L'agrégation de liens

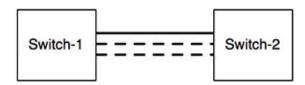
1. Introduction

Un réseau informatique se doit de remplir plusieurs exigences, dont la résilience et la modularité. Il est également souhaitable que le réseau puisse permettre une utilisation intégrale des ressources disponibles. Selon les technologies déployées, il est possible que l'architecte n'ait pas d'autre choix que d'utiliser un mécanisme actif/passif, utilisant les ressources additionnelles uniquement en cas de panne.

L'agrégation de liens permet la formation d'un lien logique répondant aux exigences de haute disponibilité, de modularité et de répartition de charge. Avant de décrire en détail l'agrégation de liens et ses apports, analysons un réseau sans utilisation de cette technologie.

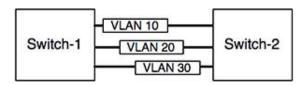
2. Sans l'agrégation de liens

Dans une configuration classique où deux commutateurs sont liés à l'aide de connexions parallèles, un seul lien sera en réalité utilisé. Cela s'explique par l'utilisation par les commutateurs d'un protocole bloquant les connexions redondantes, Spanning-Tree, évitant ainsi la formation de boucles de niveau 2 sur le réseau. La figure ci-dessous montre deux commutateurs avec un lien actif (ligne continue) et deux autres bloqués (lignes discontinues).

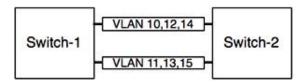


Cependant, il est possible d'utiliser des liens parallèles sans avoir pour autant recours à l'agrégation de liens. Les stratégies qui peuvent être utilisées sont :

• la répartition par lien du trafic pour un VLAN unique, comme l'illustre la figure ci-dessous :



• la répartition par lien en fonction du numéro de VLAN (VLAN pair, impair), comme l'illustre la figure ci-dessous :



Dans le premier scénario, nous n'offrons à aucun moment une quelconque redondance et la vitesse de trafic entre ces deux commutateurs sera limitée à la vitesse du lien dédié au VLAN.

Dans le deuxième scénario, la répartition étant usuellement effectuée entre deux liens, il suffirait qu'un lien ne soit plus opérationnel pour priver de connectivité plusieurs VLAN simultanément, ce qui n'est pas non plus une solution viable. Bien entendu, l'administrateur pourrait doubler ces connexions pour obtenir une résilience automatique.

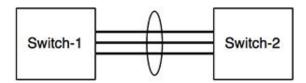
Malgré le déploiement de ressources additionnelles, la vitesse utilisable sera là aussi limitée à la vitesse d'un lien pour les utilisateurs d'un même VLAN.

Le besoin de disposer d'un mécanisme plus solide et plus fiable est donc essentiel.

Voyons maintenant comment l'agrégation de liens répond à ces différentes problématiques et pourquoi cette technologie est devenue un incontournable des architectures réseau.

3. Avec l'agrégation de liens

Un agrégat entre deux commutateurs permet d'améliorer la fiabilité de notre connexion intercommutateur tout en permettant une meilleure performance. Ce mécanisme peut être formulé sur une topologie de réseau comme cidessous :



Voyons en détail quels sont les avantages de l'agrégation de lien :

- Haute disponibilité: même en cas de panne d'un lien compris dans l'agrégat, l'agrégat reste opérationnel tant qu'un lien y est encore actif. Ainsi, un agrégat ne sera déclaré inactif que si et seulement si l'ensemble de ses membres est en état de panne. En cas de dysfonctionnement d'un lien membre de l'agrégat, le trafic est automatiquement redirigé vers les ports actifs de façon totalement transparente pour l'utilisateur, le temps de convergence est quasi nul (http://www.cisco.com/en/US/tech/tk389/tk213/technologies_white_paper09186a0080092944.shtml).
- Répartition de charge : parmi les données qui transiteront au sein de l'agrégat, toutes n'emprunteront pas le même chemin. D'ailleurs, il ne s'agit pas ici de fragmenter la trame pour l'envoyer en parties à travers plusieurs liens mais de décider quel lien empruntera la trame ! Ainsi, les commutateurs (selon leurs configurations et leurs fonctionnalités) peuvent décider de placer une trame sur un lien en fonction des critères suivants : adresse MAC (L2), adresse IP (L3), numéros de ports (L4). Ces critères peuvent être considérés des deux côtés de la communication (source et/ou destination), ensemble ou séparément. Ainsi, l'administrateur peut décider de faire passer par le même lien un flux entre une même IP source et IP de destination ou indiquer au commutateur de prendre en compte les informations de couche Transport, c'est-à-dire les numéros de ports source et destination permettant à deux flux concurrents entre deux mêmes machines d'emprunter un lien différent.
- Pas de blocage de ports : un seul lien logique étant formé, Spanning-Tree ne bloquera pas les ports membres de l'agrégat. Le protocole considérera uniquement l'agrégat et non plus les ports membres de façon individuelle.
- **Résilience automatique** : selon le mécanisme d'agrégation utilisé, il est possible de solliciter des ports « en attente » qui ne seront joints automatiquement à l'agrégat qu'en cas de panne de connexions actives. Cela est uniquement possible si nous utilisons un agrégat négocié à l'aide d'un protocole particulier.

La bande passante d'un agrégat est souvent confondue avec la capacité de transfert maximale entre deux machines, il s'agit pourtant bien de deux choses distinctes. D'un côté, la bande passante d'un agrégat est tout simplement la somme de la bande passante des ports actifs, unis dans l'agrégat. De l'autre, le débit exploitable entre deux machines dépend de la façon dont fonctionne l'agrégat mais aussi d'autres facteurs limitants, comme la vitesse des interfaces réseau des ordinateurs. Ainsi, un agrégat de 8 ports de vitesse 100 Mbit/s ne garantira en aucun cas un débit utilisable entre les machines de 800 Mbit/s pour une seule et même communication.

Il faut en réalité voir les agrégats comme des « autoroutes de l'information », je m'explique. Prenez une autoroute trois voies à laquelle nous rajouterions une voie additionnelle, représentée par le schéma ci-dessous.

