

Réseau Académique Parisien

Haute disponibilité sur le Réseau Académique Parisien

TutoJRES #9 : Haute disponibilité des réseaux

Université Paris Descartes, 24 Février 2009

rap-ds@rap.prd.fr



- ◆ Projet initié en 1999, réseau construit en 2001
- ◆ 5 POPs en anneau dans Paris
- ◆ 65 établissements
- ◆ 116 sites raccordés en optique (Ethernet 100M ou 1G)
- ◆ 24 sites en bas-débits (SDSL)

- ◆ Services déployés :
 - IPv4 unicast et multicast
 - IPv6 unicast et multicast
 - CoS
 - VPN L2 (VLAN Ethernet)
 - Services multimédias (VoD)
 - Supervision 24/7

- ◆ Renouvellement des marchés en 2006-2007
- ◆ Refonte du backbone en 2008

- ◆ Objectif : Rendre le réseau plus fiable au meilleur coût
- ◆ Constats :
 - La fiabilité moyenne du réseau est en général supérieure à 99,7%
 - Les interruptions de service sont toujours dues à une panne localisée en un point du réseau (équipement ou liaison)
 - Le réseau est une ressource de plus en plus critique
- ◆ Démarche :
 - 99,7% de fiabilité, soit 0,3% de non fiabilité = 0,003 de probabilité de panne
 - En doublant les chemins et en redondant les services, la probabilité d'interruption de service devient le produit des probabilités sur chaque liaison:
$$0,003 * 0,003 = 0,000009$$
 - 0,000009 = 0,0009% de taux de non fiabilité

=> Fiabilité = 99,9991%

Historique du raccordement fiabilisé sur RAP

Réseau Académique Parisien

- ◆ Q1 2007 : ***Double raccordement sur RENATER***
 - Redondance : 2 raccords distincts au niveau physique
 - Répartition des trafics : Mode Actif-Actif, partition du backbone en 2 zones
 - Eviter l'asymétrie de trafic

- ◆ Q2 2007 : ***Spécifications du double raccordement pour les sites***
 - Prise en compte des différents cas : (1)double raccordement à RAP, (2)raccordement secouru, (3)raccordement RAP/FAI
 - Différents types de raccordement au niveau physique
 - Différents modes : actif-actif vs actif-passif en évitant l'asymétrie de trafic
 - Le service est basé sur BGP (fiabilisation IP)

- ◆ Q4 2007 : ***Mise en oeuvre, premier site doublement raccordé***

- ◆ Q1/Q2 2008 : ***Tutoriel “Raccordement fiabilisé à RAP”*** (70 participants)

- ◆ Q1 2009 : ***Fiabilisation des niveaux 2***

Constation sur l'origine des interruptions de service

Réseau Académique Parisien

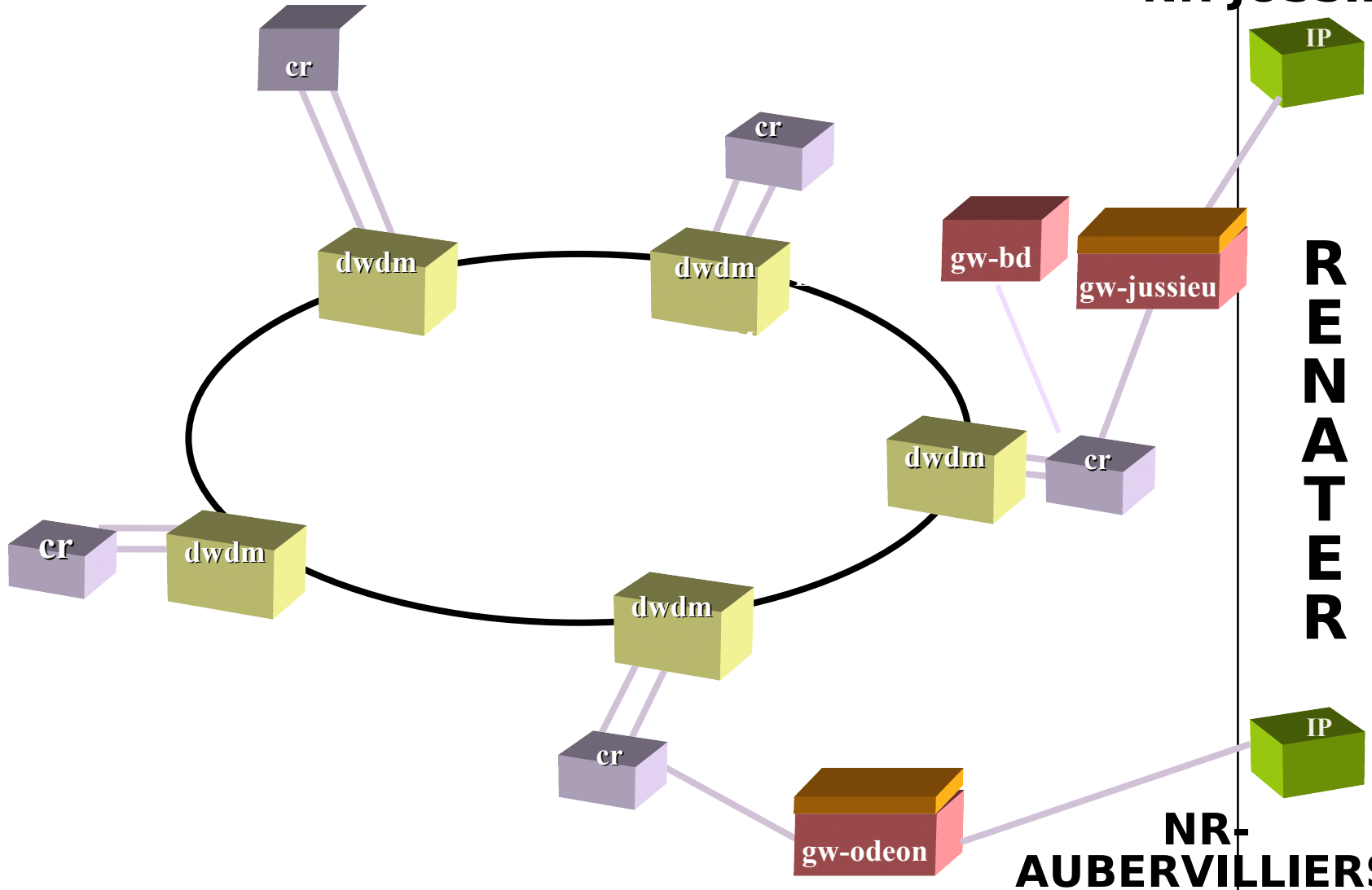
- ◆ Interruption due à un incident sur fibre optique : _____ - -
- ◆ Interruption due à un incident sur équipement actif : ___ -
- ◆ Interruption due à l'exploitation du réseau : _____ +
 - ◆ Maintenance
 - ◆ Configuration
 - ◆ Erreur
- ◆ Interruption due à l'environnement : _____ + +
 - ◆ Alimentation électrique
 - ◆ Climatisation
 - ◆ ...



