

ETHERNET/IP: TECHNIQUES FOR FAULT TOLERANT NETWORKS

Sivaram Balasubramanian, Anatoly Moldovansky and Kendal R. Harris
Rockwell Automation

Presented at the ODVA
2007 CIP Networks Conference & 12th Annual Meeting
September 18-20, 2007
Englewood, Colorado USA

Abstract

This paper describes a protocol for EtherNet/IP networks that can tolerate multiple failures in network infrastructure and network interfaces on end devices. This protocol is called Multiple Fault Tolerant Protocol or MFTP. A network based on the MFTP protocol is called an MFTP network. The MFTP network is based on ISO/IEC 8802-3 (Ethernet) and IEEE 802.1 standards and redundant infrastructure. In MFTP network, the decision to switch between infrastructures is made individually in each end device. The MFTP protocol allows realization of very large networks with 10,000 end devices or more and with worst case fault detection and switchover time of 1 ms or less for most types of faults.

Introduction

There exists a class of applications in the industrial control market where continuous system operation is required even when faults occur in control system elements. This paper describes a Layer 2 protocol for EtherNet/IP networks that can tolerate multiple failures in network infrastructure and network interfaces on end devices. Such fault tolerant networks can be used to realize highly reliable control systems. For example, it can be used in combination with voting scheme for sensors/actuators and hot-backup scheme for controllers/actuators to realize highly dependable control system.

There are several approaches to fault tolerant networks. One approach is to use two or more networks, with each fault tolerant device having a network interface connected to each network. The devices use diagnostic messages to detect faults and switch from one network to other when a fault is detected. This approach has a failure detection and reconfiguration delay of 20-30 seconds. Another approach is to use single network, but with two or more sets of network infrastructure. The devices use diagnostic messages to detect faults and a middleware component to route messages. This approach is used by Honeywell Fault Tolerant Ethernet [1] with a failure detection and reconfiguration delay of 1 second. However, neither of these approaches can be suitable for networks, such as EtherNet/IP, used in applications requiring network fault detection and recovery, or switchover, time of less than 10 ms.

This paper describes a new approach that is suitable for EtherNet/IP networks. It is based on a protocol called Multiple Fault Tolerant Protocol or MFTP. A network based on the MFTP protocol is called an MFTP network. The MFTP network is based on ISO/IEC 8802-3 (Ethernet) and IEEE 802.1 standards and redundant infrastructure. In MFTP network, the decision to switch between infrastructures is made individually in each end-device.

Characteristics of this technology are as follows:

- MFTP networks will tolerate multiple faults in network infrastructure and network interfaces on end devices
- The MFTP protocol is suitable for networks, such as EtherNet/IP, used in applications requiring network fault detection and switchover time of less than 10 ms
- The MFTP protocol allows realization of very large networks with 10,000 end devices or more and with worst case fault detection and switchover time of 1 ms or less for most types of faults. The low switchover time allows utilization of MFTP network in hard real time control systems.
- In MFTP networks it is possible to connect non-fault tolerant end devices on the same network with fault tolerant end devices and to add or remove end devices, fault tolerant or otherwise, to or from the network dynamically during run time.
- In MFTP networks it is possible to support IEC 61588 (and version 2 of IEEE 1588) time synchronization, if needed.
- It is possible to extend this approach to Gigabit Ethernet and to 3-, 4- or N-way redundancy in general.

Principle of operation

The MFTP network topology can be described as two interconnected top switches, each heading an underlying topology of star, ring, or line (daisy chain). The Figure 1 below depicts an example of MFTP star network.

The MFTP network infrastructure is built from commercial off the shelf switches compliant with IEEE 802.1D and IEEE 802.3 standards. No support of the MFTP protocol in switches is required. The number of levels of hierarchy and number of switches on each level are dependent only on application requirements.

