

Storage Spaces Direct

Le géo-cluster SQL Server simplifié

Christophe LAPORTE

SQL Server MVP / MCM



Christophe Laporte



Microsoft®
SQL Server®

~ since 1997

6.5 <= SQL Server <= 2016



christophe_laporte@hotmail.fr



<http://conseilit.wordpress.com/>



Microsoft®
CERTIFIED

Master



@conseilit



Christophe Laporte



Microsoft®
SQL Server®

~ since 1997

6.5 <= SQL Server <= 2016



christophe_laporte@hotmail.fr



<http://conseilit.wordpress.com/>

- Migrations
- Formations
- Remote DBA
- Hébergement BDD



Microsoft®
CERTIFIED

Master



@conseilit

- Conseil
 - Infrastructure / Architecture
 - Virtualisation / Cloud
 - Haute disponibilité / Montée en charge
 - Optimisation / Dépannage
- Audit

Agenda

- Rappel sur la haute disponibilité
- Le cluster de basculement
- Solution pour éviter le point de défaillance unique
 - Espaces de stockage direct / Storage Spaces Direct
 - Groupes de disponibilité / AlwaysOn Availability Groups

Pourquoi la HA ?

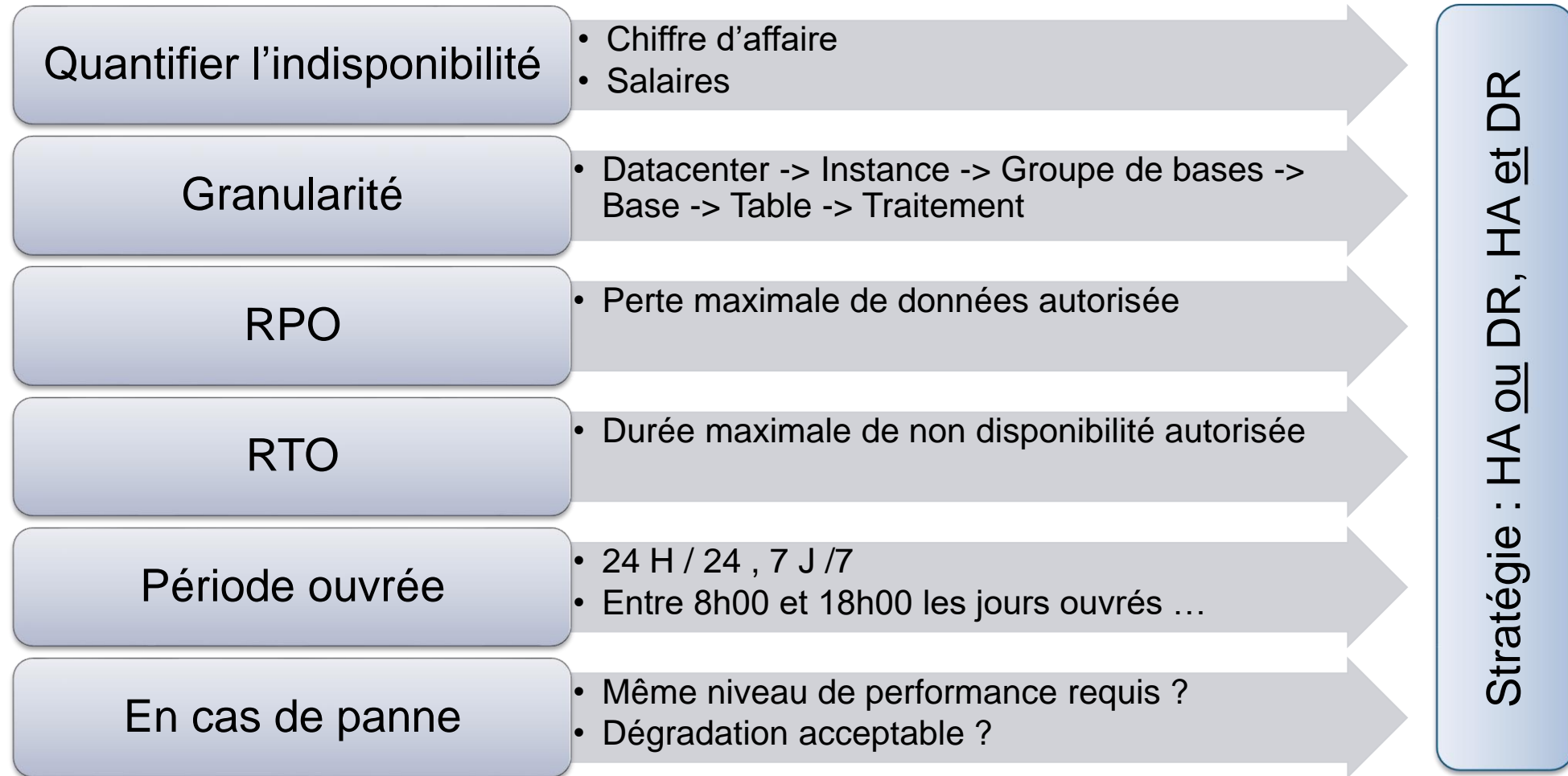
- Définition basique
 - Etre capable d'accéder à une donnée lorsque l'on en a besoin dans un laps de temps acceptable !
- BD point central dans le SI
 - Sharepoint, sites Web de paris ou commerce en ligne
 - Progiciels (RH, Compta, production, CRM)
 - Logiciels « maison »
- La non disponibilité a un coût
 - Chiffre d'affaire ...
 - Coût en temps
 - Salaires d'employés ...
- La haute disponibilité peut affecter les performances
 - Mise à jour synchrone des données sur un réplica
 - SAN Mirroring
 - Cluster (SAN: mutualisation et flexibilité)
- Ne pas confondre PCA et PRA
 - HA et DR

Disponibilité en %	Indisponibilité par année	Indisponibilité par mois	Indisponibilité par semaine
90 % (« un neuf »)	36,5 jours	72 heures	16,8 heures
95 %	18,25 jours	36 heures	8,4 heures
98 %	7,30 jours	14,4 heures	3,36 heures
99 % (« deux neuf »)	3,65 jours	7,20 heures	1,68 heures
99,5 %	1,83 jours	3,60 heures	50,4 minutes
99,8 %	17,52 heures	86,23 minutes	20,16 minutes
99,9 % (« trois neuf »)	8,76 heures	43,2 minutes	10,1 minutes
99,95 %	4,38 heures	21,56 minutes	5,04 minutes
99,99 % (« quatre neuf »)	52,56 minutes	4,32 minutes	1,01 minutes
99,999 % (« cinq neuf »)	5,26 minutes	25,9 secondes	6,05 secondes
99,9999 % (« six neuf »)	31,5 secondes	2,59 secondes	0,605 secondes

Causes de non disponibilité

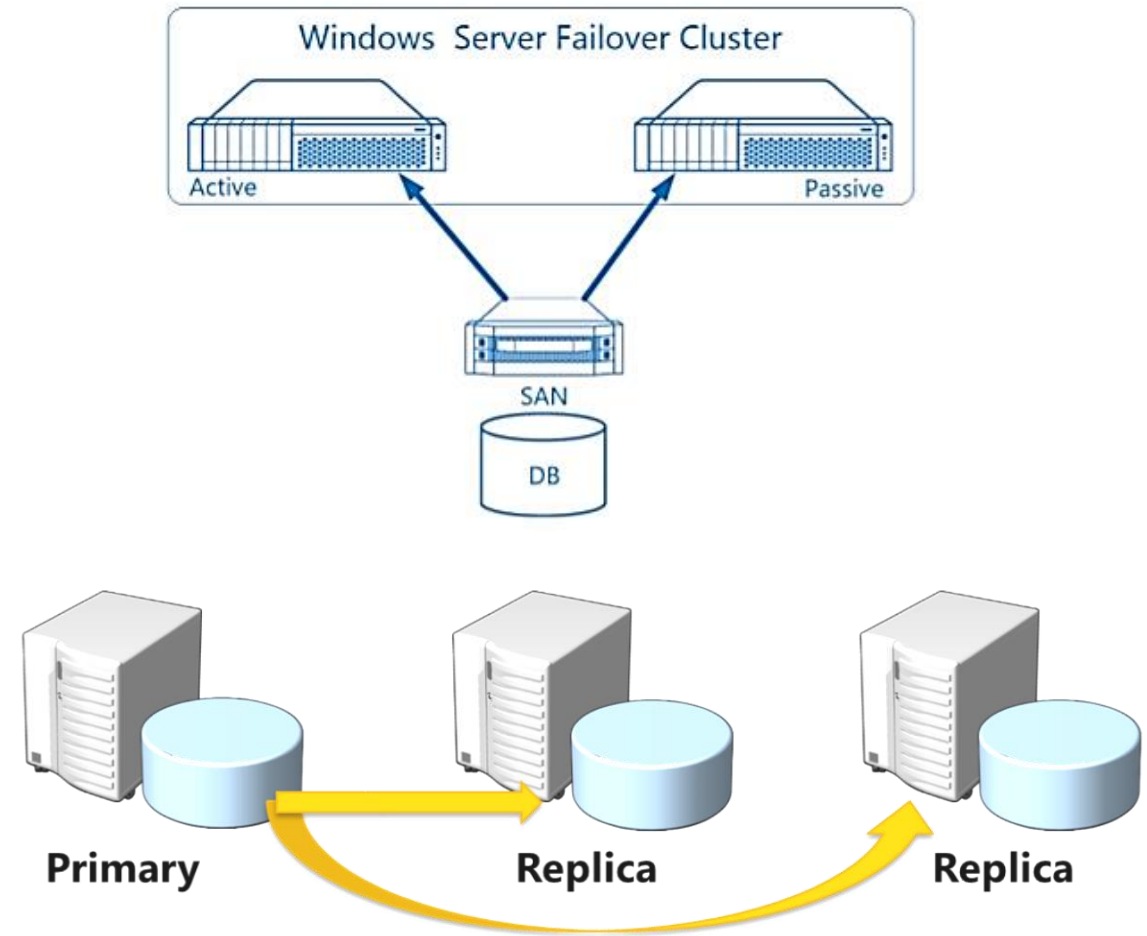
- Coupure de service planifiée
 - Changement de matériel
 - Application de Service Packs
- Coupure de service non planifiée
 - Perte du Datacenter (électricité, réseau, catastrophe naturelle)
 - Perte du serveur (alimentation, CPU, mémoire, réseau, OS crash)
 - Problème disque (corruption d'I/O, panne contrôleur disque, panne disque, panne carte RAID)
 - Problème sur le serveur / VM
 - Problème sur l'instance SQL
 - Problème sur la base
 - Problème sur une table
- Tâches de maintenance
 - Réindexation

