ONTAP configuration on ASA r2 systems

Enterprise applications

NetApp
February 10, 2026

Table of Contents

ONTAP configuration on ASA r2 systems

RAID

RAID refers to the use of parity-based redundancy to protect data against drive failure. ASA r2 uses the same ONTAP RAID technologies as AFF and FAS systems, ensuring robust protection against multiple disk failures.

ONTAP performs RAID configuration automatically for ASA r2 systems. This is a core component of the simplified storage management experience introduced with the ASA r2 personality.

Key details regarding automatic RAID configuration on ASA r2 include:

- Storage Availability Zones (SAZ): Instead of manually managing traditional aggregates and RAID groups, ASA r2 uses Storage Availability Zones (SAZs). These are shared, RAID-protected pools of disks for an HA pair, where both nodes have full access to the same storage.
- Automatic Placement: When a storage unit (LUN or NVMe namespace) is created, ONTAP automatically creates a volume within the SAZ and places it for optimal performance and capacity balance.
- No Manual Aggregate Management: Traditional aggregate and RAID group management commands are not supported on ASA r2. This eliminates the need for administrators to manually plan RAID group sizes, parity disks, or node assignments.
- Simplified Provisioning: Provisioning is handled via System Manager or simplified CLI commands that focus on storage units rather than the underlying physical RAID layout.
- Workload Rebalancing: Beginning with 2025 releases (ONTAP 9.17.1), ONTAP automatically rebalances workloads between nodes in the HA pair to ensure performance and space utilization remain balanced without manual intervention.

ASA r2 automatically uses ONTAP’s default RAID technologies: RAID DP for most configurations and RAID-TEC for very large SSD pools. This eliminates the need for manual RAID selection. These parity-based RAID levels provide better storage efficiency and reliability than mirroring, which older Oracle best practices often recommend but is not relevant for ASA r2. ONTAP avoids the traditional RAID write penalty through WAFL integration, ensuring optimal performance for Oracle workloads such as redo logging and random data-file writes. Combined with automated RAID management and Storage Availability Zones, ASA r2 delivers high availability and enterprise-grade protection for Oracle databases.

Capacity management

Managing a database or other enterprise application with predictable, manageable, high performance enterprise storage requires some free space on the drives for data and metadata management. The amount of free space required depends on the type of drive used, and business processes.

ASA r2 uses Storage Availability Zones (SAZ) instead of aggregates, but the principle remains the same: free space includes any physical capacity not consumed by actual data, snapshots, or system overhead. Thin provisioning must also be considered—logical allocations do not reflect true physical usage.

NetApp recommendations for ASA r2 storage systems used for enterprise applications are as follows:

SSD pools in ASA r2 systems

> **NetApp recommends** maintaining a minimum of 10% free physical space in ASA r2 environments. This guideline applies to SSD-only pools used by ASA r2 systems and includes all unused space within the SAZ and storage units. Logical space is unimportant; the focus is on the actual free physical space available for data storage.

While ASA r2 can sustain high utilization without performance degradation, operating near full capacity increases the risk of space exhaustion and administrative overhead when expanding storage. Running at over 90% utilization may not impact performance but can complicate management and delay provisioning of additional drives.

ASA r2 systems support storage units up to 128TB and SAZ sizes up to 2PB per HA pair, with ONTAP automatically balancing capacity across nodes. Monitoring utilization at the cluster, SAZ, and storage unit levels is essential to ensure adequate free space for snapshots, thin-provisioned workloads, and future growth. If capacity approaches critical thresholds (~ 90% utilization), additional SSDs should be added in groups (minimum six drives) to maintain performance and resiliency.

Storage Virtual Machines

Oracle database storage management on ASA r2 systems is also centralized on a Storage Virtual Machine (SVM), known as a vserver at the ONTAP CLI.

An SVM is the fundamental unit of storage provisioning and security in ONTAP, similar to a guest VM on a VMware ESX server. When ONTAP is first installed on ASA r2, it has no data-serving capabilities until an SVM is created. The SVM defines the personality and data services for the SAN environment.

