x86 Blades: Shrinking the Branch Office

Facilitating Branch Office Infrastructure Consolidation

Overview

Multisite organizations are reducing the number of servers in their branch offices by moving applications to the data center. Yet, they continue to place a few essential applications locally because of performance, availability, or compliance requirements. By making use of x86 blade servers, these lean branch offices can lower equipment and operating costs, right-size and simplify infrastructure, and improve hardware provisioning and remote management.

Introduction

More than 200 years ago the idea of interchangeable parts improved manufacturing process efficiency and ushered the era of mass production. Today, albeit with a less-spectacular outcome, modular blade servers are increasing data center efficiency and enabling mass-scale infrastructure consolidation. According to Gartner, in 2009, less than a decade after their introduction, blade servers accounted for 15 percent of all servers shipped.[1] More importantly, Gartner predicts that blade servers will be the fastest-growing server segment, with 19-percent annual growth rate through the end of 2012 .[2]

Blade servers have gained market acceptance because they address a critical and growing need in the data center: the requirement to increase server densities without increasing the size, cost, and management complexity of the underlying infrastructure. Many of the benefits that blades offer, such as reduced cabling, rapid hardware provisioning, and management automation, are also applicable to small- and medium-sized businesses (SMBs) and branch offices. Several vendors have responded to this market gap by offering a compact, multipurpose blade system for SMBs and regional branch offices. However, the need for simplicity, lower space and energy consumption, and consolidated management is not unique to large branch offices. Small- and medium-sized branch offices could also benefit from a multipurpose blade system.

Today, organizations are centralizing applications and reducing infrastructure footprint in branch offices of all sizes. Although complete application centralization is the desired goal of many such initiatives, often it cannot be achieved because of limitations imposed by the WAN. A handful of edge services such as Microsoft Active Directory Domain Service (AD DS), print services, or mission-critical line-of-business applications must remain in the branch office. Intel x86 architecture (x86) blades are ideal for hosting these less-challenging workloads. The goal of this paper is to (1) describe the ongoing infrastructure transformation in the branch office, (2) outline how x86 blades naturally complement this trend, (3) highlight the business value of x86 blades in the branch office, (4) provide a brief overview of a Cisco Integrated Services Router Generation 2 (ISR G2) x86 blade solution for small and medium-sized branch offices, and (5) describe three use cases for the Cisco ISR G2 x86 blades.

[1] http://www.gartner.com/technology/media-products/reprints/cisco/170419.html
[2] http://www.gartner.com/it/content/868800/868812/blade_servers_the_fiveyear_outlook.pdf

The Lean Branch Office

Over the last decade the data center has undergone a transformation. Small, but numerous, regional data centers have been consolidated into a few, large facilities. Today, a similar transformation is occurring in the branch office. Applications and services are being moved from local servers into the data center and delivered over the WAN. Where appropriate, WAN optimization technologies accelerate data center-bound traffic to ensure satisfactory performance of centrally hosted applications. However, server centralization has its limits, and some applications still cannot leave the branch office for one or a combination of the following reasons:

- **Performance:** Applications that cannot tolerate the latency and bandwidth limitations of a WAN link because they either require fast response times or generate large amounts of data
- **Availability:** Applications that cannot tolerate the unreliability of a WAN link because they require continuous availability
- **Compliance:** Applications that cannot be hosted outside of the branch office or require a local copy of data because they must meet security, regulatory, or internal policy requirements

The following are examples of applications that have defied most centralization efforts and continue to be deployed locally at the edge of the network:

- **Edge infrastructure services:** Microsoft AD DS, print service, Dynamic Host Configuration Protocol (DHCP) server, Domain Name System (DNS) server, and file server
- **Line-of-business client-server applications:** Point-of-sale (POS) systems, bank teller in-office control points (IOCP), and electronic-medical-record (EMR) systems
- **Branch-office management software:** Configuration management, operations management, monitoring services, update and patching services, and backup and recovery services

Most of these mission-critical edge services require modest hardware resources. Moreover, you can use server virtualization to isolate every application in a dedicated operating system environment and consolidate the resulting virtual machines onto a few physical servers. Therefore, the now-leaner branch office typically requires one or two physical servers for hosting all local applications. Table 1 illustrates hardware requirements for a 5000-employee organization with 100 branch offices and an average of 50 employees per site.

