Citrix Reference Architecture for Multi-Tenant Desktop as a Service

Using Citrix XenApp 6 to Enable the Delivery of Microsoft Windows-based Desktops and Software as a Service
Executive Summary

The Citrix® Reference Architecture for Multi-Tenant Desktop as a Service guides partners in designing the new generation of Desktop as a Service (DaaS) and Software as a Service (SaaS) services. These services leverage the Microsoft® Service Provider License Agreement (SPLA) and Citrix Service Provider (CSP) licensing programs to deliver Microsoft Windows® applications and desktops on a pay as you go basis to small and medium businesses. This reference architecture is for service providers that will manage a multi-tenant application and desktop delivery service rather than an outsourced “IT as a Service” datacenter where each end tenant manages their own resources.

Introduction and Scope

This document targets Citrix Service Providers that will provide DaaS and SaaS subscription services to a subscriber base of 50,000 or fewer active users, comprised of multiple small to medium businesses (Tenants) with around 250 (or fewer) users/subscribers, per tenant. While the systems themselves are certainly capable of scaling well above this capacity, the architectural considerations at these higher levels differ and thus will be addressed in a separate document.

The Citrix solution for DaaS and SaaS presents users with a familiar Windows desktop and applications experience, enabled by Citrix XenApp™ 6 and the Citrix Service Provider Automation Pack extensions to the Microsoft® Windows Server® 2008 R2 - Remote Desktop Services base operating system. Service Providers can deliver these DaaS and SaaS services to users of any Citrix Receiver™ enabled end-point device over secured public internet connections with Citrix NetScaler® and Access Gateway™. For subscriber locations that aggregate a number of end-points in a single setting such as a small retail business, accountant office or medical clinic, Citrix Branch Repeater™ provides a high definition user experience to all end-points over an optimized network connection.

A sample multi-tenant DaaS and SaaS environment using Citrix XenApp 6 Worker Groups for multi-tenancy is referenced throughout this document. XenApp 6 Worker Groups map to individual subscriber-specific Organization Units within a single Microsoft Active Directory Domain, or alternatively to individual subscriber specific child domains within a single Microsoft Active Directory Forest. All subscriber partitions and services are managed through centralized dashboards and monitoring systems as provided by Citrix, Microsoft and 3rd parties. This document does not provide detailed guidance for CSPs requiring multi-tenant designs where a single farm is dedicated to each subscriber, although many of the concepts within this document apply to such designs.

The sample design follows security best practices. A CSP must determine any deviations from these suggestions based on particular business needs. High Availability capabilities are inherent within many components of the architecture, whereas Disaster Recovery designs are enabled most effectively by Citrix NetScaler. Further details of additional HA and DR options are in the product documentation for the respective components at http://www.citrix.com/edocs.
The Citrix® Reference Architecture for Multi-Tenant Desktop as a Service (DaaS) presents tenants and end subscribers with a cost effective, pay per month subscription to desktop and application services provided by a 3rd party Citrix Service Provider (CSP). This section provides an overview of the Conceptual Reference Architecture as well as the core technical concepts that enable this architecture.
Core Concepts

The core concepts of end-point ubiquity, subscription-based licensing, multi-tenancy, single instance management and dynamic assembly provide much of the simplified scalability and high user acceptance of this emerging service model.

End-point ubiquity

End-point ubiquity enables the widest adoption of desktop services capable of providing high performance graphics and peripheral IO capabilities for Windows desktops on any device over any network. In the Citrix Reference Architecture for Multi-Tenant Desktop as a Service, Citrix Receiver enables this ubiquity through support for virtually every popular device and platform in the market. Citrix Receiver enables a high-definition experience (HDX™) for users of DaaS, increasing user acceptance and helping to grow this emerging services market.

Subscription-based licensing

DaaS is provided as a pay per use, pay per month model which enables subscribers to treat desktop acquisition and maintenance as an operational expense for services rather than a capital expense of owning and building their own infrastructure. The savings recognized by the economies of scale that multi-tenancy, dynamic assembly and single instance management enable further enhance the business proposition of this licensing model by keeping month to month costs relatively low for both the subscriber and service provider.

Multi-tenancy

Multi-tenancy capabilities provide large economies of scale from a single infrastructure servicing multiple organizations. This provides different levels of trade-offs regarding price and performance in reference to an individual tenant’s service level requirements.

Single instance management

Single instance management provides the most efficient life-cycle maintenance of operating system workloads and Windows-based applications. By creating a single read-only image of each critical workload, and then streaming that workload onto physical or virtual machines, CSPs can maintain
thousands of execution environments from a single source, requiring only a reboot of the individual machine to deploy the latest image. Application virtualization provides a similar solution at the application layer where a single application image is maintained and patched and then streamed into a user’s execution environment at run-time. This single application image can be securely streamed to thousands of users, across multiple tenants.

Dynamic assembly

Dynamic assembly of the operating systems, applications, and user personalization settings on a per user, per tenant basis is a critical enabler of massive scale and efficient management for Windows-based desktops and applications within a multi-tenant service provider solution. This dynamic assembly leverages single instance managed images as components that are assembled for a user based upon the services subscribed to by the tenant. Single instance images are dynamically assembled in different configurations for each tenant/user based on their individual configurations and service level agreements.

DaaS and SaaS subscription licensing

Fundamental to the concept of DaaS and SaaS is the subscriber requirement that all IT services and products are purchased as an operational expense rather than the upfront capital expense of building or upgrading a datacenter. To facilitate these subscriber requirements many software vendors have begun to introduce subscription-based licensing programs made available through service providers to subscribers. The Citrix Service Provider (CSP) licensing program is the Citrix response to this emerging model.

- **Citrix CSP licensing**
  The Citrix Service Provider (CSP) program is designed specifically for service providers who provide hosted desktop and software services to end-user customers. The CSP program addresses the service provider market for offsite, multi-tenant hosting where the end-user customer is not the licensee. The CSP program extends to service providers the “right to use” Citrix products as the underpinning of their delivery infrastructure and gives them the flexibility of a monthly “active subscriber” pricing and licensing model. Service providers always have access to the most current versions of Citrix products available in the program and only pay for actual end-user usage recorded during the previous calendar month.
• **Microsoft SPLA licensing**

SPLA (Service Provider Licensing Agreement) is the Microsoft service provider program. SPLA enables service providers and ISVs with a hosted offering to license Microsoft products on a monthly basis to provide services and hosted applications to their end customers. SPLA is a well known industry term that many service providers equate with the monthly pricing and licensing model used to charge for hosted software services. Citrix Service Provider program includes a similar CSP Program Agreement that defines service provider partner use rights.

**Overview of Reference Architecture Modules**

**Infrastructure as a Service (IaaS)**

The foundation of the service provider platform begins at the Infrastructure as a Service (IaaS) module, which has three sub-components:

1. Network infrastructure
2. Access infrastructure
3. Provisioning infrastructure

The IaaS module controls the system-wide network configuration, forest-level Active Directory management, remote access and all layers of provisioning.

**Multi-Tenant DaaS and SaaS**

The multi-tenant DaaS-SaaS module is the core component of the service provider datacenter(s)—this is what controls virtual desktops (DaaS) and applications (SaaS) delivery and the multi-tenant architecture. Within this module, applications and desktops are virtualized, subscriber partitions and Active Directory boundaries are defined, and centralized XenApp systems govern the entire system.
Dashboards and management

To successfully build and manage a service provider network, IT administrators need the correct set of management and troubleshooting tools. To do this efficiently, administrators need tools that are simple to use with a wide level of control. The Citrix Delivery Services Console, Cortex Control Panel from Citrix and other supporting management dashboards present a unified view across the entire infrastructure, from the datacenter to the device. This end-to-end view gives service providers the detailed information necessary to ensure that service level agreements for subscribers are maintained.

