# Achieving Top Web Site Performance

## Introduction

The demand for e-Business brings a unique set of challenges to network infrastructures, as well as to the managers in charge of them. Inadequate server capacity, missing content, fault-tolerant reliability issues, as well as the increasing size of E-business applications, continue to place added pressures on already over-burdened IT staffs.

Clearly, businesses that wish to maximize their return on investment (one that is designed to capture and service users) must provide high performance and high availability for their sites. Additionally, IT managers must have the tools to control the infrastructure behind these sites, including servers, networks, applications, and content. Only when a business has full control over all of these components can it guarantee users a predictable and reliable Internet experience.

### e-Business and Brand Equity

An e-Business site becomes not only a means to sell products, offer services or disseminate information, but it also represents an opportunity to build brand equity for the business itself. For example, with an E-Commerce company, every consumer in the world may know that you sell books on-line (thousands or even millions of dollars may have been previously spent to make this fact known). Yet if your site is difficult to access, or sends back error messages, you've lost an opportunity to build positive brand equity, to differentiate your products through a positive web experience. Within this scenario, the promotional dollars you've spent to attract consumers to your site may actually send them to your competitor's site if your infrastructure is unable to keep pace.

An easy to use, easy to access site therefore builds customer loyalty, and becomes that intangible extra, the successful 'ad' that can clearly differentiate your company and products within the mind of the user.

The keys to building loyalty for your site are deceptively simple: make sure that your site is always available, always working properly with the right content, and in the case of E-Commerce, always ready to take customer orders. The only way to make sure all of these things occur is to build the proper infrastructure behind your site to successfully handle the ever-increasing user demands.

# Internet Traffic and Content Management Products

So how can a business make sure that their site is always providing around-the-clock uptime to the user? The first key is to deploy the proper Internet Traffic and Content Management (iTCM) products within the infrastructure.  The most effective solutions combine both high availability and intelligent load balancing with connection and content management.

To distinguish, load balancing involves directing incoming connections to the optimal resource/site.  High availability guarantees that applications and sites are always available for the user.

F5 Networks' BIG-IP Controller, for example, is a combined high availability and intelligent load balancing solution for a single Internet site or data center, and provides fault-tolerance by eliminating single points of failure. By distributing customer requests across a group of servers, content and applications are always available.

## High Availability and Intelligent Load Balancing

To understand how F5 products provide both high availability and intelligent load balancing (as well as 24x7 uptime) it's important to first understand what can cause an web site to go down, and how an iTCM product can protect against these points of failure.

Problem: Server Failure – The server becomes unavailable due to a hardware or operating system failure.

Solution: An iTCM product, used with two or more servers, automatically routes to any server that fails or becomes unavailable. By proactively monitoring servers, the iTCM product keeps failures transparent to users. Then, once the server responds properly again, it is added back into the server farm automatically.

Problem: Software failure – Individual applications may hang up or stop responding, even though other applications are healthy.

Solution: The iTCM product detects the failure, and sends requests to another server that has that service running properly. The product can also determine that your applications are returning the "right content" to customers.

Problem: Content failure – The server and application may be working properly, but are responding to customer requests with a "404 Object Not Found" error message.

Solution: The iTCM product actively queries individual servers at the application level and, if an application is not returning the right content, will redirect requests to applications that are responding properly.

Problem: Network unavailable – The link between the server and the outside world is lost, making the server unreachable.

Solution: The iTCM product can provide high availability and intelligent load balancing over geographically distributed sites. This provides protection against network failures, failures related to a single data center, and Internet brownouts caused by slow interconnection points and

overloaded server farms. Customers and prospects receive consistent access to your site, without waiting.

# What To Look For In An iTCM Product

## High Availability

Through features such as EAV (Extended Application Verification), IT managers can develop external programs that verify the availability of applications. For example, managers can emulate what a customer would see and do while visiting a site. EAV proactively tests the transaction process, simulating a session that logs onto multiple accounts, places items into a shopping cart just as a customer would, verifies that the credit card authorization is working properly, and completes the on-line financial transaction.