Table 1. Example of Application and Services Resource Requirements for 50-User Branch-Office Services in a 5000-Employee Organization

Category	Application	Memory	x86 per core CPU Speed	Storage
Edge infrastructure services	Microsoft AD DS	512 MB	1.4 GHz	20 GB
	Microsoft Print Service	64 MB	1.4 GHz	1 GB
	Microsoft DHCP Server	128 MB	1.4 GHz	40 GB
	Microsoft DNS Server	128 MB	1.4 GHz	1 GB
	Microsoft File Server	8 GB	1.4 GHz	1 TB
Examples of mission-critical line-of-business applications	Microsoft Dynamics POS	2 GB	2 GHz	10 GB
	Bank teller IOCP	1 GB	2 GHz	40 GB
	Medium-sized clinic EMR system	2 GB	2 GHz	100 GB
Management systems	Microsoft System Center Configuration Manager	384 MB	1.4 GHz	5 GB
	Microsoft System Center Operations Manager	2 GB	2.8 GHz	20 GB

The reduction of branch-office applications to the essential minimum, modest local hardware requirements, and virtualization-assisted consolidation lower branch-office infrastructure footprint, equipment and operating costs, and management complexity. The resulting lean branch office provides a balance between a fully centralized architecture with its dependence on a slow and unreliable WAN link and a distributed architecture with its high cost and management complexity. Table 2 lists the benefits and trade-offs of the various architectures and shows how the lean branch office fits into the branch-office infrastructure continuum.

Table 2. Branch-Office Architectures

Attribute	Centralized	Lean	Distributed
Performance	Limited by WAN link bandwidth, latency, and cost	Optimized	High
Survivability	Limited by WAN reliability, redundancy, and cost	Optimized for mission-critical applications	Full
Equipment cost	Minimal	Low	High
Operating cost	Minimal	Low	High
Management	Simple	Varies	Complex
Infrastructure utilization	High	Medium	Low
Application time to market	Fast	Varies	Slow

In practice, the transition to lean branch-office architecture often translates into a deployment of a WAN optimization controller (WOC) and one or two rack or tower servers. With x86 server blades organizations can achieve additional reduction in cost and complexity while increasing infrastructure utilization and improving operational flexibility.

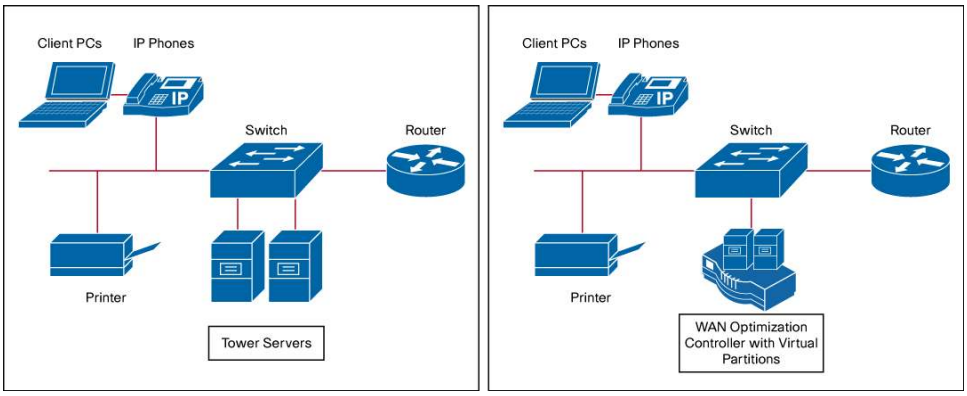
Optimizing Branch-Office Infrastructure with x86 Blades

In most organizations, revenue-generating and customer-facing employees are located in the branch office. These employees require a range of applications to effectively perform their day-to-day functions. The key to a successful server-reduction initiative is to determine which of these applications are required to reside in the branch office, which can be accessed over the WAN, how to preserve LAN-like performance for the centralized applications, and how to provide sufficient local infrastructure that can be managed remotely - all while creating a substantial cost savings. In short, branch-office infrastructure must be optimized to minimize costs and maximize simplicity without sacrificing application performance and availability.

