

DELL EMC HPC System for Life Sciences v1.2

Designed for genomics sequencing analysis, bioinformatics and computational biology

Dell EMC Engineering
February 2017

Revisions

Date	Description
February 2017	Initial release

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.
Copyright © <2016> - <2017> Dell Inc. All rights reserved. Dell and the Dell EMC logo are trademarks of Dell Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

Table of contents

Revisions.....	2
Executive summary.....	4
Audience	4
1 Introduction.....	5
2 System Design	6
2.1 Hardware Configuration.....	6
2.1.1 Master Node Configuration.....	6
2.1.2 Login Node Configuration.....	7
2.1.3 Compute Node Configuration	7
2.1.4 Common Internet File System (CIFS) gateway.....	8
2.1.5 Fat Node	8
2.1.6 Accelerator Node Configuration	8
2.2 Network Configuration	9
2.2.1 Management Network	9
2.2.2 High-Speed Interconnects	10
2.3 Storage	11
2.3.1 NSS 7.0 HA	11
2.4 Software Configuration	14
2.4.1 Bright Cluster Manager.....	14
2.4.2 BioBuilds.....	14
3 Sample Architectures	16
3.1 Case 1: PowerEdge C6320 compute subsystem with Intel® OPA fabric	16
3.1.1 Solution summary.....	17
3.1.2 Cabling information.....	17
3.2 Case 2: PowerEdge FC430 compute subsystem with IB FDR fabric	18
3.2.1 Solution summary.....	19
3.2.2 Cabling information.....	19
4 Conclusion.....	20

Executive summary

In October 2015, Dell Technologies introduced Genomic Data Analysis Platform (GDAP) v2.0 to answer the growing necessities of rapid genomic analysis due to the availability of next-generation sequencing technologies. Upon the successful implementation of GDAP v2.0, which is capable of processing up to 133 genomes per day while consuming 2 kilowatt-hour (kWh) per genome, we started to explore the life science domains beyond genomics. In addition to NGS, other lab equipment used in the life sciences also produces much richer data than before, such as mass spectrometers, electron microscopes and so on. Often these data from various lab equipment are image based and need heavy computational work to put the raw images into a useful format. However, these data can be highly variable in workloads, types of analysis and tools and make it difficult to optimize systems. Hence, Dell EMC offers the Dell EMC HPC System for Life Sciences v1.2, a high performance computing environment architected for flexibility and optimized to address various analysis workflows.

Audience

This document is intended for organizations interested in accelerating genomic research with advanced computing and data management solutions. System administrators, solution architects, and others within those organizations constitute the target audience.

1 Introduction

The Dell EMC HPC System for Life Sciences is a pre-integrated, tested, tuned and purpose-built platform, leveraging the most relevant of Dell EMC's high performance computing line of products and best-in-class partner products due to the high diversity in life sciences applications. It encompasses all the hardware resources required for various life sciences data analysis, while providing an optimal balance of compute density, energy efficiency and performance from Dell EMC's Enterprise server line-up.

The Dell EMC HPC System for Life Sciences v1.1 provides higher flexibility for the solutions. A platform is available in five variants, depending on the cluster interconnects selected, which can be either 10 Gigabit Ethernet (GbE), Intel® Omni-Path (OPA), InfiniBand® (IB) EDR or IB FDR. In this version, the following options are available:

- PowerEdge C6320 compute subsystem with Intel® OPA fabric
- PowerEdge C6320 compute subsystem with IB EDR fabric
- PowerEdge C6320 compute subsystem with 10 GigE fabric
- PowerEdge FX2 compute subsystem with IB FDR fabric
- PowerEdge FX2 compute subsystem with 10 GigE fabric

The solutions are nearly identical for IB, Intel® OPA and 10 GigE versions, except for a few changes in the switching infrastructure and network adapters. These differences are outlined in Network Components section. The solution ships in a deep 48U rack enclosure, which was chosen because of its ease of mounting PDUs and for effortless cable management. This rack houses the compute, storage and networking modules of the solution. Also, there are software modules which deploy, manage and maintain the cluster.

Not only can each sub-field of life sciences benefit immensely from the massive growth in HPC power, but also life sciences as whole can use HPC as an essential tool to integrate all the necessary data, biochemistry, genomics, proteomics, biophysics and anatomy of cells, tissues and organs into a single frame of work. However, academia and industry are not yet ready for this enormous task. Due to the highly diverse nature of life sciences research, researchers have different needs and have a difficult time picking a system that would work for everyone. Hence, Dell EMC HPC System for Life Sciences was introduced as a solution to the challenge, 'one system does not fit all' in life sciences. It was designed to be a plug-and-play turnkey solution so that researchers can spend more time working on matters in their domain, rather than concerning themselves with the computer science aspect of getting the system to function, which deals with cluster deployment and maintenance.

This reference architecture describes working solutions for diverse life sciences applications, including molecular dynamics simulation, with a flexible architecture, as well as improvement to the performance of the genomics data analysis platform.

