Design Guide: Transforming Call Centers with Pooled VDI

XenDesktop 7.6 Feature Pack 2 vSphere 6

TABLE OF CONTENTS

About FlexCast Services Design Guides 2
Project overview 2
Objective 2
Assumptions 3
Conceptual architecture 4
Detailed architecture 5
User layer 5
Access layer 5
Resource layer 7
Control layer 8
Hardware layer 9
Validation 12
Next steps 14
Appendix: Authentication and Enumeration Process. Error! Bookmark not defined.

About FlexCast Services Design Guides

Citrix FlexCast Services Design Guides provide an overview of a validated architecture based on many common scenarios. Each design guide relies on Citrix Consulting best practices and in-depth validation by the Citrix Solutions Lab to provide prescriptive design guidance on the overall solution.

Each FlexCast Services Design Guide incorporates generally available products and employs a standardized architecture, allowing multiple design guides to be combined into a larger, all-encompassing solution.

Project overview

Cost effective call centers often pose operational and technical challenges for organizations. To effectively run a call center, organizations must

- Maintain adequate staffing, which fluctuates daily and seasonally
- Focus on cost containment from a real estate and technology infrastructure perspective
- Utilize technology that provides call center agents with the right information to improve customer satisfaction

Citrix XenDesktop solves these challenges with a scalable solution allowing organizations to enable dozens, hundreds or thousands of call center agents to work onsite or remotely on a schedule tailored to the actual demand without needing office space and equipment sized for peak times.

Note: *The goal of this design guide is not to test the latest hardware available, but to provide a framework for which admins can use to customize their own design and implementation based on different hardware. As the physical servers continue to increase in processing power, each server will be able to support greater numbers of users, but the overall architecture of the solution will remain fairly constant.*

Objective

The objective of the FlexCast Services Design Guide is to construct and demonstrate a cost-effective way of delivering a call center computing environment with integrated telephony to internal and external users.

This is the challenge impacting WorldWide Corporation (WWCO), a hypothetical organization that would like to provide increased customer service with proper staffing but without large investments in real estate and technology.

To address these challenges, IT decided to implement a Citrix XenDesktop 7.6 Feature Pack 2 environment to deliver a call center virtual desktop to internal and external call center agents. To properly validate the solution, IT identified a 500-user division for the project.

WWCO business objectives

- Provide an integrated solution for local and remote call center agents
- Deliver a standard computing environment with all the necessary call center agent tools
- Provide employees with an exceptional user experience while on the LAN and WAN.
- Quickly scale up/down to accommodate temporary, external call center agents during seasonal staffing fluctuations
- Centrally manage a single master image to all users to help reduce troubleshooting and support incidents

WWCO technical objectives

- Support access to a Windows desktop from employee-owned devices with different form factors, including tablets, phones, desktops and laptops, and different operating systems, which include iOS, Mac, Android, Linux and Windows.

- Build a solution that scales from a few hundred users to thousands with minimal changes to the infrastructure
- Implement an N+1 highly available solution without large cost increases
- Centrally manage and control employee access and permissions
- Utilize virtualized components, where possible, to reduce costs

Assumptions

The following assumptions played a role in defining the overall strategy for WWCO:

- All resources (physical servers, virtual servers, Windows applications) will be hosted from a single datacenter running VMware vSphere 5.5.
- High availability is required for all critical components in N+1 mode, where enough spare capacity will be built into the system to allow for the failure of one component without impacting user access.
- WWCO's existing Microsoft Active Directory and DNS/DHCP will be reused.
- The master image will consist of call center applications.

Conceptual architecture

Figure 1, based on the overall business and technical objectives for the project as well as the assumptions, provides a graphical overview of the solution architecture.

