XFEL DAQ and control network infrastructure

Minutes of the 27th March 2008 meeting (Revised 3.4.2008)

1. Present

DESY-IT: Kars Ohrenberg, Thomas Witt, XFEL: Thomas Hott. WP76: Sergey Esenov and Christopher Youngman

2. Aims

Startup discussion of DAQ and control (WP-76) network related infrastructure requirements. The aim is to define what is required so that this information can be used to plan the hall (XHEXP1) and photon beam line tunnel installation requirements (space, power, cooling, cable & fibre routing, etc.), involve and inform other related groups such as DESY-IT, and eventually allow a cost estimate to be made.

Infrastructure information is required on the following time scales. For space in the hall (expt. floor and rooms) soon, for power and cooling by June 2008, for cable/fibre routing by end of 2008, and for the budget soon.

3. Associated information

Slides showing building and tunnel plans, known milestones and anticipated DAQ and control network requirements were used to guide the discussion. These slides are included in an associated document. Some of the numbers used in the discussion section are explained in the slides.

4. The discussion

The following model for network connectivity was assumed. That the Schenefeld hall and DESY-IT sites are connected by optical fibres. That the hall acts as a hub for all beam line and diagnostic systems in the tunnels whose connections run back to the hall. That within the hall and beam lines copper connections would be used when the distance restriction of 90m (category 7) is satisfied. That 10 Gbit/s (10 GE) ethernet connections are standard.

4.1. DESY-Schenefeld connection

There are no fundamental problems associated with connecting XHEXP1 with DESY-IT. The total length is considerably less that 10 km (XTIN-XS1 and XS1-XHEXP1 distances are 2.1 and 1.4 km, respectively) and single mode fibre can be used. XFEL requires connections for the XHEXP1 office network and for the DAQ systems for control and data archiving, both of which are discussed later.

EDMS Nr.: D00000001071551 Rev: A Ver: 1 Status: Released Dat.: 08.04.2008

In the XTIN-XS1 tunnel it is planned to install 1200 pipes into which fibres can be installed (blown) when required, this installation is shielded from radiation by the concrete floor. To acquire an allocation within the 1200 requires negotiating with other groups, known contact names are: T. Hott, K.Rehlich, A.Winter and L.Frohlich. As of Nov. 2007 space remained in the allocation.

The route connecting XS1-XHEXP1 has to be defined as more than one tunnel is present. SASE1 might be the best candidate as it will be completed first and does not have additional undulator sections. It is not planned to have under floor concrete shielded installation paths in the XS1-XHEXP1 tunnels and wall mounted cable trays should be used. Whether additional shielding is required needs to be investigated, other tunnel users like the undulator group may have useful information.

A guesstimate of the bandwidth required for DAQ archiving is 50GB/s. The number of fibres needed for 100GB/s transfer bandwidth today is 100 x 10Gbit/s lines = 200 fibres without or ~50 fibres with wavelength-division multiplexing (WDM/DWDM) technology, see section 7 . In 2012/3 this should be 10 x 100Gbit/s lines = 20 fibres without or ~2 fibres with WDM/DWDM. As WDM is expensive today's solution would raise the question of whether to install archiving equipment in the XHEXP1 hall rather than remotely at DESY-IT.

The time scale for fibre installation is driven by when the buildings are defined as ready for installation. The current dates, which may change, are: XTD Apr.2011, XTD6-10 Feb.2012, XHEXP1 underground May 2012 and XHEXP1 surface Dec.2012, with first SASE1 beams end 2014.

4.2. Beam line connectivity

It is assumed that all beam line and diagnostic systems require network connections for control and DAQ. A cost efficient and flexible mechanism for connecting these systems would be to connect them to switches located at periodic intervals in the beam lines using copper cables. Each switch spans 180m of tunnel and is connected to the hall via fibre directly (or possibly cascaded to the next switch). Cable trays mounted on the walls would be used to route cables to the switches and to route the fibres between the switches and hall end points, see below.

The following points need clarifying: the space required for the switch installation and whether additional shielding is required. On a longer timescale the following need clarification: the locations of the beam line and diagnostic systems, their data bandwidth requirements, the locations of the switches, and the number of ports required.

