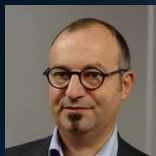


# Conseils pour le sizing d'une nouvelle infrastructure CEPH



## Réunion groupe X/Stra du 05/11/2020



*Jean-Marie Petry*  
*Pre Sales Engineer*



*Christophe Couty*  
*Account Executive*

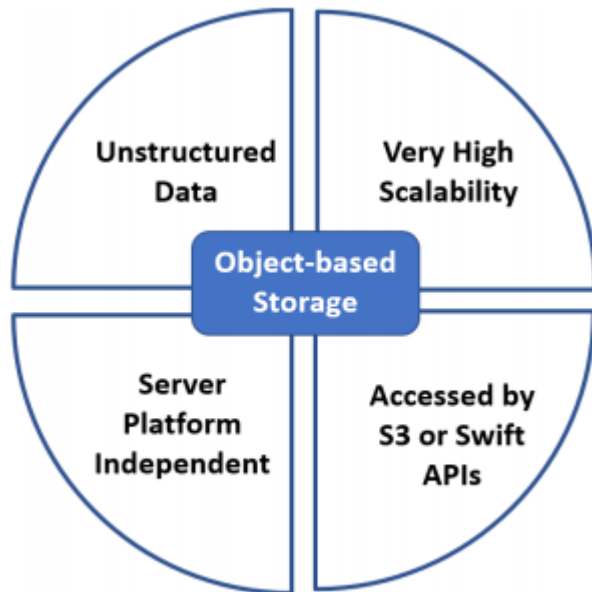
**DELL**Technologies

# Sommaire



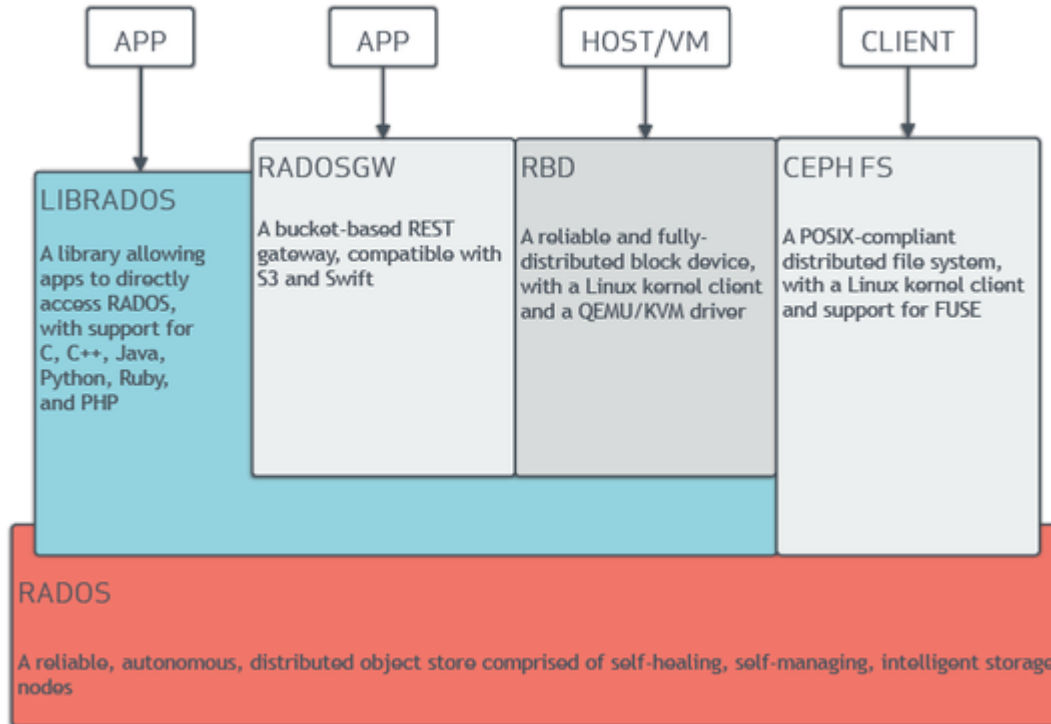
- Rappel de l'approche du stockage objet avec CEPH
- Rappel des briques architecturales CEPH
- Présentation de l'offre DELL EMC Ready Node Architecture pour CEPH
- Recommandations CPU/Mémoire/Stockage
- Exemples d'architectures cibles par capacité de stockage nécessaire
- Focus sur le nouveau PowerEdge XE7100
- Focus sur les solutions à base de stockage SSD NVMe

