

Infrastructure as a Service Technical Guidance for Hosting Service Providers

Published:

June 14, 2013

For the latest information, see <http://www.microsoft.com/hosting/en/us/default.aspx>. If you have any feedback on this document, please send e-mail to SolutionsFeedback@Microsoft.com and reference the title above.

Copyright

This document is provided "as-is". Information and views expressed in this document, including URL and other Internet Web site references, may change without notice.

Some examples depicted herein are provided for illustration only and are fictitious. No real association or connection is intended or should be inferred.

This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes. You may copy and use this document for your internal, reference purposes.

© 2013 Microsoft Corporation. All rights reserved.

Microsoft, Active Directory, Hyper-V, SQL Server, SharePoint, Windows, Windows Azure, Windows PowerShell, and Windows Server are trademarks of the Microsoft group of companies. All other trademarks are property of their respective owners.

Contents

- 1.0 Introduction..... 4
- 2.0 Problem Definition..... 5
- 3.0 Envisioning..... 5
 - 3.1 Solution Definition 6
 - 3.2 Solution Requirements 6
 - 3.2.1 IaaS Services Requirements 7
 - 3.2.2 Service Delivery Processes Requirements 10
 - 3.2.3 Infrastructure Technical Capabilities Requirements..... 11
 - 3.2.4 Service Operations Processes Requirements..... 13
 - 3.2.5 Management and Support Technical Capabilities Requirements..... 15
- 4.0 Conceptual Design..... 19
 - 4.1 Reference Model..... 19
 - 4.2 Architectural Principles..... 21
- 5.0 Physical Design..... 22
 - 5.1 IaaS Services 23
 - 5.2 Management and Support Technical Capabilities 24
 - 5.2.1 Consumer and Provider Portal..... 25
 - 5.2.2 Authentication 27
 - 5.2.3 Usage and Billing..... 29
 - 5.2.4 Management Foundation 30
 - 5.3 Infrastructure Technical Capabilities..... 36
- 6.0 Summary..... 39

1.0 Introduction

Windows Server hosting is one of the largest and fastest growing cloud services opportunities for hosting service providers (HSPs). The Microsoft vision of the cloud platform is to deliver a consistent platform and customer experience regardless of the application or operating system deployment location. Customers can host business applications in their own data center, in Windows Azure, or in an HSP data center to benefit from common technologies and experiences across these environments. In private enterprise and HSP data centers, the cloud infrastructure platform vision is realized through Windows Server 2012 and Microsoft System Center 2012. HSPs are primary participants in the cloud ecosystem and can take advantage of new services specifically designed for them. Together, the Microsoft platform features enable HSPs to offer great experiences to their customers while capitalizing on the growth opportunity in cloud services.

This document provides technical guidance and key considerations for HSPs that want to provide infrastructure as a service (IaaS), both multi-tenant and dedicated hardware hosting, to their customers. In this document, IaaS is defined as:

“The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls).”

This definition of IaaS comes from the National Institute of Standards and Technology (NIST) publication, [The NIST Definition of Cloud Computing](#). The NIST definition also defines the following “essential characteristics” of cloud computing:

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

Throughout this document, the terms “IaaS,” “IaaS services,” and “cloud services” implies services that exhibit these essential characteristics. In the remainder of this document, the term

consumer from the IaaS definition is used synonymously with the word *customer*, since the consumers of services offered by HSPs are their customers. Though the scope of this document is to help you understand how you can provide IaaS to your customers, it is assumed that your customers manage the *arbitrary software such as operating systems and applications* (from the IaaS definition) themselves.

The audience for this document is any individual that designs or implements IaaS services in an HSP organization. The guidance in this document complements several other existing guidance sources that were primarily written for the enterprise IT audience, by augmenting them with further guidance that is either unique, or specifically for, the HSP audience.

The remaining content in this document falls into two categories:

- **Vendor-agnostic:** This includes the [Problem Definition](#), [Envisioning](#), and [Conceptual Design](#) sections of the document. This guidance establishes a foundation to evaluate products from any vendor to provide IaaS services to your customers.
- **Vendor-specific:** The entire [Physical Design](#) section describes design considerations and provides example designs to implement the information that is found in the vendor-agnostic guidance with Microsoft products.

As is the case with designing solutions for any problem, solution design is an iterative process. Although the remaining content in this document is written to be read from top to bottom, it's assumed that you will iterate through the following content sections *many* times before you arrive at a final solution design for your IaaS offering.

2.0 Problem Definition

As customers evaluate HSPs, they find that each HSP provides various services, features, service levels, and prices. As the number of customer HSP options grows, it's critical for HSPs to define the "right" mix of services, features, service levels, and prices to attract and retain their target customers as cost-effectively as possible.

3.0 Envisioning

To envision the solution that you want to provide to your customer you must start by defining a high-level solution definition and scope that is followed by more detailed requirements for that solution.

3.1 *Solution Definition*

The high-level architectural definition for the IaaS offering that you want to provide likely includes a list such as the following:

- A description of the virtual machine, network and storage services that you want to provide to your customers that includes a list of features for each service. The features list will include some features that are provided by other HSPs, and other features that differentiate your service from other HSPs.
- Technical architecture should be informed first by a well-defined, customer centered hosting offer that is informed by customer market opportunity, marketplace pricing and software licensing considerations. As an example, pricing for each of the services should be cost-competitive or reflect additional value that you might provide relative to other HSPs
- An HSP billing portal that enables tracking and billing of customer provisioning and usage.
- A customer portal that enables customers to:
 - Sign up for new service
 - Provision virtual machine, network, and storage resources
 - Check their consumption and billing costs each month and pay their bill
 - Review actual service levels achieved versus the stated service level agreement (SLA) metrics to determine if the service met its SLA

After the high-level solution is defined, you can define more detailed requirements for the solution.

3.2 *Solution Requirements*

Before you create a detailed design for your solution, you must first define a number of requirements for the IaaS services that you want to provide to your customers. Some of these requirements specify the functionality that your services will provide, others specify the service levels that you want your services to meet and the requirements of the technical capabilities and operational processes that you'll need to support the services.

Though not a complete list, you can use the questions in the following sections as a *starting point* to define the requirements for the services that you want to provide to your customers. Even though you probably have existing technical capabilities and operational processes, the questions that follow are provided for you to either evaluate your existing processes and capabilities, or to define new processes and capabilities. The questions are separated into three groupings that include one of the following terms in their names:

- **Services:** Services exhibit *all* of the essential characteristics of cloud computing, on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. As mentioned in the [Introduction](#) section, this document focuses on IaaS services *only*. There is one group of service requirements in the following sections: [IaaS Services Requirements](#).
- **Processes:** Processes refer to operational processes that you will use in your environment to both introduce new services and manage services and the technical capabilities that support them in your environment. There are two groupings of process requirements in the following sections: [Service Delivery Processes Requirements](#) and [Service Operations Processes Requirements](#).
- **Technical capabilities:** Technical capabilities are provided by hardware, software, or both. Individual technical capabilities do *not* exhibit all of the essential characteristics of cloud computing on their own, although they are combined to provide IaaS services that *do* exhibit the essential characteristics. There are two groupings of technical capabilities requirements in the following sections: [Infrastructure Technical Capabilities Requirements](#) and [Management and Support Technical Capabilities Requirements](#).

3.2.1 IaaS Services Requirements

The purpose of this document is to help you provide IaaS services. Remember that in this document, infrastructure services *do* exhibit the essential characteristics of cloud computing, and while infrastructure technical capabilities are necessary to provide infrastructure services, infrastructure technical capabilities themselves *do not* exhibit all of the essential characteristics.

First you need to define which specific virtual machine, network, and storage infrastructure services that you want to provide to your customers. A typical multi-tenant IaaS virtual machine service often includes an operating system and is charged for by the hour with various processor, memory, and operating system disk configurations. After defining the services you will provide, you can use the questions in the table below to help you define the features that your services will provide.

The questions below are grouped by type of infrastructure service, but you'll want to answer similar questions for each individual infrastructure service that you will provide your customers.

Service	Questions you can ask to define the requirements
Virtual machine	<ul style="list-style-type: none"> • Will you offer a fixed number of standardized virtual machine

Service	Questions you can ask to define the requirements
	<p>hardware configurations, will you provide custom virtual machine hardware configurations, or both? When you answer this question, consider how many processors, how much memory, how much storage, and how much bandwidth each virtual machine configuration will support.</p> <ul style="list-style-type: none"> • How will customers release virtual machines they no longer need? Will you make it a one-step process where the virtual machine and associated storage are removed from the pool or will the user have to remove the virtual machine and storage separately? • When customers remove virtual machines, will you charge them for virtual machine service usage separately from storage and networking service usage, or will you charge for virtual machine and storage service usage together? • Will you charge one price for when the virtual machine is online and a different price for when it's offline? • Will you make it possible for your customers to dynamically expand the required amount of memory or processors based on resource utilization of a specific virtual machine, or will the customer have to provision a new virtual machine and attempt to scale out instead of scale up? • Will your customers have the ability to upload existing virtual machines that run on-premises to your service, or will they only be able to provision new virtual machines through your service? Are there any special configuration requirements for existing virtual machines before a customer can move them to your service? • Will you support high availability for your customer's virtual machines such as Windows Server 2012 Failover Clustering, or not? If so, will you automate this configuration and expose it as an option in your customer portal, will you or your customer have to perform some manual configuration, or will you provide it through some combination of the two? • Will you support scale-out load balancing for customer virtual machines and if so, how? • Will you enforce password complexity for your customer's virtual machines and if so, how? • How will host names be assigned to customer virtual machines to ensure no duplication for name resolution? • What virtual disk types will you support? Will you provide a disk conversion service for customers if you don't support the version that they want to move to your service? • What operating systems will you support in customer virtual machines? If the operating systems must be licensed, how will you license them from the vendor and/or accommodate the customers' ability to license them or both? For example, all Windows Server operating system instances need to be properly licensed via the Microsoft Service Provider Licensing Agreement.

