1 Overall Principle for High Availability in NFV

The ultimate goal for the High Availability schema is to provide high availability to the upper layer services.

High availability is provided by the following steps once a failure happens:

Step 1: failover of services once failure happens and service is out of work.

Step 2: Recovery of failed parts in each layer.

1.1 Framework for High Availability in NFV

Framework for Carrier Grade High availability:

A layered approach to availability is required for the following reasons:

- fault isolation
- fault tolerance
- fault recovery

Among the OPNFV projects the OPNFV-HA project's focus is on requirements related to service high availability. This is complemented by other projects such as the OPNFV - Doctor project, whose focus is reporting and management of faults along with maintenance, the OPNFV-Escalator project that considers the upgrade of the NFVI and VIM, or the OPNFV-Multisite that adds geographical redundancy to the picture.

A layered approach allows the definition of failure domains (e.g., the networking hardware, the distributed storage system, etc.). If possible, a fault shall be handled at the layer (failure domain) where it occurs. If a failure cannot be handled at its corresponding layer, the next higher layer needs to be able to handle it. In no case, shall a failure cause cascading failures at other layers.

The layers are:

Service	**End customer visible service**
Application	VNF's, VNFC's
NFVI/VIM	Infrastructure, VIM, VNFM, VM
Hardware	Servers, COTS platforms

The following document describes the various layers and how they need to address high availability.

1.2 Definitons

Reference from the ETSI NFV doc.

Availability: Availability of an item to be in a state to perform a required function at a given instant of time or at any instant of time within a given time interval, assuming that the external resources, if required, are provided.

Accessibility: It is the ability of a service to access (physical) resources necessary to provide that service. If the target service satisfies the minimum level of accessibility, it is possible to provide this service to end users.

Admission control: It is the administrative decision (e.g. by operator's policy) to actually provide a service. In order to provide a more stable and reliable service, admission control may require better performance and/or additional resources than the minimum requirement. Failure: deviation of the delivered service from fulfilling the system function.

Fault: adjudged or hypothesized cause of an error

Service availability: service availability of <Service X> is the long-term average of the ratio of aggregate time between interruptions to scheduled service time of <Service X> (expressed as a percentage) on a user-to-user basis. The time between interruptions is categorized as Available (Up time) using the availability criteria as defined by the parameter thresholds that are relevant for <Service X>.

According to the ETSI GS NFV-REL 001 V1.1.1 (2015-01) document service availability in the context of NFV is defined as End-to-End Service availability.

Service Availability refers to the End-to-End Service Availability which includes all the elements in the end-to-end service (VNFs and infrastructure components) with the exception of the customer terminal. This is a customer facing (end user) availability definition and it is the result of accessibility and #admission control (see their respective definitions above).

Service Availability=total service available time/
(total service available time + total restoration time)

Service continuity: Continuous delivery of service in conformance with service's functional and behavioral specification and SLA requirements, both in the control and data planes, for any initiated transaction or session until its full completion even in the events of intervening exceptions or anomalies, whether scheduled or unscheduled, malicious, intentional or unintentional.

The relevant parts in NFV-REL: The basic property of service continuity is that the same service is provided during VNF scaling in/out operations, or when the VNF offering that service needs to be relocated to another site due to an anomaly event (e.g. CPU overload, hardware failure or security threat).

Service failover: when the instance providing a service/VNF becomes unavailable due to fault or failure, another instance will (automatically) take over the service, and this whole process is transparent to the user. It is possible that an entire VNF instance becomes unavailable while providing its service.

Service failover time: Service failover is when the instance providing a service becomes unavailable due to a fault or a failure and another healthy instance takes over in providing the service. In the HA context this should be an automatic action and this whole process should be transparent to the user. It is possible that an entire VNF instance becomes unavailable while providing its service.

Failure detection: If a failure is detected, the failure must be identified to the component responsible for correction.

Failure detection time: Failure detection time is the time interval from the moment the failure occurs till it is reported as a detected failure.

