White Paper

Intelligent Traffic Management with the F5 BIG-IP Platform

Communications service providers (CSPs) are struggling with an explosion of data traffic resulting from increased mobile usage, streaming video, resource-intensive over-the-top (OTT) applications, and smart device proliferation. At the same time, CSPs must deploy new network technologies and services to stay competitive and satisfy customers. F5 delivers context-aware solutions that enable CSPs to intelligently manage their data traffic, conserve network resources, and quickly deliver profitable services.

by Andrew Hendry
Manager, Solution Marketing

Contents

The Data Boom Challenge **3**

Specific Challenges with VAS Platforms 3

Intelligent Traffic Steering to VAS Platforms **5**

The F5 BIG-IP Platform: A Path to Intelligence **6**

Consolidating Point Products to Boost ROI 8

Increasing Control with iRules and iApps 9

Improving TCP Optimization 9

Enhancing Efficiency with Dynamic Service Management 11

Implementing Context-Aware Policy Enforcement 12

Conclusion **13**

The Data Boom Challenge

With the advent of smartphones and tablets, more and more users are demanding data-intensive content over both wireline and wireless broadband connections. Although the resulting explosion of data traffic has been unprecedented, it was a well-forecasted trend for more than five years, and while CSPs may not own the content or the services directly, they do own the delivery of that content to their customers. Their efforts to cope with the data boom have highlighted several challenges that threaten their futures: an inability to scale effectively, weakening revenues and operating margins, and uneven network performance.

Despite their best efforts, CSPs have struggled to expand their network infrastructures and add new capabilities at a pace that meets customers' expectations. Wireless service providers, for example, have added significant capacity to their network facilities and systems, from the radio access network (RAN) to the packet core. At the same time, CSPs are rapidly evolving their networks into all-IP environments. New network technologies such as LTE, data center server virtualization, and policy management are also being introduced into the service architecture. CSPs are also deploying specialized value-added service (VAS) platforms in their networks. Wireless service providers use platforms that optimize mobile video content. Fixed service providers use platforms that filter and cache Internet content. And all CSPs enable personalized content services for their subscribers.

This constant evolution within the CSP network—more capacity, new technologies, service optimization, and customized subscriber policies—adds complexity and cost to traffic management. Looking to the future, as customers continue to rely upon mobile broadband for work and play, the challenges will only multiply, requiring CSPs to take a fresh, strategic approach.

Specific Challenges with VAS Platforms

CSPs may use a number of separate VAS platforms, depending on whether they are wireline or wireless providers, to enhance the customer experience and mitigate the impact of data traffic growth. These platforms typically include the following functions:

- **Video optimization:** Minimizes video content file size and bandwidth delivery over mobile networks through transcoding and transrating. Provides delivery management to optimize the user experience.

- **WAP gateways:** Connect and convert requests between WAP-based mobile devices such as feature phones and standard HTTP content on the web. Provide additional capabilities such as header enrichment and reporting.
- **Parental controls and URL filtering:** Store and apply content filtering and purchase blocker capabilities for wireline and wireless providers on a per-user-account basis, based on user-provided settings or age verification services. Provide targeted traffic steering to such systems based on the destination URL or IP address.
- **Transparent caching:** Provides subscriber content caching at the network edge, based on content type, destination URL or IP address, to improve performance and network utilization for wireline providers.
- **Ad insertion services:** Insert ads to targeted customers or for targeted services. Provide reporting and analytics for fulfillment.

Over time, CSPs have deployed VAS platforms from a variety of different vendors to meet business needs. In many cases, these VAS platforms are being used in ways that add network complexity, increase deployment and operating costs, and impede the deployment of new services. Existing layer 3 and layer 4 equipment, such as policy-based routers (PBRs), routes all data traffic traversing the network to these VAS platforms without regard to relevance. As a result, all VAS platforms must inspect that traffic, determine whether to apply specific traffic policies or take action, and balance the traffic load across multiple server platforms. This is a duplication of functionality that resides across the VAS platforms. Not only does every VAS platform have to develop and maintain these functional requirements, which sacrifices development of more relevant features and adds cost, but every VAS platform must now scale to process all traffic rather than all relevant traffic.