4.3. Experiment connectivity

The floor area in XHEXP1 underground is 50m long in the direction of the beam line. Instruments can be connected by copper cables running to a switch. The location of the switch could be close to the beam entry point into the building, which allows connections from devices in the last 90 m of tunnel, or at the back of the hall. In either case the cables should run along a ceiling hung cable tray, which additionally provides a route for the tunnel switch connection fibres.

Two types of detector are foreseen: small systems like a camera or a strip pixel detector requiring a few network connections, or the large detectors like the 2D-pixel detectors where the number of connections will depend on their backend implementations. Two backend solutions are being discussed, an intermediate hardware, e.g. ATCA crate train building layer in front of a smaller scale PC farm, or a large PC farm train builder. The intermediate layer requires a few tens of connections, whereas the full PC farm would require 200-1000 connections. In both cases the PC farms should not be located in the experimental hall but somewhere else in a dedicated room where all farms can be located, shared and maintained. The space required for the farm depends on the implementation, it might be a single rack, but might be more.

Given the potentially large number of switch ports and the point-to-point connection requirement a cable/fibre shaft must be foreseen as a flexible, i.e. can accommodate changes in the requirements over the lifetime of XFEL, and efficient way of connecting the required rooms in the various floor levels.

The following points need clarifying: the locations of the switches (front or back balcony) and the cable tray path. On a longer timescale the following need clarifying: the number of ports required (which depends on the implementation of the large 2D-pixel detector backend), and the space required for the switch installation.

4.4. Hall beam line end points

All fibres coming from a given beam line and its associated experiment area terminate in a switch rack installation located on the balcony on the front or back wall of the hall.

The number of racks required depends on the number of switch ports handled. A 1000 port installation requires two cable patch panel racks and one switch rack. The racks are wider and deeper (1.5m) than standard 19 inch electronics racks and sufficient space has to be provided on the balconies for them and any work performed on them. A full switch rack typically generates 2 kW of heat, which can be air cooled, but might require water cooling if the generated heat is not allowed to be released into the hall.

4.5. Schenefeld local site data cache

A data cache for staging at least two days of data is required to cover non availability of the principle store. It is currently assumed that a single cache provides data storage for all beam lines.

Knut Woller (DESY-IT) provided the following possibilities summary for this type of installation with 20GB/s (3500 TB of disk storage). Today's solutions (SunFire x4500, Nexsan Satabeast) provide approximately 350 TB of storage in a rack using 1TB disks. Similar density is now available with better throughput and lower price. The industry forecast for 2013 is for 50 TB disk capacity. It is probably safe to assume 10-20 TB will be deliverable, which means that 3500 TB (3.5 PB) will be storable in a single rack with an additional rack required for file servers and high speed networks. The racks used are water cooled as the heat load is large. Additionally sufficient space has to be provided for access, cabling and installation.

4.6. Hall office network

For the office network infrastructure DESY-IT requires a 15 m^2 room per floor, i.e. three rooms for three floors. The usual office infrastructure: cable paths to each room and wireless installation are also needed.

4.7. Network partitioning

A simple solution would be to have an office and a control and DAQ partition. The final number of partitions can be defined later.

5. Conclusions

The network infrastructure model described in at the beginning of the discussion appears to be realizable. Note that no safety margins or backups have been assumed.

The installation of switches spaced periodically in the tunnels to connect beam line and diagnostics instruments is flexible and efficient, it does not require knowing where the instruments and what their bandwidths will initially be within limits. The space and shielding requirements need clarifying, it is assumed that the switches are air cooled.

The hall beam line end point switch racks act as connection points for the tunnels switches and the instruments in the respective hall beam line. The space requirement for each end point is 3 racks per 1000 port beam line installation and enough space has to be provided on the (front) balcony for this. A decision is required as to whether the racks have to be water cooled.

Cable trays are required in the tunnels and hall beam lines to connect switches and experiment instruments to the end points.

Space is required for a data cache system. Projecting to 2013 the cache can be implemented in two water cooled racks.

Space for PC farms has to be planned and a room will be required, this might be the same room as used by the data cache as the cooling requirements will be similar.

The space required for the general office network infrastructure is one 15 m^2 room per floor.

Assuming that the data cache and PC farm rooms are as close to the experimental hall as possible and that the IT general infrastructure rooms are above each other a cable shaft linking them and the to the hall would be sensible.

