

DataFrogs

Les groupes de disponibilité
par la pratique

@conseilit



Microsoft®
SQL Server®



Christophe LAPORTE

 /in/christophelaporte

 conseilit.wordpress.com

 @conseilit

 conseilit@outlook.com

Audit – Conseil – Formation – Remote DBA



Nous n'aborderons **PAS**

- Pourquoi mettre en place une solution de haute disponibilité
- Les stratégies de haute disponibilité
- Les causes de non disponibilité
- Le coût de non disponibilité
- Et les sacrosaints RPO et RTO



Terminologie

- Cluster : groupe / grappe de serveurs
- Nœud / réplica : un serveur participant au cluster
 - Actif : propriétaire du groupe de ressource, supporte le charge
 - Passif : prêt à prendre le relai en cas de problème
- Une ressource : un service ou un device (disque, adresse IP) accédé directement ou indirectement par une application ou un utilisateur
- Un groupe de ressource : plusieurs ressources liées par des dépendances
- Failover : basculement d'un groupe de ressource vers le nœud secondaire
- Failback : retour automatique du groupe de ressource sur le nœud d'origine

Terminologie

- FCI : Failover Cluster Instance
- AG : Availability Group
- AD : Active Directory
- CNO : Cluster Name Object
- VCO : Virtual Computer Object
- SMB : Server Message Block
- FSW : File Share Witness
- WSFC : Windows Server Failover Cluster

Mise en œuvre du cluster Windows

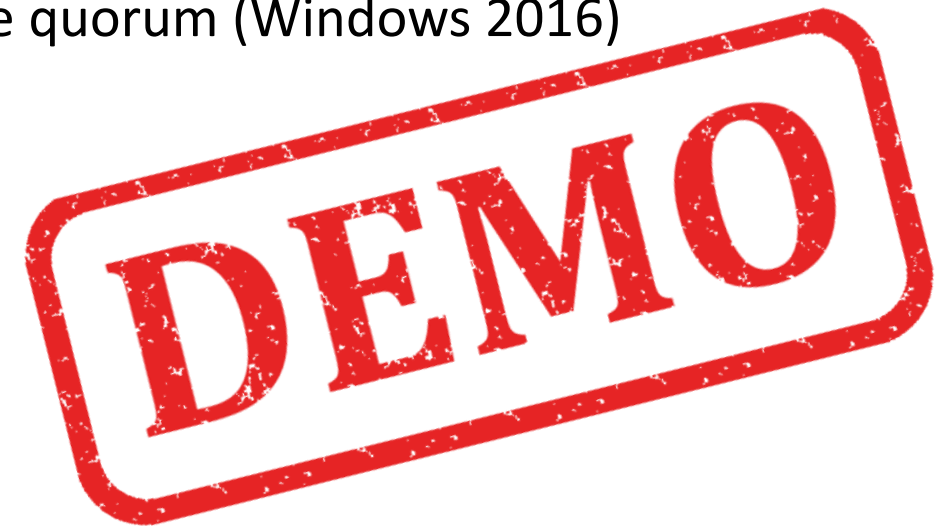
- Etapes principales
 - Ajout de la fonctionnalité Windows
 - Test du cluster
 - Création du Cluster
- Mise en œuvre
 - IHM
 - PowerShell



Quorum

Définir un consensus, suffisamment de votes (membres disponibles) pour assurer le service

- Basé sur des votes
 - Le cluster est fonctionnel quand un certain nombre de votes est atteint
 - Quorum = $\text{Floor}(N/2 + 1)$: Plus grand entier inférieur ou égal
- Quorum dynamique (Win 2012)
 - Ajuster les votes des nœuds dynamiquement
- Témoin dynamique (Win 2012R2)
 - Ajuster dynamiquement le vote du Quorum Witness
- Types de Quorums
 - Node majority
 - Node and disk majority
 - Node and file share majority
 - No majority : disk only
 - Azure quorum (Windows 2016)



Cluster – Ajustement de paramètres

ClusterLog

- Version agglomérée des ETW
(C:\Windows\System32\winevt\Logs)
 - fichier C:\windows\Cluster\reports\cluster.log

```
PS > Get-clusterlog
```

- Modifier la rétention
 - 8 MB < Taille < 2048 MB

```
PS > (Get-Cluster).ClusterLogSize
```

```
PS > Set-ClusterLog -Size 1024
```

