Keeping Up with Multi-Service Applications

Overview

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. They have also increased in complexity by becoming stateful with embedded Web services that are multi-tiered for execution, using next generation technologies such as VoIP, XML, and so on.

For organizations deploying these types of multi-service applications, ensuring application availability and manageability are key to staying in business. Downtime and management overhead can not only cause loss of revenue, but can also cause significant delays in the delivery of products and services, jeopardizing customer goodwill.

As the sites hosting multi-service applications increase in number and complexity, end-user experience 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.

The threat of DNS attacks and the lack of centralized management for multi-service applications can impede an organization's ability to be productive and profitable.

This white paper describes how F5's BIG-IP® Global Traffic Manager can solve the challenges of a multi-service application infrastructure to provide maximum availability, reduce infrastructure costs and management overhead, and improve application performance for a superior end-user experience.

Challenges

Applications are no longer monolithic entities running on a megaserver serving content to users. They have evolved into a complex set of disparate and collaborating services. These services are usually stateful, have interdependencies, and are often distributed across geographically-dispersed sites. Organizations are capable of managing only individual services and are not able to holistically deliver multi-service applications. Some of the challenges facing organizations that are delivering multi-service applications include:

- **Lack of visibility into application health** — In a multi-service application environment, organizations lack the tools to track the health and interdependencies between services to accurately determine the health of these applications. When these services go down, some organizations are faced with a manual failover process that involves tracking the services that have dependencies and then shutting them down one by one. This process is not only labor intensive, but prone to error especially when organizations have a large number of applications and services not to mention the cost of downtime.

  Organizations also need to know if these applications are delivering the right content since the up/down status of an application isn't enough to guarantee high availability.

- **Sub-optimal user experience** — As organizations distribute their mission-critical applications to multiple sites, they have to manage their global traffic so that end-users are always routed to the best performing site. Multi-service applications

must track and remember end-user transactions across services especially if they have dependencies between them. Organizations fail when they try to deliver individual services to users without taking into account the state of these services, their interdependencies, and the identity of the end user. End-users are easily frustrated with unavailable sites, broken transactions, and lost data and will quickly click to another site, giving their business to someone else. Gartner conducted a survey that found out,
"Internet users will wait no more than 20 seconds to see content."

- **Application performance degradation** — Traffic conditions in the network are very dynamic and change unpredictably. This can cause a surge of traffic at a website that can overwhelm the infrastructure and the application, causing a severe degradation in performance. Network administrators need the ability to adapt and evolve with changing network conditions to dynamically distribute traffic across sites. A proactive approach is to define custom traffic distribution policies that align with their business goals.

- **Maintenance outages** — Network administrators are continually faced with the challenge of performing planned maintenance on distributed sites without causing an interruption to application availability. With a multi-service application infrastructure, this becomes an even bigger challenge since only the services that are part of an application undergoing maintenance need to be turned off. Too often, organizations have no choice but to shut down the entire data center to do their upgrades.

- **Global impact of management mistakes** — As applications become more distributed, network management of multi-service applications across multiple sites poses a huge challenge. Any errors or mis-configurations are now felt at a global level. Organizations need management tools that can provide a global view of their infrastructure while giving them an easy way to manage the network and their business policies while protecting application availability. Integrating multi-vendor management solutions is not only expensive, but can easily turn into a management nightmare. And with the emergence of IPv6, network administrators must manage their infrastructure to scale and support both IPv4 and IPv6 records.

- **DNS management** — Managing DNS is very cumbersome, and most administrators do so from the command line. As the number of DNS zones grow, adding, deleting and moving DNS records via a command line interface can easily lead to errors. DNS administrators need an easy, flexible, and powerful user interface that scales with the DNS infrastructure and simplifies the task of managing DNS.

  Administrators also want the seamless integration of global load balancing with their existing DNS infrastructure. They don't want to change their existing policies and DNS records especially when managing millions of zones.

- **Security vulnerabilities** — Organizations need a holistic and integrated approach to secure the network and applications against potential threats and attacks. 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.

Solution

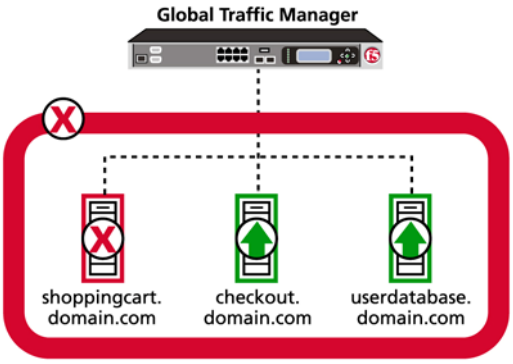
The BIG-IP Global Traffic Manager

F5's BIG-IP Global Traffic Manager is built on TMOS that provides organizations with a holistic way to solve the challenges of multi-service applications, including performance, security, high availability, and management. The BIG-IP Global Traffic Manager is the only solution that provides organizations with the flexibility and scalability they need to adapt to evolving application requirements while delivering the best quality of service to their end users.