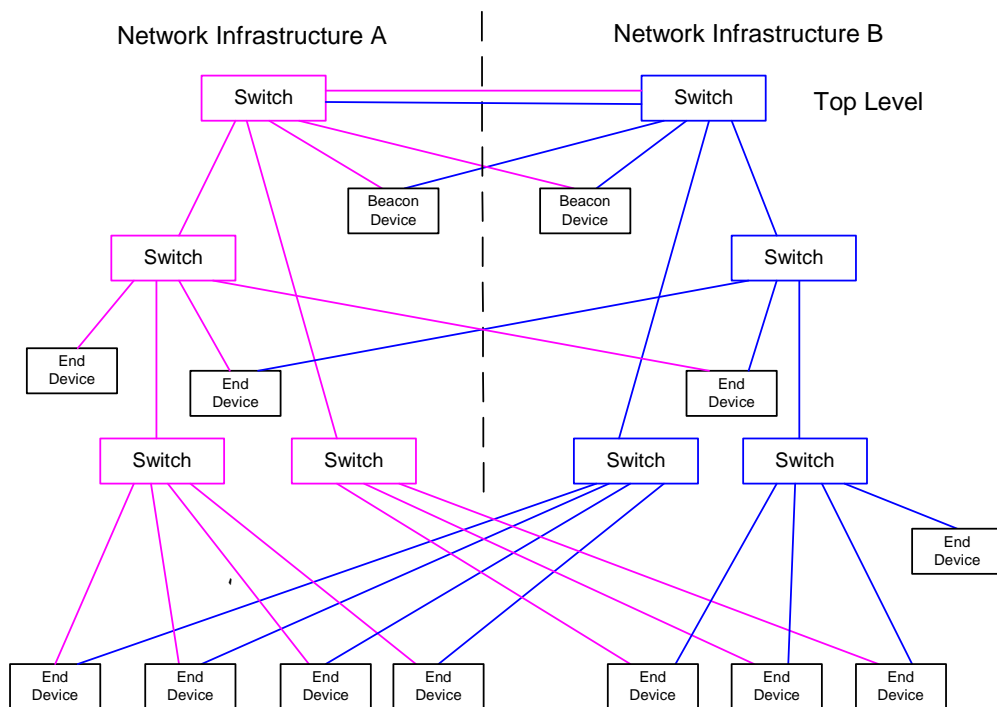


Figure 1: MFTP Network Example

Figure 1 shows an example of a MFTP network in the 2-way redundancy mode. It uses two sets of network infrastructure A and B (shown in two different colors). Even with three levels of hierarchy it is possible to construct very large networks. For example, a network infrastructure built from switches with eight regular ports and one uplink port can contain over 500 end devices and switches with 24 regular ports and one uplink port can contain more than 10,000 devices. Though symmetrical arrangement of switches can provide better non-overlapping sub-trees for fault tolerance, it is not a requirement.

Two backbone switches at top level are connected to each other with two or more IEEE 802.3 aggregated links [2]. With link aggregation capability traffic is shared among bundle of links and failure of one link will not bring the network down. With such an arrangement network infrastructures A and B form a single network.

Two types of end devices can be connected to the MFTP network infrastructure: doubly attached and singly attached. A doubly attached end device can function as a MFTP end device or as a MFTP beacon end device. Though doubly attached end devices have two network ports to interface with two network infrastructures A and B, they will use only one MAC address and one IP address. Singly attached end devices may also be connected to MFTP network but they do not support the MFTP protocol.

At any given point in time MFTP end device will actively communicate through only one of its ports, while blocking all traffic on its other port, with the exception of beacon messages and failure notification messages. An MFTP end device will continuously monitor status of immediate network links connected to it and arrival of beacon messages on both its ports. Fault tolerance is achieved in a distributed fashion by MFTP end devices switching between their ports from inactive to active mode and vice versa. Several examples are described in sub-sections 2.1 to 2.5.