Alarm: Alarms are notifications (not queried) that are activated in response to an event, a set of conditions, or the state of an inventory object. They also require attention from an entity external to the reporting entity (if not then the entity should cope with it and not raise the alarm).

Alarm threshold condition detection: Alarm threshold condition is detected by the component responsible for it. The component periodically evaluates the condition associated with the alarm and if the threshold is reached, it generates an alarm on the appropriate channel, which in turn delivers it to the entity(ies) responsible, such as the VIM.

Alarm threshold detection time: The threshold time interval between the metrics exceeding the threshold and the alarm been detected.

Service recovery: The restoration of the service state after the instance of a service/VNF is unavailable due to fault or failure or manual interruption.

Service recovery time: Service recovery time is the time interval from the occurrence of an abnormal event (e.g. failure, manual interruption of service, etc.) until recovery of the service.

SAL: Service Availability Level.

1.3 Overall requirements

Service availability shall be considered with respect to the delivery of end to end services.

- There should be no single point of failure in the NFV framework.
- All resiliency mechanisms shall be designed for a multi-vendor environment, where for example the NFVI, NFV-MANO, and VNFs may be supplied by different vendors.

- Resiliency related information shall always be explicitly specified and communicated using the reference interfaces (including policies/templates) of the NFV framework.

1.4 Time requirements

The time requirements below are examples in order to break out of the failure detection times considering the service recovery times presented as examples for the different service availability levels in the ETSI GS NFV-REL 001 V1.1.1 (2015-01) document.

The table below maps failure modes to example failure detection times.

Failure Mode	Time
Failure detection of HW	<1s
Failure detection of virtual resource	<1s
Alarm threshold detection	<1min
Failure detection over of SAL 1	<1s
Recovery of SAL 1	5-6s
Failure detection over of SAL 2	<5s
Recovery of SAL 2	10-15s
Failure detection over of SAL 3	<10s
Recovery of SAL 3	20-25s

2 Hardware HA

The hardware HA can be solved by several legacy HA schemes. However, when considering the NFV scenarios, a hardware failure will cause collateral damage to not only to the services but also virtual infrastructure running on it.

A redundant architecture and automatic failover for the hardware are required for the NFV scenario. At the same time, the fault detection and report of HW failure from the hardware to VIM, VNFM and if necessary the Orchestrator to achieve HA in OPNFV. A sample fault table can be found in the Doctor project. (https://wiki.opnfv.org/doctor/faults) All the critical hardware failures should be reported to the VIM within 1s.

Other warnings for the hardware should also be reported to the VIM in a timely manner.

2.1 General Requirements

- Hardware Failures should be reported to the hypervisor and the VIM.
- Hardware Failures should not be directly reported to the VNF as in the traditional ATCA architecture.

- Hardware failure detection message should be sent to the VIM within a specified period of time, based on the SAL as defined in Section 1.
- Alarm thresholds should be detected and the alarm delivered to the VIM within 1min. A certain threshold can be set for such notification.
- Direct notification from the hardware to some specific VNF should be possible. Such notification should be within 1s.
- Periodical update of hardware running conditions (operational state?) to the NFVI and VIM is required for further operation, which may include fault prediction, failure analysis, and etc.. Such info should be updated every 60s.
- Transparent failover is required once the failure of storage and network hardware happens.
- Hardware should support SNMP and IPMI for centralized management, monitoring and control.

2.2 Network plane Requirements

- The hardware should provide a redundant architecture for the network plane.
- Failures of the network plane should be reported to the VIM within 1s.
- QoS should be used to protect against link congestion.

2.3 Power supply system

- The power supply architecture should be redundant at the server and site level.
- Fault of the power supply system should be reported to the VIM within 1s.
- Failure of a power supply will trigure automatic failover to the redundant supply.

2.4 Cooling system

- The architecture of the cooling system should be redundant.
- Fault of the cooling system should be reported to the VIM within 1s.
- Failure of the cooling system will trigger automatic failover of the system.