RAP (avant)

Réseau Académique Parisien

NR-JUSSIEU

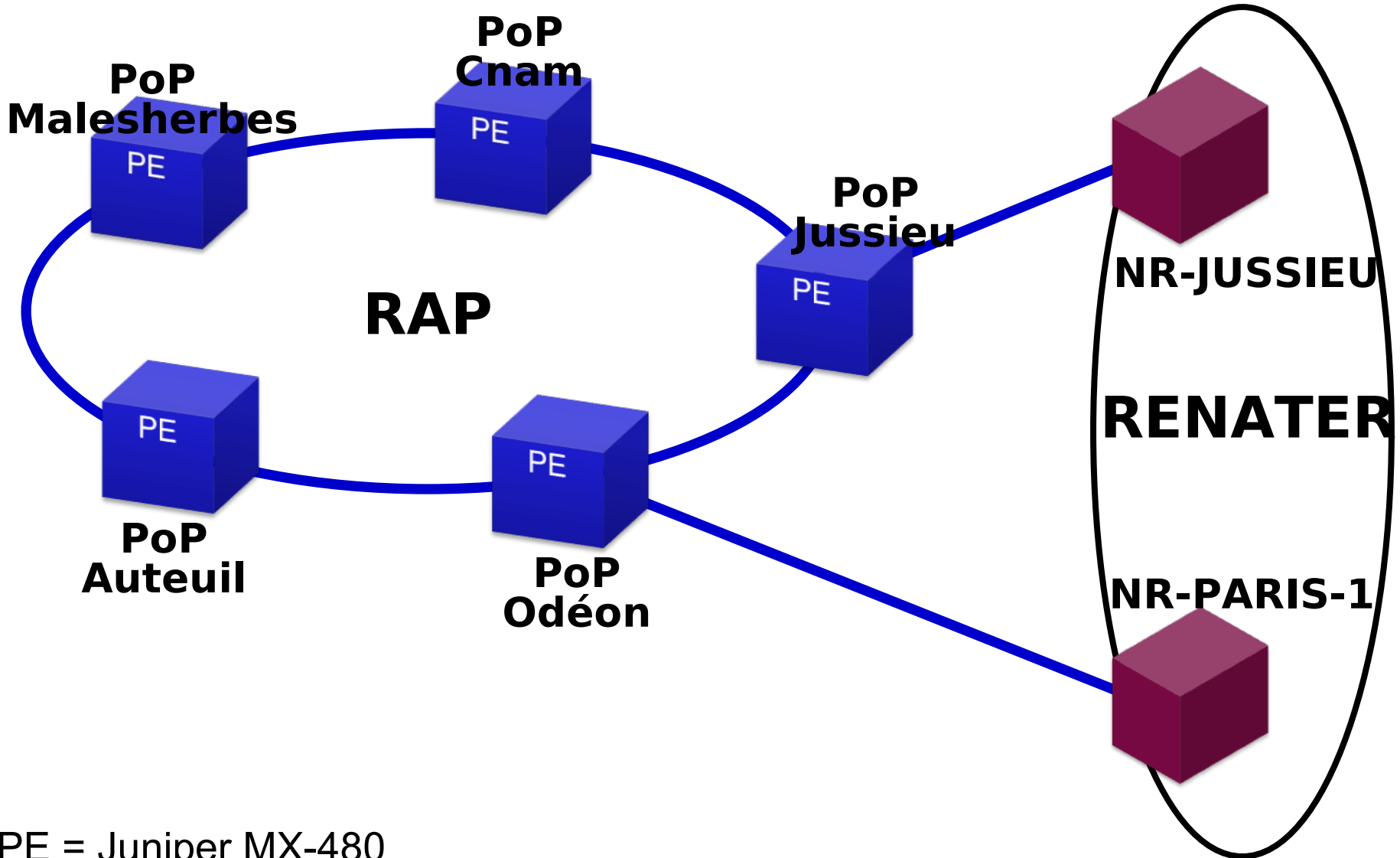


RENATER

NR-AUBERVILLIERS

RAP (maintenant)

Réseau Académique Parisien



PE = Juniper MX-480

Double attachement sur RENATER

Réseau Académique Parisien

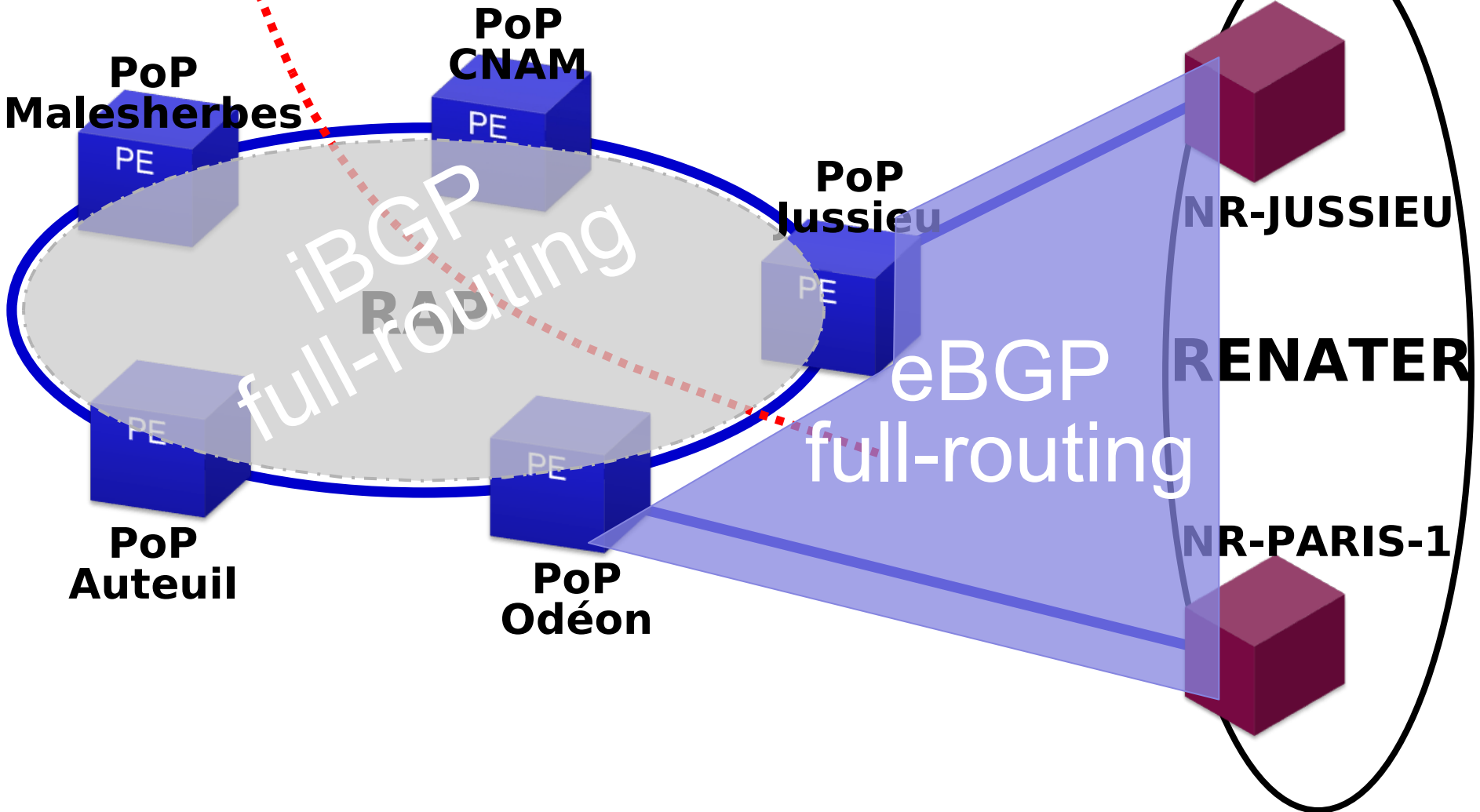
- ◆ Redondance d'accès :
 - 2 liens physiques distincts :
 - PoP Jussieu - NR-JUSSIEU
 - PoP Odéon - NR-PARIS-1
- ◆ Répartition des charges :
 - Partition du backbone en 2 zones : Odéon et Jussieu
 - Utilisation des communautés BGP définies par RENATER :

