F5 White Paper

Service Delivery Networking

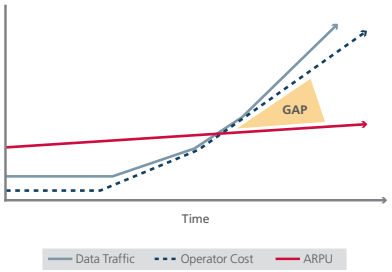
Service providers must constantly balance the cost of providing differentiated services, maintaining average revenue per user (ARPU), and the demands of subscribers for a higher quality of experience, even in the face of explosive growth and demand. Service Delivery Networking presents a new way to address many of these unique challenges.

by KJ (Ken) Salchow, Jr.
Manager, Technical Marketing

Contents

Introduction

The changes in the demands on service provider networks over the past several years have been drastic and rapid. The explosive growth of mobile IP devices and data usage, not to mention the rapid adoption and penetration of wireless voice, has put a perpetual strain on service providers who must constantly balance the cost of providing differentiated service, maintaining average revenue per user (ARPU), and the demands of subscribers to provide a higher quality of experience. The highly competitive nature of the business doesn't allow for a single failure.



The gap between increasing operator costs and ARPU.

The opportunities for misstep are numerous, particularly in the mobile realm where the constant battles between bandwidth and performance are directly at odds with the technology required. But even in the wireline business there are multiple points within service delivery where operators want more control and capability to appropriately manage bandwidth in relation to the services being offered and the subscribers using them (contextual policy and charging are among their top needs). More immediately, the promise of cost savings for a converged IP network delivering voice, data, and video via a consolidated and integrated architecture is not being realized due to the complexity and demands of the various services (especially when compared to the intelligence of packet-based equipment). Because of this, service providers often deploy similar functionality—such as NAT, IPv6, DNS, AAA, optimization, traffic steering, and load balancing—separately for each service offering.

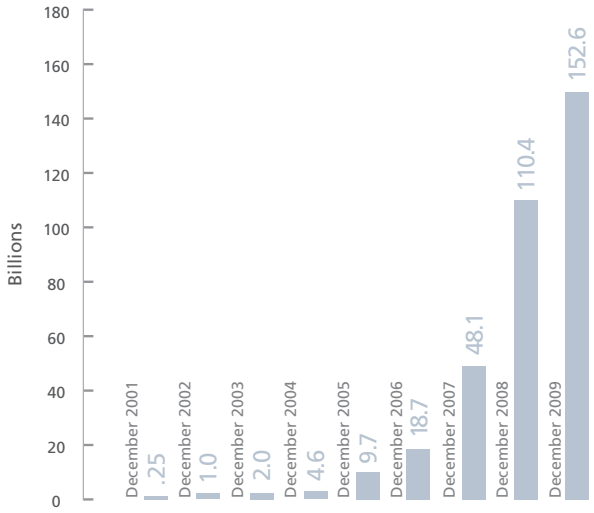
Even as service provider networks evolve to address these demands, they face challenges. The most pressing of these challenges is the appropriate way to migrate users from the mixed circuit-switched/packet-switched networks to the next-generation IP core and access networks that promise greater capacity, reduced OpEx, and faster delivery of differentiated, value-added services. Yet

as service providers struggle with converged IP networks, the IP network itself is undergoing a change from IPv4 to IPv6, presenting another complex issue of backwards compatibility versus the many benefits of the new protocol. At the same time, there is no magical switch that can be flipped, giving all subscribers and services the ability to run on that IP core or to utilize a full IPv6 core. The act of achieving simplification generates complexity.

The good news is that the move to an all IP core has many significant benefits. While service providers have unique needs and challenges, there is an established market of IP-based solutions. The F5 Service Delivery Network architecture, augmented with carrier-grade, NEBS-compliant platforms, presents an exciting alternative to not only assist service providers with today's challenges, but also to build an architecture that is ready for tomorrow.

