The Journey to Unified Computing

An ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) White Paper
Prepared for Cisco

April 2010

EMA™ IT MANAGEMENT RESEARCH, INDUSTRY ANALYSIS AND CONSULTING

Table of Contents

Executive Summary

The demands of today's data center are vastly different from those even five short years ago. Advances in virtualization have finally brought the vision of a data center that is truly adaptive to changing business needs within grasp. At the same time, virtualization has also forced a sea change in data center management. It is no longer acceptable to run various aspects of compute, storage and networking domains as separate silos. All three technology "stacks" must now be managed by a centralized, policy-based orchestrator that is virtualization aware, automatically determining the right mix of hardware and software required in order to meet a given business need.

Virtualization has also introduced additional challenges to the data center in terms of capacity optimization. Many servers today are no longer monolithic, single-purpose application servers constrained by processor power. Thanks to virtual machines (VMs), servers now run multiple VMs in order to maximize hardware utilization, and advances in virtualization now allow VMs to be easily moved from machine to machine. This has resulted in total available memory now becoming the key resource constraint—often long before compute power.

Cisco's Unified Computing System, or UCS, is an innovative set of technologies that set the stage to deliver on the requirements of the data center—today and tomorrow. Organically designed and built from the ground up by Cisco in partnership with industry leaders like BMC, Citrix and Intel, UCS shows high promise for delivering on Cisco's "Data Center 3.0" vision while allowing customers to deploy the solution incrementally while leveraging existing multi-vendor investments and obtaining a solid ROI from each stage in the process.

Cisco's Unified Computing System, or UCS, is an innovative set of technologies that set the stage to deliver on the requirements of the data center—today and tomorrow.

Introduction

Economic conditions, coupled with ever-increasing competitive pressures, continue to drive IT to increase data center efficiencies while decreasing overall expenditures. "Green IT" initiatives that aim to decrease an organization's carbon footprint by reducing data center power and cooling requirements conveniently align with corporate financial goals by saving money and increasing efficiencies in the process.

Virtualization has emerged as a core technology for achieving these goals, allowing IT to simultaneously increase efficiencies while reducing expenditures. Virtual storage, for example, removed the necessity for large data stores on each discrete server and drastically simplified the backup and recovery process, reducing both capital and operational expenditures.

Virtual Machines, or VMs, encapsulate an entire business application environment inside of a self-contained package that can be easily and quickly moved around the data center, abstracting much of the application dependencies from the server hardware. VMs are emerging as the new virtualization "atomic unit," serving as the cornerstone of many enterprise and service provider virtualization strategies. Vendors have responded by replicating VM technologies originally designed for mainframes, including hypervisors that serve as an intermediate abstraction and management layer between VMs and the underlying hardware. VMware in particular has led the industry with its VMotion technology, which provides the ability to move running VMs from server to server without service interruption.

Many IT organizations, faced with an unmanageable morass of virtualization technologies, attempted to create their own mechanisms to manage and automate the various silos of virtualization. These organizations, which attempted to stitch together solutions from multiple vendors, generally experienced poor results that made it extremely difficult to derive full value from their investments.

What is needed is an integrated architecture to span the virtualization silos and provide centralized command-and-control across the data center. This architecture also needs to be capable of defining and implementing policies, capturing and storing system state, and automating many of the routine management tasks that plague data center personnel on a day-to-day basis. A system based on this architecture automatically determines the correct mix of resources required to support an application running in the data center, and as business demand ebbs and flows, would be capable of intelligently deciding which resources to allocate to which applications, moving VMs around the data center and provisioning hardware as dictated by pre-defined policies.

In this perfect world, a harmony would exist between meeting business objectives and maximizing the cost/performance ratio of data center assets, keeping performance to a maximum while extracting every available dollar of business value from data center investments. This is the vision of a "private cloud," which until recently, was more vaporware than reality.

Cisco has invested heavily in taking a bold step toward achieving this vision through its "Data Center 3.0" initiative. Unified Computing, a major component of this initiative, was introduced in March 2009. It included many lessons learned from examining mistakes made by other vendors in the past and improving on them. Examples of these mistakes include "accidental architectures" resulting from multiple management stacks being bolted together, utilization of closed, proprietary technologies and interfaces, plus limited memory and network I/O capabilities that plagued many early attempts at integrated data center architectures.