Two beacon end devices are always connected to top level backbone switches. Both of them are always active. A MFTP beacon end device is a MFTP end device that is configured to perform beacon function. Beacon end devices transmit a short multicast beacon message on the network periodically. Similarly to a regular MFTP end device, beacon end device at any given point of time will communicate through only one of its ports, while blocking all traffic while blocking all traffic on its other port, with the exception of failure notification messages. All MFTP messages are transmitted at highest priority on network using IEEE 802.3 tagged frames.

The beacon interval is a configurable parameter chosen for a specific system. For example, for a 3 switch level system the beacon interval is 450 microseconds. This interval is arrived based on worst case delay for a beacon message to travel from beacon end device to the farthest end device. The end devices use a timeout mechanism to declare fault on a port when no beacon messages are received on that port for given timeout period. As an example, for 450 microseconds beacon interval, the timeout period can be 950 microseconds. The timeout period is chosen such that at least 2 beacon messages from each beacon device will have to be lost before fault is declared. Upon detecting a fault condition on one port the end device will switch to other port, provided a double fault has not been detected.

Commercial Ethernet switches implement a variety of traffic control features such as learning, filtering and IGMP snooping. Broadcast messages are not usually affected by any of these features and will be propagated throughout the network. Hence they have no issues with respect to rapid reconfiguration. For fast reconfiguration, multicast filtering features in the switches should be disabled. The multicast traffic will therefore be treated as the broadcast traffic.

Unicast packets are affected by learning and filtering features. After end device port reconfiguration switches will have invalid knowledge. A switch implementing learning will update its database when a packet with a learned MAC address in source field is received on a different port from learned port stored in database. When a MFTP end device switches to the inactive port, the first thing it will do is to send a short multicast message, called learning update message, through its newly enabled port. As this message propagates through the network, the switches will update their learning resulting in rapid reconfiguration of the unicast traffic. This message is of no interest to other end devices in the network and will be dropped by them.

Figure 2 shows the MFTP stack architecture. It is applicable to both MFTP and beacon end nodes.

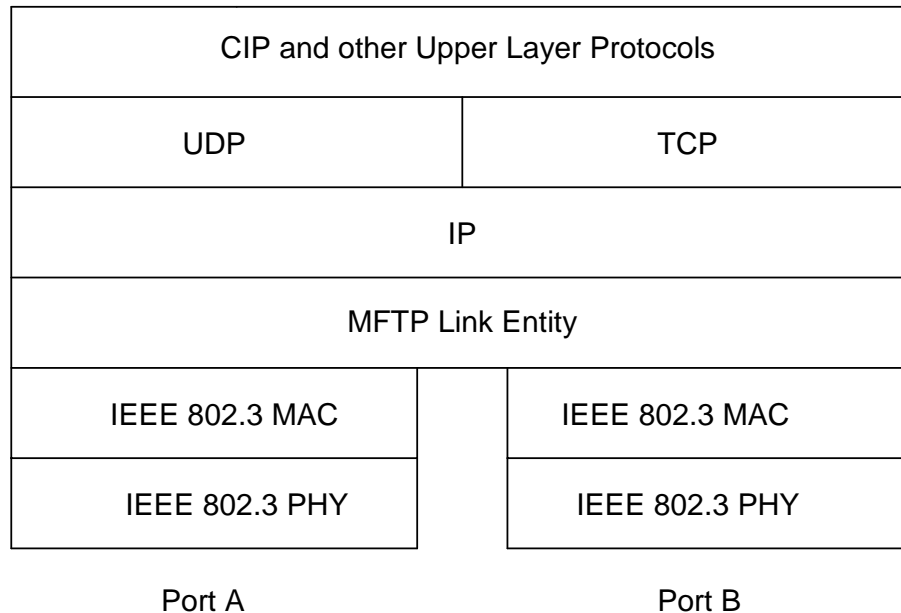


Figure 2: MFTP Stack Architecture

The MFTP stack contains two identical ISO/IEC 8802-3 ports, identified here as ports A and B, connected to the network infrastructure. These ports interface with the MAC sub-layer compliant with the ISO/IEC 8802-3 standard.

The MFTP Link Entity continuously monitors link status of two ports. All traffic flows only through the active port and is blocked on the inactive port with the exception of received beacon messages.