This architectural weakness may not have been a huge consideration in the past, but with the tremendous growth in data traffic, CSPs are beginning to struggle with scalability and the costs to maintain these platforms. The situation gives CSPs a huge opportunity to optimize the overall architecture of the VAS platforms by centralizing these critical traffic management functions on a dedicated platform. Those who do so will achieve higher performance, reduce costs, gain flexibility, and improve service availability.

Intelligent Traffic Steering to VAS Platforms

Traffic steering is the ability to direct user requests to the correct and appropriate content source or VAS platforms based upon pre-established network policies. That content source may be branded CSP services like a web portal or commerce site, OTT content from outside the CSP network, or VAS platforms that optimize services. The policies for how and where to direct users can be based on many parameters, such as device type, access network, type and status of customer account, user location, roaming status, and availability of VAS systems.

CSPs use disparate suppliers to support their VAS platforms. Traffic management capabilities are often components of any given VAS solution to balance the load among servers. While some suppliers bundle a partner solution, others may offer a limited subset of traffic management capabilities themselves. Traffic steering can, however, become the strategic point of control for connecting users to desired services in a way that optimizes both customer experiences and network resources.

CSPs that provide for a centralized traffic management function in front of the VAS platforms as part of their service architecture can realize a number of advantages. A traffic steering solution at a central network node can simplify traffic management and control compared to the same functionality distributed throughout the VAS platforms. With centralized steering, only relevant traffic is passed to individual VAS platforms, which reduces the need to scale those systems and significantly drives down total costs. For example, rather than allowing a PBR to pass all port 80 traffic to a mobile video optimization platform, an intelligent traffic steering platform can identify any video traffic and steer only that traffic to the optimization platform. Similarly, traffic can be steered to transparent caching platforms in wireline networks by content type or unique resource identifier (URI).

This model not only eliminates redundant functionality on VAS platforms but also can reduce by 50 to 75 percent the amount of traffic those platforms must process. Traffic can be identified by each individual TCP connection and HTTP request to be steered to the relevant VAS platforms. This enables the chaining of multiple VAS platforms within a single session to deliver even more efficiency. The CSP also gains operational flexibility to add or swap out VAS platforms without compromising service availability or functionality, because the centralized traffic steering function

can be set up to balance the load away from any affected systems during a transition.

In addition, for most VAS platforms, any traffic management capability is secondary to the core mission. With a central traffic steering function in place, CSPs can take advantage of a best-in-class system that delivers the highest performance and scalability and the most innovative feature set. For example, once they begin to virtualize these VAS platforms into application clouds, this centralized function can provide dynamic service bursting to adjust platform capacity on demand. Finally, after such a system is in place, CSPs can phase in incremental service functions such as large-scale network address translation (NAT), IPv6 gateways, and security firewall capabilities rather than having to deploy yet more single-solution products.

The F5 BIG-IP Platform: A Path to Intelligence

The F5® BIG-IP® platform offers intelligent traffic management solutions that help CSPs solve challenges resulting from traffic growth and complex network evolutions. The platform, featuring a dedicated, carrier-grade controller on NEBS-compliant hardware, establishes a centralized traffic management function in front of VAS and other components of the architecture. Unparalleled system capacity and scalability in throughput, simultaneous connections, and transactions per second enable BIG-IP controllers to offload processing-intensive layer 4 through layer 7 functionalities from legacy platforms such as Gateway GPRS Support Node (GGSN), packet gateways, and network firewalls.