Définition d'une stratégie PCA/PRA



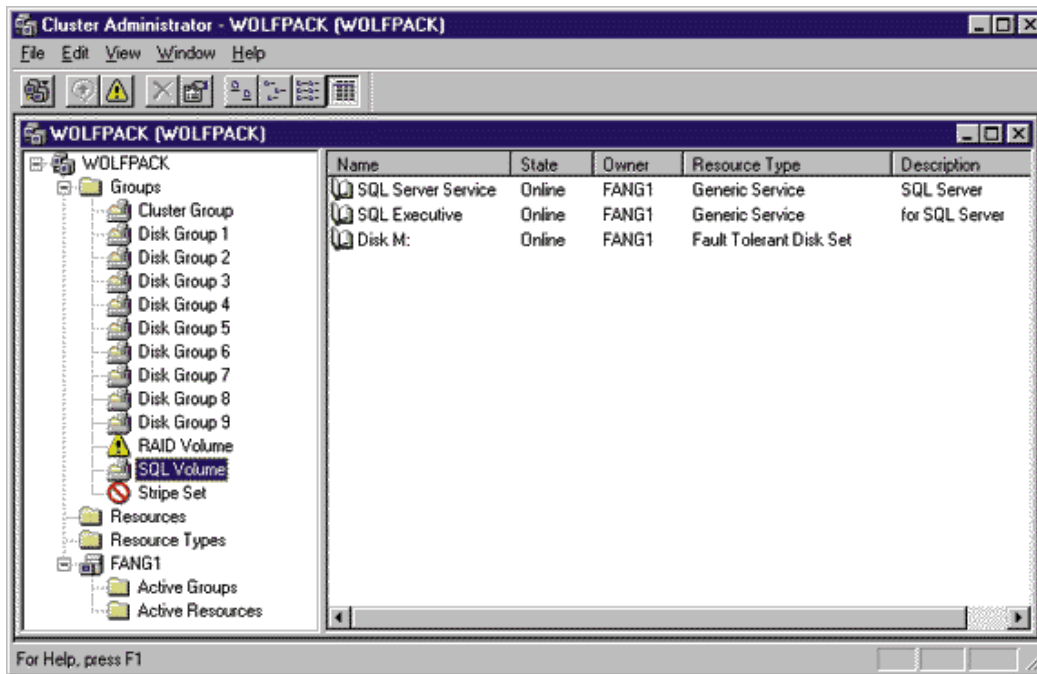
Les solutions intégrées dans SQL Server

- Cluster de basculement
- Groupes de disponibilité



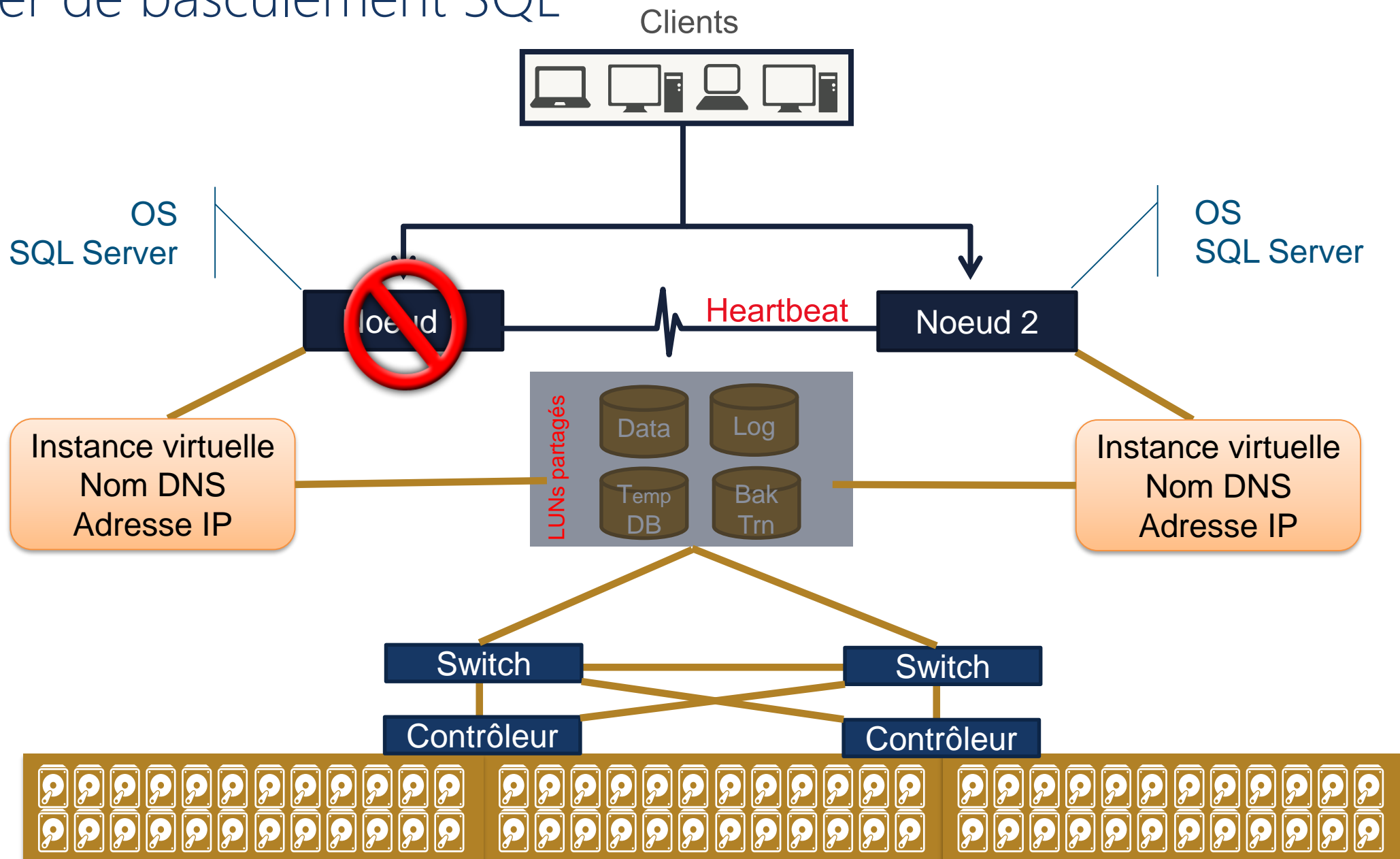
WolfPack ...

- 20 ans déjà
 - Windows NT 4 – Wolfpack
 - SQL Server 6.5
 - HCL réduite



- Cluster de basculement
 - Stockage partagé (historiquement)
 - SAN FC / FCoE
 - SAN iSCSI
 - Nœuds
 - Réseau privé
 - Réseau public
 - Services (Rôles) en HA
 - Serveur de fichiers
 - DHCP
 - SQL Server
 - Chaque rôle dispose de
 - Ressources propres
 - Point d'accès (@IP, Nom réseau)

Cluster de basculement SQL



Faílover Cluster Instance

PROS | CONS

- 
- | PROS | CONS |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none">• Granularité instance• Point d'accès réseau<ul style="list-style-type: none">• Nom DNS & Adresse IP• Tolérance à la panne<ul style="list-style-type: none">• Matérielle• Logicielle | <ul style="list-style-type: none">• Réputation de complexité• Stockage centralisé<ul style="list-style-type: none">• Point défaillance unique• Ne facilite pas le Géo Cluster• Performances discutables |

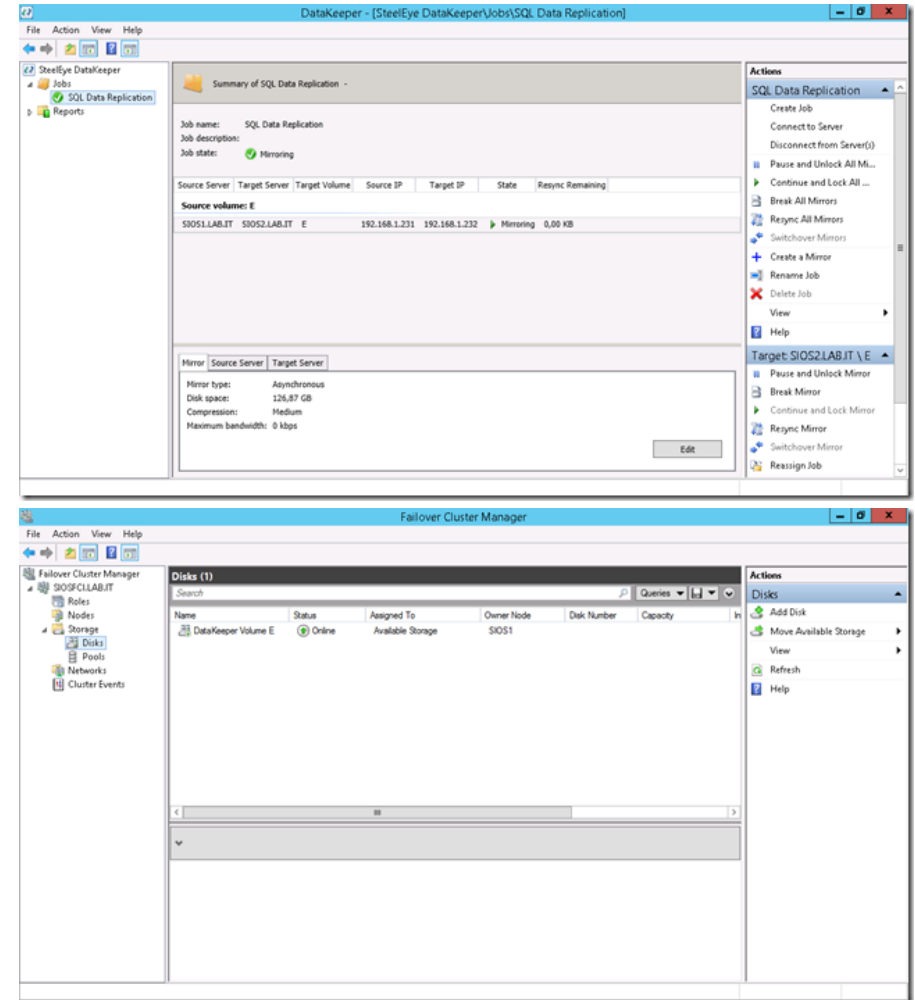
Pourquoi préférer le cluster de basculement ?

- AGs comblent les lacunes
- Mais
 - Modèle de récupération complet seulement
 - Seulement les bases utilisables
 - Agent SQL Server
 - Seulement les bases de données
 - MSDTC ?
 - Nombre de work
 - Nécessité d'un réplica en lecture seule ?
 - Edition entreprise (BAGs equivalent Database Mirroring)

FCI encore d'actualité malgré
le stockage centralisé !!!

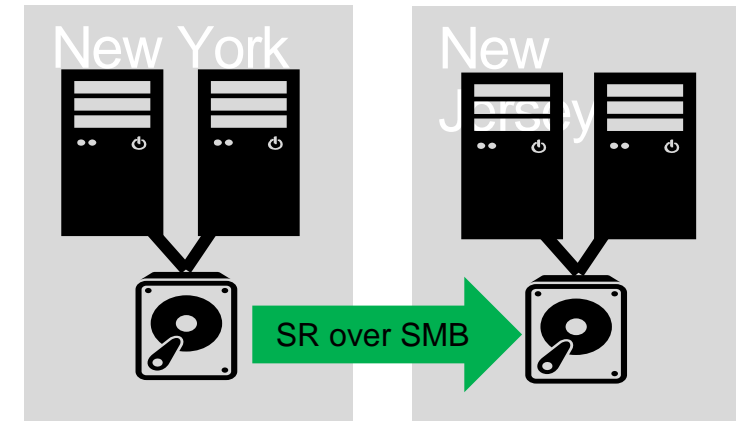
Les solutions pour contourner le stockage centralisé

- Utiliser un stockage local
 - + mécanisme de "replication"
- Solutions constructeur
 - San Mirroring
 - EMC, HP, Unisys, ...
- Editeurs tiers
 - DoubleTake
 - SIOS Datakeeper



Les solutions pour contourner le stockage centralisé

- Stretch Cluster
 - Storage Replica (Windows 2016)
- (Convergence)
 - Partage SMB3 sur SOFS
 - Virtualisation : Shared VHDX
- Hyper convergence
 - Storage Spaces Direct (Windows 2016)