Service	Questions you can ask to define the requirements
	<ul style="list-style-type: none"> • Will you support customer virtual machines that run non-uniform memory access (NUMA)-aware applications? Will your fabric management system place virtual machines based on available NUMA nodes? • Will you enable customer virtual machines to be collocated or explicitly not collocated across geographies? • Will you support dedicated physical servers for a customer, where only their virtual machines run, that aren't combined with those of other customers?
Network	<ul style="list-style-type: none"> • Will you enable your customers to physically or virtually connect their own networks to your networks? If so, how will you enable your customers to do so? • Do you plan to encrypt your customers' network data, or will you enable them to encrypt their data if they choose? If so, how will you enable your customers to do so? • Will you enable network level access controls between your network and the customer's on-premises network? If you do, will you have a good understanding of the protocols that are required in both directions? How will you discover what protocols are required? Will you have someone who can provide network analysis?
<p>Storage</p> <p>Note: This storage is <i>separate</i> from "local" storage that you might include with a virtual machine in the virtual machine requirements.</p>	<ul style="list-style-type: none"> • Will you impose storage limits on customer virtual machines and their data drives, operating system drives, or both? If so, what limits will you impose for each and how will you impose them? • Will you enable read/write caching for your storage service and if so, will you enable your customers to turn it on or off? • Will you provide a caching drive for customer virtual machines? If you do not, will you provide any other storage-based optimizations? • Will you provide tiered-storage to customers, so that high average number of I/O operations per second (IOPS) workloads can benefit from more performant storage? • Will you provide commodity storage for customer storage of backups and disaster recovery files? • Will you maintain redundant copies of you or your customers' information? If so, how, and how will you move from primary storage to a storage copy if necessary? • Do you encrypt customer data on disk, or enable the customers to do so? If not, what security measures will you apply to storage to protect it from being stolen outside of facilities-based security measures?

Table 1: IaaS services requirements

3.2.2 Service Delivery Processes Requirements

After you've defined which services and features you'll provide to your customers, we recommend that you then define the service level requirements for the services and the processes that you'll use to ensure that the services adhere to the requirements. Answers to the questions in the following table can serve as your service delivery processes requirements. These questions should be answered for every service that you provide to your customers.

Process	Questions you can ask to define the requirements
Capacity management	<ul style="list-style-type: none"> • How much capacity will you start with? • Will you impose scaling limits for each service? If so, will the limits be hard limits that the customer cannot exceed, or soft limits that the customer can exceed outside of the standard provisioning functionality provided by your services? If the customer wants to exceed these limits, what procedures will you have a plan in place to enable them to do so? • What is the projected capacity requirement over the next year and how do you plan to measure capacity utilization over time to ensure that you continue to meet demand? • How long does it take for you to add new physical resources such as compute, network, storage, and facilities to meet projected capacity requirements? • What operational processes and technical capabilities will you require to meet these requirements?
Availability and continuity management	<ul style="list-style-type: none"> • What level of availability do your customers require for the service? <ul style="list-style-type: none"> ○ Are they willing to pay different prices for different levels? ○ Will the availability level(s) that you offer also incorporate downtime for a disaster, for example a whole data center failure, or not? ○ How will you measure availability for each service, and what will your definition of unavailable be? Your definition of unavailable is what you'll monitor and report back to your customers, and what will define whether or not you met your SLA. • What operational processes and technical capabilities will you require to meet these requirements?
Information security management	<ul style="list-style-type: none"> • What levels of security isolation between tenant resources will you provide? • Will your infrastructure support encryption of customer data both when it is stored and in transit? If so, will this encryption be part of your standard service offering or will it be an option that you charge a premium price for? • What operational processes and technical capabilities will you require to meet these requirements?
Regulatory policy and compliance	<ul style="list-style-type: none"> • Will your services need comply with any regulatory or compliance policies and/or support your customers' needs to comply with them? If so, which policies, and how will you comply with them?

Process	Questions you can ask to define the requirements
management	<ul style="list-style-type: none"> • If you will host customer data that is subject to regulatory or compliance policies, how will you provide access to the infrastructure so that auditors can evaluate your infrastructure for certification purposes? • What operational processes and technical capabilities will you require to meet these requirements?
Financial management	<ul style="list-style-type: none"> • What price will you charge for each of your services and in what units will it be charged for? For example, the unit of charge for a virtual machine service might be minutes or hours that the virtual machine was used, the unit of charge for a network service might be megabytes of data transferred in and out of your data center, and the unit of charge for a storage service might be the number of gigabytes of data stored. • Will your prices be competitive with other HSPs? If they are not, what value differentiators will you offer to compensate for the price difference? • How will you provide usage and charge data to your customers? • What operational processes and technical capabilities will you require to meet these requirements?
Service level management	<ul style="list-style-type: none"> • How will you manage services for compliance against the SLAs that you provide your services with? • How will you measure customer satisfaction with your service over time and make adjustments to improve customer satisfaction? • How will you decrease your cost required to provide a service over time? • What operational processes and technical capabilities will you require to meet these requirements?
Service lifecycle management	<ul style="list-style-type: none"> • How will you decide when and what to change about the features offered by your services? • How will you decide to retire an existing service or begin offering a new one? • How will new services be created and tested to ensure they comply with all functional service level requirements before they're made available to customers? • What operational processes and technical capabilities will you require to meet these requirements?

Table 2: Service delivery processes requirements

3.2.3 Infrastructure Technical Capabilities Requirements

Infrastructure technical capabilities are the core foundation for providing IaaS services to customers, but as already mentioned, do *not* exhibit all of the essential characteristics of cloud computing on their own. They must be combined with management and support technical capabilities to provide IaaS. The answers to the questions in the following table can serve as your infrastructure technical capabilities requirements. Requirements for management and support technical capabilities are discussed in the next section.

Technical capability	Questions you can ask to define the requirements
<p>Common requirements</p> <p>These requirements must be defined for the remaining capabilities in this table.</p>	<ul style="list-style-type: none"> • From which suppliers will you purchase the hardware and software to provide the technical capability? What are the suppliers' ordering processes and lead times? Will you negotiate volume discounts with them? Will you establish fulfillment SLAs with them? How will you ensure that your capacity requirements and supplier fulfillment SLAs are aligned? • What standards will you define for the hardware and software that you will use to provide the capability? How much will you be able to standardize the hardware or software or both to provide the capability? • What process will you put in place to update hardware and software to new versions when they become available? Will you define a process to acquire non-standard hardware and software? • How will you implement each capability to support your regulatory compliance, access management, and information security management processes?
<p>Compute</p>	<ul style="list-style-type: none"> • What form-factor servers will you use, for example blade chassis, rack units, and others? • Will you purchase individual servers or purchase them pre-configured in racks or in blade chassis, or will you purchase both? • What processor, memory, disk, storage interface, and network interface will the servers you purchase have?
<p>Network</p>	<ul style="list-style-type: none"> • What on-premises networking equipment will you require to connect to your customers' networks or the Internet or both? • What carriers will you use and what are their SLAs, bandwidth, and latency characteristics and costs? • What protocols or other capabilities or both will you require your customers to have to connect to your service? • Will you support high availability network mechanisms such as the Border Gateway Protocol (BGP) on your customer's end of a physical or virtual network connection? • Will you use port access control lists (ACLs) or a service level firewall to enable access controls between tenants and between tenants and their on-premises infrastructure? • Will you provide Internet Protocol security (IPsec) support for tenants who want to use it? If so, will you enable hardware offloads to support it? • Will you provide quality of service (QoS) for tenant network traffic? If so, how will you provide it?
<p>Storage</p>	<ul style="list-style-type: none"> • Will you enable customers to integrate their storage systems with yours for backup, snapshots, and replication? • Will your networked-storage require a Fibre Channel or Ethernet network? If Ethernet, will it utilize iSCSI, FCoE, or SMB protocols? The answers to these questions
<p>Virtualization</p>	<ul style="list-style-type: none"> • What will your strategies be for planned and unplanned downtime of your hypervisor servers? • Will you your hypervisor support oversubscription of processor and

Technical capability	Questions you can ask to define the requirements
	<p>memory resources? If so, how will it support oversubscription and how will you use it?</p> <ul style="list-style-type: none"> • What types of storage interfaces will you make available to virtual machines? Examples are IDE, SCSI, and Fibre Channel. Will the virtual machine you provide to your customers have access to physical disks, virtual disks, or both? • Will you use the virtual switch capabilities provided by the hypervisor, or by extending its capability with third-party functionality?

Table 3: Infrastructure technical capabilities requirements

3.2.4 Service Operations Processes Requirements

To deliver reliable IaaS services to customers on an ongoing basis, you must have clearly-defined operational processes that you consistently follow. A variety of information technology service management (ITSM) frameworks exist to help you define operational processes. You will likely define the standard process, and then for each individual service, define how it will be applied. Answers to the questions in the following table can serve as your service operations processes requirements.

Process	Questions you can ask to define the requirements
Request fulfillment	<ul style="list-style-type: none"> • Will you have a clearly-defined request fulfillment process and how will you ensure that it's consistently followed? • How will you fulfill customers' requests for service? Will you fully automate processes to create or remove capacity at customer request or will you fulfill requests with a combination of automated and manual processes? • What service levels metrics will you provide for fulfilling requests?
Asset and configuration management	<ul style="list-style-type: none"> • Will you have clearly-defined asset and configuration management processes and how will you ensure that both are consistently followed? • How will you manage physical assets from acquisition to retirement in your environment? • Which configuration items (CIs) of the compute, network, and storage technical capabilities in your environment will you enable your customers to configure? You will likely not enable customers to configure all CIs of the technical capabilities, but instead, will enable them to configure only the CIs necessary to support the features of the services you provide.
Change management	<ul style="list-style-type: none"> • Will you have a clearly-defined change management process and how will you ensure that it's consistently followed? • Will you have a change advisory board that approves or denies critical or major changes? If so, who are the members, when does