- Eventuellement changer le niveau de détail

```
PS > Set-ClusterLog -Level X
```

Level	Error	Warning	Info	Verbose	Debug
0 (Disabled)					
1	X				
2	X	X			
3 (Default)	X	X	X		
4	X	X	X	X	
5	X	X	X	X	X

Paramètres réseau

- Tolérance à des erreurs réseau transitoires
- [Tuning Failover Cluster Network Thresholds - Microsoft Tech Community](#)

```
PS > (get-cluster).SameSubnetThreshold = 10
```

```
PS > (get-cluster).CrossSubnetThreshold = 20
```

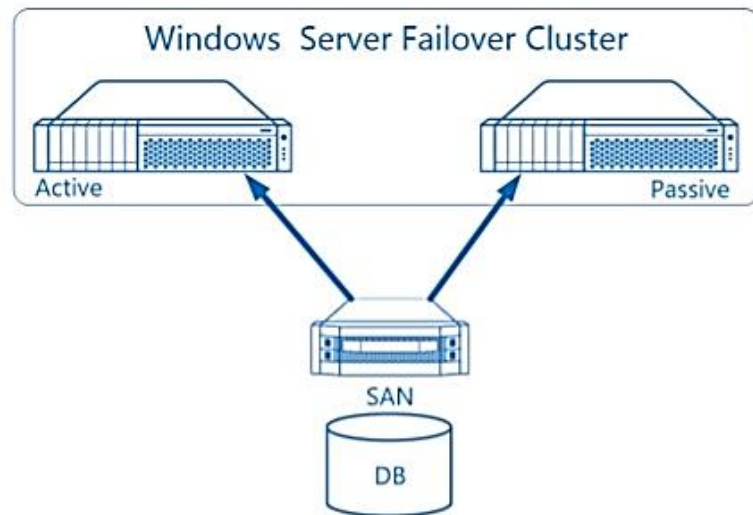
```
PS > (get-cluster).RouteHistoryLength = 40
```



La HA avant SQL Server 2012

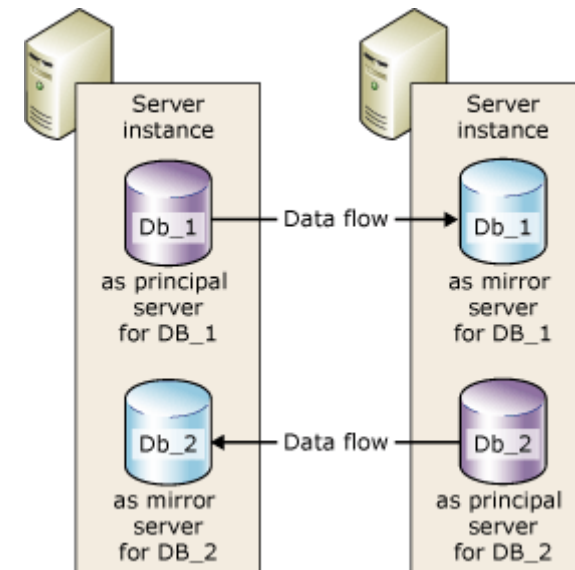
Cluster de basculement

- Granularité instance
- Nom réseau et @IP unique
- Setup spécifique
- Edition standard et entreprise



