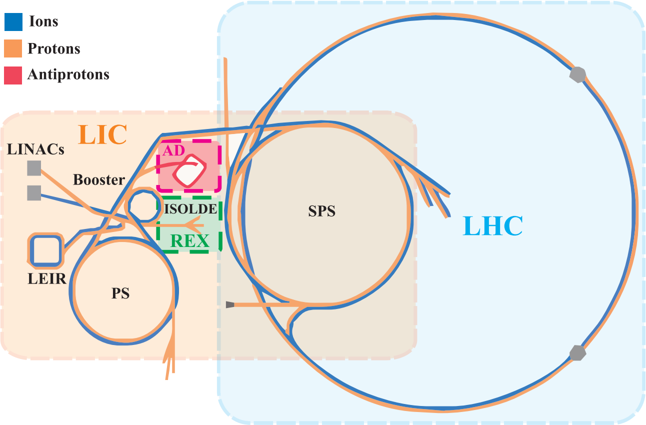
White Rabbit Overview

Maciej Lipiński

Hardware and Timing Section
European Organization for Nuclear Research, CERN

24 November 2017

CERN accelerator complex

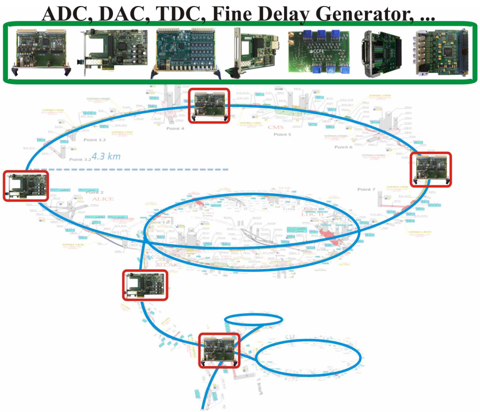


LHC : **L**arge **H**adron **C**ollider
LINAC : **LIN**ear **AC**celerator
PS : **P**roton **S**ynchrotron
SPS : **S**uper **P**roton **S**ynchrotron

LIC : **L**HC **I**njection **C**hain
AD : **A**ntiproton **D**ecelerator
ISOLDE : **I**sotrope **S**eparator **O**n**L**ine **DE**vice
REX : The **R**adioactive beam **Ex**periment@ISOLDE

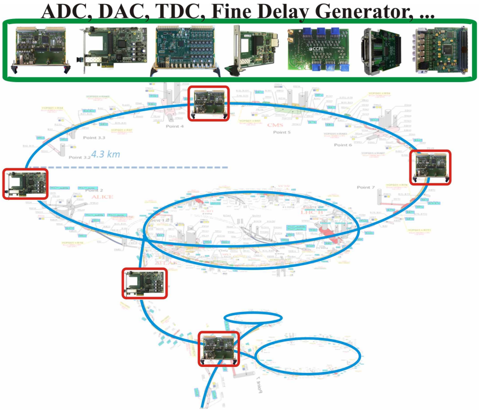
Time-triggered distributed system

- Each accelerator consist of hundreds of subsystems



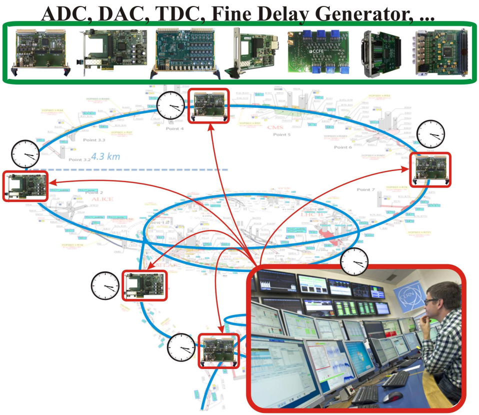
Time-triggered distributed system

- Each accelerator consist of hundreds of subsystems
- Each subsystem performs time-triggered actions



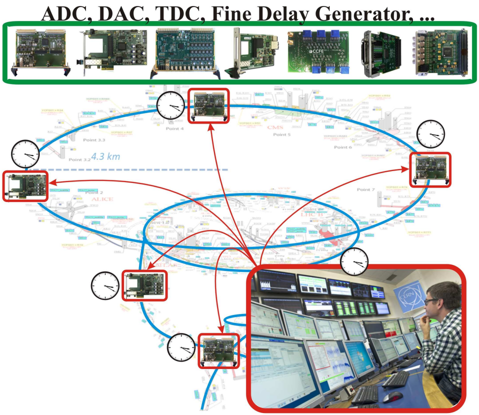
Time-triggered distributed system

- Each accelerator consist of hundreds of subsystems
- Each subsystem performs time-triggered actions
- All actions are orchestrated by the **General Machine Timing**:
  - UTC time
  - Actions for next milisecond



Time-triggered distributed system

- Each accelerator consist of hundreds of subsystems
- Each subsystem performs time-triggered actions
- All actions are orchestrated by the **General Machine Timing**:
  - UTC time
  - Actions for next milisecond
- Accelerator control is based on:
  - Accurate device synchronization
  - Deterministic data transmission



General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008

General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008
- Stringent requirements

Requirement	Value(s)
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs

General Machine Timing renovation – White Rabbit

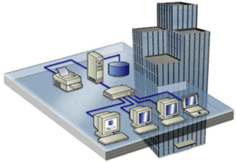
- Decision to renovate in 2008
- Stringent requirements
- Name & logo

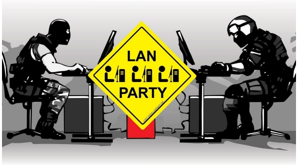
Requirement	**Value(s)**
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs

General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008
- Stringent requirements
- Name & logo
- Open & based on standards:
  - Bridged Local Area Network
  - 1 Gbit Ethernet
  - Precision Time Protocol

Requirement	Value(s)
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs

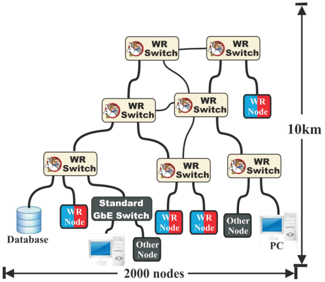




General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008
- Stringent requirements
- Name & logo
- Open & based on standards:
  - Bridged Local Area Network
  - 1 Gbit Ethernet
  - Precision Time Protocol

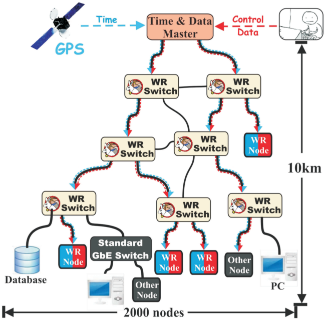
Requirement	Value(s)
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs



General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008
- Stringent requirements
- Name & logo
- Open & based on standards:
  - Bridged Local Area Network
  - 1 Gbit Ethernet
  - Precision Time Protocol
- Extends standards to meet CERN requirements with two services:

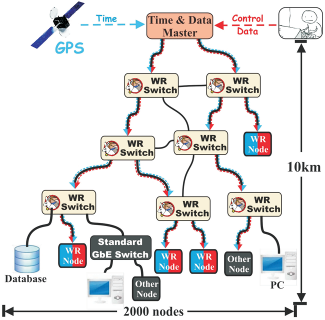
Requirement	Value(s)
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs



General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008
- Stringent requirements
- Name & logo
- Open & based on standards:
  - Bridged Local Area Network
  - 1 Gbit Ethernet
  - Precision Time Protocol
- Extends standards to meet CERN requirements with two services:
  1. **Sub-ns synchronization**

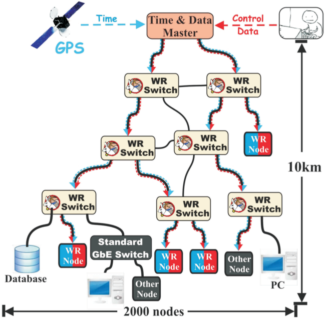
Requirement	Value(s)
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs



General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008
- Stringent requirements
- Name & logo
- Open & based on standards:
  - Bridged Local Area Network
  - 1 Gbit Ethernet
  - Precision Time Protocol
- Extends standards to meet CERN requirements with two services:
  1. **Sub-ns synchronization**
  2. **Deterministic and low-latency data transmission**

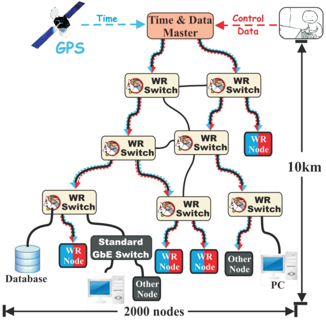
Requirement	**Value(s)**
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs



General Machine Timing renovation – White Rabbit

- Decision to renovate in 2008
- Stringent requirements
- Name & logo
- Open & based on standards:
  - Bridged Local Area Network
  - 1 Gbit Ethernet
  - Precision Time Protocol
- Extends standards to meet CERN requirements with two services:
  1. **Sub-ns synchronization**
  2. **Deterministic and low-latency data transmission**
- Foreseen for non-radiation areas
  (in radiation: WorldFIP – now, Powerlink – in future)

Requirement	Value(s)
Network size:	10km & 2000
Accuracy	sub-ns
Message size	1.2–5 kB
Msgs lost per year	1
Network max latency	1ms
Switch max latency	10μs



White Rabbit applications

- Particle accelerators
  - CERN
  - GSI (Germany)
  - JINR Dubna (Russia)

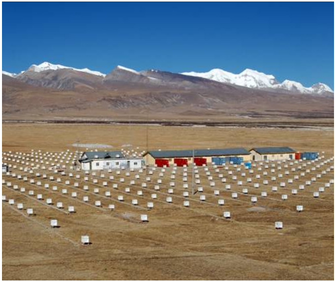
European Organization for Nuclear Research, CERN

All users: `www.ohwr.org/projects/white-rabbit/wiki/WRUsers`

White Rabbit applications

- Particle accelerators
  - CERN
  - GSI (Germany)
  - JINR Dubna (Russia)
- Cosmic ray & neutrinos detectors
  - LHAASO (China)
  - HiSCORE (Siberia)
  - KM3NET (at the bed of Mediterranean)

The Large High Altitude Air Shower Observation

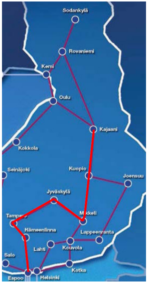


All users: `www.ohwr.org/projects/white-rabbit/wiki/WRUsers`

White Rabbit applications

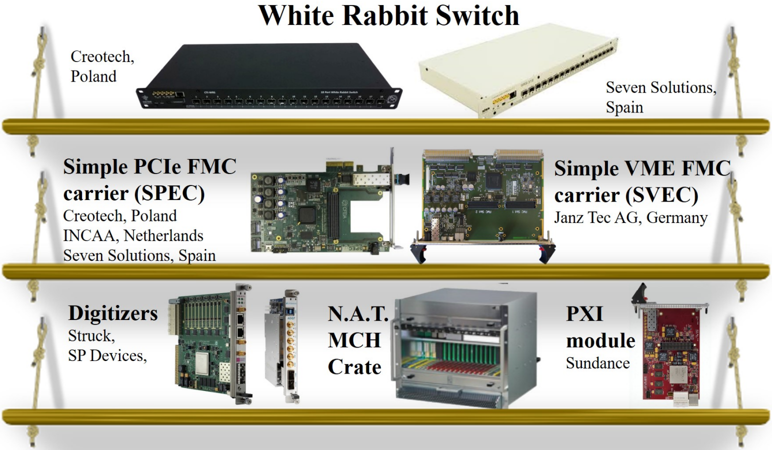
- Particle accelerators
  - CERN
  - GSI (Germany)
  - JINR Dubna (Russia)
- Cosmic ray & neutrinos detectors
  - LHAASO (China)
  - HiSCORE (Siberia)
  - KM3NET (at the bed of Mediterranean)
- National Metrology Institutes
  - MIKES (Finland)
  - VSL (Netherlands)
  - LNE-SYRTE (France)

Finish National Metrology Institute



All users: `www.ohwr.org/projects/white-rabbit/wiki/WRUsers`

Open **and** commercially available off-the-shelf



Companies selling White Rabbit:

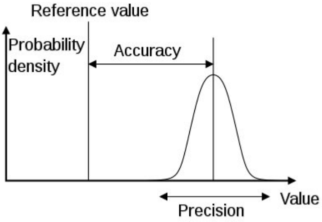
www.ohwr.org/projects/white-rabbit/wiki/wrcompanies

Outline

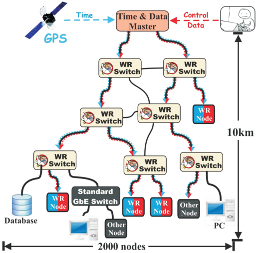
Outline

Sub-ns synchronization

- Synchronization performance:
  - **Sub-ns** accuracy: $max(|TE|) < 1ns$
  - **Sub-50 ps** precision: $sdev(TE) < 50ps$

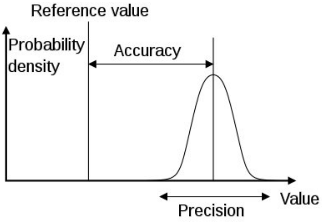


Value = Time Error (TE)

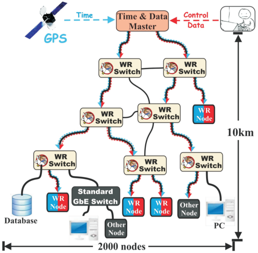


Sub-ns synchronization

- Synchronization performance:
  - **Sub-ns** accuracy: $max(|TE|) < 1ns$
  - **Sub-50 ps** precision: $sdev(TE) < 50ps$
- Building blocks:
  - Precision Time Protocol (PTP, IEEE1588)
  - Layer 1 syntonization
  - Digital Dual Mixer Time Difference
  - Link delay model

