



GRAND DUCHY OF LUXEMBOURG  
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa  
Infrastructure Trust Fund

# Conception d'un Point d'Echange Internet (IXP)

Conception technique de l'IXP,  
ressources techniques et services à  
valeur ajoutée



# Remerciements et attribution

**Le contenu et informations de cette présentation sont initialement développés et maintenus par l'(les) organisation(s)/individu(s) ci-dessous et mis à la disposition du project AXIS de l' Union Africaine**

**Cisco ISP/IXP Workshops**

**Philip Smith: - [pfsinoz@gmail.com](mailto:pfsinoz@gmail.com)**

 **APNIC** [www.apnic.net](http://www.apnic.net)

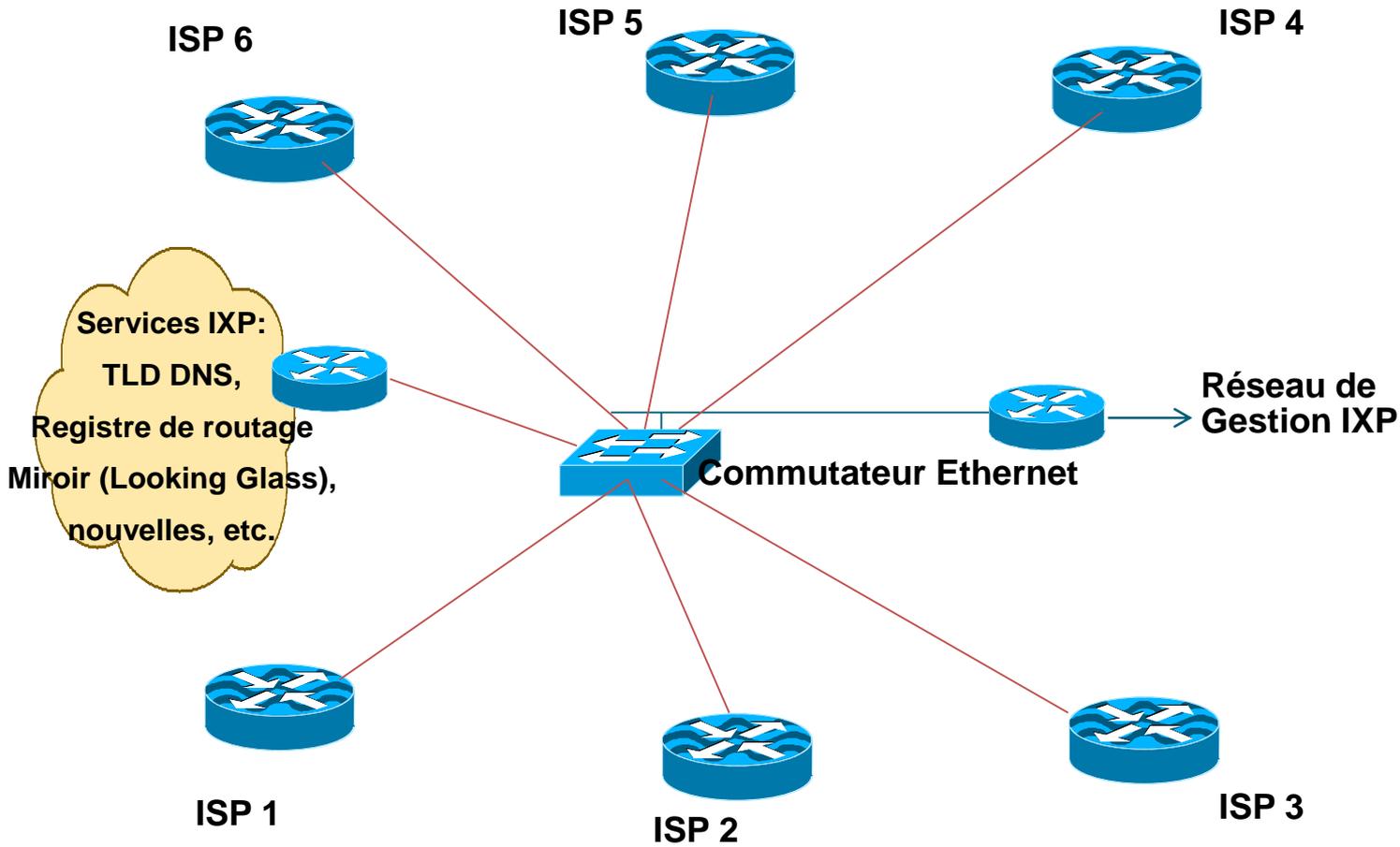
# Conception IXP

- Point d'échange Niveau 2
- Considérations sur la conception
- Collecteurs de Route et Serveurs
- Qu'est-ce qui peut mal tourner?