End-points and offices

When desktops and applications are delivered as a service, the user is the ultimate judge of the experience being delivered. Service providers must deliver a great, consistent experience across all networks to any device to capitalize on the largest subscriber base. How does a service provider do this when they do not manage the endpoint or the network? Citrix Receiver and HDX are the strategic components that make this a reality.

With Citrix Receiver, CSPs have complete control over security, performance and user experience with no need to own or manage the physical device or its location. Users simply install Citrix Receiver on their own device to gain access to their desktop and all of their business, web and SaaS applications.
The complete implementation of the reference architecture leverages multiple vLANs, security zones and a combination of physical and virtual server installations to enable a highly scalable, high performance and easily managed multi-tenant DaaS solution for service providers targeting small to medium businesses. This section describes the design, the various components and configurations used, and the best practices approach employed to build a solution capable of supporting 50,000 desktops, representing thousands of subscribing businesses, within a single CSP management scope.
Multi-Tenancy Design Considerations

Looking at the complete virtualization stack that enables DaaS, one notices the several potential layers where multi-tenancy capabilities can be introduced within the environment. All of them have their advantages and disadvantages to a DaaS business. Note the following trade-offs of implementing multi-tenancy at a specific layer:

<table>
<thead>
<tr>
<th>Layer</th>
<th>Example</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Network</td>
<td>Completely separate datacenters or networks</td>
<td>Greatest level of design flexibility and tenant isolation</td>
<td>Highest cost per tenant because all infrastructure is replicated per tenant</td>
</tr>
<tr>
<td>Virtual Network</td>
<td>Co-located datacenters using vLANs as the primary isolation layer</td>
<td>Very high level of design flexibility and tenant isolation</td>
<td>Only slightly less expensive than physical network isolation</td>
</tr>
<tr>
<td>Physical Server</td>
<td>Co-located datacenters renting dedicated physical servers to subscribers</td>
<td>Relatively high level of design flexibility</td>
<td>Network layer security can be compromised; still relatively high cost; intra-server communications can be cumbersome to design</td>
</tr>
<tr>
<td>Virtual Machine</td>
<td>Dedicated Virtual Hosts; co-located VMs within a single datacenter and network</td>
<td>Relatively high level of design flexibility; lower per machine cost for provider and tenant because hardware is</td>
<td>Network layer security can be compromised; intra-server communications can be cumbersome to design</td>
</tr>
<tr>
<td>Session Layer Virtualization</td>
<td>Hosted Shared Desktops</td>
<td>Highest subscriber density; lowest infrastructure and management costs</td>
<td>Lower design flexibility per tenant than machine and lower layer partitioning</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Application Virtualization</td>
<td>Microsoft App-V, Citrix Application Streaming</td>
<td>Enables dynamic assembly of OS and Applications, on-demand for each user; enables single instance management for application workloads across tenants; enables offline access to CSP managed applications</td>
<td>Requires repackaging of applications for streaming and virtualization; some applications still not deliverable in this manner</td>
</tr>
</tbody>
</table>
Citrix Service Providers leverage different multi-tenant approaches deploy and manage a single instance of Citrix infrastructure (thus saving costs) while continuing to meet the individual tenant expectations. A CSP must take into account several considerations to determine the best approach for their business. This section discusses some of those considerations, as well as three example multi-tenant approaches used in the market today.

Although a specific CSP business model can consider more parameters, five primary considerations will influence most multi-tenant DaaS designs.

- **Isolation**: Isolating users of one tenant from users of other tenants to prevent leakage of sensitive information or from being affected by activities of other tenants.

- **Performance Guarantees**: Ensuring that performance of one tenant is not negatively affected by activities of other tenants.

- **Customized Experience**: Providing unique environmental, security or performance aspects to an individual tenant based on their specific service level agreement (SLA) within the multi-tenant environment.

- **Self-Service Administration**: Providing the ability for tenants to perform some level of administration for their specific services.

- **Cost**: Delivering the correct mix of the above capabilities at an appropriate cost.

Citrix DaaS infrastructure, based upon XenApp 6, supports several options for multi-tenancy that provide different blends of isolation, performance management, customization, self-service and cost. CSPs can determine which of these options meet the needs of their customers and develop offerings and price-points accordingly.

The three most widely employed models of XenApp multi-tenancy are shared (session layer isolation), partial isolation (machine layer isolation), or full isolation (network layer isolation).
Shared model

The shared (session layer isolation) model has one XenApp farm with infrastructure components and session hosts shared between tenants. Although not recommended from a best practice or security perspective, this is a common model in use for smaller providers today, particularly for those CSPs offering basic, standard desktop services where cost—not security—is the most significant business concern.

The key characteristics of this model are shown in the sliding scale to the right:

- Users from multiple tenants can have sessions on a single XenApp session host. This requires appropriate lockdown of XenApp session hosts to minimize the possibility of a user of one tenant negatively affecting the users of another tenant. However, there is still a chance that a user could compromise a server (thus affecting users of another tenant).
- User performance guarantees established by using the XenApp CPU Utilization Management feature.
- A separate Web Interface site can provide custom branding for each tenant. In addition, Windows and Citrix policies in Active Directory can provide a highly customized experience to users, for example, wallpaper, theme, Citrix HDX settings, and so on.
- This method of multi-tenancy is extremely cost effective because a CSP can spread all of the infrastructure costs across multiple tenants.
Partial isolation model

The partial isolation (machine layer isolation) model has one XenApp farm with shared infrastructure components, but the session hosts are not shared. More and more CSPs are moving to this method thanks to the new XenApp 6 architecture. Although it does not provide the extreme security found in the fully isolated model, this model provides arguably the most optimal blend of isolation, performance, customization, self-service admin and cost for most tenants—which translate into a very competitive price-point.

The key characteristics of this model are shown in the sliding scale to the right:

- Each tenant has a dedicated pool of session host servers so that users from different tenants do not have access to another tenant’s session hosts. XenApp Worker Groups and Worker Group Preference help easily create this deployment. As a best practice, administrators should still always lock down individual session hosts.
- Because users from one tenant can have sessions only on their designated XenApp servers, a user cannot negatively impact the performance of another tenant’s users. Administrators further guarantee performance to users by using the CPU Utilization Management feature in XenApp.
- In addition to the customization capabilities mentioned in the shared deployment, each tenant can have custom machine images for XenApp session hosts.
- CSPs can allow tenants to perform some level of administration for their pool of session hosts, for example, helpdesk activity for viewing which users are logged onto which servers, shadowing a session or resetting a session.
Though each tenant has dedicated session hosts, the costs might not be much higher than that of the shared model. This deployment method offers a blend of multi-tenancy capabilities at a very reasonable cost.

**Full isolation model**

In the full isolation (network layer isolation) model, one XenApp farm with dedicated infrastructure is deployed per tenant. None of the DaaS components are shared and in most cases each farm will reside on a dedicated vLAN or physical network. This model is best suited for tenants with stringent confidentiality and security requirements, such as federal agencies, healthcare, and so on, or those with heavy-duty performance or customization needs. These capabilities come at a cost, but most CSPs typically charge a premium for this type of service. It is understandably less common to see deployments of this nature – but important to understand that the option exists. This option is recommended for those environments where the tightest possible security, regardless of cost, is the primary requirement.