A Physical Layer failure of immediate link connected to the active port or a failure of the switch on the other end of this link can be detected rapidly and the MFTP Link Entity will reconfigure ports, provided the other port was not in the fault mode as well. After reconfiguration all traffic will start flowing through newly activated port.

Some messages may be lost during the reconfiguration process, but upper layer protocols will function correctly, since this will be no different from messages lost due to other network errors. If duplicate messages will be received, they will be detected by upper layer protocols as well.

The MFTP Link Entity also monitors arrival of beacon messages on both ports. When a beacon message fails to arrive on the active port for a configured timeout period, the port is declared to be in the fault mode. The MFTP Link Entity will reconfigure ports, provided the other port was not in the fault mode as well. After reconfiguration all traffic will start flowing through newly activated port.

It is possible for certain rare faults to occur in transmit path direction that is opposite to flow of beacon messages. If such a fault manifests itself in Physical layer, it will be detected by end devices or switches adjacent to fault. This will either result in a MFTP end device reconfiguring its ports immediately or will result in traffic being blocked on affected link. The latter will lead to loss of beacon messages on downstream end devices and they will reconfigure themselves on beacon timeout.

For the small subset of transmit path faults that do not manifest themselves in Physical layer a separate mechanism is used. When frames fail to arrive on a receiving end device from a transmitting end device of interest within associated timeout period, the receiving end device CIP connection will timeout. The timeout period is based on lowest requested packet interval of all connections from transmitting end device (producer) of interest to receiving end device (consumer).

Upon detection of CIP connection timeout, the receiving end device will send a failure notification message to associated transmitting end device. When failure notification message is received through active or inactive port of transmitting end device, that end device will send a path check request message to beacon end devices. On receiving path check request message, beacon end devices will respond with path check response message. If path check response message is not received within configured timeout period, the end device that sent the path check request message will declare fault on its active port and will switch over to inactive port, provided a double fault has not been detected.

The MFTP beacon end devices also behave in same way to detect connection timeouts and to send failure notification messages. When beacon end devices receive failure notification messages themselves, they will verify their own transmit path failures by sending path check request messages to a preconfigured set of end devices. On receiving path check request message, end devices will respond with path check response message. If path check response message is not received within configured timeout period, the beacon device that sent the path check request message will declare fault on its active port and will switch over to inactive port, provided a double fault has not been detected.

Following sub-sections provide different single fault examples that will aid in understanding how different multiple faults examples will be handled by MFTP end devices.

Single Fault Example 1

Figure 2 shows single fault example 1. Assume that all fault tolerant end devices have port 1 active that is connected to network infrastructure A, prior to fault occurring. Since the fault is on immediate link connected to end device, it will be detected immediately through IEEE 802.3 link fault detection mechanism and the end device will reconfigure its active port 1 to inactive and inactive port 2 to active, with a total delay from fault occurring to reconfiguration of 10 microseconds or less. Immediately, this end device would also transmit multicast learning update message on network through newly active port 2 to update learning in network switches. Subsequently, the network will operate with the said device communicating through infrastructure B, while other fault tolerant devices continuing to communicate through infrastructure A.

