Cisco on Cisco: Cisco Unified Communications on Cisco Unified Computing System

How Cisco achieved scalability for voice communications at Research Triangle Park.

EXECUTIVE SUMMARY

Cisco IT

- High Tech, Cisco on Cisco
- Research Triangle Park
- 5000 employees

Challenge:

- Provide reliable voice communication services to rapidly growing number of users and increasing devices per user
- Scale Unified Communications (UC) infrastructure beyond the limits of its existing MCS server backbone
- Avoid any disruption in service during migration

Solution:

- Migrate existing voice clusters to UC on Unified Computing System (UCS) in two brownfield data centers
- Replace eight dedicated MCS servers with two UCS B-series blades and two UCS C-series

Results:

- Consolidated eight dedicated MCS servers to two B-200 blades and two C-210 Series, 50 percent loaded
- Reduced 300 servers for telephony to fewer than 100 UCS servers
- Lowered power consumption by 75 percent

A decade ago, supporting the voice communication needs of the Cisco workforce meant two things: connecting one physical device, such as an office phone, and one “soft” device, such as a laptop, for each employee on the network. As the number of users on the voice network has grown, so too has the number of devices they use. Workers today expect to do business anywhere, anytime, on any device. They need a phone at work and one at home, plus mobile support in between. They work from laptops, tablets, and video devices using multiple types of software phones: whatever it takes to help ensure they can communicate with colleagues, customers, and partners at all times.

Few of these workers think about what it means to support their voice requirements on the back end, that is, except for the team whose job it is to keep the infrastructure running at maximum capacity. With more than 220,000 phones alone to support, the situation became dire at the company headquarters in San Jose in the middle of 2008. When the voice infrastructure reached the ceiling of what its existing processors could support, it was time for a change.

“There were services that we could not deploy because we had run out of processing capacity, and we couldn’t add more physical servers because our cluster was already so large that is was beyond the Cisco Services architecture. Our only option was to limit the number of services we could offer to our users,” says Chuck Churchill, Unified Communications (UC) program director for Cisco IT.

Cisco® Unified Communications Manager Version 8.0 provided enhanced scalability capabilities, but only if run on a more powerful compute platform. Cisco IT found a solution in its own backyard: a virtualization strategy enabled by a migration to the Cisco Unified Computing System® (UCS).

Challenge: Explosive Growth in Users and Devices

Rapid growth in Cisco’s employee user base combined with new ways of working put additional processing demand on the IT department’s communication services. New devices and technologies increased compute requirements, because they deliver richer features with higher, more complex rule-based processing

Rigorous testing confirmed that UCS had what it took to support the company’s growing appetite for communication services. Cisco IT decided to migrate its entire voice communication system, more than 460 Media Convergence Server (MCS) servers arranged in 22 clusters for telephony, voicemail, telepresence, conferencing, and call center services, to the new Cisco UCS platform.

As a result of the migration, decision makers intended to reduce cost through lower space and power needs, improve service redundancy and recovery capabilities, and reduce the cycle time needed for capacity changes. They also wanted to demonstrate the capabilities of the UCS platform for running a 100,000+ employee communications system.

With a new strategy in place, the question became, “Where to begin?” Cisco IT needed to choose an appropriately large and busy data center to prove the reliability of UCS. Its Research Triangle Park (RTP) campus emerged as a promising candidate. The site has a high concentration of both UC and UCS-savvy IT engineering staff. And with 5000 users, the RTP campus helped ensure that the first deployment would be more than just an academic exercise. Adding to its appeal, RTP had been through previous technology field trials, so its user base had experience in adjusting to the deployment of new products.