The key characteristics of this model are shown in the sliding scale to the right:

- Tenants are completely isolated, including dedicated brokering operations.
- Performance guarantees are similar to the partial isolation model.
- The customized experience aspects remain the same as that of the partial isolation model.
- Service providers have the option to allow the tenant to perform a much higher level of self-service administration for example, helpdesk activity, managing session hosts, managing applications, and so on.
The costs are higher for this model because the infrastructure components are not shared between tenants.

Creating the IaaS Layer

This section discusses the basic design considerations for the IaaS foundation that enables a secure and scalable DaaS solution. It begins at the lowest layer of the design and builds up from there, starting with the CSP datacenter network layout and ending with the virtual and application provisioning designs.

The layers addressed in this section are:

- Network boundaries, vLANs, and basic network services
- Active Directory and Organization Unit considerations
- Virtual Provisioning of DaaS workloads
- Application Provisioning

Note: We address the NetScaler, Branch Repeater, Access Gateway, and Web Interface designs in the DaaS Secure Access and Acceleration section after discussing DaaS subscriber partitions.
Network Boundaries and vLANs

Network design

The Citrix Reference Architecture for Multi-Tenant Desktop as a Service uses several secured zones to enable a flexible and scalable solution that provides for tight security at critical boundaries between the management layers and tenant partitions. These networks and zones are:

DMZ - Demilitarized zone

The DMZ is a network area isolated by firewalls on either side, providing a barrier between the public internet and the CSP datacenter operational environment. This network is the outside facing perimeter to the environment. The machines in this zone host the NetScaler and Access Gateway servers.

Tenant network

The tenant network is behind the DMZ and further isolated from the CSP network by a standard firewall. The firewall at this layer ensures that management communications flow between the CSP and tenant networks in a manner that ensures the multi-tenant isolation and CSP delegated administration appropriate for each tenant. The tenant network is further divided into separate vLANs:
**Application vLAN**

This vLAN contains the back-office applications that enable web, mail, collaboration and line of business application backend services. Leverage these vLANs and services across tenants or dedicate them to a specific tenant, as determined by the CSP and subscriber service level agreements.

**DaaS – tenant vLAN**

When tenants require a separate network segment for their dedicated XenApp servers, use per tenant vLANs. Service providers might choose to charge subscribers in these vLANs a slight premium for this level of security because of the slightly higher infrastructure and management costs to the CSP.

**DaaS – shared vLAN**

This vLAN presents a single network segment to XenApp SaaS and DaaS worker servers within the farm for those tenants that do not require the sense of security that a separate vLAN can provide. At the service provider’s discretion, subscribers in this vLAN might recognize a slight advantage regarding cost of service because much of the infrastructure, including the network, is shared among tenants and service providers can pass these savings to the subscriber.

**CSP network**

The CSP network is the area hosting all of the CSP global provisioning, authentication, management services and dashboards.

**Provisioning vLAN**

The provisioning vLAN provides a secure network for the single instance management images and infrastructure services that enable administrators to quickly provision XenApp and other tenant workloads from a shared read-only vDisk. This approach enables dynamic scaling of large-scale environments while using the minimal amount of backend storage and provides relatively simple life-cycle management as compared to most other methods.
Authentication vLAN
The authentication vLAN contains the Microsoft Active Directory Forest and Domains within a secured network. Only specific secured access to Active Directory capabilities are enabled in this vLAN for specific users, administrators and machines from the other networks.

Management vLAN
The management vLAN contains many of the foundational network services necessary in a hosted environment such as Domain Name Services, NTP and SNMP as well as other services provided by the CSP.
Physical connections and additional networks

Two additional networks are enabled specifically for performance reasons across the entire infrastructure: One network is dedicated to storage and the other to hypervisor management, in this example Citrix XenServer®.

**Storage** – The storage zone is an un-routable network that separates the storage and network data within the environment. All XenServer hosts have a direct connection to this zone for SAN access.

**XenServer Management** – The XenServer management zone separates XenServer-to-XenServer network traffic from the data and storage traffic.

A Cisco Catalyst 6513 layer-3 switch is physically connected to these networks to provide connectivity across them. The switch configuration secures these distinct networks through prescriptive route and access-list tables. The Cisco Catalyst 6513 is directly connected to a pair of Cisco Catalyst 3750G switches that provide connectivity to the individual XenServers. The above figure presents a view of the connections from the physical server to the switch.

This design delivers the foundation for a highly scalable system while providing physical separation necessary to meet the security requirements of a multi-tenant DaaS environment.

**Note:** The details of any particular 3rd party product are not provided as an endorsement or recommendation for the use of that product. These details are provided solely as a reference.
regarding the hardware and software used within the Citrix Cloud App Delivery Group lab environment when this document was written.

Each network in this sample uses a simple IP addressing scheme with a private network address set by RFC1918 and RFC4193. This design provides ample capacity for individual vLANs or security domains within the networks and eases the overhead of managing the routing and access-list tables.
Storage Configuration

In a Citrix DaaS environment, the availability and performance of the storage infrastructure are critical because thousands of users are impacted by storage outages or performance issues. Thus the storage architecture must provide the level of availability and performance typical for business-critical applications. Citrix recommends that CSPs choose a storage vendor that has software and hardware solutions to address the availability and performance requirements for large, scalable Citrix DaaS environments.

The Citrix Cloud App Delivery Group lab implementation of the Citrix Reference Architecture for Multi-Tenant Desktop as a Service employed a NetApp FAS 3170 SAN with dual controllers. Each controller is individually connected to the storage shelves and clustered to provide the maximum amount of redundancy in case of a controller failure. From the raw storage, a single aggregate (Agg0) was created for each controller. From the aggregates, NFS volumes were created and presented to the XenServers as storage repositories.

Hypervisor Host Configuration

The Citrix DaaS solution is hypervisor agnostic; a service provider can use any of the following hypervisors for the infrastructure or the XenApp servers that deliver the shared hosted desktops.

Citrix XenServer

Citrix XenServer is a complete, managed server virtualization platform built on the powerful Xen® hypervisor. Xen technology is widely acknowledged as the fastest and most secure virtualization software in the industry. XenServer is designed for efficient management of Windows and Linux® virtual servers and delivers cost-effective server consolidation and business continuity. More information on XenServer is available at [www.citrix.com/xenserver](http://www.citrix.com/xenserver).
Microsoft Hyper-V
Microsoft Windows Server 2008 R2 with Hyper-V® builds on the architecture and functions of Windows Server 2008 with Hyper-V by adding multiple new features that enhance product flexibility. Hyper-V is available in a Standard, Server Core and free Hyper-V Server 2008 R2 versions. More information on Hyper-V is on the company website.

VMware vSphere
VMware vSphere consists of the management infrastructure or virtual center server software and the hypervisor software that virtualizes the hardware resources on the servers. It offers features such as Distributed resource scheduler, vMotion, HA, Storage vMotion, VMFS, and a multipathing storage layer. More information on vSphere is on the company website.

The following server types were used to host the infrastructure for the entire Citrix DaaS solution.

