

R6.B.05 – OPTIMISATION DE SERVICES COMPLEXES

PARTIE III

Edward Staddon

Edward.Staddon@univ-ubs.fr

Université Bretagne Sud, IUT de Vannes, Département Informatique



PLAN DU COURS

- Optimisation de l'accès aux ressources
- Réseau de diffusion de contenus
- Introduction à la qualité de service

PLAN DU COURS

- Optimisation de l'accès aux ressources
- Réseau de diffusion de contenus
- Introduction à la qualité de service



INTRODUCTION À LA QUALITÉ DE SERVICE

QUALITÉ DE SERVICE

→ QoS

- Utilisation de mécanismes / technologies réseau pour
 - Contrôler le trafic
 - Assurer la performance des applications critiques
 - Capacité réseau limitée
- Permet l'ajustement du trafic réseau global
 - Hiérarchie des applications et trafic
 - Bande passante dédiée, gigue contrôlée, latence faible, etc...
- Généralement appliqué aux réseaux de grande envergure
 - Trafic a forte consommation de ressources
 - Petit réseau → moins de trafic → moins d'efficacité
- Approche requise pour plusieurs services courants
 - IPTV, jeux en ligne, streaming, Visio, VOD, VoIP, etc...



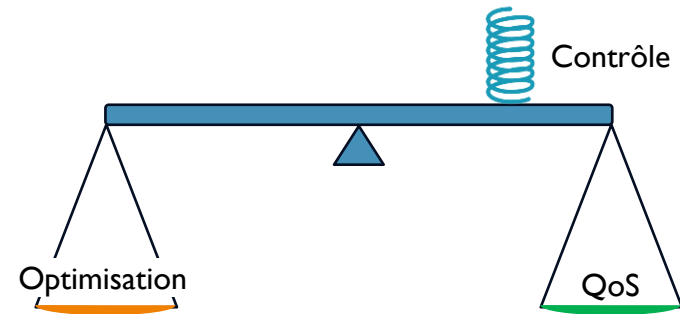
Variation du délai
de transmission

POURQUOI ?

- Ressources réseau limitées
 - Débit des liens, capacité de stockage des routeurs, etc...
- Utilisation de trafics variés
 - Temporellement et quantitativement
- Besoin d'optimisation des ressources
 - Multiplexage statistique
 - Allocation d'un débit inférieur au débit crête à chaque connexion
 - Supposition de probabilité faible d'émissions concourantes
- Problème de congestion
 - Impossible au niveau des liens
 - Contrôle d'accès organisé par la politique d'ordonnancement des équipements (scheduling)
 - Possible au niveau des routeurs
 - Espace de stockage limité (buffer)
 - Encombrement → retard → délai
 - Débordement → perte → taux de perte élevé

CONTRÔLE DE QUALITÉ

- Besoins contradictoires
 - Usagers et applications
 - Garantie de la qualité de transfert des données (QoS : taux de perte, délai, débit, etc.)
 - Opérateurs et réseau
 - Optimisation de l'utilisation des ressources
- Compromis géré par le contrôle



- Propriétés des mécanismes de contrôle :
 - Flexibilité (s'adapter au trafic)
 - Efficacité (faible complexité, peu de ressources)
 - Robustesse (continuité du service)

CONTRÔLE DE QUALITÉ

Plusieurs difficultés lors du contrôle

- **Haut débit**
 - Contrôles réactifs sont peu efficaces
 - Pendant le temps d'allé retour → beaucoup de données peuvent arriver → submersion du réseau
 - Lié à la capacité du réseau
 - Débit x délai → capacité
- **Services multiples**
 - Besoins très variés des applications
 - Taux de perte nul, faible, etc ...
 - Délai de transmission constant, variable, infini, etc...
 - Débit constant, sporadique, continument variable / par palier, etc...
- **Type de flux**
 - Définition des paramètres nécessaires pour décrire les différents services

NIVEAUX D'ANALYSE

■ Session

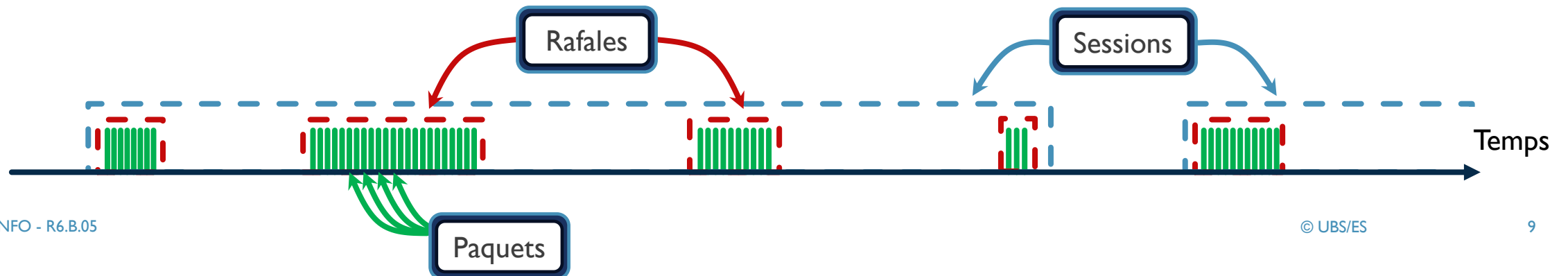
- *Nature* → variable, rafale, constante, etc...
- *Bande passante requise*
- *QoS* → taux d'erreur admissible, délai maximum, variation de la gigue, etc...
- *Échelle de temps* → secondes à jours)

■ Rafale (Burst)

- Fréquence, longueur, intensité, sporadicité
- *Un message* → plusieurs paquets
- *Échelle de temps* → milliseconde

■ Paquet

- *Échelle de temps* → microseconde



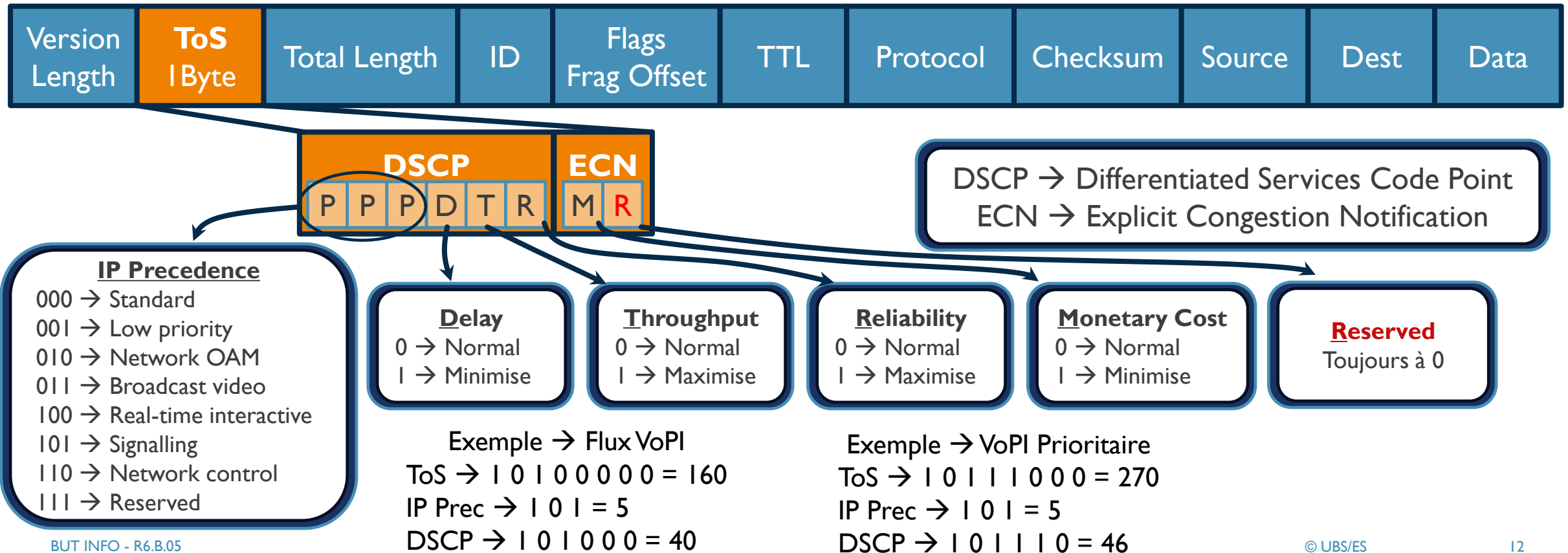
TYPES DE TRAFIC RÉSEAU MESURÉS

- **Bande passante** → *Vitesse de liaison*
 - Indique au routeur comment utiliser la bande passante
 - Ex. attribution d'un volume donnée de bande passante a différentes files d'attente pour différents types de trafic
- **Délai** → *Temps nécessaire à un paquet d'arriver à sa destination*
 - Souvent affecté par le délai de la file d'attente
 - Lié aux périodes de congestion → longues périodes d'attente
 - QoS permet d'éviter les temps d'attente → création de files d'attente prioritaires pour certains types de trafic
- **Perte** → *Quantité de données perdues suivant une perte de paquets*
 - Généralement lié à la congestion
 - QoS permet de définir quels paquets abandonner
- **Gigue** → *Vitesse irrégulière des paquets en raison d'une congestion*
 - Peut entraîner l'arrivée tardive et désordonné de paquets
 - Peut provoquer une distorsion ou des lacunes dans des fichiers audio / vidés

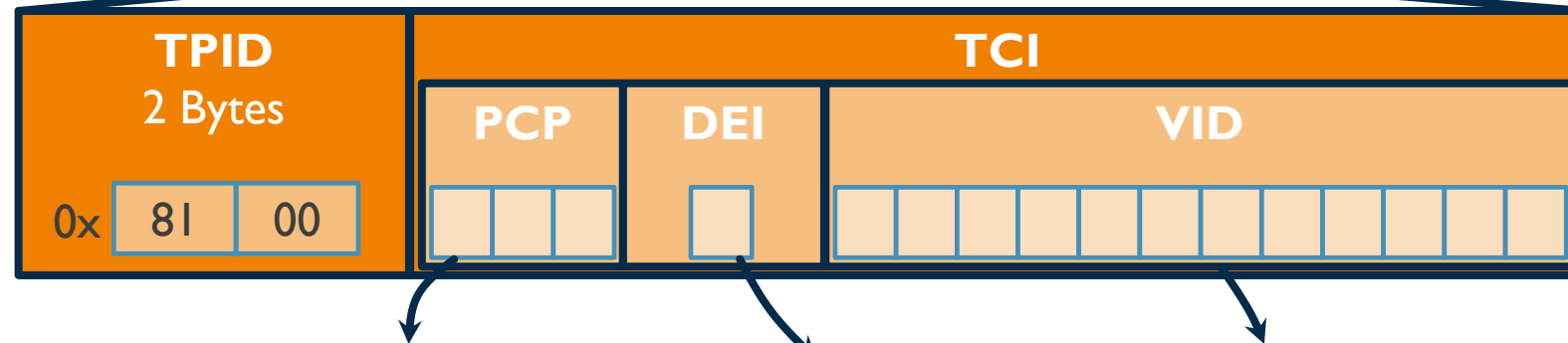
COMMENT EST-CE QUE CELA FONCTIONNE ?

- Marquage des paquets
 - Permet d'identifier le type de service concerné
- Pris en compte au niveau des routeurs
 - Création de files d'attente virtuelles distincts
 - Classement des applications par priorité
- Réservation de la bande passante
 - Aux applications sensibles
 - Aux sites web prioritaires
- Marquage → « classification »
- Plusieurs méthodes
 - Utilisation d'un champ de l'entête couche réseau
 - Ex. adresse destination IP, label du paquet MPLS (MultiProtocol Label Switching), etc...
 - Ensemble de champs sur plusieurs couches
 - Ex. champs *Protocol* et *Port Number* IP et TCP pour déterminer le type de paquet
 - analyse couteuse en profondeur

MARQUAGE IP



MARQUAGE ETHERNET



PCP

000 → Best effort (default)
001 → Background (lowest)
010 → Excellent effort
011 → Critical applications
100 → Video < 100 ms latency and jitter
101 → Voice < 10 ms latency and jitter
110 → Internetwork control
111 → Network control (highest)

DEI

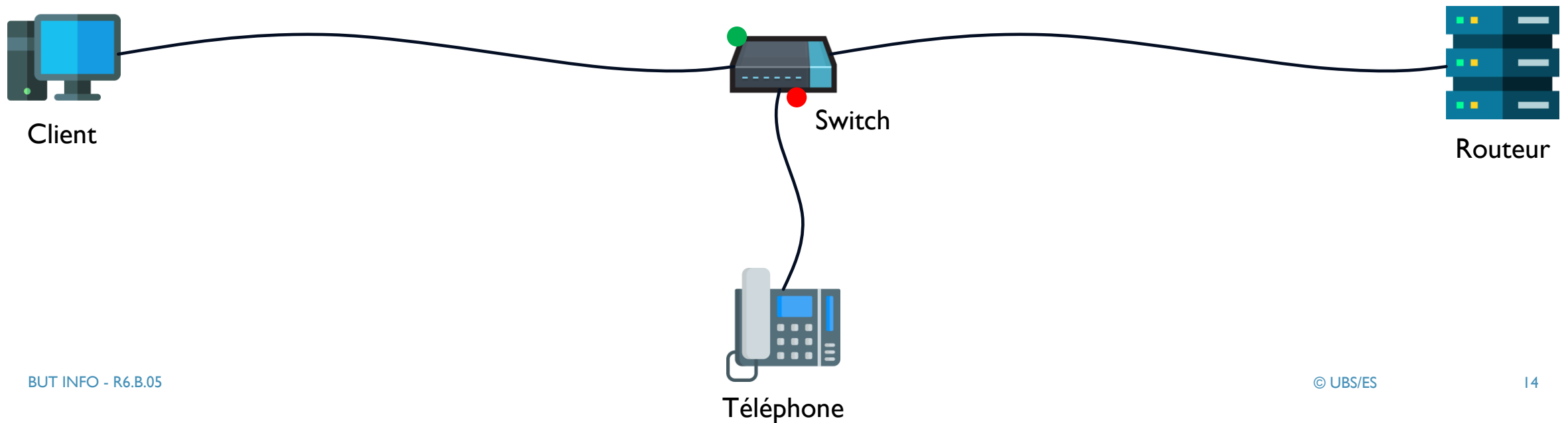
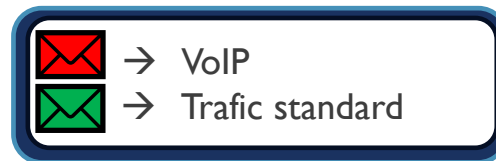
0 → No drop
1 → Drop if congestion

VID

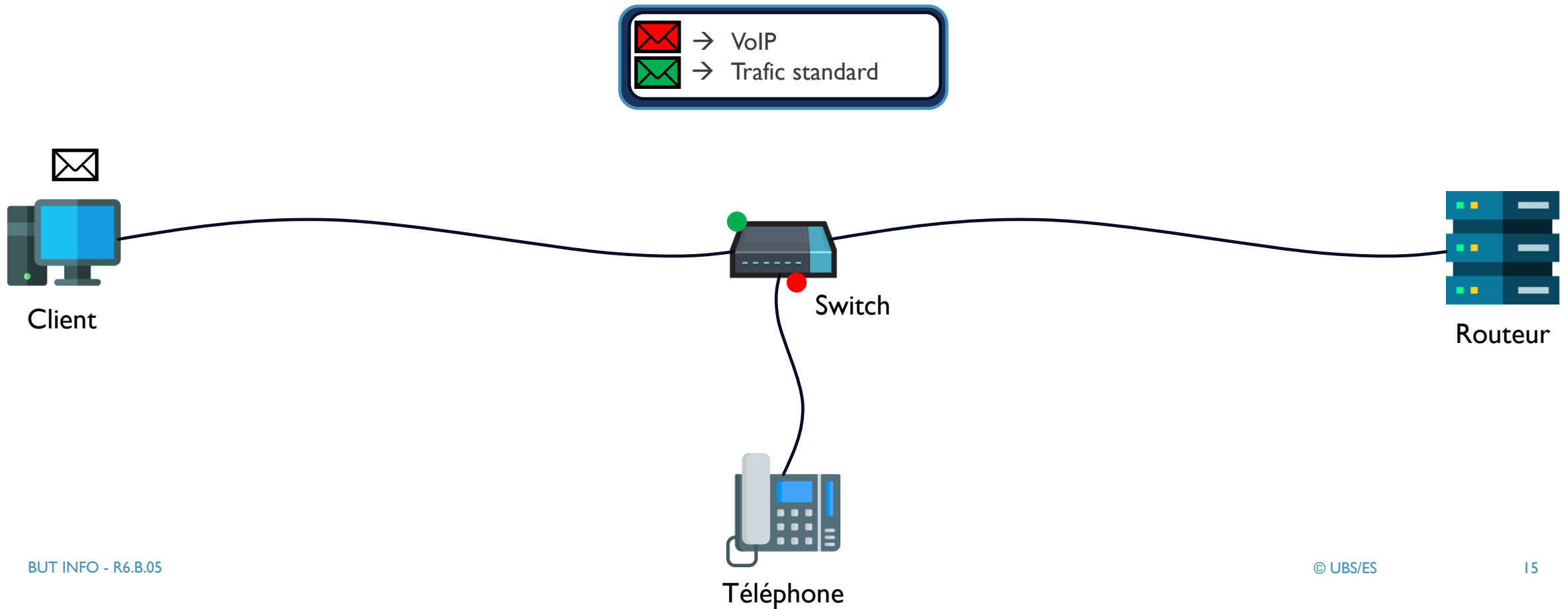
0x000 → No VLAN
0xFFF → Reserved
4 094 VLANs

