

Qualité de service dans les réseaux

Problématique, solutions et challenges

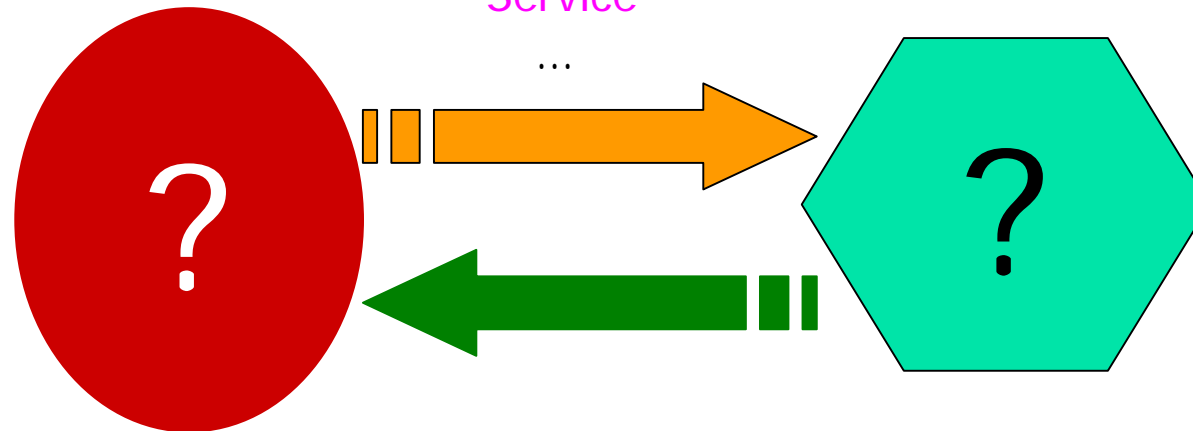
Zoubir Mammeri



Appelant
Demandeur
Client
Utilisateur
...

Opération
Fonction
Travail
Service
...

Appelé
Demandé
Serveur
Fournisseur
...



Qualité

- Exigée, souhaitée, implicite/explicite, convenue à l'avance, Faire confiance...
- Mesurable (Qualitative/Quantitative) ou non
- Mesurable en ligne / hors ligne

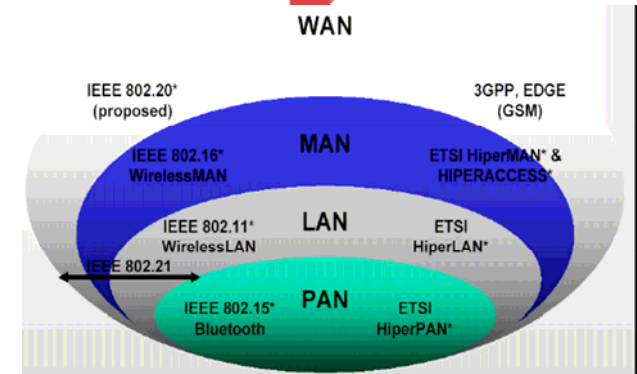
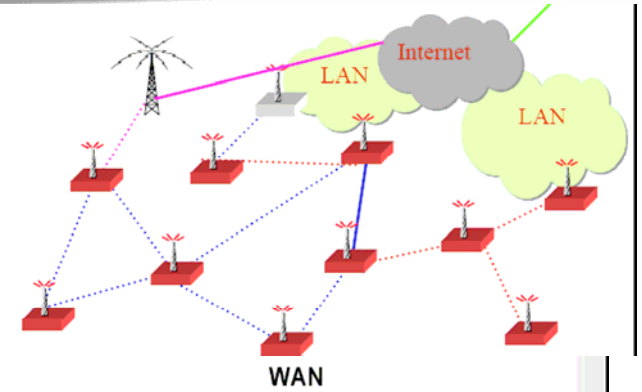
Plan

1. Introduction - Contexte
2. Définitions et Concepts
3. Panorama des fonctions de gestion de QoS
4. Modèles de trafic
5. Contrôle d'admission
6. Allocation de ressources
7. Routage
8. Ordonnancement
9. Contrôle de congestion
10. Autres fonctions
11. Mapping de QoS
12. Politique de QoS
13. Conclusion

1/6

Diversité des Réseaux

- Réseaux publics, privés, dédiés...
- Filaires - Sans fil (ondes radio, IR, Satellite)
- PAN, LAN, MAN, WAN
- Industriels, embarqués, bureautiques...
- Site : salle de TP, gare, train, avion, voiture...
- 1 domaine, n domaines
- Administration : centralisée, répartie, autonomie
- Environnement : montagne, tunnel, chaleur, humidité, rayonnement...
- Réseaux : invisibles, intelligents, self-organizing, accessibles partout...
- ...



Diversité du public concerné

■ Genres d'intervenants (utilisateurs)

- Personne, robot, capteur, objet logiciel...
- Informaticien, automaticien, opérateur de télécom, fournisseur d'accès, militaire, industriel ... grand public
- Très exigeant ('le réseau est sensé tout faire'), on prend ce qu'offre le réseau ('compréhensif')
- Accepte de négocier, veut tout préconfiguré...
- Accepte un coût : très élevé, ..., modeste, gratuit
- ...

■ Vues des intervenants

- Développement de réseaux et de services
- Contenu et sa diffusion
- Contenu et son utilisation
- Transport de bout en bout
- Transport sur un domaine, un routeur ou une antenne
- ...

Diversité des Applications

■ Secteurs/domines d'activité

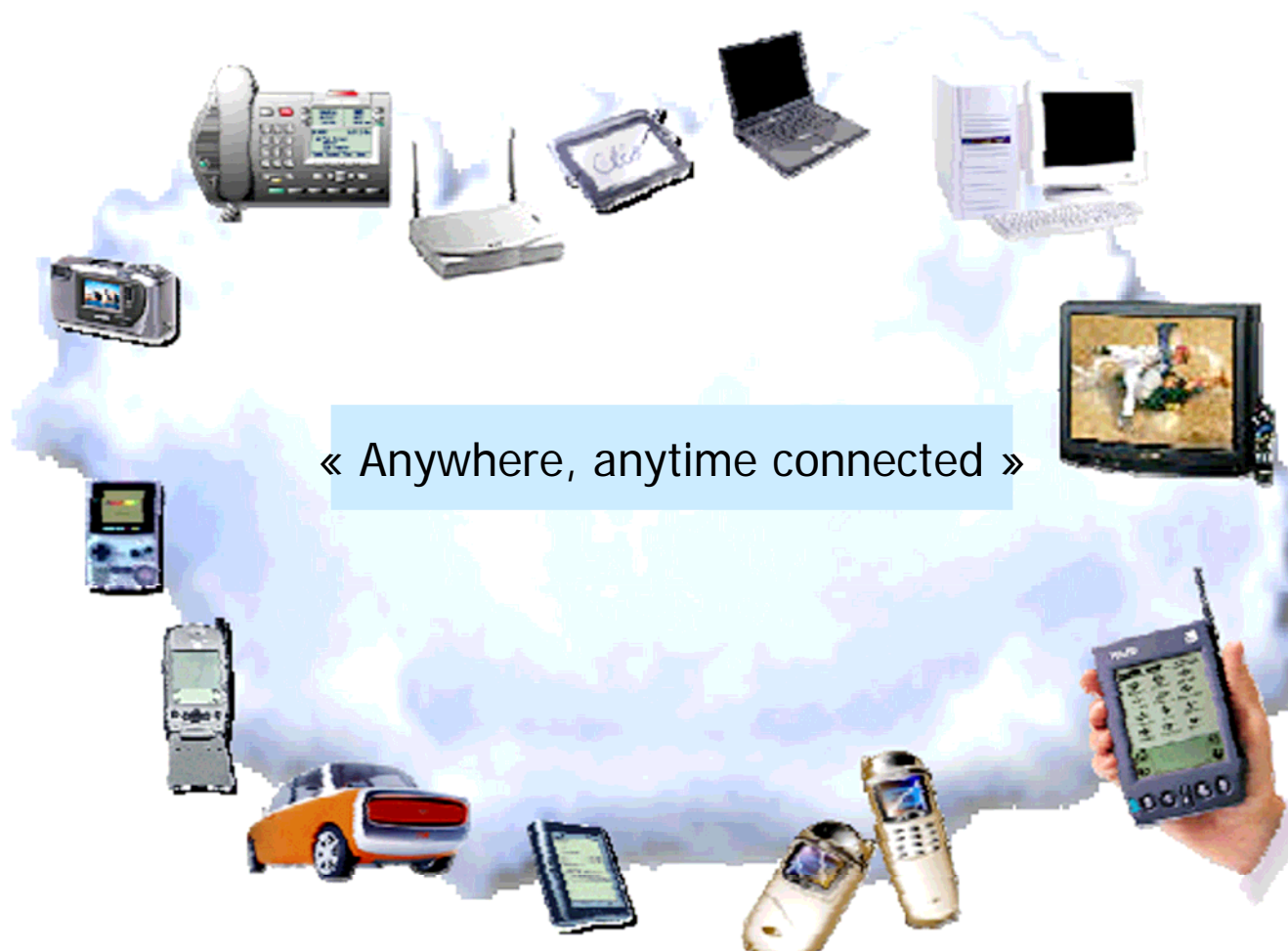
- Commande/supervision de centrales nucléaires
- Organisation de commandement militaire
- Santé
- Transport
- Vidéo surveillance, identification de personnel
- Contrôle-commande
- Commerce électronique
- Loisirs, Musique, Jeux
- ...

■ Nature des échanges

- Critiques ou non
- Applications : Transactionnelles, Réactives, Interactives...
- 1 vers 1, 1 vers m , m vers 1, n vers m
- ...

1. Introduction – Contexte 4/6

Diversité des Equipements



Diversité des Equipements

- Grand public, privé, spécialisé
- Fixe, mobile (mobilité lente ou rapide)
- Cher/pas cher
- Disparaît après utilisation ('sensor') ou non
- Ecoute : toujours à l'écoute, dormant, émetteur, récepteur...
- Avec contraintes de batterie (rechargeable ou non)
- Localisable : à la demande, toujours, de manière intelligente
- Equipement intelligent ou non
- ...