Zone RAP	NR-PARIS-1	NR-JUSSIEU
Jussieu Site raccordé au PoP <u>CNAM</u> ou <u>Jussieu</u>	2200:590 <i>RAP-Community-Backup</i>	2200:610 <i>RAP-Community-Main</i>
Odéon Site raccordé au PoP <u>Odéon</u> , <u>Auteuil</u> ou <u>Malesherbes</u>	2200:610 <i>RAP-Community-Main</i>	2200:590 <i>RAP-Community-Backup</i>

Double attachement sur RENATER

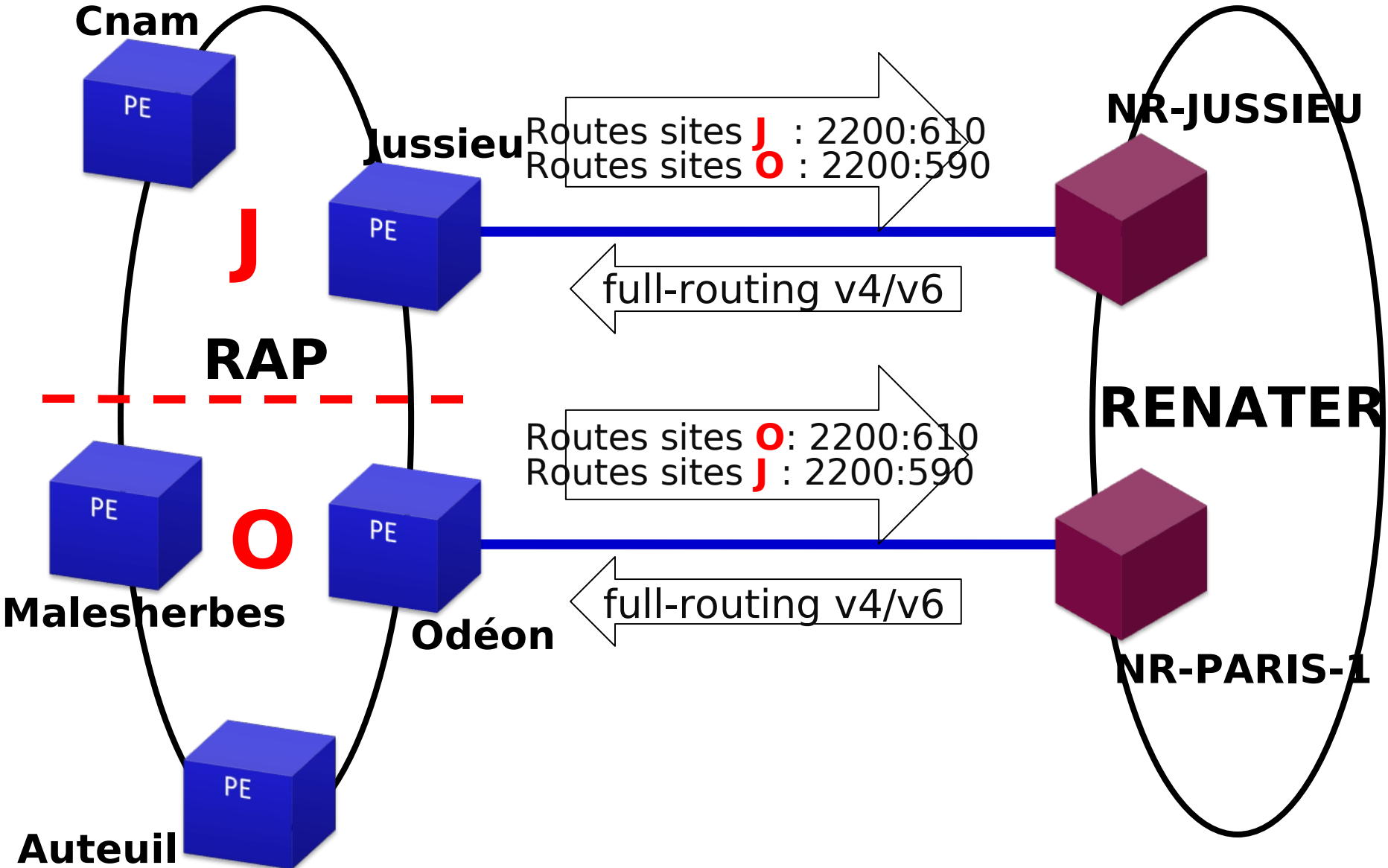
Réseau Académique Parisien

Zone Odéon Zone Jussieu



Double attachement sur RENATER

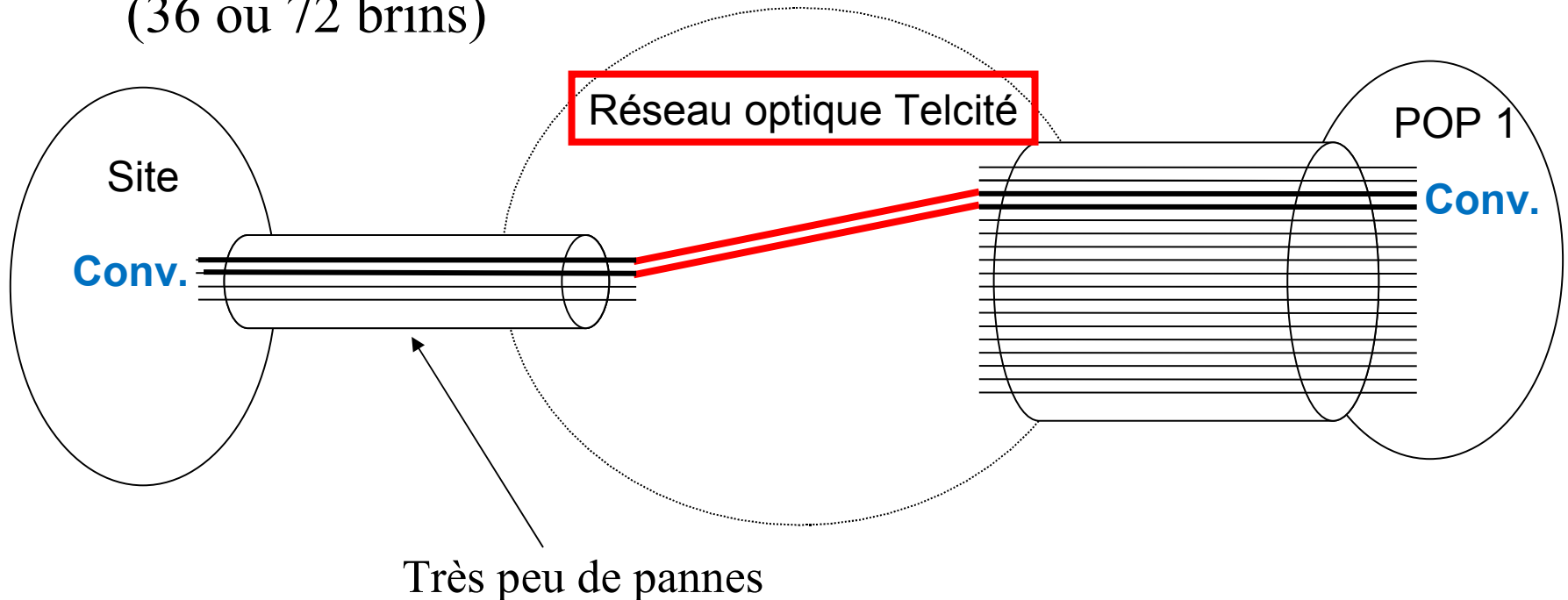
Réseau Académique Parisien



Composition d'une liaison typique

Réseau Académique Parisien

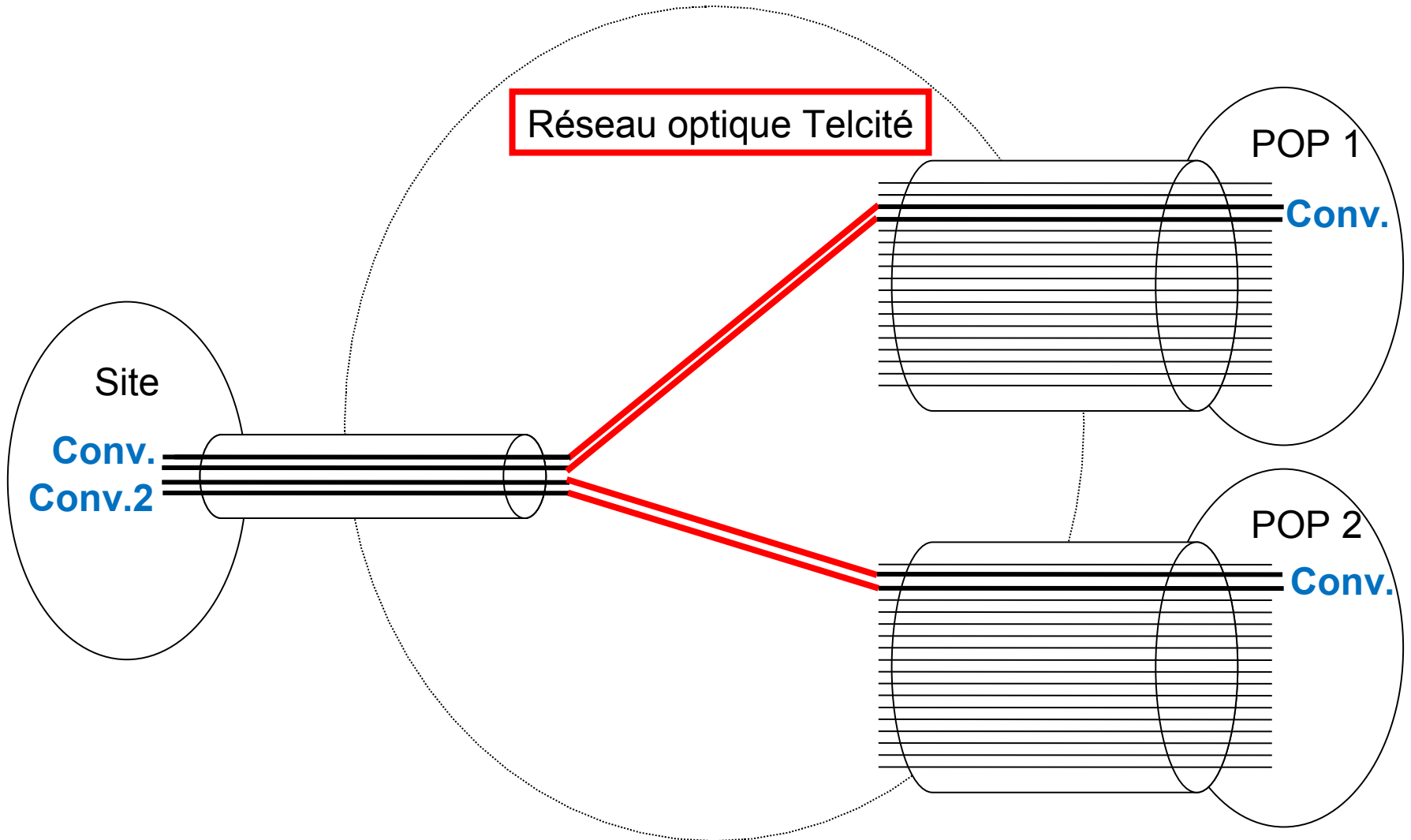