TPID → Tag Protocol Identifier
TCI → Tag Control Information
PCP → Priority Code Point
DEI → Drop Eligible Indicator
VID → VLAN identifier

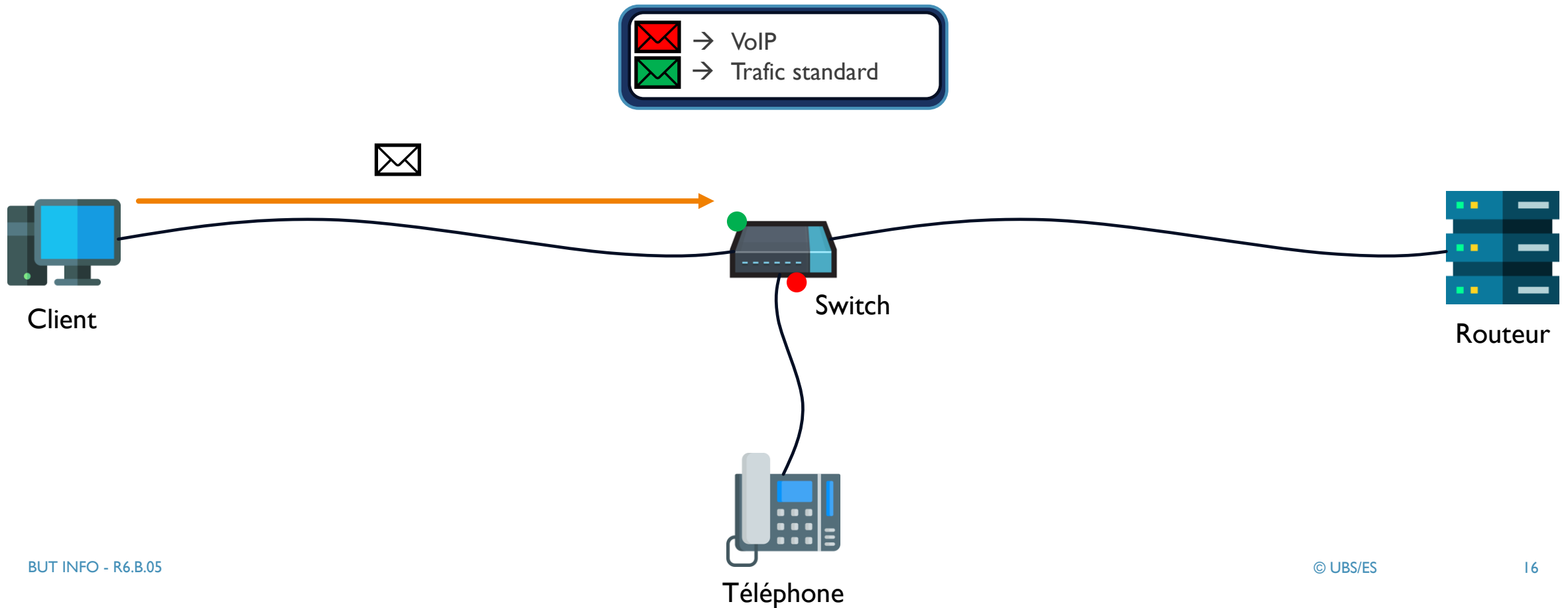
UTILISATION DU MARQUAGE



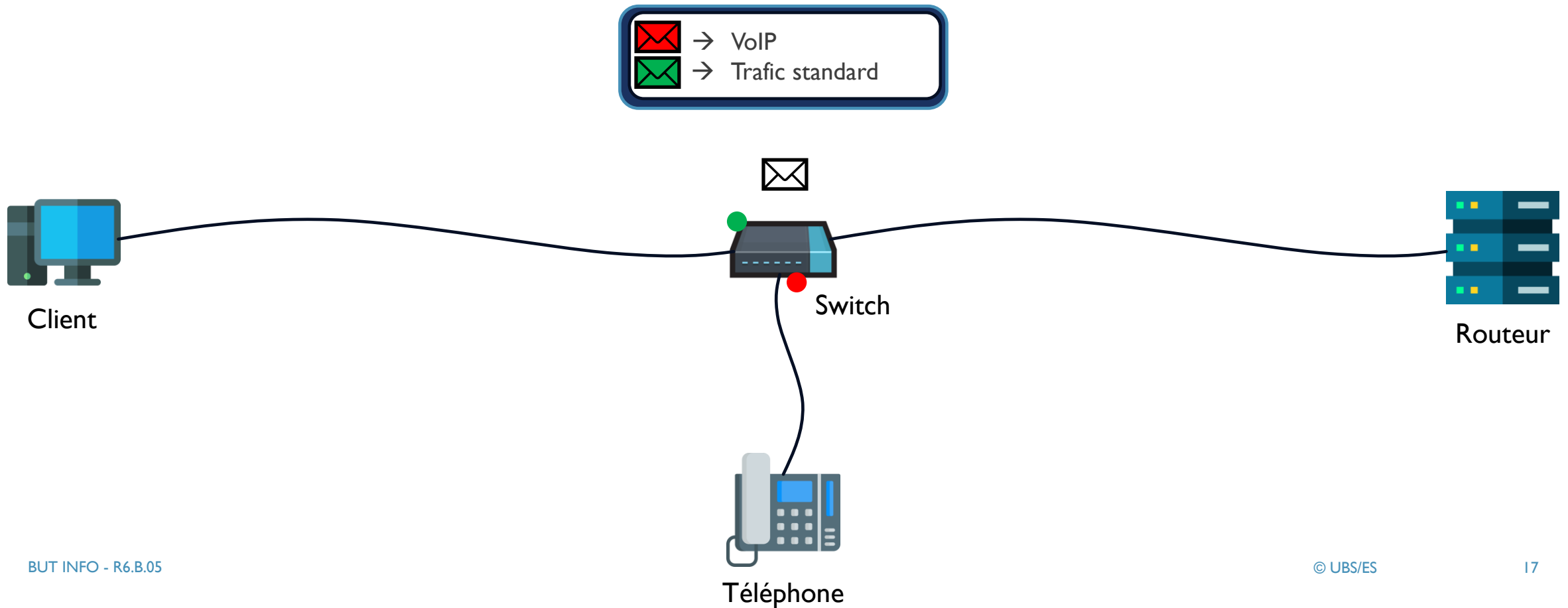
UTILISATION DU MARQUAGE



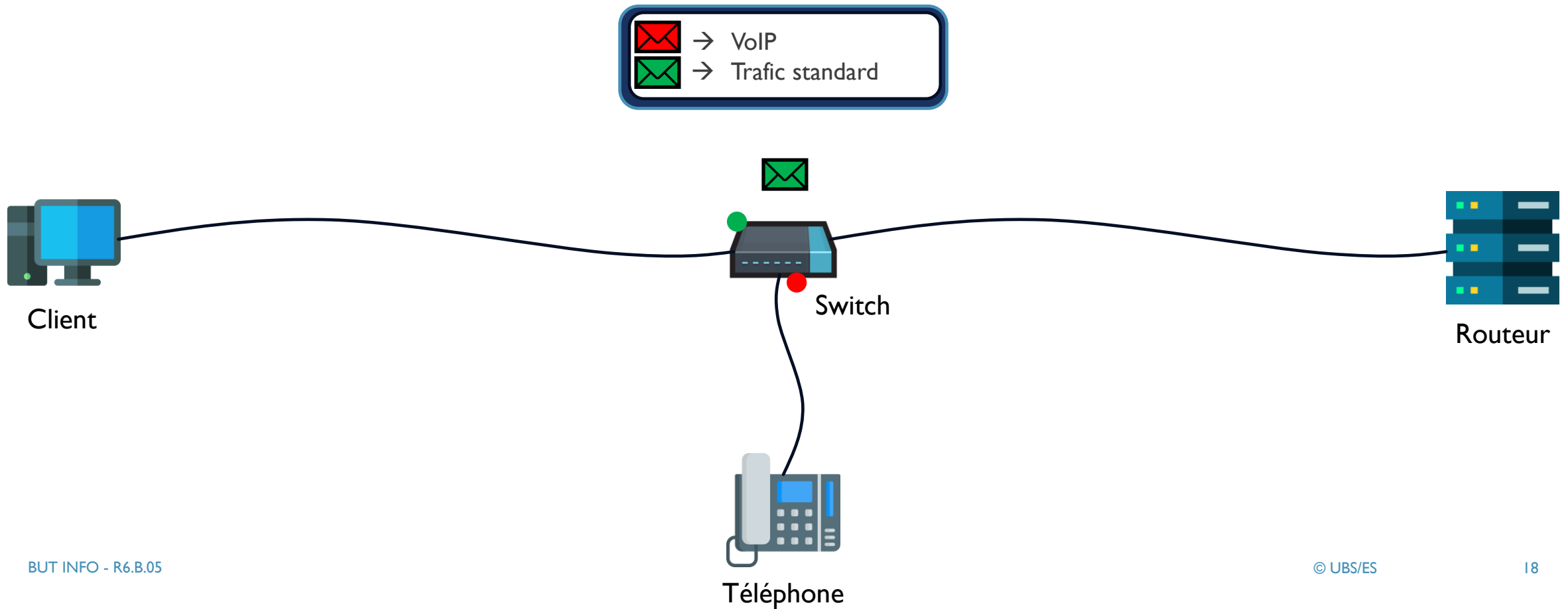
UTILISATION DU MARQUAGE



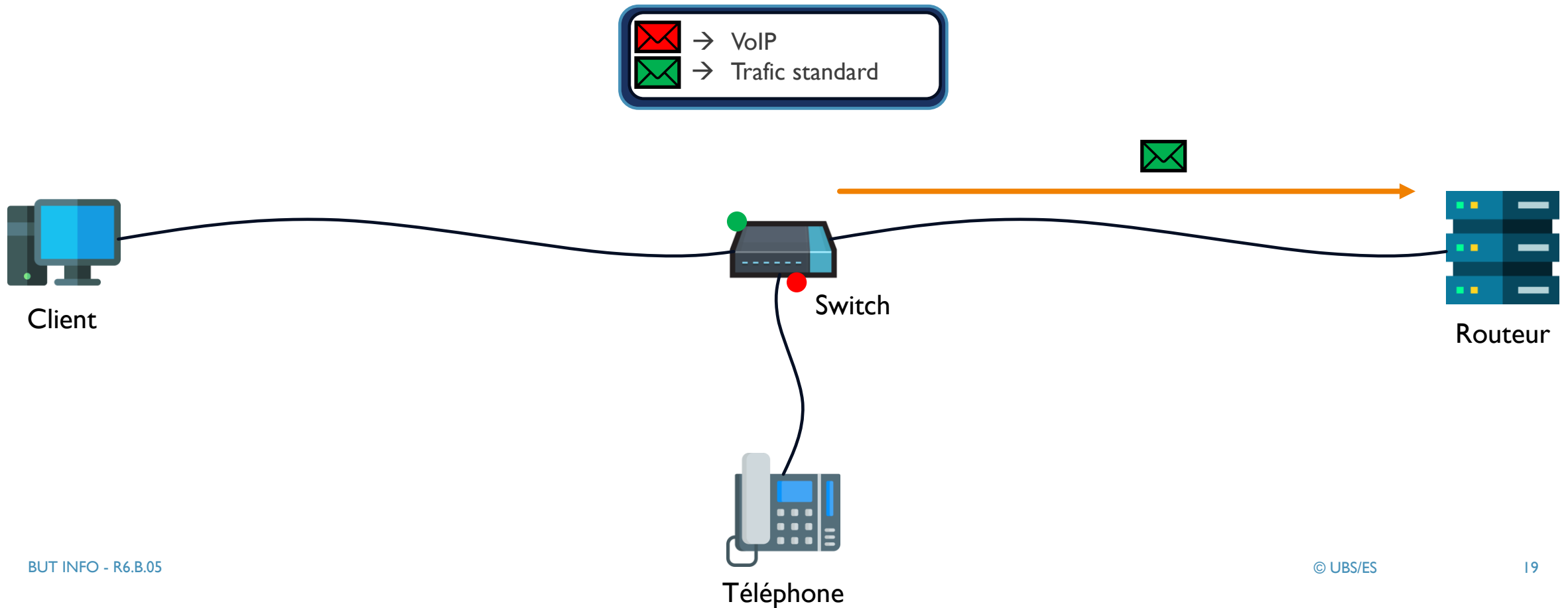
UTILISATION DU MARQUAGE



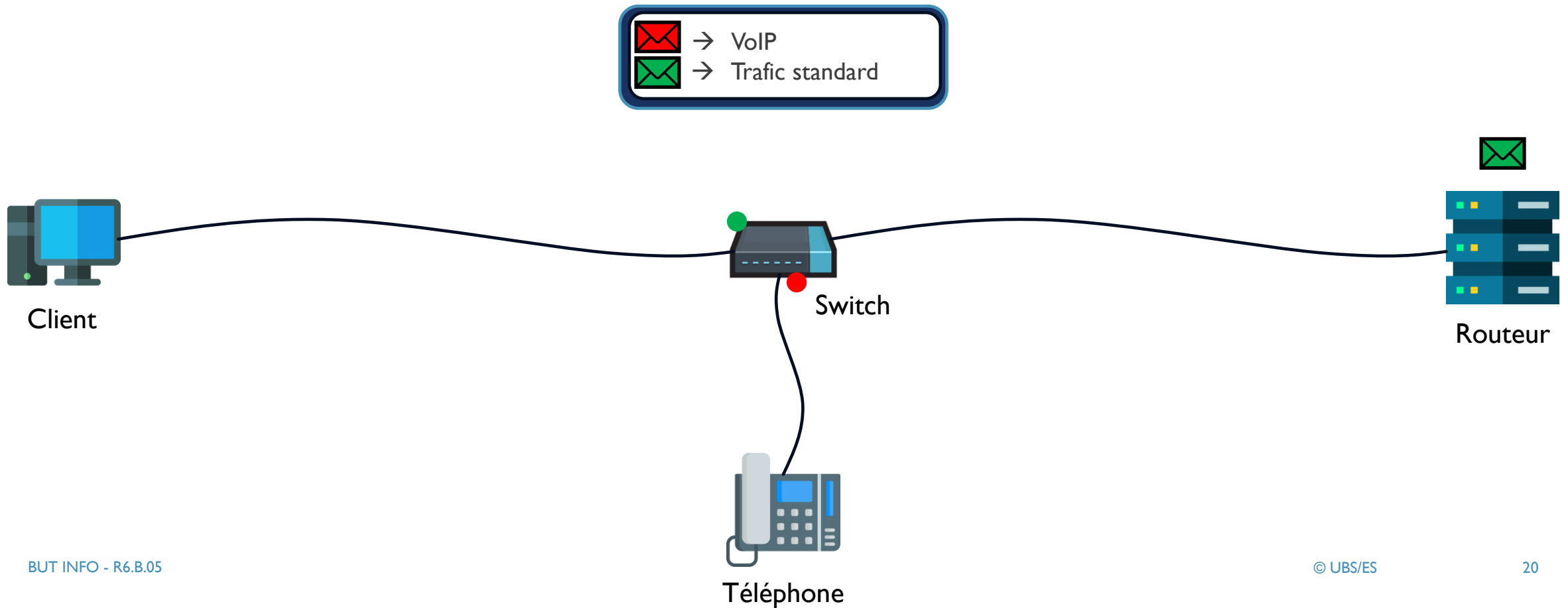
UTILISATION DU MARQUAGE



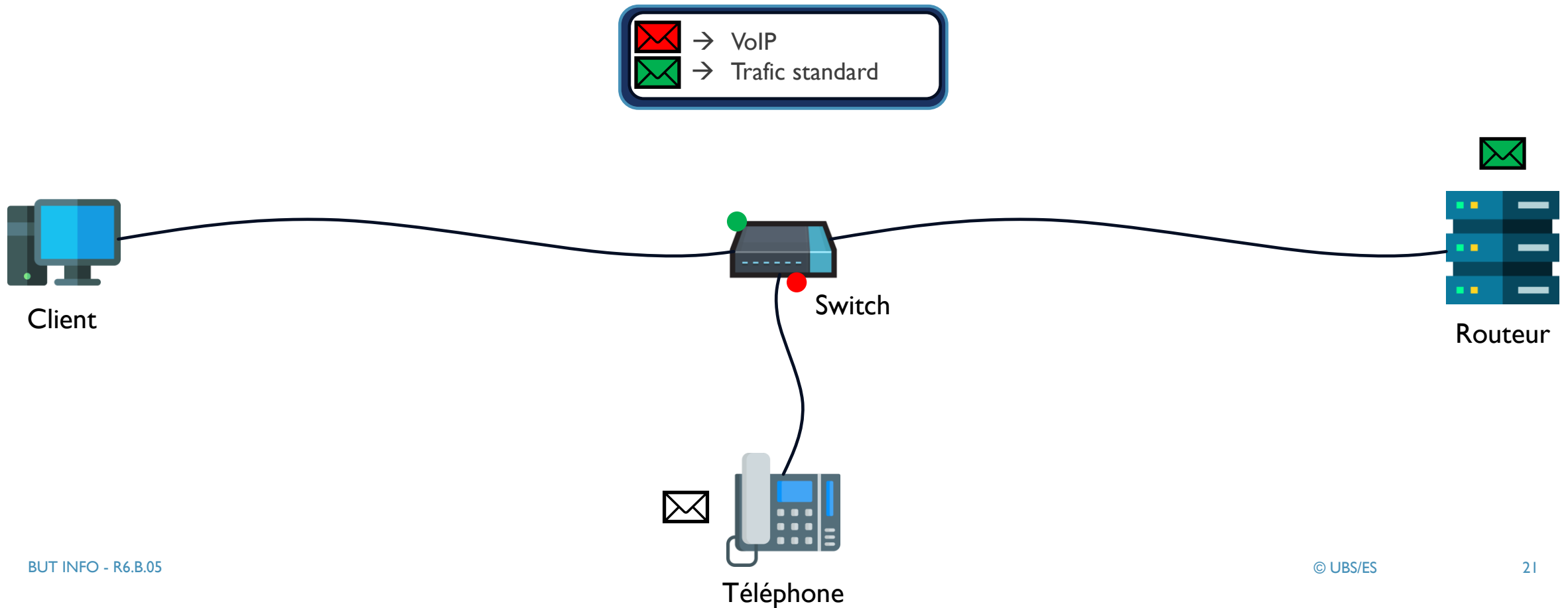