Process	Questions you can ask to define the requirements
Release and deployment management	<p data-bbox="641 243 1356 275">the board meet, and what are the board's evaluation criteria?</p> <ul data-bbox="597 289 1396 583" style="list-style-type: none"> <li data-bbox="597 289 1356 380">• Will you have clearly-defined release and deployment management processes and how will you ensure that they're both consistently followed? <li data-bbox="597 394 1396 485">• How will you define a release? How will you test and deploy it? How will you rollback the release in the event there is a problem with the deployment of the release? <li data-bbox="597 499 1356 583">• What deployment process will you use for the underlying technical capabilities of the services that you provide to your customers?
Access management	<ul data-bbox="597 604 1404 1178" style="list-style-type: none"> <li data-bbox="597 604 1404 663">• Will you have a clearly-defined access management process and how will you ensure that it's consistently followed? <li data-bbox="597 678 1404 737">• What access policies and restrictions will you define for your IaaS services to protect both you and your customers? <li data-bbox="597 751 1356 810">• Who will have what access to the resources within your own organization? <li data-bbox="597 825 1404 915">• What levels of access will your customers have to your resources? How will you assign this access to them and manage it? <li data-bbox="597 930 1372 1041">• Will you maintain records of who in your own organization accessed the technical capabilities that you use to provide the services to your customers? Do you and can you provide this information to your customers if they request it? <li data-bbox="597 1056 1404 1178">• Will you maintain records of who in your customer's organizations accessed the services that you provide to your customers? Do you and can you provide this information to your customers if they request it?
Systems administration	<ul data-bbox="597 1199 1404 1423" style="list-style-type: none"> <li data-bbox="597 1199 1404 1289">• Who in your organization will be responsible for administering each of the technical capabilities that you will use to provide IaaS services to your customers? <li data-bbox="597 1304 1404 1362">• What administration tasks must be completed for your technical capabilities on a daily, weekly, or as-needed basis? <li data-bbox="597 1377 1404 1423">• Which administration tasks will you perform manually, which will you automate, and how will you automate them?
Knowledge management	<ul data-bbox="597 1444 1404 1780" style="list-style-type: none"> <li data-bbox="597 1444 1404 1503">• Will you have a clearly-defined knowledge management process and how will you ensure that it's consistently followed? <li data-bbox="597 1518 1372 1640">• How will you manage knowledge that pertains to the services that you provide to your customers, both for your own employees to correctly manage the service, but also for your customers to manage their use of your service? <li data-bbox="597 1654 1388 1713">• How will you make this knowledge available to your employees and customers? <li data-bbox="597 1728 1372 1780">• Will you have a knowledge management system that supports your process, and if so, what is it, and how will it be used?
Incident and problem management	<ul data-bbox="597 1801 1404 1879" style="list-style-type: none"> <li data-bbox="597 1801 1404 1879">• Will you have clearly-defined incident and problem management processes and how will you ensure that they're both consistently followed?

Process	Questions you can ask to define the requirements
	<ul style="list-style-type: none"> • How will incidents be handled for the services that you provide? How will your customers file incident requests, and how will you resolve those requests? • What processes or tools or both will you use to manage incidents and their resolutions? • What resolution targets will you have for incidents? Will you provide different resolution targets for different levels of service or price? For example, you might define gold or bronze levels of service where each level has different resolution targets. What will you charge for each level? • Will you notify customers in the event of a major incident to your service, and if so, how? • How will you determine the root cause of incidents in your environment? • Will you proactively try to identify problems with services before customers experience them and how will you identify them?

Table 4: Service operations processes requirements

3.2.5 Management and Support Technical Capabilities Requirements

In addition to infrastructure technical capabilities, you must have several technical capabilities that support your service operations processes, but that also manage and support the infrastructure technical capabilities used to provide your IaaS services. Answers to the questions in the following table can serve as your management and support technical capabilities requirements.

Technical capability	Questions you can ask to define the requirements
Consumer and provider portal	<ul style="list-style-type: none"> • Will you provide a web-based, or some other type of graphical user interface customer portal, an application programming interface (API), or both to your customers so that they can provision and de-provision accounts and resources on your services? • Will your customer portal or API allow your customers to view SLA reporting and their service usage data? • How will your customer and provider portals integrate with all of your other management and support technical capabilities? • How will your employees manage the services that you offer? Will they do so via a provider portal or some other interface or both?
Usage and billing	<ul style="list-style-type: none"> • How will you bill your customers? Will you bill them electronically, or through paper invoices? • For what service utilization units, such as VM hours, gigabytes of data stored, and gigabytes of network bandwidth used, will you collect usage data to bill your customers? How often will you collect usage data, and how long will you maintain it? • What forms of payment will you accept from your customers? • How will you resolve billing issues with your customers? • What will you do if and when a customer is delinquent on their payment to you? • How will you collect usage data for billing? Will the data be collected through manual processes, automated processes, or some combination of the two?
Service reporting	<ul style="list-style-type: none"> • How will you provide service reporting to your customers? At a minimum, you should be able to report on actual and target service levels for each of the services that you provide to your customers. • What information will you provide in your customer reports? • How long will you maintain reporting data for customers? • How will this capability integrate with your service monitoring capability?
Service monitoring	<ul style="list-style-type: none"> • What service level metrics will you monitor for each of the technical capabilities that you use to provide services to your customers and how will you monitor these metrics? At a minimum, you must monitor enough metrics to provide service level reporting data to your customers that supports your SLAs. • How will the monitoring data be provided to your service reporting capability? • How will this capability integrate with your usage and billing capability?
Service management	<ul style="list-style-type: none"> • What type of service management system will you use to manage the IaaS services that you provide to your customers? • Which of your operational processes will your service management capability support and how? • How will your service management capability integrate with your

Technical capability	Questions you can ask to define the requirements
	<p>service monitoring, configuration management, and change management technical capabilities?</p> <ul style="list-style-type: none"> • How will your service management capability maintain information that supports your change management, knowledge management, asset management, and incident and problem management processes?
Configuration management	<ul style="list-style-type: none"> • Will you have a configuration management technical capability? If so, will it support your asset and configuration management processes? Will it include a configuration management database (CMDB)? • Will your configuration management technical capability enable you to maintain current, desired and proposed values for your CIs? • Will your configuration management technical capability allow you to remediate current states to desired states when they don't match?
Authorization	<ul style="list-style-type: none"> • Who in your organization will be authorized to access the resources that are used to provide your services? What will they be authorized to do with those resources? • Who in your customers' organizations will be authorized to access the resources that you provide, and what are they authorized to do with them? • How will your authorization capability integrate with your authentication and directory capabilities? • Will you enable your customer to federate authorization mechanisms that they might use in their own organization with yours to enable them to automatically revoke authorization to services on your systems that an employee who might leave their company had access to? This scenario is in contrast to them having to replicate authorization policies across both their own systems and yours.
Authentication	<ul style="list-style-type: none"> • How will your employees authenticate to your technical capabilities and services? • What authentication mechanisms will you provide for your customers to authenticate to the services you provide? • How will your authentication capability integrate with your authorization and directory capabilities? • Will you enable customers to federate authentication mechanisms that they might use within their own organization with yours to support customer single sign-on strategies? If so, how will you support such strategies?
Directory	<ul style="list-style-type: none"> • What type of directory service(s) will be used to maintain customer information? What information will be maintained? • How will your directory capability integrate with your authentication and authorization capabilities?
Data protection	<ul style="list-style-type: none"> • How will you protect the data that is required to provide your services, for example, billing, reporting, CI, incident, and so on? • How will you protect your customers' data? Will you keep copies

Technical capability	Questions you can ask to define the requirements
	<p>of their data for data protection, will you provide backup and recovery services that enable the customer to restore a point-in-time backup, or both?</p>
Deployment and provisioning	<ul style="list-style-type: none"> • How will new customer virtual machines, virtual network, and virtual storage be provisioned or de-provisioned on your fabric? • Will you enable your deployment and provisioning mechanism to integrate with any systems that your customers might have in their own environments? If so, how?
Process automation	<ul style="list-style-type: none"> • How will automation of tasks that are invoked by users through the customer and provider portal capability be executed and managed? • Which of the tasks executed by other management and support technical capabilities will be automated, and how will the process automation capability manage the automation of those tasks? • Will you support customers who might want to integrate their automation capabilities with some or all of your automation capabilities? If so, how will you support your customers' efforts to do so?
Fabric management	<ul style="list-style-type: none"> • All of the infrastructure technical capabilities used to provide IaaS services are often collectively referred to as a "fabric," when they're managed like a single resource. Will you have a fabric management technical capability that allows you to collectively provision and manage all of the infrastructure technical capabilities that support your IaaS services? • How will your fabric management technical capability: <ul style="list-style-type: none"> ○ Provision fabric resources when a customer requests new capacity from your services? ○ Monitor resource utilization and ensure that your infrastructure capabilities are efficiently utilized? For example, it might automatically consolidate customer virtual machines on hypervisor servers so that unused servers could be shut down to save power and costs. ○ Update hypervisor servers? Will the fabric management capability relocate virtual machines automatically as hypervisor servers are updated? Will that process incur any downtime of customer virtual machines?
Network support	<ul style="list-style-type: none"> • How are IP addresses assigned and managed for both your infrastructure technical capabilities, and for the resources that you enable your customers to provision? • Will your customers be able to utilize their existing IP addresses, will they need to use your IP addresses, or both? • How will IP hostnames be resolved at the fabric level, and how will they be resolved for your customers' virtual machines?

Table 5: Management & support technical capabilities requirements

4.0 Conceptual Design

After you determine the requirements for providing your IaaS services, you can begin to design your solution. Before you create a physical design, it helps to first define a reference model and some architectural principles, which work together to serve as a foundation, and guide your physical design.

4.1 *Reference Model*

Reference models are created to help people clearly and consistently understand a problem domain. Reference models often decompose problem domains into various entities, and relationships between entities. Reference models are typically vendor and product-agnostic. Among other purposes, reference models serve as a common taxonomy to evaluate different vendors' product capabilities. The following figure shows a reference model that was created with "Cloud Services Foundation" as its problem domain.