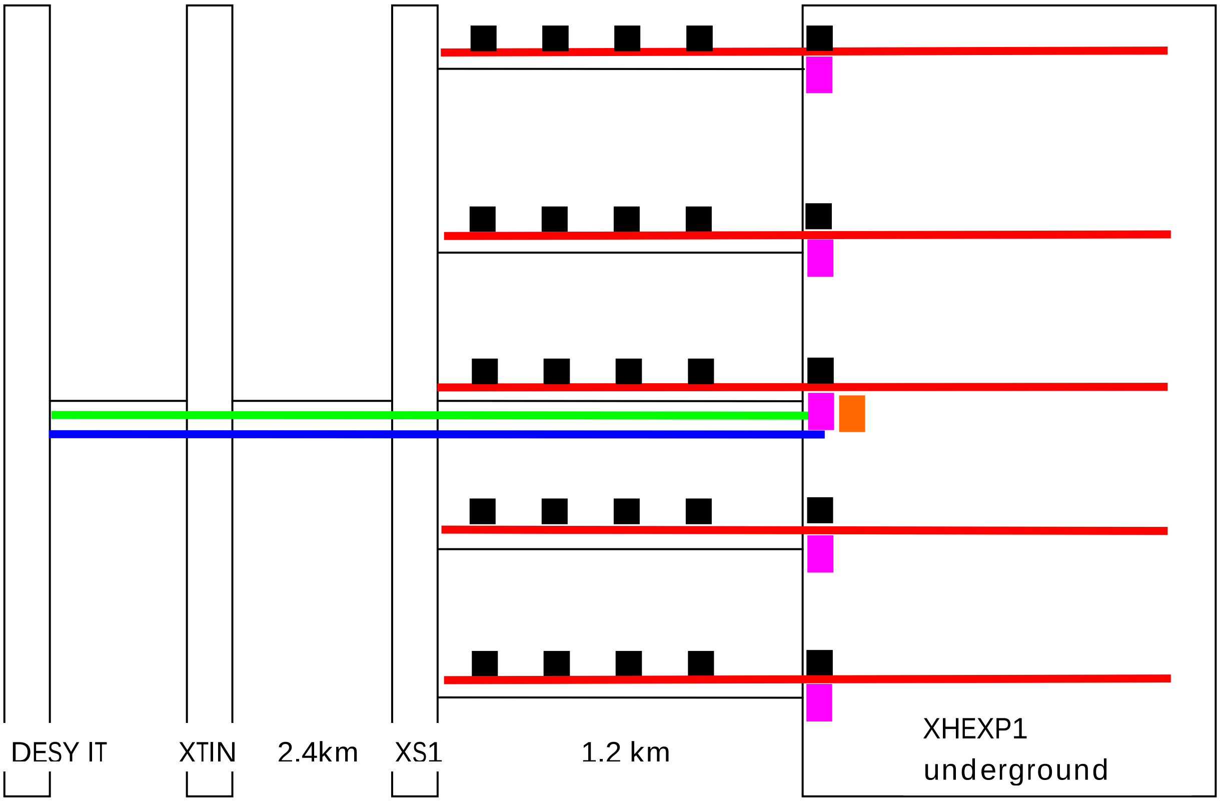
6. Schematics derived from layout discussed

The network layout foreseen is shown schematically in the following sections. Note that the room space required by the data cache and PC farms has to be determined.

