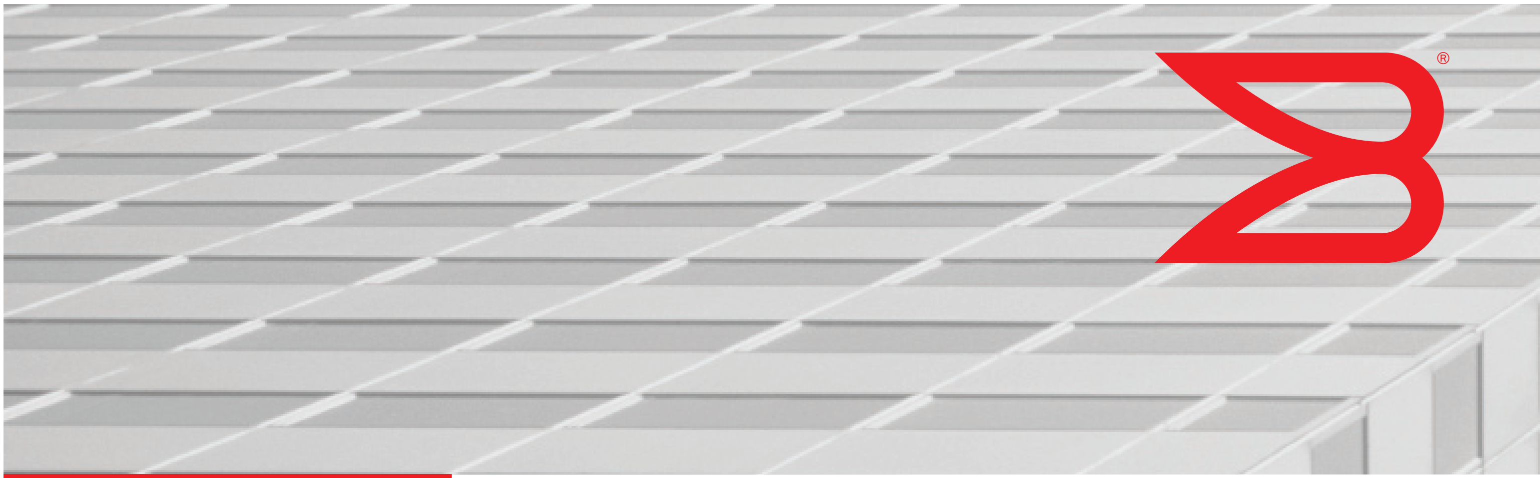

IP NETWORK

Multi-Chassis Trunking for Resilient and High-Performance Network Architectures

Multi-Chassis Trunking is a key Brocade technology in the Brocade One architecture that helps organizations build scalable and resilient network infrastructures to support business needs today and scale to meet future needs.

BROCADE

EVOLVING NETWORK DEMANDS

We are living in an information age where the interaction among organizations, their employees, and their customers is continuously evolving. Organizations are communicating with their customers using tools such as eCommerce, e-mail, YouTube, and video conferencing-- making a reliable network and high-performance infrastructure key business enablers. Disruption to this key infrastructure leads to lost productivity and in turn, loss of customer confidence and revenue. Furthermore, today's organizations demand flexible network infrastructures capable of handling ever-increasing data traffic and optimized for virtualization technologies to deliver a wide range of on-demand services.

The Brocade® One™ architecture is designed to help these organizations transition smoothly to a world in which information and applications can reside anywhere. Multi-Chassis Trunking (MCT) is a key Brocade technology enabling this network architecture. It helps organizations build scalable and resilient network infrastructures for tomorrow, so that they can continue to leverage their existing assets and still establish new standards for uptime with the introduction of future-proof technologies.

Layer 2 networks are ubiquitous in Metropolitan Area Networks (MANs), in the data center, and in enterprise network infrastructures. These networks are dependent on loop prevention protocols such as Spanning Tree Protocol (STP) and variations. However STP protocols suffer from many drawbacks, the most significant being slow convergence and lack of load balancing on redundant paths. They operate by blocking ports and create unused capacity in the network.

Network architects have overcome slow convergence using technolog.ies such as Virtual Switch Redundancy Protocol (VSRP), Metro Ring Protocol (MRP), and Ethernet Ring Protection (ITU G.8032), all of which offer sub-second convergence. However, they works on the same link-blocking technology and do not provide the desired active-active load balancing. The onus of efficient network operation lies with the network architect and advance planning using load balancing on a per-VLAN basis. Further, the network operator needs to continuously monitor the network to rebalance.