- ◆ Un segment entre le site et la plus proche station de métro (4 brins) construit en égout lors du raccordement du site
- ◆ Un segment sur le réseau Telcité (continuité de deux brins optiques)
- ◆ Un segment entre la station de métro proche du POP et le POP (36 ou 72 brins)





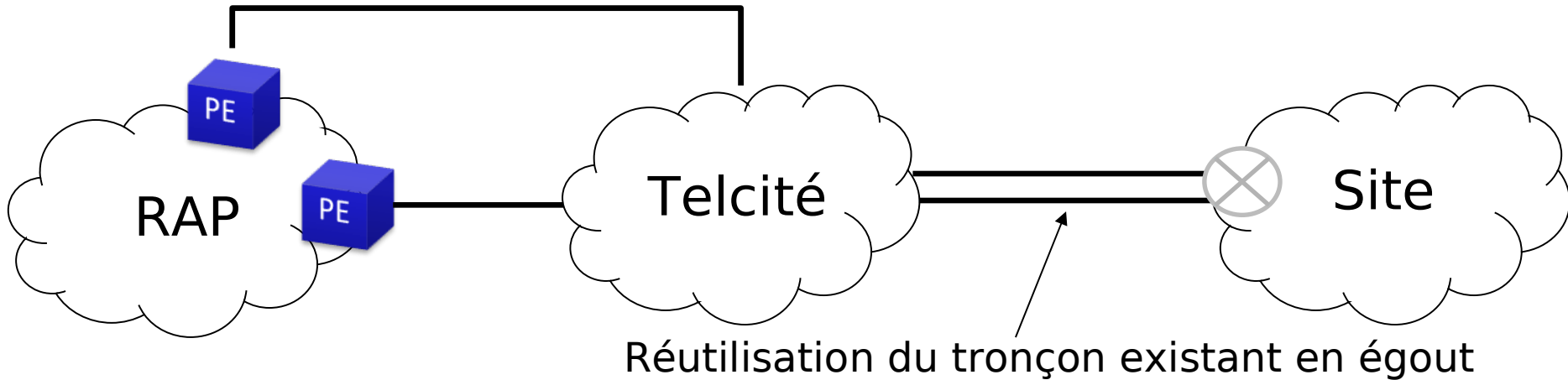
Exemple d'un double raccordement d'un site