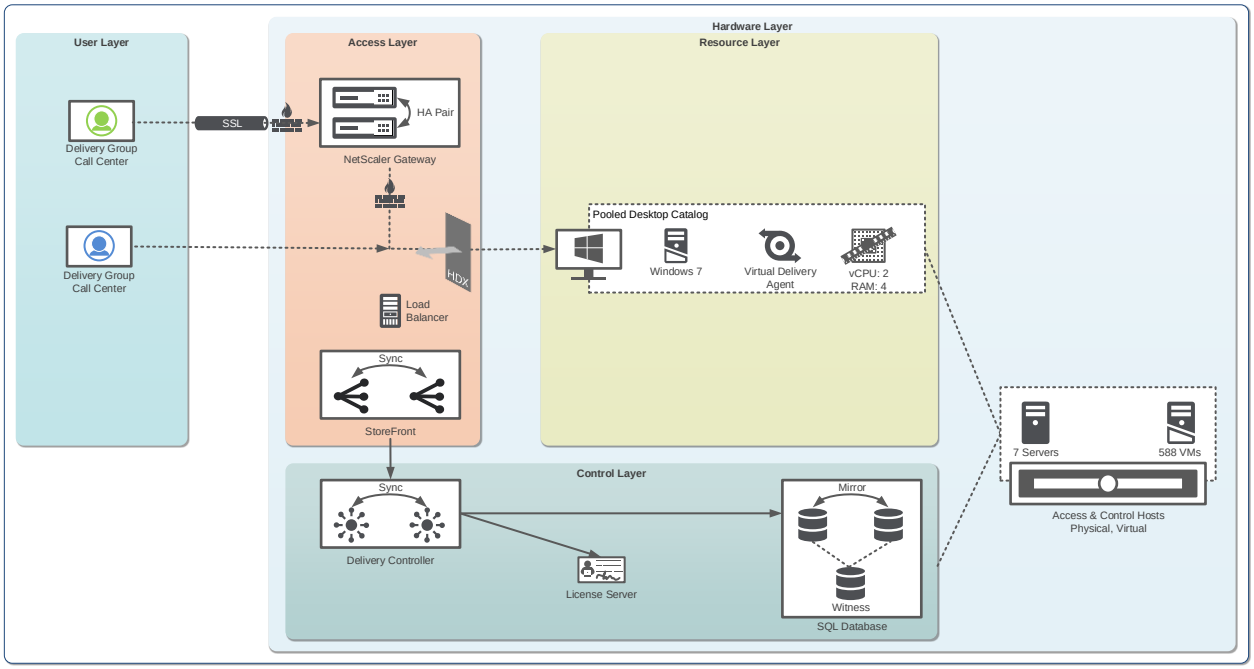

Figure 1: Conceptual architecture

This architecture is suitable for 500 users requiring local and remote access to a standardized call center configured virtual desktop.

At a high level, the following information can be ascertained from the conceptual architecture:

- The 500-user division used in the first phase of the rollout is called Call Center. This group will utilize personal and thin client endpoints to connect to the environment from the local office and home offices (for temporary employees).
- The allocated resources for the Call Center user group is a single, non-persistent desktop pre-configured with a set of standardized, call-center applications.
- The base operating system, Windows 7, is delivered to the appropriate virtual machines via Machine Creation Services.
- User customization is denied in order to keep users in a standard environment, helping to reduce potential issues.
- The total hardware allocation requirement for the solution is 7 physical servers.

Each layer of the architecture diagram and the relevant components are discussed in greater detail below.

Detailed architecture

The overall solution for WWCO is based on a standardized five-layer model, providing a framework for the technical architecture. At a high level, the 5-layer model comprises:

1. User layer. Defines the unique user groups and overall endpoint requirements.
2. Access layer. Defines how user groups will gain access to their resources. Focuses on secure access policies and desktop/application stores.
3. Resource layer. Defines the virtual resources, which could be desktops or applications, assigned to each user group.
4. Control layer. Defines the underlying infrastructure required to support the users in accessing their resources.
5. Hardware layer – Defines the physical implementation of the overall solution with a focus on physical servers, storage and networking.

