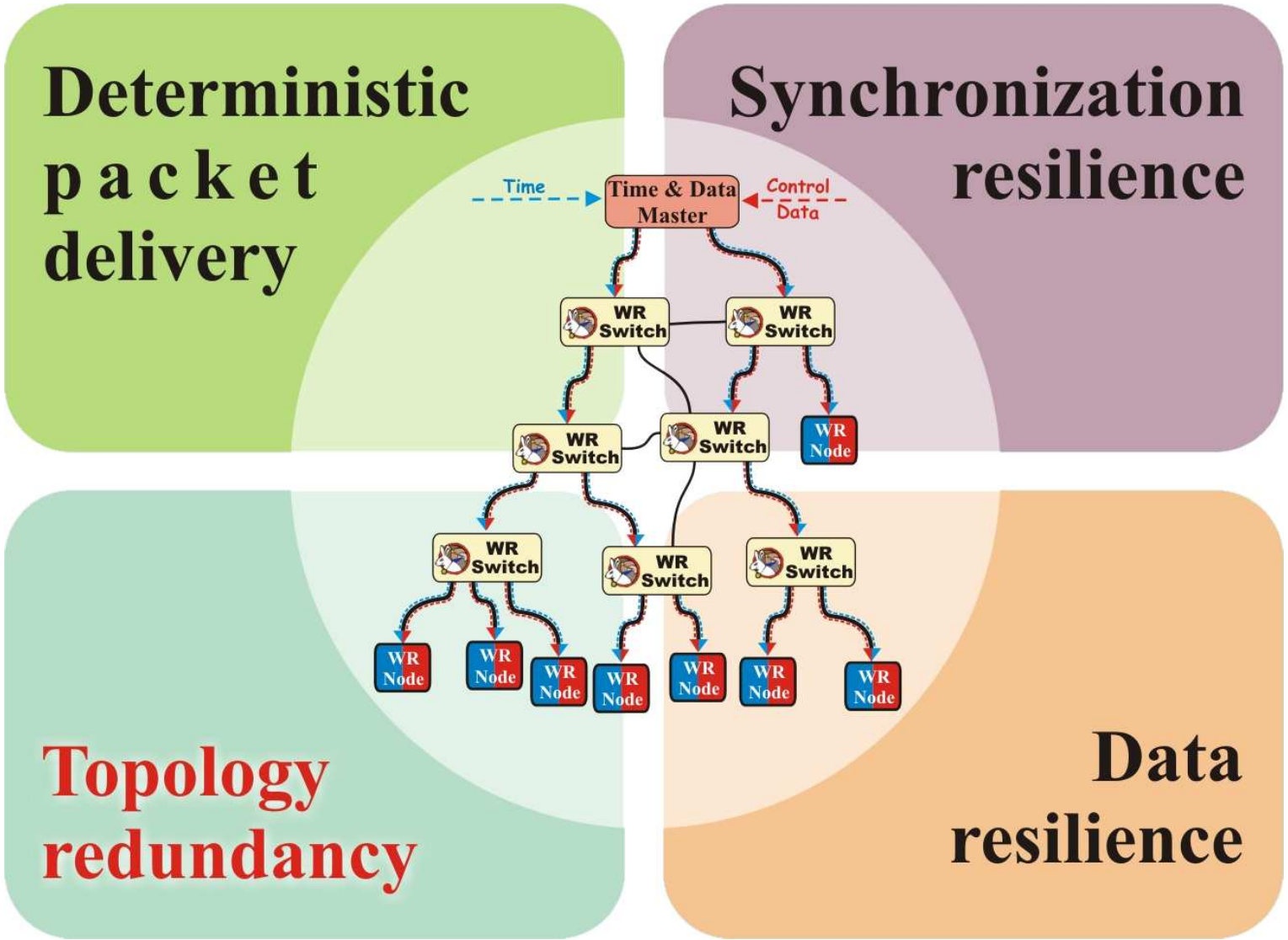
White Rabbit Network
Status and Plans

Maciej Lipiński
Hardware and Timing Section
CERN

Cesar Prados
Bel Division
GSI

March 2012

White Rabbit Network

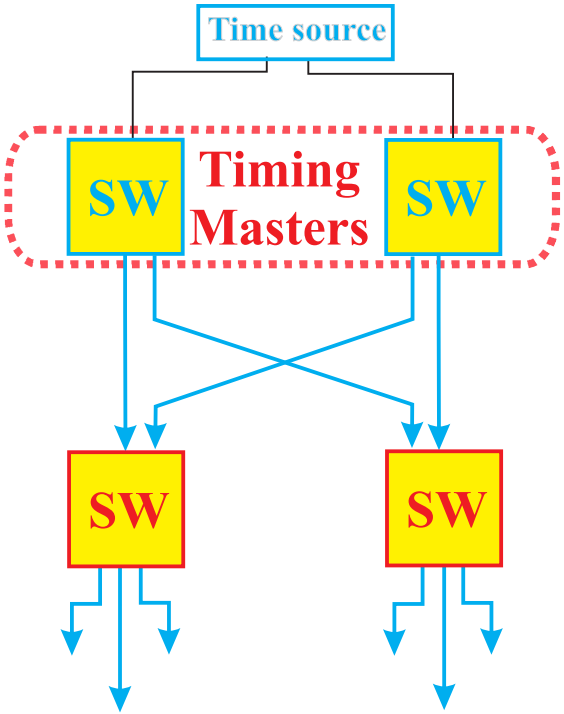


Outline

1 Introduction
2 Synchronization
3 Determinism
4 Data
5 Topology
6 Networks
7 Status and Plans

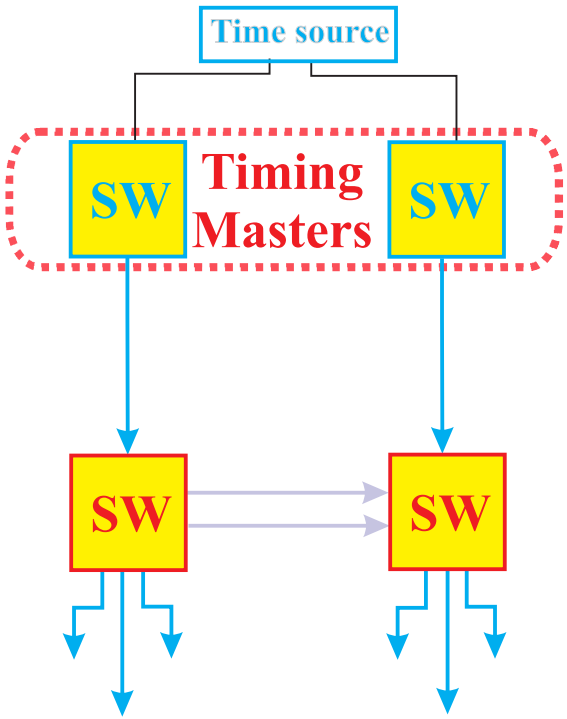
Time Distribution

- Network-wide **sub-ns** synchronization
- Constant performance in:
  - **stable** conditions
  - **transient** conditions
- Support for network redundancy
  - WRPTP ✓ (possibly further improve)
  - Hardware



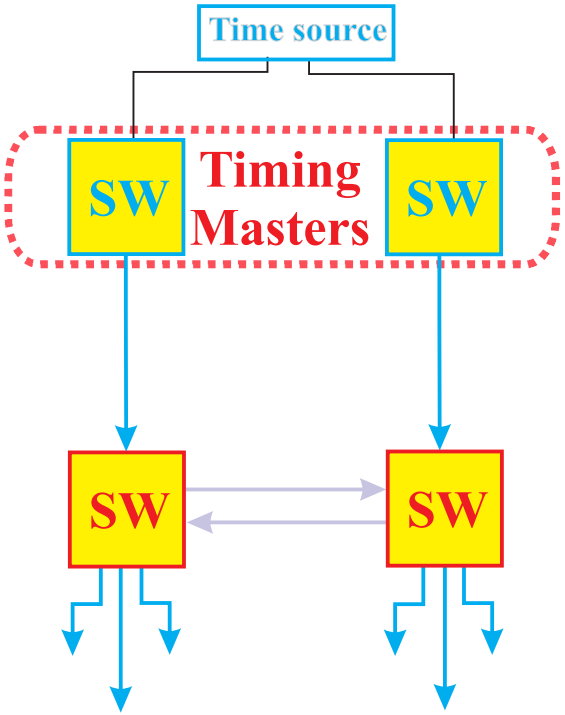
Time Distribution

- Network-wide **sub-ns** synchronization
- Constant performance in:
  - **stable** conditions
  - **transient** conditions
- Support for network redundancy
  - WRPTP ✓ (possibly further improve)