2.5 Disk Array

- The architecture for the disk array should be redundant.
- Fault of the disk array should be reported to the VIM within 1s
- Failure of the the disk array will trigger automatic failover of the system

support for protected cache after an unexpected power loss.

- Data shall be stored redundantly in the storage backend (e.g., by means of RAID across disks).
- Upon failures of storage hardware components (e.g., disks services, storage nodes) automatic repair mechanisms (re-build/re-balance of data) shall be triggered automatically.
- Centralized storage arrays shall consist of redundant hardware.

2.6 Servers

- Support precise timing with accuracy higher than 4.6ppm.

3 Virtualization Facilities (Host OS, Hypervisor)

3.1 Requirements on Host OS and Hypervisor and Storage

Requirements:

- The hypervisor should support distributed HA mechanism.
- Hypervisor should detect the failure of the VM. Failure of the VM should be reported to the VIM within 1s
- The hypervisor should report (and if possible log) its failure and recovery action. And the destination to whom they are reported should be configurable.
- The hypervisor should support VM migration.
- The hypervisor should provide isolation for VMs, so that VMs running on the same hardware do not impact each other.
- The host OS should provide sufficient process isolation so that VMs running on the same hardware do not impact each other.
- The hypervisor should record the VM information regularly and provide logs of VM actions for future diagnoses.
- The NFVI should maintain the number of VMs provided to the VNF in the face of failures. I.e. the failed VM instances should be replaced by new VM instances

3.2 Requirements on Middlewares

Requirements:

- It should be possible to detect and automatically recover from hypervisor failures without the involvement of the VIM.
- Failure of the hypervisor should be reported to the VIM within 1s.
- Notifications about the state of the (distributed) storage backends shall be send to the VIM (in-synch/healthy, re-balancing/re-building, degraded).
- Process of VIM runing on the compute node should be monitored, and failure of it should be notified to the VIM within 1s.
- Fault detection and reporting capability. There should be middlewares supporting in-band reporting of HW failure to VIM.
- Storage data path traffic shall be redundant and fail over within 1 second on link failures.
- Large deployments using distributed software-based storage shall separate storage and compute nodes (non-hyperconverged deployment).
- Distributed software-based storage services shall be deployed redundantly.
- Data shall be stored redundantly in distributed storage backends.
- Upon failures of storage services, automatic repair mechanisms (re-build/re-balance of data) shall be triggered automatically.
- The storage backend shall support geo-redundancy.

4 Virtual Infrastructure HA – Requirements:

This section is written with the goal to ensure that there is alignment with Section 4.2 of the ETSI/NFV REL-001 document.

Key reference requirements from ETSI/NFV document:

[Req.4.2.12] On the NFVI level, there should be a transparent fail-over in the case of for example compute, memory, storage or connectivity failures.

- The virtual infrastructure should provide classified virtual resource for different SAL VNFs. Each class of the resources should have guaranteed performance metrics.
- Specific HA handling schemes for each classified virtual resource, e.g. recovery mechanisms, recovery priorities, migration options, should be defined.
- The NFVI should maintain the number of VMs provided to the VNF in the face of failures. I.e. the failed VM instances should be replaced by new VM instances.

4.1 Compute

VM including CPU, memory and ephemeral disk.

Requirements:

- Detection of failures must be sub 1 second.
- Recovery of a failed VM (VNF) must be automatic. The recovery must re-launch the VM based on the required initial state defined in the VNFD.
- On evacuation, fencing of instances from an unreachable host is required.
- Resources of a migrated VM must be evacuated once the VM is migrated to a different compute node, placement policies must be preserved. For example during maintenance activities.
- Failure detection of the VNF software process is required in order to detect the failure of the VNF sufficiently. Detection should be within less than 1 second.

4.2 Network

4.2.1 Virtual network

Requirements:

- Redundant top of rack switches must be supported as part of the deployment.
- Static LAG must be supported to ensure sub 50ms detection and failover of redundant links between nodes. The distributed virtual router should support HA.
- Service provided by network agents should be highly available (L3 Agent, DHCP agent as examples).
- L3-agent, DHCP-agent should clean up network artifacts (IPs, Namespaces) from the database in case of failover.