Difficultés de parler de QoS

- Multiforme (temps, sécurité, coût...)
- Différentes vues (grand public,..., Informaticien)
- Différents niveaux (application, réseau, physique...)
- Différents mécanismes et moyens

Exposé limité au Réseau (transport de données)

2. Définitions et Concepts 1/3

■ Définition de l'IETF

« La qualité de service désigne la manière dont le service de livraison de paquets est fourni et qui est décrite par des paramètres tels que la bande passante, le délai de paquet et le taux de perte de paquets »

■ Classes de service

■ Garantie absolue (déterministe)

- Une valeur (délai < 10 ms)
- Un intervalle de valeurs (délai dans $[80 \dots 100]$)

■ Probabiliste / Statistique

- Avec une probabilité P (délai < 100 ms à 90%)
- Conforme à une certaine distribution stochastique
- Prédictive, charge contrôlée, Molle, Meilleure que le meilleur effort

■ A meilleur effort

2. Définitions et Concepts 2/3

Paramètres de QoS

- **Aspects temporels** : Temps de transfert, latence, délai, gigue, temps de réponse, temps d'aller-retour, temps d'établissement/fermeture de connexion...
- **Volume** : Bits/s, paquets/s, pourcentage de bande passante...
- **Fiabilité/disponibilité/robustesse** : Taux de disponibilité, MTBF, MTTR...
- **Paramètres d'erreurs** : Taux d'erreur, de perte, de désordre de paquets...
- **Coût** : Coût (€ ...), pénalité, bonus...
- **Sécurité** : Capacité du contrôle d'accès, capacité du chiffrement, surcoût des mécanismes de sécurité...

2. Définitions et Concepts 3/3

SLA (Service Level Agreement)

- Gestion orientée connexion vs Gestion orientée SLA
- SLA = Contrat User-Provider
- SLA Statique ou Dynamique

■ SLA = {

QoS requise,

Spécification de trafic,

Règles de traitement de paquets,

SLS

Coûts, pénalisation, bonus,

Aspects juridiques

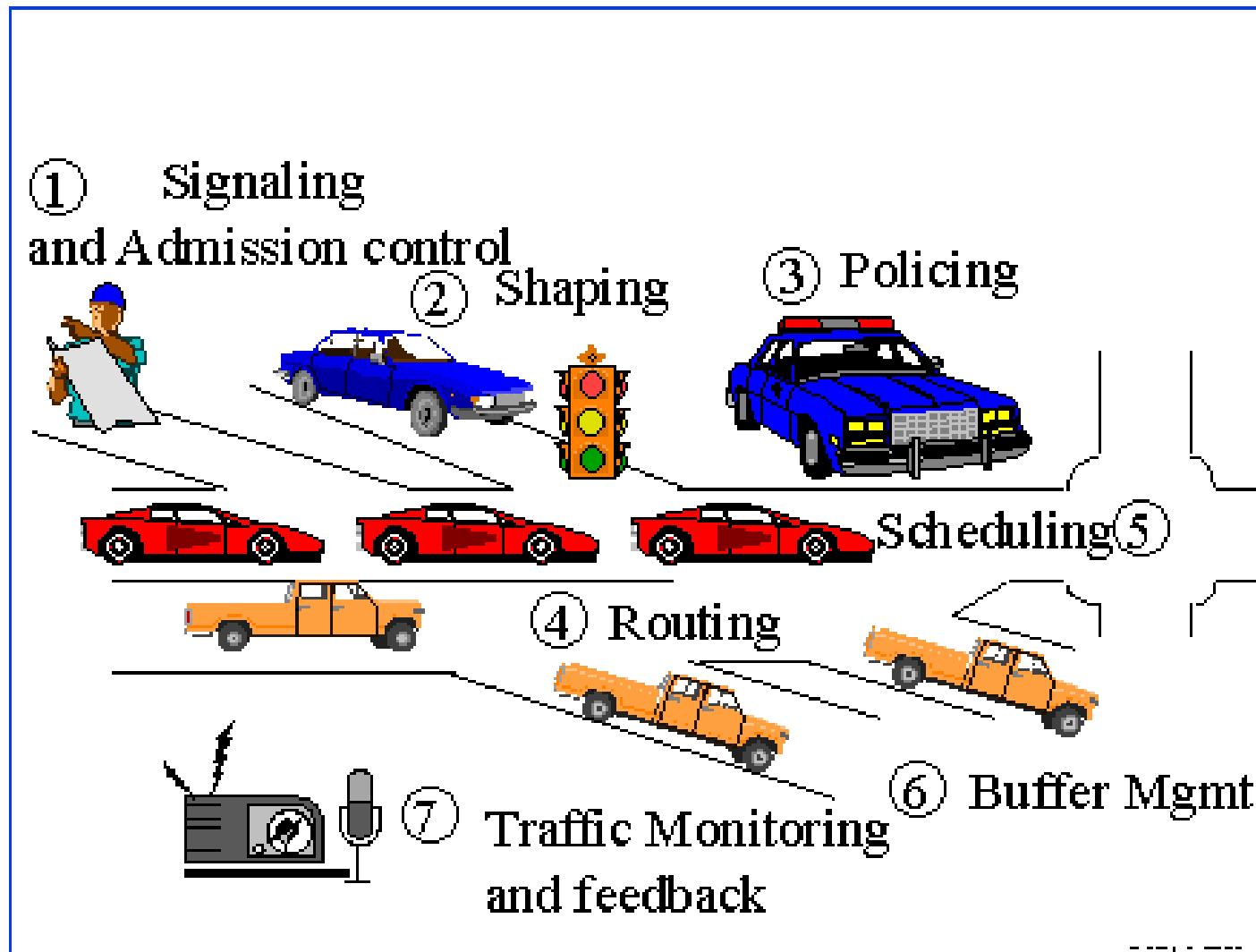
...

