



Distributing Applications for Disaster Planning and Availability

Managing applications in multiple data centers in real time can be a challenge, especially when the data centers are geographically distributed. F5 products such as BIG-IP Global Traffic Manager, BIG-IP Link Controller, and BIG-IP WAN Optimization Module help you manage and distribute applications between data centers.



WHITE PAPER

Distributing Applications for Disaster Planning and Availability

Introduction

Site outages due to natural disasters, attacks, or application infrastructure failures are a major cause of lost revenue, diminished customer satisfaction, and reduced user productivity.

Very few organizations-typically only large, globally distributed companies or web companies that have a reason for local geographic presence-have the foresight and the budget to prepare for disaster by building data centers in varied geographic sites. More often, organizations either opt for regional secondary physical data centers, setting up their secondary sites in an active-standby configuration with a manual recovery process; or they leverage the cloud to provide a virtual, on-demand data center. Both options can be costly, error prone, and slow if deployed haphazardly. When they're most needed, these secondary data centers can experience broken transactions between data centers; cause customer dissatisfaction with user-facing applications; and incur downtime costs that severely disrupt business and decrease profitability.

Creating secondary data centers as part of a complete application delivery solution doesn't have to be difficult or complicated. F5 offers a suite of tools that enable you to leverage all the benefits of secondary sites. BIG-IP Global Traffic Manager (GTM) can holistically manage data center applications and user access in an active-active configuration across multiple sites; BIG-IP Link Controller can maintain upstream connectivity for data centers between multiple providers; and the BIG-IP WAN Optimization Module (WOM) can optimize all connections between data centers, such as user connections and site-to-site data replication.

Keeping Your Data Center Healthy

Maintaining applications in multiple data centers can cause a whole host of issues, from keeping the applications in sync, to managing user connections to those applications, to ensuring everything is up and running. Application availability is a paramount concern when dealing with disaster recovery. Keeping your applications available across multiple data centers poses a variety of challenges:

- **Lack of visibility into data center and application health.** How do you gauge the health of the data center and application?
- **Sub-optimal user experience.** When organizations deliver applications, how do they handle broken sessions, retrieve lost data, and secure personal information?
- **Maintenance overhead.** Too often, organizations have no choice but to shut down the entire data center to perform upgrades, and site-to-site data replication across the WAN can eat up valuable time. If you're an e-commerce site, can you imagine the lost revenue?
- **DNS management and security.** Domain Name System (DNS) is possibly



WHITE PAPER

Distributing Applications for Disaster Planning and Availability

the most critical and pervasive networking technology businesses use, yet it continues to be one of the most vulnerable. What happens when DNS management errors break your entire application infrastructure? Older BIND versions are more susceptible to attacks and are difficult to upgrade without the proper management tools. New security threats such as zone file tampering, DNS pharming, DoS, and SYN floods are constantly emerging. Unfortunately, DNS is often misunderstood and mismanaged within the enterprise, which can lead to configuration and architecture errors that expose vulnerable points in the network.

Disaster Recovery Management

Planning for a disaster while keeping applications up and running across sites can easily turn into a perpetual loop of fixing broken transactions, maintaining user satisfaction, and juggling downtime. Using a manual process to solve these challenges can be costly, error prone, and slow-disrupting business and decreasing profitability.

Organizations need a solution that enables them to solve these challenges by providing:

- Superior application availability and performance.
- Reduced management overhead.
- Improved operational efficiency.

The ultimate solution would give organizations an intelligent way to manage data centers and applications under any condition or circumstance. It would be able to detect data center and application health-including internal and external web services-from a single management platform. In the event of a problem, this solution needs to automatically and transparently reroute users to the appropriate location while maintaining application and service availability.

Maintaining application availability between sites and during a disaster (or even during peak loads) is not simply relying on ping and keeping DNS records updated. It requires a much more in-depth architecture and an application delivery architecture-a complete solution-that can provide the following:

- **Holistic application monitoring.** It's not enough to check whether the application is up or down. The solution's application monitoring should include checking the application and factoring in all dependencies. Automating the failover process eliminates management overhead, minimizes costly downtime, and removes the guesswork involved in tracking interdependencies.
- **Service management and maintenance.** By following good management guidelines, the solution should be able to intelligently track and manage dependencies in a multisite application infrastructure. The most helpful management tool would facilitate the identification and monitoring of the application infrastructure dependencies from a single location for at-a-glance



WHITE PAPER

Distributing Applications for Disaster Planning and Availability

operational efficiency.