Voie 1 - 130 km/h	
Voie 2 - 130 km/h	
Voie 3 - 130 km/h	
Voie 4 - 130 km/h	

Dans ce cas, nous permettrions à des véhicules d'emprunter cette voie, augmentant de fait la capacité d'admission de l'autoroute sans pour autant en modifier la vitesse limite. En outre, sachiez que les ports doivent disposer de la même vitesse pour joindre un même agrégat.

Pour chaque communication transitant par l'agrégat, le commutateur désignera un lien par lequel l'information transitera, par le biais d'un calcul de hash XOR (ou exclusif). Comme évoqué précédemment, il appartient à l'administrateur de définir les critères de répartition de charge, parmi les suivants :

Couche 4 - Transport		Couche 3 - Réseau				Couche 2 - Liaison de données		
Port source	Port destination	Adresse source	ΙP	Adresse IP of destination	е	Adresse source	MAC	Adresse MAC de destination

Parmi les scénarios paramétrables par l'administrateur, on peut en citer quelques un avec leurs avantages et inconvénients :

Critère(s) paramétré(s)	Avantage(s)	Inconvénient(s)		
MAC source ET MAC de destination	Utile lorsqu'il s'agit de répartir des accès au sein d'un même sous-réseau.	Considérant un trafic entre un poste et sa passerelle par défaut (utilisée pour sortir du réseau local dans lequel se trouve le poste), le lien sélectionné sera toujours le même.		
IP source ET IP de destination	Qu'elles soient locales ou non, les communications peuvent être réparties efficacement entre les machines.	Plusieurs flux entre les mêmes adresses utiliseront le même lien, limitant le débit à la vitesse d'un lien de l'agrégat.		
Port source ET port de destination	Les flux entre deux machines peuvent emprunter un lien différent et ainsi profiter de la vitesse des liens de l'agrégat.	Risque de saturation de l'agrégat plus important. Un flux est toujours limité à la vitesse maximale d'un lien.		

L'agrégation de liens dans l'univers Cisco s'appelle l'EtherChannel, qu'il soit statique ou dynamique. Avec un commutateur type Catalyst 2960, nous pouvons configurer jusqu'à 16 ports dans un agrégat, et jusqu'à six agrégats différents au sein d'un même commutateur.

Justement, il est temps de parler maintenant des différents modes d'agrégations possibles.

4. Types d'agrégation

La configuration de l'agrégation de lien peut être de deux natures : statique ou dynamique.

Dans une configuration statique (c'est-à-dire sans négociation d'agrégation), un administrateur peut définir jusqu'à huit ports qui joindront l'agrégat et qui seront actifs dans ce dernier. Cependant, en cas de panne, le statut de

l'agrégat dépendra du statut des ports inclus statiquement dans l'agrégat.

Si un agrégat dynamique est configuré, la formation de l'agrégat sera dynamiquement négociée entre les commutateurs, validant ainsi les paramètres de configuration et ses membres. En utilisant un commutateur Cisco, vous avez deux choix concernant le protocole de négociation : **PAgP**, protocole propriétaire Cisco ou **LACP**, protocole standardisé.

LACP, au-delà d'être un standard IEEE, nous permet d'inclure dans un agrégat jusqu'à 16 ports au total, 8 ports actifs et 8 en « standby », permettant en cas de panne d'un des liens actifs de l'agrégat, l'utilisation automatique (et potentiellement immédiate selon les plateformes) d'un des ports en attente d'utilisation.

En conclusion, vous comprendrez donc que s'agissant d'une quelconque négociation utilisant un commutateur virtuel VMware, nous ne pourrons faire appel qu'au protocole standard, LACP. Voyons comment les deux options de configurations peuvent être satisfaites du côté du commutateur physique et du commutateur virtuel.

5. L'agrégation de liens avec les hyperviseurs vSphere

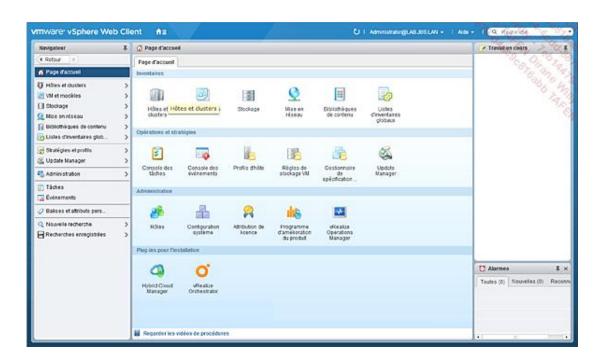
Les commutateurs virtuels VMware ne fonctionnant pas comme un commutateur physique classique, nous n'aurons pas les limitations que nous décrivions avec Spanning-Tree, même en cas de non-agrégation des liens. Mais comme vu précédemment, il est utile de déployer cette technologie, en statique ou dynamique. vSphere ainsi que vCenter prennent d'ailleurs en charge ces deux formules de déploiement.

Cependant, les options offertes et la configuration ne seront pas identiques, c'est ce que nous allons voir !

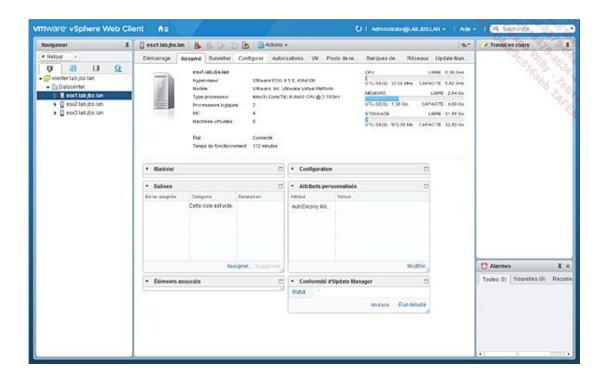
6. Configuration d'un agrégat statique avec un vSphere Standard Switch

La configuration d'un agrégat statique (et donc sans négociation) peut être effectuée sur un vSS (*vSphere Standard Switch*) à l'aide du vSphere Web Client selon la procédure suivante.

