excelfore

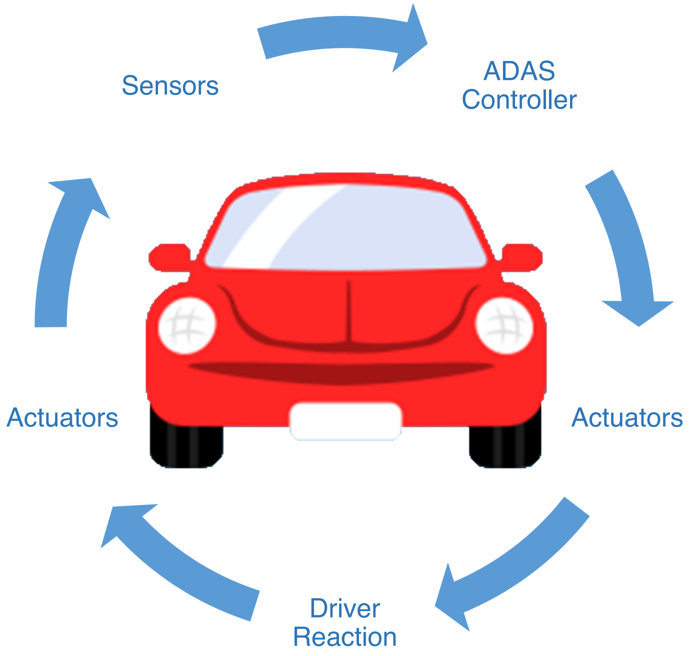
# Using Redundant Data Paths and Clock Domains in Ethernet TSN

## for Mission-Critical Network Reliability


Presented by: Shrikant Acharya
Chief Technology Officer, Excelfore Corp.

Contributing Authors: Anoop Balakrishnan, Excelfore Corp.
Shiro Ninomiya, Excelfore Corp.


# Mission-Critical Automotive Networking




Everything
Working
Together

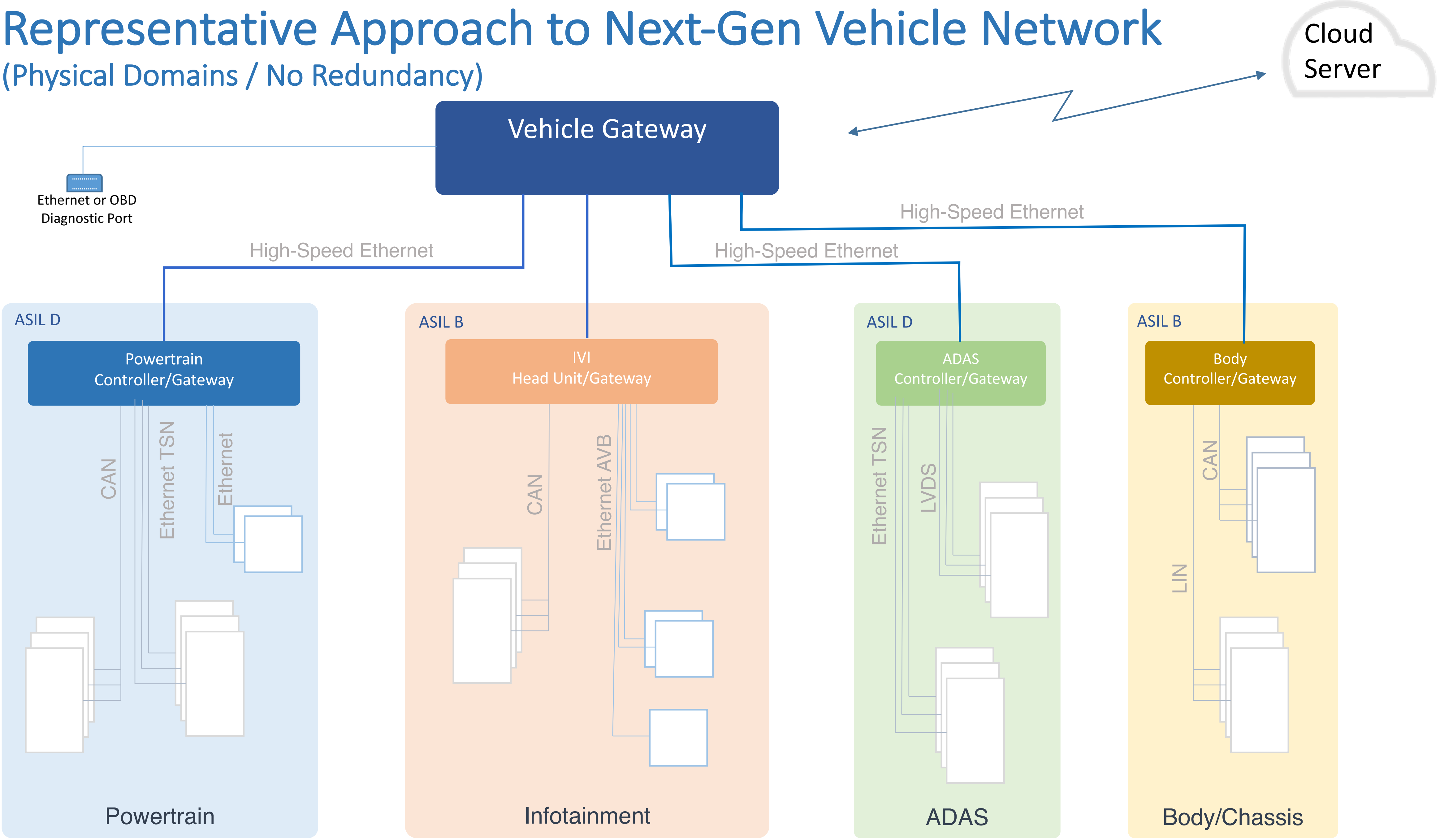
Enhanced Safety

Network
Failure

Problems!

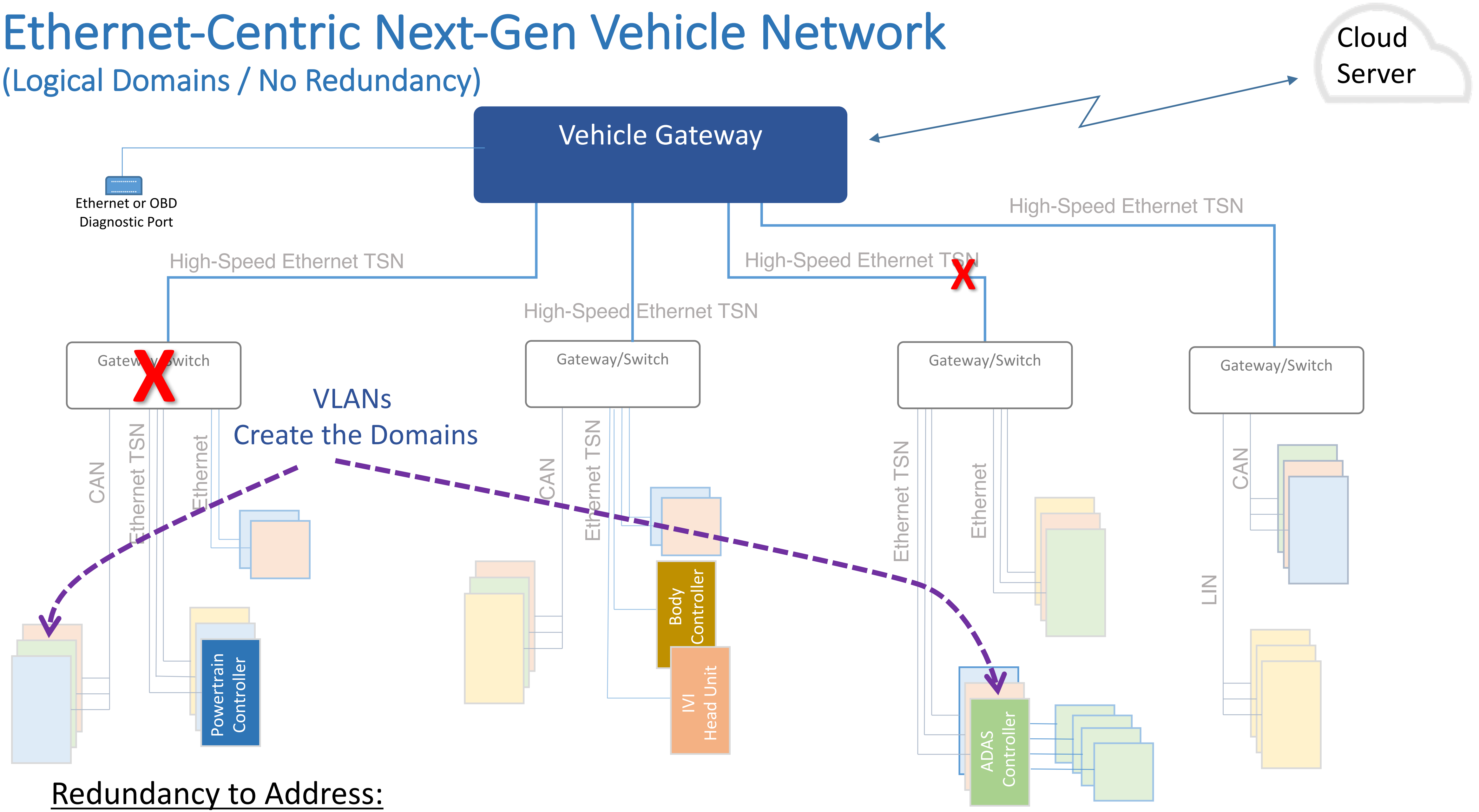
# Representative Approach to Next-Gen Vehicle Network