4.2.2 vSwitch

Requirements:

- Monitoring and health of vSwitch processes is required.
- The vSwitch must adapt to changes in network topology and automatically
- Support recovery modes in a transparent manner.

4.3.3 Link Redundancy

Requirements:

- The ability to manage redundant interfaces and support of LAG on the compute node is required.
- Support of LAG on all interfaces, internal platform control interfaces, internal platform storage interfaces, as well as interfaces connecting to provide networks.
- LACP is optional for dynamic management of LAG links.
- Automated configuration LAG should support active/standby and balanced modes. Should adapt to changes in network topology and automatically support recovery modes in a transparent manner.
- In SR-IOV scenario, link redundancy could not be transparent, VM should have two ports directly connect to physical port on host. Then app may bind these two ports for HA.

5 VIM High availability

The VIM in the NFV reference architecture contains all the control nodes of OpenStack, SDN controllers and hardware controllers. It manages the NFVI according to the instructions/requests of the VNFM and NFVO and reports them back about the NFVI status. To guarantee the high availability of the VIM is a basic requirement of the OPNFV platform. Also the VIM should provide some mechanism for VNFs to achieve their own high availability.

5.1 Architecture requirement of VIM HA

The architecture of the control nodes should avoid any single point of failure and the management network plane which connects the control nodes should also be redundant. Services of the control nodes which are stateless like nova-API, glance-API etc. should be redundant but without data synchronization. Stateful services like MySQL, Rabbit MQ, SDN controller should provide complex redundancy policies. Cloud of different scale may also require different HA policies.

Requirement:

- In small scale scenario active-standby redundancy policy would be acceptable.
- In large scale scenario all stateful services like database, message queue, SDN controller should be deployed in cluster mode which support N-way, N+M active-standby redundancy.
- In large scale scenario all stateless services like nova-api, glance-api etc. should be deployed in all active mode.

- Load balance nodes which introduced for all active and N+M mode should also avoid the single point of failure.
- All control node servers shall have at least two network ports to connect to different networks plane. These ports shall work in bonding manner.
- Any failures of services in the redundant pairs should be detected and switch over should be carried out automatically in less than 5 seconds totally.
- Status of services must be monitored.

5.2 Fault detection and alarm requirement of VIM

Redundant architecture can provide function continuity for the VIM. For maintenance considerations all failures in the VIM should be detected and notifications should be triggered to NFVO, VNFM and other VIM consumers.

Requirement:

- All hardware failures of control nodes should be detected and relevant alarms should be triggered. OSS, NFVO, VNFM and other VIM consumers can subscribe these alarms.
- Software on control nodes like OpenStack or ODL should be monitored by the clustering software at process level and alarms should be triggered when exceptions are detected.
- Software on compute nodes like OpenStack/nova agents, ovs should be monitored by watchdog. When exceptions are detected the software should be restored automatically and alarms should be triggered.
- Software on storage nodes like Ceph, should be monitored by watchdog. When exceptions are detected the software should be restored automatically and alarms should be triggered.
- All alarm indicators should include: Failure time, Failure location, Failure type, Failure level.
- The VIM should provide an interface through which consumers can subscribe to alarms and notifications.
- All alarms and notifications should be kept for future inquiry in VIM, ageing policy of these records should be configurable.
- VIM should distinguish between the failure of the compute node and the failure of the host HW.
- VIM should be able to publish the health status of the compute node to NFV MANO.

5.3 HA mechanism of VIM provided for VNFs

When VNFs deploy their HA scheme, they usually require from underlying resource to provide some mechanism.

This is similar to the hardware watchdog in the traditional network devices. Also virtualization introduces some other requirements like affinity and anti-affinity with respect to the allocation of the different virtual resources.