}

3. Panorama des fonctions de gestion de QoS 1/4

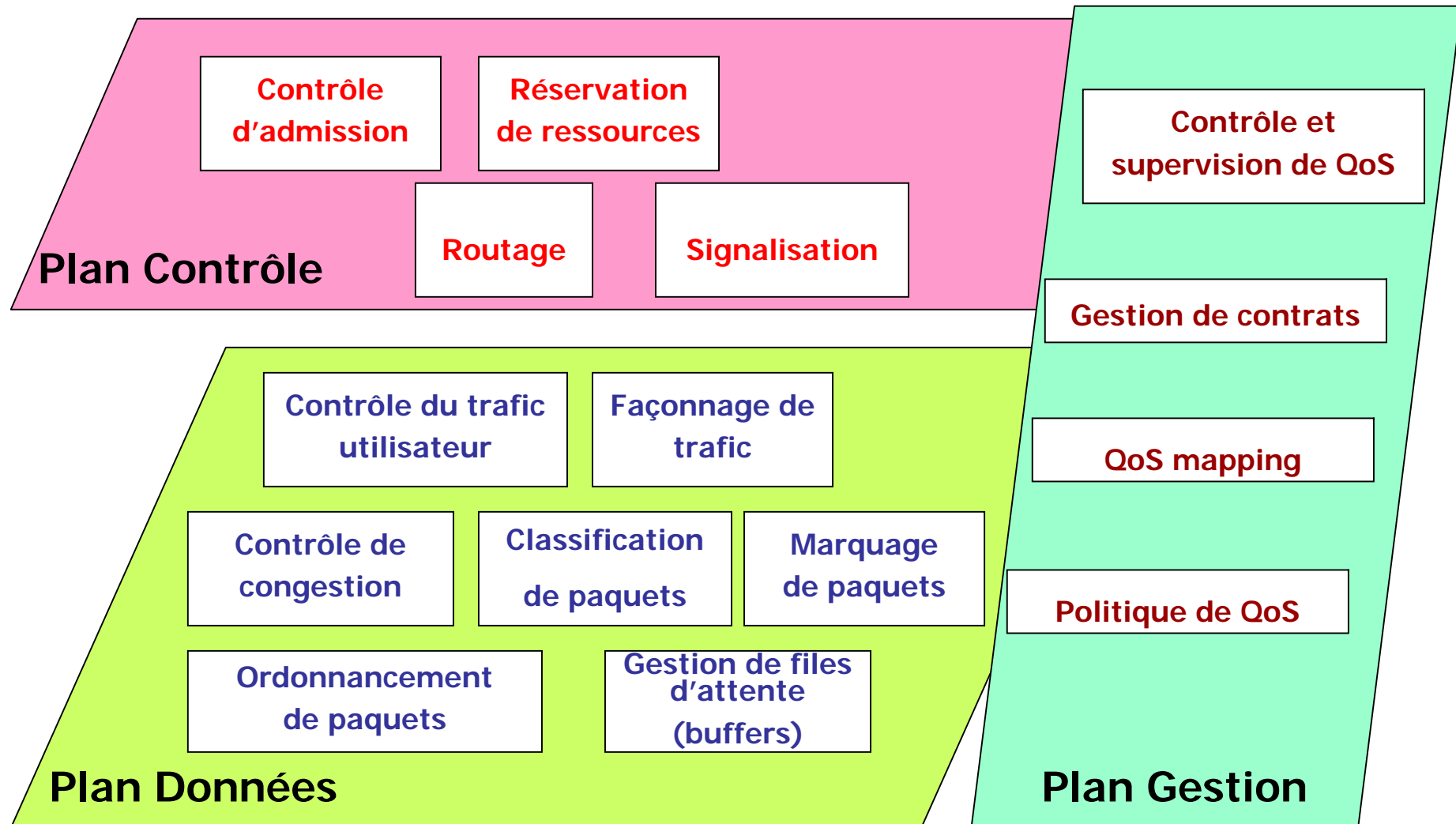


3. Panorama des fonctions de gestion de QoS 1/4

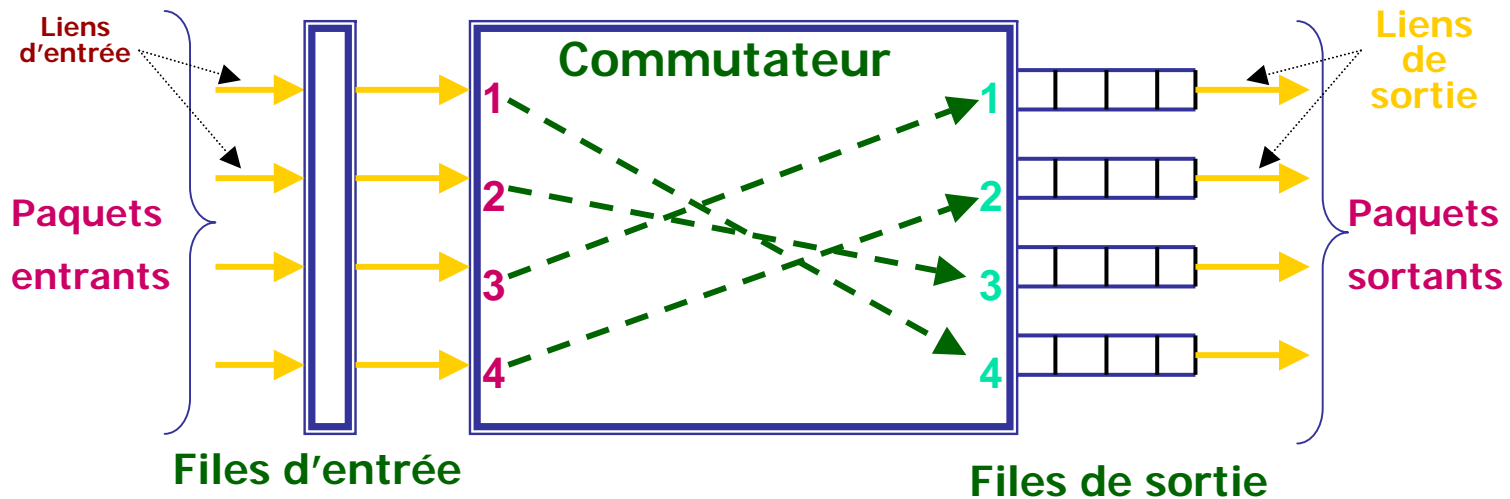


3. Panorama des fonctions de gestion de QoS 2/4

Fonctions mises en œuvre pour la garantie de QoS

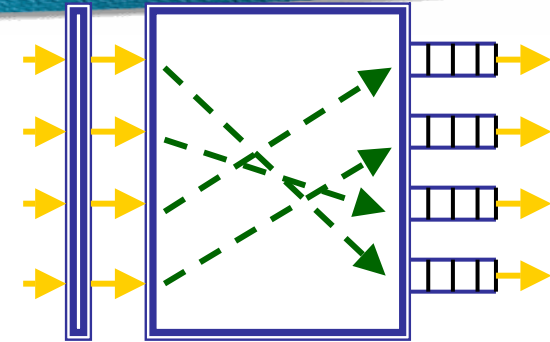


Fonctionnement simplifié d'un commutateur/routeur



3. Panorama des fonctions de gestion de QoS 4/4

Composants du délai de bout-en-bout



→ Délais d'attente dans les files d'entrée

→ Délais de construction de paquets

→ Délais de commutation

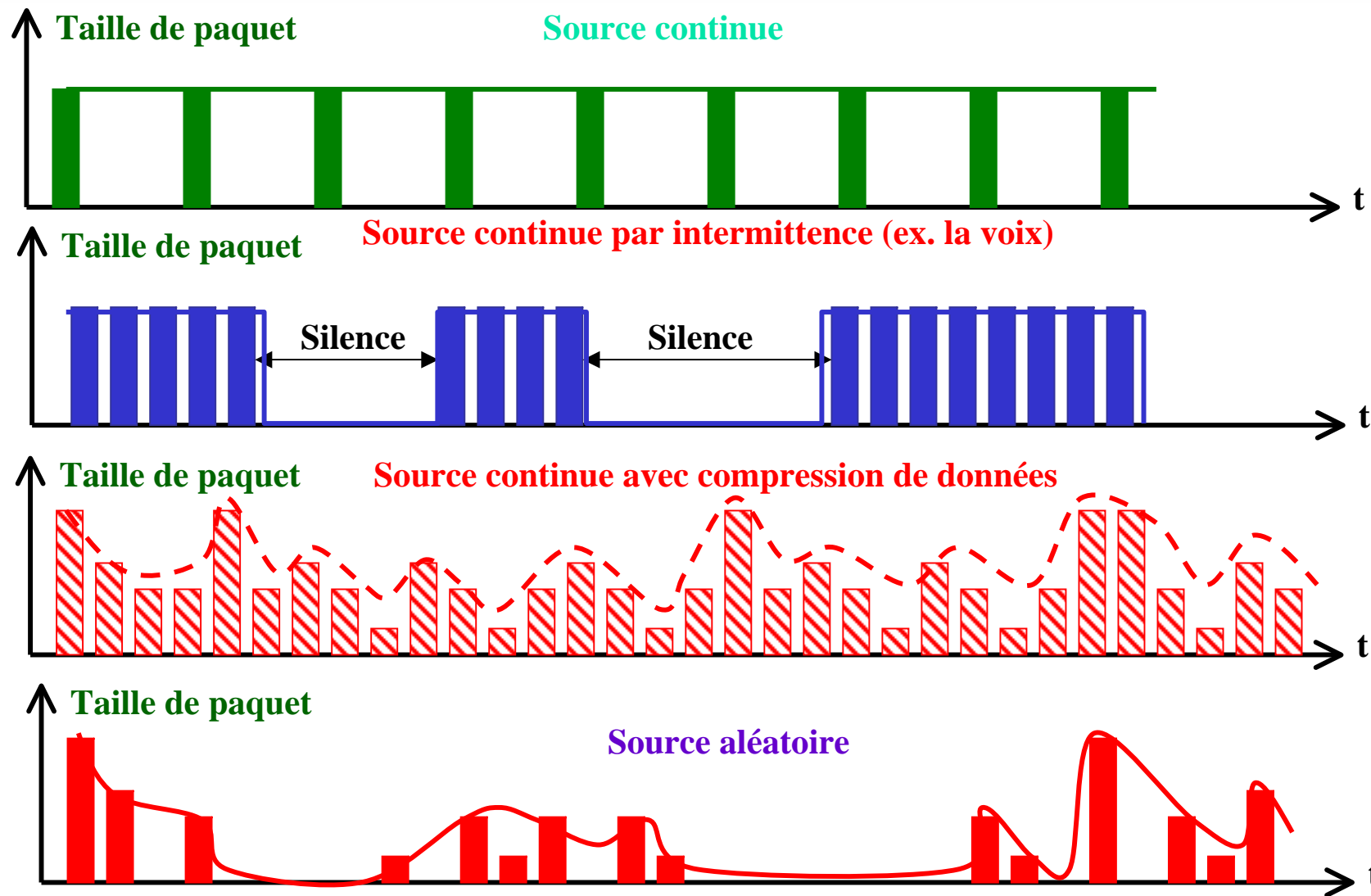
→ Délais d'attente dans les files de sortie

→ Délais de transmission

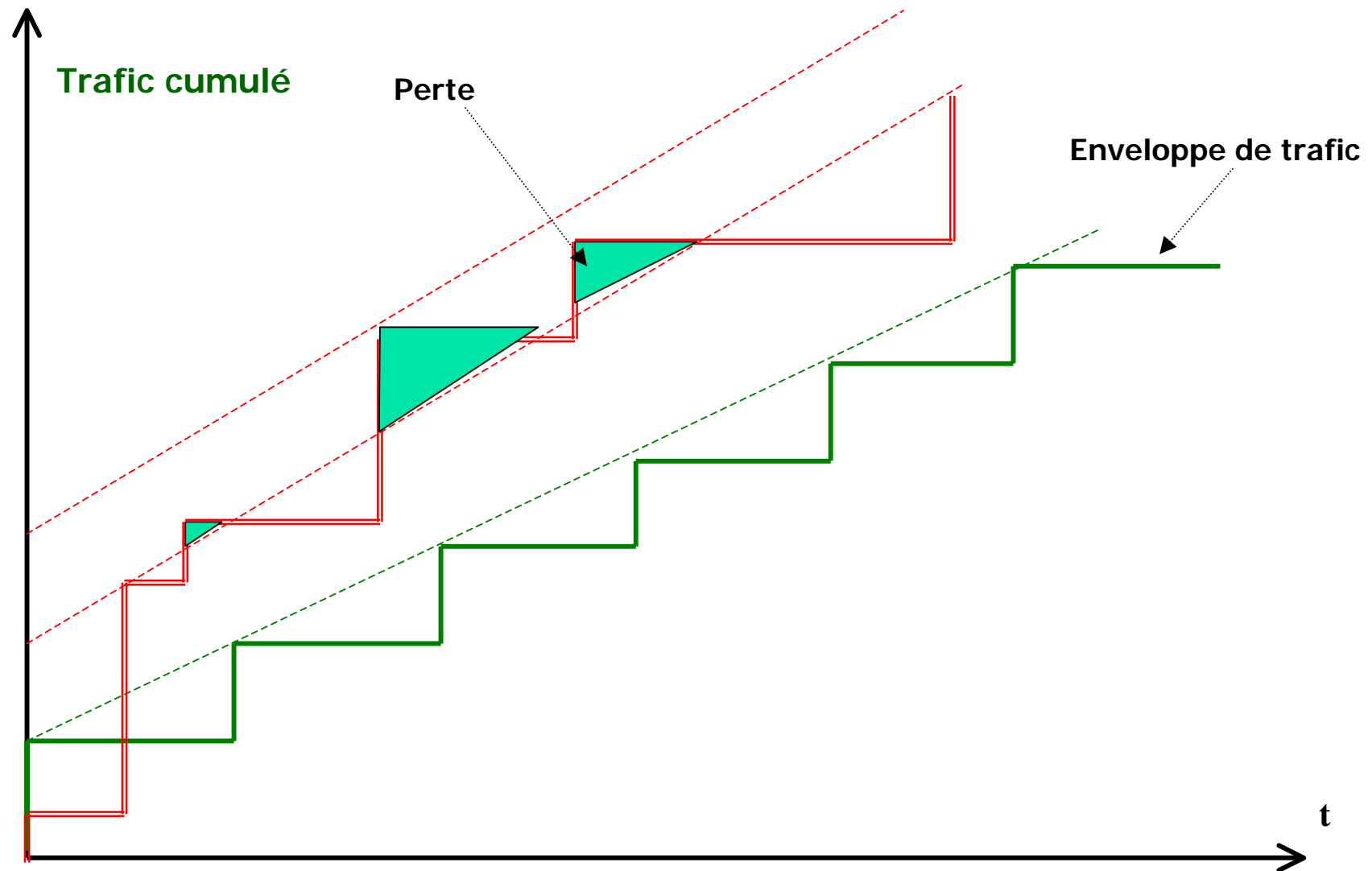
→ Délais de propagation

Délais négligeables et/ou constants

4. Modèles de trafic 1/6



4. Modèles de trafic 2/6



4. Modèles de trafic 3/6

Modèles fréquents

- **Modèle périodique**

Période, Longueur max de paquet

- **Modèle sporadique**

X_{min} (intervalle de temps min entre deux messages successifs)

X_{ave} (intervalle de temps moyen entre deux messages successifs)

I (intervalle de temps sur lequel X_{ave} est calculé).