Réseau Académique Parisien

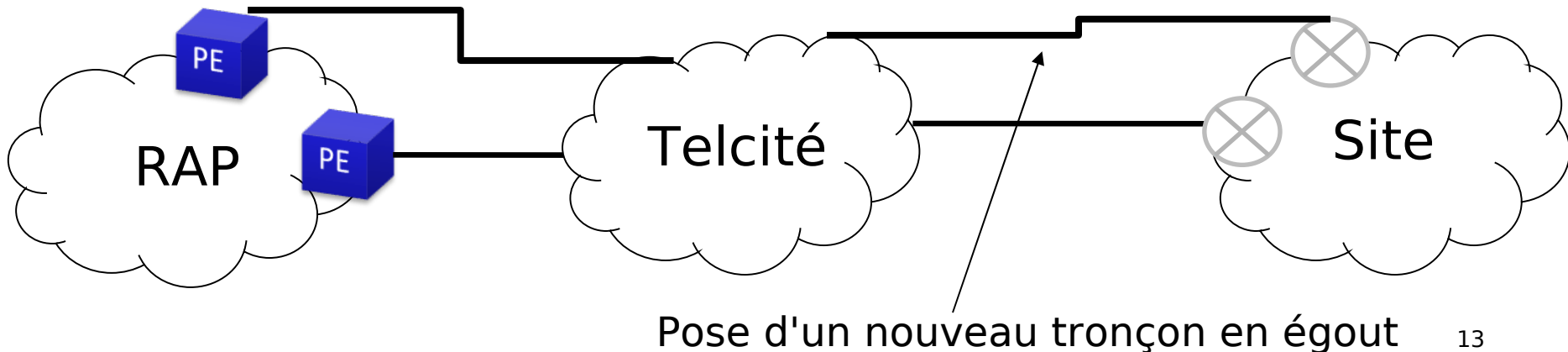


Types de raccordements fiabilisés au niveau physique

◆ Raccordement fiabilisé « standard » :



◆ Raccordement fiabilisé sur câble séparé :



Fiabilisation des services IPv4/IPv6

Réseau Académique Parisien

- ◆ Raccordement statique ou dynamique :
 - Routage statique ou peering eBGP
 - Toutes les routes des sites sont injectées dans l'iBGP de RAP
- ◆ Double attachement d'un site, 2 peerings eBGP :
 - PoP nominal
 - PoP de secours
 - Répartition des charges par préfixe IP
- ◆ Une communauté nominale par PoP : pour le rendre propriétaire
- ◆ Tableaux des correspondances ***communauté SITE - localpref RAP*** :
 - Policy en entrée
- ◆ Trafic en sortie du site vers Internet : OSPF fait son travail
- ◆ Trafic en entrée de l'Internet vers le site :
 - Tableau des correspondances ***localpref RAP - communauté RENATER***
 - Policy en sortie
- ◆ Pour un site statique :
 - Ses routes sont marquées de la localpref nominal du PoP auquel il est raccordé

Fiabilisation des services IPv4/IPv6

Réseau Académique Parisien

Communauté	POP 1 = PoP nominal, 2 = PoP de secours				
	Jussieu	Odéon	Auteuil	Malesherbes	CNAM
100	1	2	2	2	2
200	2	1	2	2	2
300	2	2	1	2	2
400	2	2	2	1	2
500	2	2	2	2	1

1 = PoP nominal, 2 = PoP de secours (3,4,5 : généralisation à 5 accès)

