

Network Working Group
Internet-Draft
Intended status: Informational
Expires: December 30, 2015

Y. Cheng
China Unicom
D. Zhang
Aliababa Group
JF. Tremblay
Viagenie
S. Zhu
Alibaba Group
J. Bi
Tsinghua University
L. M. Contreras
Telefonica I+D
June 30, 2015

Use Case for Distributed Data Center in SUPA
draft-cheng-supa-ddc-use-cases-08

Abstract

Large scale Distributed Data Centers (DDCs) can provide various services and usually consist of numbers of internal and external links where various VPNs are built upon. The Service provisioning and network connectivity configurations could be complex and dynamic, and manual configuration is onerous and error-prone. This draft analyzes the use cases in DDCs, in which some VPN scenarios are covered, and the applicability of Simplified Use of Policy Abstractions (SUPA) data models which can be used for better and automated resource usage and easy service/network deployment/configuration.

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."

This Internet-Draft will expire on December 6, 2015.

Copyright Notice

Copyright (c) 2015 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.

Table of Contents

1. Introduction	2
2. Requirements Language	3
3. Terminology	3
4. Challenges Faced by Data Center ISPs	4
5. SUPA Benefits	4
6. Scenarios	5
6.1. Scenario:Inter DC Connectivity	5
6.2. Scenario:vDC Connectivity	8
6.3. Scenario:Dynamic Link Configuration for DC	11
6.4. Scenario:DC Connectivity for Virtual Private Clouds (VPC)	12
7. Security Considerations	14
8. IANA Considerations	15
9. Acknowledgements	15
10. Contributors	15
11. References	15
11.1. Normative References	15
11.2. Informative References	15
Authors' Addresses	16

1. Introduction

The SUPA (Simplified Use of Policy Abstractions) work aims at providing data models, including network service data models, policy data models, to easily, accurately, and efficiently select and use the available communication network capabilities. An example of the data model can be found in [I-D.zaalouk-supavpn-service-management-model]. Service Manager (SM) is used by an a communications service provider and/or operator to deploy and manage services on top of network facilities. An example of SM is a set of applications used by an Operational Support System (OSS) entity to perform network configuration. Several SUPA

use cases have been introduced in the problem statement document. This document reviews various scenarios for Distributed Data Center (DDC) use case.

A large-scale Distributed Data Center (DDC) operator may have to maintain multiple data centers and the inter-connecting networks in order to provide server hosting, leased line, and value-added services to enterprises and ISPs, and have multiple data centers. In this DDC network, traffic at each site is routed via configuring policy routes and the operator may be able to adjust routes prioritization to choose an outgoing link. This type of static provisioning comes with high costs and poor operability; and as a result the bandwidth resources intra/inter-data centers are not efficiently utilized.

In quite a few scenarios, the links between DCs are VPNs (MPLS, SSL, IPsec, and etc.). SUPA will be mainly used for those VPN configurations. Although there may be some cases where physical links are used, but those are out of the scope of this draft.

DC and network may belong to different operators. If a DC operator needs to configure network connectivity for DCs, it may need to cooperate with network operators providing such connectivity. Network operators can define and provide data models to enable this.

This document illustrates several distributed datacenter (DDC) applications and explains how an operator could use SUPA to provide these applications.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Terminology

The terminology used in the SUPA problem statement draft [I-D.zhou-supra-framework] and [I-D.karagiannis-supra-problem-statement] apply also to this draft.

DC	Data Center
DDC	Distributed Data Center
ECA	Event Condition Action
NM/NC	Network Manager / Controller

OSS	Operational Support System
SM	Service Manager
SUPA	Simplified Use of Policy Abstractions
TTM	Time to Market
VAS	Value Added Service
vDC	virtual Data Center
VPC	Virtual Private Cloud (PC)

4. Challenges Faced by Data Center ISPs

There are many challenges in traditional DDCs:

1. The infrastructure and network links rent from other operators, are normally dependent on manual planning and design, which leads to inflexibility and low efficiency of resource usages. SUPA can benefit the DDC operators to manage the network resources rent from others in a more flexible way.
2. Service expansion is limited in a single physical DC. Each DC resource is isolated, so service and resource can only be deployed in one single DC.
3. VAS (Value Added Service) is provisioned via static configuration, which brings complex training, long service TTM time and poor flexibility. This could not meet the requirements of complex use cases, e.g., lot of VAS devices, significant differences between various services.