- **Modèle avec rafale (seau percé)**

Débit moyen d'écoulement du seau (ρ) et la taille maximale du seau (σ).

- **Modèle de trafic de l'IETF (RFC 2215)**

$$A(T) \leq \min(M + pT, \sigma + \rho T)$$

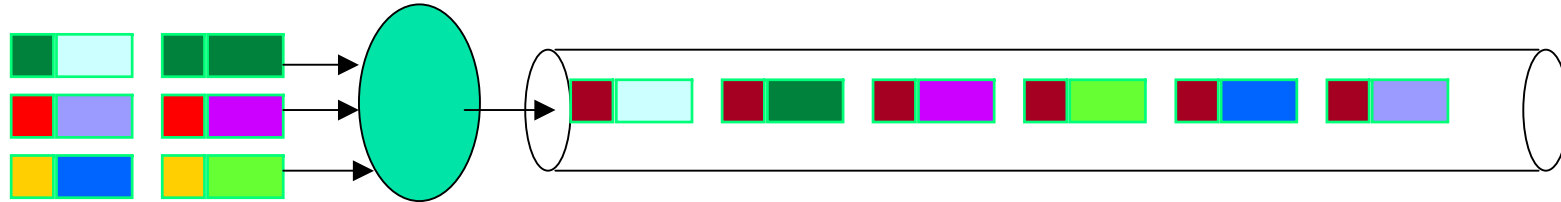
$A(T)$: borne supérieure de trafic par intervalle de temps T

σ, ρ : taille et débit du seau percé, p : débit max, M : taille max de paquet

- **Autres modèles : statistiques, probabilistes**

4. Modèles de trafic 4/6

Agrégation de flux



■ Besoins

- Plusieurs flux à transporter avec un même niveau de QoS
- Multiplexage

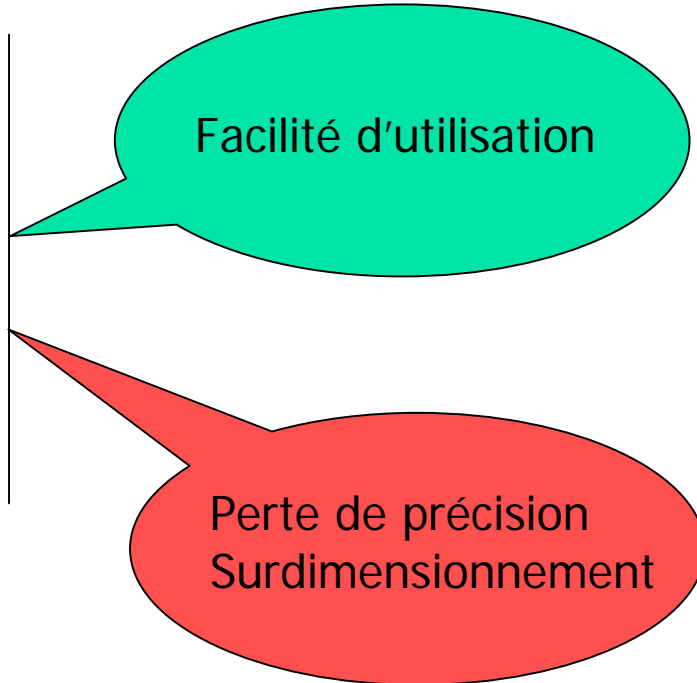
■ Difficultés

- Choix des flux à agréger
- Précision du trafic agrégé vs complexité de l'agrégation
- Modèles de dégradation de QoS pour les flux individuels

4. Modèles de trafic 5/6

Propriétés des modèles de trafic

- **Simplicité d'expression**
- **Facilité de vérification et de test**
- **Surcoût d'implantation faible**



Facilité d'utilisation

Perte de précision
Surdimensionnement

4. Modèles de trafic 6/6

Des problèmes à résoudre

- Modèles statistiques efficaces
- Combinaison de modèles pour l'agrégation de flux
- Compromis : Complexité/Précision/Surdimensionnement

5. Contrôle d'admission 1/6

■ Objectif

- Est-ce que le nouveau flux peut affecter la QoS des flux déjà acceptés ?
- Est-ce que le nœud peut offrir la QoS requise par le nouveau flux ?
- Est-ce que le nouveau flux a le droit d'utiliser les ressources du nœud ?
- Est-ce que tous les nœuds à traverser acceptent le nouveau flux ?

■ Informations utilisées

- Caractéristiques du nouveau trafic et de la QoS demandée
- Etat et historique du réseau
- Dates de fin des trafics déjà acceptés
- Perturbations éventuelles de la QoS des trafics déjà acceptés
- Politique d'utilisation des ressources

■ Le CA peut se faire sur la base de connexion ou de SLA

5. Contrôle d'admission 2/6

Exemples de Contrôle d'admission

■ **CA pour WFQ** $r_{new} + \sum_{i=1}^n r_i \leq r$

■ **CA pour DEDD** $D_{new} \geq \sum_{i=1}^m DL_i$

■ **CA statistique**

$$BP_{duLien} - \left(\sum_{i=1}^n \rho_i \right) + \left(\left(\sqrt{\sum_{i=1}^n \sigma_i^2} \right) \left(\sqrt{-2 \ln(\varepsilon) - \ln(2\pi)} \right) \right) \geq BP_{demandée}$$

5. Contrôle d'admission 3/6

■ Coût et performance du CA : dépendent des trafics

■ Caractérisation d'un trafic

- Trafic périodique : aisé
- Trafic apériodique
 - Distribution des instants d'arrivée selon quelle loi (poisson, ...) ?
 - Taille maximale des avalanches ?
 - Durée minimale d'avalanche ?
 - Distribution de la taille des avalanches ?
 - Distribution des pertes de messages ?
 - Corrélation entre les paquets (pour autoriser les pertes) ?
 - Souvent difficile à modéliser :
 - choix de paramètres pour "convenance mathématique"
 - Reste beaucoup à faire pour modéliser le trafic aléatoire/sporadique

5. Contrôle d'admission 4/6

Propriétés

(à prendre en compte durant la conception d'un CA)

- Décisions incrémentales (ne pas toujours considérer tous les flux)
- Exactitude (compliquée à cause des phénomènes aléatoires)
- Complexité
 - Problème de la diversité des modèles de flux
 - Utilisation en ligne sans surcoût important
- Flexibilité
 - Problème de la diversité des modèles de flux
- Passage à l'échelle

5. Contrôle d'admission 5/6

Contrôle d'admission basé sur les mesures

- Type de CA de plus en plus étudié
- Données de départ
 - Beaucoup d'utilisateurs incapables de décrire clairement leurs trafics
 - Beaucoup d'utilisateurs ne demandent pas des garanties absolues
- Principe du CA
 - Estimer progressivement le modèle de trafic
 - Utiliser les mesures pour raffiner le modèle
- Problèmes
 - Que faut-il mesurer ? Quand ? Où ?
 - Comment définir progressivement des modèles de trafic ?
 - Comment évaluer l'apport par rapport au CA sans mesure ?

5. Contrôle d'admission 6/6

Des problèmes à résoudre

- Caractérisation approximative des flux et complexité du CA
- Compromis entre complexité et performance
- Elaboration et maîtrise de CA statistique
- CA adapté aux réseaux sans fil
- Contrôle d'admission interdomaine

6. Allocation de ressources 1/3

Besoins et stratégies

- **Ressources : Bande passante, mémoire, CPU...**
- **Allocation de ressources** \Rightarrow Politique d'allocation (droits, coûts...)
- **Stratégies d'allocation de ressources**
 - Non statistiques (statiques)
 - Allouer une capacité maximale
 - Adaptée au trafic CBR
 - Risque de sous-utilisation du réseau
 - Statistiques (dynamiques)
 - Allocation non basée sur le débit maximum de la connexion
 - La somme des débits des connexions acceptées peut être supérieur à celui des ports de sortie du nœud
 - Adaptée à des flux variables
 - Difficulté de prédire la garantie de QoS

6. Allocation de ressources 2/3

Négociation

■ Stratégie sans négociation

- rigide (tout ou rien)
- sûre

■ Stratégie avec négociation

- au moment d'établissement de connexion...
- flexible
- complexe

■ Stratégie avec renégociation

- s'adapter au réseau à tout moment
- transmettre à moindre coût
- très complexe

6. Allocation de ressources 3/3

Des problèmes à résoudre

- Réserveation et utilisation de ressources pour les flux agrégés
- Techniques de réserveation de ressources dans les réseaux mobiles
Comment réserver des ressources sur des chemins dont les nœuds sont mobiles ?
- Spécification et signalisation des possibilités d'adaptation des applications
- Protocoles de négociation efficaces

