Cisco Mail: High availability design for your most mission-critical application

IT managers agree that one of their most crucial job responsibilities is making sure their company's email service is available around the clock. Cisco Mail is an enterprise-grade email and collaboration service that gives companies of all sizes the high availability and disaster recovery capabilities that, until now, only very large enterprises could afford. Furthermore, all employees benefit from high availability email, not just a select few. At the same time, Cisco Mail offers on-demand scalability for adding capacity as needed and native Microsoft® Outlook® support, eliminating adoption barriers and user training.

This white paper explains the resilient architecture behind the Cisco Mail service. The unique combination of scalable Cisco® cloud infrastructure, Linux®-based product architecture, fault-tolerant service configuration, and operational expertise makes it possible for us to offer a Service Level Agreement-backed 99.9% uptime guarantee and a 4 hour recovery time objective (RTO).

The Cisco Collaboration Cloud

Cisco Mail delivered over the Cisco Collaboration Cloud http://www.cisco.com/en/US/prod/ps10352/collaboration_cloud.html is built on a high availability architecture that provides continuous service and rapid recovery for business continuity.

The Cisco Collaboration Cloud is a private, global network owned and operated by Cisco that has been delivering fast, reliable and highly secure SaaS since 1995. This purpose-built communications infrastructure employs multiple fully redundant data centers strategically located near major Internet access points around the world, and uses dedicated high-bandwidth fiber to route traffic around the world.

Data centers are equipped with generators, uninterruptible power supplies, physical security, and redundant cooling systems. Experienced Cisco personnel provide logistical security, operational, and change management support, as well as continuous monitoring, 24x7, 365 days per year. Highly trained Cisco engineering staff members are always available to quickly address any issues within each data center.

Cisco Mail high availability architecture overview

The foundation of the Cisco Mail service is a compartmentalized, fault tolerant architecture that delivers fast disaster recovery and ensures business continuity.

Compartmentalized architecture

Cisco Mail employs a unique "pod" design, which is a sub-divided architecture that contains problems, such as component failures or application corruption, and prevents them from impacting a large number of users. Each pod incorporates a specific set of network, front-end, and back-end components that is easily replicated, as shown in Figure 1.

Within each pod, the front-end components that interface with the Internet are separated from the back-end processing and data storage systems. The combination of the pod design and the separation of front-end and back-end systems speeds issue isolation for fast disaster recovery.

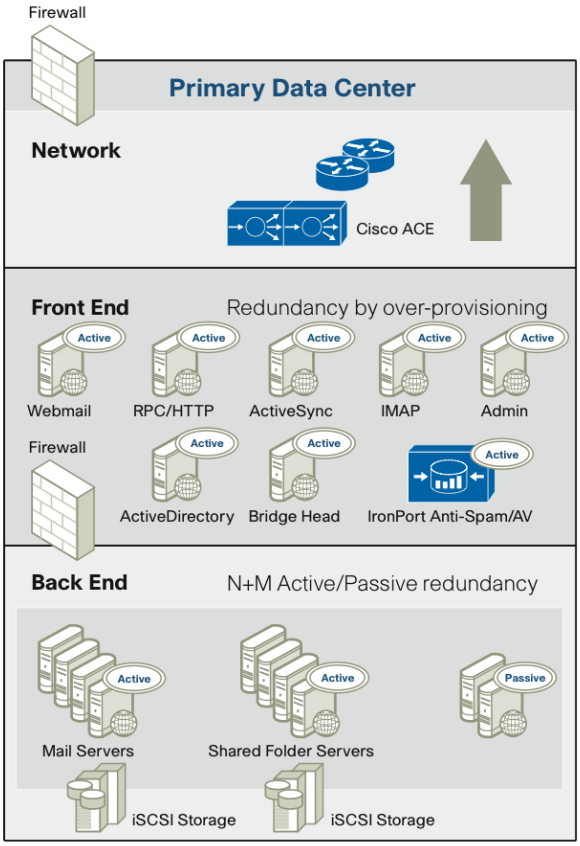

Figure 1: Pod design

Geographic redundancy

Historically, geographic redundancy has been beyond the reach of many companies because of cost constraints. Cisco's economies of scale allow us to offer customers full local and geographic redundancy. Each system component is fully redundant within a pod and also across geographically dispersed sites. This helps to ensure service continuity even in the case of a broad scale catastrophic event. Data is replicated in real time between the primary and backup datacenters to protect against data loss.