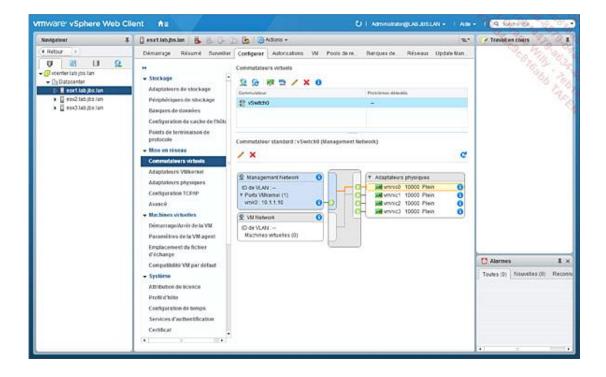
Positionnons-nous sur la partie Hôtes et clusters pour démarrer la configuration de l'hôte.



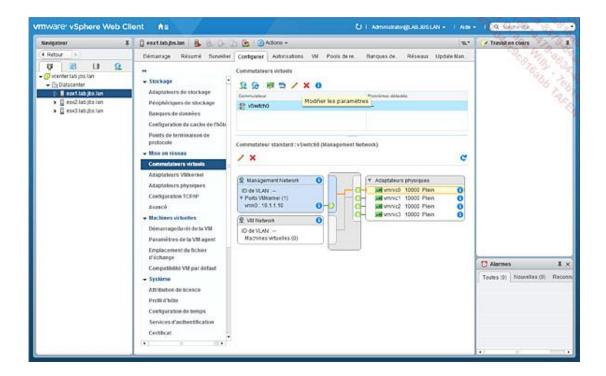
Voici la page décrivant le statut et la configuration de l'hôte. Cliquez sur l'onglet Configurer.



Puis rendez-vous dans la partie Commutateurs virtuels pour configurer le commutateur virtuel standard, vSwitch0.



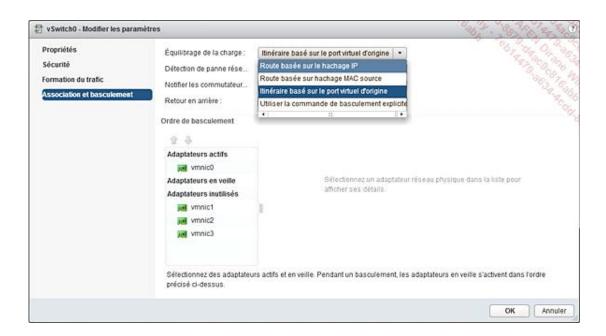
Sélectionnez l'icône en forme de crayon au niveau du vSwitch pour en modifier les paramètres.



Dans la section Association et basculement, on peut constater qu'il y a un seul adaptateur utilisé, vmnic0.



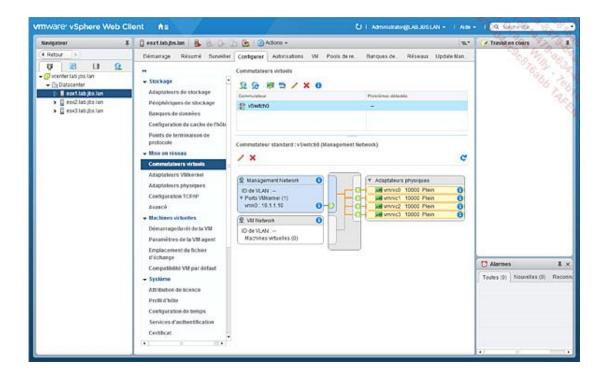
Sélectionnons tout d'abord le moyen de répartition de la charge, ici l'équivalent de « IP Hash », **Route basée sur le hachage IP**, changeant le paramètre par défaut basé sur le port virtuel d'une machine virtuelle.



À l'aide des flèches, positionnez les quatre adaptateurs comme actifs. Enfin, validez la configuration à l'aide du bouton **OK**.



L'administrateur peut constater que tout fonctionne à l'aide de la console mise à jour.



Notez que ces paramètres peuvent être également configurés au niveau des groupes de ports, rendant possible (malgré cette configuration au niveau du commutateur virtuel) l'utilisation d'un autre adaptateur physique pour un groupe de machines virtuelles.

Cette configuration se doit d'être reflétée sur le commutateur connecté à l'hyperviseur, voici une configuration exemple pour un commutateur Cisco :

```
port-channel load-balance src-dst-ip
!
int range fa0/1-4
switchport trunk encapsulation dot1q
switchport mode trunk
channel-group 1 mode on
!
int Port-Channel1
switchport trunk encapsulation dot1q
switchport mode trunk
```

La première ligne étant importante pour définir la répartition par lien en fonction de l'adresse IP source et de destination, en cohérence avec ce que nous avons configuré sur le commutateur virtuel, sous l'option **IP Hash**. Par défaut, un commutateur Catalyst 2960 utilise une répartition de charge basée sur l'adresse MAC source (http://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst2960/software/release/12-2 53 se/configuration/guide/2960scq/swethchl.html#25587).

Les lignes de configuration suivantes précisent la configuration du lien, dépendant de ce que vous souhaitez faire, ici un trunk dot1q de 4 ports. Prêtez attention à ce que la configuration (sauf la commande **channel-group**) soit reflétée à la fois sur les ports membres et le port logique d'agrégat, ici Port-Channel1.

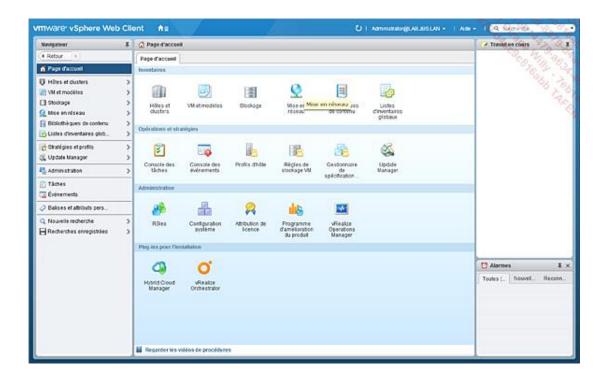
Concernant les lignes précisant l'encapsulation à utiliser (dot1q), il est à noter que cette commande est nécessaire sur les commutateurs supportant à la fois Dot1q et d'autres protocoles, tels que son équivalent propriétaire Cisco, ISL.