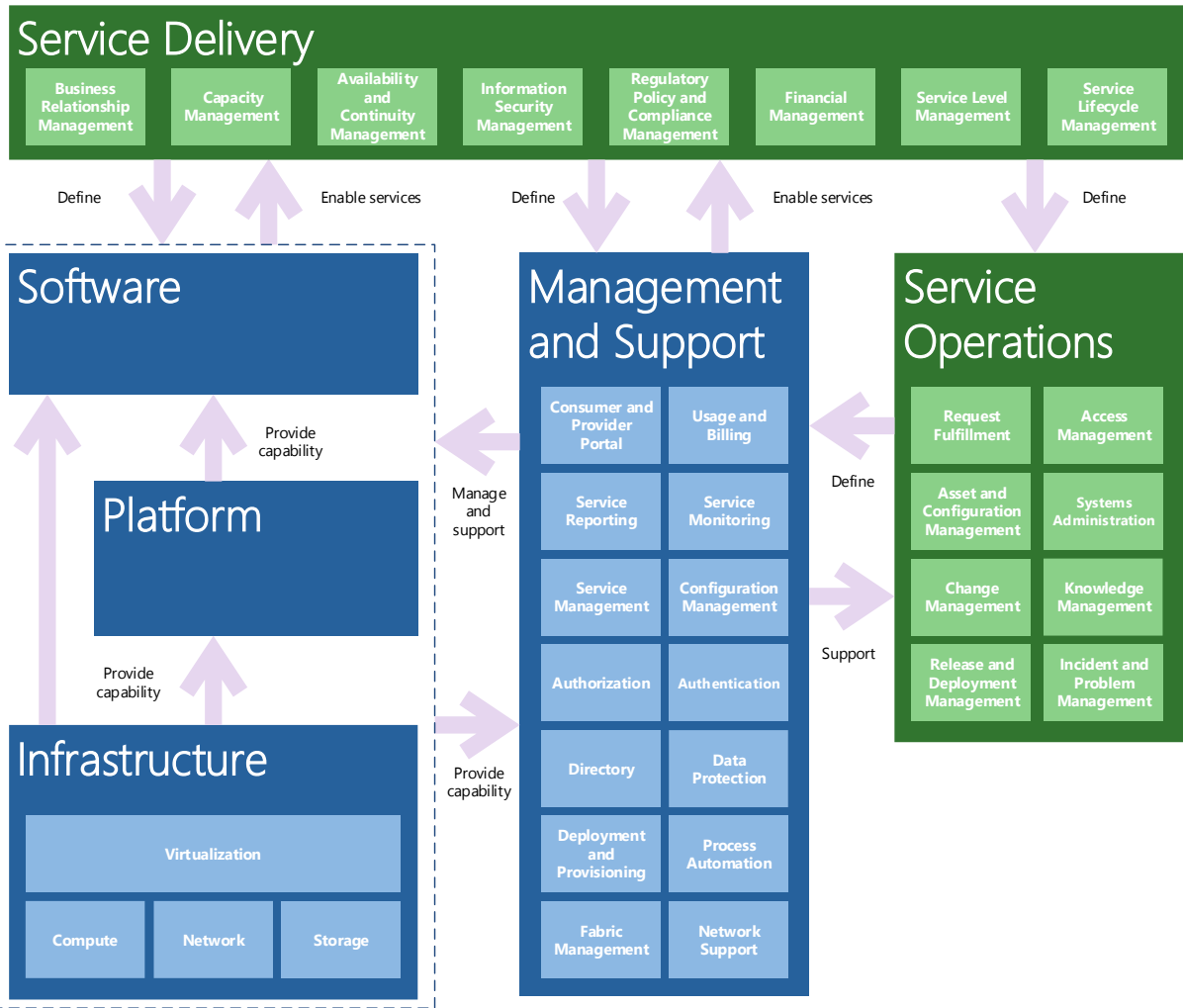


Figure 1: Cloud Services Foundation Reference Model

Figure 1 includes vendor-agnostic terms that can serve as a common taxonomy for creating a foundation on which to provide cloud services. Reference models like the one above also provide details for each entity within them, the relationships between those entities, and what each color represents. That level of information is *not* provided in this document, because it is provided separately in the Microsoft [Cloud Services Foundation Reference Model](#) (CSFRM) document, which is available as part of the Microsoft [Cloud Services Foundation Reference Architecture](#) (CSFRA) guidance set.

You are encouraged to read the CSFRM document for a greater understanding of the CSFRM but doing so is not required to understand the remaining content in this document. Most of the sections within the [Solution Requirements](#) and [Physical Design](#) sections of this document are aligned to the CSFRM. The alignment of these sections with the CSFRM illustrates another key use for a reference model, to serve as a framework for designing solutions within the

same, or related problem domains for which a reference model was created. Since the CSFRM is a foundation for providing cloud services, and IaaS services are cloud services, the CSFRM provides a framework for designing IaaS services.

4.2 *Architectural Principles*

While reference models are helpful to decompose a problem domain, they are even more useful when they are coupled with architectural principles. Architectural principles are enduring, rarely amended, and inform and support the way solutions in a problem domain fulfill their missions and goals. They strive to be compelling and aspiring, and might not ever be fully realized, but serve as a goal to move towards.

The principles listed in the [Architectural Principles, Concepts, and Patterns](#) document in the Microsoft [Cloud Services Foundation Reference Architecture](#) guidance set details a number of vendor-agnostic principles that are identified as commonly used across multiple HSPs. A subset of the principles described in that document are repeated in the following list, but you're encouraged to read the document in its entirety for a more detailed explanation of each principle, as well as to understand the concepts and patterns guidance that the document also contains, even though they're not discussed in this document. Achieving any of the principles listed below requires you to have all of the technical capabilities and processes defined in the CSFRM. You might choose to adopt none or all of these principles, but regardless of which you choose to adopt, we recommend that you define your own list of principles that will guide the design of the IaaS services that you will provide to your customers.

- **Perception of infinite capacity:** Though no HSP has infinite capacity, one of the reasons that customers select an HSP is because they expect it to meet their capacity needs, regardless of what those needs are. Well-implemented technical capabilities and processes from the CSFRM will enable you to provide services that your customers perceive as having infinite capacity.
- **Perception of continuous service availability:** Few, if any HSPs provide a 100 percent availability guarantee to their customers, but customers expect the service to be available continuously. Well-implemented technical capabilities and processes from the CSFRM will enable you to provide services that your customers perceive as continuously available.
- **Optimization of resource usage:** The more you can optimize your resource utilization, the more inexpensively you can provide the service to your customers, by your own organization, or both.

- **Increase customer satisfaction through measured continuous improvement:** It's critical to adhere to the service levels that you've committed to meet for your customers, but how well you do that can always improve. In addition, improvement efforts often enable you to increase customer satisfaction rates and lower your cost to deliver services over time.
- **Take a holistic approach to availability design:** When you provide an at-scale service, broad resiliency strategies can be employed that might be more inexpensive than more narrow design strategies that are based on server-by-server availability. Define what types of customer workloads your IaaS offering will support to select the best, most cost-effective hardware implementation to support those customer workloads. This evaluation and definition is critical for business success as it determines operational efficiency relative to contributions margins for a given IaaS service.
- **Minimize human involvement:** Providing IaaS services typically requires little or no human involvement between the provider and the customer, so that customers can self-provision, manage, and deprecate their own virtual machines. In addition to supporting customer self-service, automation allows for tasks to be completed faster, and with less errors than tasks completed by humans.
- **Create a seamless user experience:** As more and more organizations start to use computing capacity that HSPs provide, one of their HSP evaluation criteria is how seamlessly they can incorporate the resources of their HSPs into their own on-premises environments. Which HSPs enable them to federate their authentication mechanisms with their own existing mechanisms for single sign-on? Which HSPs enable them to use the same management tools that they use to manage their own infrastructure to seamlessly manage their HSPs infrastructure? How seamless it is to integrate your resources into your customer's environments can be a competitive differentiator between HSPs.

5.0 Physical Design

With defined requirements, like those listed in the [Solution Requirements](#) section of this document, and a reference model and principles, like those defined in the [Conceptual Design](#) section of this document, you can then select the products and technologies that best meet your requirements and align to your conceptual design. The following table lists the hardware vendor-agnostic products and Microsoft products and technologies that can be used to implement the infrastructure and management and support technical capabilities that are shown in *Figure 1: Cloud Services Foundation Reference Model*.

CSFRM entity	Product/technology
--------------	--------------------

Consumer and Provider Portal	Windows Azure Services for Windows Server or your own custom portal that uses the Service Management API
Usage and Billing	System Center 2012 - Operations Manager, System Center 2012 - Virtual Machine Manager, Service Management API, and independent software vendor (ISV) control panel solutions from Citrix, Ensim, ExtendASP, IntY, Cloud Cruiser, Parallels and other ISVs that use the Service Management API.
Service Reporting	System Center 2012 - Operations Manager
Service Monitoring	System Center 2012 - Operations Manager, System Center 2012 - Virtual Machine Manager, and Windows Server 2012 Failover Clustering
Service Management	System Center 2012 - Service Manager
Authorization	The Service Management API and Windows Server Active Directory
Authentication	Windows Server 2012 Active Directory, Active Directory Federation Services (AD FS) and the Service Management API
Directory	Windows Server 2012 Active Directory service and the Service Management API
Data Protection	System Center 2012 - Data Protection Manager (DPM)
Deployment and Provisioning	System Center 2012 - Virtual Machine Manager and the Service Management API
Process Automation	System Center 2012 - Orchestrator
Fabric Management	System Center 2012 - Virtual Machine Manager
Network Support	Windows Server Domain Name System (DNS) and Windows Server 2008 Dynamic Host Configuration Protocol (DHCP) Server service
Network	Your choice of network hardware
Compute	Your choice of server hardware
Storage	Your choice of storage hardware, which could include Windows Server 2012 File Server servers
Virtualization	Windows Server 2012 Hyper-V

Table 6: Products/technologies for CSFRM entities

After selecting the products and technologies to implement the IaaS infrastructure, you can continue with your physical design. The remainder of this document assumes that you're using the technologies listed in *Table 6* to implement the corresponding reference model entities that are also listed in that table.

5.1 IaaS Services

This document assumes that you'll provide IaaS services such as virtual machines, network, and storage to your customers. Although it is beyond the scope of this document, you might also choose to provide services other than infrastructure services to your customers such as:

- **Platform services:** Platform service functionality is often consumed by software services. Examples of platform services are:
 - Database hosting with SQL Server 2012
- **Software Services:** Software services are often consumed by end-users. Examples of software services are:
 - Email hosting via Exchange 2013
 - Content and collaboration hosting with SharePoint 2013
 - Unified communications hosting including instant messaging, conferencing and voice with Lync 2013
 - Website and E-commerce hosting with Windows Server 2012 and Windows Azure Services for Windows Server Web Sites Cloud
 - Customer relationship management, enterprise resource planning, or other line-of-business application hosting

You can learn more about providing these services at www.microsoft.com/hosting. Though the remaining information in this document helps enable you to provide all types of cloud services to your customers, its focus continues to be on enabling you to provide infrastructure services.

The remaining sections are aligned to entities from *Figure 1: Cloud Services Foundation Reference Model*. The design for each section should align to your reference model, your architectural principles, and the solution requirements that you defined for the relevant entities within your reference model. The guidance provided in each section includes design considerations and options for reference model entities, and also design guidance for many of them.