7. Routage 2/3

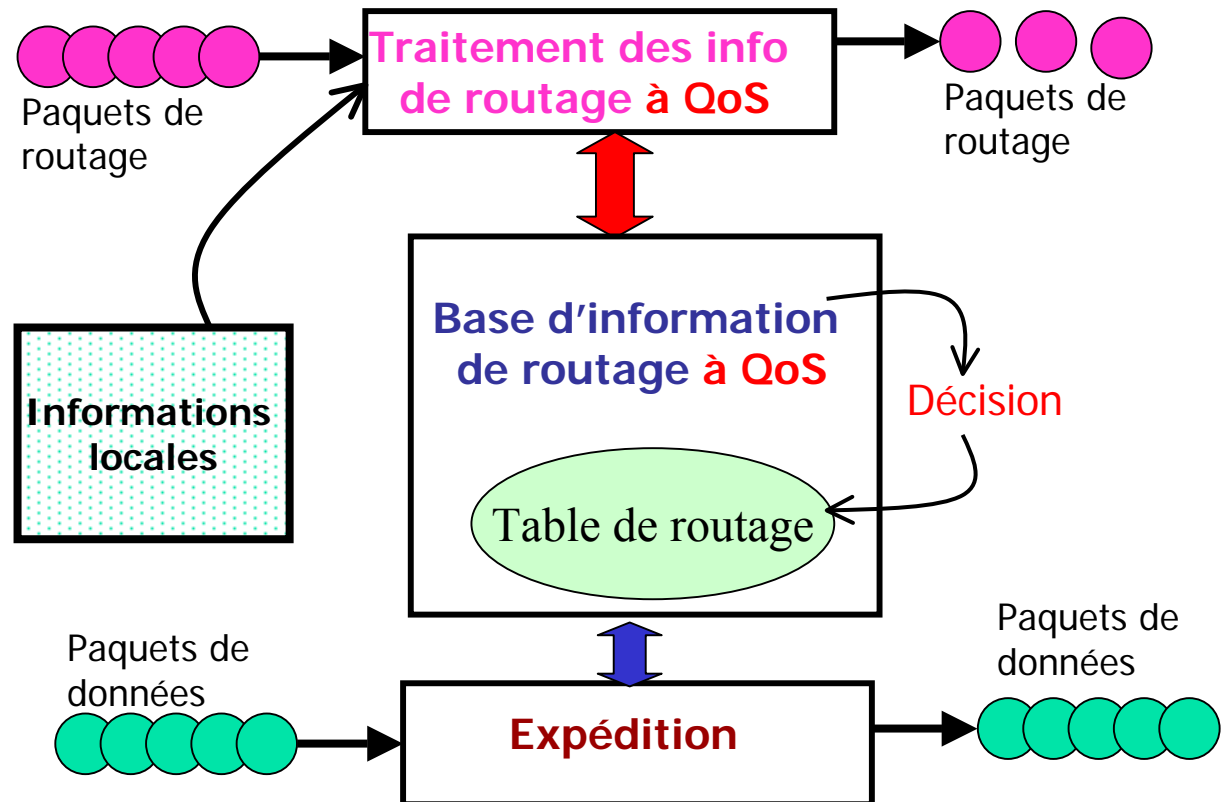
Fonctions et composants du routage

- Collecter des informations sur l'état du réseau
- Trouver le **meilleur chemin** pour un nouveau flux

- **Protocole de routage**

- **Base d'information de routage**

- **Algorithme de routage**



7. Routage 3/3

Algorithmes de routage

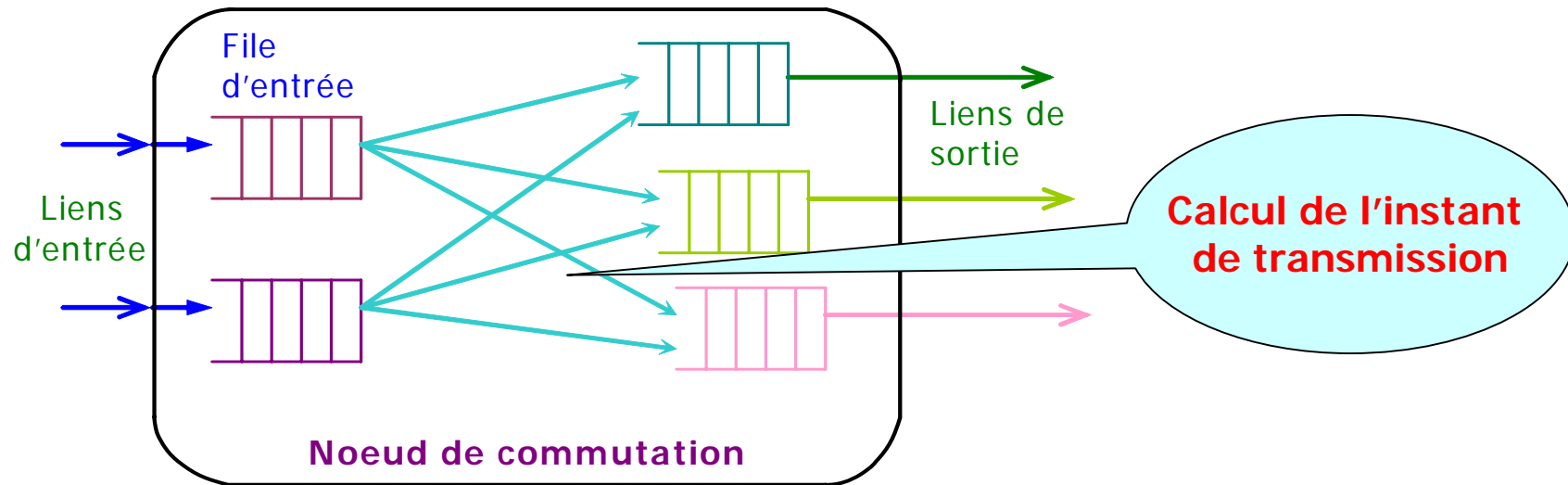
- De nombreux algorithmes de routage existent
- Classes d'algorithmes de routage
 - Contrôle : par la source | distribué | hiérarchique
 - Récepteurs : Unicast | Multicast | Anycast
 - QoS : Oui | Non
- Critères de classement d'algorithmes de routage
 - Contraintes prises en compte (délai, gigue, bande passante, ...)
 - Stratégie du routage (par la source, distribué, hiérarchique)
 - Complexité de l'algorithme
 - Complexité de la communication pour maintenir les informations d'état

7. Routage 3/3

Des problèmes à résoudre

- Métriques à considérer (*sensibles à la QoS*)
- Précision de la topologie et de la charge du réseau
- Minimisation du surcoût (*compromis : efficacité/coût*)
 - Echange d'informations d'état (*fréquence adéquate*)
 - Traitement (*pré-calcul de chemins*)
 - Table de routage (*hiérarchique*)
- Maîtrise et prédiction des congestions
- Complexité et efficacité du protocole de routage (QOSPF...)
- Lien routage – réservation de ressources
- Routage inter domaines
 - Représentation commune des états, Politique d'allocation des ressources
 - Interopérabilité des protocoles/algorithmes de routage
- Routage dans les réseaux sans fil
- Intelligence/adaptabilité du routage

8. Ordonnancement 1/6



■ Propriétés

- Types de garantie : meilleur effort, statistique, déterministe
- Caractéristiques de QoS : délai, BP, perte...
- Isolation des flux
- Équité : distribution de la bande passante excédentaire
- Complexité
- Scalability

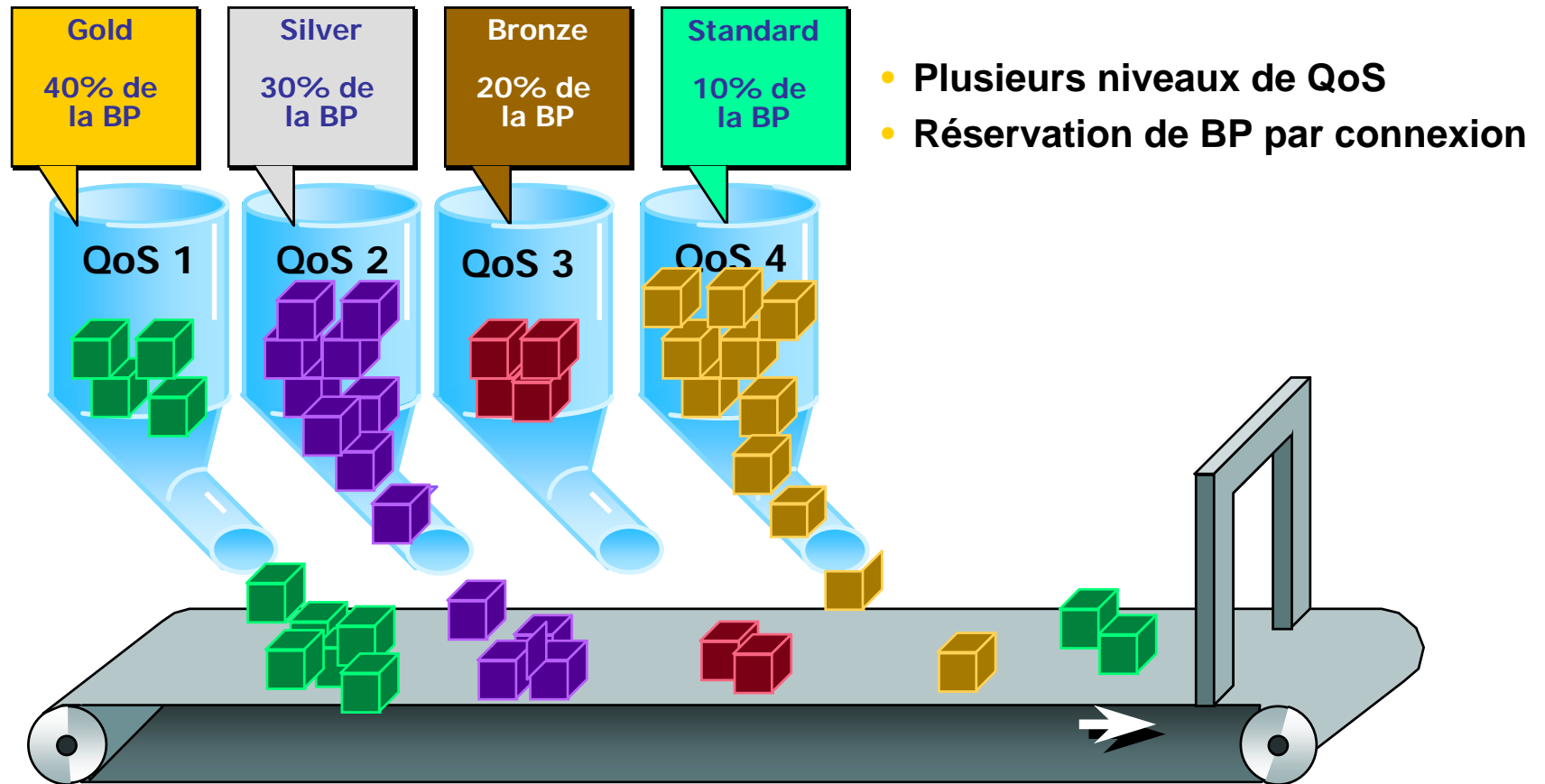
8. Ordonnancement 2/6

Algorithmes d'ordonnancement (disciplines de service)

- WFQ « Weighted Fair-Queueing »
 - Plusieurs extensions de WFQ existent (FWFQ, SCFQ, SFQ, WFWFQ, ...)
- Stratégies Round Robin
 - RR simple, WRR (weighted RR) , HHR (Hierarchical RR)...
- Delay Earliest Due-Date
- Jitter Earliest Due Date
- Class Based Queuing
- Autres disciplines