7. Configuration d'un agrégat statique avec un vSphere Distributed Switch

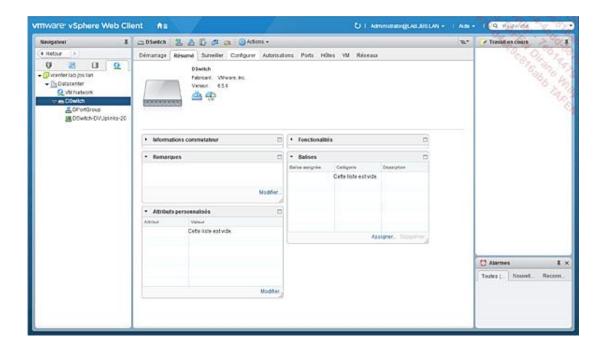
Configurer un agrégat statique sur un switch distribué peut se faire au niveau du « port group » ou d'un port mais pas au niveau du commutateur virtuel, à la différence d'un vSS.

Nous allons faire la configuration d'un hôte en deux temps : commencer par la configuration des « uplinks » puis la création d'un groupe de ports utilisant l'agrégat. La procédure à suivre est la suivante :

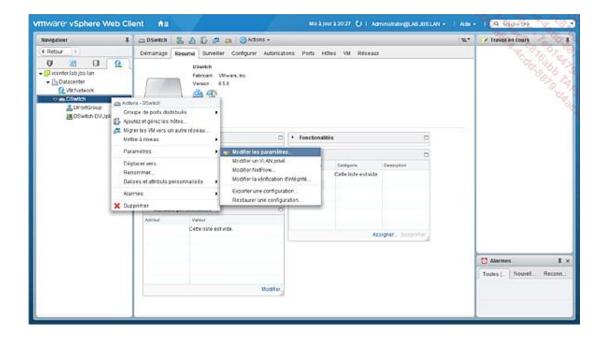
Rendez-vous dans la partie Mise en réseau pour commencer.



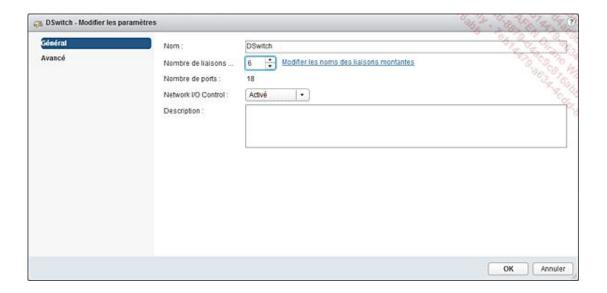
Voici la page de résumé du commutateur distribué déjà existant sur notre infrastructure.



Effectuez un clic droit sur ce dernier pour en modifier les paramètres comme suit.



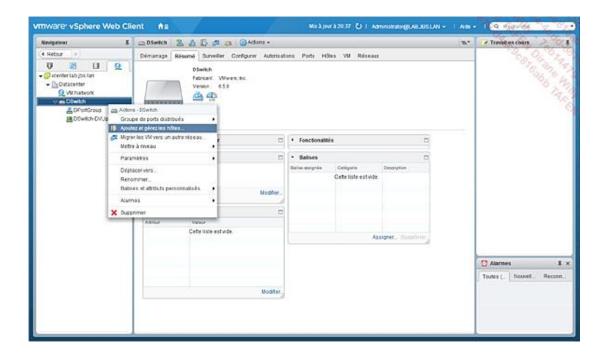
Augmentez le nombre de ports pour y rajouter les trois « uplinks » de l'hôte qui ne sont pas inclus dans la configuration. La valeur précédente était ici de 3, un port physique pour chaque hôte.



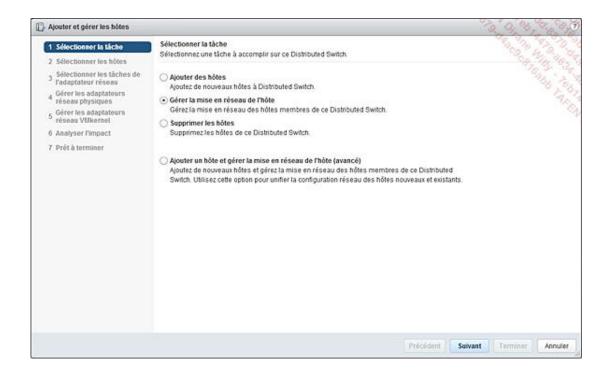
Nommez les ports pour les identifier aisément, puis validez.



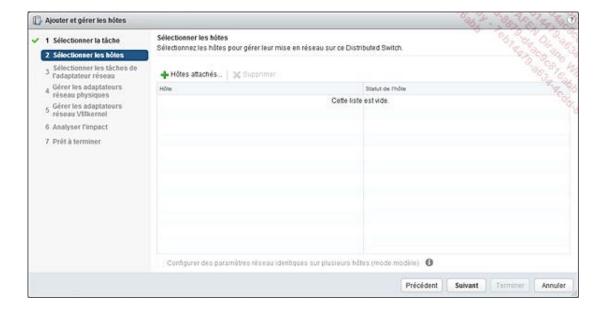
Mettez à jour la configuration de l'hôte réseau concerné par le changement.



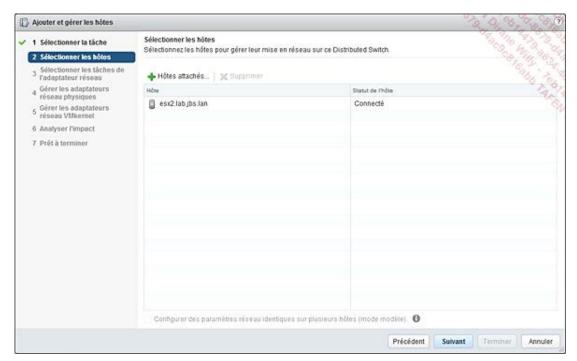
Sélectionnez l'option **Gérer la mise en réseau de l'hôte**.



