PowerEdge Product Group

Direct from Development

PowerEdge MX7000 Chassis Management Networking Cabling

Tech Note by:

Brian Claggett

Michael E. Brown

SUMMARY

The purpose of this whitepaper is to describe the new and improved features of the MX7000 chassis management network, provide recommended network cabling diagrams, and show the fault tolerant capabilities of the network design in the Dell EMC MX7000 chassis.

The MX7000 chassis features dual redundant management modules, with each management module featuring two management network ports, for a total of 4 management network ports on the chassis. The management network is meant to provide network connections for chassis management separate from the customer data network. There are several new and improved design points for this network compared to previous generation chassis that significantly impact how these new chassis should be cabled and managed.

MX7000 Chassis Overview

The MX7000 chassis enclosure contains server sleds, IO Management Modules (IOMs), power supplies and Management Modules (MMs). Each Management Module provides two ethernet ports for access to the management interfaces of the components in the chassis. This document describes the network cabling best practices to connect the management modules to a management network, as well as best practices for cabling together multiple chassis into stacks. When chassis are cabled together following best practices, they achieve the best resilience against various faults.

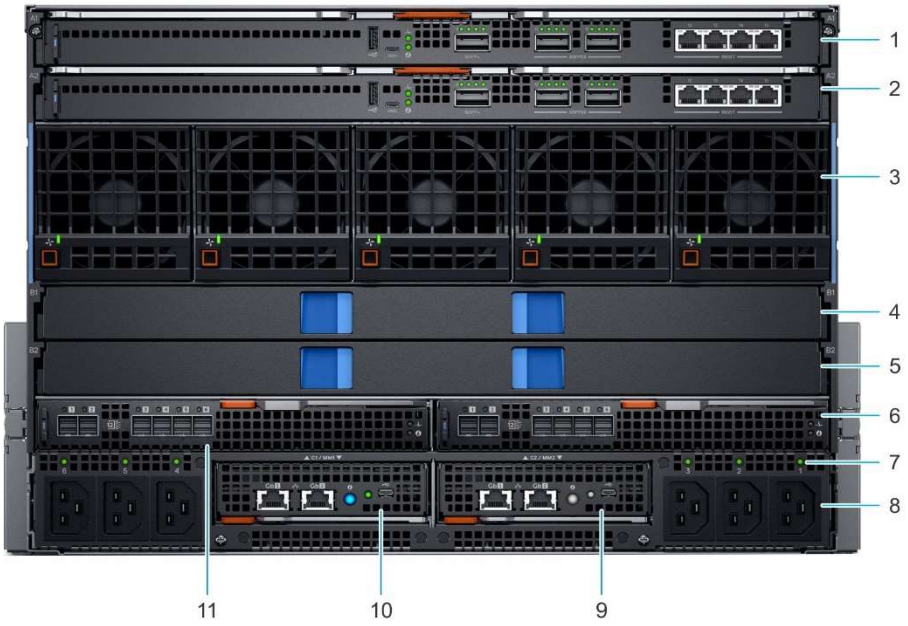

Figure 1 – Back View of the MX7000 enclosure

1. Slot for Fabric A1
2. Slot for Fabric A2
3. Rear fans (5)
4. Slot for Fabric B1
5. Slot for Fabric B2
6. Slot for Fabric C2
7. Power status LED
8. C22 Power inlet (6)
9. Management Module 2
10. Management Module 1
11. Slot for Fabric C1

Cabling an Individual Chassis

When cabling an individual chassis, connect one network cable from each management module to the data center top of rack switch. Ensure that both ports on the top of rack switch are enabled and on the same network and VLAN.

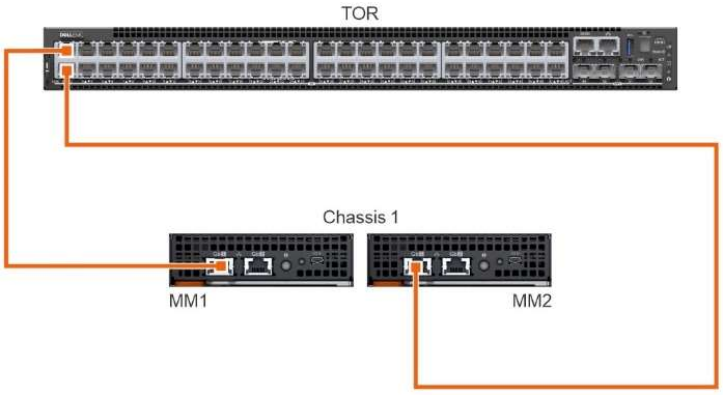

Figure 2 – Single chassis connection diagram

This configuration provides redundant network access to the chassis, protecting against link failure or management module failure. You can cable the redundant connections to the same top of rack switch or to two redundant switches, just be sure to configure both connections the same and have them on the same network and VLAN. There is no requirement to prefer either Gb 1 or Gb 2 for these connections.

