F5 Technical Brief

Enhancing Application Availability with F5 and NetApp

Leverage F5 and NetApp technologies to improve VMware application availability, enhance disaster avoidance, and simplify storage migration via long distance live migration.

by Matt Quill
Principal Solutions Engineer, F5

and Bikash Roy Choudhury
Solutions Architect, NetApp

Contents

Introduction

As the adoption of virtualization technologies and cloud computing proliferates, IT enterprises face the significant challenge of delivering applications, storage, and computing power to remote or regional offices, geographically dispersed data centers, or even hybrid private/public clouds. To adapt to these ever changing circumstances, IT personnel need to devise solutions that closely integrate storage, application, and networking technologies for rapid deployment and mobility.

In a previously published paper,[1] F5 presented a model for long distance live migration with VMware that enabled administrators to move applications across any distance. This paper will address the extension of this functionality by describing a joint solution developed by F5 and NetApp that utilizes VMware long distance live migration. F5® BIG-IP® WAN Optimization Module™ (WOM), running on BIG-IP® Local Traffic Manager™ (LTM) with BIG-IP® Global Traffic Manager™ (GTM), in combination with NetApp FlexCache enables LAN-like read performance of applications over the WAN and ensures that applications and storage can be migrated more securely and faster than ever. Together, F5 and NetApp offer rapid application mobility with sustained performance. This paper will review the results of extensive interoperability testing performed at F5's San Jose test facilities.

With F5 BIG-IP WOM and NetApp FlexCache, enterprises can now accelerate applications and storage to transparently redirect users to new sites. Adopting a solution that combines the FlexCache storage acceleration technology with BIG-IP WOM acceleration can significantly extend the distance enterprises can migrate their virtual machines—far beyond metropolitan areas to regional and even global locales.

The Benefits of Long Distance Live Migration

Consistent User Experience

When an enterprise moves an application, administrators can maintain access and performance, and both in-memory applications and storage can be accelerated to data centers across regions or continents.

Disaster Avoidance

Together, BIG-IP WOM and NetApp FlexCache enables enterprises to rapidly redeploy resources when known events could cause business disruption. For example, when a data center is notified of an approaching storm, the IT staff can redeploy mission-critical applications without sustaining an outage. Providing IT staff with this crucial capability enables migration not only between data centers but also to remote hosted infrastructures. With the ability to move applications and storage over distances exceeding five thousand miles without disruption, businesses can maintain operations even during large-scale weather events.

Rapid Application Re-Deployment

Together, BIG-IP WOM and FlexCache enable applications and virtual images to be re-distributed beyond regions and even across continents, so administrators can redeploy applications closer to users.

Business Continuance

The combination of BIG-IP WOM and BIG-IP GTM and FlexCache enables data mobility across regions or continents without application downtime. Administrators can transport live applications and storage anywhere in the world while maintaining availability.

BIG-IP WOM Deduplication and Compression Reduces Bandwidth Requirements

Using BIG-IP WOM with adaptive compression and symmetric deduplication ensures that applications can realize LAN-like performance over the WAN so application performance is not degraded while migration operations take place. For example, if a user requests data from a remote data center's cold cache (managed by FlexCache), BIG-IP WOM ensures that the remote data center's read operations are accelerated.

Network Design

Since migration is a simple transfer of a snapshot of a particular server, all IP information is encapsulated within the snapshot. The migration of a virtual image over distance needs to encapsulate the IP address and login information.

The Ether-IP feature of the BIG-IP system enables administrators to create layer 2 tunnels as they migrate applications between sites. New client connections are automatically redirected via BIG-IP GTM so that the movement to the new site is transparent to clients. During the entire testing process, application availability must be maintained without downtime and with minimal degradation in performance. The long distance live migration must transparently redirect users to the new site with little or no effect on the user. Finally, BIG-IP GTM handles DNS updates and redirects to the new site gracefully and without manual intervention. BIG-IP GTM ensures that new sessions are automatically redirected to new sites so that failover is transparent.

