Dependability Entering Mainstream IT Networking Standards (IEEE 802.1)

64th Meeting of the IFIP 10.4 Working Group on
Dependable Computing and Fault Tolerance

Visegrád, Hungary, June 27-30, 2013

Wilfried Steiner, Corporate Scientist
wilfried.steiner@tttech.com

Background

IEEE 802.1 is standardizing general architectures for local area networks (LANs) and metropolitan area architectures (MANs).

Together with IEEE 802.3 they are the main working groups working standards for Ethernet switches.

Efficient utilization of the communication bandwidth and plug-and-play capabilities are topmost requirements in IEEE 802.1.

With AVB, IEEE 802.1 moved into the area of real-time communication.

With TSN, IEEE 802.1 moves into the area of dependable communication.

Upcoming mainstream IT equipment aims to provide real-time and dependable communication features (to a significant higher degree than today).

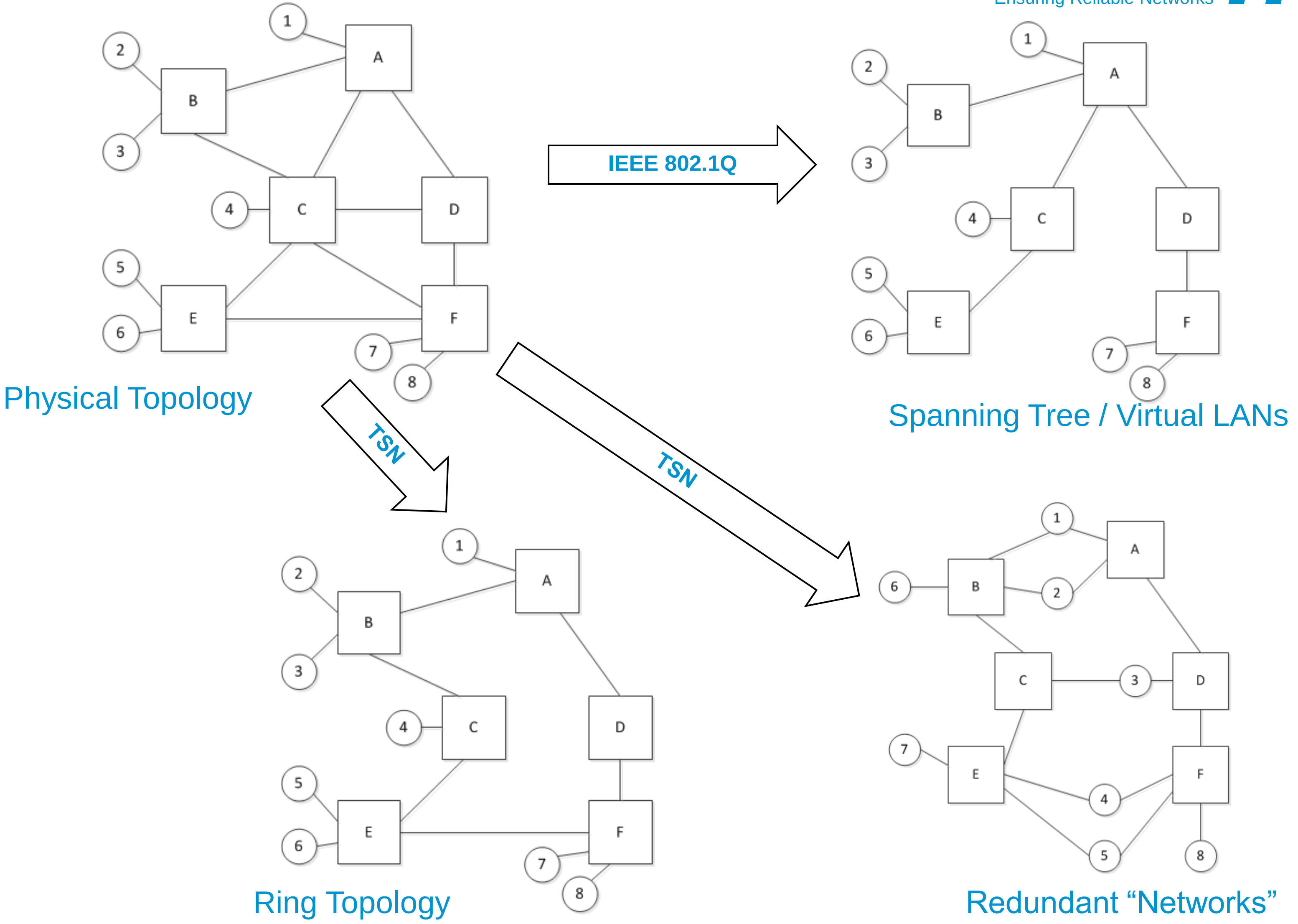
AVB – Audio/Video Bridging

802.1AS Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks: a protocol and technique to synchronize local clocks in the network to each other.

