F5 White Paper

Using Strategic Points of Control to Optimize the Application Infrastructure

Application performance depends on much more than speeding communication between the browser and the server. Modern application infrastructures include storage, the WAN connection, and the minute-by-minute workload on both. If the application front end is accelerated but the behind-the-scenes architecture is slowing application response time, then much of the value of any front-end investment is lost. By optimizing strategic points of control in the network, F5 products can help an enterprise leverage those front-end investments, accelerate the entire application infrastructure, and introduce a layer of agility that is just not possible with traditional architectures.

by Don MacVittie
Technical Marketing Manager

Contents

Introduction

Application delivery only becomes more complex. As the amount of data involved continues to grow, the timely delivery of applications from the server to users increasingly depends on the ability of products like F5 BIG-IP® Local Traffic Manager™ (LTM) and BIG-IP® WebAccelerator™ to accelerate connections between users and servers and to balance connection loads between multiple servers.

The success of such products is good news for users—because they spend less time waiting—and good news for the organization, which protects its relationships with the users of public-facing applications and/or increases its productivity by speeding internal applications.

But a real-world application does not exist in a vacuum, and nearly all web applications rely upon storage services to deliver content to users. Storage services are critical to internal file sharing and the use of productivity applications such as the Microsoft Office suite of products. Moreover, the amount of data any enterprise must store—and access—only grows. In 2010, 1.2 zetabytes of information were generated worldwide, with 90 percent of that data unstructured. This ever-increasing volume of data means that no matter how quickly a given web application may be delivered, the entire enterprise can be slowed to a crawl if storage resources are not similarly load balanced and optimized.

Similarly, replication—for both files and databases—relies heavily on the corporate WAN connection for reliability and timeliness. A slow, high-loss, or over-burdened WAN connection can prevent replication from ever completing, or at least result in recovery point objectives (RPOs) being set farther in the past than is best for the needs of the organization.

There are points on the network where storage, replication, and the connections required for both can be managed intelligently. With the right tools, IT staff can activate these strategic points of control to optimize application delivery and data storage. BIG-IP LTM sits at one strategic point of control, between users and servers, introducing a layer of abstraction that enables administrators to change how a web application or service is delivered with complete transparency for users. In addition, an infrastructure typically offers two other strategic points for controlling data storage.

The second strategic point of control—between consumers of storage and the storage systems themselves—is very similar to the one between users and servers. At this strategic point of control, an F5 ARX® file virtualization device introduces a layer of abstraction between the users and the files to improve access times,

redistribute workloads, streamline management of the storage environment, and potentially reduce storage costs, all without affecting users. In addition, this control point can make it easy to integrate lower-cost or capacity-optimized storage technologies into existing environments, and then can automatically place data on the most appropriate platform.

The third strategic point of control is at the corporate WAN connection. Placing F5 BIG-IP® WAN Optimization Manager™ (WOM) at both ends of a data center-to-data center (DC-to-DC) or data center-to-cloud WAN connection enables administrators to change the bandwidth being transferred over the WAN, prioritize that traffic, and optimize tasks such as replication to ensure their timely completion.

A truly optimized network takes advantage of not just one but all three strategic points of control, accelerating application delivery, managing data storage, and optimizing the WAN communications necessary for both.

64% of surveyed organizations with F5 BIG-IP report a payback period of 18 months or less.

Source: TechValidate Survey TVID: C0D-031-1FF

52% of IT organizations reduced their overall annual storage budget by an average of 20% or more by deploying F5 ARX.

Source: TechValidate Survey TVID: 3F9-1DA-708

Problem Overview

Even when an organization uses an Application Delivery Controller to speed the responsiveness of applications and implement high availability, one place in the corporate network typically remains a bottleneck for delivering applications in a timely manner. That bottleneck is the storage network, which poses constraints both in terms of storage location and I/O performance. All applications utilize the file system at several levels for configuration of file storage—including both temporary and long-term, unstructured data storage—and for database connectivity. But for most organizations, the location of a file or set of files depends more on the application or user who created it than on the need for access to that file. This dependency on application or user source creates clutter on even the fastest data center storage subsystems, and in most cases it causes underutilization of slower data storage subsystems.

