BROCADE
A Broadcom Company

Technical Brief

Fabric Operating System 9.0 Traffic Optimizer

Abstract

In general, three parameters dictate network quality: latency, throughput, and errors. Traffic Optimizer optimizes all three. In an ideal lossless network, response time and performance variations among the traffic flows sharing resources are minute. Traffic Optimizer organizes flows such that those with similar performance characteristics share the same dedicated resources. Traffic Optimizer optimizes the listed performance characteristics by sorting traffic into performance groups (PGs):

- Latency – Frame propagation delay or I/O response time
- Throughput – Port bit rate or IOPS
- Errors – Frame delivery, bit errors, or I/O failures

Specific performance attributes, such as destination port speed and protocol, are used to classify traffic. The performance groups are bound to dedicated virtual channels (VCs) within the fabric. Each virtual channel has three properties that Traffic Optimizer utilizes: isolation, dedicated resources for flow control, and independent prioritization. The autonomous control of segregated flows based on performance characteristics prevents lower performance groups from blocking higher performance groups. Traffic Optimizer technology improves network performance by lowering I/O response times and increasing overall SAN performance. The clean, clear flow of like speed traffic reduces contention and associated errors.

Technology evolves, and Brocade has optimized VC efficiency by enhancing effectiveness through targeted flow performance characteristics. Demands on an enterprise SAN have never been more significant due to other complementary technology evolutions such as Non-Volatile Memory Express (NVMe), all-flash array (AFA), and host virtualization. Flows compete for fabric resources, and head-of-line-blocking (HoLB) is not an option. Blocking is efficaciously dealt with using Traffic Optimizer technology.

Congestion

What is congestion in a Fibre Channel SAN? Congestion is a condition where the demand for a particular network resource is higher than what is available. Fibre Channel congestion does not result in buffer overflows and subsequent frame loss. Nevertheless, buffers in a SAN do fill; they just do not overflow because ingress traffic yields to full buffers. The following are the crucial SAN resources involved with congestion. Traffic Optimization primarily focuses on the first resource, port bandwidth:

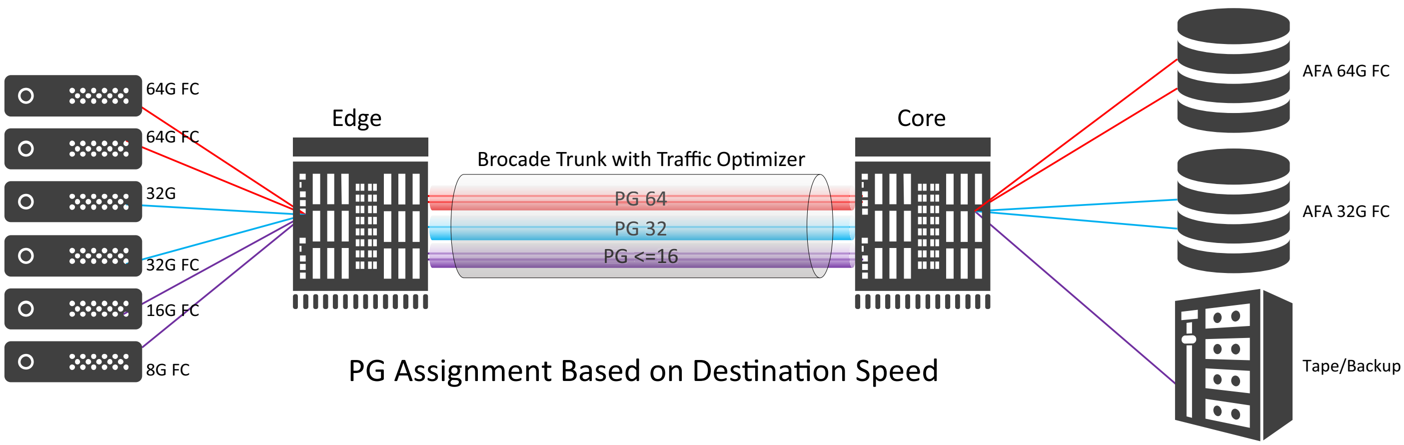
- Port bandwidth – Lack of port bandwidth results in oversubscription.
- Credits – Lack of credits results in credit-stalled devices.
- Buffers – Buffer starvation results in an inadequate number of credits and droop.

Fibre Channel is designed to avoid packet drops through flow control in congestion scenarios, unlike other lossy protocols. However, sustained and severe congestion can lead to a performance impact and frame drops as a last resort. All three of these situations result in fabric congestion; Traffic Optimizer focuses on the first point by segregating flows to devices with different destination port speeds:

- The destination port is physically slower than the sender's port speed (oversubscription).
- The receiving port is slower than the aggregate of multiple sending ports (oversubscription).
- The destination device cannot accept data as quickly as it arrives (credit stall, an insufficient number of credits).

Buffer credits are returned at the rate at which a receiving device can accommodate, which may reduce the rate of many other flows that share a port along the path. A 32G device is much less likely to run at 8G compared to a 16G device. Ultimately, the real benefit is separation based on inherent capability; Traffic Optimizer uses speed as a means to this end. The fabric knows the destination port speeds and can classify flows into performance groups (PGs) accordingly; in turn, PGs are assigned to virtual channels. See the following figure.

Figure 1: Host to Storage via Traffic Optimizer VCs



FC Flow Control

Fibre Channel (FC) technology continues to evolve through innovation and enhancements in performance, management, automation, NVMe, Extension, FCR, analytics, and better flow control strategies. Traffic Optimizer is Brocade's next-generation adjunct to Fibre Channel flow control, improving network performance by lowering latency and increasing throughput.