802.1Qat Stream Reservation Protocol (SRP): a protocol that allows applications to dynamically reserve bandwidth in the network.

802.1Qav Forwarding and Queuing Enhancements for Time-Sensitive Streams: an enhancement over strict priority based forwarding and queueing mechanisms that establishes fairness properties for lower priority traffic in the network.

802.1BA: definition of profiles for AVB systems.



Physical Topology

IEEE 802.1Q

Spanning Tree / Virtual LANs

TSN

TSN

Ring Topology

Redundant “Networks”

TSN – Time-Sensitive Networks

802.1ASbt Timing and Synchronization: Enhancements and Performance Improvements

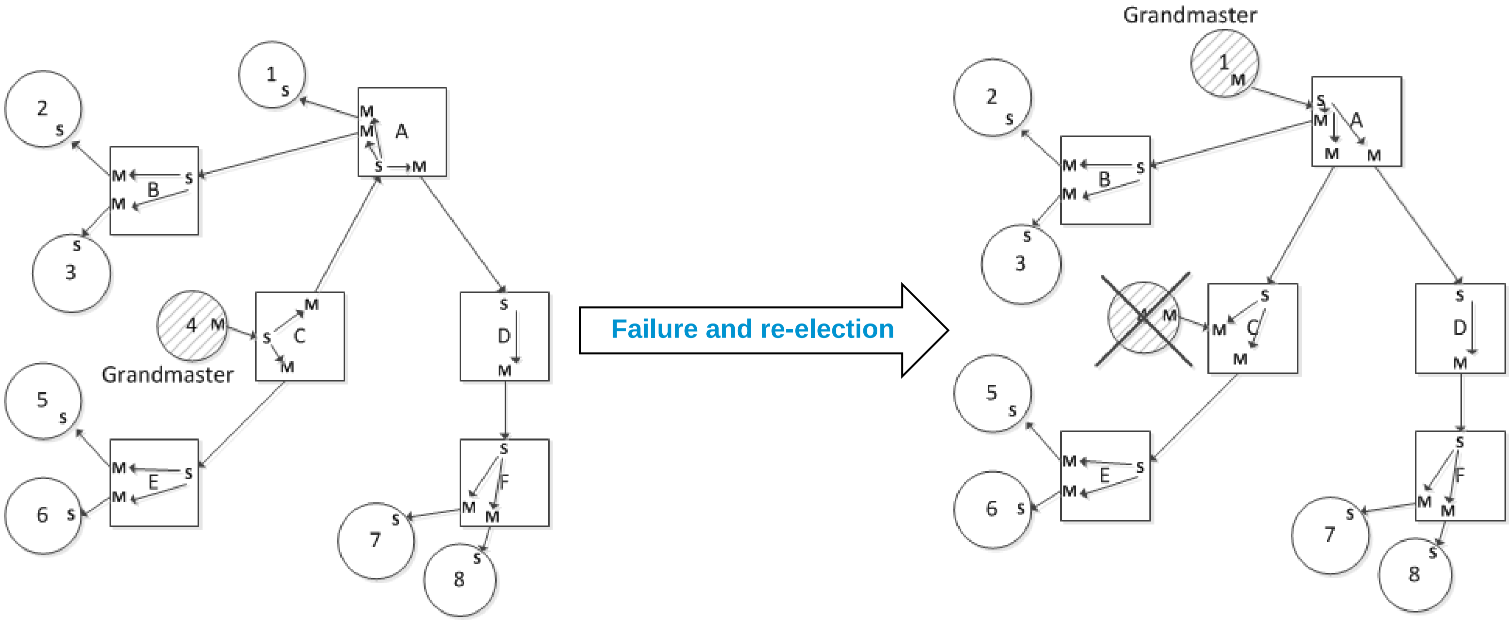
802.1Qbv Enhancements for Scheduled Traffic: a basic form of time-triggered communication

802.1Qbu Frame Preemption: a mechanism that allows to preempt a frame in transmit to intersperse another frame.

802.1Qca Path Control and Reservation: protocols and mechanisms to set up and manage the redundant communication paths in the network.

802.1CB Frame Replication and Elimination for Reliability: to eliminate redundant copies of frames transmitted over the redundant paths setup in 802.1Qca.

