

R6.B.05 – OPTIMISATION DE SERVICES COMPLEXES

PARTIE I

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

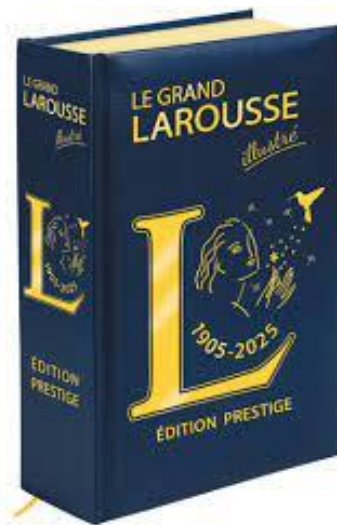
PLAN DU COURS

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



OPTIMISATION DE L'ACCÈS AUX RESSOURCES

OPTIMISATION ?

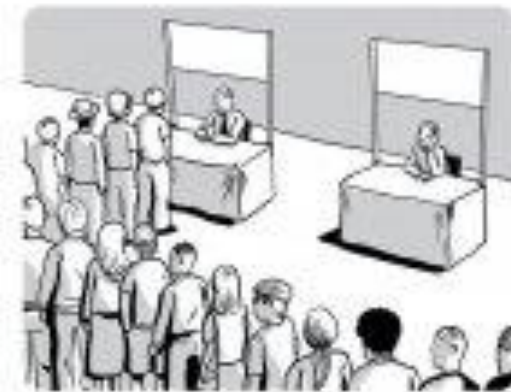
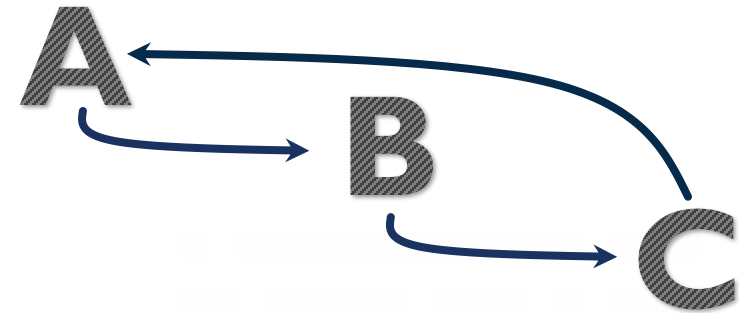


*Action d'optimiser ou d'optimaliser ;
fait d'être optimisé ou optimalisé.*

*Démarche consistant à rendre
optimal le fonctionnement d'un
système.*

POURQUOI OPTIMISER ?

- Fonctionnement des systèmes informatique très logique
 - Pas de « réflexion » extérieur ou d'analyse situationnelle
 - Suivi « bête » d'actions
- Souvent monopolisation ou mauvaise gestion des ressources
 - ➔ Perte d'efficacité



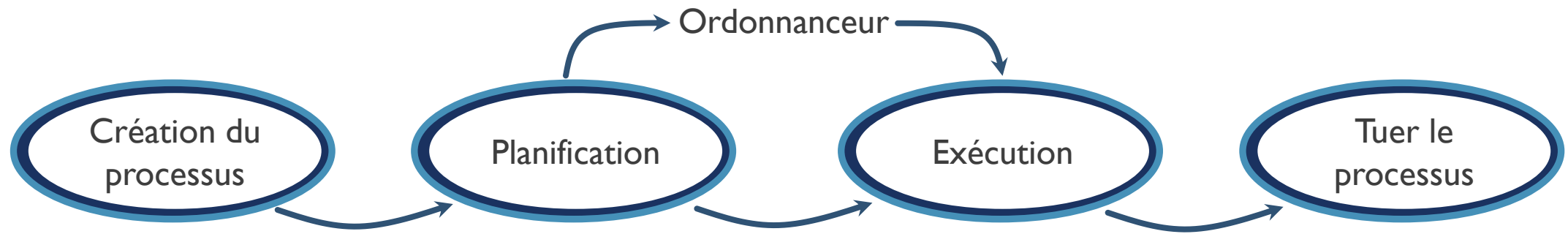
COMMENT OPTIMISER ?

- Mise en place de systèmes de gestion et d'organisation
- Systèmes tiers ayant une vue d'ensemble
 - ➔ Ne traite pas les informations, simplement les orienter
- Plusieurs niveaux d'application
 - ➔ Matérielle
 - CPU, RAM, Disque, etc...
 - ➔ Applicatif
 - Web, mail, jeux, etc...



