IRONPORT

IronPort Core MTA Architecture

WHITE PAPER

TABLE OF CONTENTS

Executive Summary

The Revolutionary Queuing Engine that Powers the IronPort Security Gateway Appliances

Email has become the dominant form of business communication, rivaling if not exceeding the importance of voice networks. Indeed, email has had such an extraordinary impact that, like the fax and ATM, it's hard to imagine life before its widespread adoption over the last decade. The very power of the medium has also attracted a disturbingly large and growing number of security threats – spam, fraud, viruses, regulatory violations and intellectual property theft.

Email infrastructure must scale to meet the growth of the medium, and at the same time respond to the growing email related threats. Throughput at the mail gateway has been growing at nearly 100 percent per year driven by three compounded factors: (1) more email users (2) increased number of emails sent per user per day, and (3) rapidly growing message sizes due to the increasing adoption of HTML email and large attachments. A more disturbing trend has also fueled the growth in perimeter email volumes – the exponential and unending growth in volume of illicit messages and attacks.

Traditional mail gateways are built on architectures that are as much as 20 years old. These old architectures have inherent limitations with queue design and implementation that make them unable to meet the growing mail throughput requirements and the demand for more sophisticated spam, virus and policy filtering.

To address this need, IronPort® Systems developed a completely new MTA design. Using advanced queuing techniques, the IronPort MTA delivers a massive performance breakthrough – as much as 20x greater throughout than traditional UNIX-based systems. This performance leap allows the IronPort security appliances to handle the growing spikes in mail volume associated with a spam or virus outbreak, and still have power left to perform advanced message filtering algorithms. The robust design of the queuing engine also allows the IronPort appliances to act as a shock absorber between the private network and the Internet, absorbing the ups and downs of the public Internet and providing a steady stream of mail into or out of the groupware servers, thereby enhancing overall messaging system stability and availability.

THE FOUNDATION: ASYNCOS

Many of the limitations of a traditional UNIX-based gateway program lie not in the application itself, but in the way the applications interact with the underlying operating system. To address these limitations, IronPort has developed a unique operating system called AsyncOS™, specifically optimized for the asynchronous task of relaying email messages.

Email is a connection intensive medium. Any reasonable sized network may easily have thousands of simultaneous mail connections coming in or going out. These connections are often relatively slow as they may be connected to a busy mail server at the other end of the Internet. A traditional MTA has difficulty dealing with a large number of simultaneous connections. Most traditional MTAs running on general-purpose operating systems such as UNIX or Windows are limited to 100 or maybe 200 simultaneous connections because the operating system limits the number of threads that can be open at the same time. This is because the traditional threading model requires a dedicated memory stack for each thread, and the system cannot provide more memory to open new threads. IronPort's AsyncOS features a stackless threading model that does not require a large memory stack for each thread, allowing the IronPort MTA to support a massive concurrency – 10,000 simultaneous connections – 100 times greater than a traditional MTA.

This massive concurrency ensures that for all practical purposes the IronPort MTA will never be connection bound. Solving the concurrency bottleneck means that the bottleneck shifts to I/O. Since all messages in an MTA must be safely written to disk, the MTA is an I/O intensive application. The I/O bottleneck is addressed in AsyncOS in two ways. The first is through IronPort's I/O driven scheduler. AsyncOS takes advantage of the asynchronous nature of messages to process messages in any order that is optimal. If a thread is actively using I/O, the system allows it to finish its I/O transaction, and will not incur the overhead of a context switch induced by a time-based scheduler. This increases the efficiency of the I/O system dramatically. The second

I/O optimization is in the AsyncOS file system. Traditional MTAs use the file system to maintain the state of the application. If a major receiving domain becomes unavailable and a queue starts to grow, the overhead associated with a traditional file system begins to drag down the overall throughput of the machine. So when the receiving mail domain comes back online, the MTA needs to resume delivery and clear the queue. But at the moment the MTA needs maximum throughput to clear the queue, the file system overhead actually makes the throughput minimum. So the queue grows, causing more overhead, which in turn results in a bigger queue, until the system finally grinds to a halt, requiring administrator intervention.

ADVANCED QUEUE DESIGN AND CONNECTION MANAGEMENT

On top of this heavily optimized operating system, IronPort has developed a completely new MTA architecture. The IronPort contains a unique independent queue design. The system maintains a separate queue for every destination domain. The system also maintains an awareness of the state of all receiving domains. If a major domain such as Hotmail goes down, the system marks the domain as being down, and all new messages from the groupware servers are placed in the queue for that domain, but queuing a message for a down domain will not initiate a separate retry cycle for each new message received. Instead the IronPort security appliance parks all messages for the down domain, and does a single, global retry on that domain. When the receiving domain comes back up, all messages are delivered. This solves a very common problem for traditional MTAs. They frequently become paralyzed by large numbers of retries on a popular host that is down. Similarly, the IronPort MTA has the ability to set the retry schedule on a per-domain basis. This solves another very common MTA problem, large numbers of bounced spam messages that plug the system queues. Spam attacks frequently have high rates of invalid email addresses. These bounce messages are often going to a domain that never accepts mail in the first place, sending a traditional MTA into a fit of retries for mail that was junk to begin with. It usually requires a system administrator to intervene, sort through the queue, and destroy or remove all messages bound for the offending domain. By adjusting the retry on a per-domain basis, administrators can set the retry to zero for suspect domains and allow the IronPort to clear these messages automatically. These capabilities allow the IronPort appliance to act as a “shock absorber” in front of the groupware servers, queuing messages gracefully without manual attention.

