

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2014

L. Xia
Q. Wu
Huawei
D. King
Lancaster University
H. Yokota
KDDI Lab
October 21, 2013

Use cases and Requirements for Virtual Service Node Pool Management
draft-xia-vsnpool-management-use-case-01

Abstract

Network edge appliances such as subscriber termination, firewalls, tunnel switching, intrusion detection, and routing are currently provided using dedicated network function hardware. As network function is migrated from dedicated hardware platforms into a virtualized environment, a set of use cases with application specific requirements begin to emerge. These use cases and requirements cover a broad range of capability and objectives, which will require detailed investigation and documentation in order to identify relevant architecture, protocol and procedure solutions.

This document provides an analysis of the key management requirements for applications that may be hosted within a virtualized environment. These engineering requirements are based on a variety of uses cases and goals , which include: virtual application security, reliability, scalability, performance, operation and automation.

Note that this document is not intended to provide or recommend protocol solutions.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

Internet-Draft

VSNPool

October 2013

This Internet-Draft will expire on April 24, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Internet-Draft

VSNPool

October 2013

Table of Contents

1.	Introduction	4
2.	Terminology	6
3.	Virtual Service Node (VSN) Overview	7
3.1.	Reliability of VSN, VServer, VSNP	8
3.2.	Reliability of Network Connectivity	9
4.	Use Cases	10
4.1.	Virtualized Mobile Core Network and Service Systems	10
4.2.	Resilience for Stateful Service	11
4.3.	Auto Scale of Virtual Network Function Instances	12
4.4.	Reliable Network Connectivity between Network Nodes	14
4.5.	Existing Operating Virtual Network Function Instance Replacement	15
4.6.	Reliable vCDN	16
4.7.	VSN Cluster	16
4.8.	VSN Resilience Classes	18
4.9.	Reliable Traffic Steering	18
5.	IANA Considerations	21
6.	Security Considerations	22
7.	References	23
7.1.	Normative References	23
7.2.	Informative References	23
	Authors' Addresses	24

1. Introduction

Network virtualization technologies are finding increasing support among network and Data Center (DC) operators. This is due to demonstrable capital cost reduction and operational energy savings, simplification of service management, potential for increased network and service resiliency, and service and traffic elasticity.

Within traditional DC networks, varied middleware boxes including FW (Fire Wall), NAT (Network Address Translation), LB (Load Balancers), WoC (Wan Optimization Controller), etc., are being used to provide network applications (services), traffic control and optimization. Each function is an essential part of the entire operator and DC network, and overall service chain (required traffic path for users). Combined these functions and capabilities can be termed as service nodes.

In terms of virtualizing network functions, a significant amount of service nodes and function instances within the service nodes can be migrated into virtualized environments, in essence the middleware capability is implemented in software on commodity hardware using well defined industry standard servers. Thus allowing the creation, scaling, migration, modification, and deletion of single or groups of functions, across few or many service nodes.

These virtual service nodes may be location independent, i.e., they may exist across distributed or centralized DC hardware. This architecture will pose new issues and great challenges to the automated provisioning across the DC network, while maintaining high

availability, fault-tolerant, load balancing, and plethora of other requirements some of which are technology and policy based.

Today, architecture and protocol mechanisms exist for the management and operation of server hardware supporting applications, these hardware resources are known as server node pools, which may be accessed by other servers and clients. These server node pools have a well-established set of requirements related to management, availability, scalability and performance. Within this document we refer to virtualization of server node pools as Virtual Service Node Pool (VSNP).

[VNF-PS] provides an overview of the problem space related to service node reliability. This document provides an analysis of the key applications that may be hosted within a virtualized environment. These engineering requirements are based on a variety of objectives related to virtual application security, reliability, scalability, performance, operation and automation.

This document is not intended to provide or recommend solutions. The intention of this document is to present an agreed set of objectives and use cases for VSNPs, identify requirements and present architecture framing.

2. Terminology

Broadband Network Gateway (BNG): IP Edge Route where bandwidth and QoS policies may be applied, to support multi-service delivery [[TR-101](#)].

Call Session Control Function (CSCF): A function that is used to manage the mobile IP Multimedia Subsystem (IMS) signaling from users to services and network gateways.

Hypervisor: Software running on a server that allows multiple VMs to run on the same physical server. The hypervisor manages and provide network connectivity to Virtual machines [[NV03-FWK](#)].

IP Multimedia Subsystem (IMS): The IP Multimedia Subsystem used

within mobile core networks.

Network Functions Virtualization (NFV): Moving network function from dedicated hardware platforms onto industry standard high volume servers, switches and storage.

Residential Gateway (RGW): A device located in the home network performing gateway function.

Set-top Box (STB): This device contains audio and video decoders and is intended to connects to a variety of home user devices media servers and televisions.

Virtual Machine (VM): Software abstraction of underlying hardware.

Virtualized Server (VServer): A virtualized server runs a hypervisor supporting one or more VMs [[NV03-FWK](#)].

Virtualized Service Node (VSN): A virtualized network function instance implemented in software on Virtualized Server.

Virtual Service Node Pool (VSNP): Virtualized Server resources supporting a variety of network functions..

[3.](#) Virtual Service Node (VSN) Overview

Shifting towards virtualization of hardware function presents a number of challenges and requirements, this document focuses on those related to network function availability and reliability. In large DC environments, a Virtual Service Node (VSN) may need to deal with traffic from millions of hosts. This represents a significant scaling challenge for VSN operation.

