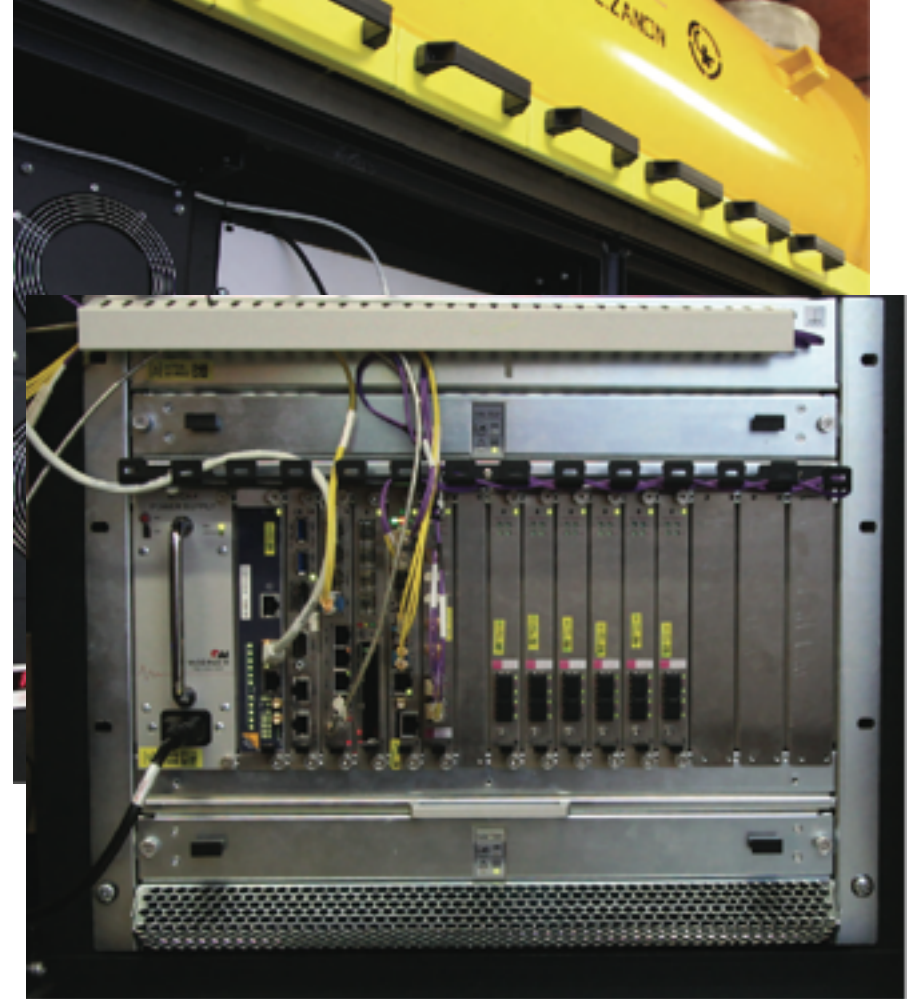


PCIexpress in MicroTCA.4 System at XFEL

150 NAT-MCH-PHYS and 50 NAT-MCH-PHYS80



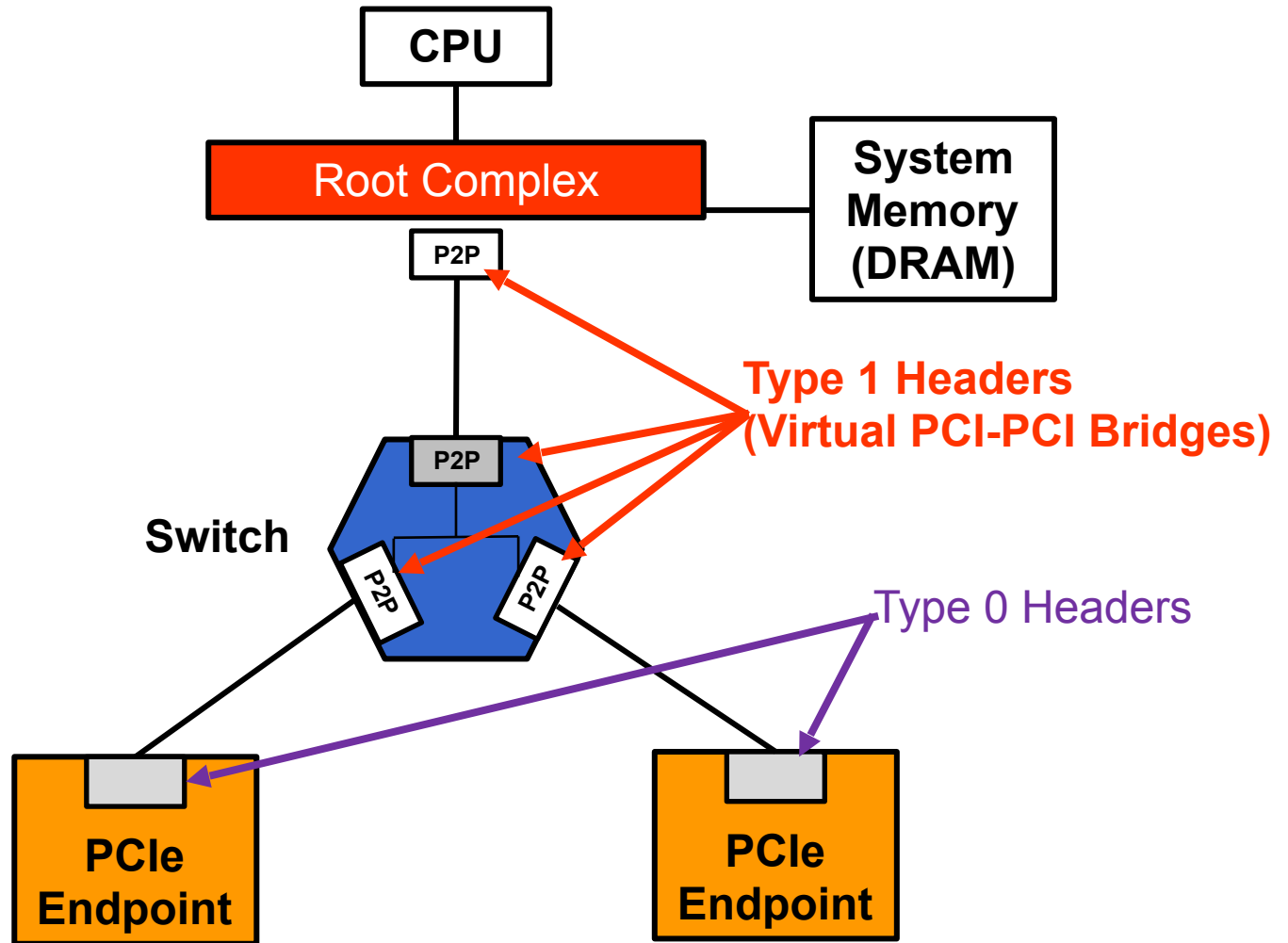
PCIexpress in MicroTCA Systems

Agenda



- PCIe: Topology, device tree, root complex
- Sub-dividing a system with N root complexes
- Hot-swap

lspci -s bb:dev.func Configuration Space



PCIexpress devices & endpoints

root@nat: #lspci



lspci

00:00.0 Host bridge: Intel Corporation 2nd Generation Core Processor Family DRAM Controller (rev 09)

00:01.0 PCI bridge: Intel Corporation Xeon E3-1200/2nd Generation Core Processor Family PCI Express Root Port (rev 09)

00:01.1 PCI bridge: Intel Corporation Xeon E3-1200/2nd Generation Core Processor Family PCI Express Root Port (rev 09)

00:02.0 VGA compatible controller: Intel Corporation 2nd Generation Core Processor Family Integrated Graphics Controller (rev 09)

00:16.0 Communication controller: Intel Corporation 6 Series Chipset Family MEI Controller #1 (rev 04)

00:19.0 Ethernet controller: Intel Corporation 82579LM Gigabit Network Connection (rev 04)

00:1a.0 USB Controller: Intel Corporation 6 Series Chipset Family USB Enhanced Host Controller #2 (rev 04)

00:1c.0 PCI bridge: Intel Corporation 6 Series Chipset Family PCI Express Root Port 1 (rev b4)