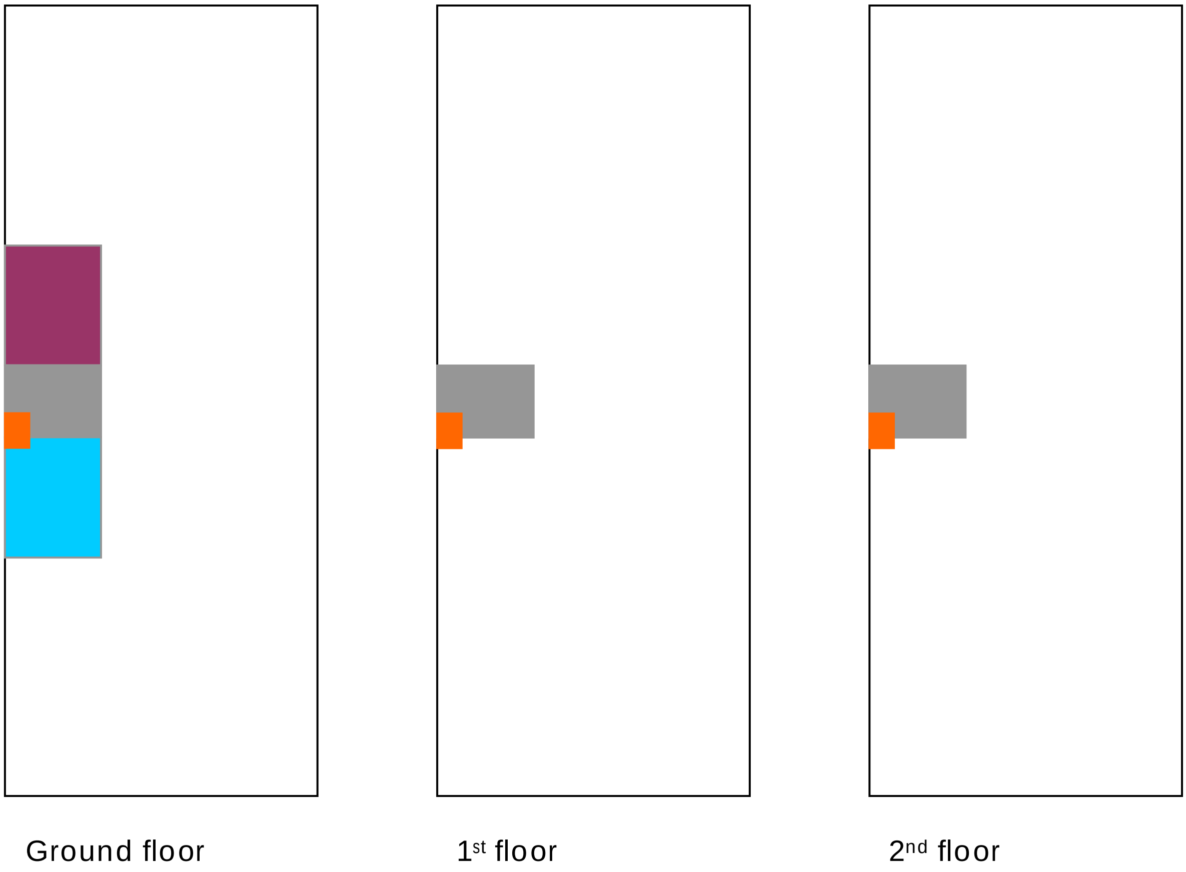
6.1. Key

- switch
- switch and patch racks
- Cable connections
- Fat data archive
- Hall office
- Cable/fibre shaft
- Data cache room
- PC farms room
- IT office infrastructure room (15

6.2. Tunnels and hall underground

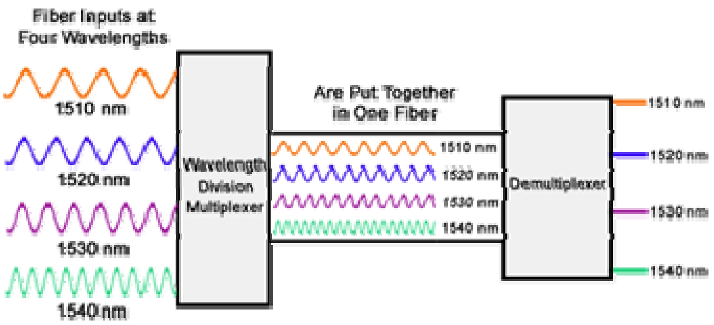


6.3. XHEXP1 upper floors



7. What are WDM and DWDM?

WDM and DWDM involve combining multiple wavelengths of light into a single fiber. For each "channel", a separate source laser is required that emits a slightly different wavelength. Each signal or "channel" carries a separate voice or data transmission. Once the signals are generated, they must be combined into a single fiber. There are several different devices that have been developed such as arrayed waveguide gratings (AWGs), graded index lenses, planar waveguide integrated circuits, diffraction gratings, Bragg filters and multiplayer bandpass filters.



Source:http://www.williams-adv.com/markets/wdm-and-dwdm.php