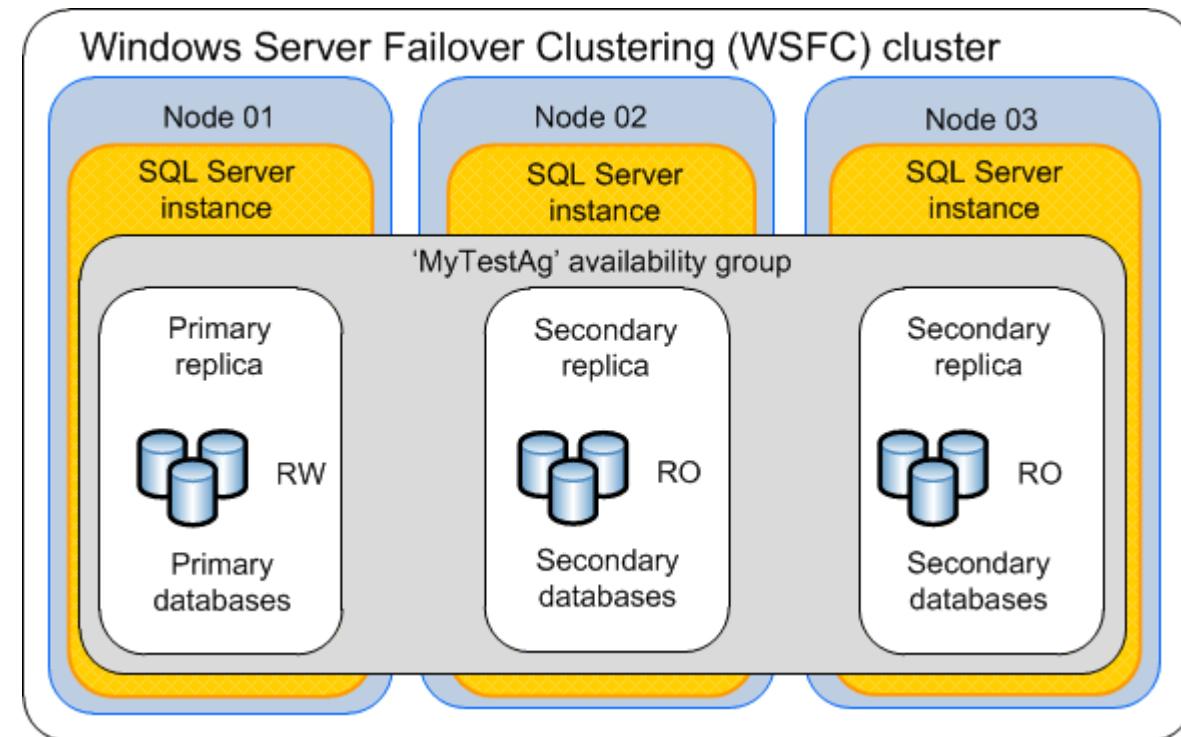
Mise en miroir

- Granularité database
- Répartition de charge
- Standalone instances
- Edition standard et entreprise