8. Ordonnancement 3/6

Exemple : WFQ



8. Ordonnancement 4/6

■ Temps virtuel de transmission du paquet k de la connexion i

$$F_i^k = \max(F_i^{k-1}, V(a_i^k)) + \frac{L_i^k}{r\omega_i}$$

r : débit du lien, L_i^k : taille du paquet, ω_i part de BP allouée à la connexion i.

$$V(t + \tau) = V(t) + \frac{r \times \tau}{\sum_{i \in B_j} \omega_i}$$

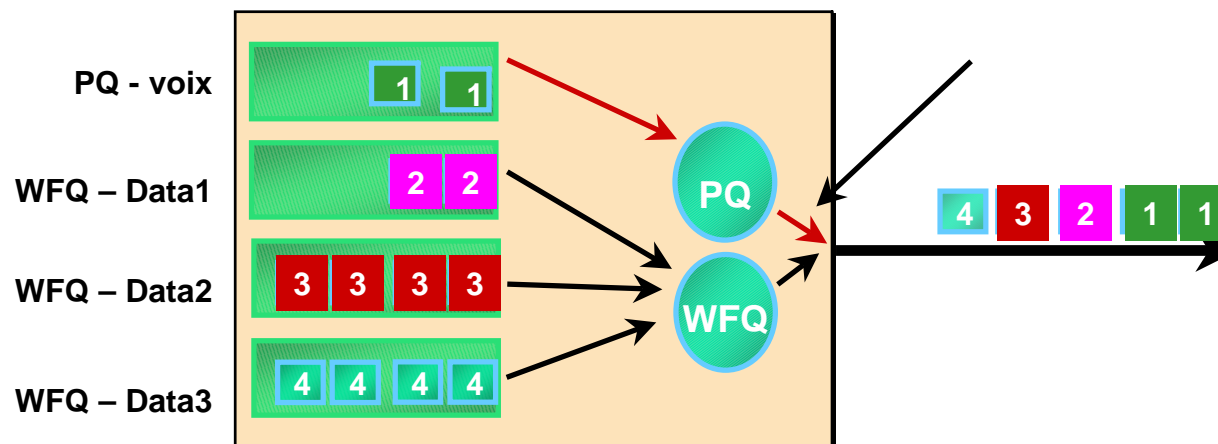
B_j ensemble de connexions actives durant l'intervalle de temps $[t .. t + \tau]$

■ Avantages

- Garantie de débit r_i pour la connexion i : $\frac{r \times \omega_i}{\sum_k \omega_k}$
- Garantie de délai r_i pour la connexion i : $\frac{L_i}{r_i} + \frac{L_{\max}}{r}$
- Isolation des flux
- Équité pour la distribution de BP excédentaire

8. Ordonnancement 5/6

Combinaison WFQ-PQ



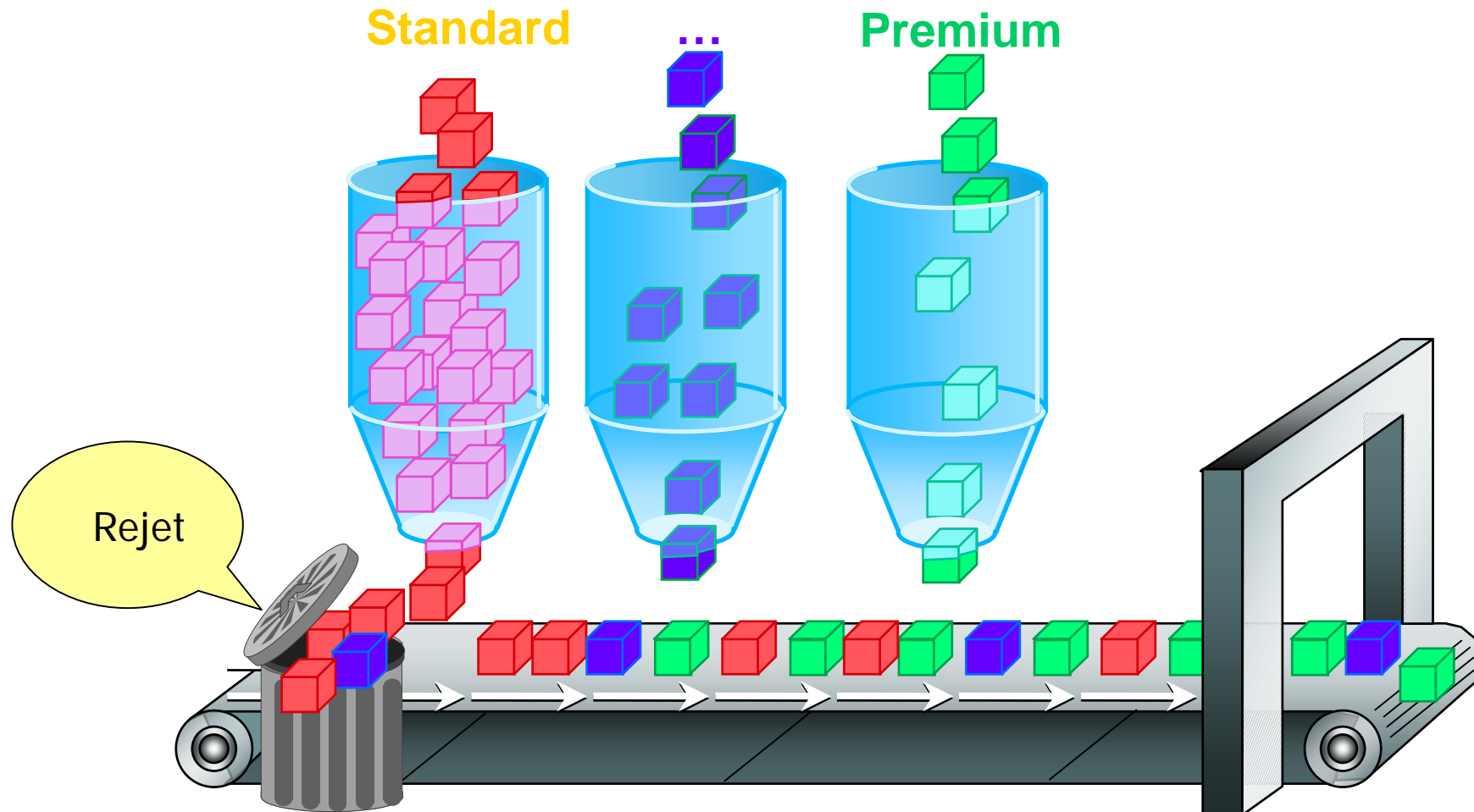
8. Ordonnancement 6/6

Des problèmes à résoudre

- Complexité des algorithmes en cas de contraintes multiples
Maîtrise des heuristiques
- Modèles analytiques de bornes
- Adaptation des algorithmes aux changements de trafic
- Modèles de conception et d'analyse de composants
Routage/ordonnancement/CAC/Réservation de ressources
- Algorithmes d'ordonnancement et réseaux sans fil
Effets des capacités de lien dynamiques
- Effets des mécanismes de sécurité sur les performances de l'ordonnancement

9. Contrôle de congestion 1/4

- Flux aléatoires + Mémoire limitée \Rightarrow Possibilité de congestion



9. Contrôle de congestion 2/4

Stratégies

- Politique de suppression de paquets

Quand supprimer des paquets ?

Quels paquets supprimer ? Dans quelles conditions ?

- Méthodes : réactives vs préventives

- Techniques de contrôle de congestion

- RED ("Random Early Detection"), WRED (Weighted RED)
- ECN ("Explicit Congestion Notification")
- Autres

9. Contrôle de congestion 3/4

Technique RED (Random Early Detection)

■ RED : technique la plus populaire pour l'évitement de congestion

■ Principes de base de RED

– Estimation d'une taille moyenne de queue Q :

$$Q_n = (1-\alpha)Q_{n-1} + \alpha q_n \quad \alpha \ll 1$$

– Rejet probabiliste en fonction de la taille moyenne de queue

- Si $Q < \text{Seuil_min}$: pas de rejet
- Si $Q > \text{Seuil_max}$: rejet de tous les paquets
- Si $\text{Seuil_min} \leq Q \leq \text{Seuil_max}$: rejet avec une probabilité $p(Q)$

$p(Q)$ croît selon :

$$p(Q) = P_{\max}(Q - \text{Seuil_min}) / (\text{Seuil_max} - \text{Seuil_min})$$