  - Hardware



Time Distribution

- Network-wide **sub-ns** synchronization
- Constant performance in:
  - **stable** conditions
  - **transient** conditions
- Support for network redundancy
  - WRPTP ✓ (possibly further improve)
  - Hardware



Deterministic data delivery

Special Traffic (optional)

- Priority 7 (broadcast)
- Quality of Service (QoS)
- Determinism of data delivery

Normal Traffic

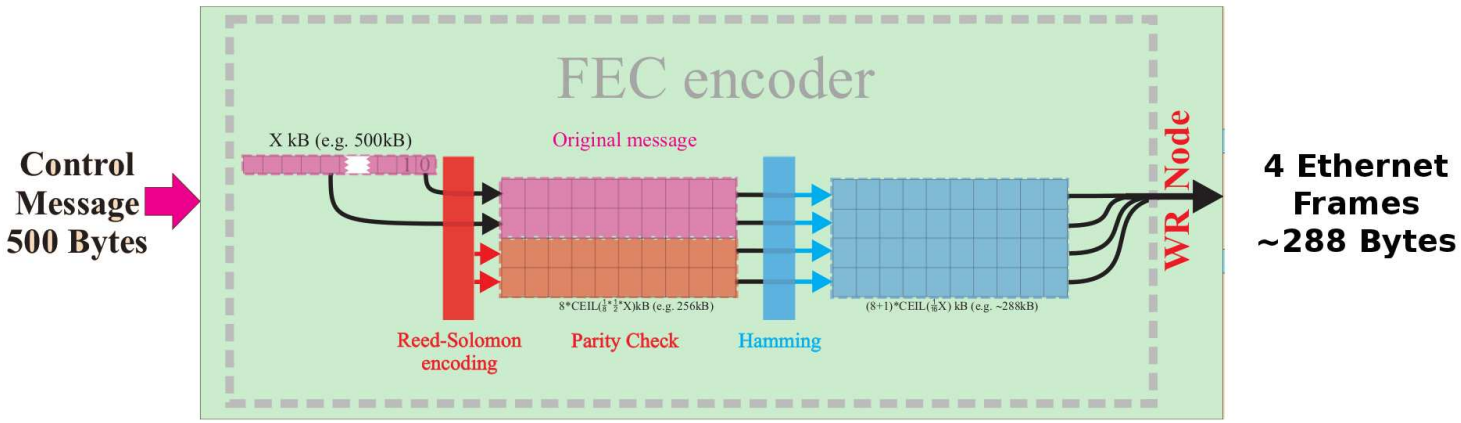
- Any non priority 7 traffic
- Less stringent determinism of data delivery

Quality of Service (WR Switch)

- Cut-through ✓
- Separate resources for Special Traffic
- Output queuing
- Switch firmware optimization (SWcore + RTU)

Forward Error Correction

- Well-researched, concepts agreed
- Encoder ✓ not optimized for resources
- Decoder (MSc thesis):
  - Hardware-accelerated LM32
  - Eventually to be encoder-decoder