Démo

- Ajout de disques locaux
- Ajout des fonctionnalités
- Test du cluster
- Création du cluster
- Ajout du File Share Witness
- Activation de Storage Spaces Direct

```
$VhdPath = "C:\Hyper-V\Virtual Hard Disks"
$VMNames="SRV1","SRV2"
foreach ($VmName in $VMNames) {
    Get-VM -Name $VmName | Stop-VM -Force
    Add-VMScsiController -VMName $VmName

    for ($i=1; $i -le 4; $i++){
        New-VHD -Path $VhdPath"\$VmName-Data$i.vhdx" -SizeBytes 10GB
        Add-VMHardDiskDrive -VMName $VmName -Path $VhdPath"\$VmName-Data$i.vhdx" -ControllerType SCSI -ControllerNumber 1
        Write-Host "Disk $VhdPath\$VmName-Data$i.vhdx Added to $VmName"
    }
    Get-VM -Name $VmName | Start-VM
}
```

Add Features

```
Install-WindowsFeature -Name Failover-Clustering,File-Services -IncludeAllSubFeature -IncludeManagementTools -ComputerName srv1
Install-WindowsFeature -Name Failover-Clustering,File-Services -IncludeAllSubFeature -IncludeManagementTools -ComputerName srv2
```

Test cluster, else SQL Server won't install

```
Test-Cluster -Node srv1,srv2 -Include "Storage Spaces Direct", "Inventory", "Network", "System Configuration"
```

Create the Cluster and add a File Share Witness

```
New-Cluster -Name ClustS2D -Node srv1,srv2 -NoStorage -StaticAddress 10.0.0.10
Start-Sleep -s 2
Get-Cluster | Set-ClusterQuorum -FileShareWitness "\\AD\FSW"
```

enable Storage Spaces Direct

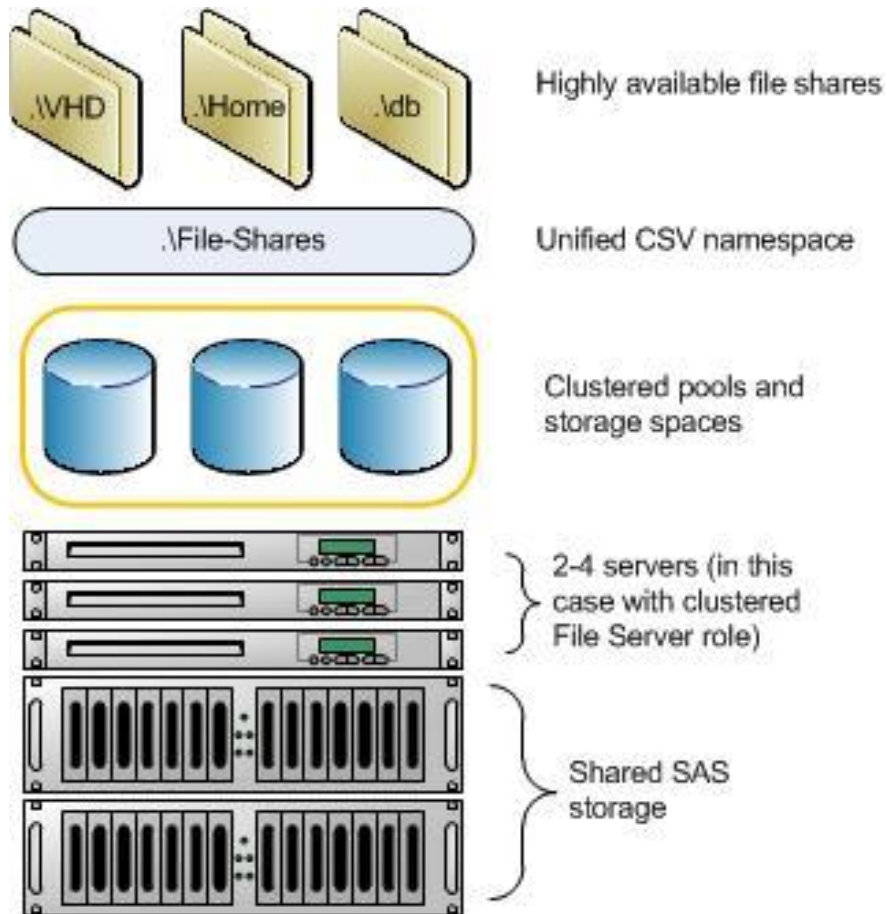
```
Enable-ClusterS2D
```

Storage spaces



- Storage pool
 - Collection de disques
 - Combiner HDD et SSD
- Storage spaces
 - Solution de stockage hautement disponible et flexible
 - Fournit la "tolérance de panne" aux "disques" virtuels, créés à partir d'espace libre du storage pool
- Volume
 - Espace de stockage où l'on dépose des fichiers
 - Visible dans c:\ClusterStorage

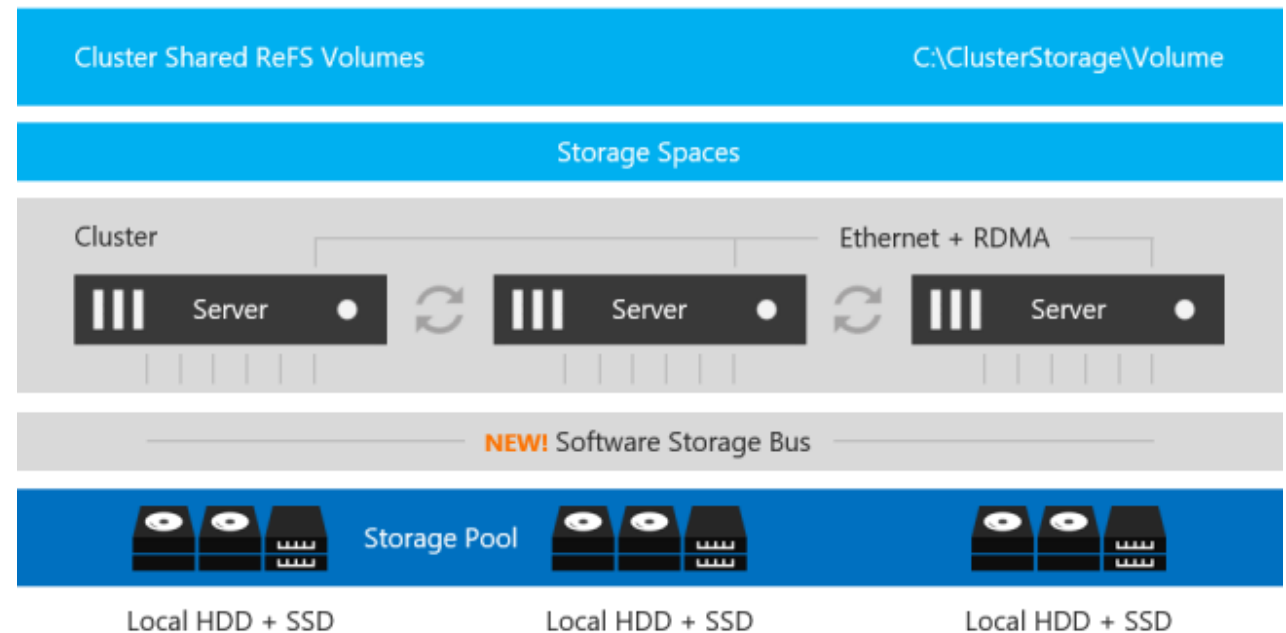
Clustered Storage Spaces



- Combinaison des fonctionnalités
 - Cluster de basculement
 - Storage spaces
 - => Tous les volumes sont accessibles depuis tous les nœuds à tout moment
- Prérequis
 - SAS ou JBOD attaché aux nœuds du cluster
- Scénarii
 - Cluster de fichier hautement disponible (SOFS)
 - Pas de coupure en cas de bascule (SMB3)

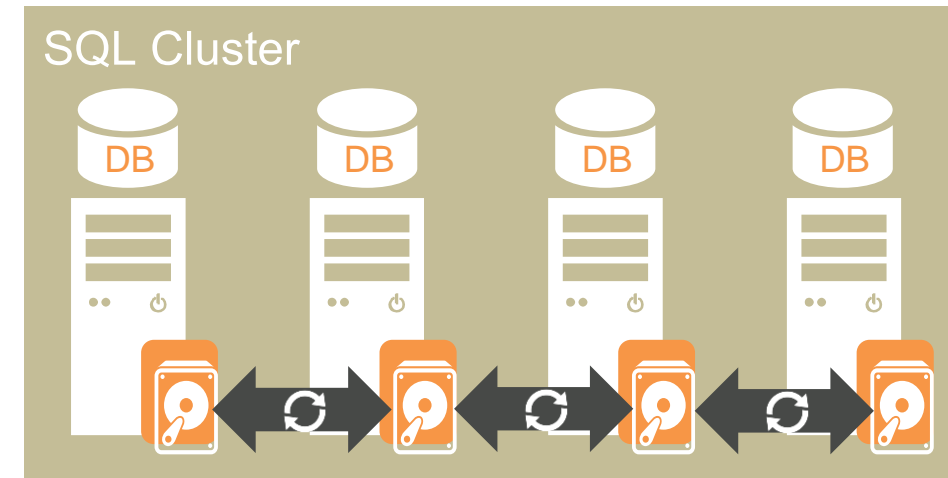
Windows 2016 - Storage spaces direct

- Approche alternative au problème du stockage centralisé
 - Première solution réelle de type Software-Defined Storage
- Système hyper convergé
 - Nœuds multi rôles
 - Stockage
 - Calcul
 - Acteurs du marché
 - Nutanix
 - VMware Virtual San
 - Simplivity (HPE)
- Idée maitresse
 - Eliminer le stockage centralisé
 - Totalement indépendant du matériel
 - SAS limite le Scale-Out par le coût
 - SAS n'est pas adapté aux grandes distances
 - Utilisation de matériel "standard" voire peu onéreux



Windows S2D & SQL Server

- Cluster sans baie de disque ...
 - Baie de disque virtuelle
 - Stockage local
 - Répliqué sur tous les noeuds
 - Volume(s) disponible dans le cluster
- Fiabilité
 - Pas de point de défaillance unique
- $2 \leq \text{Serveurs} \leq 16$
- CPU / RAM
- Réseau inter nœuds rapide
 - Ethernet 10 Gb+ + RDMA
 - SMB3



Azure :

- S2D supporté
- SIOS DataKeeper

Démo

- Création d'un volume
- Installation SQL Server

The first screenshot shows the 'Specify the size of the virtual disk' step. The 'Specify size' option is selected, with a size of 20 GB. The 'Resiliency: Two-Way Mirror' is highlighted with a red box. The second screenshot shows the 'Confirm selections' step, where the virtual disk properties are confirmed: Name: SQLServer, Storage tiers: Enabled, Capacity Tier - resiliency: Two-Way Mirror, Capacity Tier - media type: HDD, Capacity Tier - size: 20,0 GB, Total requested size: 20,0 GB.

```
# 2 way mirroring
New-Volume -StoragePoolFriendlyName 'S2D on ClustS2D' `
-FriendlyName 2wmVolume `
-ResiliencySettingName 'Mirror' `
-PhysicalDiskRedundancy 1 `
-FileSystem CSVFS_REFS `
-Size 20GB
```

```
# attach iso to VM
Set-VMdvdDrive -VMName SRV1 -Path $SQLServerISO
Set-VMdvdDrive -VMName SRV2 -Path $SQLServerISO
```

```
Invoke-Command -Credential $Cred -VMName SRV1 {
    . ([ScriptBlock]::Create($Using:FunctionDefs))

    # Create a new virtual instance
    CreateSQLCluster
}
```

```
Invoke-Command -Credential $Cred -VMName SRV2 {
    . ([ScriptBlock]::Create($Using:FunctionDefs))

    # Add a second node
    AddNodesQLCluster
}
```

The screenshot shows the 'Cluster Disk Selection' step of the 'Install a SQL Server Failover Cluster' wizard. It lists various configuration options on the left, including 'Cluster Disk Selection'. On the right, it shows the 'Available shared disks' table with one entry: 'Cluster Virtual Disk (SQLServer)' with a green checkmark in the 'Qualified' column.