IEEE 802.1AX Link Aggregation (LAG) technology has solved this using multipathing at Layer 2 and flow-based load balancing. However, the protocol constrains the network to a node-to-node topology. Organizations require a Layer 2 multipath solution that can provide dynamic flow-based load balancing to multiple network nodes. Brocade MCT is designed to address these requirements for today's resilient and high-performance networks.

MCT OVERVIEW

Multi-Chassis Trunking is technology that allows multiple switches to appear as single logical switch connecting to another switch using a standard LAG. Since the technology is an enhancement to the standard LAG protocol, a single MCT-unaware server or switch using a standard LAG trunk can connect to two MCT-aware switches--and the traffic is dynamically load balanced.

MCT inherits the all benefits of LAG by providing multiple physical links to act as a single logical link; the new bandwidth available is an aggregate of all the links in the group. The traffic is shared across the links in the group using dynamic flow-based load balancing and traffic is moved to a remaining link group in tens of milliseconds in the event of a failure in one of the links. While standard LAG provides link- and module-level protection, MCT adds node-level protection, while maintaining failover times better than sub-200 milliseconds for uplink failure. Further, it works with existing switches and servers that connect to MCT and does not require any rip-and-replace of existing infrastructure.

MCT is available on all Brocade NetIron® XMR, CER, and CES and Brocade MLX and MLXe Series platforms.

MCT OPERATION

MCT is designed to achieve the desired active-active topology and efficient Layer 2 multipathing, while ensuring that the network scales effectively. Hence the load on the Inter-Chassis Link (ICL), the link connecting the MCT peers, is kept to a minimum and the protocol relies on direct forwarding of traffic to the client wherever possible. To understand how MCT functions, the following sections explain the key elements of MCT.

MCT Elements

- **MCT peers.** A pair of MCT nodes that appear as a single logical node. A LAG from a server or a switch can be spread across the MCT peers and they will appear as a single logical endpoint.
- **MCT client.** A switch or server connected to the MCT peers using a LAG. The client device is only required to support IEEE 802.1AX LAG; it could be static or dynamic using Link Aggregation Control Protocol (LACP). In a multi-tier topology, the client could be a pair of MCT devices.
- **Inter-Chassis Link (ICL).** A physical link or a LAG group connecting the two MCT peers. ICL is an important link and it is recommended as a LAG group for protection against link or module failure.
- **Keep-alive VLAN.** An optional VLAN for continuity check messages and failure handling in the event of ICL failure.

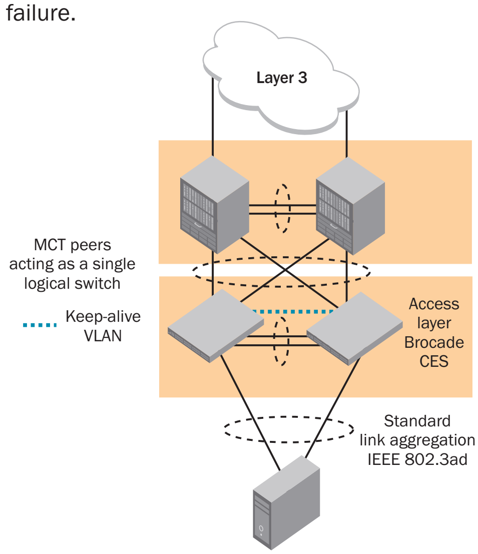

Figure 1.
Typical MCT topology.

Glossary of MCT Terms

- Cluster Communication Protocol (CCP): A reliable protocol that maintains state and MAC table synchronization between the MCT peers
- Inter Chassis Link (ICL): Physical link or LAG group connecting the MCT peers
- MCT client: A switch or server connected to MCT peers using a LAG
- MCT peers: Pair of MCT nodes that appear as a single logical node
- Port Loop Detection (PLD): A protocol to detect and repair loops in a Layer 2 networks
- RBridge ID: A unique ID associated with each bridge that includes MCT peers or clients

MCT functionality can be divided into to two main parts: LAG operation and Cluster Communication Protocol (CCP). MCT clients perform only the LAG operations defined in IEEE 802.1AX. The LAG can be a static or a dynamic LACP trunk. Cluster Communication Protocol is a reliable protocol that runs between the MCT peers over the ICL, and it maintains state and MAC table synchronization between the two peer nodes.

