Object Storage for The FIRE Archive

Introduction

EMBL-EBI is seeking to increase its object storage capacity supporting its FIRE archive; Exploring the Infrastructure of EMBL's European Bioinformatics Institute (EMBL-EBI). This storage will replace 10PB of end of life storage, and provide capacity for the substantial data growth expected over the coming year.

Year	Daily Growth	Yearly Growth	Archive Size	Data Size
2021	62 TB	22 PB	67 PB	201 PB
2022	94 TB	33 PB	100 PB	301 PB
2023	140 TB	50 PB	151 PB	452 PB
2024	211 TB	75 PB	226 PB	677 PB
2025	316 TB	113 PB	338 PB	1015 PB

Specification

Vendors should respond using the Object Storage for FIRE - Technical Requirements Response Sheet and provide any other documents and diagrams as requested.

Validated Solutions

Vendors may only bid object storage products validated against the FIRE archive. The FIRE archive has to date been validated against the following 2 appliance based object storage platforms:

- Dell ECS - EX5000D
- Quantum X200 - Twin controllers

EMBL-EBI will validate other object storage platforms in future.

Mandatory Requirements

General

1. The solution must be conformant with the Amazon S3 API
2. The solution must be fully licenced for use by EMBL-EBI without limitation
3. The solution must be an appliance based solution, not a software defined solution
4. All hardware, software, upgrades and licencing components of the system must be covered under a 5 year support direct to the hardware vendor, including next business day onsite-support for all hardware components
    a. The 5 year support period will commence following the acceptance of the system as detailed under Delivery & Acceptance below
5. Serial numbers for each hardware component must be provided ahead of delivery, broken down by rack and intended location (RU)

Delivery & Acceptance

6. Detailed proposed racking and connectivity diagrams for the solution must be provided as part of the tender response, including details of the count and type of connections used

    a. The vendor or integrator will conduct an on-site survey to validate the racking and connectivity diagrams and ensure compatibility with the environment, the final design must be validated and approved in consultation between EMBL-EBI and the vendor before delivery or installation can commence
7. Acceptance of the solution will occur after:
    a. The total usable capacity of the solution as described in the proposal has been installed and commissioned
    b. The system has been upgraded to the latest stable firmware and operating system versions
    c. All failed components and disks have been replaced
    d. EMBL-EBI completes performance testing of the solution using the performance metrics as defined in Object Storage for FIRE - Technical Requirements Response Sheet, validating that the solution performs to with 80% of the values quoted, or to the limit of EMBL-EBI network connectivity, whichever is lower
    e. The monitoring described in the technical requirements is delivered and confirmed by EMBL-EBI
    f. Delivery, installation and acceptance of the solution must occur before January 31st 2023
8. The vendor or integrator must name a domestically based project manager that will oversee the entire fulfilment process from order to acceptance, coordinating deliveries and ensuring work is completed on time
9. Installation and configuration processes are to be completed by domestically based engineering teams wherever possible, the vendor and integrator must identify as part of the proposal which engineering teams will be assigned to deliver the solution and where they are located

Capacity & Performance

10. The solution must be deployed in a 3 GEO (3 geographically separated data centres) configuration
11. The solution must be tolerant to the failure of any of the 3 data centres by leveraging geo-distributed erasure coding
12. The solution as bid supports the creation of at least 500 buckets
13. The solution as bid supports the creation of at least 100 mil. objects per bucket
14. The solution as bid supports the creation of at least 500 mil. objects globally
15. The solution can scale to support the storage of 100 petabytes of data or more
16. The solution supports object sizes over 4TB
17. The solution must provide real-time bucket level quotas, enforceable based on total bucket size
18. The solution must provide a minimum of 33PB usable capacity
19. The solutions is provided in one of the following EC configurations:
    a. X200: 18/7
    b. EX5000D: 12+4

Networking & Connectivity

The vendor must provide detailed and complete racking and networking diagrams as part of their response.