Qualified	Disk	Message
✓	Cluster Virtual Disk (SQLServer)	

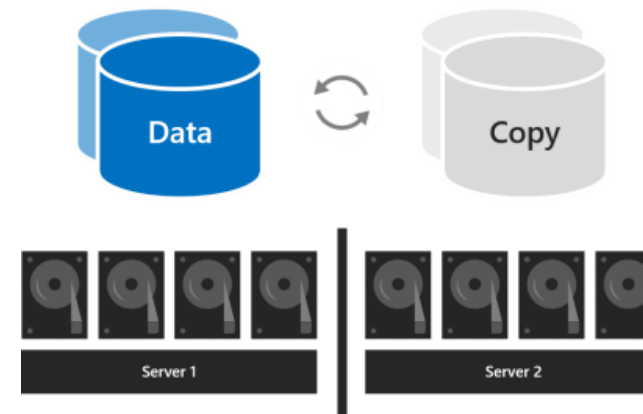
Résilience – Miroir 2 nœuds



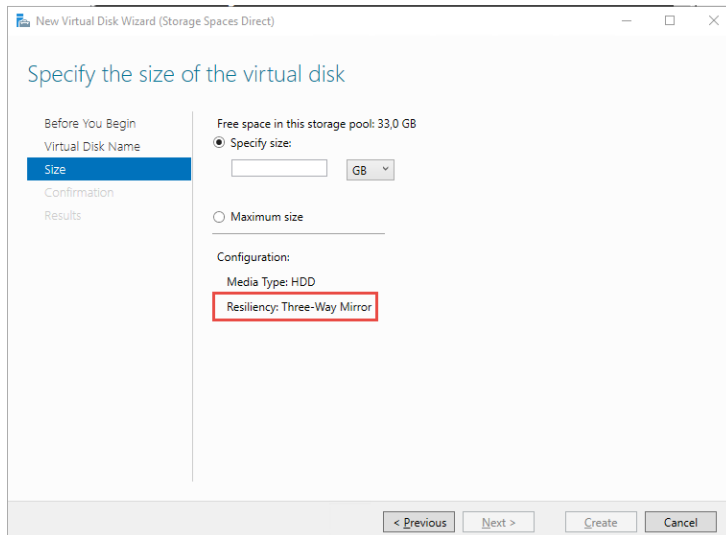
2 way mirroring

```
New-Volume -StoragePoolFriendlyName 'S2D on ClustS2D' `
-FriendlyName 2wmVolume `
-ResiliencySettingName 'Mirror' `
-PhysicalDiskRedundancy 1 `
-FileSystem CSVFS_REFS `
-Size 20GB
```

- 2 way mirroring
 - Equivalent Raid 1
 - Nombre de nœuds : 2
 - Défaillances possibles : 1
 - Efficacité de stockage : 50%



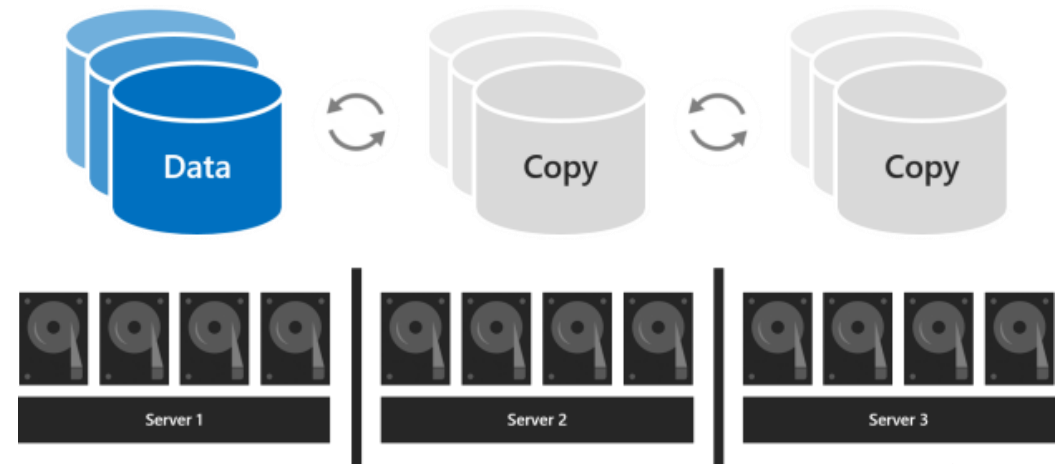
Résilience – Miroir 3 nœuds



3 way mirroring

```
New-Volume -StoragePoolFriendlyName 'S2D on ClustS2D' -  
-FriendlyName 3wmVolume -ResiliencySettingName 'Mirror' -  
-PhysicalDiskRedundancy 2 -FileSystem CSVFS_REFS -Size 20GB
```

- 3 way mirroring
 - Nombre de nœuds : 3
 - Défaillances possibles : 2
 - Efficacité de stockage : 33%



Résilience – Simple parité

Pas
d'assistant !

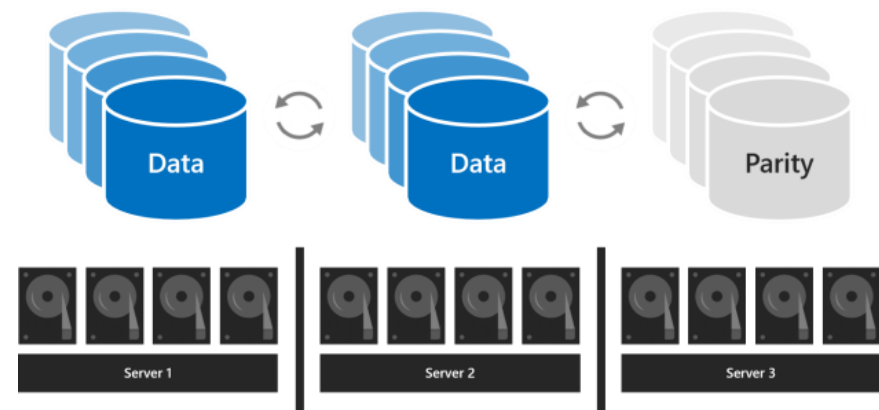
Intérêt de la
solution ?

```
# Single Parity
New-Volume -StoragePoolFriendlyName 'S2D on ClustS2D' `
-FriendlyName spVolume `
-ResiliencySettingName 'Parity' `
-PhysicalDiskRedundancy 1 `
-FileSystem CSVFS_REFS `
-Size 20GB
```

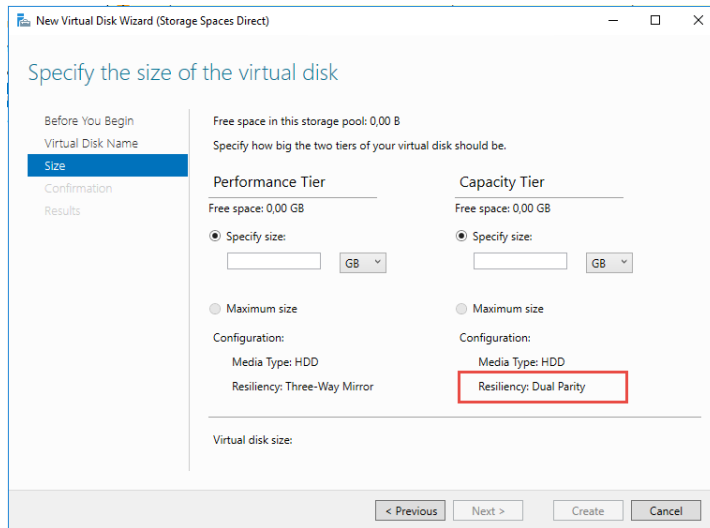
⚠ Warning

We discourage using single parity because it can only safely tolerate one hardware failure at a time: if you're rebooting one server when suddenly another drive or server fails, you will experience downtime. If you only have three servers, we recommend using three-way mirroring. If you have four or more, see the next section.

- Single parity
 - Equivalent Raid 5
 - Nombre de nœuds : 3
 - Défaillances possibles : 1
 - Efficacité de stockage : $\geq 66\%$ $\leq 87\%$



Résilience – Double parité

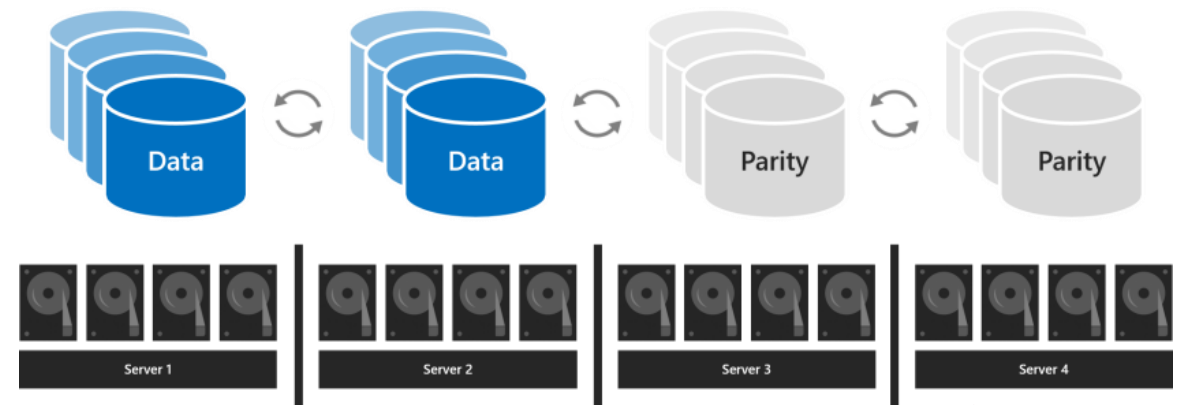


■ Dual parity

- Equivalent Raid 6
- Nombre de nœuds : ≥ 4
- Défaillances possibles : 2
- Efficacité de stockage : $\geq 50\%$ $\leq 80\%$

Dual parity

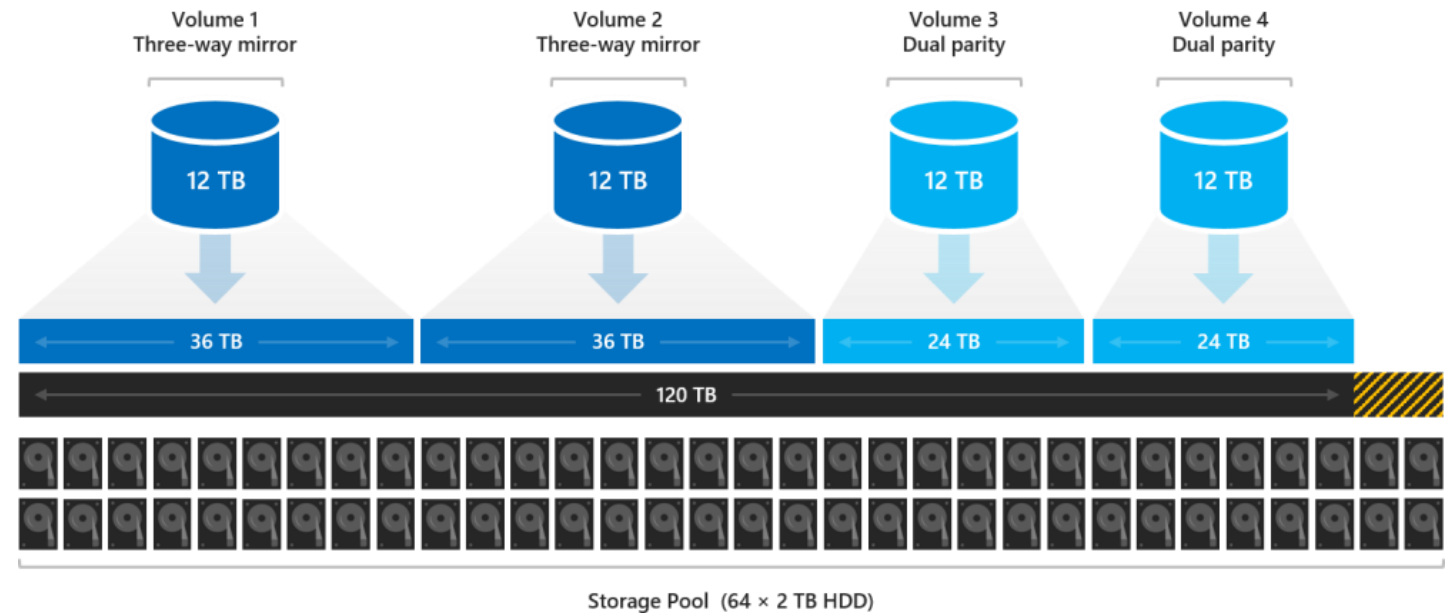
```
New-Volume -StoragePoolFriendlyName 'S2D on ClustS2D'`
            -FriendlyName dpVolume`
            -ResiliencySettingName 'Parity'`
            -PhysicalDiskRedundancy 2`
            -FileSystem CSVFS_REFS`
            -Size 20GB`
```