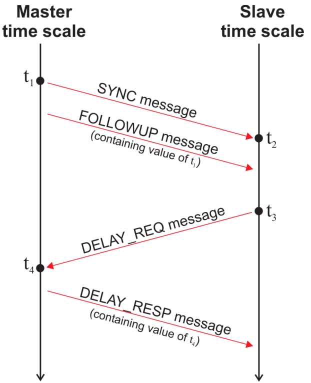


Value = Time Error (TE)



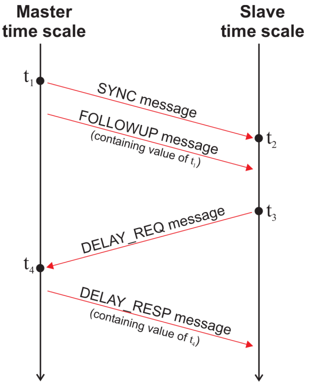
Outline

Precision Time Protocol (PTP, IEEE1588)



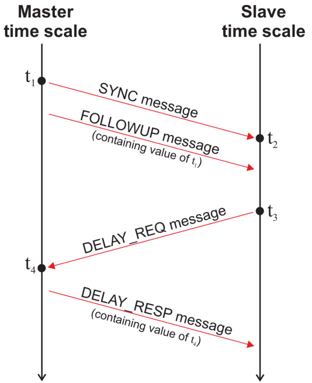
- Packet-based synchronization protocol

Precision Time Protocol (PTP, IEEE1588)



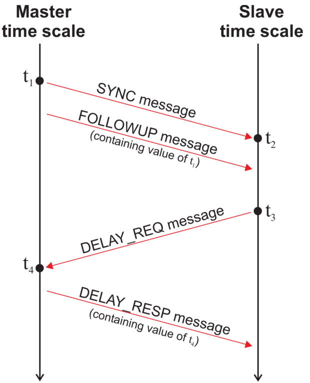
- Packet-based synchronization protocol
- Simple calculations:
  - Link *delay*$_{ms}$ $\delta_{ms} = \frac{(t_4 - t_1) - (t_3 - t_2)}{2}$
  - Clock *offset*$_{ms}$ $= t_2 - t_1 + \delta_{ms}$

Precision Time Protocol (PTP, IEEE1588)



- Packet-based synchronization protocol
- Simple calculations:
  - Link $delay_{ms}$ $\delta_{ms} = \frac{(t_4 - t_1) - (t_3 - t_2)}{2}$
  - Clock $offset_{ms} = t_2 - t_1 + \delta_{ms}$
- Sub-μs synchronisation

Precision Time Protocol (PTP, IEEE1588)

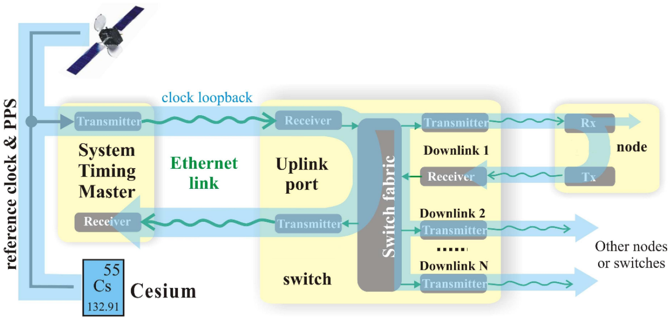


- Packet-based synchronization protocol
- Simple calculations:
  - Link *delay*$_{ms}$ $\delta_{ms} = \frac{(t_4 - t_1) - (t_3 - t_2)}{2}$
  - Clock *offset*$_{ms}$ $= t_2 - t_1 + \delta_{ms}$
- Sub-μs synchronisation
- Limitations:
  - Free-running oscillators
  - Timestamping precision
  - Medium asymmetry

Outline

Layer 1 Syntonization

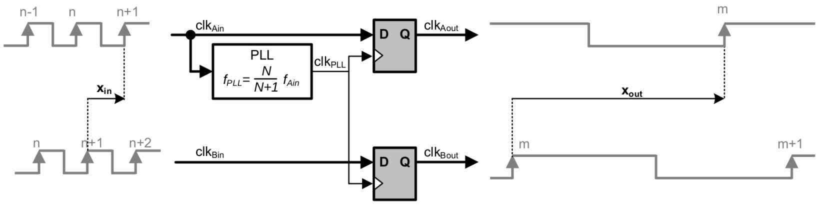
- All network devices use the same physical layer clock
- Clock is encoded in data by master and recovered by slave
- Clock loopback and phase detection allow precise timestamps



Outline

Digital Dual Mixer Time Difference (DDMTD)

- Clever implementation of a phase detector in an FPGA
- Uses D-flip-flops to zoom-in phase offset
- Allows for phase measurements at picosecond level

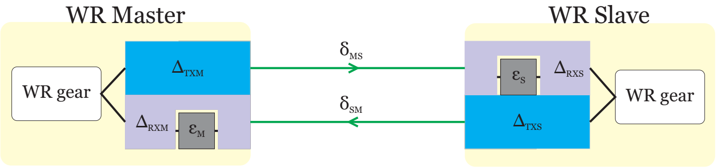


www.cern.ch/white-rabbit/documents/DDMTD_for_Sub-ns_Synchronization.pdf

Outline

Link delay model

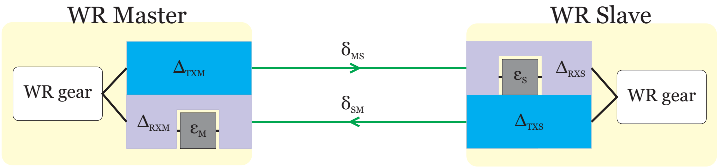
- Hardware delays:
- Link asymmetry:



www.cern.ch/white-rabbit/documents/WR_Calibration-v1.1-20151109.pdf

Link delay model

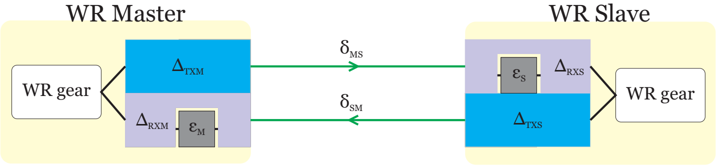
- Hardware delays:
  - Calibrate static delays: Δ_{TXM}, Δ_{RXM}, Δ_{TXS}, Δ_{RXS}
  - Measure semi-static delays: ϵ_M, ϵ_S
- Link asymmetry:



www.cern.ch/white-rabbit/documents/WR_Calibration-v1.1-20151109.pdf

Link delay model

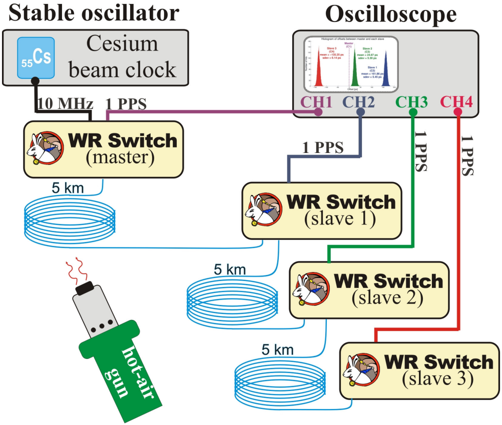
- Hardware delays:
  - Calibrate static delays: Δ_{TXM}, Δ_{RXM}, Δ_{TXS}, Δ_{RXS}
  - Measure semi-static delays: ϵ_M, ϵ_S
- Link asymmetry:
  - Single fibre for two-way communication
  - Fibre asymmetry coefficient: $\alpha = \frac{\delta_{MS} - \delta_{SM}}{\delta_{SM}}$



