

FEATURE BRIEF

Dynamic Quality of Service

Businesses are looking to reduce CAPEX and OPEX by consolidating their server and storage infrastructure and running more applications on them. Flash arrays and hyperconverged infrastructure are rapidly becoming the solutions of choice as customers modernize their data centers. However, many are disappointed when they encounter inconsistent performance caused by resource contention, software overhead, controller bottlenecks, and other system constraints. Most Quality of Service (QoS) features are not designed to address these issues, because they only set caps on performance, and are unable to prioritize I/O or data protection operations.

Pivot3’s Dynamic QoS allows business to meet application SLAs with easy to manage policies that provision performance, prioritize workloads, and manage data placement and protection. With the ability to automate and schedule granular QoS settings, from a LUN/datastore level down to an individual VM/VMDK, Dynamic QoS significantly reduces the time it takes to manage performance and data protection.

Pivot3 Dynamic QoS

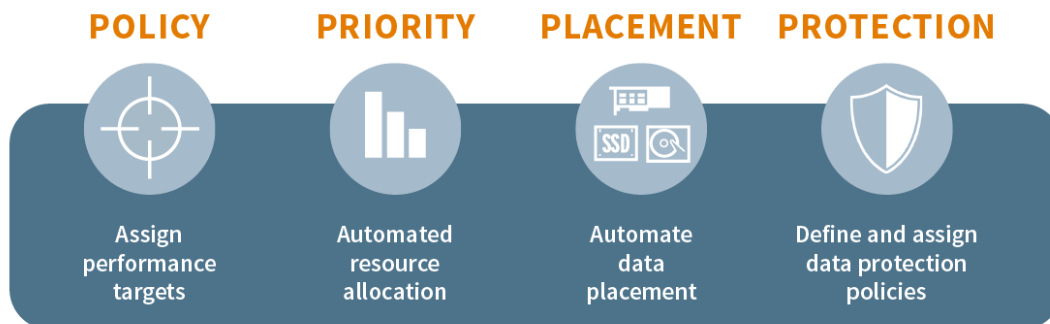


Figure 1: Dynamic QoS Capabilities all work together in real-time to achieve SLAs

Policy

Each business workload is unique, which is why Pivot3 allows customers to assign an appropriate performance level to each one. QoS performance policies define performance targets for each workload that manage minimum IOPS and throughput, and maximum latency. QoS policies are assigned when a volume is created and can be changed on-the-fly as business needs require. By automating policy changes, QoS scheduling give IT the agility to support the business as application priorities and workloads change.

Priorities

Each QoS policy has a corresponding Service Level that automatically governs how each one is prioritized when the system is under load. There are three service levels built into the QoS engine: Mission-Critical, Business-Critical, and Non-Critical, and each is pre-assigned to QoS policy. They instruct the system on how important it is to maintain each QoS policy's targets. For example, if there's resource contention, QoS policy 1 (Mission-critical) will be maintained by prioritizing its I/O requests over Non-Critical workloads first, and Business-Critical workloads, if necessary.

Placement

Dynamic QoS manages where data is stored in real-time (RAM, PCIe flash, SSD or HDD) to ensure performance Service Levels are met. The caching and tiering algorithms of the system are directly tied to the QoS policies and priorities in order to ensure that the right data is placed in the appropriate storage medium to deliver on the specified performance targets. The data is moved between tiers in real-time as workloads change, ensuring predictable performance and the most efficient use of system resources.

Protection

Data protection QoS provides policy-based management and automation of snapshots, replication and retention to best align with data protection needs. Policies can be applied to volumes or groups of volumes, and can be scheduled, allowing customers to set pre-defined schedules for both performance and data protection policies that can change automatically at pre-determined times. A queuing mechanism ensures that mission-critical jobs go to front of the queue to be executed first, business-critical jobs next, and finally, non-critical jobs.

Meet SLAs in Three Simple Steps



Dynamic QoS is available as part of the all-inclusive software feature set in the Pivot3 vSTAC SLX Hyperconverged Infrastructure and N5 Flash Arrays.