Abstract

Virtual machine migration is a critical requirement in today’s data centers to support services dynamically deployed in distributed infrastructures. To keep the uniform environment required by application software, large number of networking based solutions have been presented to provide data center interconnection for virtual resource management. Aiming at simplifying networking complexities in off-site service deployment, this document presents a framework based on content delivery network for supporting virtual machine migration over the legacy wide area network.

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 RFC 2119 [RFC2119].

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 April 16, 2013.

Copyright Notice
Copyright (c) 2012 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 .................................................. 3
2. Terminology .................................................. 4
3. CDN based Migration Architecture ............................. 4
   3.1. Functional Architecture .................................. 4
   3.2. Protocols ................................................. 6
   3.3. Mapping Tables in RMG ..................................... 7
4. Use Cases ....................................................... 8
   4.1. Batch Migration ......................................... 8
   4.2. Multi-nodes Backup ....................................... 9
   4.3. Virtual Desktop Migration ............................... 11
5. Implementation Consideration ................................ 11
   5.1. Discussion .............................................. 11
   5.2. Experimental Work ...................................... 12
6. IANA Considerations ........................................... 12
7. Security Considerations ...................................... 12
8. Acknowledgements ............................................. 12
9. Normative References ......................................... 12
Authors’ Addresses .............................................. 12
1. Introduction

In an infrastructure-as-a-service (IaaS) system, IT infrastructure is deployed in a provider’s data center as virtual machine (VM). Managing a VM’s full life-cycle mainly includes setting up networks dynamically for groups of VMs and their storage requirements, such as VM disk image deployment or on-the-fly software environment creation. With IaaS cloud growing popularity, VM migration technology has been widely used in off-site database backup, new services deployment and server system maintenance or upgrade.

Basically, VM migration is categorized into two types: live-migration (real-time) and off-line migration (non real-time). The former migration procedure is typically controlled by VM manager, which involves real-time status synchronization, including networks, memory and storage, between two hosts where VMs locate. Due to strict requirement to network performance, most former methods (application-level or process-level) are only limited within a data center instead of data center interconnection (DCI) scope to prevent poor migration performance or system exception. Not migrating VM in run-time, the later VM migration can be implemented by sending VM image files (static disk image and dynamic snapshot) from source host to destination. That is to say, file-level data transfer between hosts is efficient enough for non real-time VM migration when network quality is favorable. It also means dependence on expensive and sophisticated DCI solutions is largely reduced. As a fast file delivery and cache system, content delivery network (CDN) provides not only high performance data transfer channel but also easy-to-go control plane. Thus, DCI solution should take CDN based scheme into consideration for VM migration.

This document only focus on overall framework, scenarios and implementation consideration for CDN based VM migration. Completed comparisons between VM migration approaches and related networking issues are outside the scope of this document. Motivation behind this document can be concluded as follows:

1. Provide a simple transition scheme before evolving to future’s cost efficient DCI solution in large scale, and give a simple option to those who don’t want to take risk of complex DCI in current phrase.

2. Protect investment on legacy infrastructure. Typically, for some Asia telecom carries, who not only own broadband network and data center infrastructure but also a national wide CDN. Comparing CDN system upgrade, it is a big burden to construct a new DCI specified data plane over underlying broadband network in the near future.
2. Terminology

Virtual Resource Node (VRN): Data centers in which VM migration happens are abstracted as VRN. VRN consists of host, storage and distributed file system in a data center. Before migrating from one to another, virtual resource corresponding to VM is stored in the format of VM image files in VRN, and then is delivered from source VRN to destination VRN across network.

Cloud Virtual Machine Manager (CVMM): Different from typical virtual machine management (VMM) located in data center internally, CVMM is an infrastructure manager that deploys virtualized services on both local pool of resources and external IaaS cloud to support virtual resource life-cycle management.

CDN Management Node (CMN): CMN provides content management for cache node in CDN network, including content adding, content routing, content delivery, resource discovery function. Only one CMN in a CDN system.

CDN Cache Node (CCN): CCN carries VM images for migration received from source VRN, and deliver those images to destination VRN via its high speed data channel provided by CDN interconnected network which is typically constructed over legacy L3 network. Any cache node in legacy CDN system could be a CCN which could receive/deliver image files from/to VRN. Thus, the CCN is also partitioned into source CCN and destination CCN.

Resource Migration Gateway (RMG): RMG is a control system that supports virtual resources in a cloud to migrate across wide area network via CDN system. Its interface with CVMM parses migration instruction and reports related migration status. While the interface with CMN is used to apply for data channel and cache space from CDN for arriving VM images. RMG also finishes some conversion or mapping operation, including content/image format, content/image ID, content location and content routing, between CDN system and VM file system.