- **User client continuity.** The solution should be able to direct users to the appropriate data center based on the state of the data center, the application, any web service dependencies, and user-based information such as location and credentials. Tracking the application state based on user continuity and application monitoring is essential to ensuring that the right content, without broken sessions or lost data, is delivered to users.
- **DNS management.** The best solution should make managing DNS simple and error-free, because one minor configuration error can bring down an entire application infrastructure. Simple fixes to this problematic scenario include an easy-to-use user interface, DNS error checking, and automatic reverse lookups.
- **Security.** Organizations need a holistic and integrated approach to securing the network and applications against potential threats and attacks.

Disaster Preparedness with BIG-IP Devices

The BIG-IP family of Application Delivery Controller (ADC) devices provides high availability, maximum performance, and centralized management for applications running across multiple data centers. Built on F5's unified, modular, and scalable TMOS architecture, BIG-IP ADC devices manage and distribute user application requests and application traffic according to business policies and data center, network, application, and web service conditions to ensure maximum availability.

BIG-IP GTM

BIG-IP Global Traffic Manager (GTM) is the cornerstone of distributing applications across multiple data centers. Beyond basic DNS, BIG-IP GTM provides granular application delivery management as users and applications move between data centers under normal or adverse circumstances.

Holistic health monitoring

Acting as the global application traffic cop, BIG-IP GTM checks the health of the entire application infrastructure across all data centers, eliminating single points of failure and routing traffic away from poorly performing sites. By collecting performance and availability metrics from data centers (through BIG-IP Local Traffic Manager), ISP connections, servers, caches, and even user content, BIG-IP GTM ensures high availability and adequate capacity.

Application-centric monitoring



WHITE PAPER

Distributing Applications for Disaster Planning and Availability

Today's applications are more sophisticated and require intelligent health checking to determine stability and availability. Instead of relying on a single health check, BIG-IP GTM aggregates multiple application monitors so state can be verified across many levels. This results in higher availability, improved reliability, and the elimination of false positives, which reduces management overhead.

Only BIG-IP GTM provides pre-defined, out-of-the-box health monitoring support for more than 18 different applications, including SAP, Oracle, LDAP, MySQL, and more. BIG-IP GTM performs targeted monitoring of these applications-and any application where a custom monitoring profile is attached-to accurately determine health, reduce downtime, and improve the user experience.

BIG-IP GTM tracks the health of applications that are dependent on one another and marks all related objects down if the health check of one object in that group fails. For example, if the SharePoint web service is up and answering requests but it can't access the SharePoint data store, BIG-IP GTM will mark the SharePoint web service as down and unavailable because it can't access data, even though the web service is actually up. This enables you to align and monitor application objects according to business logic and profitability, build scalable traffic distribution policies, and better manage application dependencies.

In the event of a disaster, or when any application is moved from one data center to another, BIG-IP GTM applies this deep application monitoring to determine which data center should handle specific user requests. As applications become more diversified between multiple data centers and the cloud, the ability to correlate all application tiers into one management and distribution system will become critical.

