Expert Reference Series of White Papers

Data Center Basics: Cisco Unified Computing System

Data Center Basics: Cisco Unified Computing System

Joe Rinehart, MBA, CCIE #14256, CCNP/DP/VP

Introduction

I began my career in Cisco networking in the traditional Routing and Switching technology areas. At the time, this was one of the foremost areas of engineering expertise, prompting me to pursue the Cisco Certified Network Associate (CCNA), the Cisco Certified Network Professional (CCNP), and finally the Cisco Certified Internetwork Expert (CCIE). Technical expertise, like anything else, has a distinct lifecycle, and in order to stay in touch with the cutting edge of technology, I aggressively pursued Voice technologies.

In 2009, I was caught by surprise when Cisco announced their entry into the server market, since the proverbial landscape was already crowded by long-standing players such as IBM and HP. The new product was the Cisco Unified Computing System, which launched with the 5100 blade server chassis and 6120-XP Fabric Interconnects (pictured below). While rack-mounted equivalents (the C-series) were introduced later, the original conception was the real game-changer. In the next few paragraphs you will find the more compelling technical aspects of this particular Cisco innovation.

Figure 1: Cisco UCS 5100 B-Series Chassis

Figure 2: Cisco UCS 6120-XP Fabric Interconnects

Review of Server/PC Architecture

To grasp the bigger picture of what Cisco UCS offers, we need to hit the proverbial "rewind" button and understand the architecture of computers in general. If you've been around the industry a long time, this might be very familiar and somewhat mundane, but if not, these hidden secrets may be invisible to you but still important. (See Figure 3.)

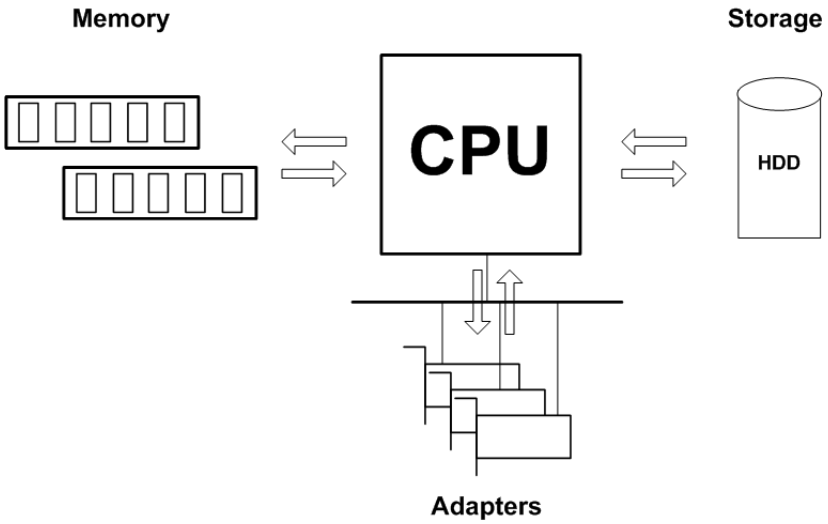

Figure 3: Server/PC Architecture Elements

Central Processing Unit (CPU)

While the term CPU is used freely in both technical publications and popular culture, many individuals may not fully grasp exactly what this component actually does. Often perceived as the "brain" of the computer, this device could actually be considered the "heart" as well. If you've ever tried to activate a computer without a CPU installed, you certainly realize that without it nothing happens. Composed of numerous electronic circuits, this component performs calculations, executes instructions, and brings life to the computer in which it is installed.

The CPU has evolved substantially over the years, with new form factors, sockets, speeds, and capabilities emerging over time. The CPU plugs into a large printed circuit board called the **motherboard**, using a connector called a **socket**. Today, multiple sockets, and by extension, multiple processors, are commonly installed on the board itself. To further complicate the computing equation, a processor may have multiple **cores**, which means that one socket can actually have multiple processors in the same physical package.

Memory

Some time ago, Bill Gates was once reportedly quoted as saying that 640K of RAM was all that anyone would ever need in a personal computer or server. My own personal laptop today came with 4 GB of memory, a far cry from the standard of 640K. The memory I am referring to here is **Random Access Memory**, sometimes referred to as **RAM**, **DRAM**, and a host of other, constantly-changing designations (I will use the general acronym RAM for convenience).

RAM also plugs into the motherboard, and interacts with the CPU and operating system to perform various functions, and store data during normal system operations. If you've ever tried to remember a phone number long enough to dial it, then you have a rough idea of how this type of short-term memory works. When the system is shut off, all contents of RAM are lost because the chips cannot hold the information without an active power source. It's worth noting that the amount of memory required for optimal operation continues to grow, depending on the role of the machine and the operating system being used.