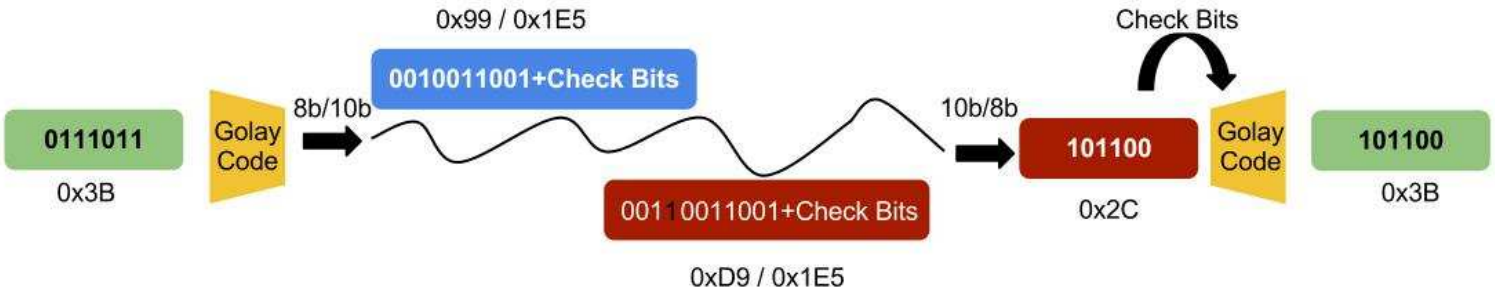


FEC: Linear Code and 8b/10b I

- The FEC scheme is **Linear Code** for the Binary Erasure Channel and a **Block Code** for the Packet Erasure Channel.
- The numbers told us that we need to repair just a bit... Hamming was the option.
- Gigabit Ethernet uses **8b/10b** in the physical layer.
- One bit error in a 10 bits encoded code word may result in a multi-bit error $> 2 \rightarrow$ **Hamming is not enough!**.
- The best Linear Code able to recover 4 bit error and detect up to 8 is the **Golay Code Family**.

FEC: Linear Code and 8b/10b II

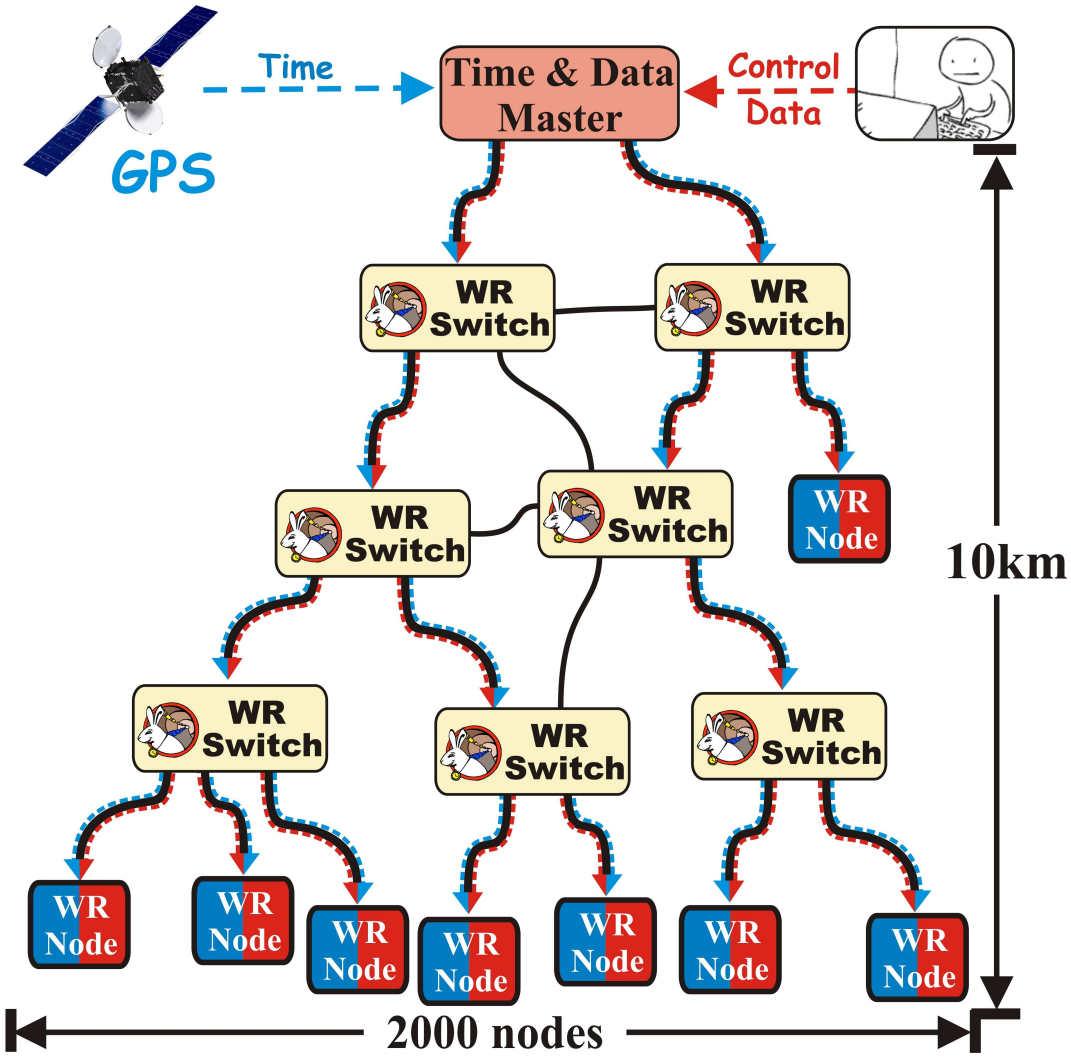
An bit error in the cable could mean more than 2 errors after 10b/8b, One bit error it is not enough!!!!