ASA r2 systems use a SAN-only ONTAP personality, which is streamlined to support block protocols (FC, iSCSI, NVMe/FC, NVMe/TCP) and removes NAS-related features. This simplifies management and ensures that all SVM configurations are optimized for SAN workloads. Unlike AFF/FAS systems, ASA r2 does not expose options for NAS services such as home directories or NFS shares.

When a cluster is created, ASA r2 automatically provisions a default data SVM named svm1 with SAN protocols enabled. This SVM is ready for block storage operations without requiring manual configuration of protocol services. By default, IP data LIFs in this SVM support iSCSI and NVMe/TCP protocols and use the default-data-blocks service policy, which simplifies initial setup for SAN workloads. Administrators can later create additional SVMs or customize LIF configurations based on performance, security, or multi-tenant requirements.

> Logical interfaces (LIFs) for SAN protocols should be designed based on performance and availability requirements. ASA r2 supports iSCSI, FC, and NVMe LIFs, but note that automatic iSCSI LIF failover is not enabled by default because ASA r2 uses shared networking for NVMe and SCSI hosts. To enable automatic failover, create iSCSI-only LIFs.

SVMs

As with other ONTAP platforms, there is no official best practice for the number of SVMs to create; the decision depends on management and security requirements.

Most customers operate a single primary SVM for day-to-day operations and create additional SVMs for special needs, such as:

- A dedicated SVM for a critical business database managed by a specialist team
- An SVM for a development group with delegated administrative control
- An SVM for sensitive data requiring restricted administrative access

In multi-tenant environments, each tenant can be assigned a dedicated SVM. The limit for the number of SVMs and LIFs per cluster, HA pair, and node are dependent on the protocol being used, the node model, and the version of ONTAP. Consult the NetApp Hardware Universe for these limits.

> ASA r2 supports up to 256 SVMs per cluster and per HA pair starting with ONTAP 9.18.1 (previously 32 in earlier releases).

Performance management with ONTAP QoS on ASA r2 systems

Safely and efficiently managing multiple Oracle databases on ASA r2 requires an effective QoS strategy. This is especially important because ASA r2 systems are all-flash SAN platforms designed for extremely high performance and workload consolidation.

A relatively small number of SSDs can saturate even the most powerful controllers, so QoS controls are essential to ensure predictable performance across multiple workloads.
As a reference, ASA r2 systems such as the ASA A1K or A90 can deliver hundreds of thousands to over a million IOPS with sub-millisecond latency. Very few single workloads would consume this level of performance, so full utilization typically involves hosting multiple databases or applications. Doing this safely requires QoS policies to prevent resource contention.

ONTAP QoS on ASA r2 works the same way as on AFF/FAS systems, with two primary types of controls: IOPS and bandwidth. QoS controls can be applied to SVMs and LUNs.

IOPS QoS

IOPS-based QoS limits the total IOPS for a given resource. In ASA r2, QoS policies can be applied at the SVM level and to individual storage objects such as LUNs. When a workload reaches its IOPS limit, additional I/O requests queue for tokens, which introduces latency. This is expected behavior and prevents any single workload from monopolizing system resources.

> Be cautious when applying QoS controls to database transaction/redo log data. These workloads are bursty, and a QoS limit that seems reasonable for average activity may be too low for peak bursts, causing severe performance issues. In general, redo and archive logging should not be limited by QoS.

Bandwidth QoS

Bandwidth-based QoS limits throughput in Mbps. This is useful when workloads perform large block reads or writes, such as full table scans or backup operations, which consume significant bandwidth but relatively few IOPS. Combining IOPS and bandwidth limits can provide more granular control.

Minimum/guaranteed QoS

Minimum QoS policies reserve performance for critical workloads. For example, in a mixed environment with production and development databases, apply maximum QoS to development workloads and minimum QoS to production workloads to ensure predictable performance.