20. The network configuration of the solution must be redundant, and able to cope with the failure of any one switch within the solution with no impact to connectivity
21. The vendor must provide all switches, cables and transceivers required to complete connectivity of the system as described in the bid, and to clear acceptance testing
22. The solution must be uplinked to the EMBL-EBI general purpose network with a minimum of 100Gbit aggregate bandwidth per site
    a. General purpose network switches are Arista (7000 series)
        i. Where required Arista branded or guaranteed Arista compatible transceivers must be supplied to complete the uplink of the system to EMBL-EBI switches, for reference the correct part number is 25G part - SFP-25G-SR
23. Any out of band management interfaces provided within the solution must be connected to the EMBL-EBI management network
    a. These out of band management interfaces are assumed to be 1G RJ45 connections

b. These management interfaces must be aggregated within rack to a management switch provided by the vendor in each rack
c. Any transceivers required to effect 1 x 1G RJ45 uplink connection from each in rack management switch to the EMBL-EBI management switches (RJ45) are to be provided by the vendor

Housing & Power

The equipment will be housed at three locations. 2 sites at EMBL-EBI's leased data centre at Harlow (KAO Data) and one in the on campus data centre at Hinxton. Please refer to the Racking Requirements in the appendices for further details of the racking environments and restrictions.

24. Harlow (2 nodes)
    a. Both nodes at Harlow (KAO Data) are to be housed into existing racks
    b. The racks at the Harlow site are not wide format, the vendor must deliver the solution with in-rack horizontal PDUs with an RJ45 management interface accessible to EMBL-EBI via the out of band management network described above
        i. For reference the current model in use by EMBL-EBI is Geist 30020L
            1. Other vendors and models of PDU are acceptable, the vendor must ensure the port socket count and types are appropriate and that power is appropriately balanced across power phases
        ii. The vendor must ensure that the power cable length and connectors on the PDUs are appropriate
25. Hinxton (1 node)
    a. The node at Hinxton can be delivered in vendor provided racks
    b. These racks must include monitored PDUs (IPMI) with a visual display of load, with an RJ45 management interface accessible to EMBL-EBI via the out of band management network described above
26. No more than a 20 kW total power consumption per rack at either site
27. All mounting rails or other hardware required to mount the hardware into the racks must be provided where required (rack mounting kit)
28. Front to back cooling distribution is mandatory
    a. Consideration must be given to networking hardware when mounted to the rear
29. Provider must rack and provision all hardware, including connecting all power cables, networking or other related equipment
30. Racks must be populated from the bottom up for weight distribution and loading purposes, especially storage nodes with disk
31. If any inter-rack cabling is required, please specify and diagram in detail as part of the response
32. Please specify if each rack can be housed independently if contiguous rack space is required and the dependencies between such racks
33. Specific space within racks is reserved for EMBL-EBI networking and patch cabling, please refer to the Racking Requirements under appendices
    a. Should additional space, power or other physical requirements emerge please contact EMBL-EBI to discuss
34. All racking and cabling are to be completed to EMBL-EBI Cabling Guidelines, please refer to the appendices

Costing

35. Itemised indicative costs for all elements of the solution, including all hardware, software, updates or other costs required to operate the solution must be included in the response
36. Indicative pricing for support of all elements of the solution, including all hardware, software and updates, beyond the initial 5 year period must be provided

Budget

The total allocated budget for this purchase is 4m GBP, solutions should bid at this price with maximum capacity and performance.

Scored Requirements

	Scored Technical Requirements	Weighting 100%
1	Total usable capacity in PiB	40%
2	System performance, resilience, scalability, quotas etc.	25%
3	Total power consumption of the solution under peak load per usable PiB of storage capacity delivered	20%
4	Monitoring capabilities and features	10%
5	Extended warranty or other additional support services	5%