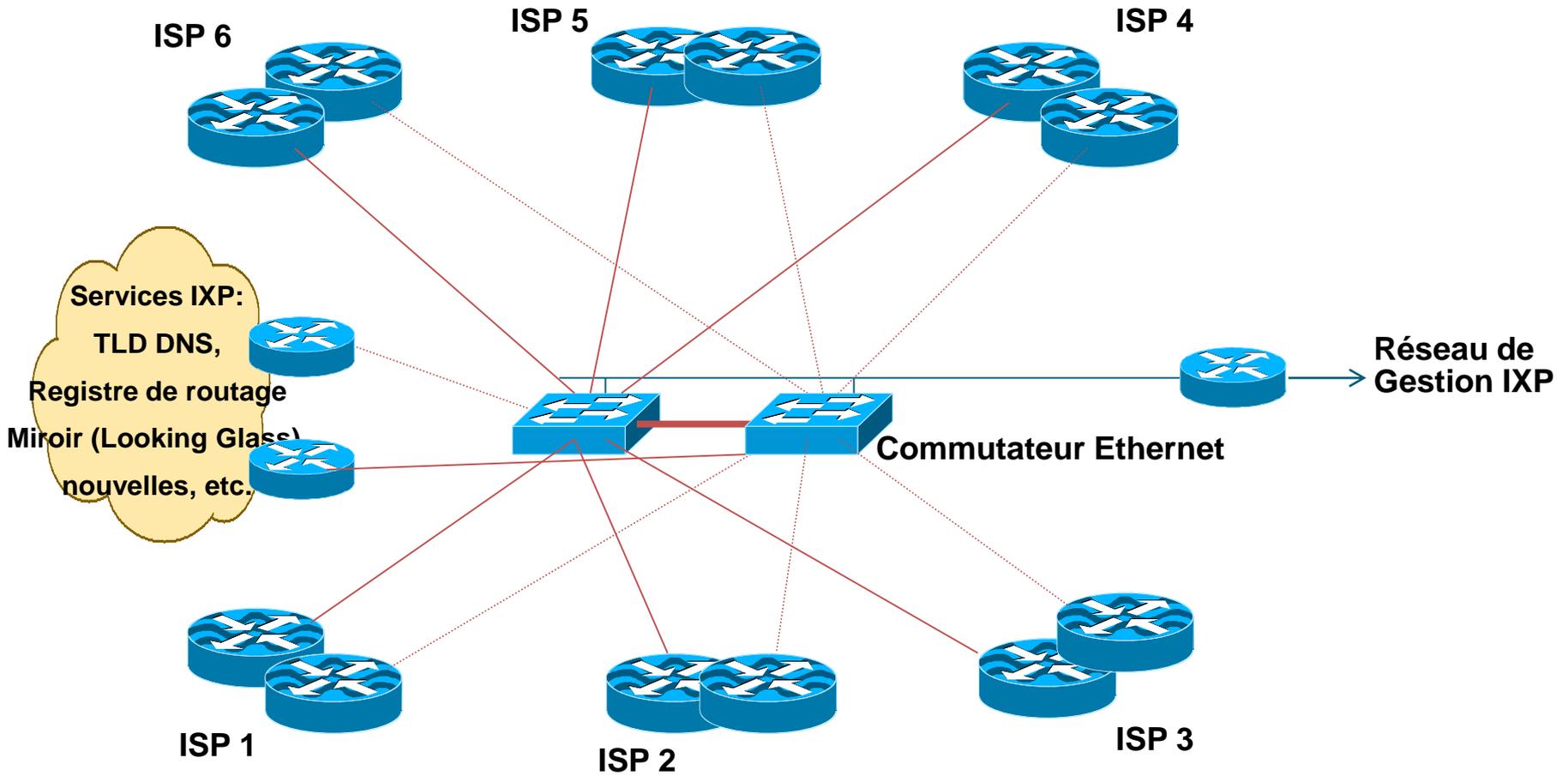
# Échange de Niveau 2

L'IXP traditionnel

# Échange de Niveau 2



# Échange de Niveau 2



# Échange de Niveau2

- Deux commutateurs pour la redondance
- Les ISP utilisent des routeurs pour la redondance ou pour le partage de charge
- Offre des services pour le «bien commun»
  - Portails Internet et moteurs de recherche
  - DNS TLD, Nouvelles, serveurs NTP
  - Registre de routage et miroir

# Échange de Niveau 2

- Exige une gestion neutre de l'IXP
  - généralement financé à parts égales par les participants de l'IXP
  - Couverture, support, services à valeur ajoutée 24 heures sur 24, 7 jours sur 7
- Emplacement sûr et neutre
- Configuration
  - Espace d'adressage privé, si pas de transit et de services à valeur ajoutée
  - Les ISP exigent AS, les IXP de base n'en ont pas besoin

# Échange de Niveau 2

- Considérations de sécurité des réseaux
  - Le commutateur LAN doit être configuré de manière sécurisée
  - Les routeurs de gestion nécessitent une authentification TACACS +, une sécurité vty
  - Les services IXP doivent être derrière un/des routeur(s) avec des filtres puissants

# Couche 2 par rapport à la Couche 3

- Niveau 3
  - L'équipe IXP nécessite une bonne connaissance de BGP
  - S'appuie sur une partie tierce pour la configuration BGP
  - Moins de liberté sur qui s'apparie avec qui
  - Habituellement entre en concurrence avec l'adhésion à l'IXP
  - Tend à être distribuée sur une zone large

# Couche 2 par rapport à la Couche 3

- Niveau 2
  - L'équipe IXP n'a pas besoin de connaissances sur le routage
  - Facile à démarrer
  - Plus compliqué à distribuer sur une vaste zone
  - Les ISP sont libres de mettre en place des accords d'échange de trafic (peering) entre eux, comme ils le souhaitent

