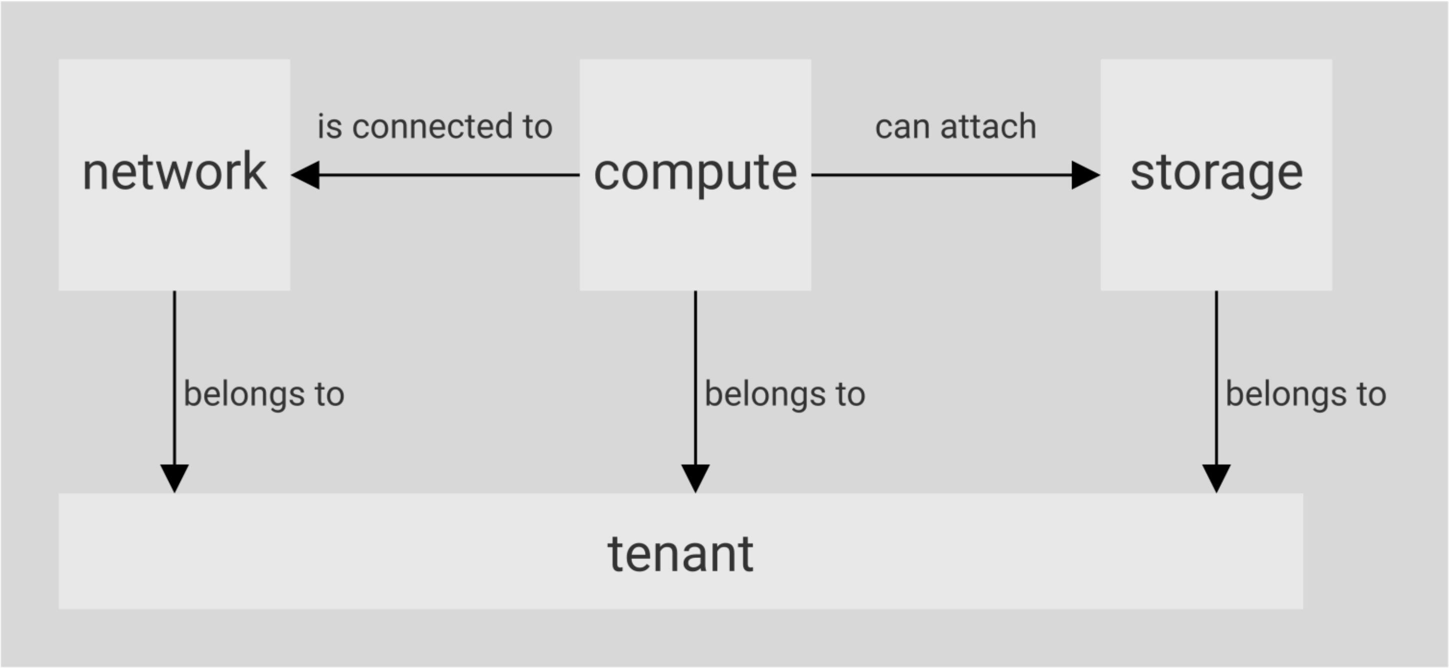
Section 3.1 – Modeling Abstraction for the NFVI



- **Physical Infrastructure Resources**: physical hardware components such as servers, (including random access memory, local storage, network ports, and hardware acceleration devices), storage devices, network devices, etc. and the basic input output system (BIOS).
- **NFVI Software**: both the host Operating System (OS) as well as the virtualization/containerization technology
- **Virtual Infrastructure Resources**: all the infrastructure resources (compute, storage and networks) which the NFVI provides to the workloads.
  These virtual resources can be managed by the tenants and tenant workloads directly or indirectly via an application programming interface (API).
- **Workloads (VNFs/CNFs)**: workloads such as virtualized and/or containerized network functions that run on top of a VM or as a Container.