The network design and configuration is fairly simple. Each MCT peer and client is assigned a unique ID, known as the RBridge ID. The MCT peer nodes are connected using an ICL and configured as a cluster. Once the peering relationship is established between the two nodes, CCP runs over them to make them appear as a single logical node for LAG purposes. Adding a client is Justas straightforward as associating client RBridge ID with the connected port on the MCT peers. The CCP protocol manages the rest and keeps the network operation simple.

Traffic from the client is load balanced using a hashing algorithm over the LAG ports. The MCT nodes forward the traffic to the destination directly. The CCP ensures that the MAC table between the two nodes is always synchronized, hence in steady state, and ICL use is kept to a minimum to limit overhead. Downstream traffic on MCT nodes is also directly sent to the client.

LAYER 3 RESILIENCY

A simple MCT topology addresses resiliency and efficient multipathing in Layer 2 network topologies. To interface with a Layer 3 network, MCT is configured with Virtual Router Redundancy Protocol (VRRP) to add redundancy in Layer 3. The standard VRRP mode is active-standby and all traffic is forwarded through the master. The Brocade NetIron platforms also support VRRP-E server virtualization (see Figure 2), in which multiple VRRP standby nodes are supported and each node is capable of forwarding to an upstream Layer 3 network. This provides efficient deployment for both Layer 2 and Layer 3 networks.

Metro Ring Protocol (MRP) was designed to create a scalable, Layer 2, loop-free ring topology and is a faster alternative to spanning tree protocols. When used with MCT, the MCT peers act as a logical MCT pair and participate in the MRP ring. The client nodes can then use standard LAG and achieve active-active dual-homing to the MRP ring. This provides greater flexibility in network design and helps connect MCT clusters. Typical use cases are dual-homing in metro networks, as shown in Figure 3, and joining MCT clusters in medium to large data centers, shown in Figure 4.

FAILURE HANDLING

A link or node failure in the MCT network is handled in the same way as a failure in a standard LAG. In the event of a failure in any of the client links, the client instantly rebalances the upstream traffic over the rest of the links in a LAG group. Most clients use a hashing mechanism based on the packet header to ensure that each flow is maintained on a single link and load balancing occurs on a flow basis.

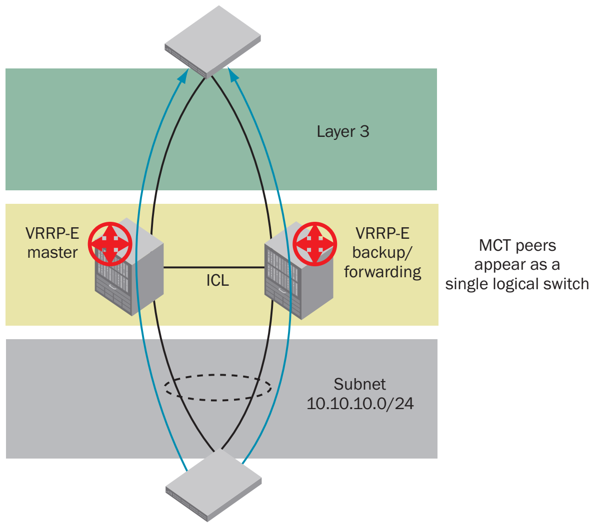

Figure 2.
MCT with VRRP for Layer 3 resiliency.

An MCT node with a failed link communicates LAG state change to the peer using CCP. Any downstream traffic to the client from this node will use the remaining links in the LAG group. If no alternate client links are available on this MCT node, traffic will be transmitted over the ICL and use the downlinks on the peer node to ensure traffic delivery. Downstream traffic arriving on the peer node will be sent directly to the client, just as it would have without the failure.

