



The F5 Application Availability Between Hybrid Data Centers Reference Architecture

Today's F5-powered hybrid data centers enable organizations to minimize downtime, scale on demand, and increase the ability to reliably deliver new applications and functionality.



Introduction

The path to successful application delivery has been a long and winding road for many companies.

Back in the days of Y2K and the dot.com bust, applications were often delivered out of a physical data center. This usually consisted of a dedicated raised-floor room at the corporate headquarters or leased colocation space from one of the web hosting vendors—or both.

Global organizations and ecommerce sites then started to distribute their applications and deploy them at multiple physical data centers to address geo-location, redundancy, and disaster recovery (DR) challenges. This was an expensive endeavor even without the networking, bandwidth, and leased line costs.

Enter the cloud

When server virtualization emerged and organizations realized that they had the ability to divide resources for different applications, content delivery was no longer tethered 1:1 with a physical device. Content could live anywhere. With virtualization technology as the driving force, cloud computing formed and offered yet another avenue to deliver applications.

As cloud adoption grew, along with the software, platforms, and infrastructures enabling it, organizations were able to quickly, easily, and cost effectively distribute their resources around the globe. This allowed organizations to place content closer to the user depending on their location, and provided some fault tolerance in case of a data outage. Cloud also offers organizations a way to manage services rather than boxes along with just-in-time provisioning rather than over provisioning, which can be costly. Cloud enables IT as a Service and the flexibility to scale when needed.

Today, there is a mixture of options available to deliver critical applications. Many organizations have private, owned, on-premises data center facilities. Others lease resources at a dedicated location.

Staying a step ahead

In order to achieve or even maintain continuous application availability and keep up with the pace of new application rollouts, many organizations are looking to expand their data center options, including cloud, to ensure application availability. This is important since 84 percent of data centers had issues with power, space, cooling capacity, assets, and uptime that negatively impacted business operations [according to IDC](#). That translates into application rollout delays, disrupted customer service, or unplanned expenses for emergency fixes.

Many organizations have found that operating multiple data centers is no easy task. New data center deployments or even the integration of existing data centers can cause havoc for visitors, employees, and IT staff alike. Public web properties, employee access to corporate resources, and communication tools such as email require security and back-end data replication for content consistency. On top of that, maintaining control over critical systems spread around the globe is always a challenge.



WHITE PAPER

The F5 Application Availability Between Hybrid Data Centers Reference Architecture

Simplify. Scale. Secure.

The F5® BIG-IP® platform provides organizations with global application services for DNS, federated identity, security, SSL offload, optimization and application health and availability.

Together, they create an intelligent, cost-effective, resilient global application delivery infrastructure across a hybrid mix of data centers. As companies simplify, secure, and consolidate across multiple data centers, they mitigate the impact to users or applications, minimize downtime, ensure continuous availability, and have on-demand scalability as needed.

Hybrid Data Center Planning

While cloud computing is top of mind for CIOs that want to increase their data center capacity and availability, it also offers a path to reducing system management and maintenance as well as IT infrastructure investment. It is not so much about the cloud specially but what the cloud offers: agility, flexibility, elasticity, and IT as a service. Many organizations have the means by which they can achieve this today.

Extending the traditional data center using a hybrid model means being able to integrate cloud-hosted resources. For most organizations this means expanding elasticity and contracting capacity by adding and removing resources to an application, including remote resources. Flexibility and cost savings drive this model, and when done right it can realize the benefits of cloud without requiring a wholesale data center re-architecture.

Old applications, new privacy concerns

There are a number of challenges for organizations as they try to expand to a hybrid data center model. Many legacy applications in a traditional/primary data center are not built for cloud consumption and are difficult to move. In some cases, an Infrastructure as a Service deployment enables an organization to integrate compute and storage resources into the data center, whether it's cloud-based, traditional, or simply highly virtualized.

While most organizations covet the cloud, many are also very cautious about putting mission critical, sensitive applications in the public cloud. Instead, organizations are looking to build out their own highly virtualized data center, also known as a private cloud.

Hybrid Requirements

The ability to scale to a hybrid model requires two things: resources (physical or virtual) and a means to distribute load across them. In the world of application delivery, we call the resources “pools” and the means to distribute them an application delivery controller (ADC), previously known as a load balancer.