www.cern.ch/white-rabbit/documents/WR_Calibration-v1.1-20151109.pdf

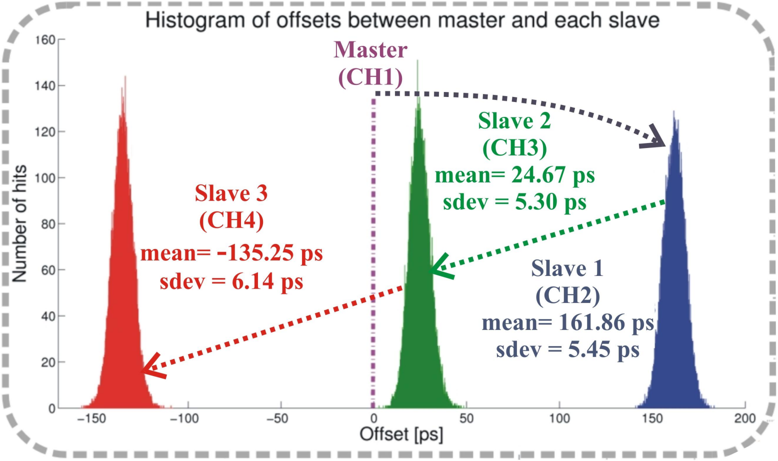
Outline

WR Synchronization performance



www.cern.ch/white-rabbit/documents/White_Rabbit-a_PTP_application_for_robust_sub-nanosecond_synchronization.pdf

WR Synchronization performance



www.cern.ch/white-rabbit/documents/White_Rabbit-a_PTP_application_for_robust_sub-nanosecond_synchronization.pdf

Outline

WR standardization in IEEE1588

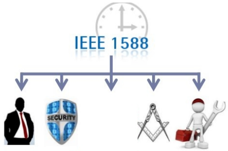
- IEEE standards are revised periodically

WR standardization in IEEE1588

- IEEE standards are revised periodically
- The IEEE1588 is revised by industry/academia

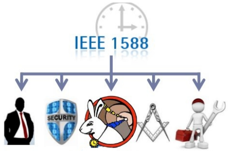
WR standardization in IEEE1588

- IEEE standards are revised periodically
- The IEEE1588 is revised by industry/academia
- Revision performed in 5 sub-committees
  (`https://ieee-sa.imeetcentral.com/1588public/`)



WR standardization in IEEE1588

- IEEE standards are revised periodically
- The IEEE1588 is revised by industry/academia
- Revision performed in 5 sub-committees (`https://ieee-sa.imeetcentral.com/1588public/`)
- High Accuracy sub-committee dedicated to WR



WR standardization in IEEE1588

- IEEE standards are revised periodically
- The IEEE1588 is revised by industry/academia
- Revision performed in 5 sub-committees (https://ieee-sa.imeetcentral.com/1588public/)
- High Accuracy sub-committee dedicated to WR

WR standardization in IEEE1588

- IEEE standards are revised periodically
- The IEEE1588 is revised by industry/academia
- Revision performed in 5 sub-committees (https://ieee-sa.imeetcentral.com/1588public/)
- High Accuracy sub-committee dedicated to WR
  - Generalization of WR methods
  - Inclusion of the generalizations

WR standardization in IEEE1588

- IEEE standards are revised periodically
- The IEEE1588 is revised by industry/academia
- Revision performed in 5 sub-committees
  (`https://ieee-sa.imeetcentral.com/1588public/`)
- High Accuracy sub-committee dedicated to WR
  - Generalization of WR methods
  - Inclusion of the generalizations
- Revised IEEE1588 expected in 2018/2019

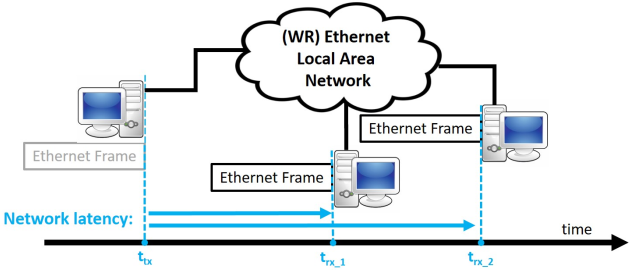
Outline

Determinism and network latency

- **Determinism**
  A deterministic system is predictable: it provides calculable and consistent characteristics of operation that are required by the application, e.g. network latency of data transmission.

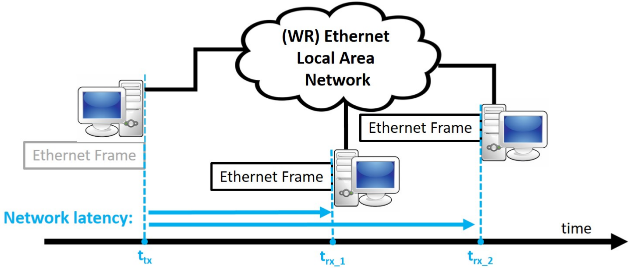
Determinism and network latency

- **Determinism**
  A deterministic system is predictable: it provides calculable and consistent characteristics of operation that are required by the application, e.g. network latency of data transmission.
- **Network latency**



Determinism and network latency

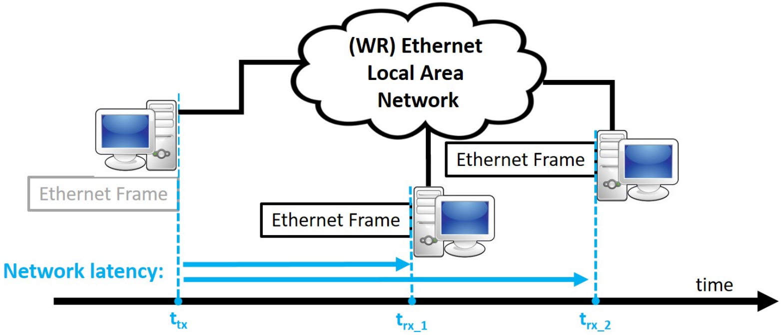
- **Determinism**
  A deterministic system is predictable: it provides calculable and consistent characteristics of operation that are required by the application, e.g. network latency of data transmission.
- **Network latency**