About NetApp FlexCache

NetApp FlexCache provides cache architecture at the storage protocol layer. Similar to the way a cache in the memory architecture of a compute system improves performance, FlexCache improves performance in NFS environments by scaling out cache volumes for increased input/output per second, bringing data closer to hosts for decreased latencies, and offloading overburdened storage controllers. As a storage performance accelerator, FlexCache alleviates the bottleneck of the origin systems, scales up the aggregate throughput to the clients, and reduces latency across the WAN by caching data close to the clients. It can also be used as a simple tiered storage deployment that moves data between storage tiers without tedious administrative overhead. The current supported protocol of FlexCache includes NFS v2 and v3.

FlexCache architecture supports caching on both memory and disk. Data is pulled from the origin system on demand to memory in the caching system and then stored on local disk. Any write from the client is proxied back to the origin so no dirty data stays in the caching system. FlexCache can be deployed as an add-on software license at no additional cost to existing NetApp FAS or V-Series systems or through dedicated storage accelerator systems that consist of the following:

- **Origin volume:** A NetApp Data ONTAP volume that is mapped to a FlexCache caching volume as the original source of data.
- **Cache volume:** A Data ONTAP volume that is created as a FlexCache caching volume.

The key use cases for FlexCache in this long distance live migration are:

- Immediately access data on the remote site after a migration is performed on the VM.

- Keep the NFS mounts for the VMware ESX servers at different data centers that are local to each location. This avoids stretching the NFS protocol and makes it simple and easy to manage and troubleshoot.
- Further leverage the ESX server resources in the new location by migrating VMDK files to the new storage location, rather than causing resource constraints or bottlenecks on the ESX servers at the original location.

For more information about NetApp FlexCache, refer to the technical report, FlexCache Caching Architecture.

About F5 BIG-IP WOM

BIG-IP WAN Optimization Module (WOM) overcomes network and application issues on the WAN to ensure that application availability, data replication, and disaster recovery requirements are met. These services are integrated directly on existing BIG-IP devices and include superior compression, encryption, and traffic control capabilities that dramatically reduce data replication times and enable more efficient use of existing bandwidth.

The Massive Cost Savings of WAN Acceleration

WAN acceleration technology holds the promise of providing LAN-like application performance over WAN circuits. Organizations have historically responded to slow application performance over the WAN by increasing their WAN network capacity. However this has several downsides:

- Application performance problems over the WAN are frequently the result of latency and protocol "chattiness" as opposed to raw throughput.
- The cost of upgrading WAN capacity can be substantial, especially at the upper end of the bandwidth spectrum.
- In typical replication or transfer scenarios, the same data is retransmitted repeatedly, so data transfers over the WAN that do not use compression or deduplication technologies frequently send substantially more data than is required.

WAN acceleration technologies have been designed to address this challenge. Instead of investing substantial dollars in upgrading circuit capacity, WAN acceleration appliances combine deduplication, compression, TCP optimization, and application-specific tuning to ensure that WAN capacity is used to the fullest extent. Deduplication ensures that only unique data blocks are transferred over

the network, compression adds an additional level of data reduction, and TCP optimization addresses the challenge of latency and chatty applications over distance. The result is that WAN acceleration can frequently deliver 10× or greater performance without new investment in network capacity.

About F5 BIG-IP GTM

Deploying multiple data centers is a big step toward protecting a business from site outages and improving application performance. But to fully achieve these goals, enterprises need an efficient way to monitor infrastructure and application health, and to control this distributed infrastructure according to business needs.

BIG-IP Global Traffic Manager (GTM) provides a more intelligent way to respond to DNS queries than simple load balancing among multiple data centers. BIG-IP GTM distributes user application requests based on business policies, data center conditions, network conditions, and application performance. This gives administrators holistic control over global traffic to ensure high availability and maximum performance for applications running across multiple dispersed data centers. The result is better application performance, less downtime, and simplified management.