Adaptive QoS

Adaptive QoS adjusts limits based on the size of the storage object. While rarely used for databases (because size does not correlate to performance needs), it can be useful for virtualization workloads where performance requirements scale with capacity.

Efficiency

ONTAP space efficiency features are fully supported and optimized for ASA r2 systems. In almost all cases, the best approach is to leave the defaults in place with all efficiency features enabled.

ASA r2 systems are all-flash SAN platforms, so efficiency technologies such as compression, compaction, and deduplication are critical for maximizing usable capacity and reducing costs.

Compression

Compression reduces space requirements by encoding patterns in data. With SSD-based ASA r2 systems, compression delivers significant savings because flash eliminates the need for overprovisioning for performance. ONTAP adaptive compression is enabled by default and has been thoroughly tested with enterprise workloads, including Oracle databases, with no measurable performance impact—even in environments where latency is measured in microseconds. In some cases, performance improves because compressed data occupies less cache space.

> ⓘ Temperature-sensitive storage efficiency (TSSE) is not applied on ASA r2 systems. On ASA r2 systems, compression is not based on hot (frequently accessed) data or cold (infrequently accessed) data. Compression begins without waiting for data to become cold.

Adaptive compression

Adaptive compression uses an 8KB block size by default, matching the block size commonly used by relational databases. Larger block sizes (16KB or 32KB) can improve efficiency for sequential data such as transaction logs or backups but should be used cautiously for active databases to avoid overhead during overwrites.

> ⓘ Block size can be increased up to 32KB for quiescent files such as logs or backups. Consult NetApp guidance before changing defaults.

> ! Do not use 32KB compression with deduplication for streaming backups. Use 8KB compression to maintain deduplication efficiency.

Compression alignment

Compression alignment matters for random overwrites. Ensure correct LUN type, partition offset (multiple of 8KB), and filesystem block size aligned to database block size. Sequential data such as backups or logs does not require alignment considerations.

Data compaction

Compaction complements compression by allowing multiple compressed blocks to share the same physical block. For example, if an 8KB block compresses to 1KB, compaction ensures that the remaining space is not wasted. This feature is inline and does not introduce performance penalties.

Deduplication

Deduplication removes duplicate blocks across datasets. While Oracle databases typically yield minimal deduplication savings due to unique block headers and trailers, ONTAP deduplication can still reclaim space from zeroed blocks and repeated patterns.

Efficiency and thin provisioning

ASA r2 systems use thin provisioning by default. Efficiency features complement thin provisioning to maximize usable capacity.

> Storage units are always thinly provisioned on ASA r2 storage systems. Thick provisioning is not supported.

QuickAssist Technology (QAT)

In NetApp ASA r2 platforms, Intel QuickAssist Technology (QAT) provides hardware-accelerated efficiency that differs significantly from software-based Temperature-Sensitive Storage Efficiency (TSSE) without QAT.

QAT with hardware acceleration:

- Offloads compression and encryption tasks from CPU cores.
- Enables immediate, inline efficiency for both hot (frequently accessed) and cold (infrequently accessed) data.
- Significantly reduces CPU overhead.
- Delivers higher throughput and lower latency.
- Improves scalability for performance-sensitive operations such as TLS and VPN encryption.

TSSE without QAT:

- Relies on CPU-driven processes for efficiency operations.
- Applies efficiency only to cold data after a delay.
- Consumes more CPU resources.
- Limits overall performance compared to QAT-accelerated systems.

Modern ASA r2 systems therefore deliver faster, hardware-accelerated efficiency and better system utilization than older TSSE-only platforms.

Efficiency best practices for ASA r2

NetApp recommends the following:

ASA r2 defaults

Storage units created on ONTAP running on ASA r2 systems are thin provisioned with all inline efficiency features enabled by default, including compression, compaction, and deduplication. Although Oracle databases generally do not benefit significantly from deduplication and may include uncompressible data, these defaults are appropriate for almost all workloads. ONTAP is designed to efficiently process all types of data and I/O patterns, whether or not they result in savings. Defaults should only be changed if the reasons are fully understood and there is a clear benefit to deviating.