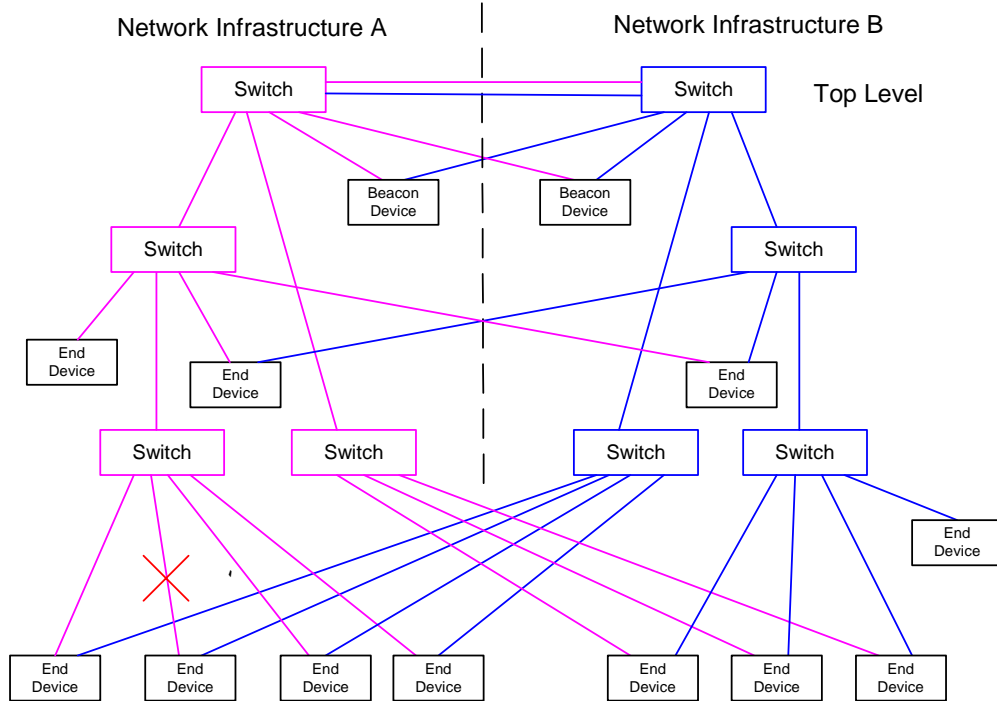


Figure 2: Single Fault Example 1

Single Fault Example 2

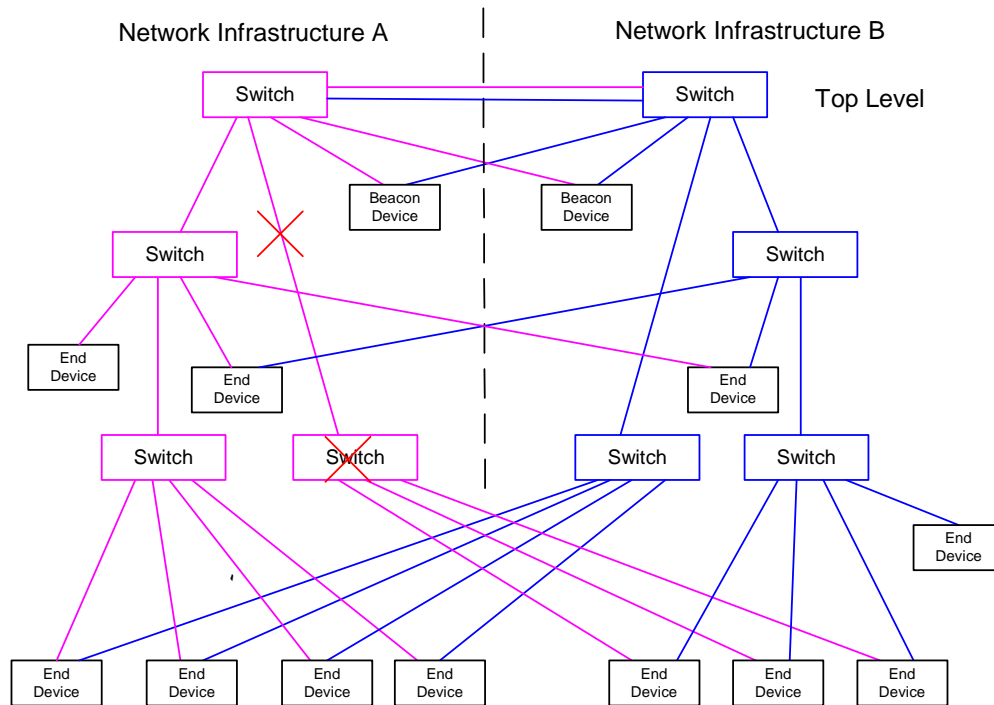


Figure 3: Single Fault Example 2

Figure 3 shows single fault example 2. In this case, the fault may have occurred on switch or on its uplink. Assume that all fault tolerant end devices have port 1 active that is connected to network infrastructure A, prior to fault occurring. If the fault occurred on switch and is detectable on the Physical layer, the three affected downstream end devices will detect it immediately will reconfigure their ports within a period of 10 microseconds or less. If the fault occurred on switch but didn't manifest on the Physical layer, or if the fault occurred on uplink, the three affected end devices will fail to receive beacon messages within beacon timeout period. Consequently, those end devices will reconfigure their ports in less than 1 millisecond. Immediately, they would also transmit learning update message through their newly active