



# Deploying the BIG-IP System to Enable Long Distance Live Migration with VMware vSphere vMotion

---

# Table of Contents

|  |           |
|--|-----------|
| Introducing the BIG-IP and VMware vMotion deployment guide   |           |
| Prerequisites and configuration notes .....  | 1         |
| Product versions and revision history .....  | 2         |
| Configuration details .....  | 2         |
| <b>Configuring the local and remote BIG-IP systems .....</b>   | <b>5</b>  |
| Performing the initial configuration tasks .....   | 5         |
| Creating a self IP .....   | 6         |
| Configuring the WAN optimization module .....  | 6         |
| Creating the iSession profile .....  | 9         |
| Creating the WAN Optimization policy .....   | 10        |
| <b>Configuring the BIG-IP GTM .....</b>  | <b>12</b> |
| Creating the data centers .....  | 12        |
| Creating the monitor .....   | 12        |
| Creating the GTM servers .....   | 12        |
| Creating the GTM pool .....  | 13        |
| Creating a wide IP on the GTM .....  | 14        |
| <b>Managing VM hosts, VM storage and client traffic between data centers during vMotion events .....</b> | <b>15</b> |
| Components to be managed by automation .....   | 15        |
| Minimizing client downtime .....   | 17        |
| Modifying your application virtual server to reference the iRule .....                                   | 21        |
| <b>Configuring the VMware infrastructure .....</b>   | <b>22</b> |
| Modifying the VMware ESX configuration .....   | 22        |
| <b>Appendix A: Test results .....</b>  | <b>25</b> |
| Testing methodology .....  | 25        |
| <b>Appendix B: Frequently asked questions and deployment considerations ....</b>                         | <b>27</b> |

---

# Introducing the BIG-IP and VMware vMotion deployment guide

Welcome to the BIG-IP system deployment guide for VMware vSphere™ vMotion™. This guide provides step by step instructions for configuring the BIG-IP system with the WAN Optimization Module (WOM) for VMware long distance live migration with vMotion.

With this implementation, long distance live migration with vMotion becomes possible between two geographically disparate data centers while the virtual machines being moved are active. Through the use of BIG-IP Global Traffic Manager (GTM) and BIG-IP iRules, user web application traffic can also follow to the new data center.

VMware vMotion technology (deployed in production by 70% of VMware customers according to a VMware customer survey from October 2008), leverages the complete virtualization of servers, storage and networking to move an entire running virtual machine with no downtime from one ESX server to another.

For more information on VMware vSphere vMotion, see <http://www.vmware.com/products/vmotion/>

For more information on the F5 devices included in this guide, see <http://www.f5.com/products/>.

You can also visit the VMware page of F5's online developer community, DevCentral, for VMware forums, solutions, blogs and more: <http://devcentral.f5.com/Default.aspx?tabid=53&view=topics&forumid=46>.

To see test results of this deployment guide configuration, see *Appendix A: Test results*, on page 25.

For information on the *F5 Management Plug-In™ for VMware vSphere*, see [www.f5.com/pdf/deployment-guides/f5-management-plug-in-vsphere-dg.pdf](http://www.f5.com/pdf/deployment-guides/f5-management-plug-in-vsphere-dg.pdf)

To provide feedback on this deployment guide or other F5 solution documents, contact us at [solutionsfeedback@f5.com](mailto:solutionsfeedback@f5.com).

## Prerequisites and configuration notes

The following are general prerequisites for this deployment.

- ◆ You must have the WAN Optimization module licensed and provisioned on your BIG-IP systems.
- ◆ This guide includes configuration for the BIG-IP GTM. If you want to take advantage of the benefits of the BIG-IP GTM, you must have the GTM module licensed and provisioned on the BIG-IP system.
- ◆ Virtual IP Addresses in DNS controlled by Global Traffic Manager should have their time-to-live (TTL) records set to a minimum number of seconds in order to facilitate data center failover. Our recommendation is 15 seconds, but the times will vary depending on individual DNS architectures.

- ◆ There must be two BIG-IP systems running the WAN Optimization module, one as the local endpoint (primary) and one as the remote endpoint (secondary).
- ◆ Must have ESX vMotion and Storage vMotion licenses.
- ◆ VMware vMotion uses TCP port 8000. This port must be allowed to be initiated between the two data centers and vMotion between the two ESX servers.
- ◆ BIG-IP iSessions use TCP port 443. This port must be allowed to be initiated between the two data centers.
- ◆ VMware vMotion preserves all network settings of a virtual machine. While moving a machine between two data centers, the IP and network configuration for the migrated hosts between the two data centers must be identical. However, the infrastructure does not need to be part of the same layer 2 broadcast domain.
- ◆ See *Appendix B: Frequently asked questions and deployment considerations*, on page 27 for more information.

## Product versions and revision history

Product and versions tested for this deployment guide:

| Product Tested                                 | Version Tested |
|--|----------------|
| BIG-IP system with the WAN Optimization module | v10.1          |
| VMware vSphere vMotion                         | v4             |

| Document Version | Description  |
|------------------|--|
| 1.0              | New deployment guide   |
| 1.1              | Changed the term “long-distance vMotion” used in this guide to “long distance live migration” or “long distance live migration with vMotion” |