Client continuity

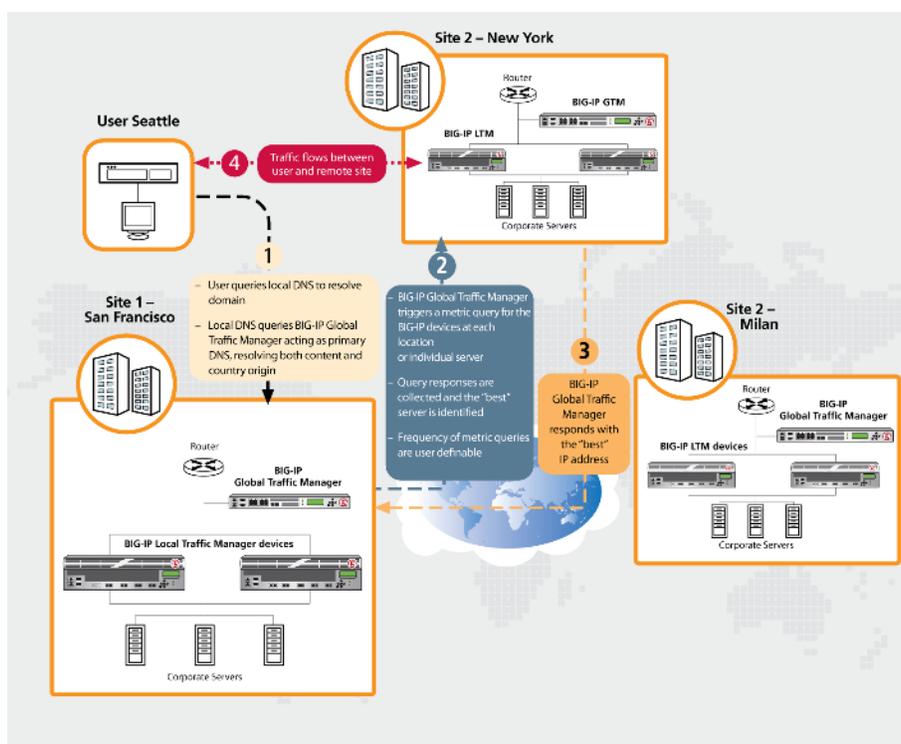
To deliver a superior user experience, BIG-IP GTM not only monitors applications, but also tracks user state as users make requests to applications within the data center. Users can persist across applications and multiple data centers and be transparently routed to the appropriate data center or server based on application and user state. If a user begins a session with an application in one data center and that application is moved to another data center, session integrity is always maintained-there are no broken sessions or lost or corrupted data. Organizations gain improved infrastructure scalability, better TCO, and reduced support calls by relying on BIG-IP GTM to manage user access to applications across multiple data centers.

DNS Management and Security

WHITE PAPER

Distributing Applications for Disaster Planning and Availability

BIG-IP GTM is a global DNS solution providing name services at the very edge of the application delivery network. By employing geographic location services, BIG-IP GTM can direct users to the best application delivery data center based on their physical location. Working in concert within the data center, BIG-IP Local Traffic Manager (LTM) can load balance local and recursive DNS services creating a fault-tolerant architecture from the mobile edge through to the application. BIG-IP GTM also creates a more secure DNS environment by providing protection against DNS-based attacks and by supporting DNSSEC.



BIG-IP GTM detects availability and performance problems across data centers to automatically reroute user application requests to the best-performing site.

BIG-IP Link Controller

BIG-IP Link Controller is a link load balancer that manages multiple upstream network connections and distributes application traffic across the appropriate link. BIG-IP Link Controller is often used for QoS deployments where certain applications are routed across specific network links for the highest performance. BIG-IP Link Controller is also used as a fault-tolerance and disaster recovery tool, guaranteeing that at least one upstream network connection is always available for application traffic.

ISP redundancy



WHITE PAPER

Distributing Applications for Disaster Planning and Availability

Any data center deployment that maintains only one link to the public network exposes a single point of failure, which in turn exposes serious network reliability issues. Typically used with multiple upstream ISP links, BIG-IP Link Controller uses sophisticated monitors to detect errors across an entire link, providing end-to-end reliability and availability over multiple WAN links. BIG-IP Link Controller intelligently manages bi-directional traffic flows from a data center to create a fault-tolerant and optimized Internet connection for all applications.

BIG-IP Link Controller also eases multi-homed deployments to protect the data center's external network from ISP failures. DNS-based technology removes the dependency on routing technologies such as border gateway protocol (BGP) to provide failover capabilities. By using this technology, BIG-IP Link Controller eliminates multi-homed problems such as latency, high update overhead, and the inability to manage application traffic over the public Internet. You can also aggregate inexpensive links; data about performance, costs, and business policies will help you determine which link to use.

BIG-IP WAN Optimization Module

Using dispersed data centers comes with two key concerns. One is managing access to applications and data during a disaster. The other is initially moving applications and their data to those data centers, and keeping all data in sync. Keeping data centers in replicated sync comes with two main challenges: the massive amount of data needed for true replication, and having to rely on latent WAN links that are typically part of the public Internet. The BIG-IP WAN Optimization Module (WOM) helps manage both of these challenges by overcoming network and data constraints to ensure that applications and their data can flow between data centers.