Section 3.1 – Modeling Virtual infrastructure resources

- **Tenants:** represent an independently manageable logical pool of compute, storage and network resources
- **Compute resources:** represent virtualised computes for workloads and Operating and other Systems as necessary
- **Storage resources:** represent virtualised resources for persisting data
- **Network resources:** represent virtual resources providing layer 2 and layer 3 connectivity



Section 3.1 - Modeling Tenant

An NFVI needs to be capable of supporting multiple tenants and has to isolate sets of infrastructure resources dedicated to specific workloads from one another. Tenants represent an independently manageable logical pool of compute, storage and network resources abstracted from physical hardware.

Example: a tenant within an OpenStack environment or a Kubernetes cluster.

Attribute	Description
name	name of the logical resource pool
type	type of tenant (e.g. OpenStack tenant, Kubernetes cluster, ...)
vcpus	max. number of virtual CPUs
ram	max. size of random access memory in GB
disc	max. size of ephemeral disc in GB
networks	description of external networks required for inter-domain connectivity
metadata	key/value pairs for selection of the appropriate physical context (e.g. location, availability zone, ...)

Section 3.1 - Modeling Compute

A virtual machine or a container/pod belonging to a tenant capable of hosting the application components of workloads (VNFs). A virtual compute therefore requires a tenant context and since it will need to communicate with other communication partners it is assumed that the networks have been provisioned in advance.

Example: a virtual compute descriptor as defined in TOSCA Simple Profile for NFV.

Attribute	Description
name	name of the virtual host
vcpus	number of virtual cpus
ram	size of random access memory in GB
disc	size of root disc in GB
nics	sorted list of network interfaces connecting the host to the virtual networks
acceleration	key/value pairs for selection of the appropriate acceleration technology
metadata	key/value pairs for selection of the appropriate redundancy domain

Section 3.1 – Modeling Storage

A block device of a certain size for persisting information which can be created and dynamically attached to/detached from a virtual compute. A storage device resides in a tenant context and exists independently from any compute host.
Example: an OpenStack cinder volume

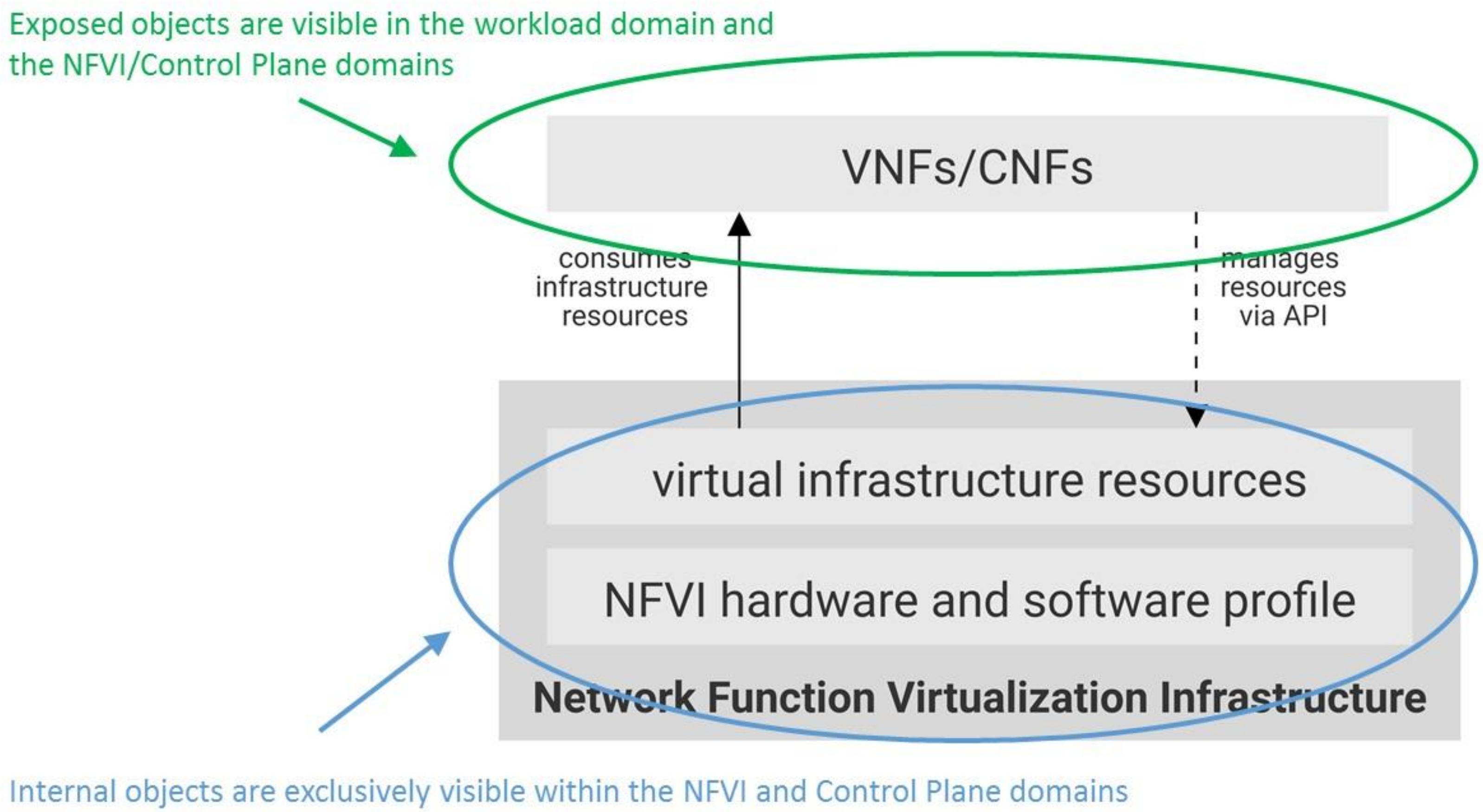
Attribute	Description
name	name of storage resources
size	size of disc in GB
attachments	list of compute hosts to which the device is currently attached
acceleration	key/value pairs for selection of the appropriate acceleration technology
metadata	key/value pairs for selection of the appropriate redundancy domain

