Disaster Recovery: Not Just Planning for the Worst

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

Congratulations to you if you had the foresight and the budget to build another data center in a different geographic site. Unfortunately, most organizations set up their secondary site in an active-standby configuration with a manual recovery process that can be costly, error prone, and slow. They experience broken transactions, customer dissatisfaction, and downtime costs that severely disrupt their business and decrease profitability.

This paper describes how businesses can use F5 Networks' BIG-IP Global Traffic Manager to leverage all the benefits of their secondary site in an active-active configuration to holistically manage their applications across multiple sites and provide:

- Application high availability by tracking and managing multiple data centers, applications, and any web services that work together within a composite application. This involves tracking dependencies to transparently redirect users to an available site, eliminating the cost of downtime.

- Application state and persistence tracking to ensure that users persist to the application in the same site to eliminate broken sessions and corrupted data.

- Application maintenance to ensure that applications can be brought up or down across multiple sites, eliminating downtime and any errors or guesswork without management overhead.

This paper also describes how you can use F5 Networks' BIG-IP Link Controller to maintain ISP link connectivity and WANJet™ to accelerate site-to-site data replication across the WAN.

Challenges   Keeping your applications available across multiple data centers pose 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 do their upgrades, or what about the time it takes to do site-to-site data replication across the WAN. If you're an e-commerce site, can you image the lost revenue?

- **Impact of DNS mistakes –** Domain Name Server (DNS) management continues to be the least understood and the most pervasive networking technology used by businesses. What happens when DNS management errors break your entire application infrastructure?

- **Security vulnerabilities –** Older BIND versions are more susceptible to attacks and are hard to upgrade without the proper management tools. New security threats are constantly emerging such as Zone file tampering, DNS Pharming, DoS, SYN floods, etc. Recent attacks against Akamai and the root DNS servers in the United States and pharming attacks against .com DNS servers, confirm that DNS-level attacks are growing in number. Unfortunately, DNS is often poorly understood, exposing susceptible points in the network because of configuration/architecture errors.

The Ultimate Solution

All too often, keeping applications up and running across sites consist of running the treadmill of fixing broken transactions, minimizing customer dissatisfaction, and juggling downtime. Using a manual process to solve these challenges can be costly, error prone, and slow, disrupting business and decreasing profitability.

What organizations need is a solution that enables them to solve these challenges and provide:

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

The ultimate solution gives organizations an intelligent way to manage their data centers and the applications they host. They need a way to detect the data center's and application's health including any web services in a composite application. In the event of a problem, the solution needs to automatically and transparently reroute the user to a functional application.

Sounds great, but what exactly does a solution like this need?

- **Holistic Monitoring** — It's not enough to check if the application is up or down. The solution must take a holistic approach, checking both the application and factoring in all dependencies. Automating the failover process eliminates management overhead, minimizes the cost of downtime, and removes the guesswork involved in tracking interdependencies.

- **Client Continuity –** The solution should be able to direct users to the appropriate data center based on the state of the data center, application, web service dependencies, and user identity. Tracking the application state is essential to making sure that the users are delivered the right content without broken sessions or lost data. An intelligent solution should also be able to maintain the user's session by resolving the user back to the same data center, tracking the user's identity, transaction history, and the dependencies between services.

- **Service Management and Maintenance –** Following good management guidelines, the solution should be able to intelligently track and manage dependencies in a multi-site application infrastructure. The most helpful management tool would facilitate the identification and monitoring of the application infrastructure dependencies from a single locale for at-a-glance operational efficiency.

- **DNS Management –** The best solution should make the job of managing DNS simple and error free, especially because one minor configuration error can bring down an entire application infrastructure. Easy 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 secure the network and applications against potential threats and attacks.

Solution   F5 Networks' BIG-IP Global Traffic Manager provides high availability, maximum performance, and centralized management for applications running across multiple data centers. Built on F5's modular and scalable TMOS architecture, the BIG-IP Global Traffic Manager distributes user application requests according to business policies and data center, network, application, and web service conditions to ensure the highest possible availability. The following sections describe how the BIG-IP Global Traffic Manager accomplishes these tasks.

Holistic Health Monitoring

In an active-active configuration, the BIG-IP Global Traffic Manager checks the health of the entire application infrastructure, eliminating single points of failure and routing traffic away from poorly-performing sites. By collecting performance and availability metrics from each data center, ISP connection, servers, caches, and even user content, the BIG-IP Global Traffic Manager ensures high availability and adequate capacity prior to directing traffic to a site.