Cabling Multiple Chassis

With the automatic uplink detection and network loop prevention, you can connect multiple chassis together and use fewer data center network ports. When multiple chassis are cabled in this manner it is known as a “stack”. When cabled correctly, you can maintain access to all chassis in the stack, even in the face of faults such as bad cables, bad management modules, or unplugged/unpowered chassis.

Note: While the automatic uplink detect and network loop prevention allow for chassis to be interconnected without network issues, stacked chassis are not automatically part of a management group.

See the MX7000 Multi Chassis Management feature for group management. This feature (not covered in this whitepaper) lets you manage multiple chassis through one management module and requires the network stacking arrangement described here.

Recommended Topologies for Multiple Chassis

While the auto sensing loop prevention algorithms allow for virtually any cabling arrangement and still provide access to all components in a chassis stack, there are optimum cabling topologies. These topologies provide redundant connections into the stack and continue to provide connectivity to every stack member even in the face things like firmware updates, management module failure, or cabling faults. The cabling diagrams shown will protect against any single point of failure in the entire stack.

(continued next page)

With this configuration the group management will survive the following failures and retain full management network access to all nodes in the stack:

- Any single network cable failure
- Any single management module failure
- Power loss to any single chassis in the stack

It is important to note that these are the minimum failures that group management was designed to handle. It is possible to have more failures than the above and still have full functionality, but it would depend heavily on where those failures occurred. Also note in diagrams below: we show only the redundant management modules and not the entire chassis to reduce visual clutter of the diagrams. Each set of two side-by-side management modules represents one individual chassis.

Two Chassis Stack Example

For consistent access to all chassis and their components from a management network, it is important that both ports connected on the top of rack switch are connected to the same layer 3 network. A very common misconfiguration that has been seen from customers is accidentally connecting the two redundant uplink ports to different VLANs. If the chassis are connected to different VLANS, when the primary link fails, access to the management network will appear to be lost, as the redundant connection is brought up with the same IP configuration on a different layer 3 network.

In the 2 chassis stack, each chassis has a connection to the top of rack switch as well as redundant connections to the other chassis. The third connection between chassis is not technically needed, but for consistency as this stack is expanded to more chassis, it is illustrated and recommended so that all the various configurations are uniform.

It is very important to note for all configurations that only a single connection to the top of rack is ever active at one time. The multiple connections do not increase management bandwidth out of the stack, but rather they provide redundancy in case of a link failure, management module failure, or unpowered chassis. No customer management network traffic will ever flow through the chassis stack unless it is going to a device in that stack.

