

# Dell PowerFlex Rack with PowerFlex 3.x

## Architecture Overview

## Notes, cautions, and warnings

 **NOTE:** A NOTE indicates important information that helps you make better use of your product.

 **CAUTION:** A CAUTION indicates either potential damage to hardware or loss of data and tells you how to avoid the problem.

 **WARNING:** A WARNING indicates a potential for property damage, personal injury, or death.

<b>Chapter 1: Introduction.....</b>	<b>4</b>
<b>Chapter 2: Revision history.....</b>	<b>5</b>
<b>Chapter 3: System overview.....</b>	<b>6</b>
System features.....	6
System components.....	7
Storage schemas.....	8
Network architecture.....	9
<b>Chapter 4: Base configuration and scaling.....</b>	<b>16</b>
PowerFlex rack storage-only deployment.....	16
PowerFlex rack two-layer deployment.....	16
PowerFlex rack hyperconverged deployment.....	17
<b>Chapter 5: Additional references.....</b>	<b>18</b>
Network components.....	18
Storage components.....	18
Virtualization components.....	19

# Introduction

This document describes the high-level design of PowerFlex rack.

The target audience for this document includes customers, sales engineers, field consultants, and advanced services specialists who want to deploy a high-performance, scalable, and flexible infrastructure using PowerFlex rack.

The dvswitch names are for example only and may not match the configured system. Do not change these names or a data unavailable or data lost event may occur.

Dell PowerFlex rack was previously known as Dell VxFlex integrated rack. Similarly, Dell PowerFlex Manager was previously known as Dell VxFlex Manager, and Dell PowerFlex was previously known as Dell VxFlex OS. References in the documentation will be updated over time.

PowerFlex rack architecture is based on Dell PowerEdge R650, R750, R6525, R7525, R640, R740xd, and R840 servers.

PowerFlex Manager provides the management and orchestration functionality for PowerFlex rack. References to PowerFlex Manager in this document apply only if you have a licensed version of PowerFlex Manager. For more information, contact Dell Technologies Sales.

For additional PowerFlex rack documentation, go to [PowerFlex rack technical documentation](#).

## Revision history

Date	Document revision	Description of changes
January 2024	10.5	Added information about <i>Storage schemas</i>
November 2023	10.4	Added support for Cisco Nexus 93180YC-FX3
August 2023	10.3	Added support for Multi-subnet and multi-VLAN configurations
March 2023	10.2	Updated the content for aggregation switch list
February 2023	10.1	Added support for PowerFlex R7525 nodes
November 2021	10.0	Added support for: <ul style="list-style-type: none"> <li>PowerFlex management controller 2.0, an R650-based controller that uses PowerFlex storage and a VMware ESXi hypervisor</li> <li>PowerFlex R650/R750/R6525 nodes</li> </ul>
July 2021	9.0	Includes information about the following features: <ul style="list-style-type: none"> <li>New Cisco management, access, and leaf-spine switches</li> <li>New management aggregation switch</li> </ul>
December 2020	8.0	Includes information about the following features: <ul style="list-style-type: none"> <li>VMware vSphere 7.0</li> <li>SDC access to SDS using Layer 3 routing</li> <li>Continuation of 100 Gb networking</li> </ul>
August 2020	7.0	Updated <ul style="list-style-type: none"> <li>System features</li> <li>System components</li> </ul>
June 2020	6.2	Updated information about <ul style="list-style-type: none"> <li>Network architecture</li> </ul> Removed information about <ul style="list-style-type: none"> <li>Compute</li> <li>Storage</li> <li>Virtualization</li> <li>Management</li> <li>Sample PowerFlex rack configurations</li> </ul>

# System overview

## System features

PowerFlex rack is a modular hyperconverged platform that enables extreme scalability and flexibility for next-generation cloud applications and mixed workloads.

PowerFlex rack is a scale-out solution that enables you to add PowerFlex rack nodes with various CPU, memory, and drive options. PowerFlex rack is designed for deployments involving large numbers of virtualized and bare metal workloads.

Together, the components offer balanced CPU, I/O bandwidth, and storage capacity relative to the compute and storage arrays in the PowerFlex rack. All components have N+1 redundancy.

These resources can be scaled up as necessary to meet increasingly stringent application workload requirements. To scale up storage or compute resources, add nodes to the environment. Optionally, add expansion cabinets with more resources - this is the default for the PowerFlex rack network design.

PowerFlex rack provides the following features:

- Integrated scalable network architecture with a choice of the following:
  - Topologies depending on scalability needs can be a leaf-spine or aggregation and access architecture
  - Network hardware can be either Cisco Nexus switches, or Dell PowerSwitch switches
  - Bandwidth depending on your needs
- Complete flexibility in designing the system with the following options:
  - PowerFlex hyperconverged nodes
  - PowerFlex storage-only nodes
  - PowerFlex compute-only nodes
- For the PowerFlex management controller 2.0, a three-node or more PowerFlex management controller supporting VMware ESXi and PowerFlex is available for shared storage layer, providing high availability for the VM workloads on the PowerFlex management controller. The PowerFlex management nodes are based on PowerFlex R650 nodes and a HBA355i controller.
- Supports 10 GbE, 25 GbE or 100 GbE port bandwidth for backend connectivity
- External servers can access data in the PowerFlex rack
- Optional support for multi-subnet or multi-VLAN configuration for the following network types:
  - OOB Management
  - ESXi Management (Includes support on NSX-T nodes)
  - PowerFlex management controller vMotion, PowerFlex vMotion, and NSX-T vMotion
  - PowerFlex management controller storage data server management and PowerFlex storage management
  - Replication data
- (Optional) Dual network environment using your existing software-defined network (SDN), such as Cisco ACI
- Supports native asynchronous replication between sites
- Scale beyond 1000 nodes with a leaf-spine network design
- PowerFlex nodes support SSD and NVMe drive technologies
- Improved node maintenance using protected maintenance mode (PMM)
- Supports self-encrypting drives (SEDs) enabled by Dell CloudLink
- Secure and scalable storage using PowerFlex:
  - Seamlessly increase storage capacity or performance with horizontal node scaling