Test Scenarios

Test Environment Configuration

For the purposes of this test, performed in F5's San Jose facilities, F5 built an environment with a pair of NetApp FAS2040 arrays running Data ONTAP version 7.3.2. Because BIG-IP WOM is a symmetric deployment, F5 had a pair of BIG-IP 6900 devices running BIG-IP WOM and a separate BIG-IP 6900 device running BIG-IP GTM to redirect traffic between sites. To simulate WAN latency, F5 used a LANforge device. They created a 200 GB volume in each location and configured a separate 200 GB NetApp FlexCache volume. Although the size of the FlexCache volume is typically smaller than the actual production volume, in this instance F5 simulated larger data sets. F5 chose SharePoint 2010 as the application to test, and to model a high transaction environment they used LoadRunner as the test utility. The specifics of the combined NetApp and F5 test-bed are outlined in the tables below.

Hardware	Functionality	Quantity
BIG-IP 6900 device	Application delivery and WAN optimization	3
NetApp FAS2040	File storage	2
LANforge	WAN simulation	1

Table 1: F5 test lab hardware configuration

Software	Version	Notes
BIG-IP WOM	10.2	RTM build
NetApp Data ONTAP	7.3.2	NetApp OS
LoadRunner	9.50	Load simulation
SharePoint	2010	Application test
LANforge	5.0.8	WAN simulator
BIG-IP GTM	10.2	Global Traffic Manager

