



FCoE Addressing & High Availability

T11/07-547v2, December 2007
JR Rivers



Introduction

SAN and LAN redundancy schemes are different

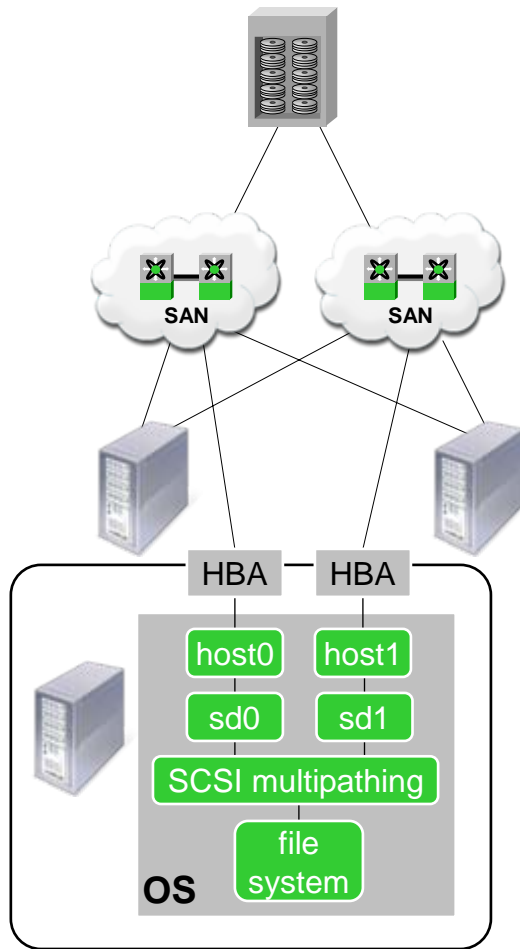
- SAN uses redundant physical networks
- LAN uses single logical network with redundant physical links/devices

FCoE has large return on investment when deployed at the server access layer

- Reduces adapters, switches, and cabling expenses

FCoE must fit into existing network and OS structures

SAN Redundancy Overview



Distinct networks (fabric) based on redundant components

- SAN A/B

Redundancy in each network is based on Shortest Path First routing

- Spanning tree for multicast

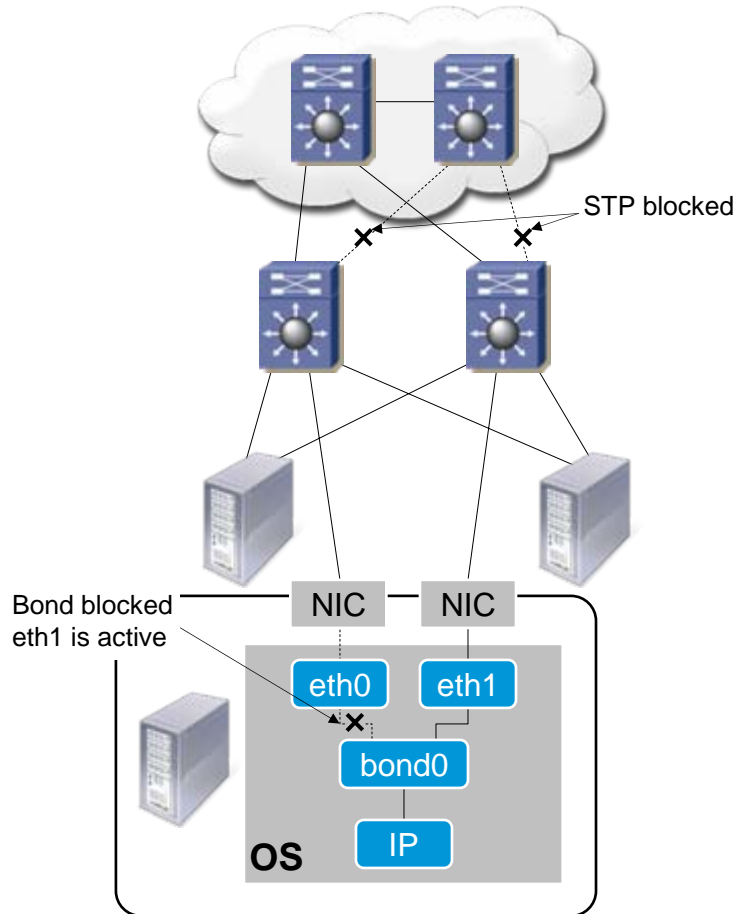
Hosts are attached redundantly to each network

