OpenStack Cloud Architectures	
General purpose	
Technical considerations	General purpose clouds are expected to include these base services: • Compute • Network • Storage Each of these services have different resource requirements. As a result, you must make design decisions relating directly to the service, as well as provide a balanced infrastructure for all services. Take into consideration the unique aspects of each service, as individual characteristics and service mass can impact the hardware selection process. Hardware designs should be generated for each of the services.
Architecture	Increasing server density means sacrificing resource capacity or expandability, however, increasing resource capacity and expandability increases cost and decreases server density. As a result, determining the best server hardware for a general purpose OpenStack architecture means understanding how choice of form factor will impact the rest of the design.
Compute focused	
Technical considerations	Compute-focused clouds are a specialized subset of the general purpose OpenStack cloud architecture. A compute-focused cloud specifically supports compute intensive workloads. Compute intensive workloads may be CPU intensive, RAM intensive, or both; they are not typically storage or network intensive.
Operational considerations	A compute-focused cloud should have an emphasis on server hardware that can offer more CPU sockets, more CPU cores, and more RAM. Network connectivity and storage capacity are less critical.
Architecture	When designing a compute-focused OpenStack architecture, you must consider whether you intend to scale up or scale out. Selecting a smaller number of larger hosts, or a larger number of smaller hosts, depends on a combination of factors: cost, power, cooling, physical rack and floor space, support-warranty, and manageability.
Storage focused	

Technical considerations	Input-Output performance requirements require researching and modeling before deciding on a final storage framework. Running benchmarks for Input-Output performance provides a baseline for expected performance levels.
Operational considerations	Scaling storage solutions in a storage- focused OpenStack architecture design is driven by initial requirements, including IOPS, capacity, bandwidth, and future needs.
Architecture	Storage-focused OpenStack clouds must address I/O intensive workloads. These workloads are not CPU intensive, nor are they consistently network intensive. The network may be heavily utilized to transfer storage, but they are not otherwise network intensive. The selection of storage hardware determines the overall performance and scalability of a storage-focused OpenStack design architecture.
Network focused	
Technical considerations	All OpenStack deployments depend on network communication in order to function properly due to its service-based nature. In some cases, however, the network elevates beyond simple infrastructure. All OpenStack deployments depend on network communication in order to function properly due to its service-based nature. In some cases, however, the network elevates beyond simple infrastructure. These architectures depend on the network infrastructure and require network services that perform reliably in order to satisfy user and application requirements. These networks include Content delivery network, Network management functions, Network service offerings, Web portals or web services, High speed and high volume, transactional systems, High availability, Big data, Virtual desktop infrastructure (VDI), Voice over IP (VoIP), Video Conference or web conference, High performance computing (HPC)
Operational considerations	Network latency through slow page loads, degraded video streams, and low quality VoIP sessions impacts the user experience. Users are often not aware of how network design and architecture affects their experiences. Both enterprise customers and end-users rely on the network for delivery of an application. Network performance problems can result in a negative

	experience for the end-user, as well as productivity and economic loss.
Architecture	A wide variety of factors can affect a network-focused OpenStack architecture. While there are some considerations shared with a general use case, specific workloads related to network requirements influence network design decisions. One decision includes whether or not to use Network Address Translation (NAT) and where to implement it. If there is a requirement for floating IPs instead of public fixed addresses then you must use NAT.
Multi-site	
Technical considerations	There are many technical considerations to take into account with regard to designing a multi-site OpenStack implementation. An OpenStack cloud can be designed in a variety of ways to handle individual application needs. A multi-site deployment has additional challenges compared to single site installations and therefore is a more complex solution. Inter-site link capacity describes the capabilities of the connectivity between the different OpenStack sites. This includes parameters such as bandwidth, latency, whether or not a link is dedicated, and any business policies applied to the connection. The capability and number of the links between sites determine what kind of options are available for deployment.
Operational considerations	Multi-site OpenStack cloud deployment using regions requires that the service catalog contains per-region entries for each service deployed other than the Identity service. Most off-the-shelf OpenStack deployment tools have limited support for defining multiple regions in this fashion

Architecture	The majority of OpenStack components are designed to run within the context of a single region. The OpenStack Compute service is designed to manage compute resources within a region, with support for subdivisions of compute resources by using availability zones and cells. The OpenStack Networking service can be used to manage network resources in the same broadcast domain or collection of switches that are linked. The OpenStack Block Storage service controls storage resources within a region with all storage resources residing on the same storage network. Like the OpenStack Compute service, the OpenStack Block Storage service also supports the availability zone construct which can be used to subdivide storage resources. The OpenStack dashboard, OpenStack Identity, and OpenStack Object Storage services are components that can each be deployed centrally in order to serve multiple regions.