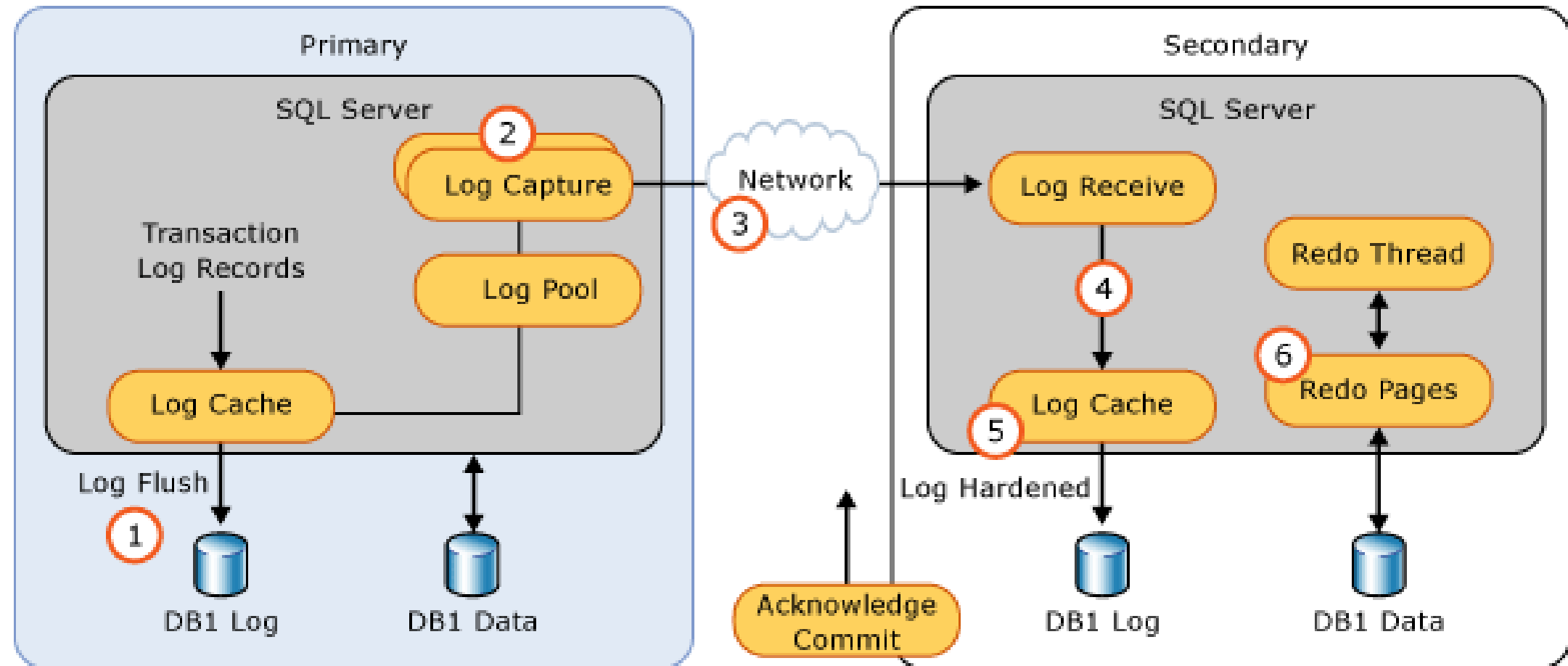


Les groupes de disponibilité

- Le meilleur des 2 mondes
- Cluster
 - Listener (@IP, Nom réseau)
 - Basculement automatique
- Database Mirroring
 - Un groupe de 1..N bases de données
 - Stockage asymétrique
 - Plusieurs réplicas
 - Synchrones ou asynchrones
 - Réplica éventuellement en lecture seule

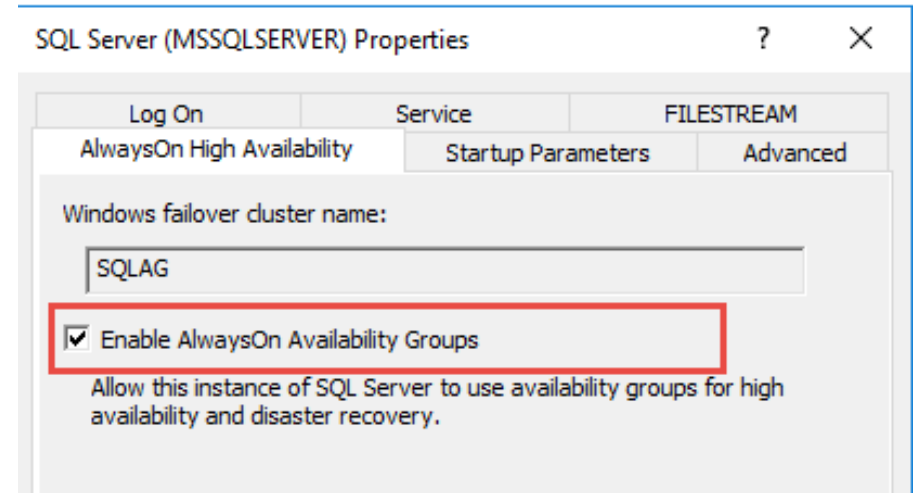


Fonctionnement d'un AG



Groupes de disponibilité : préparatifs

- Système
 - Activation de la fonctionnalité
- SQL Server
 - Création des points de terminaison TCP
 - Autoriser les connexion sur les points de terminaison
- Comment ?
 - Assistant dans SSMS
 - T-SQL
 - PowerShell



Groupes de disponibilité : implémentation

- Création d'un groupe de disponibilité 2 réplicas
 - Joindre un réplica secondaire au groupe de disponibilité
 - Joindre une base à un groupe de disponibilité
 - Ajouter un Listener à un groupe de disponibilité
- Création d'une routing list
 - Connexion sur un réplica secondaire (read intent)
- Comment ?
 - Assistant dans SSMS
 - T-SQL
 - PowerShell



Groupes de disponibilité : les limites

- Modèle de récupération **complet** seulement
- Seulement les bases
 - Utilisateur
 - En lecture écriture
 - Multi utilisateurs
- Synchronisation
 - Des Jobs SQL
 - Des SID des Logins SQL
- Versions identiques
 - Sauf si upgrade ...
- Editions identiques
- Architecture disque
 - Si possible identique
 - /!\ Support SQL2016-