# Couche 2 contre couche 3

## Résumé

- Couche 2 est un **RÉEL** point d'échange internet
- Couche 3 est un concept marketing utilisées par les ISP de Transit
  - N'est pas un IXP réel

# Considérations sur la conception d'IXP

# Conception de point d'échange

- Le noyau de l'IXP est un commutateur Ethernet
- A supplanté tous les autres types de périphériques réseau pour un IXP
  - Du moins cher et plus petit commutateur 12 ou 24 ports 10/100
  - Au plus grand commutateur 192 port 10GigEthernet

# Conception de point d'échange

- Chaque ISP participant à l'IXP apporte un routeur à l'emplacement de l'IXP
- Besoins des routeurs:
  - Un port Ethernet pour se connecter au commutateur IXP
  - Un port WAN pour se connecter au media WAN menant en retour au backbone ISP
  - Être en mesure d'exécuter BGP

# Conception de point d'échange

- Commutateur IXP situé dans un rack d'équipement dédié à IXP
  - Inclut également d'autres équipements opérationnels IXP
- Des routeurs des ISP participants se trouvant dans des racks avoisinants/adjacents
- Des connexions en cuivre (UTP) faits pour des connexions 10 Mbps, 100 Mbps ou 1Gbps
- Fibres utilisées pour 10 Gbps et 40 Gbps

# Échange de trafic (Peering)

- Chaque participant doit exécuter BGP
  - Ils ont besoin d'avoir leur propre numéro AS
    - **ASN** public, **PAS** d'ASN privé
- Chaque participant configure le BGP externe directement avec les autres participants de l'IXP
  - Échanger du trafic avec tous les participants ou bien
  - Échanger du trafic avec un sous-ensemble de participants

# Échange de trafic (plus)

- Echanges de trafic multilatéraux obligatoires (MMLPA)
  - Chaque participant est tenu de s'apparier avec tous les autres participants dans le cadre de leur adhésion IXP
  - **N'a pas d'antécédents de succès** — la pratique est fortement déconseillée
- Echanges de trafic multilatéraux (MLP)
  - Chaque participant s'apparie avec tous les autres participants (généralement aidé par un serveur route)
- Échanges de trafic bilatéraux
  - Les participants mettent en place des échanges de trafic entre eux en fonction de leurs propres besoins et relations commerciales
  - Il s'agit de la situation la plus courante au niveau des IXP aujourd'hui

# Routage

- Les routeurs frontières des ISP au niveau des IXP ne devraient généralement PAS être configurés avec un itinéraire par défaut ou porter la table de routage Internet complète
  - Transporter par défaut ou la table complète signifie que ce routeur et le réseau ISP sont ouverts à des abus par des membres non-peering de l'IXP
  - La configuration correcte consiste à seulement porter des routes offertes à des pairs IXP sur le routeur de peering de l'IXP
- Remarque: Certains ISP offrent du transit à travers des tissus IX
  - Ils font cela à leurs risques et périls - voir plus haut

# Routage (suite)

- Les routeurs frontières ISP au niveau de l'IXP ne doivent pas être configurés pour transporter le réseau LAN de l'IXP au sein de IGP ou iBGP
  - Utilisez le concept next-hop-self de BGP
- Ne pas générer des agrégats préfixes d'ISP sur le routeur d'échange de trafic (peering router) de l'IXP
  - Si la connexion du backbone au routeur de l'IXP s'arrête, le basculement normal de BGP sera alors réussi

# Espace d'adressage

- Certains IXP utilisent des adresses privées pour les LAN IX
  - Un espace d'adressage public signifie que le réseau IXP peut être divulgué à Internet, ce qui peut être indésirable
  - Parce que la plupart des ISP filtre l'espace d'adressage RFC1918, cela évite le problème
- Certains IXP utilisent des adresses publiques pour les LAN IX
  - Espace d'adressage disponible par les RIR
  - Les termes de participation IXP souvent interdisent le LAN IX d'être transporté dans le backbone de membres ISP

# Politique AfriNIC sur l'espace d'adresses IXP

- La politique d'affectations d'utilisateurs finaux répond à l'espace d'adressage public IXP sous infrastructures essentielles
- Cela nécessite que l'IXP soit un IXP de couche 2 avec 3 ISP connectés et avoir une politique claire et ouverte pour l'adhésion
- L'allocation minimum pour une infrastructure essentielle est /24 pour l'IPv4 et /48 pour l'IPv6

# Matériel (Hardware)

- Essayez de ne pas mélanger les vitesses de ports
  - Si des connexions 10 Mbps et 100 Mbps sont disponibles, terminez sur des commutateurs différents (IXP L2)
- Ne mélangez pas les transports
  - Si terminaison ATM PVCs et G/F/Ethernet, terminez sur des appareils divers
- Insister pour que les participants IXP apportent leur propre routeur
  - déplace le problème de tamponnage hors de l'IXP
  - la sécurité est de la responsabilité du ISP, pas de l'IXP

# Services offerts

- Les services offerts ne devraient pas rivaliser avec les ISP membres (IXP de base)
  - par exemple, l'hébergement web à un IXP est une mauvaise idée à moins que tous les membres y consentent
- Les opérations IXP devraient rendre les statistiques de performance et de débit disponibles aux membres
  - Utilisez des outils tels que MRTG pour produire des graphiques de débit IX servant comme informations aux membres (ou au public)