## (Physical Domains / No Redundancy)



# Ethernet-Centric Next-Gen Vehicle Network

## (Logical Domains / No Redundancy)




**Redundancy to Address:**

Failure of a Network Link
Failure of a Device on the Network

# Mission-Critical Network Redundancy

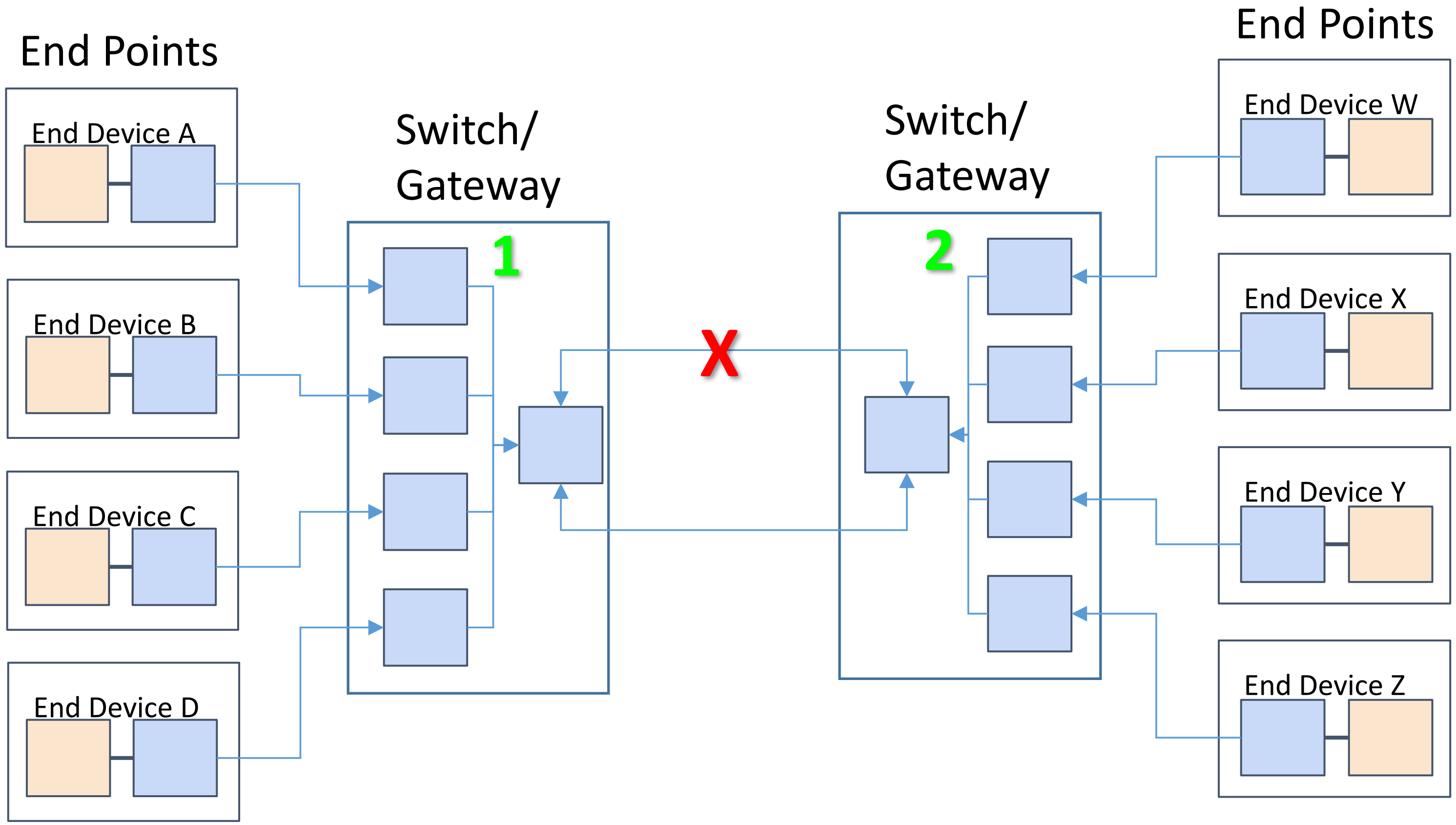
Key Software Concepts for Redundancy in TSN Networking:

A. Redundant Data Paths – IEEE802.1CB
B. Timing and Synchronization – IEEE802.1AS / 802.1BA
C. Redundant Clock Domains – IEEE802.1ASrev

Three Levels of Hardware Redundancy:

1. Redundant Links Between Network Gateway/Switches
2. Daisy Chaining End Devices to a Network Gateway/Switch
3. Daisy Chaining End Devices to Redundant Network Gateway/Switches

# Redundant Links Between Switches




<u>Positive Attributes:</u>

- Protection from Failure of Network Link on Highspeed Backbone
- Maximum of 2 Switch Hops Retains TSN Guaranteed Latency (< 2ms on 100Mbps Ethernet)

<u>Shortcomings:</u>

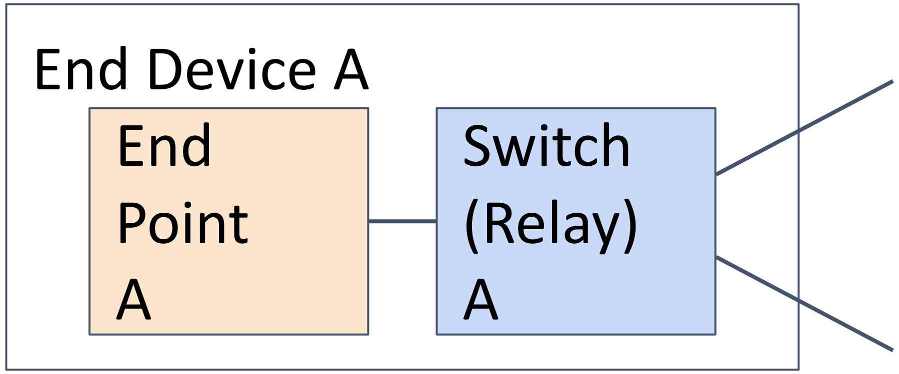
- No Protection from Failure of Network Link to End Devices
- No Protection from Gateway/Switch Device Failure

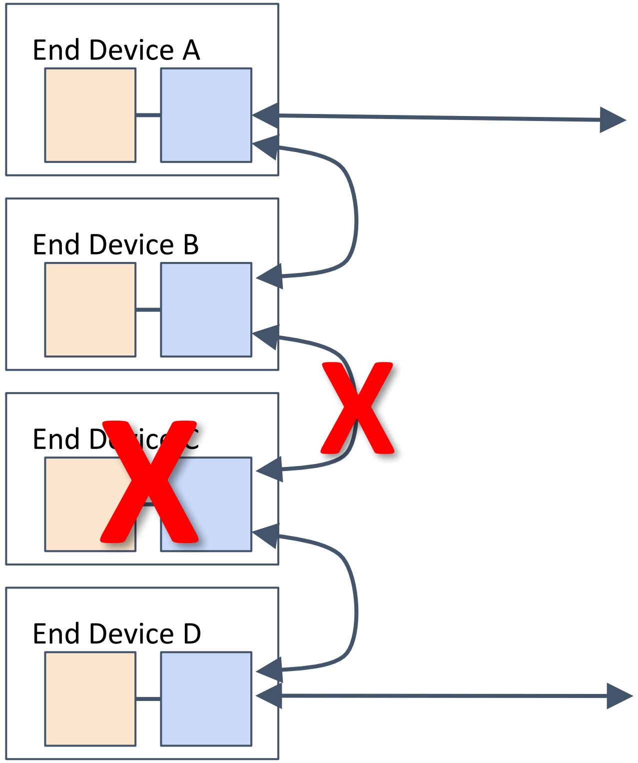
# Dual Ethernet Nodes: Key Hardware Feature