General recommendations

- Disable Compression for Encrypted or App-Compressed Data: If files are already compressed at the application level or encrypted, disable compression to optimize performance and allow more efficient operation on other storage units.
- Avoid Combining Large Compression Blocks with Deduplication: Do not use both 32KB compression and deduplication for database backups. For streaming backups, use 8KB compression to maintain deduplication efficiency.
- Monitor Efficiency Savings: Use ONTAP tools (System Manager, Active IQ) to track actual space savings and adjust policies if needed.

Thin provisioning

Thin provisioning for an Oracle database on ASA r2 requires careful planning because it involves configuring more logical space than is physically available. When implemented correctly, thin provisioning delivers significant cost savings and improved manageability.

Thin provisioning is integral to ASA r2 and closely related to ONTAP efficiency technologies because both allow more logical data to be stored than the physical capacity on the system. ASA r2 systems are SAN-only, and thin provisioning applies to storage units and LUNs within Storage Availability Zones (SAZ).

> ASA r2 storage units are thin provisioned by default.

Almost any use of snapshots involves thin provisioning. For example, a typical 10 TiB database with 30 days of snapshots might appear as 310 TiB of logical data, but only 12 TiB to 15 TiB of physical space is consumed because snapshots store only changed blocks.

Similarly, cloning is another form of thin provisioning. A development environment with 40 clones of an 80 TiB database would require 3.2 PiB if fully written, but in practice consumes far less because only changes are stored.

Space management

Some care must be taken with thin provisioning in an application environment because data change rates can increase unexpectedly. For example, space consumption due to snapshots can grow rapidly if database tables are reindexed, or wide-scale patching is applied to VMware guests. A misplaced backup can write a large amount of data in a very short time. Finally, it can be difficult to recover some applications if a LUN runs out of free space unexpectedly.

In ASA r2, these risks are mitigated through **thin provisioning**, **proactive monitoring**, and **LUN resize policies**, rather than ONTAP features like volume-autogrow or snapshot-autodelete. Administrators should:

- Enable thin provisioning on LUNs (`space-reserve disabled`) - this is the default setting in ASA r2

- Monitor capacity using System Manager alerts or API-based automation
- Use scheduled or scripted LUN resize to accommodate growth
- Configure snapshot reserve and automatic snapshot deletion via System Manager (GUI)

> Careful planning of space thresholds and automation scripts is essential because ASA r2 does not support automatic volume growth or CLI-driven snapshot deletion.

ASA r2 does not use fractional reserve settings because it is a SAN-only architecture that abstracts WAFL-based volume options. Instead, space efficiency and overwrite protection are managed at the LUN level. For example, if you have a 250 GiB LUN provisioned from a storage unit, snapshots consume space based on actual block changes rather than reserving an equal amount of space upfront. This eliminates the need for large static reservations, which were common in traditional ONTAP environments using fractional reserve.

> If guaranteed overwrite protection is required and monitoring is not feasible, administrators should provision sufficient capacity in the storage unit and set snapshot reserve appropriately. However, ASA r2's design makes fractional reserve unnecessary for most workloads.

Compression and deduplication

Compression and deduplication in ASA r2 are space efficiency technologies, not traditional thin provisioning mechanisms. These features reduce the physical storage footprint by eliminating redundant data and compressing blocks, allowing more logical data to be stored than the raw capacity would otherwise permit.

For example, a 50 TiB dataset might compress to 30 TiB, saving 20 TiB of physical space. From the application perspective, there is still 50 TiB of data, even though it occupies only 30 TiB on disk.

> The compressibility of a dataset can change over time, which may increase physical space consumption. Therefore, compression and deduplication must be managed proactively through monitoring and capacity planning.