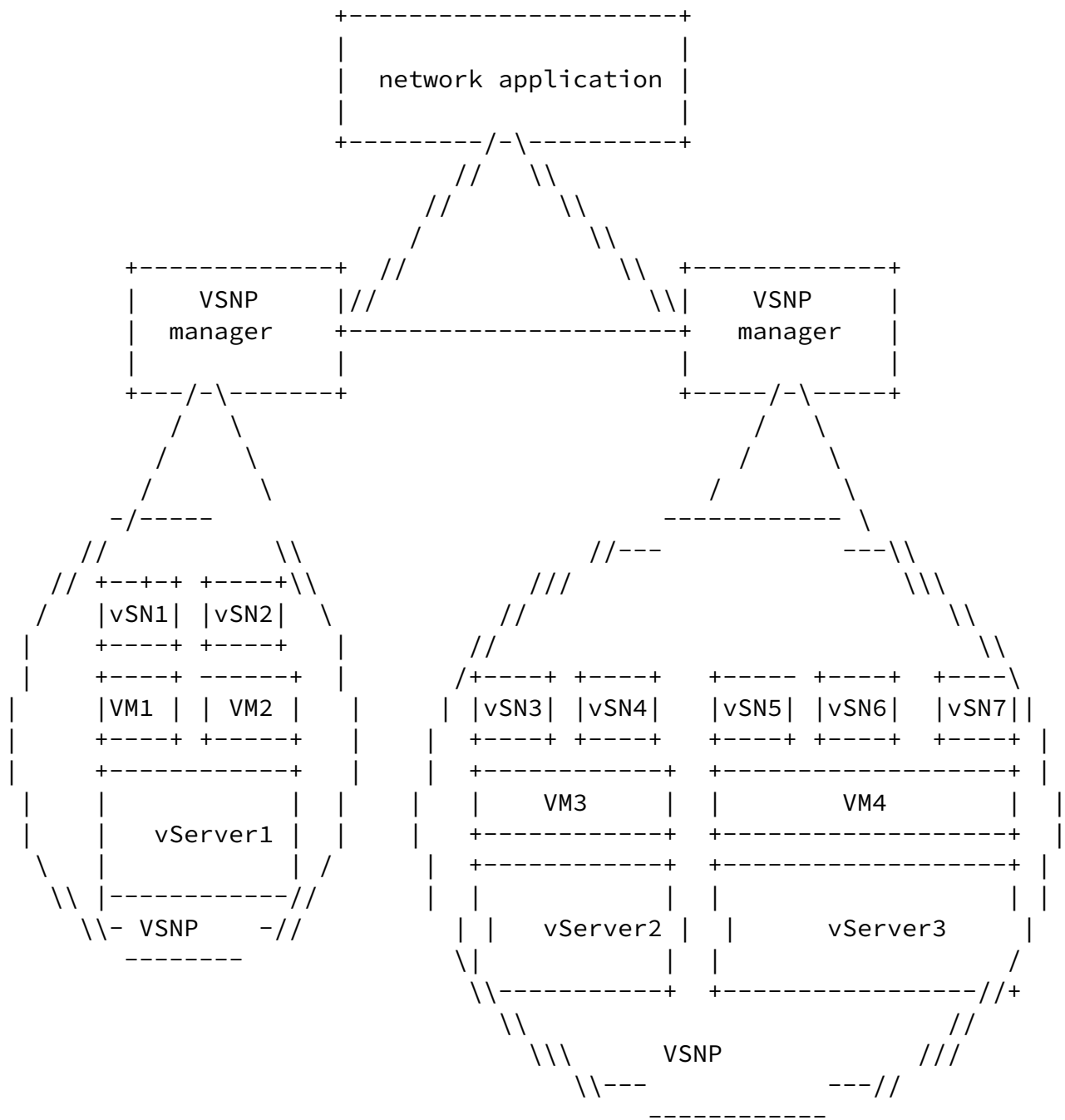


Figure 1: Overall Architecture of VSNP

As shown in Figure 1, the overall architecture of VSNP includes VS, VSN, VSNP, VSNP manager and the connectivity between any two VSs, and the connectivity between VSN and VSNP manager. Rserpool [\[RFC5351\]](#) has the similar architecture to provide high-availability and load balancing, However Rserpool are only used to manage physical servers and can not deal with virtualized function instance when it was designed.

Note that VSNP and VSNP manager also can be used to manage traditional service node.

[3.1.](#) Reliability of VSN, VServer, VSNP

The VSN, VServer and VSNP components are implemented in different network layers and should be considered as different hardware or logical elements.

Multiple VSNs can be provided on one or more VServers for increased reliability. If a VServer detects the failure of the VSNs, it should take the appropriate action for failover and ensures the service continuity.

In order to manage server virtualization across a set of VServers and provide fault tolerant and load sharing across VServers, the VSNPs may be initiated and established as logical element(e.g., a set of VSN providing the same service type), facilitating the migration of a large number of VSNs running on different hypervisors and belonging to different VServers to register into and deregister out. In case of VSN failure or VServer overloading, such VSNPs can be used to support both traditional and virtualized service node replacement or service node adding. However when VSNPs is used to support the operation of traditional service nodes, this doesn't scale very well.

Considering the reliability requirements, VSNP architecture should support several key points detailed below:

- o Application resource monitoring and health checking;
- o Automatic detection of application failure;
- o Failover to another VServer or VSNP;
- o Transparency to other VSNs, VServers or VSNPs;
- o Isolation and reporting of failures;

- o Replication of state for active/standby network functions.

[3.2.](#) Reliability of Network Connectivity

The other category of reliability requirements concerns the network connectivity between any two VSNs, or any two VSNP managers and the network connectivity between VSN and VSN manager.

The connectivity between VSNs is used to deliver service through a set of VSNs to meet the service requirements.