Requirement:

- VIM should provide the ability to configure HA functions like watchdog timers, redundant network ports and etc. These HA functions should be properly tagged and exposed to VNF and VNFM with standard APIs.
- VIM should provide anti-affinity scheme for VNF to deploy redundant service on different level of aggregation of resource.
- VIM should be able to deploy classified virtual resources to VNFs following the SAL description in VNFD.
- VIM should provide data collection to calculate the HA related metrics for VNFs.
- VIM should support the VNF/VNFM to initiate the operation of resources of the NFVI, such as repair/reboot.
- VIM should correlate the failures detected on collocated virtual resources to identify latent faults in HW and virtualization facilities
- VIM should be able to disallow the live migration of VMs and when it is allowed it should be possible to specify the tolerated interruption time.
- VIM should be able to restrict the simultaneous migration of VMs hosting a given VNF.
- VIM should provide the APIs to trigger scale in/out to VNFM/VNF.
- When scheduler of the VIM use the Active/active HA scheme, multiple scheduler instances must not create a race condition
- VIM should be able to trigger the evacuation of the VMs before bringing the host down when **maintenance mode** is set for the compute host.
- VIM should configure Consoleauth in active/active HA mode, and should store the token in database.
- VIM should replace a failed VM with a new VM and this new VM should start in the same initial state as the failed VM.
- VIM should support policies to prioritize a certain VNF.

5.4 SDN controller

SDN controller: Distributed or Centralized.

Requirements:

- In centralized model SDN controller must be deployed as redundant pairs.
- In distributed model, mastership election must determine which node is in overall control.
- For distributed model, VNF should not be aware of HA of controller. That is it is a - logically centralized system for NBI(Northbound Interface).
- Event notification is required as section 5.2 mentioned.

6 VNF High Availability

6.1 Service Availability

In the context of NFV, Service Availability refers to the End-to-End (E2E) Service Availability which includes all the elements in the end-to-end service (VNFs and infrastructure components) with the exception of the customer terminal such as handsets, computers, modems, etc. The service availability requirements for NFV should be the same as those for legacy systems (for the same service).

Service Availability =total service available time / (total service available time + total service recovery time)

The service recovery time among others depends on the number of redundant resources provisioned and/or instantiated that can be used for restoring the service.

In the E2E relation a Network Service is available only of all the necessary Network Functions are available and interconnected appropriately to collaborate according to the NF chain.

General Service Availability Requirements:

- We need to be able to define the E2E (V)NF chain based on which the E2E availability requirements can be decomposed into requirements applicable to individual VNFs and their interconnections.
- The interconnection of the VNFs should be logical and be maintained by the NFVI with guaranteed characteristics, e.g. in case of failure the connection should be restored within the acceptable tolerance time.
- These characteristics should be maintained in VM migration, failovers and switchover, scale in/out, etc. scenarios.

- It should be possible to prioritize the different network services and their VNFs. These priorities should be used when pre-emption policies are applied due to resource shortage for example.
- VIM should support policies to prioritize a certain VNF.
- VIM should be able to provide classified virtual resources to VNFs in different SAL.

6.1.1 Service Availability Classification Levels

The [ETSI-NFV-REL_] defined three Service Availability Levels (SAL) are classified in Table 1. They are based on the relevant ITU-T recommendations and reflect the service types and the customer agreements a network operator should consider.

[ETSI-NFV-REL] ETSI GS NFV-REL 001 V1.1.1 (2015-01): http://www.etsi.org/deliver/etsi_gs/NFV-REL/001_099/001/01.01.01_60/gs_NFV-REL001v010101p.pdf