Server hardware used in the Citrix lab implementation

2 x HP DL360
Dual – Quad core Intel E5335 @ 2.00 GHz
16GB – PC3 1066MHz
6 x 1GbE NICs

10 x HP DL380
Dual – Quad core Intel E5335 @ 2.00 GHz
96GB – PC3 1066MHz
6 x 1GbE NICs
## Securing the Networks

Administrators can secure communications between the various CSP networks through firewall configurations that deny all traffic to those networks other than that necessary to provide and manage secure DaaS services across multiple tenants.

<table>
<thead>
<tr>
<th>Port</th>
<th>Source to Destination</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP: 443</td>
<td>Public Internet to DMZ</td>
<td>End-point to Access Gateway SSL connectivity</td>
</tr>
<tr>
<td>TCP: 1494/2598</td>
<td>DMZ to Tenant Network</td>
<td>Access Gateway to XenApp Session, HDX/ICA® connectivity for endpoints. This protocol is encapsulated in 443/SSL from the end-point to Access Gateway in the line item above.</td>
</tr>
<tr>
<td>TCP: 80/443</td>
<td>DMZ to Tenant Network</td>
<td>Citrix Web Interface connections</td>
</tr>
<tr>
<td>TCP/UDP: 389/636/3268</td>
<td>DMZ to Management Network</td>
<td>Active Directory communications and authentication</td>
</tr>
<tr>
<td>TCP: 53</td>
<td>DMZ to Management Network</td>
<td>DNS</td>
</tr>
<tr>
<td>TCP: 123</td>
<td>DMZ to Management Network</td>
<td>NTP</td>
</tr>
<tr>
<td>TCP: 161-162</td>
<td>DMZ to Management Network</td>
<td>SNMP</td>
</tr>
<tr>
<td>Port</td>
<td>Source to Destination</td>
<td>Purpose</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>UDP: 6901-30 / 10802-3</td>
<td>Tenant Network to Management Network</td>
<td>Provisioning Server vDisk streaming</td>
</tr>
<tr>
<td>SMB: 445</td>
<td>Tenant vLANs to Application vLAN</td>
<td>File Shares and Citrix Streaming</td>
</tr>
<tr>
<td>TCP/UDP: 389/636/3268</td>
<td>Tenant Network to Management Network</td>
<td>Active Directory communications and authentication</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Port</th>
<th>Source to Destination</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP: 53</td>
<td>Tenant to Management</td>
<td>DNS</td>
</tr>
<tr>
<td>TCP: 123</td>
<td>Tenant to Management Network</td>
<td>NTP</td>
</tr>
<tr>
<td>TCP: 161-162</td>
<td>Tenant to Management Network</td>
<td>SNMP</td>
</tr>
<tr>
<td>TCP: 1433-1434</td>
<td>Tenant to Management Network</td>
<td>SNMP</td>
</tr>
<tr>
<td>TCP: 80/443/9035</td>
<td>Tenant to Management Network</td>
<td>Citrix EdgeSight Services</td>
</tr>
<tr>
<td>TCP: 11161</td>
<td>Tenant to Management Network</td>
<td>Citrix Power and Capacity Management</td>
</tr>
<tr>
<td>TCP: 2512</td>
<td>Tenant to Management Network</td>
<td>Citrix Independent Management Architecture communications</td>
</tr>
<tr>
<td>TCP: 135 / 27000</td>
<td>Tenant to Management Network</td>
<td>Citrix Licensing Services</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Software Used</th>
<th>Hardware Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Cisco IOS</td>
</tr>
<tr>
<td>Hardware Used</td>
<td>Cisco Catalyst 6513</td>
</tr>
<tr>
<td></td>
<td>Cisco Catalyst 4900M</td>
</tr>
<tr>
<td></td>
<td>Cisco Catalyst 3750G</td>
</tr>
</tbody>
</table>
Active Directory and Organization Unit Considerations

Windows-based DaaS and SaaS are inherently based upon Microsoft Windows principals and technologies. As a result, Microsoft Active Directory plays a critical role in many aspects of the Citrix DaaS designs, most notably as the central account authority and as a foundational element to organizational and multi-tenant capabilities.

The above graphic highlights the complementary structures within the Citrix XenApp 6 Delivery Services Console (DaaS/SaaS publishing and management) and the Microsoft Active Directory Console (Account Authority and organizational structure). Sub-nodes in each tree view map to tenant specific groups. In fact, the Citrix nodes are logical administrative folders that map directly to the Active Directory Organization Units and/or child domains that provide the basic tenant partitioning and security from an account perspective for this solution.

This section focuses solely on the Active Directory design; the XenApp complement is discussed in the Hosted Applications section later in this document.
Active Directory Organization Units and Group Policy Objects

The Active Directory (AD) considerations for the DaaS Reference Architecture follow Microsoft recommended best practices for AD designs. We have also followed current industry best practice implementations for the SMB focused Citrix Service Providers as understood through conversations with providers across the spectrum over the last few years. The result of this combined set of considerations points to the Organization Unit (OU) as the preferred tenant partitioning mechanism within AD for most instances.

As discussed in the earlier section on multi-tenancy, some subscribers might be better served by a dedicated child domain, or in extreme cases their own isolated AD forest and all other infrastructure. Those two scenarios represent the edge case in the SMB-focused service provider market, and as such are not addressed in detail in this reference architecture. The context of this document and all further design considerations within this document use a single AD OU structure as its basis.

As shown in the graphic to the right, we created a separate OU for each tenant, labeled Tenant1, Tenant2, and so on. This enables the leverage of AD Group Policy Objects (GPOs) as a foundational mechanism for assigning properties to each tenant. Doing so ensures easy, quick administration and enforcement of Service Level Agreements (SLAs) regarding the configuration and security of each tenant.

Subordinate OUs to each tenant enable further granularity of management with regards to separate applications, desktops and users/user groups within a single tenant. This OU structure provides a clearly navigable solution for troubleshooting individual tenant configurations, while also leveraging the cascading aspects of Group Policy.

The advantage of cascading Group Policy in a multi-tenant environment is that it enables the least amount of customization with the greatest impact per tenant or SLA. For example, a CSP might begin their service offering with a generic GPO configuration that is applied at the CSP OU (CSPDemo.com in the graphic). This GPO is applied to all subordinate OUs (Tenants) by default. Any tenant that requires a modification to this generic base GPO is then assigned an additional GPO for their OU container that contains only those configuration customizations that augment the generic offering. Troubleshooting a particular tenant configuration thus usually only requires consideration of the GPO for that tenant.

**Active Directory infrastructure**

<table>
<thead>
<tr>
<th>Software Requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Microsoft Windows Server 2008 R2</td>
</tr>
<tr>
<td>Installed Roles</td>
<td>Active Directory Services</td>
</tr>
<tr>
<td>Additional Software</td>
<td>None</td>
</tr>
<tr>
<td>Hardware Used</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Hypervisor</td>
<td>XenServer 5.6 SP1</td>
</tr>
<tr>
<td>Virtual Server Specifications</td>
<td>4 vCPU w/ 4GB RAM</td>
</tr>
</tbody>
</table>

The sample implementation includes two AD domain controller virtual machines provisioned with the global catalog role in the management zone to provide authentication services to the tenants. Each virtual machine is sized to support up to 10,000 users. Additional domain controllers are required if the CPU within the virtual machine exceeds 50%.

