

Architecture de Réseaux Redondants

AfNOG 2011

Jean Robert Hountomey

“Le technicien de surface a tiré la prise...”

- Pourquoi avait-il accès près de l'équipement?
- Pourquoi ne s'en est on aperçu qu'après?
- Pourquoi cela a pris 6 semaines ?
- Pourquoi l'alimentation n'était pas sécurisée ?
- Pourquoi le réseau n'était pas redondant?



Design Réseau et Architecture



- **...cela peut être critique**
- **...cela peut contribuer au succès du réseau**
- **... cela peut contribuer à sa faillite**

La loi de Ferguson en Architecture Réseau



**“No amount of magic knobs will
save a sloppily designed network”**



**Paul Ferguson—Consultant,
Cisco Systems**

Qu'est ce qu'un réseaux bien architecturé



- **Principaux facteurs à prendre en considération :**
 - Infrastructure physique**
 - Topologie/protocole hiérarchique**
 - Redondance**
 - Agrégation d'adresses (IGP et BGP)**
 - Dimensionnement**
 - Implémentation de politique (cœur/périphérie)**
 - Management/maintenance/exploitation**
 - Coût**

Un tabouret à trois pieds

- **Design de réseau en pensant à la résilience**
- **Utiliser la technologie pour identifier et supprimer les points faibles**
- **Mettre des procédures en place pour diminuer les risques d'erreurs humaines**
- **Tous ces éléments sont nécessaires et interagissent**



Design



Technologie

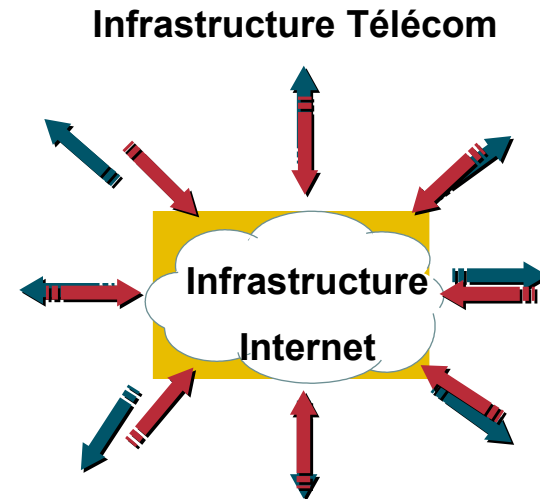


Procédure

Vieux Monde contre Nouveau monde



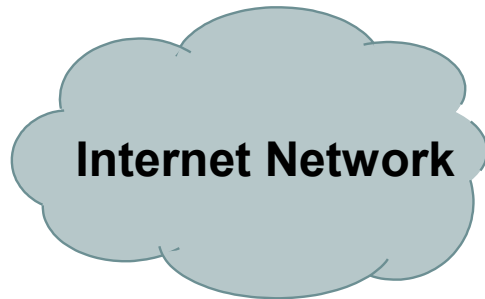
- En dépit du changement de relation Client-Fournisseur, les bases de la construction d'un réseau n'ont pas changées
- Il y a des leçons apprises en 100 ans d'expérience que les **ISPs** peuvent apprendre des **Opérateurs Télécom** et les **Opérateurs** peuvent apprendre de l'expérience de croissance de +100% par an acquise par les **ISPs**



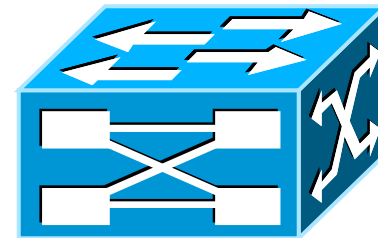
Vieux Monde contre Nouveau monde



- **Internet/réseaux niveau 3**
Construit la redondance dans le **système**
- **Opérateurs voix et réseaux niveau 2**
La boîte est redondante



contre



Comment y arrive-t-on ?

“In the Internet era, reliability is becoming something you have to build, not something you buy. **That's hard work, and it requires intelligence, skills and budget. Reliability is not part of the basic package.”**

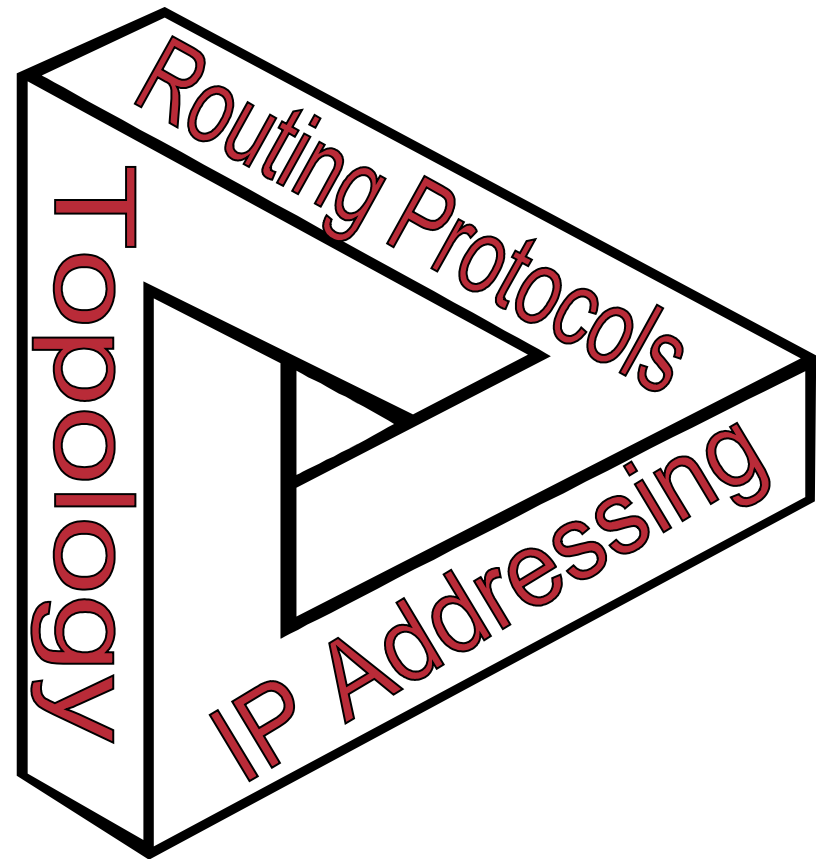
Joel Snyder – Network World Test Alliance 1/ 10/00
“Reliability: Something you build, not buy”



**Outils conceptuels pour
réseaux ISP qui affectent la
topologie**

Concepts de base de scalabilité pour ISP

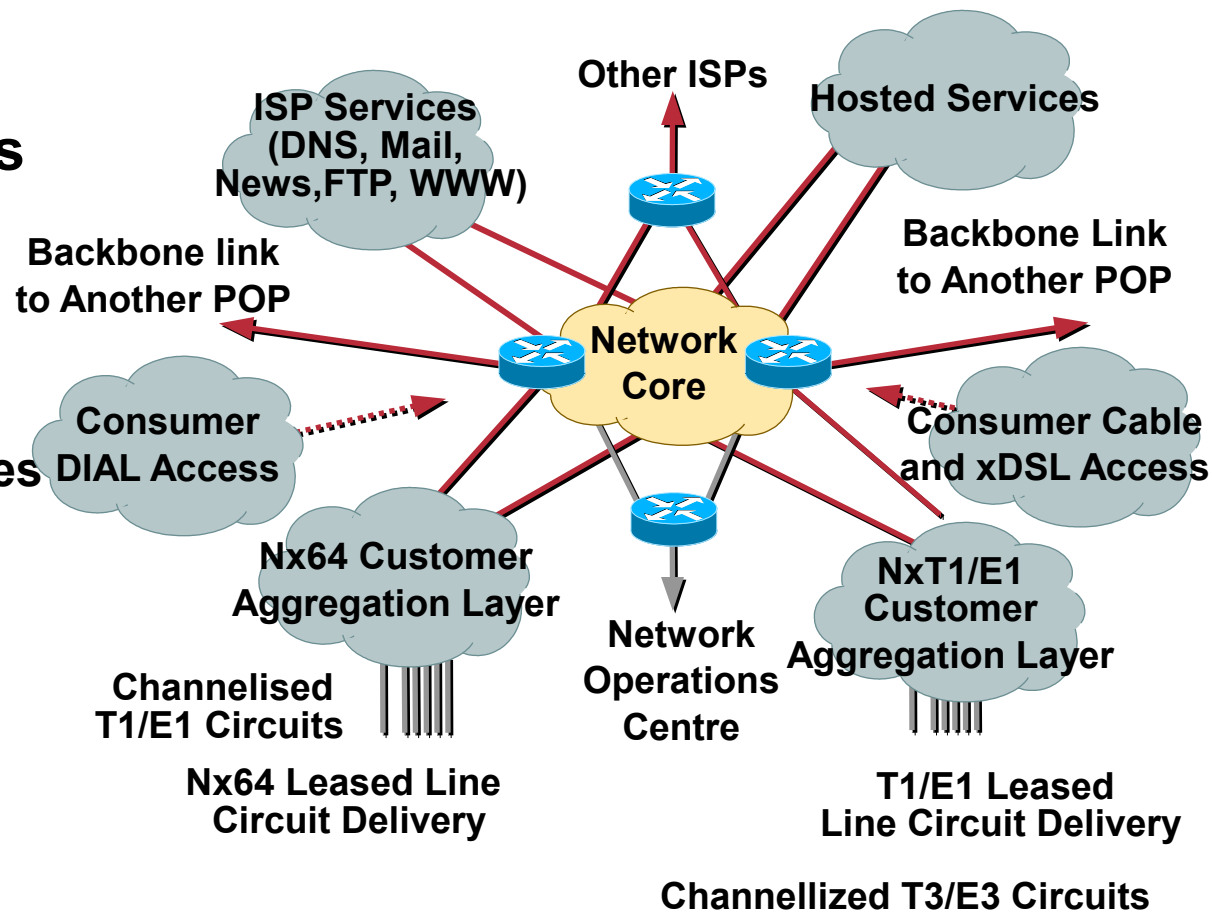
- Design Modulaire et Structuré
- Design Fonctionnel
- Design par tiers/hiérarchique