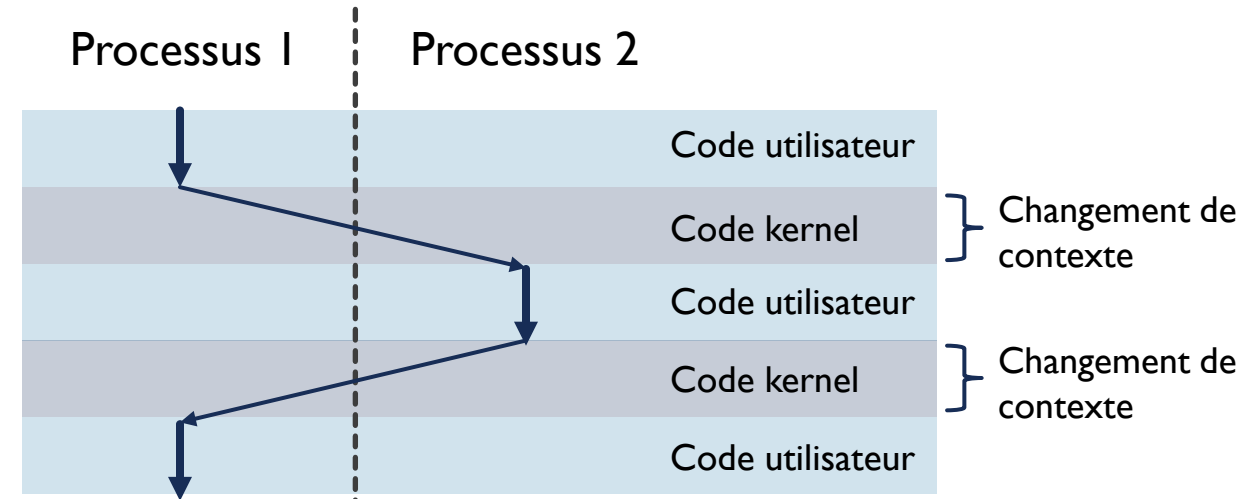
UTILISATION MATÉRIELLE

- **Objectif : réduire le temps d'attente d'accès aux ressources demandées**
 - Calcul, stockage court terme, stockage long terme
 - Très important dans des systèmes multi-process (tel les OS)
- Utilisation d'approches comme la « Gestion des processus »



GESTION DES PROCESSUS

- Chaque processus possède un mode de fonctionnement
 - Défini sa priorité, ses droits d'accès
 - Certains processus possèdent une priorité plus élevée (Kernel > tout)
- Changement de contexte lorsqu'un processus rencontre un évènement
 - Passe à l'état « attente », reçoit une interruption, etc.
 - Gestion du contexte seulement via le Kernel



ALGORITHMES DE PLANIFICATION DES PROCESSUS

- Plusieurs algorithmes de planification existent
 - **First-Come, First-Served (FCFS)**
 - Approche simplifié ordonné (*premier arrivé, premier servi*). Un processus en cours d'exécution continue jusqu'à sa fin, ou en attente d'un I/O
 - **Shortest Job First (SJF)**
 - Approche proactive sélectionnant le processus avec le temps d'exécution le plus court (burst)
 - **Round Robin (RR)**
 - Approche proactive avec une durée d'exécution fixe pour l'ensemble des processus, assurant une distribution équitable des ressources CPU. Un processus n'ayant pas fini son exécution est bloqué et mis à la fin de la file d'attente.
 - **Planification Prioritaire**
 - Approche associant un niveau de priorité à chaque processus (type, importance, besoins en ressources) dont le plus élevé est exécuté en premier.
 - **File d'Attente Multi-Niveau**
 - Approche divisant la file d'attente en plusieurs sous-files avec chacun un niveau de priorité. Chaque processus est mis dans la file appropriée, chacun avec son propre algorithme de planification.

AVANTAGES ET INCONVÉNIENTS

Avantages

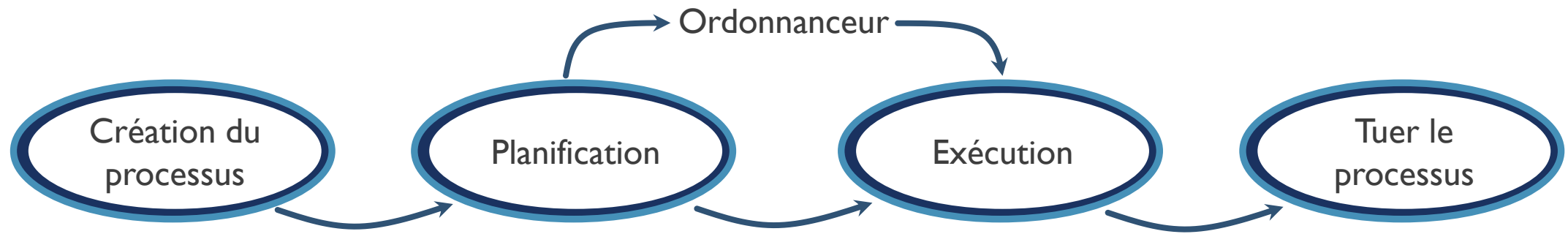
- Exécution parallélisée de plusieurs programmes
- Isolation de processus
 - Pas d'interférences
- Utilisation équitable des ressources
- Changements d'états facilité