5. SUPA Benefits

There are requirements from DDC operators to optimize and automate the service deployment. Also, a tenant may expect the traffic steering capabilities in order to use the rent links more efficiently. SUPA can show its advantages in above cases since it is able to:

- o Support an open network architecture: standardizaed data models enable an open architecture and make it possible for unified service / network planning, which can interconnect with third party cloud platform, supporting fast service innovation.

- o Support overall DC resource integration: SUPA data models can be used for network resource virtualization; inter-DC resource, virtual DC (vDC) resource, etc, can be integrated and controlled by a centralized functional entity.
- o Support automatic E2E service delivery: Network (including virtual network), computing, inter-DC management of storage resource, automatic service delivery, automatic VPN connection configurations between DCs, which improves operation efficiency.
- o Improve DDC network usage by means of Intelligent scheduling of DDC traffic, improving link usage.
- o Support VAS service on-demand provisioning automatically: Create or delete VAS nodes on-demand, based on various service requirements; network forwarding policy based on the VAS routing, to achieve automatic draining and automatic configuration of VAS device policy.

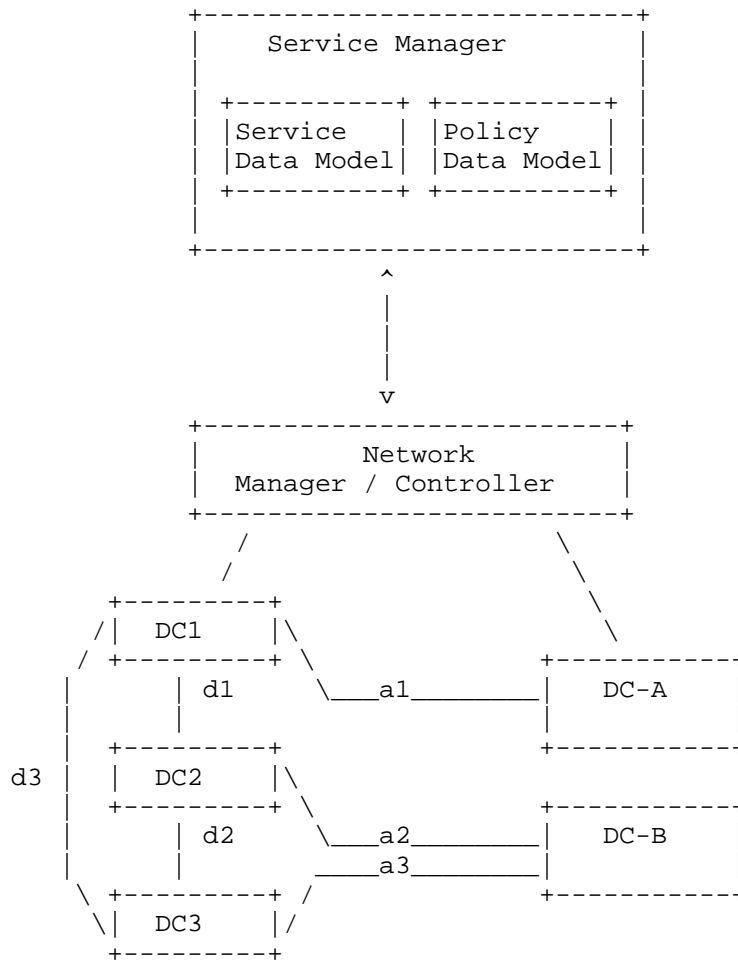
Please refer to [I-D.zhou-supra-framework] and other SUPA related documents for more details of SUPA features.

The following sections will illustrate three typical cases in distributed data center which could benefit from SUPA architecture.

6. Scenarios

In the following uses, Service Manager (SM) is used for service and policy definition; and Network Manager (Controller) is used for network topology maintenance and mapping data models to detail network configurations, as defined in [I-D.zhou-supra-framework].

