Solution Guide – Azure Native Qumulo for Azure Virtual Desktop profile services

A leading financial services company was looking to retire their end-of-life remote desktop platform. At the same time, they were hiring new employees to satisfy a pandemic-driven increase in demand. With 8,000 remote users providing critical business services to their customers and distributed across both the East and West Coasts, and with the expectation that they would continue to expand that number, the customer opted to move their remote desktop services to Azure Virtual Desktop (AVD), using Nerdio Manager for Enterprise to manage AVD resources and services, and using FSLogix to manage remote user profiles for all users in both deployments.

Solution requirements

Having already managed a remote-user solution at a smaller scale, and having learned what architectures and management practices worked or didn't, the customer defined the following requirements for their updated virtual desktop environment:

Scalable simplicity

The customer's previous solution stack was not able to scale to support up to 4,000 users per region within a single volume, or namespace. As the number of remote users increased, the customer would need to provision new shares on their existing file-data service to accommodate the increased demand for both capacity and IOPS.

Customer benefits

Enhanced user productivity

Compared with the old system, users were able to connect and disconnect from the Azure Virtual Desktop environment in a fraction of the time, leading to fewer help desk calls and more time serving customers.

Service elasticity

A system that automatically scales throughput in response to high demand helps keep users productive even through peak periods.

Reduced costs

Using Azure Native Qumulo for virtual desktop profile storage resulted in a net cost savings of $325,000 per year relative to their prior service.

What the customer found as they expanded was that each new share added to their operational burden: first by requiring administrative time to monitor the share's operational status, utilization levels, and performance; and second by requiring a dedicated share to be provisioned in the other Azure region to serve as a failover volume in the event of a regional outage. The previous file service did not offer native replication tools, so mirroring each share to the other region involved a complicated system of third-party tools, regular manual checks, and troubleshooting when replication failed for any of a number of reasons.

With 4,000 users in each region, and with the potential for adding new remote employees in response to continuously changing business demands, the customer needed a solution that not only offered to scale well beyond the initial number of 4,000 remote users per region, but one that also supported replication of user data at any scale to the other Azure region. Ideally, the customer wanted a solution that scaled seamlessly to any size within a single namespace in order to minimize the operational complexity of the overall solution.

Optimized for peak performance at minimal cost

Within each region, up to 4,000 remote users connecting to the solution at the same time combine to create a heavy load on the solution every morning and again every evening when they log off. An undersized storage system could struggle to support so many simultaneous requests, leading either to excessive login/logout times for each user – or failing entirely to connect some users, forcing them to attempt to reconnect. In either case, the net result is a loss of user productivity, degraded service to the organization's customers, and an undue burden on internal IT staff to manage slowdown events when they occur.

At the same time, a service that's sized to meet the throughput demands of a 30-minute login window in the morning and a 30-minute disconnect window in the evening can mean the customer pays for bandwidth that isn't necessary the other 23 hours every day.

A key requirement for the profile storage system was the ability to support the peak throughput demands generated by thousands of users all connecting to the system at the same time, but which would not incur charges for throughput that wasn't being used.

Highly available

The solution needed to remain available in the event of not just a local service interruption such as a hardware or network issue within a given region, but also to be able to rapidly recover all related services and data in the event of a region-level failure within Azure.

As part of their new AVD solution, the customer wanted to minimize the risk of service disruption, ensuring that in the event of a region-wide failure in one region, all affected users could quickly reconnect to AVD services and data in the other region to return to productivity.

Storage requirements

The customer's previous AVD profile storage service had subjected their IT team to performance bottlenecks that could not be easily resolved, and the service's lack of easy expandability had led to an unacceptably high administrative burden just to maintain normal service levels. On top of that, the service's high transaction costs were making it difficult to justify expanding the service, even as demand for users was rising.

To provide storage services for their remote user profiles, the customer considered several factors when evaluating the available options. Besides the above-mentioned requirements for seamless scalability, elastic throughput (both IOPS and servicewide bandwidth), and native replication support to a second Azure region, the customer's storage selection process also factored in the following:

- The cost per user per GB per month for baseline storage of the user's profile
- The cost per user per month for ensuring that each user's profile was replicated to the secondary region. The customer needed to consider both the cost for the capacity on the remote service, as well as the cost of replicating data between Azure regions.
- The cost for ensuring that enough throughput would be available to support each user during peak periods

Solution

The customer had already decided on supplementing their AVD service with Nerdio Manager for Enterprise for image, infrastructure and resource management. To ensure a consistent user experience in an ephemeral desktop environment, they chose FSLogix, a Microsoft service that uses a back-end file storage platform to provide user portability across both physical and virtual desktops.