- Optional data compression feature that improves storage efficiency
- Optional Data at Rest Encryption (D@RE) feature provides data security using CloudLink as a software encryption layer, or as a key manager for self-encrypting drives
- Engineered system with end-to-end lifecycle management with PowerFlex Manager
- Optional support for multitenancy and routed reachability between PowerFlex compute-only nodes and PowerFlex storage-only nodes on leaf-spine-based PowerFlex rack
- Optional support for management aggregation switches
- Support for fiber channel host bus adapters (HBA) that are connected to storage devices outside of the standard
- PowerFlex supports multi-subnet or multi-VLAN for all network types other than data, vSAN, and NSX overlay
- PowerFlex supports multiple Cisco and Dell switch models with different firmware and operating system versions

## System components

PowerFlex rack contains compute, network, storage, virtualization, and management resources.

The following table shows supported and optional components of PowerFlex rack:

Resource	Components
Compute	PowerFlex R650/R750/R6525/R7525/R640/R740xd/R840 nodes
Network	<p>Management switch options:</p> <ul style="list-style-type: none"> <li>● Dell PowerSwitch S4148T-ON</li> <li>● Cisco Nexus 92348GC-X</li> </ul> <p>Management aggregation switch option:</p> <ul style="list-style-type: none"> <li>● Cisco Nexus 93180YC-FX</li> <li>● Cisco Nexus 93180YC-FX3</li> </ul> <p>Aggregation switch options:</p> <ul style="list-style-type: none"> <li>● Cisco Nexus 9336C-FX2</li> <li>● Dell PowerSwitch S5232F-ON</li> </ul> <p>Access switch options:</p> <ul style="list-style-type: none"> <li>● Cisco Nexus 93240YC-FX2</li> <li>● Cisco Nexus 93180YC-FX</li> <li>● Cisco Nexus 93180YC-FX3</li> <li>● Dell PowerSwitch S5248F-ON</li> </ul> <p>Spine switch options:</p> <ul style="list-style-type: none"> <li>● Cisco Nexus 9336C-FX2</li> <li>● Cisco Nexus 9364C-GX</li> </ul> <p>Leaf switch options:</p> <ul style="list-style-type: none"> <li>● Cisco Nexus 93240YC-FX2</li> <li>● Cisco Nexus 9336C-FX2</li> <li>● Cisco Nexus 9364C-GX</li> </ul> <p>Border-leaf switch option:</p> <ul style="list-style-type: none"> <li>● Cisco Nexus 9336C-FX2</li> </ul> <p>Customer access switch (optional ACI switches Dell provided in a dual network installation)</p> <ul style="list-style-type: none"> <li>● Cisco Nexus 93240YC-FX2</li> <li>● Cisco Nexus 93180YC-EX</li> </ul>
Storage	PowerFlex software-defined storage
Virtualization	VMware vSphere ESXi

Resource	Components
Management	PowerFlex Manager software
Management controller	PowerFlex management controller 1.0 based on vSAN PowerFlex management controller 2.0 based on PowerFlex
Data encryption (optional)	Dell CloudLink or self-encrypting drives (SEDs)

## Storage schemas

### Protection domains

A protection domain (PD) is a group of nodes or storage data servers that provide data isolation, security, and performance benefits. A node participates in only one protection domain at a time. Only nodes in the same protection domain can affect each other, nodes outside the protection domain are isolated. Secure multi-tenancy can be created with protection domains since data does not mingle across protection domains. You can create different protection domains for different node types with unequal performance profiles. All the hosts in the domain must have the same type and configuration. A PowerFlex hyperconverged node should not be in the same protection domain as a PowerFlex storage-only node. The node configuration must match, which includes the drives, CPU, and memory. Any difference in the node configuration leads to an unknown performance impact.

### Storage pools

Storage pools are a subset of physical storage devices in a protection domain. Each storage device belongs to one (and only one) storage pool. The best practice is to have the same type of storage devices (HDD versus SSD or SSD versus NVMe) within a storage pool to ensure that the volumes are distributed over the same type of storage within the protection domain.

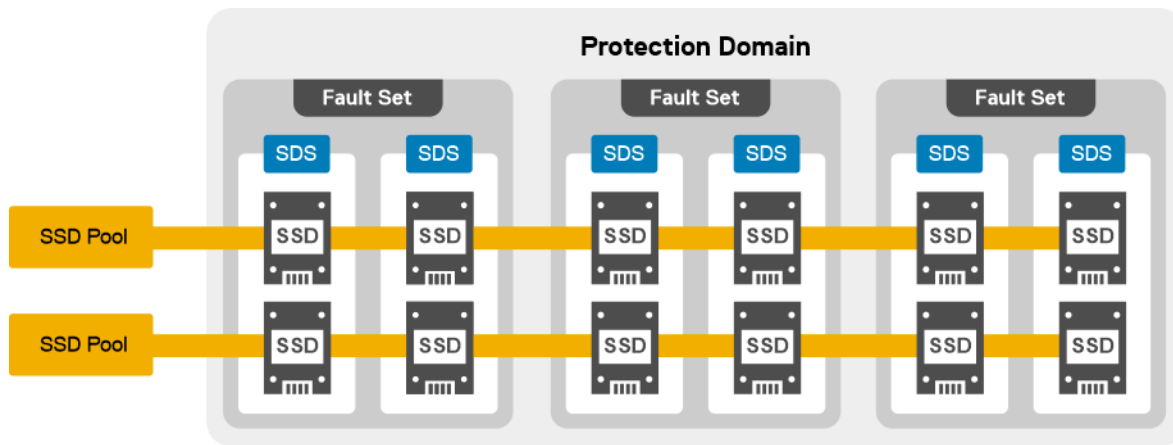
PowerFlex supports two types of storage pools. You can choose between both layouts. A system can support both fine granularity (FG) and medium granularity (MG) pools on the same storage data server nodes. Volumes can be non-disruptively migrated between the two layouts. Within an fine granularity pool, you can enable or disable compression on a per-volume basis:

- **Medium granularity:** Volumes are divided into 1MB allocation units, distributed, and replicated across all disks contributing to a pool. MG storage pools support either thick or thin-provisioned volumes, and no attempt is made to reduce the size of user-data written to disk (except with all-zero data). MG storage pools have higher storage access performance than fine granularity storage pools but use more disk space.
- **Fine granularity:** A space efficient layout, with an allocation unit of just 4 KB and a physical data placement scheme based on log structure array (LSA) architecture. Fine granularity layout requires both flash media (SSD or NVMe) as well as NVDIMM to create an fine granularity storage pool. fine granularity layout is thin-provisioned and zero-padded by nature, and enables PowerFlex to support in-line compression, more efficient snapshots, and persistent checksums. FG storage pools use less disk space than MG storage pools but have slightly lower storage access performance.

### Fault sets

A fault set is a logical entity that contains a group of storage data servers within a protection domain that have a higher chance of going down together; for example, if they are all powered in the same rack. By grouping them into a fault set, PowerFlex mirrors data for a fault set on storage data servers that are outside the fault set. Thus, availability is assured even if all the servers within one fault set fail simultaneously.



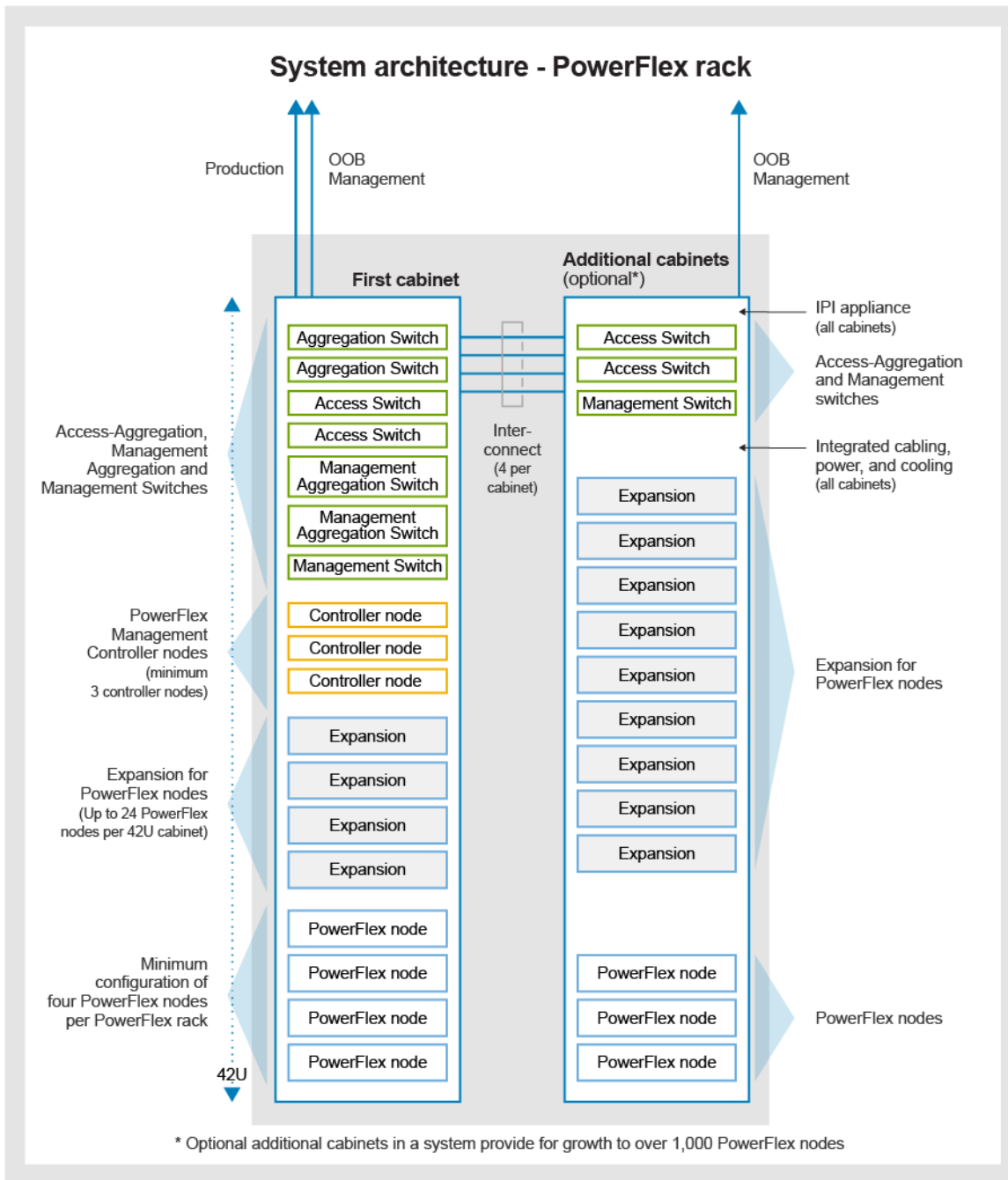


## Network architecture

PowerFlex rack is available as two different network architectures to meet the requirements for different performance and scaling requirements.