6.1. Scenario:Inter DC Connectivity



Inter DC Connectivity

There can be a number of links between data centers, and the configuration of such links could be complex. As shown in Figure 1, service data models and policy data models can be defined to automate the configuration procedures. The service data model for connectivity is used for specifying attributes of (virtual) links, e.g. the end points of links, bandwidth, QoS and availability parameters, etc. The policy model can specify some high level requirements to the links, like routing strategy (via and not via) and price/cost strategy. The policy data model can also define the policy rules that drive the security requirements.

An application example is, the links interconnecting two DCs together should guarantee a minimum bandwidth, certain QoS parameters, and provide availability guarantees.

Another service policy example in Figure 1, for traffic from DC2 to DC-B, if the load on a link exceeds a threshold (e.g., 90%), some (new) traffic can be redirected to another link.

Requirements and configurations derived from above application scenarios can be described by service data model and policy data model.

```

module: ietf-supa-ddc
  +--rw ddc-services
    +--rw .....                               other possible attributes
    +--rw ddc-service* [name]
      |   +--rw name                               string
      |   +--rw connection-type?                   enumeration
      |   +--rw connection-name                     string
      |   +--rw bandwidth                           uint32
      |   +--rw latency                             uint32
    +--rw .....                               other possible attributes

```

Service Data Model for Inter DC Connectivity

The above service data model can be used to describe links attributes for a VPN, including bandwidth , latency, etc.

```

module: ietf-supa-policy
  +--rw supa-policy
    +--rw .....                               other possible attributes
    +--rw supa-policy-atomic
      |   +--rw supa-ECA-policy-rule
      |     +--rw policy-rule-name?                 string
      |     +--rw has-policy-events?                 boolean
      |     +--rw has-policy-conditions?             boolean
      |     +--rw has-policy-actions?                 boolean
    +--rw .....                               other possible attributes

```

ECA Policy Data Model

The above policy data model can be used to describe the requirement that when the load on a link exceed a threshold. In this case,

"event" is the bandwidth of a link, "condition" is "load >= 80%",
 "action" is "redirect some traffic to another link".

```

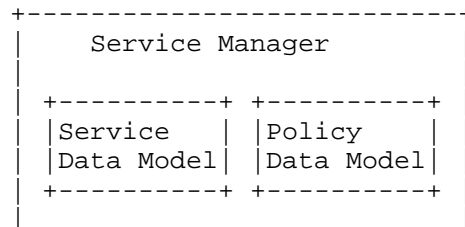
module: ietf-supa-policy
  +--rw policy-set
    +--rw .....                               other possible attributes
    +--rw policy-rule
      +--rw rule-name?                         string
      +--rw rule-type?                         enumeration
      +--rw policy-rule-priority?              uint8
      +--rw intent-policy-rule
        +--rw desired-state
          +--rw constraint?                     string
          +--rw constraint-priority?            uint8
        +--rw behavior-constraint
          +--rw constraint?                     string
          +--rw constraint-priority?            uint8
      +--rw .....                               other possible attributes
  