The application delivery tier, where the load balancing service resides topologically in the data center, is responsible not only for distributing load across resources but for being able to mitigate failure without disrupting the application service. That goes for elasticity too. It is possible to add and remove resources (intentionally through provisioning processes or unintentionally through failure) from a given pool without disrupting the overall application service. This allows organizations to provision services as they happen rather than needing to over-provision equipment.

This is the heart of the business and operational value the BIG-IP platform brings to an organization: load balancing services that result in non-disruptive, seamless, and transparent elasticity.

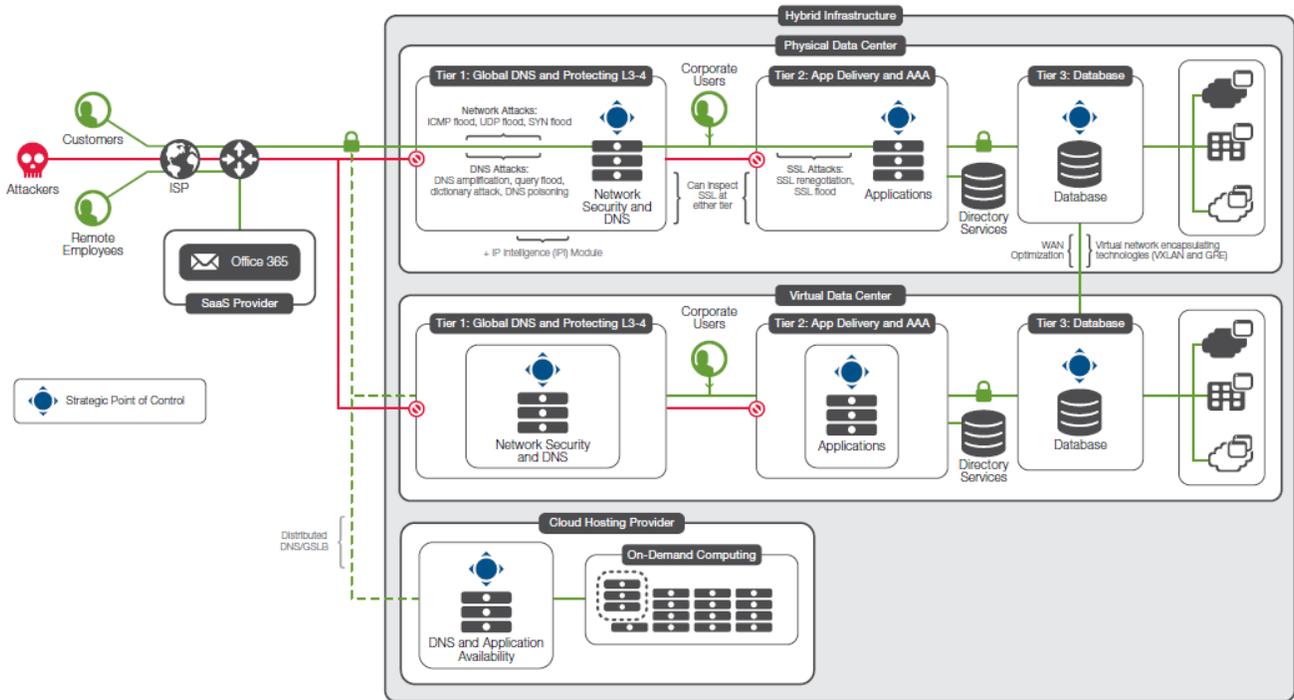


Extend and integrate

Since most organizations’ data centers already have the building blocks of cloud, it isn’t hard to see how this elasticity can extend to integrate cloud-hosted resources. Such resources are either non-disruptively added to or removed from the load balancing service’s “pool” of resources. To the BIG-IP platform, it does not matter whether the resources in the pool are local or remote, traditional or cloud, or physical or virtual—resources are resources.

Whether the data center is still very traditional (physical-based), has moved into a highly virtualized state, or has gone all the way to cloud is really not relevant to the BIG-IP platform. All resources can be operationally managed consistently by the BIG-IP platform and F5 BIG-IQ®—an intelligent framework that saves time and reduces costs by simplifying management of application delivery services.

Delivery to multiple infrastructures



Integrating cloud-based resources into the architecture requires only one thing: connectivity, along with awareness of how those resources are being used.