Gestion de la capacité

(2x) (1) (50%)	Two-way mirroring
(3x) (2) (33%)	Three-way mirroring
(3x) (1) (66.7-87.5%)	Single parity
(4x) (2) (50-80%)	Dual parity
(4x) (2) (33.3-80%)	Mixed

Two-way mirroring	Performance Tier
Three-way mirroring	
Single parity	Capacity Tier
Dual parity	
Mixed	



Baie de disques- Performances

Operation ▼	Duration ▼	IOSize ▼	IOType ▼	PendingIO ▼	FileSize ▼	IOPS VNX7500 ▼	MBs/Sec ▼	Min_Lat(ms) ▼	Avg_Lat(ms) ▼	Max_Lat(ms) ▼
Read	30	8	Random	8	16000	28 578,81	223,27	0	1	170
Write	30	8	Random	8	16000	17 697,72	138,26	0	3	579
Read	30	8	Sequential	8	16000	16 509,39	128,97	0	3	123
Write	30	8	Sequential	8	16000	26 184,05	204,56	0	1	198
Read	30	32	Random	8	16000	23 369,09	730,28	0	2	44
Write	30	32	Random	8	16000	7 250,18	226,56	0	8	499
Read	30	32	Sequential	8	16000	18 001,49	562,54	0	3	40
Write	30	32	Sequential	8	16000	7 156,32	223,63	0	8	84
Read	30	64	Random	8	16000	10 570,81	660,67	0	5	156
Write	30	64	Random	8	16000	5 294,93	330,93	0	11	108
Read	30	64	Sequential	8	16000	10 215,73	638,48	0	5	27
Write	30	64	Sequential	8	16000	3 511,60	219,47	1	17	41
Read	30	128	Random	8	16000	6 047,07	755,88	0	10	171
Write	30	128	Random	8	16000	2 882,85	360,35	0	21	119
Read	30	128	Sequential	8	16000	5 113,71	639,21	0	12	47
Write	30	128	Sequential	8	16000	1 762,32	220,29	2	35	82
Read	30	256	Random	8	16000	3 040,23	760,05	0	20	182
Write	30	256	Random	8	16000	1 478,31	369,57	0	42	152
Read	30	256	Sequential	8	16000	2 535,46	633,86	0	24	125
Write	30	256	Sequential	8	16000	880,84	220,21	8	72	143

Performances & capacité : mix de résiliences

- Pour un seul et même volume
 - Miroir : un cache performant
 - Parité : stockage capacitif

Two-way mirroring	Performance Tier
Three-way mirroring	
Single parity	Capacity Tier
Dual parity	
Mixed	

New Virtual Disk Wizard (Storage Spaces Direct)

Confirm selections

Before You Begin
Virtual Disk Name
Size
Confirmation
Results

Confirm that the following are the correct settings, and then click Create.

VIRTUAL DISK LOCATION

Server: Win16S2D1
Subsystem: Clustered Windows Storage
Storage pool name: S2D on ClustS2D4
Status: OK
Free space: 95,5 GB

VIRTUAL DISK PROPERTIES

Name: SQLServerVolume
Storage tiers: Enabled

Performance Tier - resiliency: Three-Way Mirror
Performance Tier - media type: HDD
Performance Tier - size: 2,00 GB
Capacity Tier - resiliency: Dual Parity
Capacity Tier - media type: HDD
Capacity Tier - size: 10,0 GB
Total requested size: 12,0 GB

< Previous Next > Create Cancel

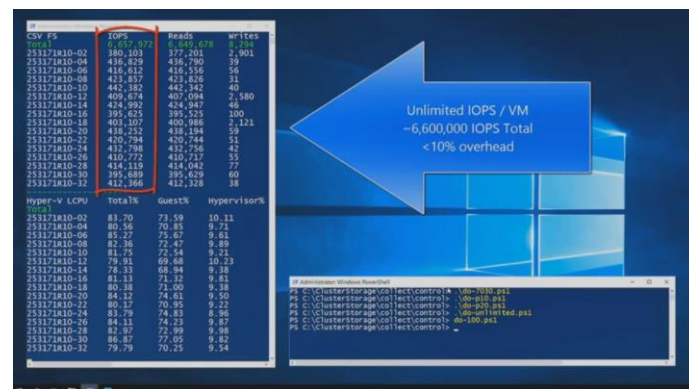
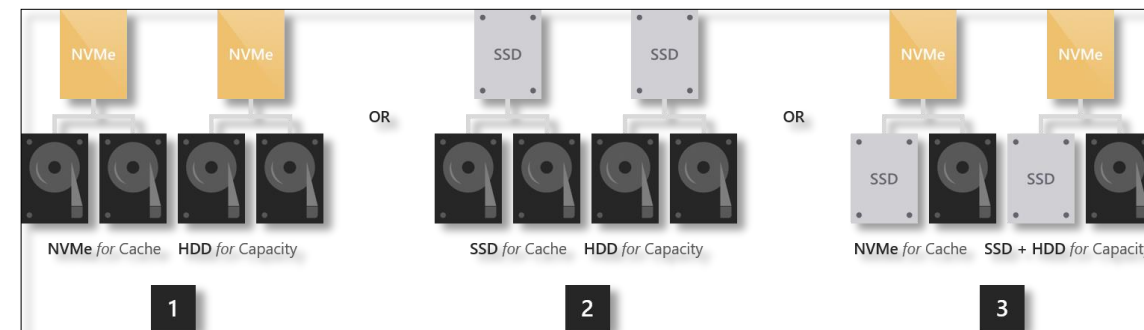
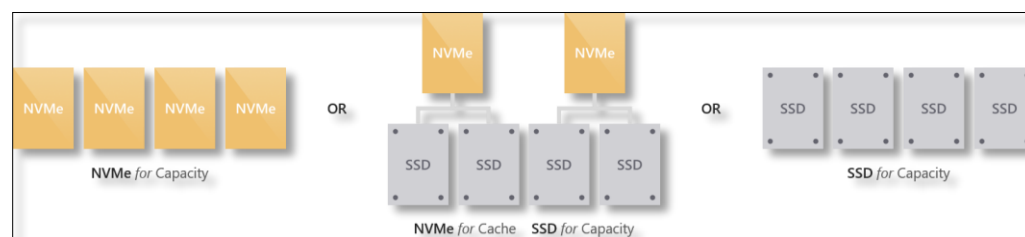
Performances & capacité : mix de résiliences

- Maximiser la performance

- Uniquement des disques Flash
- Charge de travail intensive
 - 1Tb/sec – 6 000 000 IOPS

- Balance entre capacité et performance

- Possibilité d'utiliser du cache
 - 2 disques par serveur
 - Ratio de 1:2 disque capacité



- 160 GB /Sec
- EN Wikipedia : 11,5 GB
- => 14 transferts /sec

CSV FS	IOPS	Reads	Writes	BW (MB/s)	Read	Write	Read Lat (ms)	Write Lat
total	306,427	306,402	25	160,460	160,460	1		
ignite1	26,300	26,299	2	13,605	13,605		4.123	1.467
ignite10	24,572	24,570	1	12,882	12,882		4.381	0.744
ignite11	26,584	26,583	1	13,937	13,937		4.063	0.682
ignite12	27,785	27,782	2	14,566	14,566		3.896	0.906
ignite2	23,950	23,946	4	12,555	12,555		4.528	0.640
ignite3	24,876	24,873	3	13,041	13,041		4.365	0.624
ignite4	26,403	26,399	4	13,841	13,841		4.102	0.742
ignite5	24,468	24,464	4	12,826	12,826		4.419	0.616
ignite6	25,341	25,340	1	13,286	13,286		4.283	1.127
ignite7	23,902	23,900	2	12,531	12,531		4.579	1.110
ignite8	25,853	25,852	1	13,554	13,554		4.197	0.936
ignite9	26,393	26,393	1	13,837	13,837		4.140	0.862

La fin des disques rotatifs ?

- 1956
 - IBM
 - “Super” computer 305 RAMAC
 - HDD 5 MB
 - > 1 000 kg
 - > 50 000 \$
- Actuellement
 - SSD Samsung : 16 TB
 - HDD Seagate : 10 TB



Disques SSD

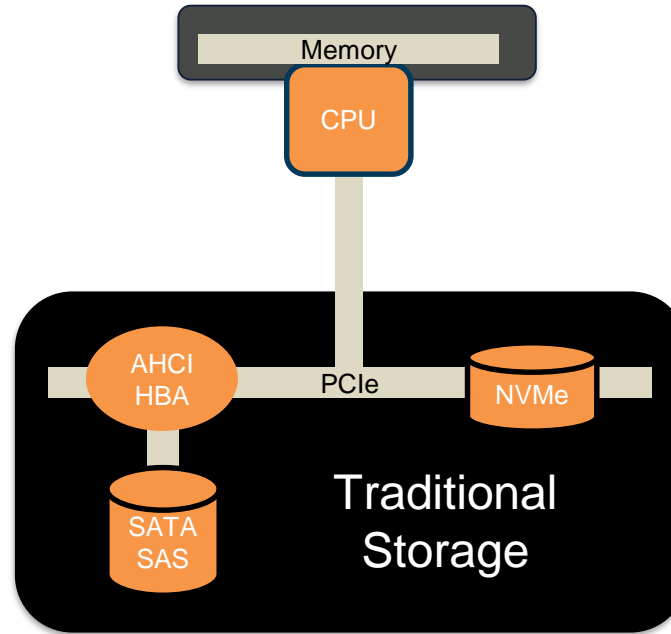
- Interface

- SAS
- SATA
- PCI
- NVMe

- Repères

- Ram : $6 \text{ ns} = 6 \times 10^{-9} \text{ sec}$
- CPU à 3,5 GHz : 10^{-9} sec
- HDD rotatif : $7 \text{ ms} = 7 \times 10^{-3} \text{ sec}$
 - 1 000 000 de fois plus lent !!!!
 - 1 IO prends autant de temps que 1 000 000 cycle CPU
- SSD : $50 \mu\text{s} = 10^{-6} \text{ sec}$
 - 1 000 fois plus lent que RAM
 - 1 000 fois plus rapide que HDD rotatif ...
- \approx escargot (0,0275 m/s) et guépard (28 m/s)

[http://fr.wikipedia.org/wiki/Ordre_de_grandeur_\(vitesse\)](http://fr.wikipedia.org/wiki/Ordre_de_grandeur_(vitesse))