port 2 to update learning in network switches. Subsequently, the network will operate with the said devices communicating through infrastructure B, while other fault tolerant devices continuing to communicate through infrastructure A.

Single Fault Example 3

Figure 4 shows single fault example 3. In this case fault may have occurred on top level backbone switch of network infrastructure A or on one of aggregated links connecting top level backbone switches. If the fault occurred on one of aggregated links, it will have minimal impact on network operation. Some packets may be lost for brief period, but the system operation will continue with remaining link(s).

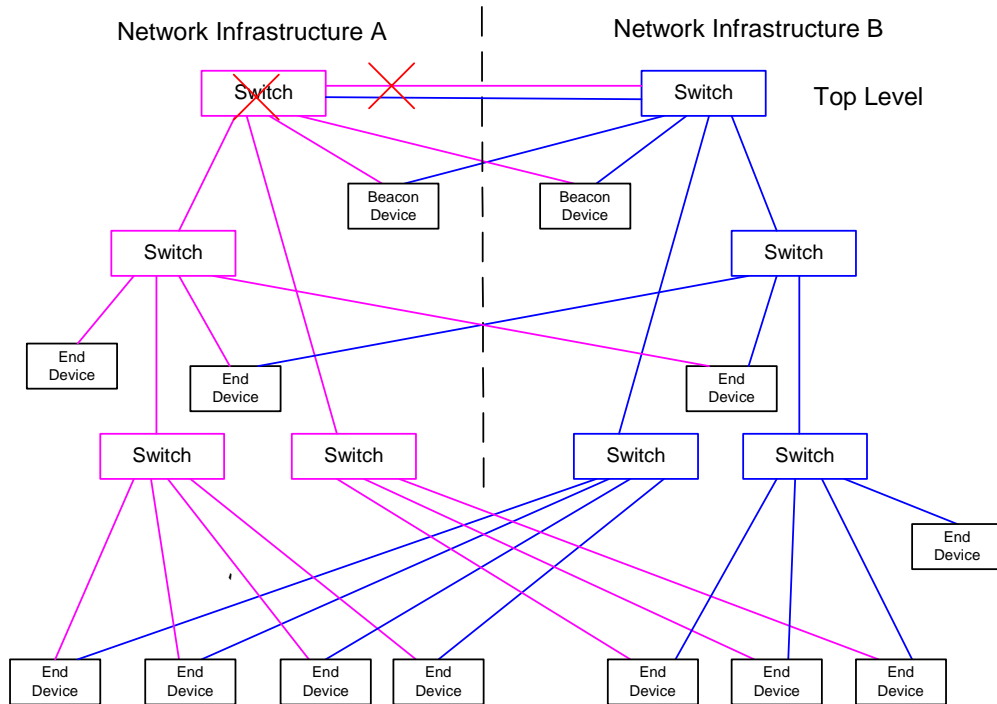


Figure 4: Single Fault Example 3

If the fault occurred on top level backbone switch of network infrastructure A, then all MFTP end devices including beacon end devices will switch over to infrastructure B. The system will continue to operate, so long as a double fault doesn't occur in infrastructure B.