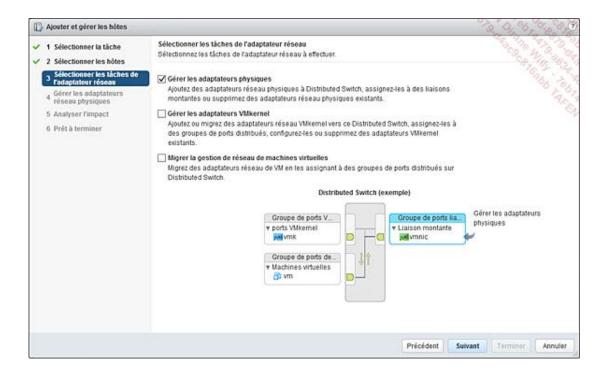
Ajoutez ici l'hôte concerné par vos changements, ici ESX2, à l'aide du bouton Hôtes attachés.



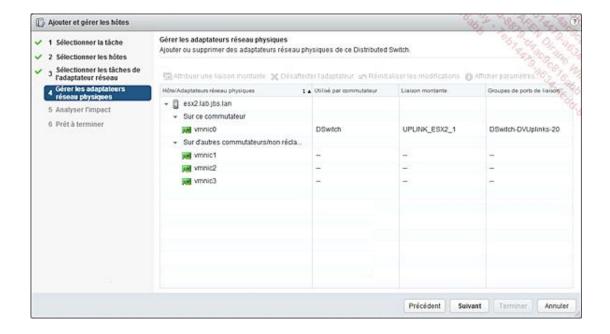




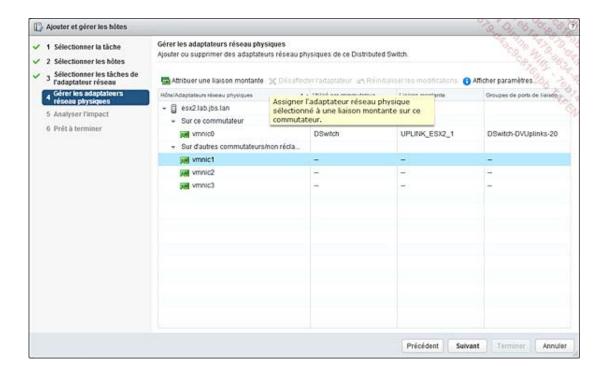
Sélectionnez ici uniquement l'option pour apporter des changements aux adaptateurs physiques.



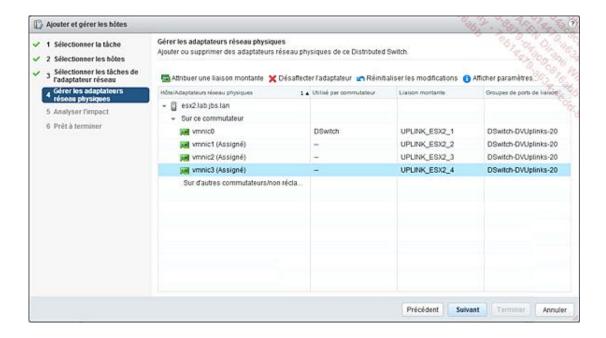
Seul un vmnic est associé. Associez les autres ports avec les appellations que nous avons définies précédemment.



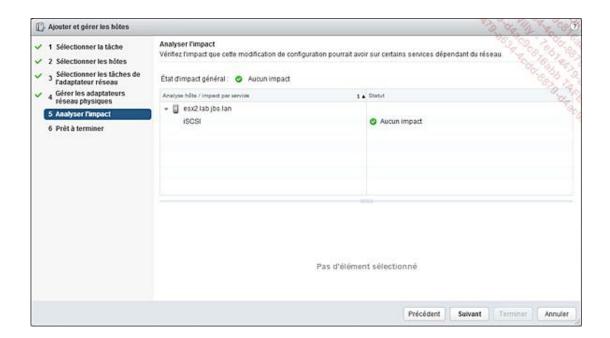
Pour associer un port physique à une liaison montante, il suffit de sélectionner le port physique et de cliquer sur le bouton **Attribuer une liaison montante**.

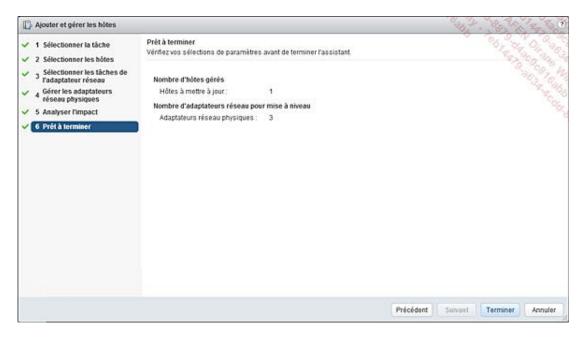


Vous devriez obtenir un résultat comme montré ci-dessous une fois cette étape terminée.

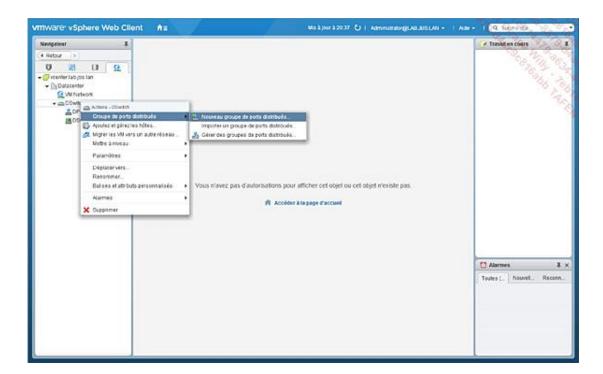


Puis terminez l'assistant.

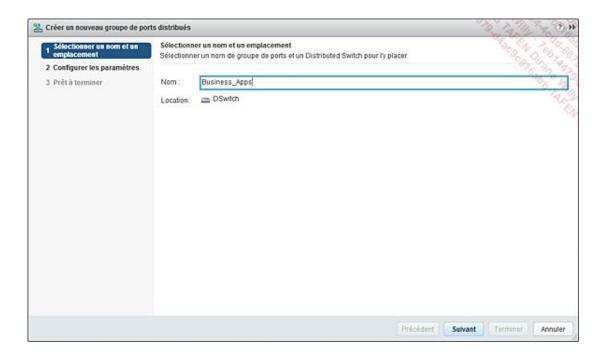




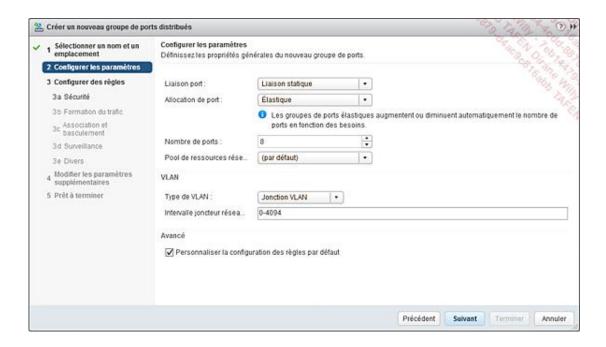
Créez maintenant un groupe de port distribué qui utilisera cette liaison agrégée.