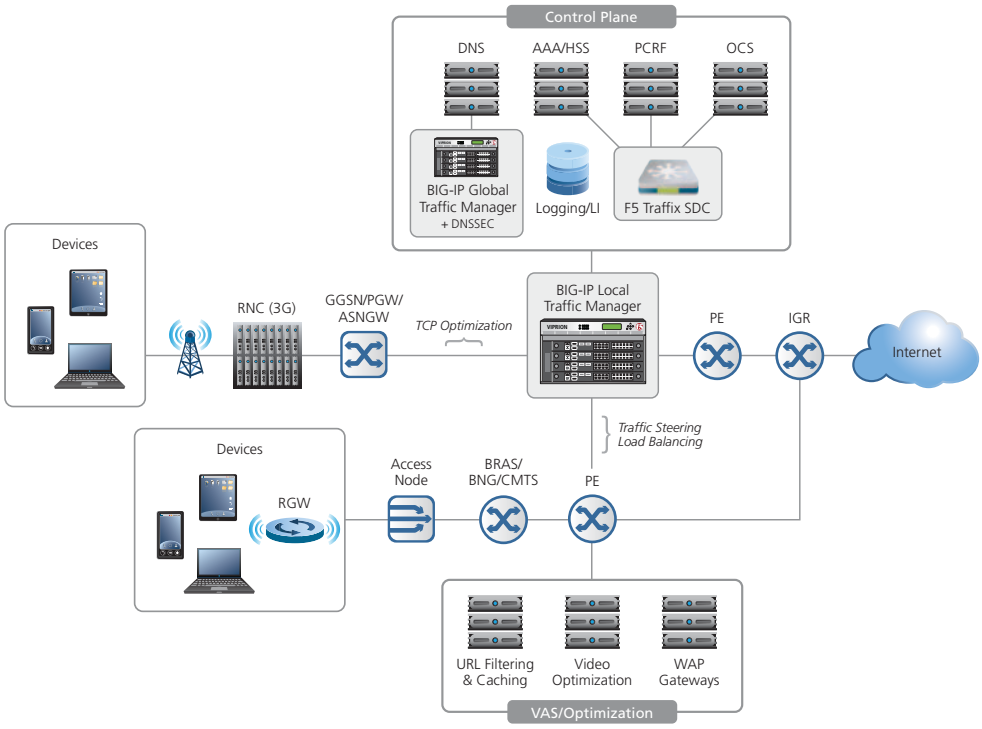

Figure 1: Intelligent traffic management with F5

From this strategic point of control, BIG-IP products provide:

- Sophisticated, subscriber-aware and context-aware traffic management with session-based data inspection in real time—providing critical traffic visibility, control, and analytics based on parameters such as subscriber profile, device or content type, location, and network conditions.
- The ability to direct targeted content to VAS platforms or to content sources based on subscriber or contextual data.
- Support across heterogeneous systems for balancing traffic loads, monitoring system health, and steering traffic based on availability to improve service reliability under the most demanding conditions.
- A full proxy architecture and any-service-on-any-blade IP capabilities to read and modify all traffic in any direction, prioritize critical traffic during bursts, and filter out unwanted traffic.
- HTTP header enrichment to provide intelligence to content providers and other third-party providers.

The BIG-IP platform helps CSPs scale capacity and offer superior end-to-end performance. A unique, patent-pending TCP optimization methodology, F5® TCP Express™, improves application performance. Customizable configuration with the F5® iRules® scripting language, which enables rapid policy creation and automation, and with F5® iApps® templates, which help CSPs provision, analyze, and control application services across the network.

F5 simplifies the network architecture and provides the performance, reliability, and flexibility CSPs need to manage network resources efficiently and cost-effectively, deploy new services quickly, and manage changes more gracefully.

Managing the Explosion of Data

Voluminous data can be managed to take back control of the network, as shown in the F5 Service Provider Series video, Traffic Steering.

Consolidating Point Products to Boost ROI

To respond to the growth and innovation in data networks, CSPs have expanded existing legacy platforms and added new ones without a holistic view of the network architecture. In many cases, this has resulted in needlessly complex networks that cannot readily be scaled, increase deployment and operating costs, and reduce the ability to add or adapt new services.