Application-Centric Monitoring

Today, applications are more sophisticated and require intelligent health checking to determine availability. Instead of relying on a single health check, the BIG-IP Global Traffic Manager aggregates multiple monitors so you can check application state at multiple levels. This results in higher availability, improved reliability, and the elimination of false positives to reduce management overhead.
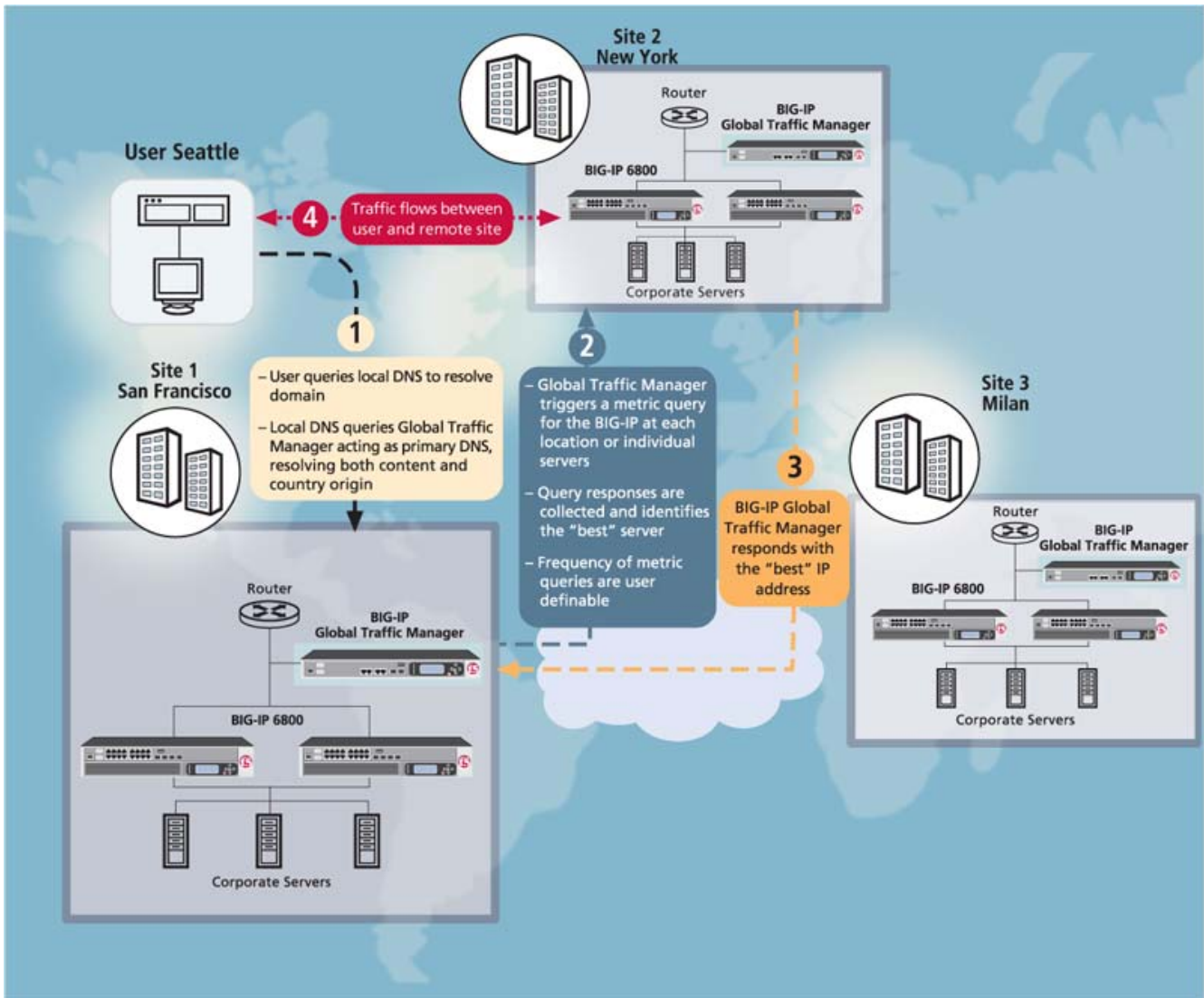
Only the BIG-IP Global Traffic Manager provides pre-defined, out-of-the-box health monitoring support for over 18 different applications, including SAP, Oracle, LDAP, mySQL, and more. The BIG-IP Global Traffic Manager performs targeted monitoring of these applications to accurately determine their health, reduce downtime, and improve the user experience.