2 System Design

The first step in designing the system is to decide upon the following four basic considerations:

- Type of workload
 - Genomics/NGS data analysis only
 - General purpose and Genomics/NGS data analysis
 - Adding molecular dynamics simulation capacity
- Parameter for sizing
 - Number of compute nodes
 - Genomes per day to be analyzed
- Form factor of servers
 - 2U shared infrastructure of high density that can host 4 compute nodes in one chassis (C6320)
 - 2U shared infrastructure of very high density that can host 8 compute nodes in one chassis (FC430)
 - 4U rack mount servers (R930)
 - 1U rack mount servers (R430)
 - 2U rack mount servers that can host up to 4 accelerators per node (C4130)
- Types of interconnect: All three options are available for any server except FC430. Mellanox ConnectX-3 (IB FDR) is only high speed interconnect option for FC430.
 - Intel® Omni-Path Host Fabric Interface (HFI) 100 series card
 - Mellanox ConnectX-4, Single Port, VPI EDR, QSFP28 Adapter
 - Mellanox ConnectX-3, Single Port, VPI FDR, QSFP+ Adapter, Low Profile

Following are the technical specifications of the servers that are considered for the Dell EMC HPC System for Life Sciences.

2.1 Hardware Configuration

There are several considerations while selecting the servers for compute node, login node, fat node and accelerator node, which are the components of the Dell EMC HPC System for Life Sciences. While 1U form factor PowerEdge R430 is the recommended server for master node, login node and CIFS gateway, the 4U form factor PowerEdge R930 is recommended as the appropriate server for a fat node and the 1U form factor PowerEdge C4130 is recommended to host accelerators; two options are available for selecting the compute nodes. These servers are not offered as standard off-the-shelf components. Although all the servers mentioned below are customizable, the configuration of servers that best fit life sciences applications are mentioned below.

2.1.1 Master Node Configuration

The master node is responsible for managing the compute nodes and optimizing the overall compute capacity. Hence it is also known as a “head” node. Usually, the master and login nodes are the only nodes that communicate with the outside world and act as a middle point between the actual cluster and the outside network. A master node is also referred to as the front end node because it provides the point of access and testing of the programs you want to run on a cluster system.

The master node controls the OS imaging and administration of the cluster. One master node is default, and high availability of the master node is optional.

The configuration of Power Edge R430 which is the recommended server for a master node is provided below:

- PowerEdge R430, 2-Socket server with Intel Xeon E5-2680 v4 processors
- 8 x 16GB RDIMM, 2400MT/s, Dual Rank
 - DDR4 DIMMs at up to 2400MT/s
 - 12 DIMM slots: 2GB/4GB/8GB/16GB/32GB
- RAID 1 for H330/H730/H730P (2 HDDs or SSDs)
- PERC H730 Integrated RAID Controller, 1GB Cache
- 2 x 500GB 7.2K RPM NLSAS 6Gbps 2.5in Hot-plug Hard Drive, 3.5in HYB CARR, 13G
- Interconnect options
- iDRAC8 Enterprise with OpenManage Essentials
- DVD+/-RW SATA Internal for 4HD Chassis

2.1.2 Login Node Configuration

Login nodes are where you login for access to the compute nodes. These nodes will be used for all access, compilation and job submission needs. Login node and high availability are optional. Power Edge R430 is the recommended server; for its configuration please refer to section 2.1.1.

2.1.3 Compute Node Configuration

Higher number of cores is more important than CPU speed for the applications in NGS data processing. Power Edge FC430 offers the ultimate density and is the choice for NGS data processing. Another option for the customers who also use other applications in addition to the NGS applications is PowerEdge C6320.

PowerEdge C6320 – 1U, Half-width

- PowerEdge C6300 Enclosure, 2.5 in Hard Drives with Backplane: houses four PowerEdge C6320 servers
- PowerEdge C6320, 2-Socket server with Intel Xeon E5-2697 v4 processors
- 1TB 7.2K RPM SATA 6Gbps 2.5in Cabled Hard Drive
- 8 x 32GB RDIMM, 2400MT/s, Dual Rank
 - DDR4 DIMMs at up to 2400MT/s
 - 16 DIMM slots: 4GB/8GB/16GB/32GB/64GB
- iDRAC8 Enterprise

PowerEdge FC430 – 1U, Quarter-width

- PowerEdge FX2 chassis: houses 8 FC430 servers
- PowerEdge FC430, 2-Socket server with Intel Xeon E5-2680 v4 processors
- 1.8" Backplane with 1 Hard Drive and 1 IB Port
- PowerEdge FC PCIE Mezzanine Adapter

- 8 x 32GB RDIMM, 2133MT/s, Dual Rank
- 200GB Solid State Drive uSATA Mix Use Slim MLC 6Gbps 1.8inHot-plug Drive
- Interconnect: Mellanox ConnectX-3 FDR Mezzanine adapter
- iDRAC8 Enterprise

PowerEdge C6320p - 1U, Half-width

- Intel Xeon Phi 7230. 64 cores, 1.3 GHz
- 96 GB at 2400 MT/s
- Internal 1.8" 240GB Solid State Drive 6Gbps

2.1.4 Common Internet File System (CIFS) gateway

A Dell EMC PowerEdge R430 is used as the CIFS gateway for transferring data generated by the next-generation sequencing machines into the storage. The configuration for CIFS gateway is provided below:

- PowerEdge R430 Server, 2-Socket server with Intel Xeon E5-2680 v4 processors
- 6 x 8GB RDIMM, 2400 MT/s, Dual Rank
 - DDR4 DIMMs at up to 2400MT/s
 - 12 DIMM slots: 2GB/4GB/8GB/16GB/32GB
- RAID 1 for H330/H730/H730P (2 HDDs or SSDs)
- PERC H730 Integrated RAID Controller, 1GB Cache
- 2 x 500GB 7.2K RPM NLSAS 6Gbps 2.5in Hot-plug Hard Drive,3.5inHYB CARR,13G
- Intel Ethernet X540 DP 10GBASE-T Server Adapter, Low Profile
- iDRAC8, Basic
- DVD+/-RW SATA Internal for 4HD Chassis

2.1.5 Fat Node

A fat node is a node with a large memory. Power Edge R930 is offered as an optional fat node with up to 3TB of memory. All the features benefit *de novo* genome sequence assemblies like Velvet. The size of the data on which these applications operate is massive. Hosting this data in memory with 64 cores eliminates the overhead caused by interconnects, disk look-ups and swapping, resulting in a speedup in time-to-results.