The Brocade devices also support Port Loop Detection (PLD), a protocol that can be run along

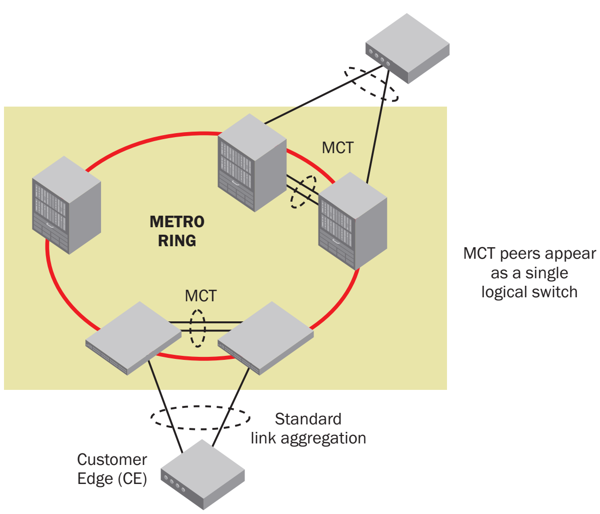

Figure 3.
MCT with VRRP for Layer 3 resiliency.

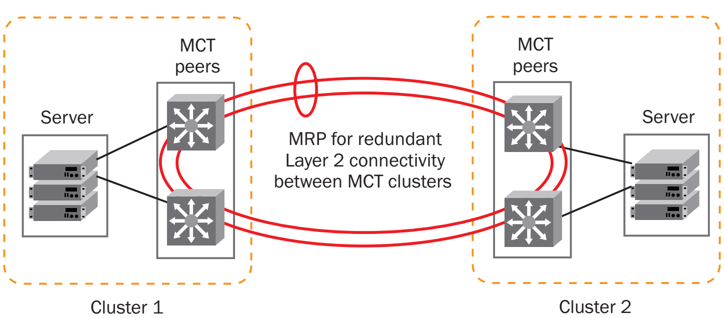

Figure 4.
Connected MCT clusters in the data center.

with other loop prevention protocols to quickly detect and repair loops. Loops are typically introduced due to network misconfiguration or an errant device in the network. ICL links are usually configured as a LAG group, which can scale up to 64 x 10 Gigabit Ethernet (GbE) to provide resiliency and design flexibility. In addition, ICL links are protected with a backup keep-alive VLAN. This helps the network to continue to function in the unlikely event of multiple failures, including an ICL failure.

BENEFITS

MCT helps create a network topology that is easy to configure and manage by eliminating the need for STP from server to core in a data center. MCT offers five times the improvement over alternate approaches by providing fast link and node failover times. Stranded network capacity can be recovered by migrating to MCT and overall network utilization can be improved by as high as 100 percent. All links from server to core are forwarding traffic and Brocade's patent pending hashing algorithms enable efficient link utilization. Further, with this approach, the network can can appear as a single logical device with over 30 Terabits per second (Tbps) of switching capacity.

MCT coupled with VRRP-E server virtualization extends these benefits to an upstream Layer 3 network and preserves efficient network utilization even when Virtual Machines (VMs) and VM mobility are deployed. MCT helps provide investment protection by unlocking the available bandwidth and seamlessly interoperating with third-party client devices.

In a metro network, MCT allows active-active dual-homing to a pair of access switches, where the customer edge device needs to support a standard LAG. This can help a service provider offer more resilient and/or high-bandwidth services than are possible using current technologies. A majority of metro networks are ring topologies. And MCT can be used in conjunction with Metro Ring Protocol to offer a resilient and a more efficient network.

SUMMARY

Multi-Chassis Trunking helps organizations build resilient and high-performance network architectures and allows them to transition smoothly to a world in which information and applications can reside anywhere. It provides service providers the flexibility to offer new and differentiated services. The resulting network architectures are fast and reliable and can scale to meet the needs of customers or applications. A focus on multipathing and efficient network usage helps contain capital budgets and because it does not use STP, operating costs can be kept low. Availability of MCT on the Brocade NetIron XMR, CES, and CER and Brocade MLX and MLXe Series products offers organizations the flexibility to use the right product for an their specific needs

ABOUT BROCADE

Brocade provides innovative, end-to-end network solutions that help the world's leading organizations transition smoothly to a virtualized world where applications and information can reside anywhere. These solutions deliver the unique capabilities for a more flexible IT infrastructure with unmatched simplicity, non-stop networking, optimized applications, and investment protection. As a result, organizations in a wide range of industries can achieve their most critical business objectives with greater simplicity and a faster return on investment.

For more information about Brocade products and solutions, visit www.brocade.com.

Corporate Headquarters
San Jose, CA USA
T: +1-408-333-8000
info@brocade.com

European Headquarters
Geneva, Switzerland
T: +41-22-799-56-40
emea-info@brocade.co

Asia Pacific Headquarters
Singapore
T: +65-6538-4700
apac-info@brocade.com

BROCADE