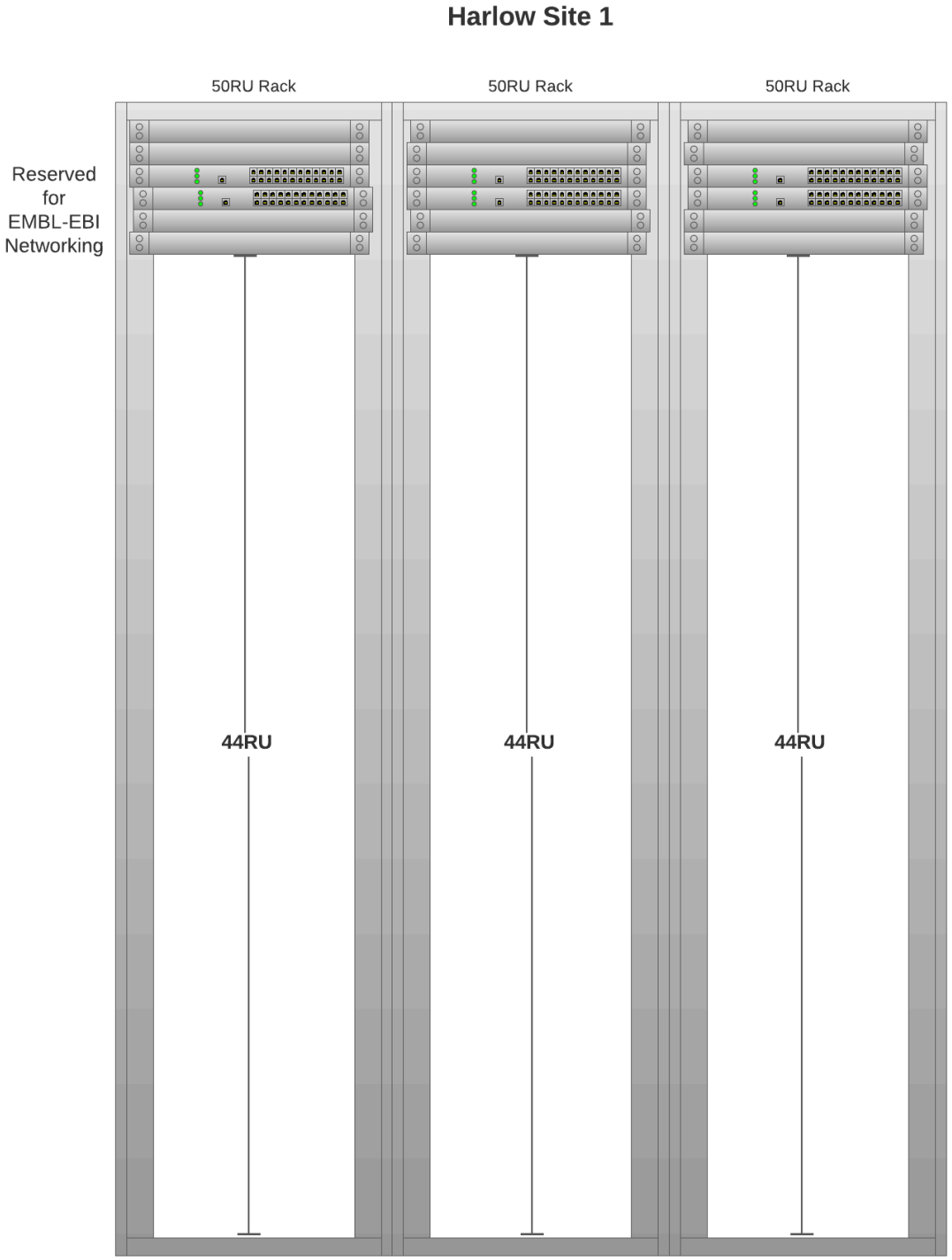
Appendices

Racking Requirements

The following rack environments are available, should further racks be required please immediately contact EMBL-EBI to discuss requirements.