The connectivity between a data center and the cloud is generally referred to as a cloud bridge. The cloud bridge connects the two data center worlds securely and provides a network compatibility layer that “bridges” the two networks. This provides a transparency that allows resources in either environment to communicate without concern for the underlying network topology.



WHITE PAPER

The F5 Application Availability Between Hybrid Data Centers Reference Architecture

How this is accomplished varies from solution to solution. Emerging “virtual network encapsulation” technologies such as VXLAN and GRE are designed to make this process even smoother. A bridge only connects to places. BIG-IP does more than that—it routes application traffic across the best lane of the bridge at the best time. BIG-IQ Cloud handles application delivery management across multiple infrastructures, including data center; public; and private cloud, including AWS and vCHS.

Once a connection is established and network bridging capabilities are in place, resources provisioned in the cloud can be non-disruptively added to the data center-hosted pools. From there, load is distributed per the BIG-IP platform’s configuration for the resource, such as an application.

The Road to Hybrid

There is a perception in the market that you aren’t going to get to a hybrid data center until you have private cloud. The road to hybrid doesn’t require that you completely revamp a data center to be cloud-based before you can begin taking advantage of public cloud resources.

In fact, a hybrid approach that integrates a public cloud into your existing data center provides an opportunity to move steadily in the direction of cloud without being overwhelmed by the transformation that must ultimately occur. This also allows you to properly plan which workloads are suitable for cloud consumption.

A traditional, hybrid cloud-based approach allows your organization to build the skill sets necessary and define the appropriate operational roles that it will need. This approach also enables you to fully understand the fundamental differences in operational models required to implement the automation and orchestration that ultimately bring all cloud benefits to the table, as opposed to just cheaper resources.

A diversified hybrid solution can be as simple as a SaaS application with a direct connection to your primary data center or as large as an active-active disaster recovery platform running within a highly virtualized, cloud-based data center.

Three Scenarios for Application Availability Between Hybrid Data Centers (AAHDC)

The deployments in these customer scenarios are not mutually exclusive and can be collocated at the same site. They do not include all potential use cases.

AAHDC scenarios	Benefits
High availability	Fault tolerance for public web properties
Office 365/Exchange	Anytime access to critical business productivity tools
GoldenGate data replication	Optimized data replication to ensure data is consistent across systems



Use Case #1: High Availability

Achieving the “five nines” of website availability (around five minutes of downtime a year) has been a goal of many organizations since the beginning of the Internet era. There are several ways to accomplish this but essentially a few principles apply:

- Eliminate single points of failure by adding redundancy so if one component fails, the entire system still works
- Have reliable crossover to the duplicate systems so they are ready when needed
- Have the ability to detect failures as they occur so proper action can be taken

If the first two are in place, you may never see a failure. Keep in mind, however, that in addition to these principles maintenance is a must.

The BIG-IP platform high-availability functionalities, such as connection mirroring, configuration synchronization, and network failover, allow core system services to be available for the BIG-IP platform to manage in the event that a particular application instance becomes unavailable. Organizations can synchronize BIG-IP system configurations across data centers to ensure the most up-to-date policy is being enforced throughout the entire infrastructure. In addition, the BIG-IP platform itself can be deployed as a redundant system either in active/standby or active/active mode.

Avoiding a single point of failure

Web applications come in all shapes and sizes, from static to dynamic, simple to complex, and specific to general. No matter the size, availability is important to support the customers and the business. The most basic high-availability architecture is the typical three-tier design. A pair of ADCs in the DMZ terminates the connection. They in turn intelligently distribute the client request to a pool (multiple) of application servers which then query the database servers for the appropriate content. Each tier has redundant servers so in the event of a server outage, the others take the load and the system stays available.¹

This is a tried and true design for most operations and provides resilient application availability within a typical data center. But fault tolerance between two data centers is even more reliable than multiple servers in a single location, simply because that one data center is a single point of failure.

Maintain seamless delivery

A hybrid data center approach not only allows organizations to distribute their applications when it makes sense but also provides global fault tolerance to the overall system. Depending on how an organization’s disaster recovery infrastructure is designed, this can be an active site, a hot standby, a leased hosting space, a cloud provider, or some other contained compute location. As soon as that server, application, or even location starts to have trouble, an organization can seamlessly maneuver around the issue and continue to deliver its applications.

Advantages for a range of industries

The various combinations of hybrid data center types can be as diverse as the industries that use them.