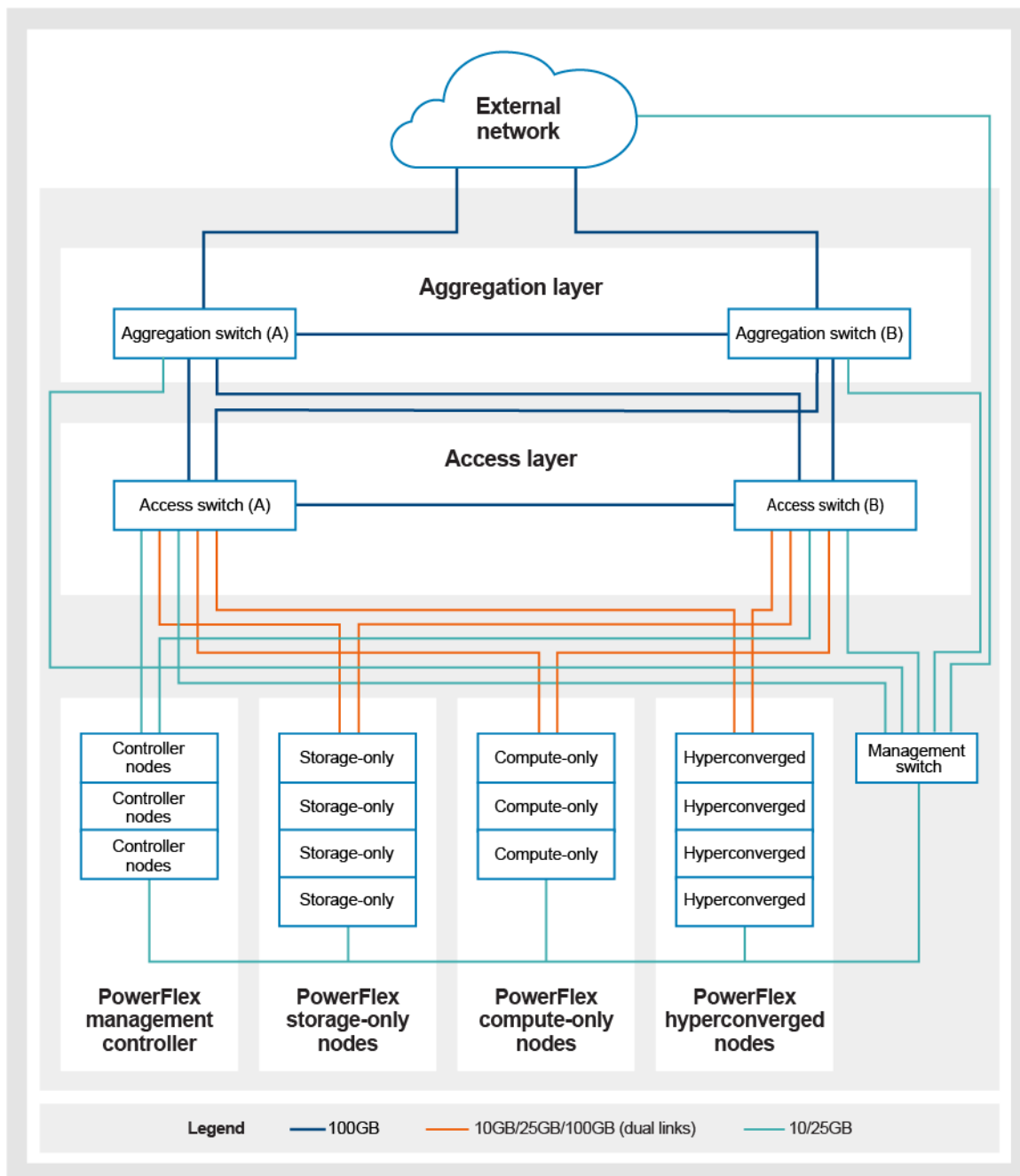
### Access and aggregation architecture

The following figure is a high-level, system layout of the PowerFlex rack using an access and aggregation architecture. This figure is not specific to a particular system, but is a generic representation.