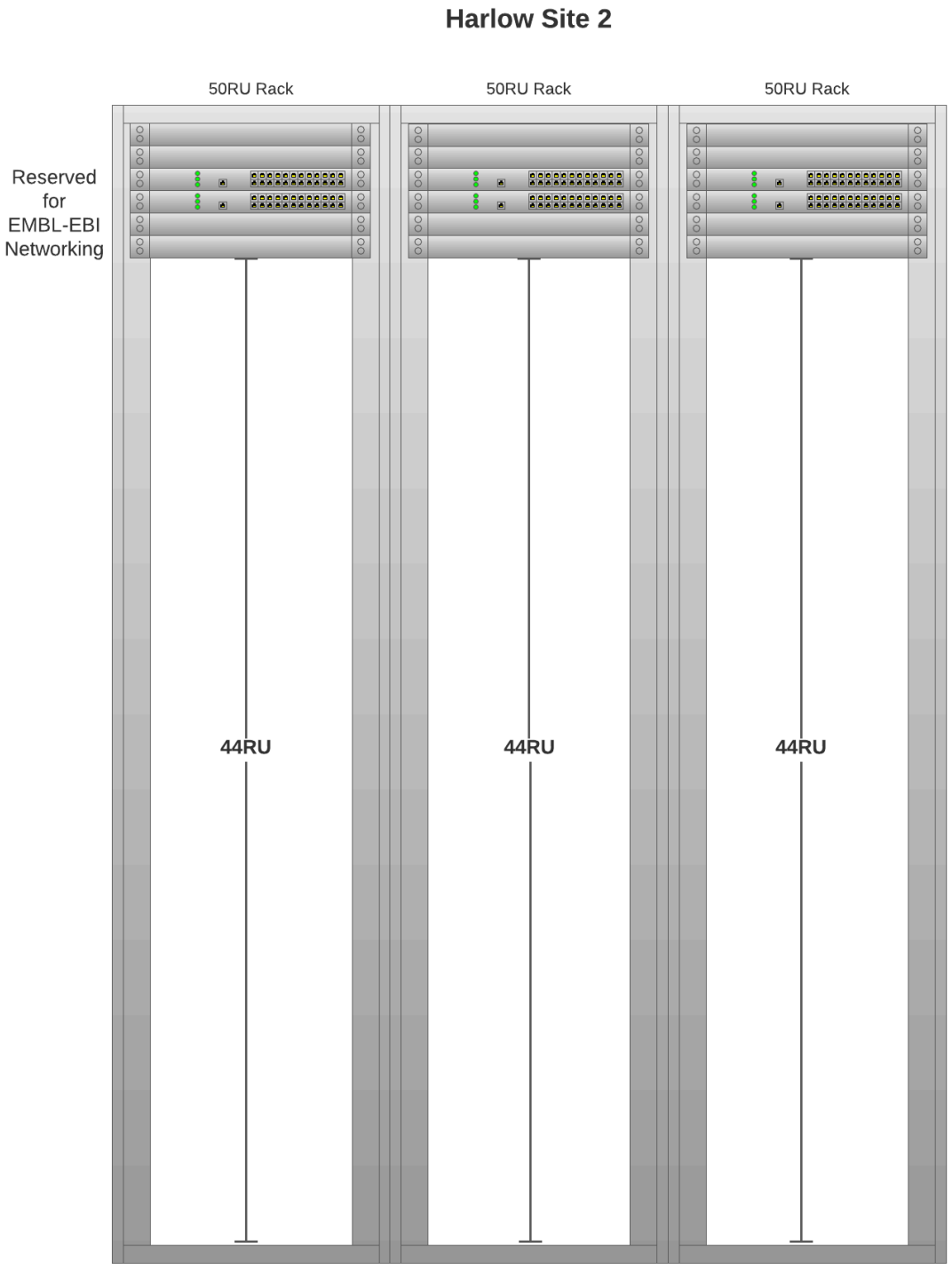
Harlow Site 1

3 contiguous industry standard 600mm wide Minkell racks – 50u, 1200mm deep