Additionally, by using EAV, administrators can guarantee that customers will always be sent to a server that is providing the right, timely answers to their requests.

## HTTP Header Load Balancing

iTCM products must also be able to make granular and intelligent load balancing decisions. F5's BIG-IP Controller, for example, can identify specific traffic based on HTTP header information or IP address to determine where a customer request should be sent -- directing traffic to a set of servers or devices that can best service the request.  For example, BIG-IP can recognize whether a customer is "gold" (a frequent buyer) or "bronze" (only an occasional buyer). A "gold" customer's service request can be load balanced to a pool of servers reserved for similar high priority customers - ensuring that these "paying" customers receive the best service possible.

HTTP header load balancing is one of BIG-IP's numerous methods of intelligent traffic control that gives e-Businesses nearly unlimited flexibility and manageability over the flow of IP-based traffic.  HTTP Header Load Balancing not only provides URL Parsing, but also much more by examining:

- HTTP Cookie Header
- Client Source Address
- HTTP URI (or Universal Resource Indicator)

Additionally, BIG-IP recognizes and provides high availability load balancing to any HTTP header, including HTTP Version, HTTP host field (also known as URL), and the HTTP method being used in the request (get, put, etc.).

## Response Time

Obviously, the longer that customers must wait to access your site, the more impatient they become.  Response time, therefore, is a critical measure for the success of a site.  iTCM products allow site administrators to determine and set server response time thresholds.  That way, customers will never be directed to a server that is not meeting these response time thresholds, ensuring quicker access to the site.

## Persistence Features

*Persistence* is required to seamlessly process customer requests while keeping the link between your site and your customer unbroken. E-Commerce applications especially pose their own special challenges for persistence, specifically in regard to performance of the load balancing device.

That's why iTCM products must be able to provide intelligent features such as SSL Session ID Tracking to ensure customers stay connected to a single server while completing a transaction. This is invaluable for environments that receive a lot of traffic from sources such as AOL, where numerous users can be assigned the same IP address, confusing the typical load balancing product and causing traffic to be concentrated upon a single server in the array. Instead, the iTCM product your choose should read specific Session ID's from an SSL transaction, assuring the user is uniquely identified and delivered accurate and timely content until the transaction is complete.

Additionally, the iTCM product should be able to load balance both inbound and outbound traffic for devices like firewalls or routers to ensure that users return to the same device.

## Traffic Surge Protection

To prevent users or potential customers from abandoning your site, or worse, clicking to a competitor's site, sudden surges (or spikes) in traffic must not slow an e-Business site to a crawl. A successful site must be able to immediately scale service in response to these surges.

Most load balancing products, of course, provide some form of this protection. However, many are not capable of providing true high availability as well. Passive in nature, they perform no active verification of availability of servers or content on those servers. Instead, they rely on the failure of actual traffic to detect that a server was incapable of responding. Failure of traffic simply translates to a customer who is unable to reach your site.

iTCM products must also provide dynamic load balancing algorithms in addition to the static algorithms. For example, F5 Networks' BIG-IP Controller features more load balancing algorithms than any other competing product, which simply translates into the more choices for the types of traffic you need to handle.   In Observed Mode, connections go to the server with the best balance of fewest connections and fastest response time. Predictive Mode reacts by sending connections to the currently improving server with the best performance ranking (as determined by BIG-IP). In short, BIG-IP *intelligently* distributes spikes in customer demand to the most available server, giving your business the means to provide customers with the quick response they expect.

## Fault Tolerance

iTCM products must provide fault tolerance by eliminating single points of failure.  By distributing end user requests across a group of servers, content and applications are always available. For example, dual BIG-IP Controllers provide additional fault tolerance with automatic fail-over in less than a second.  BIG-IP features:

- Session State Mirroring  (ensures persistent transactions won't get dropped during failover).