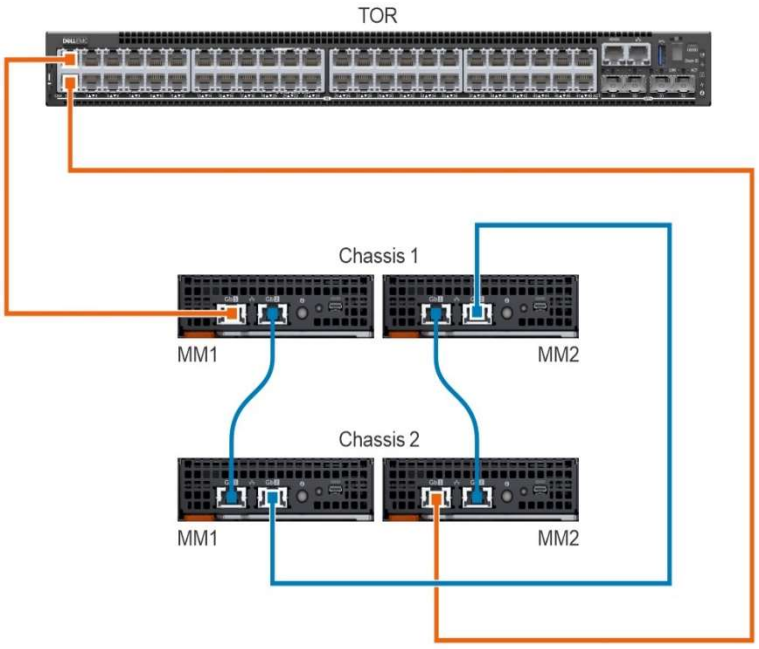

Figure 3 – Two chassis stack connection diagram

Ten Chassis Stack Cabling Example

The MX7000 chassis allows up to twenty chassis to be stacked and managed as a group. When stacking more than two chassis, the chassis that are not directly connected to the top of rack switch are connected only to other neighbor chassis in a redundant fashion. The two chassis that are connected to the top of rack switch are connected to each other to provide a redundant connected path. This redundant path is important to allow the stack to survive power loss to a chassis in the middle of the stack.

Additional chassis may be added to the stack by following the connections as shown below for those chassis that are not connected the top of rack switch.