There are two broad architectural alternatives for including application hosting capabilities in a lean branch-office infrastructure (Figure 1):

- **Tower or rack server(s) hosts virtualized or nonvirtualized applications:** This design provides a general-purpose platform for a wide variety of applications, flexibility to easily accommodate changes in application requirements, and hardware performance and capacity for any branch-office application. However, it creates wire-once infrastructure that requires onsite visits to be modified; consumes extra cabling, space, energy, and ancillary components; and leads to hardware underutilization because of overprovisioning and the standalone nature of each device.
- **Dedicated partition(s) on the WOC hosts virtualized applications:** This design option takes advantage of virtualization to consolidate lightweight branch-office applications, provides more flexible remote management, and reduces infrastructure and energy expenses. However, the small amount of available hardware resources limits the scalability and the type of applications that can be hosted, the management of applications requires product-specific expertise, and the lack of enterprise-class virtualization capabilities often leads to unpredictable service levels due to unregulated sharing of hardware resources.

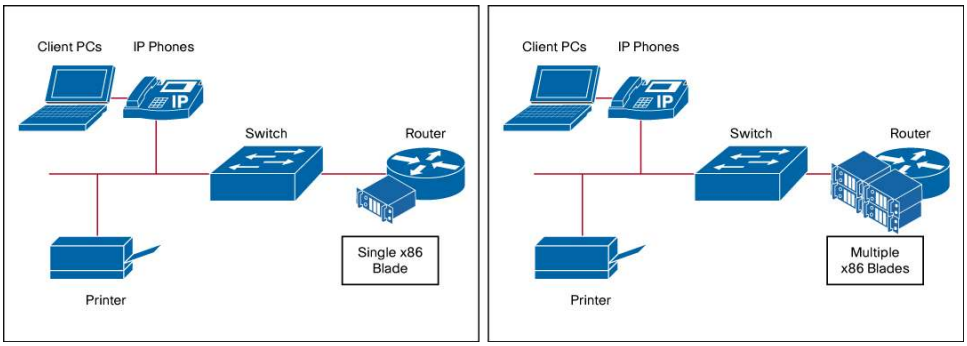
Figure 1. Traditional Lean Branch-Office Application Hosting Options



A general-purpose x86 blade housed in a branch-office networking device - which must be present in the branch office anyway - eliminates the worst of each while offering the best of both application hosting scenarios. x86 blades integrated into a networking device provide the following application hosting options (Figure 2):

- **Single dedicated general-purpose x86 blade integrated with a networking device:** This option provides performance, capacity, flexibility, and versatility of a tower or rack server while providing all-in-one device, cost-effective, and remotely configurable infrastructure.
- **Multiple dedicated general-purpose x86 blades integrated with a networking device:** In addition to the benefits listed previously, this setup provides scalability, application availability, and disaster recovery for a marginal increase in cost.

Figure 2. Networking and x86 Blade Integrated Application Hosting Options



These application hosting options bring the following IT benefits to the branch office:

- **Infrastructure right-sized for the needs of the branch office:** General-purpose x86 blades offer the performance, capacity, and versatility of a tower or rack server while sharing power, connectivity, and physical enclosure with other branch-office devices. The richer hardware configuration helps ensure that the infrastructure can scale with future business needs. The use of shared components and server virtualization increases utilization of the infrastructure. Improved scalability and utilization in turn generate a higher return on infrastructure investment.
- **Higher infrastructure flexibility and feature velocity:** By integrating networking and computing, a wire-once physical infrastructure that needs frequent onsite support is replaced by a soft-wired virtual infrastructure that is easier to manage remotely. Hardware provisioning is reduced to a quick plug-and-play process that a nontechnical branch-office employee can perform. The use of server virtualization further accelerates the velocity with which IT delivers new features to the branch office. Improved flexibility and feature velocity in turn allow faster response to new business opportunities or competitive threats.

- **Compact, simplified, all-in-one device infrastructure:** Server blades consolidated and integrated into a single branch-office networking device reduce physical space; lower energy consumption; improve energy management; reduce cabling and wiring; and eliminate duplicated components such as multiple power supplies, Ethernet ports, or patch panels. Moreover, consolidation reduces the number of hardware support contracts and simplifies management. Less equipment and complexity in turn creates infrastructure and operating cost savings.