UTILISATION DU MARQUAGE



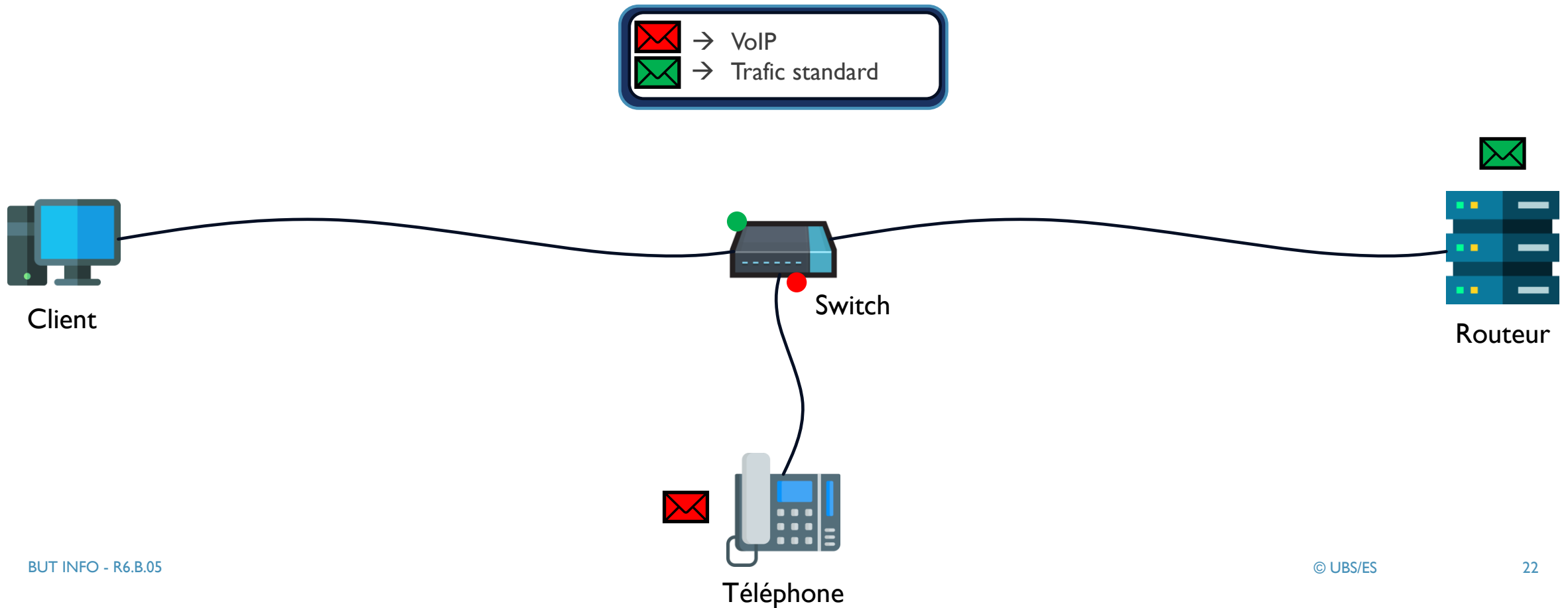
UTILISATION DU MARQUAGE



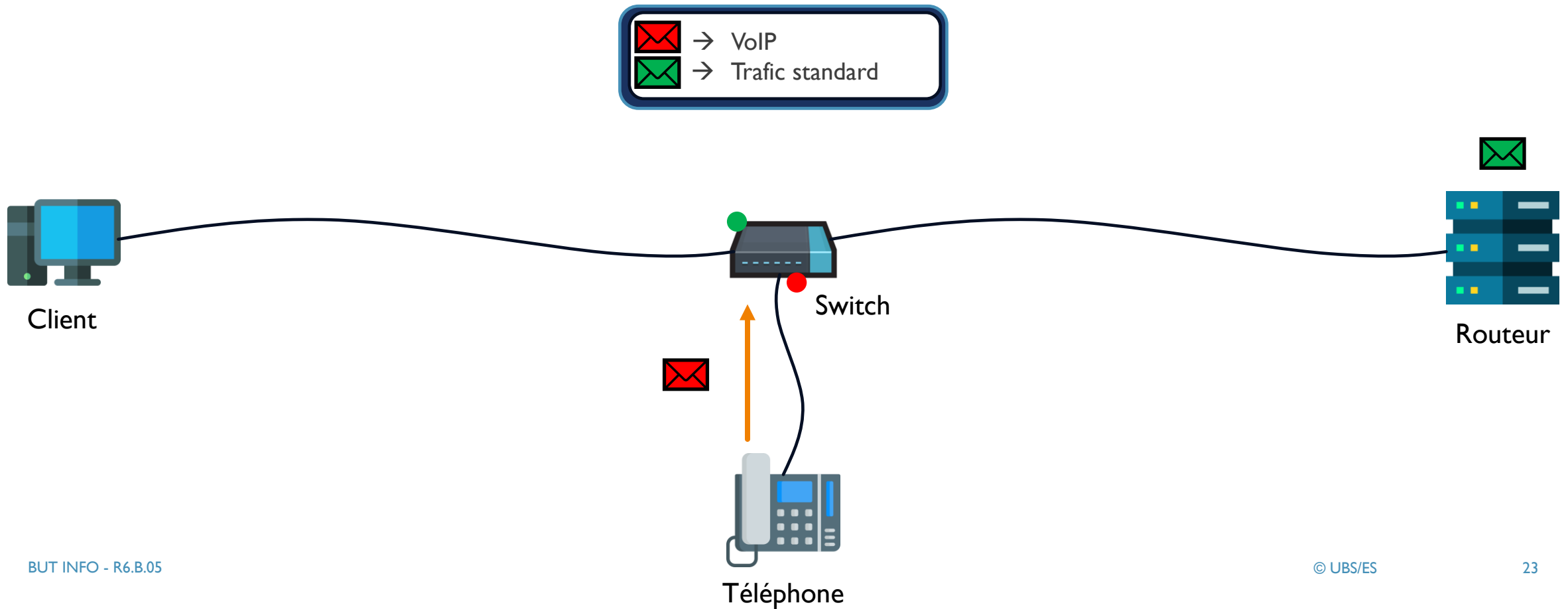
UTILISATION DU MARQUAGE



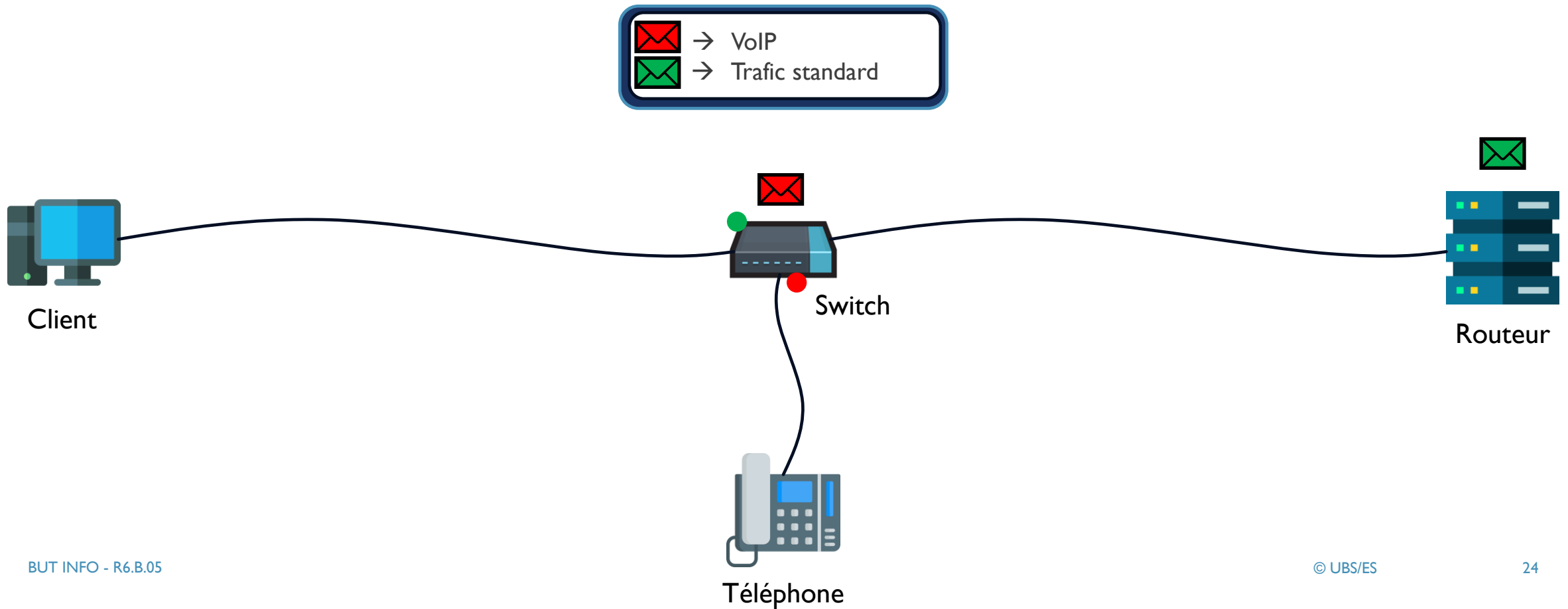
UTILISATION DU MARQUAGE



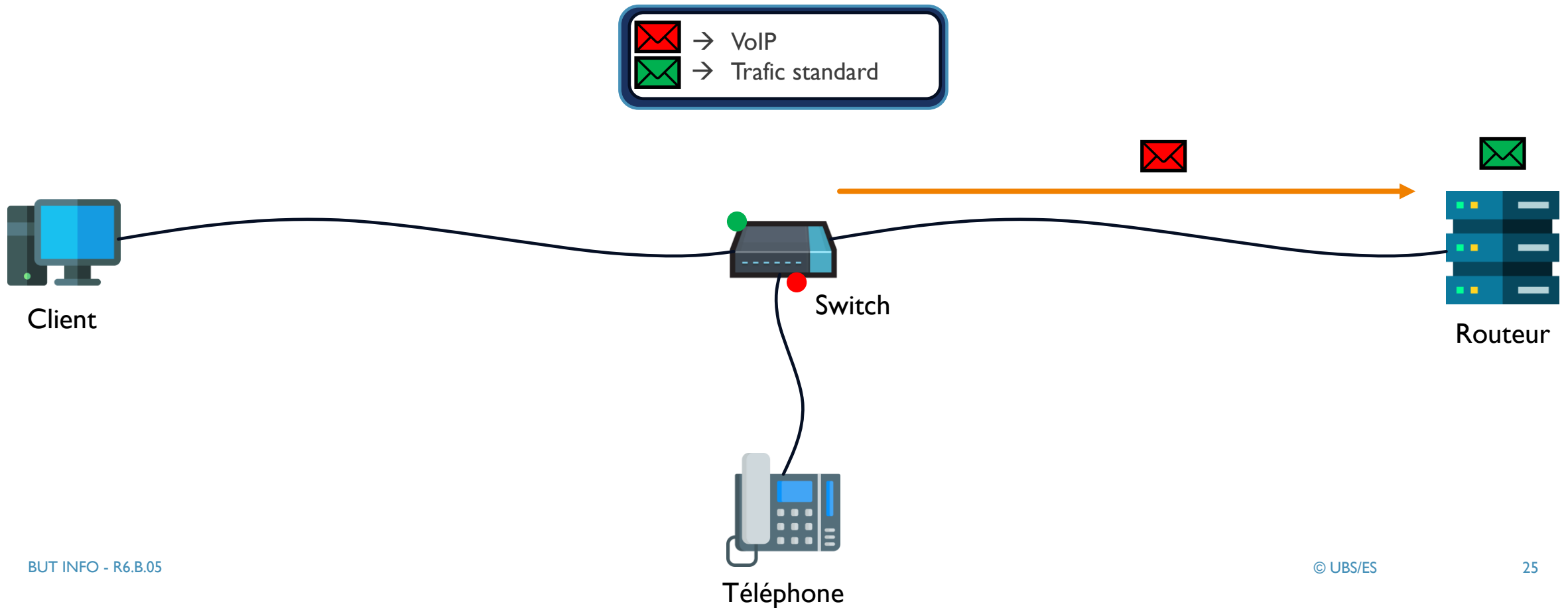
UTILISATION DU MARQUAGE



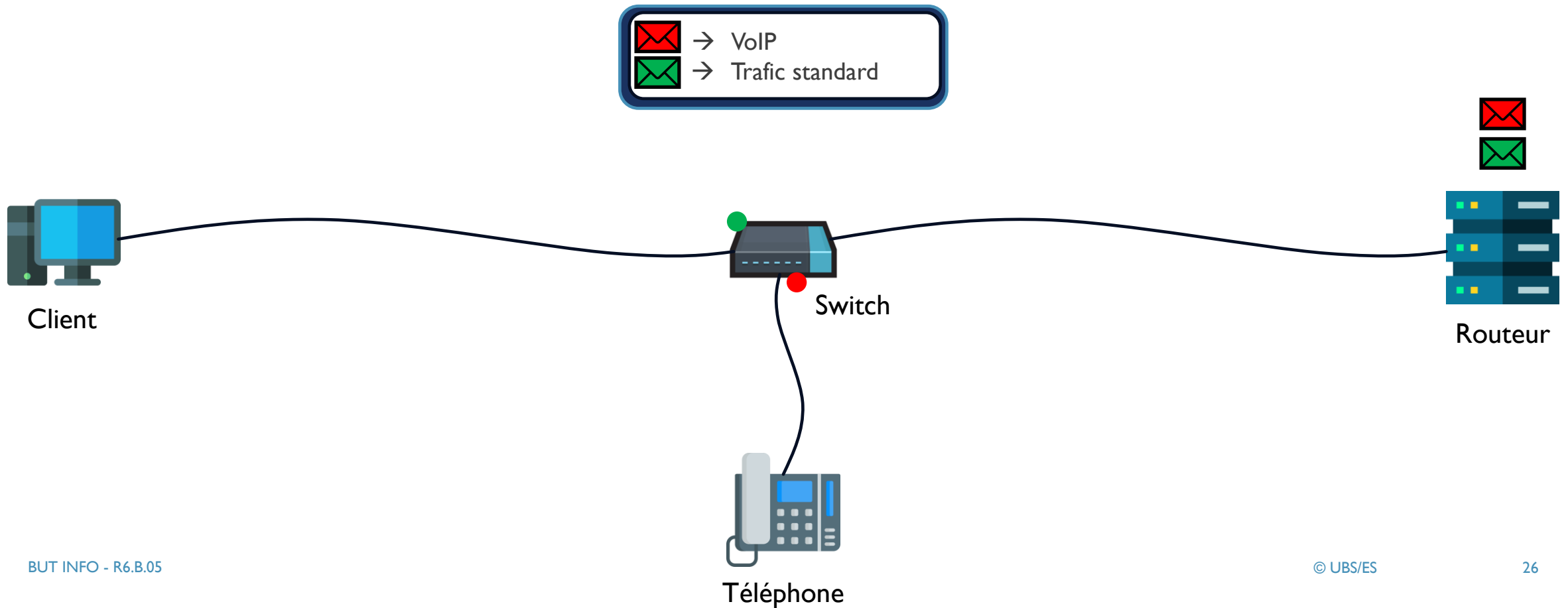
UTILISATION DU MARQUAGE



UTILISATION DU MARQUAGE

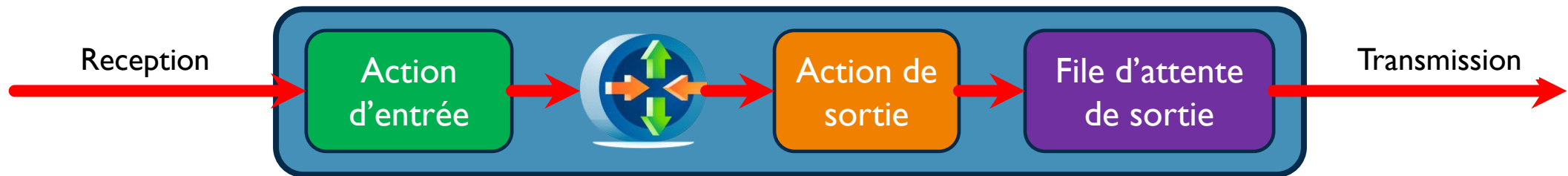


UTILISATION DU MARQUAGE



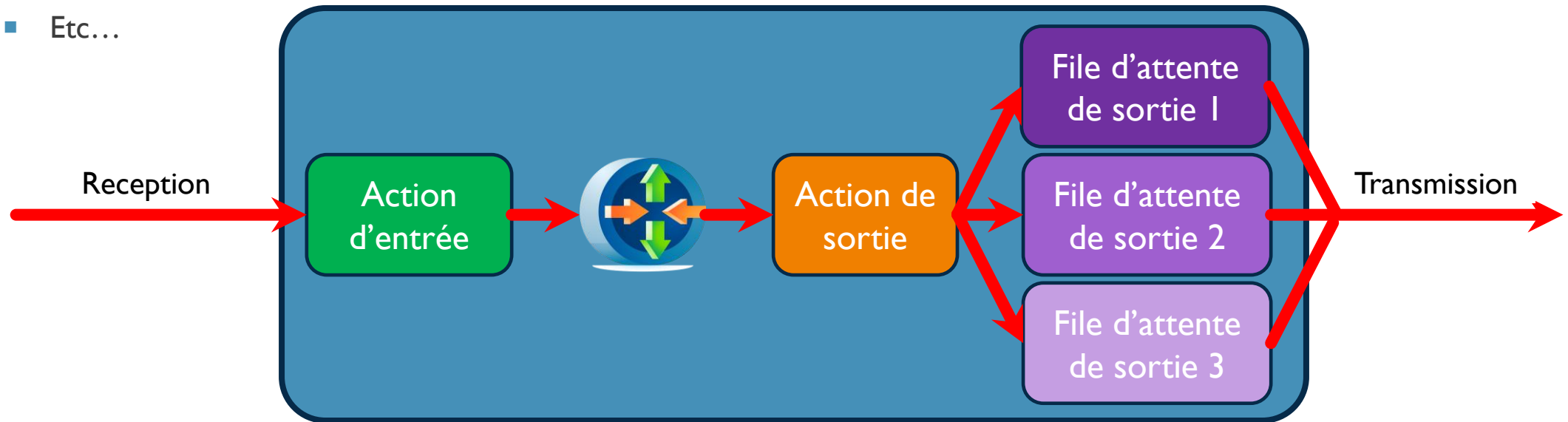
LES FILES D'ATTENTE