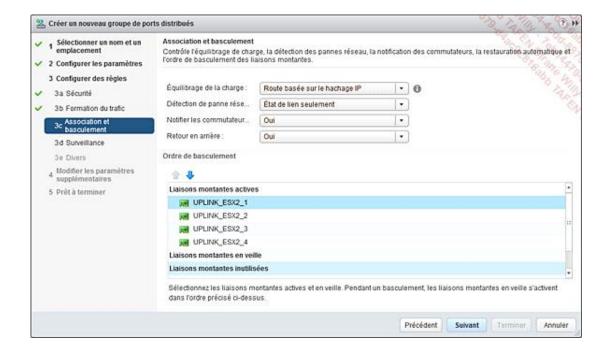
Donnez-lui un nom.



Puis définissez-le comme un groupe de port qui laissera les machines virtuelles faire du trunking de VLAN.



Puis, associez les liaisons montantes au groupe de port, en n'oubliant pas de définir la répartition de charge à un équivalent type « IP Hash ».



Une fois cette étape effectuée, validons la création du groupe de ports. Les machines virtuelles s'exécutant sur ESX2 pourront bénéficier de l'agrégation de ports configurés sur le commutateur.

La configuration côté commutateur physique reste la même ici.

8. Configuration d'un agrégat dynamique avec un vSphere Distributed Switch

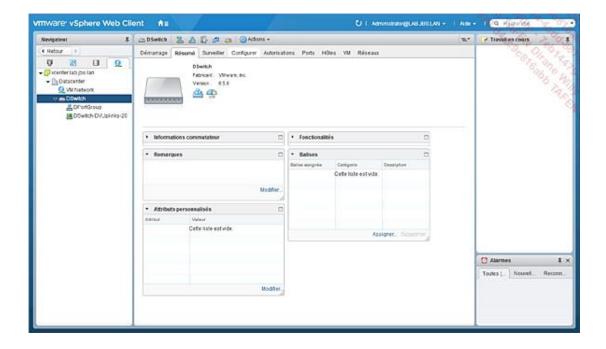
Un agrégat dynamique peut être formé entre un hyperviseur et un commutateur physique, à l'aide d'un vDS.

La configuration se fait comme suit, à l'aide du vSphere Web Client.

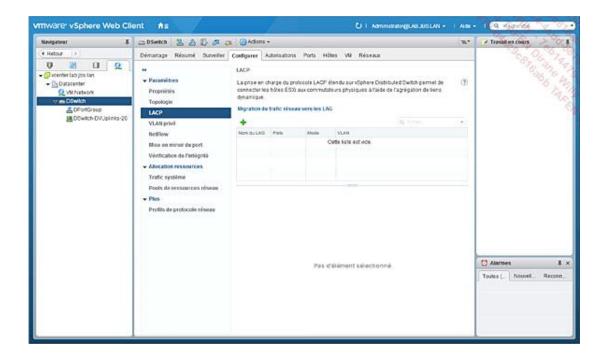
Commencez par vous rendre dans la partie réseau.



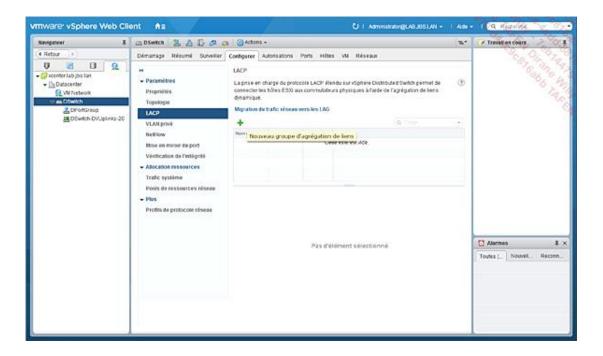
Puis, positionnez-vous sur le switch distribué que vous souhaitez configurer, puis cliquez sur l'onglet Configurer.



Dans la section **LACP**, vous allez pouvoir créer l'agrégat ainsi qu'en spécifier ses paramètres.



Cliquez sur le bouton + pour commencer.

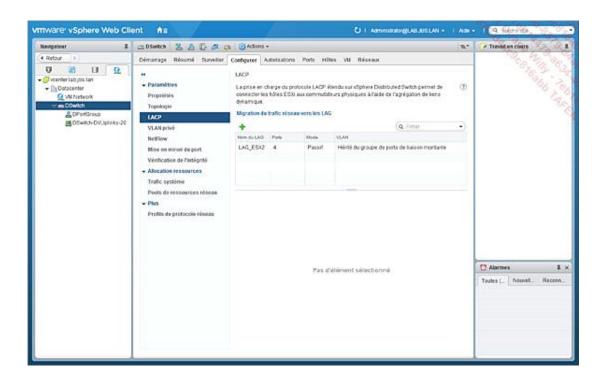


Spécifiez ici les caractéristiques de l'agrégat qui sera négocié, ici :

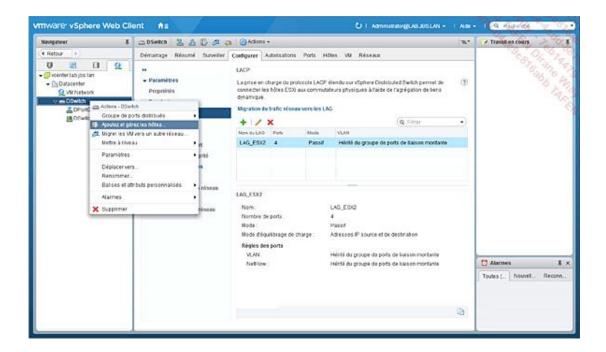
- Le nom : « LAG_ESX2 ».
- Le nombre de ports.
- Le mode : ici spécifier « passif » car le commutateur physique jouera le rôle de négociateur dans l'établissement de l'agrégat. Vous trouverez plus de détails sur le sujet en fin de section.
- Le mode de répartition de charge : ici l'équivalent « IP Hash ».