9. Contrôle de congestion 4/4

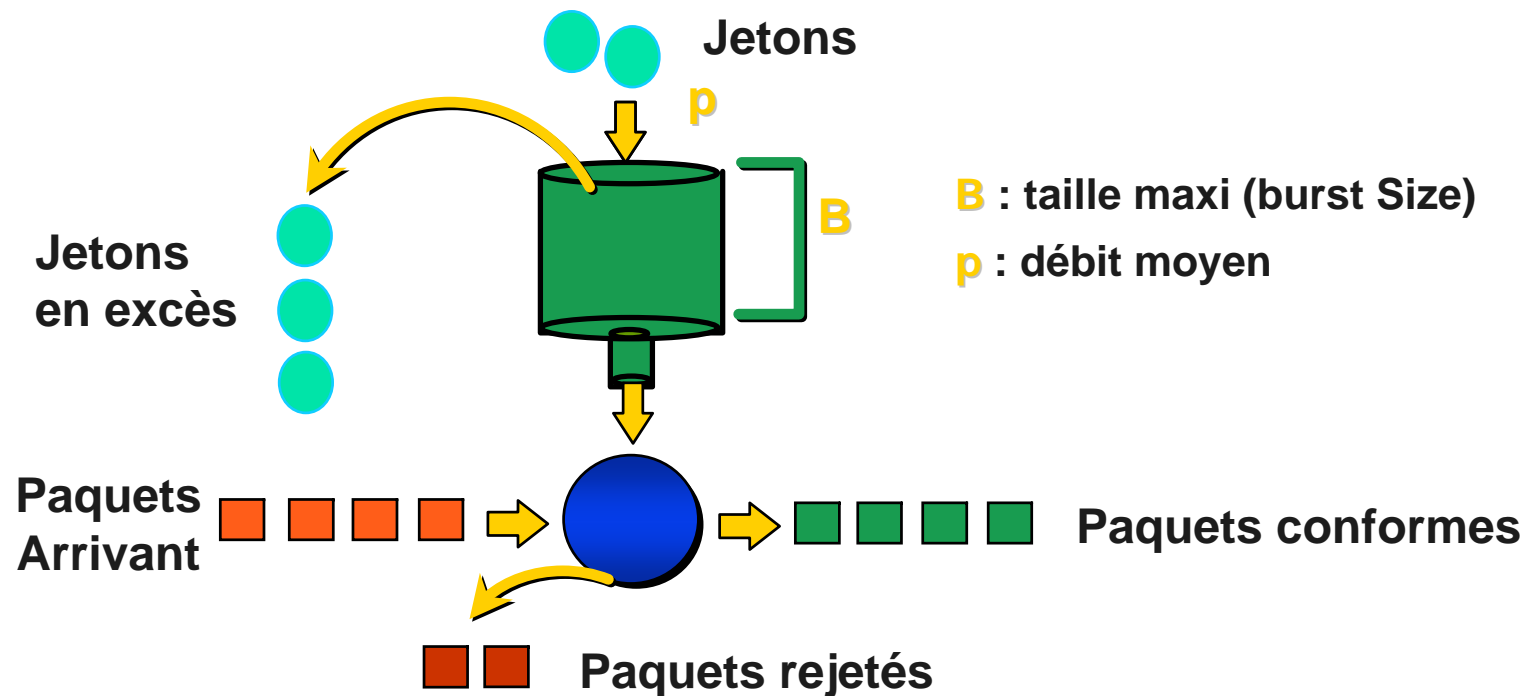
Des problèmes à résoudre

- Utilisation des techniques de prévention
 - Choix des indicateurs de charge et leur mesure
 - Choix des seuils de congestion (statique, dynamique)
 - Agrégation et dégradation de QoS pour les flux individuels
 - Maîtrise de l'analyse statistique de la charge dynamique
- Applications adaptatives à l'écoute du réseau
- Gestion de buffers adaptive en fonction des applications
- Combinaison : CA/Ordonnancement/Contrôle de congestion

10. Autres fonctions 1/5

Contrôle de trafic

- Protection du réseau contre les abus des utilisateurs
- Contrôle de conformité du trafic de l'utilisateur
- Deux techniques : seau percé et seau à jetons



10. Autres fonctions 2/5

Classification et marquage de paquets

■ Classification de paquets

- Mettre le paquet arrivant dans la bonne file d'attente
- Critères de classification (@IP, n° de port, ...)
- Coût de la classification

■ Marquage de paquets

- Marquer le paquet arrivant pour qu'il obtienne une certaine QoS
- Fonction au cœur de DiffServ
- Marquer une seule fois à l'entrée du réseau OU à chaque nœud OU ...
- Coût du marquage

10. Autres fonctions 3/5

Façonnage de trafic (« traffic shaping »)

■ Objectif

Contrôle de la quantité et du rythme d'arrivée des informations au réseau

■ Deux techniques d'implantation du façonnage du trafic

- Seau percé (« leaky bucket »)
- Seau à jetons (« token bucket »)

■ Problèmes

- Surcoût du « traffic shaping »
- Où placer les « shapers » ? À l'entrée ? Partout ?
- Agrégation de flux et shaping

10. Autres fonctions 4/5

Signalisation

- Echange d'informations d'état et de réservation
- Quelles informations échanger ? A quelle fréquence ?
- **Protocoles de signalisation**
 - RSVP (ReServation Protocol)
 - SBM (Subnet Bandwidth Manager)
 - LDP (Label Distribution Protocol)

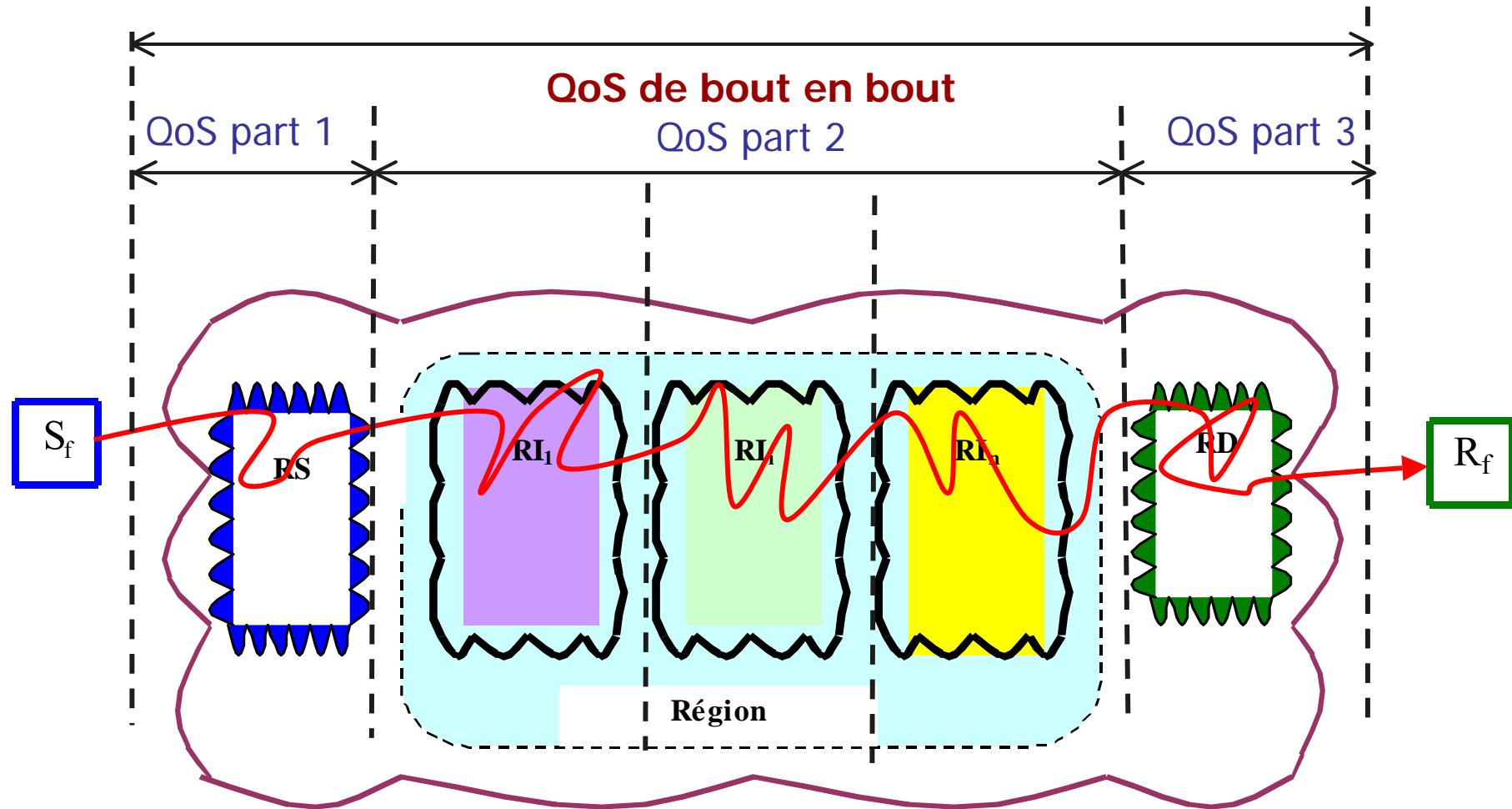
10. Autres fonctions 5/5

Contrôle et supervision de la QoS

- Mettre en place des indicateurs de suivi de la QoS fournie
- Comparaison de la QoS fournie avec celle requise/promise
- Adaptation des ressources réservées
- Notification/Alerte de l'utilisateur et déclenchement de la renégociation

11. Mapping de QoS 1/3

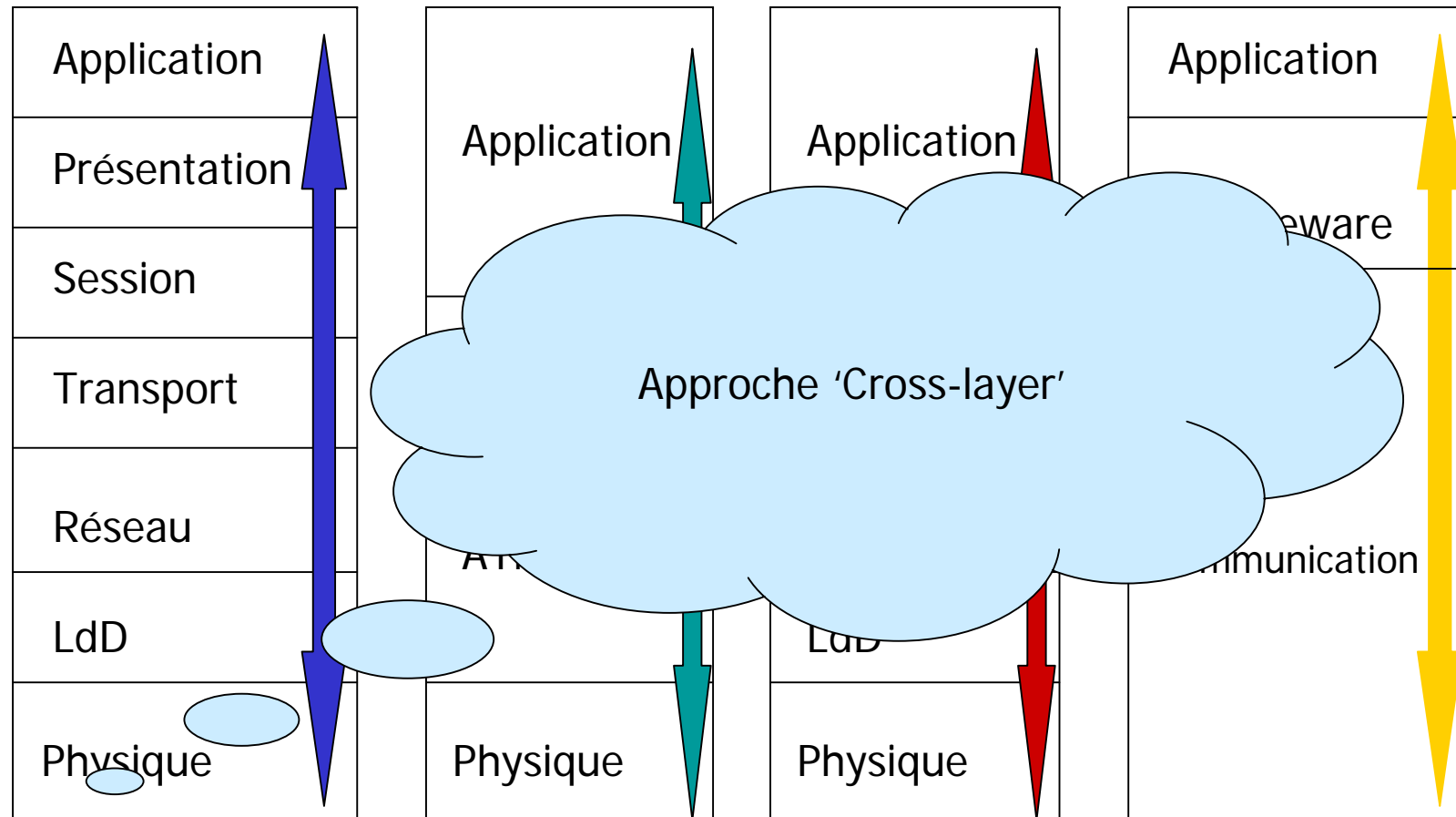
Mapping Inter-Domaines



■ Chaque Composant a sa propre vue de la QoS \Rightarrow besoin de mapping

11. Mapping de QoS 2/3

Mapping Inter-couches



■ Chaque **Couche** a sa **propre vue** de la QoS \Rightarrow **besoin de mapping**

