



Antelope in an HA Cluster Environment

An Avenue to
Reliability and Transparency
Case Study: DPC

*Stefan M. Radman, Consultant - Senior ICT Engineer,
Open Systems & Services, Kinematics Inc.*

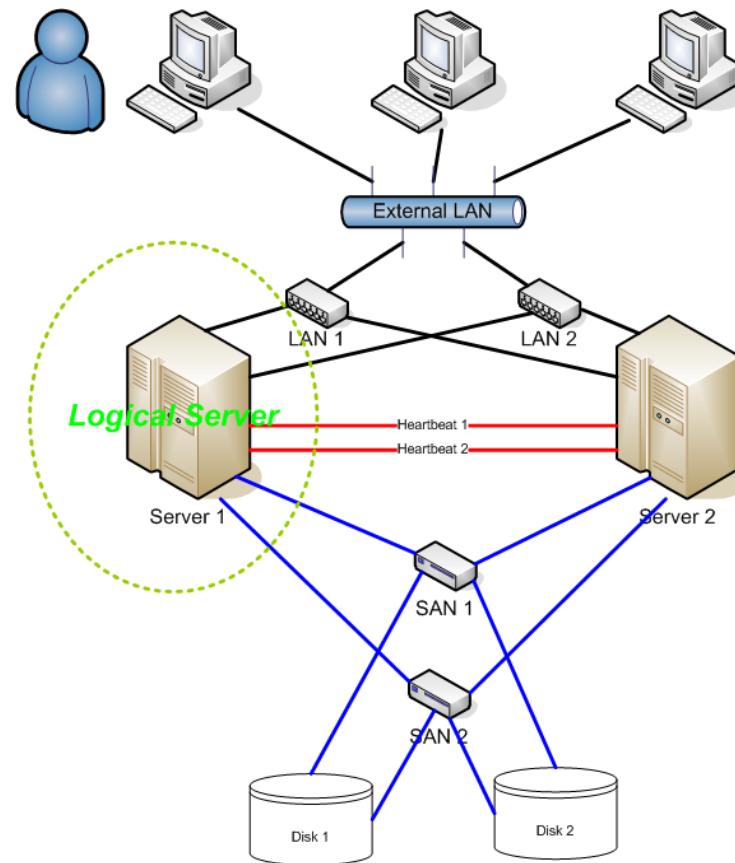


PROTEZIONE CIVILE
Presidenza del Consiglio dei Ministri
Dipartimento della Protezione Civile



What's an HA cluster?

- **High-availability clusters** (also known as **HA clusters** or **fail over clusters**) are groups of computers that support server applications that can be reliably utilized with a minimum of down-time.