Aside from the inconsistency in access speeds for data and documents with similar corporate importance and similar access requirements, the typical scenario described above can also create a highly uneven distribution of the load across the overall storage infrastructure, with some devices perpetually near full while others sit nearly idle, storing little data. The most common solution to this problem exacerbates data distribution inequity, since most organizations will decide to "use that system for more, so the load is evened out." But "more" usually translates to offering up disk space, not intelligently managing what files go where.

In addition, there is an interesting difference between the optimization of networking and storage infrastructures. Network optimization manages how data flows from one point to the next, treating data as transient; it moves from one place to the other, is used and discarded. On the other hand, storage optimization addresses both data flow and how data is stored over the long term.

The typical storage network situation in most organizations also creates unnecessarily large backup windows or replication timeframes. By mixing files of varying importance, age, and access requirements, typical storage scenarios result in much more backup or replication of data than would be necessary in a highly optimized environment. Not only does this unnecessary backup and replication contribute to the constant struggle to meet backup windows and achieve RPOs and RTOs, but when backups or replicas are maintained off-site, it increases the amount of traffic flowing over the WAN.

These issues cannot be addressed simply by building storage capacity. The addition of capacity involves a management cost that includes both the increase in overall storage-environment complexity and the greater management requirements of heterogeneous environments. Migrating data from an existing infrastructure to new hardware also can absorb a lot of staff time. Such difficulties are serious enough to make organizations delay upgrading or implementing newer technology, even when the benefits of the change are obvious. Yet unstructured storage and the requirements for it continue to grow at an amazing pace, with no letup in sight, so administrators have no choice but to manage that growth as efficiently as possible.

Enabling a Dynamic Storage Infrastructure

As with the network strategic point of control, the benefits of exercising the storage strategic point of control extend well beyond the obvious reasons for placing a device there. For instance, in many cases, approximately 80 percent of the data stored on Tier 1 devices is aged or inactive, yet organizations continue to back up this unchanged data week after week. Allowing the enterprise to dictate when and if files are backed up is impossible in a non-tiered environment because the unstructured data is dispersed across disks without grouping by importance or business value. But an organization with automated storage tier management can establish a policy that automatically moves unchanged data from Tier 1 to a deduplication platform, SATA disk, or the cloud. Backups and/or replication can

be streamlined with the now-much-smaller volumes of Tier 1 data used by the organization on a daily basis. Similarly, mission-critical data can be backed up more frequently, while other data can be backed up or replicated less often.

This strategic point of control also enables administrators to level disk usage across various NAS devices. Placing those devices, and access to them, behind the strategic point of control means that files can be automatically moved as needed to best use the total available space without impacting users in any obvious way. Users and applications neither confront the capacity restrictions of the NAS devices, nor do they see any change in how they access the files. The strategic point of control can continue to present the files as being in the same location regardless of where they are actually stored in the NAS infrastructure.

Three Strategic Points of Control

Most data center managers and systems administrators are aware of the *strategic points of control* within their network. The point between applications and their users is the most obvious one. From here, IT staff can manage connections to applications, optimize data being sent to clients, and group actual servers behind a virtual IP to ensure peak-time service of client requests and help maintain uptime.

Other strategic points of control may be less obvious. Easing the bottleneck created by single-server processing, and thus increasing the amount of front-end traffic, can quickly reveal secondary bottlenecks, or make them more painful. Fortunately, the right tools deployed at other strategic points of control can alleviate the resulting pain.

Storage Strategic Point of Control

A normal data center environment includes configured NAS devices with drives to those devices mapped on servers, whether virtual or physical. This setup gives the servers access to storage space, but limits how that space is used, since the connections are server-based, not file-based. Although it is possible to somewhat mitigate those limitations manually, doing so requires administrators (and in the broader context, corporate users) to take extra steps when the servers are deployed and to adjust with each reconfiguration of storage.

A better, technical solution to this problem does not require administrators or corporate users to manage files in a particular manner that may or may not be clear and that requires extra time and effort. This solution leverages a second strategic point of control to increase the storage system's agility and resilience.

