

Risk

++++technology overview+++++

takers

BY TRACY THOMPSON

Illustration by Adam McCauley

Summer, 2001. The dot-com implosion had already ravaged the stock market and forced hundreds of start-ups into bankruptcy. In Seattle, the fallout was still painfully clear: Businesses like Mylackey.com, which picked up your dry cleaning or rented movies for you, now had Office Space for Rent signs in their front windows. Whole floors of buildings—indeed, a few whole buildings—went dark.

And F5, the company where I work and which had survived the dot-com disaster, and had later begun to flourish by shifting its business model from dot-coms to large enterprise customers, decided to go ahead and redesign its best-selling local traffic management product, BIG-IP.

The risk was huge. Not only were we planning to redesign a product that was responsible for almost 80% of our yearly sales at the time, but we'd also be shifting 80% of our research and development budget into a new product with a new architecture. We risked losing not only time—about three years in development time alone—but also market share to a couple of major competitors.

But if we did this redesign correctly, nobody else would be able to touch it, and we could offer a new architecture that nobody

Why F5 bet 80% of its R&D budget to redesign a product that already worked just fine.

could match and that would pay off for our customers for years to come.

We decided to go for it. And although the rest is history, more history is about to be written, and there's a lot more to the story.

The old standby

F5 used to be known primarily as a load-balancing company. That is, in their most basic form, our systems (Intel-based hardware with our own proprietary software) sat in front of web servers at, say, Virgin Airlines' website. When you wanted to purchase a ticket online, our product, behind the scenes, directed you to the best server so you'd get the fastest response. *The Wall Street Journal* called F5 a "traffic cop" for Internet commerce, and its analogy was reasonably accurate.

Granted, the concept and practice of basic load balancing is still a great thing—and I believe we continue to do it better than anyone else. But around 2001, our customers, many of them major enterprises, began to request that our devices do more than just load-balance simple HTTP traffic between their servers. They wanted to apply the model of virtualizing web servers to provide high-performance, cost-effective scaling and security to other application types beyond HTTP. They wanted to compress their traffic. They wanted to look deeper into the application data stream so they could control where to send it. And they wanted to add security services.

"The management of our devices became critical—the ability to provide version control, fast software updates, licensing schemes, and audits," said Erik Giesa, F5's vice president of product management and product marketing. "All of that would be wickedly hard to do if the underlying operating sys-

tem and architecture were different for each of our products.”

A new way of thinking was needed, a new architecture to handle the many new demands. That architecture, introduced in September of 2004, was called TMOS.

TMOS—Driving application-level intelligence into the network

Developed over three years, TMOS is a foundation for load balancing, application firewalling, data compression, local caching, protocol and WAN optimization, QoS assurance, and other services that secure, accelerate, and optimize enterprise applications. TMOS eliminates the drawbacks of handling traffic when these services are performed in different devices, and it also enables enterprises to scale up to meet growing needs.

TMOS essentially enables organizations to move the work that had to be done exclusively in the application onto the network, resulting in better control, lower cost, and faster time to market. “It frees up the servers and applications, so that server CPUs can be dedicated to processing the application functions,” says Giesa. “The only other alternative for customers was to deploy separate devices in front of their servers and caches.

With TMOS, they can remove that entire layer—abstract it up one layer, if you will.”

“It provides a common services layer for modular forms of load balancing, SSL processing, data compression, WAN optimization, application-specific firewalls, and other modules,” says Steve Steinke, an analyst for the New York-based 451 Group. “In order to perform these functions, devices need to analyze TCP/IP packets, parse HTTP and XML streams, decrypt SSL rapidly, inspect packets for malware, and identify

traffic otherwise counter to policy. Having multiple appliances or even separate blades for these functions ensures duplication of effort as well as management complexity and potential unintended side effects.”

Packet-by-packet versus full-proxy

One of the key differentiators of the TMOS design versus other designs is how TMOS handles the traffic. A packet-by-packet design is where a device in the middle of a stream of communications is not an endpoint for those communications, but instead just passes packets of data through. A full-proxy design is the opposite of a packet-by-packet design. Instead of having a minimal understanding of the communications streaming through the device, a full proxy completely understands the protocols, and is itself an endpoint and an originator for the protocols. F5’s BIG-IP v9 with TMOS is a full-proxy design.

So what does this all mean? Well, a full-proxy device has the advantage of more easily supporting application-level protocols.

That is, by being able to completely “speak” protocols, it affords a much greater degree of flexibility in actively optimizing those protocols (as opposed to a packet-by-packet design, which supports protocols passively, for the most part).

“TMOS was designed to enable high performance when running multiple services concurrently—SSL, compression, content switching, caching,” notes Giesa. “Because businesses have a mix of traffic, you’re doing multiple things at once. In order to address multiple functions beyond just the web, we had to become a true, full, IP-generic proxy. With TMOS, we own the connections.”

Functionality others aren’t built to match

So can other solutions offer what F5’s TMOS brings to the table? When you take a closer look, the answer is no. Back in 2001, we knew we were going to acquire other companies to round out our product portfolio in order to build what we call the application delivery network. We also knew that we needed an architecture that would give us the flexibility to support multiple application types while being fast, scalable, and modular. We started with the right foundation—TMOS.

“It’s like constructing a building,” says Giesa. “If you want a good structure, you start with a good architectural plan. Then you build a foundation. But some of our competitors are doing this backwards.”

For example, one of F5’s competitors features a load-balancer that is based on Linux, a WAN optimization solution that is based on VxWORKS, and a network firewall that is primarily based on an ASIC architecture. On top of that, the vendor has its own operating system. There’s no common foundation for sharing services or delivering a common management framework, and no way to extend the functionality or layer the technologies it has acquired without tacking on more systems or devices.

“F5’s TMOS architecture [provides] the ability to handle traffic only once and the ability to extend existing capabilities with software modules,” says Steinke. “In fact, BIG-IP comes with software modules already installed. Installing the add-ons is simply a matter of paying for and enabling the modules.

“Compare this with Cisco Systems’ IOS,” Steinke continues, “which is about as far from modular as can be imagined. For all practical purposes, Cisco will continue to have difficulty adding fully integrated performance and security features—hardware as well as software—to its boxes and blades.”

Worth the risk

The question remains, though: Has the initial risk of this redesign paid off?

“In the last year alone,” answers Giesa, “we’ve surpassed Cisco in market share for application delivery products, moved into the application delivery leader position in the Magic Quadrant for Gartner, and shipped three new products on the TMOS platform, including those we had acquired. We’ve also surpassed revenue and earnings expectations, and we’ve delivered on the promise of integrated products that are now forming the world’s first complete application delivery network.

“Risk,” he adds, “has its rewards.” ✱

Additional resources

Data sheet: *BIG-IP Local Traffic Manager v9 with TMOS* (www.f5.com/products/bigip/ltm/tmos.html)

Case study: *BorderWare Creates iRule* (www.f5.com/solutions/success/borderwareirule_ss.html)

F5’s DevCentral Developer Community (<http://devcentral.f5.com>)