- Distinct access switches and cables

Host uses SCSI multi-pathing to access similarly connected storage

- devmapper, PowerPath, DMP, etc.
- Active-active or active-passive based on host and storage device capabilities

LAN Redundancy Overview



Single network with redundant components and links

Spanning Tree Protocol eliminates loops to create single active tree

- Router redundancy provided by HSRP, VRRP, etc.
- Forwarding database populated by learning from traffic

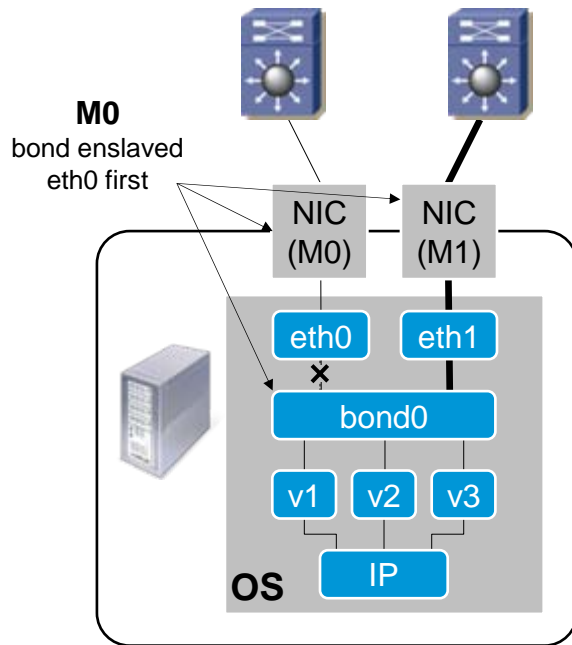
Hosts attach to network redundantly

- Distinct access switches and cables

Bonding driver selects active path

- Only one active at a time
- Selected by link state, remote connectivity probe (ping), configuration

More on LAN Bonding



Available with modern OSes

- Linux used in this description

NICs are “enslaved” into a bond

- Bond uses MAC address of first enslaved device
 - All devices to use bond MAC address

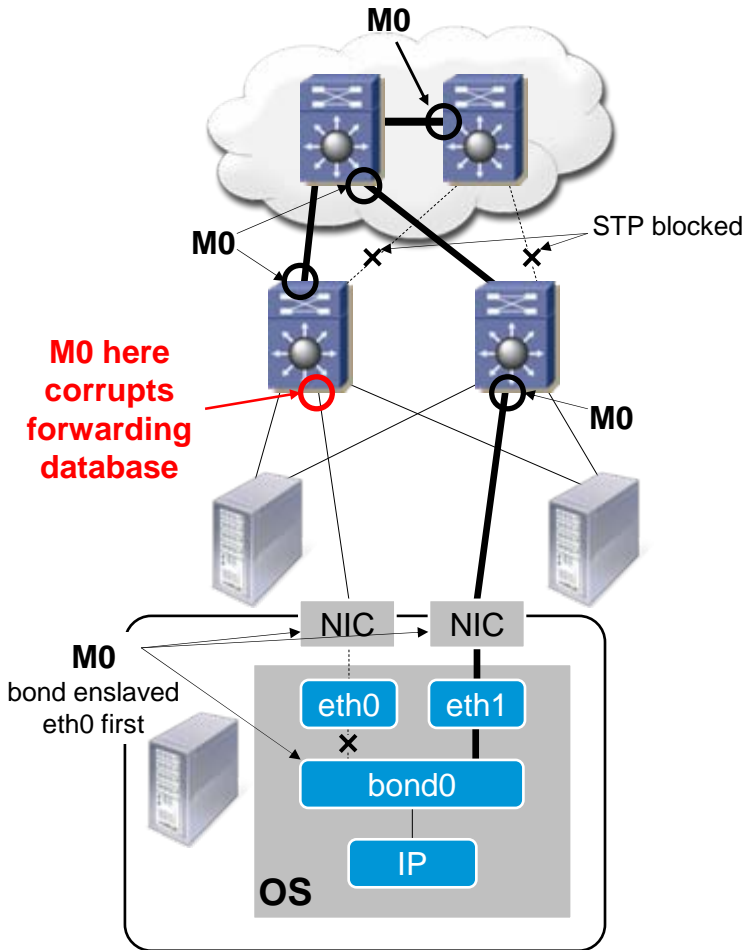
Redundant bond is “active-backup”