The BIG-IP Global Traffic Manager 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. 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.

Client Continuity

The BIG-IP Global Traffic Manager is the only solution that tracks application state and provides the intelligence to deliver a superior user experience. Users can persist across applications and data centers and be transparently routed to the appropriate data center or server based on application state. Session integrity is always maintained, with no more broken sessions, lost, or corrupted data. Organizations gain improved infrastructure scalability, better TCO, and reduced support calls.

The following figure shows how the BIG-IP Global Traffic Manager detects availability and performance problems across data centers to automatically reroute user application requests to the best-performing site.

Wide Area Persistence

The BIG-IP Global Traffic Manager provides sophisticated modes of persistence to ensure that users are directed to the right resources. It intelligently distributes traffic to the same site to maintain consistency for applications or transactions. The BIG-IP Global Traffic Manager synchronizes persistence information across all devices, ensuring that users are directed back to the same site regardless of their entry point. Finally, it propagates the desired persistence information to local DNS servers to reduce the required frequency of synchronizing back-end databases.

Composite and Shared-Services Applications

Applications have evolved from being business enablers to being the core business of organizations. Gone are the days when all users would connect to one centralized mega site to access information. Applications have become mission critical, extending their reach to a global market. With the move to a Service Oriented Architecture, applications have also increased in complexity by becoming stateful with embedded Web services that are multi-tiered for execution.

As the sites hosting composite or shared-service applications increase in number and complexity, performance also takes a toll. In this situation, organizations could end up losing business and revenue due to unsatisfied customers. Zona Research reported that over $25 billion dollars are lost every year due to poor web performance. In the case of financial firms, downtime impacts revenue, delivery, and government compliance that can lead to severe penalties. Finally, the threat of DNS attacks and the lack of centralized management for shared-service applications can also impede an organization's ability to be productive and profitable.

The BIG-IP Global Traffic Manager tracks the dependencies of web services within an application. If any service is not available, the BIG-IP Global Traffic Manager automatically marks the entire application down and transparently reroutes user requests for that application to an alternative site. This scheme automates your failover process to eliminate management overhead, the cost of downtime, and the guesswork involved in tracking web service interdependencies.

The following example shows an online e-commerce application that consists of the shopping cart web service (shoppingcart.domain.com), a checkout web service (checkout.domain.com), and a user database web service (userdatabase.domain.com). The user adds items to a shopping cart and comes back after an hour to check out the items. If the user were resolved to data center 2 for the checkout process, the session would break because all the user's information (items purchased, login information) resides in data center 1. The transaction would fail and the user would have to start all over again.
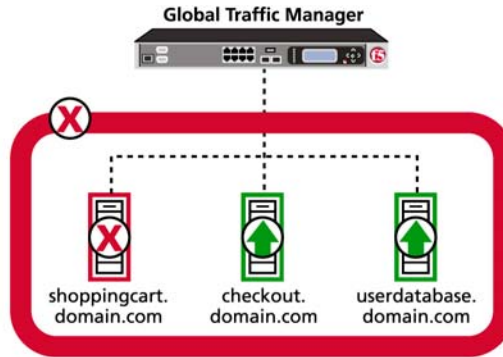
Load Balancing

The BIG-IP Global Traffic Manager includes the industry's most advanced traffic distribution capabilities to match the needs of any organization or globally-deployed application. These include:

Application Availability	LDNS Persistence	Ratio
Dynamic Ratio	LDNS Round Robin	Round Robin
Geography	Least Connections	Round Trip Time
Global Availability	Packet Completion Rate	User-defined QoS
Hops	Packets Per Second	Virtual Server Capacity
Kilobytes Per Second		

The BIG-IP Global Traffic Manager routes users to the best global resource based on comprehensive site and network metrics. For example, the Quality of Service (QoS) load balancing mode includes a hops coefficient, based on the number of hops between the user and the local DNS. Managers can use hop rate to send the user to the data center that has the fewest hops between the user and the data center, ensuring more rapid access. Dynamic Ratio load balancing mode solves the problem of "winner takes all" common to other global traffic management systems. Dynamic Ratio sends a portion of traffic to the best performing site, second best performing site, and so on in proportion to the health and performance of the network and server resources.