The Explosive Growth of Mobile Usage

According to IDC[1], U.S. Internet traffic more than doubled from an estimated 8.4 petabytes (PB) per day in 2006, to more than 18PB per day in 2009, estimating another doubling to 36PB per day by 2013 (an estimated 13.1 exabytes per year) with an average growth rate of roughly 50 percent year-over-year. While still only a small percentage of the volume, mobile Internet usage is rising much faster, estimated to increase 39 times current rates by 2014[2]. While that seems amazing in itself, it does not begin to truly quantify the challenges of the service provider as it only measures the amount of data that forays onto the Internet. For instance, according to CTIA[3], wireless providers reported handling 1.5 trillion SMS messages in 2009 for the U.S. alone; with flat sequential growth (11.1%, the lowest in half a decade) this will approach 2 trillion messages in 2010. This, when compared to the reported 80 billion sent in 2005, is an 1830% increase in 5 years. This also doesn't account for the reported 34.5 billion SMS messages in 2009, a number that grew over 160% year-over-year. The growth of the number of messages is only half the story; MMS messages are now being pushed to larger sizes and migrating entirely to the IP network; and mobile devices are capable of taking >5MP pictures and full HD video, increasing the size of the average MMS message as well. This is all on top of the web browsing, Facebook updates, Twitter, Flickr, and YouTube traffic reported in mobile Internet usage—a combined 107.8 billion MB, or 107.8 PB of data in the last 6 months of 2009 in the U.S. alone.



Monthly Text/SMS Traffic Volumes: December 2001–December 2009[4]

This constant requirement for growth, and the associated investment to continue to meet user needs, flies in the face of ARPU[5], which has dropped from $71.75 in the first half of 1993 to $47.74 for the last half of 2009. Service providers continually have to increase the capacity and reliability of their system while simultaneously deriving less and less revenue per user in this highly competitive business. Finding a way to manage growth while simultaneously extracting efficiencies is a key driver behind the development of new network designs and best practices.

Next Generation IP Core and Access Networks

One of the goals behind IP Multimedia Subsystem (IMS) design was the notion of consolidation and integration; components, like subscriber authentication, authorization and accounting (AAA), could be used for multiple services instead of the traditional model that included custom, unique AAA components for every service offering. Deploying services once and leveraging them across multiple offerings is a key success factor in reducing the cost and complexity of the service provider network. Moving to an IP core has certainly enabled a significant amount of consolidation and corresponding reductions in CapEx and OpEx, but not nearly as much as many had hoped.

While new services and offerings can be deployed using the new IP core networks, legacy user equipment on legacy architectures that use combined packet-switched and circuit-switched technology cannot simply take advantage of these offerings. Instead, it requires the continued maintenance of the circuit switched infrastructure to handle the non-integrated voice traffic. The result is that while legacy IP-based data services can be moved into the IP core, many legacy systems remain necessary for the foreseeable future. Because of this, service providers are faced with only partial consolidation.

An additional challenge is the lack of visibility into the services that are carried over the IP network. Typical packet-based solutions, even those that are considered deep packet inspection (DPI) devices, are limited in their ability to differentiate one service from another. While they can use standardized packet header information, that doesn't provide visibility into the actual service data being transmitted. The standardized nature of this inspection renders it incapable of understanding non-standard headers that might be used by service providers to identify end-to-end quality of service (QoS) requirements. These devices also generally lack the capability of associating different services together as a single "user session." Because of this, enforcing bandwidth and QoS restrictions across services for the same user is extremely difficult and complex. As voice becomes integrated as an IP service as well, this inability to differentiate correctly and the need to maintain separate flows becomes a serious drawback. Consequently, many service provider networks, while talking about consolidation and integration, still tend to create unique flows for voice, video, and data simply because they cannot provide the same control over an integrated, single packet flow solution.