Performance DAS : SSD SAS

Operation	Duration	IOSize	IOType	PendingIO	FileSize	IOPS RAID SSD	MBs/Sec	Min_Lat(ms)	Avg_Lat(ms)	Max_Lat(ms)
Read	10	8	Random	8	20000	161 955,10	1 265,27	0	0	47
Write	10	8	Random	8	20000	61 224,10	478,31	0	1	36
Read	10	8	Sequential	8	20000	168 954,60	1 319,95	0	0	26
Write	10	8	Sequential	8	20000	101 239,40	790,93	0	0	25
Read	10	32	Random	8	20000	65 572,70	2 049,14	0	1	4
Write	10	32	Random	8	20000	29 222,20	913,19	0	3	89
Read	10	32	Sequential	8	20000	67 405,39	2 106,41	0	1	10
Write	10	32	Sequential	8	20000	21 447,70	670,24	0	5	94
Read	10	64	Random	8	20000	35 722,19	2 232,63	1	3	24
Write	10	64	Random	8	20000	14 956,80	934,80	0	8	150
Read	10	64	Sequential	8	20000	35 990,60	2 249,41	2	3	24
Write	10	64	Sequential	8	20000	16 425,11	1 026,56	0	7	163
Read	10	128	Random	8	20000	18 734,99	2 341,87	3	6	35
Write	10	128	Random	8	20000	8 152,81	1 019,10	0	15	330
Read	10	128	Sequential	8	20000	19 285,48	2 410,68	1	6	18
Write	10	128	Sequential	8	20000	8 745,72	1 093,21	0	14	326
Read	10	256	Random	8	20000	9 989,40	2 497,35	2	12	21
Write	10	256	Random	8	20000	4 381,64	1 095,41	1	28	610
Read	10	256	Sequential	8	20000	8 394,52	2 098,63	2	14	38
Write	10	256	Sequential	8	20000	4 183,70	1 045,92	1	30	599

Performance DAS : SSD PCI

Operation ▼	Duration ▼	IOSize ▼	IOType ▼	PendingIO ▼	FileSize ▼	IOPS Fusion IO ▼	MBs/Sec ▼	Min_Lat(ms) ▼	Avg_Lat(ms) ▼	Max_Lat(ms) ▼
Read	60	8	Random	8	20000	117 814,33	920,42	0	0	33
Write	60	8	Random	8	20000	120 789,28	943,66	0	0	4
Read	60	8	Sequential	8	20000	166 328,68	1 299,44	0	0	38
Write	60	8	Sequential	8	20000	120 665,68	942,70	0	0	1
Read	60	32	Random	8	20000	43 461,68	1 358,17	0	2	44
Write	60	32	Random	8	20000	30 691,76	959,11	0	3	5
Read	60	32	Sequential	8	20000	44 263,06	1 383,22	0	2	29
Write	60	32	Sequential	8	20000	30 348,49	948,39	0	3	6
Read	60	64	Random	8	20000	22 055,58	1 378,47	0	5	30
Write	60	64	Random	8	20000	15 181,68	948,85	0	7	10
Read	60	64	Sequential	8	20000	22 253,76	1 390,86	1	5	31
Write	60	64	Sequential	8	20000	15 268,36	954,27	0	8	19
Read	60	128	Random	8	20000	11 114,45	1 389,30	1	11	65
Write	60	128	Random	8	20000	7 571,43	946,42	0	16	18
Read	60	128	Sequential	8	20000	11 158,85	1 394,85	0	10	82
Write	60	128	Sequential	8	20000	7 519,90	939,98	1	16	26
Read	60	256	Random	8	20000	5 568,75	1 392,18	0	22	114
Write	60	256	Random	8	20000	3 828,96	957,24	1	32	35
Read	60	256	Sequential	8	20000	5 579,71	1 394,92	0	22	110
Write	60	256	Sequential	8	20000	3 811,18	952,79	0	33	44

Performance DAS : SSD PCI NVMe

Operation	Duration	IOSize	IOType	PendingIO	FileSize	IOPS P3700	MBs/Sec	Min_Lat(ms)	Avg_Lat(ms)	Max_Lat(ms)
Read	30	8	Random	8	5000	293 653,32	2 294,16	0	0	3
Write	30	8	Random	8	5000	131 564,20	1 027,84	0	0	9
Read	30	8	Sequential	8	5000	306 275,90	2 392,78	0	0	3
Write	30	8	Sequential	8	5000	134 100,13	1 047,65	0	0	8
Read	30	32	Random	8	5000	72 990,33	2 280,94	0	1	8
Write	30	32	Random	8	5000	33 383,44	1 043,23	0	3	12
Read	30	32	Sequential	8	5000	81 879,53	2 558,73	0	1	10
Write	30	32	Sequential	8	5000	33 786,90	1 055,84	0	3	15
Read	30	64	Random	8	5000	35 157,48	2 197,34	0	3	10
Write	30	64	Random	8	5000	16 764,46	1 047,77	0	7	23
Read	30	64	Sequential	8	5000	41 034,10	2 564,63	0	2	10
Write	30	64	Sequential	8	5000	16 919,96	1 057,49	0	7	22
Read	30	128	Random	8	5000	17 898,67	2 237,33	0	6	20
Write	30	128	Random	8	5000	8 459,16	1 057,39	0	14	36
Read	30	128	Sequential	8	5000	20 528,22	2 566,02	0	5	15
Write	30	128	Sequential	8	5000	8 454,67	1 056,83	0	14	37
Read	30	256	Random	8	5000	9 006,46	2 251,61	1	13	24
Write	30	256	Random	8	5000	4 248,64	1 062,16	0	29	52
Read	30	256	Sequential	8	5000	9 765,34	2 441,33	1	12	24
Write	30	256	Sequential	8	5000	4 220,13	1 055,03	0	29	51

Coûts en baisse ...

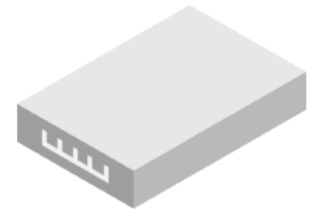
- Disques
 - Sous système le plus lent
 - Disques lent = système peu optimisé
 - => Impossible de rentabiliser les licences
- Privilégier les technologies flash
- **Toujours** choisir en fonction de la performance et non du volume
 - IOPS vs GB



Intel P3700
800GB NVMe SSD

\$1,999

\$6.16 / 1K IOPS



Intel S3700
800GB SATA SSD

\$1,425

\$14.90 / 1K IOPS

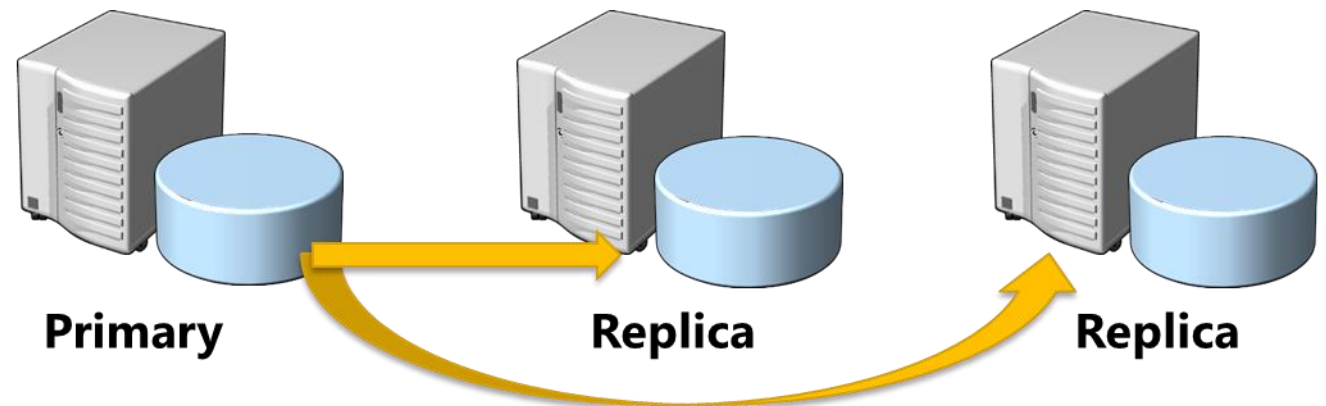
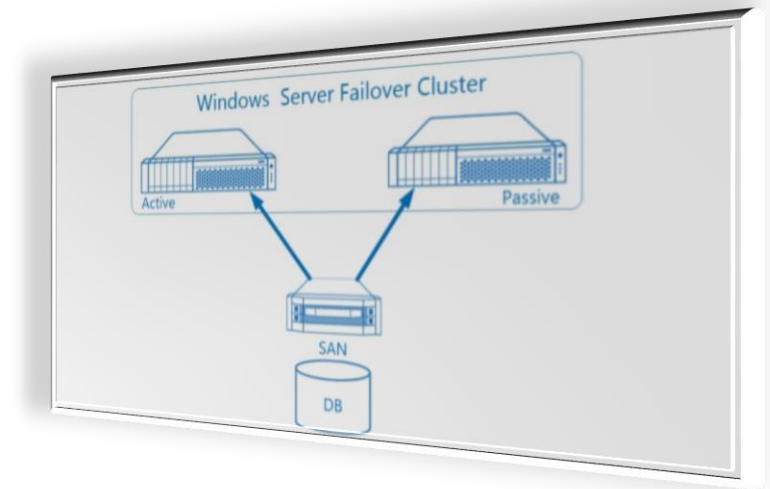
Vs

Avantages du S2D ?

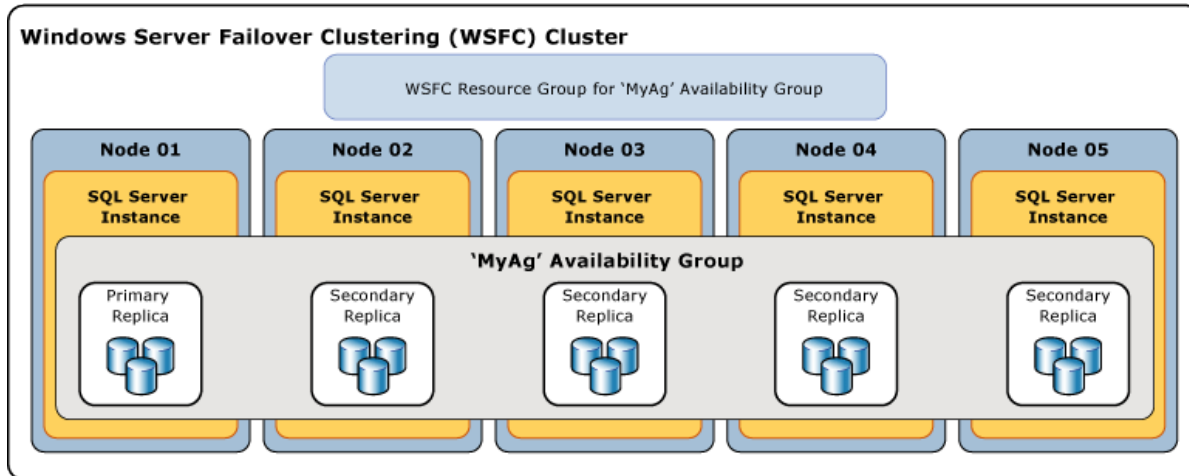
- Forte tolérance à la panne
- Simplification du Géo Clustrer
- Bascule du cluster plus rapide qu'un SAN
- Performance disque : élimine le problème du "voisin bruyant"
- Migration facilities
- Pas de nécessité de lettre de volume
 - Pas de dépendance sur une ressource de type disque
 - Plus de limite à 25 instances virtuelles
 - Plusieurs instances sur un seul volume