Design modulaire et structuré

- Organiser le réseau en modules séparés et and répliquables

- Coeur
- POP
- Hosting services
- ISP Services
- Support/NOC



Design modulaire et structuré

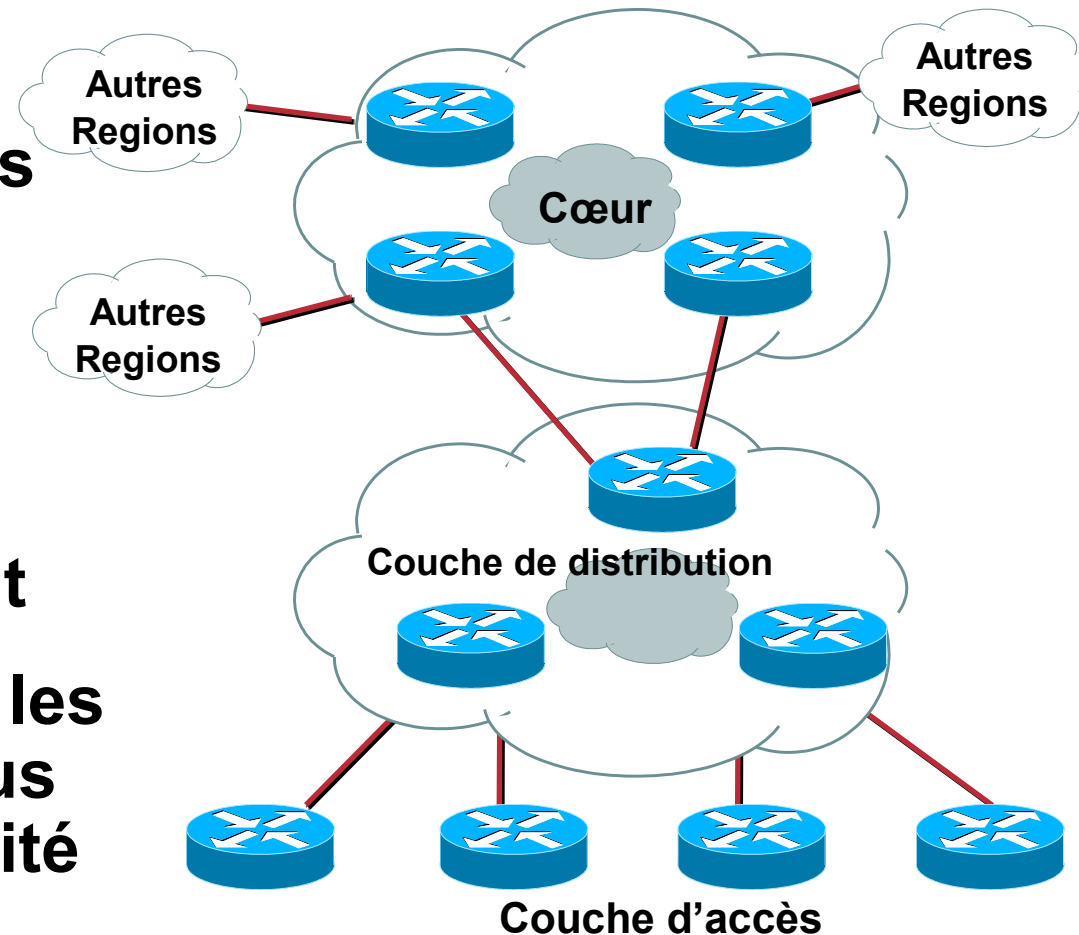
- **La modularité rend un réseau plus dimensionnable**
 - **Design de petite unité de réseau qui sont branchées les unes aux autres**
 - **Chaque module est construit pour une fonction spécifique**
 - **Upgrader consiste à redimensionner un seul module, pas le réseau**

Design Fonctionnel

- **Une boîte ne peut pas tout faire**—(même si des gens ont cherché à le faire)
- Chaque router/switch dans le réseau a une fonction bien définie
- Les différentes **boîtes** interagissent ensemble
- Les équipements sont sélectionnés et fonctionnellement placés dans le réseau en fonction de leurs points forts

Design réseau par tiers et hiérarchique

- **Plat**—les topologies maillées ne scale pas
- **La hiérarchie est utilisée pour le dimensionnement**
- **Bon concept, mais les contours sont plus flous dans la réalité**



Multiple niveaux de redondance

- **Redondance de POP à 3 niveaux**

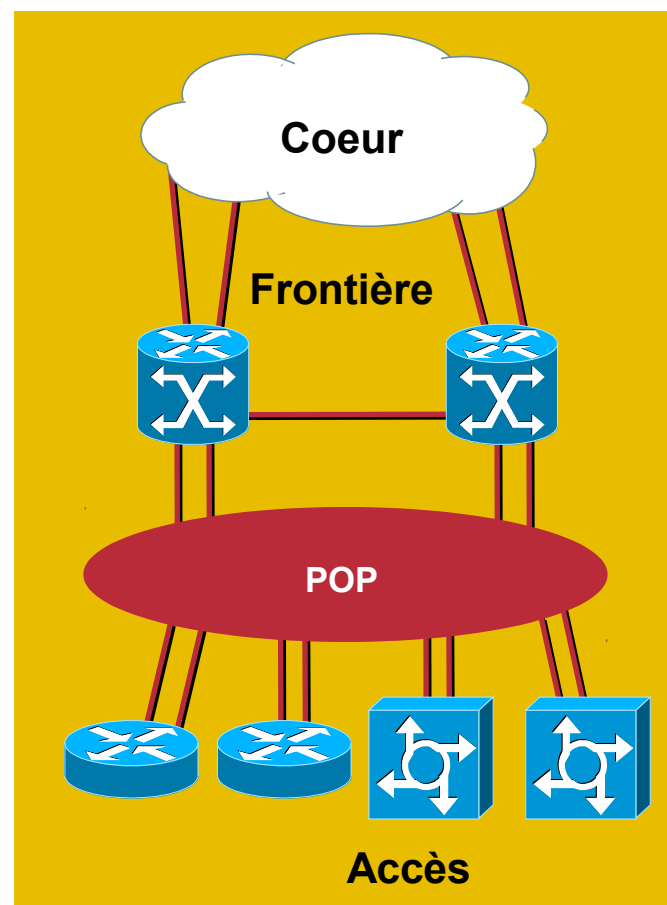
Les faillites de bas niveaux sont supportables