Rapid disaster recovery

Fault tolerant Cisco Mail architecture automatically fails over to redundant systems within the primary data center or to the remote backup data center, as required. Once the cause of failure has been identified and rectified, service is switched back to the primary system. In addition, snapshots are taken every 8 hours at both the local and remote data storage subsystems, so that, even in the event of application failure, data can still be recovered. Five snapshots are kept at both locations, so there are 10 restore points in total, ensuring the highest level of data protection recoverability.

The Cisco Mail platform is a Linux-based system that employs standard Linux file system components and tools that have been proven through years of use by thousands of organizations. The Linux file system provides exceptional flexibility, enabling block-level replication and granular restores down to the single-message level.

Monitoring and maintenance

Cisco Mail data center Network Operations Centers (NOCs) are staffed around the clock, 365 days a year, by experts who can identify and resolve issues quickly. Sophisticated monitoring systems continuously check the health of individual components, along with the performance of the service.

Server-level monitoring tracks server availability and network responsiveness. Monitored resources include CPU, memory, and network interface, as well as file system, process, and application availability.

Application-level monitoring watches every application within the Cisco Mail infrastructure for proper functionality and performance. Monitored functions include, but are not limited to, message delivery, storage and retrieval, end-user interfaces, and email security.

Performance-level monitoring tracks key performance indicators for webmail, MAPI, BlackBerry and Microsoft ActiveSync®, including email round-trip delivery time within the Cisco Mail infrastructure, throughput, and user performance.

If any function, process, or performance falls below threshold levels, an alert is generated and sent to NOC personnel.

High availability Cisco architecture and round-the-clock monitoring and maintenance enable you to focus on core projects rather than spending time firefighting email problems.

Cisco Mail architecture deep dive

The Cisco Mail architecture provides multi-layer redundancy throughout the network, front-end, and back-end layers, as illustrated in Figure 2. Additionally, each layer has its own resilience so that an outage in one area of the infrastructure will not impact the others. Firewalls between the Internet and network layer protect the network and front-end components from external threats, including Denial-of-Service (DoS) attacks.