Geographic Load Balancing

The BIG-IP Global Traffic Manager resolves IP addresses down to the country, increasing topological control for managing global traffic. For sites maintaining content in different languages, this ensures that users around the world get the information they need in their own language.

Custom Topology Mapping

The BIG-IP Global Traffic Manager offers organizations deploying Intranet applications the ability to set up custom topology mappings. By defining and saving custom region groupings, you can set up topology based on traffic distribution policies that match your internal infrastructures.

Intelligent Traffic Routing Control

Only the BIG-IP Global Traffic Manager features an intelligent programming language, iRules, which you can use to customize the dynamic distribution of global traffic. The BIG-IP Global Traffic Manager looks deep inside Domain Name Server (DNS) messages to distribute application traffic to the desired data center, pool, or virtual server. This capability reduces

latency, increases protection against malicious attacks, and improves application performance. Because iRules is based on an easy-to-use TCL-based scripting language, development and administrative costs are nominal.

The following iRule looks at every incoming DNS request and checks for the LDNS address. If it is the desired LDNS address, then the appropriate pool (3DNS_pool) is used. For DNS requests from all other LDNS servers, if the resource record type is type "A" and the www.domain.com is the requested site, the request is redirected to a another site.

```
rule CLIENT_RULE {
    when DNS::REQUEST {
        if ( IP::remote_addr == 10.10.10.10 )
            use pool 3dns_pool1
        }
        else if ( RRTYPE == "A" && RRNAME == "www.domain.com" )
            use cname www.redirect.domain.com
        }
    }
}
```

ISP Redundancy

As organizations increase their use of the Internet to deliver Web applications, maintaining only one link to the public network exposes a single point of failure, which poses a serious network vulnerability. The BIG-IP Link Controller monitors the availability and performance of multiple WAN ISP connections to intelligently manage bi-directional traffic flows to a site, providing fault tolerant and optimized Internet access.

The BIG-IP Link Controller uses sophisticated monitors to detect errors across an entire link to provide end-to-end, reliable WAN connectivity. In the event of a failure, traffic is transparently directed across other available links so users stay connected.
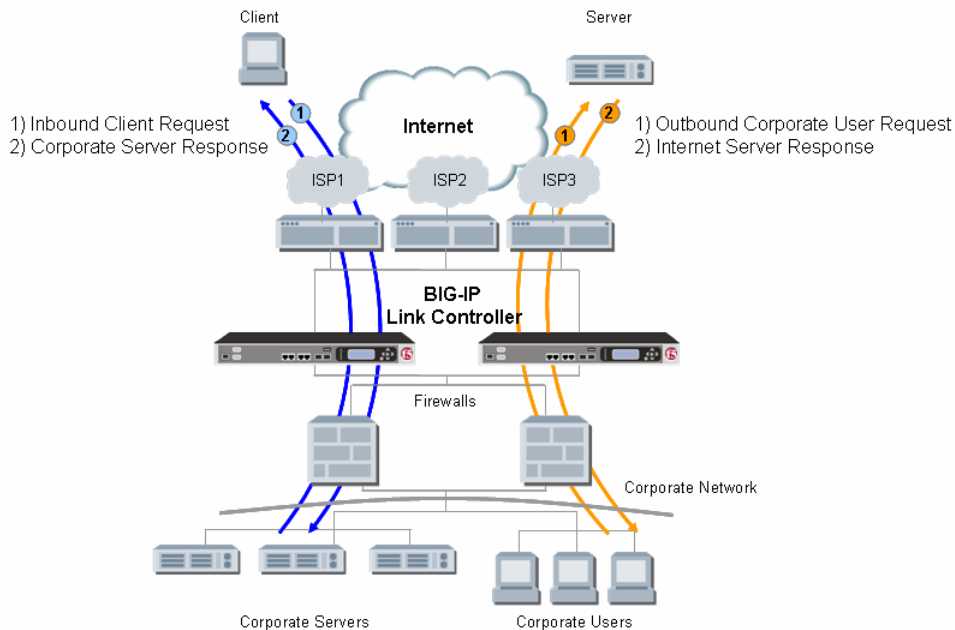
The BIG-IP Link Controller also eases multi-homed deployments so you no longer need ISP cooperation, large bandwidth connections, designated IP address blocks, ASNs, or high-end routers to protect your network from ISP failures. Using DNS-based technology that removes the dependency on BGP to provide failover capabilities, the BIG-IP Link Controller eliminates multi-homed problems such as latency, high update overhead, and inferior traffic management.