The storage strategic point of control is that point between stored-resource users and the resources themselves. Just as access to servers and applications can be virtualized at the network strategic point of control, access to storage resources can be virtualized at the storage strategic point of control. By inserting a device into this point of control and defining virtual IP addresses that act as virtual NAS boxes, storage or network administrators can insert, remove, reconfigure, and replace the physical NAS boxes behind the point of control without compromising user access or efficiency. More importantly, this strategic point of control can be used to direct unstructured data to the ideal location without the user knowing that file A is on one NAS device while file B is stored on a completely different device.

The strategic point of control in front of the organization's NAS storage allows administrators to implement intelligent automated policies. A prime example is storage tier management, which can be implemented painlessly and heterogeneously. Set policies for files to be moved between tiers on the back end, while maintaining the apparent storage location on the front (client) end, to abstract the tiers from user control and consign them to automated handling by the system. This abstraction and automation reduces tier errors and eliminates the need to remember or look up the location of file X because it can be moved behind the strategic point of control without users being aware of or affected by any change. This is much the same strategy as moving an application from one server to another behind the network strategic point of control while the application's users see no change, and it delivers similar benefits.

WAN Strategic Point of Control

The WAN connection is another strategic point of control that, depending on the network architecture, can have a massive impact on application availability and reliability. With multiple data centers or the use of cloud computing at any level, the volume of data passing over the WAN connection can conceivably create a bottleneck that slows many applications, or even the organization's entire suite, to a crawl.

Consequently, the edge of the WAN is a critical point to insert tools that can empower administrators to manage WAN traffic, to optimize WAN connections, and to delay or eliminate costly bandwidth upgrades. The WAN strategic point of control enables the deduplication, compression, and prioritization of outbound traffic and the reversal of those tasks as data flows back into the data center.

Given the correct tools, the WAN strategic point of control also offers a point at which data can be encrypted on its way out of the building, offloading this CPU-

intensive work from the servers. The need to offload encryption is growing again as virtualization puts many more servers on each set of hardware. When only a single application or application server ran on each set of physical hardware, encryption placed a burden on the CPU, but it did not bury the CPU in encryption tasks. As virtualization grows, the number of "servers" requesting encryption by the CPU also grows, burdening the hardware much more than in a non-virtualized environment and running the risk of overwhelming it. To maintain a high virtual server density, it's critical to offload encryption to the place where it is needed, and not before.

The nature of the optimizations offered by the WAN strategic point of control requires a matching WAN optimization device at the other end of the connection. When all other things are equal, hardware optimizations are faster than software-based optimizations. This implies the value of an appliance at the WAN strategic point of control, but for organizations sending data out to the cloud or another hosted environment, hardware may not be a viable solution. So both hardware and software solutions are needed to make the most of this strategic point of control—a virtual machine for deployment in cases where the hardware is owned by a third party, and an appliance for deployment where the infrastructure is under the aegis of the IT department. This system of alternatives offers speed and high volume when an appliance can be deployed and flexible deployment models when one can't. Conveniently, the appliance in a data center can also act as one end of several connections, as may be necessary when the organization relies on several cloud vendors or a WAN connection to a remote data center.

Once the WAN strategic point of control has been implemented, much added flexibility can be implemented, too. For instance, shift loads or increase availability by routing requests to any of the locations serviced by the strategic point of control. Another example is the ability to use rate-shaping to control how much data of a given type is allowed to pass through the connection, which can guarantee bandwidth to mission-critical applications.

Leveraging Strategic Points of Control with F5 Products

It is a reality of the current IT environment that the amount of unstructured data—files and folders—that any organization must store tomorrow will be greater than it is storing today. For some organizations the amount of disk space used by unstructured data is massive. According to International Data Corporation (IDC), growth in

unstructured data will average more than 60 percent per year. That's a doubling in volume every 20 months. This rate of growth brings with it more than just a lot of terabytes of disk space. It also brings inefficiency in storage utilization, increasing backup times, more disruptive data migrations, rising costs, and access issues as array after array is added to the mix.

F5 ARX: Intelligent File Virtualization

F5 ARX is a file virtualization device that puts a layer of abstraction between clients and mass storage to enable management with the ARX strategic point of control. This layer of abstraction is very thin—stored files maintain their names and storage mechanisms—yet it introduces a way to strategically control data, access that data, and even manage how backups and replication are handled.