Inconvénients

- Charge élevée
 - Besoin d'une utilisation des ressources de l'OS pour mieux structurer et suivre les files
- Grande complexité
- Possibilité de blockage
 - Utilisation d'approches tel semaphores et mutex, qui peuvent se retrouver bloquer
- Beaucoup de changements de contexte
 - Changements fréquents avec stockage et chargement des états → utilisation des ressources

GESTION DE LA MÉMOIRE

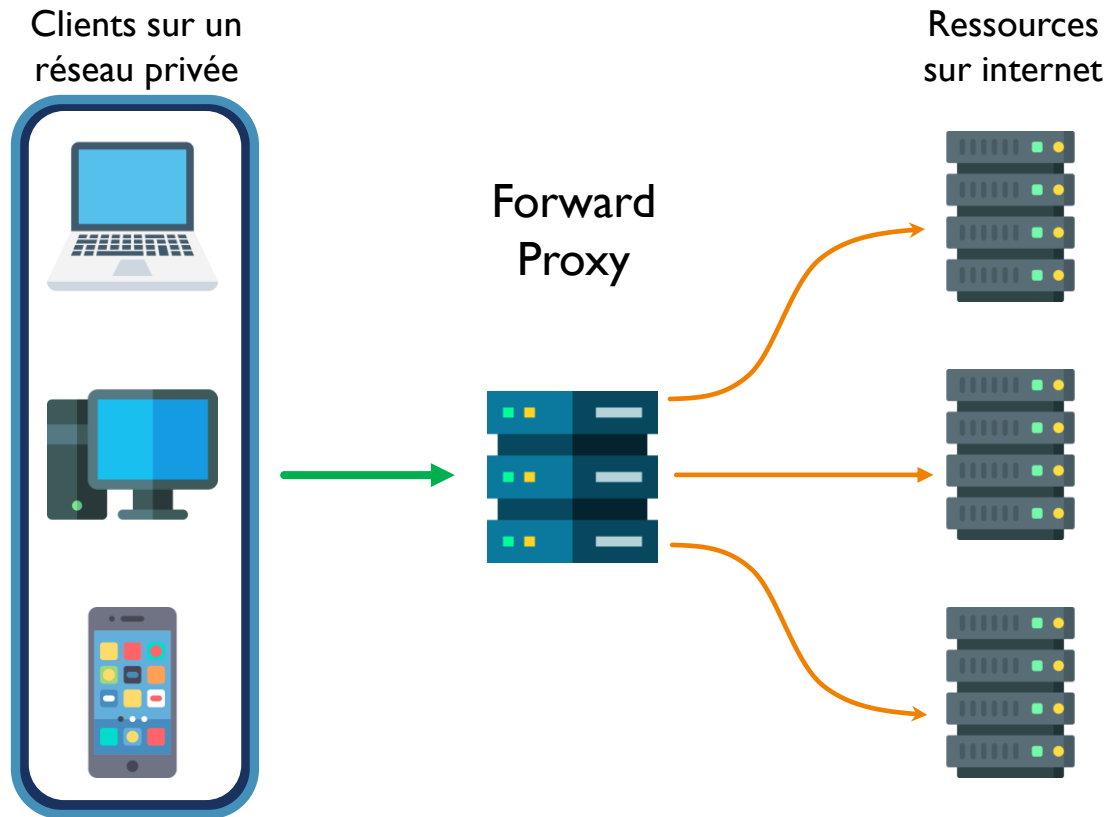
- **Objectif : réduire le temps d'attente d'accès aux ressources demandées**
 - Calcul, stockage court terme, stockage long terme
 - Très important dans des systèmes multi-process (tel les OS)
- Utilisation d'approches comme la « Gestion des processus »