```

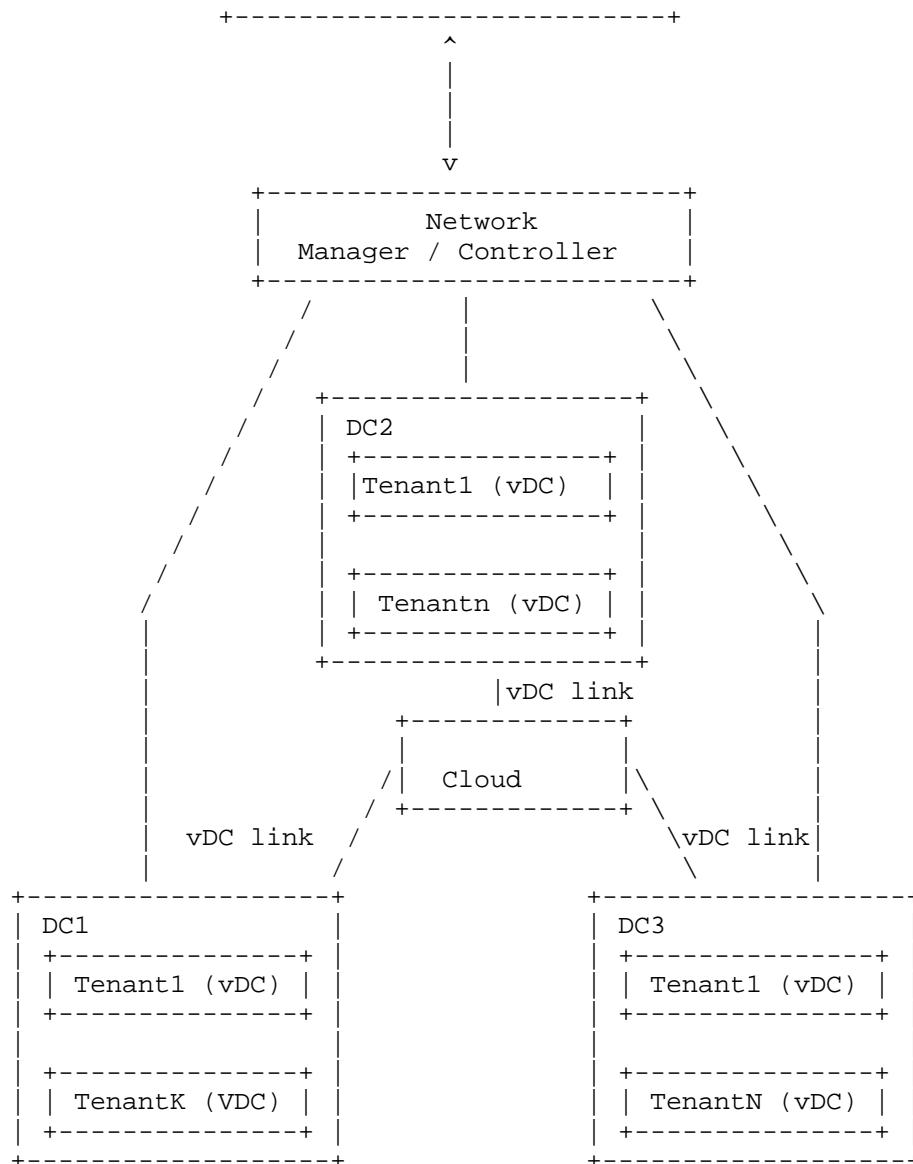
Policy Data Model for high level requirements

The policy model shown above can be used to express some sophisticated requirements, e.g. the number of hop of any link should be less than 5, or any links should not share any network nodes in between and should be completely independent to each other so as to achieve high availability in case of network node failure.

Inter DC connections can be classified into two types: connections within a single administrative domain and connections across multiple administrative domains. Links d1, d2 and d3 are within an administrative domain; and links a1, a2 and 3 are across domains. The difference between them is that connections across multiple administrative domain require extra negotiation and authentication/authorization, which can be achieved via communications between SMS. Data models for this purpose should also be defined.

6.2. Scenario:vDC Connectivity





vDC Connectivity

A DC tenant may have resources, e.g. network, computing, storage, etc, in multiple physical DCs. DC operators will provide internal network connectivity for these distributed resources, and make it

look like one seamless entity, which can be called as virtual DC (vDC).

The internal links for vDC can be implemented via tunneling overlay technologies, e.g. VPN or VxLAN, etc. The tunnels need to be dynamically established, managed and released.