The Cisco Unified Computing System, or UCS, is a well-designed hardware and software architecture that effectively combines management of virtualized compute, network, and storage access with a centralized policy-based manager (known as the Unified Computing Manager) that gets Cisco even closer to fully realizing the Private Cloud vision. This paper will examine Cisco's Unified Computing System as a solution to the requirements of today's highly complex data centers and will describe the journey that organizations that invest in UCS can expect to experience.

The Journey to Unified Computing

Many organizations delayed hardware refresh cycles over the past few years due to economic uncertainties. Today, these organizations now face long overdue upgrades, and they need to ensure that every dollar spent on upgrades will yield substantial business value.

Cisco's Unified Computing System (UCS) strategy is designed to be gradually deployed in three manageable stages, each of which provides rapid and substantial business value.

Cisco's Unified Computing System (UCS) strategy is designed to be gradually deployed in three manageable stages, each of which provides rapid and substantial business value. The first step, deploying virtualization-optimized server hardware, enables server consolidation. The second step involves establishing a unified network fabric, fully virtualizing discrete network stacks and standards across a common high-speed,

low latency network fabric. The third and final step involves deploying a centralized, policy-based management system that is capable of managing not only the technologies deployed in the first two steps, but other "legacy" systems that will continue to exist across the data center.

UCS combines the latest advances in virtualization, network, server and management technologies in a way that may well revolutionize the data center in the same fashion that Cisco revolutionized networking in the data center (and beyond).

UCS combines the latest advances in virtualization, network, server and management technologies in a way that may well revolutionize the data center in the same fashion that Cisco revolutionized networking in the data center (and beyond). As one might expect, tightly integrated solutions like UCS do require a certain amount of UCS investment before realizing the full potential of the architecture. While UCS does require Cisco UCS server and networking hardware, it is far from a 100% lock-in strategy, as Cisco consciously designed UCS around a number of open standards to ensure integration with incumbent data center investments. Cisco also worked closely with a number of industry leaders to design UCS, including Intel, VMware and BMC, ensuring UCS compatibility with a wide variety of third-party technologies.

UCS Building Blocks

As mentioned previously, UCS was designed to be deployed incrementally so that each stage can provide substantial business value on its own. The "building blocks" that form the foundation of this strategy are virtualization-friendly servers, unified network fabric, and a centralized UCS manager. It is important to note that, unlike many competitive offerings, each of the UCS building blocks can be deployed independently and in parallel with existing technology investments.

UCS Servers

The first cornerstone in the UCS foundation is a set of innovative, virtualization-friendly servers. All UCS servers share a number of common characteristics, including the ability to communicate via *Cisco Unified Fabric* for high-speed network access, support for *VN-Link Virtual Interface Cards* that abstract virtual machine network configuration from the physical hardware, and *Cisco Extended Memory* technology to provide two to three times the amount of physical memory than competitive offerings.

Figure 1: Cisco B-Series Blade and C-Series Rack Mount Servers

Cisco's UCS Servers are available either in a blade/chassis configuration, known as "B-series," or in a rack mount configuration, known as "C-series." UCS servers support Cisco Unified Fabric, described below, providing high-speed (10 Gbps) interconnects to LAN, storage and compute networks via a single media type. They also leverage Cisco Extended Memory Technology, which allows UCS two-socket servers to support as much as 384GB of RAM and four-socket servers to support up to 512GB.

UCS servers also embed VN-Link virtual interface cards, providing a direct, fully managed connection between VM virtual Network Interface Cards (vNICs) and VN-Link enabled Cisco switches, via the Cisco Unified Fabric.

One benefit of all of this virtualization is the ability to abstract the hardware from the applications running on it, all the way down to the network interface. When a new server is added to the system, UCS Manager virtually eliminates manual configuration, automatically detects the new hardware and provisioning it according to pre-defined policies. This is true "bare metal" provisioning.

Unified Fabric

Cisco's Unified Fabric is a huge leap forward in data center networking technology. Classic data center connectivity required separate physical networks for LANs (Ethernet), Storage (Fibre Channel), and management (such as server clustering). Add second redundant adapters for each connection type, and each blade or rack mount server can demand up to six ports—a veritable cabling nightmare. Add to that all of the upstream switching and bridging technologies required to make all of this work together, plus the extra power and cooling required, and it's easy to see why data center networking expenditures are so high.