00:1d.0 USB Controller: Intel Corporation 6 Series Chipset Family USB Enhanced Host Controller #1 (rev 04)

00:1f.0 ISA bridge: Intel **Corporation QM67 Express Chipset Family LPC Controller** (rev 04)

00:1f.2 IDE interface: Intel Corporation 6 Series Chipset Family 4 port **SATA** IDE Controller (rev 04)

00:1f.3 SMBus: Intel Corporation 6 Series Chipset Family SMBus Controller (rev 04)

00:1f.5 IDE interface: Intel Corporation 6 Series Chipset Family 2 port SATA IDE Controller (rev 04)

01:00.0 PCI bridge: Integrated Device Technology Inc. Device 808f

01:00.2 System peripheral: Integrated Device Technology Inc. Device 808f

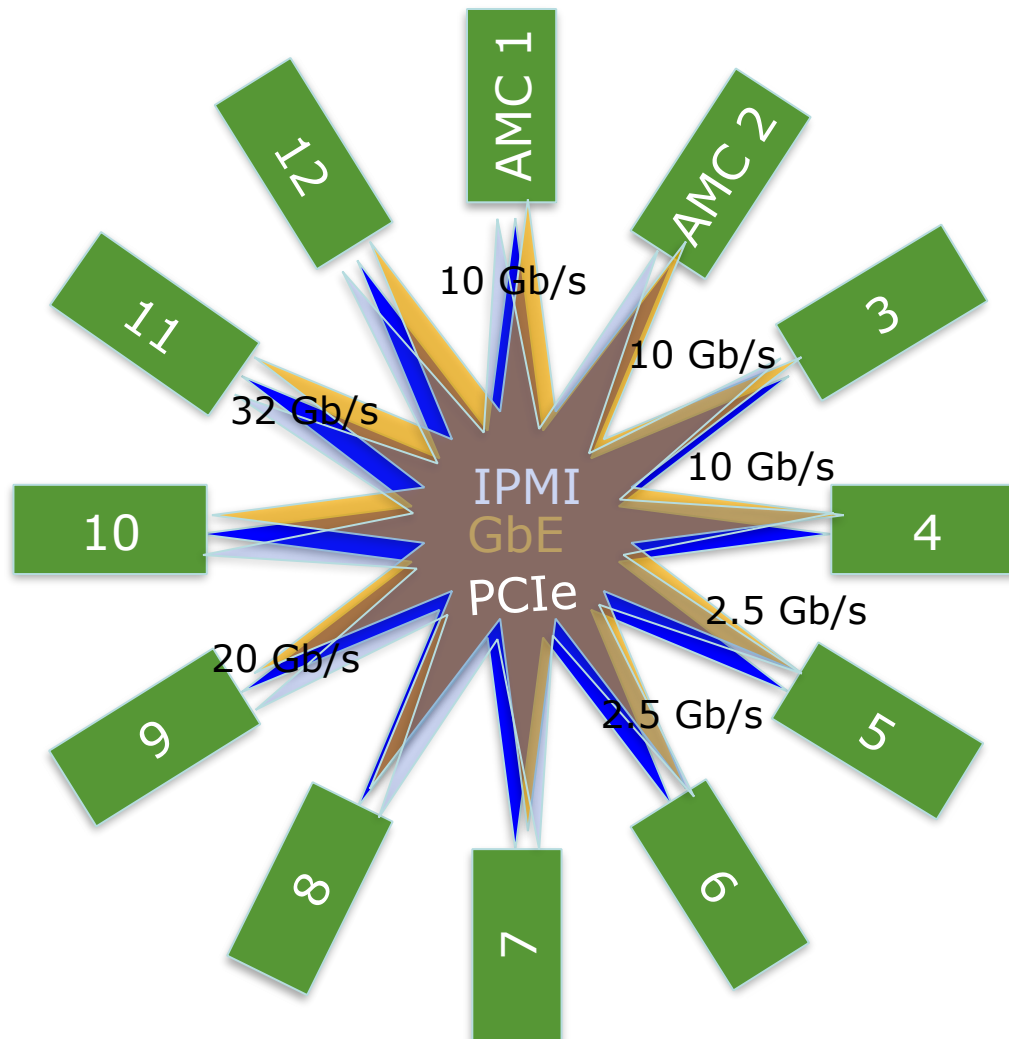
02:08.0 PCI bridge: Integrated Device Technology Inc. Device 808f