Lastly, the explosive growth of mobile computing requires service providers to begin implementing IPv6. IPv6 provides significantly more address space to handle the massive number of devices and includes several other advances over IPv4. However, this implementation also presents significant challenges. First, even if the fact that most subscriber content still runs on the predominantly IPv4 Internet, many of the user devices, packet-switch network devices, and applications are not IPv6-enabled. This produces yet another set of infrastructure and service offerings that must be created and maintained in parallel with existing systems: one for IPv4 systems, one for IPv6 systems, and potentially another complete infrastructure to provide interoperability between them.

This overlay of new and old adds a complexity to the service provider network that marginalizes the benefits of, and the ability to implement a truly integrated IMS infrastructure; thus, it continues to erode existing profitability and makes it difficult to capitalize on the new revenue sources and services that could positively impact ARPU.

Managing and Measuring Bandwidth

While there has been plenty of negative press concerning the ideas of net-neutrality and the efforts of service providers to provision, manage, and ration services and bandwidth, the reality is that unless service providers manage to control the explosive growth then no one will be able to access any of these services with any reliability. Bandwidth, like any resource, is finite and must be managed appropriately. This is a difficult concept for subscribers to understand, especially in developed markets where access to basic utilities are often viewed as limitless. Like any utility, the only real mechanism service providers have is to charge based on use; the more you use, the more you pay. Making this even more complicated is the fact that, depending on where in the network you look, service providers aren't always able to enforce the elements of control they desire. Between user equipment, finite radio spectrum, backhaul connections, and the content being requested (much of which is provided by external parties on public networks), service providers have limited control over end-to-end performance, absolute aggregated bandwidth needs, and availability of services. Service providers need to be able to measure, monitor, and manage bandwidth by service and user regardless of whether they own the service, host the service, or simply provide a conduit to the service.

However, because many service provider networks maintain segmented traffic—those for voice, video, and data—it can be very difficult to enforce anything other than very high-level bandwidth management, such as ensuring that a certain percent of bandwidth is reserved for voice traffic, video traffic, and generic data access. Since bandwidth—and more specifically combined bandwidth—is a finite quantity, service providers increasingly want to segment the individual services by user and service level. They need the ability to manage bandwidth for individual users and provide differentiated services for those willing to pay more. Instead of simply being able to put limits on total consumption of voice, video, and data, service providers want the ability to create differentiated levels of service in real time. Even heavy data users might have some data traffic that is more important to them than other data traffic and would be inclined to pay for guaranteed delivery. For instance, many users would be willing to pay for a premium email service or online gaming service that guarantees specific levels of bandwidth or a premium video service that guarantees high-quality, real-time streaming.

Implementing opt-in premium services requires greater capability to monitor, manage, and steer traffic not just based on service type, but also on subscriber, device, network capacity, and other contextual information. It also requires the ability to adequately report usage for appropriate billing of those services and enforcement of specialized service policy.

The Ideal Solution

What service providers need is a common IP architecture of reusable services that can be applied regardless of traffic type, whether voice, data, or video. They require a set of services—like large scale NAT, DNS, AAA brokering and scale, optimization, and security—that become part of the converged IP architecture maintaining the context of the session, providing the necessary visibility, and easily acting on the traffic regardless of the type. The requirements of such a system must include deep session intelligence; service integration management; programmatic system interfaces that provide links into back-end systems for subscriber policy information and billing; and adaptability to solve future challenges and overcome incompatibility of divergent systems.

Traditional solutions force a choice between intelligence and adaptability.

Deep Session Intelligence

Deep session intelligence, or layer 7 intelligence, goes beyond traditional IP layer 2-4 packet switching, providing insight into the actual services being requested and delivered. This intelligence gives the service provider the ability to understand the context of services requests—what type of service is being requested, what device is requesting the service, which subscriber owns the device, and what policies are in place for that service. It's absolutely critical to include these contextual elements when developing new, differentiated service offerings or when implementing policy enforcement and payment. In addition, deep session intelligence brings new capabilities to the service provider network by enabling a two-sided business model.