Once established in a position to steer traffic to VAS platforms, the BIG-IP platform enables CSPs to consolidate several incremental network functions to increase network efficiency and ROI. F5 products provide a number of additional service functions, including security, translation, processing offloading, optimization, and policy enforcement.

BIG-IP Product Function	Description
Gi and Data Center Firewall	Protect the entire CSP infrastructure with an ICSA-certified firewall
CGNAT/IPv6 Solutions	Mitigate IPv4 address depletion, support both IPv4 and IPv6 simultaneously, and enable migration to all-IPv6 networks
WAP Offload	Provide WAP 2.0 offloading support to relieve VAS platforms from scalability issues
Deep Packet Inspection (DPI)/ Policy Enforcement	Provide intelligence and policy control with session-based packet inspection at wire speed

Figure 2: Service functions that can be consolidated with F5 products

By consolidating multiple solutions within the same high-performance platform, the BIG-IP product family establishes efficient, strategic points of control across the network. This reduces deployment and operating costs, creates a simpler, more easily managed network, and improves application performance by reducing the

number of hops in the end-to-end network. CSPs also gain flexibility to deal with changing technologies and network conditions and can speed deployment of new services.

Increasing Control with iRules and iApps

The BIG-IP platform provides two valuable tools to improve flexibility and control: iRules and iApps. The iRules scripting language, a powerful and flexible feature of BIG-IP systems, provides CSPs with a customizable configuration toolkit based on the F5® TMOS® architecture. iRules provides unprecedented control to enable direct manipulation and management of any IP application traffic. With its easy-to-learn scripting syntax, iRules enables CSPs to customize how they intercept, inspect, transform, and direct inbound or outbound application traffic.

iRules can be applied to any IP application or protocol, enabling new degrees of application optimization and security. Furthermore, iRules can be invoked and manipulated via the F5® iControl® API, allowing the network to modify web services in ways that previously required changes in the applications.

iRules helps CSPs to rapidly build and automate network management policies. With iRules, network administrators can read and modify all traffic in any direction, prioritize critical traffic during traffic bursts, and filter out unwanted traffic, all without impacting real-time services. This is especially useful to solve unexpected issues, meet unanticipated future requirements, and deal with changing network conditions without a typical platform software development and upgrade cycle.

iApps templates provide a new way to architect and provision application delivery. iApps can unify, simplify, and control an entire application delivery network, providing a contextual view and advanced statistics about the application services supporting the CSP's business. An application-centric view means deploying application services that reside in the network—such as authentication, data protection, traffic management, and acceleration—and aligning them to the applications for which they're being used.

Improving TCP Optimization

Inherent TCP/IP inefficiencies, coupled with access network congestion that lead to latency and packet loss, adversely affect application performance for subscribers. These inefficiencies increase response times for applications and significantly reduce bandwidth efficiency—in other words, the ability to "fill the pipe."

The BIG-IP platform offers a native TCP/IP stack called F5® TCP Express™ with numerous optimizations and RFC extensions. These proprietary optimizations go beyond standard TCP enhancements to dramatically improve application performance across real-world subscriber access networks. The improvements are not seen in typical packet-blasting test harnesses; rather, they are designed to respond to real-world client and Internet conditions.

While regular server TCP stacks are optimized for high bandwidth and low latency conditions, the 2G and 3G networks of CSPs feature exactly the opposite. TCP Express eliminates the need for clients and servers to negotiate the lowest common denominator for communications in congested networks. Client and server connections are isolated, controlled, and independently optimized to provide the best performance for every connecting device. The BIG-IP system intermediates on behalf of the client via a capability called stack brokering, using TCP Express to optimize client-side delivery while maintaining server-optimized connections inside the network.

Independent testing tools and customer experiences have shown TCP Express to deliver up to double the performance gain for users and quadruple the improvement in bandwidth efficiency with no change to application servers or client devices. The combination of a full proxy TMOS architecture and TCP Express works to dramatically improve performance for all TCP-based applications.