03:00.0 PCI bridge: **PLX Technology Inc. Device 8748** (rev ba)

04:00.0 PCI bridge: PLX Technology Inc. Device 8748 (rev ba)

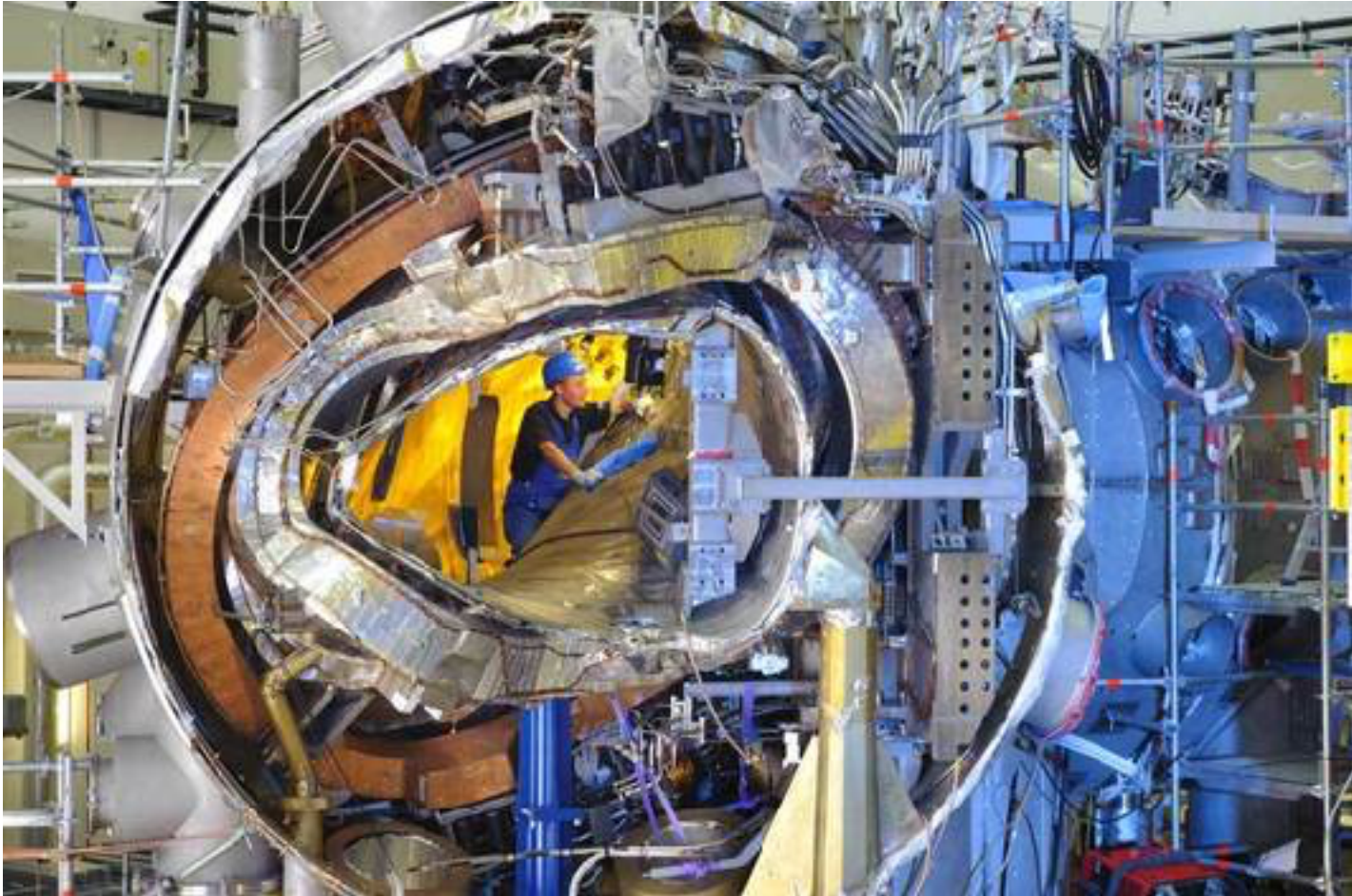
Star Topology

Any Slot can be Root Complex



Application Example: MPI Greifswald

12 * 4 PCIe lanes 2 * 16 PCIe lanes



- Setup**
- Base Configuration
 - Switch **HA300-1024P**
 - Age Time
 - Port on/off
 - Port VLAN
 - 802.1Q VLAN
 - 802.1X
 - 802.1p
 - Port Mirroring
 - Jumbo Frame
 - Link Aggregation
 - Rapid Spanning Tree
 - Link Status
 - PCMP2005 counters
 - Configure PCIe Virtual Switches**

