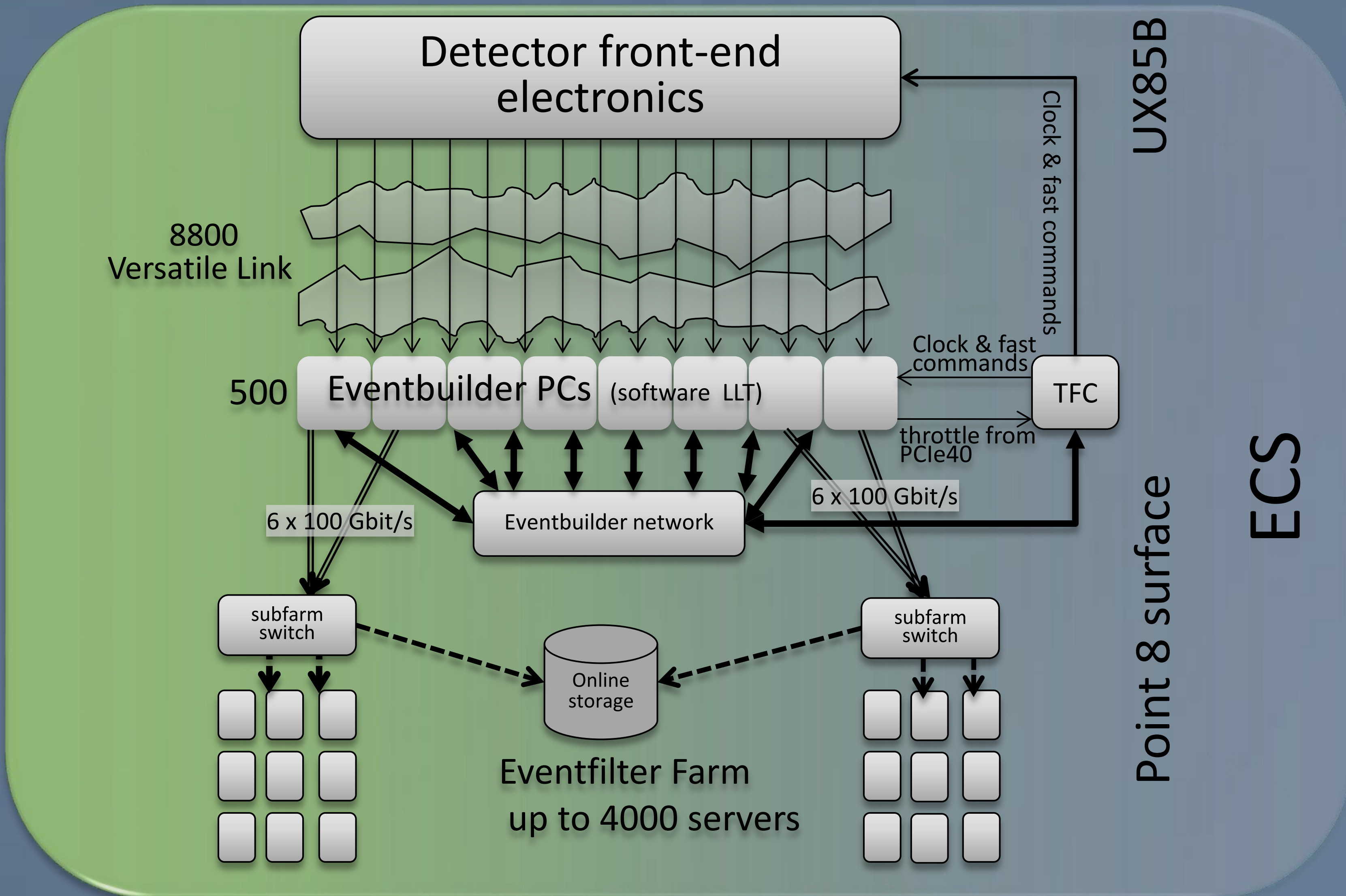
Online system
Infrastructure for LHCb Upgrade workshop – Session II: Cooling, transfer lines, gas, piping
20/2/2015

Niko Neufeld PH/LBC

LHCb ONLINE

Online Architecture



Online system upgrade in numbers

- 9000 links from detector
- Eventbuilder system of 500 – 600 servers and O(10) switches
- Eventfilter farm of up to 4000 servers, will start with ~ 1000 servers, O(100) switches
- Experiment Control System infrastructure, O(100) servers, storage O(10) Petabyte

Item	Power
Eventbuilder server	500 W
Eventbuilder switch	5 kW
Eventfilter server	up to 350 W
Eventfilter switch	300 W
Controls server	300 W
Storage	25 kW
Total	800 – 2000 kW