On one side, subscribers might opt in for premium services (HD video, segmented data rates, or premium online gaming) if the service provider can accurately segment traffic based on its type, specific service, and the subscriber requesting the service. Deep session intelligence provides insight into the traffic, device type, and the subscriber and provides the ability to make intelligent, policy-based decisions about how to handle the traffic. Generic, non-differentiated traffic might be passed on as best-effort. Premium, differentiated traffic can be passed to value-added services even if, on the wire or in the air, the only difference between these requests is the subscriber making them. This provides unique, new services with ease, lowering costs and simultaneously increasing ARPU.

On the other side is the ability to use deep session intelligence to deliver targeted advertising or premium content based on the user profile, device type, and current geolocation of the subscriber. Service providers can then offer advertisers targeted campaigns based on the location of the subscriber, the capabilities of the device, the service or content preferences of the subscriber, or even targeted device upgrade offers that are uniquely qualified to the subscriber. These premium advertising options give service providers another unique way to provide differentiated offerings and, potentially, new revenue streams to mitigate the decline in average local monthly bill per subscriber (ALMB) brought on by competitive pressures.

Integrated Services Management

The proliferation of devices within the service provider network is due in part to the lack of deep session intelligence to understand and differentiate the various types of traffic and how to manage it. It is also the result of point-solutions designed for a single purpose and focused on a specific traffic flow. Integration of services

management, in combination with session intelligence, provides a simplified, more easily managed solution that ultimately improves overall performance, while reducing CapEx and OpEx. With integrated services management, user sessions can be managed together, in a unified manner, instead of as disparate individual requests across different services.

Incorporating services like bi-directional IPv6 support, large scale NAT, DPI (for traffic management), and AAA services along with caching, compression, and DNS enables the service provider to apply the most appropriate service to each unique traffic flow. Regardless of traffic type (voice, video, or data) the source of that traffic (IPv4 or IPv6) or the destination (internal services, hosted services, or external services) each traffic flow is handled by a single, simple process that applies the most appropriate services while enforcing policy and reporting charges. The integration of multiple services enables a single traffic flow without sacrificing the ability to manage traffic, provide differentiated services, or lose backwards compatibility for legacy user devices. This simplification makes the transition of subscribers and service from legacy data paths to the next generation network a seamless and nearly transparent operation.

In addition, giving service providers a centralized strategic point of control enables the integration of highly specialized services such as premium video transformation or compression by simply using deep session intelligence to identify the correct content and/or subscriber and steering that traffic to an appropriate, available resource using load scaling technologies. Not only does this simplify the interface for adding new services and features because of the unified traffic flow, but it enables the deployment of new services that—depending on user equipment capability—could be deployed to subscribers that were not originally targeted. For instance, while a new premium video service might be provisioned for the latest and greatest handsets, it might be found that a particular model of older handset also displays the video appropriately. Because of the unified traffic flow, service providers can expand the possible subscriber base for this service. Integrated, contextual services allow the application of basic, but critical, traffic management services across the entire IP network. This helps offload redundant and non-value-add functions from the services themselves and provides a strategic point of control for simplified service provisioning.

Programmable Interface

Even with a consolidated solution where subscriber traffic follows a singular path regardless of service type, it is challenging to continually adapt to meet the needs of the business. While reduced complexity can significantly increase the reliability

and serviceability of the network, the service provider network still spans massive geographies and incorporates thousands of access points, localized points-of-presence, backhauls, and all the associated devices between the subscribers, the services, and content they desire. This represents a significant challenge for even the most nimble organization.