Harlow Site 2

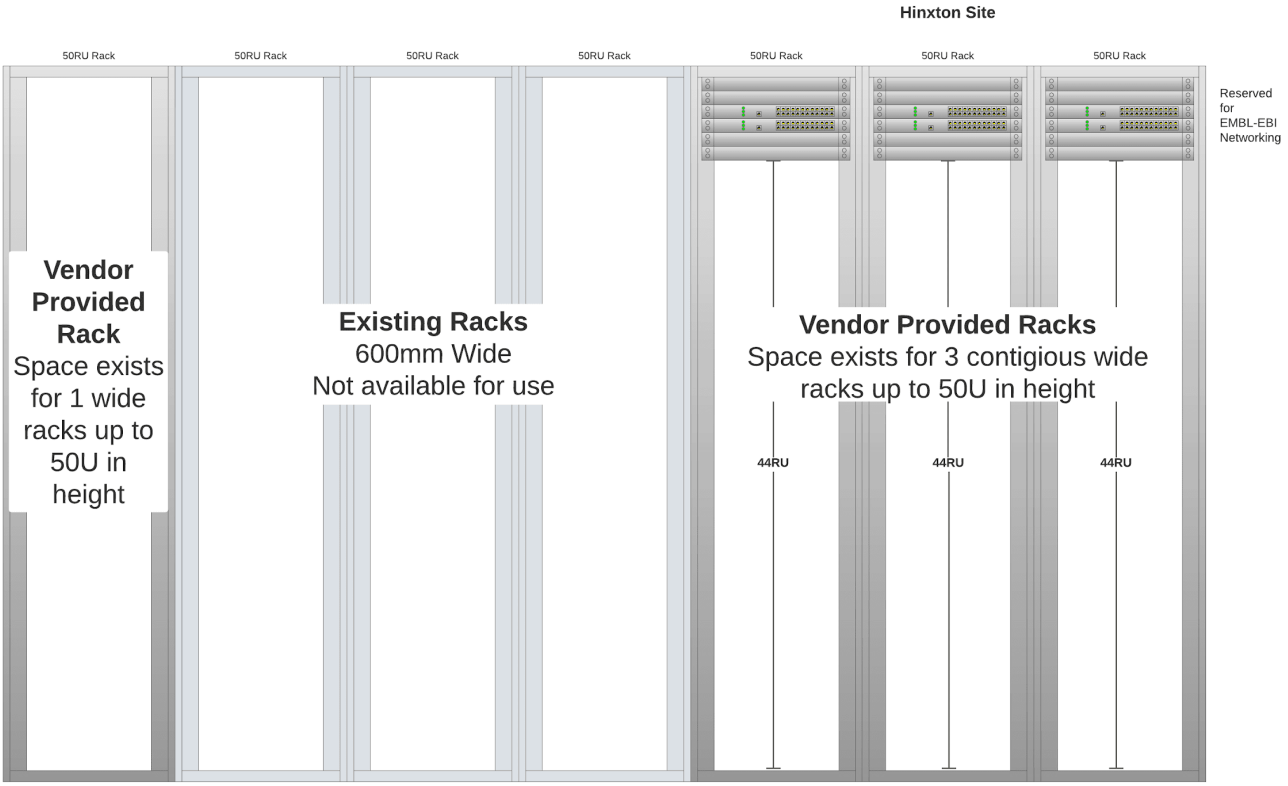
3 contiguous industry standard 600mm wide Minkell racks – 50u, 1200mm deep