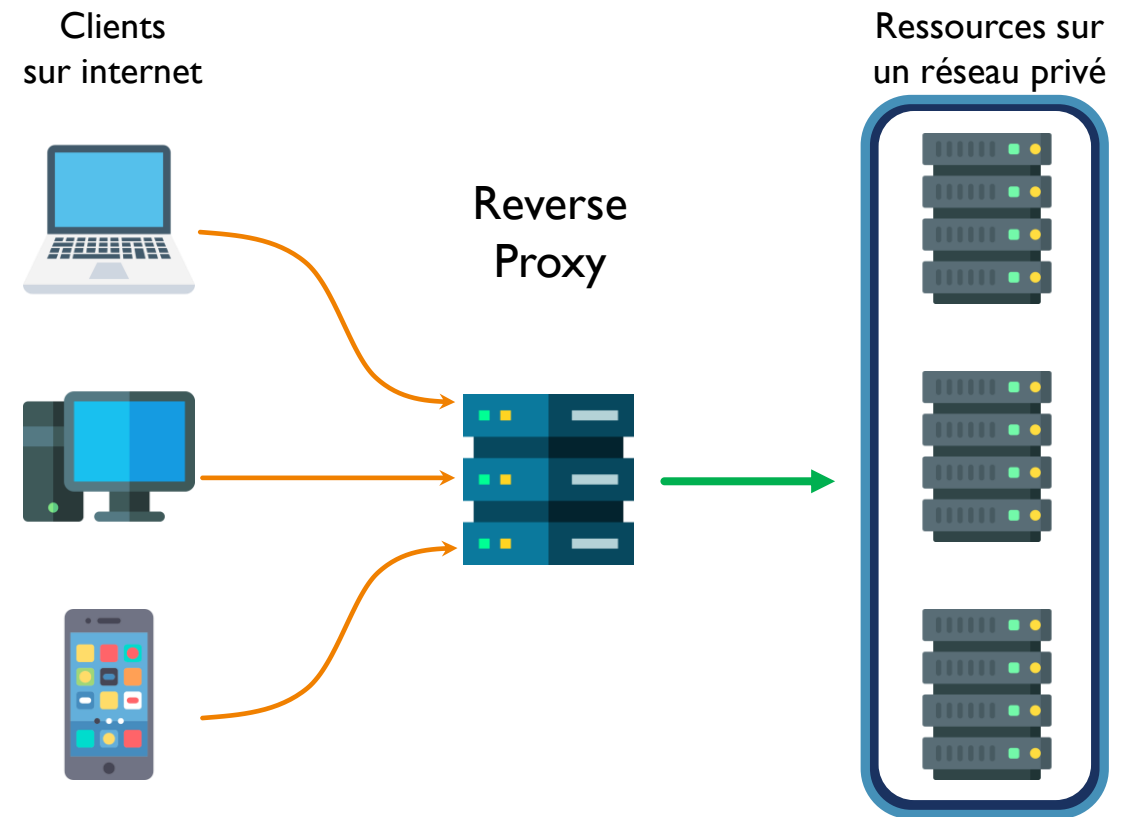
UTILISATION APPLICATIF

- Couramment utilisé dans les applications internet
- Beaucoup de demandes d'accès aux informations web, mail, etc.
- Proposition de méthodes pour gérer les politiques d'accès, le routage d'information et la sécurité des systèmes
 - **Reverse Proxy**

PROXY ?



BUT INFO - R6.B.05



© UBS/ES

14

UTILISATIONS DE SERVEURS PROXYS

Serveur proxy
==
Forward Proxy

■ Cache

- Mise en place d'un système de cache → *Partie II*

■ Anonyme

- Protège l'identité des entités derrière
 - Forward Proxy → clients
 - Reverse Proxy → serveurs

■ Sécurité

- Agit comme un pare-feu, permettant le déploiement de filtres au niveau des entreprises / FAI

■ Chiffrement & déchiffrement

- Peut chiffrer et cacher les données transmises ou l'adresse IP (NAT)

■ Logging

- Identifier les actions répétitives → peut servir pour le cache

■ Micro-services

- Utilisation récente (Proxy HTTP, TCP, Side-Car).
Déploiement à côté de l'application pour s'occuper des communications réseaux
 - *Proxy Side-Car* → passer une requête HTTP en HTTP2/HTTP3

APPORTS DES REVERSE PROXY

■ **Entrée de requêtes**

- Utilisé dans les architectures micro-services ou kubernetes
- Orientation de la demande basée sur la structure et contenu de la requête de manière anonyme
 - HTTP → vers le serveur web
 - SMTP → vers le serveur mail