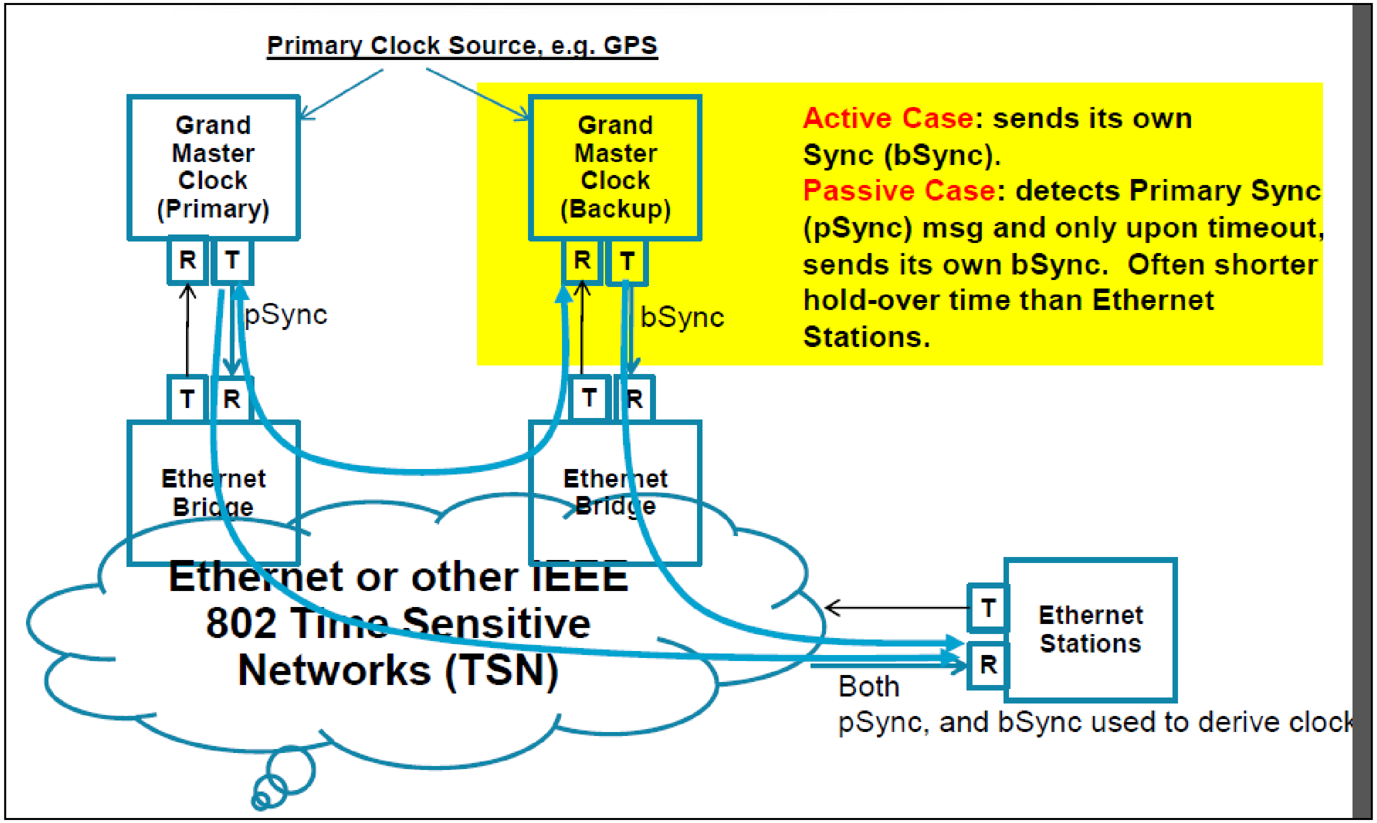
IEEE 802.1AS Clock Synchronization



Failure and re-election

The clock synchronization protocol is a classical master-slave protocol.
The master is called the “grandmaster”.
When the grandmaster fails, then a new grandmaster is elected.
Issues with this mechanism have been reported by industry.

802.1ASbt Clock Synchronization Proposals for Improvements


Primary Clock Source, e.g. GPS
Grand Master Clock (Primary)
R
T
pSync
T
R
Ethernet Bridge
Grand Master Clock (Backup)
R
T
bSync
T
R
Ethernet Bridge
Ethernet or other IEEE 802 Time Sensitive Networks (TSN)
T
R
Ethernet Stations

Active Case: sends its own Sync (bSync).
Passive Case: detects Primary Sync (pSync) msg and only upon timeout, sends its own bSync. Often shorter hold-over time than Ethernet Stations.

Both
pSync, and bSync used to derive clock

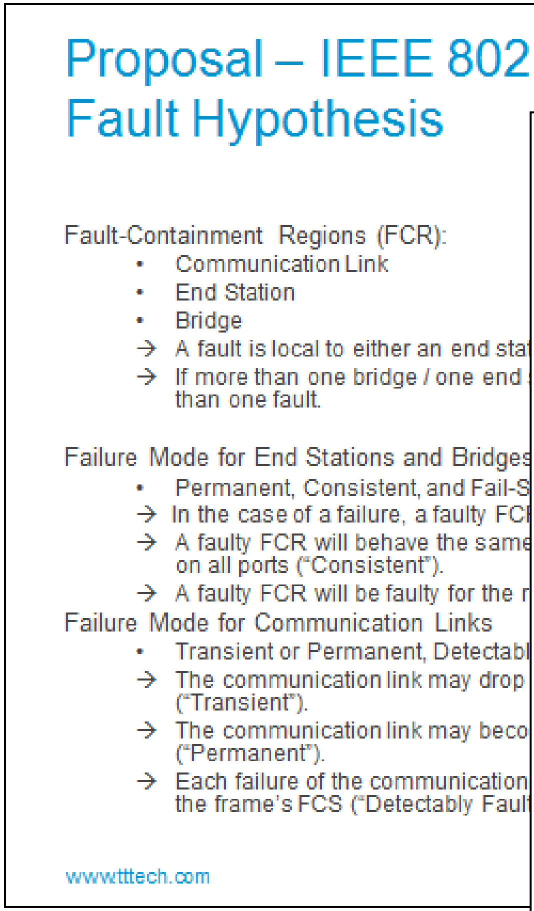
http://www.ieee802.org/1/files/public/docs2013/ASbt-Spada-Kim-Fault-tolerant-grand-master-proposal-0513-v1.pdf