Defining differentiated services based on contextual information might not be all that difficult. On the other hand, implementing the policies and creating the connections to charge appropriately for those services can be extremely difficult if service providers have to touch every device responsible for policy enforcement and control. Service providers need to be able to deploy policies simultaneously across the network while also updating billing information related to the changes in policy. The only reasonable way to achieve this is through the use of programmatic interfaces similar to those recommended by the 3rd Generation Partnership Project (3GPP), a collaborative association that unites various telecommunications standards bodies.

By implementing a programmatic interface, service delivery can automatically and dynamically react to changes in traffic and implement policy based on content, device type, or location. This is the final step to using deep session intelligence and integrated services to change the way traffic is managed across the entire system. With a simplified traffic flow, implementing changes in policy and charging via a programmatic interface can make differentiated, user-unique rate plans and service offerings a reality.

Adaptable by Nature

Having deep session intelligence and integrated services is a great start, but without the ability to adapt to changes, the service provider network could soon become as outdated and complex as today's existing infrastructure. Because the only constant in the service provider network is change, the network solution needs to be built with the ability to adapt to change.

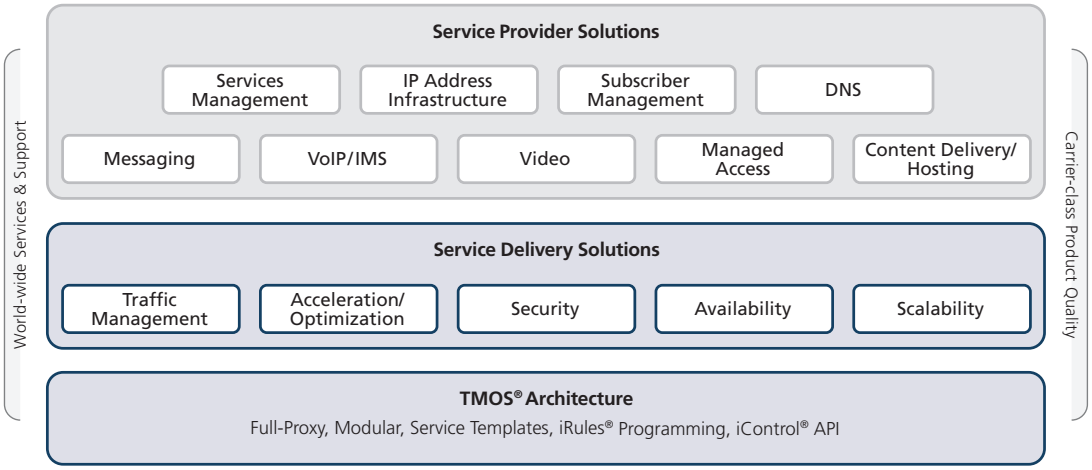
Despite standards and proposals, not all solutions deployed within the service provider network are inherently capable of working together. Differences in protocols (IPv4 versus IPv6), differences in implementation (custom QoS tags that are different from one device to another), or unexpected issues can all cause complications when deploying new services. By providing a central, programmable, bi-directional point of control, service providers can create a service bridge to ensure interoperability between different protocols, or even different implementations of the same protocol. Service providers can then ensure that the diverse services in place can function as a

single, unified architecture, and that new services can be added to the existing system seamlessly, without affecting the existing services.

As service providers undertake the migration to an all IP network, it is critical to have the ability to translate from one protocol to another, allowing the continued provisioning of legacy services while simultaneously moving to the new, integrated core network.

The F5 Service Delivery Network

F5 Networks® has been the leader in the creation of application delivery solutions for enterprise organizations who rely on IP-based business solutions since the very first, hardware-based load balancer was introduced. Over the past decade, F5 solutions have evolved and changed to meet the increasingly complex nature of the Internet and application delivery. F5 also has a longstanding relationship with the service provider market, and has tailored IT solutions to meet the more demanding environments of the service provider network. This has resulted in the creation of carrier-grade, NEBS compliant hardware platforms with the scale, capacity, and reliability that service providers require. It has also spawned Service Delivery Networking: a dedicated effort to deliver an integrated approach to delivering services.