Les faillites de bas niveau déclenchent des faillites de haut niveau

L2: deux de chaque

L3: IGP et BGP fournissent redondance et partage de charge

L4: TCP re-transmissions recovers during the fail-over



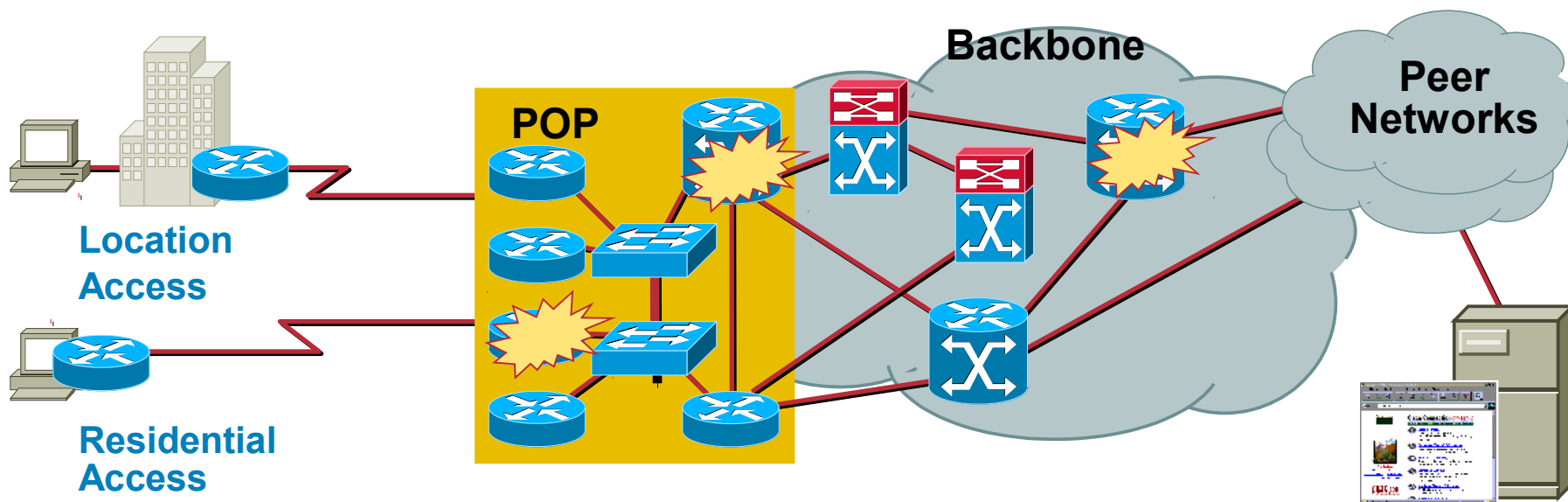
Multiple niveaux de redondance

- **Objectifs**

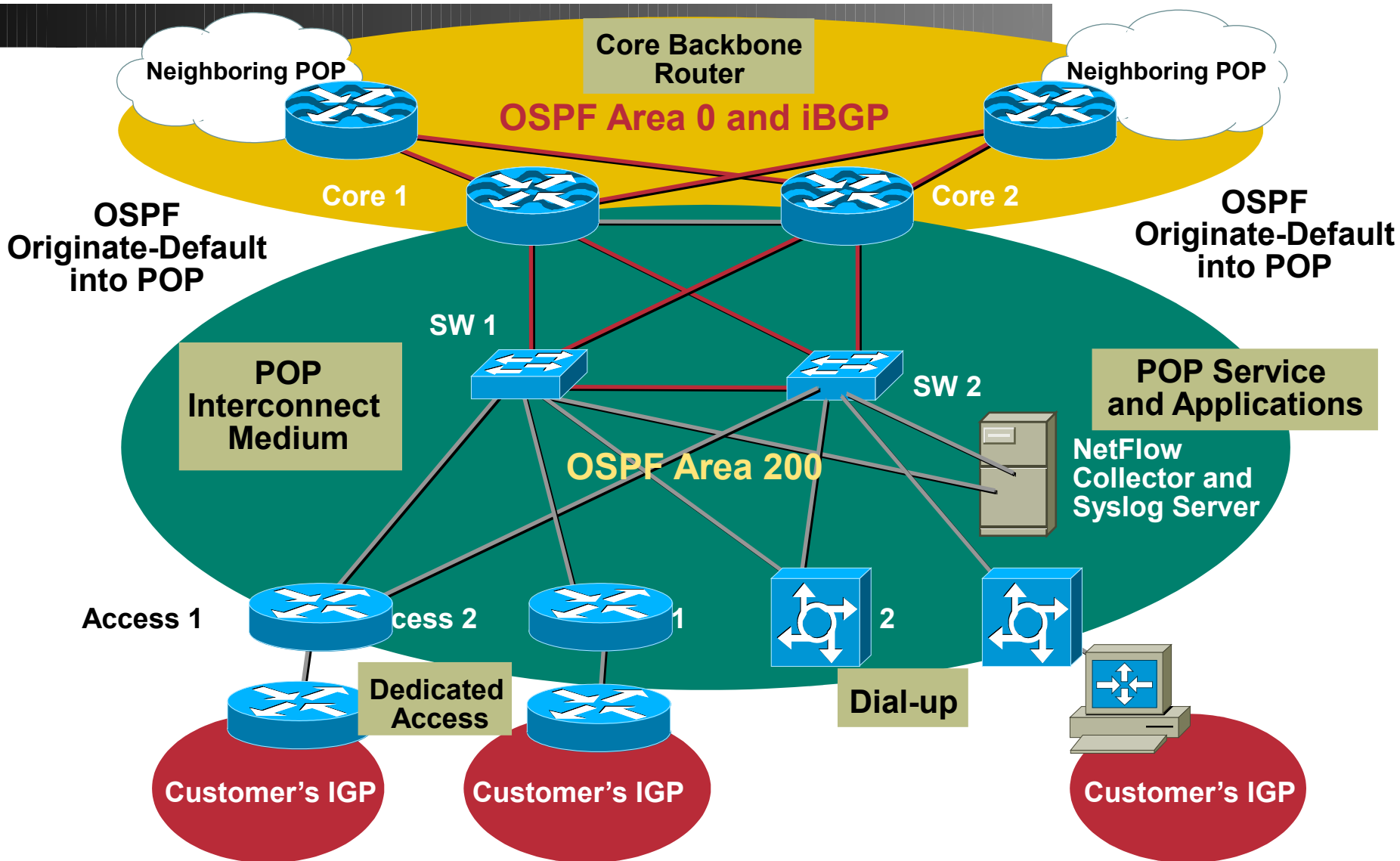
Impacter le moins possible le client final

Minimiser l'impact des fautes dans n'importe quelle partie du réseau

Le réseau doit résister à des fautes de niveau 2, 3, 4 et à des crash routeurs



Multiple niveaux de redondance



Design et Technologie

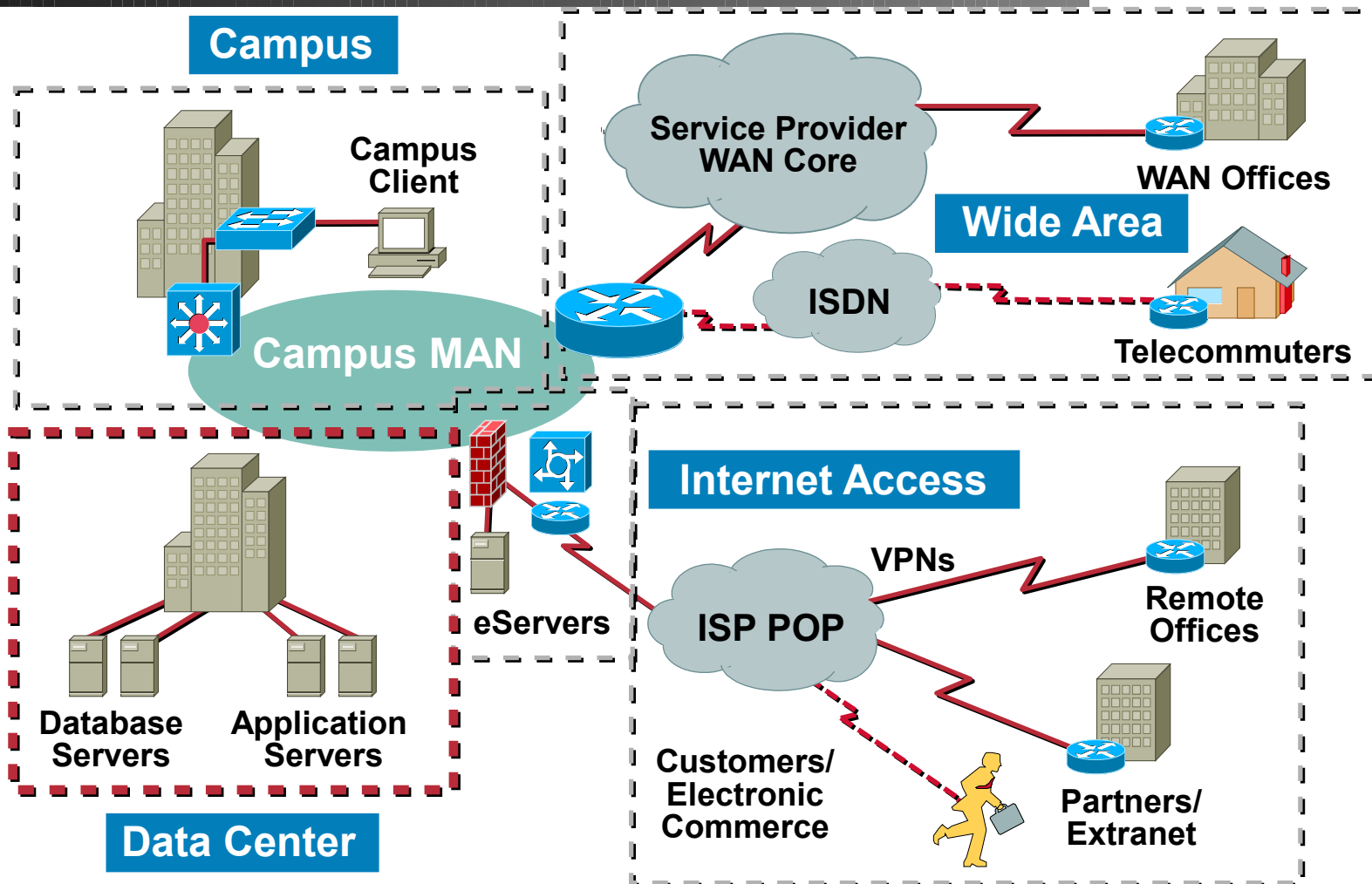


Les bases : Machines et Environnement



- **Courant sécurisé**
- **Refroidissement sécurisé**
- **1:1 or N:1 redondance de cartes**
- **Redondance de processeurs**
- **Redondance de fond de panier**
- **Contrôle de l'environnement**
- **Câblage**

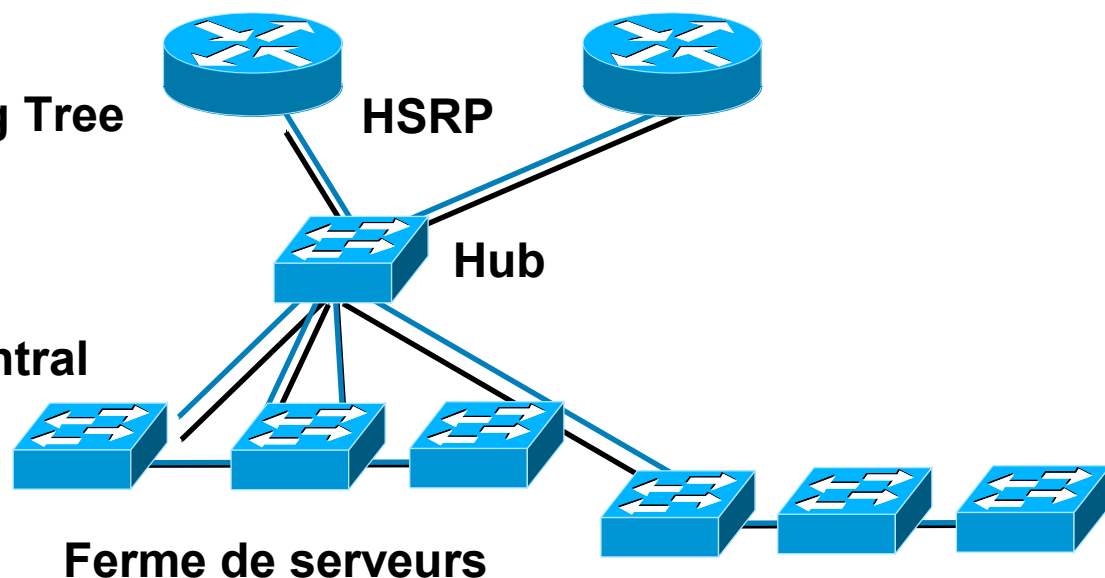
Disponibilité du Data Center



Une mauvaise architecture

Un maillon faible

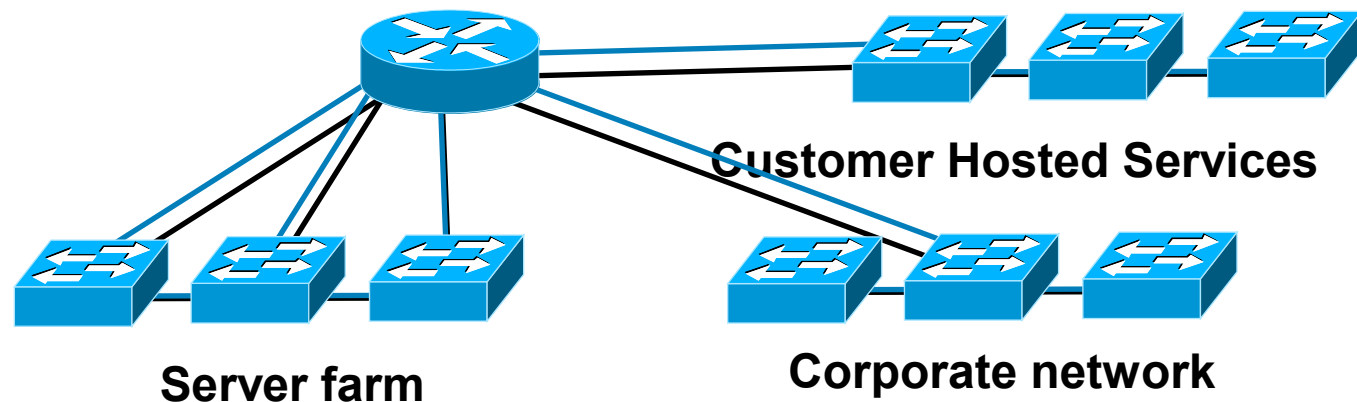
- Un domaine de collision
- Un domaine de sécurité
- Convergence du Spanning Tree
- Pas de traceroute
- Pas de backup
- Performance du switch central
- Où est le firewall ?