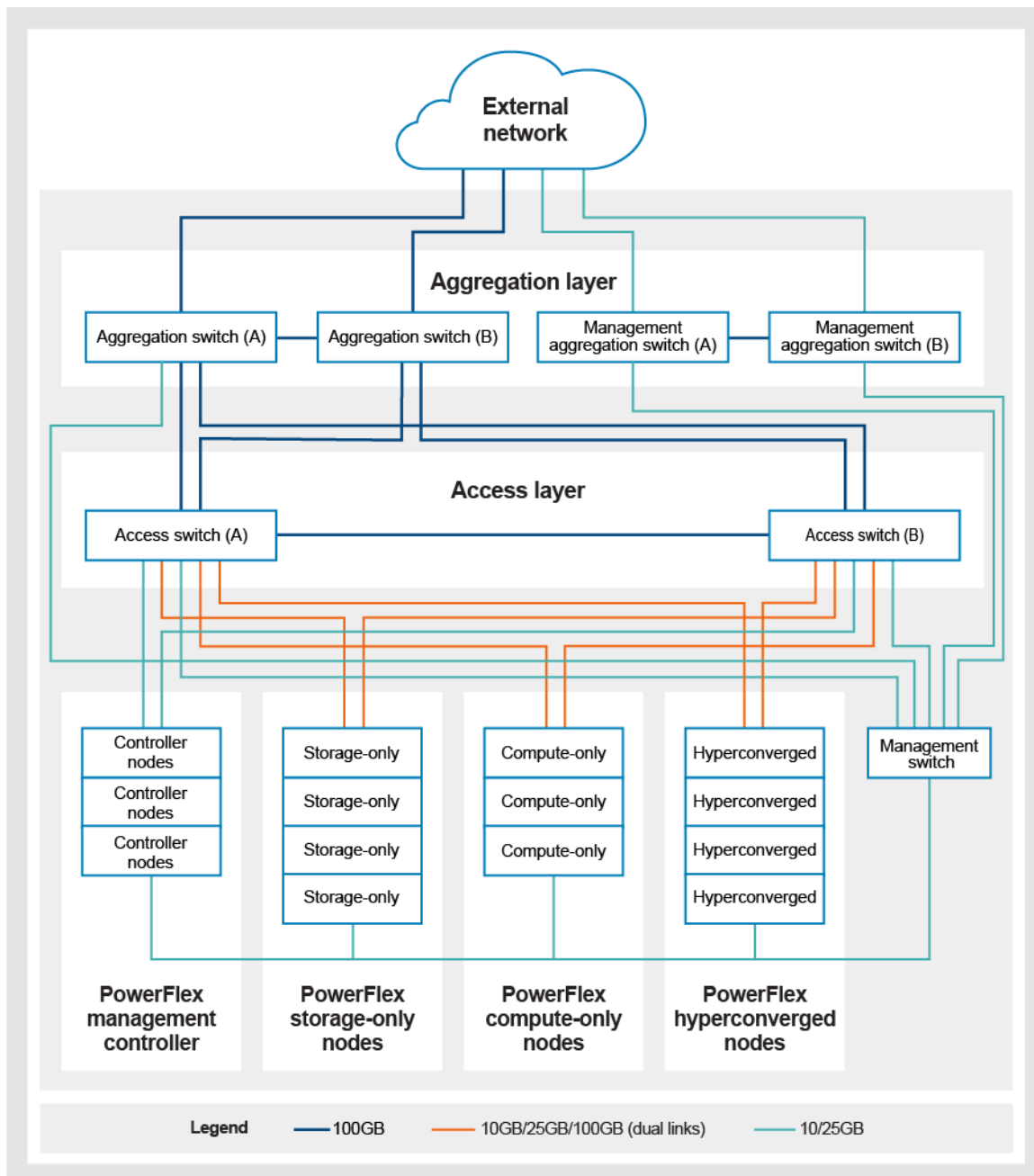
The following figure shows the logical layout of the access and aggregation architecture:

**NOTE:** There is an additional 1 Gb link from the PowerFlex controller nodes to the out-of-band management switch.



The following diagram shows the logical layout of the access and aggregation with management aggregation architecture:

**NOTE:** There is an additional 1 Gb link from the PowerFlex controller nodes to the out-of-band management switch.



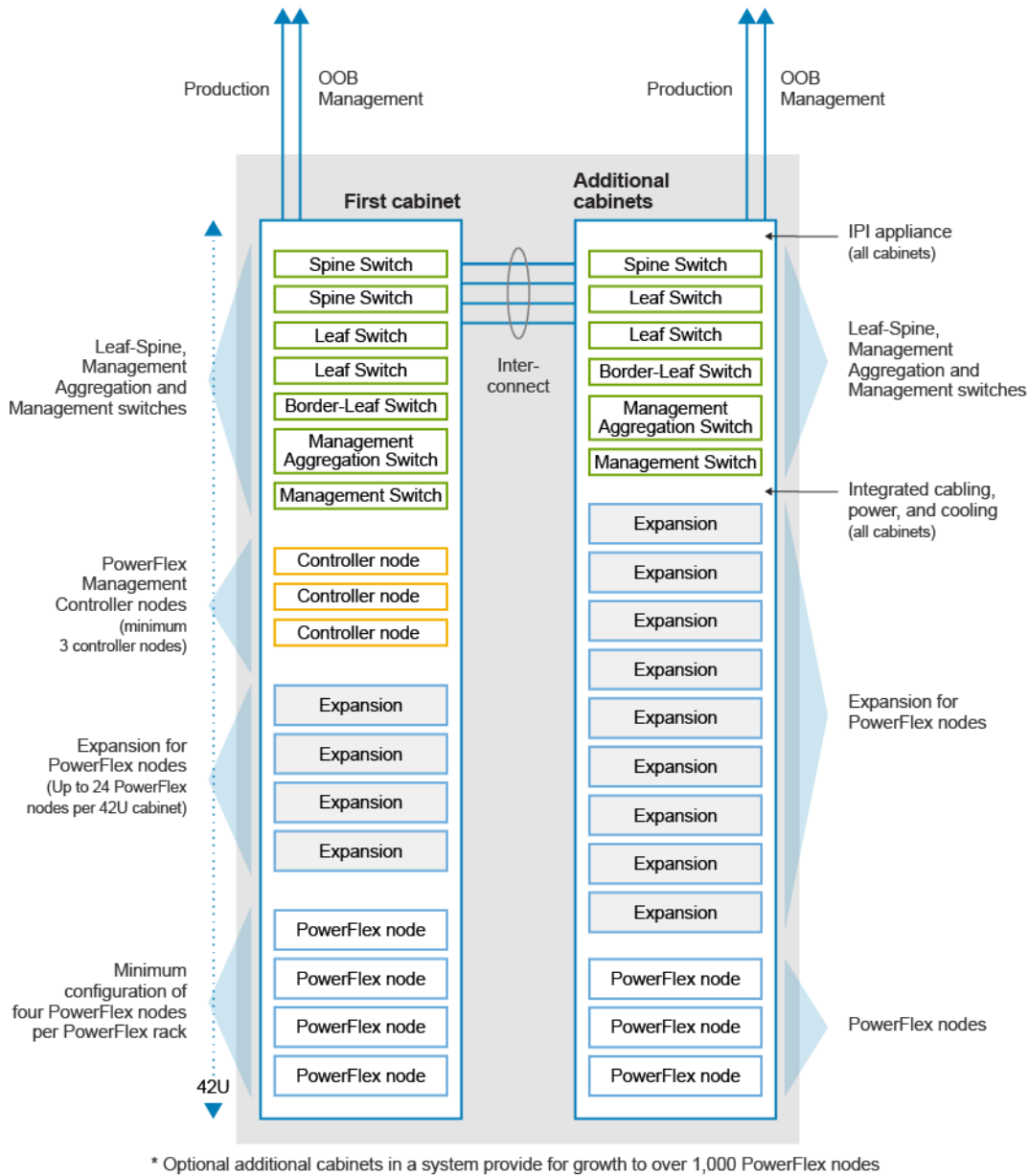
## Leaf-spine architecture

Leaf-spine architecture is 100 GbE connectivity from every spine to every leaf. Network switches provide two options to support leaf-to-node connectivity.

- 10 GbE or 25 GbE: Cisco Nexus 93240YC-FX2 switch
- 100 GbE: Cisco Nexus 9336-FX2, or 9364C or 9364-GX switches

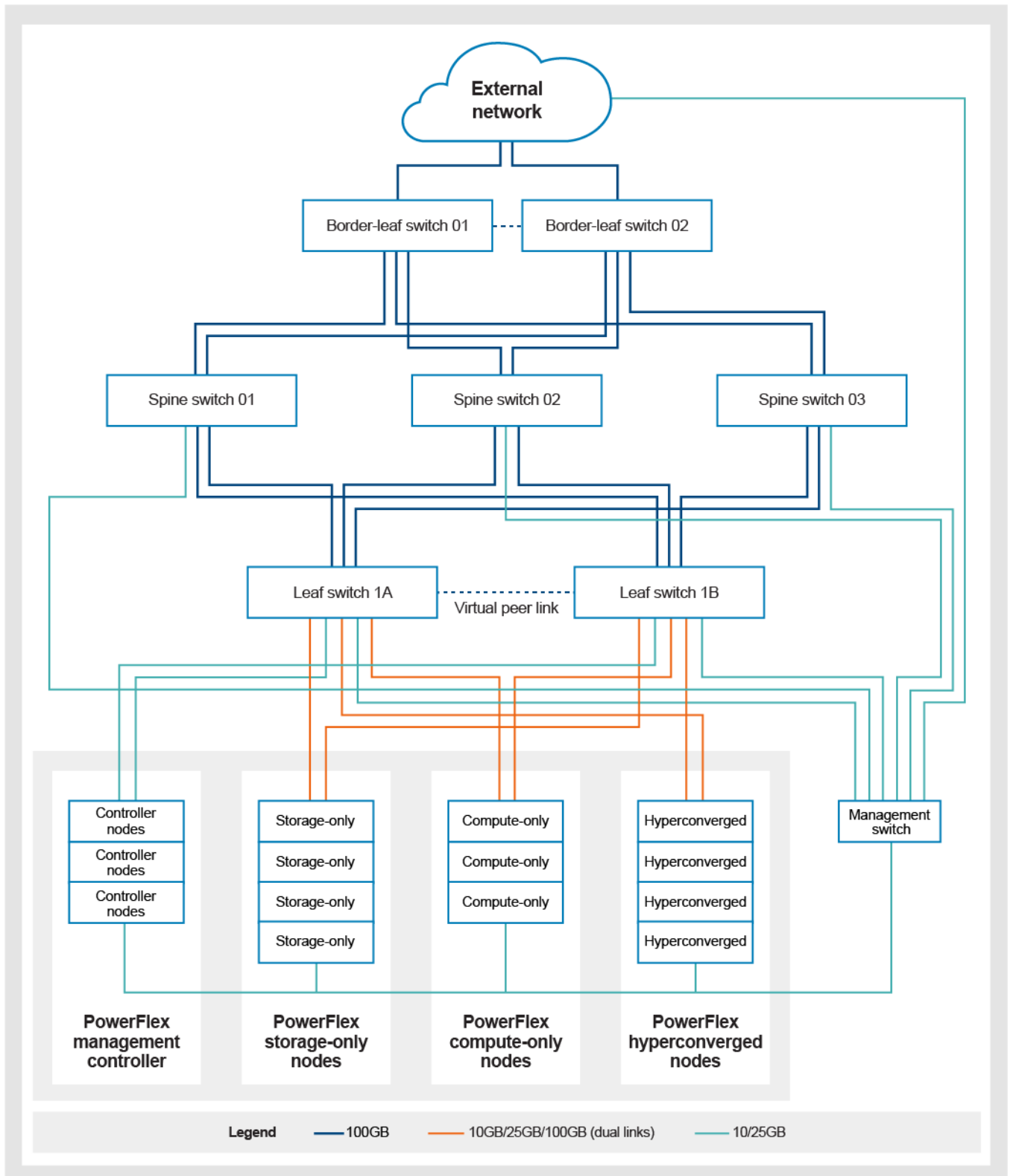
The following figure is a high-level, system layout of the physical PowerFlex rack, designed using a leaf-spine architecture:

## System architecture - PowerFlex rack



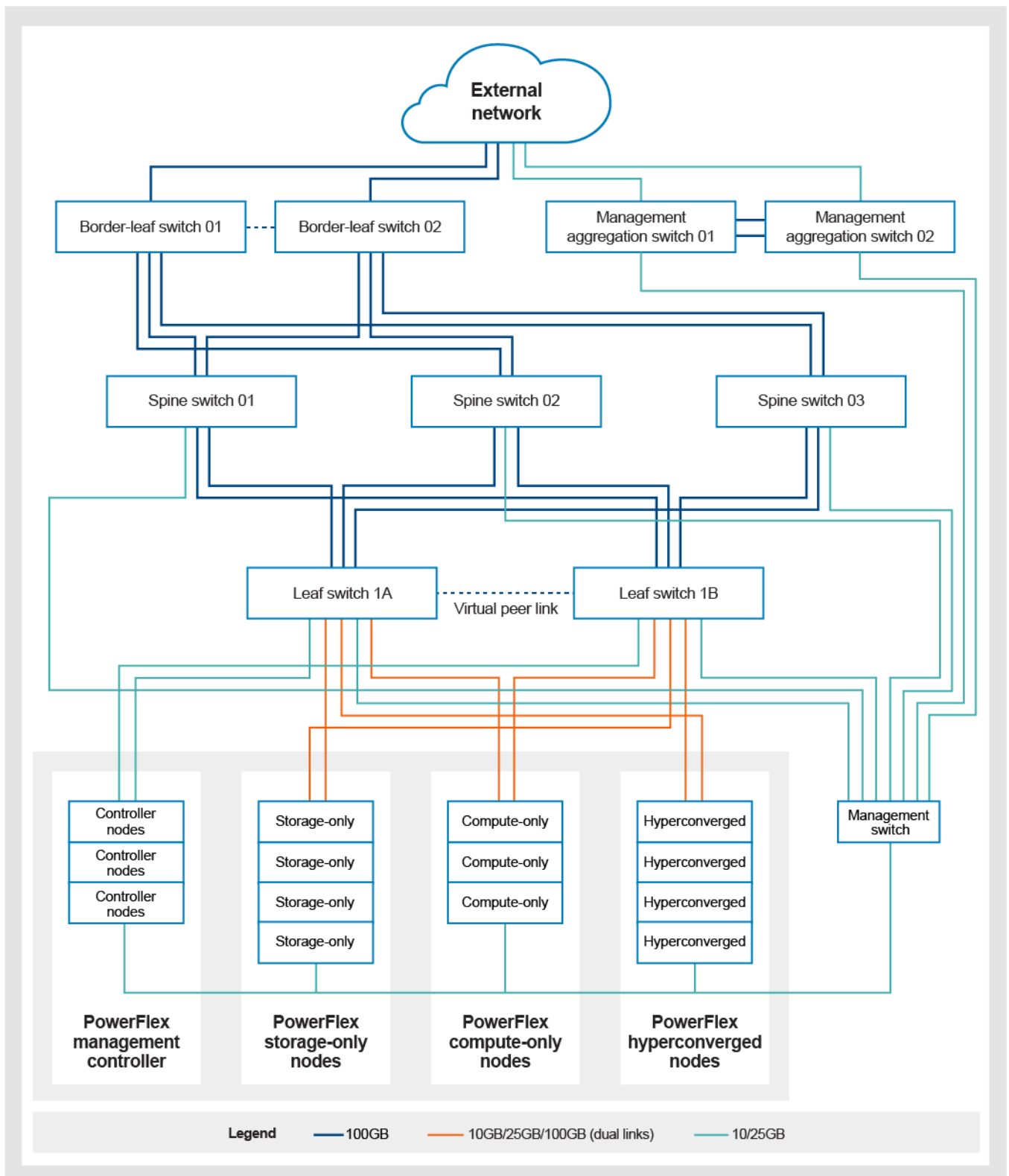
The following diagram shows the logical layout of the leaf-spine architecture:

**NOTE:** There is an additional 1 Gb link from the PowerFlex controller nodes to the out-of-band management switch.



The following diagram shows the logical layout of the leaf-spine architecture with optional management aggregation switches:

**NOTE:** There is an additional 1 Gb link from the PowerFlex controller nodes to the out-of-band management switch.




## Base configuration and scaling

### PowerFlex rack storage-only deployment

PowerFlex rack has a base configuration that is a minimum set of PowerFlex storage-only nodes and fixed network resources.


Within the base configuration, you can customize the following hardware aspects:

Hardware	Minimum set
Network	<ul style="list-style-type: none"> <li>One management switch</li> <li>One pair of access or leaf switches</li> <li>A pair of aggregation switches or three spine switches</li> <li>A pair of management aggregation switches (optional)</li> <li>A pair of border-leaf switches</li> </ul> <p> <b>NOTE:</b> Only in a leaf-spine configuration.</p>
Storage	<p>At least four PowerFlex storage-only nodes are required. However, Dell Technologies recommends using at least six nodes to build a PowerFlex storage pool.</p> <p>If storage compression is active, a minimum of two NVDIMM components per PowerFlex node are required. A recommendation is made according to the system sizing calculation.</p>
Management	Three PowerFlex management controller nodes with high availability.

### PowerFlex rack two-layer deployment

PowerFlex rack has a base configuration that is similar to a PowerFlex storage-only node deployment, but adds a minimum set of PowerFlex compute-only nodes. The minimum set of PowerFlex storage-only nodes and fixed network resources are also required.