FC flow control is a link-level buffer credit strategy that is implemented to prevent frame drops. The reception of a buffer credit from the opposite end indicates enough buffer space to accept an FC frame. When the receiver stops releasing credits to the sender, the sender stops sending. This mechanism ensures that the receiver has adequate buffer space to accept sent data, thereby preventing frame drops.

Was anything wrong with Brocade's flow control? In short, no. Was there room for optimization? Yes. FC is designed to create virtually lossless networks. Typically, mission-critical applications leverage a SAN because enterprise storage demands an ultrahigh bandwidth, ultrahigh reliability, ultralow latency, lossless network for robust operation. FC flow control meets these design goals, whereas Ethernet networks do not. This paradigm is why FC networks are ubiquitous across the world's data centers and continue to be steadfastly deployed.

The slowest destination rate becomes the rate of all flows when a common flow control is used. The rate at which a particular destination port can consume incoming data varies. Flow-controlled networks benefit from being lossless; however, this comes with an inherent challenge, primarily head-of-line-blocking (HoLB). Efficient data transmission becomes impeded when different flow rates share the same flow control, causing flows to operate at the slowest flow rate. Credits are not withheld per se; more so, they are merely returned at the rate of the slowest device. For instance, data flows through an ISL or Brocade Trunk (BT) have various destinations, either to hosts (reads) or to storage (writes). An entire link is affected if a common flow control applies to all flows. Flow-control problems are exacerbated as fabrics scale by an increased number of end devices and/or hops. This behavior is normal FC "flow control" and is characteristic of a lossless network. The network needs to be lossless. The flow control challenge is addressed through advanced Traffic Optimizer technology by Brocade.

Previously, the impact domain was reduced by assigning traffic flows to specific VCs. Binding flows to dedicated VCs limited congestion to only those flows mapped to a congested VC. Which flow was mapped to which VC was pseudorandom. Now, Traffic Optimization reduces the risk of an impact domain ever experiencing congestion since flows are intelligently grouped to reduce that probability. Flows are no longer pseudorandomly assigned to VCs; they are assigned by performance groups.

Imagine a 5-lane expressway. Traffic is not optimized on this highway. Let's say trucks travel slower than cars. Trucks and cars are randomly dispersed across all lanes, resulting in the slowing of all lanes. Effectively, no lane can go faster than the slowest vehicles, which are the trucks. This example is analogous to a pre-Brocade Gen 7 SAN in which traffic was pseudorandomly assigned to VCs, not based on any performance criteria.

Traffic optimization occurs when vehicles are assigned to specific lanes based on their speed. A Brocade Gen 7 SAN assigns virtual channels according to performance groups (PGs), which are traffic classifications based on destination port speed. Traffic destined to ports of the same speed drains from the SAN at equal rates, and flows of the same speed do not impede each other. Performance groups make for optimal performance.

Brocade Virtual Channels

Brocade VCs have addressed head-of-line-blocking with great success, demonstrating significant performance gains over many generations of FC. Groundbreaking at the time, as port speeds increased, VCs became a partial solution calling for further optimization. Brocade's VC implementation has distinct egress queues that establish virtual channels across a physical link, each channel having its own set of autonomous buffer credits and flow control. The link bandwidth is shared, but flows are entirely independent. When traffic performance characteristics such as destination port speed are uniform, VCs offer an expedited pathway. A more significant benefit is gained by sorting traffic into VCs based on the most pertinent performance characteristic, which destination speed is.

A VC services multiple flows since there is a limited number of VCs per traffic classification. Users do not configure or set up VCs; rather, they are a fundamental component of the Fabric Operating System (FOS). For five generations of FC, Brocade VCs operated within a relatively narrow range of port speeds (1, 2, 4, 8, and 16: 8–15Gb/s spread) by leveraging multiple paths with independent flow control. VCs were assigned pseudorandomly by a hash; destination port speed was not a factor. FC wire speeds have continued to increase, and Gen 7 is now 64Gb/s with a vast port speed range (8, 16, 32, and 64: 32–56Gb/s spread). Wide variations in flow speeds impose flow control problems across the fabric unless those variations in speed are accounted for.

Brocade® Traffic Optimizer technology takes VCs to the next level. Traffic Optimizer organizes and manages traffic flows according to their PG, and fabric resources are allocated per PG. Flows are assigned to a PG based on the destination port speed, and in turn, PGs are bound to specific dedicated VCs. Brocade fabrics know the destination port speed for every flow. Other classification criteria could potentially be used in future releases.

Gen 7 interoperability with existing Gen 6 fabrics is fully supported. ISL connections between Gen 7 and pre-Gen 7 platforms result in Traffic Optimizer VCs being remapped using the legacy classification.

Summary

Traffic Optimizer is a new and unique feature in Brocade Gen 7 SAN products. Traffic Optimizer adds performance groups (PGs), a powerful enhancement to virtual channels. Traffic Optimizer requires no configuration, is enabled on all Gen 7 devices, and permits connectivity with Gen 6 platforms. Performance groups classify and separate traffic into like performance characteristics and assign that traffic to specific VCs. Traffic Optimizer prevents slower destinations from impeding faster ones. Traffic Optimizer is interoperable with other FC technologies (FICON, Extension, Access Gateway, long-distance links, security, and NPV) and with Brocade Gen 6 fabrics.

BROCADE
A Broadcom Company