5.2 *Management and Support Technical Capabilities*

The design for your management and support technical capabilities from *Figure 1: Cloud Services Foundation Reference Model* can be tightly coupled with the design for your infrastructure technical capabilities, depending on which products you use to provide these capabilities in your environment. Because in this document it's assumed that you're using Microsoft System Center 2012 to provide the management and support technical capabilities in your environment, the design of these capabilities can be largely independent from the design of your infrastructure technical capabilities, because it supports multiple hypervisor and hardware technologies.

5.2.1 Consumer and Provider Portal

The customer portal is the primary interface between your customers and the IaaS services that you offer them. Through this portal, your customers create an account, provision or de-provision services, interact with the services, monitor health of their services, and pay for their services.

You may use an existing custom-developed customer portal, or you may choose to use the web-based tenant portal that Microsoft provides as part of [Windows Azure Services for Windows Server](#), shown in *Figure 2* below.

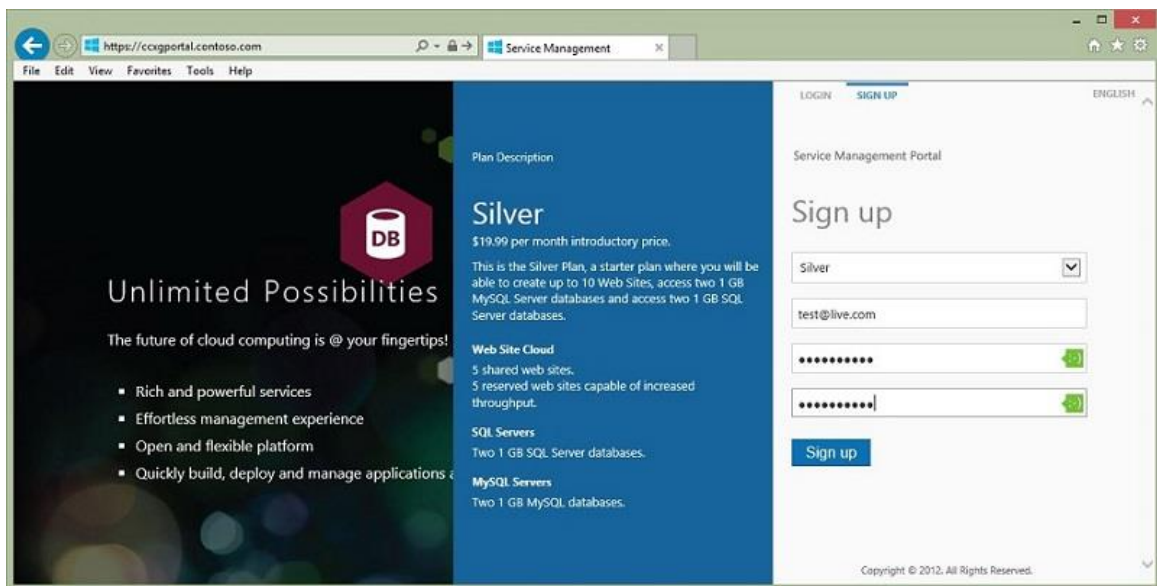


Figure 2: Tenant portal

Figure 2 shows the first step of a multi-step wizard that your customers would go through when they sign up and subscribe to plans that you configured through the admin portal shown in *Figure 3*. All pages of the wizard are shown in the [End-User Experience](#) article.

Note: The next version of Windows Azure Services for Windows Server is called the [Windows Azure Pack for Windows Server](#). You can learn more about the Windows Azure Pack by reading the [Windows Azure Pack for Windows Server Overview](#) white paper or viewing any of the [Windows Azure Pack presentations](#) from Microsoft Tech-Ed 2013.

Figure 3 shows the first step of a multi-step wizard for creating a *plan* in the admin portal. Plans include one or more provisioned services and are used to distinguish collections of service offerings that you want to make available for your customers to subscribe to. All

steps in the wizard are illustrated in the Create Plans for End Users section of the [Provision Services](#) article.

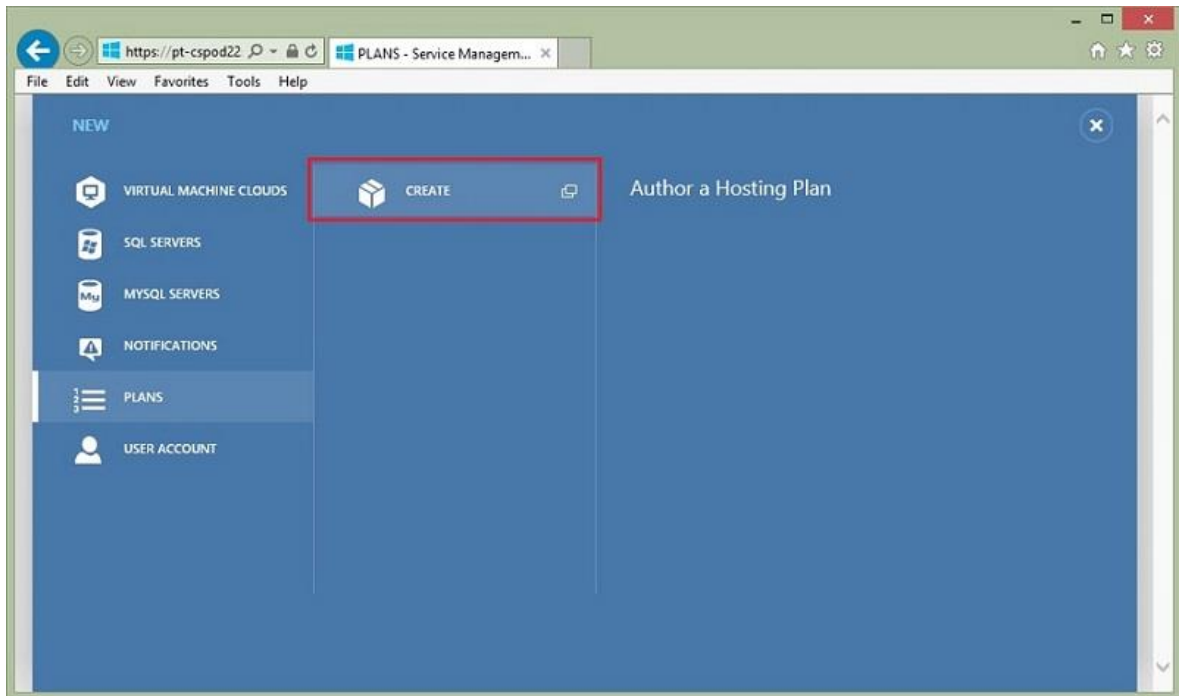


Figure 3: Admin portal

Both portals utilize the Service Management API, which is also available as part of Windows Azure Services for Windows Server. The Service Management API communicates with System Center 2012 Virtual Machine Manager through the System Center 2012 [Service Provider Foundation](#) (SPF), which is discussed further in the [Management Foundation](#) section of this document. The [Web Sites, Virtual Machines, Service Management Portal, and Service Management API Samples - Feb 2013](#) contains information for customizing and extending the portal experience by using the REST-based Service Management API to include branding, billing, custom service integration and integration into existing portals. Windows Azure Services for Windows Server is installed through the Microsoft [Web Platform Installer](#), as detailed in the [Installation Instructions](#) for Windows Azure Services for Windows Server. Prior to installing, we recommend that you read the [Installation Guide](#) and [Capacity Planning](#) articles.

For an overview of the Service Management Portal, you can view the short [Azure Services Video Demo](#). For a deeper understanding, you can view the video of the [Windows Server Virtual Machine: Adding Windows Azure Services](#) presentation from the Microsoft Management Summit 2013.

When you've defined your customer portal requirements, and decided what customer portal experience to offer your customers, consider this: In addition to hosting some of their virtual machines on your IaaS offering, some of your customers might currently provide the tenant portal included in Windows Azure Services for Windows Server to their internal consumers for provisioning services on their on-premises private cloud. Remember the "Create a seamless user experience" principle in the [Architectural Principles](#) section of this document. If you provide the tenant portal from Windows Azure Services for Windows Server to your customers, they would find your customer portal experience very similar to their own, providing your customers a seamless path to your customer portal and services.

A rich ecosystem of independent software vendors (ISVs) provide end-to-end control panel and provisioning systems that leverage Windows Server 2012 and System Center 2012. When making build vs. buy decisions, consider reviewing the numerous ISV control panel solutions available from Citrix, Ensim, ExtendASP, IntY, Cloud Cruiser, Parallels and others.

In addition to the appearance of your customer portal, you also must determine the functionality that it will provide to your customers, such as the other Management and Support Technical Capabilities from the CSFRM that are described in the following sections.

5.2.2 *Authentication*

The table below includes some authentication options that you may consider providing to your customers for authentication to your customer portal.

Option	Advantages	Disadvantages
<p>Customers authenticate to your proprietary authentication mechanism, separately from any authentication mechanisms that they might already use for on-premises systems.</p>	<p>You and your customers don't have to manage any federation or account synchronization.</p>	<ul style="list-style-type: none"> • Your customers have to log on to their own on-premises systems, and log on separately to your systems. • Your customers must manage accounts in both their own system and your system. For example, if an employee joined or left the organization, the customer would have to update their own system and your system to reflect the change.
<p>Customers can federate their on-premises authentication mechanism such as Windows Server Active Directory Domain Services with your authentication mechanisms directly by using technologies such as Windows Server Active Directory Federation Services.</p>	<ul style="list-style-type: none"> • Your customers can log on to their on-premises systems and seamlessly integrate with your customer portal without having to log on separately to your system. • Your customers only have to manage accounts in their own system and those changes are automatically reflected in your system. For example, if an employee joined or left the organization, the customer wouldn't have to update their own system and your system. 	<ul style="list-style-type: none"> • Both you and your customer must manage a federation with each other. You must manage separate federations with each of your customers, and each of your customers must manage separate federations with you, and each of the other service providers with which they host services. • You have to write the custom code to enable this capability in your custom-developed portal, because the portal provided with Windows Azure Services for Windows Server does <i>not</i> include this functionality.