Storage

Since the contents of RAM disappear when the machine is off, common sense dictates that information has to be stored in some location or everything would be a waste. This is the purpose of storage, usually involving a series of spinning disks that store information magnetically. Commonly referred to as "hard disks", these devices can come in all shapes and sizes, and they may exist in multiple numbers on one computing device. Solid state drives have begun to emerge, which makes use of all circuit-based components, much like flash drives. Larger storage arrays are available for mass storage of data as well.

Adapters

Computers would amount to little more than expensive toys of they had no ability to interact with the world outside of their own internal circuitry. In order to accomplish that worthy task, slots exist on the motherboard for specific adapters, some of which may be built into the board itself (video, keyboard, and mouse inputs and outputs). The original adapters used a now-obsolete connector called **ISA (Industry Standard Architecture)**, with the most current adapter type called **Peripheral Component Interconnect (PCI)**. Space doesn't permit a listing of the vast number of various adapters, but chances are your own personal experiences would help you list more than just a few.

Servers

Early Implementations

When I first entered the industry back in 1998, I worked at a regional Internet Service Provider in downtown Seattle, Washington, in the United States. At that time our "server farm" was nothing more than a set of warehouse shelves with tower-based personal computers connected to the data center network. When a customer required a new service to be turned up, the company would requisition a new computer, install it, and put it into production. In this environment, one service was inseparably tied to a single server. As pictured in Figure 4, our ISP had e-mail on one dedicated server, web services on another completely separate box, and Domain Name Services (DNS) on others.

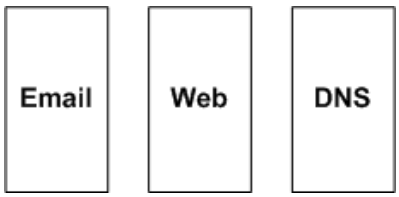

Figure 4: Single-Use Servers

Rack Mount Servers

If you recall the image of a shelf with computers sitting on it, you could easily come to the conclusion that a great deal of space was wasted. As the Internet began to grow and mature, the necessity of getting more equipment into the same amount of space began to get a lot of attention. This led to the rise of rack-mount servers, in which the same (or in most cases better) computing power got packed into compact boxes that fit into a small space in a network rack (usually expressed in **RUs** or **Rack Units**, or 1.75 inches).

Blade Servers

Continuing the push toward greater computing density, various vendors (such as HP) introduced the concept of **blade servers**. In this form factor, a large enclosure containing multiple slots provided centralized power and network connectivity to a series of servers. These servers, laid out in dense formats called blades, contain the same typical server components and insert into the enclosure. Most data center server deployments make use of Rack Mount servers, Blade Servers, or both (Figure 5).

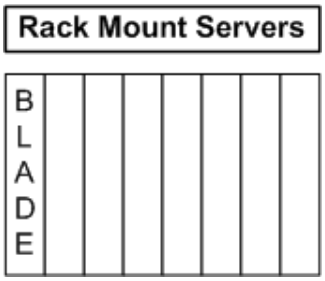

Figure 5: Rack-Mount and Blade Servers

Server Virtualization

There are few nearly universal truths of human experience, usually because of the dominance of personalities and long-held opinions, although one such truth would be the following: **all resources are finite!** In the computing and networking world, these resources would be space, power, cooling and, of course, budget. Simply making servers denser is not enough to solve the problem, in part because of server underutilization. Using the previous example of single-use servers, here is a possible usage scenario.

In this example, three physical servers, whether rack-mount or blade-form factors, are all drawing power, space, and cooling resources, and none are fully utilized (Figure 6). Virtualization software, using a construct called a hypervisor, can create three virtual machines on a single unit of hardware, reducing the required resources by 66.6 percent! The resulting configuration would look roughly like this (Figure 7).

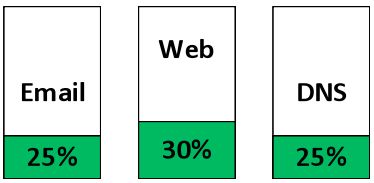

Figure 6: Server Utilization

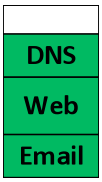

Figure 7: Virtualized Servers

Most industry experts identify VMware as the market leader in virtualization, though other vendors such as Citrix, Microsoft, Xen, and HyperV have established a presence in this fast-growing technology segment.