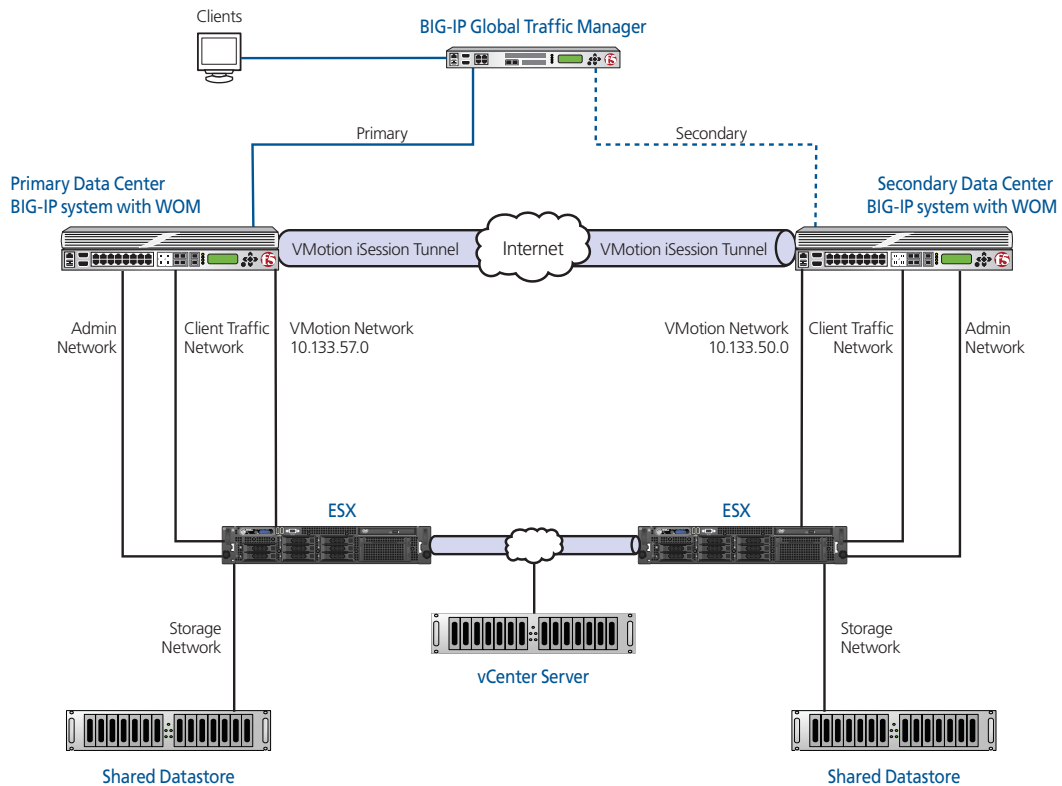
## Configuration details

Figure 1, on page 3 is a logical configuration diagram highlighting the major components of a long distance live migration deployment. In our example, we have two data centers (Primary and Secondary) connected via a WAN (the actual WAN connectivity technology is not relevant in this architecture and may be IPSec, MPLS, Point-to-Point Frame Relay, or another technology).

Each data center contains nearly identical ESX and BIG-IP infrastructures. At a minimum, this configuration requires the ESX server must be vMotion compatible and the both BIG-IPs must support iSessions within the WAN Optimizations Module (WOM). Additionally, the VLANs, port groups and other immediate IP infrastructure with connected to the virtual machine must exist on each participating host. Upstream in both data centers, router and firewall configurations must also be consistent to allow client traffic.

In each data center, the ESX servers are configured as recommended by VMware, with client traffic residing on one dedicated physical network interface, the storage network on another interface, the administrative traffic on its own interface, and finally, the vMotion network on its own interface. ***By configuring VMWare ESX in this recommended manner, the vMotion network can have a separate TCP gateway address and therefore participate in long distance encrypted and accelerated live migration.***

An iSession tunnel is established between the BIG-IP systems in each data center. No further changes are required on other routers and switches to build this tunnel. The iSession tunnel uses TCP port 443, therefore that traffic has to be allowed through the firewall.



**Figure 1** Logical configuration example

## Network Design

For long distance live migration traffic to succeed, on each ESX server we change the default gateway of the VMkernel managing vMotion to point to the BIG-IP system. This change is made through the Networking settings of the vCenter client. Optimization policies on the BIG-IP then recognize vMotion traffic and optimize it through iSession technology.

As in typical single data center deployments, vCenter Server is also used to manage Virtual Machines and vMotion in multiple data center deployments. However, unlike single data center deployments, the vCenter Server must be deployed to support two data centers in a long distance live migration deployment. There are two different modes you can use to accomplish this long distance deployment.

In the first deployment mode, the vCenter Server can control servers in both data centers. In the second deployment example (not pictured in Figure 1) each data center has its own vCenter Server. After long distance live migration, a particular host must be de-registered from the primary data center and re-registered in the secondary data center.

Finally, the last consideration is the strategy for managing Virtual Hosts. We have two recommendations in this regard. The first is to configure every server in your ESX cluster to participate in long distance live migration using vMotion. If Dynamic Resource Scheduler (DRS) is turned on, this may be not be possible.

The second recommendation is to use two dedicated migration hosts, one in each data center. These dedicated migration servers are responsible for long distance live migration and are the only ESX servers to have their configurations adjusted to participate in long distance live migration.

For this method to be successful, the migration hosts are setup to not participate in DRS (in their own cluster). To move a server from the primary data center to the secondary, first the virtual machine is moved to this ESX migration host, and then it is migrated over to the secondary data center. By using this procedure, dynamic vMotion, resource groups, and other provisioning configuration within existing ESX clusters do not need to be modified.

---

## Configuring the local and remote BIG-IP systems

In this section, we configure the BIG-IP system, including the WAN Optimization Module. Much of the configuration needs to be completed on both the local and remote (primary and secondary) BIG-IP systems.

◆ **Note**

---

*In this document, we typically refer to the **primary** and **secondary** data centers or **BIG-IP** systems. The WAN optimization module uses **local** and **remote**.*

## Performing the initial configuration tasks

In this section, we configure the BIG-IP system with required VLAN and Self IP address information. Complete these procedures only if you do not already configured these objects on the BIG-IP LTM.

### Creating a VLAN

The task is to create a VLAN on the BIG-IP LTM system.

◆ **Note**

---

*You may already have a VLAN on the BIG-IP system for the networks that will participate in vMotion. If you do, continue with **Creating a self IP**, on page 6.*

#### To create a VLAN

1. On the Main tab, expand **Network**, and then click **VLANs**.  
The VLANs screen opens.
2. Click the **Create** button.  
The new VLAN screen opens.
3. In the **Name** box, type a unique name for the VLAN. In our example we use **vmotion-vlan**.
4. In the **Tag** box, you can optionally type a tag. In our example, we leave this blank, and the BIG-IP LTM automatically assigns a tag.
5. In the Resources section, from the **Available** list, select the interface that will have access to tagged traffic, and add it to the **Untagged** box by clicking the Add (<<) button.  
In our example, we select **1.14**.
6. Click the **Finished** button.

## Creating a self IP

Self IP addresses are the IP addresses owned by the BIG-IP LTM system that you use to access the VLANs. The next step in this configuration is to create a self IP address that is used in the local endpoint BIG-IP configuration.

### To create a self IP address

1. On the Main tab, expand **Network**, and then click **Self IPs**. The Self IP screen opens.
2. Click the **Create** button. The new Self IP screen opens.
3. In the **IP Address** box, type a static IP address that resides on the VLAN you created in the preceding procedure.
4. In the **Netmask** box, type the corresponding subnet mask.
5. From the **VLAN** list, select the VLAN you created in *Creating a VLAN*. In our example, we select **vmotion-vlan**.
6. From the **Port Lockdown** list, select **Allow None**.
7. Click the **Finished** button. The new self IP address appears in the list.
8. Repeat this entire procedure on the remote endpoint BIG-IP system.

## Configuring the WAN optimization module

In this section, we configure the WAN optimization module (WOM). The WAN optimization module allows you to encrypt and accelerate data between BIG-IP devices, accelerate applications across the WAN, and much more.

One of the options in configuring the WAN optimization module is the choice to use Dynamic Discovery. The benefit of dynamic discovery is that it reduces configuration complexity. However, when dynamic discovery is used, the BIG-IP currently disables iSession routing in order to prevent inadvertent routing loops. In our environment, dynamic discovery is allowed, but care was taken to ensure iSession routing was enabled.