# Approche Stockage Objet / Ceph



- Structure à plat, indépendant de la localisation et conçu pour les données non structurées
- Architecture sans point de défaillance: No Single Point Of Failure (NSPOF)
- Métadonnées dynamiques (défini par l'utilisateur) et scalabilité jusqu'à 1000 nœuds (nœuds OSD et storage scale indépendamment)
- Accessible par application S3 ou Swift API, REST et SOAP
- Capacité de stockage en To voire Po : fichiers, vidéos, images, sauvegarde, fichiers ISO...
- Bien adapté au Cloud avec structure de fichiers à plat

# Rappel briques architecturales CEPH



# Offre Dell EMC Ready Architecture for Red Hat Platform Components



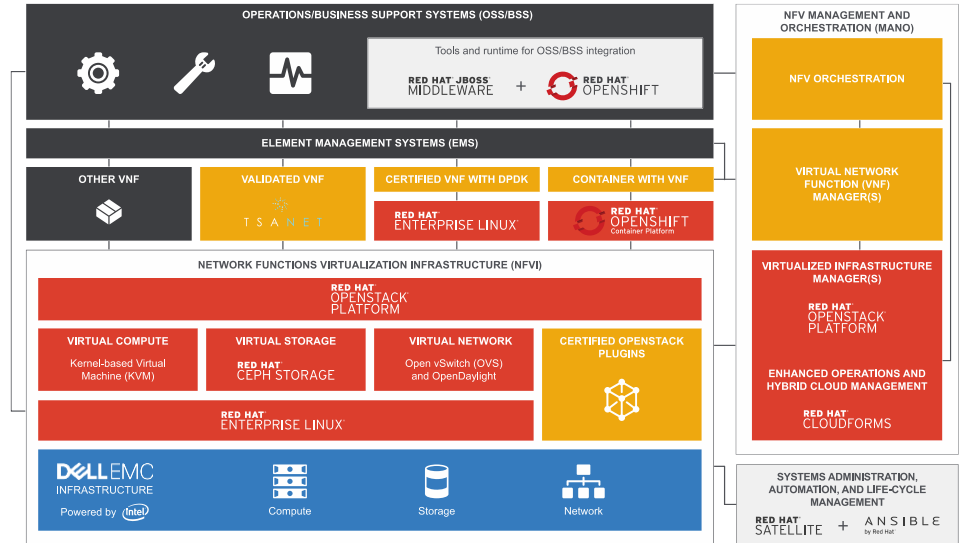
## RED HAT PLATFORM

- Kernel-based Virtual Machine
- Red Hat Ceph Storage
- Open vSwitch and Neutron plugins
- Cinder plugins
- Red Hat Enterprise Linux



## DELL EMC INFRASTRUCTURE

- Dell EMC Networking
  - S5232F-ON, S5224F-ON (25GbE)
  - S3048T-ON, S4048-ON (1GbE et 10GbE)
- Dell EMC PowerEdge XE2420 et XE7100
- Dell EMC Servers with 2nd Generation Intel® Xeon® Scalable Processor
  - R6525 / R7525 (AMD Qlogic 25GbE)
  - R640/R740 (Intel XXV710 25GbE)
  - R740XD (Intel XXV710 25GbE & MLNX CX6 DX 100GbE)
- Dell EMC Unity
  - Block, image, File system



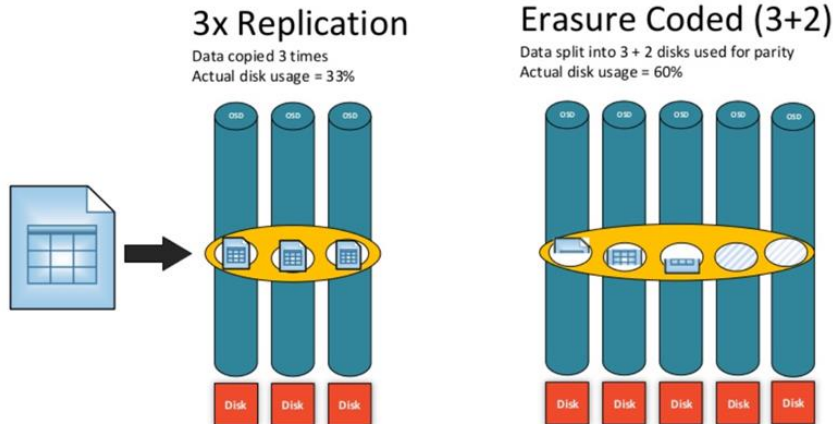
Lien vers site Web Dell dédié pour architecture Openstack & Ceph architecture:

<https://www.dell.com/support/article/fr-fr/sln310368/dell-emc-ready-and-reference-architectures-for-openstack-platform-with-general-solutions?lang=en>

# Généralités sur les configurations CEPH



- Règles générales (cf slides suivants):
  - 1 Cœur-Ghz par OSD
  - SATA/SAS SSD pour HDD ratio: 1:4 à 1:5
  - NVMe SSD pour HDD ratio : 1:17 à 1:18
  - 16GB RAM de base + 2/3Go par OSD
- Sécurisation de la données:



# Recommandations CPU et mémoire

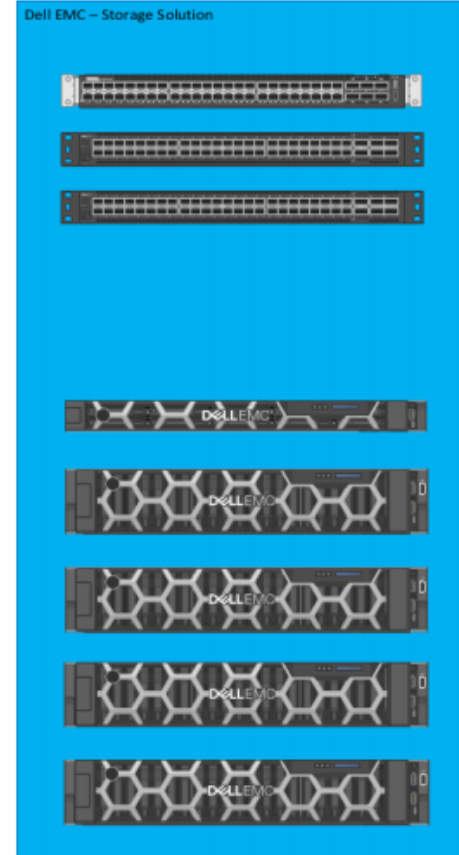
Exemple: cluster 4 nœuds

Component	Min. RAM per instance (GB)	Recommended RAM per instance (GB)	Instances	Total Min. RAM (GB)	Total Recommended RAM (GB)
Operating system	16	16	1	16	16
Ceph OSD	2	8	12	24	96
Ceph MON	1	1	1	1	1
Ceph MGR	1	1	1	1	1
Ceph RGW	1	1	1	1	1
Total	-	-	-	43	115

Component	Min. cores per instance	Recommended cores per instance	Instances	Total min. cores	Total recommended cores
Operating system	2	2	1	2	2
Ceph OSD HDD	0.5 (of 2GHz CPU)	1	12	6	12
Ceph MON	1	1	1	1	1
Ceph MGR	1	1	1	1	1
Ceph RGW	3	3	1	3	3
Total	-	-	-	13	19

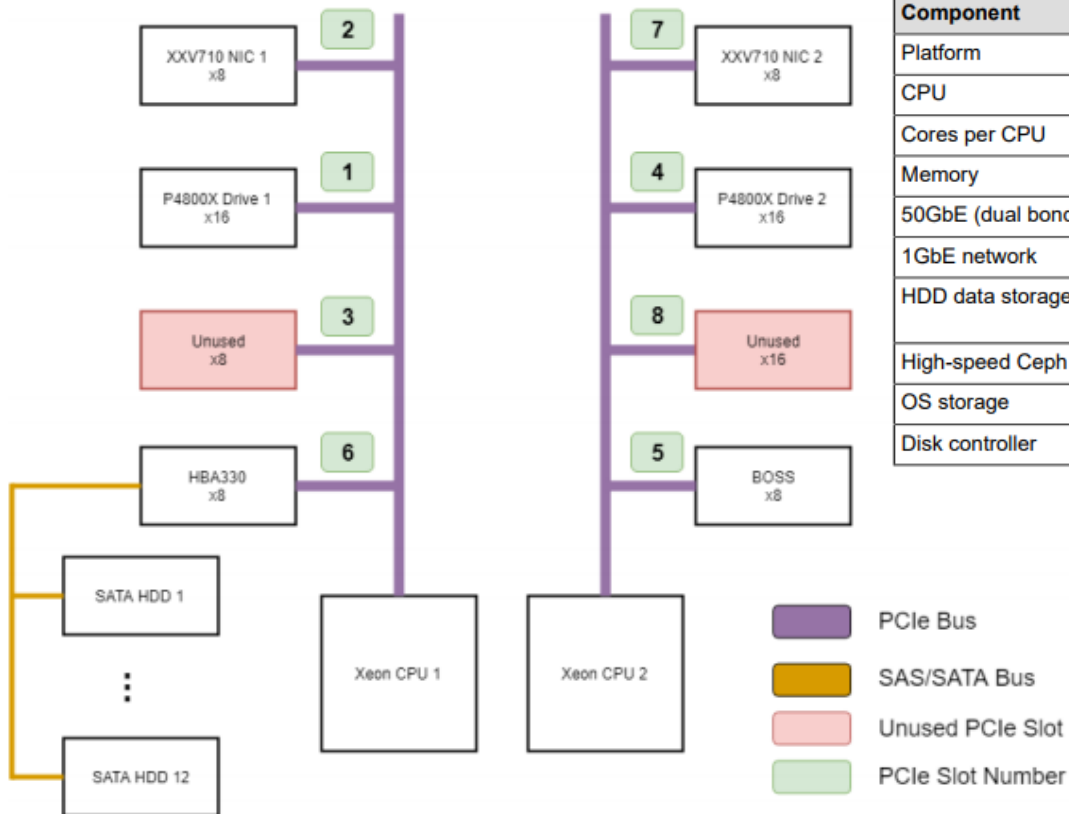
Lien vers doc Object Storage Architecture Guide:

[https://www.delltechnologies.com/resources/en-us/asset/technical-guides-support-information/solutions/red\\_hat\\_ceph\\_storage\\_v3-2\\_object\\_storage\\_architecture\\_guide.pdf](https://www.delltechnologies.com/resources/en-us/asset/technical-guides-support-information/solutions/red_hat_ceph_storage_v3-2_object_storage_architecture_guide.pdf)





# Recommandations Stockage



Component	Details
Platform	Dell EMC PowerEdge R740xd
CPU	2x Intel® Xeon® Silver (see note) 4114 2.2 GHz
Cores per CPU	10
Memory	192 GB (12x 16GB RDIMM, 2666MT/s)
50GbE (dual bonded 25GbE) network	2x Intel® XXV710 Dual Port 25GbE SFP28
1GbE network	i350 Quad Port 1GbE, rNDC
HDD data storage	12x 10TB (see caution) 7.2K RPM 3.5" SATA HDD 6Gbps
High-speed Ceph metadata storage	2x Intel® Optane® (see note) P4800X 750GB
OS storage	BOSS (2x M.2 Sticks 240G RAID 1)
Disk controller	HBA330 (JBOD)

Lien vers doc Object Storage Architecture Guide:

[https://www.delltechnologies.com/resources/en-us/asset/technical-guides-support-information/solutions/red\\_hat\\_ceph\\_storage\\_v3-2\\_object\\_storage\\_architecture\\_guide.pdf](https://www.delltechnologies.com/resources/en-us/asset/technical-guides-support-information/solutions/red_hat_ceph_storage_v3-2_object_storage_architecture_guide.pdf)



# Carte BOSS

## Qu'est-ce que c'est ?

Les solutions Software Defined Storage (SDS) imposent systématiquement une solution proposant un HW controller pour l'OS: cela se traduit par l'implémentation de cartes BOSS :

- 2 disques M2 SSD **240GB**  
ou
- 2 x M2 SSD 480GB

## Spécifications

- One or Two 110mm M.2 SATA devices
- M.2 Devices are Read Intensive 120GB/240GB
  - *Same Cost & Performance as 2.5"*
- Fixed function **Hardware RAID 1** (mirroring)
- Single x8 PCIe Gen 3 host interface
- Dual x1 SATA ports for device interfaces
- Presents single virtualized SATA device to the host
- Half height / half length PCIe adapter module

## OS Supportés

- Windows 2012R2 / 2016 ou 2019
- RHEL 6.9 / 7.3 ou supérieur / 8.x
- SLES 12 SP2
- VMWare ESXi 6.0 U3 / 6.5 / 6.7 / 7.0



# Announce PowerEdge XE7100 – Extreme Dense



- ✓ Accelerated analytics performance with up to **4x** GPUs combined with internal storage<sup>1</sup>
- ✓ Boost deep learning and inference performance with **2x** 2<sup>nd</sup> Gen Intel® Xeon® Scalable Processors, up to 52 cores<sup>2</sup> per node
- ✓ Hot plug architecture for ease of serviceability without powering down