en.wikipedia.org/wiki/High-availability_cluster

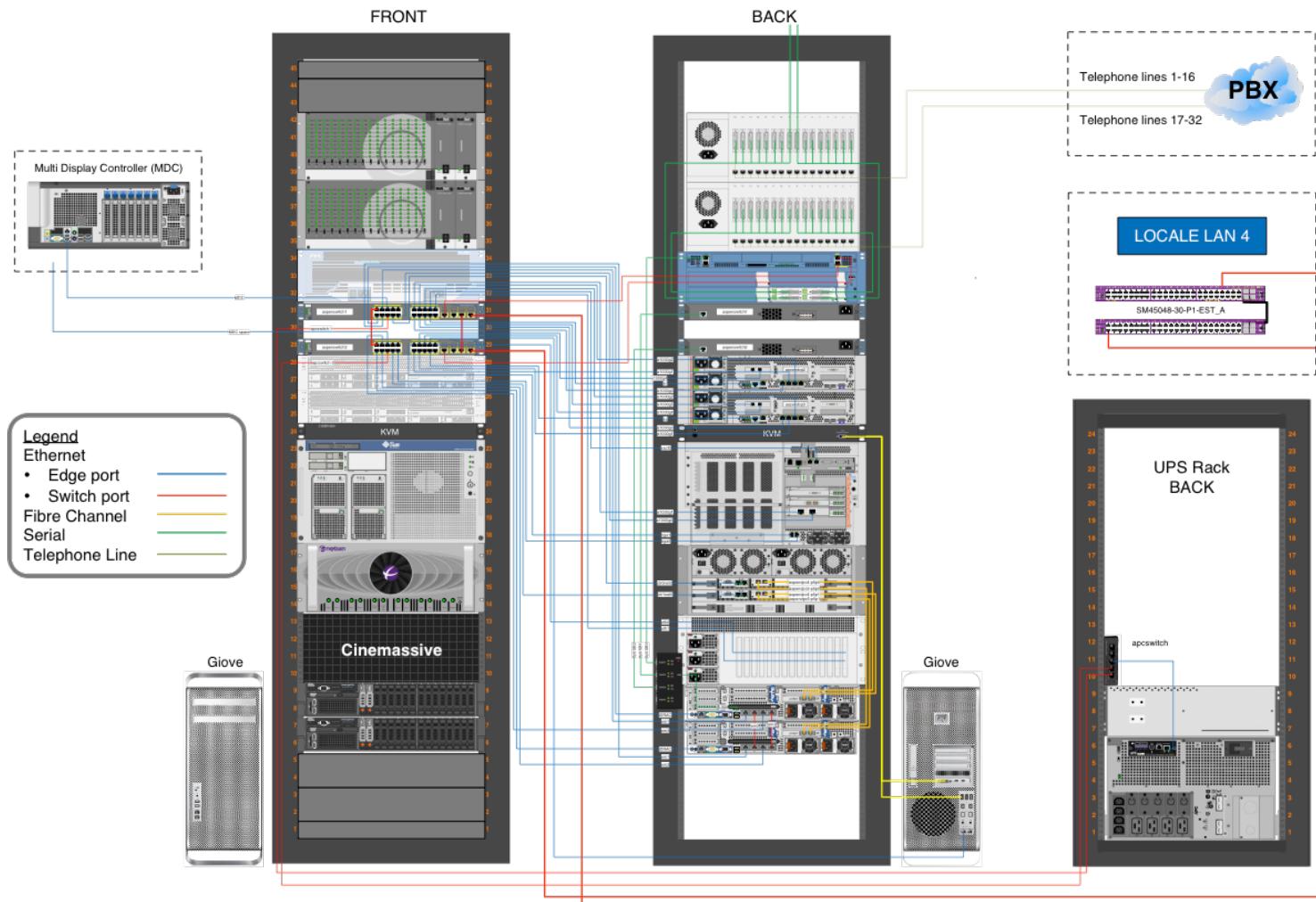
Features

- Redundant computers
- Detect hardware/software faults
- Eliminate single points of failure by design
- Resolve resource dependencies
- Restart failed services and applications
- Reduce downtime
- Increase RAM
(Reliability/Availability/Maintainability)

Infrastructure

- 2 x Dell PowerEdge R720 servers
 - Lot of memory (128G each)
 - Fiberchannel HBA
- Nexsan SATABeast
 - Redundant controller
 - 4 Fiber channel ports (direct attached)
- 2 x Cisco 2960S switches
- Managed UPS with „redundant Ethernet“