Does this sound familiar to anyone?

It seems like industry has been there XX years ago.

Raise awareness in IEEE 802.1 of dependability communities …

Ensuring Reliable Networks TTTech



Proposal – IEEE 802.1Q AVB Fault Hypothesis

Fault-Containment Regions (FCR):
- Communication Link
- End Station
- Bridge

→ A fault is local to either an end stat
→ If more than one bridge / one end than one fault.

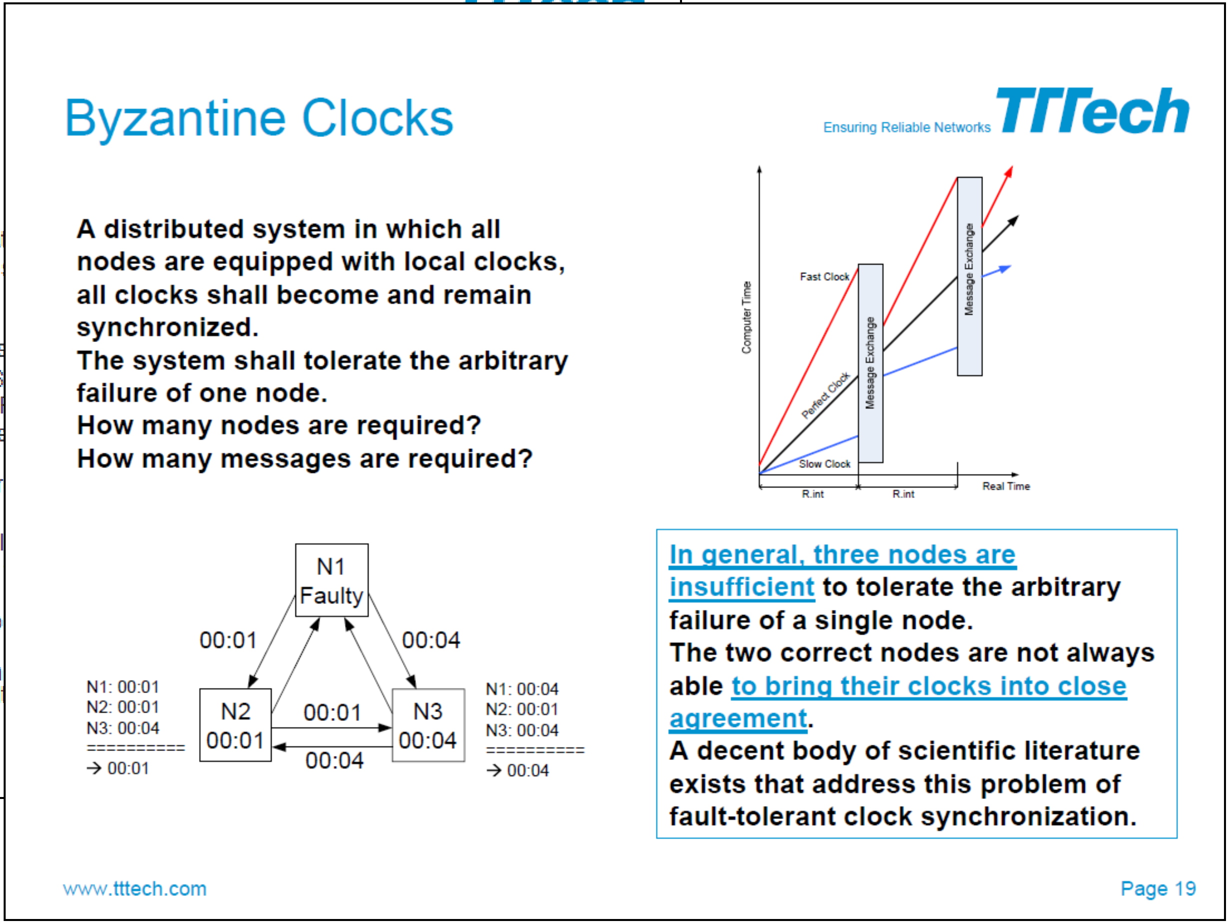
Failure Mode for End Stations and Bridges
- Permanent, Consistent, and Fail-S