Single Fault Example 4

Figure 5 shows single fault example 4. In this case, a fault occurred on a link connecting beacon end device to backbone switch or a fault occurred on one of beacon end devices. Assume that both beacon devices have port 1 active that is connected to network infrastructure A, prior to fault occurring. If the fault is on immediate link connected to beacon end device, it will be detected immediately and the beacon device will reconfigure its ports. Subsequently, the network will operate with the said beacon end device communicating through infrastructure B, while other end devices continuing to communicate through network infrastructure A. It should be noted that one beacon packet may get lost during this process, but it will have no effect on system operation.

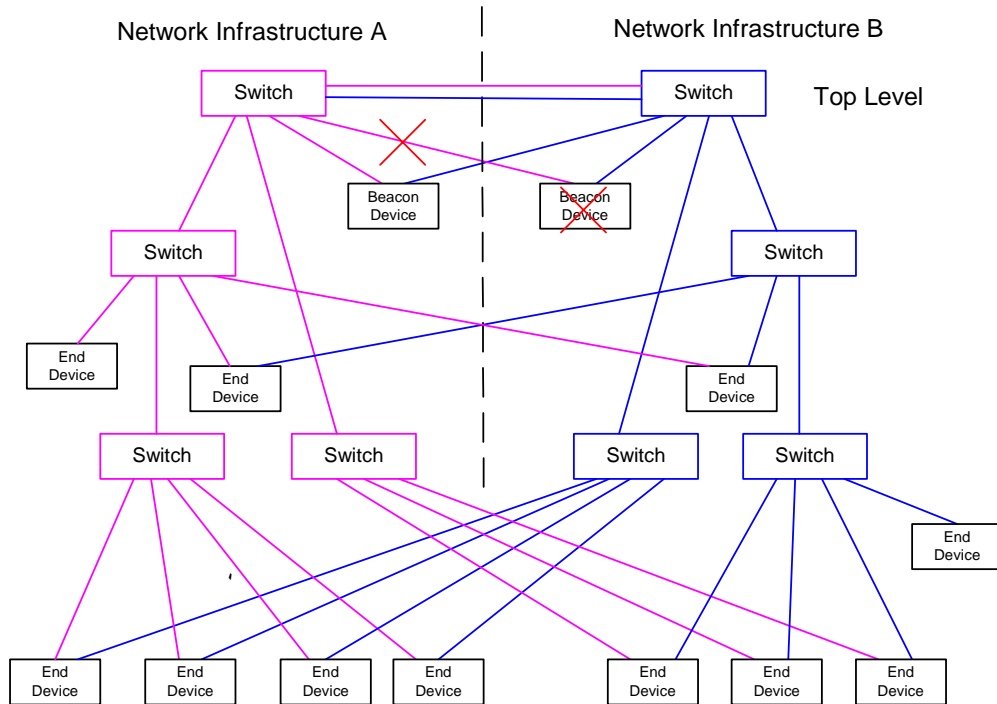


Figure 5: Single Fault Example 4

If the fault occurred on a beacon end device, since the other beacon end device is still active, the system will continue to operate without any problems.

Multiple Faults Example

Figure 6 shows one possible combination of multiple single faults. It should be clear from the discussion thus far how this case will be handled. After reconfiguration some devices will use infrastructure A, while other will use infrastructure B.

As can be observed, MFTP networks will tolerate all single faults and all possible combinations of multiple single faults. It should be noted that a 2-way redundant system is not capable of tolerating a double fault, where a double fault is two or more single faults that occur within same sub-tree. To handle double faults, a 3 or more way redundant system is required.

It is possible to apply the MFTP protocol to 3, 4 or N-way redundancy in general. For example to realize 3-way redundancy, end devices shall be provided with three ports connected to three sets of infrastructures. Three Beacon devices shall be connected to top level backbone switches and the backbone switches shall be connected with each other using multiple links in full mesh topology. Such a 3-way redundant system will tolerate all single faults, all double faults and all combinations of single and double faults.

