High Availability in the Force10 Networks E-Series

Overview

Built for reliability and high availability, the E-Series from Force10 Networks provides system control plane, manageability, serviceability and link level features.

Force10 Innovation in High Availability Design

While evolving to meet the needs of telecommunications carriers, each generation of switch and router platforms has benefited from successive improvements in hardware and software availability. Force10 has taken high availability to the next level in the design of the E-Series to keep pace with the exponentially increasing volumes of mission-critical traffic in today’s networks.

Within the E-Series architecture careful attention has been paid to the reliability and availability of the ASICs, hardware and software. The interaction between these three design areas demanded particular attention due to the E-Series’ extreme scalability and dual switch and router functionality. The results of this intense focus were multiple advances in high availability design. These include innovations such as the patented 5 Tbps passive copper backplane, the multiprocessor-based route processor module (RPM), and ASIC-based filtering and rate limiting of CPU-bound control traffic.

These capabilities – combined with manageability features, redundancy of critical components, alarms and fault reporting, ECC protected memory, and process isolation – deliver the high availability required in today’s business critical network infrastructure. Force10 high availability features include:

• System Design for High Availability
  – Passive copper backplane
  – Redundancy of critical components
  – Reliable power system design
  – Environmental monitoring

• Control Plane Features for High Availability
  – Multiprocessor control plane
  – Control plane availability
  – Route processor module failover
  – Modular FTOS

• Manageability and Serviceability Features
  – Online insertion/removal (OIR)
  – Line card persistent configuration and pre-configuration
  – Maintenance and serviceability

• Link Availability Features
  – Link aggregation (LAG)
  – Equal cost multi-path routing (ECMP)

System Design for High Availability

Non-Blocking 1.68 Tbps Switch Fabric Passive Copper Backplane

The E-Series backplane is the industry’s first high speed non-optical backplane to achieve 1.68 Tbps in a single-rack switch/router chassis. Unlike optical backplane interconnect systems or active copper backplanes, the patented E-Series backplane has no single points of failure and eliminates costly electrical-optical-electrical conversions. The resulting system simplicity means bulletproof reliability and minimum cost.

Redundancy of Critical Components

The E-Series eliminates single points of failure by providing redundancy and OIR for all critical components.

• 1+1 redundant route processor modules (RPMs) (active/standby)
• Redundant switch fabric modules (SFMs)
• Redundant DC power entry modules (PEMs)
• Redundant AC power supply modules
• Redundant fan subsystems

In case of a component failure (or removal), the redundant component is brought online immediately by the Force10 Operating System (FTOS™) software. The failure can be logged in four user-configurable ways: via console trap, via SNMP, via syslog and via the front panel major/minor alarm indicators and relay contacts.

Reliable Power System Design

The E-Series chassis support power supply redundancy with load sharing and OIR support. During normal operation the E-Series will draw power from all working
power inputs. If any single DC PEM or AC PSM fails or is removed, the hardware will switch to the remaining supplies without interruption of service.

Every card in the E-Series incorporates independent power conversion modules to eliminate the possibility that a converter failure could affect other systems beyond the card it is on. Failure of any power module is detected by the FTOS software and reported by any of the four reporting methods mentioned above.

**Environmental Monitoring**

The FTOS software continually monitors environmental conditions within the chassis, including temperature and voltage levels on each card in the system. If temperature levels increase beyond user configurable thresholds, FTOS takes action as follows:

<table>
<thead>
<tr>
<th>Condition</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>All temperatures</td>
<td>Fan speed increased to maintain temperature below minor alarm threshold</td>
</tr>
<tr>
<td>Above minor alarm threshold</td>
<td>Minor (yellow) alarm reported via RPM indicators and alarm contacts, and console/syslog/SNMP messages</td>
</tr>
<tr>
<td>Above major alarm threshold</td>
<td>Major (red) alarm reported via RPM indicators and alarm contacts, and console/syslog/SNMP messages</td>
</tr>
<tr>
<td>Above maximum 70° C threshold</td>
<td>Major alarm reported and affected component(s) powered down</td>
</tr>
</tbody>
</table>

Similarly, if power levels on any card fall below nominal operating ranges, FTOS reports the condition as a minor alarm. If power levels drop below the minimum operating voltage, FTOS powers off the affected component and issues a major alarm.

**Control Plane Features for High Availability**

**Multiprocessor Control Plane**

With three processors on each RPM and one processor on every line card, the E-Series’ system control plane is purpose-built to deliver high performance and fault tolerance to the full FTOS suite of L2 switching features and L3 routing protocols. Software processes are distributed among the processors, creating a true real-time multiprocessing environment that enables process isolation with memory protection. These features are absolutely necessary for fault tolerance and rapid convergence in large-scale enterprise and service provider networks.

**Modular FTOS**

FTOS is based on NetBSD, a highly portable, open source OS optimized for networking applications. FTOS’ modular design and process isolation deliver many key reliability and scalability benefits, including allowing the system to offload processes, such as sFlow or BFD, to line cards.

The modular structure of FTOS also eliminates many potential catastrophic system failures by restricting faults to specific processes. Because each process is modular and runs in its own protected memory space, a fault in one module cannot affect any other modules. As a result, software problems are isolated to specific processes and the rest of the system continues to operate.