Cisco Unified Fabric is an innovative "wire once" approach based on the Fibre Channel over Ethernet (FCoE) standard that combines support for Fibre Channel and Ethernet over a single high bandwidth, low latency 10 Gbps fiber media. A major benefit of Unified Fabric is that it allows server I/O configurations to change over time without the need to change physical cabling, a tremendous capital and operational cost saver. UCS Manager, described below, provides a centralized control, configuration and management point for all UCS devices connected via Unified Fabric. The cost savings multiply quickly from decreasing server cabling from six ports per server to two, plus equipment savings from the reduced network hardware requirements, plus lower ongoing maintenance costs.

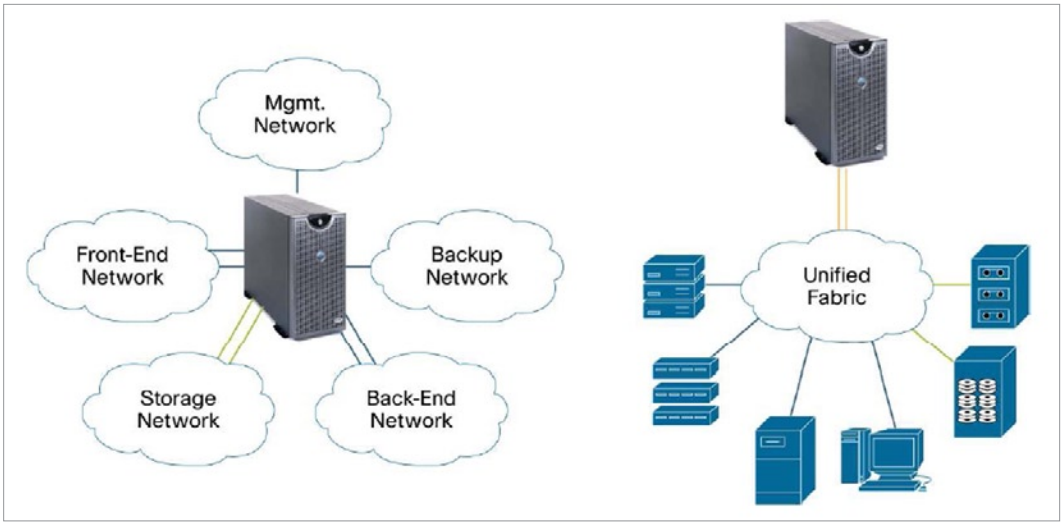

Figure 2: Classic Server Connectivity Compared with Cisco Unified Fabric

VN-Link

Cisco VN-Link is the virtual equivalent of using a cable to connect a physical server Network Interface Card (NIC) with a network port of an access-layer switch. The Cisco Nexus 1000V is a software switch on a server that delivers virtual networking capabilities to VMs hosted on that server, providing policy-driven virtual machine network capabilities. A major benefit of VN-Link is that when VMware VMotion moves a VM from server to server, network configuration also automatically moves—without service interruption or manual reconfiguration. In addition to abstracting the VM's physical network configuration, VN-Link also provides deep packet inspection capabilities all the way down to the VM. Think of it as a way to "sniff" virtual networks exactly the same way as hardware networks. Organizations that deploy VN-Link along with their virtualization servers will realize reduced network configuration and troubleshooting costs, along with higher overall server availability during VMotion events.

Extended Memory Technology

Many servers become memory-bound long before they become CPU-bound today, particularly as the popularity of memory-intensive applications like virtual machines and Online Analytical Processing (OLAP) continue to grow. Organizations that wish to support multiple VMs per server such as service providers or large enterprises deploying private clouds need to maximize the total amount of RAM per server while minimizing hardware and software licensing costs. Traditional two-socket Xeon servers support a maximum of 96GB of RAM, and prior to Extended Memory Technology, the only way to obtain more RAM was to move to a more expensive four-socket server, driving up hardware and software licensing costs.

Cisco developed and patented Extended Memory Technology to solve this problem. It allows four DIMMs to appear as a single DIMM to the Xeon processors, quadrupling the addressable memory. This translates to a maximum of 384GB of Samsung 40nm class, DDR3, high-efficiency RAM per two-socket server using 8GB DIMMs, a four-fold increase over other traditional two-socket servers, plus a low-cost option that supports 192GB using economical 4GB DIMMs. Cisco can also support up to 512GB of RAM in their new C460 four-socket server, which utilizes the new Intel Nehalem-EX architecture. Cisco Extended Memory Technology is available on both Cisco "B-class" blade and "C-series" rack mount servers.