Data Center Architecture

Now that we have reviewed the evolution of servers and data center technologies, we are ready to look more closely at the architecture of a typical data center. Pictured in Figure 8 are the various elements in a representative format, resembling some designs that I have created over the years.

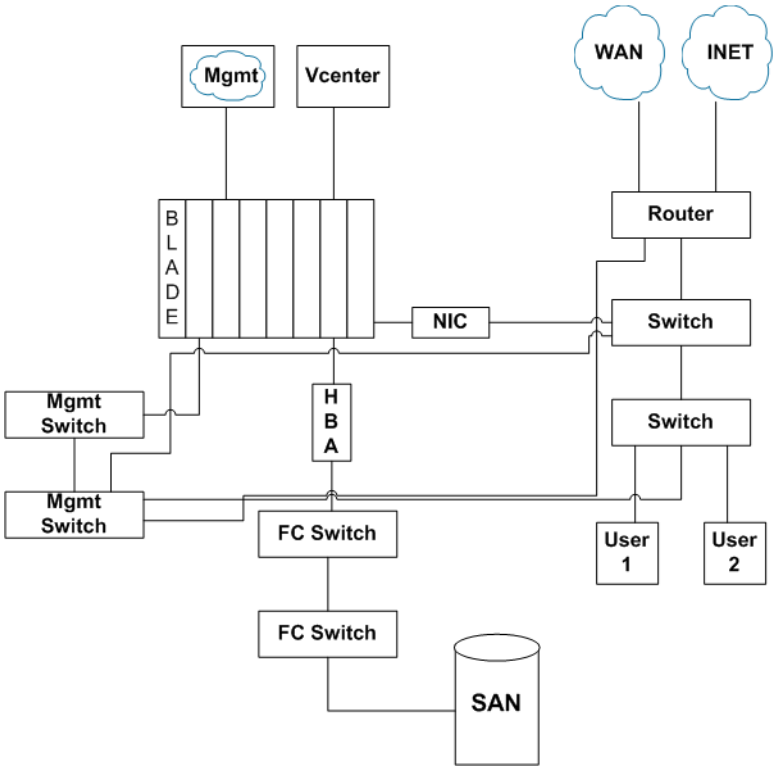

Figure 8: Typical Data Center Topology

Server Infrastructure

As discussed earlier in some detail, servers lie at the heart of the data center, performing numerous tasks and providing applications to the enterprise they serve. This may be a mix of rack-mounted and blade servers, though trends for greater density in servers would slant toward a blade-based infrastructure.

Storage Infrastructure

Mentioned only briefly earlier, large-scale storage arrays provide massive data archiving capabilities, often using completely separate networks using fiber channel protocols. Data transfer in this type of network takes place at a block level, which is intolerant of loss or delay. Using specialized fiber channel switches and unique server adapters called **Host Bus Adapters (HBAs)**, this adds another layer of complexity to the overall infrastructure. This is referred to as a **Storage Area Network (SAN)**.

Management Infrastructure

From a design perspective, the preferred method of network management is to make use of an **out-of-band (OOB)** network that is logically, or even physically, separate from the network carrying user data (production traffic). Many network devices actually have dedicated Ethernet management ports to facilitate this exact structure. In a physically separate management network, a separate set of switches connects the OOB interfaces, in essence requiring another network for this purpose.

Cisco UCS Architecture

The Cisco Unified Computing System (UCS), when inserted into the data center, actually consolidates many of these separate elements into a single environment. In the next section, all of the various threads presented thus far blend together into a new architecture with substantial benefits for all aspects of a business enterprise.