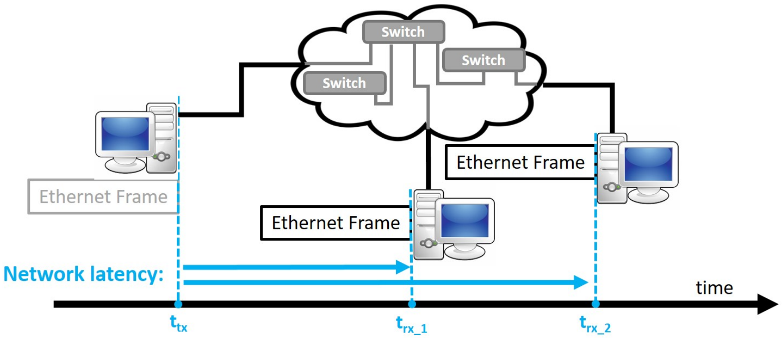
Deterministic network is a network in which we can calculate the maximum latency

Outline

Network latency contributors



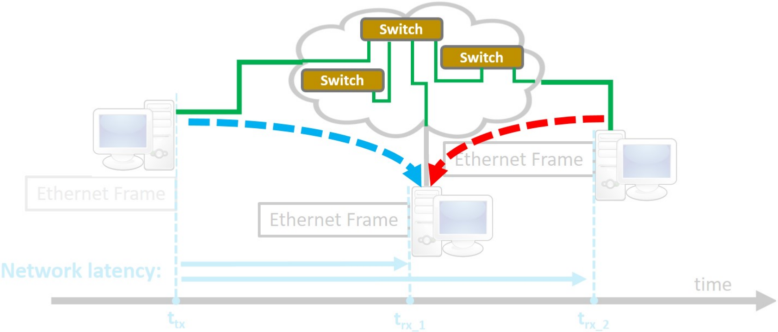
Network latency contributors



Network latency contributors

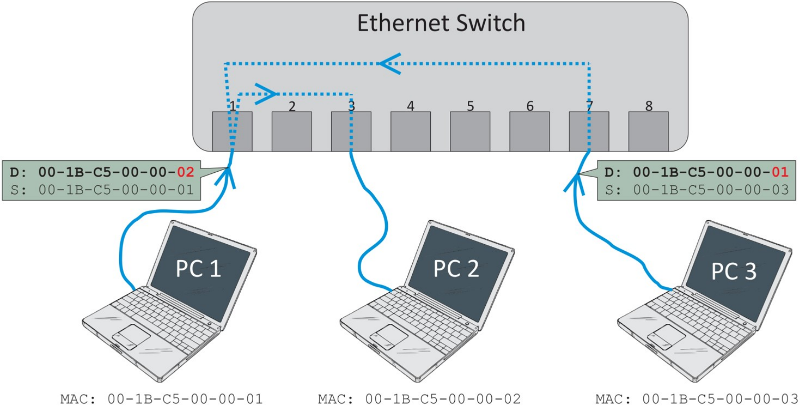
- Cables: 5us/km – we cannot do much about this
- Switch operation
- Other traffic

We can do something about this



Outline

Switch in a nutshell



Latency in WR Switch

- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind

Latency in WR Switch

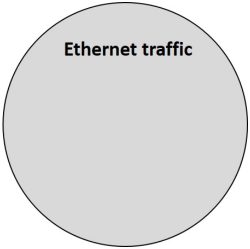
- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind
- WR Switch is a cut-through switch for all traffic:
  - Frames forwarded as soon as possible
  - Common optimization of latency in switches

Latency in WR Switch

- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind
- WR Switch is a cut-through switch for <u>all traffic</u>:
  - Frames forwarded as soon as possible
  - Common optimization of latency in switches
- WR Switch is latency-optimized for <u>selected traffic</u>:

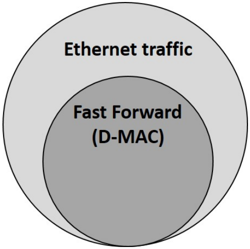
Latency in WR Switch

- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind
- WR Switch is a cut-through switch for all traffic:
  - Frames forwarded as soon as possible
  - Common optimization of latency in switches
- WR Switch is latency-optimized for selected traffic:



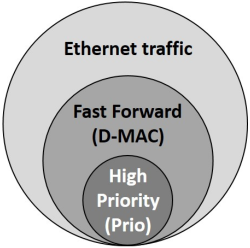
Latency in WR Switch

- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind
- WR Switch is a cut-through switch for all traffic:
  - Frames forwarded as soon as possible
  - Common optimization of latency in switches
- WR Switch is latency-optimized for selected traffic:
  - **Fast Forward (FF)**:
    - Any traffic with selected D-MAC addresses
    - Optimization of latency due to switch operation
    - Released feature



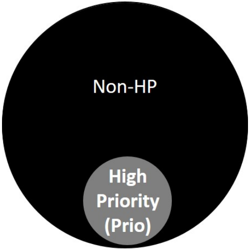
Latency in WR Switch

- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind
- WR Switch is a cut-through switch for all traffic:
  - Frames forwarded as soon as possible
  - Common optimization of latency in switches
- WR Switch is latency-optimized for selected traffic:
  - **Fast Forward (FF)**:
    - Any traffic with selected D-MAC addresses
    - Optimization of latency due to switch operation
    - Released feature
  - **High Priority (HP)**:
    - Fast Forward traffic with selected priorities
    - Optimization of latency due to non-HP traffic
    - Experimental



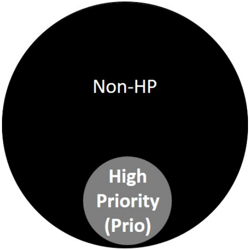
Latency in WR Switch

- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind
- WR Switch is a cut-through switch for all traffic:
  - Frames forwarded as soon as possible
  - Common optimization of latency in switches
- WR Switch is latency-optimized for selected traffic:
  - **Fast Forward (FF)**:
    - Any traffic with selected D-MAC addresses
    - Optimization of latency due to switch operation
    - Released feature
  - **High Priority (HP)**:
    - Fast Forward traffic with selected priorities
    - Optimization of latency due to non-HP traffic
    - Experimental



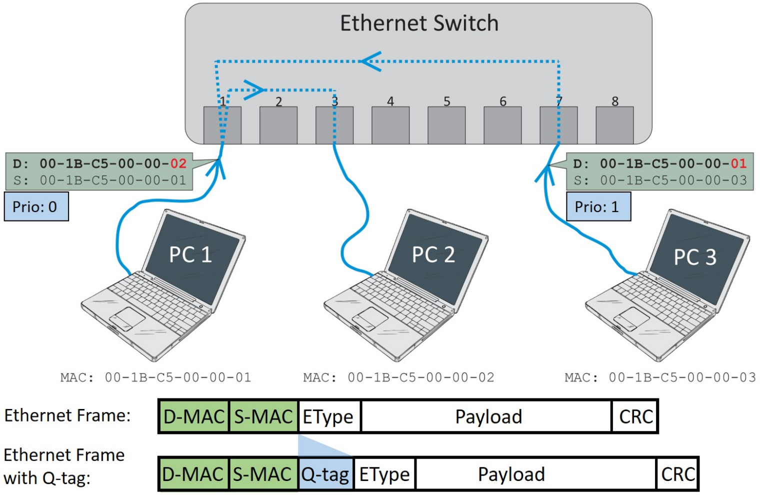
Latency in WR Switch

- WR Switch is deterministic by design:
  - Open-source, each source of latency is verifiable
  - Designed with latency in mind
- WR Switch is a cut-through switch for all traffic:
  - Frames forwarded as soon as possible
  - Common optimization of latency in switches
- WR Switch is latency-optimized for selected traffic:
  - **Fast Forward (FF)**:
    - Any traffic with selected D-MAC addresses
    - Optimization of latency due to switch operation
    - Released feature
  - **High Priority (HP)**:
    - Fast Forward traffic with selected priorities
    - Optimization of latency due to non-HP traffic
    - Experimental
- And what on Earth are priorities...?