Reseau d'entreprise

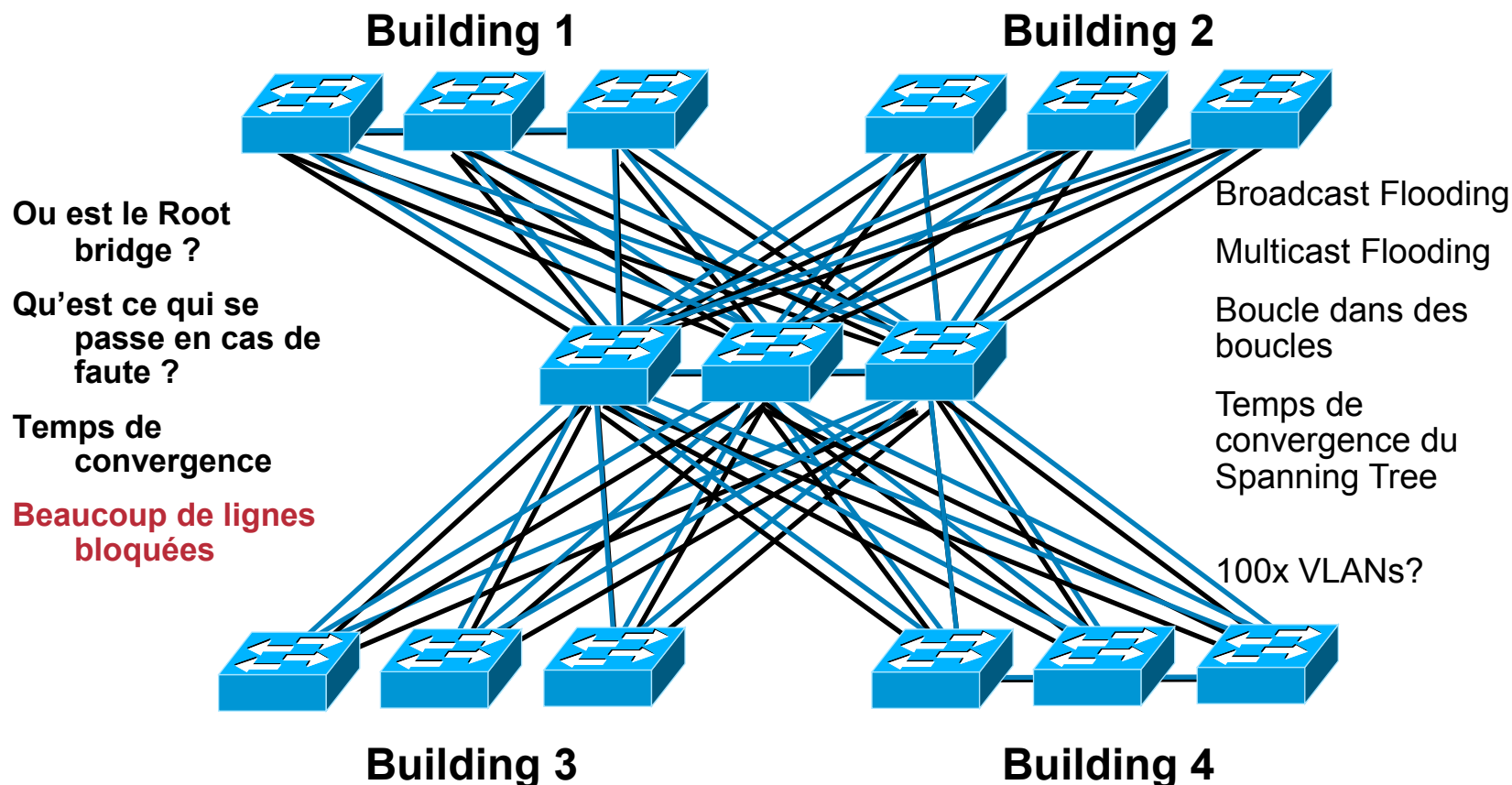
Une autre ...

- Simple à construire
- La résilience est le problème du fabricant
- Plus cher
- Aucun routeur n'est résilient aux fautes logicielles
- Vous avez toujours besoin d'un plus gros routeur

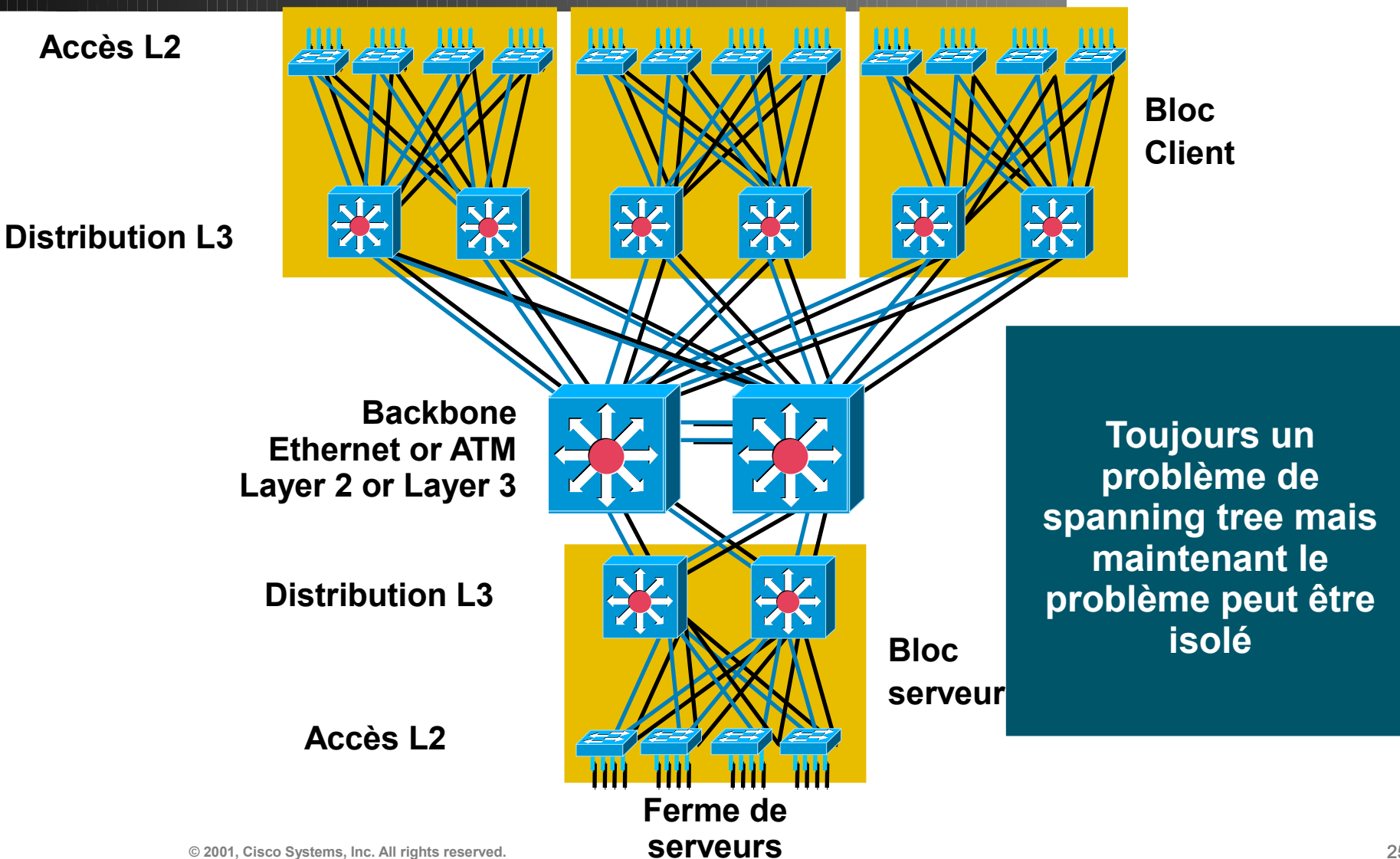


Encore pire

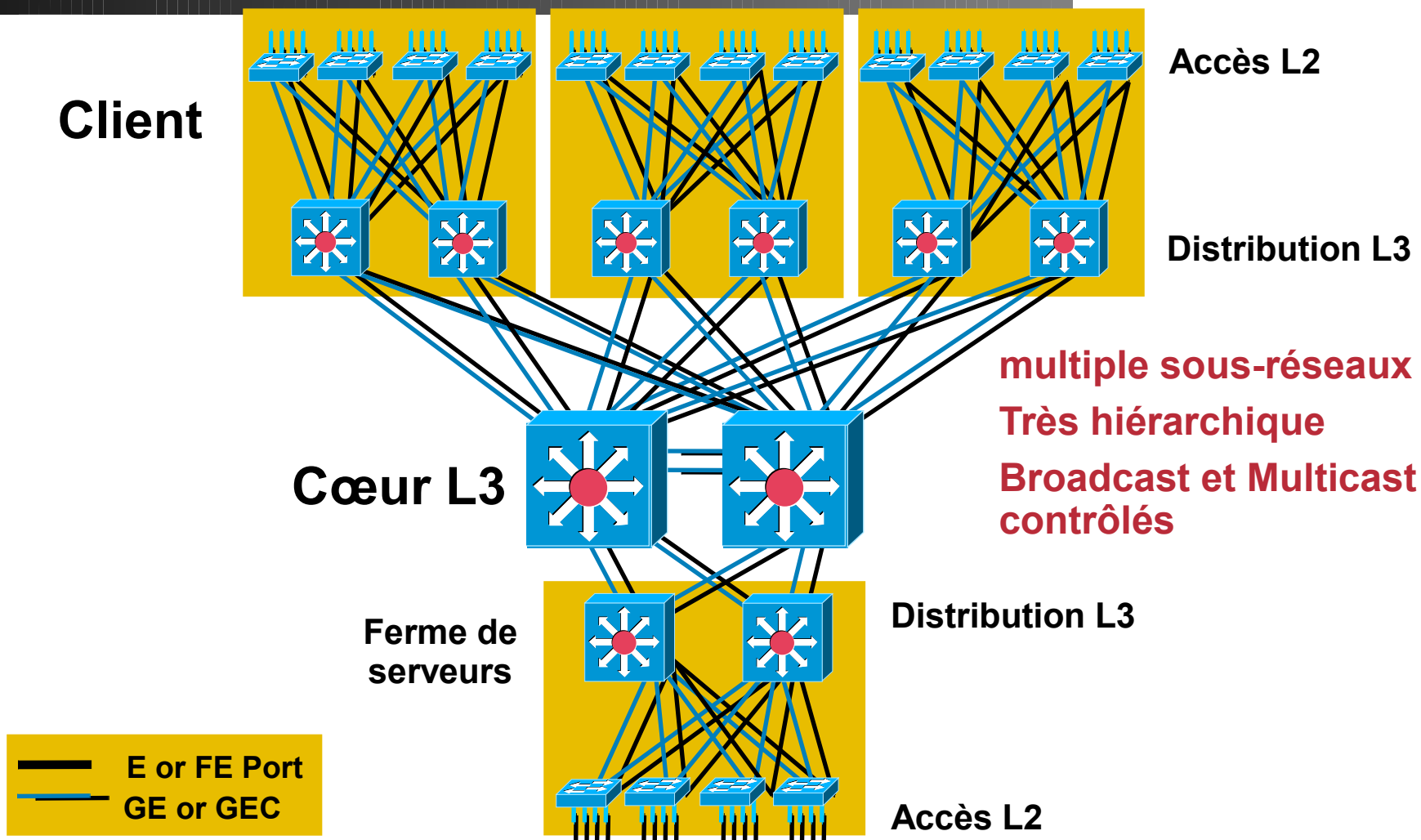
Éviter les réseaux niveau 2 Très maillé, Non-Déterministe