- Limitation of Single Node End Points
  - Redundant Paths only at Switch Nodes, not at End Points
    - Frame Replication at the Switch
    - No Frame Replication at End Point

- Enhanced Redundancy with Dual Node End Points
  - End Points can Replicate Frames from a Talker
  - Daisy Chaining of End Points Improves Redundancy
  - Daisy Chaining of End Point Improves Utilization of Switch Ports
  - Automotive Processors Support Dual Ethernet Nodes:
    - NXP i.MX6 Family
    - TI Jacinto J6 Family

# Daisy Chaining Dual Node End Devices




End Device with 2 Ports May have a 3 Port Switch:

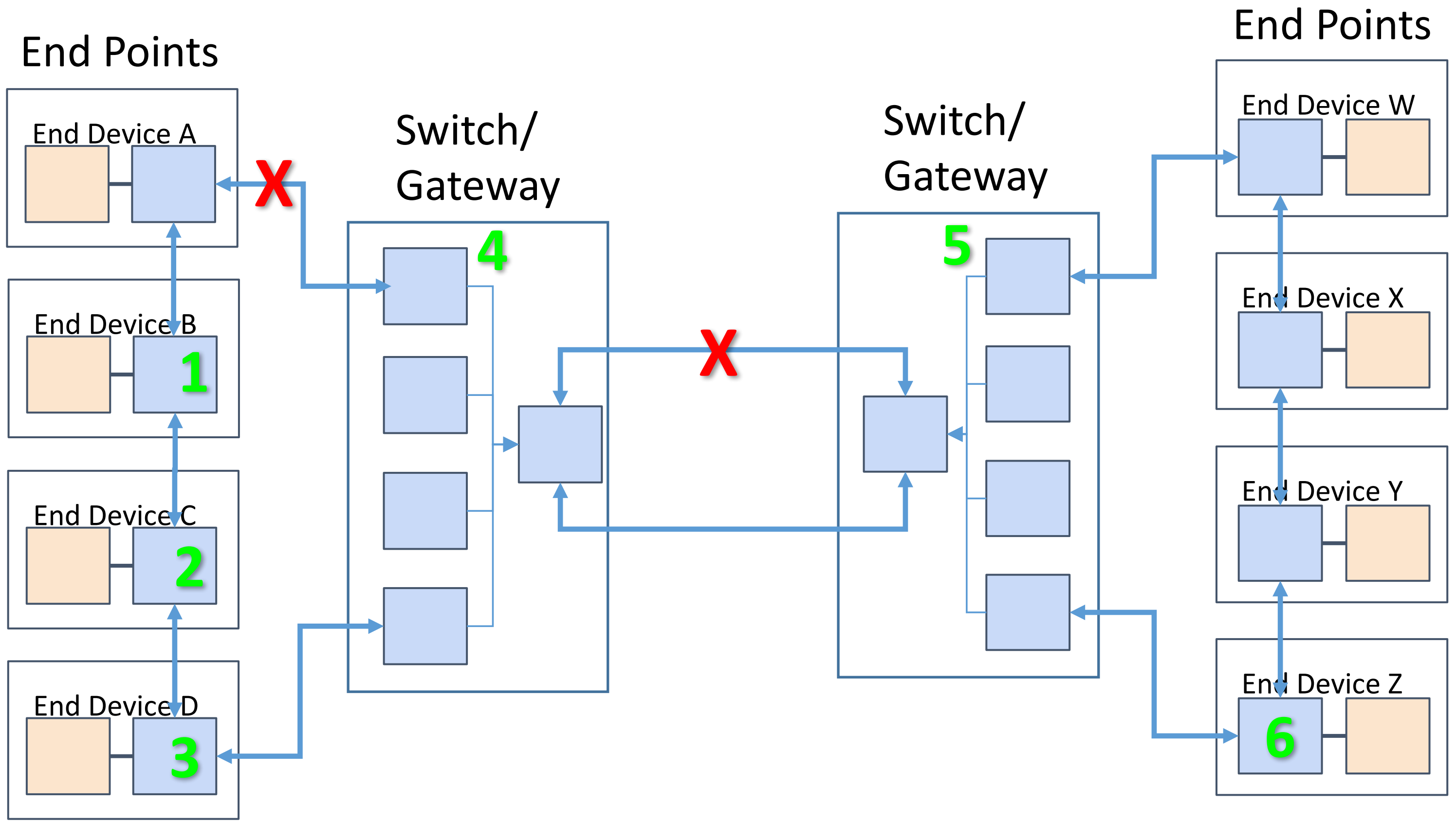
- 2 External Ports
- 1 Internal Port

One Link Failure Does not Disconnect Devices

One Device Failure Does Not Disconnect Other Devices

Careful Analysis of Switch Hops Required to Ensure Guaranteed Latency

# Redundant Links between Switches / Dual Node End Points




3 Hops from End Point to Backbone    2 Hops in the Backbone    1 Hop from Backbone to End Point

Positive Attributes:

- Protection from Failure of Any One Network Link
- Network is Still Protected from Edge Device Failure
- Better Node Utilization at the Switch
- Maximum of 6 Switch Hops (3 + 2 + 1 -or- 1 + 2 + 3)
  Retains TSN Guaranteed Latency with Any One Failure

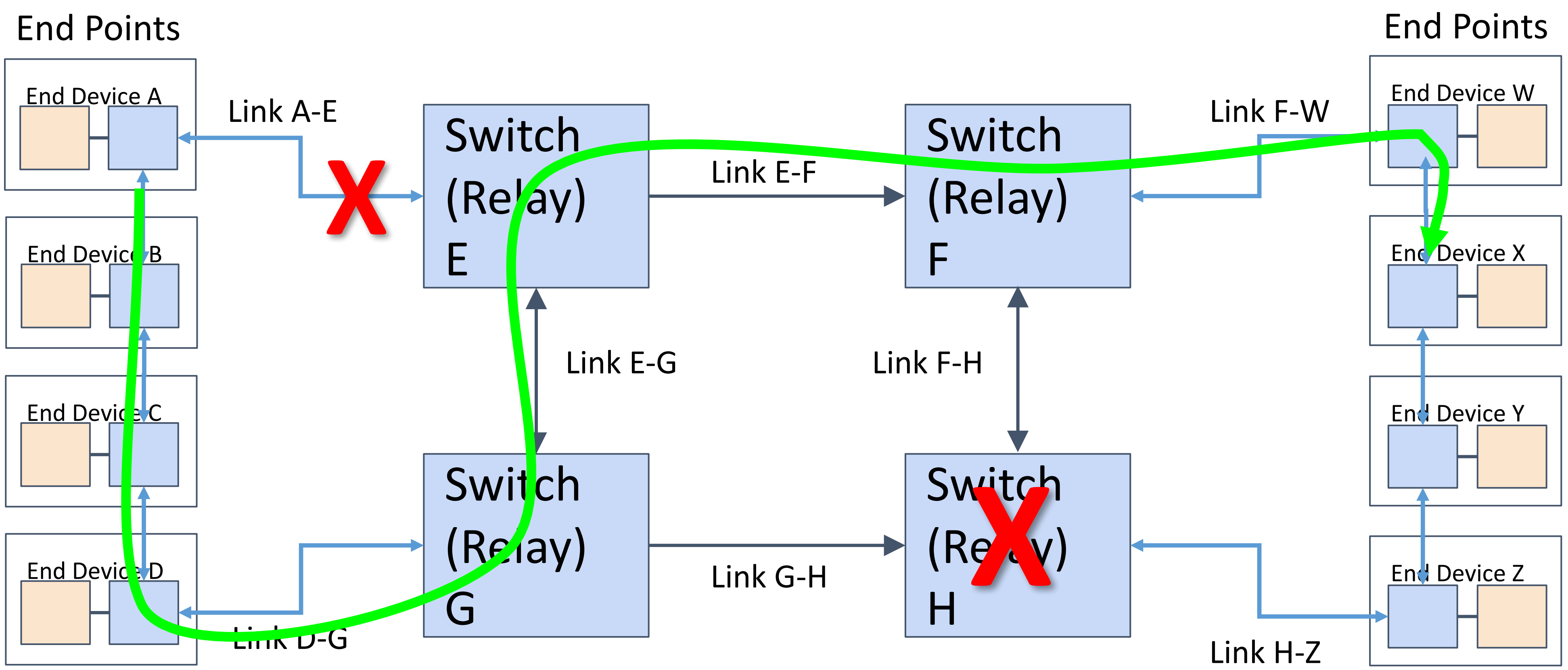
Shortcomings:

- No Protection from Gateway/Switch Device Failure

# Redundancy Impact

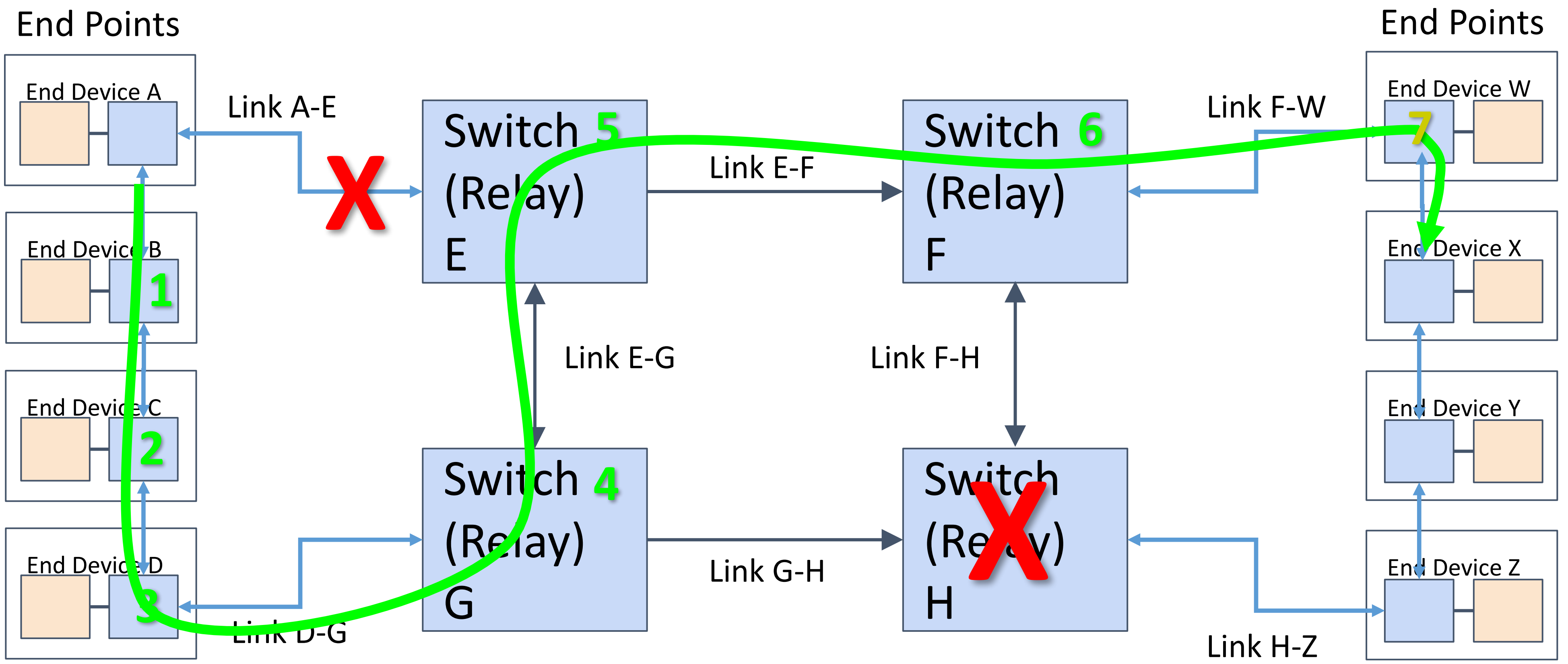
- Hardware Costs
  - End Points Need Two External Ethernet Nodes
- Software Performance (higher impact with higher payloads, utilization doubles)_
  - Overhead of Replication on the End Point
    - All packets = processing doubled
    - If overhead for packet transmission = 10%, with replication = 20%
  - Overhead of Replication on the Switch (Utilization is Doubled)
    - Depends how many packets need to be replicated to the various ports
    - Also impacted is how many deletions are happening
- Network Bandwidth
  - Aggregates Bandwidth Load of Daisy-Chained End Points
  - Overall Network Traffic on Some Links May Increase by Multiple (discussed later)
- Daisy Chaining Mitigates the Port Utilization at the Switch
- End Points with Single Nodes Can Not Daisy Chain
  - May be Appropriate for Non Mission-Critical Tasks

# Full Redundancy in End-to-End Network Connections




Loss of Any Single Network Link, or Any Network Switch, is Recoverable
Loss of Any End Point Does Not Affect Connectivity of Other End Points

# Control Latency: Analyze the Number of Hops




3 Hops from End Point to Backbone

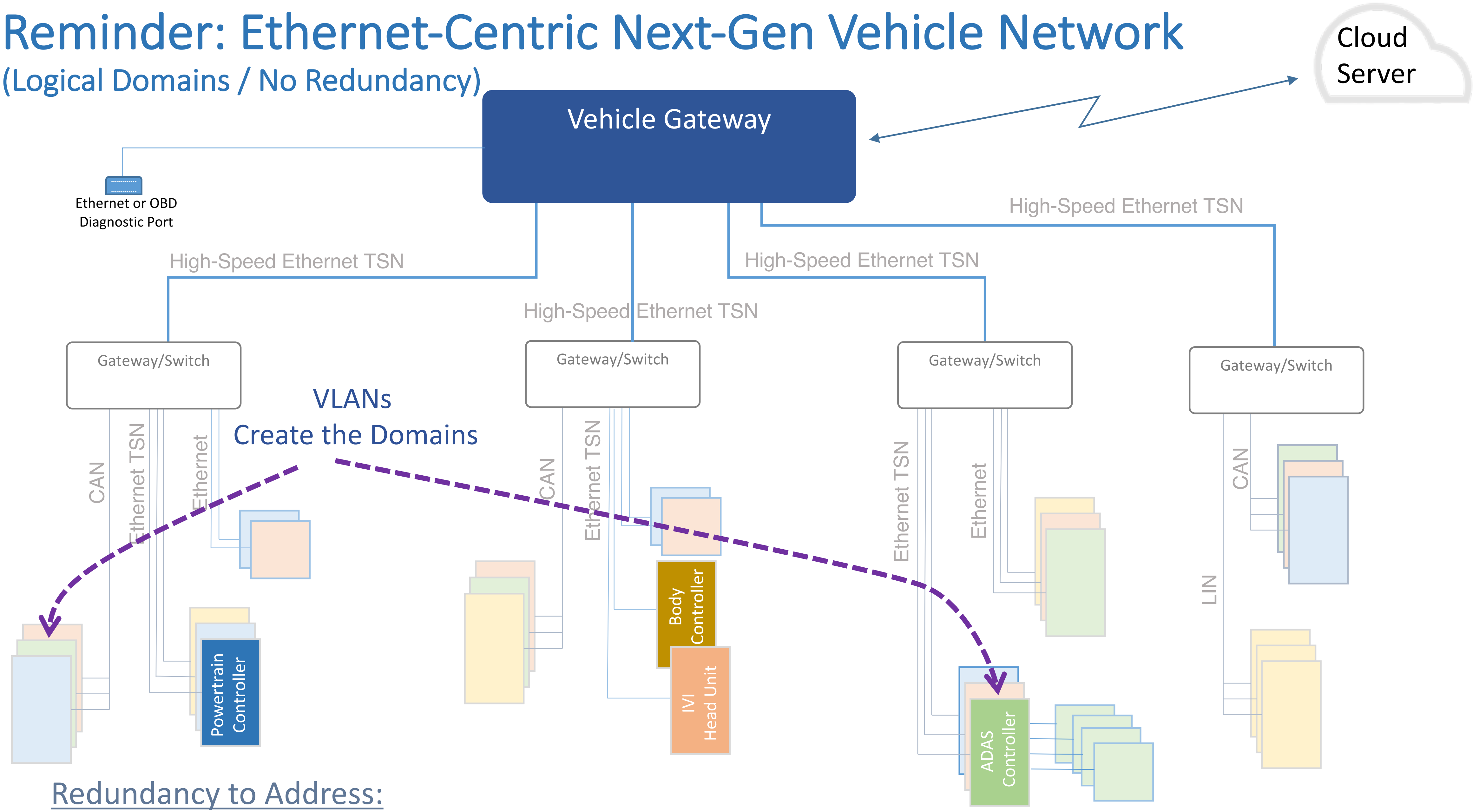
3 Hops in the Backbone

1 Hop from Backbone to End Point

2ms End-to-End Latency Guaranteed on 100Mbit Network - For Any One Failure **No More than 7 Switch Hops**