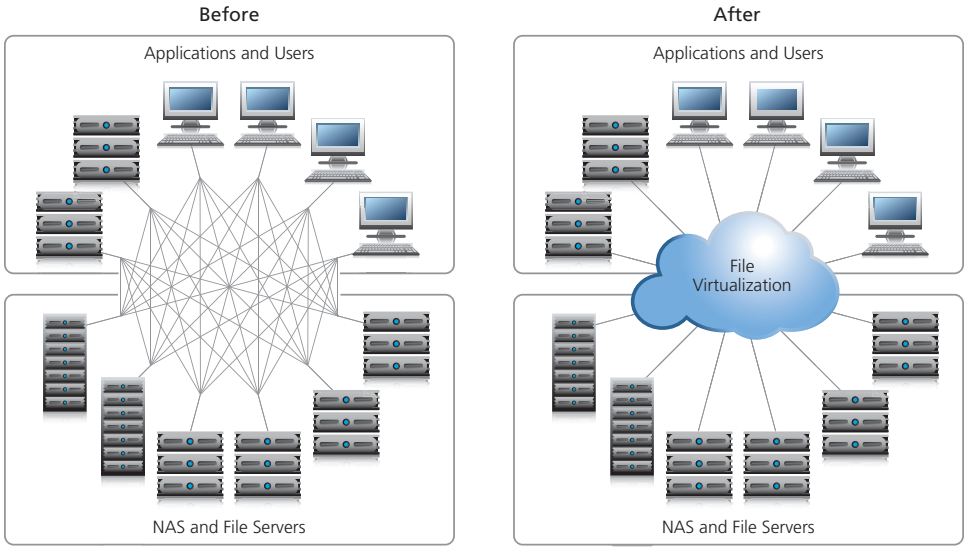

Figure 1: ARX streamlines NAS connections and inserts a strategic point of control between file services and consumers of file services.

Situated between computers and storage, mapping what is actually there to the files and folder arrangements the client expects, ARX is a powerful device for leveling storage resources and automating rules-based tier management. Rules-based tiering enables IT to establish policies that automatically place data on the appropriate tier of storage, depending upon the data's business value. By virtualizing the storage environment and using automated policies to move lower-priority data to slower disks, ARX enables IT staff to cut through storage clutter, dramatically simplify how

data is managed, and reduce costs with less expensive disk space more suitable for second-tier data.

Subsequently, ARX can improve RTOs by ensuring restoration of important data on tier one before tier-two restoration begins. The backup window can benefit significantly when less-frequently-modified files are moved from tier one to tier two or three, and then backups are reconfigured to continue nightly backups on files that change often but only occasionally back up files that don't. Reducing the frequency of backups for tier-two data will reduce the time required to complete nightly backups, in some cases by a major percentage of the overall backup window. Such tier-based backup scheduling is a quick and easy way to reduce the growing problem of large backup times.

ARX also mitigates the issues of capacity growth. With an ARX device in the network, a new storage device or cloud storage solution can be easily added by placing that solution behind ARX in a pool of available NAS devices and marking its place in the existing directory structure. Administrators can then establish a policy to automatically move files from over-utilized devices to the new one—without impacting users or applications and without IT staff interaction. In fact, this migration is so transparent; it can take place during normal business hours. Because users' hierarchies map to ARX, the storage device behind a user folder can change without the user ever noticing, unless the new system is implemented as a separate sub-folder on ARX. (Since the new storage device can be implemented as the same sub-folder that user systems already point at, however, it is unlikely that IT decision-makers would want to cause users and administrators unnecessary pain.) Regardless, if files are moved from an overburdened array, users may well notice a massive performance improvement when opening their files.

Finally, the thin layer of virtualization provided by ARX means that users won't realize that some files in a given directory are stored on high-performance NAS device X, while others are stored on a far less expensive NAS device, Y. IT staff thus gain the opportunity to tailor storage to the business value of the data stored there while increasing their ability to control costs in purchase decisions.

In all, ARX provides greater control of storage at the points that matter most—between the storage devices and the clients, and between the storage devices and the backup software.