→ In the case of a failure, a faulty FCR
→ A faulty FCR will behave the same on all ports ("Consistent").
→ A faulty FCR will be faulty for the r

Failure Mode for Communication Links
- Transient or Permanent, Detectabl

→ The communication link may drop ("Transient").
→ The communication link may beco ("Permanent").
→ Each failure of the communication the frame's FCS ("Detectably Faul

www.tttech.com



Byzantine Clocks

Ensuring Reliable Networks TTTech

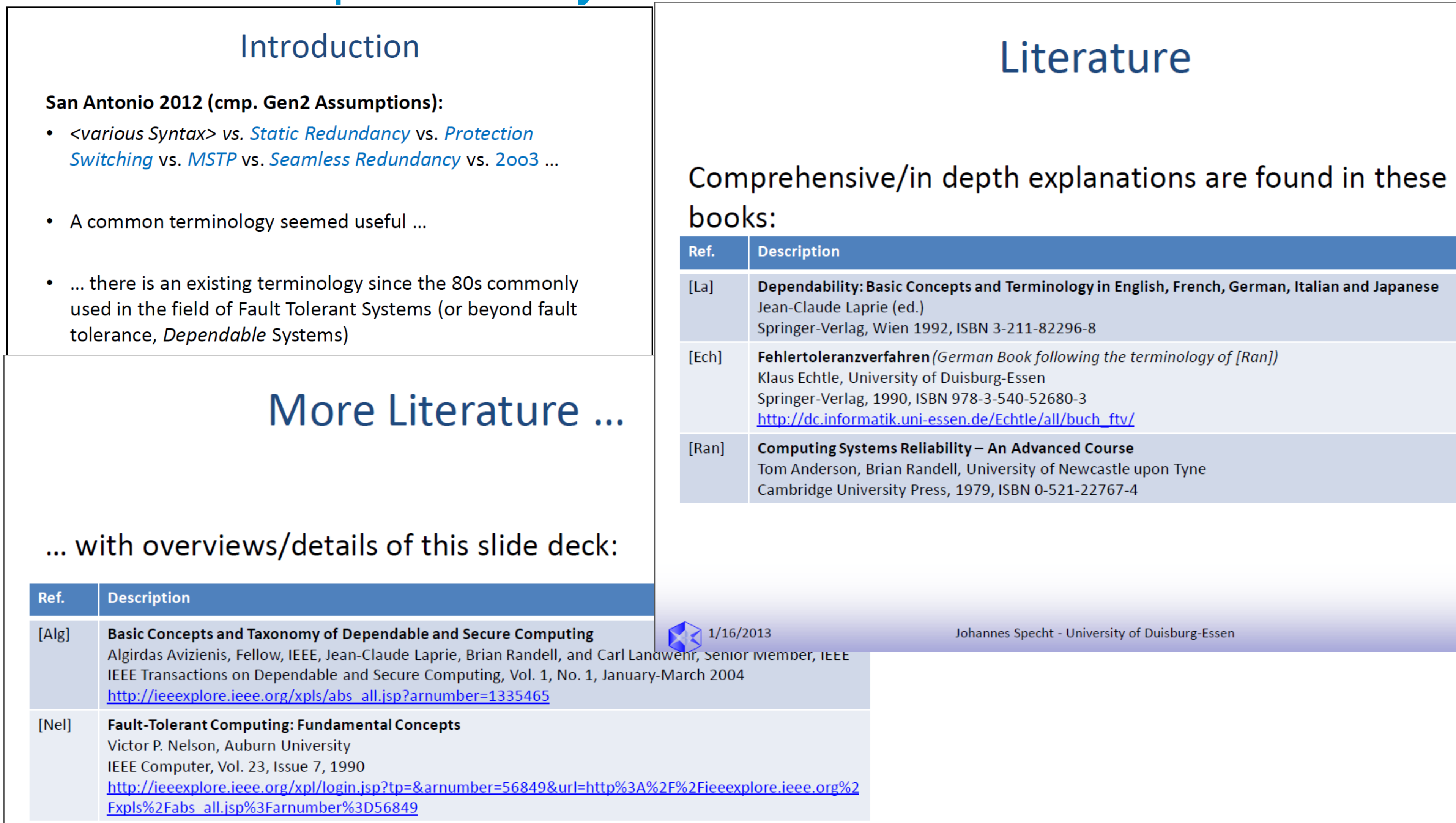
A distributed system in which all nodes are equipped with local clocks, all clocks shall become and remain synchronized.
The system shall tolerate the arbitrary failure of one node.
How many nodes are required?
How many messages are required?