L'agrégat créé apparaît dans la liste des agrégats LACP désormais.



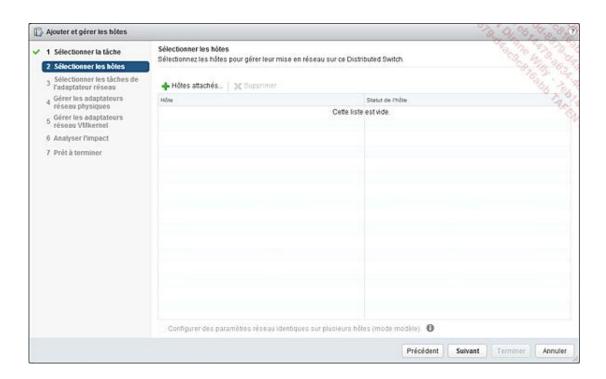
Une fois l'agrégat créé, il vous suffit de réassocier les ports physiques de l'hyperviseur à ceux de l'agrégat dynamique. Pour ce faire, lancez l'assistant de configuration des hôtes.



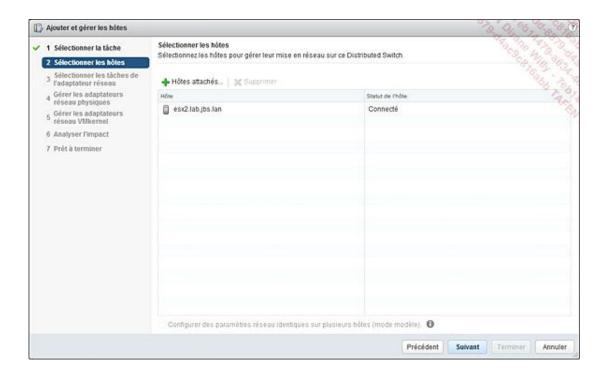
Sélectionnez l'option pour modifier la configuration réseau de l'hôte.



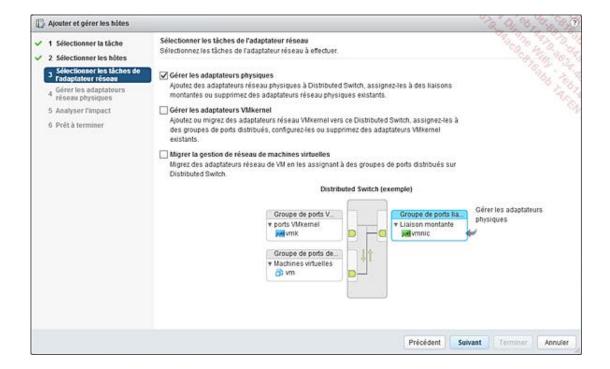
Rajoutez l'hôte que vous souhaitez configurer, ici ESX2.



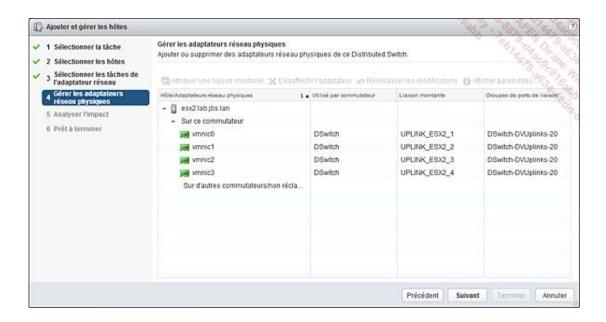




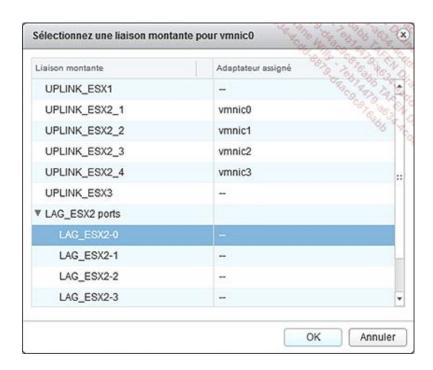
Spécifiez ce que vous souhaitez modifier, ici la configuration des adaptateurs physiques.



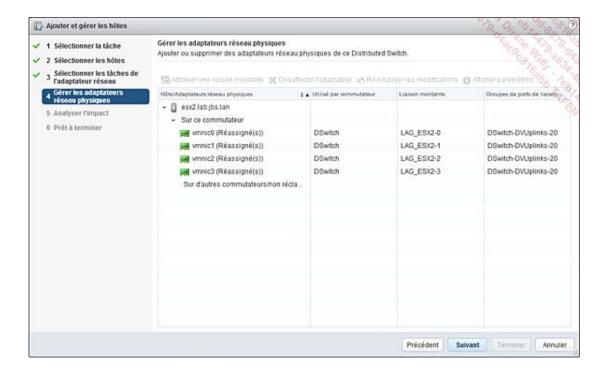
Les ports sont associés tels quels avant la configuration. Sélectionnez un lien puis cliquez sur **Attribuer une liaison montante**.



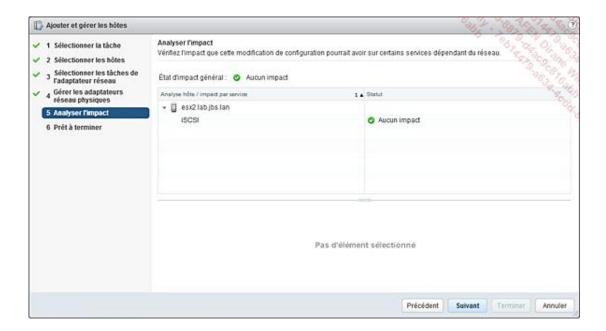
Associez le port à la liaison montante créée par la configuration de l'agrégat dynamique.