Infrastructure



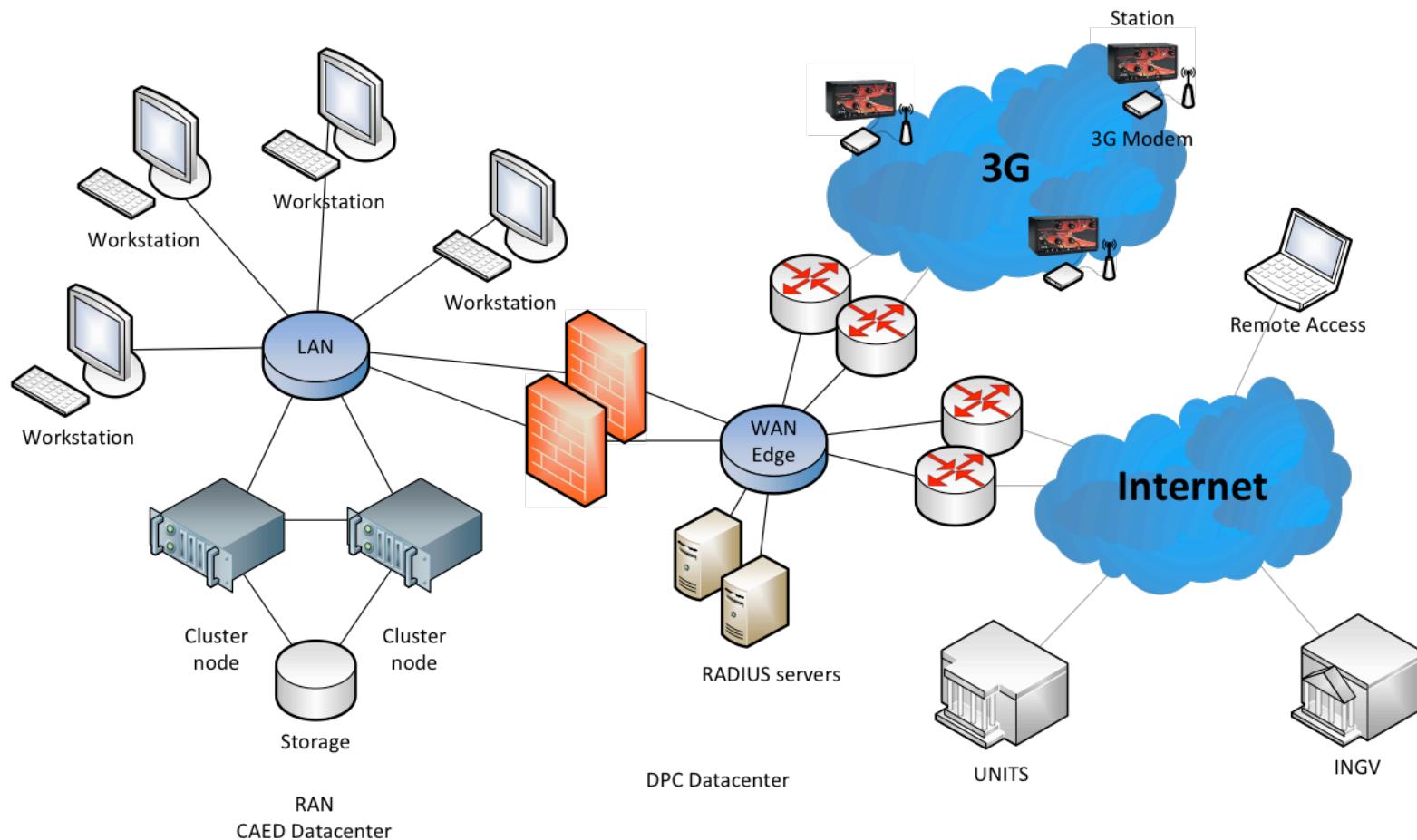
Servers and Storage

- Dual Dell PowerEdge R720 servers
 - Red Hat Enterprise Linux 6.5
 - Redundant power supplies (1+1)
 - Redundant boot subsystem (RAID1)
 - Redundant network interfaces (2 x 2)
 - Dual-CPU
 - ECC Memory
- NexSAN SATABeast
 - Redundant power supplies (1+1)
 - Redundant controllers (2)
 - Redundant fiber-channel paths (4)

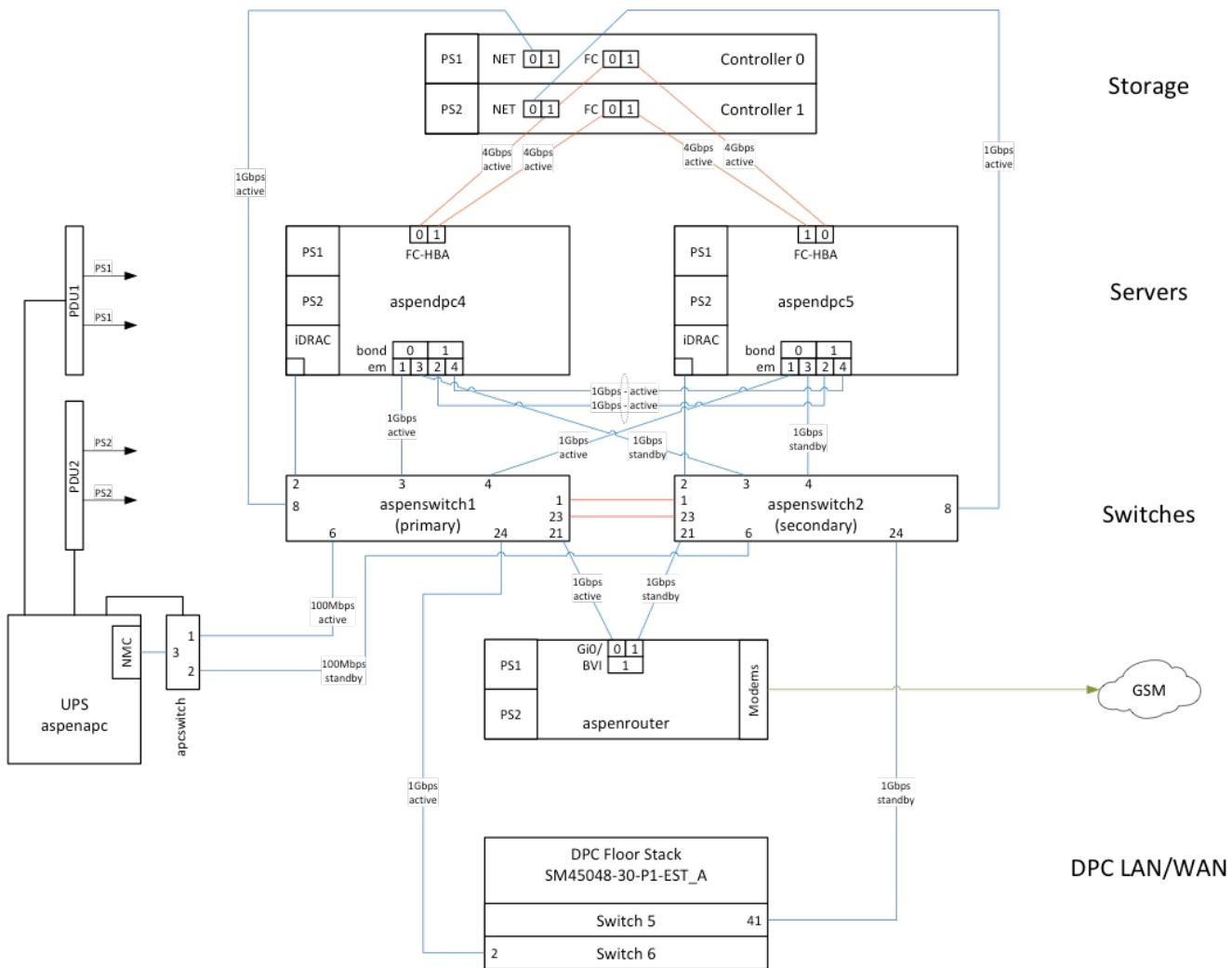
Network & Infrastructure

- Dual Cisco 2960 switches
 - Redundant paths (Rapid Spanning Tree)
 - Single power supply
- Dual firewalls
- Redundant WAN (Wide Area Network)
- Managed UPS