F5 Intelligent Service Delivery Network

To enable Service Delivery Networking, F5 offers a portfolio of products that perform common, reusable services for all traffic types—such as NAT, IPv6, firewalling, traffic steering, DNS, load balancing, AAA brokering, and scaling,

session, and service awareness for policy enforcement, optimization, and security. To make this simple to deploy, manage, and maintain, each solution has a common IP architecture based on a very fast, low latency, full IP proxy. The F5 TMOS® modular product architecture makes this happen by enabling specific services on an as-needed basis depending on the current requirements of the traffic and/or policy.

F5 provides a suite of purpose-built, carrier-grade service delivery solutions that provide the common reusable services, layer 7 intelligence, integration, programmability, and adaptability required to build a Service Delivery Network.

- Services Management:
  F5 BIG-IP® devices serve as a strategic point of control in the network, combining scalability and high availability with intelligent and versatile policy-based control over all traffic, giving carriers and service providers the ability to satisfy changing market needs profitably.
- Subscriber Management
  BIG-IP devices combine scalability and high availability with full proxy capabilities to assure interoperability between AAA services, giving carriers and service providers the ability to grow the subscriber base, while adding new, innovative services.
- Managing Large Scale, Carrier-Grade NAT
  BIG-IP products have been managing IP Network Address Translation (NAT) for years across multiple heterogeneous networking environments. Only recently has the need for core IP NAT ballooned to a scale that can only be handled by a device built to manage millions and millions of concurrent IP connections.
- Managing IPv6
  BIG-IP IPv6 functionality enables service providers to deliver applications across and manage both an IPv4 and IPv6 infrastructure simultaneously. BIG-IP devices can help lower costs of managing both networks and can help expedite the full migration from IPv4 to IPv6 for both internal services as well as external users.
- Scaling DNS Services
  BIG-IP domain name services are more than large scale distributed global DNS products; they provide a complete end-to-end DNS solution. BIG-IP products provide a more intelligent way to respond to DNS queries based on a variety of network conditions and situations, distributing user application requests and application services based on business policies, data center conditions, network conditions, and application performance.

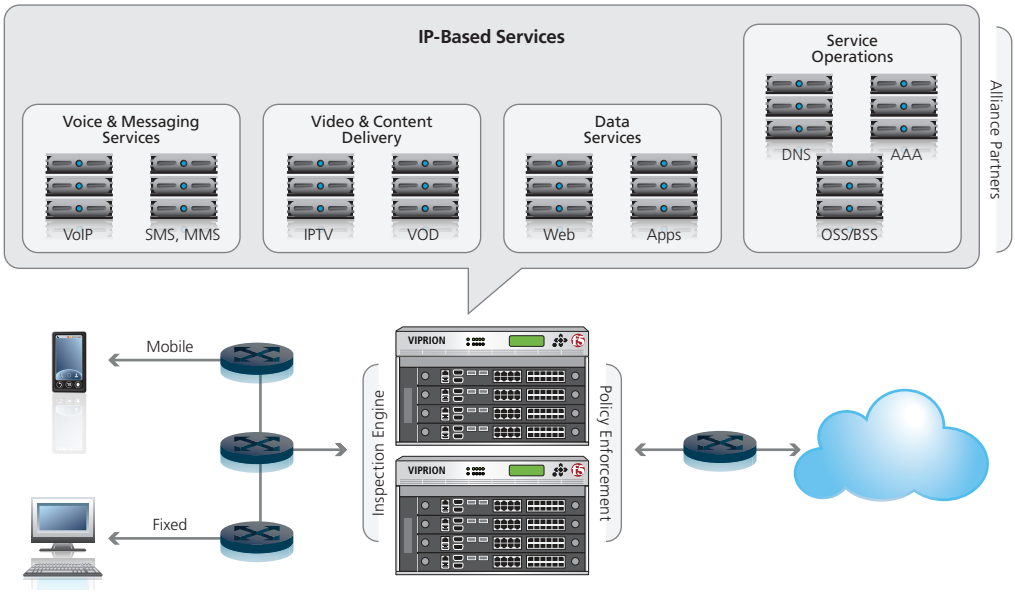
- VoIP Infrastructure
  BIG-IP devices offer scalability and flexibility for the Session Initiation Protocol (SIP) employed in IP Multimedia Subsystem (IMS) for VoIP services, giving carriers and service providers the ability to integrate new voice and multimedia services with greater ease and confidence.
- Video Traffic Steering and Service Availability
  Distributing video to mobile devices can be a challenging endeavor, making sure the appropriate content type is delivered to the specific device. BIG-IP products facilitate mobile content delivery by managing user and device access to the content servers.
- Hosted Service and Content Delivery
  As the service provider network continues to migrate to the IP core, the ability to keep individual services, traffic, and user devices segmented for privacy, billing, and management becomes more critical. BIG-IP products and solutions enable hosting and cloud providers to build a strong infrastructure foundation that maximizes the use of resources while remaining agile enough to support both traditional and cloud computing architectures.
- Managed Access
  Offering complete managed access services to customers creates a unique set of challenges for managed service providers: continually updating user databases, maintaining user state, change control, providing secure remote connectivity, and reliable application access. BIG-IP products are flexible enough to manage user access requirements in multiple locations in hosting and cloud-based user management deployments while also providing complete application delivery solutions to allow flexibility, agility, and manageability of user devices and back-end services.