### To configure the WOM module

1. On the Main tab of the local BIG-IP system, expand **WAN Optimization**, and then click **Configuration**. The Local Endpoint Configuration screen opens.
2. In the **IP Address** box, type the BIG-IP self IP address you provisioned for iSession endpoint in the data center.
3. Make sure the **Create iSession Virtual Server** list is set to **Yes**.
4. Click the **Save** button.

5. In the **Advertised Routes Configuration** section, click the **Create** button. The Advertised Route is the local subnet that the local endpoint advertises to all configured remote endpoints to which it is connected. For this vMotion deployment guide, you advertise the local vMotion subnet.
6. In the **Alias** box, type an alias for this route. This is optional. In our example, we type **vlan\_1057**.
7. In the **Subnet Address** box, type the appropriate subnet address. In our example, we type **10.133.57.0**.
8. In the **Netmask** box, type the associated netmask. In our example, we type **255.255.255.0**.
9. Make sure the **Enabled** box is checked.
10. Click the **Finished** button.
11. In the **Dynamic Discovery** section, we leave the default settings.
12. Repeat this entire procedure on the remote endpoint BIG-IP system, using the appropriate BIG-IP self IP address in step 2, and the appropriate Advertised Route information.

The screenshot shows the WAN Optimization Configuration interface. It is divided into three main sections: Local Endpoint Configuration, Advertised Routes Configuration, and Dynamic Discovery.

**Local Endpoint Configuration:** The profile is set to 'Basic'. The IP Address is 10.133.57.236, the SSL Profile (Server) is 'serverssl', and 'Create or Update iSession Virtual Server' is set to 'Yes'. There are 'Update' and 'Delete' buttons.

**Advertised Routes Configuration:** A table lists the configured routes. A 'Create...' button is visible in the top right.

| <input checked="" type="checkbox"/> | Alias     | IP Address  | Netmask       | Enabled |
|-------------------------------------|-----------|-------------|---------------|---------|
| <input type="checkbox"/>            | vlan_1057 | 10.133.57.0 | 255.255.255.0 | Enabled |

A 'Delete...' button is located below the table.

**Dynamic Discovery:** The profile is set to 'Basic'. Under 'General Configuration', both 'Allow Remote Endpoints to Discover This Local Endpoint' and 'Allow This Local Endpoint to Discover Remote Endpoints' are checked. 'Automatically Include Discovered Remote Endpoints' is set to 'Yes'. There is an 'Update' button.

*Figure 2 The WAN optimization configuration*

After performing this procedure on both BIG-IP systems, you connect your two BIG-IP systems together via an iSession tunnel by identifying each remote endpoint. If dynamic discovery was left on (as in step 11), you only perform the following procedure on one of the BIG-IP systems. If you did not, you must repeat this procedure on the remote BIG-IP system.

### To configure the remote endpoints

1. On the Main tab of the local BIG-IP system, expand **WAN Optimization**, and then click **Configuration**.
2. On the Menu bar, click **Remote Endpoints**.
3. Click the **Create** button.
4. From the **Remote Endpoint** list, select **Advanced**.
5. In the **IP Address** box, type the IP address you provisioned for remote iSession endpoint.
6. **Important:** From the **Routing** list, select **Enabled**.
7. Click **Finished**.
8. If you disabled dynamic discovery in the previous procedure, you must repeat this procedure on the remote BIG-IP system.

| WAN Optimization >> Configuration >> New Remote Endpoint...   |               |
|---|---------------|
| Remote Endpoint:  | Advanced      |
| IP Address  | 10.133.58.236 |
| SSL Profile (Server)  | None          |
| Tunnel Port   | 443           |
| Source Address  | none          |
| State   | Enabled       |
| Routing   | Enabled       |
| <input type="button" value="Cancel"/> <input type="button" value="Repeat"/> <input type="button" value="Finished"/> |               |

*Figure 3 Configuring the remote endpoint*

### Ensuring that iSession routing is enabled

As mentioned previously, if Dynamic Discovery is enabled, the BIG-IP system automatically sets remote endpoint routing to disabled. We want to ensure remote endpoint routing is enabled (as in step 6 above).

#### ◆ Important

*We recommend you check that routing is enabled after anytime the BIG-IP system reboots or hotfix/upgrade installations, as routing may revert to **Disabled** to avoid any routing loops.*

### To ensure that iSession routing is enabled

1. On the Main tab of the local BIG-IP system, expand **WAN Optimization**, and then click **Configuration**.
2. On the Menu bar, click **Remote Endpoints**.
3. Click the IP address of the appropriate endpoint.

- 
4. From the **Routing** list, make sure that **Enabled** is selected. If it is not, select **Enabled** from the list.
  5. Click the **Update** button.
  6. Repeat this procedure on the remote BIG-IP system.

## Creating the iSession profile

In this procedure, we create an iSession profile. The iSession profile tells the system how to optimize traffic. Creating a custom iSession profile for each application is a best practice. For vMotion with version 10.1 of the BIG-IP system, we disable iSession Reuse Connection and Deduplication.

### To create the iSession profile

1. On the Main tab, expand **Local Traffic**, and then click **Profiles**.
2. On the Menu bar, from the Services menu, select **iSession**.
3. Click the **Create** button.
4. In the **Name** box, type a name for this profile. In our example, we type **vmotion-isession**.
5. In the Settings section, click the Custom box for **Deduplication**, and then select **Disabled** from the list.
6. Click the Custom box for **Reuse Connection**, and then select **Disabled** from the list.
7. Leave the other settings at the default levels.
8. Click the **Finished** button (see Figure 4).

Local Traffic >> Profiles : Services : iSession >> New iSession Profile...

**General Properties**

|                |                   |
|----------------|-------------------|
| Name           | vmotion-issession |
| Parent Profile | issession         |

**Settings** Custom

|                   |           |                                     |
|-------------------|-----------|-------------------------------------|
| Mode              | Enabled   | <input type="checkbox"/>            |
| Deduplication     | Disabled  | <input checked="" type="checkbox"/> |
| Reuse Connection  | Disabled  | <input checked="" type="checkbox"/> |
| Target Virtual    | match all | <input type="checkbox"/>            |
| Port Transparency | Enabled   | <input type="checkbox"/>            |

**Compression Settings** Custom

|               |         |                          |
|---------------|---------|--------------------------|
| Adaptive      | Enabled | <input type="checkbox"/> |
| Deflate       | Enabled | <input type="checkbox"/> |
| Deflate Level | 1       | <input type="checkbox"/> |
| LZO           | Enabled | <input type="checkbox"/> |
| Null          | Enabled | <input type="checkbox"/> |

Cancel Repeat Finished

*Figure 4* Configuring the iSession profile

## Creating the WAN Optimization policy

The next task is to create the WAN optimization policy. For this configuration, we create a new optimization policy for vMotion.