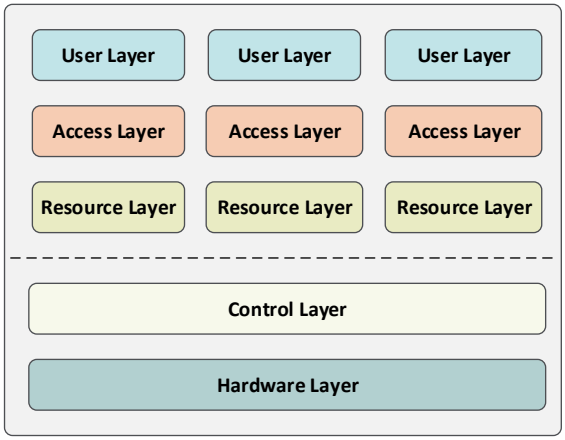

Figure 2: Five-Layer Model

User layer

The user layer focuses on the logistics of the user groups, which includes client software, recommended endpoints and office locations. This information helps define how users will gain access to their resources, which could be desktops, applications or documents.

- Citrix Receiver client. This client software, which runs on virtually any device and operating platform, including Windows, Mac, Linux, iOS and Android, must be downloaded onto user endpoints to access virtual desktops, which are hosted in the datacenter. Citrix Receiver provides the client-side functionality to secure, optimize and transport the necessary information to/from the endpoint/host over Citrix HDX, a set of technologies built into a networking protocol that provides a high-definition user experience regardless of device, network or location.
- Endpoints. The physical devices used by the internal Call Center user group are thin clients configured with Citrix Receiver while external users will be able to use their own traditional (desktop or laptop) personal device. Due to the application and job requirements, WWCO has decided that mobile devices are not acceptable endpoints.
- Location. The Call Center user group will work from local offices, over secure network connections as well as external locations, over unsecure network connections. All traffic must be encrypted.

Access layer

The access layer defines the policies used to properly authenticate users to the environment, secure communication between the user layer and resource layer and deliver the applications to the endpoints.

The following displays access layer design decisions based on WWCO requirements.

Users connecting from…	Remote, untrusted network	Local, trusted network
Authentication point	NetScaler Gateway	StoreFront

Authentication policy	Multi-factor authentication (username, password and token)	Single-factor authentication (username and password)
Session policy	Mobile Traditional	Not applicable
Session profile	No Access ICA proxy	Not applicable
User group	Call Center	Call Center

- Authentication. For internal users, single factor authentication, using a username and password, will be satisfactory. However, allowing external users to access the environment from a remote location without more stringent authentication requirements would pose security risks to WWCO. When external users access the environment, the external URL will direct requests to NetScaler Gateway, which is deployed within the DMZ portion of the network. NetScaler Gateway will accept multi-factor authentication credentials from users and pass them to the appropriate internal resources (Active Directory domain controllers and token authentication software such as RADIUS).
- Session policy. NetScaler Gateway can detect the type of endpoint device and deliver a specific access experience based on device properties and policy. WWCO policies are:
    - Mobile. When users connect with a mobile device, a separate policy will be applied to block access, based on WWCO requirements. By using the following expression within the NetScaler Gateway session policy configuration, this policy will only be applied to mobile devices:
      "REQ.HTTP.HEADER User-Agent CONTAINS CitrixReceiver"
    - Traditional. This policy will be applied to all non-mobile devices by using the following expression within the NetScaler Gateway session policy configuration:
      "ns_true"
- Session profile. As Call Center group members only require access to their respective virtual desktops from traditional endpoints, the session profile will be configured as ICA proxy instead of full VPN mode. ICA proxy allows only HDX traffic to pass from the endpoint to the user's physical desktop through NetScaler Gateway, while full VPN mode makes the endpoint act as if it is physically on the internal network. Using an ICA proxy session profile helps protect the environment by allowing only session-related traffic to pass, while blocking all other traffic.