3. CDN based Migration Architecture

3.1. Functional Architecture
The functional architecture (as depicted in the Figure 1) describes how RMG, cloud manager and CDN management element collaborate on fulfilling VM migration via CDN system which set up over legacy physical L2/L3 network. Here, one key role of the RMG is to decouple the signaling dependency between cloud system and CDN system. In terms of mechanism outlined below, the VM images are able to be transferred from upper cloud layer (SRC VRN) to CDN and network layer, and then back to the cloud layer (DEST VRN). The typical workflow is summarized as following steps:

1. CVMM receives VM migration instruction in two ways: manually requested by cloud administrator or automatically triggered by user applications. In either case, CVMM records current virtual resource status and clones necessary disk image from the physical machine that the VM is hosted.

2. CVMM identifies this migration operation belong to local or remote. If it is remote migration, CVMM sends a migration request which also contains source and destination information about VRNs.

3. RMG checks its internal VRN-CCN mapping table to get information about the CCN directly connected with SRC or DEST VRN. Then, RMG sends request to CMN with dedicated protocol which includes SRC and DEST nodes information about CCN and size or identification of VM
4. Receiving message from RMG, CMN first requires whether the DEST CCN or network is busy or not. If the DEST is congested, CMN will suggest RMG to modify the destination for VRN. Otherwise, it will check its internal database to judge if CDN system has cached such image file before. If there is an existing image copy, CMN will push the image file to DEST CCN, and then informs RMG to receive the image from the VRN connected.

5. If there is no VM image in CDN system at current phase, CMN thereafter responds the RMG’s requesting for image transport. And RMG send permitting message for VM migration to CVMM as well. Finally, CVMM informs VRNs (SRC and DEST) to start up its file transfer process and keep status for both. CVMM, RMG and CMN maintain necessary dialog during images transfer phrase to refresh status messages.

6. Once the DEST VRN successfully receives image file from DEST CCN, new VM is created in the DEST VRN’s internal physical host.

3.2. Protocols

```
+--------+   C-R     +-------+   R-C     +-------+
|  CVMM  |<--------->|  RMG  |<--------->|  CMN  |
|        | Protocol    |        | Protocol    |
+--------+   V-C     +-------+
|  VRN   |<--------->|  CCN  |
|        | Protocol    |
```

Negotiation Protocol between CVMM, RMG and CMN

```
+--------+   V-C     +-------+
|  VRN   |<--------->|  CCN  |
|        | Protocol    |
```

Transport Protocol between VRN and CCN

Figure 2 Reference Protocols

There are two type protocols in suggest CDN based VM migration scheme (see Figure 2): control protocol between CVMM, RMG and CMN, and data transfer protocol between VRN and CCN.

1. C-R Protocol

TBD

2. R-C Protocol

TBD
3. V-C Protocol

TBD

3.3. Mapping Tables in RMG

<table>
<thead>
<tr>
<th>CCN ID</th>
<th>CCN IP addr.</th>
<th>VRN ID</th>
<th>VRN IP addr.</th>
<th>HOST ID</th>
<th>VM ID</th>
<th>IMAGE ID</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SRC</td>
</tr>
<tr>
<td>SH 1#</td>
<td>10.3.1.2</td>
<td>DC 2#</td>
<td>1.1.3.2</td>
<td>west1</td>
<td>Xen-10</td>
<td>Linux-2</td>
<td>SRC</td>
</tr>
<tr>
<td>BJ 2#</td>
<td>3.3.1.2</td>
<td>DC 1#</td>
<td>1.1.3.2</td>
<td>west1</td>
<td>kvm-07</td>
<td>Win2k-1</td>
<td>DEST</td>
</tr>
</tbody>
</table>

Table 1: CDN-VRN Content Mapping Table

<table>
<thead>
<tr>
<th>SRC</th>
<th>SRC CCN IP addr.</th>
<th>SRC VRN ID</th>
<th>SRC VRN IP addr.</th>
<th>DEST</th>
<th>DEST CCN ID</th>
<th>DEST CCN IP addr.</th>
<th>DEST VRN ID</th>
<th>DEST VRN IP addr.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: CDN-VRN Location Mapping Table

Figure 3 illustrates mapping tables stored in RMG to reflect how to assign VRN’s VM images to CCN cache system

1. CDN-VRN Content Mapping Table

CCN ID: the identification of CCN supporting VMs migration, directly connecting to a VRN with VRN ID

CCN IP addr.: the IP address of CCN with CCN ID

VRN ID: the identification of VRN initiating VMs migration, connecting to a CCN with CCN ID

VRN IP addr.: the IP address of a concrete CCN with VRN ID

Host ID: the identification of host initiating VM migration in a VRN
with VRN ID

VM ID: the identification of VM to be migrated in a host with HOST ID

Image ID: the identification of image file describing a VM with VM ID

Status: source or destination node

Others: TBD

2. CDN-VRN Location Mapping Table

This table shows how to record the source and destination binding relation between VRN-CCN pairs in VM migration duration.