■ **Déploiement Canary**

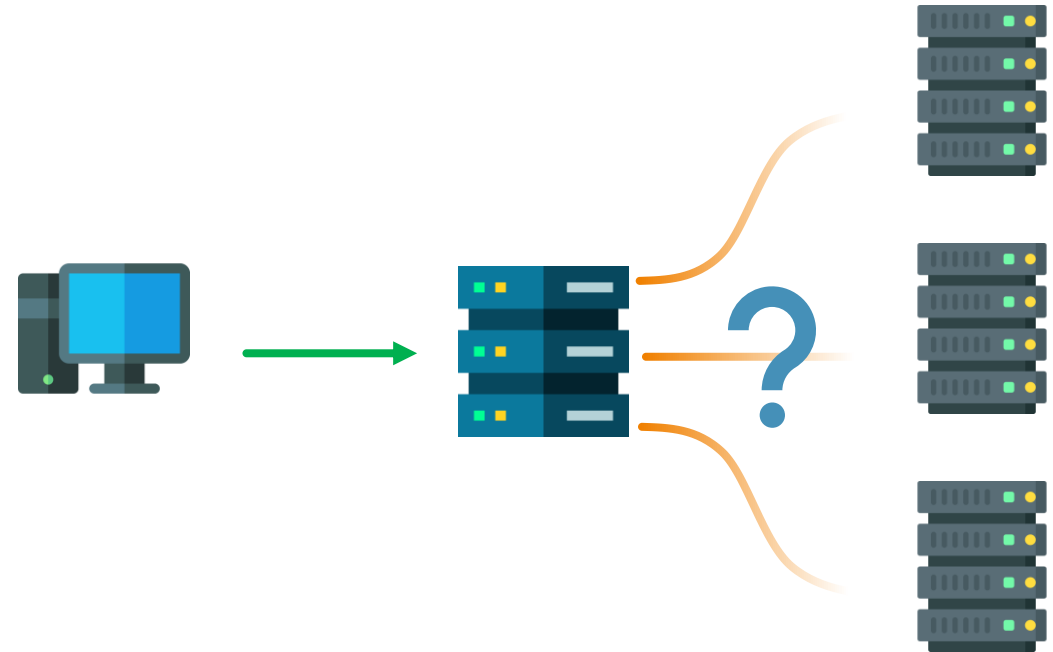
- Déploiement progressive d'une nouvelle version d'une application
 - Sous ensemble d'utilisateurs ont accès à la nouvelle version

■ **Load balancing**

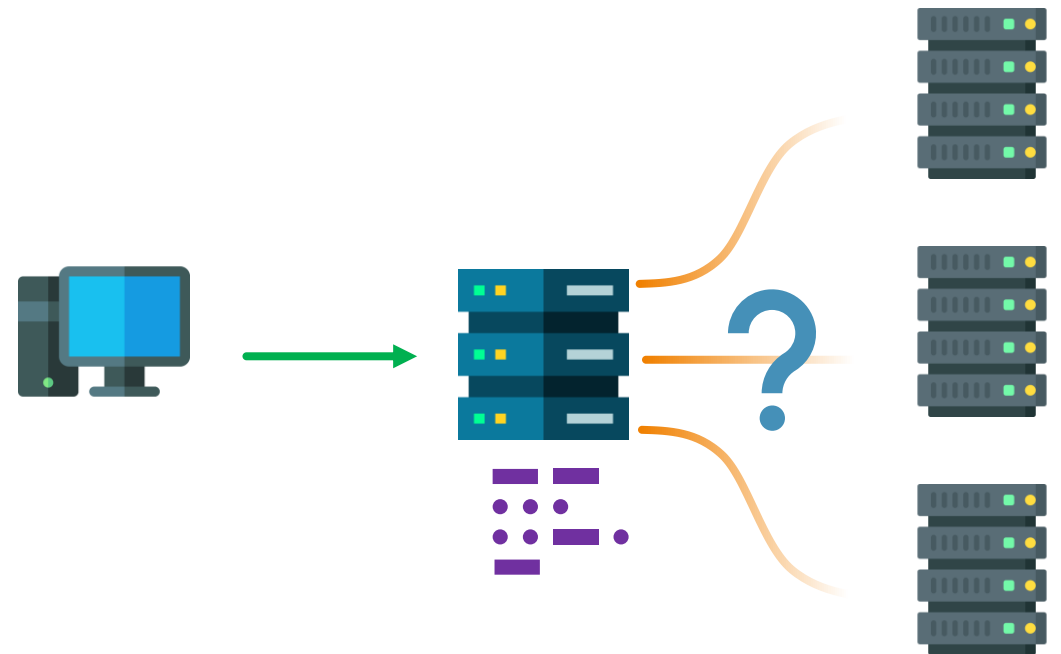
- Méthode d'équilibre de charge entre plusieurs serveurs
- Plusieurs algorithmes existent
- Permet également la compression des requêtes
 - Augmentation de la performance

LOAD BALANCING

- Approche importante lorsque plusieurs serveurs existent
 - Certains peuvent être submergés de requêtes → pas optimal
- Il existe 2 approches principaux
 - **Statique**
 - Distribution basée sur des paramètres configurés
 - **Dynamique**
 - Distribution basée sur l'état des serveurs → *peut s'adapter*