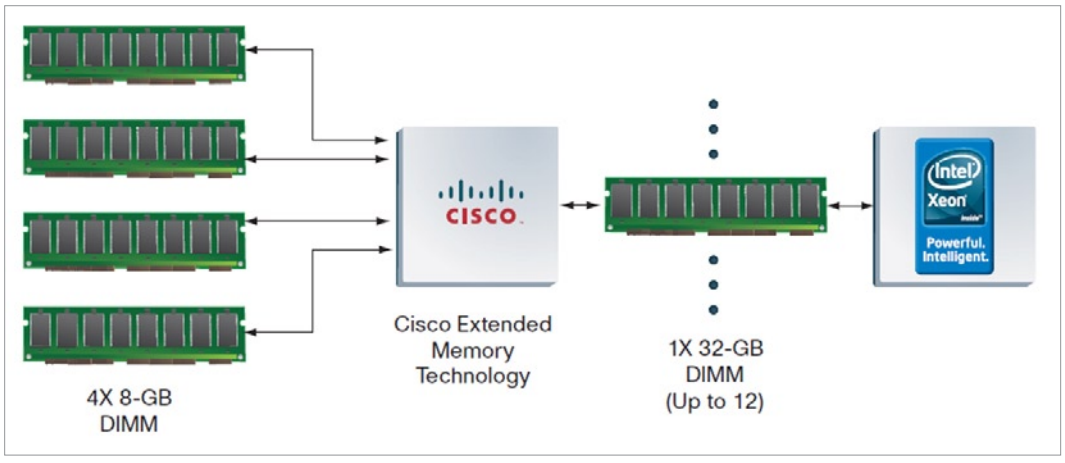

Figure 3: Cisco Extended Memory Technology

UCS Manager

In addition to the substantial hardware improvements described above, Cisco created a new UCS Manager package that ties everything together via a "single pane of glass" console. UCS Manager is more than just an orchestration server, however, as Cisco also embedded UCS management components directly into UCS servers (chassis, servers and network components), Unified Fabric, and VN-Link virtual network adapters. It is important to understand that this is more than just adding a bunch of management agents to the servers—UCS management components are built directly into the hardware. UCS Manager acts not only as a high-level orchestrator and automation tool, but also as a unified device manager for the entire UCS system. It also contains the centralized repository of system state, enabling UCS hardware components to be truly stateless—critical when combining hardware into resource pools that to enable dynamic, policy-based provisioning.

It is important to understand that this is more than just adding a bunch of management agents to the servers—UCS management components are built directly into the hardware.

UCS Manager also integrates with a number of third-party tools, including BMC BladeLogic for bare-metal provisioning, enterprise systems management suites, and Configuration Management Databases (CMDBs). This process is simplified because UCS Manager was built with a GUI management interface plus fully documented, open Command-Line (CLI) and XML-based Application Programming Interfaces (APIs).