### To create a new WAN Optimization policy

1. On the Main tab, expand **WAN Optimization**, and then click **Configuration**.
2. On the Menu bar, click **Optimization Policies**.
3. Click the **Create** button. The Common Application Optimization Policies page opens.
4. Click the **Create Custom Policy** button. The New Optimization Policy wizard opens.
5. Type a name for this virtual server. In our example, we type **vsphere-vmotion**.
6. Select **No** for the question asking if this is an iSession endpoint tunnel terminating virtual server.
7. In the **IP Address** box, type the host IP address of the remote vMotion host.

8. In the *What kind of application would you like to optimize?* box, type **8000**.  
***Important:** Make sure your firewall rules allow TCP port 8000 traffic between the vMotion networks of your ESX servers.*
9. From the *Will clients be connecting to this virtual server over a LAN* list, select **Yes**.
10. Encrypting the tunneled data is optional. In our example, we select **Yes**. VMware vMotion is an unencrypted technology, and using encryption on the BIG-IP system provides an additional layer of security (but adds a small amount of time to every transaction, see Appendix A for a list of specific differences).
11. In the VLAN section, from the **Available** list, select the appropriate VLANs and click the Add (<<) button.
12. In the Profile Settings section, from the **iSession Profile** box, select the profile you created in *Creating the iSession profile*, on page 9.
13. Click the **Finished** button.

WAN Optimization » Configuration » New Optimization Policy...

**Common Questions**

| What name would you like to use for this virtual server?        | vsphere-vmotion   |          |           |             |           |  |             |  |             |  |              |
|---|---|----------|-----------|-------------|-----------|--|-------------|--|-------------|--|--------------|
| Is this an iSession endpoint tunnel terminating virtual server? | No  |          |           |             |           |  |             |  |             |  |              |
| Which IP Address/mask should this virtual server match?         | Type: <input checked="" type="radio"/> Host <input type="radio"/> Network<br>Address: 10.133.58.50  |          |           |             |           |  |             |  |             |  |              |
| What kind of application would you like to optimize?            | 8000 Other: <input type="text"/>  |          |           |             |           |  |             |  |             |  |              |
| Will clients be connecting to this virtual server over a LAN?   | Yes   |          |           |             |           |  |             |  |             |  |              |
| Would you like to encrypt tunneled data?                        | Yes   |          |           |             |           |  |             |  |             |  |              |
| Which VLANs should this virtual server listen on?               | <table border="1"> <thead> <tr> <th>Selected</th> <th>Available</th> </tr> </thead> <tbody> <tr> <td>vm_vsphere1</td> <td>vm_client</td> </tr> <tr> <td></td> <td>vm_exchange</td> </tr> <tr> <td></td> <td>vm_foss_web</td> </tr> <tr> <td></td> <td>vm_oracle_dr</td> </tr> </tbody> </table> | Selected | Available | vm_vsphere1 | vm_client |  | vm_exchange |  | vm_foss_web |  | vm_oracle_dr |
| Selected  | Available   |          |           |             |           |  |             |  |             |  |              |
| vm_vsphere1   | vm_client   |          |           |             |           |  |             |  |             |  |              |
|   | vm_exchange   |          |           |             |           |  |             |  |             |  |              |
|   | vm_foss_web   |          |           |             |           |  |             |  |             |  |              |
|   | vm_oracle_dr  |          |           |             |           |  |             |  |             |  |              |

**Profile Settings**

|                  |                  |
|------------------|------------------|
| iSession Profile | vmotion-isession |
| CIFS Profile     | None             |
| MAPI Profile     | None             |
| FTP Profile      | None             |

Cancel Repeat Finished

*Figure 5* Configuring the WAN Optimization policy

## Configuring the BIG-IP GTM

F5's Global Traffic Manager must be configured to direct traffic to the correct LTM virtual server. In our example, we send all traffic to the local data center, unless the utilization alarms we configure are triggered.

### Creating the data centers

In this task you need to create two data centers, called Local and Remote respectively, that correspond to your physical data centers.

#### To create the data centers

1. On the Main tab of the navigation pane, expand **Global Traffic** and click **Data Centers**. The main screen for data centers opens.
2. Click the **Create** button. The New Data Center screen opens.
3. In the **Name** box, type a name for this data center. In our example, we type **Local**.
4. Complete the rest of the configuration as applicable for your deployment.
5. Click the **Finished** button. Repeat this procedure for the Remote data center.

### Creating the monitor

The next step is to create an HTTP monitor.

#### To create the monitor

1. On the Main tab of the navigation pane, expand **Global Traffic** and then click **Monitors**.
2. Click the **Create** button. The New Monitor screen opens.
3. In the **Name** box, type a name for the monitor. In our example, we type **VM-http-monitor**.
4. From the **Type** list, select **HTTP**.
5. Configure the options as applicable for your deployment.
6. Click the **Finished** button. The new monitor is added to the list.

### Creating the GTM servers

The next task is to create servers on the BIG-IP GTM system.

1. On the Main tab of the navigation pane, expand **Global Traffic** and click **Servers**. The main screen for servers opens.

- 
2. Click the **Create** button. The New Server screen opens.
  3. In the **Name** box, type a name that identifies the Local Traffic Manager. In our example, we type **Local-BIG-IP**.
  4. Configure the properties of the server as applicable for your deployment.
  5. In the Health Monitors section, from the **Available** list, select the name of the monitor you created in *Creating the monitor*, on page 12, and click the Add (<<) button. In our example, we select **VM-http-monitor**.
  6. Click the **Finished** button

## Creating the GTM pool

The next task is to create a pool on the BIG-IP GTM system that includes the LTM virtual server in the local data center, and one that includes the LTM virtual server in the remote data center. The remote data center pool should be Disabled after creation.

### To create a GTM pool

1. On the Main tab of the navigation pane, expand **Global Traffic** and click **Pools** (located under **Wide IPs**).
2. Click the **Create** button. The New Pool screen opens.
3. In the **Name** box, type a name for the pool. In our example, we type **Local\_pool**.
4. In the Health Monitors section, from the **Available** list, select the name of the monitor you created in *Creating the monitor*, on page 12, and click the Add (<<) button. In our example, we select **VM-http-monitor**.
5. In the Load Balancing Method section, choose the load balancing methods from the lists appropriate for your configuration.
6. In the Member List section, from the **Virtual Server** list, select the virtual server you created for the application, and click the **Add** button. Note that you must select the virtual server by IP Address and port number combination. In our example, we select **10.133.39.51:80**.
7. Configure the other settings as applicable for your deployment
8. Click the **Finished** button.

## Creating a wide IP on the GTM

The final step in the GTM configuration is to create a wide IP that includes both newly-created pools, and uses the fully qualified domain name (FQDN) you wish to use for the application. In our example, we use **vmhttp.siterequest.com**.