Tables 3 through 5 provide in-depth explanation of each benefit.

Table 3. Right-Sized Infrastructure Benefits Details

Current setup	• Applications are running on rack or tower server(s). • Applications are running on a WOC hosted virtual server(s).
Challenges	• Branch-office rack or tower servers are underutilized because they tend to be overprovisioned. Underutilized servers waste power, and unused capacity diverts capital from other spending priorities. • The self-contained nature of rack or tower servers underutilizes generic components (for example, cooling and power subsystems) that could be shared across multiple devices. Duplicated components lead to a costlier infrastructure. • WOC virtual servers do not scale, and they support only light workloads because of limited hardware resources. • Management of applications on WOC virtual servers requires product-specific expertise. • WOC virtual servers do not use enterprise- and production-class virtualization, leading to unpredictable application service levels.
New setup	• Dedicated, general-purpose x86 blades are housed in networking device chassis.
Results	• Generic components (for example, power and cooling) are shared with other devices. • One-for-all-purposes x86 servers are replaced by x86 blades designed for the needs of the lean branch office. • All hardware on the x86 blade is dedicated to the application(s) running on the blade. You can select hardware configuration of the x86 blade based on scalability and performance needs of the target workload. • Generic x86 blades require the same skill set and use the same management tools as general-purpose x86 servers. • Enterprise- and production-class virtualization enables granular and precise control over allocation and sharing of hardware resources.
Benefits	• Infrastructure is right-sized to the needs of the lean branch office, with lower total cost of ownership (TCO) than a comparable rack or tower server solution. • Increased utilization of hardware maximizes return on infrastructure investment. • Scalable hardware footprint future-proofs the infrastructure and protects your current investment.
Costs	• You need networking devices that can house x86 blades. The Cisco 2900 and 3900 Series Integrated Services Routers (ISRs) offer this capability by default - at no additional cost.

Table 4. Infrastructure Flexibility and Feature Velocity Benefits Details

Current setup	• Rack or tower server(s) are wired into a specific branch-office design.
Challenges	• A failed server requires an onsite visit for replacement and rewiring by a skilled technician. • Additional or refreshed servers require an onsite visit for a skilled technician to install and wire. • Wire-once infrastructure requires onsite, high-touch effort to set up and modify.
New setup	• Plug-and-play x86 blades are configured through software for specific branch-office designs.
Results	• Plug-and-play x86 blades have a three-step installation process that unskilled staff can perform at the remote branch office: ◦ Remove cover, or unclamp and slide out the existing blade. ◦ Slide in a new blade. ◦ Clamp down the blade to secure it in place. • All power cables, network wires, and operator peripherals are eliminated. • You can remotely configure traffic and security policies for virtual servers. • You can remotely install a new operating system if hardware fails.
Benefits	• Remote management is faster because integration of networking and server virtualization enables fast configuration of VLANs and more granular implementation of security policies. • Physical deployment is faster because adding or replacing a server is a matter of merely sliding in an additional blade or replacing an existing blade. • Disaster recovery is faster because you can perform reprovisioning of failed systems remotely.
Costs	• You need networking devices that can house x86 blades. The Cisco 2900 and 3900 Series ISRs can house x86 blades by default at no additional cost.

Table 5. Simplified, All-in-One Device Benefits Details

Current setup	• Tower or rack servers are wired together with networking equipment into a single infrastructure.
Challenges	• There are many duplicate wires and ancillary components (for example, patch panels, power supplies, and Ethernet ports). • The physical footprint is large. • Power consumption is high. • You have multiple hardware support contracts. • The physical attack surface is large.
New setup	• The new setup has a compact all-in-one device infrastructure with modular networking and x86 compute blades.
Results	• There are no wires or cables between networking devices and servers. • The entire branch-office infrastructure is consolidated into a device with maximum dimensions of (H x W x D) 5.25 x 17.25 x 18.75 in. (133.35 x 438.15 x 476.25 mm). • Maximum power consumption per x86 blade is reduced to 50 watts. • You have a single hardware support contract for any number of x86 blades that fit into the networking device chassis. • The physical attack surface is reduced to a single device.
Benefits	• Your capital expenditures (CapEx) are lower because there are no additional power cables, connectivity wiring, patch panels, Ethernet ports, power supplies, or rack (or physical) space. • Your operating expenses (OpEx) are lower because you have lower electricity consumption and fewer hardware support contracts, and no onsite repair or installation visits are required. • Your company's physical security is improved.
Costs	• You need networking devices that can house x86 blades. The Cisco 2900 and 3900 Series ISRs that house the x86 blades involve no additional cost.