APPROCHES STATIQUES



ALÉATOIRE



ALÉATOIRE



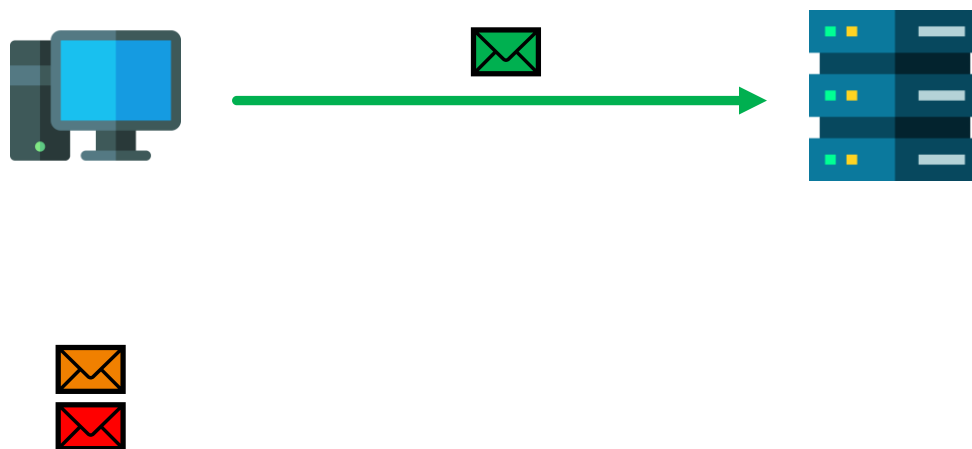
ALÉATOIRE



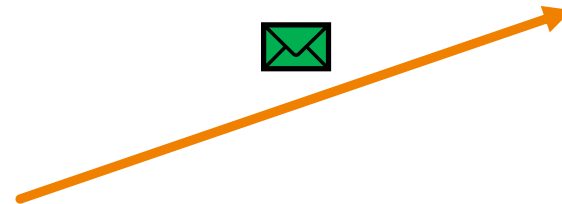
ALÉATOIRE



ALÉATOIRE



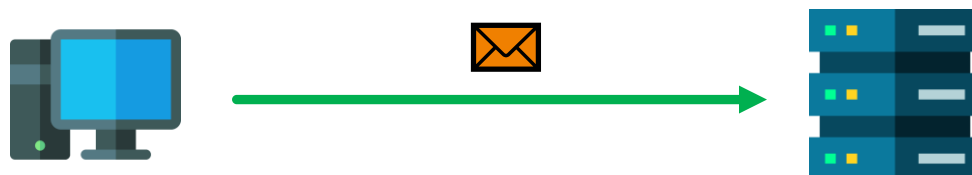
ALÉATOIRE



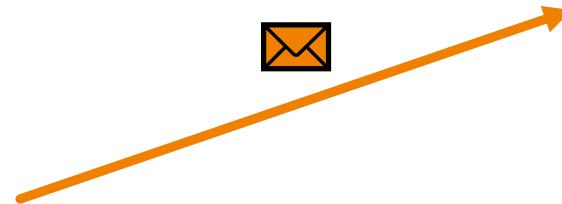
ALÉATOIRE



ALÉATOIRE



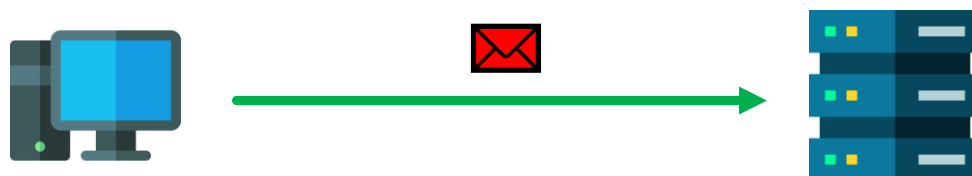
ALÉATOIRE