### To create a wide IP

1. On the Main tab of the navigation pane, expand **Global Traffic** and click **Wide IPs**.
2. Click the **Create** button. The New Wide IP screen opens.
3. In the **Name** box, type a name for the Wide IP. In our example, we type **vmhttp.siterequest.com**.
4. From the **State** list, ensure that **Enabled** is selected.
5. From the Pools section, from the Load Balancing Method list, select a load balancing method appropriate for your configuration.
6. In the Pool List section, from the **Pool** list, select the name of the pool you created in *Creating the GTM pool*, on page 13, and then click the **Add** button. In our example, we select **Local\_pool**. Repeat this step for the remote pool.
7. All other settings are optional, configure as appropriate for your deployment.
8. Click the **Finished** button.

This completes the basic GTM configuration. For more advanced GTM configuration options, see the BIG-IP GTM documentation.

---

# Managing VM hosts, VM storage and client traffic between data centers during vMotion events

This section contains information about the management of VM hosts (memory and storage) and client traffic between data centers while the vMotion events are occurring.

## Components to be managed by automation

There are three components that can be scripted to more effectively manage a long distance live migration.

- *Migrating Storage*
- *Migrating the virtual machine*
- *Using ratios to switch data center traffic with GTM*

## Migrating Storage

For this solution, we recommend customers address the Storage vMotion first. For example, the movement of a 5 Gigabyte disk takes the longest amount of time (many minutes or hours) to complete. After this storage migration is completed, the memory portion of vMotion could be completed in a matter of seconds. This order of operation results in the least amount of time when disk I/O operations would be subjected to higher latency.

Addressing storage needs has more than one possible solution:

- Using shared storage between two data centers with replication technologies from a third-party vendor.
- Using Storage vMotion and cross-mounting the shared storage between ESX hosts in both data centers.

## Migrating the virtual machine

In order to manage hosts during vMotion events, the use of scripting and orchestration is recommended. The basic components for orchestration that can be used without additional expenditure are listed below. VMware also provides VMware Orchestrator, part of the vCenter Server Suite, which may be licensed for advanced automation.

- VMware's vSphere Web Services API  
<http://www.vmware.com/support/developer/vc-sdk/>
- F5 BIG-IP's iControl API <http://devcentral.f5.com/Default.aspx?tabid=76>

## Using ratios to switch data center traffic with GTM

We recommend the use of F5 iControl to dynamically manage the ratio or the cutover point for global traffic. This ensures traffic destined for one data center does not overwhelm an increasingly smaller number of hosts. To illustrate this, the following sections examine some typical long distance live migration scenarios.

### ◆ Note

---

*Because the final implementation depends on the automation or orchestration solution used by your implementation, we do not provide detailed procedures for configuring ratios. See the product documentation or DevCentral for more information.*

## Migrating a group of vMotion servers (2 or more)

For the migration of a group of hosts, management through scripting or orchestration, plus the use of the migration iRule and Priority Pool Activation (as described in *Minimizing client downtime*) is recommended to minimize client traffic disruption. As an example, if 10 hosts are to be evacuated from the primary data center to the secondary data center, the chain of events would be as follows:

- Administrative or automated decision is made to move the pool,
- Scripting or orchestration initiates the migration of storage from the primary data center to the secondary data center for the first host.
- Once storage is completed, the memory portion of the host is moved to the secondary data center.
- This process is repeated until 50% of the hosts are migrated at which point, GTM is instructed to direct traffic to the secondary data center,
- Host migration is completed, at which point, any traffic still arriving at the primary data center because of DNS or browser cache are retransmitted to the secondary data center through the use of Priority Pool Activation.

## Migrating a single host

For the migration of a single host, the BIG-IP GTM would be switched immediately after the migration is completed through basic GTM monitoring as described in this document. The use of the migration iRule and Priority Pool Activation (as described in the next section) are not required.

The following table helps clarify when to use the following features we recommend for managing vMotion:

|                                 | Migration iRule | Priority Pool   | Orchestration (vMotion/GTM) |
|---------------------------------|-----------------|-----------------|-----------------------------|
| <b>Two or more pool members</b> | Recommended     | Recommended     | Recommended                 |
| <b>One pool member</b>          | Not recommended | Not recommended | Optional                    |

*Table 1 Recommendations for managing vMotion*

## Minimizing client downtime

Next, we describe how to configure the BIG-IP system to minimize client downtime during vMotion events. This section is optional.

First, we assume you already have the BIG-IP LTM configured for directing application traffic (including a load balancing pool and virtual server). If you do not have the BIG-IP LTM configured for an application, visit our deployment guides for specific instructions:

<http://www.f5.com/solutions/resources/deployment-guides/>.

In the following procedures, you either create a new application pool or modify the existing one, and add an iRule to the virtual server.

You must perform the following procedures on both the primary and secondary BIG-IP systems.

### ◆ Note

*You should perform all of following procedures for every application virtual server participating in vMotion where minimizing client downtime is a consideration.*

The following two sections describe the creation of the optional Priority pools and iRule when vMotion is intended for two or more pool members.

## Creating Priority pools

The first task in this optional section is to configure the application load balancing pool to use the Priority Group Activation feature. We assign a high priority to each local member of the pool, and add the application virtual server on the remote BIG-IP system to the pool with a low priority. So the BIG-IP system first sends traffic local nodes, and during migration as the final host is migrated over, client traffic is maintained through the presence of the virtual server member.

There are two options, you can create a new pool with priority group activation, or modify an existing pool. Use the procedure applicable to your configuration.

### To create a pool with priority group activation

1. On the Main tab, expand **Local Traffic**, and then click **Pools**.
2. Click the **Create** button. The New Pool screen opens.

3. In the **Name** box, type a name for the Pool.
4. In the **Health Monitors** section, select the appropriate health monitor, and click the Add (<<) button.
5. From the **Load Balancing Method** list, choose your preferred load balancing method (different load balancing methods may yield optimal results for a particular network). We typically recommend **Least Connections (node)**.
6. From the **Priority Group Activation** list, select **Less Than**. In the **Available member(s)** box, type **1**.
7. In the New Members section:
  - a) In the **Address** box:
    - For the application nodes, type the appropriate IP address of the application.
    - For the virtual server member, type the IP address of the application virtual server on the remote BIG-IP system. Note: in order to reach this virtual server, SNAT may need to be enabled. See the BIG-IP documentation or online help for configuring SNATs. You may also have to adjust your firewall rules.
  - b) In the **Service Port** box:
    - For the application nodes, type the appropriate port for the application, or select it from the list.
    - For the virtual server member, type the port for the virtual server on the remote BIG-IP system.
  - c) In the **Priority** box:
    - For the application nodes, type **5**.
    - For the virtual server member, type **1**.
  - d) Click the **Add** button to add the member to the list. Repeat this for each application node, however there should be only one virtual server member.
8. Click the **Finished** button.
9. Repeat this entire procedure on the secondary BIG-IP system. Make sure that for the virtual server pool member, you use the IP address of the application virtual server on the primary BIG-IP system.