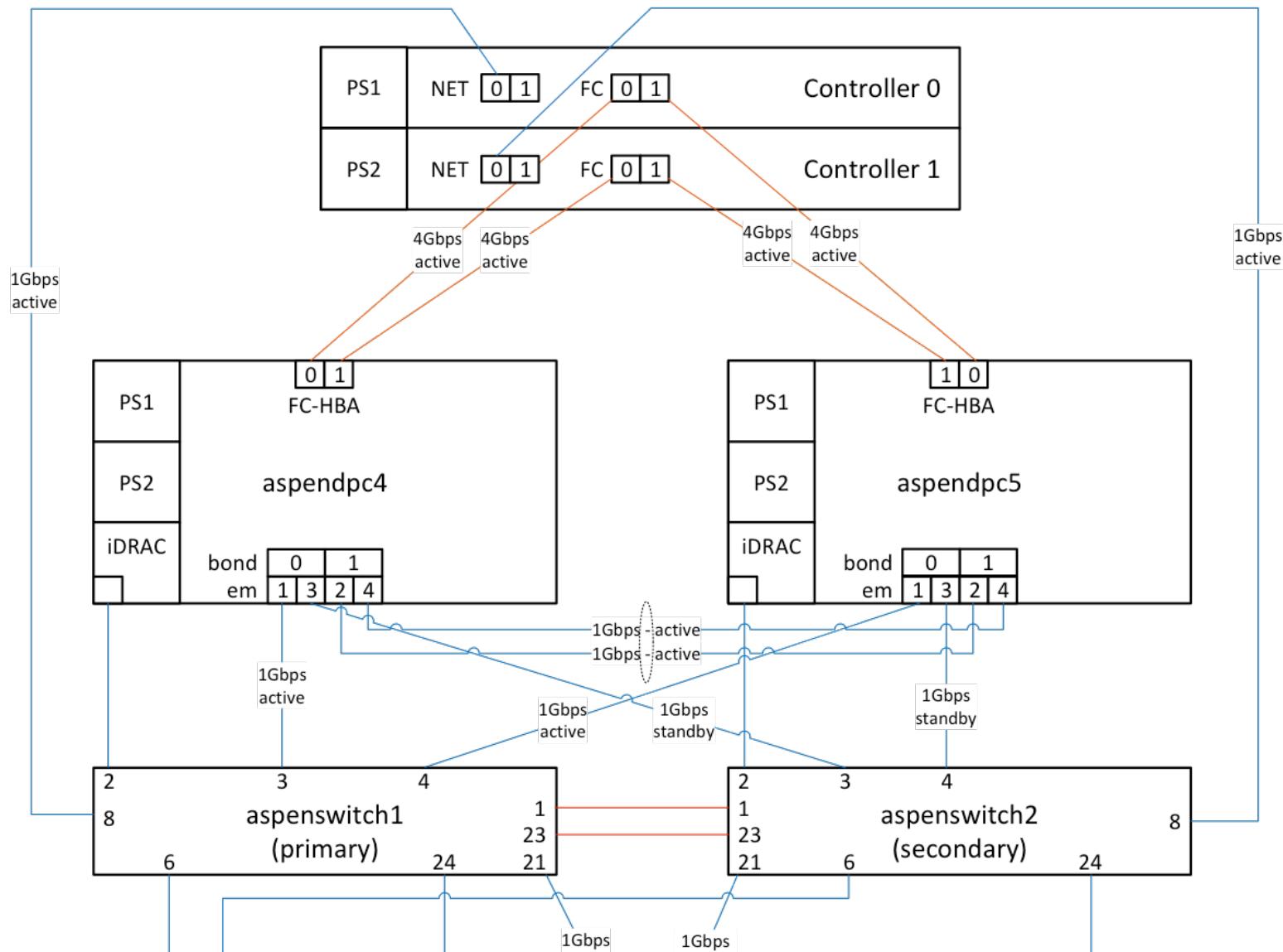
Network & Infrastructure



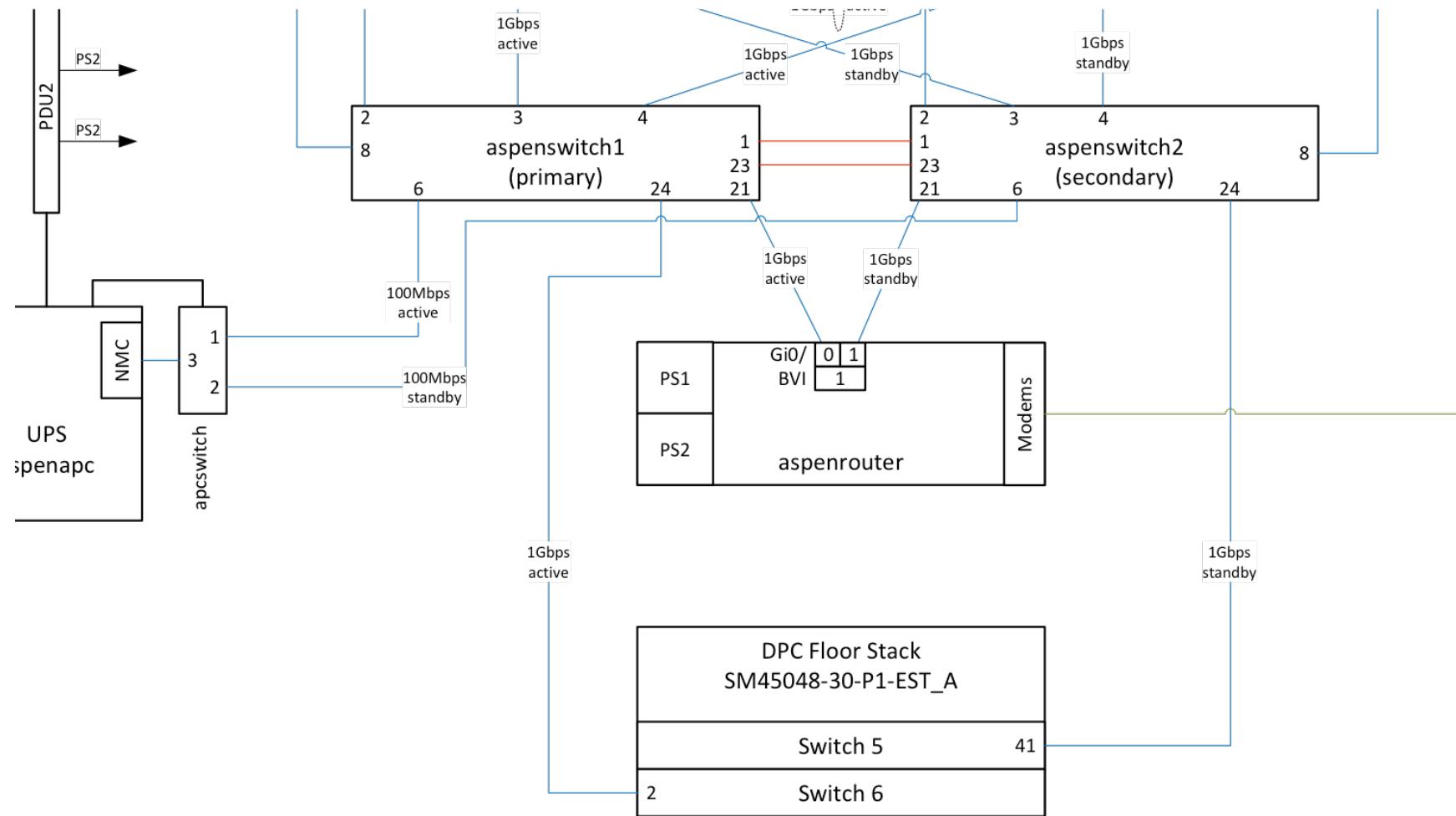
Redundancy



Redundancy



Redundancy



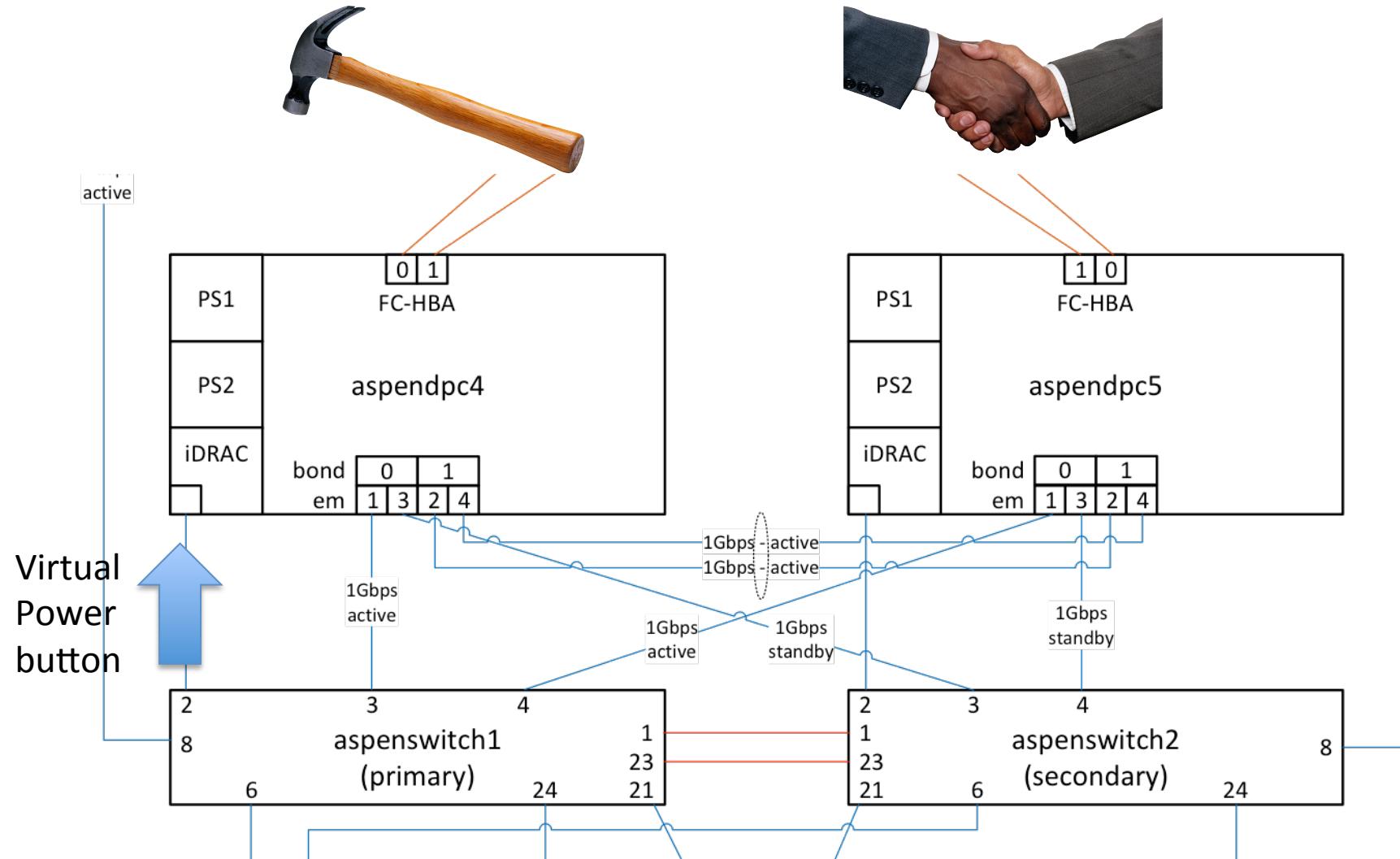
OS & Cluster software

- Red Hat Enterprise Linux 6.5 Server
- Red-Hat Enterprise Linux High-Availability Add-On
 - Corosync cluster engine
 - complex cluster framework
 - maintains communications and
 - in-memory cluster database
 - Pacemaker resource manager
 - Single configuration tool „pcs“
(pacemaker/corosync configuration system)
 - Resource agents
 - OCF (Open Cluster Framework)
 - LSB (Linux Standards Base)
 - Fencing

Fencing

- Isolate a mal-functioning node
- STONITH = **S**hoot **T**he **O**ther **N**ode **I**n **T**he **H**ead