com	lpf
100	600
200	600
300	500
400	900
500	600

com	lpf
100	500
200	500
300	500
400	500
500	1000

com	lpf
100	1000
200	500
300	500
400	500
500	500

Lpf	com
1000	610
900	590
600	590
500	610

Malesherbes

com	lpf
100	600
200	600
300	900
400	600
500	600

Auteuil

com	lpf
100	600
200	900
300	600
400	600
500	600

Cnam



Jussieu

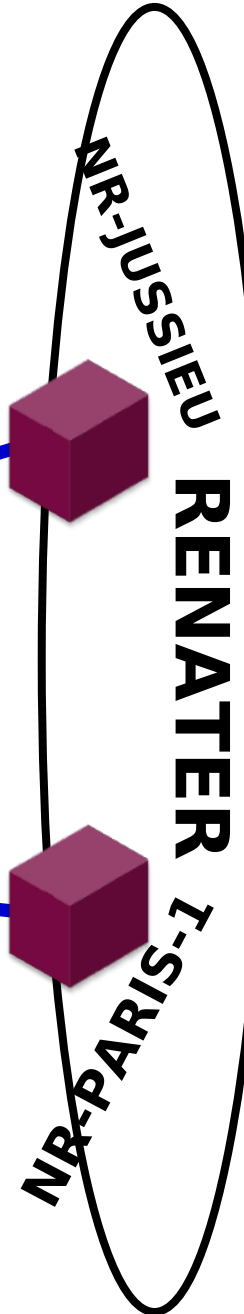


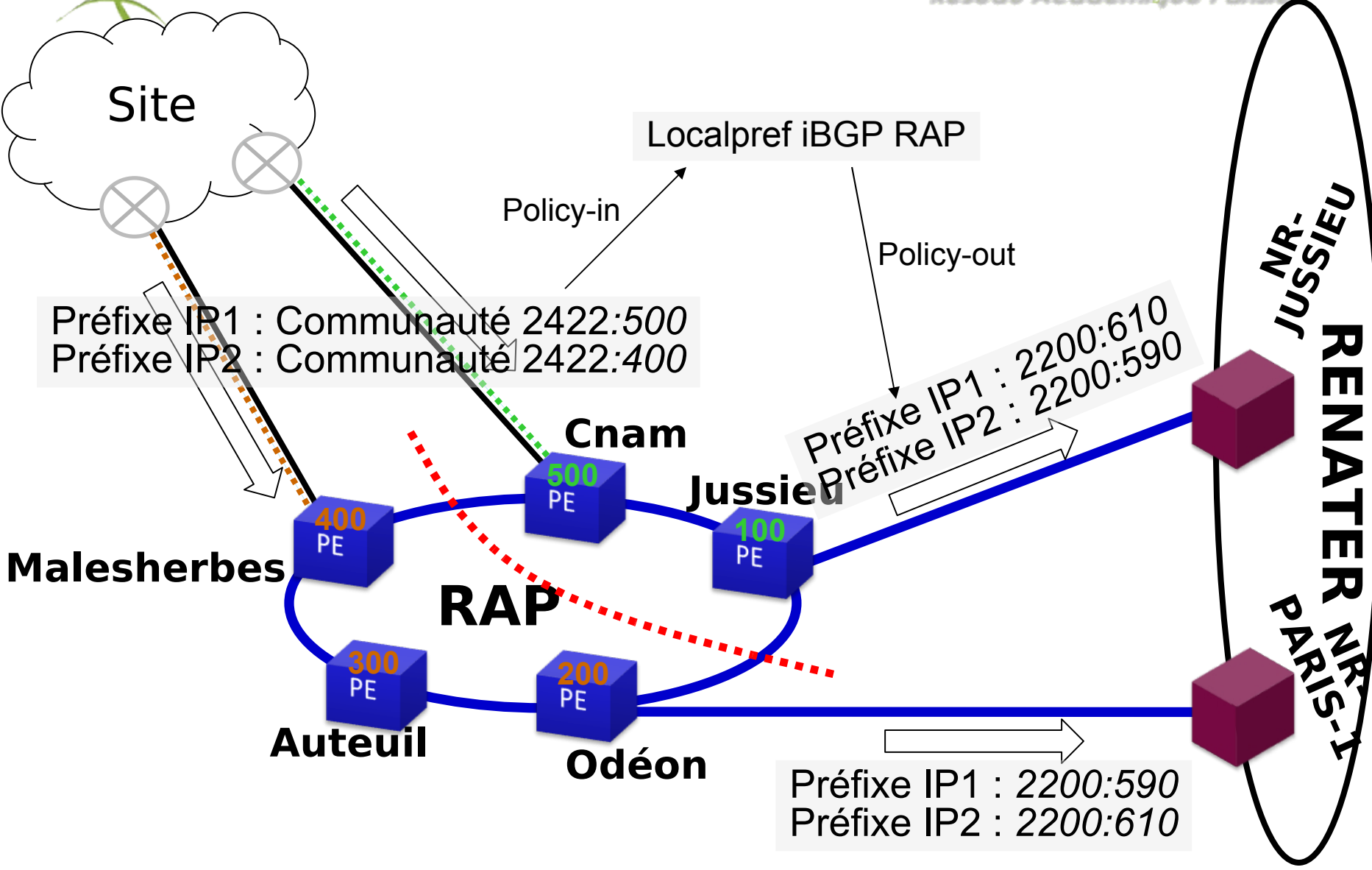
RAP



Odéon

Lpf	com
1000	590
900	610
600	610
500	590





Raccordement fiabilisé des L2VPN



Réseau Académique Parisien

- ◆ Déploiement VPLS (Q4 2008)
- ◆ Prise en compte du raccordement fiabilisé des sites (en cours)
- ◆ Ethernet OAM (à venir)

- ◆ Superviser et connaître l'état de fonctionnement du double raccordement :
 - Etat nominal ?
 - Etat non nominal fonctionnel ?
 - Etat non nominal non fonctionnel ?

- ◆ Moyens :
 - ICMP
 - depuis un serveur de supervision => attention à l'asymétrie
 - SNMP Ping : MIB ICMP
 - SNMP :
 - état du peering BGP
 - réception des routes
 - dampening ?
 - Looking-glass, prochainement sur RAP

- ◆ Métrologie
 - Courbes MRTG à adapter

- ◆ Tester les scénarios de pannes (régulièrement) !

- ◆ Conception du raccordement fiabilisé pour :
 - éviter tant que c'est possible l'asymétrie de trafic
 - tout en permettant de répartir les flux
 - sans figer « en dur » l'ingénierie du trafic

- ◆ Temps de convergence
 - en-dessous de la seconde (observé « empiriquement »)
 - réglages fins nécessaires pour le réduire encore plus

- ◆ Communautés supplémentaires
 - pour les sites raccordés à un autre FAI (3 sur RAP)

- ◆ Questions ?