5U

## TARGET WORKLOADS

## PowerEdge XE7100



### OBJECT STORAGE

Dual, ½ width nodes optimized storage-to-compute ratio



### INTELLIGENT VIDEO ANALYTICS

Full width node, accelerated analytics performance with V100S GPU<sup>1</sup>



### MEDIA STREAMING<sup>3</sup>

Full width node, balanced I/O with optimized PCIe flexibility; up to 4x T4 or NICs

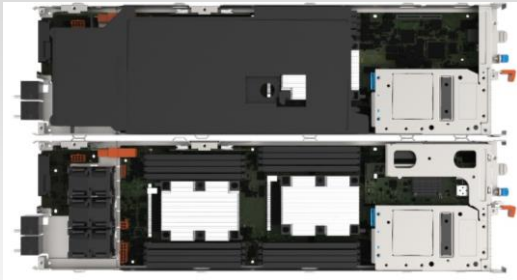
1. GPU options available post RTS ( 4 x T4 ou 2 x V100S)  
2. Optimal PCIe Flexibility configuration option will be post-RTS  
3. Support 2<sup>nd</sup> Gen Intel Xeon Scalable Processors post-RTS, with up to 26 cores/ CPU

# PowerEdge XE7100 workload-optimized configuration options



## OBJECT STORAGE

**XE7420**  
Dual, ½ width nodes

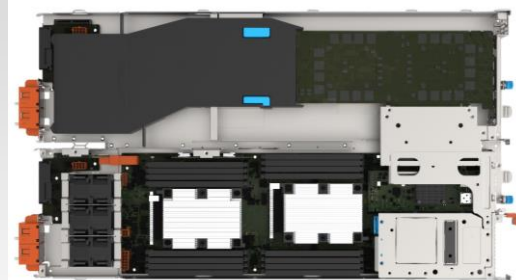


### Balanced compute-to-storage ratio

Each node has 1x LP PCIe slot

## INTELLIGENT VIDEO ANALYTICS (IVA)

**XE7440**  
Full width node



### Cost-optimized accelerator option

- 1 x FH PCIe slot to support<sup>1</sup>
- 1x GPU option (V100S super) or 1x T4, or
  - 1x FE1

## MEDIA STREAMING

**XE7440**  
Full width node



### Maximized PCIe flexibility<sup>2</sup>

- 4x LP PCIe slots to support:
- Up to 4x T4, or
  - Up to 4x NICs

1. GPU options available post-RTS  
2. Optimal PCIe Flexibility configuration option available post-RTS

# Exemples d'architectures Ceph



Capacité	Extra small	Small	Medium	High
Capacité cluster	100TB+	500TB+	1PB+	> 2 PB
Throughput optimized	4 x R740XD (8U) 16 x 6TB 1x800GB NVMe SSD 2 x 10GbE 3 x Replication	8 x R740XD (16U) 16 x 6TB 1x800GB NVMe SSD 2 x 10GbE 3 x Replication		
Cost/Capacity optimized			15 x R740XD (30U) 16 x 8TB 1x AIC NVMe SSD 2 x 10GbE 8:3 x Erasure Coding	XE7100 Storage Dense  **A venir**

Lien vers doc Object Storage Architecture Guide:

[https://www.delltechnologies.com/resources/en-us/asset/technical-guides-support-information/solutions/red\\_hat\\_ceph\\_storage\\_v3-2\\_object\\_storage\\_architecture\\_guide.pdf](https://www.delltechnologies.com/resources/en-us/asset/technical-guides-support-information/solutions/red_hat_ceph_storage_v3-2_object_storage_architecture_guide.pdf)

# Focus Solutions NVMe



## Serveur R640 1U :

Jusqu'à 10 disques NVMe 2,5 pouces

\*8 NVMe sans partage de PCIe Lanes (x4 per drive) // max 48 PCIe lanes Intel Cascade Lake



## Serveur R740XD 2U :

Jusqu'à 12 ou 24 disques NVMe 2,5 pouces (attention: partage de PCIe lanes selon le design)