In general, three nodes are insufficient to tolerate the arbitrary failure of a single node.
The two correct nodes are not always able to bring their clocks into close agreement.
A decent body of scientific literature exists that address this problem of fault-tolerant clock synchronization.

www.tttech.com Page 19

http://www.ieee802.org/1/files/public/docs2012/new-avb-wsteiner-fault-tolerant-clock-synchronization-0112-v01.pdf

… core dependability research …



Introduction

San Antonio 2012 (cmp. Gen2 Assumptions):

- *<various Syntax> vs. Static Redundancy* vs. *Protection Switching* vs. *MSTP* vs. *Seamless Redundancy* vs. 2oo3 …
- A common terminology seemed useful …
- … there is an existing terminology since the 80s commonly used in the field of Fault Tolerant Systems (or beyond fault tolerance, *Dependable* Systems)

Literature

Comprehensive/in depth explanations are found in these books:

Ref.	Description
[La]	**Dependability: Basic Concepts and Terminology in English, French, German, Italian and Japanese** Jean-Claude Laprie (ed.) Springer-Verlag, Wien 1992, ISBN 3-211-82296-8
[Ech]	**Fehlertoleranzverfahren** *(German Book following the terminology of [Ran])* Klaus Echtle, University of Duisburg-Essen Springer-Verlag, 1990, ISBN 978-3-540-52680-3 http://dc.informatik.uni-essen.de/Echtle/all/buch_ftv/
[Ran]	**Computing Systems Reliability – An Advanced Course** Tom Anderson, Brian Randell, University of Newcastle upon Tyne Cambridge University Press, 1979, ISBN 0-521-22767-4

1/16/2013 Johannes Specht - University of Duisburg-Essen

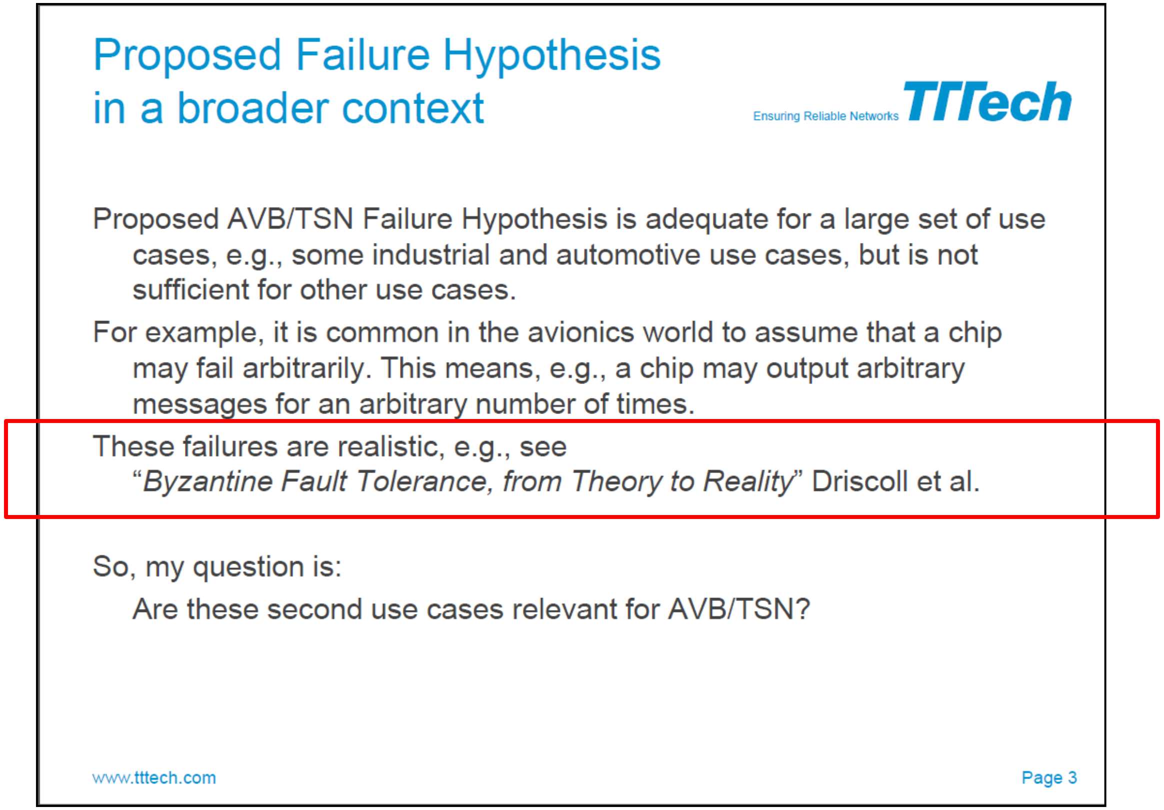
More Literature …

… with overviews/details of this slide deck:

Ref.	Description
[Alg]	**Basic Concepts and Taxonomy of Dependable and Secure Computing** Algirdas Avizienis, Fellow, IEEE, Jean-Claude Laprie, Brian Randell, and Carl Landwehr, Senior Member, IEEE IEEE Transactions on Dependable and Secure Computing, Vol. 1, No. 1, January-March 2004 http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1335465
[Nel]	**Fault-Tolerant Computing: Fundamental Concepts** Victor P. Nelson, Auburn University IEEE Computer, Vol. 23, Issue 7, 1990 http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=56849&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D56849

http://www.ieee802.org/1/files/public/docs2013/new-tsn-specht-redundancy-terminology-20130115-v01.pdf

… and the practical relevance.