Table 2: F5 test lab software configuration

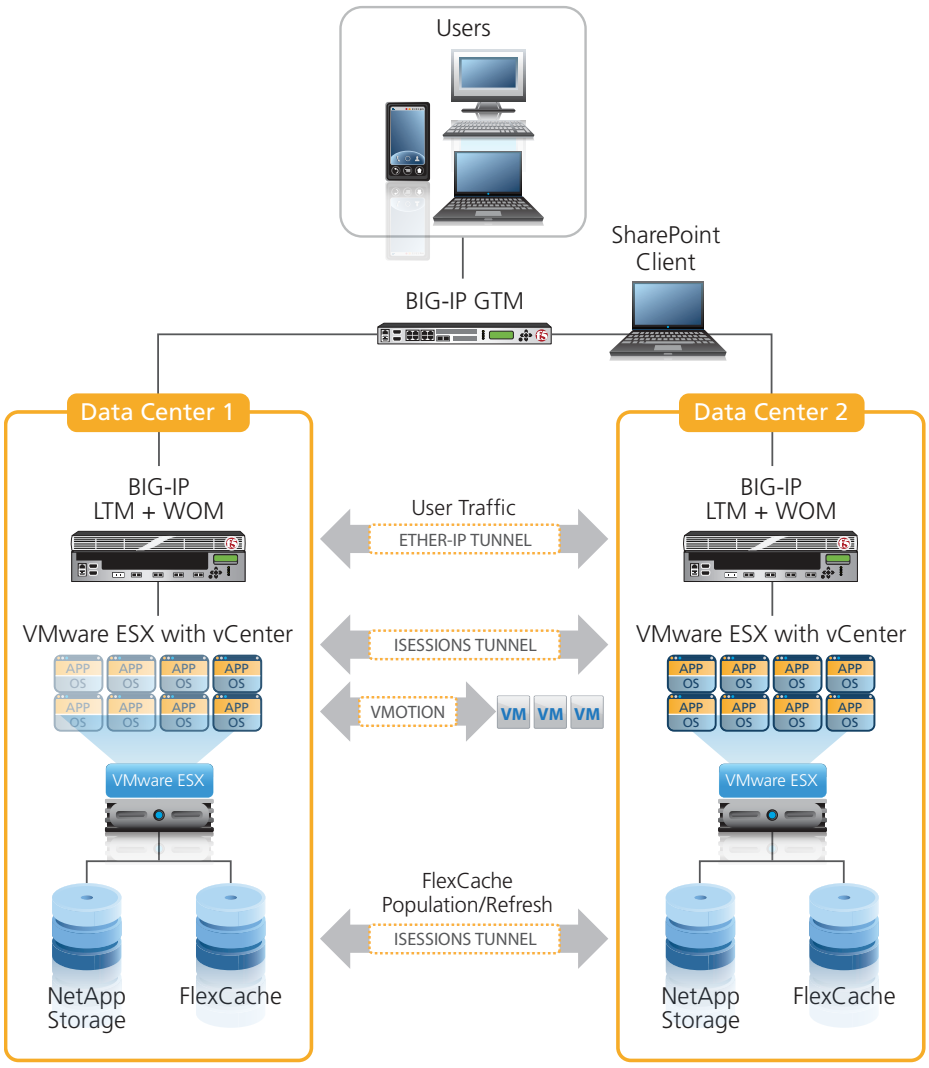


Figure 1: BIG-IP and NetApp FlexCache environment

Outline of Testing Scenarios

F5 and NetApp joined forces to test two separate scenarios. The first scenario migrated only the VM and not the storage. Long distance live migration of the OS alone would demonstrate that application performance could be maintained over distance with the combination of FlexCache and BIG-IP WOM. FlexCache accelerated read performance from remote storage while BIG-IP WOM deduplicated and accelerated traffic. In the second scenario, a combination of FlexCache and BIG-IP WOM accelerated storage migration by combining BIG-IP WOM deduplication

and a warm cache. Combining the read acceleration of FlexCache with the WAN optimizations of BIG-IP WOM ensured that both applications and storage could be rapidly moved over distance with minimal effect on the user. The link speed tested was a single GbE with a LANforge device to introduce WAN latency. An initial baseline test was performed over the GbE link with zero latency and BIG-IP WOM and FlexCache disabled. After the baseline test was run, latencies of 10, 20, 30, and 40 ms were introduced.

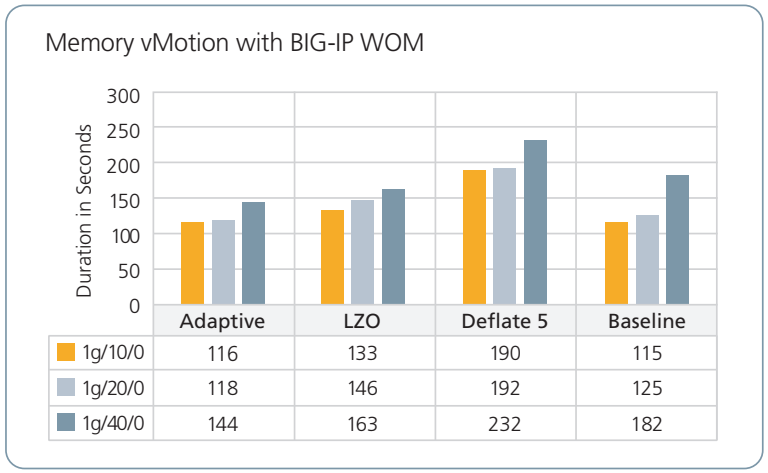
In the baseline test without BIG-IP WOM and FlexCache, the long distance live migration failed at latencies above 10 ms. The second scenario, BIG-IP WOM without FlexCache, tested long distance live migrations of both the application and storage was successful. Adding FlexCache increased the number of successful transactions and accelerated movement of storage data to the secondary data center.

Adopting both technologies improved overall performance up to 30 percent more than using BIG-IP WOM alone. Each available compression method was tested during the migration, with adaptive compression providing up to 30 percent more performance benefit than no compression. After the 40 ms test was performed, a final 300 ms long distance live migration test demonstrated that that an application could be potentially relocated to extreme distances. The graphs below demonstrate the results of F5 and NetApp combined caching, compression, and deduplication products.