# Reminder: Ethernet-Centric Next-Gen Vehicle Network

## (Logical Domains / No Redundancy)

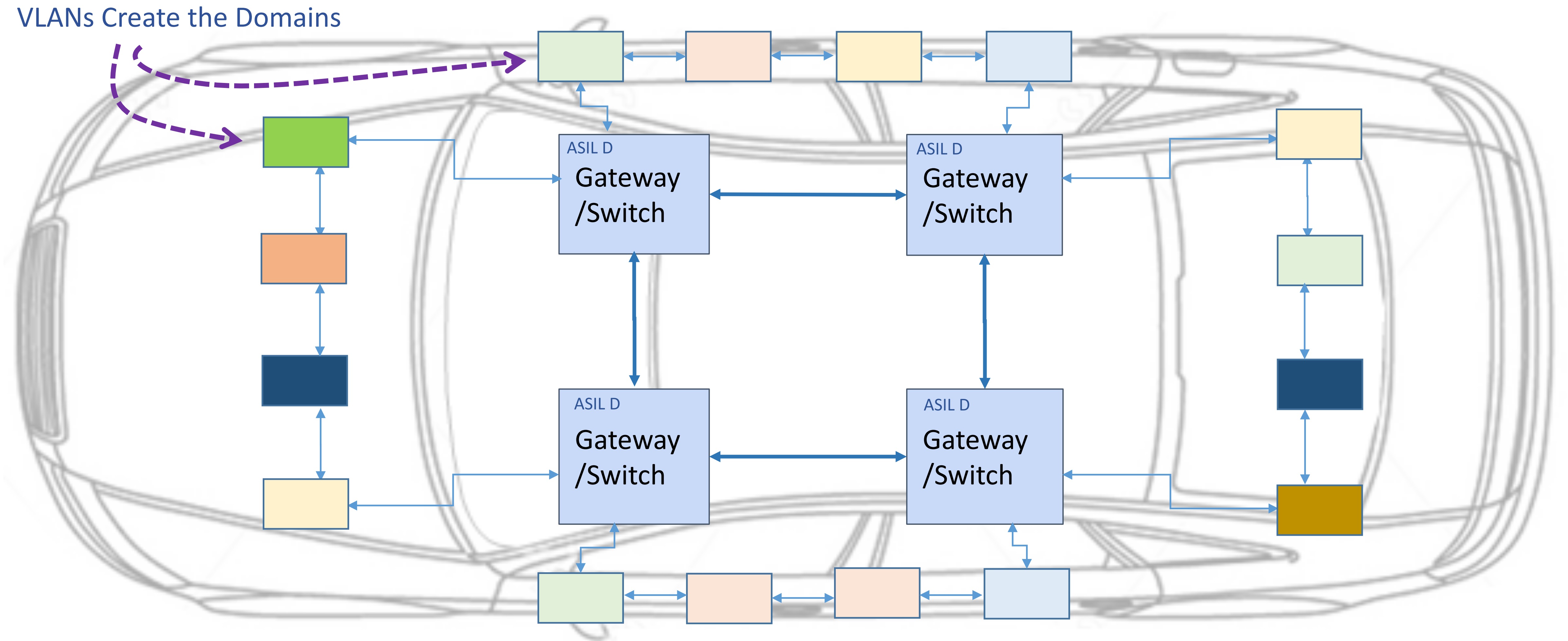



Redundancy to Address:

Failure of a Network Link

Failure of a Device on the Network

# Full Redundancy in End-to-End Network Connections




Loss of Any Single Network Link, or Any Network Switch, is Recoverable

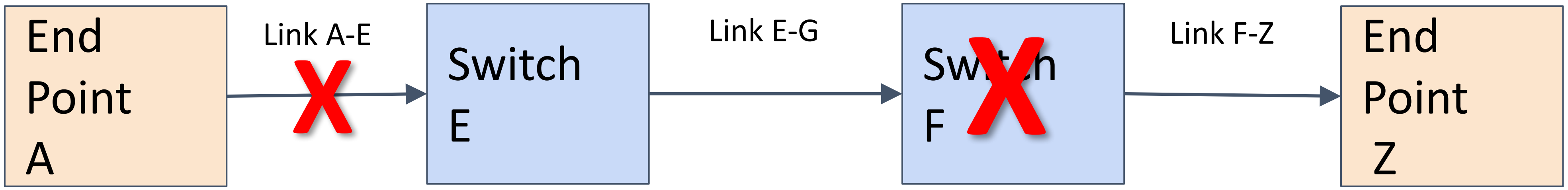
Loss of Any Single Network Link or Switch Preserves Guaranteed Latency

Loss of Any End Point Does Not Affect Connectivity or Latency of Other End Points

# Software Implications of Redundant Network Paths

## Frame Replication and Elimination for Reliability
## IEEE 802.1CB

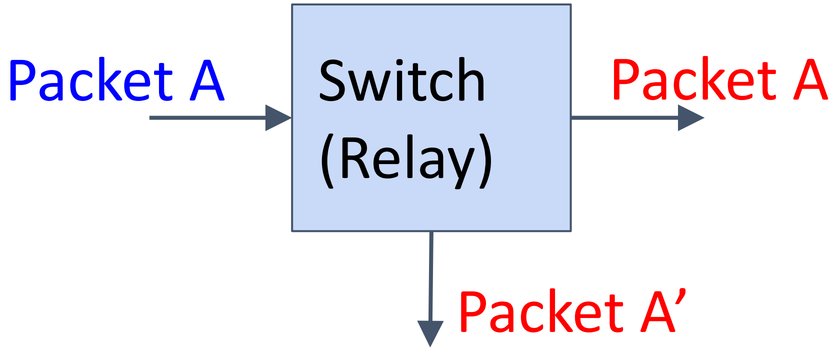
# Simple End-to-End Network Connections
## (No Redundancy)