The connectivity between VSNP manager and VSN is used by the VSNP manager to provide registry service to the VSN belonging to different VServer and provide failover of the VSN. A set of VSNP managers can be configured to provide reliable registration. When one VSN cannot obtain a register response from one VSNP manager, it can go to another VSNP manager for registration. This connectivity can also be used by VSNP to monitor the work status of VSNs periodically.

The connectivity between VSNP managers is used to maintain synchronization of data between VSNs located in different VSNP. This allows every VSNP to acquire and maintain overall information of all VSNs and provide protection for each other.

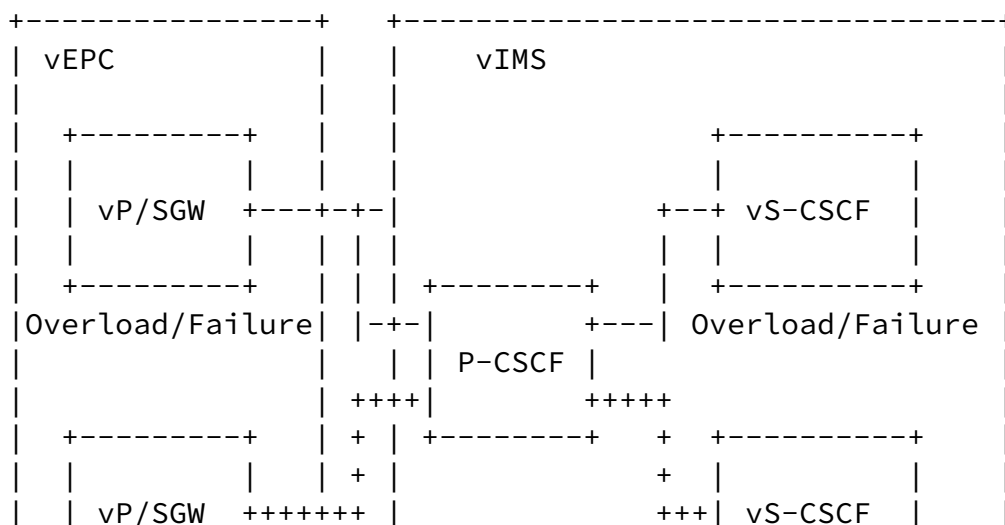
For all types of network connectivity discussed previously, the key reliability requirements stay consistent and include:

- o Automatic detection of link failure;
- o Failover to another usable link;
- o Automated routing recovery.

4. Use Cases

4.1. Virtualized Mobile Core Network and Service Systems

A key use case for NFV is the virtualization of key mobile core network functions. The ETSI NFV use case [NFV-ISG-UC] describes requirements for server and packet gateways (S/P-GW) used for Packet Data Network (PDN) connections and IMS session (see Figure 2: Virtualized mobile core network and IMS). These services are typically time dependent and may require a large number of computing resources in proportion to the number of users and/or service requests. Therefore it is desirable to scale them according to their specific computing requirements. The virtualization can be applied to the Evolved Packet Core (EPC) and the IMS to provide end to end service with service availability and resilience. When those virtualized service nodes(e.g., virtualized S/P-GW and IMS functions) are failed or overloaded, dynamic relocation of those VSNs can be performed, the relocation of the managed sessions and/or connections must be accordingly managed. It also should be noted in [NFV-REL-REQ]that the traffic in the original VSN must be routed to the new location and it is desirable that the movement of the VSN is transparent to other VSN and or physical network entities such as client application on the UE. That is to say the other VSNs don't require to take any special action to this movement.



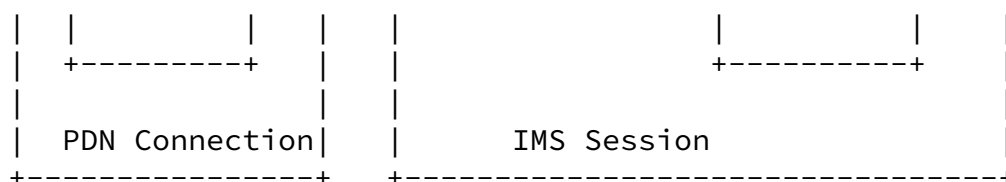


Figure 2: Virtualized Mobile Core Network and IMS

In this use case, the following requirements need to be satisfied:

Xia, et al.

Expires April 24, 2014

[Page 10]

Internet-Draft

VSNPool

October 2013

- o Resource scaling - elastic service aware resource allocation to network functions;
- o State maintenance - network and network function state management during VSN relocation, replication, and resource scaling;
- o Monitoring/fault detection/diagnosis/recovery - appropriate mechanism for monitoring/fault detection/diagnosis/recovery of all components and their states after virtualization, e.g. VNF, hardware, hypervisor;
- o Service Availability - achieving the same level of service availability for the end-to-end virtualized mobile core network as in non-virtualized networks with reduced cost;
- o Minimum impact on other relevant functions.

[More detailed description needs to be discussed.]

[4.2.](#) Resilience for Stateful Service

In the service continuity use case provided by the European Telecommunications Standards Institute (ETSI) Network Function Virtualization (NFV) Industry Specification Group (ISG) [[NFV-REL-REQ](#)], which describes virtual middlebox appliances providing layer-3 to layer-7 services may require maintaining stateful information, e.g., stateful vFW. In case of hardware failure or processing overload of VSN, in addition to the replacement of VSN, it is necessary to move its key status information to new VSN for service continuity. See Figure 3 (Resilience for Stateful Service) for clarification.

In case of multiple vFWs on one VM and not enough resources are

available at the time of failure, two strategies can be taken: one is to move as many vFws as possible to a new place according to the available resources, and the other is to suspend one or more running VSNs in the new place and move all vFws on the failed hardware to it.