4. Use Cases

4.1. Batch Migration

(*v1, *v2): Two original VM image "v1" and "v2" in a host

(*v1’, *v2’): VM image copied from images "v1" and "v2"

Figure 4 Batch migrating VM Images from City A to City B

Typically, a service deployed in a data center is hosted in different servers or VMs which might belong to different administrative domain. And these virtual or physical resources are orchestrated by cloud manager. Hence all requirements shown in this scenario for service migration are able to be abstracted at the granularity of VMs. Specifically speaking, it is a batch process for multiple VM images transfer via high speed network.

This scenario demonstrates how to implement batch migration for VMs images when dedicated service moves from one data center to another geographically separated in two cities. As depicted in Figure 4, image data "v1"and "v2"are cloned in the Src VRN as "v1‘"and "v2’". Image copies and then are fast send to Dest VRN over CDN system in which the high speed links for cache nodes are pre-constructed and carefully optimized over physical network. Obviously, VM image transfer over CDN system is better than VRN to VRN connected network in quality and bandwidth.

Most steps for negotiation protocol and data transfer are the same as workflows mentioned in Section 3.1.

6. Once the DEST VRN successfully receives image file from DEST CCN, new VM is created in the DEST VRN’s internal physical host.

4.2. Multi-nodes Backup
(*v): Original VM image to be migrated in a host
(*v’): VM image copied from (*v)

Figure 5 VM Image Delivered to Multiple VRNs
(Off-site Backup Service)

This case illustrates an off-site disaster backup service - that is, how to simultaneously deliver VM image to multiple data centers in different cities.

All major steps for control protocol are mentioned in Section 3.1. According to steps in Section 3.1, the management platform will negotiate mutually to discover VRN-CCN pairs matched in terms of mapping table in the RMG before a real migration operation occurs. Once the VRN-CCN pairs are determined, as shown in Figure 5, Src VRN which locates in city A will receive backup instruction from CVMM. It clones the corresponding VM image as "v’" and transfer it to Src CCN connected directly.

Even if the image "v’" should be sent to two destination nodes, Src VRN just send "v’" to Src CCN one time only. Because Src CCN "understands" this migration operation with a fork destination. Therefore, the "v’" image will be duplicated inside Src CCN and delivered to corresponding Dest CCNs along with the routes pre-configured by CDN system. Finally, Dest VRN fetches the same image "v’" respectively from Dest CCNs connected. The more backup nodes in the backup services, the higher transfer efficiency will be gained from CDN based migration scheme mentioned in this document.
4.3. Virtual Desktop Migration

This section illustrates a experience enhancement application for virtual desktop user, which migrates a remote virtual desktop (VM instance) to a local host existing in local CCN.

TBD

5. Implementation Consideration

5.1. Discussion

1. Status Synchronization

The effort to implement this scheme mainly exists in control protocols between CVMM and RMG, RMG-CMN, and necessary modification on CVMM and CMN system. Such modification may include revising status synchronization mechanism in VMM and re-defining content discovery flow in CMN. Since the scheme we presented is an off-line oriented solution, modification work for migration operations in a CVMM are kept in application level rather than revising hypervisor level.

2. Live-migration or offline-migration

One may argue that the scheme mentioned in this document is only suitable for offline-migration cases in which active VMs must be shut down by cloud manager before migration initiates. In view of poor performance in WAN environment and complexity for DCI technologies, offline-migration across WAN is acceptable choice. Considering existence of diverse migration levels, such as progress-level or memory-level, a revised approach for this CDN based scheme is to use live-migration technologies for real status synchronization as supplement. While those static big image files are still transferred in term of our scheme.

3. Routing, Addressing and Content Discovery

This scheme just maintains a simple table to configure static mapping between content and routing manually. It cannot deal with dynamic adding/removing either CDN nodes or VRNs. Moreover, maintaining two resource systems information and related location/routing information, RMG is easy to become the bottleneck in the whole system.

4. Engineering issues
This scheme is only suitable for those hosting service providers or carriers who retain their own CDN to improve utilization of CDN system and to avoid large scale investment on DCI devices at current phrase.

Another issue is that not all data centers are close enough to CDN cache nodes. Thus, connections between data centers and CDN node will become a burden for operators.

5. Others

TBD

5.2. Experimental Work

TBD (Part of initial experiment work have been done to validate mechanism of the scheme, which is based on open source software for CDN and cloud management.)

6. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

7. Security Considerations

TBD

8. Acknowledgements

9. Normative References

Authors’ Addresses

Liang Ou
China Telecom
109, Zhongshan Ave. West
Guangzhou,
China
Email: oul@gsta.com

Huamin Jin
China Telecom
109, Zhongshan Ave. West
Guangzhou,
China
Email: jhm@gsta.com

Mach(Guoyi) Chen
Huawei Technologies Co., Ltd
No. 3 Xinxi Road, Shang-di, Hai-dian District
Beijing  100085
China
Email: mach@huawei.com