In sizing up their solution requirements, the customer calculated that the storage system supporting FSLogix needed to scale to over 400TB of total capacity (each user's unique profile required an average of 50GB of disk space), be able to support 40,000 IOPS of sustained usage per region, as well as to accommodate burst throughput up to 60,000 IOPS per region during logon and logoff windows every day.

After evaluating the available Azure data services for storage profiles, the customer chose Azure Native Qumulo (ANQ), a cloud-native file-storage service, recognizing that of the storage services reviewed, only ANQ was able to meet all the customer's objectives for scalability, performance and simplicity.

The Azure Native Qumulo advantages

Having experienced a number of service interruptions and slowdowns due to their prior profile storage solution, and looking to avoid the management complexity and high transaction costs that came with it, the customer decided that Azure Native Qumulo would provide a much simpler storage service, at lower cost, than any of the other Azure based storage alternatives.

Of the storage options considered, Azure Native Qumulo offered a number of advantages, which made it the clear choice for delivering profile data services, including:

- Only Azure Native Qumulo offered cloud-native elasticity that enabled the service to provide the required throughput – including both IOPS and overall service bandwidth – completely independently of capacity. The other services all tied their services' available IOPS directly to provisioned capacity, forcing the customer in some situations to provision and pay for more capacity than they actually needed in order to reach the target levels for burst IOPS per user.
- Only Azure Native Qumulo was able to support a single namespace for all users in each region regardless of how much capacity was required. While the customer's 4,000 AVD users in each region could nominally have all shared a single volume on any of the other storage services, the customer would have to create and manage multiple volumes/shares per region in order to support the required IOPS levels.
- Azure Native Qumulo was the only service to include cross-regional replication as a core feature of the monthly subscription fee. Of the other available storage services, one did not offer replication as a core feature at all, and the other required an additional license cost, as well as a per-GB replication charge for all data that replicated in either direction.
- Only Azure Native Qumulo offered an economy of scale that reduced the per-user cost of the service as more users were added. The per-user-per-month fee for the other services remained fairly constant no matter how many users the solution supported.

Besides the management and licensing simplicity of ANQ, its cloud-native architecture meant that the customer could also leverage its capacity and throughput scalability to accommodate new users to virtually any scale and to deliver the necessary IOPS to sustain user activity, minimizing service slowdowns even during peak windows, all at a fraction of the cost of any of the other cloud-file storage services.

Architecture

The customer's Azure Virtual Desktop solution was ultimately deployed with the following components:

- Azure Native Qumulo Scalable File Service (ANQ) to host the individual VHD-based profiles of each desktop user. A separate ANQ instance has been deployed in each region.
- Azure Virtual Network
- VNet Injection to connect each region's ANQ instance to the customer's own Azure subscription resources
- Azure Virtual Desktop, deployed in two Azure regions, with a separate pool of users assigned to each region's AVD resources as their primary site, and each region set up as the secondary site for the other region in the event of a regional service interruption.
- Nerdio Manager to simplify and streamline the process of managing AVD-related services: resource pools, connectivity, security, desktop images, applications, and service monitoring.
- FSLogix Profile Containers to connect each AVD user to their assigned profile on the ANQ storage as part of the login process
- Qumulo Continuous Replication, configured to replicate user profile data from each region's local ANQ cluster to the ANQ instance in the other region, ensuring that user profile services will still be available in the event of a regional failover.

Topology

The customer's AVD solution was deployed in the Azure East US 2 and Azure West US 2 regions, with users evenly divided between the two. Each remote user connects to the region closest to their physical location.

To enable the entire AVD service to remain online in the event of a failure in one of the hosting Azure regions, each region is configured as a failover domain from the other Azure region. In the event of a region-wide outage, the complete set of AVD services will come online in the remaining region.

User profiles are replicated bidirectionally from the local Azure Native Qumulo service instance in each region to the remote ANQ target. This ensures that user profile data will also be available in either region in the event of a regional outage.