# Services à offrir

- ccTLD DNS
  - l'IXP du pays pourrait accueillir le DNS du plus haut niveau du pays
  - par exemple «SE». TLD est hébergé à Netnod IXes en Suède
  - Offrez la sauvegarde des ccTLD DNS d'autres pays
- Serveur racine
  - Des instances Anycast de I.root-servers.net, F.root-servers.net etc. sont présents à de nombreux IXes
- Nouvelles Usenet
  - Les nouvelles Usenet sont de gros volumes
  - pourrait économiser la bande passante à tous les membres IXP

# Services à offrir

- Collecteur de routes
  - Le collecteur de route affiche les informations d'accessibilité disponibles à l'échange
  - Détails techniques décrites plus loin
- Looking Glass
  - Une façon de rendre les routes du collecteur de route disponibles pour une vision globale (par exemple: [www.traceroute.org](http://www.traceroute.org))
  - Accès public ou réservé aux membres

# Services à offrir

- Redistribution/mise en cache de contenu
  - Par exemple, Akamised service de distribution de mise à jour
- Network Time Protocol
  - Localiser une source temps strate 1 (récepteur GPS, horloge atomique, etc.) à l'IXP
- Registre de routage
  - Utilisé pour enregistrer la politique de routage de l'adhésion IXP (plus, plus tard)

# Introduction aux collecteurs de route

Quelles routes sont disponibles à l'IXP?

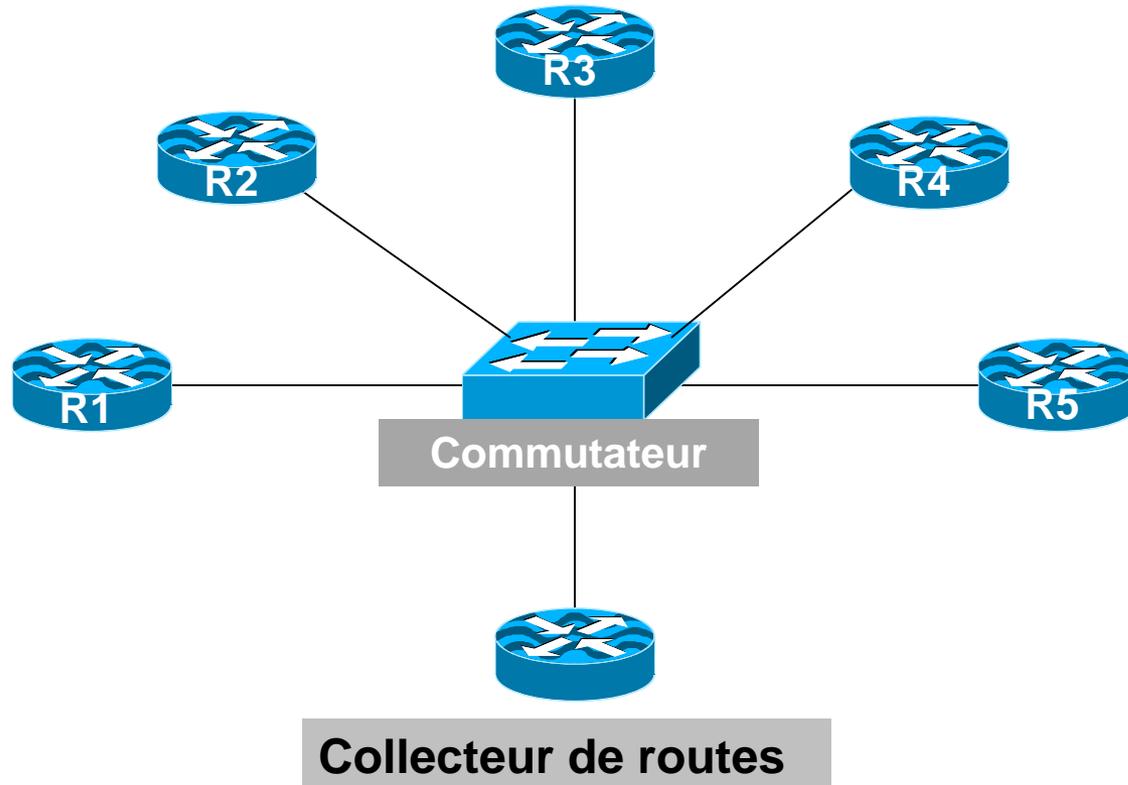
# Qu'est-ce qu'un collecteur de route?