Un meilleur cœur de réseau



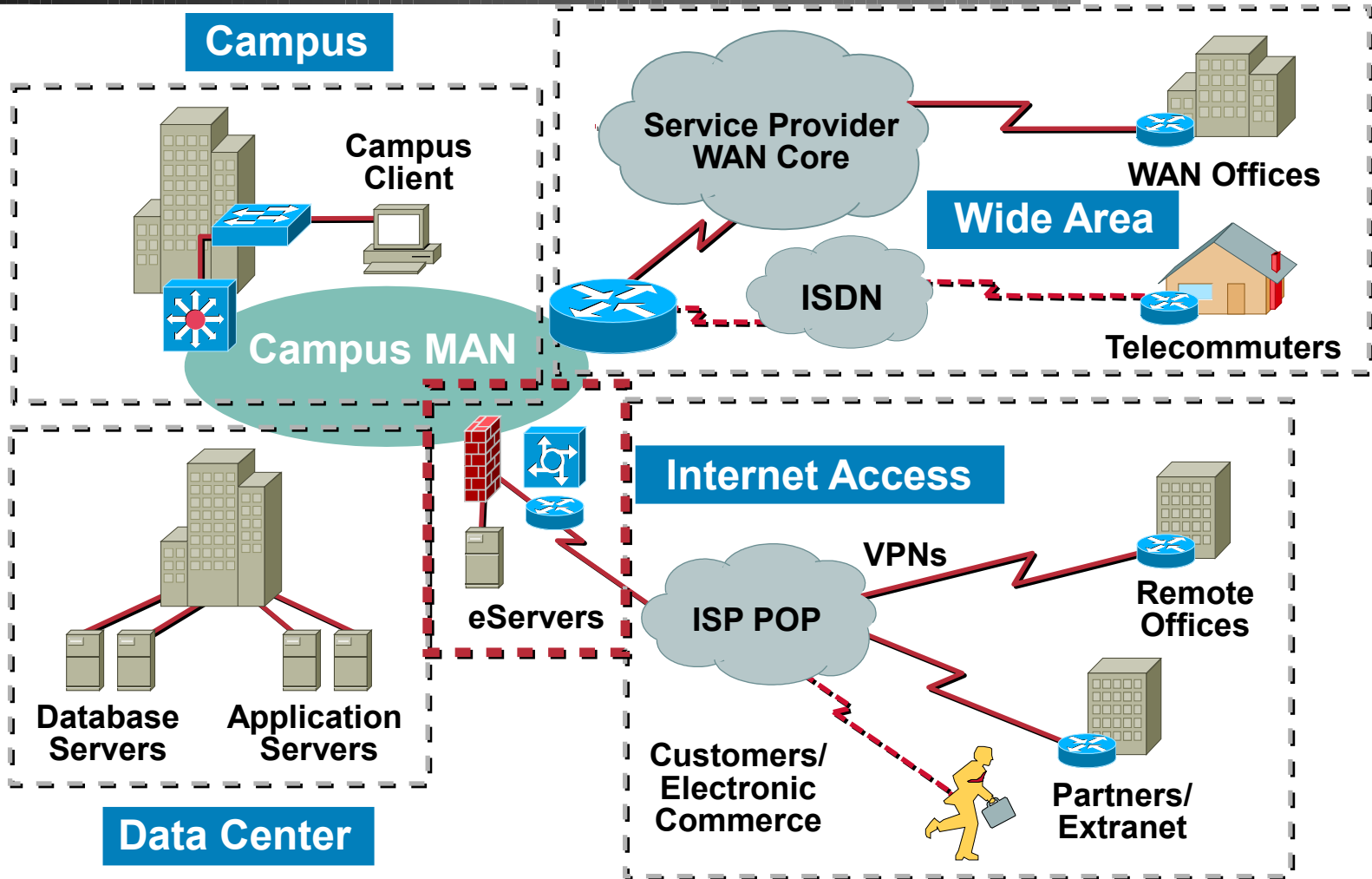
La meilleur architecture



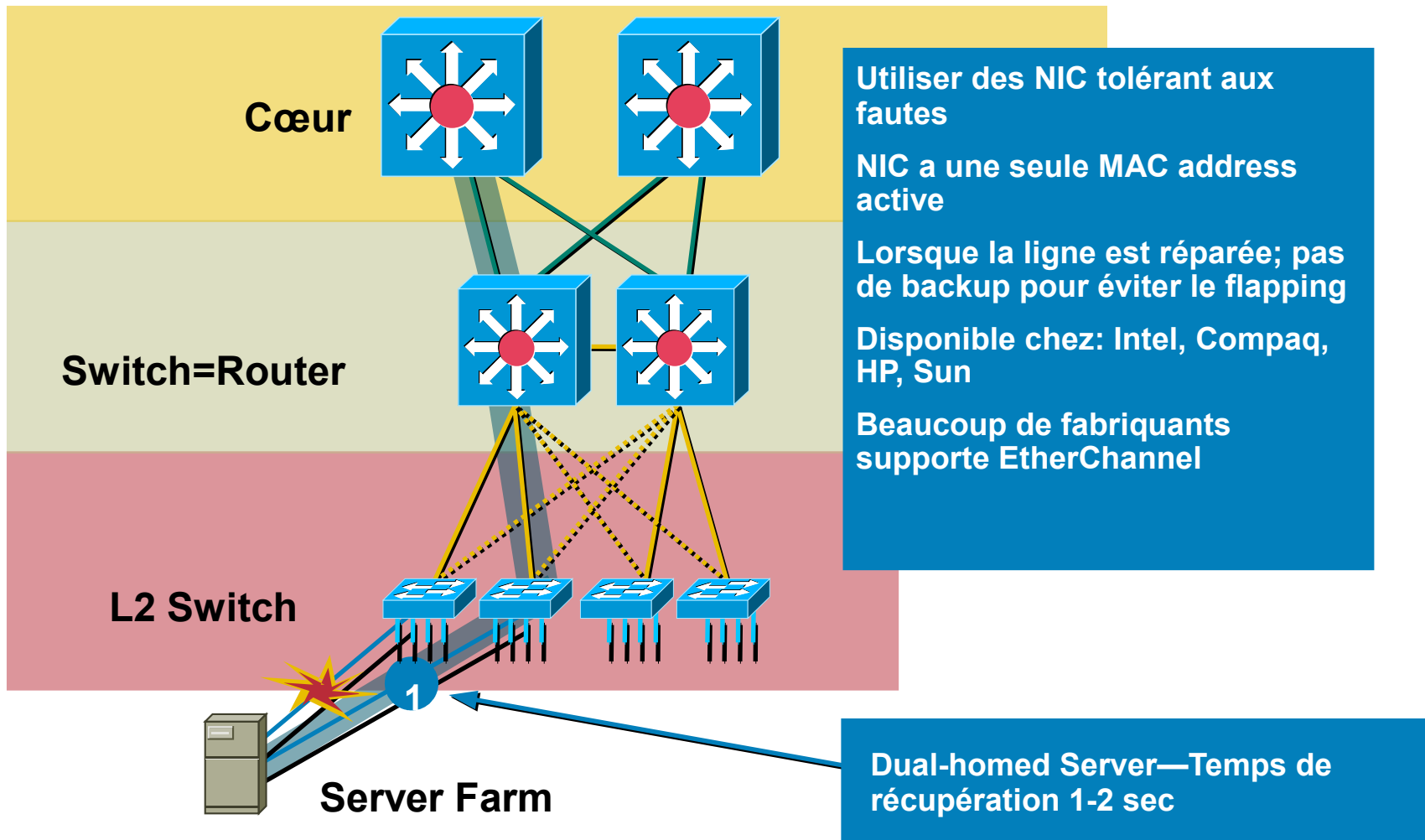
Avantage d'un backbone de niveau 3

- **Control du trafic multicast PIM**
- **Partage de charge**
- **Pas de liens bloqués**
- **Convergence rapide EIGRP/OSPF**
- **Meilleure scalabilité globale**
- **Adjacences de routers diminuées**