## Serveur R6515/R6525 EPYC2 1U :

Jusqu'à 10+2 x Direct access NVMe / disques NVMe 2,5 pouces



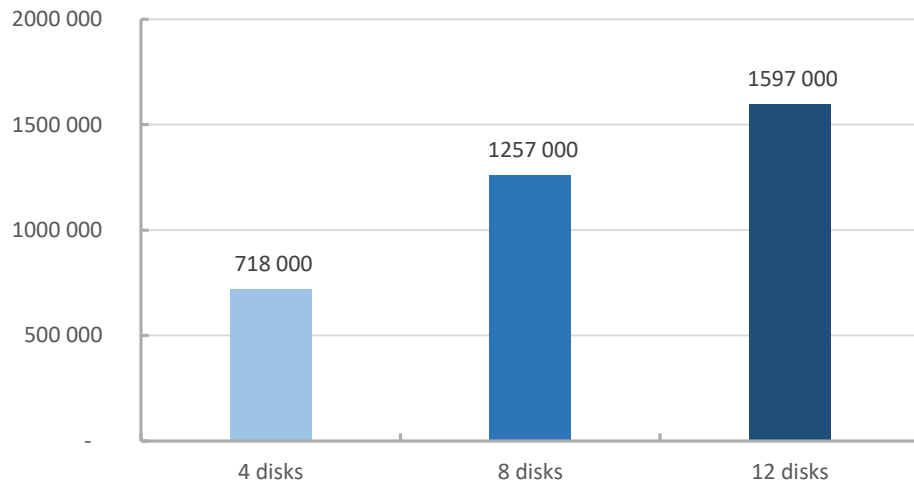
## Serveur R7515/R7525 EPYC2 2U : 160 PCIe Lanes Gen4

Jusqu'à 24 x Direct access NVMe / disques NVMe 2,5 pouces

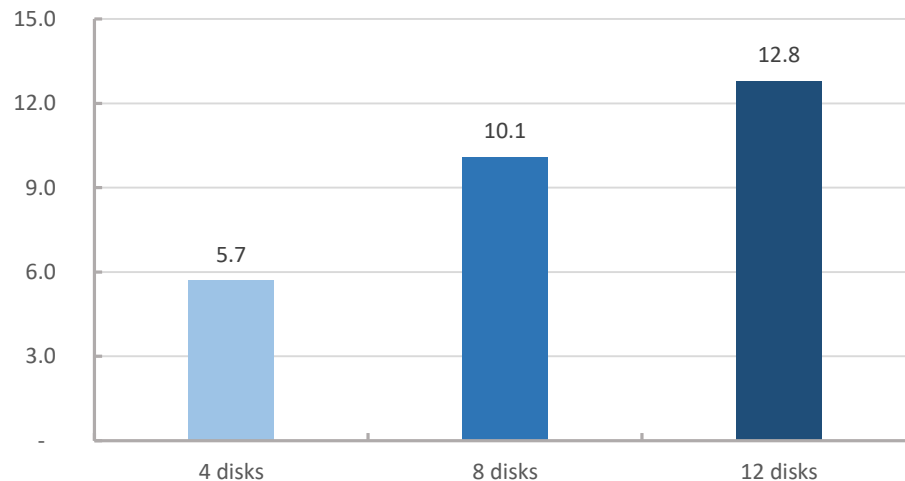
# Performance NVMe



Average IOPS : OLTP Profile



Average throughput (GBps) : OLTP Profile



OLTP profile : 8KB transfer size – 70% read / 30% write – 100% random  
The solution consistently provided average response time below 1ms

Source : A principled technology report : Get the server I/O capacity you need for heavy mixed, write, and read workloads,  
[http://www.principledtechnologies.com/Dell/PowerEdge\\_R740xd\\_SSD\\_performance\\_0717.pdf](http://www.principledtechnologies.com/Dell/PowerEdge_R740xd_SSD_performance_0717.pdf)

# Ressources documentaires



Lien vers site Dell pour architecture Cloud Openstack & Ceph architecture:

<https://www.dell.com/support/article/fr-fr/sln310368/dell-emc-ready-and-reference-architectures-for-openstack-platform-with-general-solutions?lang=en>

Lien vers doc référence RedHat pour configuration Ceph:

[https://access.redhat.com/documentation/en-us/red\\_hat\\_ceph\\_storage/3/html/configuration\\_guide/index](https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/3/html/configuration_guide/index)

Gamme PowerEdge XE

<https://www.delltechnologies.com/en-us/servers/specialty-servers/poweredge-xe-servers.htm>

Caractéristiques techniques PowerEdge XE7100 5U

[https://topics-cdn.dell.com/pdf/poweredge-xe7100\\_reference-guide\\_fr-fr.pdf](https://topics-cdn.dell.com/pdf/poweredge-xe7100_reference-guide_fr-fr.pdf)

**DELL**Technologies