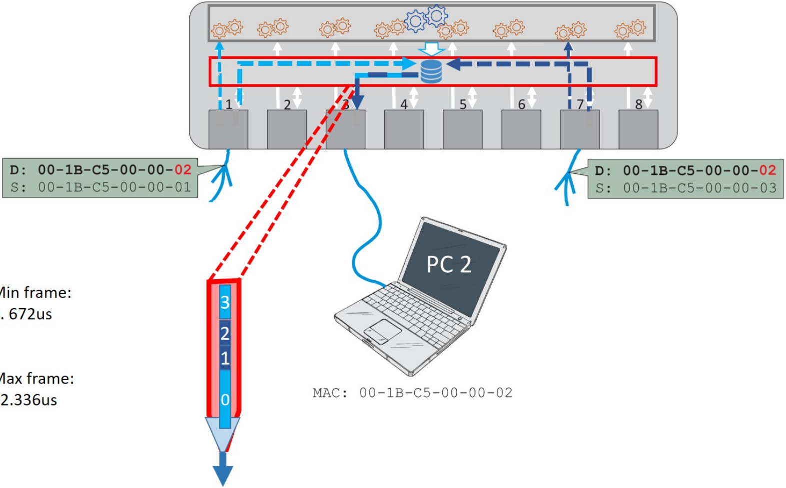
Outline

Priorities - standard extension of Ethernet Frame



Standard and optional extension of Ethernet Frame (IEEE802.1Q)

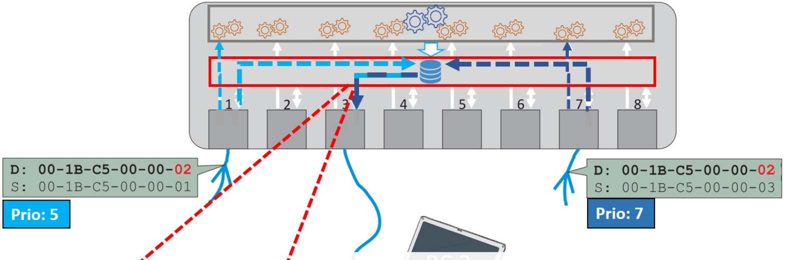
WR Switch Not Using priorities



Min frame:
0. 672us

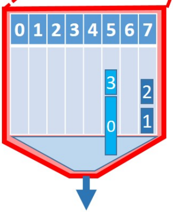
Max frame:
12.336us

WR Switch Using Standard Priorities



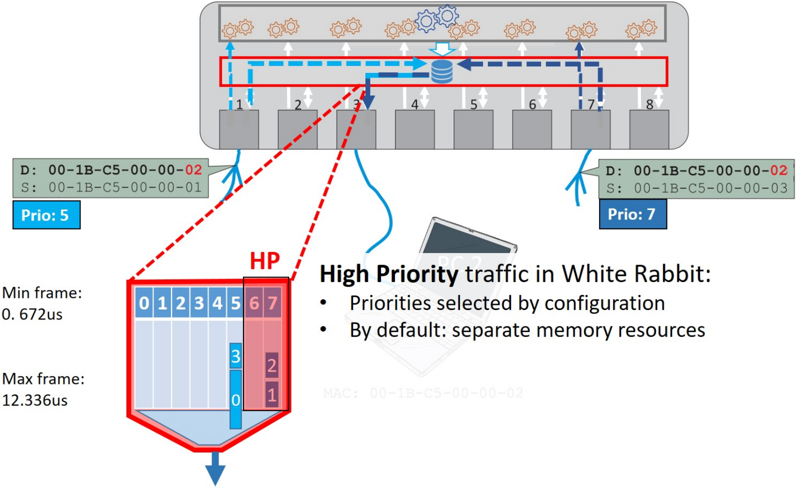
Min frame:
0. 672us

Max frame:
12.336us



- Assignment of priorities – standard option
- Two problem remain:
  - Memory resources
  - Frames being transmitted

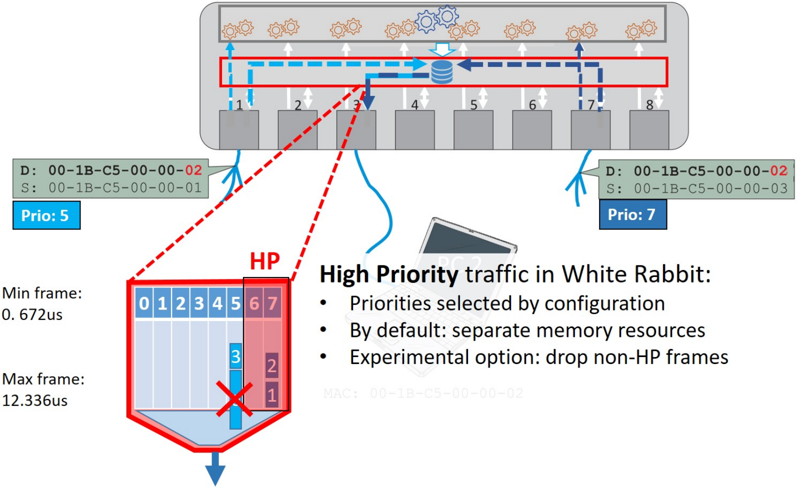
WR Switch Using Standard Priorities



High Priority traffic in White Rabbit:

- Priorities selected by configuration
- By default: separate memory resources

High Priority

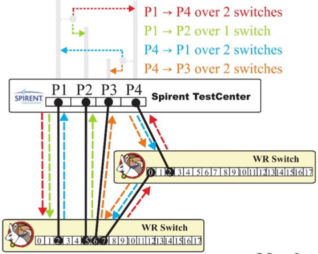


High Priority traffic in White Rabbit:

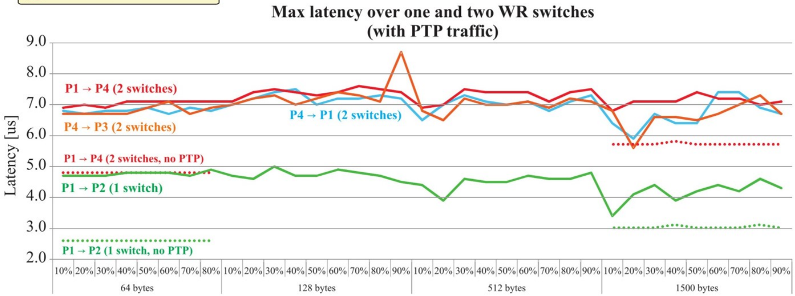
- Priorities selected by configuration
- By default: separate memory resources
- Experimental option: drop non-HP frames