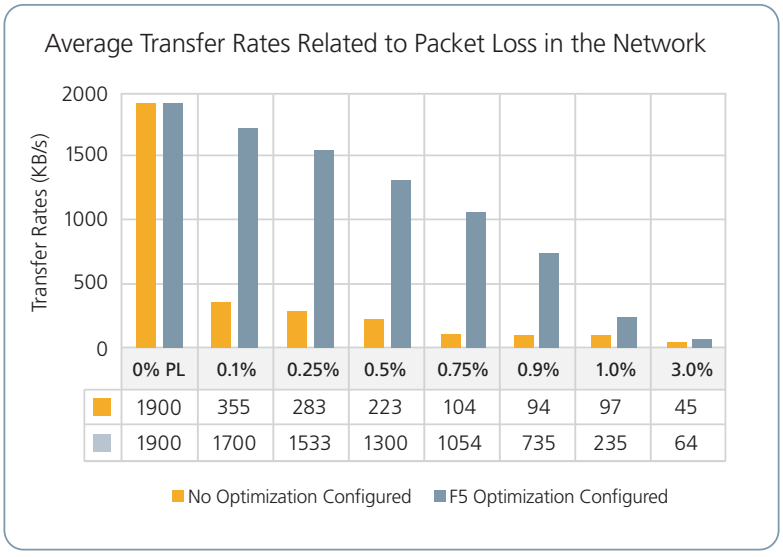

Figure 3: Improving the subscriber experience with TCP optimization from F5

While TCP Express is automatic and requires no modifications, the BIG-IP system gives administrators advanced control of the TCP stack to tune TCP communications according to specific business needs. This includes the ability to select optimizations and settings at the virtual server level per application. Administrators can use a TCP profile to tune a number of TCP variables, including:

- TIME/WAIT recycle.
- Delayed ACKs.
- Deferred accept.
- Selective ACKs.
- Limited transmit recovery.
- Bandwidth delay.
- Proxy buffer.

These TCP optimizations are easy to activate and configure in the BIG-IP management console, thus reducing the chances of operator error.

Enhancing Efficiency with Dynamic Service Management

As cloud-based deployments gain traction, CSPs are more closely evaluating a new capability called dynamic service management that can improve the efficiency of VAS platforms. Dynamic service management monitors VAS platforms and client requests in real time to determine the load, automatically coordinate the addition or removal of virtual system resources, and provide integrated load balancing to VAS/optimization platforms.

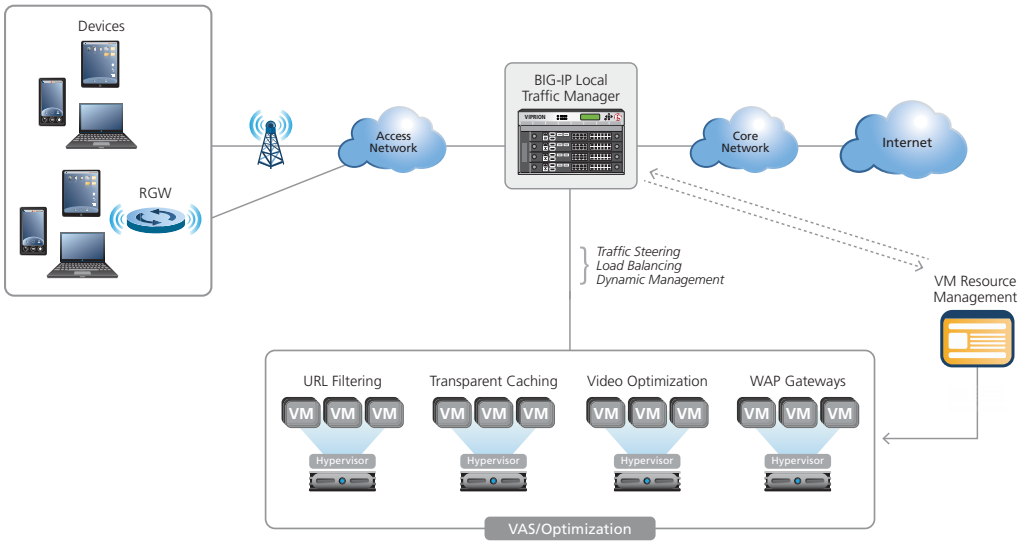

Figure 4: Dynamic service management from F5 products improves the efficiency of cloud-based deployments.