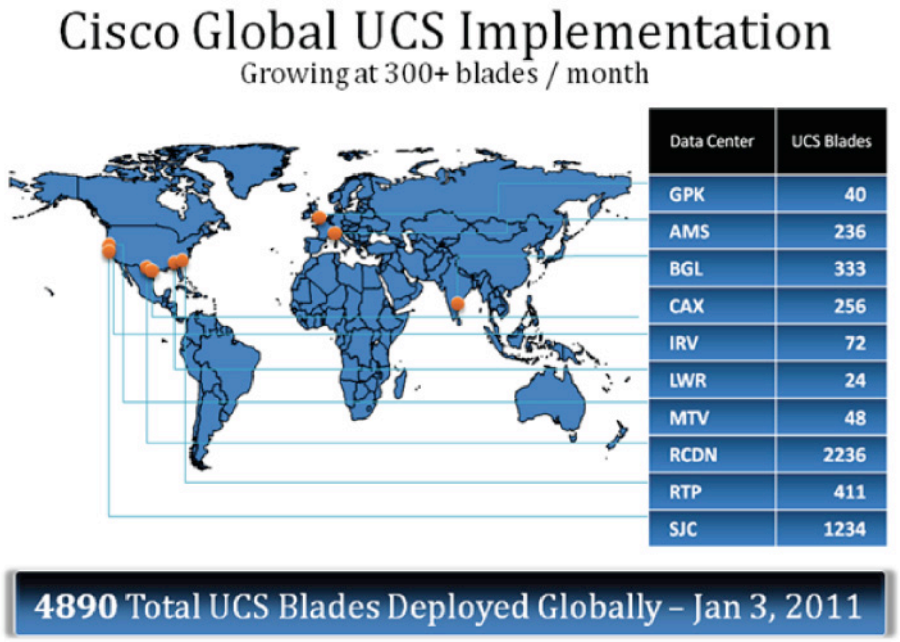
The RTP campus has two existing data center facilities that provide local infrastructure services for the RTP campus. Historically, these data centers also have provided geographical redundancy to the company’s two primary San Jose headquarters data centers for critical applications and services. As part of Cisco IT’s Global Data Center Strategy, all critical business applications services are moving to a pair of “active:active” data centers in Richardson, Texas, and a new greenfield data center is coming online in RTP to function as the disaster recovery site. Focusing on the voice clusters in the two brownfield data centers, Cisco IT had found a match for its first UC-on-UCS deployment.

> “The deployment of Cisco UCS in our major data centers provides the foundation of our private cloud architecture. This has been a transformational phase of our IT as a Service strategy. By deploying UC on UCS, we’re moving closer to our vision of delivering UC services to our business in a Communications as a Service model.”
>
> **— John Manville, VP, Cisco IT**

Cisco UCS: Standard Compute Platform Companywide

The Cisco voice team was not alone in deciding to move business-critical applications over to UCS. Support for hardware-level availability, simplified management, and unified I/O features have made Cisco UCS the standard compute platform for Cisco IT. It is foundational to the company’s data center strategy and core to its virtualization and cloud initiatives. Accordingly, Cisco IT has more than 5000 Cisco UCS blade servers deployed today in eight data centers across the globe, the largest concentration of which are in the Richardson, Texas primary data centers. Meanwhile, the company is in the process of migrating more than 2000 applications from legacy data centers and compute platforms to Cisco UCS.

Figure 1. Cisco Global USC Implementation



However, the plan to migrate UC services to Cisco UCS blade servers in an existing and constrained brownfield data center presented a number of unique challenges to the IT and facilities teams. For one thing, the teams needed free physical space to build out new Cisco UCS implementations in parallel to existing infrastructure, clearly a challenge when an existing data center is near or at physical capacity.

Constrained by Space in Brownfield Environment

Due to the lack of space and power, for the 18 months leading up to the UCS deployment, Cisco IT had been running strict policies to avoid any unnecessary infrastructure installation requests at RTP. All new installs had to be a net-zero effect on space, power, and cooling so as not to use up what little spare capacity remained. IT teams needed to plan for the decommissioning, consolidation, or virtualization of existing infrastructure before plans were approved to install the new UC on UCS infrastructure.

The strength of the data center floor gave another cause for concern. The existing raised flooring of the data centers was not strong enough to support the same UCS blade system design that was developed for greenfield data centers. In the end, Cisco IT modified the design with less densely populated racks as a cheaper and faster solution to retrofitting the floors.

Similarly, power, cooling, and network cabling all required careful consideration to help ensure adequate support for the UCS blade system design.

But the greatest challenge of all came from the users themselves. The RTP campus cluster hosts some of Cisco IT's most demanding clients. For example, the Cisco Technical Assistance Center (TAC) provides customer support to the Cisco customer base and is extremely intolerant of any service disruptions that could compromise its ability to provide an optimal customer support experience. Any changes to infrastructure that supports the TAC organization are heavily scrutinized before, during, and after execution by IT and business executives, because any service outage during the server migration could not be tolerated. Viewed from any angle, the pressure and stakes were high.