- Maintenance**
- Board Information
 - System Information
 - Reboot NAT-MCH
 - Update MCH
 - Change Password
 - N.A.T. Webpage
 - Home

Welcome to the HTML based NAT-MCH configuration tool.

Setup Functions:

- Base Configuration:** - Changes Base Configuration.
- Age Time:** - MAC Table setup: set the aging of the MAC Table Entries.
- Port VLAN:** - Port based VLAN setup and port enable/disable.
- 802.1Q VLAN:** - 802.1Q VLAN setup.
- 802.1X:** - 802.1X security setup.
- 802.1p:** - 802.1p Quality of Service setup.
- Port Mirroring:** - Mirroring of the inbound and outbound traffic on a port
- Jumbo frames:** - Support of the Jumbo frames on a port
- Link Aggregation:** - Support of up to four the Link Aggregation groups
- Rapid Spanning** - Support of the Rapid Spanning Tree by 10GbE-Switch
- IGMP Snooping** - Support of the IGMP Snooping by 10GbE-Switch (FM4000 only)
- Link Status:** - Show the current status of the Ethernet links
- Counter Statistic:** - Show the counter statistic of the Ethernet switch

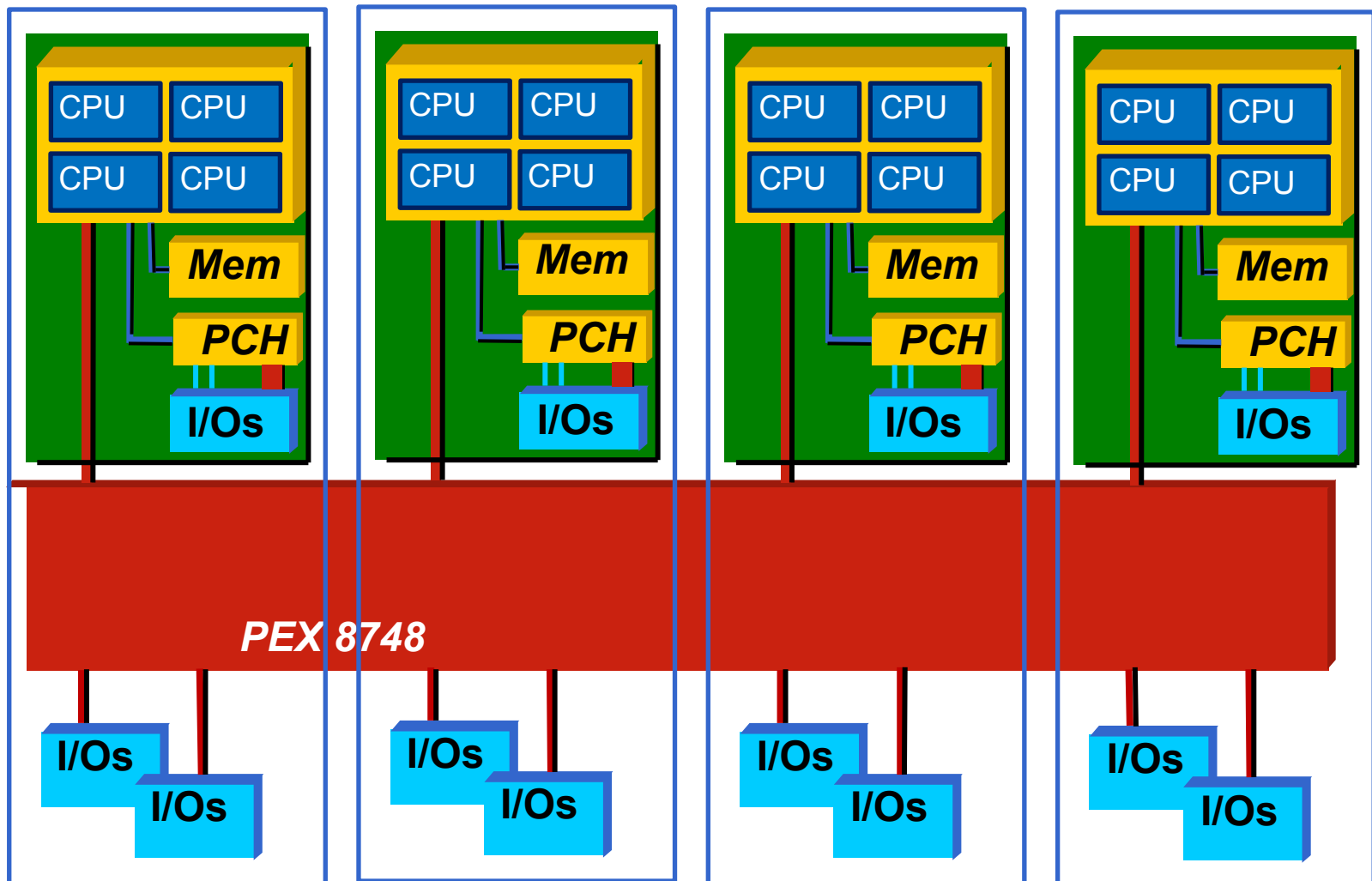
Maintenance Functions:

- Script Management:** - Backup/Restore settings to/from flash memory or file.
- Board Information:** - Provides hardware information of this NAT-MCH.
- System Information:** - Collect hardware information of this system.
- Reboot NAT-MCH:** - Allows rebooting over the Web-Interface.
- Update MCH:** - Allows updating several components over the Web-Interface.
- Change/Reset Password:** - Allows changing or resetting of the MCH Password over the Web-Interface.
- N.A.T. Webpage:** - Opens the N.A.T. webpage in a new browser window.
- Home:** - Shows this page.



PEX8748

Multi-Host Configuration: up to 6 Cluster







NAT-MCH by N.A.T.

Setup

Base Configuration

Switch **BASE:1000**

Age Time

Port on/off

Port VLAN

802.1Q VLAN

802.1X

802.1p

Port Mirroring

Jumbo Frame

Link Aggregation

Rapid Spanning Tree

Link Status

BCM5396 counters

Configure PCIe Virtual Switches

Maintenance

Board Information

System Information

Reboot NAT-MCH

Update MCH

Change Password

N.A.T. Webpage

Home

PCIe Virtual Switch configuration

Select Host AMCs (Upstream) for each virtual switch that shall be enabled first.

Select Host AMCs (Non-Transparent Upstream) for each virtual switch that shall be enabled afterwards.

Select which AMCs shall be connected to each virtual switch as downstream in the end.

Virtual Switch	Upstream AMC	NT-Upstream AMC	AMC 1 4..7	AMC 2 4..7	AMC 3 4..7	AMC 4 4..7	AMC 5 4..7	AMC 6 4..7	AMC 7 4..7
none			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Virtual Switch 0	RTM	- none -	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Virtual Switch 1	AMC 0_4		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Virtual Switch 2	- none -		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Virtual Switch 3	- none -		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Virtual Switch 4	- none -		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Virtual Switch 5	- none -		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Max. Link Speed			8.0 Gbit/s	8.0 Gbit/s	8.0 Gbit/s	8.0 Gbit/s	8.0 Gbit/s	8.0 Gbit/s	8.0 Gbit/s