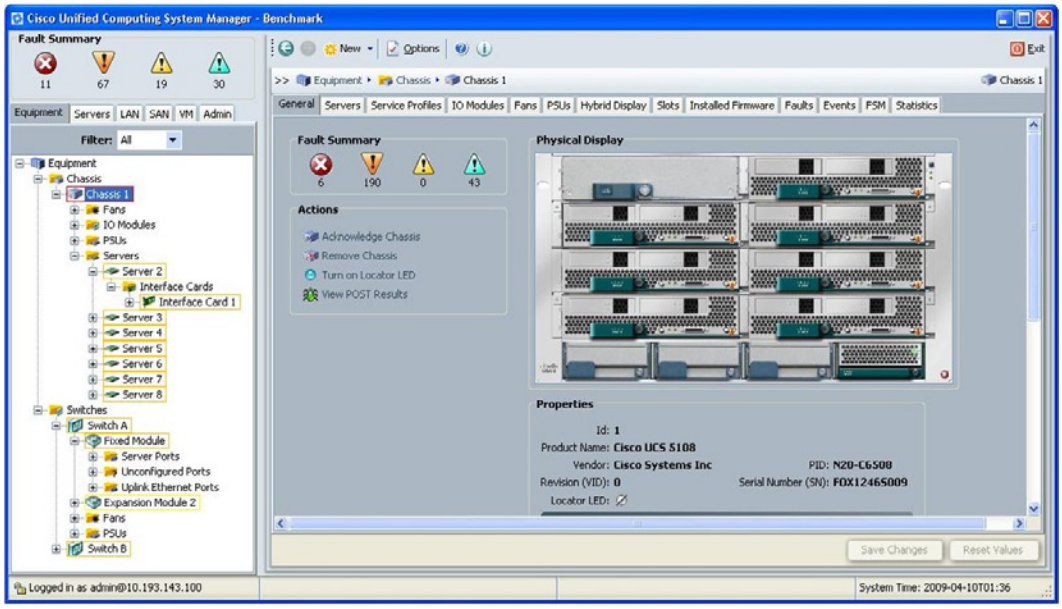

Figure 4: Cisco UCS Manager

UCS Service Profiles

UCS Service Profiles are the glue that ties UCS together, automating the configuration of bare metal servers and network connectivity and wrapping it all up with portable policies. Traditional management approaches, which treated hardware components as individually managed elements, required a tremendous amount of manual care and feeding. UCS Service Profiles solve this problem by serving as a centralized repository of system policy and configuration for all Unified Computing devices. This is significant, as they abstract all configuration and state information from the hardware, critical when enabling a truly virtualized environment.

Traditional approaches, which were unable to manage application environments holistically, managed system state on a component-by-component basis. Servers were classified into specific application pools, such as Web servers or databases, and extra "hot" spares were assigned to each pool for failover and demand relief purposes. For example, a traditional application stack with a Web server, database, storage, and VMs required a minimum of one extra server per server type for failover, plus extras to serve during times of peak capacity—a total of 7 extra servers that generally went unused. Service Profiles relieve the need to classify server hardware by application, allowing spares to be placed into a generic server resource pool that can be provisioned at a moment's notice to serve as any server type. Three spares can therefore do the work of the 7 above, a four server reduction that translates to 22% CapEx savings.

UCS Service Profiles provide a neat, portable package of state and configuration data that travels with the application around the data center. This eases the process of automatically moving applications from server to server and considerably reduces the amount of manual intervention required.

EMA Perspective

Cisco has continued its tradition of bringing disruptive technologies to market through its Unified Computing Strategy. Cisco wisely designed UCS to allow customers to incrementally and modularly deploy UCS components, realizing business benefit at each stage in the process. This strategy, clearly designed to expand Cisco's data center footprint from networks to servers and management, will require a long-term commitment from its customers in order to derive full benefit. Customers that choose to gradually invest in portions of the UCS "stack," however, will likely still receive high business benefit. In Cisco's defense, it would be impossible to pull off the high levels of integration and functionality provided by UCS without having complete and total control over the hardware. To Cisco's credit, they worked diligently to design UCS around many open interfaces and standards. For example, the UCS Manager API is an open, published, bi-directional XML interface and SDK, freely available through the Cisco Developer Network.

UCS is clearly well suited for large enterprises and service providers, and as with most large capital expenditures, it will require time before customers realize full returns on their investments. Fortunately, many of the UCS components can be gradually implemented through regular hardware refresh cycles, as opposed to requiring "fork lift" upgrades, and a rapid ROI for certain aspects of UCS, such as Unified Fabric, is highly probable. When compared with competitive offerings, particularly in terms of total cost of ownership at scale, UCS is likely less expensive in the long run.

Enterprise Management Associates (EMA) believes that Cisco customers that choose to invest deeply in the UCS solution will reap substantial total cost of ownership benefits. UCS stands to reduce hardware costs in terms of less server and network CapEx to support the same workload as competitive systems, lower operational costs in terms of providing a highly automated system that requires much less manual cabling effort, plus power savings due to reduced network hardware requirements and highly efficient server hardware design. The virtualization-aware characteristics of UCS also will likely increase overall system availability and performance, allowing data centers to become much more adaptable to changing business demands.

Enterprise Management Associates (EMA) believes that Cisco customers that choose to invest deeply in the UCS solution will reap substantial total cost of ownership benefits.

UCS is well positioned to be a highly disruptive large enterprise data center technology, and organizations already invested in Cisco network fabric, plus others that are refreshing older data center hardware, should take a hard look at Cisco. UCS could very well turn data center computing upside down in the next few years, and with Cisco's ~$35B war chest fueling future development, combined with a legion of loyal customers that will make the UCS investment, the sky (or the Cloud) is the limit.

About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that specializes in going "beyond the surface" to provide deep insight across the full spectrum of IT management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help its clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise IT professionals and IT vendors at www.enterprisemanagement.com or follow EMA on Twitter.

Corporate Headquarters:
5777 Central Avenue, Suite 105
Boulder, CO 80301
Phone: +1 303.543.9500
Fax: +1 303.543.7687
www.enterprisemanagement.com

2079.042710