In order to support the access layer design, the following components are required:

Parameter	NetScaler Gateway	Load Balancer	StoreFront
Instances	2 virtual servers	2 virtual servers	2 virtual servers
CPU	2 vCPU	2 vCPU	2 vCPU
Memory	2 GB RAM	2GB RAM	4 GB RAM
Disk	3.2 GB	3.2 GB	60 GB
Citrix product version	NetScaler VPX for vSphere 11 Build 62.10	NetScaler VPX Express for vSphere 11 Build 62.10	StoreFront 3
Microsoft product version	Not applicable	Not applicable	Windows Server 2012R2 Standard
Network ports	443	443	443
Redundancy	High-availability pair	High-availability pair	Load balanced via NetScaler Express

Resource layer

The resource layer defines the underlying image, how to deliver the image to the associated virtual machines, which applications to deliver and how to provide the right level of personalization for the respective user group.

Criteria	Decision
Operating system	Windows 7
Delivery	Machine Creation Services
CPU	2 vCPU
Memory	4 GB RAM
Image size	30 GB
Disk cache	Differencing Disk: 5 GB
Page file size	2 GB (contained within differencing disk)
Application(s)	Call center routing (VoIP) Call center agent Call center chat
Profile	Mandatory profile
Policy(s)	Hi-def experience Secure Optimized for WAN
Peripherals	USB headphones
User group	Call Center

- Machine Creation Services is not limited by scale, but rather by the type of delivery target: physical or virtual machine. As the project is based on resource delivery to virtual machines, Machine Creation Services is the ideal solution. Machine Creation Services does not require additional hardware or resources as it simply utilizes the hypervisor and local storage to create unique, thin, provisioned clones of a master image, resulting in a solution that is simple to deploy and easy to scale.
- Because call center agents only require a small set of predefined/preconfigured applications, personalization will be prevented with the use of mandatory profiles. Each call center agent using the environment will have an environment that is identical to another agent. This allows each agent to work through the structured task list as quickly as possible.
- WWCO must create a user experience that can accommodate voice communication for internal and external users while keeping the environment secure. As the network link between external users and resource is dynamic and uncontrolled, policies are needed to optimize the user experience for the WAN. Based on these requirements, the following policies will be used for the environment

Policy	Settings	Applied to…
Optimized for WAN	Based on the template “Optimized for WAN”	Any user connecting through NetScaler Gateway
Secure resources	Based on the template “Secure and Control”	Delivery group

Hi-Def experience	Based on the template “High Definition User Experience”	Any user not connecting through NetScaler Gateway

Control layer

The control layer of the solution defines the virtual servers used to properly deliver the prescribed environment detailed in the user, access, and resource layers of the solution, including required services, virtual server specifications and redundancy options.

The decisions for the Call Center group are met by correctly incorporating and sizing the control layer components, which include delivery and infrastructure controllers.

Delivery controllers

The delivery controllers manage and maintain the virtualized resources for the environment. In order to support the resource layer design, the following components are required:

Parameter	Delivery Controller
Instances	2 virtual servers
CPU	2 vCPU
Memory	4 GB RAM
Disk	60 GB
Citrix product version	XenDesktop 7.6 Feature Pack 2
Microsoft product version	Windows Server 2012R2 Standard
Network ports	80, 443
Redundancy	Load balanced via NetScaler VPX Express

A single delivery controller can easily support the load of 500 users. However, to provide N+1 fault tolerance, a second virtual server will provide redundancy in case one virtual server fails.

Infrastructure controllers

In order to have a fully functioning virtual desktop environment, a set of standard infrastructure components are required.