- Last resort for unresponsive nodes or resources
- Establish a well-known-good state (after boot)
- Utilize built-in server hardware (iDRAC,BMC)

Fencing



Antelope Failover

- LSB 3.2.0 compliant init script
(LSB = Linux Standards Base)
- Start/Stop/Status
- Customized S99antelope script
- Pacemaker resource agents
 - Virtual IP address (ocf::heartbeat:IPAddr2)
 - Storage (ocf::heartbeat:LVM)
 - Filesystem (ocf::heartbeat:Filesystem)
 - Antelope (lsb:antelope-*)

Sample Antelope Subsystem (antelope-proc)

- /etc/init.d/antelope
- /etc/init.d/antelope.pl
- /etc/init.d/antelope-proc
- /etc/sysconfig/antelope-proc
- Pacemaker Resource Group: RANPROC
 - ranproc-IP (ocf::heartbeat:IPaddr2)
 - VolGroup02 (ocf::heartbeat:LVM)
 - VolGroup02-Aspen (ocf::heartbeat:Filesystem)
 - antelope-proc (lsb:antelope-proc)

Antelope Failover

- Pacemaker checks Antelope subsystem status
- When Antelope stops it is restarted, if it fails too often it is restarted on another node
- All resources needed by Antelope subsystem are moved with Antelope (resource group)
- If a resource (group) cannot be stopped the node is fenced (rebooted/reset) before the Antelope subsystem is started on the other node.

Cluster Caveats

- Fencing is a (useful) beast
 - Timeouts
 - Turn off STONITH during testing
 - Avoid power off fencing (try reboot instead)
- Resource groups
 - Keep it simple
 - Avoid large resource groups (long chain)