Apply

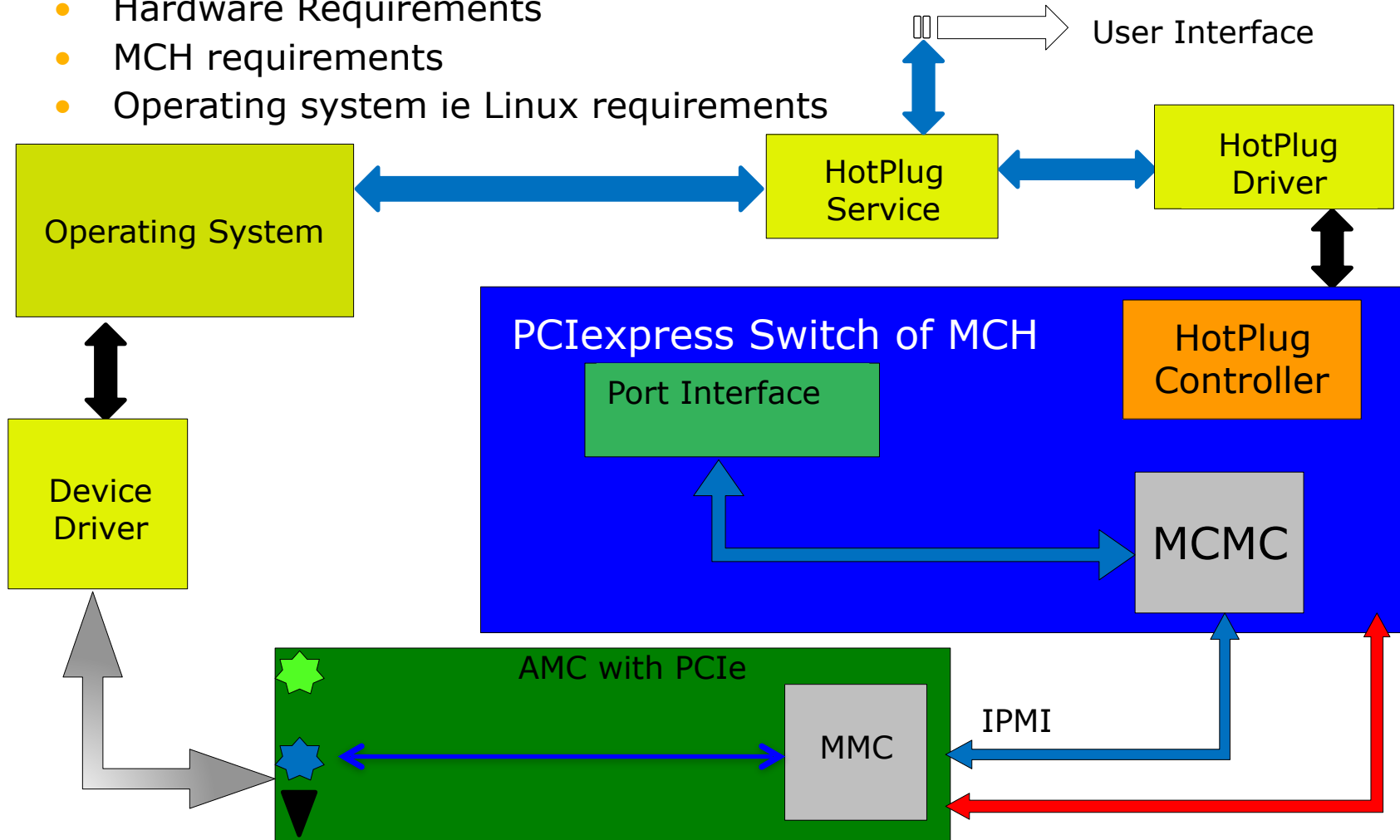
Note: You need to click apply before you can save your changes to EEPROM.