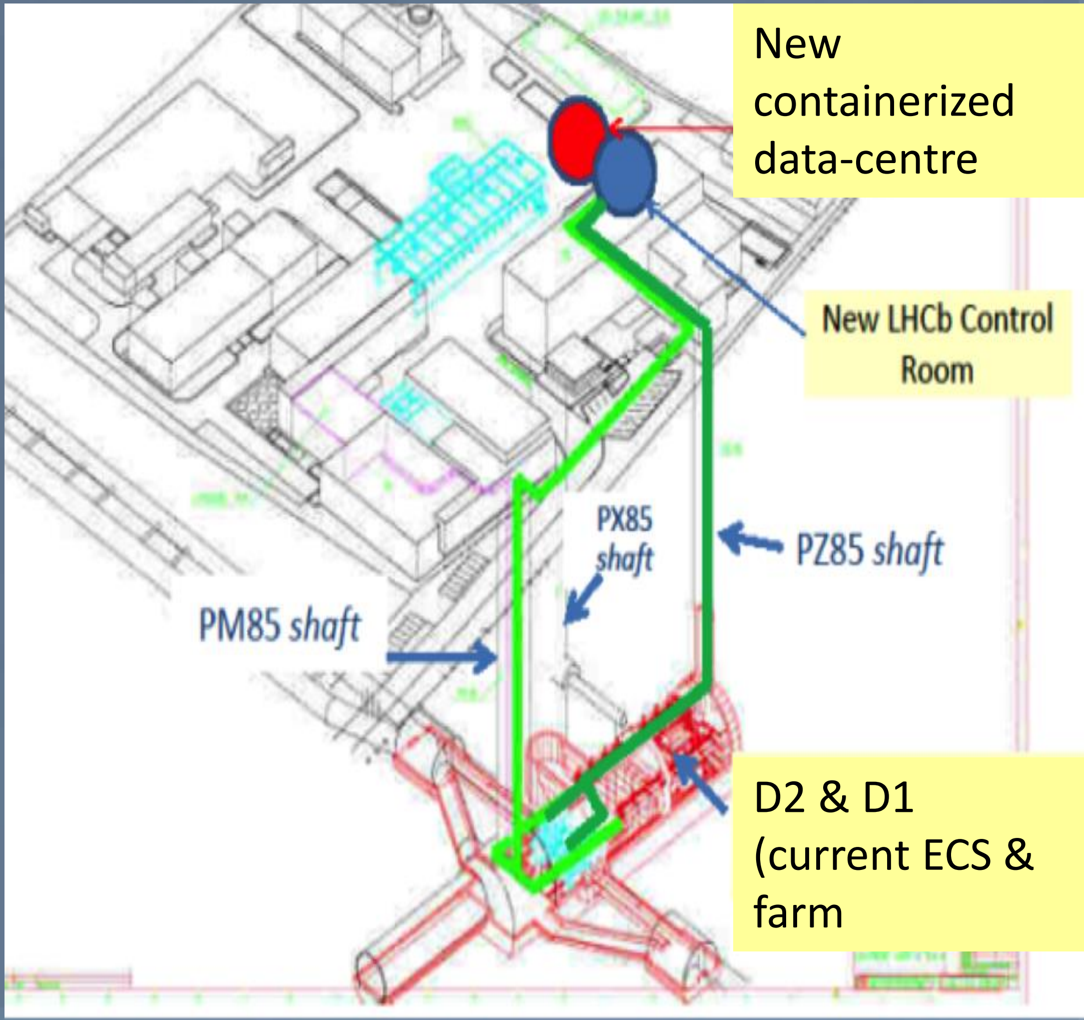
Design considerations in the Online system

- Online system cost is driven by
  - number and type of interconnects
    - shorter → cheaper
    - faster → cheaper per unit of data transported
  - power density
    - higher → fewer racks, shorter distances, more compact systems
  - spatial density
    - higher → fewer containers
- Data-centre operation much easier on the surface in a non-controlled area
  - Current LHCb data-centre is in UX85A
- Data-centre initial cost is definitely lowest for pre-fabricated ("container") solution
- *Data-centre operating cost depends on cooling solution* (and to some extent on CPU technology)

LHCb ONLINE

Location of Online system from LS2

- Most compact system achieved by locating all Online components in a single location
- Power, space and cooling constraints allow such an arrangement only on the surface: containerized data-centre



Cooling needs in UX85A for Online

- Baseline scenario
  - Very little – some ECS infrastructure in D2 / D3 barracks
  - Will continue with heat-exchanger doors using mixed water (max 200 kW) + some small amount of air-conditioning (mostly to keep the air reasonably dry)