- Power Edge R930, 4-Socket server with Intel Xeon E7-8890 v4 processors
- 96 DIMM slots, DDR4 memory
- 8GB/16GB/32/64GB DDR4 up to 2400MT/s, maximum 6 TB
- Up to 24 x 2.5" hot-plug 12Gb/6Gb SAS HDD or SAS/SATA SSD
- Up to 8 front-accessible Express Flash NVMe PCIe SSD
- PCI Express Generation 3 slots
- iDRAC8, Enterprise with OpenManage Essentials

2.1.6 Accelerator Node Configuration

Accelerators are used to speed up computationally intensive applications, such as Amber, HOOMD-blue and NAMD for molecular dynamics simulations. PowerEdge C4130 offers five different NVIDIA GPU configurations.

PowerEdge C4130 - 1U - Up to 4 accelerators per node

- PowerEdge C4130, 2-socket server with Intel Xeon E5-2690 v4 processors
- 8 x 16GB RDIMM, 2400MT/s, Dual Rank
 - 16 DIMM slots, DDR4 Memory
 - 4GB/8GB/16GB/32GB DDR4 up to 2400MT/s
- Up to 2 x 1.8" SATA SSD boot drives
- Optional 96-lane PCIe 3.0 switch for certain accelerator configurations
- iDRAC8, Dell EMC OpenManage Essentials
- Available GPUs
 - 4 x NVIDIA K80
 - 4 x NVIDIA P100 - two versions are available
 - SXM2 module for NVLink-optimized servers provide the best performance and strong-scaling for hyperscale and HPC data centers running applications that scale to multiple GPUs, such as deep learning.
 - PCIe allows HPC data centers to deploy the most advanced GPUs within PCIe-based nodes to support a mix of CPU and GPU workloads.

2.2 Network Configuration

The Dell EMC HPC System for Life Sciences is available in Intel OPA and two IB variants. There is also a Force10 S3048-ON GbE switch used in both configurations, whose purpose is described here. In one of the IB variants, the Dell EMC PowerEdge FC430 sleds have 2:1 blocking FDR connectivity to the top of rack FDR switch. The other IB variant is 1:1 non-blocking EDR network for C6320.

2.2.1 Management Network

Management traffic typically communicates with the Baseboard Management Controller (BMC) on the compute nodes using IPMI. The management network is used to push images or packages to the compute nodes from the master nodes and for reporting data from client to the master node. Dell EMC Networking S3048-ON and PowerConnect 2848 Switch are considered for management network

Dell EMC Network S3048-ON

High density, 1U 48-port 1000BASE-T + with four 10GbE uplinks, non-blocking rate-line performance, featuring the Open Networking Install Environment (ONIE).

The port assignment of the Dell EMC Networking S3048-ON switch for the Intel® OPA or IB versions of the solution is as follows.

- Ports 01-04 and 27–52 are assigned to the cluster's private management network to be used by Bright Cluster Manager® connecting master, login, CIFS gateway and compute nodes. The PowerEdge C6320 server's ethernet and iDRAC constitute a majority of these ports.
- Ports 06–09 are used for the private network associated with NSS7.0-HA.
- The rest of the port 05 and ports 12–26 are allocated to the Lustre solution for its private management network

- Port 10 and 11 are used for the PDUs.



Figure 1: Dell EMC Networking S3048-ON switch

Note: Installation of four SFP+ to RJ45 transceivers in Dell EMC Networking S3048-ON switch ports 49-52 is required.

2.2.2 High-Speed Interconnects

In high performance computing, application performance depends on the number of CPU/GPU cores, memory, interconnect, storage performance and so on. For a server to perform better, lower latency and higher bandwidth is needed for these systems to communicate with each other. The type of network chosen for computational traffic depends upon the latency, bandwidth, packet size at peak bandwidth and the message rate. The types of computational network and the switch through which it is offered are tabulated below:

Table 1: Interconnect switches

Type of Interconnect	Switch
Fourteen Data Rate (FDR)	Mellanox SX6036
Extended Data Rate (EDR)	Mellanox SB7790
10Gb Ethernet	S4820-T, N4032-F, S6000, Z9500
Intel® Omni-Path (OPA)	Dell EMC Networking H1048-OPF

Mellanox SX6036

- 36 FDR (56Gb/s) ports in a switch
- Port-to-port latency 200ns
- Up to 4Tb/s aggregate switching capacity
- 19" rack mountable chassis, 1U with optional redundant power supplies and Fan units
- On-board SM for fabrics up to 648 nodes.