Threads

- Nombre de worker threads utilisés

		# secondary replicas						
		1	2	4	5	6	7	8
# NB active databases	1	3	4	6	7	8	9	10
	5	11	16	26	31	36	41	46
	10	21	31	51	61	71	81	91
	20	41	61	101	121	141	161	181
	50	101	151	251	301	351	401	451
	75	151	226	376	451	526	601	676
	100	201	301	501	601	701	801	901
	250	501	751	1 251	1 501	1 751	2 001	2 251
	500	1 001	1 501	2 501	3 001	3 501	4 001	4 501

cores	Worker threads	Hadr pool size max
4	512	472
6	544	504
8	576	536
10	608	568
12	640	600
14	672	632
16	704	664
18	736	696
20	768	728
22	800	760
24	832	792
28	896	856
32	960	920
36	1 024	984
40	1 088	1 048
48	1 216	1 176
64	1 472	1 432
128	2 496	2 456
256	4 544	4 504
512	8 640	8 600

Maintenance des bases dans un AG

- Sauvegardes
 - **Emplacement réseau**
 - Complète
 - Principal
 - Secondaire en copy_only
 - Différentielle
 - Principal
 - Journal de transaction
 - Sur tous les réplicas
 - Log chain préservée
 - [sys.fn_hadr_backup_is_preferred_replica](#)
- Contrôle d'intégrité
 - Sur tous les réplicas ???
- Maintenance des index
 - Passage en mode asynchrone ?
 - log burst scenarios (online reindex, bulk load)



Ola Hallengren scripts

Problèmes courants

- Création de listener
 - Autorisations AD
 - Conflits d'@IP ou de nom DNS
 - Création d'un Client Access Point dans WSFC
- Connexion à un listener (multi subnet)
 - Resource listener Offline
 - Problème de Logins
 - Chaine de connexion (Application Intent, Intial Catalog)
- Read-Only routing list
 - Pas d'URL présente
 - Mauvaise configuration de port TCP
 - Problème de resolution de nom

Problèmes courants

- Erreur de synchronisation des secondaires
 - Architecture disque différente
 - Erreur d'authentification sur le EndPoint
 - Certificats absent sur secondaire (TDE)
 - Antivirus (Host intrusion)
 - Erreur dans les URL
 - Erreur de port HADRON (multi instance)
 - Règles Firewall
- Ajout de base dans un AG
 - Manque Backup Database
 - Backup log durant l'ajout de la base dans l'AG

Problèmes courants

- Attentes de type WRITE LOG / LOG BUFFER sur primaire
 - Configuration NUMA / Hidden scheduler NUMA NODE 0
- Croissance journal de transaction
 - Pas de truncate de la log tant que le REDO n'a pas été effectué sur tous les réplicas
 - log_reuse_wait_desc -> AlwaysOn
- Failovers inexplicables ???
 - Perte de Quorum
 - Lease Timeout
 - HealthCheck Timeout
 - SQL Dumps

Problèmes courants

- Bascule impossible
 - Une ou N base non synchronisée
 - Réplica secondaire déconnecté
 - WSFC : erreur de connexion à SQL Server
 - Nombre de panne maximal dépassé
- Durée de bascule importante
 - Beaucoup de transaction en attente de REDO
 - Performance / Matériel sur serveur secondaire
- Oubli du « Resume » après un failover manuel (Async)