Outline

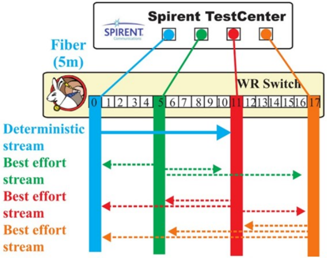
Latency of WR Switch for Fast Forward



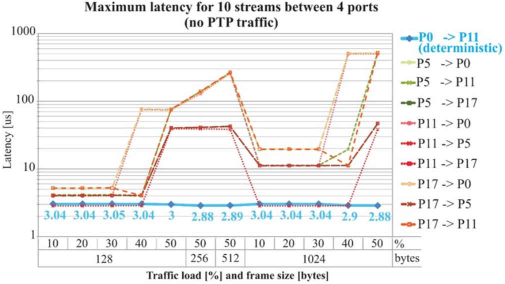
Intervening traffic	Latency [us]			
	One switch		Two switches	
	Max	Pk-pk	Max	Pk-pk
No	3.1	0.3	5.8	0.5
WR-PTP	5.6	2.8	8.7	3.9
Non-HP traffic	3.1	0.2	N/A	N/A



Latency of WR Switch for High Priority



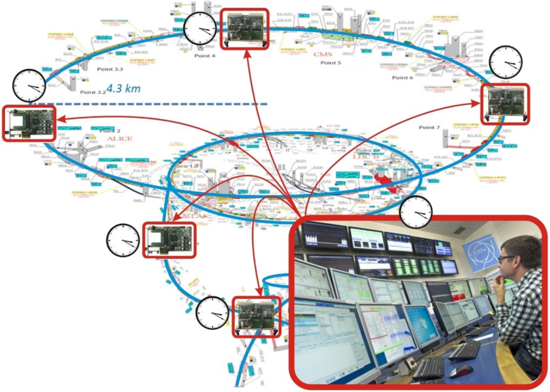
Intervening traffic	Latency [us]			
	One switch		Two switches	
	Max	Pk-pk	Max	Pk-pk
No	3.1	0.3	5.8	0.5
WR-PTP	5.6	2.8	8.7	3.9
Non-HP traffic	3.1	0.2	N/A	N/A



Outline

Outline

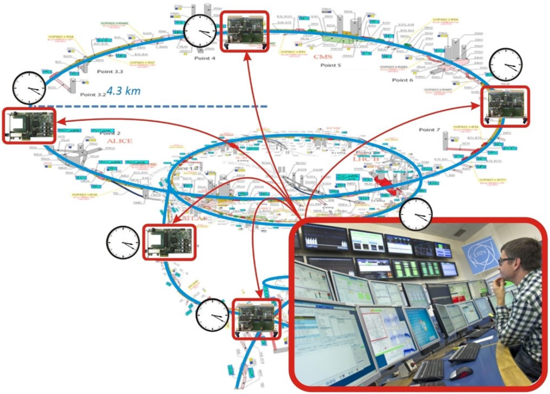
Using synchronization and determinism



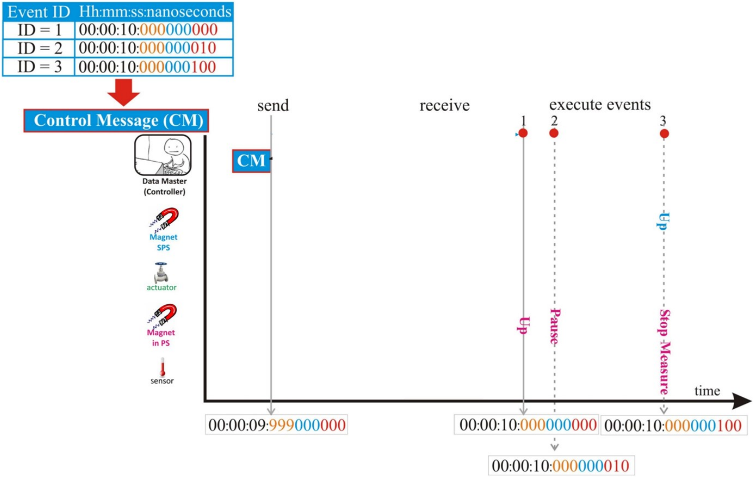
Using synchronization and determinism

Event ID	Hh:mm:ss:nanoseconds
ID = 1	00:00:10:000000000
ID = 2	00:00:10:000000010
ID = 3	00:00:10:000000100

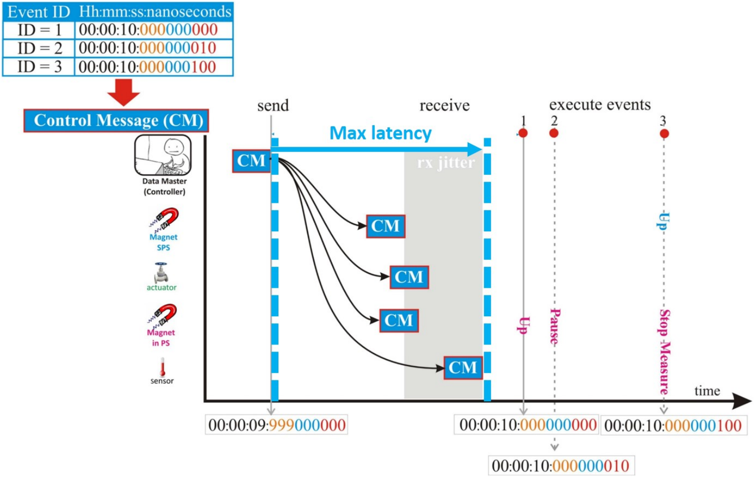
Control Message (CM)



Using synchronization and determinism



Using synchronization and determinism



Outline

Summary

- Open source (H/W & S/W) with commercial availability & support

Summary

- Open source (H/W & S/W) with commercial availability & support
- More applications than ever expected

Summary

- Open source (H/W & S/W) with commercial availability & support
- More applications than ever expected
- A versatile solution for controls & acquisition in non-radiation
  (WorldFIP is used at CERN in radiation currently, Powerlink in the future)

Summary

- Open source (H/W & S/W) with commercial availability & support
- More applications than ever expected
- A versatile solution for controls & acquisition in non-radiation
  (WorldFIP is used at CERN in radiation currently, Powerlink in the future)
- Standard-compatible and standard-extending
  - White Rabbit solutions to be part of IEEE1588 (PTP) standard
  - Compatible enhancement of IEEE802.1Q, similar mechanism in TSN

Summary

- Open source (H/W & S/W) with commercial availability & support
- More applications than ever expected
- A versatile solution for controls & acquisition in non-radiation
  (WorldFIP is used at CERN in radiation currently, Powerlink in the future)
- Standard-compatible and standard-extending
  - White Rabbit solutions to be part of IEEE1588 (PTP) standard
  - Compatible enhancement of IEEE802.1Q, similar mechanism in TSN
- The most accurate Ethernet-based synchronisation in the World