Management & Monitoring

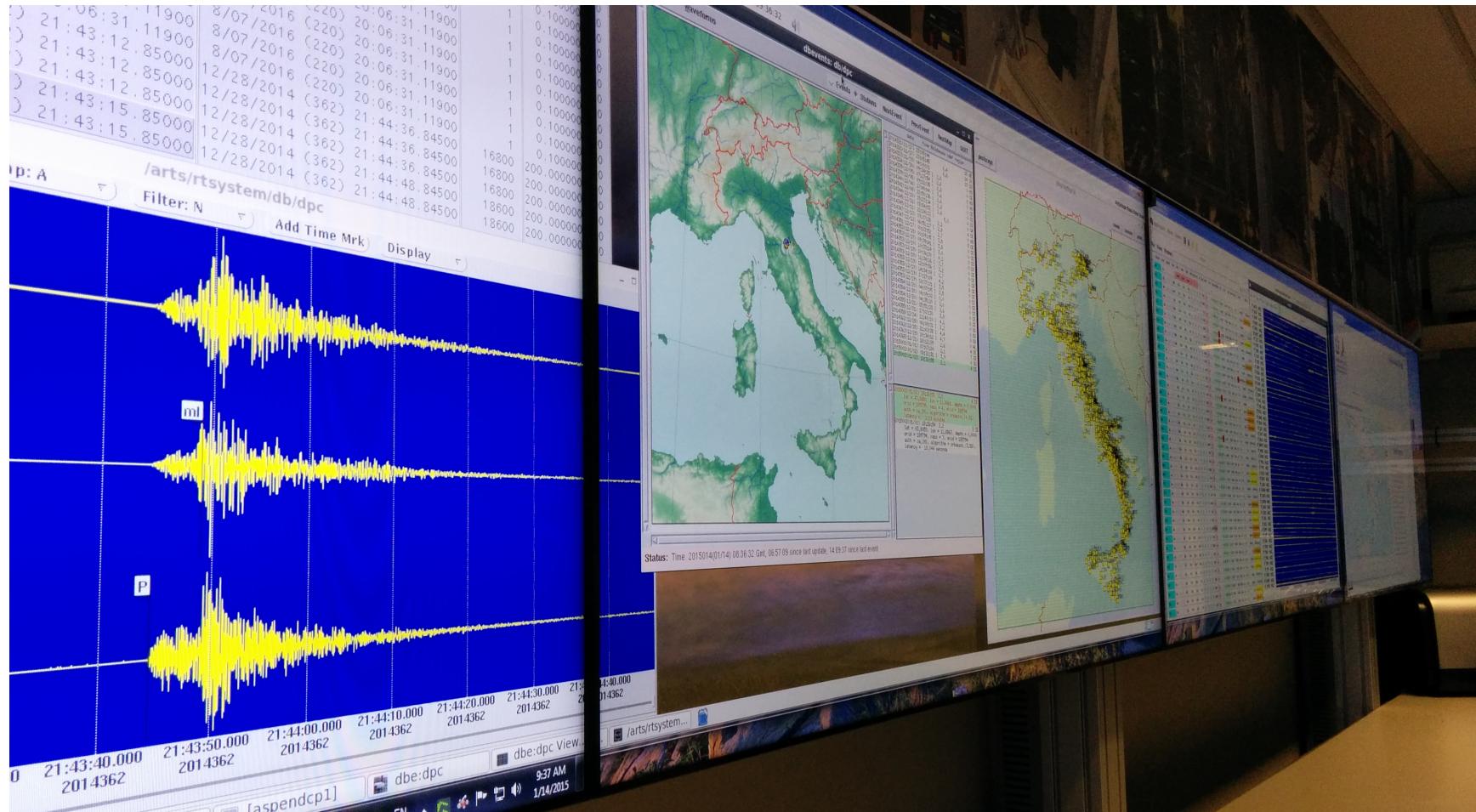
- Web-based component managers
 - Dell Remote Access Controller
 - Storage controller
- OpenManage Server Administrator
- Zabbix
 - IPMI and SNMP based hardware monitoring
 - System monitoring
- Alerting (E-mail)
- PowerChute Network Shutdown

Management & Monitoring

The collage displays several software interfaces used for system management and monitoring:

- ZABBIX Monitoring Dashboard:** Shows performance overview graphs for RAN Servers, including CPU Utilization (SAR) for ASPENDPC4 and ASPENDPC5 over a 1m period.
- Nexsan Technologies UPS Configuration:** UPS configuration interface showing UPS Blower status, Controller temperatures (35°C), and Fan Tray Blowers.
- PowerChute Agents:** UPS configuration interface for APC Smart-UPS RT 10000 XL, displaying battery capacity (100.0%), load in watts (33.0%), and recent UPS events.
- Dell OpenManage Server Administrator:** System configuration interface for PowerEdge R720, showing virtual disk properties for PERC H710P Mini (Embedded).
- Cisco Configuration Professional Express:** Catalyst 2960G switch port status and configuration interface.
- CAES Data Acquisition:** Real-time user monitoring interface showing CPU, Memory, and Disk usage.
- dmon:** A terminal window showing system monitoring data, including processes and network traffic.

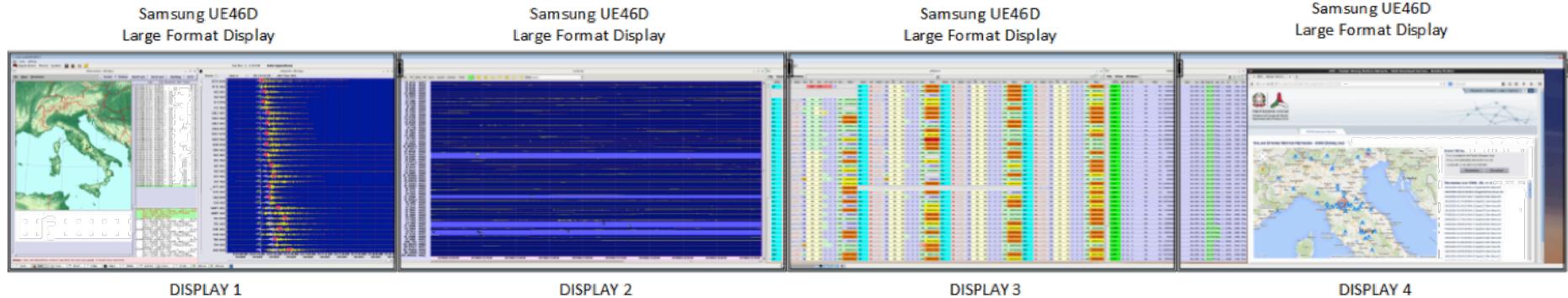
Seismic Monitoring Center



Seismic Monitoring Center

- COTS hardware & software
- Windows Workstation
- X-Windows virtual desktops
- Web-based information systems
- Flexible videowall solution
- Monitoring & control facility