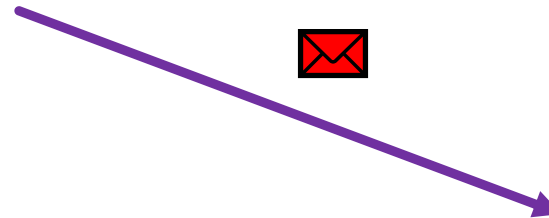
ALÉATOIRE



ALÉATOIRE



ALÉATOIRE



ALÉATOIRE



ROUND ROBIN



1



1



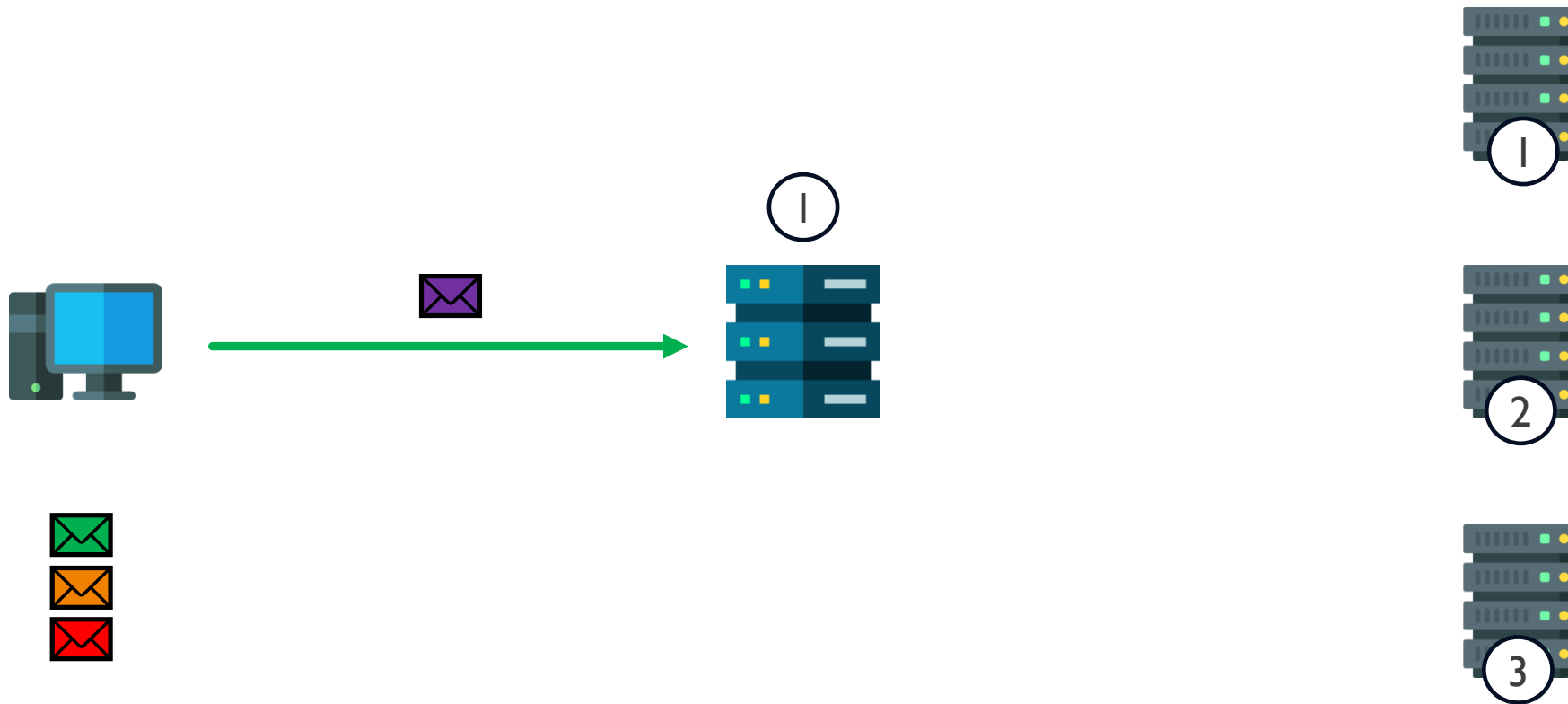
2



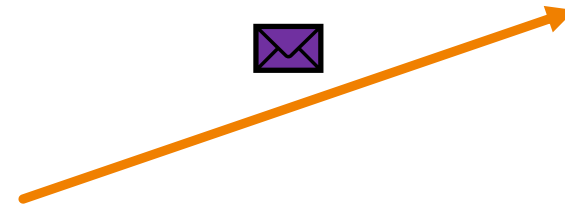
3



ROUND ROBIN



ROUND ROBIN



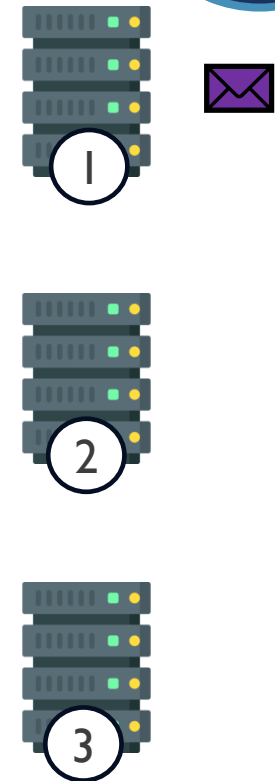
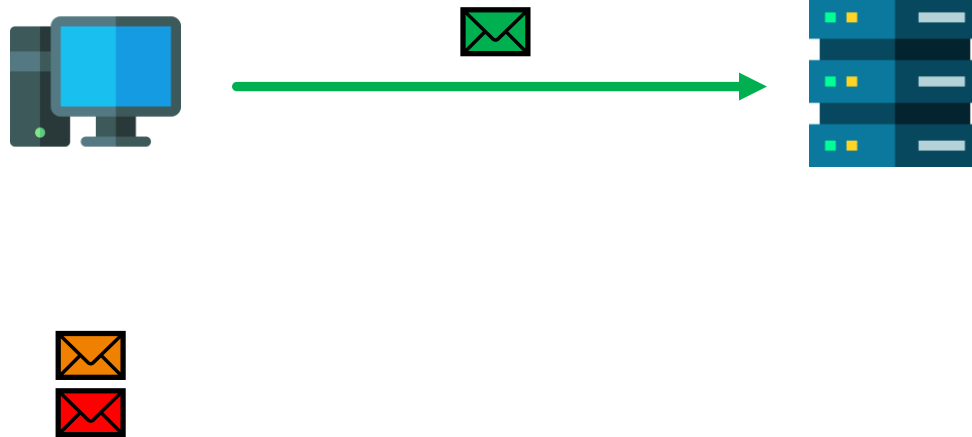
ROUND ROBIN