- Placement des paquets dans une file d'attente lorsque l'interface de sortie est occupée
 - Processus présente sur les routeurs et les switches
- Il existe plusieurs méthodes de file d'attente
 - Une file → FIFO, etc...
 - Plusieurs files → Round Robin, Prioritaire, etc...

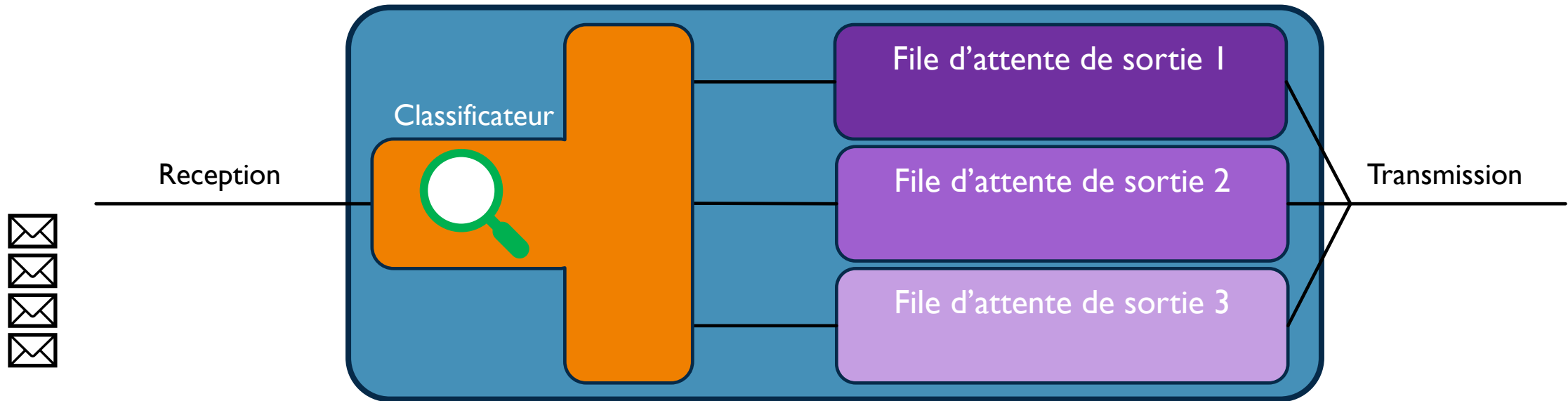


LES FILES D'ATTENTE

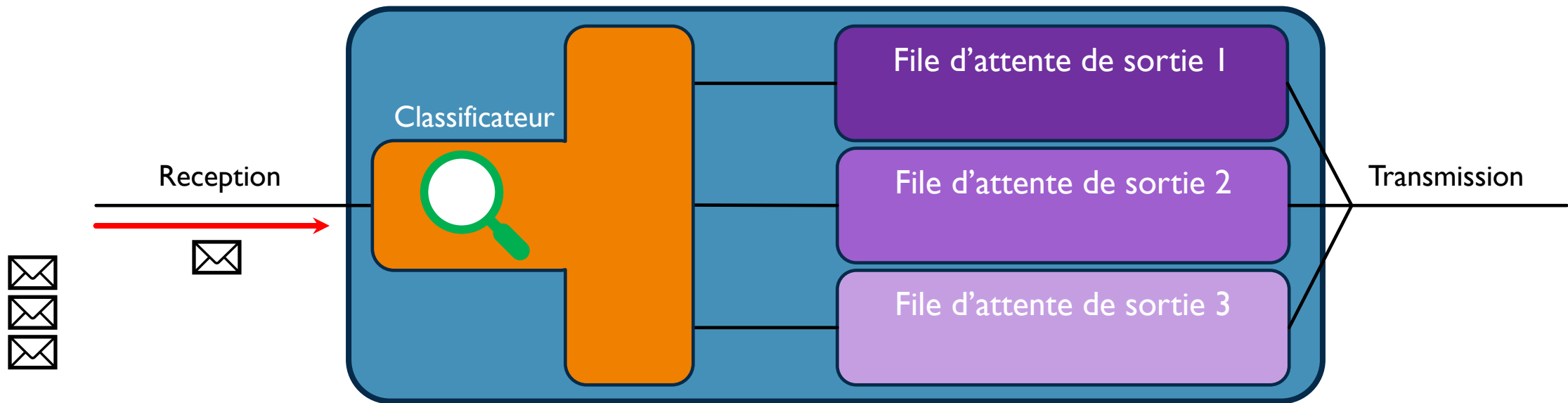
- Plusieurs approches avec multiples files
 - Séparation par niveau de priorité
 - Séparation par type d'application
 - Etc...