Proposed Failure Hypothesis in a broader context

Ensuring Reliable Networks TTTech

Proposed AVB/TSN Failure Hypothesis is adequate for a large set of use cases, e.g., some industrial and automotive use cases, but is not sufficient for other use cases.

For example, it is common in the avionics world to assume that a chip may fail arbitrarily. This means, e.g., a chip may output arbitrary messages for an arbitrary number of times.

These failures are realistic, e.g., see "*Byzantine Fault Tolerance, from Theory to Reality*" Driscoll et al.

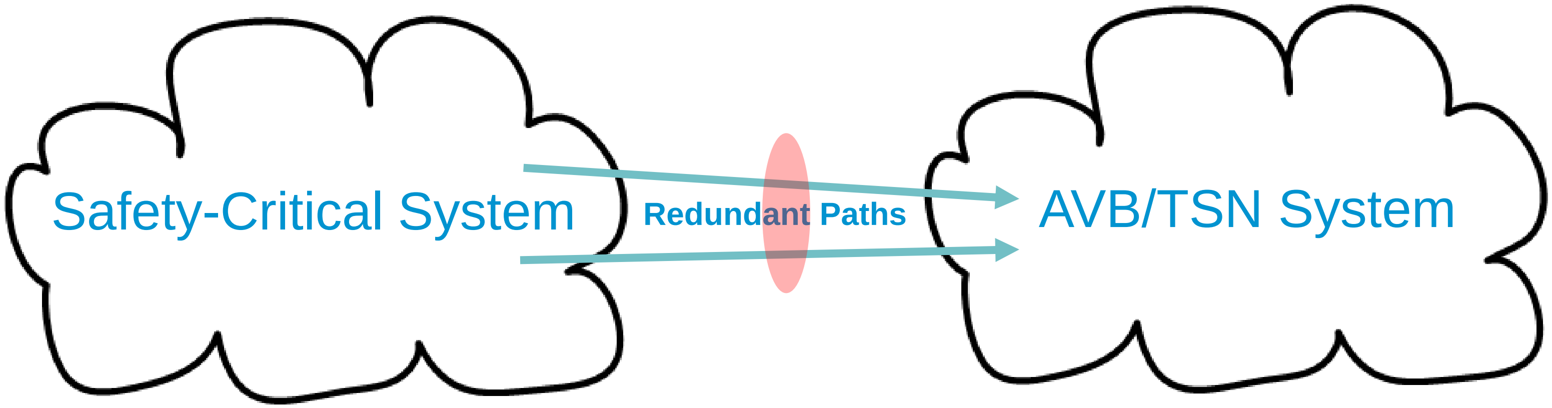
So, my question is:

Are these second use cases relevant for AVB/TSN?

www.tttech.com Page 3

http://www.ieee802.org/1/files/public/docs2013/new-avb-wsteiner-8021AS-interoperability-ft-clocksync-0514-v01.pdf

Synchronization of AVB/TSN to a Safety-Critical System



How about to standardize the quality of the clock synchronization messages (e.g., announce and synchronization messages) from redundant masters?