Option	Advantages	Disadvantages
<p>Customers can federate their on-premises authentication mechanism with yours through a federation service such as Windows Azure Active Directory.</p>	<ul style="list-style-type: none"> • Your customers can log on to their on-premises systems and seamlessly integrate with your customer portal without having to log on separately to your system. • Your customers only have to manage accounts in their own system and those changes are automatically reflected in your system. • Your customer might already be federated with Windows Azure Active Directory, particularly if they use Microsoft Office 365 or Windows Azure today, thus eliminating their requirement to set up a direct federation with you or any other HSPs they might use that are also federated with Windows Azure Active Directory. • Doesn't require you to set up direct federations with individual customers that are federated with Windows Azure Active Directory. 	<ul style="list-style-type: none"> • Both you and your customer must manage a federation with Windows Azure Active Directory. • You have to write the custom code to enable this capability in your custom-developed portal, because the portal provided with Windows Azure Services for Windows Server does <i>not</i> include this functionality. If you decide to implement this functionality we recommend that you read the Developing Multi-Tenant Web Applications with Windows Azure AD article.

Table 7: Customer portal authentication options

5.2.3 Usage and Billing

The Service Management Portal collects usage metering data via REST-based JSON queries on a scheduled basis through SPF. This data comprises all the usage metering metrics for all the virtual machines that a tenant uses, provided that those virtual machines are monitored by System Center 2012 - Operations Manager and that the data is stored in Operations Manager data warehouses. The usage data collector architecture is shown in the figure below.

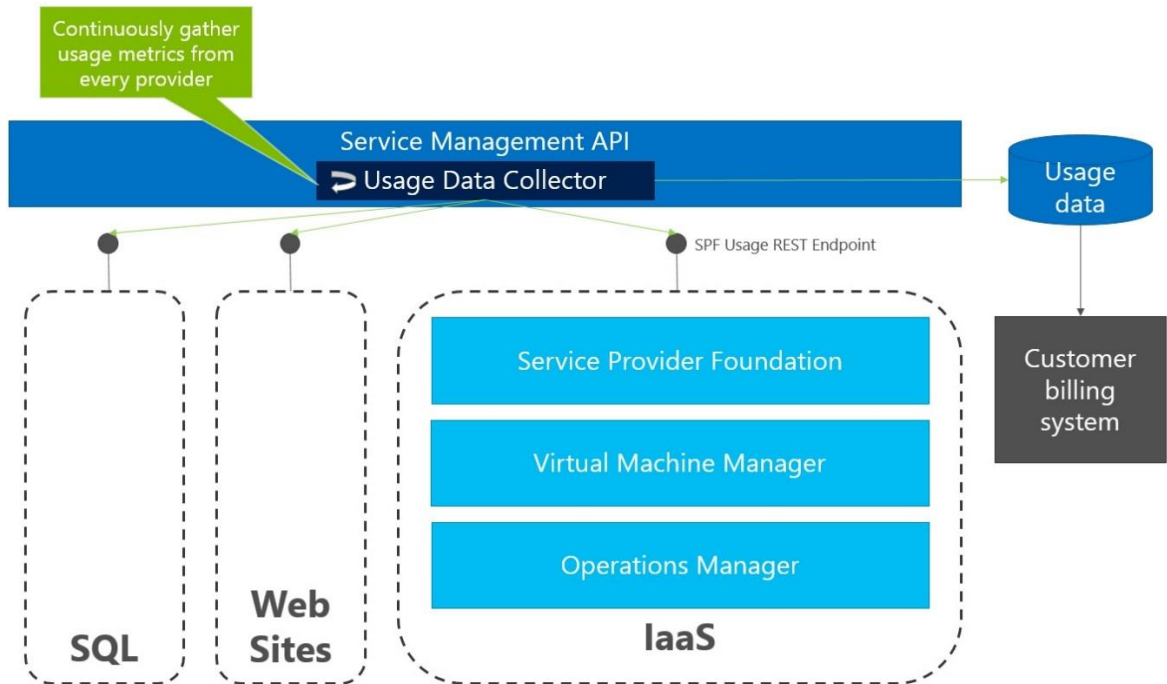


Figure 4: Usage data collector architecture

The role of SPF is to collect metrics from all the data warehouses and aggregate them for billing and analysis. The usage data that is collected is described in the [Usage Metering Data Model in Service Provider Foundation](#) article. The [Web Sites, Virtual Machines, Service Management Portal, and Service Management API Samples - Feb 2013](#) includes a sample usage adapter that you can use to consume tenant usage data and connect to your own custom billing provider, or to billing providers that are available from ISVs.

5.2.4 Management Foundation

In this document, the term *management foundation* is used to represent all of the technical capabilities in the box labeled *Management and support* in [Figure 1: Cloud Services Foundation Reference Model](#) except for the [Consumer and Provider Portal](#), customer portal [Authentication](#), and [Usage and Billing](#) technical capabilities, which were previously discussed. The management foundation capabilities are collectively provided by [Microsoft System Center 2012 SP1](#). The [Service Provider Foundation](#) (SPF), mentioned previously in this document, is a feature of the System Center 2012 - Orchestrator component. It exposes an extensible OData web service that interacts with the System Center 2012 - Virtual Machine Manager (VMM) component to enable a variety of hosting provider and customer scenarios. The SPF is shown in the following figure.

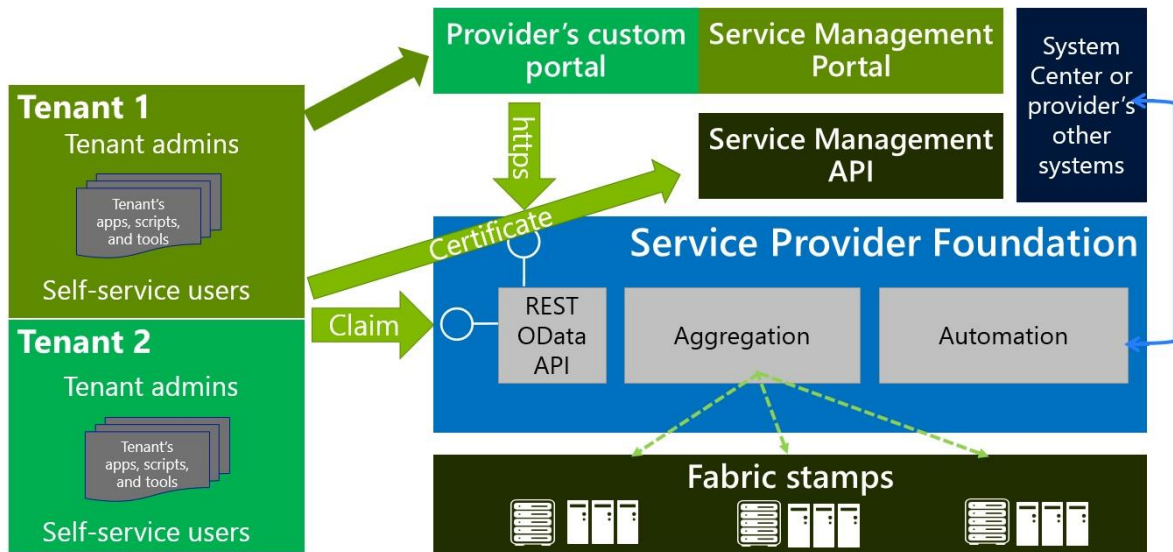


Figure 5: Service provider foundation scenarios

Figure 5 illustrates the Service Management Portal discussed in the [Consumer and Provider Portal](#) section of this document, but also shows the “Provider’s custom portal,” which can also communicate with the SPF or the Service Management API. Whether you choose to use the Service Management Portal or your own custom portal, SPF enables multi-tenancy for System Center 2012 SP1.

Authentication and Authorization

Table 7: Customer portal authentication options detailed authentication options that you might provide to your customers through your customer portal. Regardless of what authentication options you provide to your customers, System Center 2012 still uses Windows Server Active Directory accounts for authentication and authorization before executing its operations. Because your customers likely do not have accounts in the Windows Server Active Directory domain that your System Center 2012 servers are members of, SPF doesn’t use Windows Server 2012 Active Directory for authentication of your customers. Instead, it provides a claims-based authentication security model for customer or tenant access to services and resources. It registers a certificate's public key and issuer name from an issued certificate that is provided to you by your customer, and maintains that information as trusted issuer objects.

To provide secure multi-tenant operations, requests are performed in the context of a user role that maps a claim token from a tenant to a Tenant administrator user role or to a Tenant self-service user role in System Center 2012 – Virtual Machine Manager (VMM). Like other user roles in VMM, you assign Windows Server Active Directory accounts to the roles, and the roles are granted access to resources.

HSP administrators can use the SPF OData services to create the required infrastructure. For more information, see the [Service Provider Foundation SDK](#). The user authentication process between your customers and System Center 2012 is shown in the figure below.

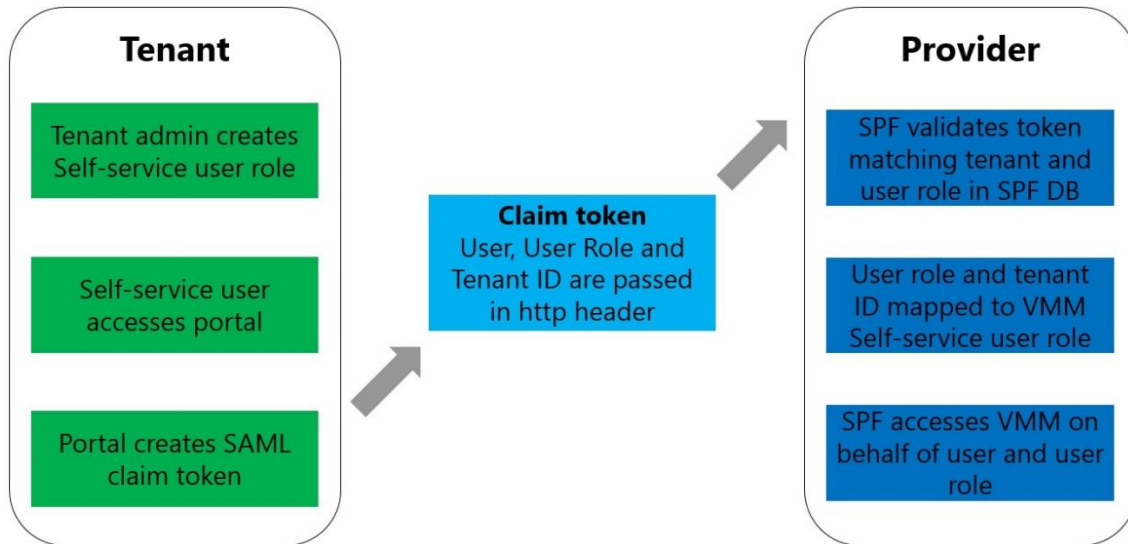


Figure 6: SPF authentication

For details about this process, see the [Managing Certificates and User Roles in Service Provider Foundation](#) article.