PCIexpress Requirements

Hot Plug and not Hotswap Solution

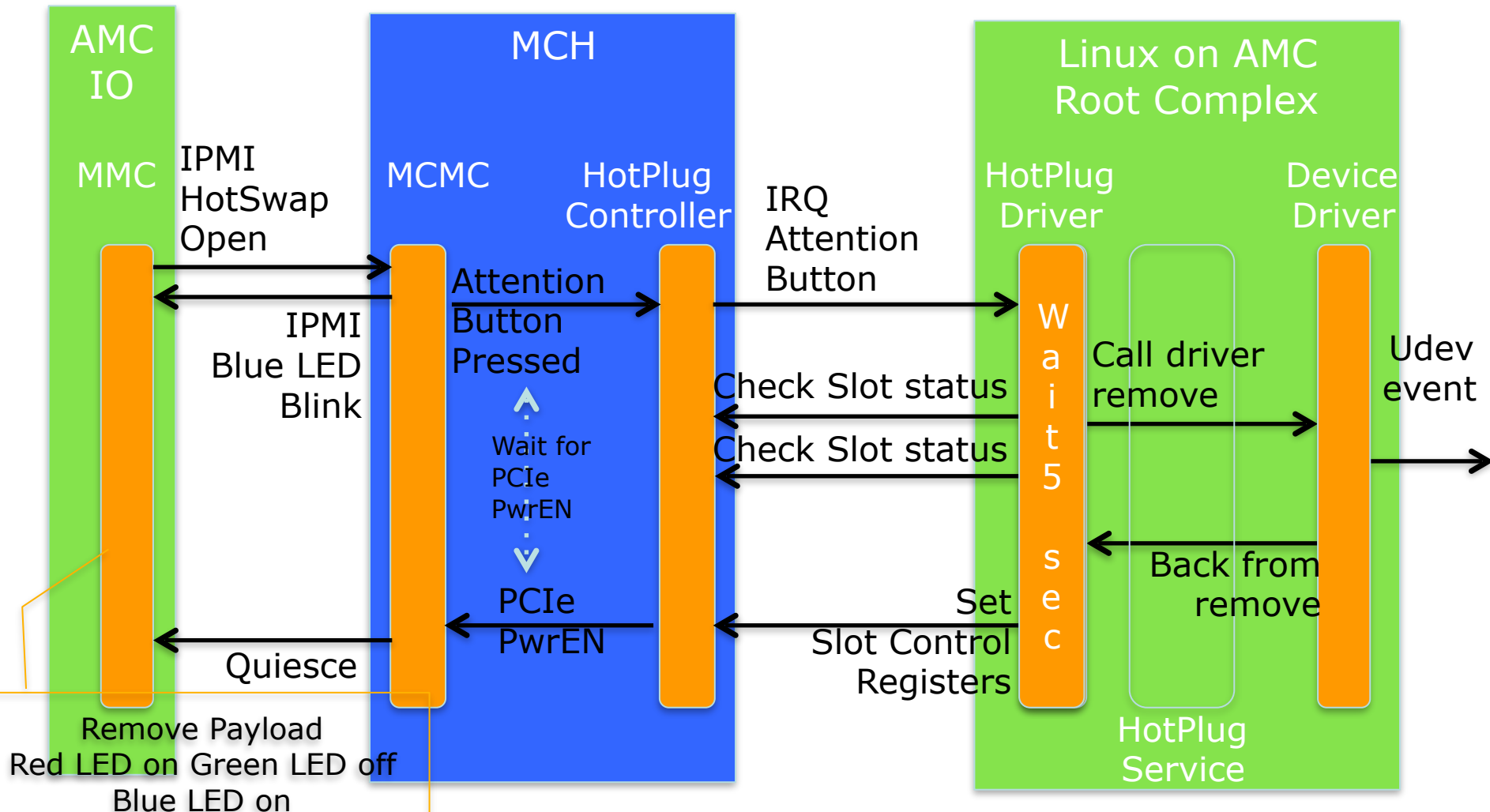


- Hardware Requirements
- MCH requirements
- Operating system ie Linux requirements



Hot Removal of AMC with PCIe

Hotswap handle pulled out: standard procedure





Testing HotPlug

Prepare **MCH** and OS

- The PCI Express Hot-Swap in MTCA depends on:
 - Linux OS Hotplug Driver (pciehp)
 - PCI Express Switch with the Hot-Plug controller on the MCH
- We have to ...
- Configure the MCH (PCI Express Switch)
 - Enable Hot-Plug Controller on the PCI Express Switch

PCIe parameter	Current Configuration
configuration flags:	
upstream slot power up delay	25 sec ←
power up delay for AMC	15 sec ←
100 MHz spread spectrum	disabled ↕
hot plug support	enabled ↕ ←
PCIe early ekey (before payload)	disabled ↕
Use PCIe on MCH RTM(disable AMC12)	no ↕

Thank you very much!

Questions?



Vollrath Dirksen
vollrath@nateurope.com



N.A.T. GmbH
Konrad-Zuse-Platz 9
53227 Bonn, Germany

www.nateurope.com

MTCA.4 Training:

techlab.desy.de/support/training

www.nateurope.com/services/support.html