FEC: Linear Code and 8b/10b III

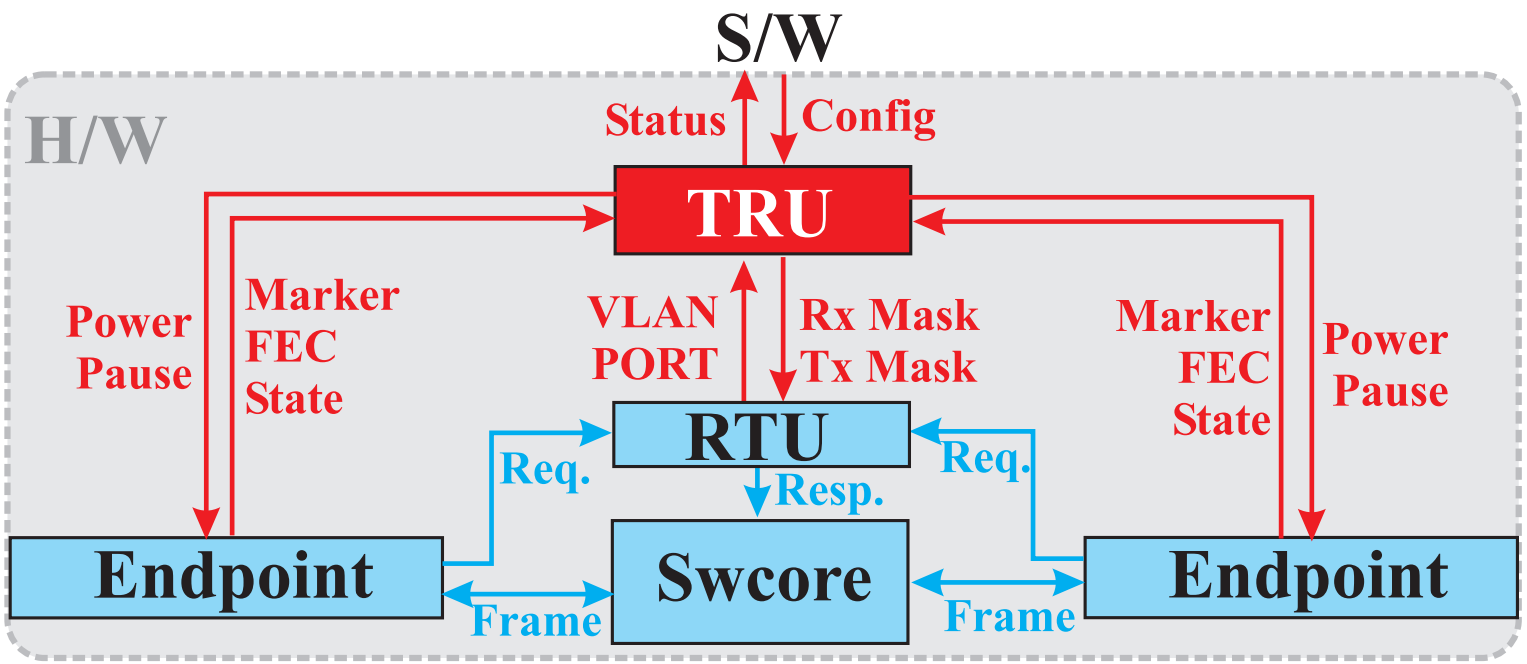
- Golay is a (24,12,8) code, rate of 50%
- **Fixed encoding time** and **upper bound** for the **decoding**
- Hardware implementation of a variant of the Golay Code:
  - Encoding in one clock
  - Decoding, best case 2 clocks, worst case 6 clocks working at 125 Mhz
- Next steps:
  - hook it to the Starter Kit and plus the Block Code, and then we have a FEC.
  - Customization of the code for our case

Topology Redundancy



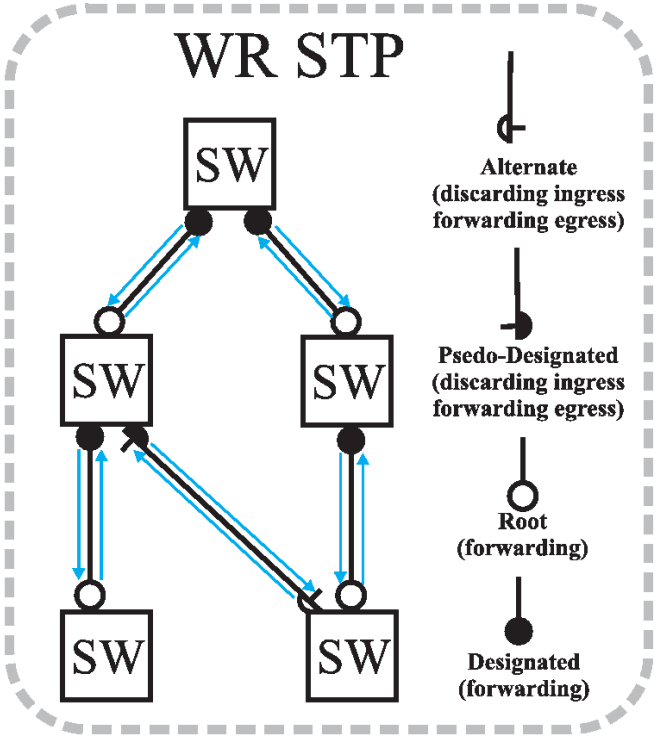
Topology Resolution

- Target: $\leq$ 2 Ethernet Frames lost @ convergence
- Extensive research of current solutions
- Two solutions considered
  - enhanced Spanning Tree Protocol (WR STP)
  - enhanced Link Aggregation Control Protocol (WR LACP)
- **Universal firmware** to support enhancements