- Active/Active Mode

- Active/Standby

- Network Based Failover

- Hardware Based Failover

## SSL Acceleration

The iTCM product should also include what is called SSL Acceleration, which is used to offload SSL (Secure Socket Layer) processing from servers - enhancing their performance while greatly improving response time and traffic management for a customer's transactions.  For example, SSL Acceleration improves the performance of E-Commerce servers and provides security, speed and traffic management during business-critical online transactions – all from a single location, and without the cost and hassle of installing additional hardware or software on each servers.

## Security

iTCM products should also provide a number of inherent security features designed to protect against attacks and provide added protection for your servers and network devices.

Specifically, these products should enrich network security by simultaneously working with firewalls and any other router-like network devices, such as transparent caches and proxy servers, without requiring any additional software.  Administration should be achieved with Secure Remote administration based on secure shell (SSH) for command line or SSL for browser-based management.

Ideally, the product should also be able to load balance firewalls in a 'firewall sandwich' configuration, enabling transparent device persistence.  This addresses the need of certain applications that require users to be sent through the same firewall to successfully and securely complete a transaction.

**In summary, the iTCM product you choose should have at least the following***:*

Multiple load balancing modes

Multiple modes for managing persistence

Security features, including:

-Default deny

-SSH/SSL management

-Trace routing

-IP address checking and port mapping

-NAT and SNAT

SSL Accelerator feature

Enhances speed and traffic management for secure transactions. Provides centralized management of SSL certificates supporting all commercial servers, eliminates the cost of installing and managing more software or hardware on each Web server. Enables cookie management for encrypted sessions, and supports 200 SSL connections per second or up to 17 million transactions per day.

Multiple Modes For Managing Failover

For example, BIG-IP stateful failover mirrors sessions and persistence tables, and is the only product that truly does not lose shopping carts

Proactive probing of applications to verify that they are working properly – BIG-IP is the only product on the market that really supports Layer 7 applications:

-Notifies administrator of application failure

-Marks server/device down

-Directs users to properly running servers/applications

-Proactively probes servers and verifies that the server is properly functioning

-Brings the server back into "live" server farm for user access

## Building Out The Foundation

There's a lot to consider when building the infrastructure of e-Business in order to control and service vast amounts of traffic. Four main elements are required: 1) server control, 2) network control, 3) control of applications and content, and 4) network management.

F5's BIG-IP Controller optimizes server availability and performance for a local site (server control):



There are three other critical "points of control" that businesses need to consider:  network control, control of applications and content, and control of network management.

### Network Control

F5's 3-DNS Controller is a high availability, intelligent load balancing solution for geographically distributed E-Commerce sites (network control). F5's 3-DNS Controller allows businesses to distribute Internet traffic around the world. Obvious benefits include better disaster-protection and better performance by being able to route traffic away from bottlenecked networks and sites at

geographically dispersed locations. 3-DNS allows e-Businesses to maximize network performance and give their customers a productive and rich web experience.



3-DNS routes a user's request using business-specific production rules to the "best" site, which may be anywhere on a e-Business' geographically distributed network. 3-DNS (which can be used with F5's BIG-IP® Controller or any other local area load balancer) renders the fastest DNS resolutions, resolving requests for web pages more than twice as fast as industry standard BIND v8.2.2. No other product can match 3-DNS' flexibility and level of intelligence to help e-Businesses meet their customer's expectations for a 'fast-loading' site. This translates to better quality of service, increased traffic, and ultimately higher sales opportunities.