Mellanox SB7790

- 36 EDR (100Gb/s) ports in a switch
- Port-to-port latency 90ns

- Up to 7Tb/s aggregate switching capacity
- 19" rack mountable chassis, 1U with optional redundant power supplies and Fan units
- On-board SM for fabrics up to 2k nodes.

Dell EMC Networking S6000

- 1U high-density 10/40GbE ToR switch with 32 ports of 40GbE (QSFP+) or 96 ports of 10GbE and eight ports of 40GbE or 104 ports of 10GbE.
- Up to 2.56Tbps of switching capacity.

Dell EMC Networking Z9500

- High-density 3U with 132 ports of 40GbE (528 ports of 1/10GbE using breakout cables)

Dell EMC Networking S4820T

- 1U high performance ToR switch provides (48) 1/10G BASE-T ports that support 100Mb/1 Gb/10Gb and four 40GbE QSFP+ uplinks.
- Each 40GbE QSFP+ uplink can be broken out into four 10GbE ports using breakout cables

Dell EMC Networking N4032F SFP Switch

- 24x 10GbE SFP + auto-sensing (10Gb/1Gb) fixed ports
- Up to 32 10GbE ports using breakout cables and optional QSFP+ module
- One hot swap expansion module bay
- Dual hot-swappable redundant power supplies (460W)

2.3 Storage

The performance requirements of HPC environments with ever-larger compute clusters have placed unprecedented demands on the storage infrastructure. The storage infrastructure consists of the following components:

- Dell EMC Ready Bundle for HPC NFS Storage (NSS 7.0-HA)
- Dell EMC Ready Bundle for HPC Lustre Storage

2.3.1 NSS 7.0 HA

NSS 7.0 HA is designed to enhance the availability of storage services to the HPC cluster by using a pair of Dell EMC PowerEdge servers and PowerVault storage arrays along with Red Hat HA software stack. The HA cluster consists of a pair of Dell EMC PowerEdge servers and a network switch. The two PowerEdge servers have shared access to disk-based Dell EMC PowerVault storage in a variety of capacities, and both are directly connected to the HPC cluster by using OPA, IB or 10GbE. The two servers are equipped with two fence devices: iDRAC8 Enterprise, and an APC Power Distribution Unit (PDU). If failures such as storage disconnection, network disconnection, and system stopping from functioning, etc., occur on one server, the

HA cluster will failover the storage service to the healthy server with the assistance of the two fence devices; and also ensure that the failed server does not return to life without the administrator’s knowledge or control.

The test used to evaluate the NSS7.0-HA functionality and performance is shown in Figure 2. The following configuration was used.

- A 32-node HPC compute cluster (also known as “the clients”) was used to provide I/O network traffic for the test bed.
- A pair of Dell EMC PowerEdge R730 servers were configured as an active-passive HA pair and function as a NFS server for the HPC compute cluster.
- Both NFS servers were connected to a shared Dell EMC PowerVault MD3460 storage enclosure extended with one Dell EMC PowerVault MD3060e storage enclosure (Figure 2 shows a 480 TB solution with the two PowerVault MD storage arrays) at the back-end. The user data is stored on an XFS file system created on this storage. The XFS file system was exported to the clients by using NFS.
- The NFS servers were connected to the clients by using the public network. This network was Intel OPA.
- For the HA functionality of the NFS servers, a private 1 Gigabit Ethernet network was configured to monitor server health and heartbeat, and to provide a route for the fencing operations by using a Dell EMC Networking 3048-ON Gigabit Ethernet switch.
- Power to the NFS servers was provided by two APC switched PDUs on two separate power buses.