Solution: Capacity on Demand

After extensive planning and negotiating, Cisco IT agreed to proceed in two steps. It would upgrade half of the RTP servers, one half of one cluster, in a first wave, completed in June 2010, and the rest in a second wave, which is in progress now. To date, the four Unified Communications Manager nodes: the Publisher, the primary Trivial File Transfer Protocol (TFTP) server, and two call-processing servers that resided in one of the RTP data center facilities were migrated to a Cisco UCS B-series system. The failover TFTP server and two additional further call-processing servers remain on MCS servers in the other RTP data center site. These servers will be migrated to Cisco UCS C-Series servers in the near future.

Why is the C-Series ideal for this situation? The IT infrastructure in the older of the RTP data centers will soon migrate to a greenfield data center that is scheduled to come online in Raleigh later this year. In the interim, no UCS B-Series clusters will be deployed into the old data center, so any compute demand must be serviced by the UCS C-Series platform. The C-Series platform provides more flexibility to handle smaller compute footprints and is simpler to integrate into a brownfield data center.

Because it was not feasible to retain the existing networking configuration during the migration, the team had to manage the reconfiguration. This process involved developing tools to identify phones and devices that used hardcoded IP addresses for the UC TFTP servers in their configuration. Users with phones configured with hardcoded addressing for TFTP servers were notified by email multiple times ahead of the change to make the necessary changes to their devices.

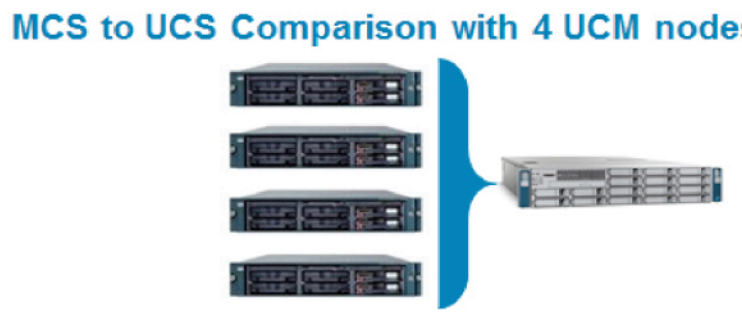
By the numbers: UC in Cisco IT

- 22 Unified Communication Manager (UCM) clusters
- More than 450 office sites supported globally
- More than 220,000 IP phones globally
- More than 2000 voice gateways globally
- Clusters vary from 5 server nodes to 19 server nodes
- Production UCM versions 7.13 / 8.02
- Outgoing server model standard: MCS 7845-H2, > 450 in production
- Remote site voice service resiliency supported with Survivable Remote Site Telephony (SRST) on the Cisco Integrated Services Router 3845
- Voicemail supported by 80 Unity Servers and 63 Exchange Servers
- UC services supported: Office and home IP phones; WebEx® conferencing; Telepresence®; extension mobility; single number reach; dual mode phones; presence; software clients

Results: Infrastructure Improvements All Around

With Cisco UCS, Cisco Unified Communications applications now run in a virtualized environment, which includes VMware, Cisco Unified Computing System servers, and Fibre Channel SAN storage. In moving to UCS, Cisco IT consolidated eight dedicated MCS servers to two B-200 blades and two C-210 Series, which were only 50 percent loaded. It reduced 300 servers for telephony to fewer than 100 UCS servers. And it has lowered power consumption by 75 percent.

Figure 2. MSC to UCS Comparison



Item	MCS	UCS
Rack Units	8 U	2 U
Power Consumption	1000 - 1200 Watts	250 - 300 Watts
Rack and Stack	2 Hrs	<1 Hrs
Cabling	20 (12 n/w + 8 pwr)	4 (2 n/w + 2 pwr)
UCM Install/Config	4 Hrs	4 Hrs + <1 Hrs VMWare

Although it is still early in the migration process, Cisco IT already has begun to realize the benefits of moving its Unified Communications applications to Cisco UCS. “From a platform perspective, the nodes on UCS are much faster than the MCS counterparts,” says C. J. Lagos, a Cisco IT engineer who has worked for the last three years designing, implementing, and supporting Cisco Unified Communications systems. “The difference is quite noticeable when installing software, booting up, and rebooting. The logistics of building a server are much easier, too. No more headaches trying to get a UCM ISO burned and into a DVD drive in a datacenter. Having complete control over this aspect is a major advantage.”