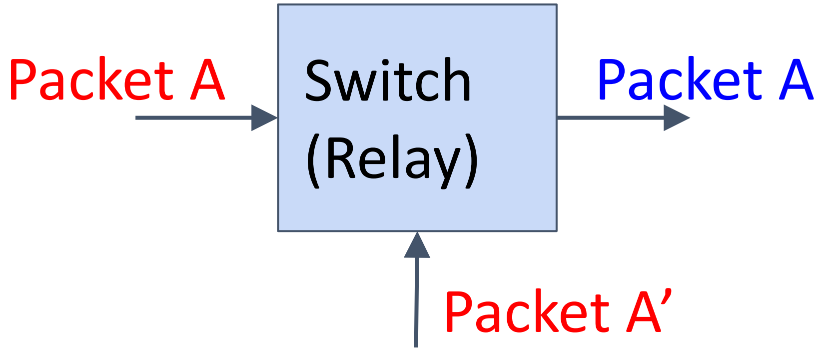



- Link A-E
- Link E-F
- Link F-Z
- Switch E
- Switch F

Failure in Any One Makes the Connection Fail

# FRER
(Frame Replication and Elimination for Reliability)



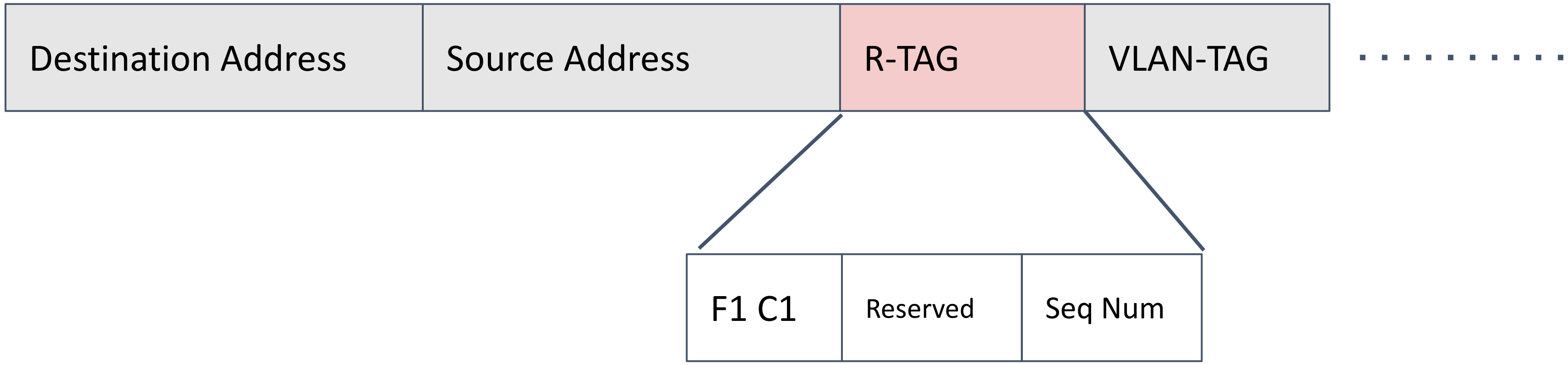




**Replication**
1x Incoming “Packet A”
“Packet A” is Replicated
2x “Packet A” Sent Out

**Elimination**
2x Incoming “Packet A”
1x “Packet A” is Eliminated
1x “Packet A” Sent Out

# Identifying "Packet A"

## Ethernet Header

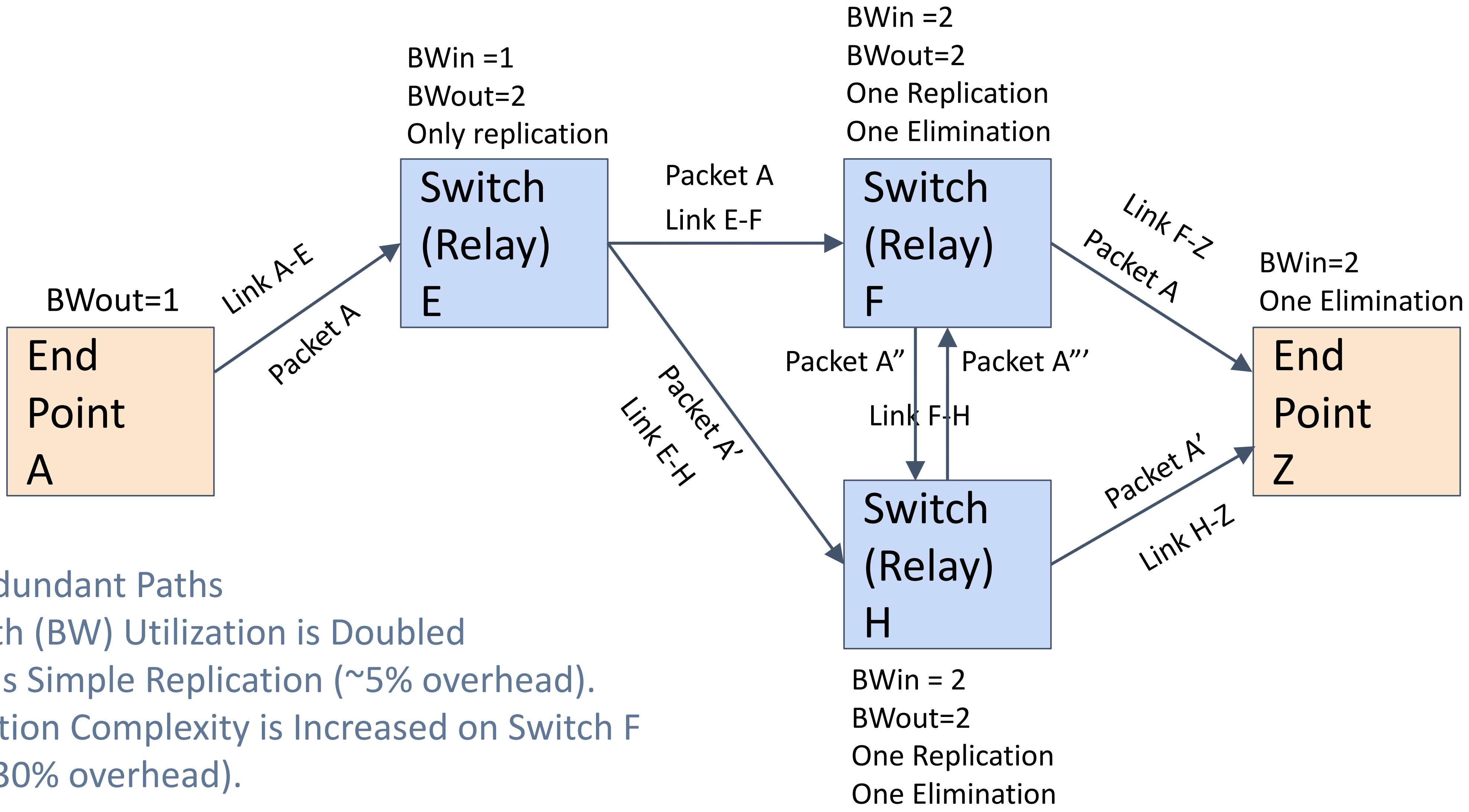



(optionally 'HSR-TAG' or 'PRP-TAG')

- Destination Address + Source Address + Vlan ID + Seq. Number can Identify the Packet

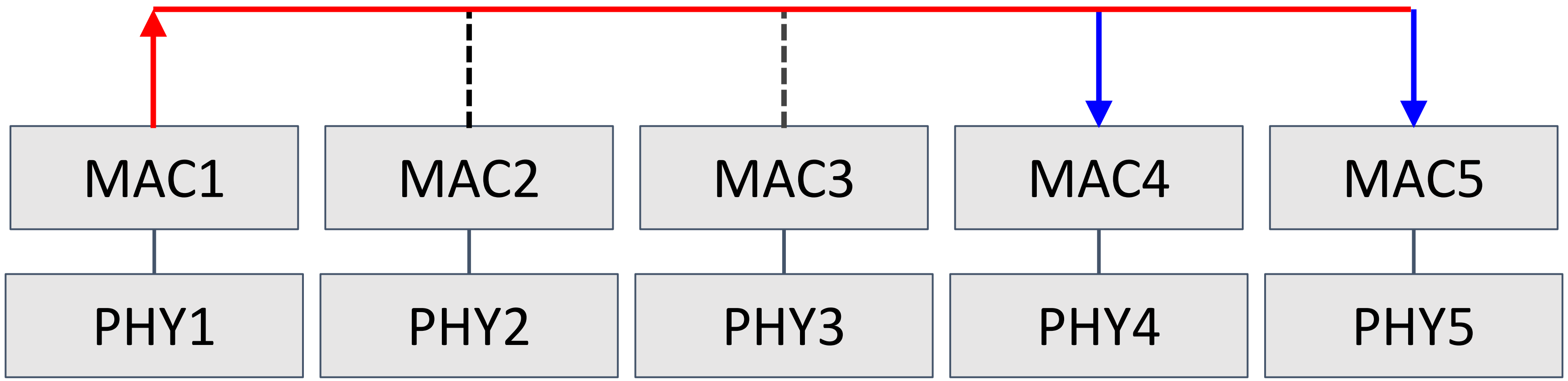
- This Packet Identification is Sufficient for Replication and Elimination by Relay System (Switch)

# Frame Elimination and Replication Explained



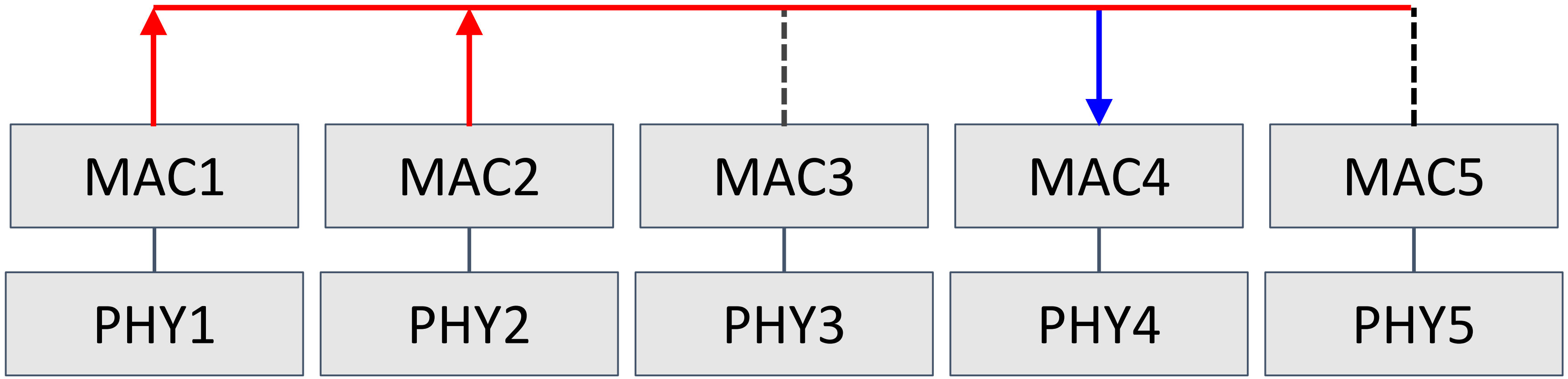

1. Many Redundant Paths
2. Bandwidth (BW) Utilization is Doubled
3. Switch E is Simple Replication (~5% overhead).
4. Computation Complexity is Increased on Switch F and H (~30% overhead).