WHITE PAPER

The F5 Application Availability Between Hybrid Data Centers Reference Architecture

Enterprises probably already have some level of hybrid, even if it is a mix of owned space plus SaaS. They typically prefer to keep sensitive assets in-house but have started to migrate workloads to hybrid data centers. Financial industries have different requirements than retail. Startups, on the other hand, might start completely with a cloud-based service and then build their own facility if one is needed.

Mobile app developers, particularly game-focused, often use the cloud for development and then bring the app in-house once it is released. Enterprises, on the other hand, have historically developed in house and then pushed out to a data center when ready. The variants of industries, situations, and challenges the hybrid approach can address are many.

The beauty of the BIG-IP platform is that it can provide the global DNS services to direct the visitor to the most appropriate data center or hybrid cloud depending on the request, context, and application health. The BIG-IP platform:

- Provides the network security to protect against inbound attacks
- Offloads SSL to improve the performance of the application servers
- Not only understands the application but also knows when it is having problems
- Ensures not only the best end user experience but also quick and efficient data replication

For maintenance situations, organizations no longer need to manually redirect traffic by configuring applications. Instead, applications are proactively redirected to an alternate data center prior to maintenance.

The BIG-IP platform is application and location agnostic, meaning the type of application or where the application lives really does not matter. As long as you tell the BIG-IP platform where to find the application, the BIG-IP platform will deliver it.

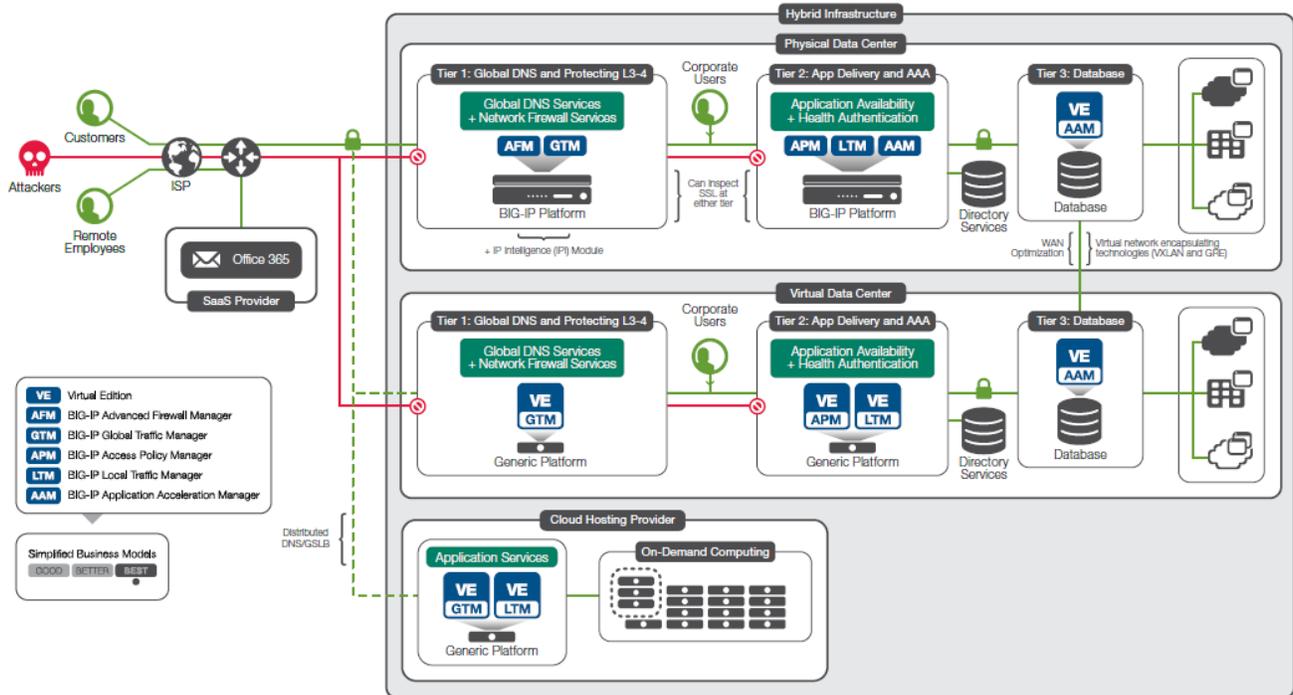
Use Case #2: Microsoft Office 365

Employee collaboration and access to communication tools are essential for workplace productivity. Organizations are increasing their use of Microsoft Office 365, a subscription-based service that provides hosted versions of familiar Microsoft applications. Most businesses choose Exchange Online as the first app in Office 365 they adopt.