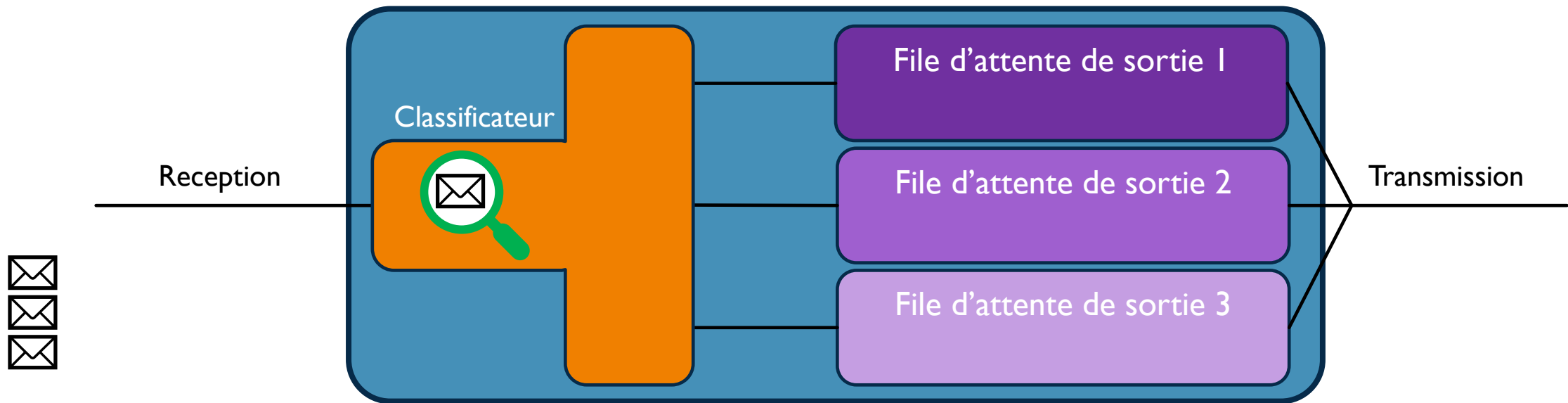
ROUND ROBIN



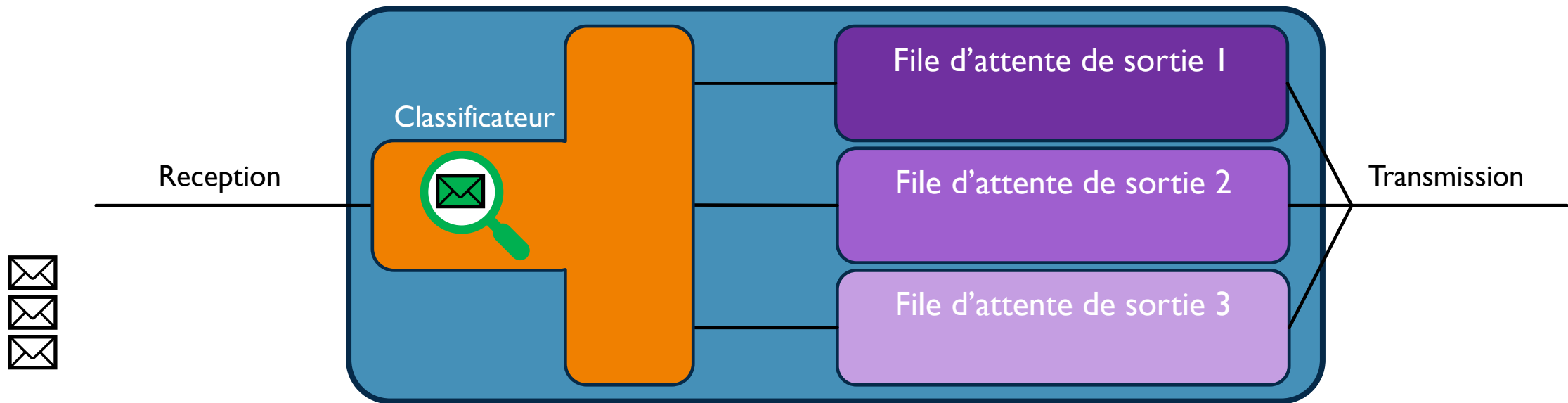
ROUND ROBIN



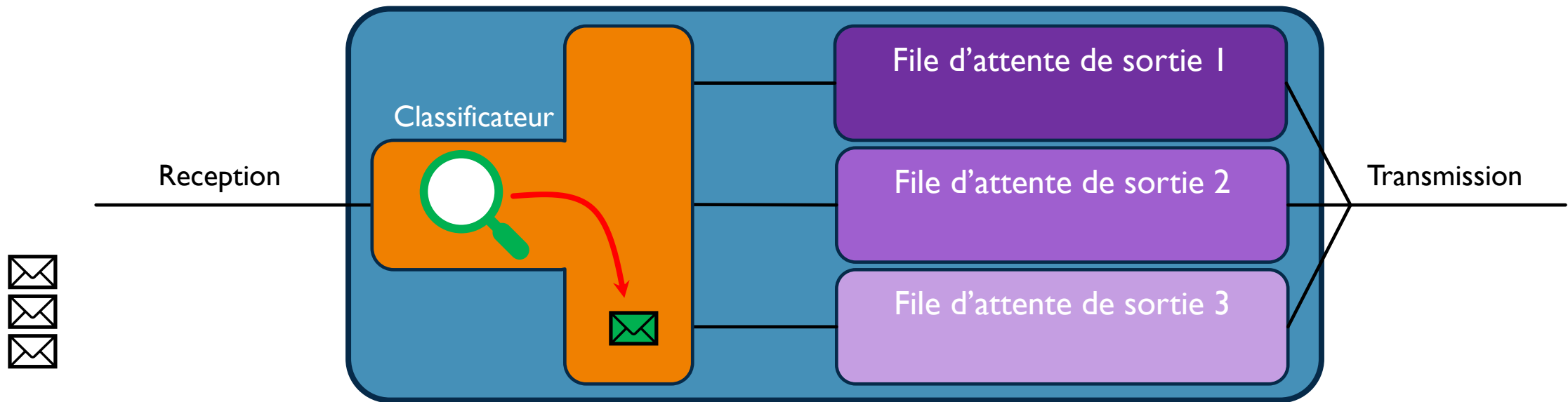
ROUND ROBIN



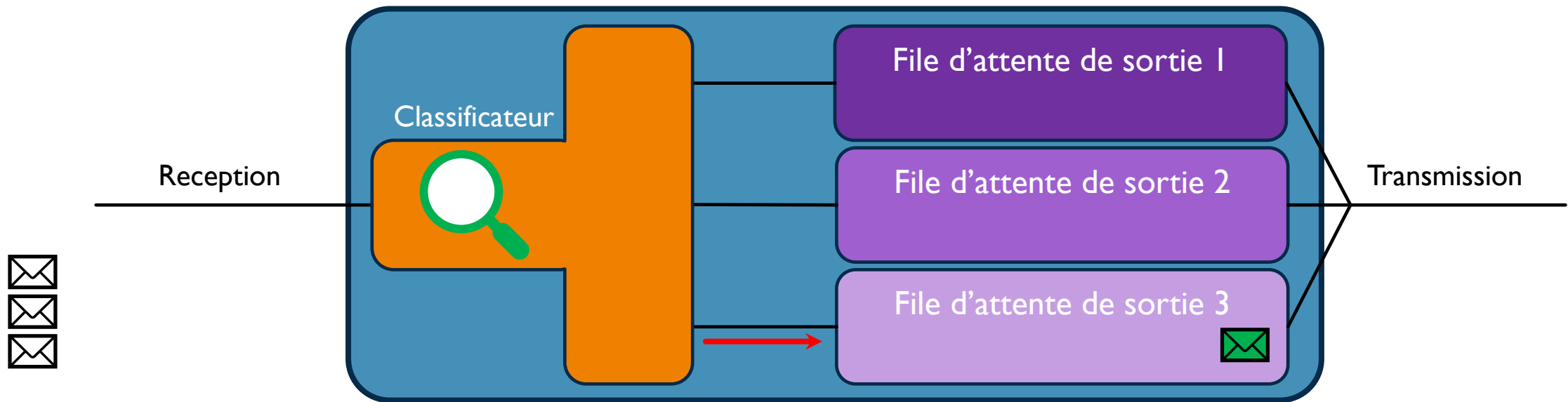
ROUND ROBIN



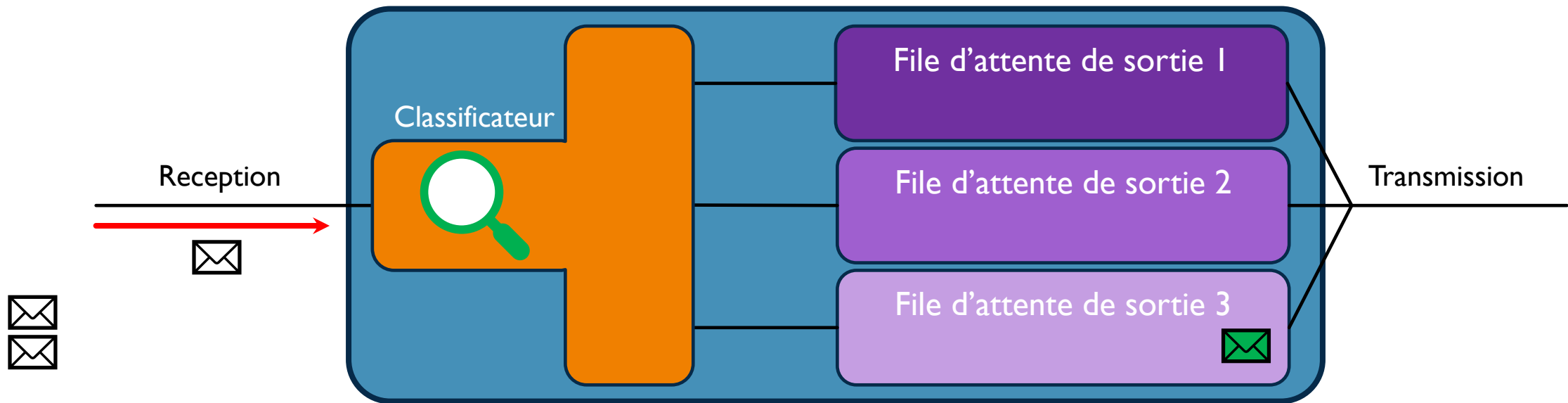
ROUND ROBIN



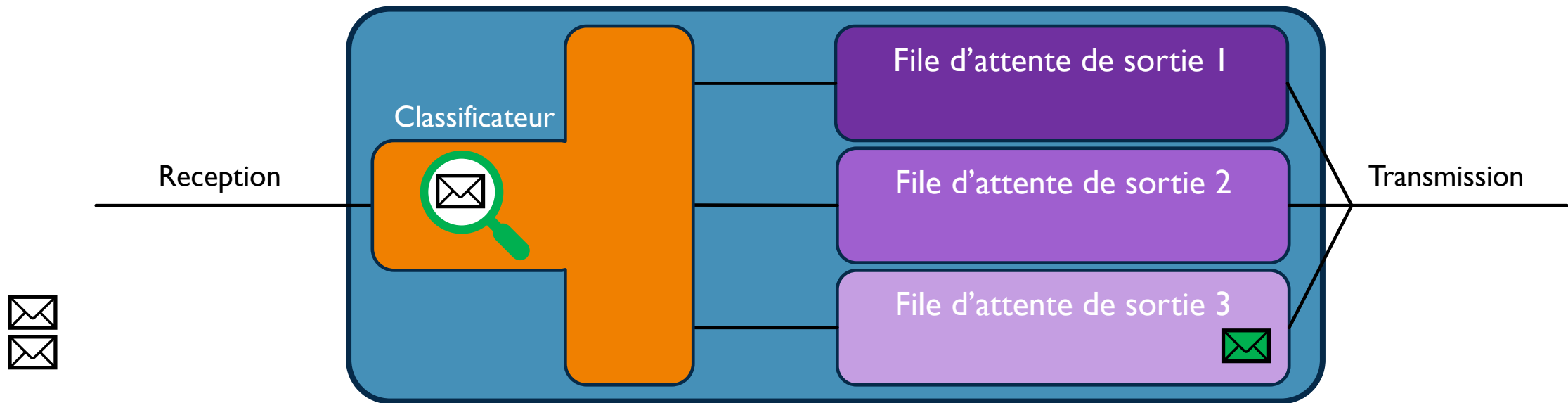
ROUND ROBIN



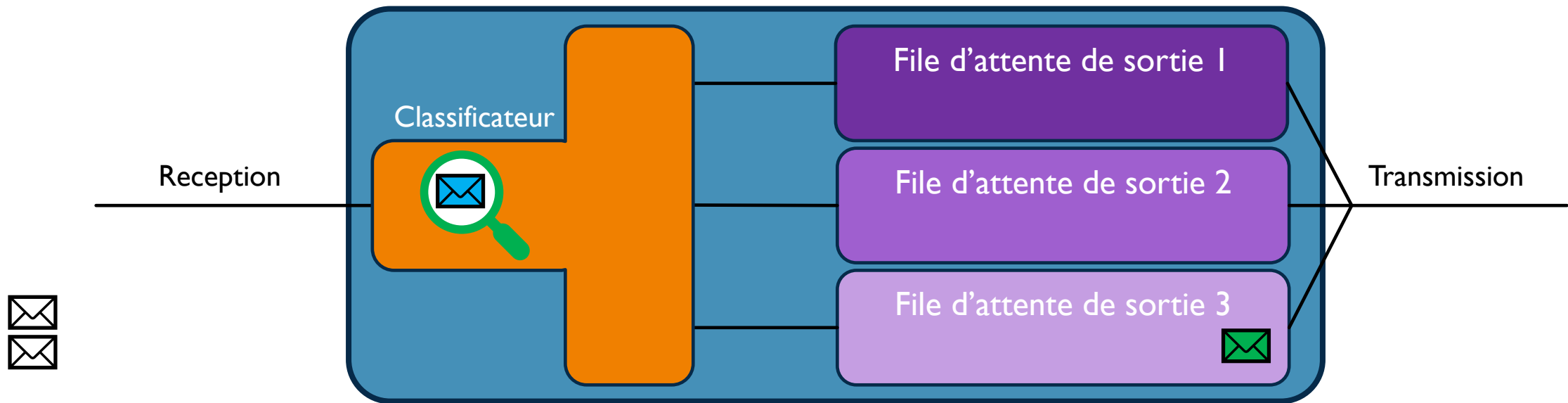
ROUND ROBIN



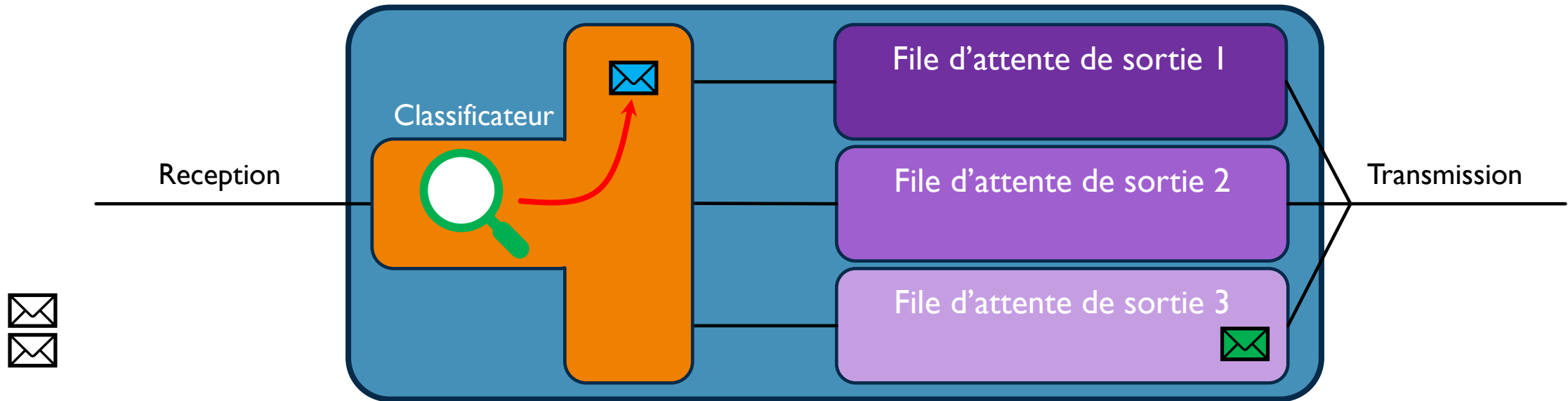
ROUND ROBIN



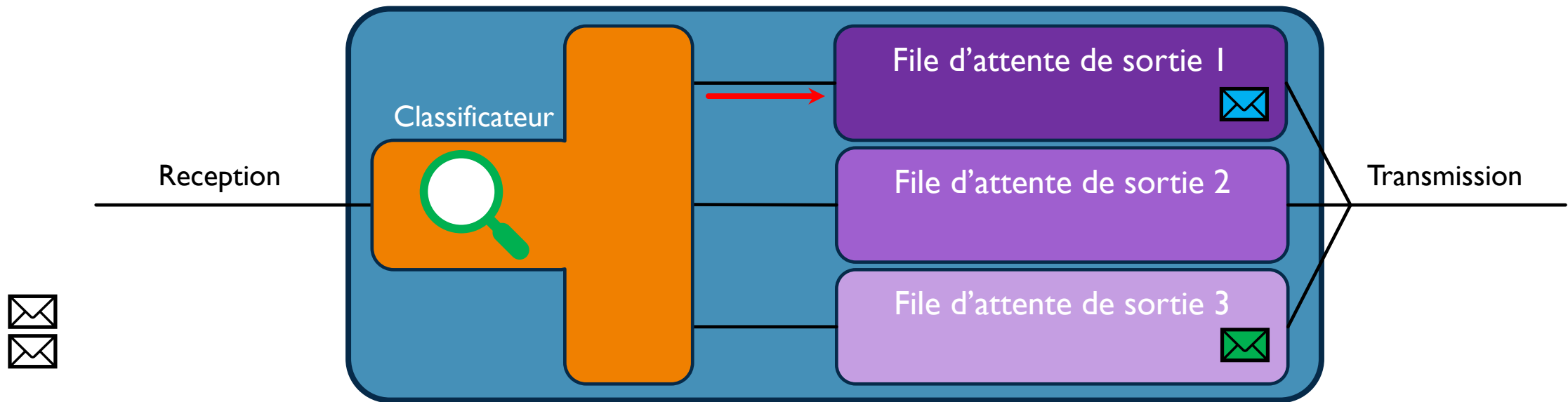
ROUND ROBIN



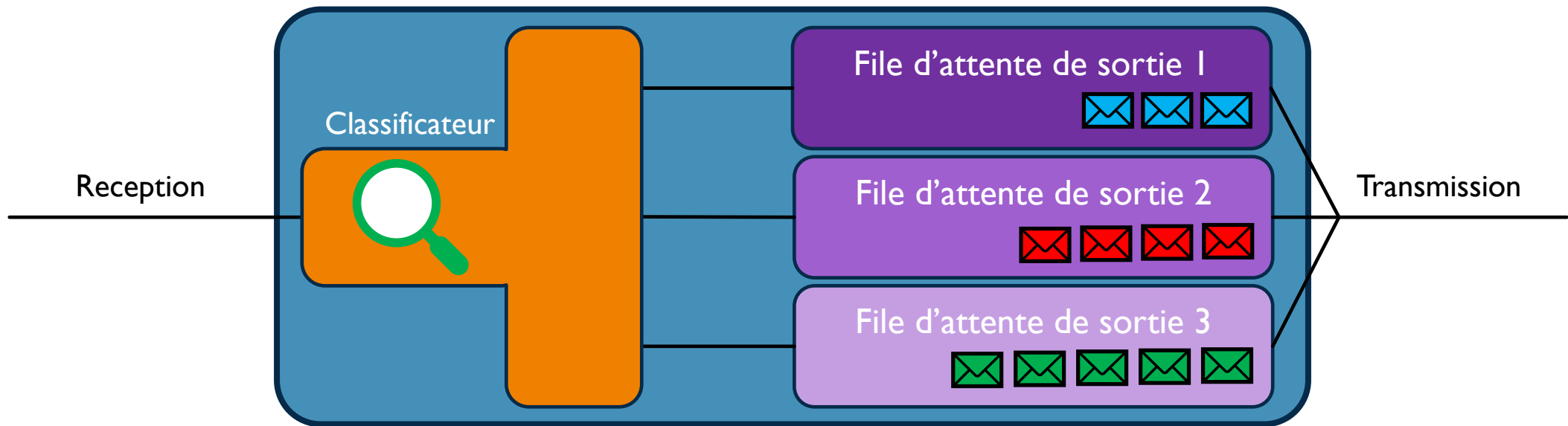
ROUND ROBIN



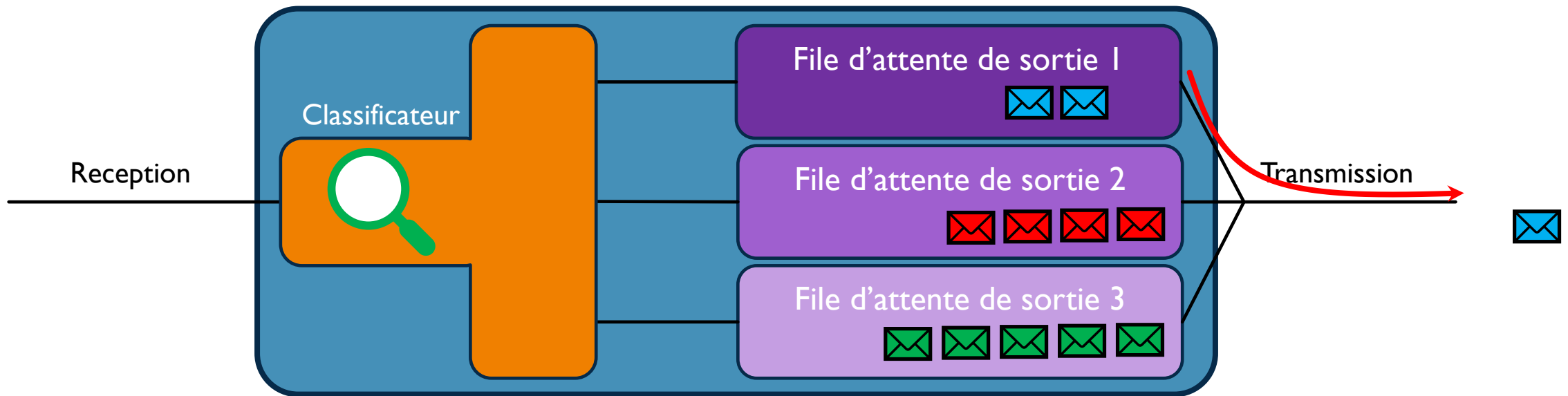
ROUND ROBIN



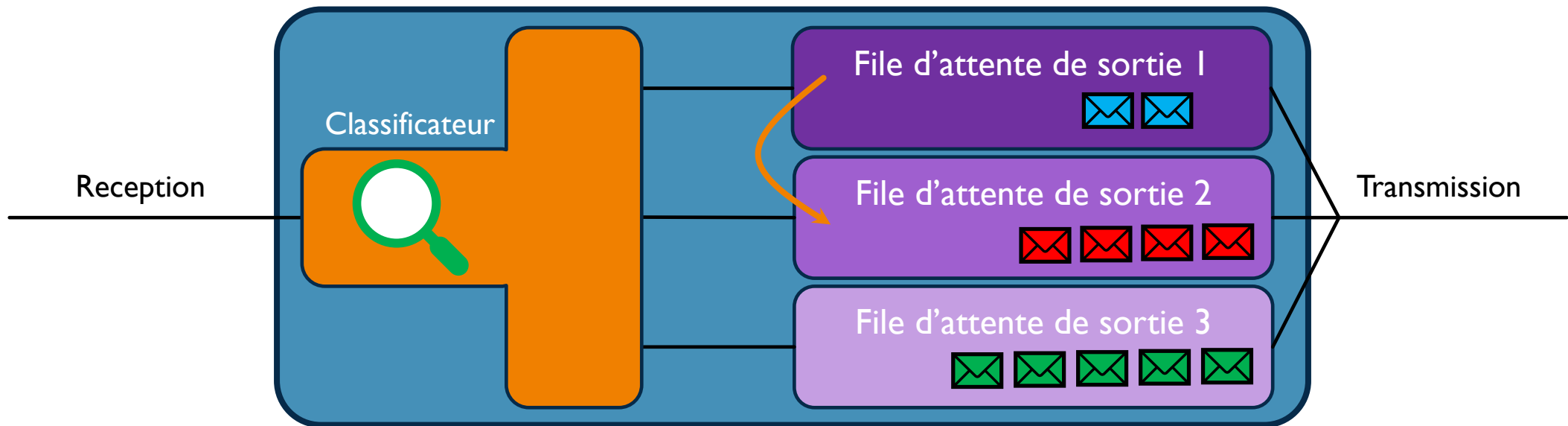
ROUND ROBIN



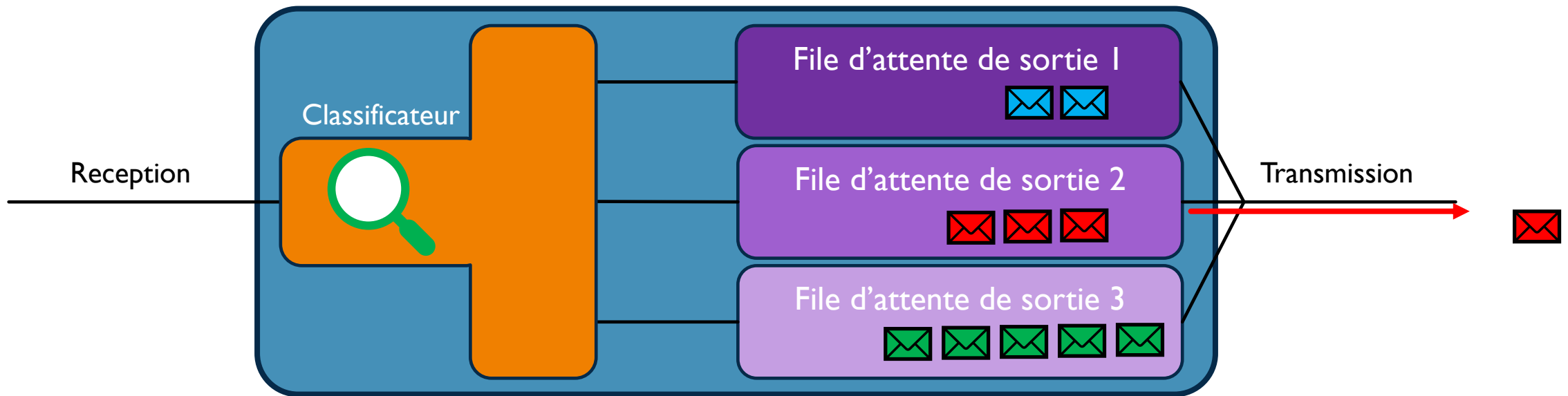
ROUND ROBIN



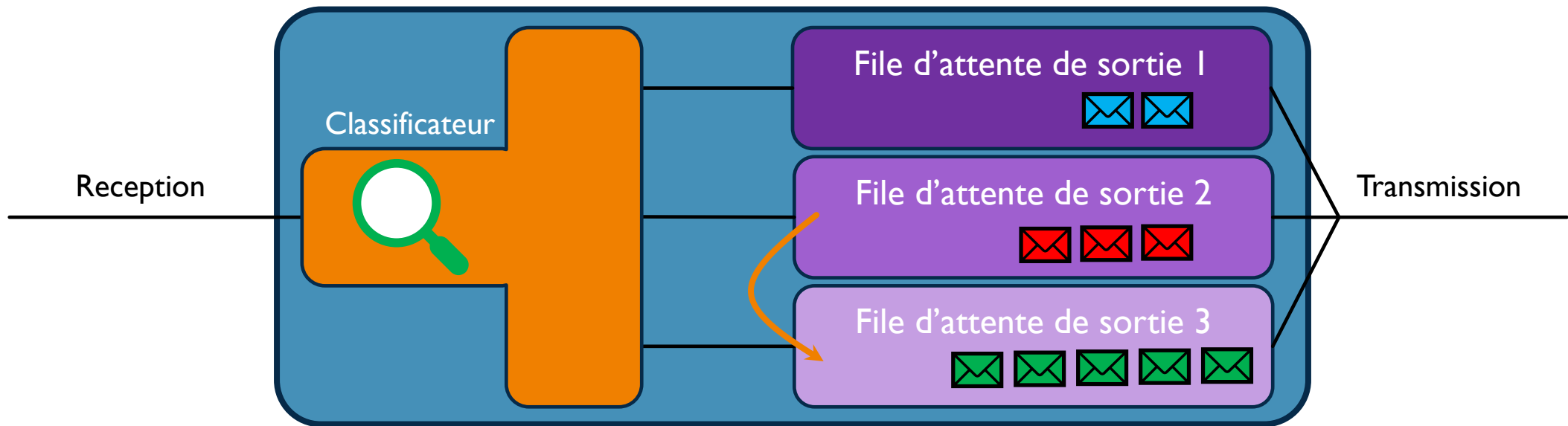
ROUND ROBIN



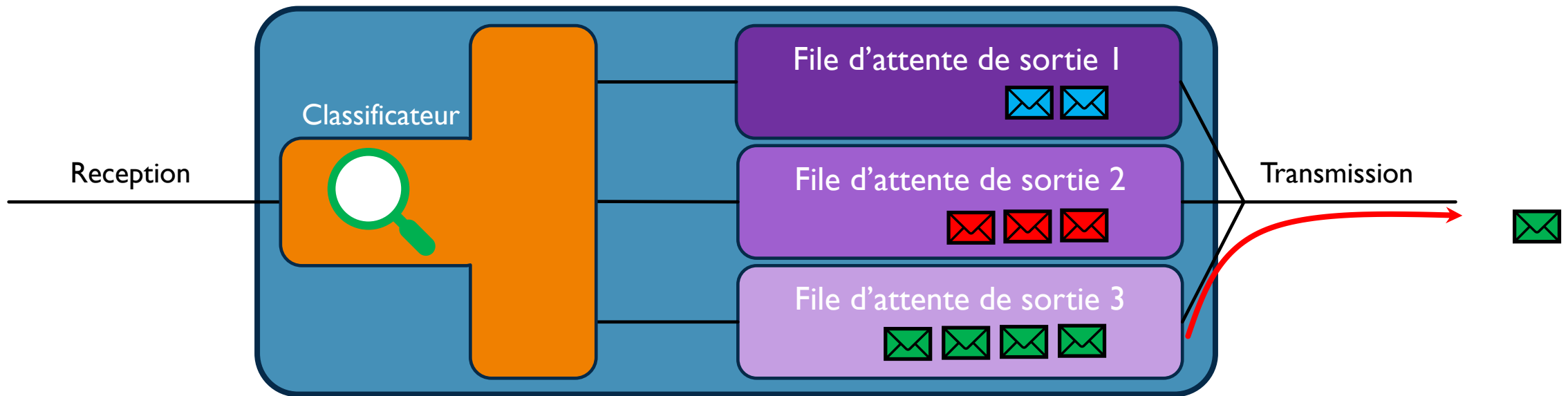
ROUND ROBIN



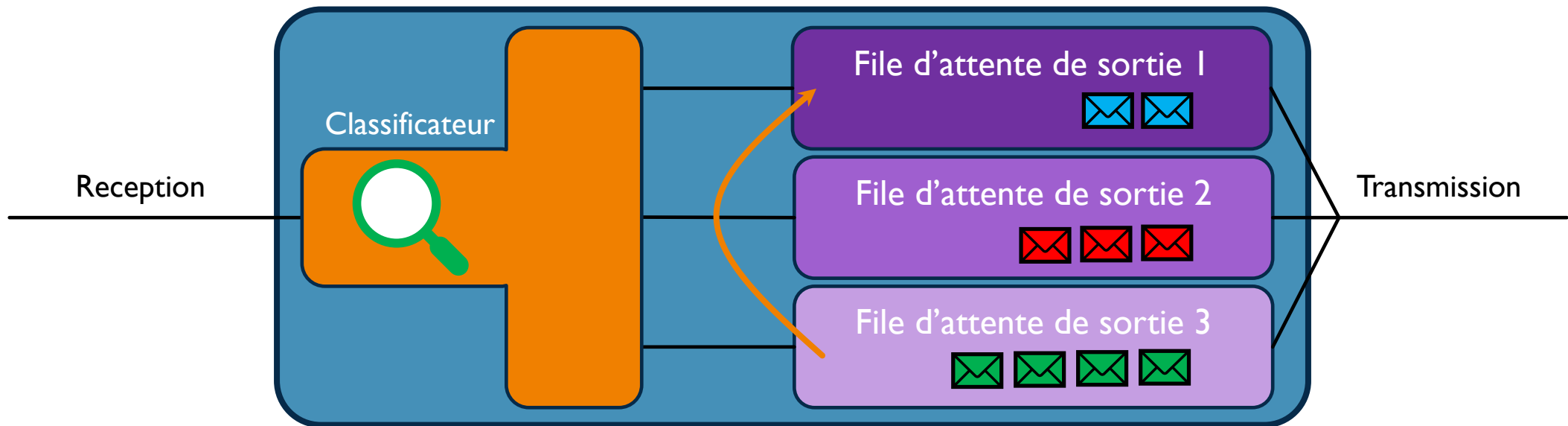
ROUND ROBIN



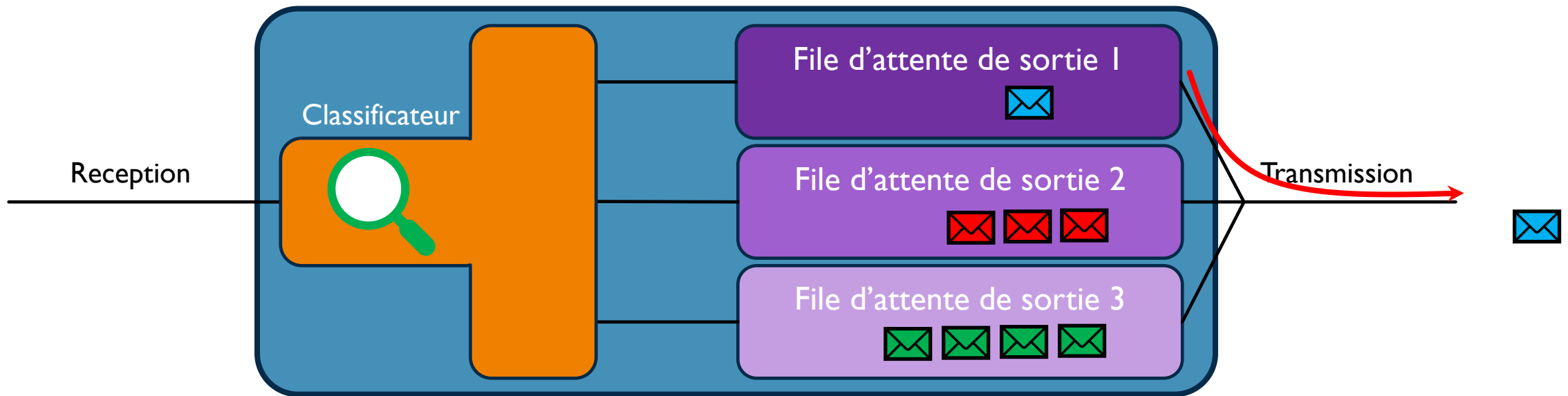
ROUND ROBIN



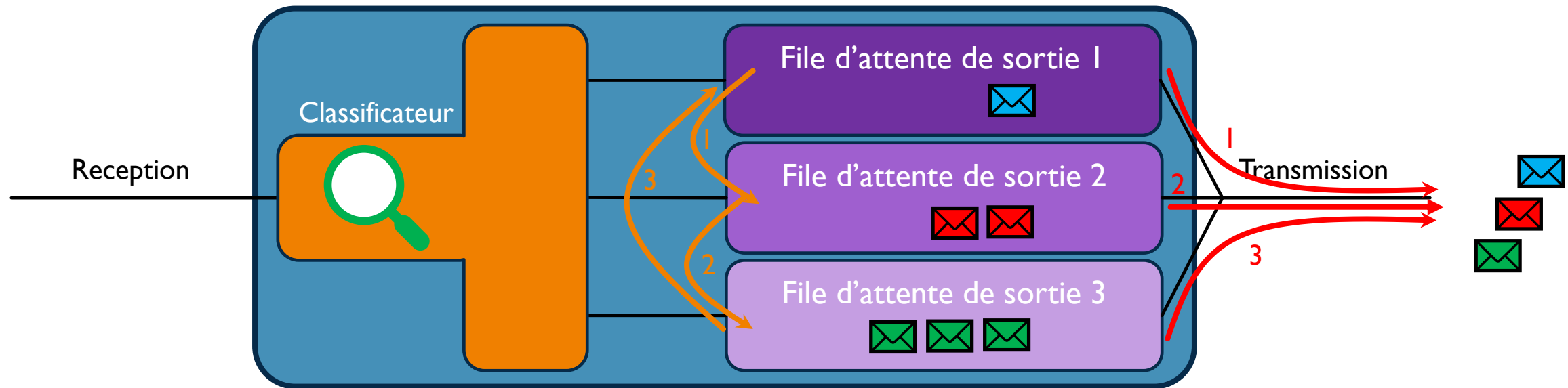
ROUND ROBIN



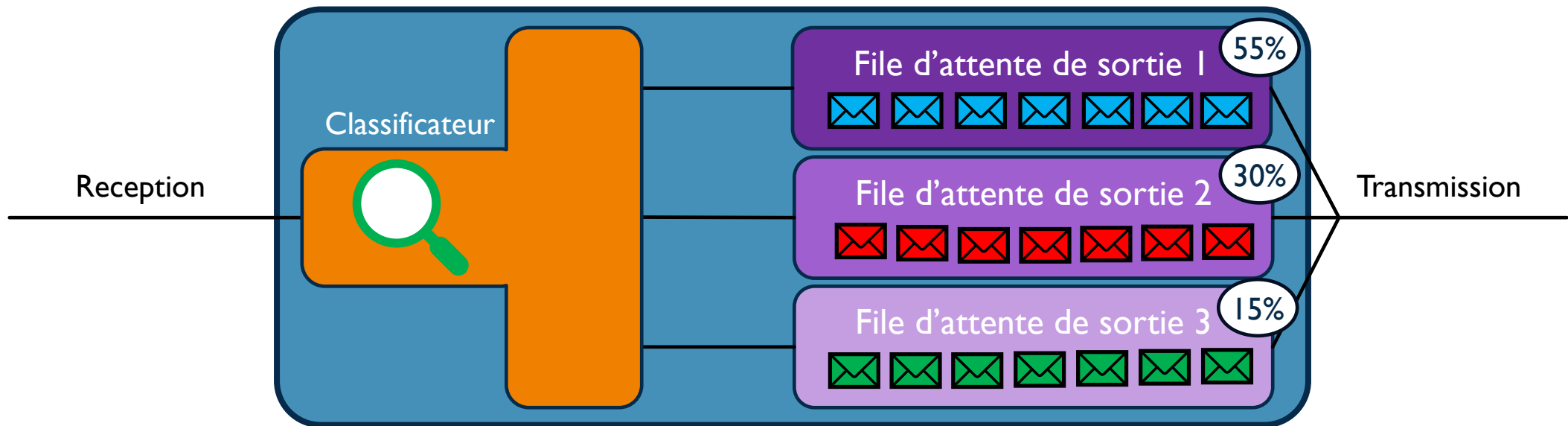
ROUND ROBIN



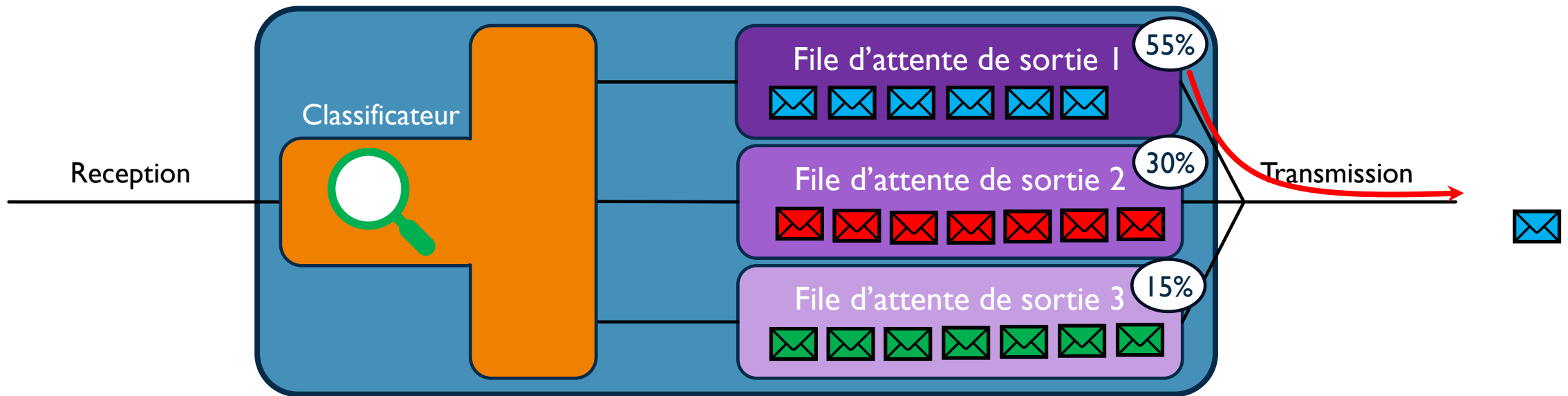
ROUND ROBIN



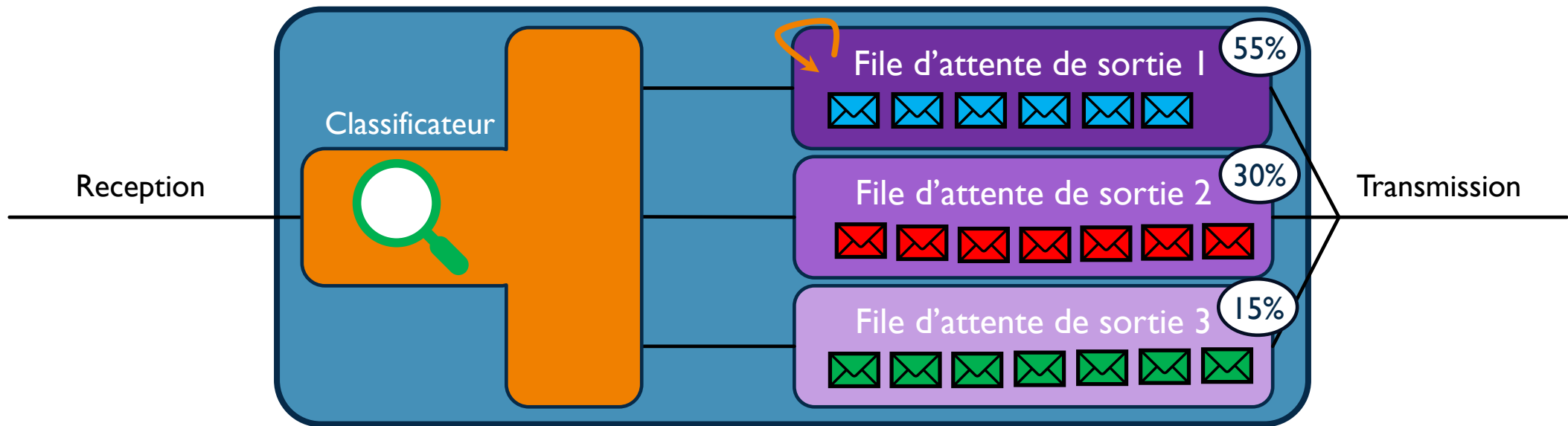