Disponibilité de serveur



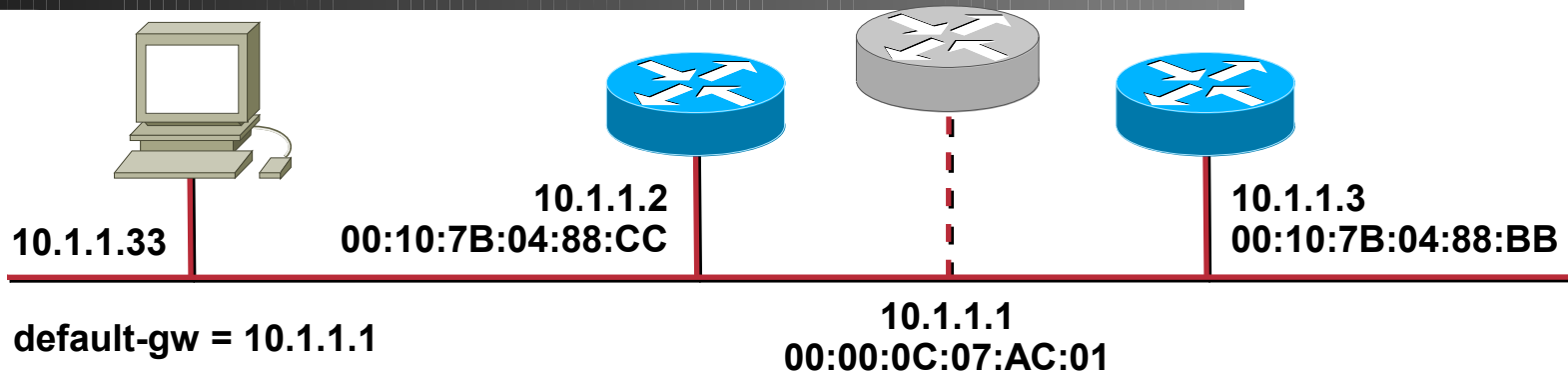
Serveurs multi-homed





HSRP—Hot Standby Router Protocol(RFC2281)

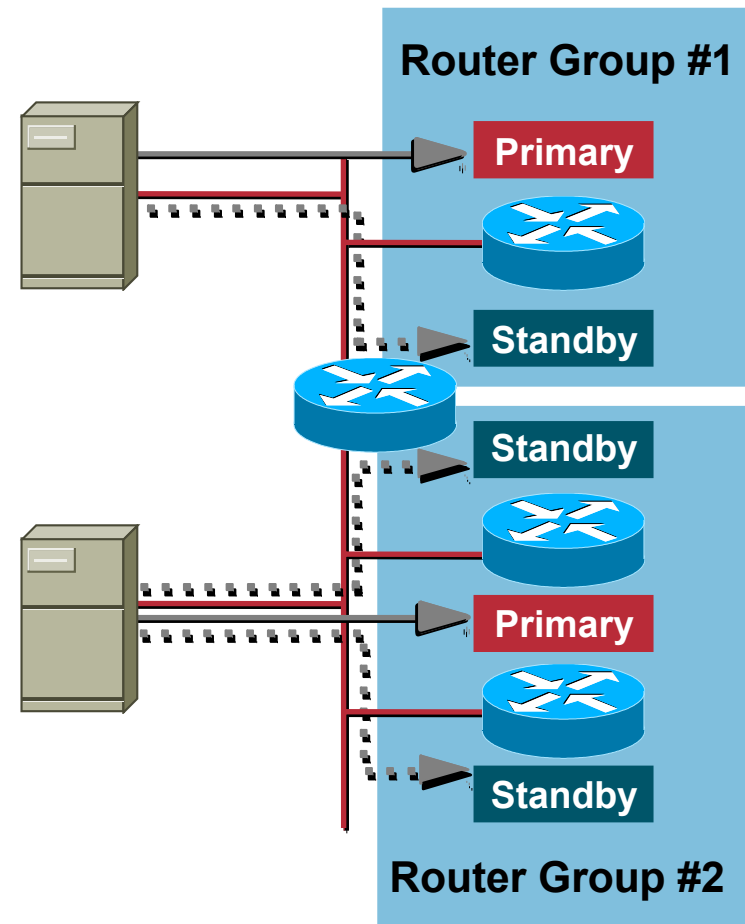
VRRP ----Virtual Router redundancy protocol (RFC3768) Technologie



- Failover transparent du routeur par défaut
- Création d'un routeur fantôme
- Un routeur est actif, il répond aux adresses fantômes de niveau 2 et 3
- L'autre surveille et prend le relais des l'adresses fantômes

HSRP

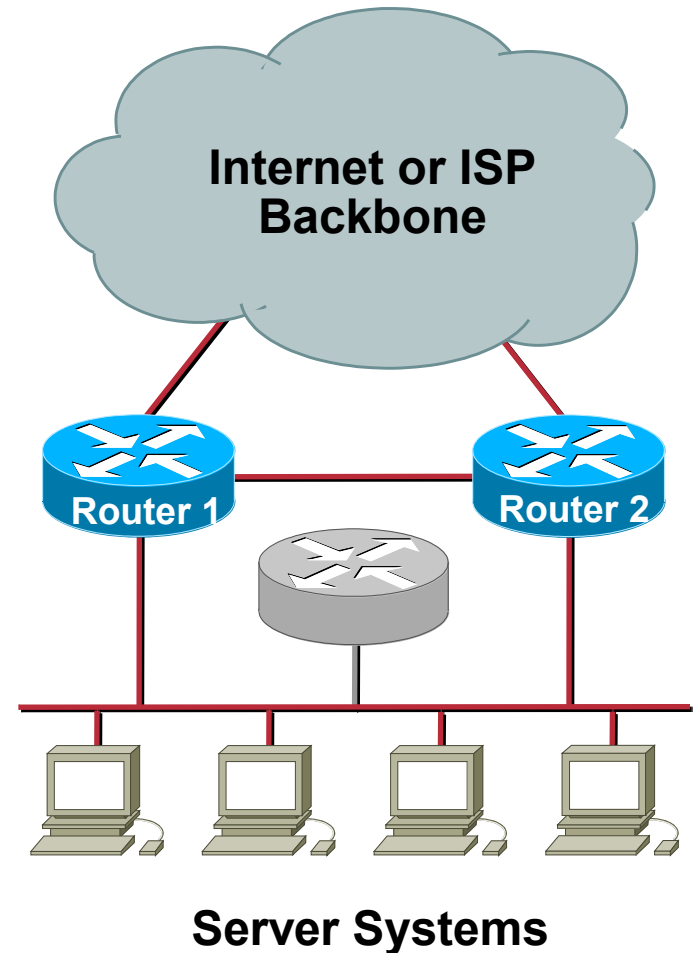
- HSRP multicast hellos toutes les 3 sec avec une priorité de 100 par défaut
- HSRP prend le control s'il a une plus grande priorité
- Si un HSRP s'aperçoit qu'il est prioritaire il prend le contrôle après un délais
- HSRP déduit 10 de sa priorité si l'interface qu'il surveille est tombé



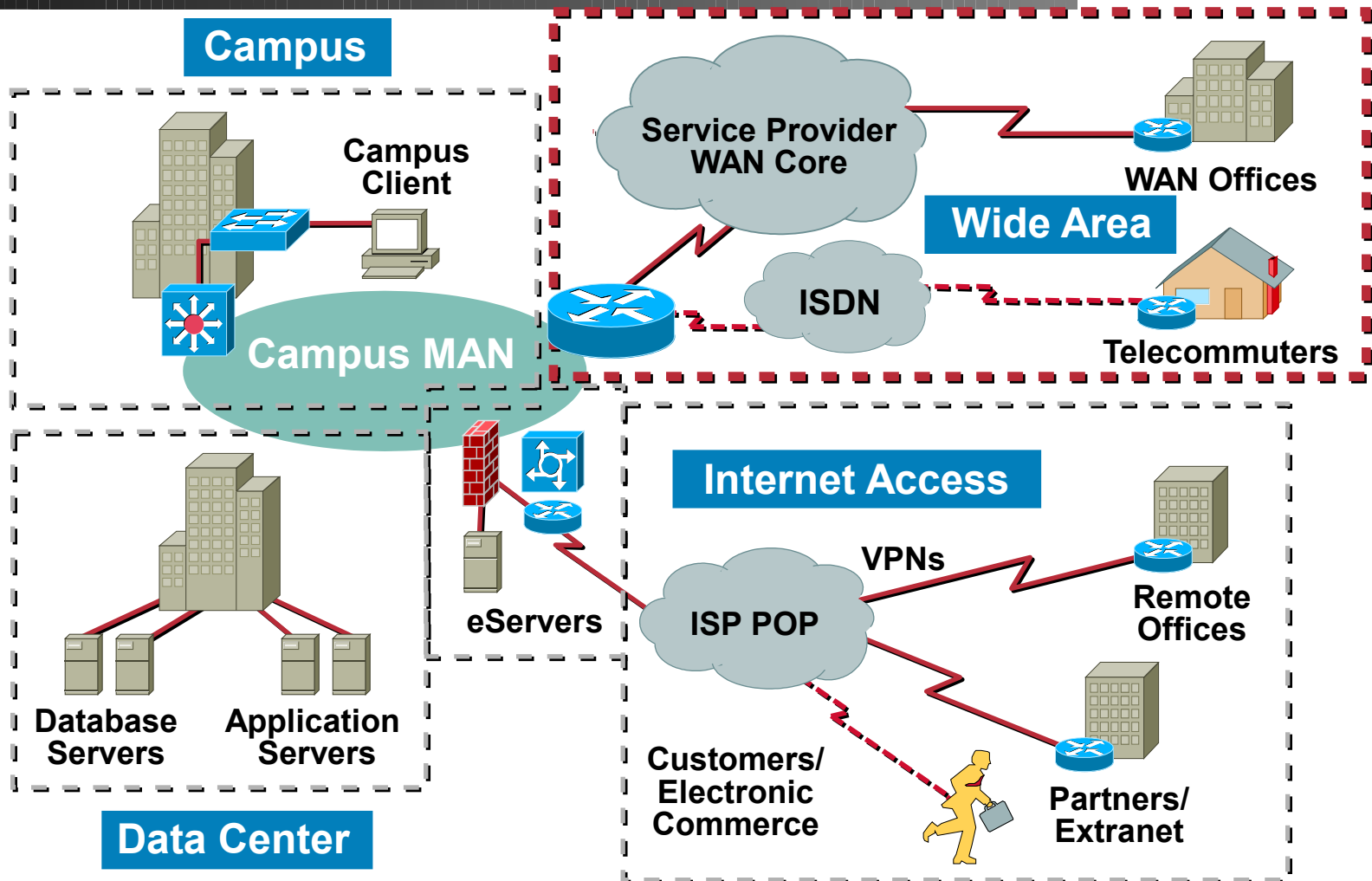
HSRP

```
Router1:  
interface ethernet 0/0  
bandwidth 128  
ip address 169.223.10.1 255.255.255.0  
standby 10 ip 169.223.10.254
```

```
Router2:  
interface ethernet 0/0  
bandwidth 1500  
ip address 169.223.10.2 255.255.255.0  
standby 10 priority 150 pre-empt delay  
10  
standby 10 ip 169.223.10.254  
standby 10 track serial 0 60
```