You can also aggregate inexpensive links, with more control over which link to use based on performance, costs, and business policies.

Service Management and Maintenance

Managing a distributed network across multiple sites from a single point is an enormous challenge. The BIG-IP Global Traffic Manager provides the tools that give you a global view of your infrastructure with the means for managing the network and business policies to ensure the high availability of your business-critical applications.

DNS Management

Only the BIG-IP Global Traffic Manager provides an integrated zone file management tool, ZoneRunner, which reduces DNS risks and streamlines DNS zone file management. ZoneRunner provides a secure environment for managing your DNS infrastructure while reducing the administrative burden by validating and error-checking zone files. Built on the newest version of BIND, ZoneRunner provides:

- Auto population of commonly-used protocols
- Validation and error checking for zone file entries
- Secure environment for DNS management
- Command line version
- Zone importation from an external server or a file
- Automatic reverse lookups
- Reduced administration for a lower TCO
- Improved infrastructure scalability

The BIG-IP Global Traffic Manager also delivers breakthrough DNS performance to handle even the busiest Internet sites. Organizations can now provide the best Quality of Service for their users while eliminating poor application performance.

Powerful, Web-Based Management Interface

The BIG-IP Global Traffic Manager provides a simple and cost-effective way for organizations to manage their global infrastructure from a centralized location, including:

- Efficient list and object management for complete visibility of global resources

- Unique naming of global objects to reduce administration and build the infrastructure around business policies

- Superior sorting and searching for fast access to global objects that can be targeted and controlled

- Streamlined setup and object creation to reduce configuration times

- Context-sensitive help for information on objects, commands, and configuration examples

- Ability to manage distributed applications as part of one collective group

Distributed Application Management

Organizations have struggled to align their applications and infrastructure with their business goals and policies. The BIG-IP Global Traffic Manager gives you the ability to implement dependencies between applications and manage them efficiently. With distributed application management, organizations can reduce administrative costs, build scalable traffic distribution policies, and improve efficiency with granular control of data center objects.

Automated Setup and Synchronization

Autosync automates setup and secure synchronization of redundant BIG-IP Global Traffic Manager devices. With Autosync, you can make configuration changes from any BIG-IP Global Traffic Manager in the network, eliminating difficult hierarchical management common to DNS.

Configuration Retrieval

VS AutoDiscovery enables the BIG-IP Global Traffic Manager to pull down configurations from any number of distributed BIG-IP systems. In large enterprises, this removes the need to repeat configurations across products, saving time.

SNMP Management Application Support

The BIG-IP Global Traffic Manager integrates its MIBs and a SNMP agent with DNS. This allows SNMP management applications (for example, HP OpenView) to read statistical data about the current performance of the BIG-IP Global Traffic Manager. SNMP management packages have an exact view of what the BIG-IP Global Traffic Manager is doing, while keeping an eye on standard DNS information.

Data Center and Sync Groups

The BIG-IP Global Traffic Manager enables the creation of logical groups of network equipment to ensure the efficient use of monitoring and metrics collection. The result is a highly scalable solution that can support the Internet's busiest sites by intelligently sharing the information with members in the logical group.

IPv6 Support

With the demand for IPv6 increasing, many sites are facing new requirements to handle IPv6 traffic. The BIG-IP Global Traffic Manager provides scalability and support for the next generation network, resolving AAAA queries with improved manageability that doesn't require wholesale network and application upgrades.

3rd Party Integration

The BIG-IP Global Traffic Manager provides the industry's most flexible solution by communicating and integrating with a broad array of network devices. This includes support for various types of remote hosts, including SNMP agents: UCD, snmpd, Solstice Enterprise, and the NT/4.0 SNMP agent.

The BIG-IP Global Traffic Manager can also communicate with 3rd-party caches, servers, routers, and load balancers to accurately diagnose the health of your network endpoints and provide a heterogeneous solution for global traffic management.