If you have an existing pool, use the following procedure to modify the pool for Priority Group Activation.

### To modify an existing pool

1. On the Main tab, expand **Local Traffic**, and then click **Pools**.

- 
2. From the Pool list, click the appropriate pool.
  3. On the menu bar, click **Members**.
  4. From the **Priority Group Activation** list, select **Less Than**. In the **Available member(s)** box, type **1**.
  5. In the Current Members section, click a member. In the Configuration section, in the **Priority Group** box, type **5**, and then click the **Update** button. Repeat this step for all existing members.
  6. Click the Current Members **Add** button.
    - a) In the **Address** box, type the IP address of the application virtual server on the remote BIG-IP system.
    - b) In the **Service Port** box, type the appropriate service port.
    - c) In the **Priority Group** box, type **1**.
    - d) Click the **Finished** button.
  7. Click the **Finished** button.
  8. Repeat this entire procedure on the secondary BIG-IP system. Make sure that for the virtual server pool member, you use the IP address of the application virtual server on the primary BIG-IP system.

## Creating the iRule

The next task is to create a new iRule. The following iRule collects any TCP connections that arrive to the primary data center after a pool member has been migrated to the secondary data center and resends these connections to the secondary data center.

This iRule is important because, even when using the BIG-IP GTM, browser caching and slow DNS updates may cause client traffic to arrive at the first data center even though BIG-IP GTM has updated DNS records to indicate that the pool member is now residing in data center 2.

### To create the iRule

1. On the Main tab, expand **Local Traffic** and click **iRules**, and then click the **Create** button.
2. In the Name box, type a name for your iRule. In our example, we use **vmotion-irule**.
3. In the Definition section, type the following iRule (you can also copy and paste, but remove the line numbers):

```

1      #timing on
2      when RULE_INIT {
3          # using cookie persistence
4
5          # Notes:
6          log local0. "version 2.0"
7          set tcp_collect_total 0
8          # release packet assuming length is 1500 max tcp size
9          set release_len 1500
10         # default limit 1M
11         set mem_limit 1048576
12         # elapse time in ms, 1000ms = 1s
13         set time_limit 100000
14         # cookie name
15         set cookie_name "bigip"
16     }
17     when CLIENT_ACCEPTED {
18         log local0. "client connected"
19         set state "standard"
20         set server_ip_addr ""
21         #set server_port ""
22     }
23     when CLIENT_DATA {
24         if {$server_ip_addr ne ""} {
25
26             switch $state {
27                 "buffer" {
28                     log local0. "buffering"
29                     set collect_time {expr [clock clicks -milliseconds] - $start_time}
30                     if {$tcp_collect_total > $mem_limit} {
31                         log local0. "memory limit reached releasing data"
32                         TCP::release $release_len
33                         set tcp_collect_total 0
34                     }
35                     if {$collect_time > $time_limit} {
36                         log local0. "memory limit reached releasing data"
37                         TCP::release $release_len
38                         set collect_time 0
39                     }
40                     TCP::collect
41                     set tcplen [TCP::payload length]
42                     set tcp_collect_total {$tcp_collect_total + $tcplen}
43                     set tflag
44                     continue
45                 }
46                 "standard" {
47                     log local0. "standard - releasing buffer"
48                     TCP::release
49                     TCP::notify request
50                 }
51                 default {
52                     log local0. "error in status change"
53                 }
54             }
55         } else {
56             log local0. "no server ip address"
57         }
58     }
59     when CLIENT_CLOSED {
60         log local0. "[IP::client_addr]"
61     }
62     when SERVER_CONNECTED {
63         log local0. "[IP::server_addr]"
64         set server_ip_addr [IP::server_addr]
65         set state "standard"
66     }
67     when SERVER_DATA {
68     }
69     when SERVER_CLOSED {
70         log local0. "[IP::server_addr]"
71         set state "buffer"
72         set start_time [clock clicks -milliseconds] }

```

4. Click the **Finished** button.

---

## Modifying your application virtual server to reference the iRule

The final step in this section is to modify the application virtual server to reference the iRule you just created.

### To modify the virtual server to use the iRule

1. On the Main tab, expand **Local Traffic**, and then click **Virtual Servers**. The Virtual Servers screen opens.
2. From the Virtual Server list, click the appropriate virtual server.
3. On the menu bar, click **Resources**.  
The Resources page for the virtual server opens.
4. In the iRules section, click the **Manage** button.  
The Resource Management screen opens.
5. From the **Available** list, select the iRule you created in the *Managing VM hosts, VM storage and client traffic between data centers during VMotion events* section, and click the Add (<<) button.
6. Click the **Finished** button.

## Configuring the VMware infrastructure

In this deployment guide, we assume you already have your VMware vMotion implementation up and running. However, there are some modifications you need to make to the VMware configuration for the configuration in this guide to work properly.

### Modifying the VMware ESX configuration

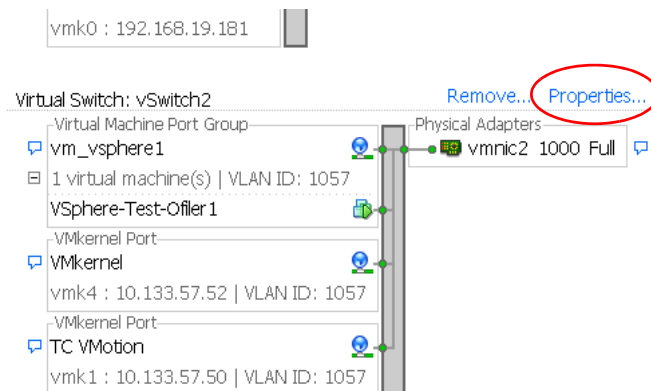
The ESX servers should be configured to have a VMkernel Port for vMotion. This VMkernel port, on an ESX virtual switch should be bound to a unique physical adapter. Each ESX server should have a shared stored device mounted via iSCSI or NFS that both ESX servers can mount. Thus, for testing, storage does not become a gating factor.

### Modifying the VMkernel default gateway

The next task is to modify the default gateway on the VMkernel for vMotion to the self IP address you created in *Creating a self IP*, on page 6.