The true power of the Service Delivery Network however isn't in how it solves common challenges that are fairly well-known and understood. The value of a reusable, integrated, service-aware, programmable, and adaptable solution is that it provides a unified architecture for solving the challenges that are yet to be discovered; these solutions are what separates one solution provider's offerings from another and implementing them quickly and efficiently is what separates them from the rest of the pack.

Conclusion

Service providers are facing many challenges—from the explosive growth of mobile computing, to the complexity of existing architectures and the move to more efficient ones, to the growing need to add contextual understanding to monitoring and management of services and the subscribers using them. In addition, every competitor is trying to solve the same challenges faster and more efficiently.

The solution, ideally, is an architectural approach that lays the foundation for a unified solution reducing complexity, CapEx, and OpEx while simultaneously enabling the quick and efficient implementation of new and differentiated services. This solution relies on deep session intelligence to provide true insight into the traffic and services being delivered; integrated services management to blend seamlessly with complimentary services and back-office systems; programmability to enable fast, efficient, and reliable policy enforcement; and adaptability to fit any service offering. Ultimately, the ideal solution reduces the cost of maintaining services, increased the reliability and can be leveraged to easily create new services that generate new revenue streams and higher ARPU.



Unified IP Services: service awareness, intelligent traffic steering, policy enforcement, TCP optimization, IPv6, NAT, DNS, security, and scaling.

The F5 Service Delivery Network, based on years of experience in the delivery of IP traffic, has been specifically augmented to meet the unique demands of the service provider market. With pre-built, production-ready solutions to common problems, a platform designed to adapt, and carrier-grade hardware, the F5 Service Delivery Network architecture can help overcome the challenges of today, prepare for the challenges of tomorrow and make the time in between more profitable.

1 April 2009, IDC #217920, Volume: 1, Tab: Markets Consumer Broadband Markets: Market Analysis

2 Morgan Stanley Research

3 Roche, R.F. & O'Neill, L (2010). Semi-annual data survey results: A comprehensive report. CTIA-The Wireless Association®.

4 Roche, R.F. & O'Neill, L (2010). Semi-annual data survey results: A comprehensive report. CTIA-The Wireless Association®.

5 Roche, R.F. & O'Neill, L (2010). Semi-annual data survey results: A comprehensive report. CTIA-The Wireless Association®.

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

CS01-00028 0111