Security

Organizations are increasingly being exploited at the DNS level with DoS attacks that compromise the security of their site. Difficulty in differentiating between legitimate DNS requests and attacks is also a very real concern. The BIG-IP Global Traffic Manager includes inherent security controls and features to protect against attacks, and to keep applications and legitimate traffic moving.

The BIG-IP Global Traffic Manager ships, by default, in a very secure mode with the following features:

- Uses packet filtering to limit or deny access to and from Web sites based on monitoring the traffic source, destination or port.

- Is a hardened device designed to resist common attacks, including:
  - Thwarts teardrop attacks
  - Protects itself and servers from ICMP attacks
  - Does not run SMTPd, FTPd,Telnetd, or any other attackable daemons

- Tolerates high levels of DNS attacks, protecting resources while maintaining maximum and continuous availability for applications and services.

- iRules can help create policies that block DNS requests from rogue sites or known sources of attacks before they can do damage.

Data Replication across the WAN

CitiStreet, one of the largest global benefits delivery firms in the United States, needed to maximize the performance of its network to backup 9 million employees' data over the WAN on a daily basis. As its existing T3 lines began to reach capacity, CitiStreet knew it had only two options: retool their infrastructure, including adding expensive additional bandwidth, or deploy WAN optimization and application acceleration technology to make the best use of what they already had. With the installation of F5 Networks' WANJet™, CitiStreet securely transferred private information 6X faster and showed a 20X improvement in throughput.

F5 WANJet was deployed at each data center to accelerate applications and maximize throughput over the WAN. WANJet also enabled CitiStreet to send the data fully encrypted between WANJet appliances without using expensive CPU cycles on CitiStreet's large UNIX servers.

After installing WANJet, CitiStreet was able to reduce the time it took to replicate 6 GB of data from 55 minutes down to 9 minutes, and reduced bandwidth consumption from 16 Mps to 4 Mbps. As a result of WANJet's unique compression technology, 6 GB of data now appears as less than 300 KB, reducing the amount of data sent by 20X while simultaneously improving application performance.

WANJet now plays a key role in the technology backbone of CitiStreet's Business Continuity Preparedness Plan, which guarantees CitiStreet customers the highest level of uninterrupted service and support.

Conclusion

The BIG-IP Global Traffic Manager provides the industry's most comprehensive solution for site failover and business continuity. In addition to performing comprehensive site availability checks, you can define the conditions for shifting all traffic to a backup data center, failing over their entire site, or controlling only those affected applications.
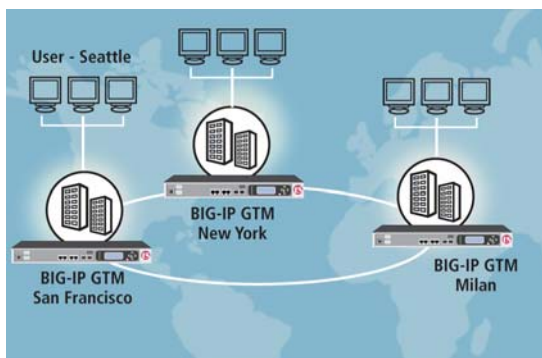
Key benefits include:

- Provides transparent delivery of applications and web services across multiple sites

- Ensures global business continuity and application availability

- Improves performance and customer satisfaction by directing users to the best site on a global basis

- Increases flexibility by directing users to the best site according to any business policy including geography, load, time of day, etc.

- Reduces management overhead with a centralized, holistic view into application and data center health

- Increases efficiency, scalability, and ROI of the global network by leveraging secondary data centers in an active-active configuration

- Enables automation of complex tasks to reduce maintenance and management overhead

Like the BIG-IP Global Traffic Manager, F5 Networks' BIG-IP Link Controller uses similar capabilities to maintain ISP link connectivity. Using unique compression technology, WANJet accelerates site-to-site data replication across the WAN.

About F5 Networks

F5 Networks is the global leader in Application Delivery Networking. F5 provides solutions that make applications secure, fast, and available for everyone, helping organizations get the most out of their investment. By adding intelligence and manageability into the network to offload applications, F5 optimizes applications and allows them to work faster and consume fewer resources. F5's extensible architecture intelligently integrates application optimization, protection for the application and the network, and delivers application reliability – all on one universal platform. Over 10,000 organizations and service providers worldwide trust F5 to keep their applications running. The company is headquartered in Seattle, Washington with offices worldwide. For more information, go to www.f5.com.