- Habituellement, un routeur ou un système Unix sur lequel tourne BGP
- Recueille des informations de routage des routeurs de fournisseurs de services au niveau d'un IXP
  - S'apparie avec chaque ISP en utilisant BGP
- Ne transmet **pas** de packets
- N'annonce **pas** de préfixes aux ISP

# But d'un collecteur de route

- Pour fournir une vue publique de l'information de routage disponible à l'IXP
  - Utile pour les membres existants pour vérifier la fonctionnalité des filtres BGP
  - Utile pour les membres potentiels à vérifier la valeur de l'adhésion à l'IXP
  - Utile pour les opérations de la communauté Internet à des fins de dépannage
    - Par exemple: [www.traceroute.org](http://www.traceroute.org)

# Collecteur de routes à un IXP



# Exigences du collecteur de routes

- Routeur ou système Unix sur lequel tourne BGP
  - Besoins en mémoire minimum - ne tient que les routes IXP
  - Exigences de transfert de paquets minimales – ne transfère aucun paquet
- Apparie eBGP avec chaque membre de l'IXP
  - Accepte tout, ne donne rien
  - Utilise un ASN privé
  - Se connecte au LAN de Transit de l'IXP
- Connexion "Back end"
  - Deuxième Ethernet globalement routé
  - Connexion au site Web IXP pour un accès public

# La mise en œuvre du collecteur de route

- La plupart des IXP met en œuvre, maintenant, une certaine forme de collecteur de route
- Avantages déjà mentionnés
- Grand outil de relations publiques
- Des exigences non-sophistiquées
  - Simplement exécute BGP

# **Introduction aux Route Servers**

**Comment mettre à échelle de très  
grands IXP**

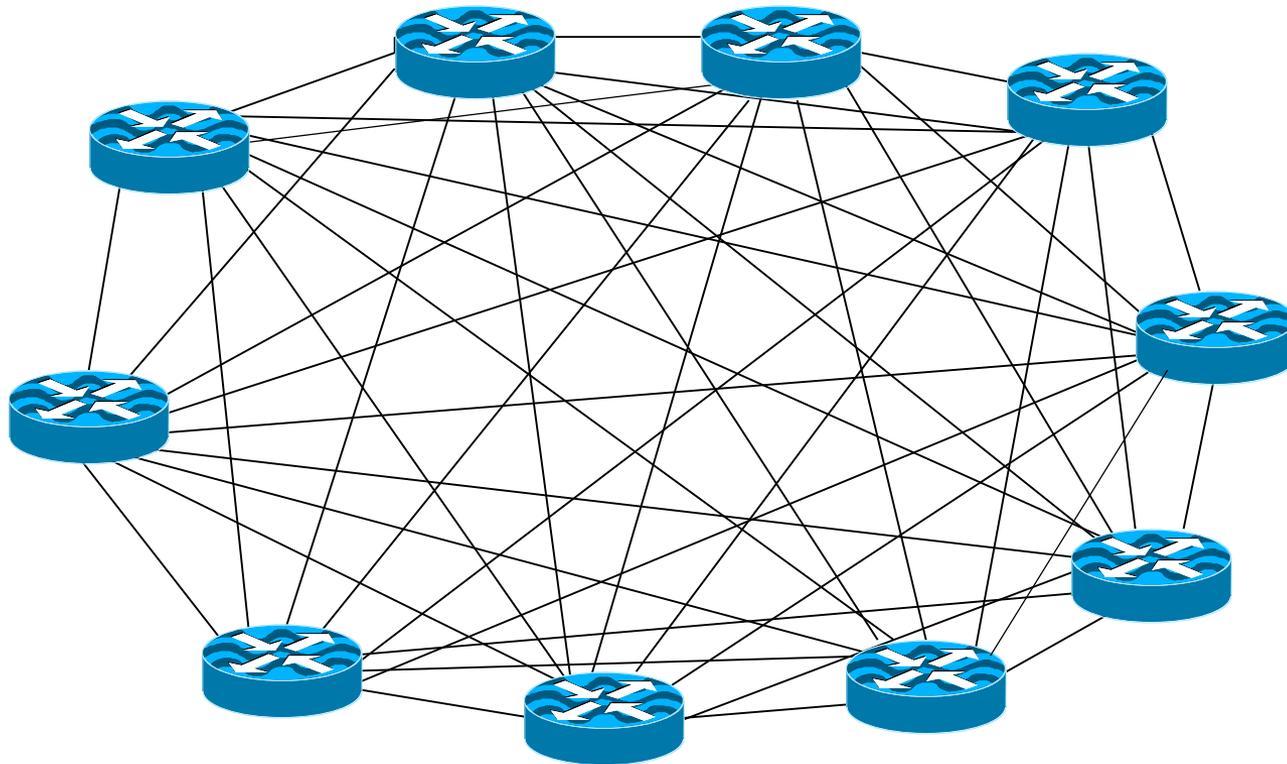
# Qu'est-ce qu'un serveur de route?

- Possède tous les caractéristiques d'un collecteur de Route
- Mais aussi:
  - Annonce des routes aux participants membres de l'IXP selon leurs définitions de politique de routage
- Mis en œuvre en utilisant les mêmes spécifications que pour un collecteur de Route