Les solutions intégrées dans SQL Server

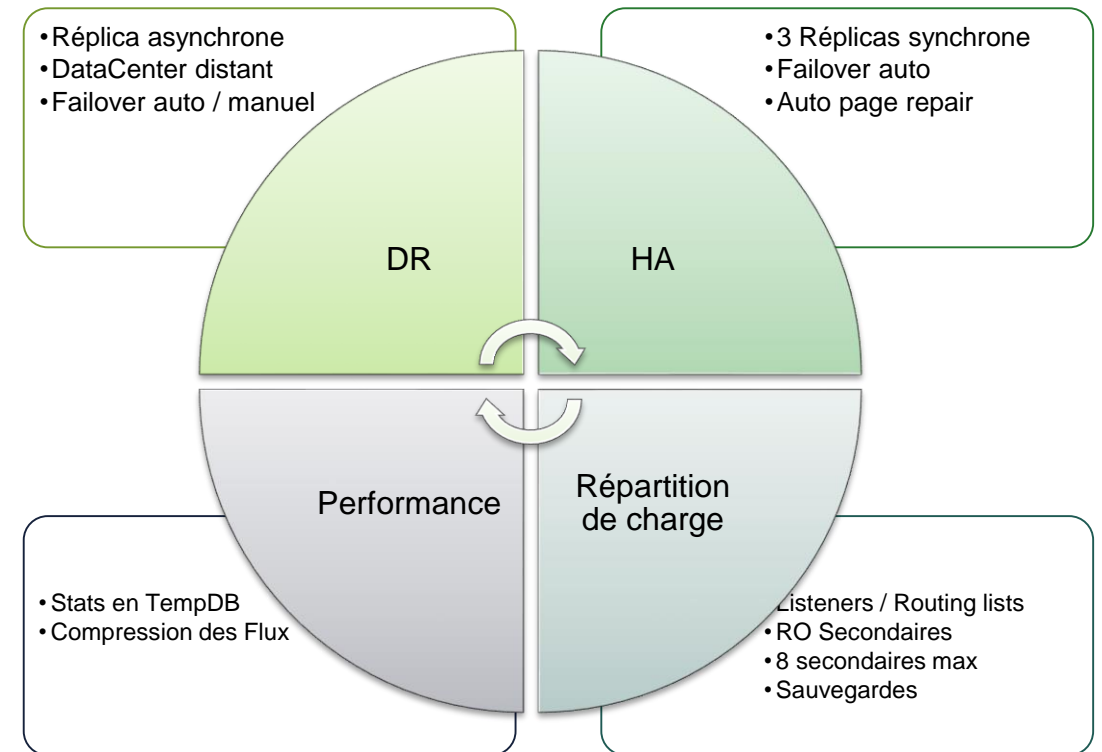
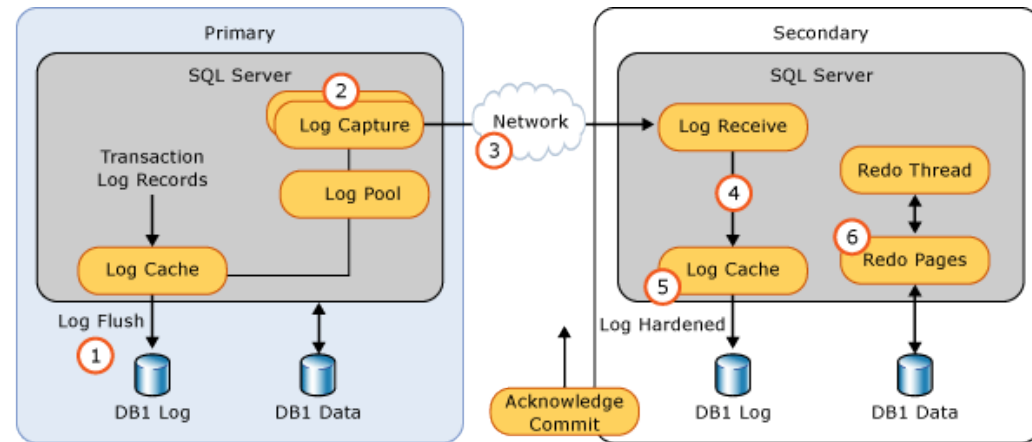
- Cluster de basculement
- Groupes de disponibilité



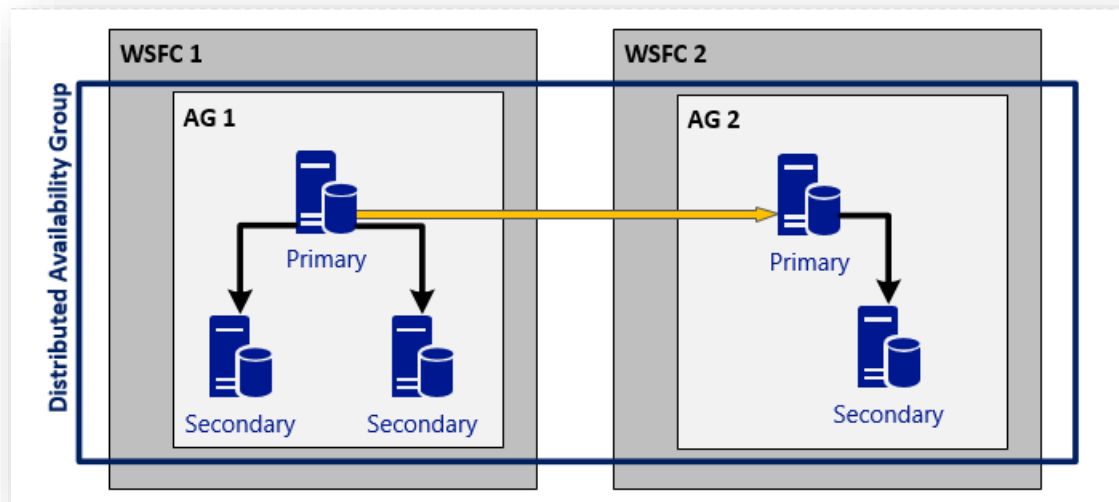
Groupes de disponibilité



- Disponible depuis SQL Server 2012 Entreprise
- Couvre les besoins en HA et DR
- Potentiellement indépendant d'un domaine AD
- Disponible en édition standard (restrictions)

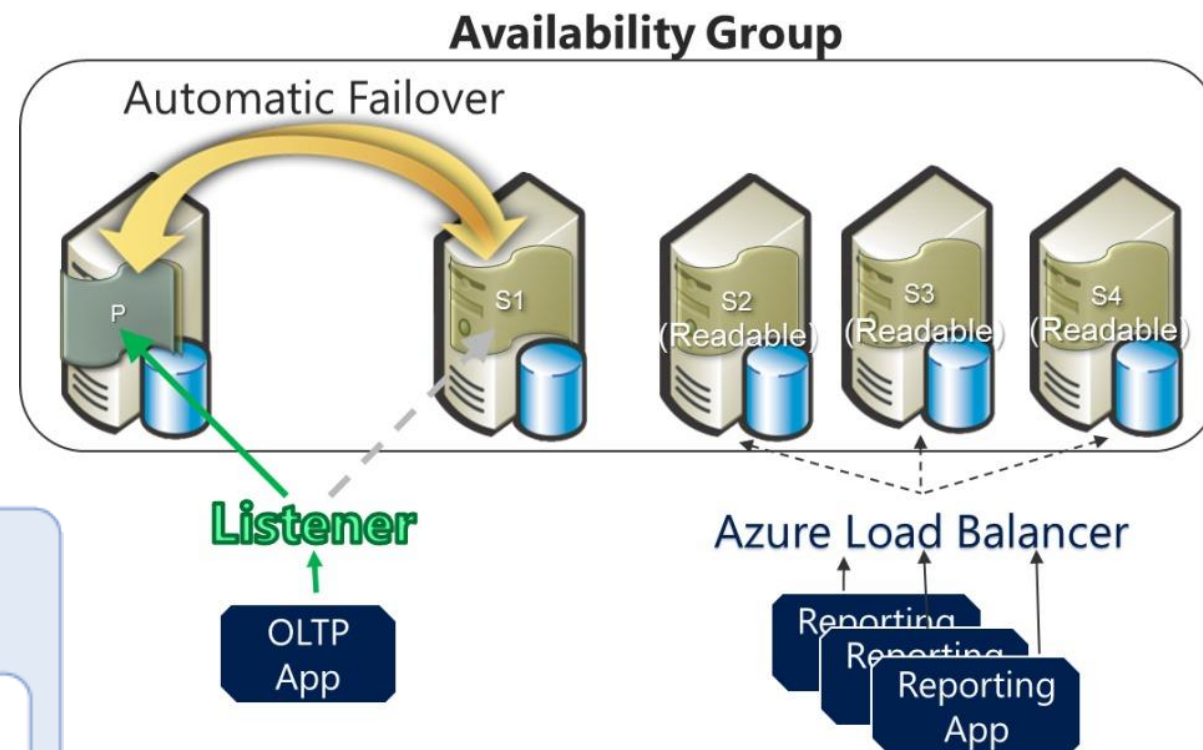
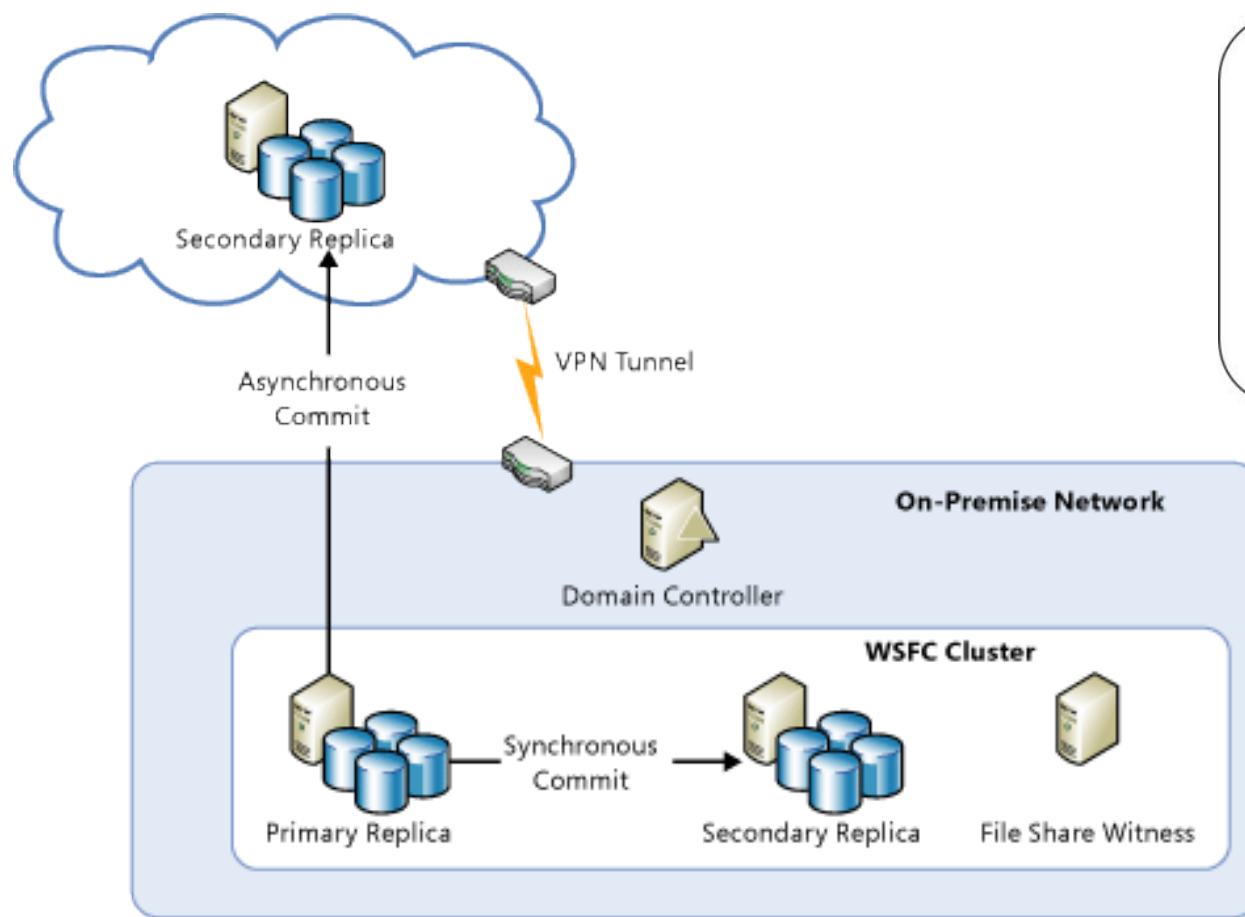