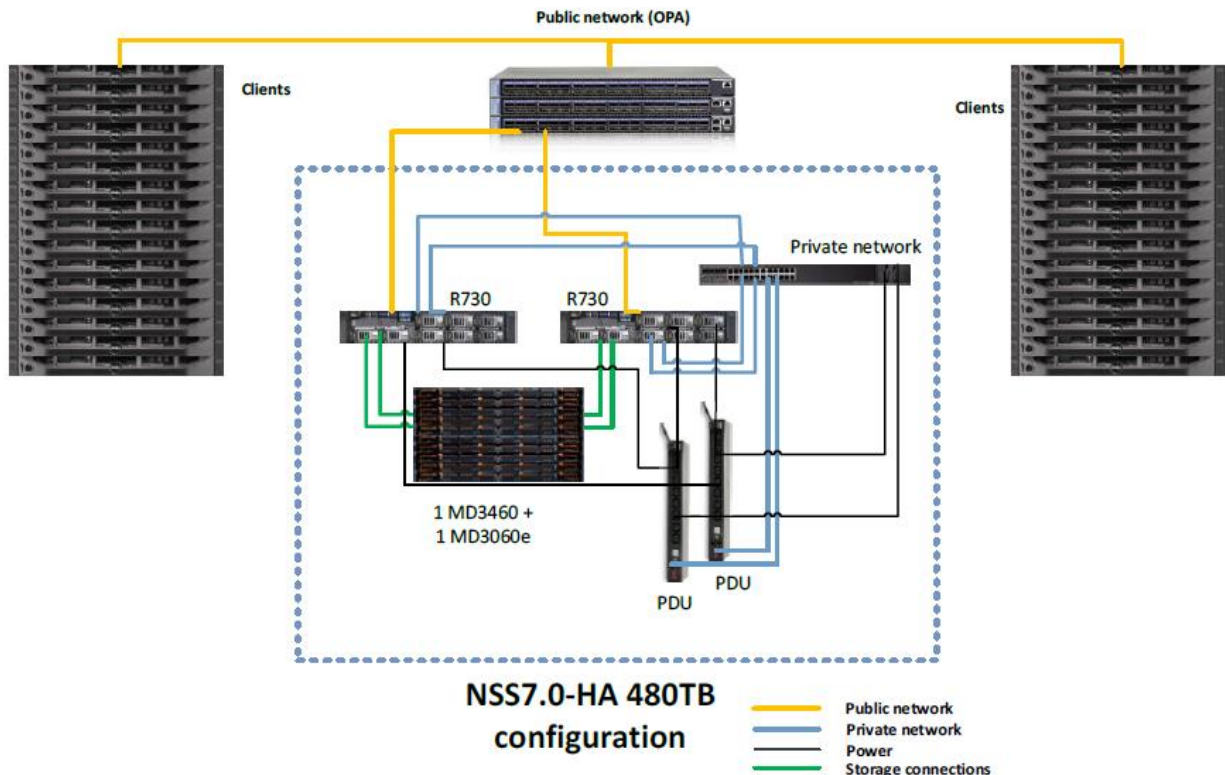


Figure 2: NSS7.0-HA test bed

Dell EMC Ready Bundle for HPC Lustre Storage The Dell EMC Ready Bundle for HPC Lustre Storage, referred to as Dell EMC HPC Lustre Storage, is designed for academic and industry users who need to deploy a fully-supported, easy-to-use, high-throughput, scale-out and cost-effective parallel file system storage solution. The solution uses the Intel® Enterprise Edition (EE) for Lustre® software v.3.0. It is a scale-out storage solution appliance capable of providing a high performance and high availability storage system. Utilizing an intelligent, extensive and intuitive management interface, the Intel Manager for Lustre (IML), the solution greatly simplifies deploying, managing and monitoring all of the hardware and storage system components. It is easy to scale in capacity, performance or both, thereby providing a convenient path to grow in the future. Figure 3 shows the relationship of the MDS, MDT, MGS, OSS and OST components of a typical Lustre configuration. Clients in the figure are the HPC cluster's compute nodes. The solution is 120 drive system configuration (60 drives per PowerVault MD3460) with 12 OSTs. Each OST consists of 10x 4TB drives. Total raw storage size is 480TB.