Parameter	SQL Server	License Server	vCenter Server 6
Instances	3 virtual servers	1 virtual servers	1 virtual server
CPU	2 vCPU	2 vCPU	2 vCPU
Memory	4 GB RAM	4 GB RAM	8 GB RAM
Disk	60 GB	60 GB	100 GB
Version(s)	Not Applicable	Citrix License Server 11.12	Not applicable

Microsoft product version	Windows Server 2012R2 Standard SQL Server 2012 Standard (x2) SQL Server 2012 Express (x1)	Windows Server 2012R2 Standard	Windows Server 2012R2 Standard
Network ports	1433	27000, 7279, 8082	80, 389, 443, 636, 902, 903, 8080, 8443, 60099, 10443, 10109, 10111
Redundancy	SQL Mirroring with Witness	None due to 30 day grace period	None

To provide fault tolerance, the following options were used:

- The XenDesktop database was deployed on an HA pair of Microsoft SQL Server 2012 servers utilizing mirroring across two virtual servers. A third virtual server running Microsoft SQL Server 2012 Express was used as a witness.
- Once active, a XenDesktop environment can continue to function for 30 days without connectivity to the Citrix License Server. Due to the integrated grace period, no additional redundancy is required.
- Only a single vCenter server is used, as the loss of the server has minimal impact on a XenDesktop environment. Without the vCenter server, only the power functions of the virtual machine are affected. All virtual servers that are currently running will continue to run, any connected user will notice no service disruption and any user who tries to connect to a session will succeed. Power functions can still be managed manually from the local console if needed.

Hardware layer

The hardware layer is the physical implementation of the solution. It includes server, networking and storage configurations needed to successfully deploy the solution.

Server

Following is the physical server implementation for the WWCO solution:

Component	Description	Quantity	Total
Server model	HP DL380P G8	7	7 servers
Processor(s)	Intel Xeon E5-2690 @2.9GHz	14	2 processors per server (16 cores)
Memory	16GB DDR3-1333	168	384 GB per server
Disk(s)	300GB SAS @ 15,000RPM	112	4.8 TB per server
Storage Array Controller	HP Smart Array P420i Controller 2 GB cache	7	1 controller per server
Microsoft product version	Windows Server 2012R2 datacenter	7	1 per server

To provide fault tolerance within the solution, the virtual servers were distributed so redundant components were not hosted from the same physical server. The virtual server allocation is depicted in Figure 3.

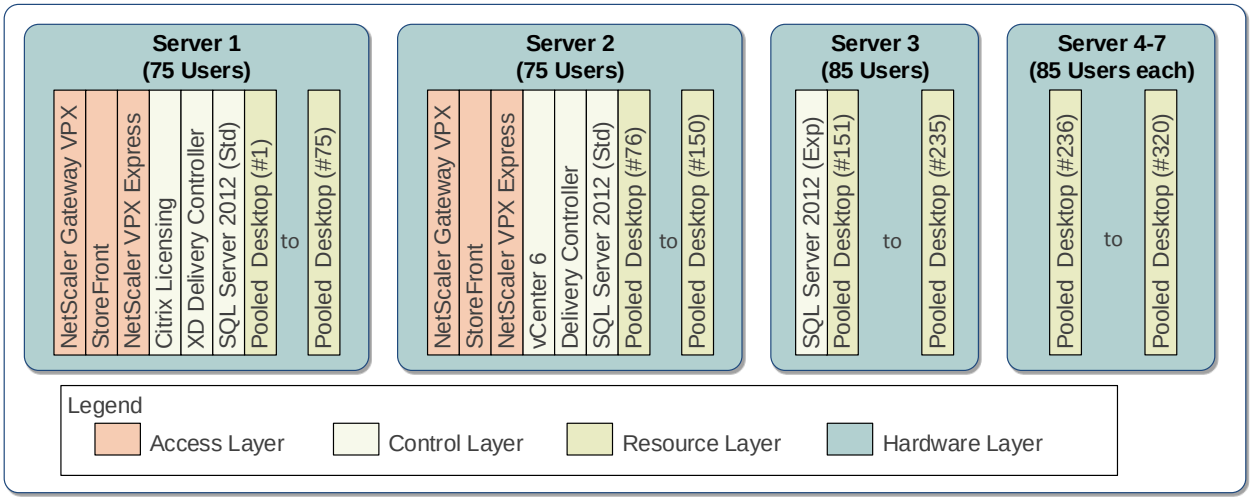

Figure 3: Virtual machine server allocation

***Note**: The resource load on the physical hardware for the access and control layer components is minimal, which is why the hosts are also able to support VDI virtual machines.*

***Note:** This design can accommodate a single server failure. However, when all servers are operational, only 500 desktops will be powered on at a time, conserving resources.*

***Note:** The entire environment can scale much higher by adding additional physical servers that mimic the configuration of Server 4.*

Storage

The storage architecture for the solution is based on the use of inexpensive local storage. To ensure an acceptable user experience, the storage architecture must have enough throughput capacity as well as fault tolerance to overcome the potential failure of a single drive.