Free space and LVM space allocation

Thin provisioning in ASA r2 environments can lose efficiency over time if deleted blocks are not reclaimed. Unless space is released using TRIM/UNMAP or overwritten with zeros (via ASMRU - Automatic Space Management and Reclamation Utility), deleted data continues to consume physical capacity. In many Oracle database environments, thin provisioning offers limited benefit because datafiles are typically pre-allocated to their full size during creation.

Careful planning of LVM configuration can improve efficiency and minimize the need for storage provisioning and LUN resizing. When an LVM such as Veritas VxVM or Oracle ASM is used, the underlying LUNs are divided into extents that are only used when needed. For example, if a dataset begins at 2 TiB in size but could grow to 10 TiB over time, this dataset could be placed on 10 TiB of thin-provisioned LUNs organized in an LVM diskgroup. It would occupy only 2 TiB of space at the time of creation and would only claim additional space as extents are allocated to accommodate data growth. This process is safe as long as space is monitored.

ONTAP failover

An understanding of storage takeover functions is required to ensure that Oracle database operations are not disrupted during these operations. In addition, the arguments used by takeover operations can affect data integrity if used incorrectly.

Under normal conditions, incoming writes to a given controller are synchronously mirrored to its HA partner. In an ASA r2 environment with SnapMirror Active Sync (SM-as), writes are also mirrored to a remote controller at the secondary site. Until a write is stored in non-volatile media in all locations, it is not acknowledged to the host application.

The media storing the write data is called non-volatile memory (NVMEM). It is sometimes referred to as non-volatile random-access memory (NVRAM) and can be thought of as a write journal rather than a cache. During normal operation, data from NVMEM is not read; it is only used to protect data in the event of a software or hardware failure. When data is written to drives, the data is transferred from system RAM, not from NVMEM.

During a takeover operation, one node in an HA pair takes over the operations from its partner. In ASA r2, switchover is not applicable because MetroCluster is not supported; instead, SnapMirror Active Sync provides site-level redundancy. Storage takeover operations during routine maintenance should be transparent, other than a brief pause in operations as network paths change. Networking can be complex, and errors are easy to make, so NetApp strongly recommends testing takeover operations thoroughly before putting a storage system into production. Doing so is the only way to ensure that all network paths are configured correctly.
In a SAN environment, verify path status using the command `sanlun lun show -p` or the operating system's native multipathing tools to ensure all expected paths are available. ASA r2 systems provide all active optimized paths for LUNs, and customers using NVMe namespaces should rely on OS-native tools, as NVMe paths are not covered by sanlun.

Care must be taken when issuing a forced takeover. Forcing a change to storage configuration means that the state of the controller that owns the drives is disregarded and the alternative node forcibly takes control of the drives. Incorrect forcing of a takeover can result in data loss or corruption because a forced takeover can discard the contents of NVMEM. After the takeover is complete, the loss of that data means that the data stored on the drives might revert to a slightly older state from the point of view of the database.

A forced takeover with a normal HA pair should rarely be required. In almost all failure scenarios, a node shuts down and informs the partner so that an automatic failover takes place. There are some edge cases, such as a rolling failure in which the interconnect between nodes is lost and then one controller fails, in which a forced takeover is required. In such a situation, the mirroring between nodes is lost before the controller failure, which means that the surviving controller no longer has a copy of the writes in progress. The takeover then needs to be forced, which means that data potentially is lost.

NetApp recommends taking the following precautions:

- Be very careful to not accidentally force a takeover. Normally, forcing should not be required, and forcing the change can cause data loss.
- If a forced takeover is required, make sure that the applications are shut down, all file systems are dismounted, and logical volume manager (LVM) volume groups are varyoffed. ASM diskgroups must be unmounted.
- In the event of a site-level failure when using SM-as, the ONTAP Mediator assisted automatic unplanned failover will be initiated on the surviving cluster, resulting in a brief I/O pause and then database transitions will continue from the surviving cluster. For more information, see the SnapMirror active sync on ASA r2 systems for detailed configuration steps.