enhanced Spanning Tree Protocol

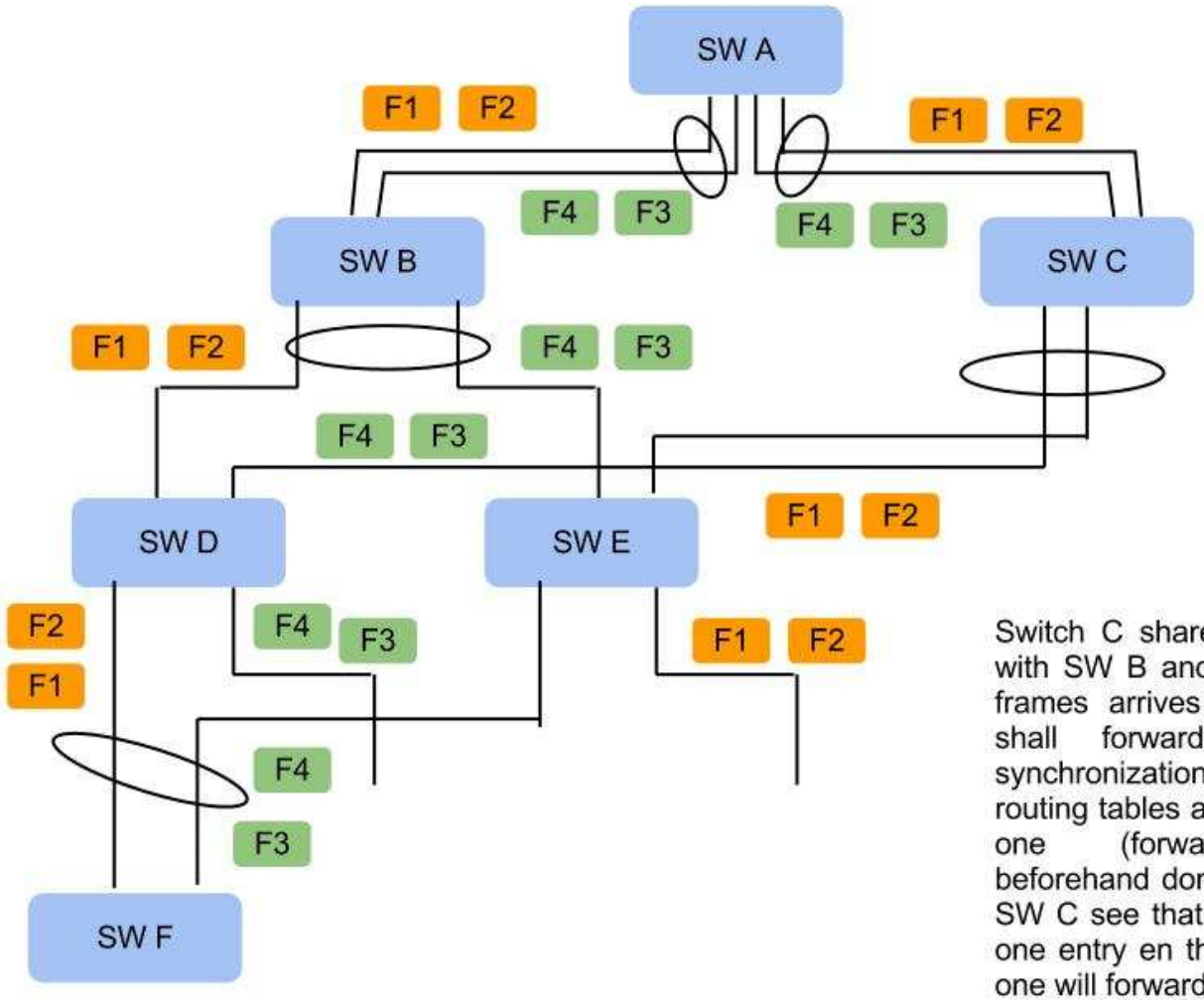
- Ultra-fast convergence ($\approx 3\ \mu s$)
- Modification of Multiple/Rapid STP
  - hardware support
  - software support
- Basic concept ✓
- Implementation



Link Aggregation Control Protocol plus FEC I

- LACP provides the means to show two networks interface as a single interfacea
- When the network interfaces are not in the same physical switch a multipath is created for a frame
- Normally the multipath is used for balancing the traffic
- On the other hand, the FEC destributes the control information in several frames
- What if we send $X\%$ of this frames in one path and $Y\%$ in the alternative path
- We could overcome with sucseed single points of failure in the network: switch breakdown and damaged fiber optic or ports.

Link Aggregation Control Protocol plus FEC II



Switch C shares the routing table with SW B and when a Broadcast frames arrives one of them only shall forward the frame, a synchronization process of the routing tables and the roles of each one (forwarder/standby) is beforehand done. Since SW B and SW C see that the SW D has only one entry en the routing table only one will forward the frame.

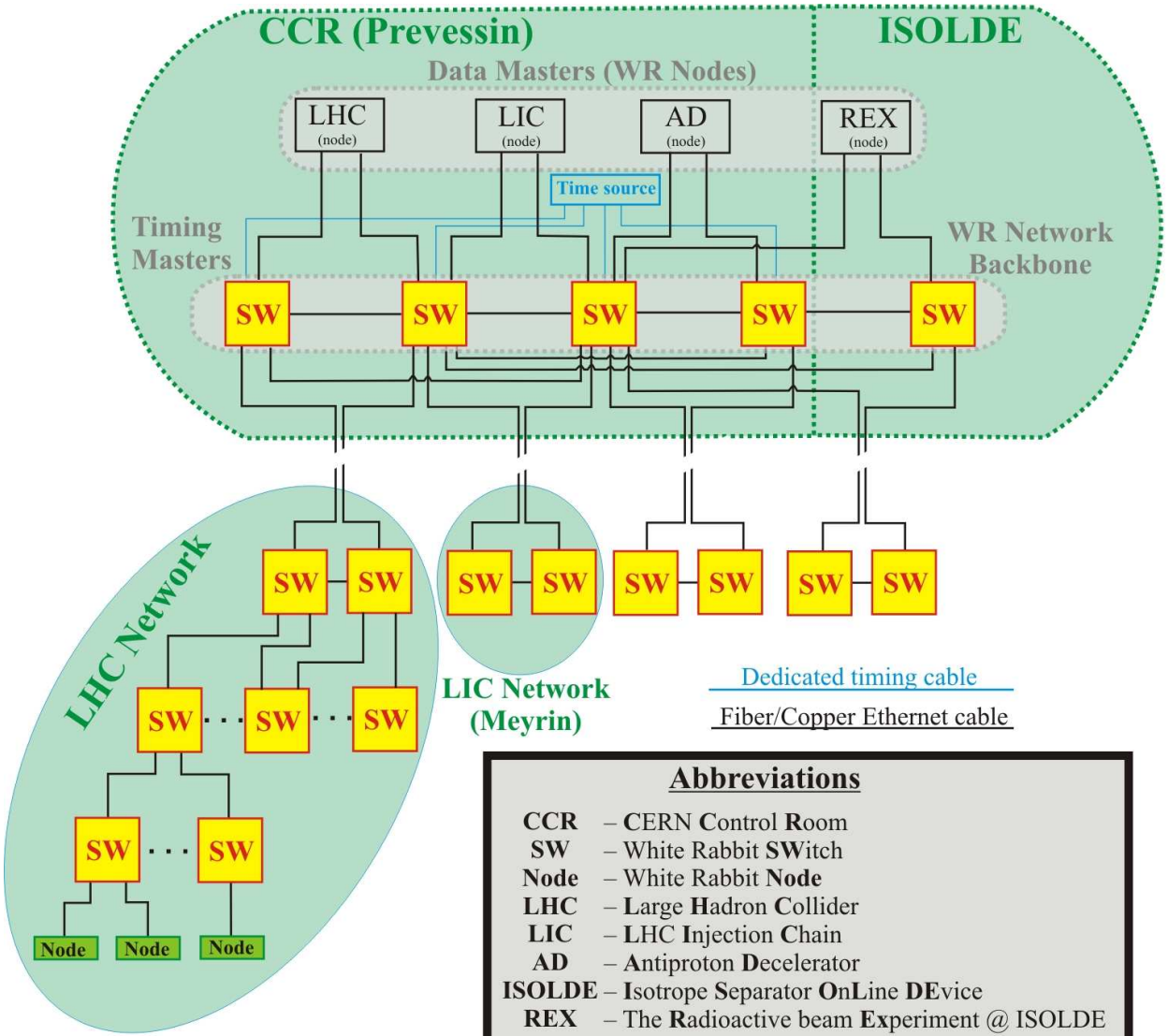
Link Aggregation Control Protocol plus FEC III