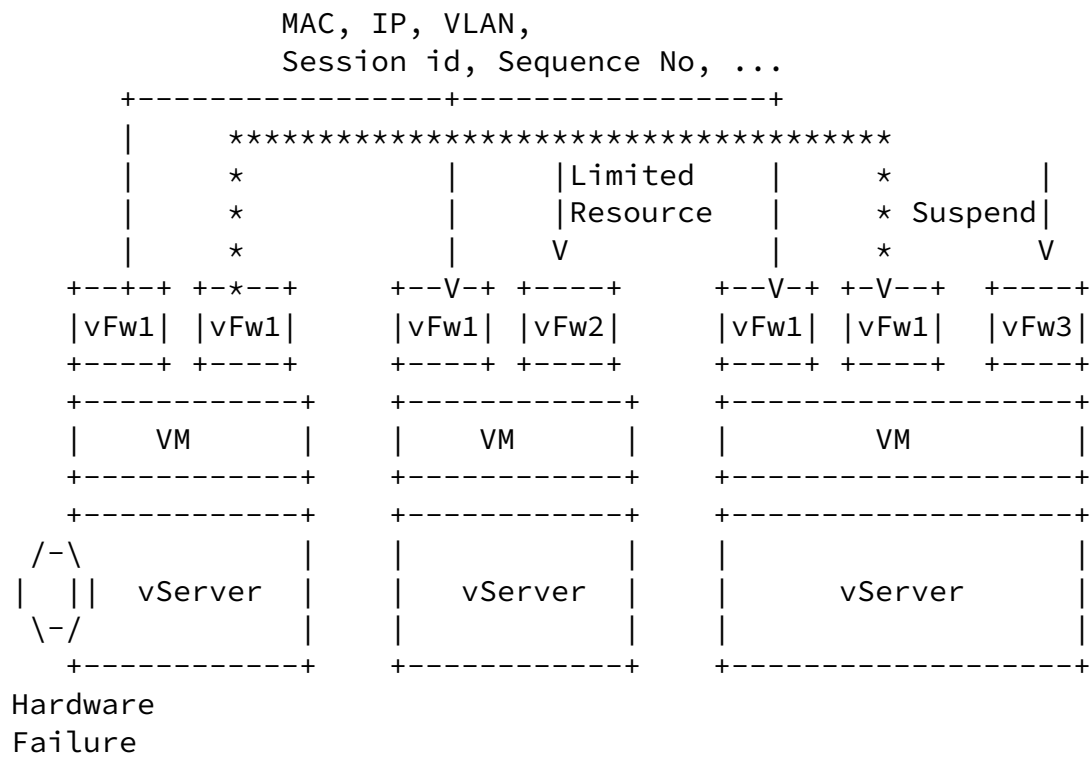


Figure 3: Resilience for Stateful Service

In both scenarios, the following requirements need to be satisfied:

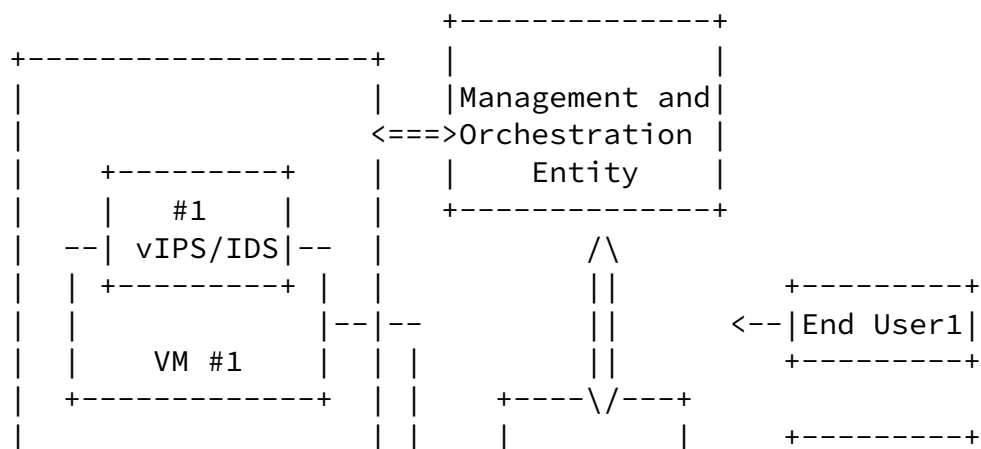
- o Supporting status information maintaining;
- o Supporting status information moving;
- o Supporting VSN moving from one VM to another VM;
- o Supporting partial VSNs moving;
- o Seamless switching user traffic to alternative VMs and VSNs.

4.3. Auto Scale of Virtual Network Function Instances

Adjusting resource to achieve dynamic scaling of VMs described in the ETSI [[NFV-INF-UC](#)] use case and [[NFV-REL-REQ](#)]. As shown in Figure 4, if more service requests come to a VSN than one physical node can accommodate, processing overload occurs. In this case, the movement of the VSN to another physical node with the same resource constraints will create a similar overload situation. A more desirable approach is to replicate the VSN and distribute service node instances ones to one or more new VSNs and at the same time distribute the incoming requests to those nodes.

In a scenario where a particular VSN requires increased resource

allocation to improve overall application performance, the network function might be distributed across multiple VMs. To guarantee performance improvement, the hypervisor dynamically adjusts (scaling up or scaling down) resources to each VSNs in line with the current or predicted performance needs.