# Software Implementation
(Replication)




- Check R-TAG in the Incoming Packets from MAC1
  If not Exit, then Insert R-TAG
- Keep Track in the Internal Table for PACKET ID
- Replicate and Send to Requested Ports (MAC4, MAC5)

# Software Implementation
(Elimination)




- Check R-TAG in the Incoming Packets from MAC1 and MAC2
- Keep Track in the Internal Table for PACKET ID
- Eliminate Replicated Packets and Send to Requested Ports (MAC4)
  If MAC4 does not Request R-TAG, Remove It

# Design Implication for Replication/Elimination

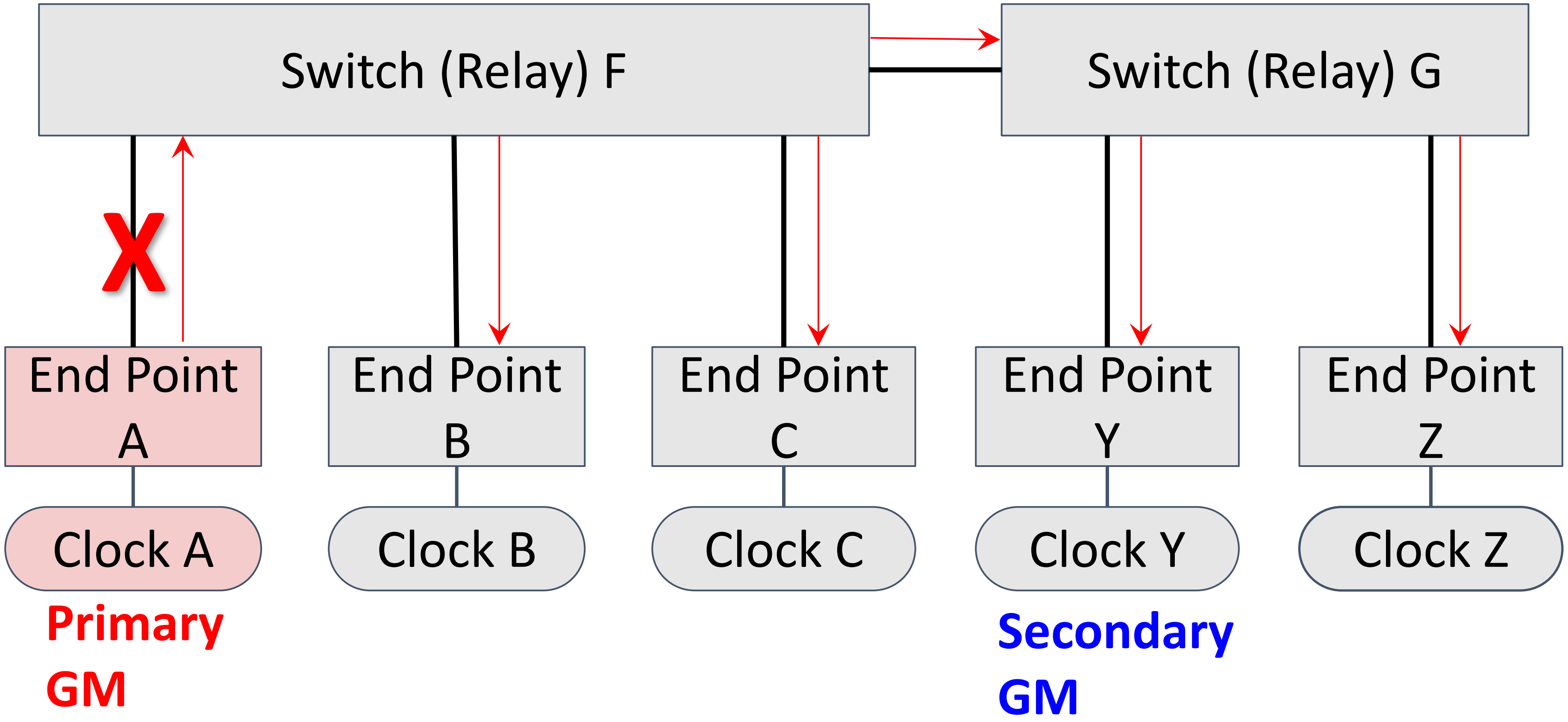
- Software Solution
  Layer 2 Software can Implement this Logic – Requires ID Check on Each Packet
  - This Impacts Latency from Additional Processing
  - Processor Utilization May Exceed Capacity Under Heavy Traffic (~40Mbits/Second of Video Data)

- Suggested Hardware Acceleration
  - R-Tag Insertion or Elimination
  - PACKET ID Look-Up Table (e.g. MAC Addr, VLANID, Sequence No.)

# Redundancy of GrandMaster Clock

## No Disruption of Network Devices by GM Failure
## IEEE 802.1AS-Rev

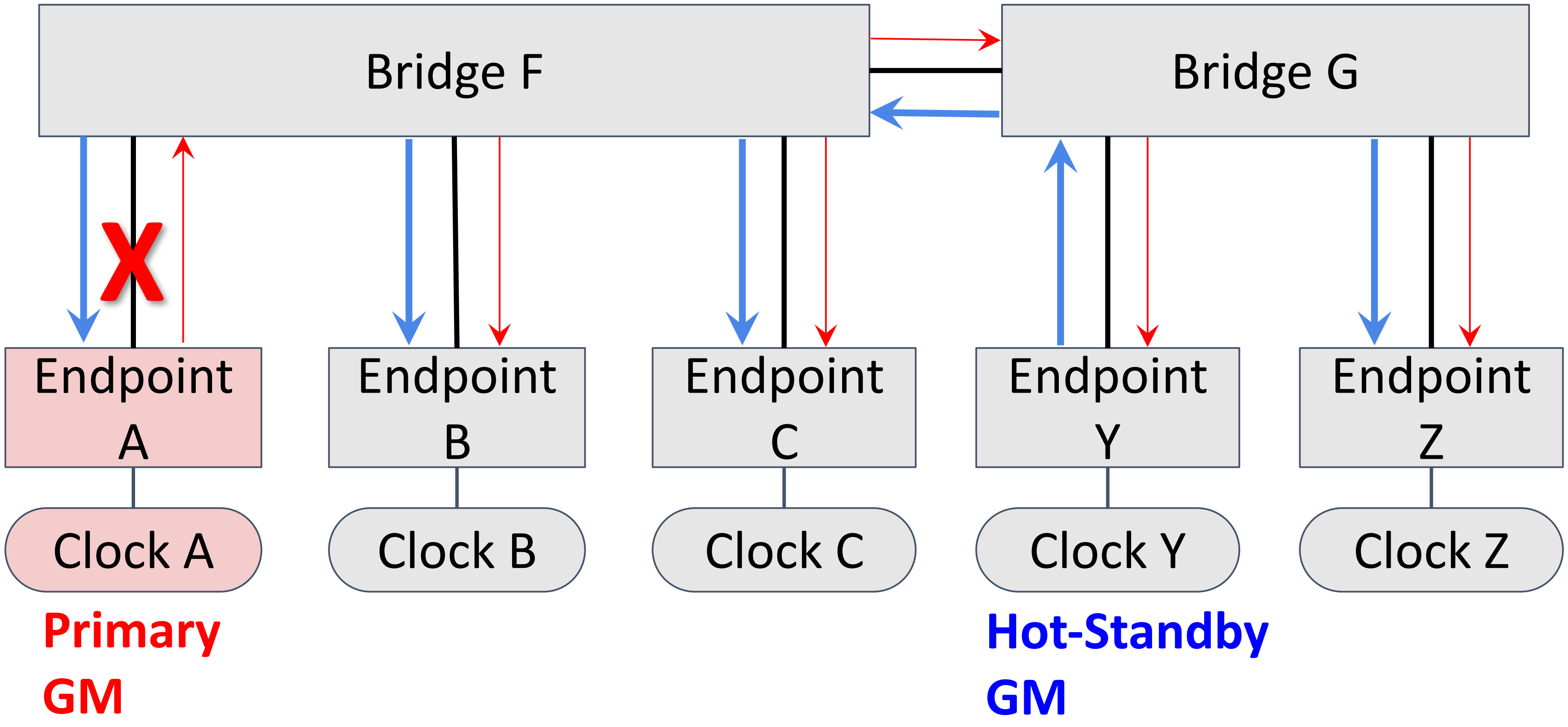
# Current Diagram for Clock Sync



# Current Procedure for Clock Sync Implementation

- End Point A Fails
  GM Clock (Clock A) is Lost on the Network
- Network Starts BMCA (IEEE1588 Best Master Clock Algorithm)
  Chooses One of among Clock B to Clock Z as New GM Clock
- Clock Y Becomes New GM Clock
- Switching GM from Clock A to Clock Y
  Procedure Requires Multiple Seconds
  All Devices Lose Synchronization During Procedure