In summary, integration of x86 blades with a networking device provides right-sized, flexible, and simplified infrastructure for hosting of lean branch-office applications. IT organizations benefit from higher productivity and faster responsiveness. But the benefits of x86 blades in the branch office extend beyond operational efficiency of the IT department. Multipurpose x86 blades generate incremental business value that more than justifies investment in the technology.

Business Value of Branch-Office x86 Blades

When making the business case for x86 blades in the branch office, TCO or short-term return on investment (ROI) is the decision-making criteria that are easiest to quantify. Although a business case can be built on cost savings alone, faster time to market for new applications and investment protection provided by higher infrastructure flexibility can also serve as a compelling justification for the x86 blade technology. Following are the key business reasons for investing in branch office x86 blades:

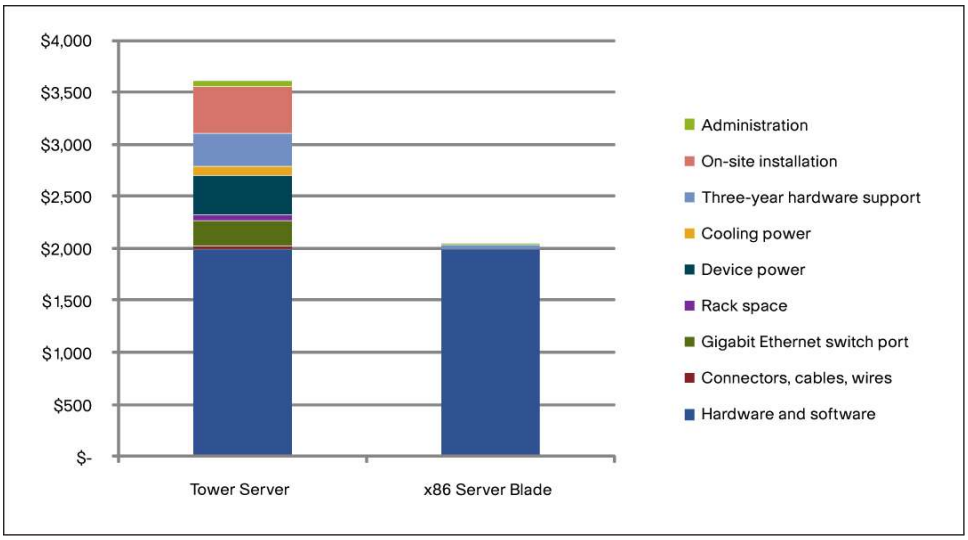
Lower Total Cost of Ownership

In comparison to tower or rack servers, x86 blades integrated into an available networking device provide the following cost savings:

- **Lower Capital Expenses:** No additional power cables, connectivity wiring, patch panels, Ethernet ports, power supplies, rack (or physical) space
- **Lower Operating Expenses:** Lower electricity consumption, less hardware support contracts, no onsite repair/installation visits

Figure 3 compares an equally priced x86 server blade and a tower server. In this example the TCO of the tower server is approximately 70 percent higher than the x86 blade integrated into an available networking device.

Figure 3. TCO Comparison of x86 Server Blade and Tower Server



Faster short-term ROI

Return on investment for a networking device integrated with x86 blades can be measured in numerous ways. The easiest metrics to quantify are savings on recurring operating expenses:

- **Lower energy consumption:** A typical x86 server blade draws 300 watts of power less than a typical tower server.
- **No additional hardware support contracts:** The hardware support services for the networking device typically cover the hardware support for the x86 blades.
- **Fewer onsite visits:** Upgrading or replacing an x86 blade is a plug-and-play process that an unskilled branch office employee can accomplish. A costly onsite visit is replaced by a shipping cost.