As shown in *Figure 5: Service provider foundation scenarios*, SPF is not only part of System Center 2012 SP1, it interacts with System Center 2012 - Virtual Machine Manager and System Center 2012 - Operations Manager components, which provide some of the technical capabilities listed in *Table 6: Products/technologies for CSFRM entities*.

In addition to the technical capabilities provided by the Operations Manager and Virtual Machine Manager components listed in *Table 6*, to provide IaaS to your customers you also need the remaining technical capabilities that are also listed in *Table 6*, which are provided by other System Center 2012 components.

The Infrastructure-as-a-Service (IaaS) Fabric Management Architecture Guide was created to help organizations develop and implement best practice fabric management architectures. The [IaaS Fabric Management Architecture Guide](#) uses the core capabilities of System Center 2012 SP1 to provide fabric management capabilities. The guide provides a lab-tested reference implementation that combines Microsoft software, hardware vendor-agnostic compute, network, and storage technologies, and additional value-added

software components, into a single, high availability System Center 2012 design. The figure below shows the IaaS fabric management architecture.

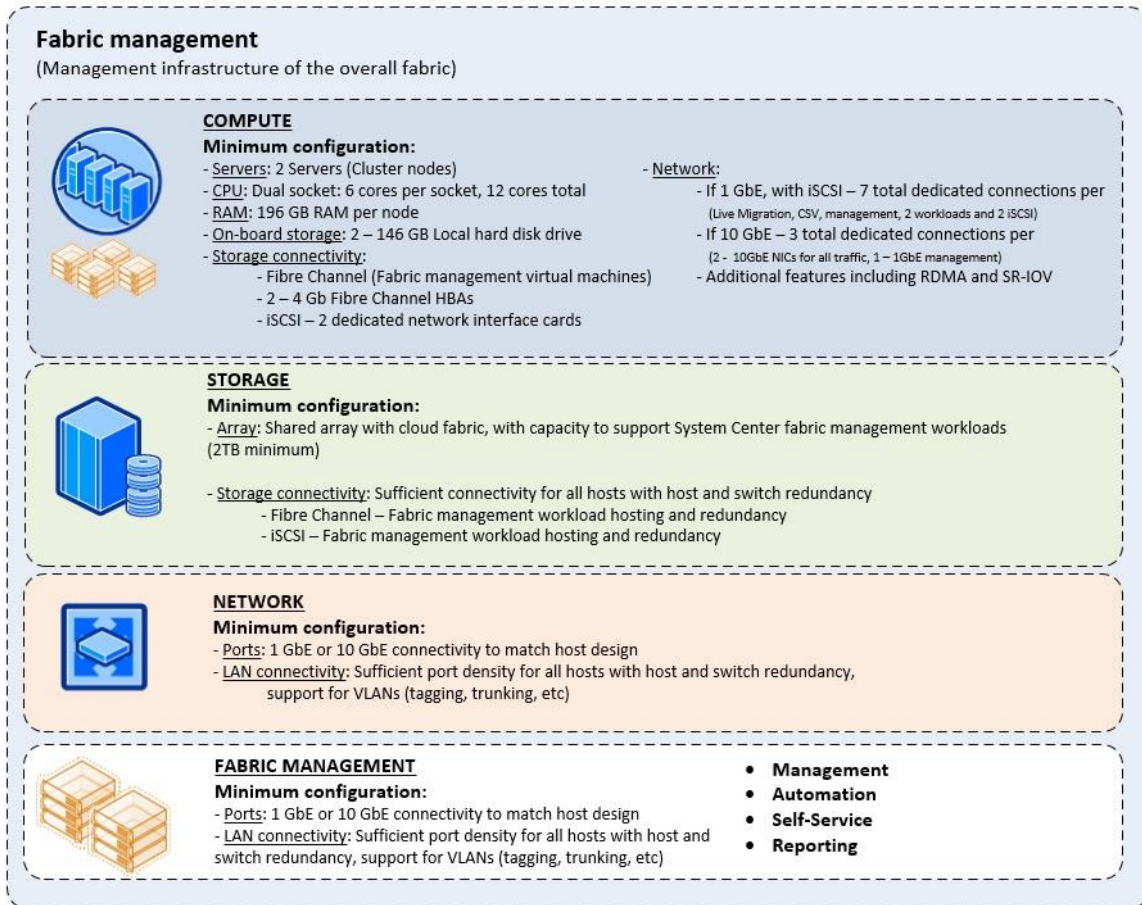


Figure 7: IaaS fabric management architecture

As noted in the "Compute" section of Figure 7, the minimum configuration consists of a two-node cluster, where both servers in the cluster are running Windows Server 2012 Hyper-V, and all of the System Center components and supporting databases execute on Hyper-V virtual machines. A more detailed view of the "Fabric Management" section of Figure 7 is shown in the figure below.

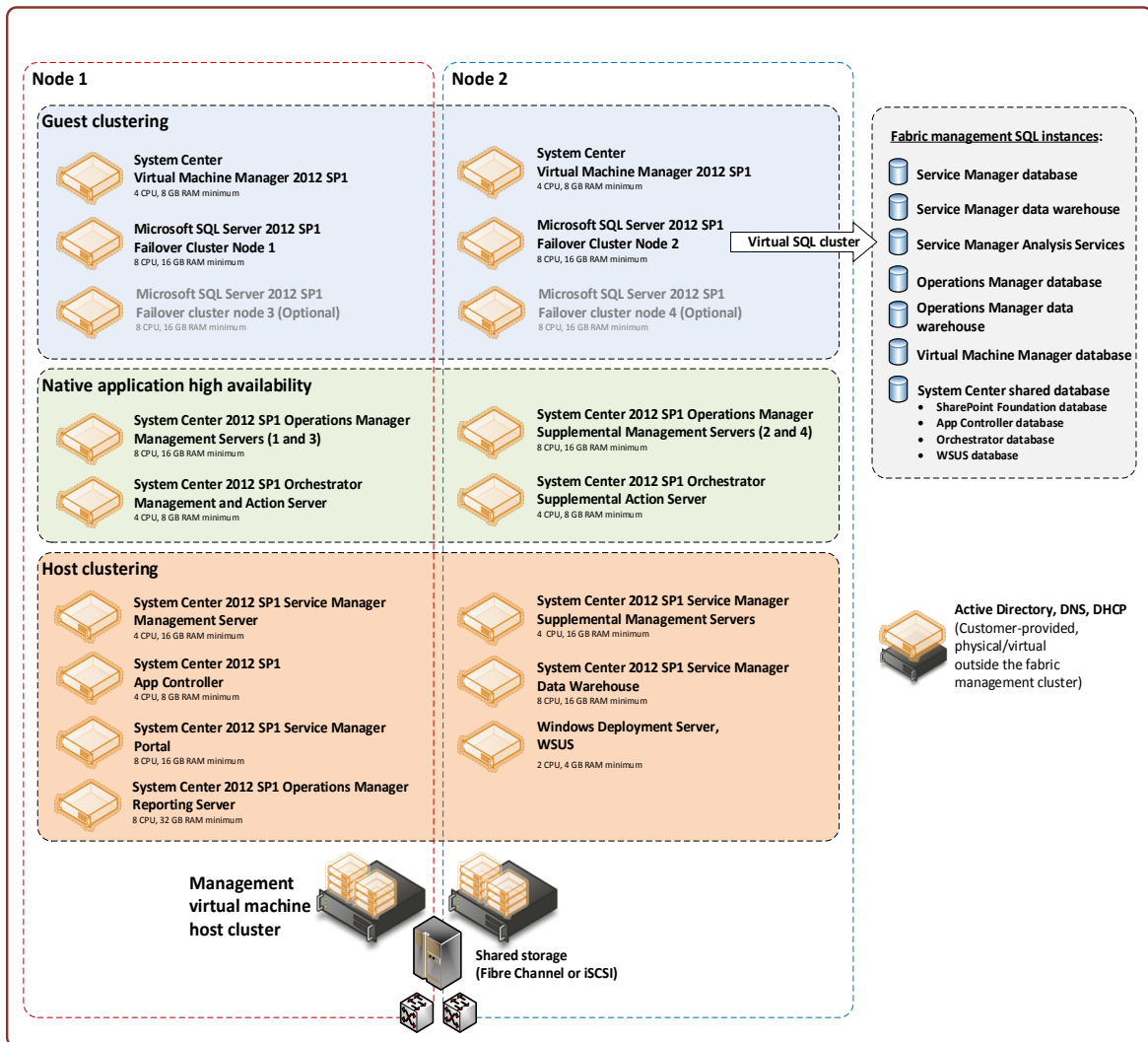


Figure 8: IaaS fabric management architecture detail

The number of physical hypervisor hosts and virtual machines that your fabric management architecture manages might require you to add additional physical hypervisor hosts to the two-node failover cluster that is represented in *Figure 7* and *Figure 8* so that virtual machines could be separated across additional nodes for scaling purposes. While *not* represented in *Figure 8*, the use of the Service Provider Foundation and Windows Azure Services for Windows Server require additional physical hosts and virtual machines to meet their unique requirements. These additional requirements are detailed separately from the [IaaS Fabric Management Architecture Guide](#) in the following documents:

- Service Provider Foundation
 - [System Requirements for Service Provider Foundation](#)
 - [Capacity Planning for Service Provider Foundation](#)
- Windows Azure Services for Windows Server

- [Service Management Portal and Service Management API Installation Guide](#)
- [Capacity Planning Windows Azure Services for Windows Server](#)

Notes:

- The [IaaS Fabric Management Architecture Guide](#) was written with enterprise IT organizations as its primary audience. As a result:
 - The term “private cloud” is used frequently in the document; however, the management platform that is described applies equally to HSP cloud infrastructures
 - There is little discussion around multi-tenancy, customer portal and billing requirements and design

When the guidance in this guide is combined with the guidance in the [Consumer and Provider Portal](#) and [Usage and Billing](#) sections of this document, and the SPF and Windows Azure Services for Windows Server guidance also in this document, this collective guidance can be used to create a complete design for the management and support technical capabilities described in *Figure 1: Cloud Services Reference Model*.