Seismic Monitoring Center



Dell Precision T3610

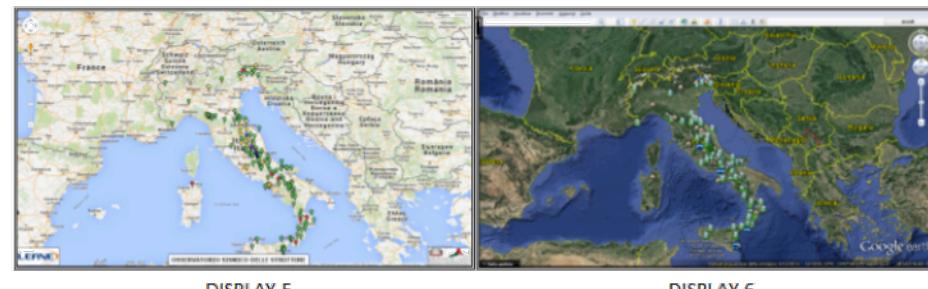


Multi Display Controller (MDC)

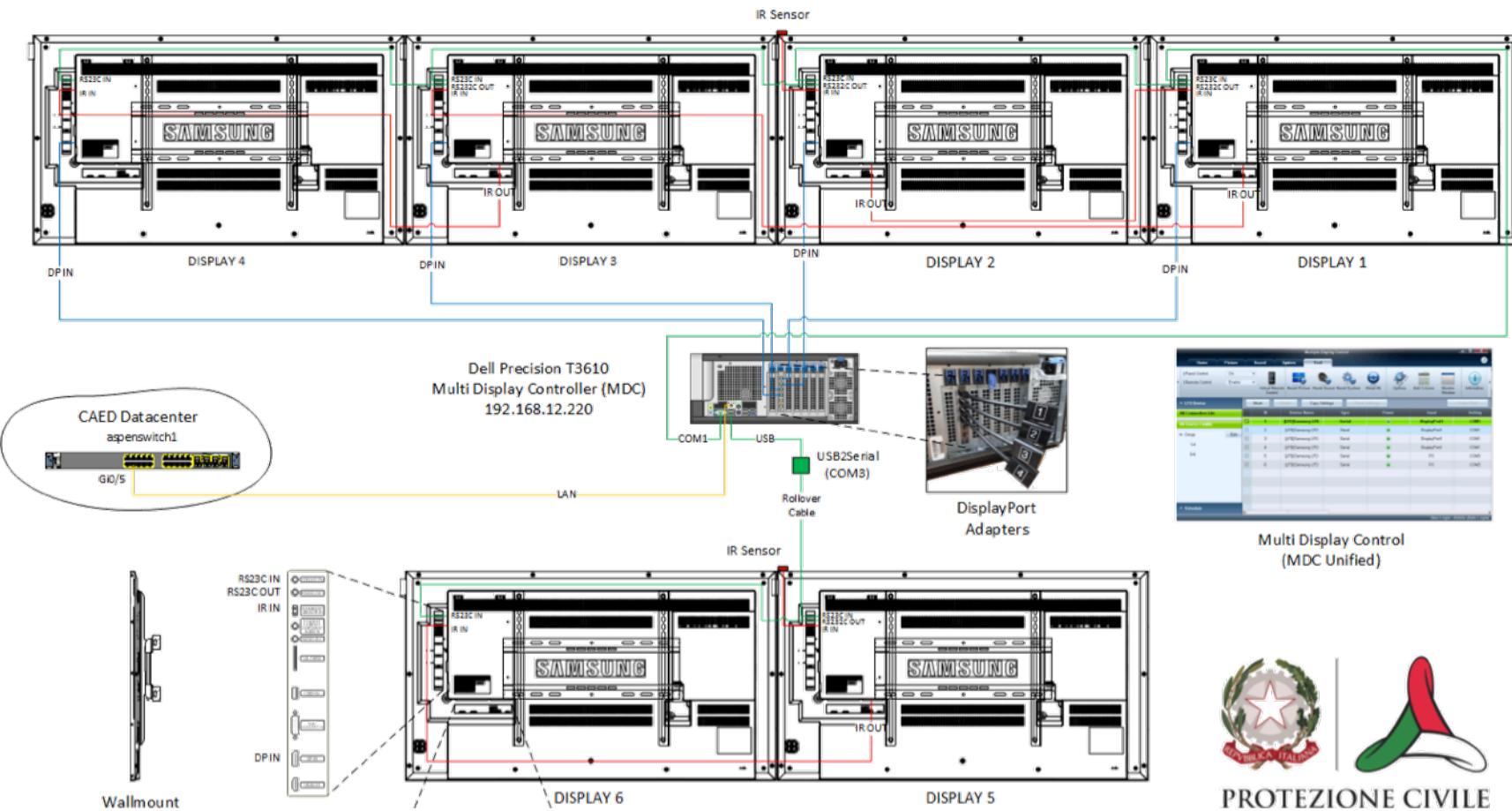


Samsung UE46D
Large Format Display

Samsung UE46D
Large Format Display



Multi –Display Control



THANK YOU



PROTEZIONE CIVILE
Presidenza del Consiglio dei Ministri
Dipartimento della Protezione Civile

*Stefan M. Radman, Consultant - Senior ICT Engineer,
Open Systems & Services, Kinematics Inc.*