Consider the full list of results, Cisco IT has experienced at RTP:

- Improved redundancy and paved the way for future high availability support
- Increased deployment flexibility through reduced server provisioning and recovery times
- Lowered total cost of ownership with fewer infrastructure components to purchase and fewer elements to manage
- Gained scalability to add more software and services without increasing cooling, power, space, or cabling costs
- Consolidated management tools, so administrators can use familiar single tools or sets of tools, reducing operating costs
- Simplified installation and upgrades through service profiles, virtual machines, and centralized management

“Moving to UCS has allowed us to scale our communication infrastructure to meet the growth of the business,” says John Manville, VP, Cisco IT. Having compared the total cost of ownership (TCO) factors involved in supporting UC on UCS using virtualization with running UC on the MCS platform, Cisco IT expects to see a 19 percent TCO saving in year-one as a result of the reduced hardware footprint and physical data center resources required to run the same compute load.

Best of all, the change was invisible to RTP Campus users.

“It was extremely rewarding for the Cisco IT team to complete this migration without disruptions to the TAC voice communication services,” Churchill says. “And even more satisfying to receive thanks from the TAC business executives after the change was completed.”

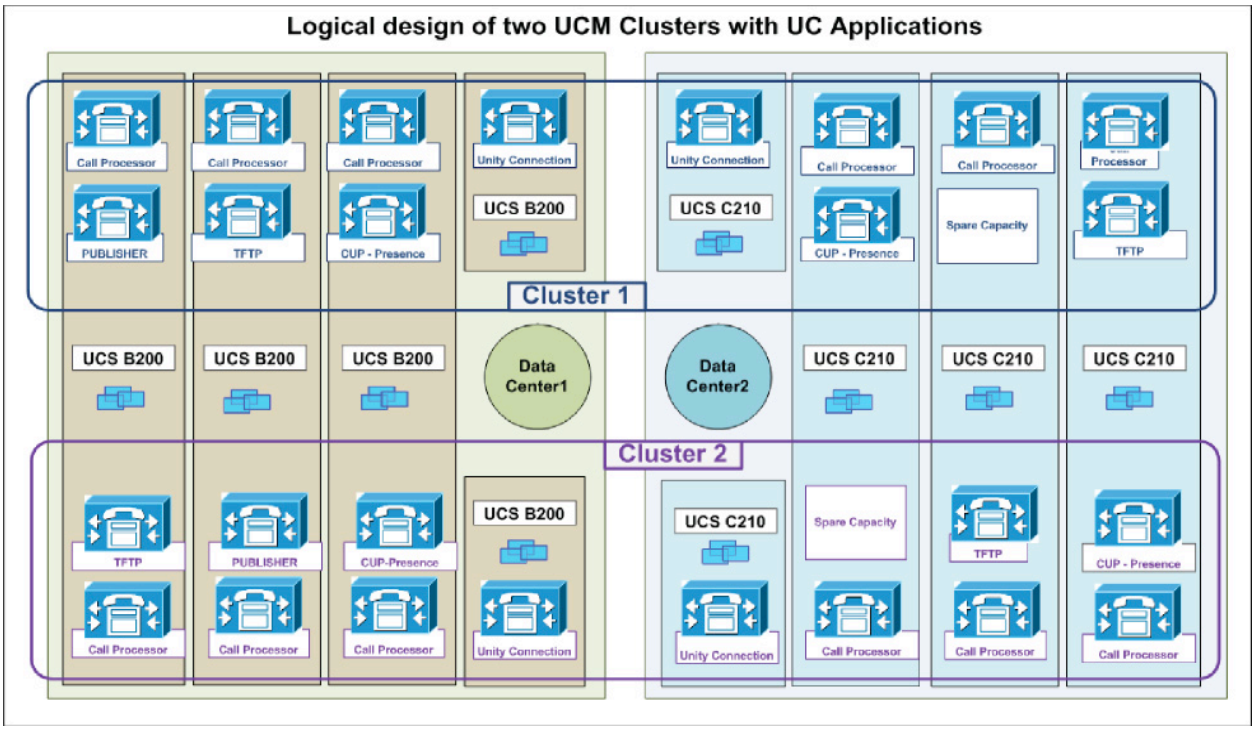
Next Steps: Completing UC Migration, and Preparing for Private Cloud

Cisco IT is in the progress of planning out the global migration of all UC servers to UCS. The next migrations will be focused on the second data center in RTP, to complete the full migration of all the RTP-based UC servers.