Parameter	Hosts
Drive count	16
Drive speed	15,000 RPM
RAID	RAID 10
IOPS per user	20
Read/write ratio	40/60
Characteristics	Random, 4K blocks

Based on tests, each user accessing a virtual desktop will generate, on average, roughly 20 IOPS during their steady state activity.

In addition to the resource layer virtual servers, the control and access layer systems generate IOPS activity. However, the impact on storage is minimal when compared to the active sessions generated by users.

As the overall solution is more write intensive, it is recommended to utilize a RAID 10 configuration across the eight hard disk drives, as RAID 10 provides fault tolerance and better write performance than RAID 5.

Networking

Integrating the solution into the network requires proper configuration to have the right components communicate with each other. This is especially important for NetScaler Gateway, which resides in the DMZ. The network is configured based on each physical server's having four network ports:

NIC instance	Function	Speed	VLAN ID
1	Management VLAN	1 Gbps	1
2	Virtual machine VLAN	1 Gbps	2
3	DMZ VLAN	1 Gbps	3
4	Disabled		

The three VLANs are divided among the physical servers, NetScaler Gateway and remaining virtual servers as shown in Figure 4.

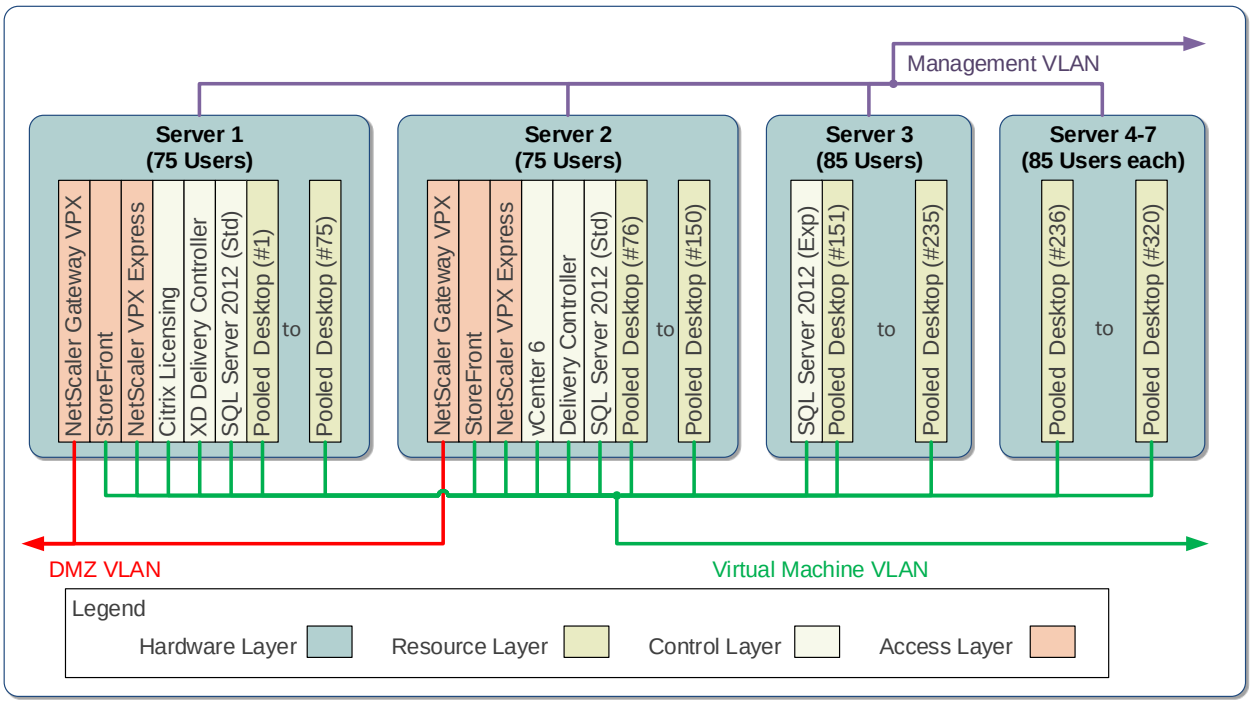

Figure 4: Networking architecture

As depicted in the diagram, the VLAN is configured as follows:

- NetScaler Gateway is configured to use the DMZ VLAN. This VLAN does not connect with any other internal networks, which helps keep the DMZ and internal traffic separated.
- The management VLAN is only connected to the physical hosts and not the virtual machines. This VLAN is for management calls to/from the physical server's hypervisor.
- The virtual machine VLAN, meant for all non-DMZ virtual machines, allows them to connect to the internal datacenter network.

Validation

The defined solution was deployed and validated by the Citrix Solutions Lab. The key findings from the validation are:

- CPU was the limiting factor in scaling out the environment.
- Each dedicated, physical server, when configured with an "Optimized for User Experience policy" supported 92 Windows 7 desktops. If the "Optimized for Scale" policy is used, each physical server can support 135 Windows 7 desktops.
- At peak, the control layer components of SQL Server, StoreFront and desktop delivery controllers consumed less than 30 percent of CPU and had over 20 percent of available memory.
- The NetScaler Gateway CPU, memory and network utilization was under 10 percent for the 500-user load.
- Based on the overall solution, a 1 Gbps switch would provide sufficient network capacity.

Figure 5 provides a graphical representation of the utilization of the control layer components as the user load increased.