Complete Application Health

The BIG-IP Global Traffic Manager provides comprehensive health monitoring of applications rather than merely reporting up/down status. The BIG-IP Global Traffic Manager delivers *complete* health checking that you can use to determine true application availability and shift traffic to better performing or available sites. With the BIG-IP Global Traffic Manager, you can reduce administrative burden, improve reliability, and eliminate the cost of downtime. Better management of distributed applications start with comprehensive application monitoring, including:

Multi-service application monitoring — The BIG-IP Global Traffic Manager can holistically monitor the health of multi-service applications, factoring in the dependencies between the various services. In the following example, the BIG-IP Global Traffic Manager automatically tracks the dependencies between the shopping cart, checkout, and user database services, marking the entire application down any one of these services fail. This helps you automate your failover process, eliminating management overhead, the cost of downtime, and the guesswork involved in tracking service interdependencies.



Application-specific monitors — The BIG-IP Global Traffic Manager provides pre-canned health monitor support for over 18 applications such as SIP, Oracle, LDAP, mySQL, etc. With the BIG-IP Global Traffic Manager, you can perform targeted application monitoring to accurately determine the health of each application, ensuring that they are delivering the right content.

Composite monitors — The BIG-IP Global Traffic Manager provides composite monitors that check the state of an application at multiple levels before determining its availability. You can combine different types of monitors in the context of an application to improve availability and eliminate loss of data. This also helps eliminate the additional overhead incurred by false positives, reducing administrative costs.

Client Continuity for Multi-Service Applications

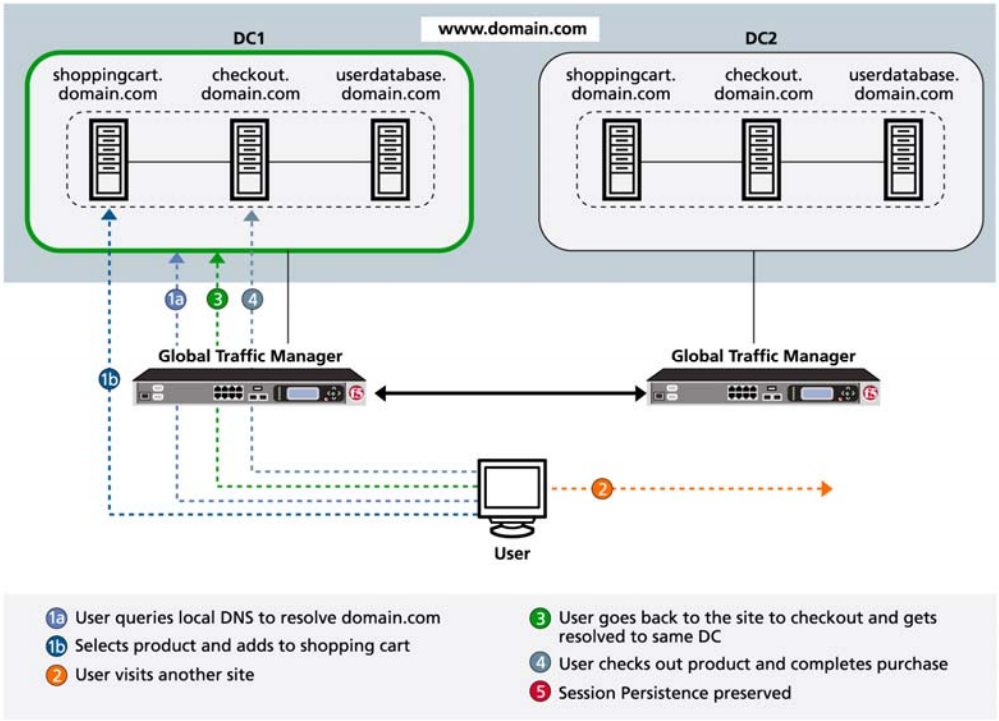
The BIG-IP Global Traffic Manager can direct end users to the appropriate service based on the:

- State of the application
- Dependencies between services
- Identity of the end user

The BIG-IP Global Traffic Manager can track application state so that end users are delivered the right content without broken sessions or lost data. This ensures application session integrity and improves the end user experience as users sessions can persist across applications, and users can be routed to the appropriate data center or server based on the state of the application.