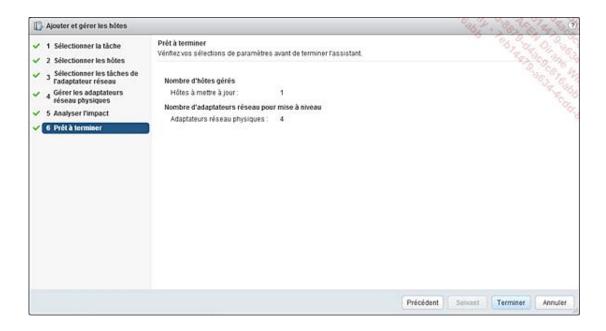
Répétez cette opération pour obtenir la configuration comme ci-dessous, puis validez.



Une analyse d'impact est effectuée avant de poursuivre, validez pour accéder au dernier écran.



Un dernier écran permet à l'administrateur de visualiser les changements et de les valider.



Avant d'appliquer ces changements et pour permettre à l'agrégat de se négocier, il est nécessaire de configurer le commutateur physique, ce que nous allons voir maintenant.

La configuration du côté d'un commutateur Cisco est la suivante :

```
port-channel load-balance src-dst-ip
!
int range fa0/1-4
switchport trunk encapsulation dot1q
switchport mode trunk
channel-group 1 mode active
!
int Port-Channel1
switchport trunk encapsulation dot1q
switchport mode trunk
```

Là aussi, il faut faire attention au mode de répartition de charge utilisé, ici nous restons sur l'équivalent « IP Hash » de la perspective de l'hyperviseur.

La ligne configurant le lien entre le port et l'agrégat « channel-group » définit également le comportement du commutateur par rapport à l'hyperviseur durant la négociation.

Ainsi, un port en mode « active » cherchera à envoyer les LACPDUs (*LACP Data Units*) pour négocier un agrégat tandis qu'un mode « passive » se contentera d'attendre que l'autre entité envoie une demande de négociation pour y répondre. Vous comprendrez ainsi que notre hyperviseur ne pourrait pas « monter » l'agrégat si la configuration côté vDS et commutateur physique avait été fixée en mode « passive ».

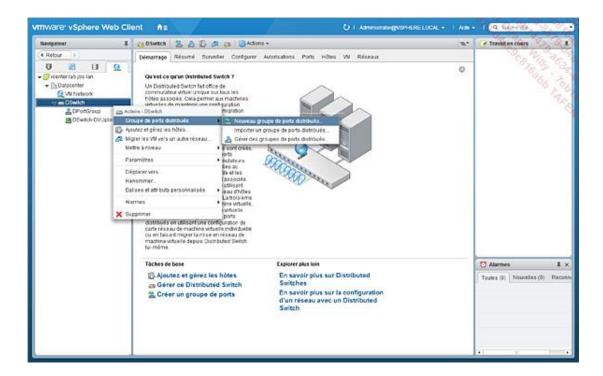
Le reste des lignes constitue la configuration classique de l'agrégat à répliquer à la fois sur l'agrégat et les ports membres.

Enfin, pour utiliser l'agrégat, il vous reste à l'associer avec un groupe de port distribué existant ou à en créer un nouveau.

0

Note importante : un groupe de ports ne peut pas contenir à la fois des liaisons montantes physiques et des agrégats, vous ne pourrez donc pas mélanger ces deux types de connexion.

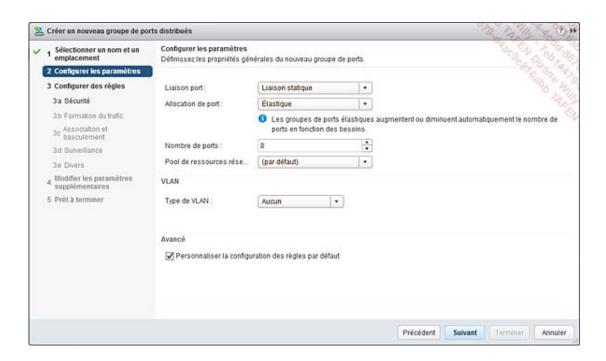
Tout d'abord, créez un groupe de port qui utilisera l'agrégat, à l'aide de la procédure suivante :



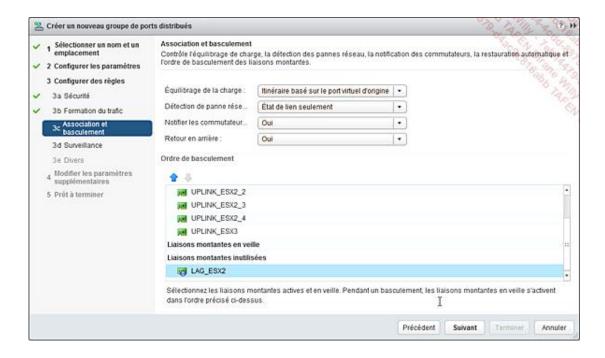
Donnez-lui un nom, ici « Business_Apps ».



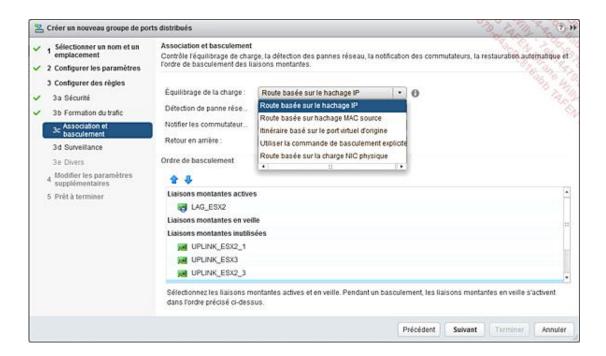
Puis cochez la case pour paramétrer les fonctionnalités du groupe de port, dont la partie **Association et basculement**.



Voici l'écran tel qu'il apparaît avant modification. Vous devez ici définir la liaison montante comme le seul lien actif ici.



Vous pouvez également définir le mode d'équilibrage de la charge. Cependant, cela n'est pas nécessaire car nous l'avons défini au niveau de l'agrégat. De plus, lorsqu'un seul agrégat figure dans une liaison montante (comme c'est le cas ici), le mode de répartition de charge de l'agrégat remplace celui défini ici.



Puis appliquez les changements.