## Control of Application and Content

F5's GLOBAL-SITE is a global data management product that manages the distribution, replication and synchronization of applications and content to geographically distributed Internet servers. Businesses that collaboratively publish content on their Internet sites can use GLOBAL-SITE to hold data in a staging repository, and then synchronize final delivery across all servers simultaneously, at one or more sites -- regardless of location. GLOBAL-SITE is smart enough to compare the new content to the prior version, and only replicates the changes – minimizing bandwidth usage. Furthermore, GLOBAL-SITE units can work in concert – a central global/SITE will post new content to remote GLOBAL-SITEs where the content is then published locally.
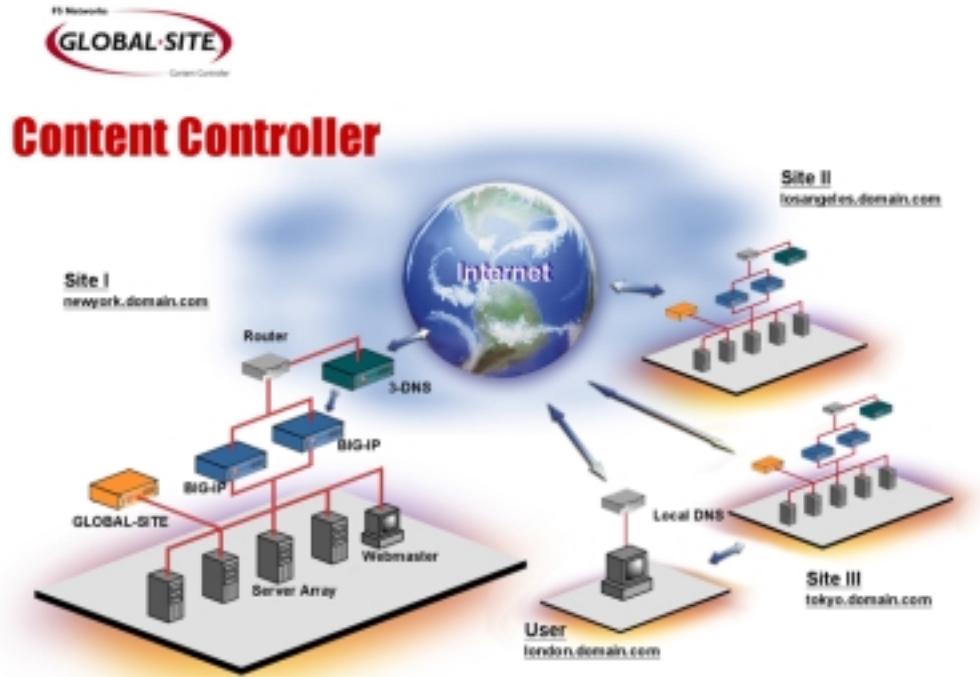
The result is a simple, efficient way to maintain dynamic content over a large set of servers. The benefits to e-Business are immediate. For example, webmasters can quickly replicate a price change on a product across all servers, automatically.  This not only saves webmasters a lot of time (by reducing the manual labor previously required) but ensures that the change is replicated everywhere – all of your customers receive the same, consistent information.

## Network Management

Network management tools can help managers spot potential problems before they occur, then visualize the future steps needed to keep e-Business sites performing as they should.

F5's SEE-IT Network Manager provides administration, monitoring and trended analysis for Internet traffic and content management based on BIG-IP and 3-DNS controller metrics. It allows businesses, in real-time, to monitor server traffic, while providing forecasting tools to assist in traffic analysis and network planning:

System administrators can perform capacity planning exercises to forecast when their infrastructure will require upgrades. The SEE-IT Network Manager also includes a system/network monitor to provide overviews on current performance, as well as an analysis package to review past operations, and help predict future functionality. For an e-Business site, SEE-IT can help administrators decide when and how to upgrade their infrastructure to meet increasing user demand, so that a potential customer need never be turned away because the site couldn't handle the traffic.

## Summary

The continued growth of e-Business applications presents many challenges to IT managers. Overcoming these challenges not only translates into increased brand equity and a more positive reflection on your business, it can help you avoid the incredible costs associated with a downed site. F5 offers total-package, end-to-end solutions that can help businesses build a solid infrastructure, maximize their return on investment, capture increasing traffic, and properly service their users – all while building positive brand equity for their company and products. In short, F5 Networks, a leader in Internet Traffic and Content Management, offers an integrated suite of high-performance products that *automatically* and *intelligently* manage user traffic and content to deliver web site integrity required for today's e-Business.