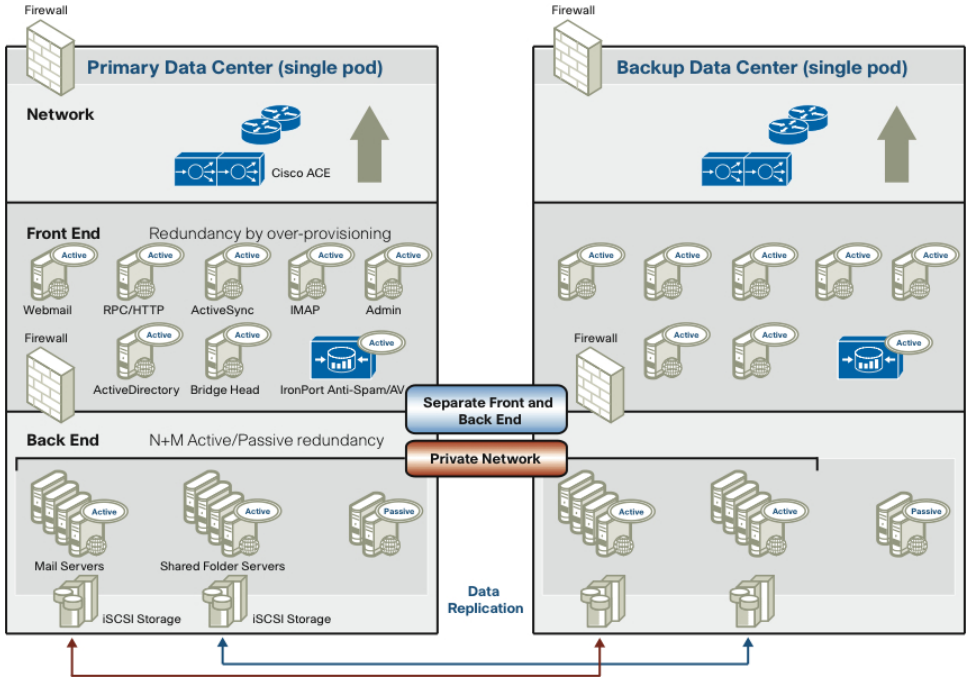

Figure 2: High availability redundant architecture

Separating functions across servers and layers provides superior performance, scalability, and security, and enables problems to be located and resolved quickly. The partitioned Cisco Mail architecture helps contain component issues within the layer and within the pod, which protects the remaining infrastructure and large groups of users from being affected. The separation of front-end and back-end components and networks combined with the use of firewalls and separate virtual LANs also protects sensitive customer data stored in the back end.

Network redundancy and failover

The network layer is an all-Cisco network that employs highly resilient Cisco Catalyst® switches and routers and Cisco Application Control Engine (ACE) global server load balancers. Network components are configured for connection redundancy and failover. The primary and remote datacenters inject the same virtual IP (VIP) into the Interior Gateway Protocol. Technologies such as route health injection (RHI) are used with both datacenters to deliver automatic failover independent of users' DNS caching. To guard against individual Internet Service Provider (ISP) service issues, Cisco uses multiple ISPs to route traffic to and from each data center.

Front-end architecture and redundancy

In addition to providing secure email access, front-end components perform mail filtering, routing and administrative duties. As traffic is received from the external network, it is filtered by Cisco IronPort™ email anti-spam and anti-virus appliances and then routed to the appropriate mail store by the mail transfer agent. After authentication, users can then access their messages using webmail, Microsoft Outlook, BlackBerry®, iPhone or Windows Mobile devices.

All front-end components are configured for active-active redundancy to deliver front-end high availability. In addition, each type of front-end server is assigned a separate VIP, with each VIP consisting of multiple servers. This approach effectively creates a dedicated, secure network for each traffic type, and enables component upgrades to be performed without disrupting services to end users.

From the external network, traffic is delivered to Catalyst switches, where ACE load-balancers intelligently route traffic between and among front-end servers based on application information and server availability. When ACE detects that a server is unavailable, traffic is redirected to other available servers in that VIP. ACE performs ongoing health checks to ensure that servers are working properly and removes malfunctioning servers from the pool.

The combination of server pools within VIPs that provide full redundancy for each traffic type and ACE load-balancing and monitoring features helps Cisco Mail provide continuous email access to users and administrative access to email administrators.

Back-end architecture and failover

The back-end layer of Cisco Mail consists of multiple mail and shared folder servers and storage pools that are firewalled from the front end. Because access is tightly controlled and the back end is not exposed directly to the external network, customer data stored there is highly secure.

Within storage pools, each storage logical unit number (LUN) stores data for a maximum of 1,000 users. By limiting the number of users in a LUN, problems can be identified and contained more quickly. Compared to some other mail systems that use very large unitary storage pools where localized issues can affect vast numbers of users, smaller Cisco Mail data stores ensure that only a small subset of users are affected in the unlikely case of an outage or data corruption.

Cisco Mail uses a combination of N+M active/passive cluster configuration, Cisco Global Site Selector (GSS) DNS, and Linux open source technologies for back-end high availability. If a primary mail server or shared folder server fails to respond, a standby server assumes the identity of the primary server and continues to serve data for seamless local failover. Cisco GSS appliances work with ACE load balancers and DNS infrastructure to continuously monitor server load and health in real time within each data center. In the event a primary data center outage or overload occurs, GSS quickly and transparently reroutes users to a geographically remote secondary data center. RHI directs traffic and serves data from whichever data center is available without user interruption.

A key contributor enabling Cisco Mail to provide rapid data recovery is the use of a standard Linux file system. This file system architecture provides simpler, more flexible and more secure data storage than relational database storage systems employed by other mail solutions, such as Microsoft Exchange.