ROUND ROBIN PONDÉRÉ



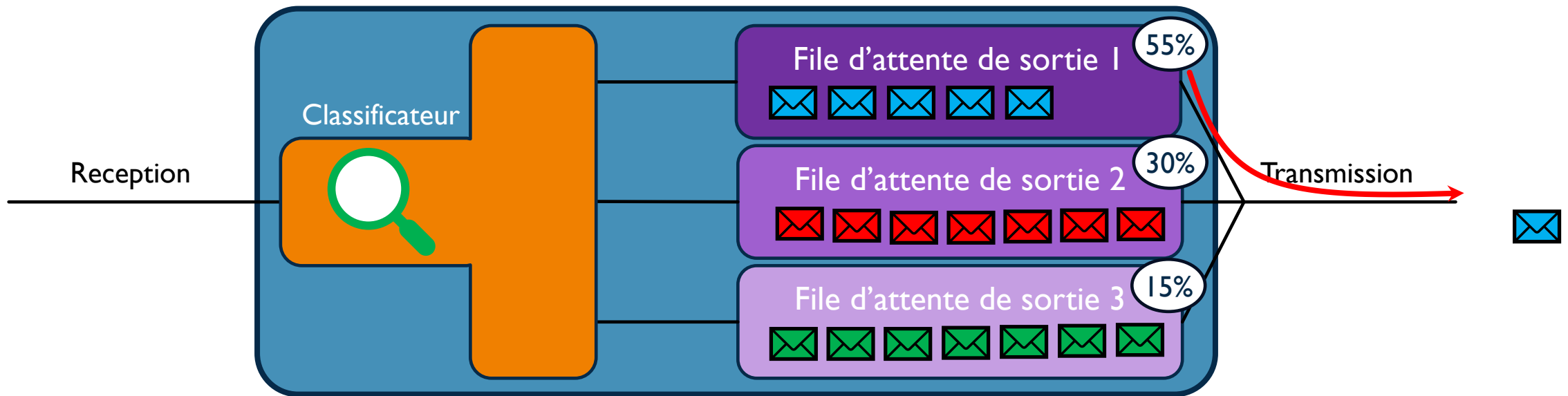
ROUND ROBIN PONDÉRÉ



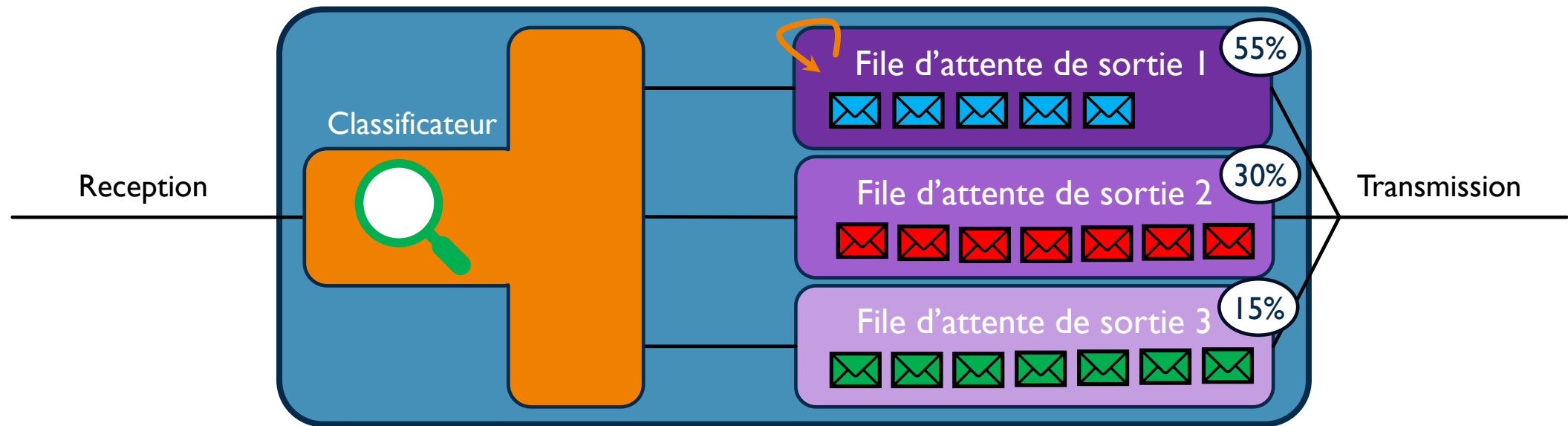
ROUND ROBIN PONDÉRÉ



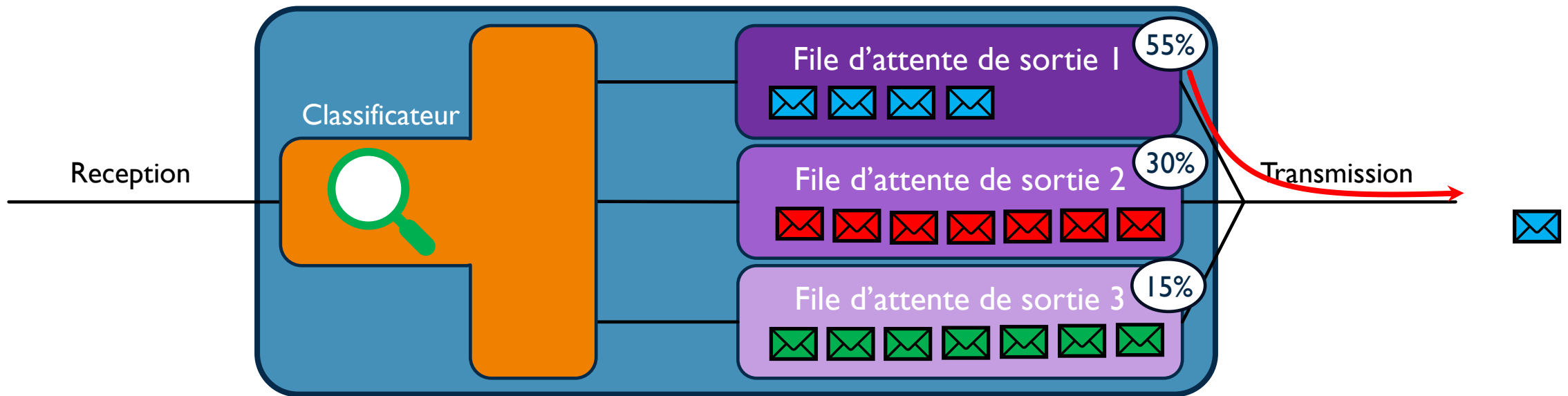
ROUND ROBIN PONDÉRÉ



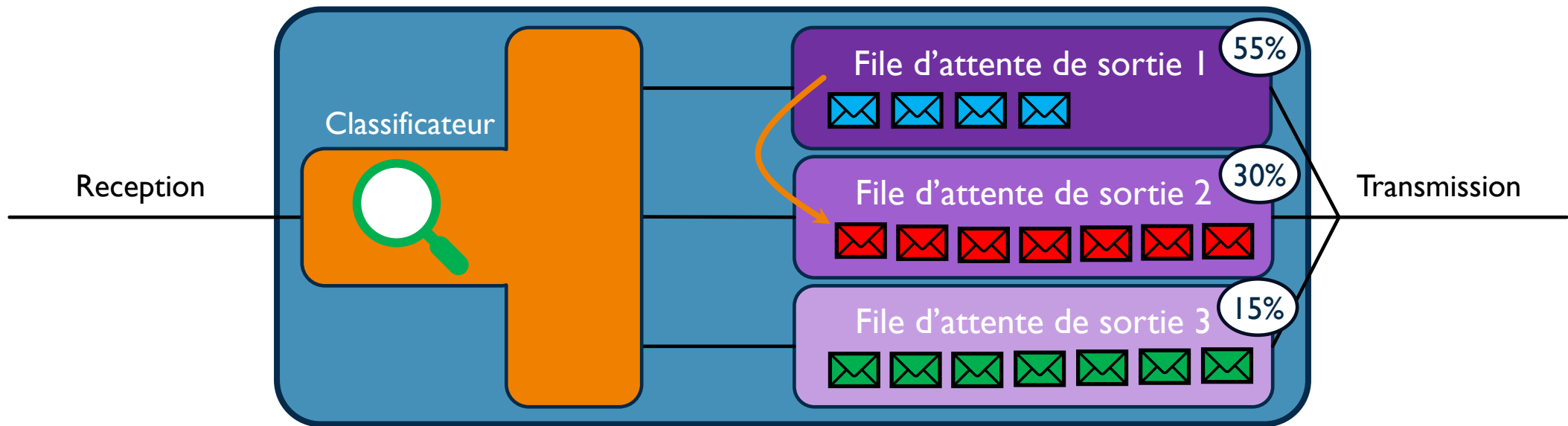
ROUND ROBIN PONDÉRÉ



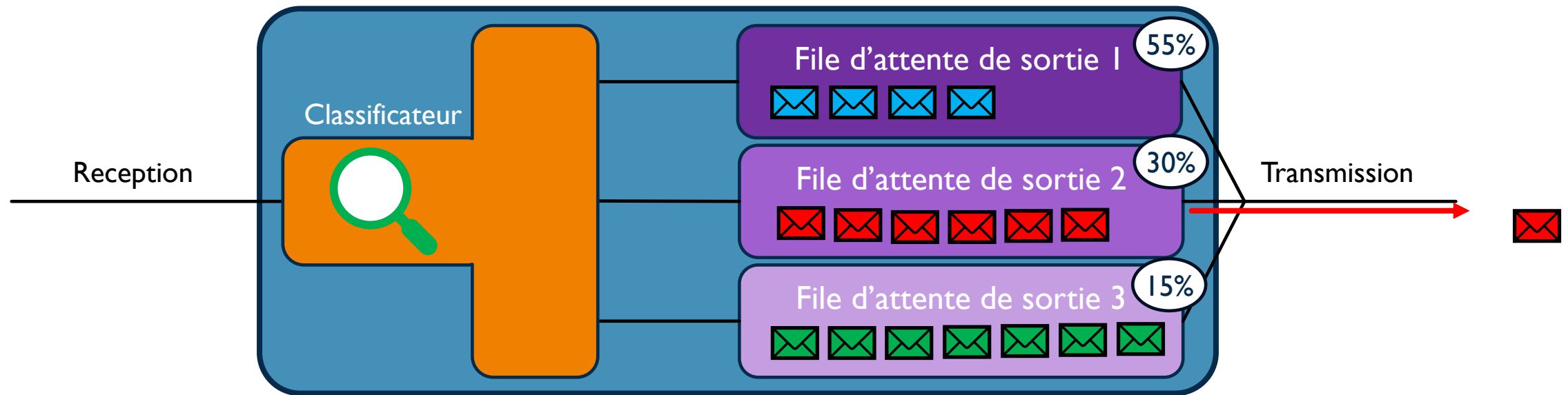
ROUND ROBIN PONDÉRÉ



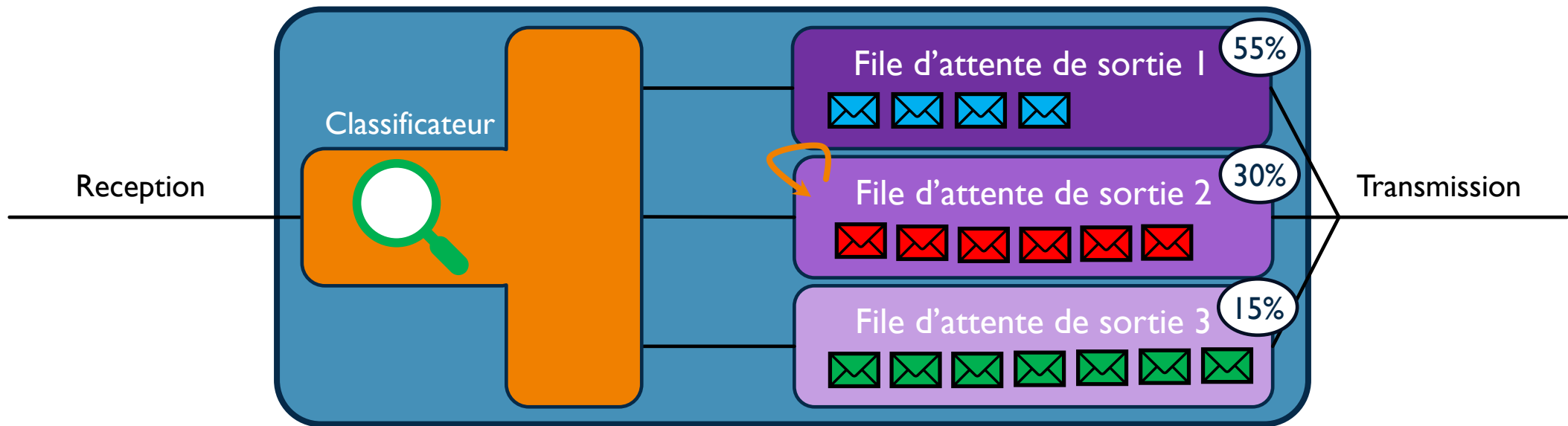
ROUND ROBIN PONDÉRÉ



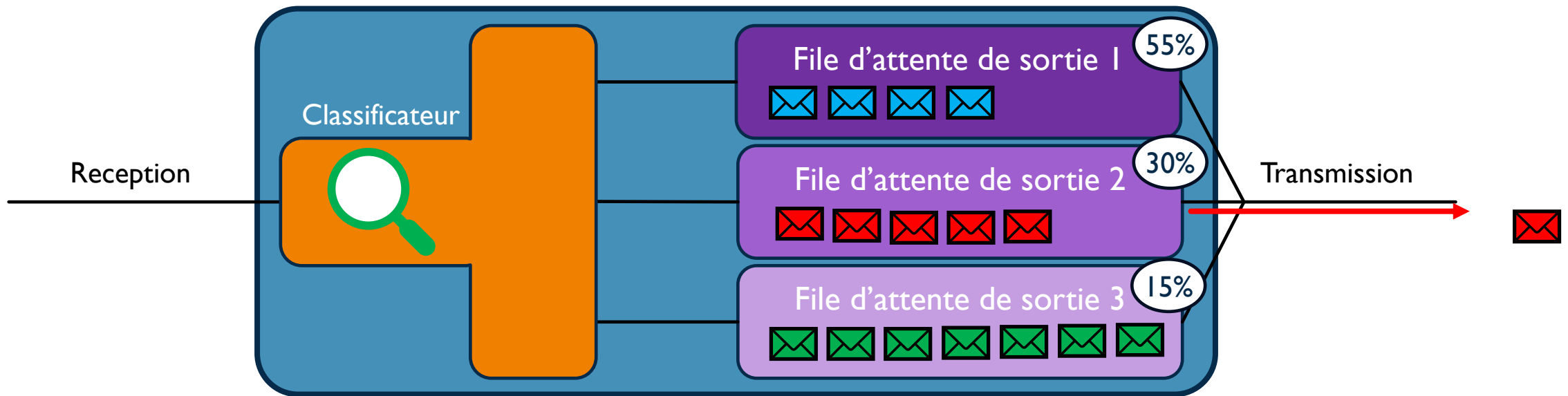
ROUND ROBIN PONDÉRÉ



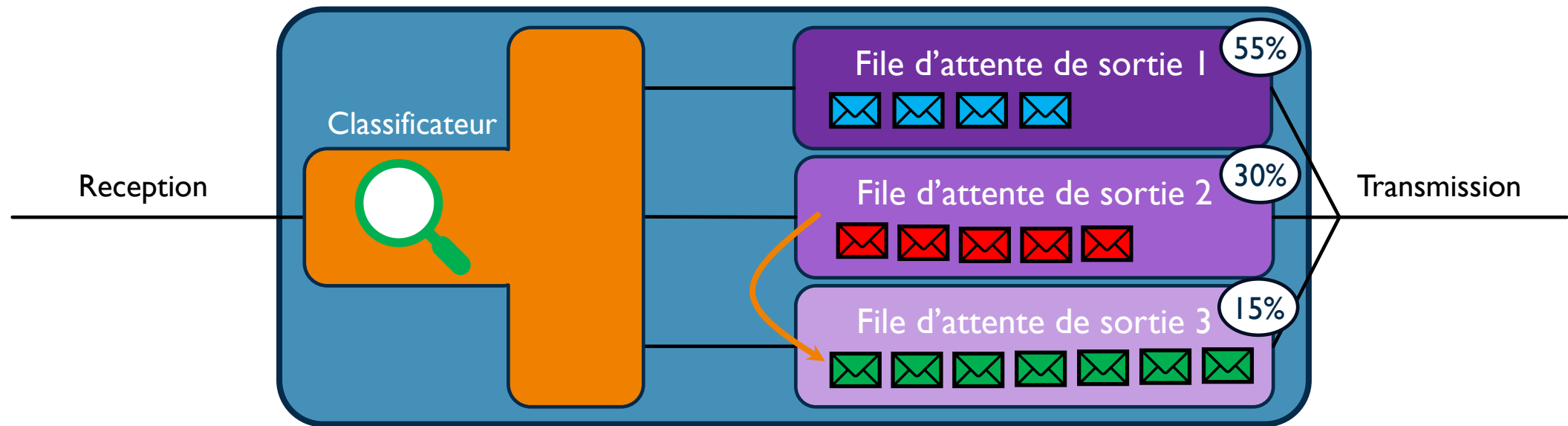
ROUND ROBIN PONDÉRÉ



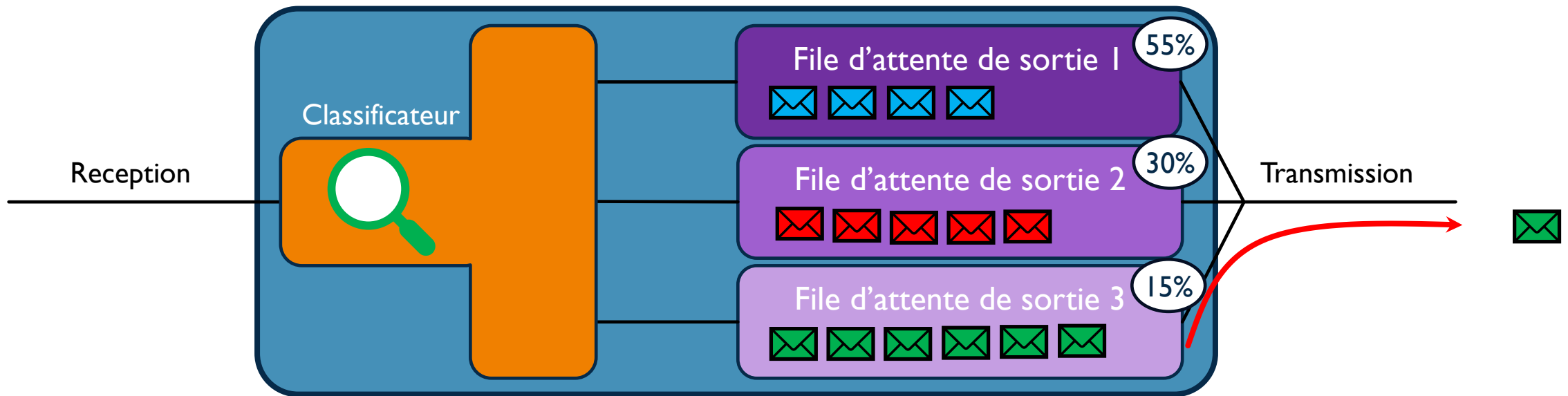
ROUND ROBIN PONDÉRÉ



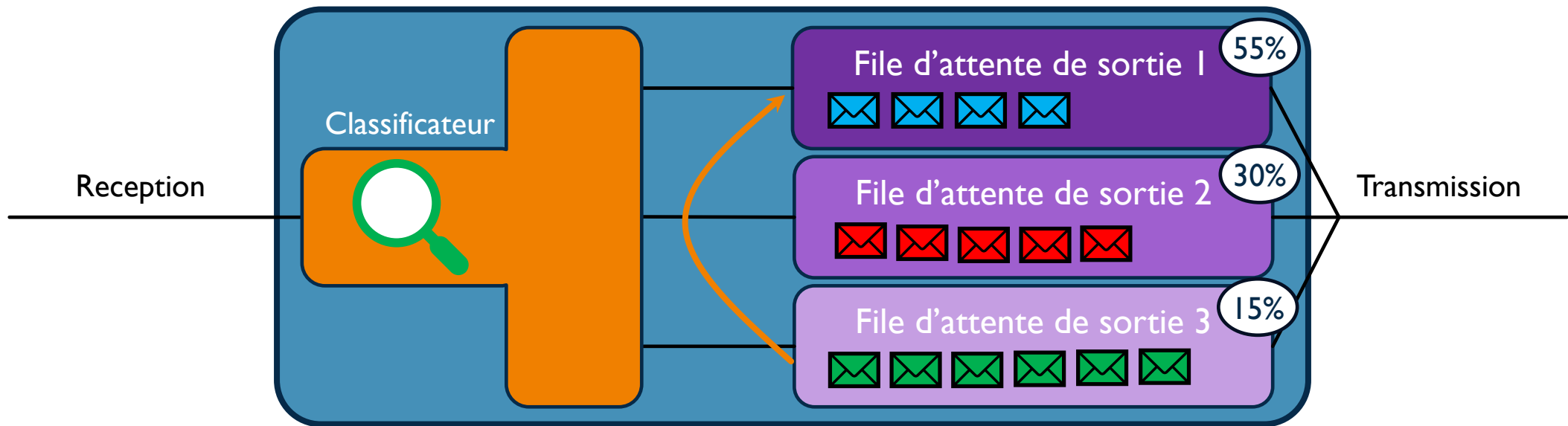
ROUND ROBIN PONDÉRÉ



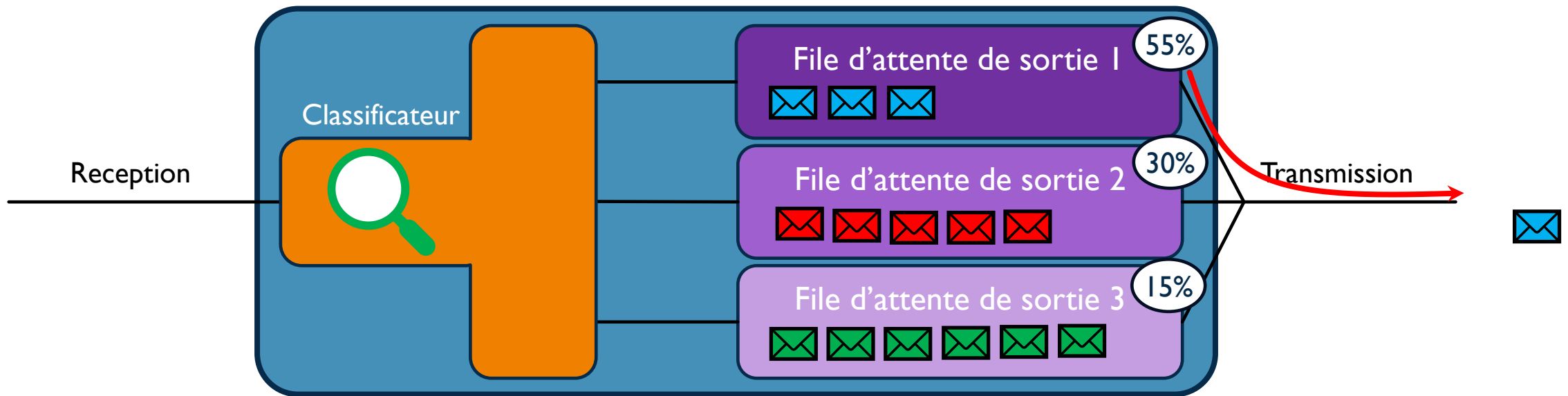
ROUND ROBIN PONDÉRÉ



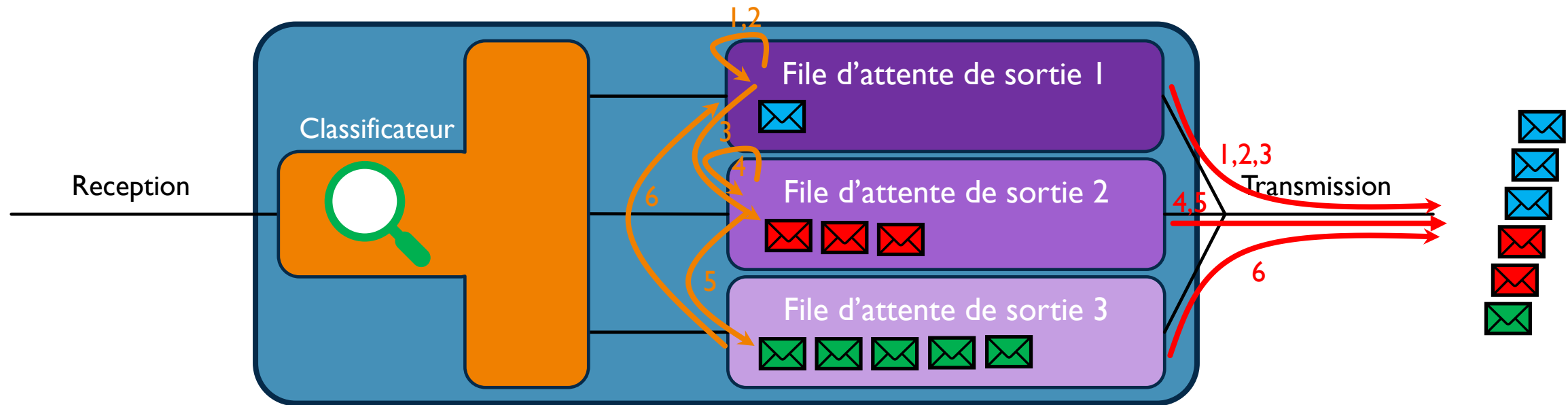
ROUND ROBIN PONDÉRÉ



ROUND ROBIN PONDÉRÉ



ROUND ROBIN PONDÉRÉ



TECHNIQUES

Hiérarchie du trafic VoIP sensible aux retards via les routeurs et commutateurs

- Plus le réseau est grand → plus de risque de congestion
 - Traitement pas assez rapide → jet de paquets
- Classification du trafic et priorisation en fonction du type
 - Utile en cas de congestion élevée