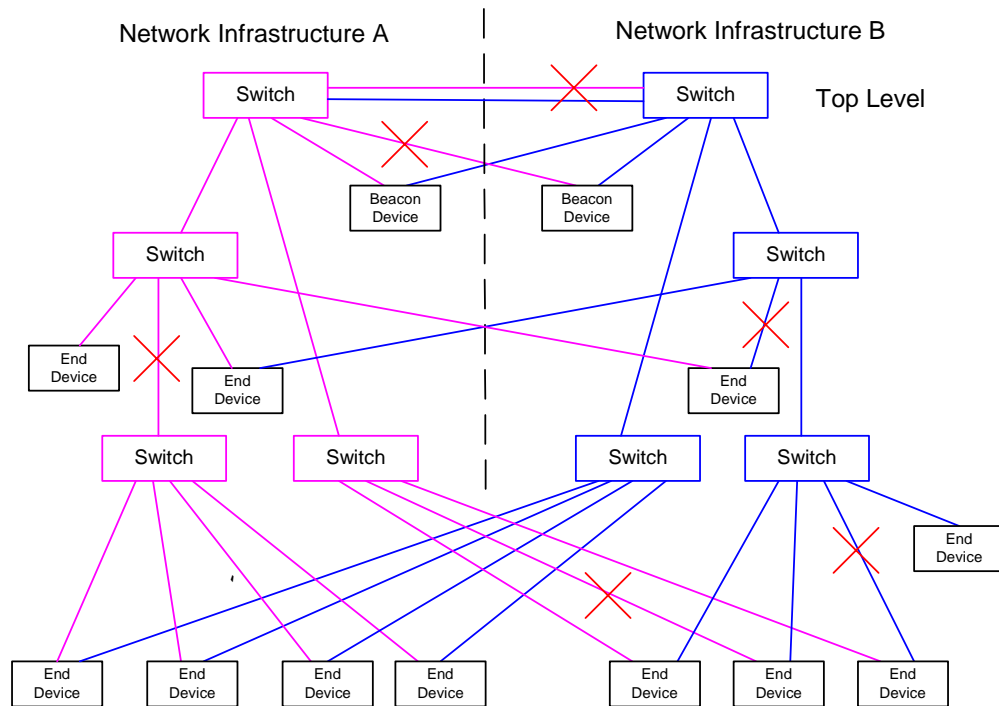


Figure 6: Multiple Faults Example

Conclusions

The MFTP protocol is suitable for EtherNet/IP networks. MFTP networks will tolerate multiple single faults in network infrastructure and network interfaces on end devices. The MFTP protocol allows realization of very large networks with 10,000 end devices or more and with worst case fault detection and switchover time of 1 ms or less for most types of faults. The low switchover time allows utilization of MFTP network in hard real time control systems. MFTP networks can be used to realize highly reliable control systems.

References

- [1] Honeywell Sales Literature PN-04-018-ENG, "Fault Tolerant Ethernet Delivers Robust Networking Solution for Experion PKS", Honeywell International Inc., Oct. 2004.
- [2] IEEE 802.3, Part 3, "Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications", 2005.

DeviceNet, DeviceNet Safety, CIP, CIP Motion, CompoNet, CIP Safety and CIP Sync are trademarks of ODVA. EtherNet/IP is a trademark of ControlNet International under license by ODVA. Other trademarks are property of their respective owners.

The ideas, opinions and recommendations expressed herein are intended to describe concepts of the author(s) for the possible use of CIP Networks and do not reflect the ideas, opinions, and recommendation of ODVA per se. Because CIP Networks may be applied in many diverse situations and in conjunction with products and systems from multiple vendors, the reader and those responsible for specifying CIP Networks must determine for themselves its suitability and the suitability of ideas, opinions, and recommendations expressed herein for intended use.

Copyright ©2007 Open DeviceNet Vendor Association, Inc. (ODVA). All rights reserved.
For permission to reproduce excerpts of this material, with appropriate attribution to the author(s), please contact ODVA on:

TEL	+1 734-975-8840
FAX	+1 734-922-0027
EMAIL	odva@odva.org
WEB	www.odva.org