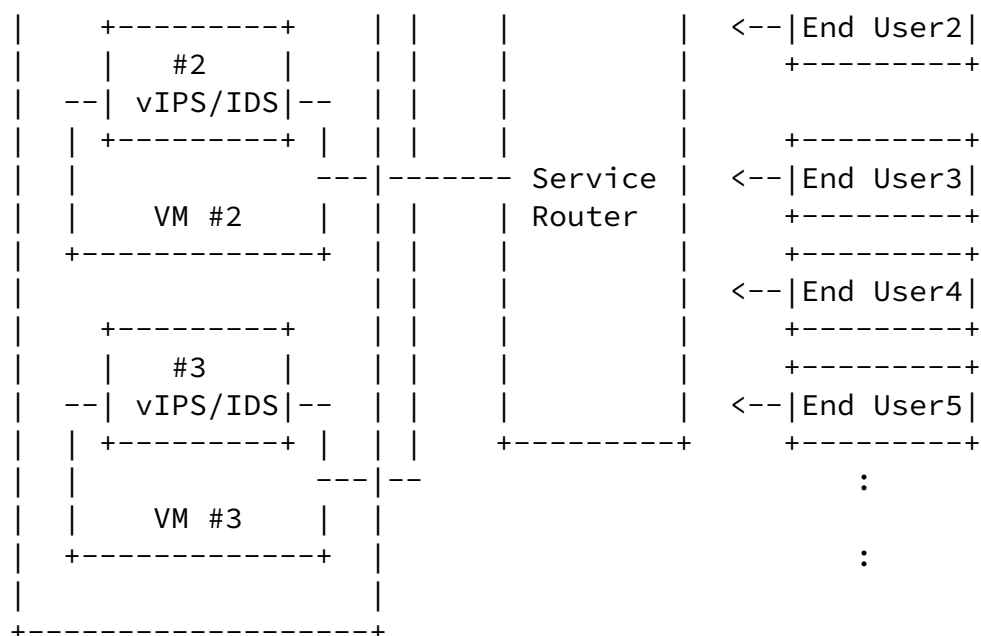


Figure 4: Auto Scaling of Virtual network Function Instances

In this case, the following requirements need to be satisfied:

- o Monitoring/fault detection/diagnosis/recovery - appropriate mechanism for monitoring/fault detection/diagnosis/recovery of all components and their states after virtualization, e.g. VNF, hardware, hypervisor;
- o Resource scaling - elastic service aware resource allocation to network functions.

4.4. Reliable Network Connectivity between Network Nodes

In the reliable network connectivity between network nodes use case provided by ETSI [[NFV-INF-UC](#)], the management and orchestration entities must be informed of changes in network connectivity resources between network nodes. For example, Some network connectivity resources may be temporarily put in power savings mode when resources are not in use. This change is not desirable since it may have great impact on reachability and topology. Another example, some network connectivity resource may be temporarily in a fault state and comes back into an active state, however some other network

connectivity resource becomes permanent in a fault state and is not available for use.

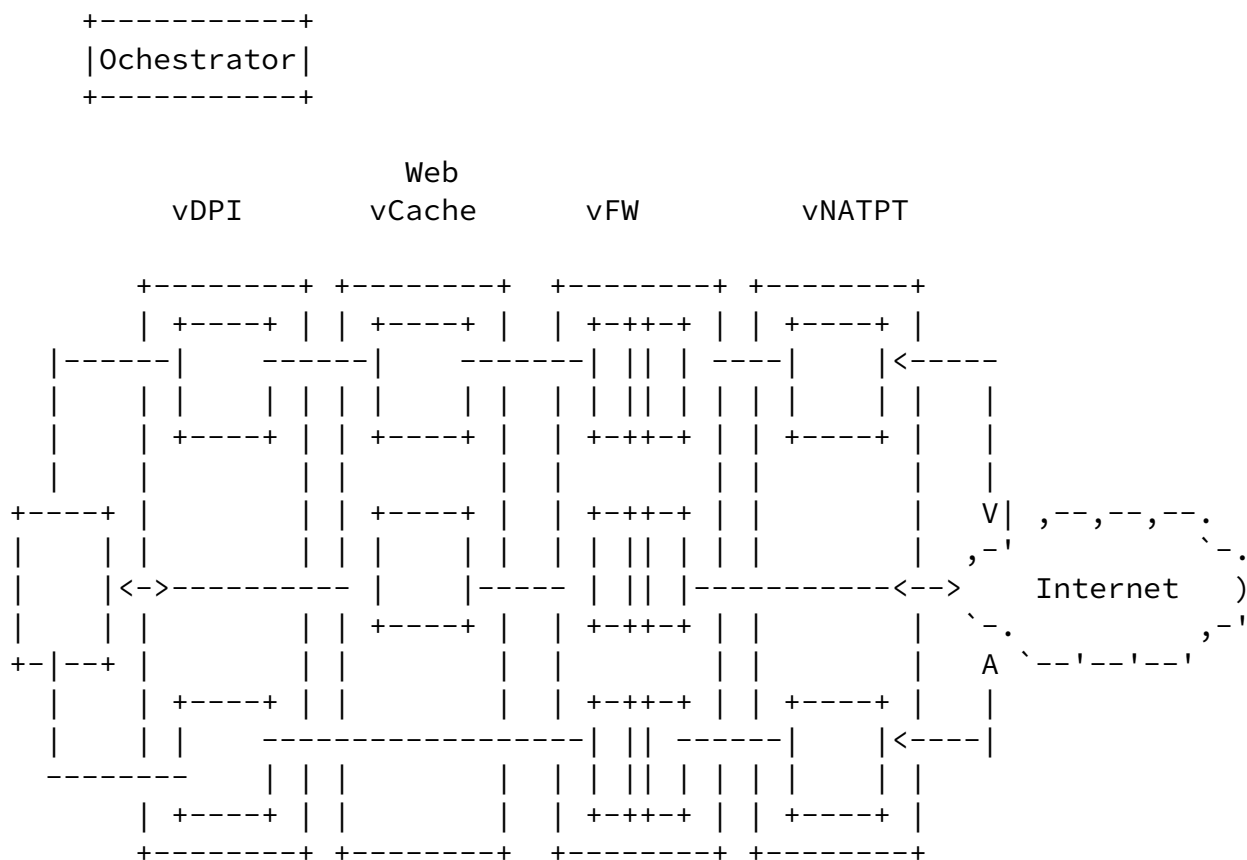


Figure 5: Reliable Network connectivity

In this case, the following requirements need to be satisfied:

- o Quick detection of link failures;
- o Adding network node instances, compute node instances and/or hypervisor instances;

- o Removing network node instances, compute node instances and/or hypervisor instances;
- o Adding or removing network links between nodes.

4.5. Existing Operating Virtual Network Function Instance Replacement

In the Replacement of existing operating VNF instance use case provided by ETSI [NFM-INT-UC] use case, the Management and Orchestration entity may be configured to support virtualized network function replacement. For example, the Network Service Provider has a virtual firewall that is operating. When the operating vFW overloads or fails, the Management and Orchestration entity determines that this vFW instance needs to be replaced by another vFW instance.

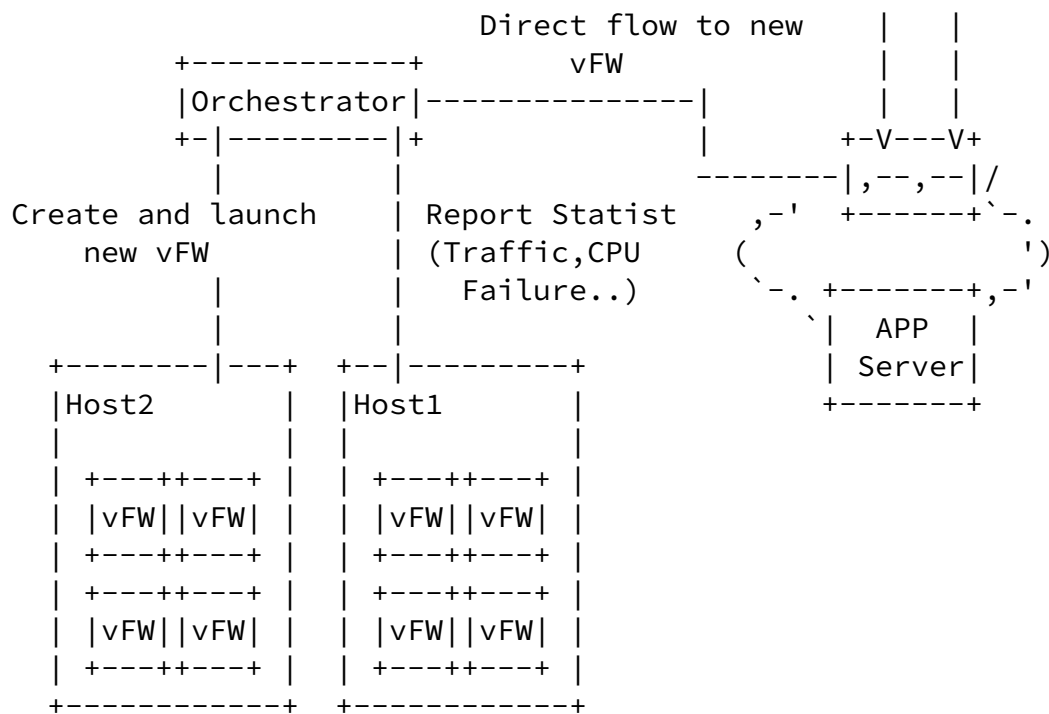


Figure 6: Existing vFW replacement

In this case, the following requirements need to be satisfied:

- o Verifying if capacity is available for a new instance of the VSN at some location;
- o Instantiating the new instance of VSN at the location;
- o Transferring the traffic input and output connections from the old instance to the new instance. This may require transfer of state between the instances, and reconfiguration of redundancy

mechanisms;

- o Pausing or deleting the old VSN instance.

4.6. Reliable vCDN

Virtualization of CDNs in the ETSI [[NFV-ISG-UC](#)] use case described the CDN controller (a centralized component such as Global Load-Balancer(GLB)) selects a Cache Node(CN), or a pool of CNs, to redirect the proper CN which will deliver user request's content. The CDN Controller and CNs may be virtualized so that the resources for the vCDN can be scaled up and down according to the volume of user requests. Then, content placed closer to the user is delivered for providing cost-effective resource utilization and network bandwidth savings.

In this case, the following requirements need to be satisfied:

- o Resource scaling (elastic virtual CN allocation according to the number of requests, proximity, etc);
- o Acceleration of network I/O (I/O centric application needs to overcome the network I/O degradation on the virtualized environment);
- o Performance monitoring (vCDN should be monitored in terms of the number of sessions, load balancing, storage usage, network throughput, etc);
- o Interoperating 3rd party DC infra (3rd party DC infra can be utilized to enhance vCDN coverage globally and to reduce infra cost for delivering the short-term international event);
- o Flexibility to fulfill specific storage density requirements, e.g. to cache a large catalog of popular content;
- o Ability of cache nodes to comply with main monitoring and reporting requirements (e.g., SNMP, syslog, etc. so that operator shall be able to manage different types of cache node for a Delivery Service).

[More description is needed.]