For more information regarding Microsoft Active Directory designs, see: http://technet.microsoft.com/en-us/library/cc268216.aspx
Virtual Provisioning of DaaS Workloads

**Hypervisor**

Virtualization is the fundamental enabler of an efficient cloud datacenter. As a best practice, enabling dynamic scale and simplified management within the Citrix Reference Architecture for Multi-Tenant Desktop as a Service, all workloads are virtualized except for Citrix Provisioning Services™ (discussed in the next section). Because the reference architecture is hypervisor agnostic, no proprietary capabilities within any of the more popular hypervisors are discussed. Each VM is simply used as a compute engine and container for workloads within the reference architecture. An example configuration as implemented within the Citrix Cloud App Delivery Group labs provides a guide for efficient network and storage configuration of the host servers.

**VM Workload provisioning**

One of the primary enablers of efficient scale and management within the DaaS solution is a robust yet simple workload provisioning system. Although there are many approaches to provisioning, from physical unattended installations to VM cloning and other more recent technologies, Citrix recommends Citrix Provisioning Services (PVS) for the most efficient scale and simplified life-cycle management of XenApp and other DaaS enabling workloads.

The remainder of this section describes the implementation within the reference architecture.
Note: Because of the intense IO requirements for vDisks in this environment Citrix recommends installing the PVS on a physical server.

Provisioning Server infrastructure

<table>
<thead>
<tr>
<th>Software Requirements</th>
<th>Microsoft Windows Server 2008 R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Microsoft Windows Server 2008 R2</td>
</tr>
<tr>
<td>Installed Roles</td>
<td>File Services (NFS Client)</td>
</tr>
<tr>
<td>Additional Software</td>
<td>Citrix Provisioning Services 5.6 SP1</td>
</tr>
<tr>
<td>Server(s)</td>
<td>2 x HP DL380 G6</td>
</tr>
<tr>
<td>Server</td>
<td>4 x 2.40 GHz hex core CPUs, 64GB RAM, 8 x 1GBit/s NICs (teamed for throughput)</td>
</tr>
<tr>
<td>Scalability</td>
<td>1,000+ target machines</td>
</tr>
</tbody>
</table>

Provisioning Services streaming technology

The Provisioning Services infrastructure is based on software-streaming technology. This technology allows computers to be provisioned and re-provisioned in real-time from a single shared-disk image. In doing so, administrators can completely eliminate the need to manage and patch individual systems. Instead, all image management is done on the master image. The local hard-disk drive of each system can be used for runtime data caching or, in some scenarios, removed from the system entirely, which reduces power usage, system failure rates and security risks.

Provisioning Services solution

Using Provisioning Services, administrators prepare a device (master target device) for imaging by installing any required software on that device, create a vDisk image from the master target device’s hard-disk drive, and saves the image to the network (on a Provisioning Server or storage device).

When the vDisk is available from the network, the target device no longer needs its local hard-disk drive to operate; it boots directly across the network. The Provisioning Server streams the contents of the vDisk to the target device on demand in real time. The target device behaves as if it is running from its local drive. Unlike thin-client technology, processing takes place on the target device.
Provisioning Services implementation

The PVS servers are directly connected to the core switch and storage network with separate network adaptors. This design separates the I/O traffic for vDisk input (loaded from the NFS share) and vDisk output (delivered to the Target Devices), providing a high level of performance (maximum IOPs), reliability (redundant devices and network connections) and scalability (increased target devices).

The implementation includes two physical servers running Windows Server 2008 R2 and the File Services Role, with Provisioning Services installed.


**Note:** Use Network Interface Card (NIC) – Teaming to increase the reliability and the I/O between the Provisioning Servers, File Server and Target Devices. Also, use dedicated NICs for loading the vDisks and for delivering the vDisks to the Target Devices.
The following diagram outlines a basic Provisioning Server and Windows Network File System (NFS) architecture:

![Diagram of Provisioning Server and NFS architecture]

The SAN is configured to present NFS volumes to the PVS servers.

This scenario compares the PVS server’s IOPS usage when streaming one XenApp server versus ten XenApp servers running a medium Microsoft Office 2007™ workload. The workload was generated using EdgeSight® for Load Testing (ESLT) version 3.6 to scale up to fifty users per server over a fifteen-minute period. The test was run with the PVS cache located on the device hard disk, the recommended deployment configuration when streaming a XenApp workload.

The data below captures the start of the workload as users initially begin to log on to the XenApp servers and continues for the duration of the test. The ESLT scripts perform the following actions:

1. Scale up to the maximum defined user count over fifteen minutes.
2. Continue to run at the maximum defined user count for five additional minutes.
3. Scale down to no users over five minutes.
The graph contains only the IOPS of the vDisk. The write cache and operating systems are located on separate physical drives to allow their IO to be analyzed independently.

There is little to no disk IO regardless of the number of XenApp servers and users involved in the test. The initial spike in IOPS was relatively small compared to the peak observed during the boot scenario.

Best Practices for Configuring Provisioning Server on a Network
http://support.citrix.com/article/CTX117374

Hotfix CPVS56SP1E029 - For Citrix Provisioning Services 5.6 SP1
http://support.citrix.com/article/CTX129381
Application Provisioning

Application provisioning within the DaaS reference architecture provides a key element to the “Dynamic Assembly” capabilities within the system. Dynamic assembly is the process by which separate elements are combined in real-time to present a user with their specific, familiar, and personalized environment of operating system, desktop, application and personalization settings.

Application virtualization and streaming

Application virtualization provides one of the key components of dynamic assembly, the separation of applications from the underlying OS. Separating the application from the OS enables management of the application and its total life-cycle as a discreet object. A further advantage to this separation of OS and application is the ability to deliver and manage a single application image across CSP tenants, personalized for each tenant’s SLA through the policies associated with that tenant’s worker group based partition.

Within the XenApp system, Citrix streaming and Microsoft® App-V combine to provide a complete and seamless DaaS and SaaS solution with mature and proven application compatibility.

Packaging applications with the profiler/sequencer
To enable single instance management of applications from a single application hub (AppHub) across tenants within a Citrix Service Provider datacenter, package the applications into a virtualized instance using either the Citrix Streaming Profiler or Microsoft App-V Sequencer. For details and best practices of the sequencing and profiling process, refer to the Citrix Streaming and Microsoft App-V integration kit documentation found in the Citrix eDocs “XenApp 6” library at http://www.citrix.com/edocs.

The resources required for both of these packaging utilities are basically the same, with only a few differences regarding the underlying machine configurations.

### Citrix Streaming Profiler Requirements

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Requirement Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Windows Server 2008 R2</td>
</tr>
<tr>
<td>Installed Roles</td>
<td>Roles as installed on the XenApp worker servers</td>
</tr>
<tr>
<td>Additional Software</td>
<td>Citrix Streaming Profiler</td>
</tr>
<tr>
<td><strong>Hardware Used</strong></td>
<td></td>
</tr>
<tr>
<td>Server</td>
<td>2 vCPU, 4GB RAM</td>
</tr>
</tbody>
</table>

### Microsoft App-V Sequencer Requirements

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Requirement Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Windows Server 2008 R2</td>
</tr>
<tr>
<td>Installed Roles</td>
<td>Roles as installed on the XenApp worker servers</td>
</tr>
<tr>
<td>Additional Software</td>
<td>Microsoft App-V Sequencer</td>
</tr>
<tr>
<td><strong>Hardware Used</strong></td>
<td></td>
</tr>
<tr>
<td>Server</td>
<td>2 vCPU, 4GB RAM</td>
</tr>
</tbody>
</table>

### The AppHub

The AppHub provides the central storage and streaming services that enable delivery of a single application image into the DaaS/SaaS worker servers. The image is assembled on-demand for each user. Within the Citrix Reference Architecture for Multi-Tenant Desktop as a Service implementation, the AppHub has two enabling technologies: Standard file storage for the images themselves and the Microsoft App-V Streaming Server.