# Caractéristiques d'un serveur Route

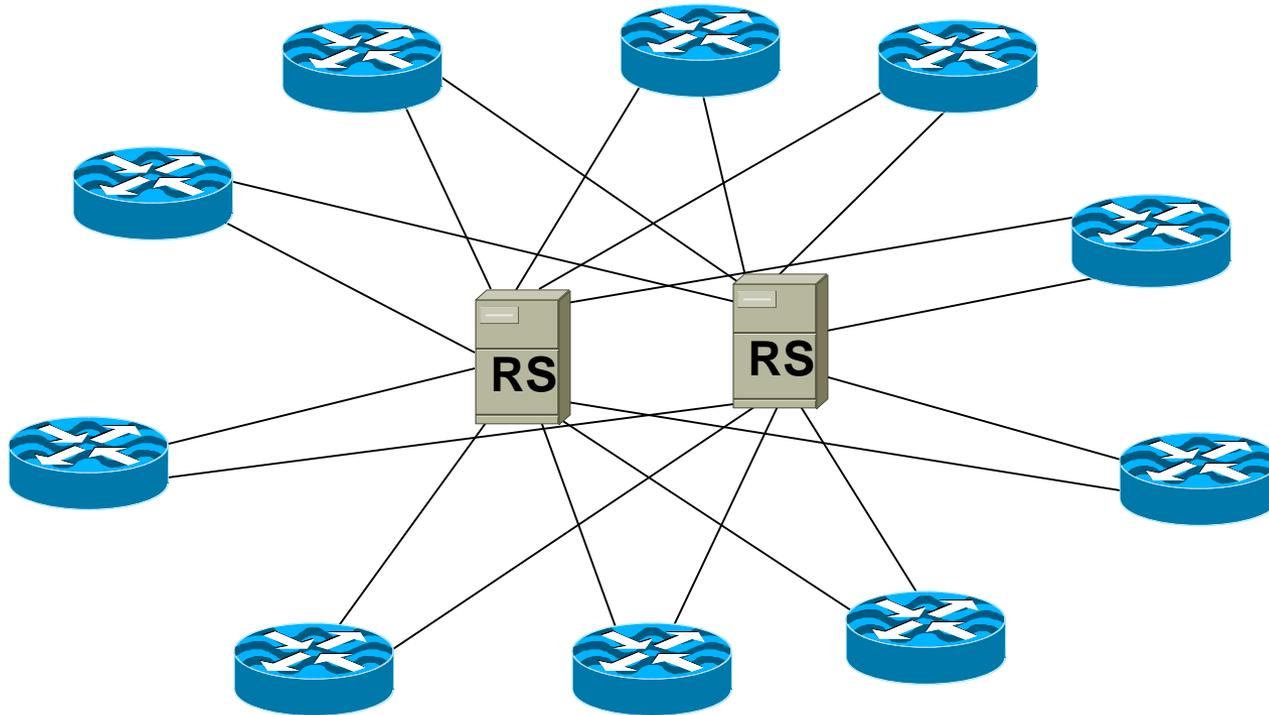
- Aide à mettre à échelle le routage pour de grands IXP
- Simplifie le processus de routage sur les routeurs ISP
- Participation facultative
  - Fourni en tant que service, n'est **PAS** obligatoire
- Résulte dans l'insertion du numéro de système autonome RS dans l'itinéraire de routage
- Utilise facultativement la politique inscrite dans IRR

# Schéma de N-squared Peering Mesh



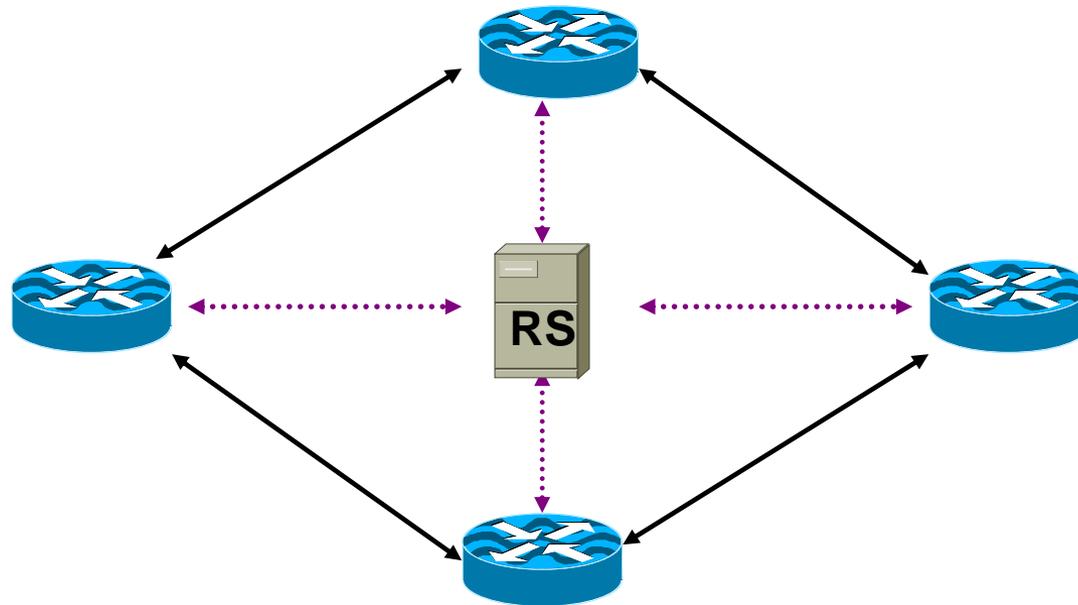
- Pour des IXP de grande taille (plusieurs dizaines de participants), le maintien d'un maillage peering (peering mesh) plus grand devient encombrant et souvent trop difficile

# Mesh avec des Route Servers



- Les routeurs ISP s'apparient avec les serveurs de Route
  - Seulement besoin d'avoir deux sessions eBGP plutôt que N