- It fullfils the GSI and CERN
- Doesn’t jepardize the deterministic delivery of information,

There is a smoothly continuation of the service there is not rush we have the **redundancy beforhand.**
Pros of LAC-FEC

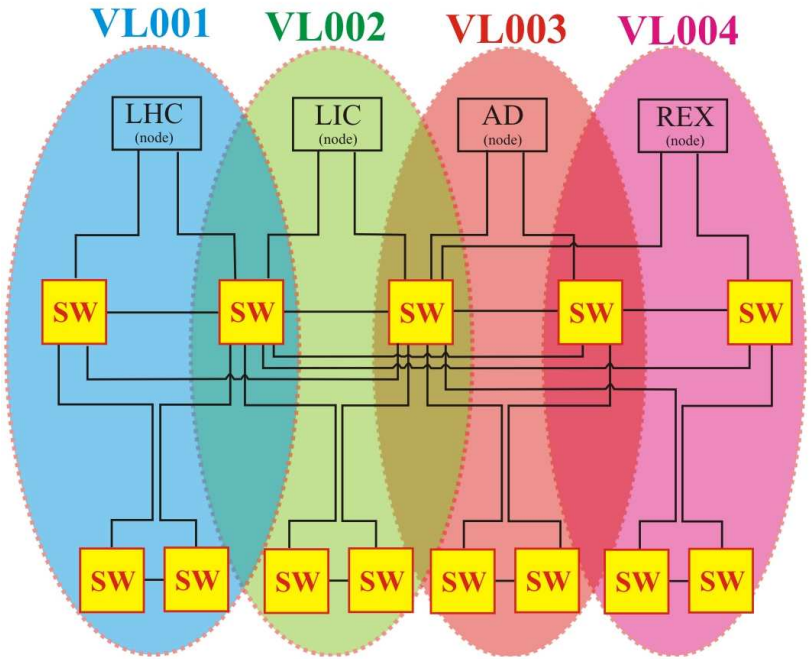
- There is not idle resources: port, fiber optic etc...
- At least duplication of the bandwidth upstream
- The proposed scheme doesn’t break the standard
- The IEEE standard is alive, present in almost all comercial switches (cisco, hp...)