Groupes de disponibilité : Evolutions

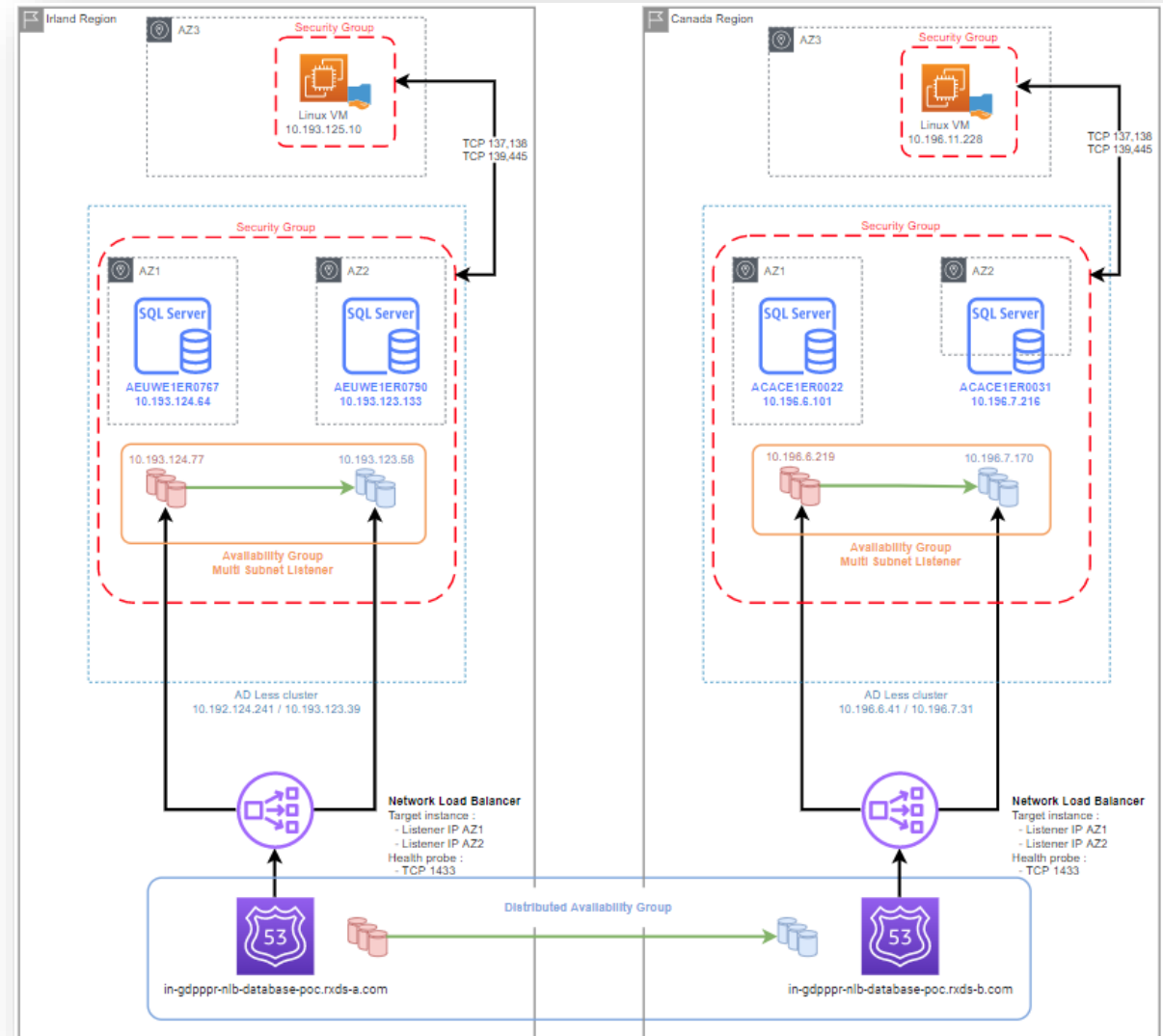
- SQL2016
 - Domaine Agnostic / AD-less AGs (Workgroup WSFCs)
 - Distributed Availability Groups, édition entreprise
 - Basic Availability Groups, édition standard
- SQL2017
 - Support de Linux
 - Reads Scale Availability Groups
 - Cluster less
 - Pas de failover automatique
 - /!\ ce n'est pas une solution de HA
- SQL2019
 - Redirection vers le réplica primaire depuis un secondaire

Groupes de disponibilité : Evolutions

	2012	2014	2016	2017	2019
<i>Total # of replicas</i>	5 (1p + 4s)	9 (1p + 8s)	2 (1p + 1s; Std) 9 (EE)	2 (Std) 9 (EE)	2 (Std) 9 (EE)
<i>Possible # of sync replicas</i>	2	3	2 (Std) 3 (EE)	2 (Std) 3 (EE)	2 (Std) 5 (EE)
<i>Possible # of auto failover replicas</i>	2	2	3	3	5
<i>DTC support</i>	No	No	Yes*	Yes	Yes
<i>Secondary Redo</i>	Single threaded	Single threaded	Parallel	Parallel	Parallel
<i>Cluster required</i>	Yes	Yes	Yes	Selected configurations	Selected configurations
<i>Active Directory Domain Services Required</i>	Yes*	Yes*	Selected configurations	Selected configurations	Selected configurations
<i>Automatic seeding</i>	No	No	Yes	Yes	Yes

Implémentation cloud

- Cloud : AWS
- 2 régions : Irlande & Canada
- Par région :
 - 1 AD-Less cluster
 - 3 zones de disponibilité
 - -> multisubnet
 - 1 AG failover automatique
 - 1 NLB pour redirection vers listener
- 1 distributed Availability Group
 - Listener_URL = NLB





[/in/christophelaporte](#)



conseilit.wordpress.com



[@conseilit](#)



conseilit@outlook.com



Conclusion

Q & A