F5 BIG-IP WOM: Replication and Remote Transfer Efficiency

When BIG-IP LTM helps to balance server utilization and ARX helps to balance storage utilization, the one element left in an enterprise architecture that could cause bottlenecks is the WAN connection—whether the bottleneck involves replication to a disaster recovery site or information synchronization between two load balanced sites. While a fully utilized WAN connection can become saturated with data pretty easily, the cost of an upgrade may be prohibitive. And the more latency and packet loss a connection suffers, the more likely saturation problems become as the send queue backs up, waiting for recovery from lost packets or for acknowledgements (ACK responses) from remote data centers.

Enter F5 BIG-IP WOM, the premiere WAN optimization product for the enterprise data center. Working as a standalone product or in conjunction with BIG-IP LTM, BIG-IP WOM compresses and deduplicates data while securing it against prying eyes and optimizing the underlying TCP connection. BIG-IP WOM capitalizes on the strategic point of control for the corporate WAN connection, managing, securing, and monitoring all that goes in and out of the building.

If ARX is already in place, the first and most visible result of a BIG-IP WOM deployment is further reduction in the enterprise backup window. When tested with some of the world's leading replication products, BIG-IP WOM improved their performance over the WAN by 2 to 59 times, depending on the product and the quality of the connection between the two locations. The more latency and packet loss that exists on the WAN connection, the more repeat data being transmitted, and the more compressible the data flowing over the connection, the more improvement BIG-IP WOM can achieve. If the data is highly compressible, involves large chunks of repeating data across multiple streams, and crosses on a poor, over-utilized WAN connection, the resulting performance improvements could be much greater than the test numbers noted above.

Unlike ARX, BIG-IP WOM products must be deployed in pairs so that compression, deduplication, and encryption can be reversed at the other end of the line. The power this symmetric deployment grants an organization is a large reduction—often by many times—in the bandwidth consumed by applications running over the WAN. Symmetric deployment also offers other, less-obvious benefits. Both data centers will reap the rewards of using this strategic point of control, which enables prioritization of in and out traffic, configuration of IP tunnels with enterprise-class security for specific applications or services, and improvement in both send and receive times.

BIG-IP WOM can also perform layer-seven rate shaping, allowing administrators to set minimums and maximums for bandwidth utilization by different applications and protocols. This feature helps IT staff to guarantee that the applications most important to the organization will get the bandwidth they need without struggling against congestion caused by less critical data.