 - Paquets avec priorité élevé → envoi en priorité avant le reste du trafic

Réservation de ressource

- RSVP → Protocole de réservation de ressources
 - Couche Transport pour réserver des ressources sur un réseau
 - Peut être utilisé pour fournir des niveaux spécifiques de QoS pour des flux de données d'application
- Permet la division des ressources réseau
 - Par trafic de différents types / origines
 - Définir des limites
 - Garantir la disponibilité de la bande passante

TECHNIQUES

Mise en file d'attente

- Processus de création de règles
 - Traitement préférentiel de certains flux de données
- Présent dans les routeurs et switch
- Utilisation de plusieurs files d'attente par priorité
 - Paquet reçu d'une priorité élevée → mis dans la file correspondante (débit élevé)
- Ex. attribution d'une règle de priorité du VoIP
 - Mise au début des files d'attente → utilisation de la majorité de la bande passante

Marquage du trafic

- Identification des besoins en termes de priorité des applications
- Utilisation de deux processus
 - Classe de service (CoS)
 - Marquage dans l'en-tête couche 2
 - Code des services différenciés (DSCP)
 - Marquage dans l'en-tête couche 3

AVANTAGES DE LA QOS INFORMATIQUE

- **Hiérarchisation illimitée des applications**

- Garanti la priorité et les ressources aux applications stratégiques

- **Meilleure gestion des ressources**

- Gestion des ressources réseau et Internet
- Réduction des coûts et la nécessité d'investissement dans des liaisons supplémentaires

- **Expérience utilisateur améliorée**

- Offrir une expérience utilisateur optimale
- Garantir des performances élevées des applications sensibles
 - Meilleure efficacité du travail des employés

- **Gestion du trafic point à point**

- Gestion essentielle dans l'acheminement du trafic
- Permet de livrer des paquets dans l'ordre sans pertes

- **Prévention de la perte de paquets**

- Se produit lors du rejet de paquets en transit
 - Défaillances, inefficacité, congestion, connexion faible, etc...
- Hiérarchisation de la bande passante des applications haute performance → réduction des pertes

- **Réduction de la latence**

- Causé par le temps de traitement des routeurs dans les files d'attente
- Réduction de la latence et l'accélération des demandes via la hiérarchisation des applications sensibles