The full solution architecture, including supporting AVD services, Nerdio Manager for Enterprise, FSLogix, and Azure Native Qumulo data services, is shown in Figure 1.
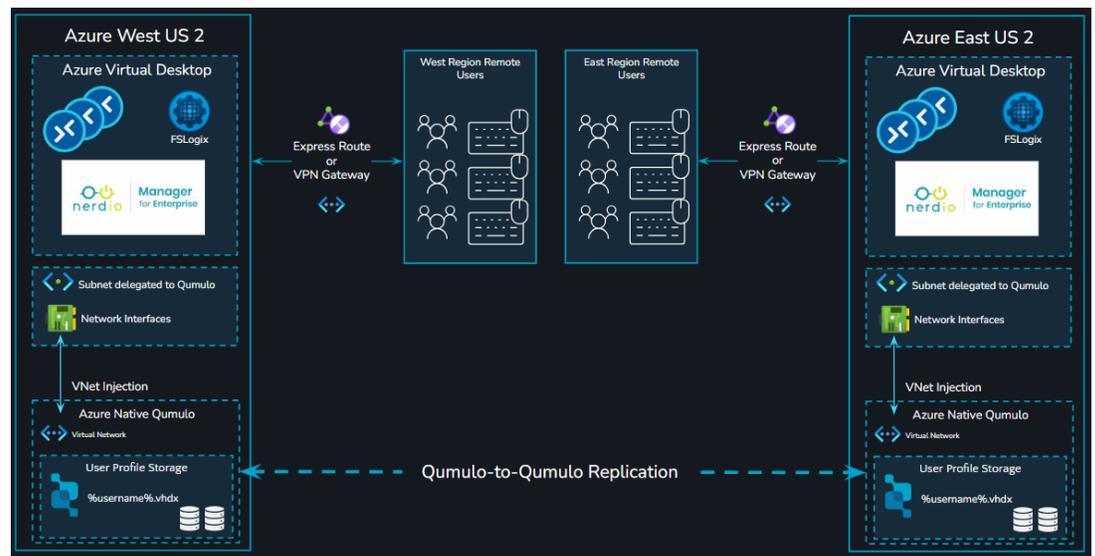
Figure 1: The customer's multi-region Azure Virtual Desktop environment

Solution benefits

As a result of choosing a multi-region remote-worker solution based on Azure Virtual Desktop, with user profile services hosted on Azure Native Qumulo, the customer was able to realize the following benefits:

Enhanced User Productivity

Compared with the customer's earlier virtual desktop deployment, the new service configuration enabled faster login times for each user every morning. The result was a more productive user base, fewer calls to the internal IT support staff, and less time spent in troubleshooting service availability and performance issues.

Service Elasticity

An undersized solution can impair user productivity during peak periods – e.g. during logon and logoff times, or under heavy utilization periods – leading to longer hold times, overburdened IT staff, frustrated customers, and potential loss of revenue. An oversized solution can incur significant operational costs if it's only fully utilized for a few hours or less per day.

As deployed, using Azure Native Qumulo to provide user-profile storage, the service was able to add IOPS and throughput capacity to the service during peak load periods, and then automatically return to normal services at all other times. This meant that the customer was not paying more than necessary for an oversized solution, nor were they causing unnecessary service slowdowns with an undersized solution.

Resiliency

Replicated user profiles are read-only under normal circumstances. The solution's RTO should include the time needed to fail over to the secondary ANQ instance (e.g. break the replication relationship and make all profiles writable) before connecting users from the remote region to AVD instances.

Cost

In fact, Azure Native Qumulo's scalability and elasticity directly resulted in the most significant benefit to the customer: an economy of scale that reduced the overall cost of services to a much lower price point per user per month than they would have experienced with any of the other profile storage services the customer considered.

By the customer's own reckoning, their choice to use Azure Native Qumulo for AVD profile data services resulted in a $325,000 cost savings per year relative to their previous solution's storage environment!

Conclusion

Based on their previous remote desktop user solution, the customer had a lot of experience with what didn't work: their prior solution was too complex, too slow, and too unreliable. In evaluating other Azure-based cloud file solutions for storing user profiles, they determined that the alternatives to Azure Native Qumulo were all of the above, as well as too expensive.

The customer realized that of all the options for Azure Virtual Desktop profile storage, only Azure Native Qumulo met the customer's requirements for single-namespace scalability, performance – delivering both sustained and burst IOPS as needed – and simplicity, even in a multi-regional deployment utilizing Qumulo's native replication features, all at a lower cost.

Related resources

[Azure Native Qumulo Scalable File Service](#)

[Qumulo Continuous Replication](#)

[Azure Native Qumulo Scalable File Service](#) (Azure Marketplace)

[Azure Native Qumulo Scalable File Service](#) (Azure blog)

[Azure Native Qumulo Scalable File Service Guide](#) (Azure Product Documentation)

[Using Failover with Replication in Qumulo](#)

[Qumulo Replication: Make Target Writable](#)