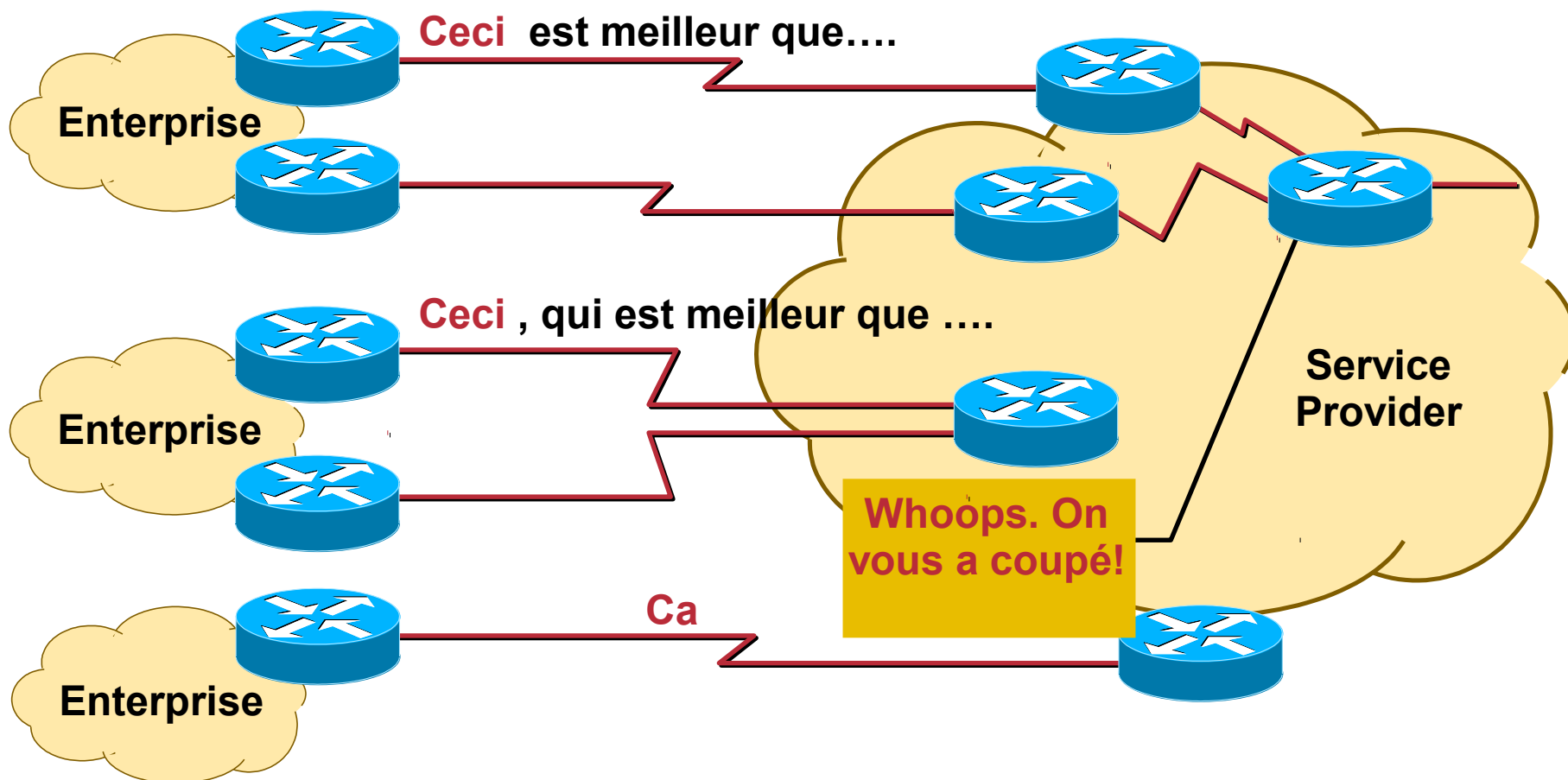
Disponibilité WAN



Diversité de circuit

- **Avoir plusieurs PVCs à travers le même port physique ne sert à rien**
- **Un port a plus de chance d'être défectueux qu'un seul PVC**
- **Utiliser des ports séparés; si possible sur des routeurs différents**
- **Essayez de demander à votre ISP de terminer vos lignes de backup sur des équipements différents**

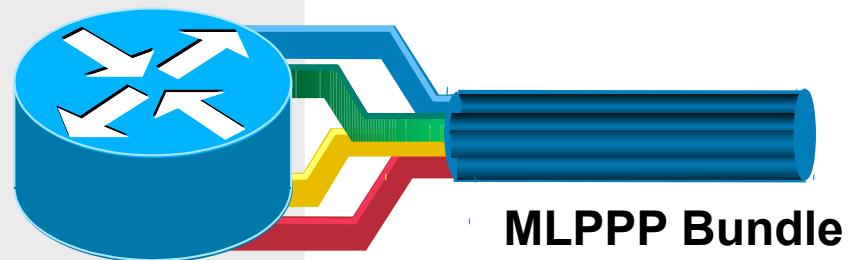
Diversité de circuit



Utilisons MLPPP

```
interface Multilink1
ip address 172.16.11.1 255.255.255.0
ppp multilink
multilink-group 1
!
interface Serial1/0
no ip address
encapsulation ppp
ppp multilink
multilink-group 1
!
interface Serial1/1
no ip address
encapsulation ppp
  ppp multilink
multilink-group 1
```

Employé avec une diversité de circuit; Multi-link PPP, fournit la redondance des lignes. Cela augmente votre bande passante



Partage de charge

- **Il y a partage de charge lorsqu'un routeur a 2 (ou plus) chemins pour atteindre la même destination**
- **EIGRP permet le partage inégale de charge**
- **Le partage de charge peut être par paquet ou par destination**
- **Le partage de charge est une technique puissante car il permet un chemin alternatif si un routeur a une déficience**

Partage de charge

- **OSPF fait le partage de charge de manière égale par défaut**
- **EIGRP fait le partage de charge de manière égale par défaut, et peut être configuré pour partager la charge de manière inégale**

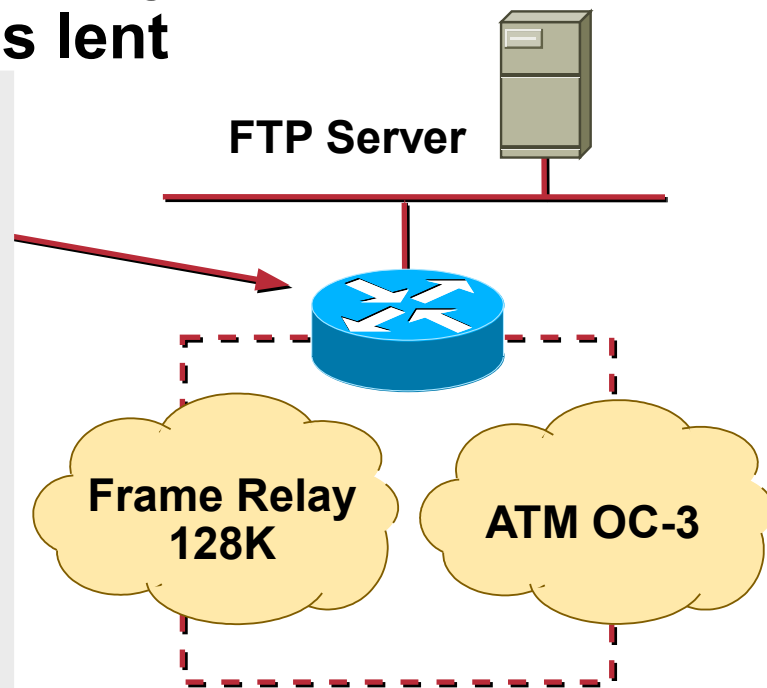
```
router eigrp 111
network 10.1.1.0
variance 2
```

- **Unequal-cost load-sharing n'est pas recommandé car il crée des problèmes de timing et de retransmissions**

Policy-based Routing

- Si vous avez des liens de coût différent et vous ne voulez pas utiliser unequal-cost load sharing, vous pouvez utiliser PBR pour envoyer le trafic basse priorité vers le lien le plus lent

```
! Policy map that directs FTP-Data  
! out the Frame Relay port. Could  
! use set ip next-hop instead  
route-map FTP_POLICY permit 10  
    match ip address 6  
    set interface Serial1.1  
!  
! Identify FTP-Data traffic  
access-list 6 permit tcp any eq 20 any  
!  
! Policy maps are applied against  
! inbound interfaces  
interface ethernet 0  
    ip policy route-map FTP_POLICY
```



Convergence

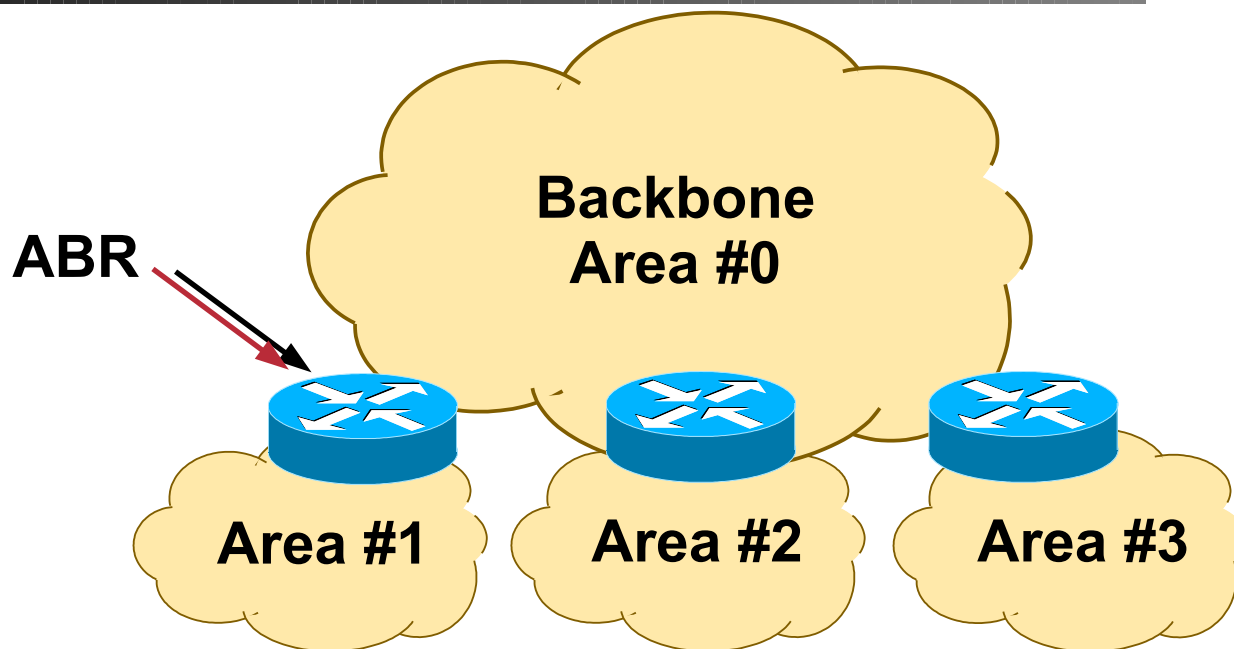
- **Le temps de convergence du protocole de routage affecte la disponibilité de votre WAN**
- **Examiner si le design niveau 2 affecte l'efficacité au niveau 3**