# Flux de routage de point d'échange basé sur RS



**FLUX DE TRAFIC**



**FLUX D'INFORMATION DE ROUTAGE**

# Avantages de l'utilisation d'un serveur de Route

- Aide à mettre à échelle le routage pour de très larges IXP
- Séparation entre le routage et le transfert
- Simplifier la gestion de configuration du routage sur les routeurs ISP

# Inconvénients de l'utilisation d'un serveur de Route

- Les ISP peuvent perdre le contrôle direct des politiques
  - Si RS est seulement pair, les ISP n'ont pas de contrôle sur les destinataires de la distribution de leurs préfixes
- Complètement dépendant sur une partie tierce
  - Configuration, dépannage etc.
- Insertion de RS ASN dans l'itinéraire de routage
  - L'ingénierie de trafic/multihoming a besoin de plus de soins
- Ceux-ci sont des inconvénients majeurs
  - Généralement l'emportent sur les avantages

# L'usage typique d'un serveur Route

- Les serveurs Route peuvent être fournis comme un **service OPTIONNEL**
  - Très courante chez les IXP de grande taille (> 50 participants)
  - Exemples: LINX, TorIX, AMS-IX, etc
- Pair d'ISP (ISPs peer)
  - Directement avec des pairs significatifs
  - Avec le serveur Route pour le reste

# Matière à réflexion ...

- L'utilisation de serveur de Route serait-elle bénéfique pour vous?
  - Utile lorsque les connaissances BGP sont limitées (mais ce n'est pas une excuse pour ne pas apprendre BGP)
  - Évite d'avoir à maintenir un grand nombre de pairs eBGP
  - Mais pouvez-vous vous permettre de perdre le contrôle politique? (Un ISP sans contrôle de ses politiques de routage, c'est quoi?)

# Qu'est-ce qui peut mal tourner

...

Les différentes façons par lesquelles  
les opérateurs d'IXP nuisent à leur  
IXP ...

# Qu'est-ce qui peut mal tourner ?

## Le concept

- Certains fournisseurs de services tentent de tirer profit de la réputation des IXP
- Vendre les services de transit Internet comme "Point d'Echange Internet"
  - "Nous échangeons des paquets avec les autres ISP, nous sommes donc un Point d'Echange Internet!"
  - Les soi-disant Échanges de Couche-3 – vraiment des fournisseurs de transit Internet
  - Des routeurs utilisés au lieu de commutateurs
  - Exemple le plus célèbre: SingTelIX

# Qu'est-ce qui peut mal tourner ?

## La compétition

- Trop de points d'échange dans une même localité
  - Les échanges concurrents défont l'objectif
- Devient coûteux pour les ISP de se connecter tous entre eux
- Un IXP:
  - n'est **PAS** une compétition
  - n'est **PAS** une entreprise à but lucratif

# Qu'est-ce qui peut mal tourner ?

## Règles et restrictions

- Les IXP essayent de rivaliser avec leurs membres
  - Offrir des services que les ISP proposeraient/proposent à leurs clients
- Les IXP fonctionnent comme un club privilégié fermé, par exemple:
  - Critères d'adhésion restrictifs (closed shop)
- Des IXP donnant accès à des utilisateurs finaux plutôt qu'aux fournisseurs de services seulement
- Des IXP interférant avec les décisions d'affaires des ISP, par exemple les échanges de trafic multilatéraux obligatoires

# Qu'est-ce qui peut mal tourner?

## Erreurs de conception technique

- Des IXP interconnectés
  - Un IXP dans une localité estime qu'il doit se connecter directement à l'IXP dans une autre localité
  - Qui paie pour l'interconnexion?
  - Comment le trafic est-il mesuré?
  - Entre en concurrence avec les ISP qui offrent déjà un transit entre les deux localités (qui ensuite refusent de se joindre à IX, nuisant ainsi à la viabilité de IX)
  - Les interconnexions Metro marchent bien (par exemple LINX)

# Qu'est-ce qui peut mal tourner?

## Erreurs de conception technique

- Les ISP font le pont entre le LAN IXP et leurs bureaux
  - “Nous sommes pauvres, nous ne pouvons pas nous permettre un routeur”
  - Les avantages financiers de se connecter à un IXP l'emportent largement sur le coût d'un routeur
  - En réalité, ça permet aux ISP de brancher n'importe quel périphérique sur le LAN de l'IXP – avec des conséquences désastreuses pour la sécurité, intégrité et fiabilité de l'IXP

# Qu'est-ce qui peut mal tourner ?

## Erreurs de conception de routage

- Serveur de Route mis en œuvre dès le premier jour
  - Les ISP n'ont aucune incitation pour apprendre BGP
  - Par conséquent, n'ont aucune incitation pour comprendre les relations, les politiques d'échange de trafic, &c
  - Entièrement dépendant sur les opérateurs RS pour dépannage, configuration, fiabilité
    - RS ne peut pas être opéré par un comité
- Le serveur de Route, c'est pour aider à faire évoluer (scale) l'échange de trafic au niveau de LARGES IXP