### Streaming Server Software Requirements

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Requirement Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Microsoft Windows Server 2008 R2</td>
</tr>
<tr>
<td>Installed Roles</td>
<td>Web Server Role</td>
</tr>
<tr>
<td>Additional Software</td>
<td>Microsoft App-V 4.6 Streaming Server</td>
</tr>
<tr>
<td><strong>Hardware Used</strong></td>
<td></td>
</tr>
<tr>
<td>Server</td>
<td>2 vCPU, 4 GB RAM, Network Attached Storage</td>
</tr>
</tbody>
</table>
Delivering applications to tenant users and groups

After the application image is uploaded to the AppHub, the next step is to publish it for subscriber access. The Citrix Delivery Services Console is the central administrative interface for configuring published applications across all tenants. The published applications are then streamed into the hosted desktop through the Offline and App-V plug-ins for Citrix Receiver. For more information about the Delivery Services Console, see “Citrix Delivery Services Console” later in this document.

See the knowledge base article: “How to Sequence an App-V Virtual Application and Stream XenApp Server to be Published Seamlessly to Users” [http://support.citrix.com/article/CTX126082](http://support.citrix.com/article/CTX126082)
Hosted applications execute on XenApp servers and are presented to the desktop through Citrix Receiver and HDX technologies. The applications themselves are either installed on the XenApp worker servers or virtualized and streamed into those servers as described in the preceding section. Because XenApp infrastructure and workers are leveraged in much the same way for both SaaS and DaaS delivery, we discuss configuration and scalability of Hosted Applications in Creating the Multi-Tenant DaaS and SaaS Layer section, next.
Creating the Multi-Tenant DaaS and SaaS Layer

The Citrix XenApp 6 farm is the primary building block for the multi-tenant DaaS and SaaS module within the reference architecture. The farm is divided into various modules and partitions to provide a scalable yet simplified management scope for over 1,000 XenApp workloads providing DaaS and SaaS for over 100,000 active users per farm/block.

The major modules and components of the Citrix Farm are:

- XenApp Data store
- XenApp Data collector
- XenApp Workers
- XenApp WorkerGroups
- Citrix License Server
- Microsoft Active Directory Data store
IMA data store

The Independent Management Architecture (IMA) data store is a central repository for all of the configuration information for the XenApp farm. This includes items such as published applications, worker groups and load evaluators. During server startup, the IMA Service queries the data store for initialization information. This is the most CPU-intensive action for the data store, as the IMA Service initialization process ensures that the local host cache (LHC) is consistent with the data store. When multiple servers boot, multiple simultaneous requests for initialization information are made to the data store.

During normal farm operation, each server accesses the data store every thirty minutes to ensure its LHC is current. The data store is also accessed if the Delivery Service Console (DSC) or other Citrix query-based utilities modify the farm configuration or request static information. However, the data store is not accessed when a user logs in, disconnects, or reconnects to the farm. All the information needed for a client to establish a connection to a XenApp server is stored in the LHC, with the exception of licensing details.

<table>
<thead>
<tr>
<th>Data Store Software Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
</tr>
<tr>
<td>Installed Roles/Features</td>
</tr>
<tr>
<td>Additional Software</td>
</tr>
<tr>
<td>Hardware Used</td>
</tr>
<tr>
<td>Server</td>
</tr>
<tr>
<td>Scalability</td>
</tr>
</tbody>
</table>

Data collector

The data collector manages all of the dynamic information in the farm. Dynamic information consists of items that change often such as connected sessions, disconnected sessions and server loads. The data collector is responsible for knowing the global state of the farm. The data collector also performs resolutions—a process where, upon user request, the data collector determines the least-loaded server that is hosting a load-balanced published application or desktop.

Sizing guidelines

The data collector stores all dynamic information in memory therefore the data collector needs enough RAM to store all of the records. Memory usage varies based on the number of published applications, number of servers and number of user sessions in the farm. The CPU plays an important role in determining the number of resolutions the data collector can process in conjunction with managing dynamic information.
### Data Collector Software Requirements

<table>
<thead>
<tr>
<th>Operating System</th>
<th>Microsoft Windows Server 2008 R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Installed Roles/Features</td>
<td>During a wizard-based XenApp installation, the Server Role Manager (using the Server Role Installer) automatically installs all prerequisite software and Windows Server roles.</td>
</tr>
<tr>
<td></td>
<td>For more information on XenApp server installation: <a href="http://support.citrix.com/proddocs/topic/xenapp6-w2k8/ps-system-requirements-w2k8-xa6.html">http://support.citrix.com/proddocs/topic/xenapp6-w2k8/ps-system-requirements-w2k8-xa6.html</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional Software</th>
<th>XenApp 6</th>
</tr>
</thead>
</table>

### Hardware Used

<table>
<thead>
<tr>
<th>Server</th>
<th>4 vCPU, 8GB RAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalability</td>
<td>1,000 XenApp Workers</td>
</tr>
</tbody>
</table>

### License Server

The license server stores and manages Citrix licenses. When users connect to a XenApp server, the XenApp server checks out a license from the license server on behalf of the client device. Subsequent connections from the same client device share the same license.

### Sizing guidelines

One of the most important considerations in determining license server requirements is processor speed. Although CPU usage is not usually high, CPU time increases as license checkout requests are made and License Management Console activity increases. The time it takes to execute these transactions is dependent on the speed of the CPU. In general, the size of the farm and the number of simultaneous client connections dictate the power of the server needed for the licensing feature.

To appropriately size the license server, determine the number of client logins per second in the farm deployment by using the Performance Monitor counters available within XenApp and the load evaluator logging feature. This analysis determines the processor speed needed for optimal license server performance.

Additionally, the license server process is single threaded, so multiple processors do not increase performance. The license server uses approximately 4.5KB of memory for every session license and 39KB of memory for every start-up license that is in use. The license server is capable of processing 248 license checkout requests per second. In a scenario where all users log in over the course of thirty minutes, a single license server can handle 446,400 users.
Citrix Delivery Services Console (DaaS and SaaS configuration and management)

The Delivery Services Console is the user interface of the Multi-Tenant DaaS and SaaS “Control Module.” A Microsoft Management Console (MMC) snap-in enables CSPs to perform a number of DaaS and SaaS management functions.

CSPs can configure…
- XenApp servers
- Server farms
- Published desktops
- Published applications
- Policies
- Printers
- Load balancing

…and monitor:
- Alerts
- Hotfix information
- Administrative changes

Install the Delivery Services Console on a standalone workstation VM or publish it from a XenApp Controller Server within the Control Module.