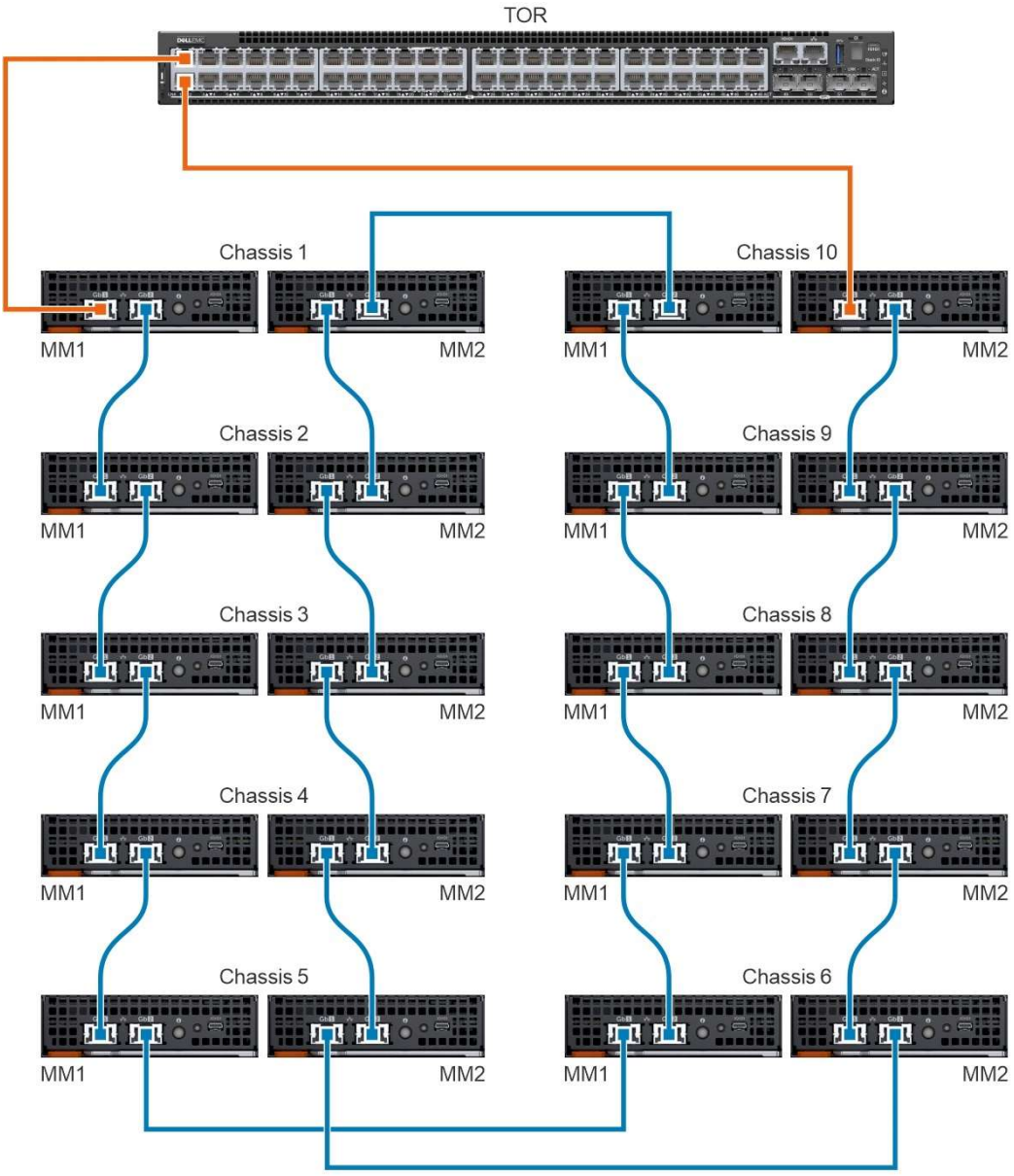

Figure 4 – Ten chassis stack connection diagram

Automatic Uplink Detection

On each management module, there are two RJ-45 ethernet connections labeled Gb1 and Gb2. When connected to a top of rack switch/router, these connections provide access to the chassis internal management network. It is not necessary to connect both Gb1 and Gb2 of a single management module when a chassis is not stacked with other MX7000s, however it is strongly recommended that each management module have its own top of rack connection. These connections will automatically detect when connected to a customer management network or when connected to another MX7000 chassis. When you connect two MX7000 chassis directly with one another, each chassis automatically senses a chassis to chassis connection and moves that link to stacking mode. There is no requirement to prefer either Gb1 or Gb2 port for stacking operations or for uplinks to top of rack. Each port has identical functionality and all connections are automatically detected and configured.

Network Loop Prevention

When cabling chassis in a stack, you will notice what appear to be connections forming network loops and redundant connections between the chassis. The management modules will automatically detect redundant connections between chassis and will use redundant links to get around failures as they are detected. The redundant connections will not cause improper network operation as only one link of a redundant set is selected to use and the other redundant connections are placed in a “blocking” state where they are ready to use at a moment’s notice when the management module detects a failure.

There are a couple points worth noting about this feature. First, when you stack chassis and cable them together, only one uplink will ever forward traffic out of the stack. The recommended cabling configuration has two links from chassis to the top of rack switch. If you have the recommended redundant connections out of the stack, as above, only one link is selected to carry traffic and the other ports are placed in a “blocking” state, ready to carry traffic as soon as a failure is detected in the primary links. When looking at switch statistics, you will see all links in a “Link Up” state, but only one will have traffic activity. If a link to the top of rack fails, if the management module with that link fails, or the chassis loses all power, the management module with the redundant network connection will take over.

Internal Spanning Tree

Each redundant management module has the concept of “Active” and “Standby” to denote that only one of the management modules handles chassis management at a time. It is important to understand that the selection of which network uplinks or stacked chassis links to use is completely independent of which module is “Active”. The system can and will route traffic over the network link on a “standby” module.

Internally, the networking algorithm uses Rapid Spanning Tree Protocol to determine a network path to every component in the management network stack. This may include forwarding paths through “Passive” MM modules. It is important to note that Spanning Tree is *not* enabled on any uplink ports to top of rack switches, only on links directly connected between chassis. Rapid Spanning Tree is automatically configured between MX7000 chassis and is not ever configured on links that are to customer network equipment (i.e. top of rack or other datacenter switches).

One important thing to be aware of are network disruptions caused by cabling or link changes in the stack. When links change, for example if a cable is pulled or added, the network may go down for approximately 30 seconds to re-compute the network topology. This affects all chassis in the stack. One thing you will notice is that every time cabling or link state changes in the stack, all servers, ioms and management modules in the stack will reconfigure their network. If the components are configured for DHCP, you will see them refresh their IP addresses. These should be rare events that do not disconnect active network sessions, though you will see delays as packets are retried until the network finishes reconfiguration.

Reserved VLANS

The MX7000 chassis internally use VLANs in the series 4000 through 4020 for communications inside the chassis and between stacked chassis. This series is reserved by the management module and chassis components and cannot be used for access through the uplink connection. The management module will filter out any traffic on these VLANs and will not allow it to enter the chassis or a chassis stack.

All other possible VLANs are available for use and VLAN tagged packets will forward through all stacked chassis.

MX7000 External Port Configuration

Each management module in the MX7000 chassis has two RJ-45 Ethernet ports labeled “Gb 1” and “Gb 2”. Each port is configured to be Auto MDI-X, which eliminates the need for an Ethernet crossover cable when connecting any MX7000 port to another chassis or top of rack switch. The external ports are set to full auto-negotiate and are capable of 1000 Mbps. When connecting a stack to a top of rack switch, it is recommended that the ports should be set to the same configuration. While the MX7000 chassis will link with ports that are configured for 100 or 10 Mbps, network speeds will have an impact when accessing the group management features or the management interfaces for the components contained within the chassis.