# Qu'est-ce qui peut mal tourner ?

## Erreurs de conception de routage

- Des réflecteurs de Route iBGP utilisés pour distribuer des préfixes entre les participants IXP
- Avantage revendiqué (1):
  - Les participants n'ont pas à connaître ou à exécuter BGP
- En fait, un désavantage
  - L'opérateur IXP doit connaître BGP
  - Ne pas connaître BGP pour un ISP est un grand inconvénient commercial
  - Les ISP qui souhaitent avoir des affaires prospères doivent être en mesure de faire du multi-home, de s'apparier avec d'autres ISP etc. – ces activités nécessitent BGP

# Qu'est-ce qui peut mal tourner ?

## Erreurs de conception de routage

### (suite)

- Réflecteurs de Route-Avantage revendiqué (2):
  - Permet à un IXP d'être commencé très rapidement
- Les faits:
  - IXP n'est qu'un commutateur Ethernet – mettre en place une maille iBGP avec des participants n'est pas plus rapide que de mettre en place une maille eBGP

# Qu'est-ce qui peut mal tourner ?

## Erreurs de conception de routage (suite)

- Réflecteurs de Route-Avantage revendiqué (3):
  - L'opérateur IXP a un contrôle total sur les activités IXP
- En fait, un désavantage
  - Les participants d'ISP abandonnent le contrôle de:
    - Leur routeur frontière ; il est situé dans l'AS de l'IXP
    - Leurs politiques de routage et d'échange de trafic
  - L'opérateur IXP est le point unique d'échec
    - S'ils ne sont pas disponibles 24 heures sur 24, 7 jours sur 7, alors l'IXP ne le sera pas non plus
    - Les erreurs de configuration BGP commises par les opérateurs IXP ont de réels impacts sur les opérations ISP

# Qu'est-ce qui peut mal tourner ?

## Erreurs de conception de routage

### (suite)

- Réflecteurs de Route-Inconvénient (4) :
  - La migration depuis un réflecteur Route pour «corriger» la configuration de routage est hautement non trivial
  - Le routeur ISP est dans l'ASN de l'IXP
    - Besoin de déplacer le routeur ISP de l'ASN de l'IXP vers l'ASN de l'ISP
    - Nécessaire de reconfigurer BGP sur routeur ISP, ajouter à l'IGP et à la maille iBGP de l'ISP, et mettre en place eBGP avec les participants IXP et/ou le serveur Route de l'IXP

# Plus d' Informations

# Point d'échange Politiques

- Les AUP
  - Politique d'utilisation acceptable
  - Règles minimales pour la connexion
- Les frais?
  - Certains IXP ne font payer aucun frais
  - D'autres IXP font payer le recouvrement de leurs coûts
  - quelques IXP sont de nature commerciale
- Personne n'est obligé de s'apparier (peer)
  - Accords laissés aux ISP, pas mandatés par IXP

# L'étiquette Point d'Echange

- Ne pas pointer une route par défaut vers un autre participant IXP
- Soyez conscient des next-hop tiers
- Annoncer seulement vos itinéraires globaux
  - Lire RIPE-399 en premier  
[www.ripe.net/docs/ripe-399.html](http://www.ripe.net/docs/ripe-399.html)
- Filtrez! Filtrez! Filtrez!
  - Et faites du reverse path check

# Exemples de point d'échange

- LINX à Londres, Royaume-Uni
- TorIX à Toronto, Canada
- AMS-IX à Amsterdam, Pays-Bas
- SIX à Seattle, Washington, États-Unis
- PA-IX à Palo Alto, Californie, États-Unis
- JPNAP à Tokyo, Japon
- DE-CIX à Francfort, Allemagne
- HK-IX à Hong Kong
- ...
- Tous utilisent des commutateurs Ethernet

# Caractéristiques des IXP (1)

- Redondance et fiabilité
  - Plusieurs commutateurs, UPS
- Le soutien
  - Le NOC pour fournir un support 24h/24, 7j/7 face aux problèmes à l'échange
- DNS, Collecteur de Route , Contenu et serveurs NTP
  - ccTLD et serveurs racines
  - Systèmes de redistribution de contenu tel que Akamai
  - Collecteur de Route - vue de table de routage

# Caractéristiques des IXP (2)

- Emplacement
  - Installations de co-location neutres
- Espace d'adressage
  - LAN d'échange de trafic
- Numéro AS
  - Si vous utilisez Collecteur de Route / Serveur
- Serveurs de Route (en option, pour les grands IXP)
- Statistiques
  - Les données de trafic - pour l'adhésion

# Plus d'info sur les IXP

- <http://www.pch.net/documents>
  - Une autre excellente ressource sur les emplacements d'IXP, documents, statistiques IXP etc.
- <http://www.telegeography.com/ee/ix/index.php>
  - Une collection d'IXP et de points d'interconnexion pour les ISP

# Résumé

- L2 IXP - le plus répandu
  - Le noyau est un commutateur Ethernet
  - ATM et d'autres technologies plus anciennes sont obsolètes
- L3 IXP - est de nos jours un concept de marketing utilisé par les fournisseurs grossistes de services Internet
  - N'offre pas la même flexibilité que L2
  - Non recommandé à moins qu'il y ait des raisons impérieuses de réglementation ou de politique
  - **Évitez !**



GRAND DUCHY OF LUXEMBOURG  
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa  
Infrastructure Trust Fund

# Conception du Point d'Echange Internet (IXP)

FIN