Test Results: BIG-IP WOM with and without FlexCache

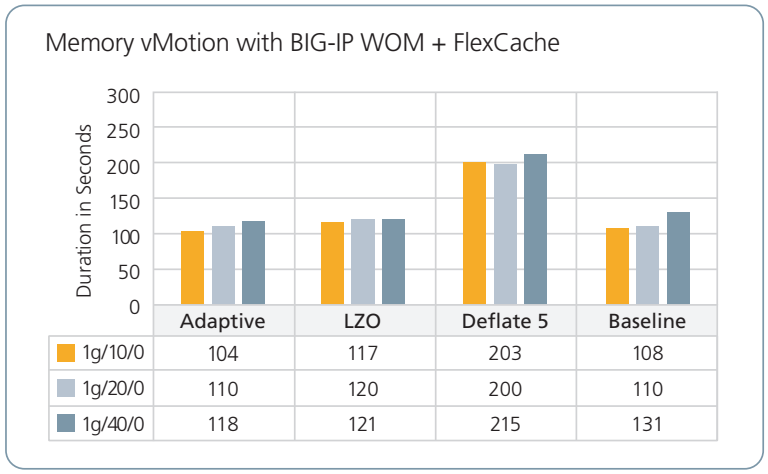
Below are the test results that demonstrate the improvements realized by running various combinations of BIG-IP WOM and NetApp FlexCache. The BIG-IP system enables users to select a compression method—Adaptive, LZO, Deflate 5, or Baseline (none)—or to allow BIG-IP to determine the most efficient compression algorithm for a particular data set. Symmetric adaptive compression typically ensures the fastest data reduction for any traffic between BIG-IP systems. Symmetric adaptive compression automatically selects and uses the appropriate Adaptive, Deflate, and LZO compression algorithms (or no compression if the data cannot be compressed) to maximize bandwidth usage and throughput. In addition, symmetric adaptive compression can use BIG-IP hardware compression where available to provide unprecedented scalability.

The first baseline test ran a simple vMotion migration without FlexCache enabled. The second test (Figure 2) performed a memory migration with FlexCache using LZO compression, and FlexCache reduced migration time by almost 25 percent.



	Adaptive	LZO	Deflate 5	Baseline
1g/10/0	116	133	190	115
1g/20/0	118	146	192	125
1g/40/0	144	163	232	182

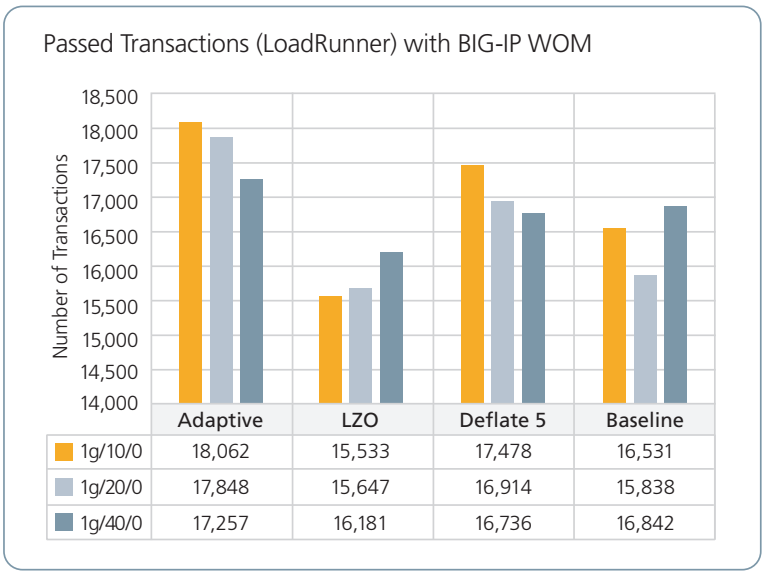
Figure 2: Memory migration: BIG-IP WOM only (1 GB connection latencies of 10, 20, and 40 ms with zero packet loss)