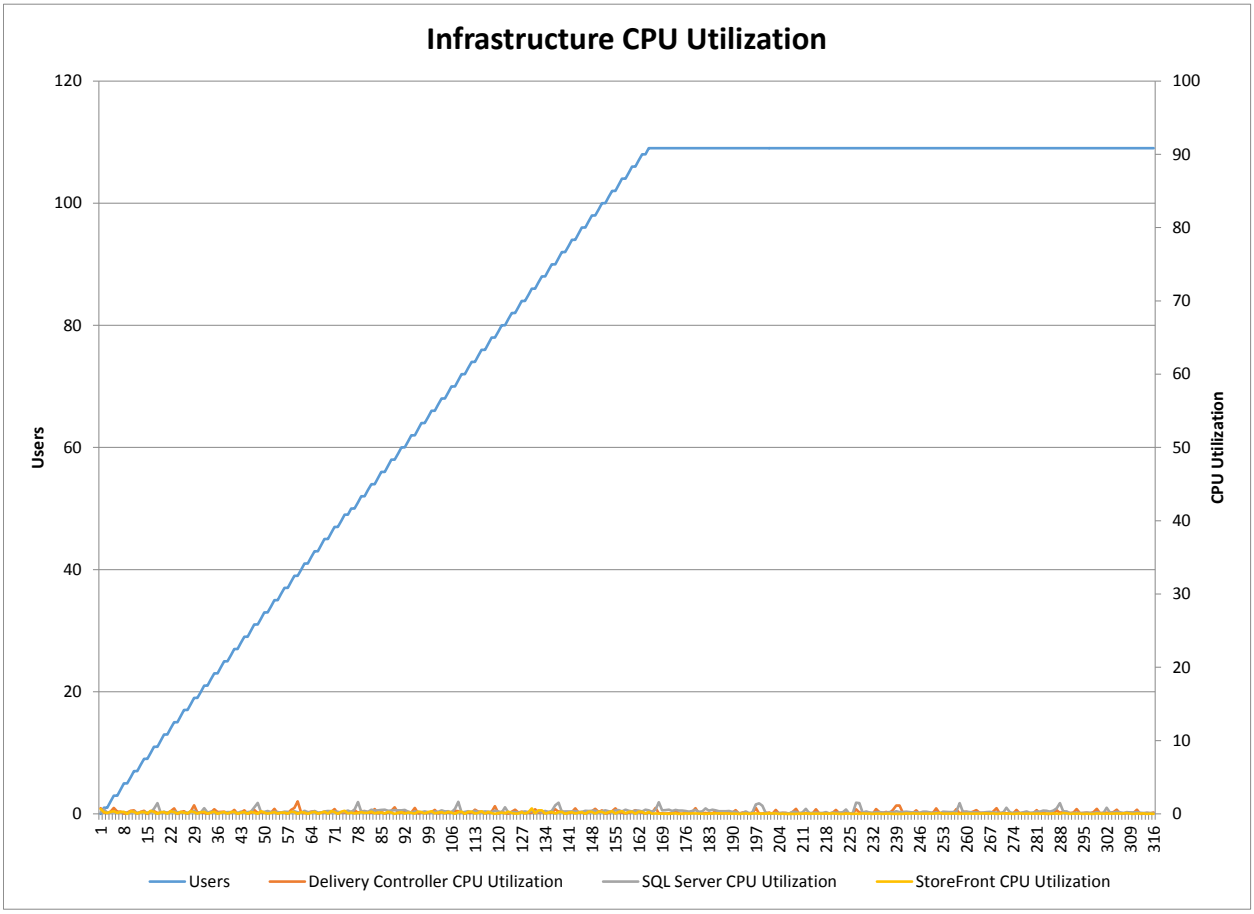

Figure 5: Processor Utilization for Control Layer Components

Although the solution was designed to only support 500 users, the control layer components, responsible for supporting and maintaining the environment, are minimally utilized and are capable of much higher user loads with the inclusion of additional physical servers.

Note: *The goal of this design guide is not to test the latest hardware available, but to provide a framework for which admins can use to customize their own design and implementation based on different hardware.*

As the physical servers continue to increase in processing power, each server will be able to support greater numbers of users, but the overall architecture of the solution will remain fairly constant.

Next steps

Organizations must constantly expand and contract their workforce to align with customer demands. This is easily observed in a call center environment where the number of call center agents must fit the anticipate call volume without requiring long hold times or idle agents. Many organizations know how many agents they need based on the seasonality of their business, but they struggle to expand and contract the desktop infrastructure to accommodate the seasonal shifts.