Table 6 compares an equally priced x86 server blade and a tower server. In this example the ROI is measured as annual operating expenses savings compared with the tower server. The payback time on investment is 3.9 years. This number represents the length of time required for the x86 server blade to pay for itself in operating costs savings.

Table 6. Operating Costs Savings Example on x86 Server Blade Compared to a Typical Tower Server

Operating Expenses Savings	
Energy savings (300 watts, $0.1 per KWh)	$790
Three-year hardware support (8x5 on site)	$310
One less onsite visit in 3 years	$450
Total OpEx Savings	$1,550
Cost of x86 Blade	$1,995
Annualized ROI	26%
Time to payback (years)	3.9

Faster time to market for new applications

Absence of physical wires and plug-and-play modularity of blade servers makes them ideal for the lean branch office, which typically lack any IT staff. In comparison to tower or rack servers, the faster deployment of x86 blade applications provides the following business advantages:

- **Faster response to new business opportunities:** Technology creates new opportunities for additional revenue streams (for example, Wi-Fi gateways and rich-media advertising displays). Deploying these new applications on server blades can reduce the roll-out time by months, in turn allowing faster recognition of the revenue.
- **Better protection against competitive threats:** Technology helps to differentiate one business from another (for example, self-serve kiosks and loyalty cards). Faster deployment of these applications allows for faster response to competition.

Future-proofing and investment protection

In comparison to virtualized applications deployed on WAN optimization controllers, x86 blades integrated into an available networking device provide higher investment protection:

- **Improved ability to accommodate for branch-office growth:** High-performance, high-capacity hardware of general-purpose x86 blades allows them to scale with growing business needs in the branch office.
- **Improved ability to reuse hardware for multiple purposes:** Some x86 server blades can be used as alternatives to rack or tower servers as well as platforms for deploying branch-office networking applications. If your business requirements change, you can quickly repurpose the x86 blade for another use without a costly onsite visit.

Cisco Unified Computing System Express x86 Blade Solution

Cisco Unified Computing System™ Express (UCS Express) is a converged computing, virtualization, and networking infrastructure for hosting WAN edge applications in the lean branch office. The Cisco UCS Express integrates multipurpose x86 blades and VMware vSphere virtualization with networking features of the Cisco IOS® Software and houses them in the second-generation Cisco ISR G2. A Multigigabit Fabric (MGF) backplane switch interconnects all x86 blades, networking modules, and other branch-office devices and facilitates integration of Layer 2 and Layer 3 networking features with VMware vSphere hosted virtual servers. The Cisco UCS Express solution enables multiple virtual instances of Microsoft Windows Server to run on dedicated general-purpose x86 blades directly inside the Cisco ISR G2.

The Cisco UCS Express solution is best suited for multisite organizations with centralized IT infrastructure that need to host a small number of essential applications locally in the branch office. Unlike rack or tower servers, the Cisco UCS Express combined with the multiservice Cisco ISR G2 provides an x86 blade system consolidated into a single branch-office device. Unlike appliance-based application hosting, the Cisco UCS Express with VMware vSphere Hypervisor provides enterprise-class virtualization hosted on high-performance, feature-rich hardware.

The Cisco UCS Express is composed of the following elements:

- **Cisco Services-Ready Engine (SRE) Multipurpose x86 Blades:** Designed to meet the requirements of lean branch-office applications, Cisco SRE x86 blades are compact, versatile, high-performance, and high-capacity servers. They offer a multicore processor, multiple storage drives, hardware Small Computer System Interface over IP (iSCSI) acceleration, Redundant Array of Independent Disks (RAID), and multiple Gigabit Ethernet interfaces. The most feature-rich configuration consumes a maximum 50 watts of power and fits into the Cisco 2911 ISR and higher without consuming any additional physical space. The Cisco SRE x86 blade operates independently of host router resources, shares only the power supply and cooling with the router, and connects to the routing engine through two interfaces: direct Layer 3 connection and Layer 2 connection through the MGF backplane switch.