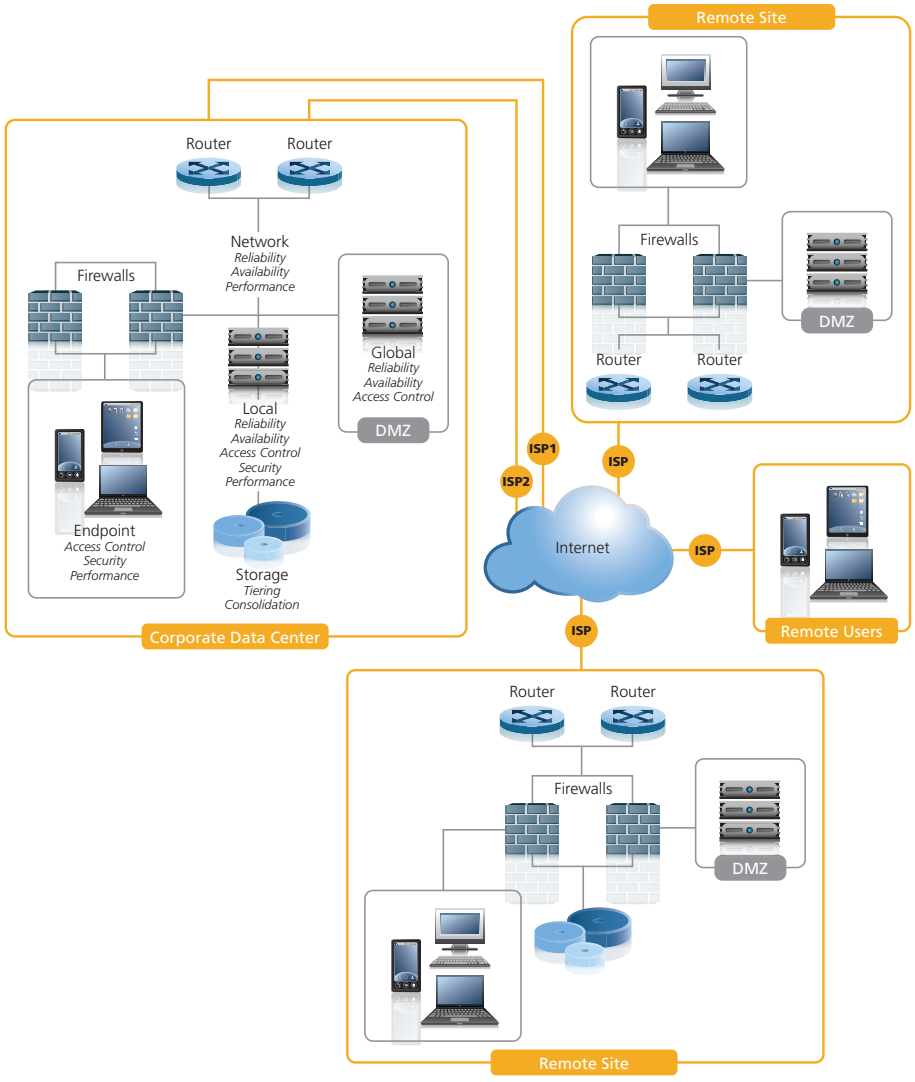

Figure 2: F5 products—including ARX, BIG-IP LTM, and BIG-IP WOM—sit at strategic points of control to help accelerate application delivery, data storage, and WAN communications while simplifying complex management architectures.

The solution to an ever-more-complex data center is to leverage strategic points of control to optimize data traffic and storage with products that can do much more than one function. For storage, the strategic point of control is between the files and the user. For networking, it sits between the user and the server, and for the

WAN, the control point lies between the servers and the public Internet. When placed in these locations, ARX, BIG-IP LTM, and BIG-IP WOM can manage network traffic, improve resource utilization, and simplify the management architecture.

Positioned between users and applications and the NAS head, ARX inserts a layer of abstraction that administrators can use to make users' lives easier and reduce their own overhead requirements. Similarly, BIG-IP LTM provides a layer of abstraction between servers and their clients that can be used to make applications more responsive and reduce administrator overhead. And when placed between the LAN and the WAN, BIG-IP WOM can reduce the bandwidth utilized between data centers and reduce administrator overhead. Taken together, these products can improve human resource utilization; improve storage, server, and WAN link utilization; and make the corporate data center more dynamic. By exerting control at strategic points, administrators can manage data center complexity without introducing one-off solutions that increase the very complexity they are intended to manage.

Synergistic Solutions

Every data center has strategic points of control where abstraction layers can be introduced to enhance automation and give IT staff increased capability to manage the network and its data. The enterprise network is busy now, and the increasing number of portable devices on the Internet will soon make it busier. Organizations that do any work "in the cloud" will have networks—or at least WAN connections on their networks—that will be even busier still. Consequently, IT staff need solutions to help manage that growing traffic without consuming all available resources.

That's where BIG-IP LTM, BIG-IP WOM, and ARX come in. While BIG-IP LTM concerns itself primarily with web-based users and BIG-IP WOM concerns itself primarily with high-volume, back-end transfers, ARX concerns itself with the storage hierarchy. When applied to individual strategic points of control, each of these F5 products is a powerful tool. Taken in concert, they optimize server response times, storage utilization, and DC-to-DC communications, lightening the load on both the LAN and the WAN while freeing administrators to focus on adding value to the business rather than scrambling after manual or piecemeal solutions to throughput, storage, and traffic management issues.

By applying the F5 portfolio of products across its network, an enterprise can cut costs, increase performance in nearly every section of the data center, and simplify administration. Corporate servers can be sped at every stage of the communications

process—file, network, and WAN I/O—and encryption can be offloaded to BIG-IP LTM or BIG-IP WOM. Furthermore, these products and their benefits are multiplicative. For instance, BIG-IP WOM can run on BIG-IP LTM, extending the functionality of the latter to remote data centers over the WAN. ARX communicates via IP, which, depending upon the local network configuration, can garner benefits from both BIG-IP LTM (if the NAS devices are local) and BIG-IP WOM (when remote NAS devices are involved). This synergy among market-leading tools, when deployed together, results in a powerful, multi-faceted solution that can optimize the data center while enhancing and streamlining its management.

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

CS01-00063 0811