Table 1: Service Availability classification levels

SAL Type	Customer Type	Service/Function	Notes
Level 1	Network Operator Control Traffic Government/ Regulatory Emergency Services	• Intra-carrier engineering traffic • Emergency telecommunication service (emergency response, emergency dispatch) • Critical Network Infrastructure Functions (e.g. VoLTE functions DNS Servers, etc.)	Sub-levels within Level 1 may be created by the Network Operator depending on Customer demands E.g.: • 1A - Control; • 1B - Real-time; • 1C - Data; May require 1+1 Redundancy with Instantaneous Switchover
Level 2	Enterprise and/ or large scale customers (e.g. Corporations, University) Network Operators (Tier1/2/3) service traffic	• VPN • Real-time traffic (Voice and video) • Network Infrastructure Functions supporting Level 2 services (e.g. VPN servers, Corporate Web/ Mail servers)	Sub-levels within Level 2 may be created by the Network Operator depending on Customer demands. E.g.: • 2A - VPN; • 2B - Real-time; • 2C - Data; May require 1:1 Redundancy with Fast (maybe Instantaneous) Switchover
Level 3	General Consumer Public and ISP Traffic	• Data traffic (including voice and	While this is typically considered to be "Best

		video traffic provided by OTT) - Network Infrastructure Functions supporting Level 3 services	Effort" traffic, it is expected that Network Operators will devote sufficient resources to assure "satisfactory" levels of availability. This level of service may be pre-empted by those with higher levels of Service Availability. May require M+1Redundancy with Fast Switchover; where M > 1 and the value of M to be determined by further study.

Requirements:

- It shall be possible to define different service availability levels.
- It shall be possible to classify the virtual resources for the different availability class levels.
- The VIM shall provide a mechanism by which VNF-specific requirements can be mapped to NFVI-specific capabilities.

More specifically, the requirements and capabilities may or may not be made up of the same KPI-like strings, but the cloud administrator must be able to configure which HA-specific VNF requirements are satisfied by which HA-specific NFVI capabilities.

6.1.2 Metrics for Service Availability

The [ETSI-NFV-REL_] identifies four metrics relevant to service availability:

- Failure recovery time,
- Failure impact fraction,
- Failure frequency, and
- Call drop rate.

6.1.2.1 Failure Recovery Time

The failure recovery time is the time interval from the occurrence of an abnormal event (e.g. failure, manual interruption of service, etc.) until the recovery of the service regardless if it is a scheduled or unscheduled abnormal event. For the unscheduled case, the recovery time includes the failure detection time and the failure restoration time. More specifically restoration also allows for a service recovery by the restart of the failed provider(s) while failover implies that the service is recovered by a redundant provider taking over the service. This provider may be a standby (i.e. synchronizing the service state with the active provider) or a spare (i.e. having no state information). Accordingly failover also means switchover, that is, an orederly takeover of the service from the active provider by the standby/spare.

Requirements:

- It should be irrelevant whether the abnormal event is due to a scheduled or unscheduled operation or it is caused by a fault.
- Failure detection mechanisms should be available in the NFVI and configurable so that the target recovery times can be met.
- Abnormal events should be logged and communicated (i.e. notifications and alarms as appropriate).

The TL-9000 forum has specified a service interruption time of 15 seconds as outage for all traditional telecom system services. [ETSI-NFV-REL_] recommends the setting of different thresholds for the different Service Availability Levels. An example setting is given in the following table 2. Note that for all Service Availability levels Real-time Services require the fastest recovery time. Data services can tolerate longer recovery times. These recovery times are applicable to the user plane. A failure in the control plane does not have to impact the user plane. The main concern should be simultaneous failures in the control and user planes as the user plane cannot typically recover without the control plane. However an HA mechanism in VNF itself can further mitigate the risk. Note also that the impact on the user plane depends on the control plane service experiencing the failure, some of them are more critical than others.

Table 2: Example service recovery times for the service availability levels

SAL	**Service Recovery Time Threshold**	**Notes**
1	5 - 6 seconds	Recommendation: Redundant resources to be made available on-site to ensure fast recovery.
2	10 - 15 seconds	Recommendation: Redundant resources to be available as a mix of on-site and off-site as appropriate. • On-site resources to be utilized for recovery of real-time services. • Off-site resources to be utilized for recovery of data services.
3	20 - 25 seconds	Recommendation: Redundant resources to be mostly available off-site. Real-time services should be recovered before data services.