IronPort security appliance also have a unique feature called Virtual Gateway™ technology. Virtual Gateway technology allows the system to identify and assign unique classes of mail to unique outbound IP addresses. This can be used to separate the outgoing mail for different organizations onto different outbound IP addresses. This is a very powerful feature for managing issues of deliverability. If any of the different mail streams cause problems with

a receiving ISP that leads the ISP to block that mail, the blockage will only be limited to the IP that caused the problem, allowing mail from the other mail streams to flow without interruption. This capability is a must for service providers that have shared infrastructure – each customer can be given their own unique IP address, ensuring no one customer will impact the mail flow of another. The other critical use of this is to separate commercial mail such as bill payments or transactions from employee generated mail. This way if there is problem with one mail stream it will not impact the operation of the others. A common application of this technology is bounce handling. Often spammers will send messages with a forged return address (otherwise known as a "joe-job") that contains a known spam trap. When the responsible corporation generates a non-deliverable bounce message, it hits the spam trap and results in blacklisting the responsible corporation. By bouncing messages on a separate virtual gateway, if a bounce hits a spam trap it will only impact the bounce mail IP address, and not interfere with corporate mail flow. The IronPort queuing engine builds separate queues for each destination domain on a per-virtual gateway basis, extending the robust queuing across multiple virtual gateways. Thus a popular receiving domain like Hotmail might have a separate queue for each virtual gateway set up, ensuring that if one virtual gateway is blocked the others continue to send mail.

In addition to advanced queuing, the IronPort MTA design has excellent connection management. The system queues and groups all messages going to a common domain. It sends multiple messages per connection, and opens multiple connections per host. Traditional MTAs will open a new connection for each message delivery, adding massive overhead to both the sending and receiving MTA.

IronPort's "Good Neighbor" algorithm senses the aggregate data rate across all connections to a given domain, and when the data rate starts flattening out it drops the newest connection, ensuring the receiving mail server does not become overloaded. It has an on-board DNS cache that is extremely high performance and matched to the throughput of the system. The cache will store the IP addresses of all MXs for a receiving domain, and spread connections across the various MXs according to the MX preference advertised by the receiver.

FLOW CONTROL

The IronPort MTA has powerful flow-control capability. This allows the MTA to regulate the rate at which it will accept messages from any given sender. This capability is linked to IronPort Reputation Filters – the system that assigns a reputation score to the IP address of the server delivering the mail. The reputation score is provided by IronPort's SenderBase® Network. SenderBase is the first and largest email and Web traffic monitoring system. SenderBase tracks a variety of network parameters about any given IP address sending mail on the Internet. These parameters include the global volume of mail

sent by any given IP address, how long that IP has been sending mail, country of origin, open proxy or open relay detection, appearance on any black- or whitelists, proper DNS configuration, ability of the sender to receive mail in return, etc. These objective, network level parameters are rolled into a score ranging from -10 to +10 that reflect the "trust worthiness" of the sender. IronPort Reputation Filters will automatically assign a mail flow policy, based on the reputation score of the sender. This policy includes message size limits, maximum simultaneous connections, keyword filtering, spam or virus filtering on/off, and most importantly the maximum number of recipients per hour the gateway will accept. This key capability enables IronPort's unique variable response to suspicious traffic. The more suspicious a sender appears, the slower it goes.

POWERFUL MANAGEMENT CAPABILITIES

In addition to very sophisticated queue design and message handling, the core IronPort MTA has powerful queue management features built in. From the Web-based IronPort Mail Flow Monitor™ interface, system administrators can easily view the activities in their queues. The top domains for which messages are queued are displayed with summary statistics. From a single page the administrator can see the number of messages in the queue, the number of connections open, successful deliveries, as well as hard and soft bounces. Drilling down on any one domain provides more detail. It shows the IP addresses of all MXs associated with that domain, the status (up/down) of the domain, last connection attempt, and the age of oldest message in the queue for that domain, and more detail on bounce messages or error codes received from the domain. This comprehensive tool allows system administrators to rapidly identify and troubleshoot problems. Alerts can also be generated when queues exceed certain thresholds. For additional troubleshooting, IronPort supports "domain debug," a special logging feature that captures every piece of the SMTP conversation only for a specified domain, allowing the administrator to identify problems without creating a massive, unwieldy log to parse. To further enable troubleshooting, specific messages that are in the queue can be redirected to any remote host or server, or simply deleted.

GET THE MESSAGE

The revolutionary design of the IronPort MTA yields the performance and functionality required to enable multiple layers of very sophisticated filtering on incoming and outgoing mail. These layers of filtering ensure that legitimate mail gets delivered without interruption, and illicit mail is stopped at the perimeter. The IronPort filtering capability is detailed in the white paper titled *Reputation-Based Email Security*. Couple this filtering capability with the world's most powerful and robust MTA, and you can be assured that your users will always "get the message".

IRONPORT

IronPort Systems
950 Elm Avenue, San Bruno, California 94066
TEL 650.989.6500 FAX 650.989.6543
EMAIL info@ironport.com WEB www.ironport.com

IronPort Systems, a Cisco business unit, is a leading provider of anti-spam, anti-virus and anti-spyware appliances for organizations ranging from small businesses to the Global 2000. IronPort appliances utilize SenderBase, the world's largest email and Web threat detection network and database. IronPort products are innovative and easy-to-use—providing breakthrough performance and playing a mission-critical role in a company's network infrastructure.

IronPort is now part of Cisco.

CISCO