- **VMware vSphere Hypervisor™ (ESXi):** The Cisco UCS Express combines the x86 blade characteristics of the Cisco SRE with the hardwarelike reliability and performance of the VMware ESXi into a server virtualization platform for the lean branch office. A jointly developed solution from Cisco Systems and VMware, the bare-metal hypervisor facilitates the consolidation of all branch-office network and application services into the Cisco ISR G2.
- **Cisco ISR G2 Multigigabit Fabric (MGF) Backplane Switch:** The MGF switch provides multigigabit Layer 2 connectivity between the Cisco SRE x86 blades and Cisco EtherSwitch® modules, facilitating gigabit-rate connectivity between the x86 blades and devices on the branch-office LAN. Virtual servers running Microsoft Windows Server can now directly take advantage of router and switch features such as VLANs, security zones, access control lists (ACLs), firewall, and intrusion prevention systems (IPSs) without affecting performance.
- **Microsoft Windows server hosted on VMware ESXi:** The VMware ESXi hypervisor allows one or multiple instances of Microsoft Windows Server to run directly on the Cisco ISR G2. This capability helps organizations with the Cisco ISR G2s to consolidate Windows applications and core Windows services such as Microsoft DHCP Server, DNS Server, and AD DS into a single device. Cisco UCS Express combined with the Cisco ISR G2 creates the industry's first and only solution for hosting all branch-office applications and services such as routing, switching, security, voice, video, wireless, computing, storage access, server virtualization, Microsoft Windows core services, and line-of-business applications.
- **Cisco Integrated Management Controller Express (IMC Express):** The Cisco IMC Express is used for management and monitoring of Cisco SRE blades through web GUI, command-line interface (CLI), or Intelligent Platform Management Interface (IPMI). The IMC Express runs on an embedded engine within the Cisco ISR G2 motherboard and thus can provide services in case the Cisco SRE blade experiences any major hardware failure or a system crash. The IMC Express is used for initial configuration of the server and to troubleshoot any problems in server operation.

In summary, Cisco UCE Express integrated with the Cisco ISR G2 provides a compact all-in-one device for hosting edge services on dedicated x86 blades. The solution is best suited for small and medium-sized branch offices with centralized applications and lean branch-office infrastructure.

Cisco UCS Express Use Case 1: Core Windows WAN Edge Services

An organization decides to implement Microsoft AD DS, DHCP Server, and DNS Server locally in the branch office for these reasons:

- **WAN reliability:** The WAN is unavailable occasionally, and a backup link is cost-prohibitive; or a second service provider is not available in a specific geographic area.
- **WAN performance:** Latency between the branch office and the nearest data center is significant and branch-office users generate constant queries to Microsoft AD DS, DHCP Server, and DNS Server.

The branch office experiences productivity slowdown when these services either are not available or perform slowly. With Cisco UCS Express, core Windows edge services can be implemented in the following ways:

- Run each service as a dedicated virtual server.
- A group of services can be co-located and run on a virtual server.

(Actual implementation will depend on the needs and requirements of the individual organization.)

This solution decreases the infrastructure footprint at the branch office.

Cisco UCS Express Use Case 2: Local Point-of-Sale Server

A retail chain cannot tolerate any downtime in its in-store POS system during shopping hours. Therefore, the retailer wants to place a POS server in the branch office to process local transactions and later synchronize the in-store inventory with the central inventory management system. With Cisco UCS Express, the local POS server can run on a dedicated virtual server hosted on VMware ESXi. This solution conserves valuable physical space in the store.

Cisco UCS Express Use Case 3: Local Autodesk AutoCAD File Share

A global engineering firm with regional branch offices requires several copies of Autodesk's AutoCAD program to run in each branch office. The AutoCAD files are relatively large, and frequent updates to the files make them time-consuming to update if stored in a central file share. Therefore, the firm wants to implement dedicated file servers at the branch offices. With Cisco UCS Express, the file share can be created on a dedicated virtual server hosted on VMware ESXi and periodically replicated to the central site. This solution reduces complexity of the remote infrastructure.

Summary

Most multisite organizations today are centralizing their branch-office applications. However, because of performance, reliability, and compliance reasons, few business-critical applications remain in the branch office. General-purpose x86 server blades integrated into available branch-office networking devices offer an ideal platform for hosting such applications and provide the following benefits over rack or tower servers:

- Lower infrastructure and operating costs
- Right-sized and simplified branch-office infrastructure
- Plug-and-play hardware provisioning with remote configuration

For more information visit: http://www.cisco.com/go/ucse/.

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 C11-613183-01 05/11