WR@CERN: overview

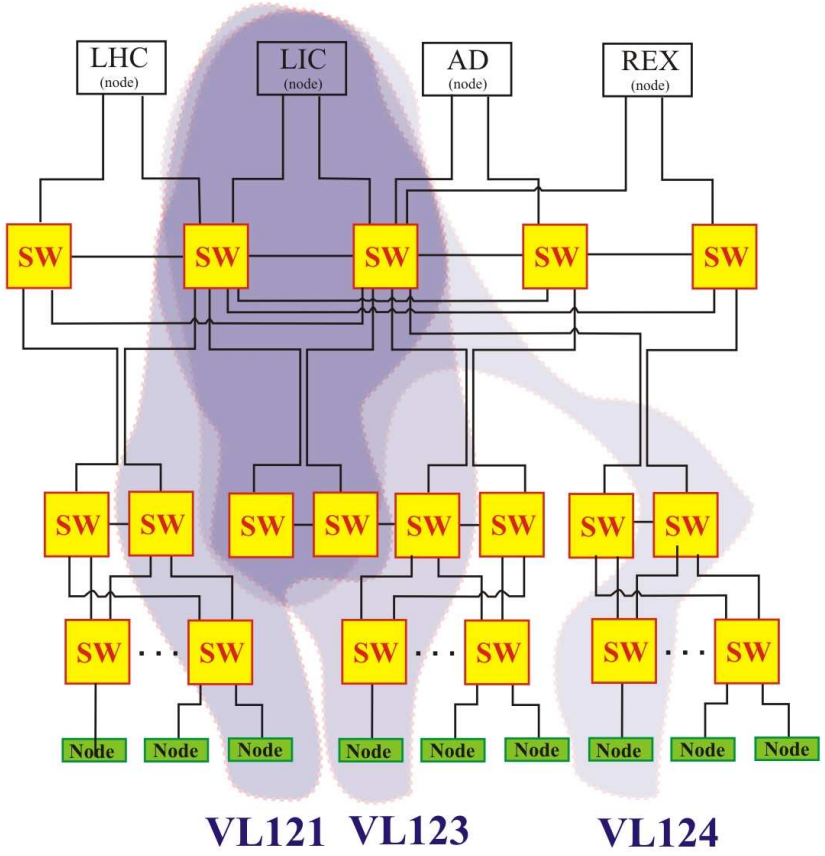


WR@CERN: VLANs

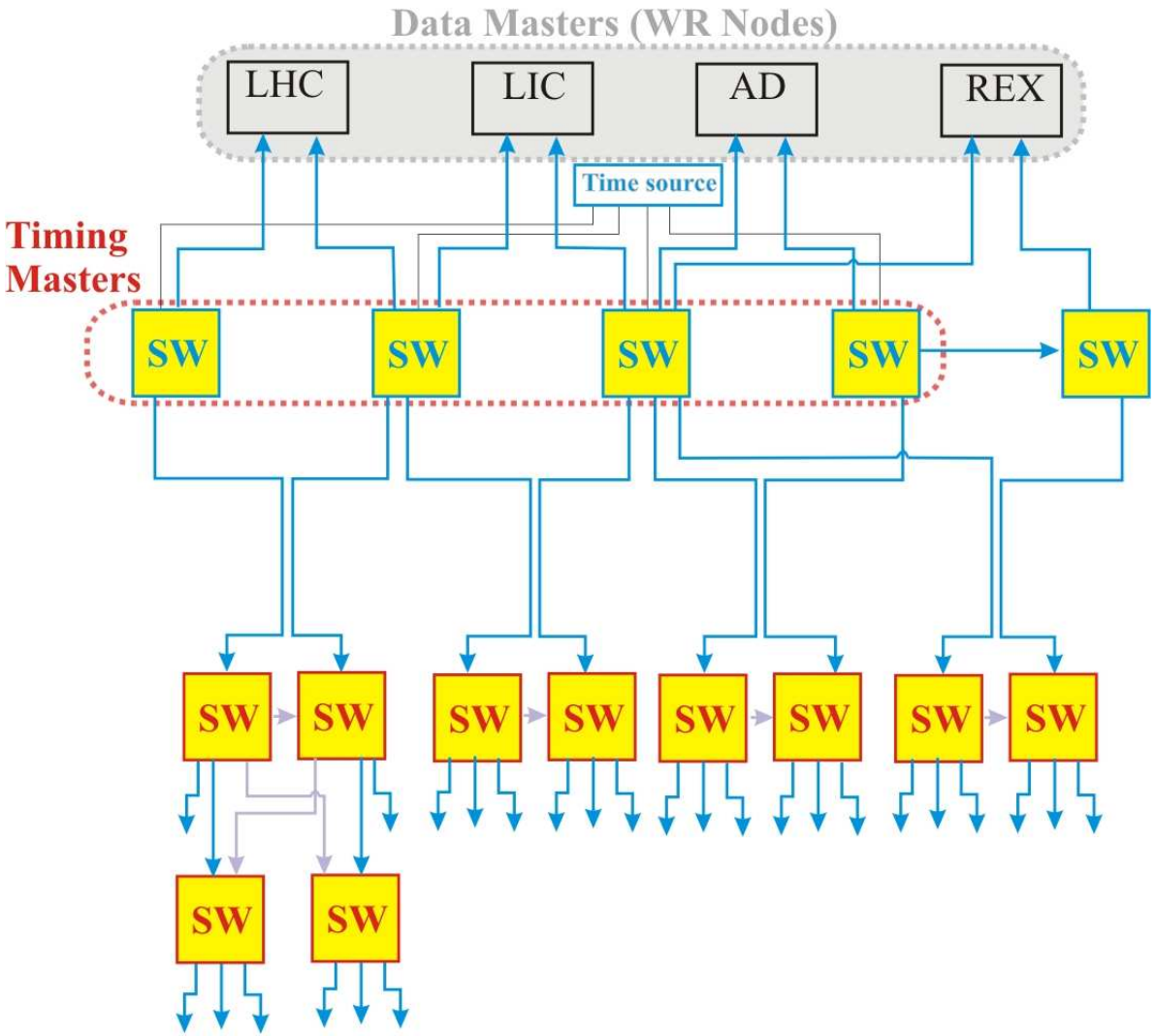
Per-accelerator VLANs



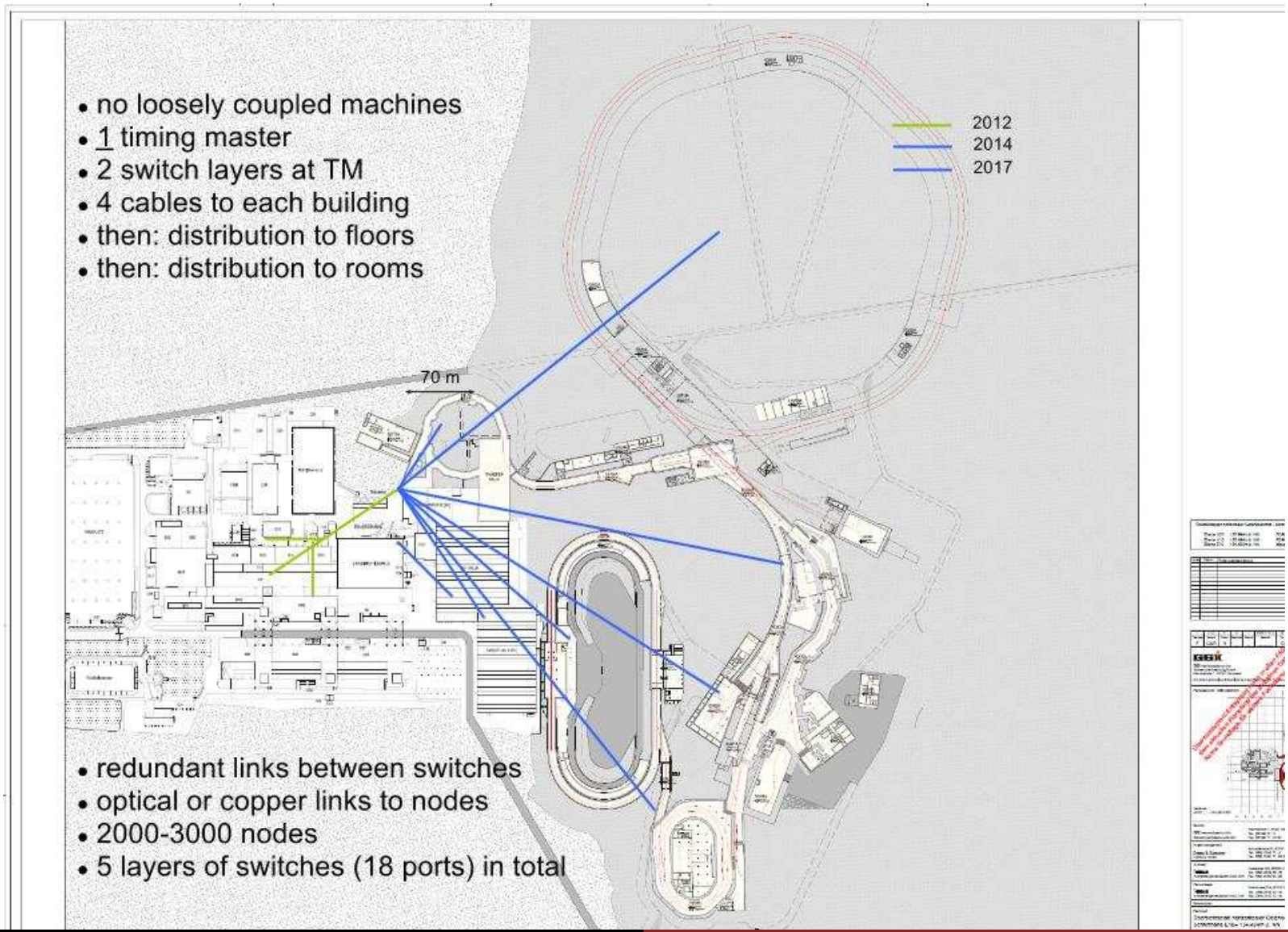
Shared accelerator VLANs



WR@CERN: Time distribution



WR@GSI: Overview



Status

Synchronization Resilience

- ✓ WRPTP support - improvements for further study
- Hardware support

Deterministic Packet Delivery

- ✓ Cut-through
- Separate resources
- Output queuing
- Optimization

Data Resilience

- ✓ FEC Encoder - more work
- FEC Decoder

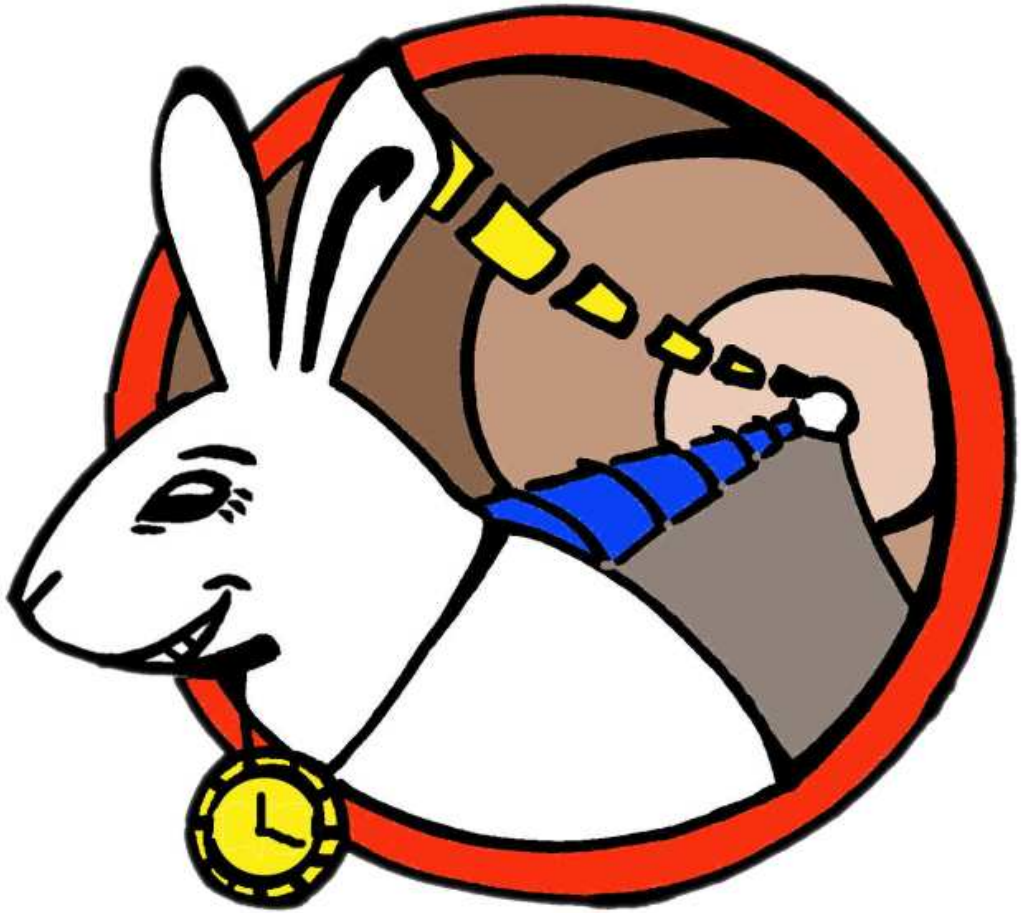
Topology redundancy

- ✓ Extensive study
- Hardware/software

Diagnostics

- Software

Plans (for the nearest future)

- **Synchronization Resilience**
  - Hardware - implementation of switch-over
- **Deterministic Packet Delivery**
  - Hardware - re-implementation of Memory Management Unit
  - Hardware - optimization of SWcore/RTU
- **Topology redundancy**
  - Hardware - implementation of Topology Resolution Unit
  - Hardware - modification of RTU
  - Software - implementation of daemon (PoC)
- **Diagnostics** - software

Thank you