Summary

- Open source (H/W & S/W) with commercial availability & support
- More applications than ever expected
- A versatile solution for controls & acquisition in non-radiation
  (WorldFIP is used at CERN in radiation currently, Powerlink in the future)
- Standard-compatible and standard-extending
  - White Rabbit solutions to be part of IEEE1588 (PTP) standard
  - Compatible enhancement of IEEE802.1Q, similar mechanism in TSN
- The most accurate Ethernet-based synchronisation in the World
- WR Switch and so WR Network
  - Deterministic by (open) design
  - Latency-optimized for selected traffic
  - Upper-bound latency for HP traffic (experimental)

Summary

- Open source (H/W & S/W) with commercial availability & support
- More applications than ever expected
- A versatile solution for controls & acquisition in non-radiation
  (WorldFIP is used at CERN in radiation currently, Powerlink in the future)
- Standard-compatible and standard-extending
  - White Rabbit solutions to be part of IEEE1588 (PTP) standard
  - Compatible enhancement of IEEE802.1Q, similar mechanism in TSN
- The most accurate Ethernet-based synchronisation in the World
- WR Switch and so WR Network
  - Deterministic by (open) design
  - Latency-optimized for selected traffic
  - Upper-bound latency for HP traffic (experimental)
- Future plans
  - Mechanisms to increase reliability
    (`http://cds.cern.ch/record/2261452`)
  - 10 Gbit Ethernet

Thank you

Thank you !

www.cern.ch/white-rabbit

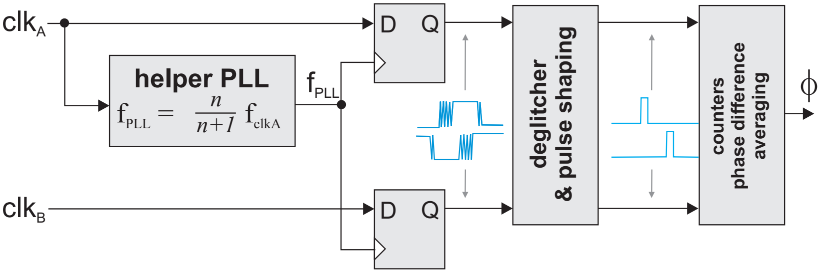
Extras

Extras

Outline

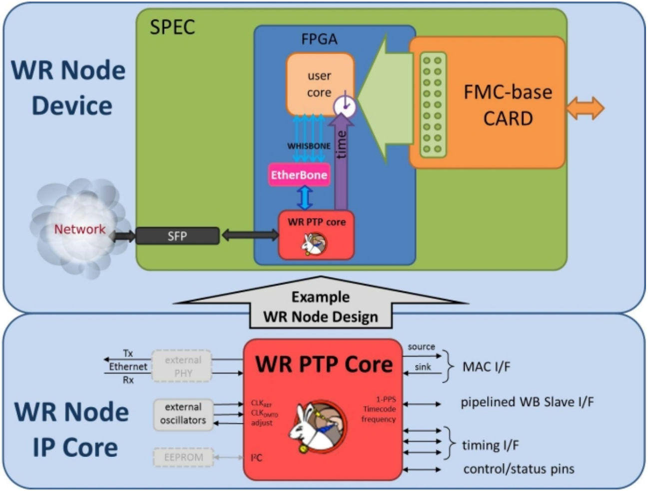
Digital Dual Mixer Time Difference (DDMTD)

- Used for precise phase measurements
- Implemented in FPGA and SoftPLL
- 62.5MHz WR clock and N=14 results in 3.814kHz output signals



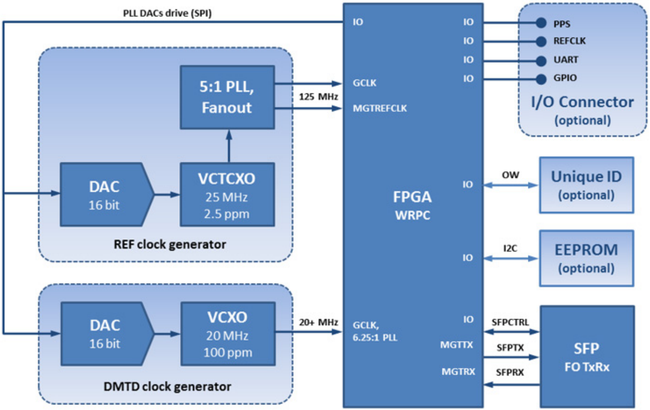
Outline

WR Node IP Core



www.ohwr.org/projects/white-rabbit/wiki/node

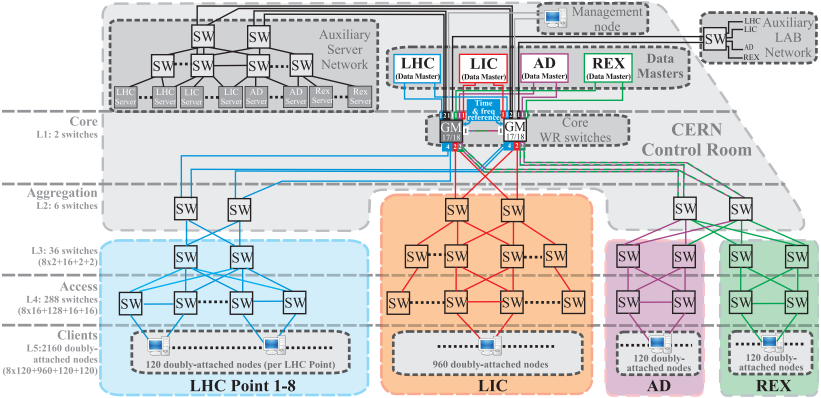
WR Node Reference Design for Hardware



www.ohwr.org/projects/white-rabbit/wiki/wrreferencedesign

Outline

WR Reference Network



Explanation:

SW – White Rabbit **SW**itch
GM – **G**rand**m**aster SW
LHC – **L**arge **H**adron **C**ollider
LIC – **L**HC **I**njection **C**hain
—— – Fiber link

AD – **A**ntiproton **D**ecelerator
ISOLDE – **I**sotrope **S**eparator **O**n**L**ine **DE**vice
REX – The **R**adioactive beam **Ex**periment @ ISOLDE
(arrow) – 10MHz, PPS and UTC

Performance of the Reference Network

Name	Requirement	In the reference WR network
Network size:	10 km & 2000 nodes	10 km & 2160 nodes
Synchronization : - accuracy over a year: - accuracy in transient: - precision: Control message - allowed size - max lost per year	 sub-ns sub-ns sub-50 ps 1200–6000 bytes 1	 0.41 *ns* 1.19 *ns* 31 *ps* 1200–6000 bytes 1 with probability R(t)
Upper-bound network latency	$< 500\ \mu s$ (derived from 1 ms)	$\leq$ 78 μs for network $\leq$ 150 μs for control message
Total reliability $R(t)$	≥ 0.98	0.9854 for $MTBF_{switch} = 40\,000\ h$ 0.9967 for $MTBF_{switch} = 100\,000\ h$ 0.9997 for $MTBF_{switch} = 650\,000\ h$