Section 3.1 - Modeling Network

A layer 2 / layer 3 communication domain within a tenant. A network requires a tenant context.
Example: a virtual compute descriptor as defined in TOSCA Simple Profile for NFV.

Attribute	Description
name	name of the network resource
subnet	network address of the subnet
acceleration	key/value pairs for selection of the appropriate acceleration technology

Section 3.2 - Object Scope (Visibility)



Key Terminology

Capability: **Refers to a feature** that may be enabled (or disabled) for a Virtual Compute instance, based on its instance type

Metric: Refers to a measurement that can be taken during operation of the NFVI

- Available RAM or Cores
- Timing to perform function x, etc.

Constraint: A parameter whose value (or range) is determined during the Reference Architecture engineering process, and which an NFVI's Metric's must remain within

* Footnote text

Section 3.3 - Capabilities & Metrics

- Capabilities
- Metrics

- May be Exposed or Internal
- Exposed are available to both

NFVI capability	Unit
CPU pinning support	Yes/No
NUMA support	Yes/No
IPSec Acceleration	[illegible]
Crypto Acceleration	
Transcoding Accele	
Programmable Acce	

NFVI capability	Unit
Percentage of vCPU cores consumed by NFVI overhead in a compute node.	% (of total available)
Percentage of memory consumed by NFVI overhead in a compute node.	% (c avai

NFVI metric	Unit
Network throughput	bits/s or packets/s
Network latency	second
Network Delay Variation	second

Section 3.5 - VIM Capabilities & Metrics

- Brand new section
- Addresses aggregate NFVI + VIM
- Mainly a place holder today

Spoiler Alert / Sample

VIM metrics	Unit
Time to create Virtual Compute for a given VNF	Max ms
Time to delete Virtual Compute of a given VNF	Max ms
Time to start Virtual Compute of a given VNF	Max ms
Time to stop Virtual Compute of a given VNF	Max ms
Time to pause Virtual Compute of a given VNF	Max ms
Time to create internal virtual network	Max ms
Time to delete internal virtual network	Max ms

Areas to be Developed and Q&A

- Clarification and cleanup
  - Clearer definitions
  - Ensure/enforce consistency
- Acceleration
- Additional metrics, capabilities and constraints as needed
  - TCAs
- Lifecycle Framework Closed-Loop Feedback
  - from work on Architecture and Implementation
  - from work with OPNFV

Thank You!