	Adaptive	LZO	Deflate 5	Baseline
1g/10/0	104	117	203	108
1g/20/0	110	120	200	110
1g/40/0	118	121	215	131

Figure 3: Memory migration: BIG-IP WOM with FlexCache

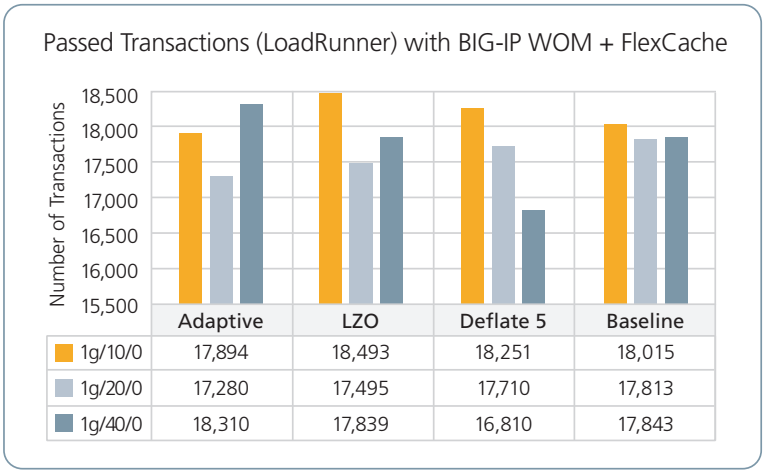
F5 ran the LoadRunner utility to measure transactions while the OS and application were being migrated to determine the additional benefits that the combination of BIG-IP WOM and FlexCache would provide. Figure 4 shows a migration without FlexCache simulating 10, 20, and 40 ms of latency.



	Adaptive	LZO	Deflate 5	Baseline
1g/10/0	18,062	15,533	17,478	16,531
1g/20/0	17,848	15,647	16,914	15,838
1g/40/0	17,257	16,181	16,736	16,842

Figure 4: Passed transactions: BIG-IP WOM only

After completing the first test, F5 re-ran the same test but with FlexCache enabled. Figure 5 shows that the overall number of successful transactions increased by 10–20 percent with FlexCache versus without.

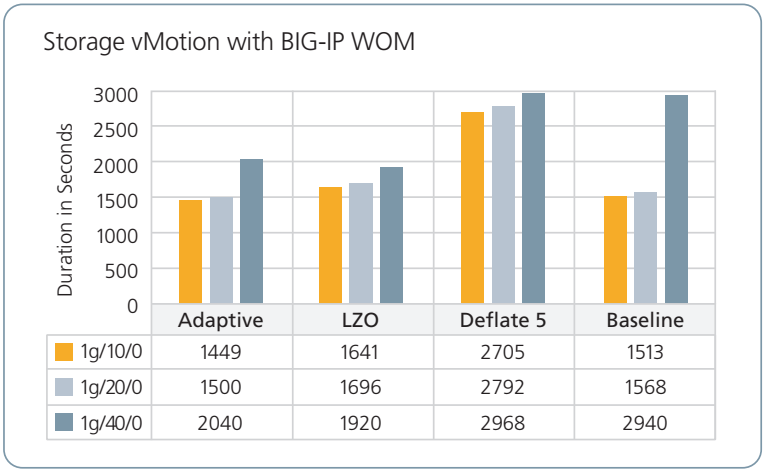


	Adaptive	LZO	Deflate 5	Baseline
1g/10/0	17,894	18,493	18,251	18,015
1g/20/0	17,280	17,495	17,710	17,813
1g/40/0	18,310	17,839	16,810	17,843

Figure 5: Passed transactions: BIG-IP WOM with FlexCache

Note that with BIG-IP WOM and FlexCache combined, the number of passed transactions went up by over 10 percent. The final test was run to determine whether BIG-IP WOM and FlexCache combined would also accelerate storage migration.