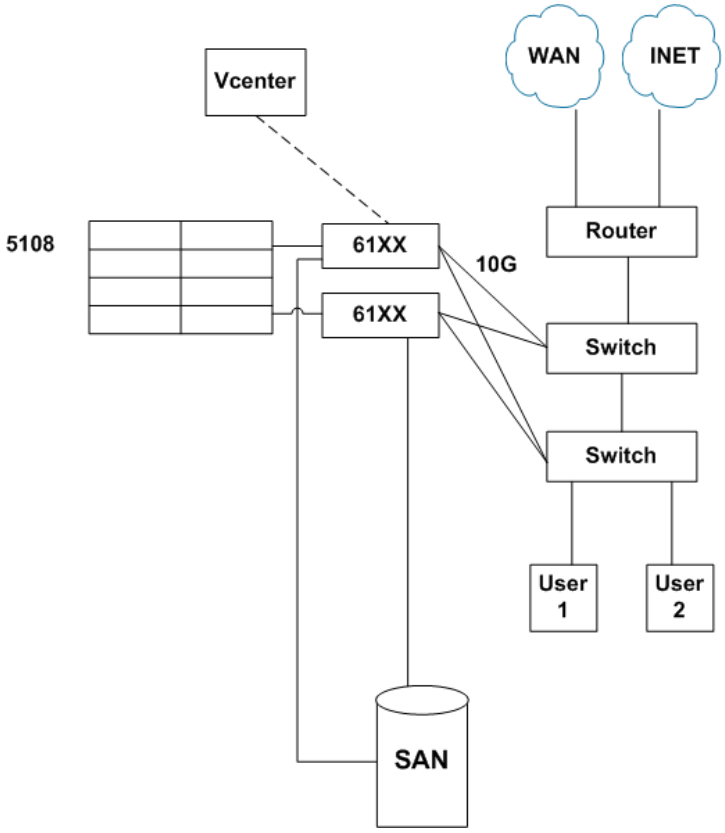

Figure 9: Cisco UCS Data Center Topology

Unified Fabric

Comparing the diagram above (Figure 9) with the previous data center diagram (Figure 8), you can see several components combined in the UCS environment. One of the first points of consolidation is the elimination of a separate Fiber Channel network, through the use of **Converged Network Adapters (CNAs)** that replace the HBA in the server platform. In addition, the Cisco 6120/40-XP Fabric Interconnects effectively transport Fiber Channel traffic using **Fiber Channel over Ethernet (FCOE)**, eliminating the need for a separate infrastructure. While the important technical details are beyond the scope of this particular discussion, it does allow for transport over a single network, in a manner similar to voice and data not long ago.

Unified Management (UCS Manager)

Embedded into the 6120/40-XP Fabric Interconnects is the UCS Manager, a web-based management software that allows for all management functions of the physical UCS infrastructure as well as interconnection with

V-Center (the VMware virtualization management software). In addition, a separate management network is also eliminated, further consolidating the infrastructure.

Service Profiles (Stateless Computing)

The specific hardware configuration elements of the blades in the server are defined in a construct termed a service profile. This profile, defined in the UCS Manager GUI, is pushed to the individual blades at run time. In addition, templates can be created to populate service profiles of differing types. Various configurations can be changed on the blades, allowing for rapid reconfiguration

Role-Based Access

Many IT organizations maintain separate personnel management structures, which may include the following.

- **Network:** Switches, Routers, and Infrastructure
- **Storage:** SAN device administration
- **Server:** Deployment of physical servers and associated applications

In many ways the consolidation of disparate network elements simplifies operational matters, but it can have the unfortunate side effect of creating proverbial "turf wars" with no simple solutions. One of the features built into the UCS Manager software is the concept of **roles**, or **role-based access**. In addition, a number of pre-defined roles already exist that can help manage access to the various aspects of the computing environment. Some of these include:

- Administrator (superuser access)
- Network Administrator (network asset access)
- Server Administrator (server access)
- Storage Administrator (storage access)

Additional roles can be defined as needed to meet business needs.

Conclusion

The Cisco Unified Computing System is a major step forward in data center development and creates numerous business and technical benefits. While complex from a technical standpoint, the system actually helps enterprises to fully realize the potential enhancements of Cisco UCS.

Images courtesy of www.Cisco.com and www.AnthonyPanda.com.

Learn More

To learn more about how you can improve productivity, enhance efficiency, and sharpen your competitive edge, Global Knowledge suggests the following courses:

DCUCS - Configuring Data Center Unified Computing

DCUCD v4.0 - Data Center Unified Computing Design

DCUCI v4.0 - Data Center Unified Computing Implementation

DCUCTS - Cisco Data Center Unified Computing System Troubleshooting v1.0

Visit **www.globalknowledge.com** or call **1-800-COURSES (1-800-268-7737)** to speak with a Global Knowledge training advisor.

About the Author

Joe Rinehart, MBA, CCIE #14256, CCNP/DP/VP is a professional trainer specializing in technology, business, and social media. He is also a successful speaker and published author, as well as a columnist for the Federal Way Mirror. He is active in the social media space, managing one of the largest groups on LinkedIn, as well as serving on the national steering committee of the Cisco Collaboration Users Group. Joe also serves as president of the Seattle Cisco Users Group, serving technology professionals throughout the Puget Sound region.

Joe Rinehart, MBA, CCIE #14256, CCNP/DP/VP

President and Chief Edutainment Officer

Gracestone Professionals, LLC

jrinehart@gracestonecompany.com

Twitter: jjrinehart