Within the base configuration, you can customize the following hardware aspects:


Hardware	Minimum set
Compute	At least three PowerFlex compute-only nodes.
Network	<ul style="list-style-type: none"> <li>One management switch</li> <li>One pair of access or leaf switches</li> <li>A pair of aggregation switches or three spine switches</li> <li>A pair of management aggregation switches (optional)</li> <li>A pair of border-leaf switches</li> </ul> <p> <b>NOTE:</b> Only in a leaf-spine configuration.</p>
Storage	<p>At least four PowerFlex storage-only nodes are required. However, Dell Technologies recommends using at least six nodes to build a PowerFlex storage pool.</p> <p>Software-defined SAN storage (uses local disks to build a PowerFlex storage pool).</p> <p>If storage compression is active, a minimum of two NVDIMM components per PowerFlex node are required. A recommendation is made according to the system sizing calculation.</p>
Management	Three PowerFlex management controller nodes with high availability.



# PowerFlex rack hyperconverged deployment

PowerFlex rack has a base configuration that is a minimum set of hyperconverged components and fixed network resources.

Within the base configuration, you can customize the following hardware aspects:

Hardware	Minimum set
Compute and storage	<p>A minimum of four PowerFlex hyperconverged nodes are required, however, six is the recommended minimum. PowerFlex hyperconverged nodes provide both storage and compute resources to the system.</p> <p>If storage compression is active, a minimum of two NVDIMM components per PowerFlex node are required. A recommendation is made according to the system sizing calculation.</p>
Network	<ul style="list-style-type: none"><li>• One management switch</li><li>• One pair of access or leaf switches</li><li>• A pair of aggregation switches or three spine switches</li><li>• A pair of management aggregation switches (optional)</li><li>• A pair of border-leaf switches</li></ul> <p> <b>NOTE:</b> Only in a leaf-spine configuration.</p>
Management	Three PowerFlex management controller nodes with high availability.

## Additional references

This section provides references to related documentation for network, storage, and virtualization components.

### Network components

Network component information and links to documentation are provided.

Product	Link to documentation
Cisco Nexus 31108TC-V	<a href="https://www.cisco.com/c/en/us/products/collateral/switches/nexus-3000-series-switches/datasheet-c78-736608.html">https://www.cisco.com/c/en/us/products/collateral/switches/nexus-3000-series-switches/datasheet-c78-736608.html</a>
Cisco Nexus 93180YC-EX	<a href="http://www.cisco.com/c/en/us/products/switches/nexus-93180yc-ex-switch/index.html">http://www.cisco.com/c/en/us/products/switches/nexus-93180yc-ex-switch/index.html</a>
Cisco Nexus 93240YC-FX2	<a href="https://www.cisco.com/c/en/us/support/switches/nexus-93240yc-fx2-switch/model.html">https://www.cisco.com/c/en/us/support/switches/nexus-93240yc-fx2-switch/model.html</a>
Cisco Nexus 9336C-FX2	<a href="https://www.cisco.com/c/en/us/support/switches/nexus-9336c-fx2-switch/model.html">https://www.cisco.com/c/en/us/support/switches/nexus-9336c-fx2-switch/model.html</a>
Cisco Nexus 9364C	<a href="https://www.cisco.com/c/en/us/support/switches/nexus-9364c-switch/model.html">https://www.cisco.com/c/en/us/support/switches/nexus-9364c-switch/model.html</a>
Cisco Nexus 93180YC-FX	<a href="https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-742284.html">https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-742284.html</a>
Cisco Nexus 93180YC-FX3	<a href="https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-744052.html">https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-744052.html</a>
Cisco Nexus 9364C-GX	<a href="https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-741560.html">https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-741560.html</a>
Cisco Nexus 92348GC-X	<a href="https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-735989.html">https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-735989.html</a>
Dell PowerSwitch S5200F-ON series switches	<a href="https://www.delltechnologies.com/asset/en-us/products/networking/technical-support/dell_emc_networking-s5200_on_spec_sheet.pdf">https://www.delltechnologies.com/asset/en-us/products/networking/technical-support/dell_emc_networking-s5200_on_spec_sheet.pdf</a>
Dell PowerSwitch S4100T-ON series switches	<a href="https://i.dell.com/sites/doccontent/shared-content/data-sheets/en/Documents/dell-emc-networking-S4100-series-spec-sheet.pdf">https://i.dell.com/sites/doccontent/shared-content/data-sheets/en/Documents/dell-emc-networking-S4100-series-spec-sheet.pdf</a>

### Storage components

Storage component information and links to documentation are provided.

Product	Description	Link to documentation
PowerFlex	Converges storage and compute resources into a single-layer architecture, aggregating capacity and performance, simplifying management, and scaling to thousands of PowerFlex nodes.	<a href="https://www.delltechnologies.com/en-us/storage/powerflex.htm">https://www.delltechnologies.com/en-us/storage/powerflex.htm</a>
NVDIMM-N	Provides high-speed DRAM performance that is coupled with flash-backed persistent storage for PowerFlex storage-only nodes and PowerFlex hyperconverged nodes.	<a href="https://www.dell.com/support/manuals/us/en/04/poweredge-t640/nvdimm-n_ug_pub/introduction?">https://www.dell.com/support/manuals/us/en/04/poweredge-t640/nvdimm-n_ug_pub/introduction?</a>

Product	Description	Link to documentation
		<a href="#">guid=guid-8884370c-5553-4089-b613-a3c570b56f0e&amp;lang=en-us</a>

## Virtualization components

Virtualization component information and links to documentation are provided.

Product	Description	Link to documentation
VMware vCenter Server	Provides a scalable and extensible platform that forms the foundation for virtualization management.	<a href="http://www.vmware.com/products/vcenter-server/">www.vmware.com/products/vcenter-server/</a>
VMware vSphere ESXi	Virtualized infrastructure for PowerFlex racks. Virtualizes all application servers and provides VMware High Availability (HA) and Dynamic Resource Scheduling (DRS).	<a href="http://www.vmware.com/products/vsphere/">www.vmware.com/products/vsphere/</a>