- Although the designs that are represented in the [IaaS Fabric Management Architecture Guide](#) and the [Service Provider Foundation](#) and [Windows Azure Services for Windows Server](#) documents listed previously were lab-tested *independently* of each other, they were **not** tested *together*. If you implement any of this guidance together in your environment, you’ll have to adjust the individual designs to support your scaling requirements and then test your collective design before providing services to your customers on the infrastructure.
- When implementing SPF, you might determine that you must add physical or virtual servers that execute the System Center 2012 - Virtual Machine Manager (VMM) component *separately* from other System Center 2012 components that are represented in *Figure 8: IaaS fabric management architecture detail*. This might be necessary because SPF scales based on what it refers to as a *stamp*. In SPF, a stamp is a reproducible collection of resources that includes a single VMM server **and** the physical servers and virtual machines that the VMM server manages. Even if you use Windows Server Failover Clustering to cluster multiple VMM servers for high availability, the cluster is still considered a single VMM server, for the stamp. Customers or tenants are assigned to stamps, and they access resources within the stamp to which they’re assigned.

In addition to the design and planning guidance for the functionality that is discussed in this section, installation guidance is also available for it, as described in the documents listed below:

- [Service Management Portal and API Installation Guide](#)
- [IaaS Deployment Guide](#)
- [How to Install Service Provider Foundation](#)

5.3 *Infrastructure Technical Capabilities*

This section focuses on the compute, network, storage, and virtualization infrastructure technical capabilities that are shown in *Figure 1*. The term “fabric” is used in this document to collectively refer to these capabilities when they’re managed by System Center 2012, as described in *Figure 8: IaaS fabric management architecture detail* of this document. Windows Server 2012 Hyper-V provides the virtualization capability, and can also be used to provide all or portions of your storage capability, depending on the storage strategy of your organization.

The Infrastructure-as-a-Service (IaaS) Fabric Architecture Guide was created to help organizations develop and implement best practice fabric architectures. The [IaaS Fabric Architecture Guide](#) uses the core capabilities of Windows Server 2012, and vendor-agnostic hardware technology to help organizations provide the infrastructure capabilities. The guide provides a lab-tested reference implementation that combines Microsoft software, hardware vendor-agnostic compute, network, and storage technologies, and additional value-added software components, into multiple fabric design patterns and storage architectures. The figure below shows the IaaS fabric architecture.

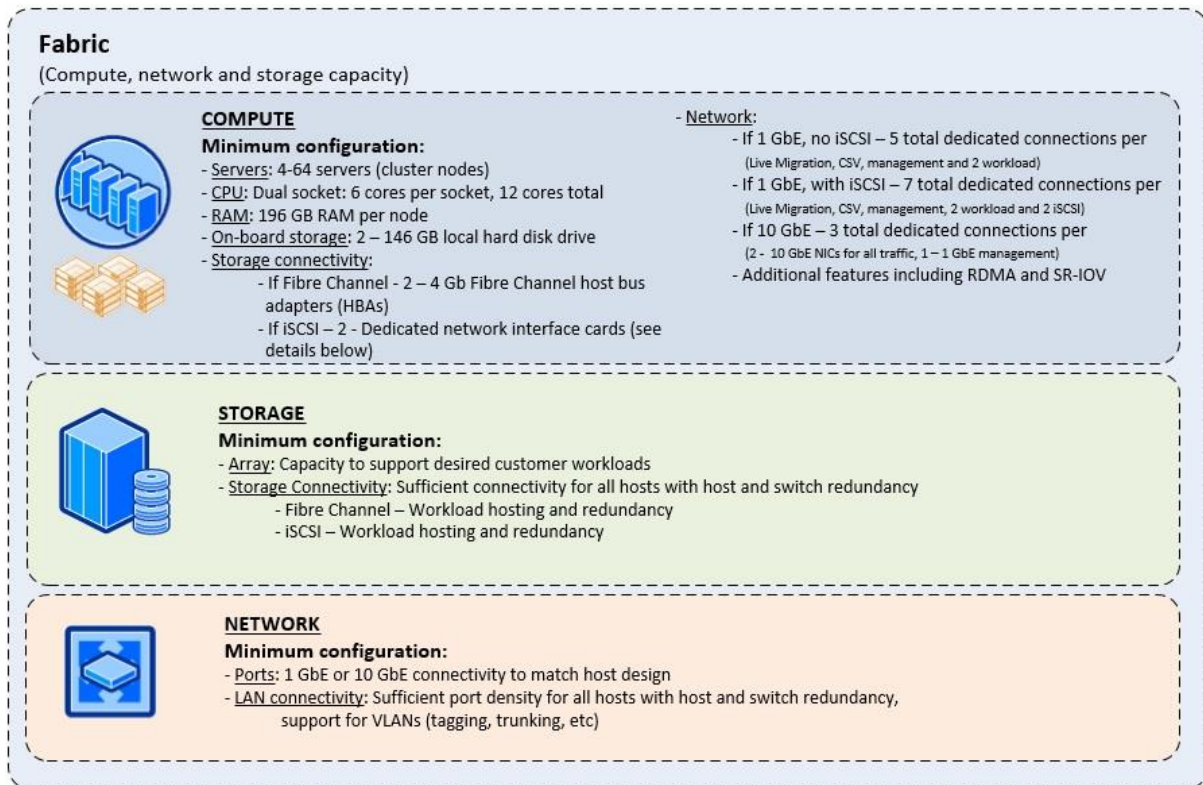


Figure 9: IaaS fabric architecture

You can assemble fabric configurations yourself, or you can choose to purchase pre-configured solutions from a number of Microsoft hardware partners through the [Microsoft Private Cloud Fast Track Program](#). Even though the term “private cloud” is used in the name of the program, the pre-configured solutions available can also be used by HSPs, if the solutions meet their requirements. The Microsoft Private Cloud Fast Track Program is a joint effort between Microsoft and its hardware partners to deliver pre-configured solutions that reduce the complexity and risk of implementing fabric and fabric management solutions. A key differentiator of the Fast Track Program is the flexibility of solutions and customer choice across the hardware vendors’ technologies.

The [IaaS Fabric Architecture Guide](#) describes multiple physical implementation patterns that can be used when you implement any of the architecture that is shown in *Figure 9*. The implementation patterns primarily differ based on the type of storage that is used to store the virtual hard disk files, and the different approaches that can be taken for the network design that transports block storage data and all other data. Windows Server 2012 introduces many new features that support storage configurations that simply were not feasible in previous versions of Windows Server. One of the patterns that is described in the document is the Continuous Availability over SMB Storage pattern. The document details three variations of that pattern:

- **Variation A:** SMB Direct that uses Shared Serial Attached SCSI (SAS)/Storage Spaces
- **Variation B:** SMB Direct that uses storage area network (SAN)
- **Variation C:** SMB 3.0-enabled storage

Note: SMB Direct is based on SMB 3.0, and it supports the use of network adapters that have Remote Direct Memory Access (RDMA) capability.

All of these patterns contain several new features that are found in Windows Server 2012 to help you provide high performance and cost-effective storage infrastructure for your IaaS offering. While this document doesn't describe each of these variations, the figure below shows Variation A.

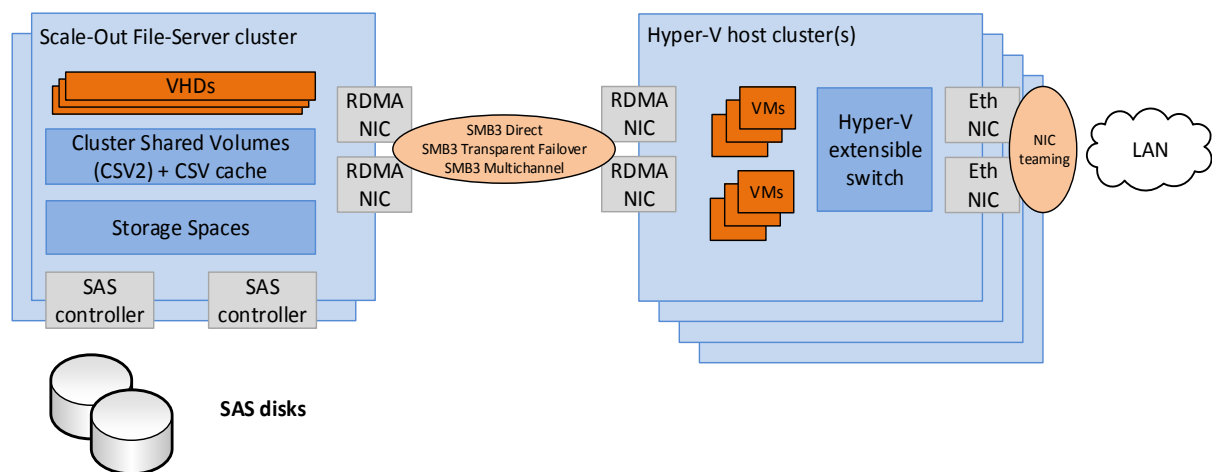


Figure 10: Continuous Availability over SMB Storage Pattern-Variation A

This pattern combines a [Scale-Out File Server](#) cluster infrastructure that uses direct-attached shared SAS storage and [Storage Spaces](#) with [SMB Direct](#) to provide a cost-effective back-end storage for virtual hard disks and SQL Server databases. This pattern has similar performance characteristics to traditional storage area network infrastructures.

Of course, instead of using Windows Server 2012 Scale-Out File Server clusters for storage in your environment, you can also use one or more of the other variations. Patterns for all of the variations and all of the configurations are described in the [IaaS Fabric Architecture Guide](#). Your answers to the questions that are listed in *Table 3: Infrastructure technical capabilities requirements* help you select the best patterns for your implementation.

6.0 Summary

This document described:

- Recommended [Solution Requirements](#) for you to define before you design a cloud services foundation on which to provide IaaS services
- A [conceptual design](#) for an IaaS infrastructure that included both a [reference model](#) and a set of [architectural principles](#) that you can use as a basis for defining your own design in your environment
- [Physical Design](#) considerations, options, and designs for the infrastructure and management and support technical capabilities that are detailed in the reference model

After you apply the information that is contained in this document and use the resources that are referenced in this document, you should be able to design a cloud services foundation with Windows Server 2012 and System Center 2012 that allows you to provide IaaS services to your customers.