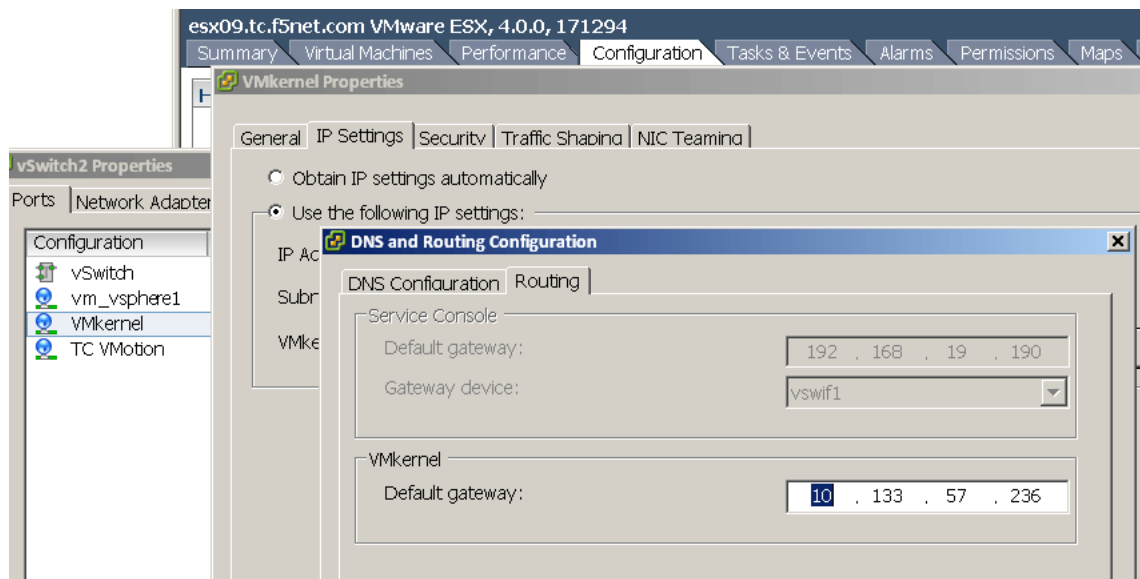
#### To modify the VMkernel default gateway

1. Open the VMware vSphere client, and select the appropriate ESX server host.
2. Click the Configuration tab.
3. In the Hardware box, click **Networking**.
4. From the Networking list, locate the Virtual Switch that contains the vMotion kernel.
5. Click the **Properties** link.



*Figure 6 Properties link of the Virtual Switch*

6. Click to highlight **VMkernel** and then click the **Edit** button. The VMkernel properties box opens.
7. Click the IP Settings tab.
8. Click the **Edit** button next to **VMkernel Default Gateway**. The DNS and Routing Configuration box opens.
9. In the **Default Gateway** box, type the IP address of the self IP on the BIG-IP device.



*Figure 7* Modifying the default gateway

10. Click the **OK** button, and then click **OK** and **Close** to close all the open windows.

The same procedure must be performed on additional ESX servers in both data centers. The VMkernel default gateway in each location should be on a local network.

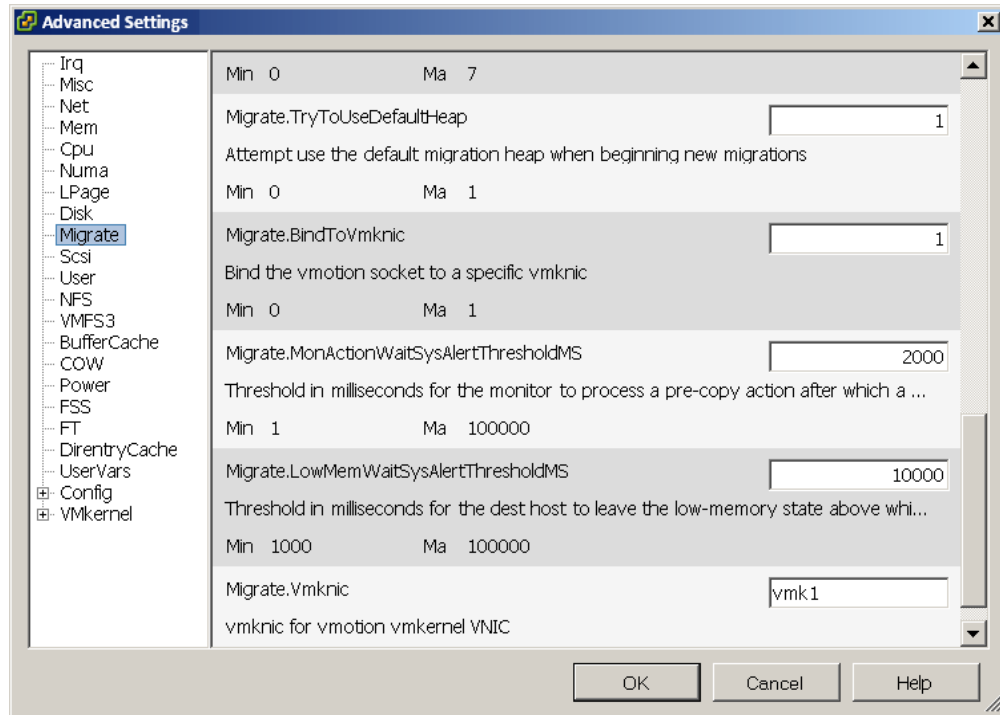
## Binding the ESX devices to a specific vmknic

The final task is to bind the ESX machine to a specific vmknic.

### To modify the VMkernel default gateway

1. Open the VMware vSphere client, and select the appropriate virtual machine.
2. Click the Configuration tab.
3. In the Software box, click **Advanced Settings**. The Advanced Settings window opens.

4. From the left navigation tree, click **Migrate**.
5. In the **Migrate.BindToVmknics** row, type **1** in the box.
6. Click the **OK** button.



**Figure 8** Modifying the *Migrate.BindToVmknics* option

This completes the deployment guide configuration.

---

## Appendix A: Test results

Testing of long distance live migration was carried out in F5's technology center using ESX Server version 4.0, BIG-IP version 10.1 and virtual machines running both Windows XP and Linux. The virtual machines were configured as follows:

- ◆ 1 Gig of RAM  
RAM fully consumed by a content management system; the machine was swapping to disk. The amount of active memory moved across the network was 1 gigabyte.
- ◆ 1 CPU  
Fully utilized (0% idle, with some processes blocked).
- ◆ 10 Gigabytes of disk space  
About 50% utilized. Note that the results quoted are primarily for RAM contents, but the same acceleration is seen on Storage vMotion.

It is important to note that the virtual machines were fully loaded (as described above) during the vMotion events. Active virtual machines take more time and resources to migrate than idle virtual machines.

## Testing methodology

vMotion testing was conducted by initiating vMotion using VMware's vSphere Web Services API while a load test was running to fully use all resources. Part of the test methodology was to insure that there was minimal or no user disruption during the HTTP based test against the content management system (as described in *Minimizing client downtime*, on page 17).

The result using various network conditions follow. The first result demonstrates a large amount of bandwidth and low latency, the second result is with relatively large bandwidth but larger latency. It is evident that even a difference of 20 ms causes large slow-downs for un-aided vMotion. Finally, in the last scenario, there is a fair amount of bandwidth but much higher latency.