As show in Figure 5, service data model and policy data model can be defined to automate the links configuration for vDCs. A service data model should specify the attributes of the tunnels, e.g., bandwidth, QoS and availability parameters. Policy systems can dynamically scale the DC resources assigned to a tenant, and the policy rules that drive the prioritization of resource assignments. The networking resources assigned to a tenant should scale proportionally to the compute resources assigned to a tenant. The traffic should be prioritized to resources owned by tenants that offer interactive services according to the time zone the DC is located in. Because a DC serving enterprise may require higher priority in working hour, and a DC providing entertainment service may need higher network priority in non-working hours.

```

module: ietf-supa-policy
  +--rw supa-policy
    +--rw .....                               other possible attributes
    +--rw supa-policy-validity-period
      |   +--rw start?                         yang:date-and-time
      |   +--rw end?                           yang:date-and-time
      |   +--rw duration?                       uint32
      |   +--rw periodicity?                     enumeration
    +--rw supa-policy-atomic
      |   +--rw supa-ECA-policy-rule
      |     |   +--rw policy-rule-name?         string
      |     |   +--rw has-policy-events?        boolean
      |     |   +--rw has-policy-conditions?    boolean
      |     |   +--rw has-policy-actions?       boolean
    +--rw .....                               other possible attributes

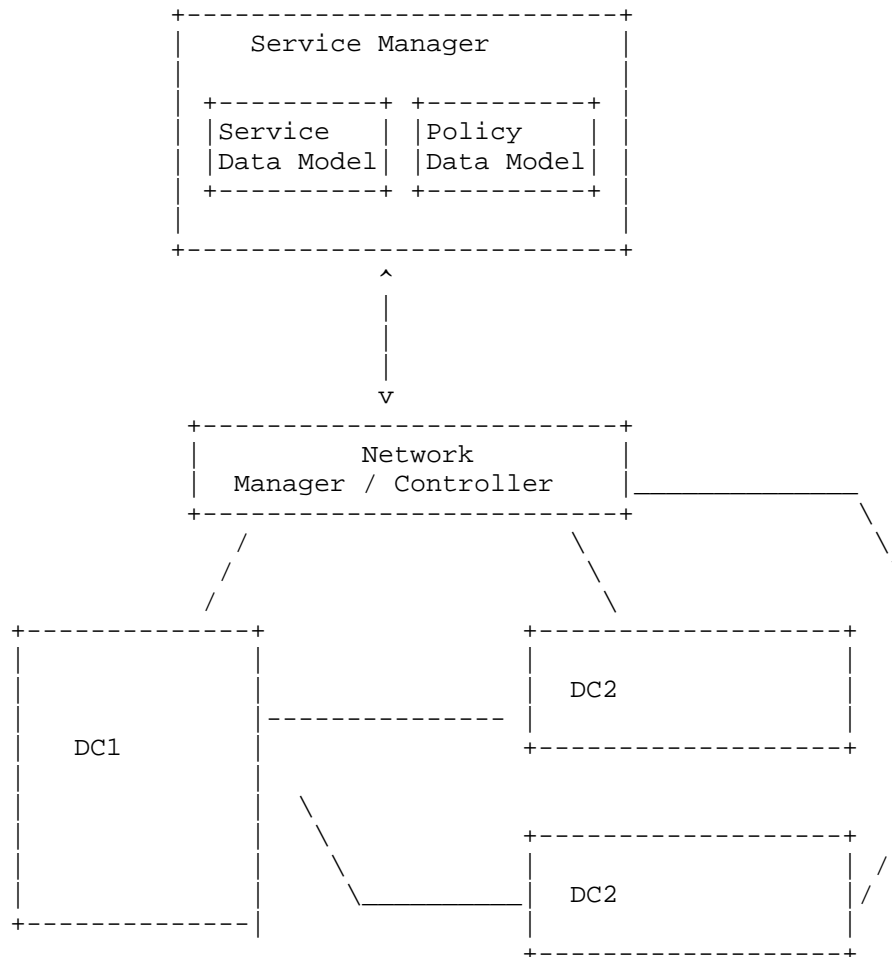
```

Policy Data Model for vDC Connectivity

In the above policy data model, events and conditions may not be necessary; the actions will be based on the time; and two actions will be required: set the VPN priority to low or high.

6.3. Scenario:Dynamic Link Configuration for DC

Static and over provisioning for DC links is not always a good solution. Sometimes dynamic configuration is necessary.