- Plan B in case of problem with long-distance versatile link
  - *No indication of any problem*
- In that case need additional 400 kW (available today)
- Mixed-water air cooling (new rear-ddor heat-exchangers)
- Or direct liquid cooling in servers with central heat-exchangers to mixed water

Cooling needs in SX85 server room

- Existing infrastructure
  - 70 kW on EOD (dual-feed, redundant)
  - Used for critical services (ECS and storage)
  - Will be kept for Run3 and beyond. A small increase (30 kW) in the same configuration is desirable
- 100 kW heat-load
- Cooling is done using mixed-water air-cooling in rear-door heat-exchangers
- System will be kept, probably with new (deeper/wider) racks, but same cooling principle
- These services are needed 24/7 → redundancy for the mixed water facility is needed
  - full redundancy with 2nd group?
  - or possibility to quickly connect mobile group?

Cooling for main-data centre

- Main options identified until today:
  - Direct Liquid Cooling (DLC)
  - Natural free cooling (NFC)
  - Water Cooled Heat Exchanger Doors (WCD)
- Racks (42U) will dissipate at least 20 kW
- Not considered:
  - Classical data-centre airco using CRAC units, hot aisle, cold aisle → not very attractive for a completely new, non-legacy, containerized data-centre
- Solution must scale to 2 MW even though not everything needs to be deployed form day 1
- Obviously want to go for a very high PUE (Power Usage Efficiency) == Power for IT / Power for data-centre
- Need cooling from 2017 as containers are being installed

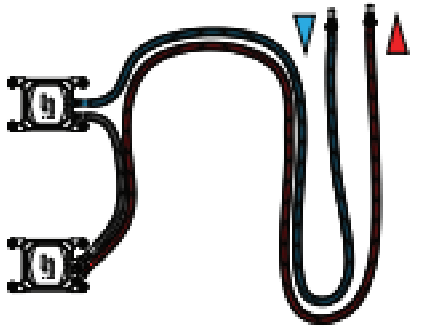
DLC pros & cons

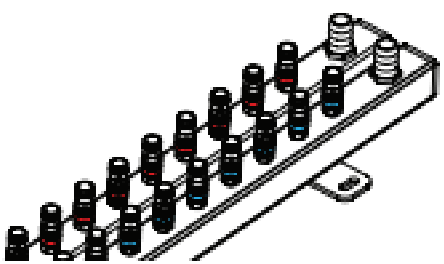
- Easy to deploy
- Can be retro-fitted to exisiting servers
- Saves power by removing / shutting off server fans
- Potential for over-clocking the CPUs (more performance for same capex)
- ‘Excellent thermal stability (custom electronics boards, PCIe40)
- Can work with facility water from 15 to 50 C
- No constraints container layout

- Cost (manifold port + heat-sink + tubes) about 250 USD/server (not including heat exchanger with facility cooling)
- Part of the infrastructure (heat-sinks, tubes) need to be changed for every server generation
- Operational reliability needs to be tested
- Some airconditioning will be required anyhow:
  - switches, memory modules, server power-supplies

Direct Liquid Cooling

- Relatively old solution (comes from the gaming market to the server)
- Prominent in proprietary solutions (e.g. HP Moonshot, IBM BlueGene)
- At least two vendor-agnostic providers: CoolIT, ASTec
- Principle: Liquid in closed loop goes through cooling plates on hot server elements (CPU, Memory, chipset, GPGPU, …)
- Connected via manifold to secondary loop