Optimizing the WAN for data replication

The amount of static and dynamic data stored in a typical enterprise data center is well into the multi-terabyte range, and moving that data is time-consuming and can place a huge burden on the WAN links between data centers. When you first deploy a new data center, you must copy that large amount of data to the new location. Then, it becomes a constant real-time updating and replication challenge to make sure each data center is the same as the next. Constant replication between data centers at best taxes a WAN link, and at worst completely consumes all available bandwidth, limiting the amount of application and user traffic that can flow between data centers.



WHITE PAPER

Distributing Applications for Disaster Planning and Availability

When connecting multiple data centers and cloud-based locations, BIG-IP WOM relies on a feature called iSessions to connect, optimize, and secure WAN links. iSessions allows BIG-IP WOM to manage the entire WAN connection between data centers at BIG-IP hardware throughput speeds, up to 12 Gbps. Once in place, BIG-IP WOM can use iSessions to optimize the WAN link and data, and then encrypt that data before it goes across the public Internet.

BIG-IP WOM and iSessions optimize both the WAN connection and the data traversing that connection to help minimize the strain on WAN links. BIG-IP WOM includes symmetric adaptive compression, which dramatically reduces and optimizes TCP traffic between links, BIG-IP devices, and data centers. BIG-IP WOM can offload services such as caching and compression from individual application servers to alleviate that processing overhead at the server level. In other words, BIG-IP WOM takes care of all application data optimization at the link level, allowing servers to serve data rather than manage data transfers.

Data de-duplication: send the data once

One of the most common issues plaguing slow WAN links is repetitive data. There is no need, for example, for a static website to send the same static images 100 times between data centers during an application migration event. The image data needs to be sent across the WAN link only once and then referenced over and over again at the destination data center.

BIG-IP WOM can alleviate repetitive data with de-duplication: symmetric caching of recurring data across the WAN. As data passes through each symmetric BIG-IP WOM device, BIG-IP WOM determines whether any part of the data was previously sent to the corresponding BIG-IP WOM device. Previously sent data blocks are replaced with much smaller data references, which are sent instead. Data referencing becomes more critical as larger blocks of identical data are sent across the WAN, for example in real-time replication and in similar data blocks contained in virtual machines.

Long distance vMotion

A relatively new addition to the disaster recovery data center landscape is the virtual machine. Virtualization and virtual machines have become a staple in the enterprise data center, but only recently has using fully packaged virtual machines for disaster recovery become feasible. With tools such as VMware's vMotion technology, local fault-tolerance in the data center has become trivial: if a piece of hardware fail, the virtual machines running on that hardware can be restored on functioning hardware. So far, this technology is limited to the local data center, however.



WHITE PAPER

Distributing Applications for Disaster Planning and Availability

With a combination of BIG-IP products-BIG-IP Local Traffic Manager (LTM), GTM, iSessions, and WOM-virtual machines running in a local data center can now easily be replicated to other data centers anywhere in the world over public WAN links, including to a cloud provider. By using BIG-IP LTM and iSessions to link multiple data centers together and BIG-IP WOM to optimize both the connection link and the data that flows over that connection, applications and data running in virtual machines can be easily migrated to multiple data centers. In the event of a disaster or other real-time need to move virtual machines, BIG-IP GTM and LTM work together to maintain user connectivity between data centers during the migration.

Conclusion

With the BIG-IP family of products, F5 provides the industry's most comprehensive solution for managing disaster recovery, site failover, and business continuity. In addition to performing comprehensive site availability checks with BIG-IP GTM, BIG-IP WOM and Link Controller can also optimize both the application data and the WAN links to maintain service levels between data centers and with user-based connectivity. BIG-IP products provide transparent delivery of applications and Web services across multiple sites, ensure global business continuity and application availability, and improve performance and customer satisfaction by directing users to the best site on a global basis. Working as a single integrated solution, BIG-IP products also reduce management overhead by providing a holistic view into application and data center health across the entire distributed network.

Whether you are building a true disaster recovery architecture, designing multiple distributed-load data centers, or even moving some application services into the cloud, BIG-IP GTM, LTM, WOM, and Link Controller provide the application delivery management you need to keep applications secure, fast, and available. During disaster or in the normal course of enterprise IT operations, the applications maintain user visibility and keep your business running.

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

Americas
info@f5.com

Asia-Pacific
apacinfo@f5.com

Europe/Middle-East/Africa
emeainfo@f5.com

Japan
f5j-info@f5.com