Dynamic service management:

- Shares the VAS infrastructure to reduce deployment and operations costs.
- Simplifies the services architecture.
- Enables real-time adaptation to changing subscriber and network conditions.
- Reduces the time to deploy new services in the network.

This valuable capability can be incorporated into the BIG-IP platform as an incremental or follow-on phase of deployment.

Implementing Context-Aware Policy Enforcement

Recently CSPs have begun to explore innovative new business models as a way to improve customer experiences, drive new revenue streams, and use their networks more efficiently. These offerings utilize customer profiles and rate plans, real-time network conditions, and content type to provide differentiated services and manage traffic through the network. For example, CSPs have migrated from the unlimited data plans of the past few years to the following:

- Tiered data plans based on GB used.
- Shared data plans based on GB shared among plan members.
- Social media data plans that zero-rate specific content (e.g., a "Facebook plan").

In these scenarios, customer traffic is allowed, denied, zero-rated, or rate limited based upon the customer's subscription plan. CSPs can build on these plans by incorporating new policies to prioritize or throttle subscriber bandwidth based on plan thresholds, network conditions, time of day, or session-based, à la carte charges.

To succeed in these new offerings, CSPs need to better understand their data traffic and integrate new policy management capabilities into the network. While many CSPs have plans to add policies, their existing equipment platforms often cannot be scaled to handle policy enforcement or would be too expensive to support if they were. Moreover, CSPs need to increase IP traffic visibility up to layer 7 as well as traffic analytics to understand network usage, all without affecting end-to-end performance.

The BIG-IP platform provides context-aware policy enforcement capabilities in addition to its sophisticated traffic management capabilities to meet these CSP needs. The BIG-IP platform works with policy and charging rules function (PCRF) products from multiple vendors to use operator-driven policies for intelligent traffic steering and shaping in real-time on a per-subscriber basis. Because the BIG-IP platform provides unprecedented capacity and scalability—in throughput, concurrent connections, and transactions per second—CSPs can have confidence that policy enforcement will not choke the network or be artificially limited by the threat of system overload.

Another current constraint in policy enforcement is the significant amount of subscriber and network data that is orphaned in disparate network elements and thus untapped to add value. Instead, CSPs must be able to use the intelligence throughout their control planes—from subscriber location to profile parameters and network conditions—to optimize the data plane traffic flow. To achieve this, the BIG-IP platform offers subscriber- and context-awareness derived from session-based data inspection as well as relation of the AAA and DNS systems.

Conclusion

CSPs today are using both intelligent traffic management solutions and optimization platforms to manage the explosion of data traffic driven largely by mobile video traffic. F5 solutions enable CSPs to manage data traffic, optimize network resources—including VAS platforms—and deploy strategic points of control to understand and profit from usage growth with innovative pricing plans and new services. CSPs can also increase efficiency, performance, and scalability by centralizing their intelligent

traffic management function with F5's highly flexible and context-aware solutions that span multiple VAS platforms. The BIG-IP product family empowers CSPs to take advantage of dynamic service management and new policy enforcement capabilities to enhance the effectiveness and profitability of their networks.

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

©2012 F5 Networks, Inc. All rights reserved. F5, F5 Networks, the F5 logo, and IT agility. Your way., are trademarks of F5 Networks, Inc. in the U.S. and in certain other countries. Other F5 trademarks are identified at f5.com. Any other products, services, or company names referenced herein may be trademarks of their respective owners with no endorsement or affiliation, express or implied, claimed by F5. CS01-00116 0912