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

The BIG-IP Global Traffic Manager maintains the client session by resolving the client back to the same data center, tracking the client's identity, transaction history, and the dependencies between services. Since the shopping cart in data center 1 is filled with items to check out, the BIG-IP Global Traffic Manager persists clients across multiple services to the same data center so that client session is preserved.



Intelligent Traffic Routing

F5 leads the industry with the ability to create custom traffic distribution policies for global traffic based on business criteria. Using iRules, a TCL-based scripting language,

you can direct the BIG-IP Global Traffic Manager to trigger off DNS events and inspect DNS messages to distribute application traffic to the appropriate data center, pool, or virtual server.

The following iRule inspects 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 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
        }
    }
}
```

Powerful 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. The BIG-IP Global Traffic Manager not only provides static load-balancing modes such as round robin, ratio, and global availability, but also gives customers the ability to route traffic on a variety of performance metrics, such as Round Trip Time, Hops, connections, dropped packets, capacity of servers, and much more.

Wide Area Persistence

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

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, the system ensures that users around the world get the information they need, in their own language.

Reduced Administration Overhead

The BIG-IP Global Traffic Manager is the only solution that provides ZoneRunner, an integrated tool for zone file management to reduce DNS risks and simplify DNS zone file management. With this tool, you can grow your DNS infrastructure in a scalable manner without the complexity and management overhead. ZoneRunner provides a secure environment to manage your DNS infrastructure and reduces administrative overhead by validating and error checking zone files. Built on the latest version of BIND, ZoneRunner provides:

- Auto population of commonly-used protocols
- Validation/Error Checking of zone file entries
- Rollback for the last transaction

- Secure environment for DNS management
- Command line version
- Importation of zones from an external server or a file
- Automatic reverse lookups
- Easy creation, editing, and searching of all records
- Reduced administration time and lower TCO
- Improved infrastructure scalability

ZoneRunner reduces risks of DNS management by providing an easy-to-use interface to manage your zone files.

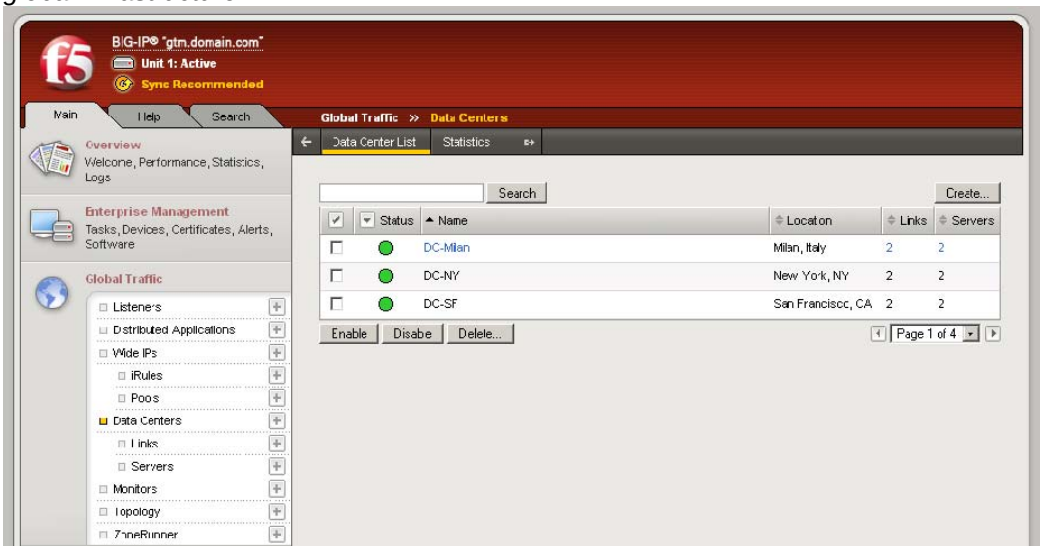
Centralized Management

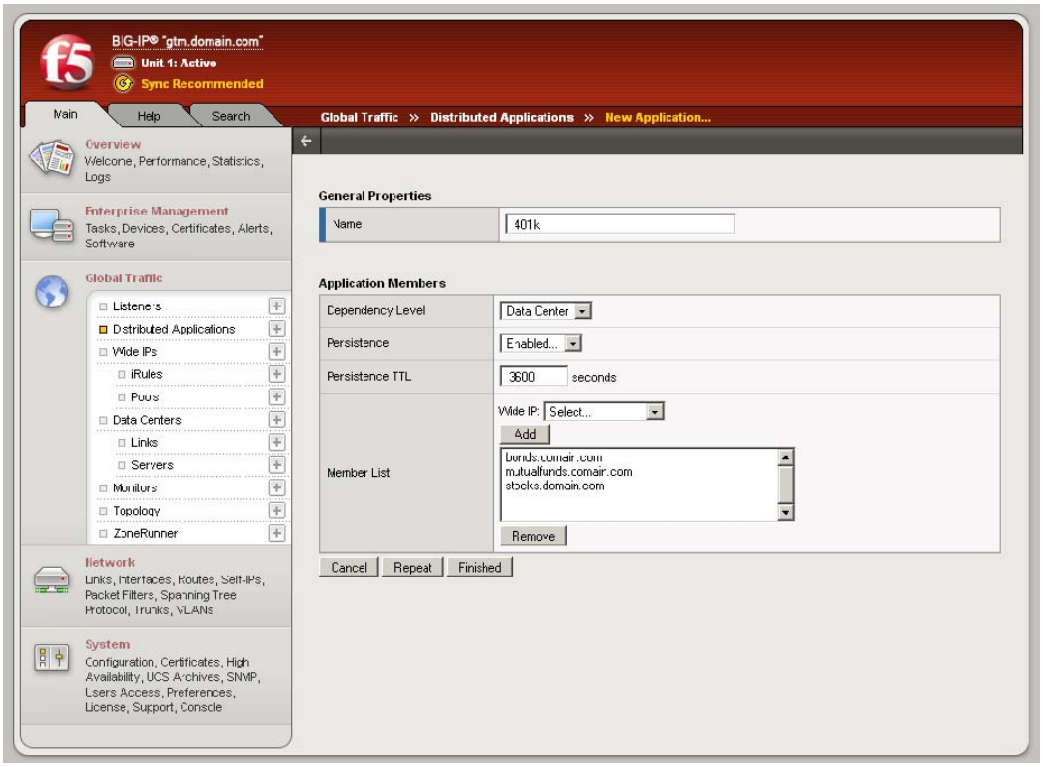
The BIG-IP Global Traffic Manager is the only vendor with a simple yet powerful way to manage your global infrastructure from a centralized location. Features include:

- Efficient list/object management
- Unique naming of global objects
- Sorting and searching capabilities
- Streamlined setup and object creation