4.7. VSN Cluster

VSN cluster is a set of VSNs which assemble together to support load balancing and high availability. It tends to be a common case in

virtual networks for the following reasons:

- o The performance of VSN is usually not as good as the appliances on dedicated hardware (e.g., physical FW, LB, etc) for VSN is realized mainly depending on software, not on dedicated hardware. VSN cluster should be supported to achieve the same performance as hardware appliance;
- o New requirements of network virtualization as well as multi-tenant support result in a large number of virtual DC network and a large amount of traffic going through them. VSN cluster can be a good choice to deal with this challenge.

There may be multiple different types of VSN clusters in one network. A large number of VSNs dispersed in the network brings difficulty to connect part of them and assemble them as an integrated network function. Also, there should be a flexible load balancing policy between all VSNs in one cluster to achieve good performance. At last, synchronization of status information between lots of VSNs in one or more clusters is more complicated than before and needs more consideration.

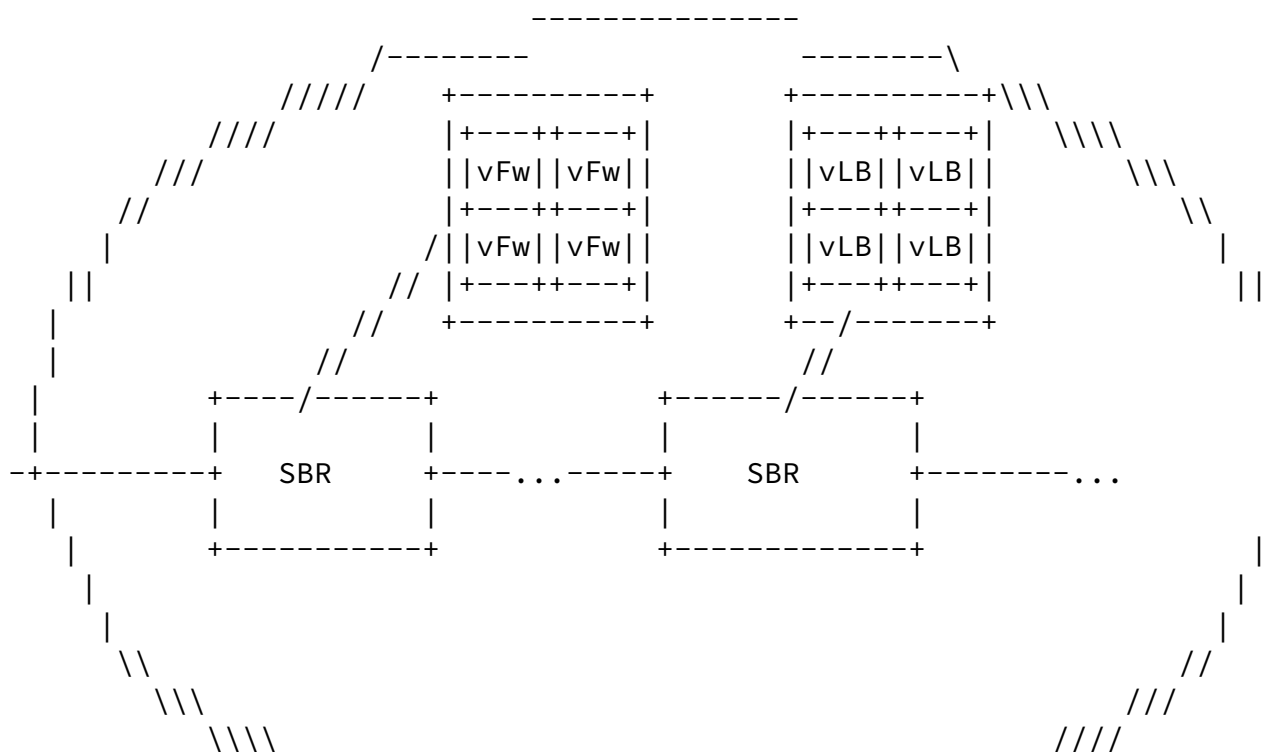




Figure 10: VSNs cluster

As shown in Figure 10, two VSNs clusters are in network, each one consists of 4 VSNs to provide the FW and LB function in a tenant

network. The service border routers connecting to them distribute different flows to each VSN for load balancing.

In this case, the following requirements must be satisfied:

- o Supporting the integration of all connecting VSNs in one cluster to provide one network function for services;
- o Improving performance by providing flexible load balancing policy between VSNs in one cluster;
- o Supporting the synchronization of status information between lots of VSNs in one or more clusters while minimizing the complication and impaction of signaling traffic.

[4.8.](#) VSN Resilience Classes

Different end-to-end services(e.g., Web, Video, financial backend, etc) have different classes of resilience requirement for the VNFs. The use of class-based resiliency to achieve service resiliency SLAs, without "building to peak" is critical for operators.

VSN resilience classes can be specified by some attributes and metrics as followed:

- o Does VSN need status synchronization;
- o Fault Detection and Restoration Time Objective (e.g., real-time, near-real time, non-realtime) and metrics;
- o Service availability metrics;
- o Service Quality metrics;

- o Service reliability;
- o Service Latency metrics for components.

[More description is needed.]

4.9. Reliable Traffic Steering

The characteristics shared by aggregation and mobile-backhaul networks, include a large number of nodes, middlebox appliances and applications providing layer-3 to layer-7 services. Connections are relatively static tunnel, that provide traffic multiplexing for many flows (see Figure 11: Reliable Traffic Steering). These networks are also known for their stringent requirements with regard to

Xia, et al.