The challenge with any SaaS application such as Office 365 is that user authentication is usually handled by the application itself, so user credentials are typically stored and managed in the cloud by the provider. The challenge for IT is to properly authenticate the employee (whether located inside or outside the corporate network) to a highly available identity provider (such as Active Directory).



Authentication without complexity



Even though Office 365 runs in a Microsoft-hosted cloud environment, user authentication and authorization are often accomplished by federating on premises Active Directory with Office 365. Organizations subscribing to Office 365 may deploy Active Directory Federation Services (ADFS) on premises, which then authenticates users against Active Directory.

Deploying ADFS typically required organizations to deploy, manage, and maintain additional servers onsite, which can complicate or further clutter the infrastructure with more hardware.

SAML (security assertion markup language) is often the enabler to identify and authenticate the user. It then directs the user to the appropriate Office 365 service location to access resources. SAML-enabled applications work by accepting user authentication from a trusted third party—an identity provider. In the case of Office 365, the BIG-IP platform acts as the identity provider.

For example, when a user requests his or her OWA email URL via a browser using Office 365, that user is redirected to a BIG-IP logon page to validate the request. The BIG-IP system authenticates the user on behalf of Office 365 and then grants access. The Office 365 environment will recognize the individual and provide their unique Office 365 OWA email environment. The BIG-IP platform provides a seamless experience for Office 365 users and with the federated identity that the BIG-IP platform enables, the IT team is able to extend SSO capabilities to other applications.



WHITE PAPER

The F5 Application Availability Between Hybrid Data Centers Reference Architecture

The benefit of using the BIG-IP platform to support Office 365 with SAML is that organizations can reduce the complexity and requirements of deploying ADFS. By default, when enabling Office 365, administrators need to authenticate those users in the cloud. If an IT administrator wants to use the corporate authentication mechanism, ADFS must be put into the corporate infrastructure. With the BIG-IP platform, organizations can support authentication to Office 365 and the ADFS requirement disappears, resulting in centralized access control with improved security.

Secure collaboration

Because email is a mission-critical application for most organizations, it is typically deployed on premises. Organizations using BIG-IP-enhanced Microsoft Exchange Server and Outlook can make it easier for people to collaborate regardless of their location. For example, if a company wanted to launch a product in Europe that had been successfully launched in the United States, it needs workers and contractors in both locations to be able to communicate and share information.²

In the past, employees may have emailed plain-text files to each other as attachments or posted them online using a web-based file hosting service. This can create security concerns since potentially confidential information is leaving the organization and being stored on the Internet without any protection or encryption. There are also concerns about ease of use for employees and how the lack of an efficient collaboration tool negatively impacts productivity.³

Internal and external availability 24/7

To solve these issues, many organizations move from the locally managed Exchange Server deployment to Microsoft Office 365. Office 365 makes it easier for employees to work together no matter where they are in the world. Employees connect to Office 365 using only a browser, and they don't have to remember multiple usernames and passwords to access email, SharePoint, or other internal-only applications and file shares.

In this scenario, an organization would deploy the BIG-IP platform in both the primary and secondary data centers. BIG-IP® Local Traffic Manager™ intelligently manages all traffic across the servers. One pair of BIG-IP devices sits in front of the servers in the core network; another pair sits in front of the directory servers in the perimeter network. By managing traffic to and from both the primary and directory servers, the F5 devices ensure availability of Office 365—for both internal and external (federated) users.

Ensuring global access

To provide for global application performance and disaster recovery, organizations should also deploy BIG-IP® Global Traffic Manager™ (GTM) devices in the perimeter network at each data center. BIG-IP GTM scales and secures the DNS infrastructure, provides high-speed DNS query responses, and also reroutes traffic when necessary to the most available application server. Should an organization's primary data center ever fail, BIG-IP GTM would automatically reroute all traffic to the backup data center. BIG-IP GTM can also load balance the directory servers across data centers to provide cross-site resiliency.

The BIG-IP platform provides the federated identity services and application availability to allow organizations to make a quick migration to Office 365, ensuring users worldwide will always have reliable access to email, corporate applications, and data.