Dynamic Link Configuration for DC

One case is virtual machine migration and large amount of data transfer between DCs. But this kind of activity does not happen frequently. A dedicated link with constant bandwidth for this purpose is too expensive. The network operator should allow the DC operator to create a link on demand when necessary. This link may

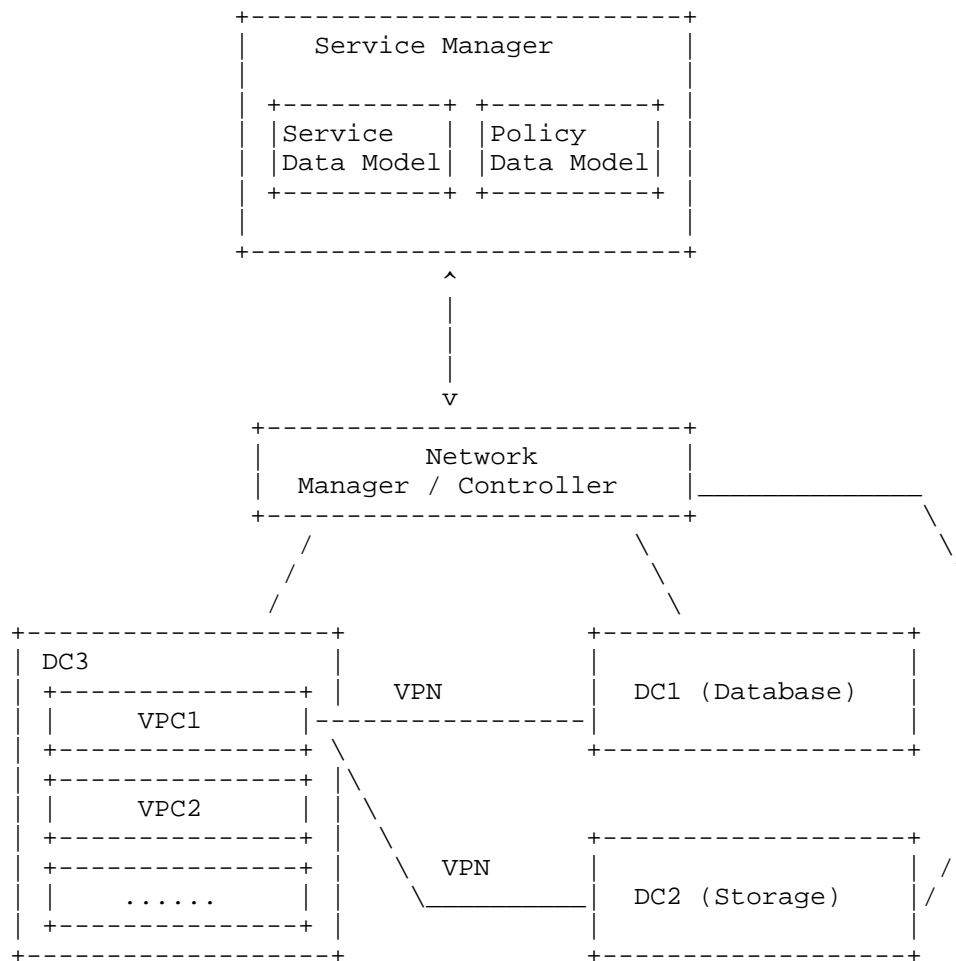
have large bandwidth but last for a limited time period. An alternative is to create short-term dedicated links for backups and migrations.

As shown in Figure 7, data models can help to automate these kind of configurations. In the data models, the attributes of links (bandwidth, QoS and availability parameters) should be specified. The policy concerning strict and soft bounds on the lifetime of such links, and the policy concerning the scheduling of dedicated links (e.g., based on the current load) and the services using the dedicated links can also be specified.

When the traffic volume between DCs exceeds a certain threshold, the policy-driven service manager requests that traffic schedules may be adjusted within bounds in order to balance load on the links (e.g., delay backups and migrations until the network has the necessary capacity).

In this case, the ECA policy model will apply, but the action is different -- change the bandwidth of link(s) with time period constraints, as shown in Figure6.