Expires April 24, 2014

[Page 18]

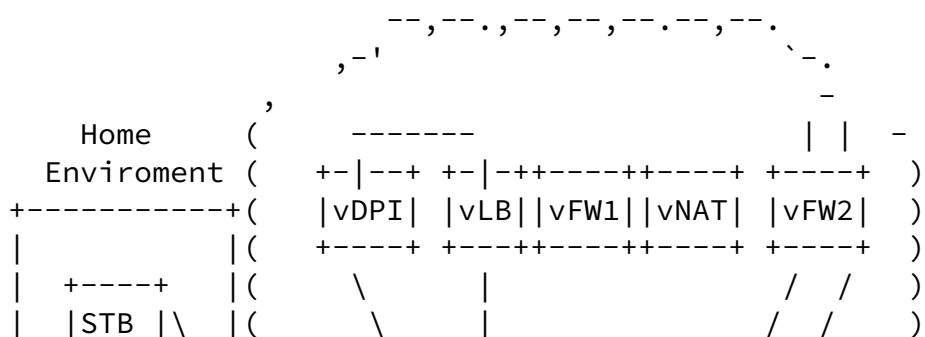
Internet-Draft

VSNPool

October 2013

reliability and short recovery times. The virtualization of the aggregation network will provide optimization of resource allocation and improved traffic forwarding.

Within the aforementioned networks subscriber traffic may be steered through more than one appliances or bypass some appliances completely. For example, traffic may pass through virtualized DPI and FW functions, However, once the type of the flow has been determined by the virtualized DPI function, the operator may decide to modify the services applied to it. For example, if the flow is an internet video stream, it may no longer need to pass the FW service, reducing traffic load on it. Furthermore, in order to reduce traffic load on some appliances or isolate fault on some appliances, after the service type has been detected, the subsequent packets of the same flow may no longer need to pass the LB service either; hence the path of the flow can be updated.



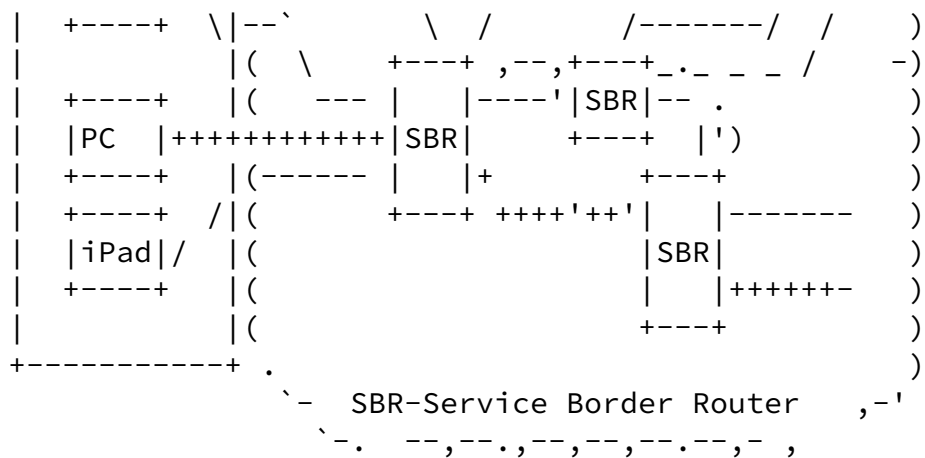


Figure 11: Reliable traffic steering

In this case, the following requirements need to be satisfied:

- o Dynamic steering traffic through a set of virtual service nodes with each providing the same or different service [[BBF-FSC-UC](#)];
- o Dynamic changes to the data path for a given traffic session/flow [[BBF-FSC-UC](#)];

- o Virtualization transparency to services - services using a network function need not know whether it's a virtual function or a non-virtualized;
- o Virtualization transparency to network control and management - network control and management plane need not be aware whether a function is virtualized or not;
- o Traffic control mechanism - data and management traffic identification/separation for non-virtualized and virtualized mobile core networks.

[5.](#) IANA Considerations

This document has no actions for IANA.

[6.](#) Security Considerations

TBD.

[7.](#) References

[7.1.](#) Normative References

- [BBF-FSC-UC]
Broadband Forum, "Flexible Service Chaining", 2013.
- [NFV-INF-UC]
"Network Functions Virtualisation Infrastructure Architecture Part 2: Use Cases", ISG INF Use Case, June 2013.
- [NFV-ISG-UC]
"Network Function Virtualisation; Use Cases;", ISG NFV Use Case, June 2013.
- [RFC5351] Lei, P., Ong, L., Tuexen, M., and T. Dreibholz, "An Overview of Reliable Server Pooling Protocols", May 2008.
- [RFC6707] Niven-Jenkins, B., "Content Distribution Network Interconnection (CDNI) Problem Statement", September 2012.
- [TR-101] Broadband Forum, "Migration to Ethernet-Based DSL Aggregation", 2006.
- [WT-317] Broadband Forum, "Network Enhanced Residential Gateway", 2013.

[7.2.](#) Informative References

- [NFV-REL-REQ]
"Network Function Virtualisation Resiliency Requirements", ISG REL Requirements, June 2013.
- [NV03-FWK]
Lasserre, M., "Framework for DC Network Virtualization", ID [draft-ietf-nvo3-framework-00](#), September 2012.
- [VNF-PS] Zong, N., "Problem Statement for Reliable Virtualized Network Function (VNF) Pool", July 2013.

Internet-Draft

VSNPool

October 2013

Authors' Addresses

Liang Xia
Huawei
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: frank.xialiang@huawei.com

Qin Wu
Huawei
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: bill.wu@huawei.com

Daniel King
Lancaster University
UK

Email: d.king@lancaster.ac.uk

Hidetoshi Yokota
KDDI Lab
Japan

Email: yokota@kddilabs.jp