11. Mapping de QoS 3/3

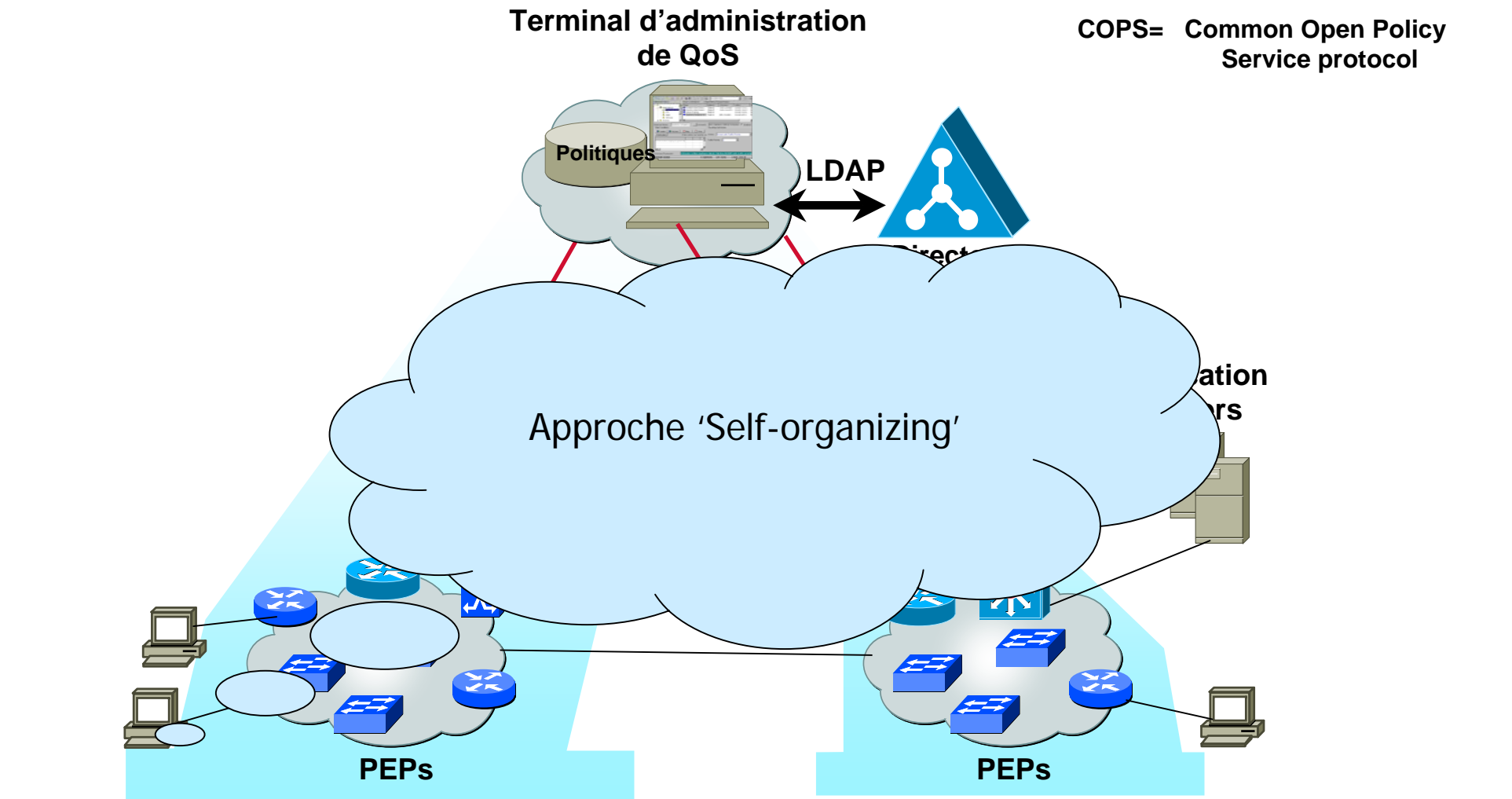
■ Des problèmes à résoudre

- Règles de mapping systématique
- Outils pour réaliser et valider le mapping

12. Politique de QoS 1/2

- Garantie de QoS → Nécessité d'une politique globale de gestion des ressources et des utilisateurs
- Répondre à des questions liées aux décisions :
 - Qui a le droit d'utiliser quel service / quelle ressource ?
 - Comment configurer/reconfigurer les équipements ?
Choix des algo de routage, d'ordo, de CA...
 - Comment faire payer les utilisateurs ?
 - Comment surveiller le réseau ?
 - Comment réagir à un excès de trafic ?
 - Comment anticiper les congestions?
 - ...
- Composants de politique de QoS (selon l'IETF)
 - BD de politiques + Points de décision (PDP) +
 - Points d'exécution de politique (PEP) + Protocoles (ex. COPS, LDAP)

12. Politique de QoS 1/2



13. Conclusion 1/2

■ Importance croissante de la QoS

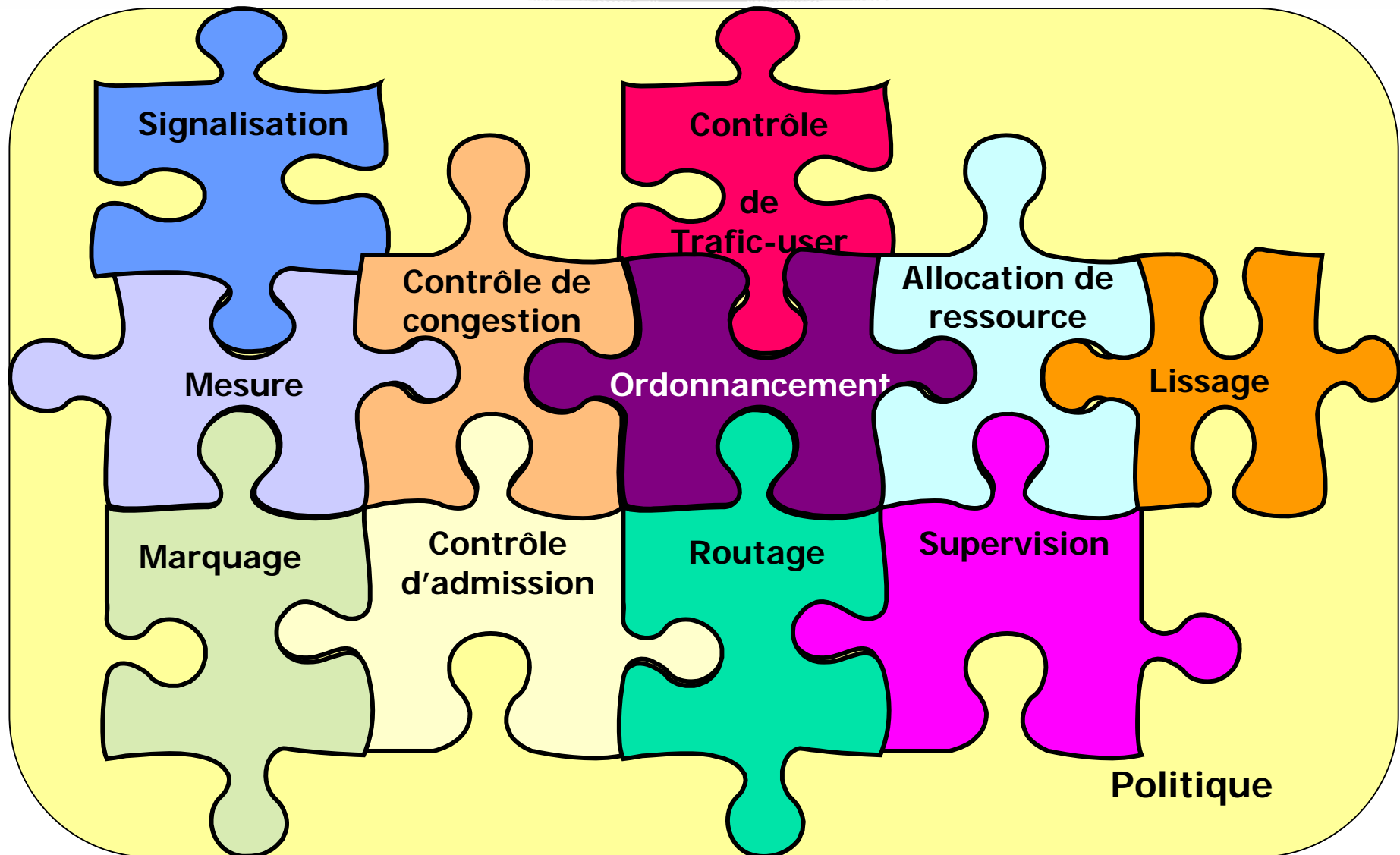
■ QoS Partout

- Demandée par les utilisateurs
- Opportunité pour les " business compagnies "
- Déploiement à grande échelle : coûte cher

■ Mise en œuvre de la QoS

- Beaucoup de mécanismes/fonctions
- Difficulté de paramétrage des fonctions
- Coût vs efficacité des mécanismes

13. Conclusion 2/2



Merci

Questions