- Context-sensitive help

The integration of the BIG-IP Global Traffic Manager with the BIG-IP Local Traffic Manager uses a single interface to manage distributed applications as part of a collective group. From this unified interface, the BIG-IP Global Traffic Manager can intelligently track and manage service dependencies in a multi-service application infrastructure. This enables organizations to easily create and manage different services and provide high availability, maintenance, and granular control with improved operational efficiency.

The BIG-IP Global Traffic Manager provides a simple yet powerful way to manage your global infrastructure.





The BIG-IP Global Traffic Manager gives you an intelligent way to manage service dependencies in a multi-service application infrastructure.

Hardened DNS Security

The BIG-IP Global Traffic Manager strengthens site security and diffuses attacks before they can start. Running on the latest version of BIND, the BIG-IP Global Traffic Manager inherits all the security protection against DNS cache poisoning that cause Pharming attacks.

By default, the BIG-IP Global Traffic Manager ships with a number of other security

features to protect against common attacks and provide additional protection for distributed applications, including:

- Packet filtering to limit or deny access to and from websites based on monitoring the traffic source, destination, or port
- A hardened device designed to resist common attacks:
    - Thwarts teardrop attacks
    - Protects itself and servers from ICMP attacks
    - Does not run SMTPd, FTPd, Telnetd, or any other attackable daemons
- ZoneRunner in protected mode to prevent any zone file tampering
- Updates via a secure (authenticated) communication channel between BIG-IP Global Traffic Manager devices to prevent dynamic DNS-based attacks
- iRules can implement DNS security policies by creating black lists of rogue sites or known sources of attacks to ignore DNS requests or responses from them
- High performance DNS tolerates high levels of DNS attacks, while delivering maximum availability for applications and services.

Conclusion

The BIG-IP Global Traffic Manager is the only application-fluent solution that solves the challenges of multi-service applications by maximizing availability, security, and performance for applications deployed across multiple data centers. Built on TMOS, the BIG-IP Global Traffic Manager evolves with an organization's changing application needs to deliver:

- Global business continuity and application availability
- Transparent delivery of applications and web services across multiple sites
- Superior application performance and client experience by directing users to the best site
- Flexibility by delivering global traffic to users according to business policies
- A holistic view into application and data center health from a single centralized location, reducing management overhead
- Efficient use of the global network by leveraging secondary data centers

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 9,000 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 http://www.f5.com.