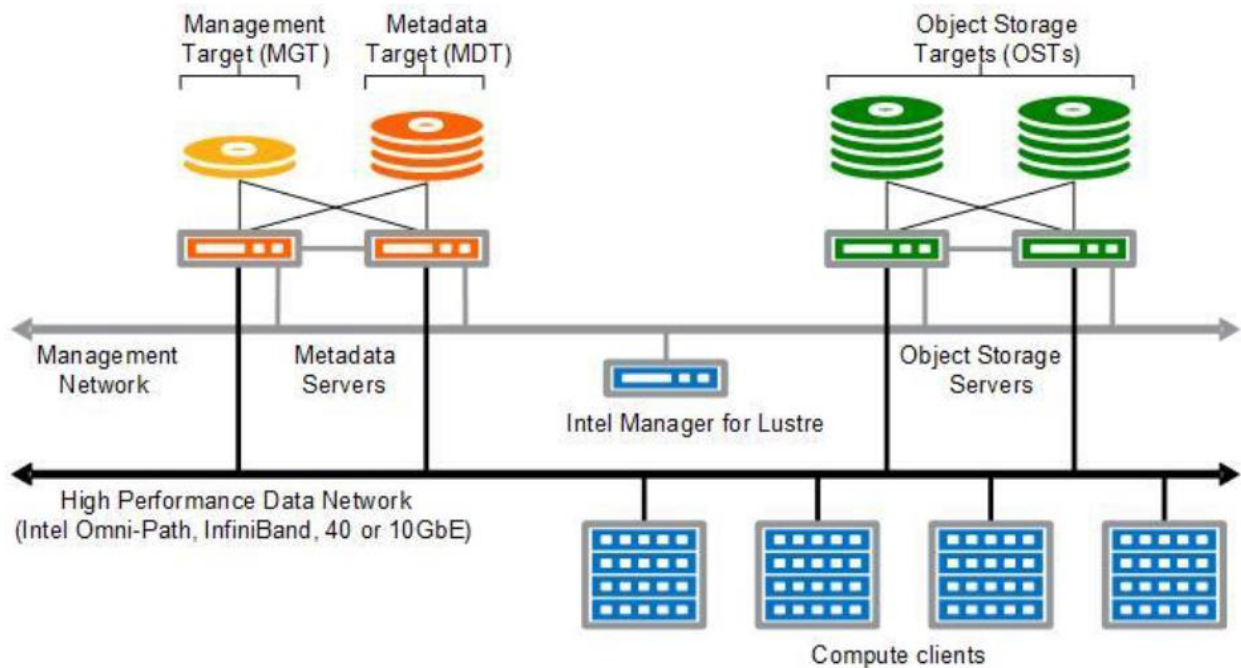


Figure 3: Lustre-based storage solution components

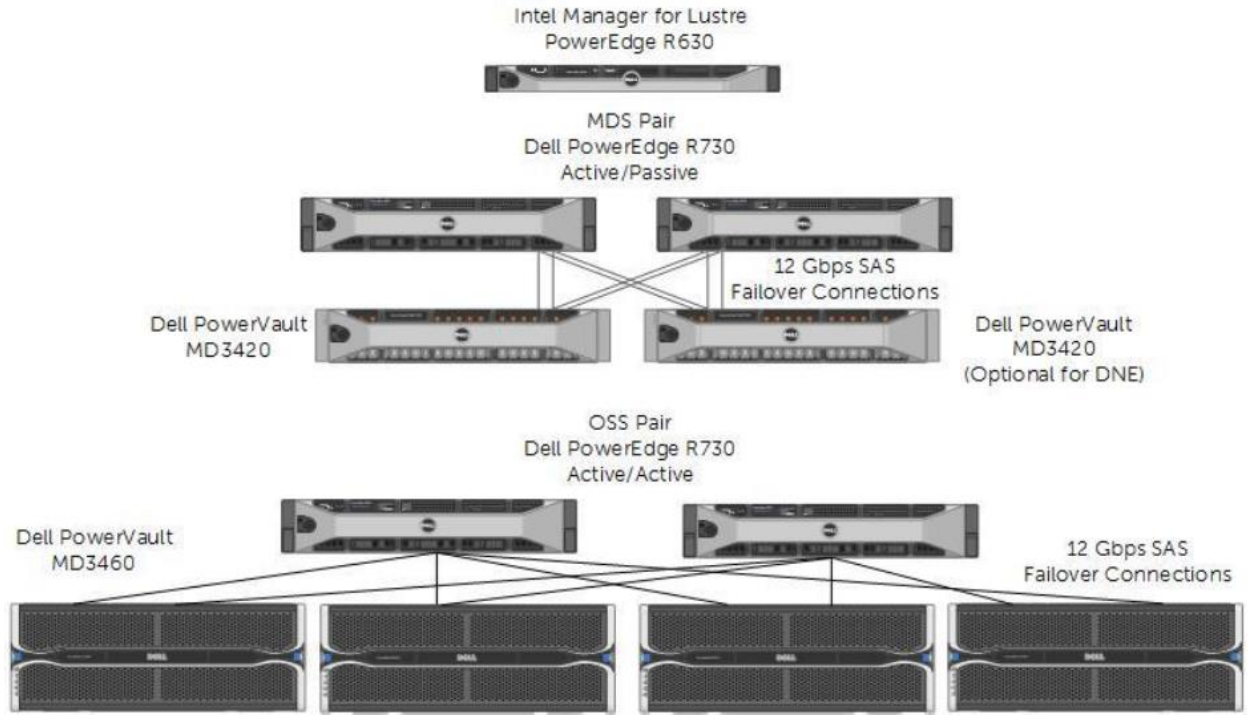


Figure 4: Dell EMC Ready Bundle for HPC Lustre Storage Components Overview

2.4 Software Configuration

Along with the hardware components, the solution includes the following software components:

- Bright Cluster Manager®
- BioBuilds

2.4.1 Bright Cluster Manager

Bright Computing is a commercial software that provides comprehensive software solutions for deploying and managing HPC clusters, big data clusters and OpenStack in the data center and in the cloud. Bright cluster Manager can be used to deploy complete clusters over bare metal and manage them effectively. Once the cluster is up and running, the graphical user interface monitors every single node and reports if it detects any software or hardware events.

2.4.2 BioBuilds

BioBuilds is a well maintained, versioned and continuously growing collection of open-source bio-informatics tools from Lab7. They are prebuilt and optimized for a variety of platforms and environments. BioBuilds tries to solve the software challenges faced by the life sciences domain.

- Imagine a newer version of a tool being released. Updating it may not be straight forward and would probably involve updating all the dependencies the software has as well. BioBuilds includes the software and its supporting dependencies for ease of deployment.

- Using BioBuilds among all the collaborators can ensure reproducibility since everyone is running the same version of the software. In short, it is a turnkey application package.