### License Server Software Requirements

<table>
<thead>
<tr>
<th>Operating System</th>
<th>Microsoft Windows Server 2008 R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Installed Roles</td>
<td>None</td>
</tr>
<tr>
<td>Additional Software</td>
<td>Citrix License Server 11.9</td>
</tr>
</tbody>
</table>

### Hardware Used

| Server | 4vCPU, 4 GB RAM |
| Scalability | 446,400 users |
Microsoft .NET Framework 3.5 SP1
• Microsoft Windows Installer (MSI) 3.0
• Microsoft Windows Group Policy Management Console
• Microsoft Visual C++ 2005 SP1 Redistributable (x64)
• Microsoft Visual C++ 2008 SP1 Redistributable (x64)
• Microsoft Visual C++ 2008 SP1 Redistributable
• Microsoft Visual C++ 2005 SP1 Redistributable
• Microsoft Primary Interoperability Assemblies 2005

Hardware Used

<table>
<thead>
<tr>
<th>WorkStation minimum</th>
<th>Varies per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalability</td>
<td>Farm-wide management</td>
</tr>
</tbody>
</table>

Active Directory Integration, XenApp Worker Groups and Policies

The release of XenApp 6 adds powerful new features for XenApp administrators through AD integration. Administrators can now manage all user and server settings through AD policies and can manage published applications, hosted desktops, load balancing and farm multi-tenancy through a new container known as a worker group.

![Worker groups and tenant silos](image)

A worker group is a collection of XenApp servers in the same farm, where administrators can associate objects such as published applications, published desktops, and policies. Worker groups allow a set of similar servers to be grouped and managed as a single entity. Worker groups are closely related to the concept of application silos. However, they streamline the creation of application silos by providing a way to synchronize the published applications and server settings across a set of XenApp servers.
Worker groups are dynamic. For example, when AD containers are associated with a worker group, changes in the AD container are automatically reflected in the server's worker group memberships.

Server groups can be added to worker groups by AD Organizational Units or Server Groups. This provides for the dynamic update of worker groups based on AD memberships of servers. That is, as servers are added or removed from the AD containers, they are automatically added to or removed from the respective worker groups.

In anticipation of future expansion, two sets of worker groups are created: One set to group servers by tenant and one set to group applications by tenant. When the administrators add capacity for an existing tenant, they do not need to modify the servers list of published applications or desktops assigned to that tenant. Instead, they simply add another XenApp server to the tenant’s OU.

With dynamic provisioning, AD use automates this step by creating a base image for XenApp with all of the applications installed. To add capacity, simply create a new instance of the base image and add it to the desired tenant OU. The server receives its server settings from AD, joins the appropriate worker groups and begins hosting published applications or desktops. Creating separate worker groups for desktops and applications gives CSPs the flexibility to easily expand their tenant base.

Worker groups and Citrix policy filters
CSP administrators can filter all XenApp server policies by worker groups, thereby restricting Group Policy Objects (GPOs) to a specific set of servers in the farm. For policies configured in the Delivery Services Console, this is the only way to assign different settings to different groups of servers because all policies are replicated to all servers, completely independent of AD.

AD GPOs are used to manage the settings in the XenApp farm. For all user and site settings GPOs can be linked to the XenApp OUs without any filters. However, if a setting is required specifically for a particular tenant’s servers, a Worker Group filter added to the policy limits it to the appropriate tenant.

Citrix policy configuration

![Active Directory Group Policy](image)

In XenApp 6 nearly all server, farm and user settings are governed by group policies.

Administrators create a GPO containing the desired Citrix policy settings and link the GPO to the appropriate tenant OUs. However, for Citrix administrators who do not have control over their AD environment or whose organizations do not use AD for directory services, XenApp 6 provides farm-based group policies through the policies node in the management console. Such policies are written to the XenApp data store and propagated to all servers in the farm.

Administrators create Citrix policies at different OU structure levels as shown in the following graphic. In this case, the priority of policies enforcement is as follows:

1. Policy created at the Default Domain Policy
2. Policy created at the top OU level
3. Policy created at the middle level OU
4. Policy created at lowest level OU
The XenApp General GPO applies to the XenApp farm as a whole. For each tenant, the CSP administrator creates new GPOs or links existing GPOs to the tenant’s OU structure. For example, the Tenant1 GPO is a general tenant GPO created to apply policies to all of the Tenant1 downstream OUs. The CTXRestrictedComputer GPO is linked to the Tenant1 computer OU (Tenant1_Computers) and CTXRestrictedUser and XASession GPOs are linked to the Tenant1 user OU.

The resultant Citrix policies applied to the Tenant1 computer and user OUs are the merged settings from all four GPOs. If there is a conflict among the policy settings of these GPOs, the settings in the Computer and User GPOs have the highest priority and overwrite the settings in the Tenant1 GPO, XenApp General GPO and Domain GPO.
DaaS Secure Access and Acceleration

For the sake of clarity in the integration flow we now return to the IaaS module to integrate NetScaler and Access Gateway within our multi-tenant DaaS and SaaS module.

Citrix NetScaler and Access Gateway

Citrix NetScaler presents a modular platform upon which several critical network security and acceleration functions are built. For those CSPs focused on the small to medium business, the Citrix Access Gateway capabilities within NetScaler are fundamental to the secure delivery of desktops and applications as a service.

Access Gateway multi-tenancy with NetScaler

The Access Gateway functionality within Citrix NetScaler VPX™ (Virtual Appliance) provides secure access to the DaaS environment over SSL (TCP 443) across the public internet. Access Gateway multi-tenant support is implemented either within a single NetScaler VPX, physical appliance HA pair, or across segregated networks and vLANs through the use of a dedicated NetScaler appliance per tenant vLAN. From a software configuration perspective all of these scenarios are fundamentally the same with regards to the integration points between the DMZ and Multi-Tenant DaaS network.

Implementation within the Citrix Cloud App Delivery lab environment

NetScaler - virtual servers

Not to be confused with a XenServer guest, the Access Gateway Enterprise Edition virtual server is an entity within a NetScaler VPX that is a representative of all the configured services available to clients. The virtual server is also the access point through which clients access these services. Configuring multiple virtual servers on a single appliance allows one Access Gateway appliance to serve multiple user communities (Tenants in our example) with differing authentication and resource access requirements.
NetScaler - profiles and policies

Configuration of authentication, authorization and accounting (AAA) allows users to log on to the Access Gateway with credentials that are recognized by either the Access Gateway or by authentication servers, such as LDAP or RADIUS, located in the secure network. Authorization policies define user permissions, determining which resources a given user is authorized to access.

The Citrix Reference Architecture for Multi-Tenant Desktop as a Service uses a single AD domain with multiple OUs; each OU represents one tenant. The figure to the right shows the AD infrastructure. The “servers” OU for each tenant contains that tenant’s dedicated XenApp servers and the “users” OU contains that tenant's user accounts. Each tenant also has an associated AAA policy that enables subscribers to securely and seamlessly log on to their DaaS and SaaS resources.

A session profile contains the settings for client connections and they are associated to session policies. You can create profiles separately from the policy using the configuration utility and then use the profile for multiple policies. Only one profile can be used with a policy.

A single Citrix XenApp farm with a load balanced pair of Microsoft Windows Server 2008 R2 servers running Citrix Web Interface supports all tenants in this implementation. Additional Web Interface servers can provide greater scale as more tenants are brought on-board.

A NetScaler HA VPX pair, enabling Citrix Access Gateway capabilities is located in the DMZ to provide secure access for tenants to their resources.

<table>
<thead>
<tr>
<th>NetScaler AG VPX Configuration</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operating System</strong></td>
<td>NetScaler VPX</td>
</tr>
<tr>
<td><strong>Hardware Used</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Server</strong></td>
<td>4vCPU, 4 GB RAM</td>
</tr>
<tr>
<td><strong>Scalability</strong></td>
<td>300 Concurrent SSL-VPN Connections</td>
</tr>
</tbody>
</table>
A firewall appliance in front of the Access Gateway provides NAT from the external Internet to the internal Access Gateway interface (virtual server) and blocks all unwanted traffic from entering the CSP environment from the public network.