# Diagram for Redundant GM Clock Sync Implementation



# Procedure for Redundant GM Clock Sync Implementation

- Primary GM is Clock A
  Secondary GM is Clock Y
- Two Domains of 802.1AS Clock are Running Separately
- Normal Circumstance:
  GM in the Secondary Domain is Not Operational
- Upon Failure of Primary GM:
  Network Seamlessly Switches to Secondary GM
- No Devices Lose their System Synchronization

Note:
Management of Multiple Domains of PTP Messages is Currently Being Defined in 802.1AS-rev

# Implementation of Redundant GM
## (Updating the gPTP Kernel)

Following Functions Must Be in Updated gPTP:

- Handling of Multiple Domains of SYNC Messages

  Our Example is Two Domains – *Could be More*

- Manage Clocks of Multiple Domains

  Keep Track of Primary GM and Secondary Stand-by GM

  Secondary GM Must Be Synchronized to the Primary GM

  (Required for Seamless Switching)

- If Primary GM Fails Each gPTP End Device Switches to Secondary GM

  No Impact from Clock Discontinuity on Any gPTP End Device

  Switching from Primary GM to Secondary GM is Seamless

# Replacement of Malfunctioning GM – a Proposal
## (Updating the gPTP Kernel)

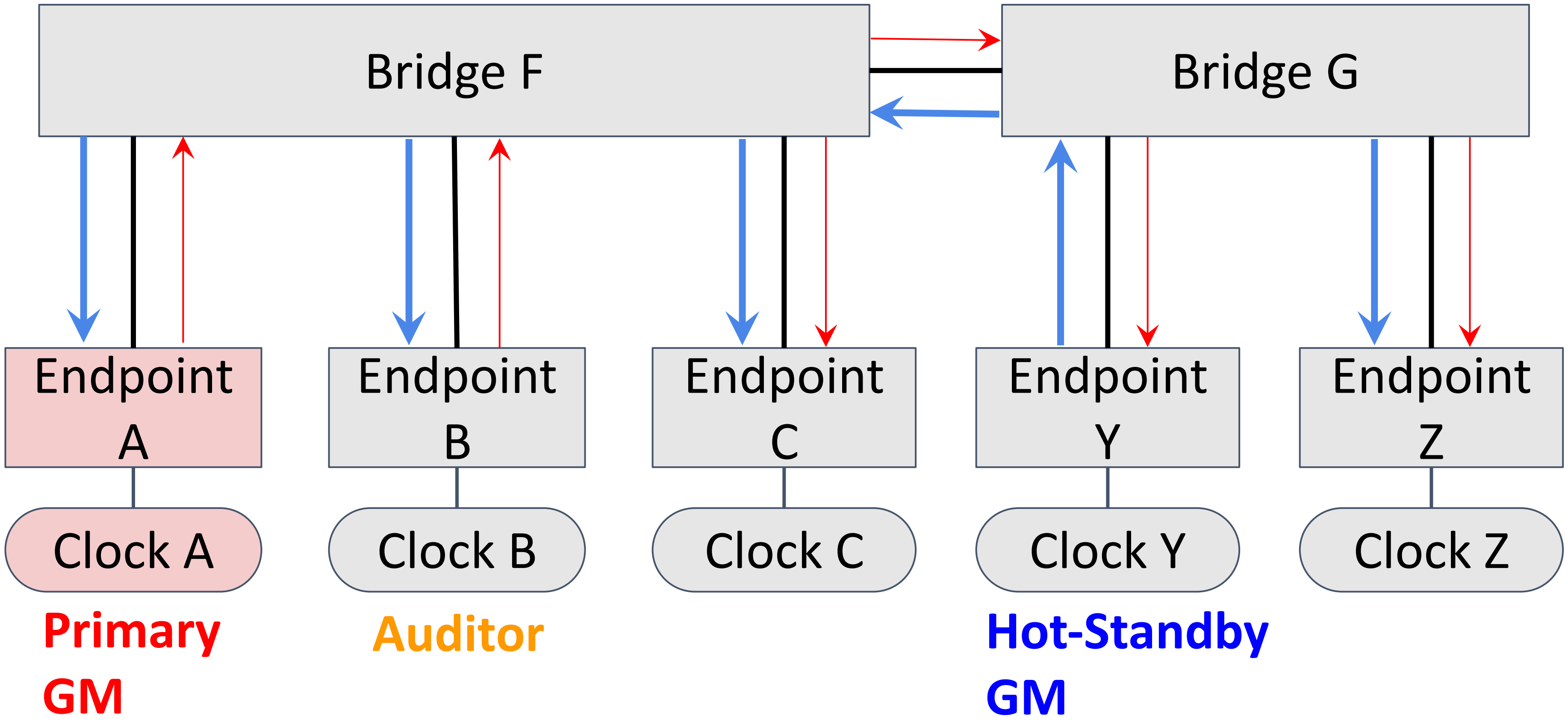
Case of a Malfunctioning GM

(Clock is degraded, but not lost)

- Two GMs Inadequate for Redundant Clock Domains with Hot Standby
  - Which GM is Correct in a Dispute?
- Requires Third GM to Audit Clock Behavior
- Implementation of the Auditor GM
  - One GM Contests That Other GM is Malfunctioning
  - Auditor Checks Status of Both GMs
  - Auditor Renders Decision and Notifies All GMs
  - Auditor Sends Malfunction Notification to GM

    It surrenders and ceases to be GM

# Diagram for Redundant GM Clock Sync Implementation



# Performance Impact of GM Clock Redundancy

- Network Traffic
  Additional ~1% Overhead in Redundant Sync Messages at 40Mbits/second

- Software Solution on Each gPTP Node
  GMAC Software Complexity will Increase
  - Each PHY/GMAC Receives 2x the Number of Sync Messages
  - Validate and Process the Secondary Sync Messages
  - Input Processing Requires More Performance in PHY/GMAC

- Suggested Hardware Acceleration
  Detection of Clock Domain ID
  Keeping Track of Separate Sync Messages and Time Stamps

# 802.1AS Rev Spec vs. Implementation

- Standard Only Warrants How Hot-Plug GM Setup Envisaged
  How to manage multiple different domains of PTP messages still under definition

- Detection of Malfunctioning GM is Not Part of the Standard
  - Left to Individual Implementation
  - Minimum: Third GM for Monitoring
    - Monitoring and Regular Review
    - Implications for Startup Time
    - Added Cost to Implement
    - Input Processing Requires More Performance in PHY/GMAC

- Cost Implication for Third GM
  - Complexity Left to System / Network Implementer

# Summary of Opportunities for Hardware Acceleration

For Frame Replication and Elimination for Reliability:

- R-Tag Insertion or Elimination
- PACKET ID Look-Up Table (e.g. MAC Addr, VLANID, Sequence No.)

For Redundancy of GrandMaster Clock:

- Detection of Clock Domain ID
- Keeping Track of Separate Sync Messages and Time Stamps

# Summary and Conclusion

- Automotive Networking Must Address Mission Critical Requirements
- Ethernet TSN Has Structures for Redundant Links to Mission Critical End Device
- Redundant Data Paths Ensure Mission Critical Network Links
- Careful Analysis of Network Hops Ensures Guaranteed Latency
- Redundant Clock Domains Could Ensure Seamless Continuity of Mission Critical Network Operation