6.1.2.2 Failure Impact Fraction

The failure impact fraction is the maximum percentage of the capacity or user population affected by a failure compared with the total capacity or the user population supported by a service. It is directly associated with the failure impact zone which is the set of resources/elements of the system to which the fault may propagate.

Requirements:

- It should be possible to define the failure impact zone for all the elements of the system.
- At the detection of a failure of an element, its failure impact zone must be isolated before the associated recovery mechanism is triggered.
- If the isolation of the failure impact zone is unsuccessful the isolation should be attempted at the next higher level as soon as possible to prevent fault propagation.
- It should be possible to define different levels of failure impact zones with associated isolation and alarm generation policies.
- It should be possible to limit the collocation of VMs to reduce the failure impact zone as well as to provide sufficient resources.

6.1.2.3 Failure Frequency

Failure frequency is the number of failures in a certain period of time.

Requirements:

- There should be a probation period for each failure impact zones within which failures are correlated.
- The threshold and the probation period for the failure impact zones should be configurable.
- It should be possible to define failure escalation policies for the different failure impact zones.

6.1.2.4 Call Drop Rate

Call drop rate reflects service continuity as well as system reliability and stability. The metric is inside the VNF and therefore is not specified further for the NFV environment.

Requirements:

- It shall be possible to specify for each service availability class the associated availability metrics and their thresholds.
- It shall be possible to collect data for the defined metrics.
- It shall be possible to delegate the enforcement of some thresholds to the NFVI.
- Accordingly it shall be possible to request virtual resources with guaranteed characteristics, such as guaranteed latency between VMs (i.e. VNFCs), between a VM and storage, between VNFs.

6.2 Service Continuity

The determining factor with respect to service continuity is the statefulness of the VNF. If the VNF is stateless, there is no state information which needs to be preserved to prevent the perception of service discontinuity in case of failure or other disruptive events. If the VNF is stateful, the NF has a service state which needs to be preserved throughout such disruptive events in order to shield the service consumer from these events and provide the perception of service continuity. A VNF may maintain this state internally or externally or a combination with or without the NFVI being aware of the purpose of the stored data.

Requirements:

- The NFVI should maintain the number of VMs provided to the VNF in the face of failures. I.e. the failed VM instances should be replaced by new VM instances.
- It should be possible to specify whether the NFVI or the VNF/VNFM handles the service recovery and continuity.
- If the VNF/VNFM handles the service recovery it should be able to receive error reports and/or detect failures in a timely manner.
- The VNF (i.e. between VNFCs) may have its own fault detection mechanism, which might be triggered prior to receiving the error report from the underlying NFVI therefore the NFVI/VIM should not attempt to preserve the state of a failing VM if not configured to do so.
- The VNF/VNFM should be able to initiate the repair/reboot of resources of the VNFI (e.g. to recover from a fault persisting at the VNF level => failure impact zone escalation).
- It should be possible to disallow the live migration of VMs and when it is allowed it should be possible to specify the tolerated interruption time.
- It should be possible to restrict the simultaneous migration of VMs hosting a given VNF.
- It should be possible to define under which circumstances the NFV-MANO in

collaboration with the NFVI should provide error handling (e.g. VNF handles local recoveries while NFV-MANO handles geo-redundancy).

- The NFVI/VIM should provide virtual resource such as storage according to the needs of the VNF with the required guarantees (see virtual resource classification).
- The VNF shall be able to define the information to be stored on its associated virtual storage.
- It should be possible to define HA requirements for the storage, its availability, accessibility, resilience options, i.e. the NFVI shall handle the failover for the storage.
- The NFVI shall handle the network/connectivity failures transparent to the VNFs.
- The VNFs with different requirements should be able to coexist in the NFV Framework.
- The scale in/out is triggered by the VNF (VNFM) towards the VIM (to be executed in the NFVI).
- It should be possible to define the metrics to monitor and the related thresholds that trigger the scale in/out operation.
- Scale in operation should not jeopardize availability (managed by the VNF/VNFM), i.e. resources can only be removed one at a time with a period in between sufficient for the VNF to restore any required redundancy.