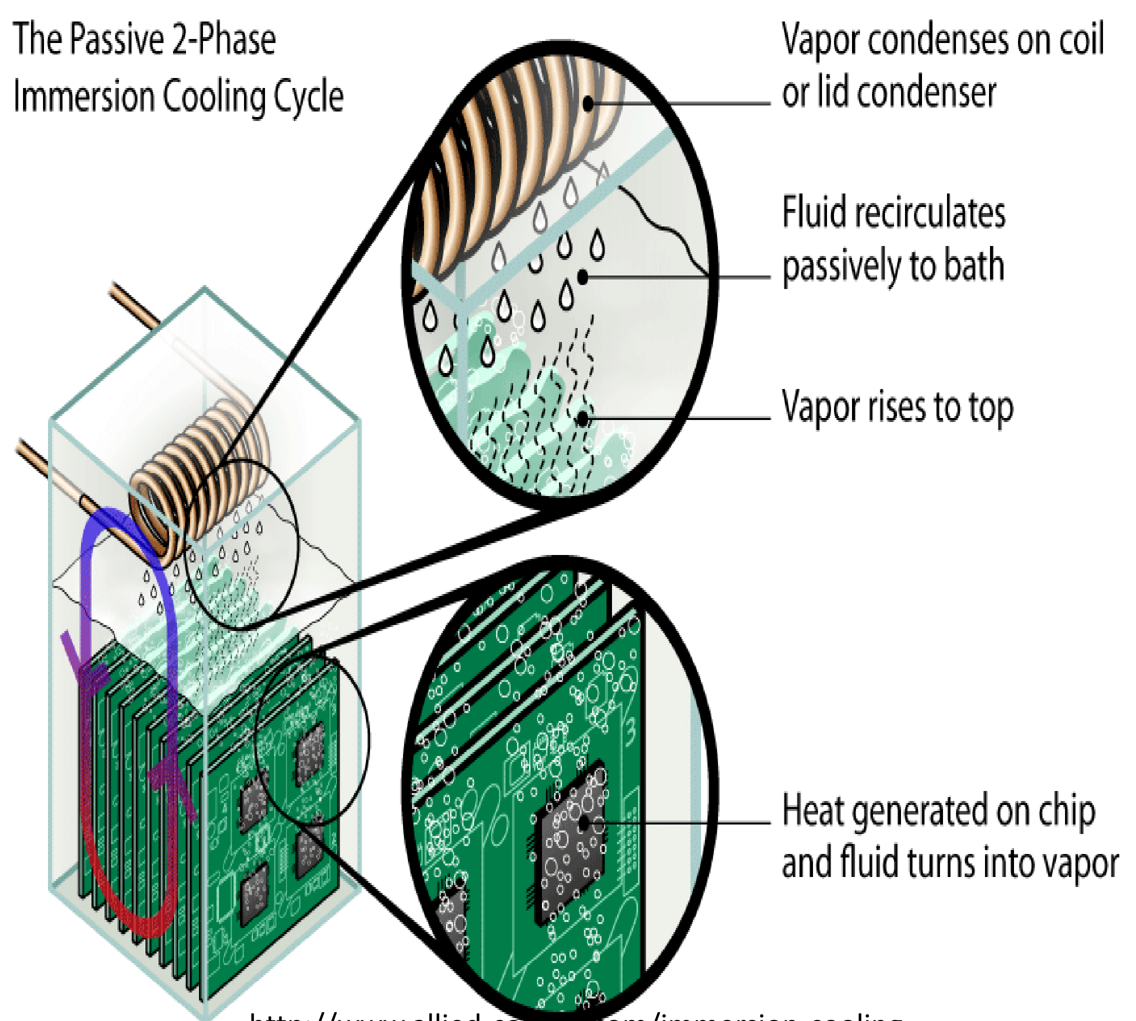




To be investigated for DLC

- Efficient connection to primary cooling at P8
  - which Tin can we get
  - which Tout is acceptable
  - flows
- Operational reliability
  - tests will start soon
- Serviceability
  - dense racks, repairs
- Detailed cost analysis
- Max in / max out temperature to evaporator towers

Immersion Cooling



http://www.allied-control.com/immersion-cooling

- Two-phase immersion cooling uses special low-boiling point fluids such as 3M Novec

Immersion cooling – the good & bad

- High density
- Non proprietary
- Novec and similar materials are non-sticky → equipment comes out dry
- Adapts to any-size (PCIe40!)
- Inherent flame-suppression
- Excellent thermal stability (overclocking possible)

- Density → only accessible from the top, incompatible with current rack-standards (OCP etc…) (this might change with wide-spread adoption)
- Cabling and other routine maintenance difficult – server must be removed, residual liquid
- Sealing of certain components vital (in particular hard-drives)
- Read also: http://www.datacenterjournal.com/it/whats-stopping-liquid-cooling/

Natural free cooling

- AKA as Direct Air Cooling
  - Air-side free cooling, where outside air is brought into the data center directly through filters or indirectly through heat exchangers.
  - Adiabatic in which the air is brought to some sort of chamber and used along with water evaporation to cool the air
- Very “hip” now in data-centres, used by many sites achieving very low PUE (Facebook etc…)

NFC – the good / the bad

- Very low operational cost
- Well suited for Geneva climate (Tmax < 35 C almost always)
- Can be bought ready integrated with many container solutions
- No modification of equipment

- High investment cost
- Will need some heating during some winter periods and certainly some drying – humidity control
- Requires careful control and monitoring (SCADA) – usually comes with the solution
- Puts constraints on rack and container layout
- Very inconvenient to work inside the racks

Summary

Location	Max. Load [kW]
UX85A D1 & D2	200
2885-R-007 (SX85)	70 + 30
Data-centre S8	2000

- The upgraded LHCb online system has significantly larger power-needs than the current one, in particular in the event-building and event-filter.
- All new power-needs are in the to-be-built new data-centre → cooling solution the most important property of this project
- For existing, re-used location current power and battery backup are sufficient