Hinxton Site

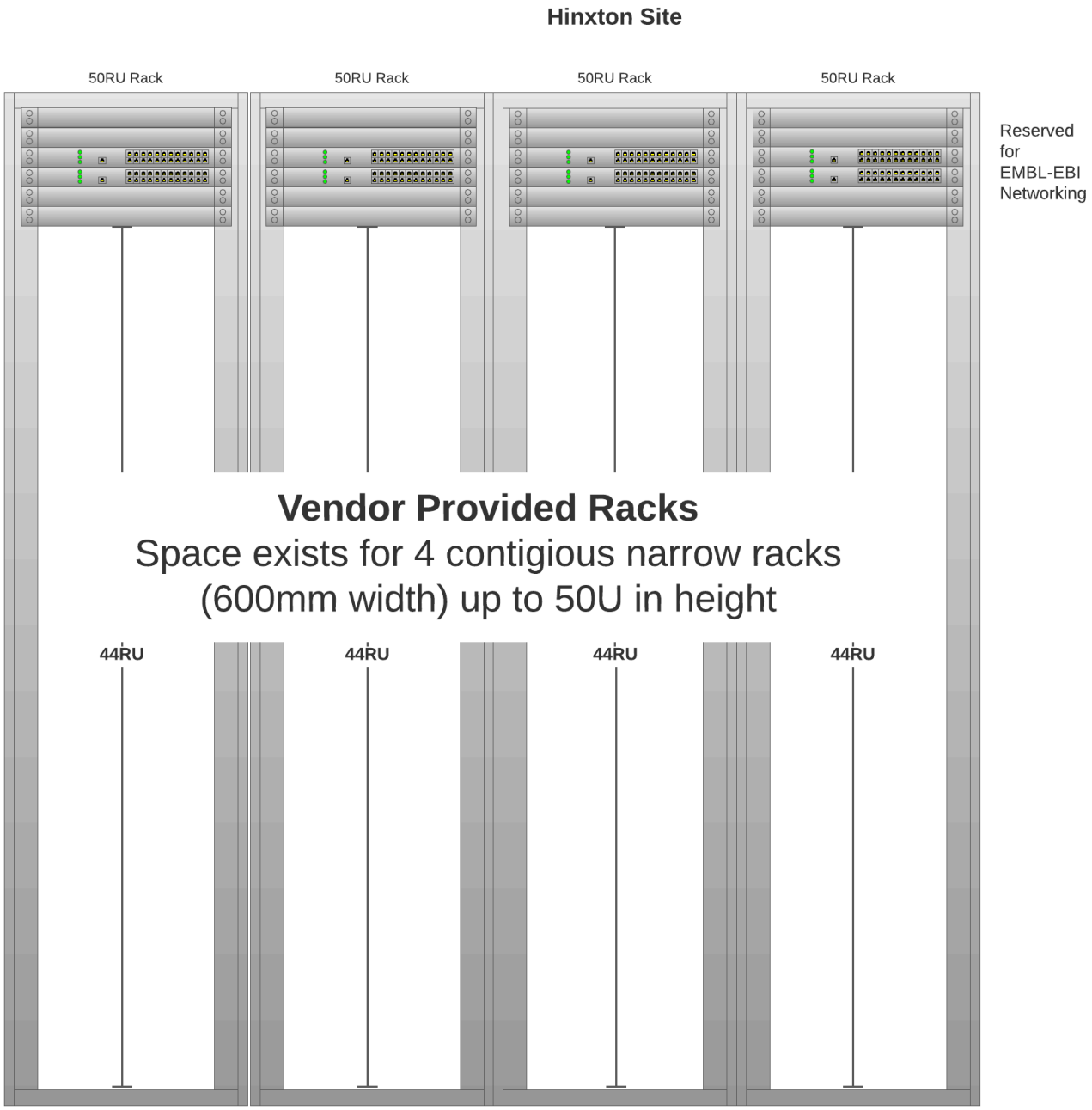
Pattern 1

4 wide racks cannot be accommodated contiguously. In the case of wide racks only 3 racks can be accommodated contiguously. An additional wide rack can be accommodated but will be separated from the first 3 racks by 3 600mm racks in the same row.



Pattern 2

Space for 4 contiguous narrow provided racks - up to 600mm wide, 1200mm deep



EMBL-EBI Cabling Guidelines

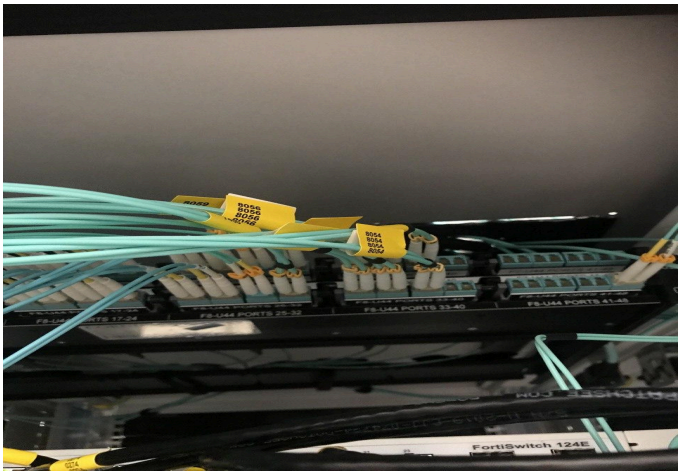
Network Installation

The cabling and labelling of all power, network and other cabling required to install the server to EMBL-EBI standards

- EMBL-EBI will provide labels, the vendor does not need to provide labels
- All cabling work must be neat, tidy and respect the flex radius of the cable types used

Label all cables at both ends with labels provided by EBI

Do not put labels on like this:

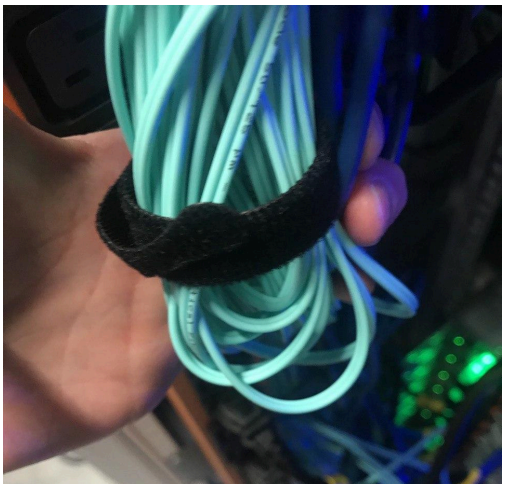


Please tightly wrap labels around the cable, roughly 10CM away from the connector:

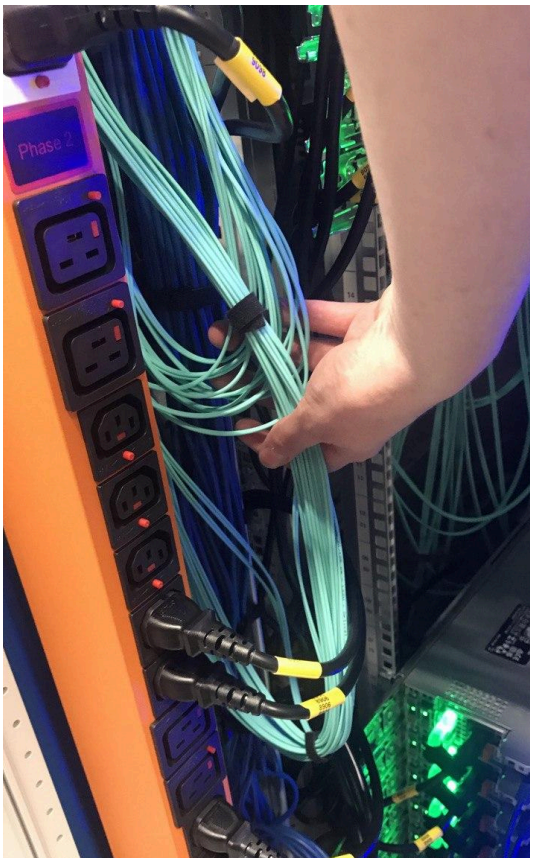
Cable Bend Radius

Do not bend or put fibre cables under strain, but allow for a decent radius. If the cable vendor has published documentation on best practices relative to bend radius, these should be taken into account and respected.

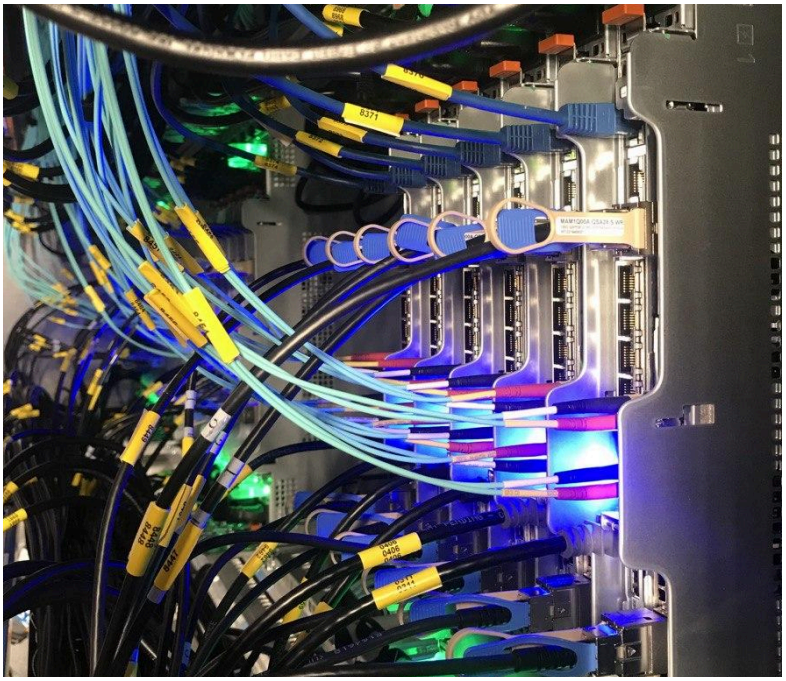
Example of a bad bend radius:



Good Bend Radius:



Cables shouldn’t be pulled tightly but arch nicely when connected to switches and or servers.



Keep copper and fibre cables in separate layers

- Fibre cables should be at the front
- Secure the separate layers to the vertical cable management
- Power cables should wrap around and be secured on the outer edge of the vertical cable management.

Cable ties

- Use velcro cable ties
- Do not use plastic cable ties