Conventional relational databases can be more sensitive to data corruption, with a single point of corruption affecting thousands or even millions of users. By contrast, a file system consists of standard directory and file objects that do not rely upon each other for integrity, and that can be independently replicated. If an error or issue does arise, standard, time-tested tools are used for rapid recovery. In addition, even if data corruption does occur, it is likely to affect only a small fraction of messages or users, as compared to database corruption that can spread and ultimately bring down the entire mail system.

The benefits associated with limiting the number of users in data storage pools combined with the use of a Linux file system are clear:

- Issues affect only a limited number of users.
- Faster problem identification and resolution
- Data loss or leakage risk is minimized as customer data is stored in a defined system and not intermingled with others' data.

Denial of Service protection

Denial of service (DoS) and Distributed Denial of service (DDoS) attacks can pose a serious threat in SaaS environments, whether directed at network equipment or servers. DoS attacks attempt to prevent authorized users or organizations access to services or resources, which can negatively impact service availability.

The Cisco Mail architecture offers multi-layer defense against DoS and DDoS attacks. At the network layer, the Cisco Application Control Engine (ACE) load balancers are equipped with sophisticated firewalls that protect against DoS attacks attempting to disrupt network performance.

Cisco IronPort anti-spam and anti-virus gateways provide an additional layer of protection against SMTP-based attacks. IronPort provides both reputation filtering and reactive filtering to protect internal mail servers. Reputation filtering relies on real-time threat assessment using IronPort SensorBase® data to identify malicious senders and pre-empt DoS attacks by either rejecting or throttling connections before messages are accepted. High performance spam and virus scanners then work to stop fast-moving virus outbreaks and botnet attacks that penetrate the initial line of defense.

In addition to these device controls, the Cisco Mail highly-redundant, distributed Linux system architecture counters DoS attacks by introducing the high capacity and resiliency necessary to handle intense traffic spikes or anomalous usage patterns. Active site failover minimizes or eliminates service disruption in the event of a DoS attack against an isolated cluster of servers. Furthermore, geographically dispersed architecture makes Cisco Mail solution a difficult target for attackers.

Finally, application security assessment and third party source code reviews ferret out application weaknesses which result in DoS vulnerabilities.

Business Continuity planning

Cisco Mail supports your Business Continuity planning by helping to ensure that mission-critical email service is always available even in the face of potentially disruptive events, such as fire, flood, office relocation, or IT department personnel changes. It also helps you deliver on goals your organization may set for the two most common business continuity metrics: Recovery Time Objective (RTO) and Recovery Point Objective (RPO).

Cisco Mail– the right choice

Cisco Mail is architected to provide reliable, continuous service with no unplanned downtime. But in the event that you need support for your Cisco Mail service, you can count on the Cisco Mail Technical Support team to resolve

your issues quickly. Available 24x7 every day of the year, Cisco technical support staff will keep you informed continuously on the status of your issue.

All Cisco support staff, including the front-line team, is highly experienced and knowledgeable about email, networking, and the needs of enterprise customers. Cisco Mail support is not a call center and is focused solely on the needs of Cisco Mail customers. The team does not support internal Cisco corporate IT needs.

Making sure that your company's employees have uninterrupted email access is crucial to your success. High availability Cisco Mail architecture gives your company continuous email service and seamless disaster recovery. Round-the-clock monitoring by expert staff reduces your maintenance burden so that you can focus on value-producing projects rather than day-to-day email support.

By choosing Cisco Mail, your company can reap the benefits of the high availability, security, and scalability that, until now, was available only to the largest enterprises—without investing in expensive data center infrastructure and equipment, high-level consulting services, and ongoing maintenance.

Cisco Mail

Cisco Mail is a corporate-grade email service that overcomes the limitations of traditional email while bridging next generation web-based collaboration. Cisco Mail eliminates the high cost and burden of in-house email management and is natively compatible with Microsoft Office Outlook, with no plug-ins to maintain or disruptions to the organization. Companies of any size can take advantage of large capacity mailboxes, native BlackBerry (BES) and Microsoft ActiveSync support, high availability for everyone, and built-in advanced security features. The flexible Cisco Mail framework—scalable cloud architecture, extensible Linux-based platform, and AJAX webmail client—enables email to be tightly integrated with a variety of collaboration approaches.

Learn more about Cisco Mail.

Visit us online at http://www.ciscomail.com

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**