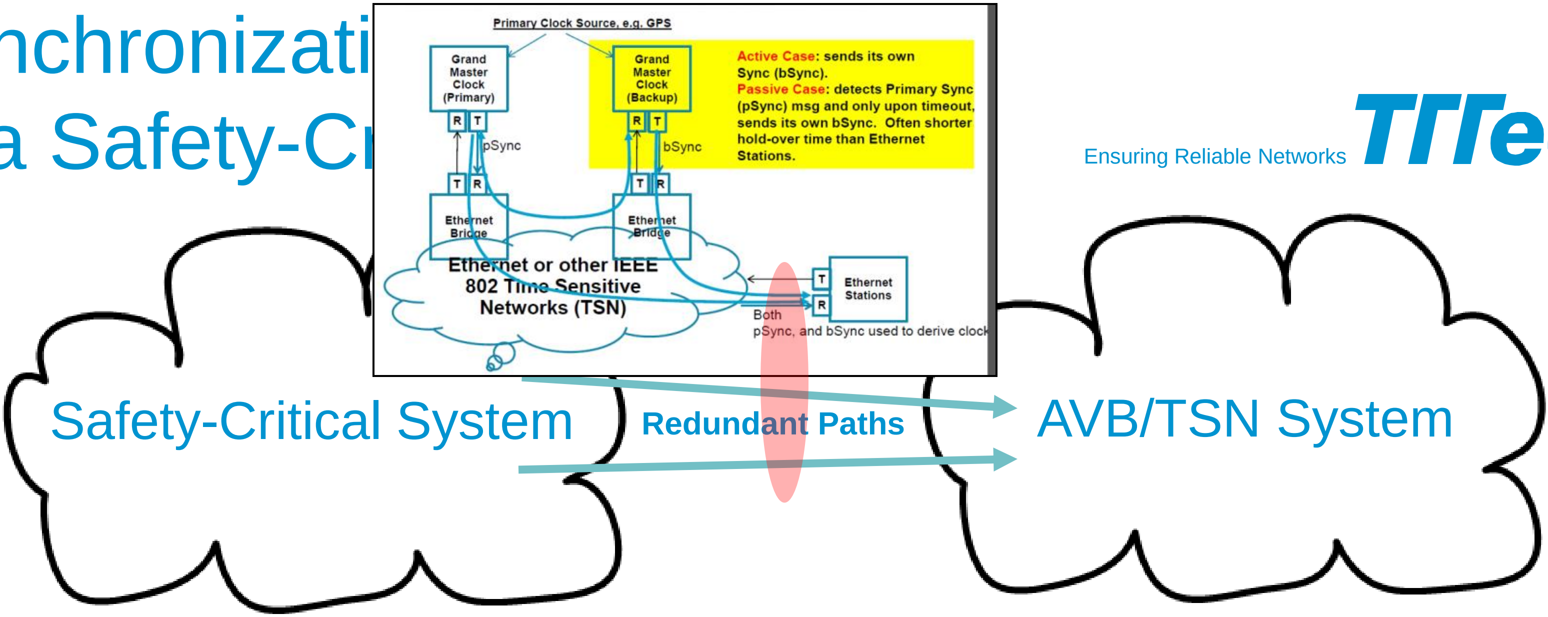
Relevant quality parameters in the fault-tolerant clock synchronization domain are, e.g.:

- Failure modes of the clock synchronization inputs to the AVB/TSN domain
- Worst-case temporal deviation of two non-faulty clocks in the system (precision)
- Temporal communication blackout characteristics
- Mean Time To Repair (MTTR)

The IEEE 1588 alternative master mechanism may be usable as a basis.

Synchronizati
to a Safety-Cr

Ensuring Reliable Networks TTTech


Primary Clock Source, e.g. GPS
Grand Master Clock (Primary)
Grand Master Clock (Backup)
R
T
pSync
bSync
T
R
Ethernet Bridge
Ethernet Bridge
Ethernet or other IEEE 802 Time Sensitive Networks (TSN)
Ethernet Stations
Both
pSync, and bSync used to derive clock

Active Case: sends its own Sync (bSync).
Passive Case: detects Primary Sync (pSync) msg and only upon timeout, sends its own bSync. Often shorter hold-over time than Ethernet Stations.

Safety-Critical System

Redundant Paths

AVB/TSN System

How about to standardize the quality of the clock synchronization messages (e.g., announce and synchronization messages) from redundant masters?

Relevant quality parameters in the fault-tolerant clock synchronization domain are, e.g.:

- Failure modes of the clock synchronization inputs to the AVB/TSN domain
- Worst-case temporal deviation of two non-faulty clocks in the system (precision)
- Temporal communication blackout characteristics
- Mean Time To Repair (MTTR)

The IEEE 1588 alternative master mechanism may be usable as a basis.

Use well-understood sync protocols

TTEthernet Executable Formal Specification

- Using symbolic and bounded model checkers *sal-smc* and *sal-bmc*
- Focus on Interoperation of Synchronization Services (Startup, Restart, Clique Detection, Clique Resolution, abstract Clock Synchronization)

Verification of Lower-Level Synchronization Functions

- Permanence Function (*sal-inf-bmc* + k-induction)
- Compression Function *(sal-inf-bmc + k-induction)*

Formal Verification of Clock Synchronization Algorithm

- First time by means of Model Checking (*sal-inf-bmc* + k-induction)

Re-use of the Formal Models to prove:

- Layered clock-rate correction algorithm (*sal-inf-bmc* + k-induction)
- Layered clock-diagnosis algorithm (*sal-inf-bmc* + k-induction)

Verification and minor corrections of the "Sparse Timebase" Concept

- Distributed computations without explicit coordination (PVS)

Work has mostly been done in the context of the Marie Curie CoMMiCS project

FP7 (FP7/2007-2013) project no. 236701

CoMMiCS

Conclusions

Dependability enters mainstream networking equipment (IEEE 802.1).

The integration of the new functionality in the existing framework is non-trivial.

- We know pretty well how to verify individual protocols on their own, but the verification of a tightly integrated set of protocols is still challenging.

The IEEE 802.1 WG not necessarily will perform dependability benchmarking and an exhaustive analysis of these measures.

Hence, a designer of a dependable system faces two issues when using the standardized IEEE mechanisms:

1. How dependable are these mechanisms and under what constraints is dependability guaranteed?
2. Are these mechanisms sufficient and how can they be improved if necessary?

We have solutions to these issues.

TTTech

Ensuring Reliable Networks

www.tttech.com