3 Sample Architectures

3.1 Case 1: PowerEdge C6320 compute subsystem with Intel® OPA fabric

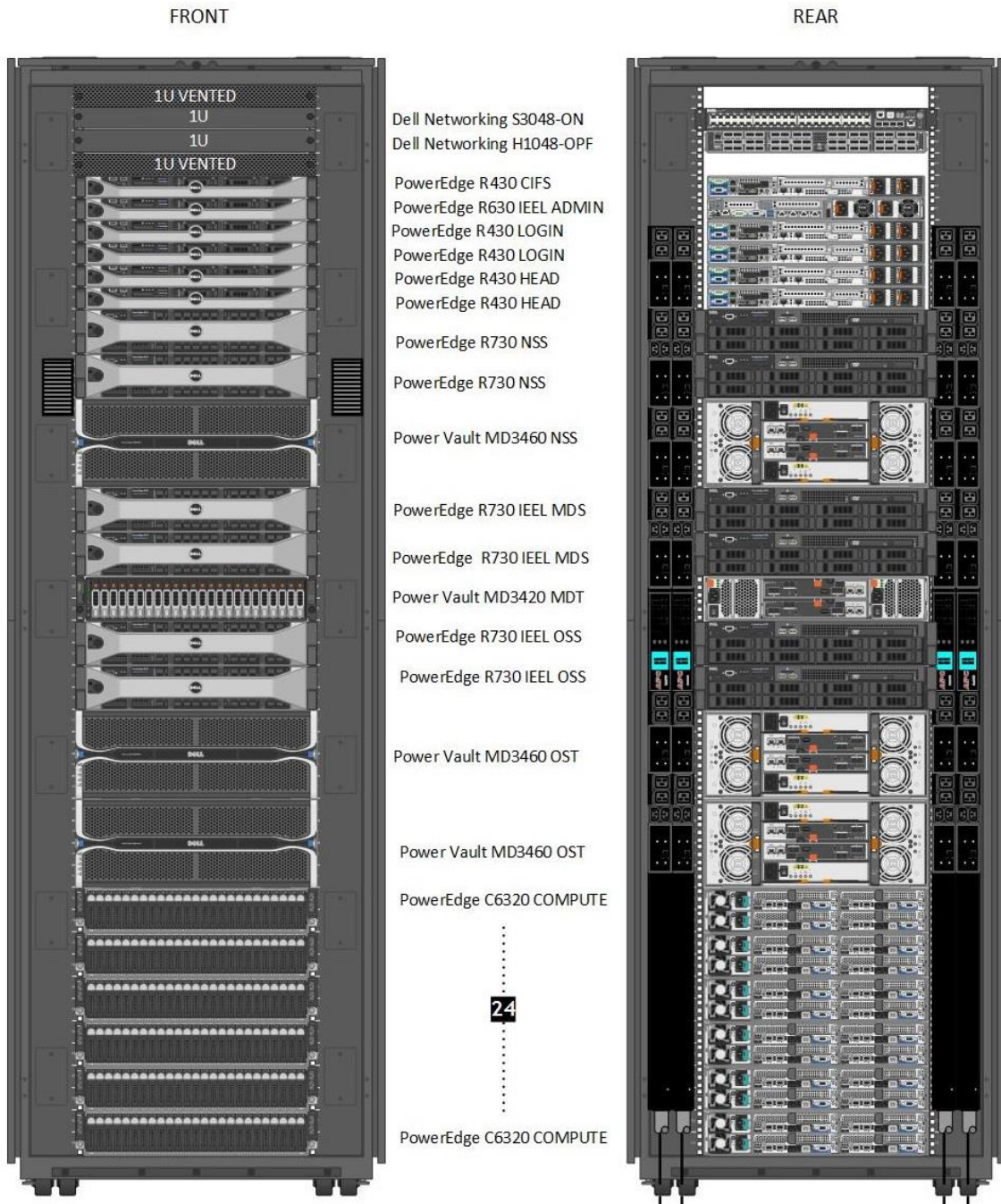


Figure 5: Dell EMC HPC System for Life Sciences with PowerEdge C6320 rack servers and Intel® OPA fabric

3.1.1 Solution summary

This solution is nearly identical to the solutions with IB EDR and 10 GbE versions, except for a few changes in the switching infrastructure and network adapters. As shown in Figure 5, this solution uses one 48U rack and requires extra deep enclosure. Bright Cluster Manager is the default tool and a proprietary software solution stack from Bright Computing. Along with automated deployment and provisioning features for the HPC cluster, it incorporates powerful and intelligent management and monitoring capabilities. The graphical user interface provides advanced configuration options.

Table 2: Hardware summary

Rack	1x 48U 600mm wide x 1200mm Deep
Master nodes	2x PowerEdge R430
Login nodes	2x PowerEdge R430
CIFS gateway	1x PowerEdge R430
Compute nodes	24x PowerEdge C6320 on 6x PowerEdge C6300 Enclosure
High speed interconnect switch	1x Dell EMC Networking H 1048-OPF
Management switch	1x Dell EMC Networking S3048-ON
Primary File System	Dell EMC NSS7.0-HA configuration (240TB)
Parallel File System	Dell EMC Ready Bundle for HPC Lustre Storage (480TB raw space)

3.1.2 Cabling information

Management interconnect and high speed interconnect require 32 cables each. Dell EMC NSS7.0-HA needs two cables for both management and high speed interconnect due to the HA configuration.

Table 3 Required cables

	Cable types	Length	Quantity
High speed interconnect cables	Copper cables	3M	24
		2M	8
Management cables	1GbE RJ-45 CAT6	10ft	24
		5ft	8