Integration with Web Interface

After the Access Gateway is configured it must be associated with a Citrix Web Interface site to provide subscriber access to DaaS and SaaS subscribed services.

In this typical scenario, tenants use a web browser or end-point running Citrix Receiver to access a secure logon point hosted by the CSP. After providing the appropriate credentials, users are directed to their organization’s customized Citrix Web Interface to access their subscribed resources.

Dashboards and Management

Administering a complete DaaS and SaaS offering can be a significantly complex consideration. The right tools and dashboard views are critical to simplifying and delegating administration for these solutions. Several core consoles enable a deep level of native control across the various components within the system. In addition, the Cortex Cloud Control Panel from Citrix enables delegated provisioning and administration across these components from a simplified interface. This section of the reference architecture provides technical data and guidance regarding some of the core native consoles, as well as a look at the Cortex Cloud Control Panel as an example of the cross product integration and delegated administration capabilities that enable simplified management of these multi-tenant systems.
Citrix EdgeSight (monitoring, reporting and troubleshooting)

EdgeSight integration provides HDX monitoring and troubleshooting capabilities. Use of EdgeSight can also simplify Citrix CSP License reporting, for use in CSP billing systems as well as for their monthly license reporting.

The EdgeSight inherent multi-tenant architecture enables CSPs to delegate certain reporting and monitoring capabilities to their tenants.

Because this component is shared across the entire CSP infrastructure, the EdgeSight Server is integrated within the Management vLAN of the CSP Network. The CSP can grant EdgeSight web console access from any client or DaaS session. All traffic to and from the EdgeSight service is secured through the Firewall and vLAN configurations as discussed at the beginning of this document.

CSP-specific report templates are available for download from the Citrix Developer Network CSP community site at http://community.citrix.com/p/csp#.

For more information you can also download the Citrix Service Providers Guide to using Citrix EdgeSight from http://www.citrix.com/skb/articles/RDY2947.
Cortex Cloud Control Panel
(provisioning, delegated administration and tenant self service)

Cortex Cloud Control Panel from Citrix provides CSPs with a single integrated interface for the management of desktops, applications and backend services.

Cortex user interface (CortexWeb)

Cortex provides a single unified interface providing both system administration and delegated administration to resellers and end-customers. The web application has three main methods of integration with other Cortex components:

- SQL Database - for configuration, users, customers, auditing and reporting.
- Web Services - for real time interaction with Active Directory and other hosted services.
- Provisioning Engine - using MSMQ, for dispatch of provisioning requests to the provisioning engine.

The Cortex web application is loosely coupled with the other Cortex components. This loose coupling provides several security benefits. The web server has no dependency on Active Directory so it can essentially operate outside of the managed domain. The website can be locked down and run with minimal administrative permissions while still allowing the Cortex system to complete administrative tasks.

Cortex system databases (CortexSQL)

SQL Server provides the backbone of the Cortex system. The database stores configuration information for all services provisioned by Cortex, as well as all customer and user details. The SQL
database also acts as a cache mechanism for Active Directory ensuring rapid user response without the need for slower AD queries.

The Cortex databases also store logging and auditing information for all provisioning transactions that pass through the system.

**Cortex provisioning engine**

The Cortex provisioning engine runs as a Windows NT Service. It monitors one or more provisioning queues for provisioning requests. When the provisioning engine receives a request, it uses provisioning rules to determine the actions required to complete the provisioning.

The provisioning rules are easily customized using a simple Windows-based graphical interface that also provides a simple way to understand specific provisioning processes, helpful for problem diagnosis.

Each provisioning action performs a reusable piece of work, typically associated with provisioning applications. Cortex includes over 100 provisioning actions. Example actions include:

- Creating an Active Directory user
- Creating a security group in Active Directory
- Creating a folder in a file system
- Creating an address list in Microsoft Exchange
- Running a shell command or a visual basic script

All Cortex provisioning processes are built using provisioning actions, enabling quick setup with little coding, while giving the service providers visibility into the processes being executed in their environment.

**Active Directory web service (ADWS)**

The Active Directory web service provides a secure and simple interface to Active Directory. The Cortex website uses this service to perform real time tasks such as user authentication and password expiry status.
Reporting

Cortex uses Microsoft SQL Server reporting services to deliver usage reporting capability through the Cortex user interface. Cortex interacts directly with the reporting services web service interface and allows controlled publishing of reports to all users of the Cortex system.
Conclusion

Companies of all sizes are looking for a smarter approach to managing the applications and data they use to run their business. More devices, more applications and more places to work means business owners have to spend an increasing amount of time on IT. Citrix Service Providers can shift the focus for their subscribers back to where it matters the most—growing the business. By offering a bundle of applications, desktops, and IT services, customers get what they want in a familiar, pay-as-you-go subscription model.

This Citrix Reference Architecture for Multi-Tenant Desktop as a Service represents a common view of those best practices as recommended by Citrix and employed by some of today’s most successful Citrix Service Providers. With core architectural innovations brought to market in XenApp 6, a reliable, scalable, and high-performance solution is now available for those CSPs wanting to provide monthly subscriptions to Windows based SaaS and DaaS services. The CSP XenApp Premium licensing program provides the foundation for aggregating over 1,000 servers across multiple tenants into a single management scope, ultimately providing a solution that enables service providers to build a flexible, scalable, and cost effective architecture to meet their customers’ and business needs.
Appendix A: Online Resources

The Citrix Service Provider Toolkit
http://community.citrix.com/kits/#/kit/734024

Top 10 Considerations for Delivering Desktops in the Cloud
http://support.citrix.com/article/CTX128899

XenDesktop Modular Reference Architecture
http://support.citrix.com/article/CTX124087

How to Deliver a Cloud Desktop using XenApp 6

Scaling Big – SaaS and DaaS Deployments for Citrix Service Providers
http://support.citrix.com/article/CTX129106

Citrix Service Providers Guide to using Citrix EdgeSight
http://www.citrix.com/skb/articles/RDY2947

Architecture planning for Citrix Service Providers
http://www.citrix.com/skb/articles/RDY2524
About Citrix

Citrix Systems, Inc. (NASDAQ:CTXS) is a leading provider of virtual computing solutions that help people work and play from anywhere on any device. More than 230,000 enterprises rely on Citrix to create better ways for people, IT and business to work through virtual meetings, desktops and datacenters. Citrix virtualization, networking and cloud solutions deliver over 100 million corporate desktops and touch 75 percent of Internet users each day. Citrix partners with over 10,000 companies in 100 countries. Annual revenue in 2010 was $1.87 billion. Learn more at www.citrix.com.

©2011 Citrix Systems, Inc. All rights reserved. Branch Repeater™, Citrix®, Citrix Access Gateway™, Citrix Merchandising™ Server, Citrix Provisioning Services™, Citrix Receiver™, EdgeSight®, HDX™, ICA®, NetScaler®, VPX™, XenApp™, XenDesktop®, and XenServer® are trademarks or registered trademarks of Citrix Systems, Inc. and/or one or more of its subsidiaries, and may be registered in the U.S. Patent and Trademark Office and in other countries. Xen® is a trademark of Citrix Systems, Inc. managed on behalf of Xen.org, and may be registered in the U.S. Patent and Trademark Office and in other countries. All other trademarks and registered trademarks are property of their respective owners.