Facteurs déterminant la convergence du protocole



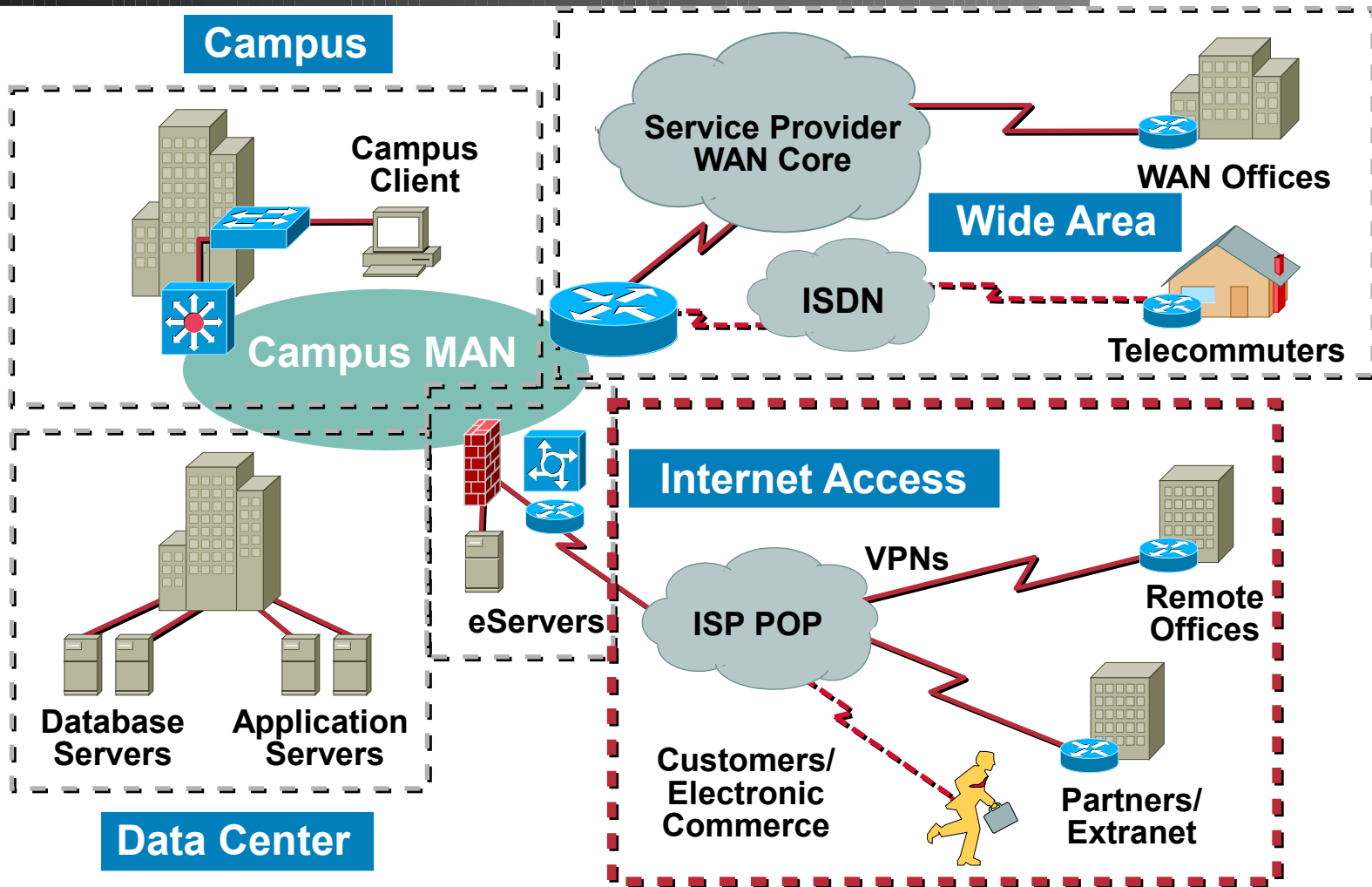
- **Taille du réseau**
- **Limitations du nombre de saut**
- **Arrangements des voisinages (cœur, bordure)**
- **Vitesse de la détection du changement**
- **Propagation des changements**
- **Design réseau : hiérarchie, summarization, redondance**

OSPF—Structure Hiérarchique

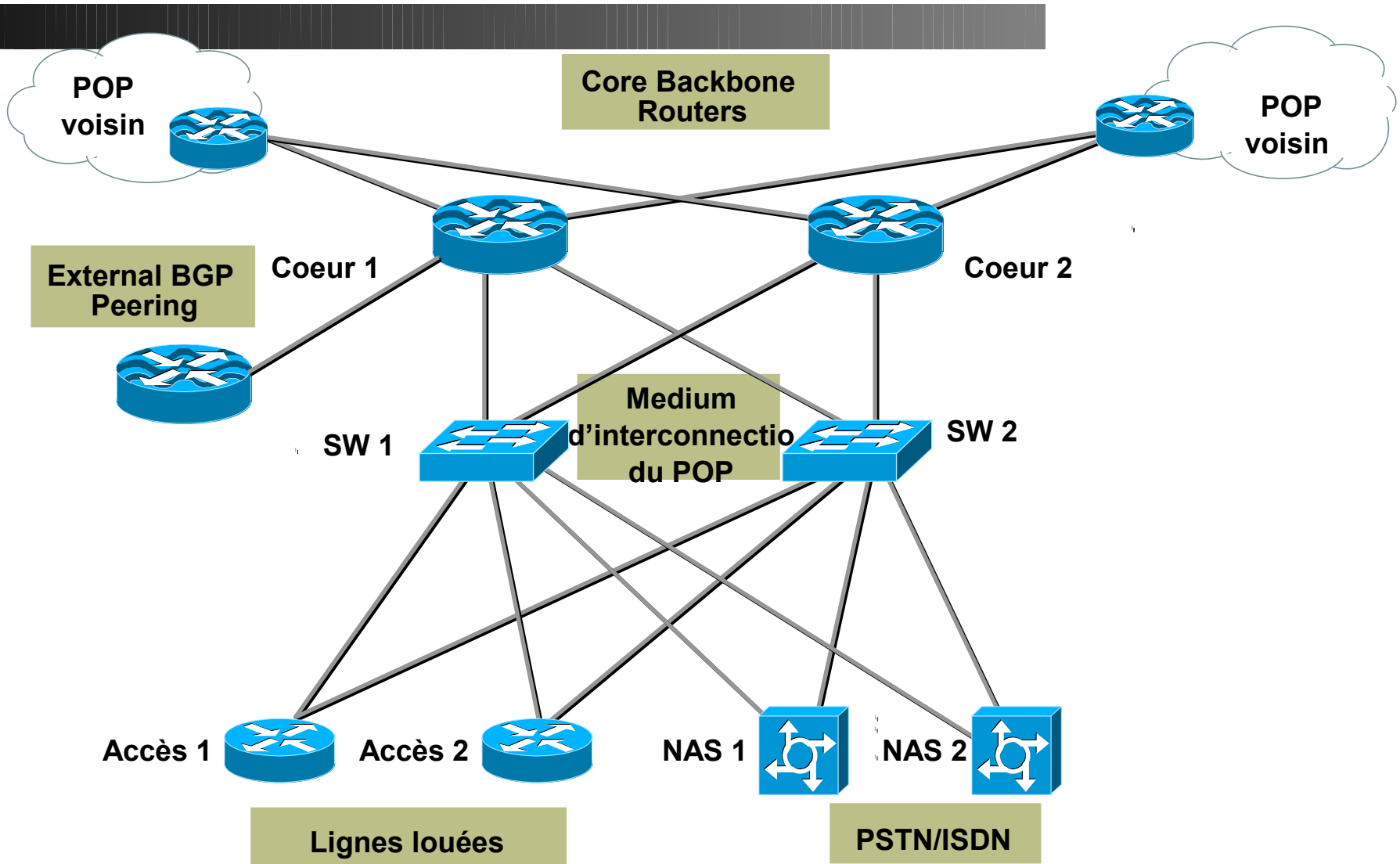


- **La topologie d'une aire est invisible hors de l'aire**
LSA flooding reste dans l'aire
Le calcul SPF se passe indépendamment dans chaque aire

Disponibilité Internet



Design de Point de Présence (POP)



Se connecter à Internet

- **Router vers Internet n'est pas significativement différent de router vers un autre WAN**
- **S'assurer de la diversité des circuits**
- **Utiliser HSRP et «track interface » pour les liens redondants**
- **Optimiser le routage avec du partage de charge**

Est ce que j'ai besoin de BGP?

Questions à poser:

Ai-je plus d'un liens vers Internet ?

Et

Est ce que pour des raisons de coût ou de sécurité ou pour raisons administratives je dois sélectionner un chemin plutôt qu'un autre

Quand ne pas l'utiliser

Lorsque vous avez un seul chemin vers Internet utiliser une route par défaut

Lorsque vous avez un plusieurs chemins vers Internet cmais vous ne voulez pas sélectionner la sortie partage de charge

“Mon ISP dis qu'il a besoin de BGP pour apprendre mes routes »

Utiliser BGP, pour envoyer vos routes mais demandez lui une route par défaut

Pour résumer

- **Implémenter des réseaux IP redondant requière une combinaison d'un bon processus, d'un bon design et d'une bonne technologie**
- **Le procédure est le plus important**



Design



Technologie



Procédure



Questions ?