6.4. Scenario:DC Connectivity for Virtual Private Clouds (VPC)



VPC to DC Connectivity

In practice, a public cloud operator can virtualize its cloud resources into different isolated virtualized private clouds and provided them for different tenants. Such a virtualized private cloud is referred to as a VPC. In a typical VPC provided by, e.g., Alibaba or Amazon, through the control portal, a tenant can establish and manage its network easily, for instance, deploying or removing virtualized network devices (e.g., virtualized routers and virtualized switches), adjusting the topology of VPC networks, specifying packet forwarding policies, and deploying or un-deploying virtual services (e.g., load balancers, firewalls, databases, DNS, etc.). The network functionalities that the tenant can accessed are virtualized and actually performed by the VMs

located on the servers connected through physical or overlay networks. Note that the servers may be located in different data centers which are geographically distributed.

The manipulation of the virtualized VPC network may also affect the configuration of physical networks. For instance, when a tenant newly deploys two VMs in its VPC which are located in different DCs, the VPC control mechanism may have to generate a VPN between two DCs for the internal VPC communication. Therefore, the control mechanism for a VPC should be able to adjust the underlying network when a tenant changes the network or service deployment of the virtual VPC network.

In many cases, a tenant may need to specify how the VPCs is connected to its enterprise cloud networks. For instance, a tenant may want to deploy multiple VPNs to connect the VPC with its private cloud networks and specify the policies to steer the traffics through different VPNs in different conditions. Note that the VPCs that the tenant may be located in different geographic regions, and the VPNs to those VPCs may need to be generated at run time.

In addition, a VPC, often provides other value added services (e.g., database Services, DNS) for VMs in certain VPCs. The VMs and the value added services could be located in different DCs, or even provided by different vendors. VPNs are configured for the VPCs to provide connection to the internal services, and to create and manage VPNs to internal services. The access of VMs to data resources should be controlled. For instance, the VMs in a VPC can access the database services only when the tenant has deployed database into its VPC through the control portal.

As shown in figure 4, service data models and policy data models can be defined to automate the configurations of links between VPC and DC where service is located. The data models should specify the policy controlling authentication and authorization concerning access to data residing in internal services.

7. Security Considerations

Security is a key aspect of any protocol that allows state installation and extracting of detailed configuration states. More investigation remains to fully define the security requirements, such as authorization and authentication levels.

8. IANA Considerations

Not applicable.

9. Acknowledgements

The authors of this draft would like to thank the following persons for review, discussion, and valuable comments: Cathy Zhou, Georgios Karagiannis, Scott O. Bradner, James Huang, Bob Natale.

10. Contributors

The following persons contribute use case and text to this draft, and are listed below:

Scott O. Bradner
sob@sobco.com

Dacheng Zhang
dacheng.zdc@alibaba-inc.com

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

11.2. Informative References

[I-D.karagiannis-supaproblem-statement]
Karagiannis, G., Qiong, Q., Contreras, L., Yegani, P., and J. Bi, "Problem Statement for Simplified Use of Policy Abstractions (SUPA)", draft-karagiannis-supaproblem-statement-06 (work in progress), March 2015.

[I-D.zaalouk-supavpn-service-management-model]
Zhang, D., Zaalouk, A., Pentikousis, K., and Y. Cheng, "VPN Service Management YANG Data Model", draft-zaalouk-supavpn-service-management-model-03 (work in progress), April 2015.

[I-D.zhou-supaframework]
Zhou, C., Contreras, L., Qiong, Q., and P. Yegani, "The Framework of Simplified Use of Policy Abstractions (SUPA)", draft-zhou-supaframework-01 (work in progress), February 2015.

Authors' Addresses

Ying Cheng
China Unicom
P.R. China

Email: chengying10@chinaunicom.cn

Dacheng Zhang
Aliababa Group

Email: dacheng.zdc@alibaba-inc.com

JF Tremblay
Viagenie

Email: jean-francois.tremblay@viagenie.ca

Shunmin zhu
Alibaba Group

Email: jianghe.zsm@taobao.com

Jun Bi
Tsinghua University
Bei Jing
China

Email: junbi@cernet.edu.cn

Luis M. Contreras
Telefonica I+D
Ronda de la Comunicacion, Sur-3 building, 3rd floor
Madrid 28050
Spain

Email: luismiguel.contrerasmurillo@telefonica.com