- Only one slave is active for transmit and receive
 - Frames from backup slave are dropped
- This is different from link aggregation bond used for increased bandwidth

Bond is a network device for upper layers

- VLANs configured on the bond
 - Bond settings apply to all VLANs
- IP interfaces to bond (or VLANs)

More on Bonding and the LAN



Ethernet “learns” from frame’s source address

- All switches in L2 domain learn all MAC addresses due to broadcast, multicast, and unknown unicast flooding

Redundant bond sends and receives frames on primary interface

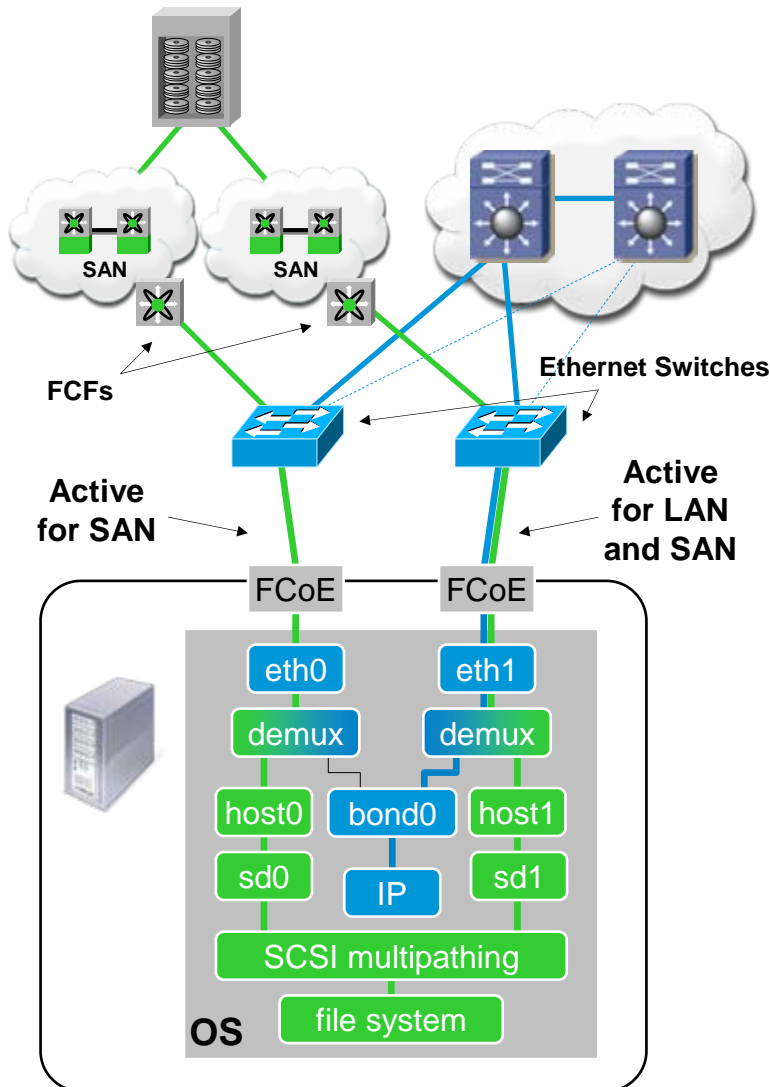
- LAN reaches bond on primary interface

Sending a frame with bond address from the bond’s backup interface is a **corrupting learning event** for the L2 domain

- Black holes (data loss)
- Out-of-order delivery
- Inconsistent security policies

Note: Switch forwarding tables are typically independent for each VLAN

Supporting Both Redundancy Models



Both server link active for SAN

- Using bond for FCoE breaks SAN A/B model

One server link active for LAN

Link usage between LAN and SAN is uncorrelated

- Selection made by different pieces of the system
- Criteria are different
 - SAN – Array disk and cache access
 - LAN – Reachability

LAN and SAN must use uncorrelated MAC addresses to avoid forwarding table corruption

Conclusions Restated

Can not use bond for Fibre Channel traffic

- Breaks SAN A/B model

LAN and SAN must use different MAC addresses

- Avoid LAN forwarding database corruption behavior
 - MAC address defined as {VLAN, MAC}
- Requires simultaneous (active) use of at least 3 different MAC addresses on the host

Mapped Address scheme supports existing LAN and SAN redundancy models



Thank You