There are two clusters currently in RTP: the RTP campus cluster, supporting the second-largest Cisco campus, and the centralized call-processing (CCP) cluster. This CCP cluster supports all Cisco telephony traffic across the eastern United States and Central and South America. Both clusters, with the supporting UC application hosts, comprise a total of 33 servers. Migrating to the UCS platform will reduce the total number of physical servers by more than 50 percent to 15 (although there will still be the same number of virtual machines in the two clusters).

Good design practice for high-availability call processing requires servers within a cluster to be split into two physically diverse data centers. Cisco IT was able to maximize virtual-to-physical server density, while providing a high-availability design by sharing physical blades between the two clusters that are served out of the RTP data centers. A logical resiliency schema was employed to distribute the virtual servers across physical servers within a data center and between the two data center sites. A baseline resiliency schema for Cisco IT designs is illustrated below. The initial host distribution within each data center may be changed based on analysis of observed service usage and performance. This analysis should take place as part of regular performance management review processes. Virtualization makes it easy to redistribute the virtual server nodes between physical blades to optimize both system performance and utilization to achieve maximize the value on the infrastructure investment.

Figure 3. Logical Design of Two UCM Clusters with UC Applications



In addition to the RTP work, projects for two of the largest of Cisco IT's UC clusters based in the San Jose headquarters are well under way and scheduled for migration during the first half of CY2011. These two clusters alone support more than 30,000 users and more than 70,000 devices. The first San Jose cluster supports all Cisco telephony traffic for the western United States and Mexico and will migrate to the new data center in Texas. The second cluster will remain in the San Jose area, because it supports Cisco employees working within the 50 buildings that make up the Cisco headquarters site. This second cluster currently requires 19 servers and is the largest cluster Cisco IT operates, not counting the additional 15 servers supporting presence, management and other phone

services. Migrating to a UCS platform will enable Cisco IT to reduce this number of servers by more than 60 percent, from 34 servers to 13 servers in San Jose.

After the largest clusters are complete, the other clusters will follow through the end of CY2011. Cisco IT has 22 voice clusters in 12 sites around the world dedicated to CCP voice, IP Contact Center services, and Extranet voice services.

At the same time as it migrates to UCS, Cisco IT also is migrating its voicemail servers from Unity® Enterprise 7.0.2 to Unity Connection 8.5. Today, the Unity Enterprise voicemail service requires 186 servers in 14 locations worldwide, many of which are Exchange servers for voicemail storage. Migrating to Unity Connection 8.5 requires an estimated 32 servers, spread across the same 14 global locations. Migrating these servers from MCS to UCS does not provide Cisco with much of a reduction in physical servers, because the Unity Connection clusters are already rather small, and need a minimum of two servers per location for redundancy. (The 14 voicemail nodes will be exactly two servers per location, for a total of 28 UCS servers in all once the migration is complete.) However, the move to UCS for voicemail is motivated not only by the reduction in servers. An administrative benefit will also accrue from having all UC services and applications running on one standard operating system and virtualized platform.

UC on UCS Paves Way to Private Cloud

In the long term, the move to UC on UCS fits into the internal private cloud architecture that Cisco is deploying across all of its data centers. This strategy enables greater agility in the deployment of new services, management of service capacity, and retirement of obsolete services. The private cloud architecture offers new capabilities in terms of service resiliency and recovery that will provide greater service availability to users without the need for dedicated duplicated infrastructure.

“The deployment of Cisco UCS in our major data centers provides the foundation of our private cloud architecture,” says Manville. “This has been a transformational phase of our IT as a Service strategy. By deploying UC on UCS, we’re moving closer to our vision of delivering UC services to our business in a Communications as a Service model.”

For More Information

To find out more about Cisco UC on UCS, go to: http://www.cisco.com/en/US/netsol/ns1067/index.html.

To learn more about Cisco UC, go to: http://www.cisco.com/en/US/netsol/ns151/networking_solutions_unified_communications_home.html.

To learn more about Cisco UCS, go to: http://www.cisco.com/en/US/netsol/ns944/index.html.

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.**

Printed in USA C36-651082-00 2/11