In addition, the modular FTOS supports hitless forwarding in addition to rapid convergence. Should any redundant component fail – whether a line card, switch fabric, route processing module (RPM) or power supply – FTOS ensures that all packets are forwarded, preventing application disruption.

**Control Plane Availability**

The hardware control plane includes ECC-protected memory systems and a separate control communications path from the RPM to each line card. ECC protection maintains the integrity of the memory systems in the event of single-bit errors. The E-Series separates user traffic and control traffic onto different paths. This ensures that heavily congested user traffic cannot slow configuration and forwarding table updates to the line cards. In addition, dedicated 100 Mbps switched paths – from the RPMs to every line card – eliminate sluggish forwarding table updates – events that could jeopardize network stability.

**Control Traffic Rate Limiting**

The E-Series’ RPMs provide innovative control traffic rate limiting and filtering capability. This feature empowers network administrators to suppress harmful denial of service (DoS) attacks and prevent unwanted traffic from flooding the network and unnecessarily burdening control processors. Independent limits for different control packet types give preferential treatment to routing updates, for example, and limit incidental control traffic such as ICMP echo request.
**Route Processor Module Failover**

RPM failover capabilities in the FTOS software are supported to enable hitless forwarding without traffic interruption and are built on several high availability mechanisms.

**Fast Failover with Configuration Synchronization**
The automatic synchronization of configuration information between redundant RPMs minimizes recovery time in the case of an RPM failure. FTOS supports two levels of synchronization: full synchronize and persistent synchronize. In persistent synchronize mode, the startup configuration is copied from the primary to the secondary every time a user saves the running configuration. In full synchronize mode, the system will synchronize the startup configuration, error messages and task manager status. Synchronizing the task manager minimizes the time needed to bring up a line card after the system fails over to the standby RPM. Synchronization gives the standby RPM knowledge of what line cards are in the system and the tasks that are running on each one, and enables features such as hitless LACP.

**Hitless Forwarding**
Hitless forwarding for systems with dual RPMs ensures that the system continues to forward traffic in case of an RPM failure. The line cards maintain state during an RPM failure, allowing Layer 2 and Layer 3 traffic to forward without interruption. Hitless forwarding for OSPF, BGP and PIM through graceful restart mechanisms are also supported. System state is maintained for Layer 2 protocols such as STP and LACP, enabling them to continue running during an RPM failover.

**Hitless Software Upgrades**
Hitless software upgrades enable users to load a new version of the FTOS software on the standby RPM, and then manually failover the system from the primary to the standby RPM without disrupting traffic forwarding. Hitless software upgrades are an extension of hitless forwarding capabilities, and use these during a software upgrade.

**Redundancy Protocols**
The E-Series supports several chassis protocols to provide link level redundancy and rapid convergence. Bidirectional forwarding detection (BFD) can be used at the link layer to enable rapid failure detection that is much faster than protocol timeouts, providing sub-second failure notification. The Force10 resilient ring protocol (FRRP) can be used to deploy a self-healing Layer 2 ring topology that converges faster than STP, and the virtual router redundancy protocol (VRRP) can be used to provide Layer 3 redundancy.

**Manageability and Serviceability Features**

**Online Insertion/Removal (OIR)**
The E-Series supports OIR of all critical components to minimize system downtime for repairs and scheduled maintenance.

- RPMs
- SFMs
- DC PEMs
- AC PSMs
- Fan trays
- Line cards

**Line Card Persistent Configuration and Pre-Configuration**
To reduce system downtime during the replacement of line cards, Force10’s FTOS software supports persistent configuration and pre-configuration of line card slots. In the case where a line card is removed, FTOS stores the line card type, MAC address assignments and configuration information. FTOS will also automatically reconfigure a replacement card once it is inserted in the slot.

Pre-configuration of line card slots minimizes system down time by enabling the system administrator to configure an empty slot as if a line card were present. Pre-configuration eliminates the need to enter configuration commands after the line card is up and running. As a result, it dramatically reduces the time needed to add and provision new ports in the chassis.

In both cases, FTOS can also be set to report an error if an incompatible card type is inserted into the slot. As an example, such an error would occur if a 48-port GbE card was pulled and a 4-port 10 GbE card inserted in its place.
Maintenance and Serviceability

The E-Series simplifies routine maintenance and management with a number of features that reduce the duration of scheduled and non-scheduled maintenance:

• Front side access to all cabling
• Integrated cable management
• Removable air filter
• Redundant 10/100 management ports
• Alarm reporting to syslog, SNMP, console and RPM alarm LEDs
• Inline system health checks that monitor control plane and data plane availability

Link Availability Features

Link Aggregation (LAG)

By using LAG in the E-Series, Force10 enables customers to put up to 16 Ethernet links in one logical group and balance traffic across every link. If any link in the group fails, LAG automatically distributes all traffic to the remaining links within 100 ms. LAG is an active/active redundancy scheme since it uses all functioning links to send traffic.

Equal Cost Multi-path Routing (ECMP)

By using ECMP in the E-Series, Force10 enables customers to put up to 16 IP links in one logical group and balance traffic across every link. If any link in the group fails, ECMP automatically distributes all traffic to the remaining links. ECMP is an active/active redundancy scheme since it uses all functioning links to send traffic.