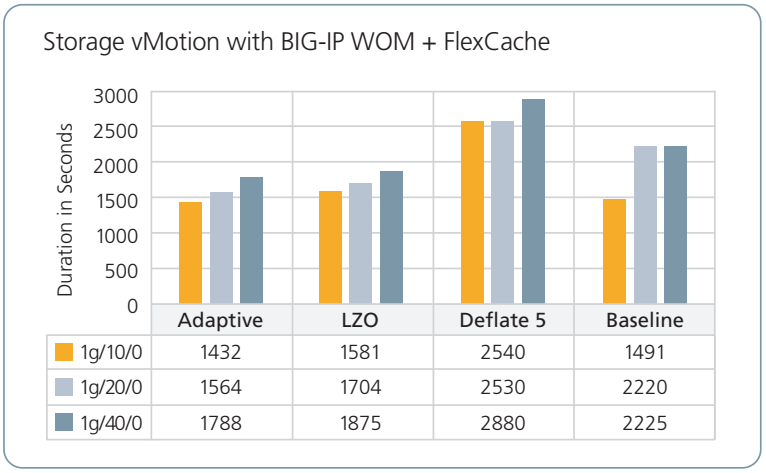
In the next scenario (Figure 6) a storage migration was initiated with only BIG-IP WOM enabled.



	Adaptive	LZO	Deflate 5	Baseline
1g/10/0	1449	1641	2705	1513
1g/20/0	1500	1696	2792	1568
1g/40/0	2040	1920	2968	2940

Figure 6: Storage migration: BIG-IP WOM only

Figure 7 shows the same scenario as Figure 6, but with FlexCache enabled. It improves the time to move the storage anywhere from 5 to 30 percent depending on the compression method used. The only exception was the 20 ms RTT Baseline or Adaptive compression scenario.



	Adaptive	LZO	Deflate 5	Baseline
1g/10/0	1432	1581	2540	1491
1g/20/0	1564	1704	2530	2220
1g/40/0	1788	1875	2880	2225

Figure 7: Storage migration: BIG-IP WOM with FlexCache

Conclusion

Enterprise IT environments have experienced revolutionary changes in recent years. Technologies such as virtualization, once just enablers of test and development environments, have evolved into full-scale, mission-critical applications. Separating applications and operating systems from dedicated physical resources not only allows enterprises to efficiently use compute resources, but also to enhance application mobility. By handling migration operations within the data center, administrators can address resource constraints and increase uptime. Extending this functionality beyond the metropolitan area can protect against localized events. F5 and NetApp have a proven solution that enables application mobility across regions and even continents. Combining the application acceleration and symmetric deduplication capabilities of F5 BIG-IP WOM with the read cache acceleration in NetApp FlexCache results in a simple, flexible solution that enables full application availability.

Using F5 BIG-IP WOM and BIG-IP GTM in combination with NetApp FlexCache can:

- Enable organizations to perform live migrations of live applications and storage over distances exceeding 3,000 kilometers.
- Substantially improve application availability and responsiveness during live migrations.
- Provide rapid disaster avoidance capabilities without taking applications and servers offline.
- Accelerate traffic over the WAN to overcome latency and packet loss issues.
- Seamlessly redirect new requests to the secondary data center via BIG-IP GTM.

[i] *Enabling Long Distance Live Migration with F5 and VMware vMotion*

F5 Networks, Inc. 401 Elliott Avenue West, Seattle, WA 98119 888-882-4447 www.f5.com

F5 Networks, Inc.
Corporate Headquarters
info@f5.com

F5 Networks
Asia-Pacific
apacinfo@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa
emeainfo@f5.com

F5 Networks
Japan K.K.
f5j-info@f5.com

CS01-00046 0511