Citrix XenDesktop solves these challenges with a scalable solution allowing organizations to enable dozens, hundreds or thousands of call center agents to work onsite or remotely on a schedule tailored to the actual demand without needing office space and equipment sized for peak times.

To help you learn more about the potential benefits that XenDesktop 7.6 Feature Pack 2 can provide, Citrix has prepared the following resources: .

- XenDesktop 7.6 Blueprint: A layered solution for all successful designs and deployments, focusing on the common technology framework and core decisions
- Reviewer's Guide: Prescriptive guide for deploying the solution to five or 10 users quickly and easily in a non-production environment
- XenDesktop Deployment Options: Multiple deployment options are available for XenDesktop including on-premises, cloud-hosted, Citrix Workspace Cloud and Citrix Service Provider. Learn how each one impacts the XenApp architecture

About Citrix
Citrix (NASDAQ:CTXS) is the cloud company that enables mobile workstyles—empowering people to work and collaborate from anywhere, easily and securely. With market-leading solutions for mobility, desktop virtualization, cloud networking, cloud platforms, collaboration and data sharing, Citrix helps organizations achieve the speed and agility necessary to succeed in a mobile and dynamic world. Citrix products are in use at more than 260,000 organizations and by over 100 million users globally. Annual revenue in 2012 was $2.59 billion. Learn more at www.citrix.com.