Use Case #3: GoldenGate Data Replication

While delivering applications from multiple data centers is a business goal, companies still need to ensure the data at each location is current. The problem for most organizations is the size of the database and the amount of data they need to replicate keeps increasing. At the same time, the available bandwidth between locations typically doesn't change or is used by other traffic or applications. Thus, there is less available bandwidth for a growing amount of data.

Oracle's [GoldenGate](#) is database replication software. Just like the Golden Gate Bridge moves people from one side of San Francisco Bay to the other, GoldenGate moves data from one database to another. Often, this is done over very high-latency links, which can cause very slow data transfers and extremely long replication times. Generally, organizations are not moving a few megabytes of data—they're moving terabytes or exabytes. This is on top of any other data that is being transferred. Replications can happen either in real time or at certain intervals.

At the primary data center is the source database server. GoldenGate extracts data from the primary source database into Trail files, and a process called the Data Pump moves the data across the network to the target database server. Then the GoldenGate replicate process reads the Trail files and applies them to the target database. F5® BIG-IP® Application Acceleration Manager™ (AAM) with an F5® iSession® tunnel accelerates the Data Pump network portion of the process.⁴

Up to 33 times faster

The Data Pump is a single TCP connection and opening more of that TCP pipe, so to speak, moves more data faster. With BIG-IP AAM on both ends and only BIG-IP TCP optimization enabled, there is a six-fold performance improvement over the baseline. By adding encryption and compression, the BIG-IP platform realized 24 times the throughput over the baseline. The BIG-IP platform can even overcome packet loss and achieve 33 times the throughput of the GoldenGate software.

Say you had a data center in Boston and a data center in San Jose. Imagine they were getting a whopping 200MB connection. Adding the BIG-IP platform on both sides of that connection would result in a 20–50 percent improvement in throughput.

Faster in both directions

With any WAN, greater distance, more latency, congestion, or packet loss can significantly reduce throughput. If you have networks that are under provisioned or do not have enough bandwidth, or you are trying to move data over long distances (latency), the BIG-IP platform can overcome the limitations and improve the throughput to enable faster database backups, copies to the DR site, or recovery from a DR site. And the BIG-IP platform works in either direction so you can accelerate the amount of data you need to move. In addition, the BIG-IP platform can also manage traffic to the database servers themselves for greater performance.

The important thing to remember about data replication is that it is critical to use the bandwidth available. You don't need to use a chunk of the pipe all the time for the same application. You can have multiple applications going across, and you don't need to dedicate a portion of the pipe just for replication. The Oracle GoldenGate/F5 BIG-IP platform offers better performance for all the applications.



WHITE PAPER

The F5 Application Availability Between Hybrid Data Centers Reference Architecture

Conclusion

The pressure to “go cloud” is often immense. However, IT still needs the opportunity to evaluate both data center and cloud environments for appropriateness and to put into place the proper policies and governance structure around the use of cloud resources.

The application availability between hybrid data centers enables IT to begin taking advantage of cloud resources without a wholesale rip-and-replace of existing technology. It also provides the breathing room IT needs to ensure that the move to cloud will be a smooth one. This process ensures benefits will be realized without compromising on the operational governance required to ensure availability and security of network, data, and application resources. Best of all, it means IT can manage services instead of boxes.

However a hybrid data center is constructed, it is a transformational journey for both IT and the business. It is not something that should be taken lightly or without a long-term strategy in place. When done properly, F5-powered application availability between hybrid data centers can bring significant benefits to an organization and its people.

1 http://support.rightscale.com/12-Guides/Designers_Guide/Cloud_Solution_Architectures/Designing_and_Deploying_High-Availability_Websites

2 <http://www.networkworld.com/article/2449080/software/avoid-the-gotchas-of-office-365-migrations.html>

3 <https://f5.com/solutions/customer-stories/george-weston-optimizes-office-365-deployment-usin>

4 <http://gavinsoorma.com/2010/02/goldengate-concepts-and-architecture/>

F5 Networks, Inc.
401 Elliott Avenue West, Seattle, WA 98119
888-882-4447 www.f5.com

Americas
info@f5.com

Asia-Pacific
apacinfo@f5.com

Europe/Middle-East/Africa
emeainfo@f5.com

Japan
f5j-info@f5.com