Groupes de disponibilité distribués



- Disponible depuis SQL Server 2016 Entreprise
- Mise en oeuvre
 - Création de l'AG1
 - Création de l'AG2
 - Création d'un "super" groupe de disponibilité
 - Dont les bases sont des AG
- Tire profit de l'amorçage direct (Direct Seeding)
- Chaque cluster possède
 - Son propre quorum
 - Sa propre configuration de vote
- Transferts réseau efficace
 - Un seul envoi des données sur le réseau
- Scénario de migration
 - Chaque cluster peut avoir des versions d'OS différentes
- Seul le basculement manuel est supporté

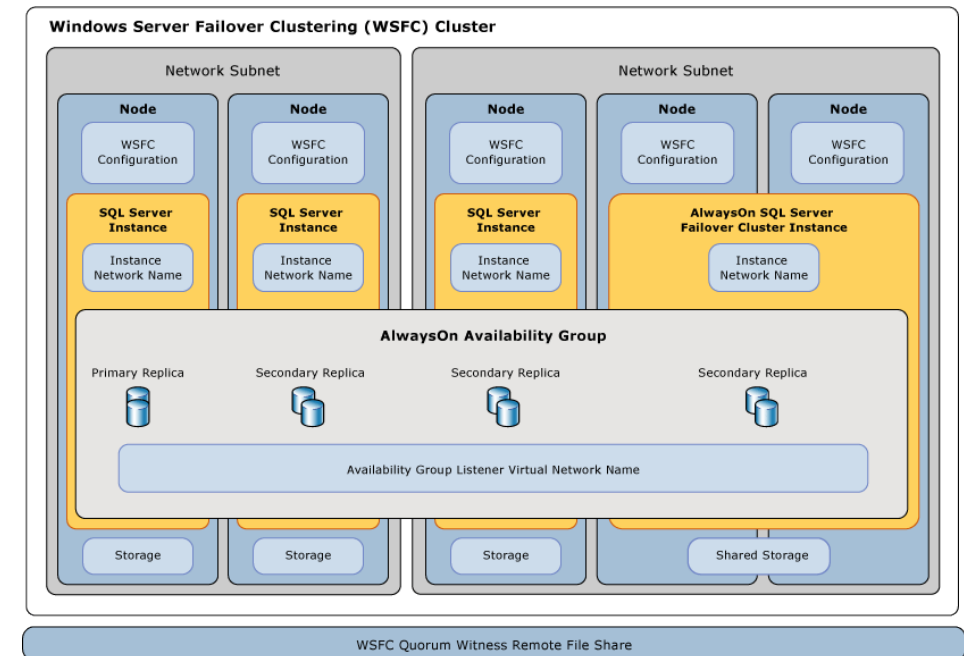
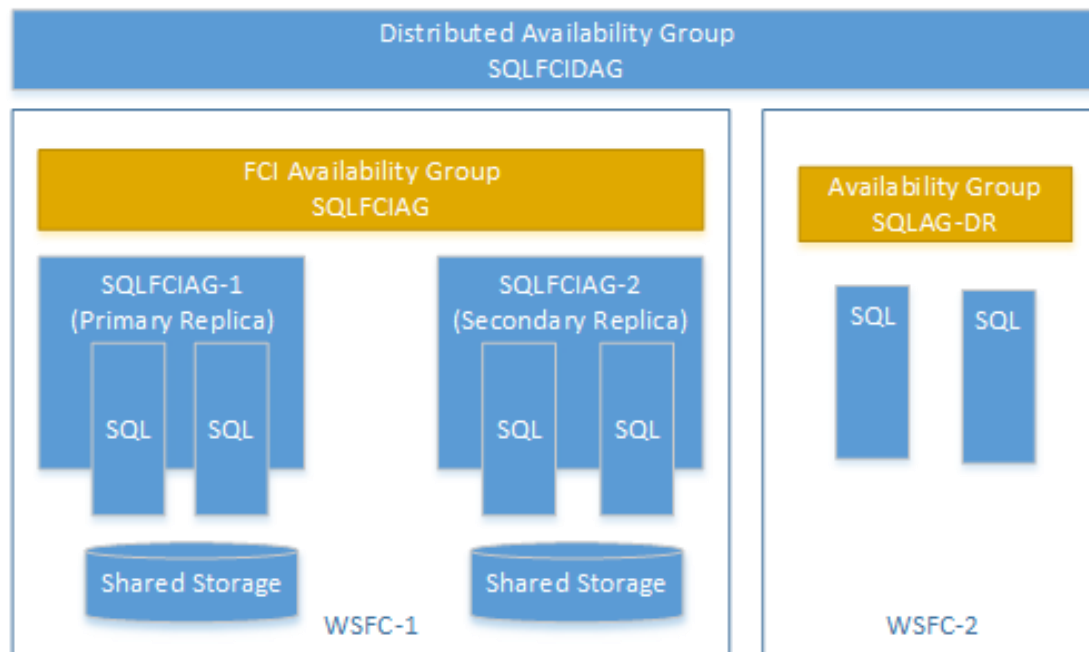
Azure



- Réplica asynchrone
- Scénario disaster recovery
- Scénario reporting / sites internet
- Assistant intégré à SSMS

Mixer les solutions

- Un groupe de disponibilité peut s'étendre
 - Sur un ou plusieurs FCI
 - Et sur une ou plusieurs instance SQL



Conclusion

- Différentes techniques de haute disponibilité
 - Peuvent être combinées
 - Evolution constante
- Points de défaillance unique
 - Tendent à disparaître pour SQL
 - Ne pas négliger l'application
- Ne remplace pas une stratégie
 - De sauvegarde
 - D'archivage

Windows 2012 SQL Server 2012

- WSFC
 - Quorum dynamique
 - Multi subnet failover cluster
- FCI
 - TempDB locale
 - Data sur share SMB3
- Groupes de disponibilité
 - Edition entreprise
 - 1 + 4 replicas
 - 2 replicas failover auto
 - 3 replicas synchrones
 - Backup on secondary

Windows 2012 R2 SQL Server 2014

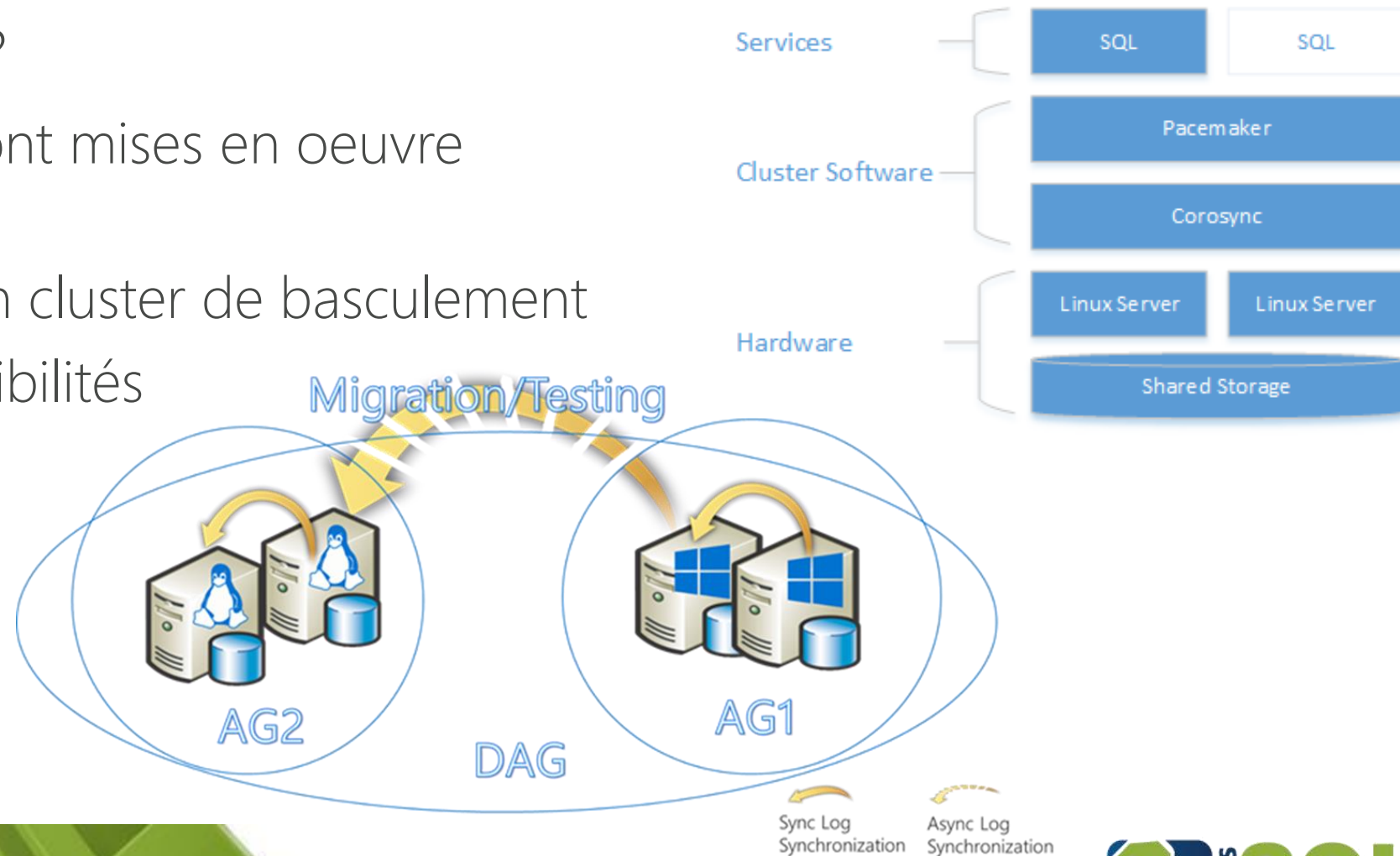
- WSFC
 - AD detached
- FCI
 - Témoin dynamique
 - Data sur disque CSV
- Groupe de disponibilité
 - 1 + 8 replicas
 - Replica sur Azure
 - Secondary readable even if primary unavailable

Windows 2016 SQL Server 2016

- WSFC
 - Domain Agnostic
 - Site awareness
 - Cloud witness
 - Rolling upgrade
- FCI
 - Storage spaces direct
 - Stretch Cluster
- Groupe de disponibilité
 - Distributed AGs
 - DTC Support
 - Auto seeding
 - 3 replicas auto failover
 - RO Load balancing
 - AG Standard edition
 - Performances

Conclusion

- SQL Server v.Next ?
 - Les 2 techniques sont mises en oeuvre
 - Gage de pérennité
 - Instance SQL sur un cluster de basculement
 - Groupes de disponibilités



Q & A

- Q & A
- Merci pour votre attention