3.2 Case 2: PowerEdge FC430 compute subsystem with IB FDR fabric

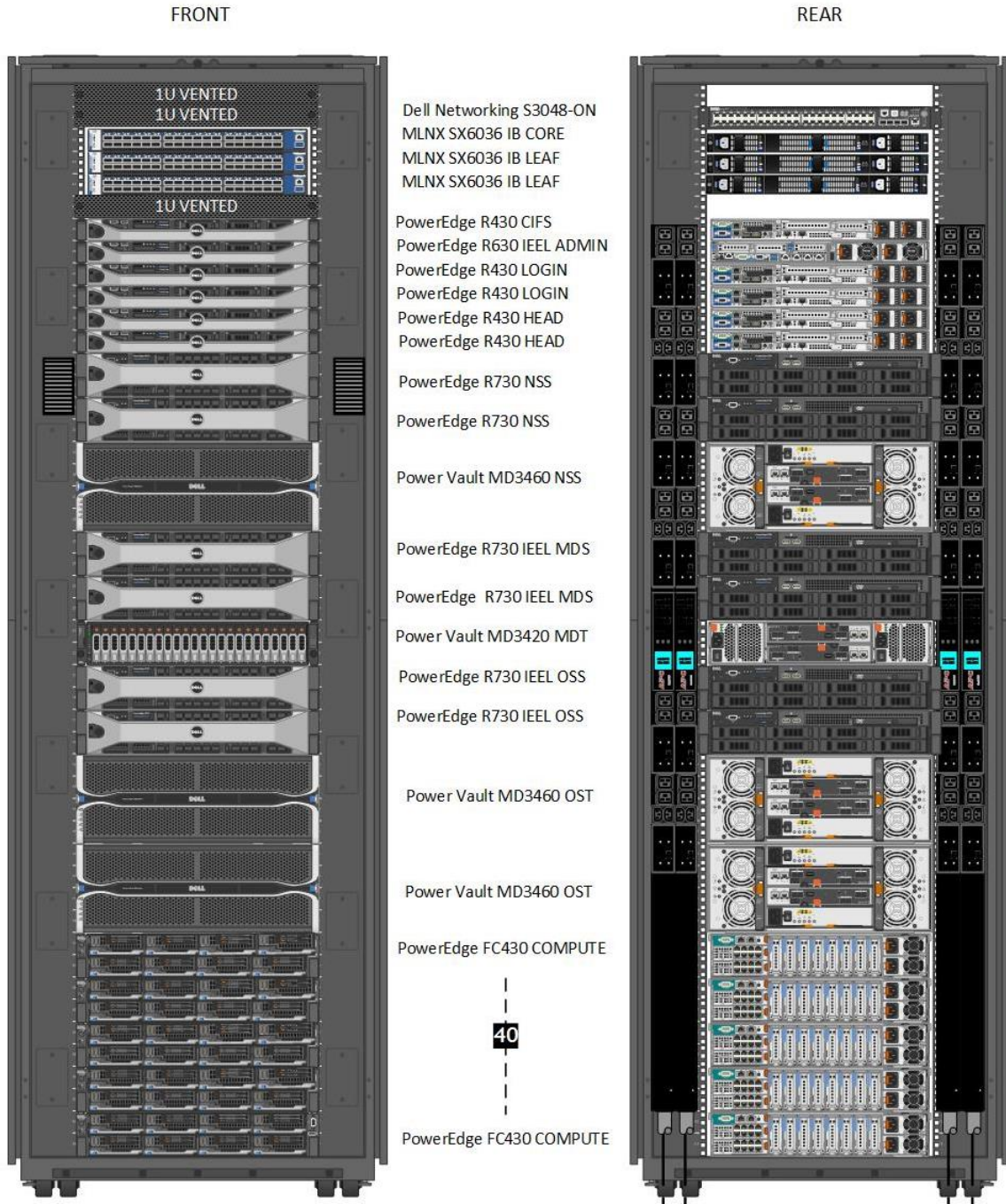


Figure 6: Dell EMC HPC System for Life Sciences with PowerEdge FC430 rack servers with IB FDR fabric

3.2.1 Solution summary

The FC30 solution with IB FDR interconnect is nearly identical to the 10 GbE version, except for a few changes in the switching infrastructure, and network adapters and have 2:1 blocking FDR connectivity to the top of rack FDR switch.

- The port assignment of the Dell EMC Networking S3048-ON switch for the Intel® OPA or IB versions of the solution is as follows.
- Ports 01-04 and 27–52 are assigned to the cluster’s private management network to be used by Bright Cluster Manager® connecting master, login, CIFS gateway and compute nodes. The PowerEdge C6320 server’s Ethernet and iDRAC constitute a majority of these ports.
- Ports 06–09 are used for the private network associated with NSS7.0-HA.
- The rest of the port 05 and ports 12–26 are allocated to the Dell EMC Ready Bundle for HPC Lustre Storage for its private management network
- Port 10 and 11 are used for the PDUs.

For the 10GbE configuration, the deployment and management of the cluster is done over the 10 Gigabit Ethernet network by using the Dell EMC Networking S4820T switch. So, the first virtual LAN on the S3048-ON, from ports 0–16, is not used. The other two virtual LANs are still used for the same purpose as in the IB FDR configuration.

Table 4 Hardware summary

Rack	1x 48U 750mm wide x 1200mm Deep
Master nodes	2x PowerEdge R430
Login nodes	2x PowerEdge R430
CIFS gateway	1x PowerEdge R430
Compute nodes	40x PowerEdge FC430 on 5x PowerEdge FX2 chassis
High speed interconnect switch	3x MLNX SX6036
Management switch	1x Dell EMC Networking S3048-ON
Primary File System	Dell EMC NSS7.0-HA configuration (240TB)
Parallel File System	Dell EMC Ready Bundle for HPC Lustre Storage (480TB raw space)

3.2.2 Cabling information

Table 5 Required cables

	Cable types	Length	Quantity
High speed interconnect cables	Fiber optic cables	3 M	40
	Copper cables	2 M	8
Management cables	1GbE RJ-45 CAT6	10 ft	15
		5 ft	10
		1ft	5

4

Conclusion

The HPC System Builder for Life Sciences provides the minimum architecture that could achieve the targeted NGS workload, informed decision making and increased efficiency.

However, the configuration provided by the Dell EMC HPC System Builder for Life Sciences tool is intended to be used as a starting point only. Dell EMC suggests that you contact your technical sales representative to review this quote for completeness and to include other variables not included as input to the tool.