The decision on which Storage vMotion to use depends on the type of application, the allowable latencies for users and the distance between the two data centers.

The 622 (OC12) and 1000 results, shaded in the following table, are network conditions that VMware has also tested.

| Network Conditions |                    |                    | BIG-IP with iSessions - average time in seconds | No acceleration - average time in seconds |
|--------------------|--------------------|--------------------|---|---|
| <i>Mbps</i>        | <i>RTT Latency</i> | <i>Packet Loss</i> |   |   |
| 45 (T3)            | 100 ms             | 0%                 | 215 seconds                                     | 823 seconds                               |
| 100                | 25 ms              | 0%                 | 78 seconds                                      | 370 seconds                               |
| 155 (OC3)          | 100 ms             | 0%                 | 209 seconds                                     | 805 seconds                               |
| 622 (OC12)         | 40 ms              | 0%                 | 117 seconds                                     | 357 seconds                               |
| 1000               | 20 ms              | 0%                 | 38 seconds                                      | 132 seconds                               |

**Table 2** Typical migration times of an active virtual machine with 1 gigabyte of memory

Notes:

- For the Gigabit LAN case tests, Jumbo Frames were not turned on.
- Mbps is Megabits per second
- RTT = Round Trip Time

---

## Appendix B: Frequently asked questions and deployment considerations

***Q: Doesn't vMotion require ESX hosts to share a layer 2 bridge? How does this work over a traditional WAN, like the Internet?***

**A:** A common misconception is that vMotion requires a layer 2 bridge to work across a WAN. However, the only technical requirement from a network standpoint for vMotion to succeed is that the network from the guest VM perspective remain identical. This means the guest IP address remains unchanged, and all port groups which touch the guest exist in the source and target ESX hosts. The vMotion traffic itself uses the VMkernel port of ESX, and this does not have to be identical on each host. It is through the VMkernel port and default gateway (which is not shared by the guest VM) that we route traffic across the iSession tunnel.

***Q: Can you summarize what IP addresses are different, and why?***

**A:** The following tables summarize the key network IP addresses.

| IP address                   | Description  | Different for each data center?   |
|------------------------------|--|---|
| LTM virtual server           | Public IP which is used for client connections between clients and the LTM. The GTM determines which LTM virtual server to direct clients to when they initialize new application connections. | Yes   |
| VM guest IP                  | The IP address the guest uses to receive and respond to client requests.   | No  |
| ESX VMkernel IP              | Used for vMotion and Storage vMotion traffic.  | Yes   |
| ESX VMkernel default gateway | The gateway used to route vMotion traffic between hosts.   | Yes. Specifically, this value will be a self IP on the local LTM of each data center. |

***Q: The guide mentions use of a dedicated migration host in each data center to transition from one vCenter to another. Can you elaborate?***

**A:** An alternative deployment scenario would leverage a dedicated host in each data center to be used for long distance live migration with vMotion, analogous to a dedicated host in VMware Fault Tolerant deployments. This host would not be a member of any DRS or HA resource pools, nor would it use any Distributed Virtual Switches you may have in your cluster. In this scenario, the dedicated hosts in each data center would only have to be able talk to the other hosts on the Service Console network (or Management Network on ESXi), and have the same port groups which are accessed by the VM configured on the standard switch(es) of that host (as required by

vMotion). The work flow of migrating a virtual machine from, for example, a DRS cluster from one data center into a DRS cluster in the other data center would work as follows:

1. Both transition hosts would be initially managed by the vCenter in the primary data center.
2. Configure the hosts with access to a datastore in both the primary and secondary data centers.
3. vMotion the VM from the DRS resource pool to this dedicated host.
4. Storage vMotion, then vMotion the VM from the local dedicated host/datastore to the remote dedicated host/datastore.
5. De-register the secondary host from vCenter in the primary site.
6. Register this host with the vCenter in the secondary site.
7. vMotion the VM from this host into the local DRS resource pool in the target data center.

***Q: Do any MAC addresses change during vMotion? What about ARPs and reverse ARPs?***

**A:** No. A key principle of vMotion is that the networking stack from the guest perspective remain exactly the same before and after a vMotion. In a typical LAN scenario, a reverse ARP is issued at the completion of the vMotion in order to tell the local router that traffic bound for the VM has moved guests, and to direct that traffic to the new MAC address of the destination host. This is necessary because in a non-F5 enhanced vMotion event, both hosts are on the same broadcast domain.

However, in this solution, the guest has moved from a host in one physical data center to another. The task of routing traffic bound for the guest is managed not by a single local router, but by GTM and LTM. When the guest arrives in the secondary data center, inbound connections continue to reach the guest VM, because GTM and the LTMs are aware of the guest's location dynamically, and will route those inbound connections to the correct data center where the guest lives at that moment in time. MAC addresses do not change on the guest nor the hosts.

***Q: What are the optimal use cases for F5's long distance live migration solution?***

**A:** A key element of this solution is the transparent redirection of inbound connections between clients and application VMs. Migrating web applications between data centers is the ideal use case for this type of solution, as web applications have many short lived connections between clients and servers. Web applications are almost always session based, meaning once a user begins using a web application, it is important that all requests from that user persist to the same VM. Should that VM migrate from one data center to another, requests from an existing user session must continue to reach the same VM. The F5 solution meets these requirements transparently and effectively, making it an ideal use case.

---

Applications that have long-lived persistent connections, such as SSH, telnet, or streaming media, are not good use cases. Similarly, any applications that are highly transactional, such as database applications, are not good use cases for the solution. Attempting to perform a Storage vMotion (whether local or over long distance) of a database is not recommended and such use cases are better addressed using database replication solutions from storage vendors, which are purpose built for moving and synchronizing highly transactional data between sites.

***Q: What are some suggested strategies for automating this solution?***

**A:** One of the key benefits of both VMware and F5 solutions is the ability to automate complex tasks through published APIs. Automation decreases the risk of human error, simplifies complexity, and reduces operating costs associated with a task by streamlining workflow.

Fortunately, this solution lends itself quite well to automation.

Many organizations already have workflow engines in their environment to automate tasks. Others develop scripts in-house for common functions. In either scenario, the discreet steps of executing a long distance live migration with vMotion can be programmatically executed using the VMware vSphere Web Services API:

1. Execute the Storage vMotion
2. Execute the vMotion
3. (optionally) De-register host with vCenter in the primary data center
4. (optionally) Register host with vCenter in secondary data center.

For further discussion on VMware or vMotion, visit the VMware forums on DevCentral:

**[devcentral.f5.com/Default.aspx?tabid=53&view=topics&forumid=46](http://devcentral.f5.com/Default.aspx?tabid=53&view=topics&forumid=46)**