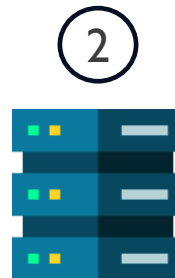
2



ROUND ROBIN



ROUND ROBIN



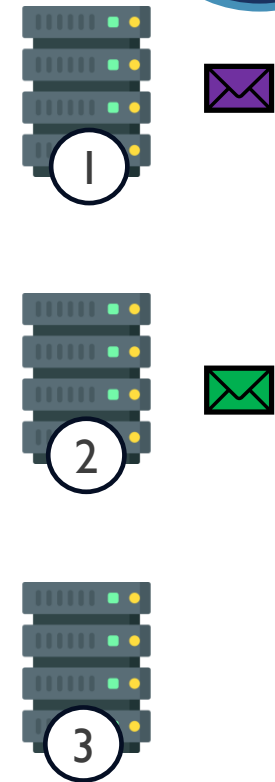
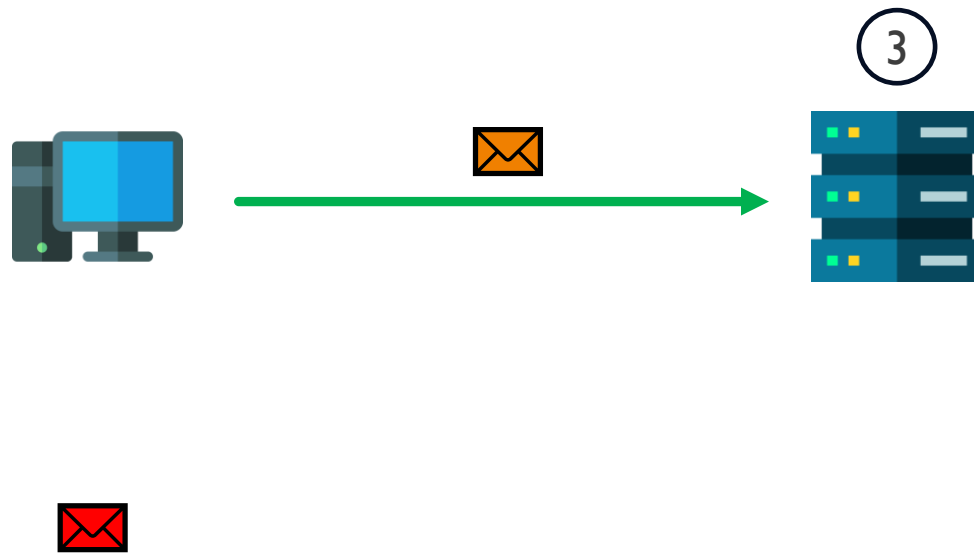
ROUND ROBIN



3



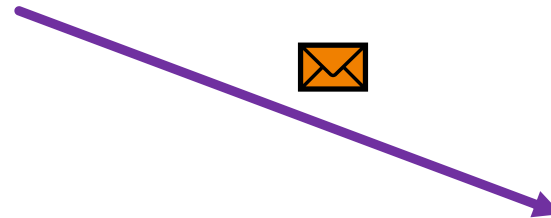
ROUND ROBIN



ROUND ROBIN



3



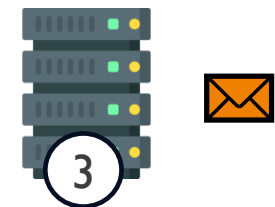
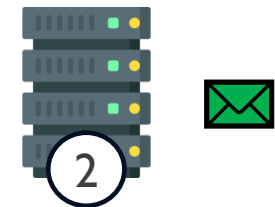
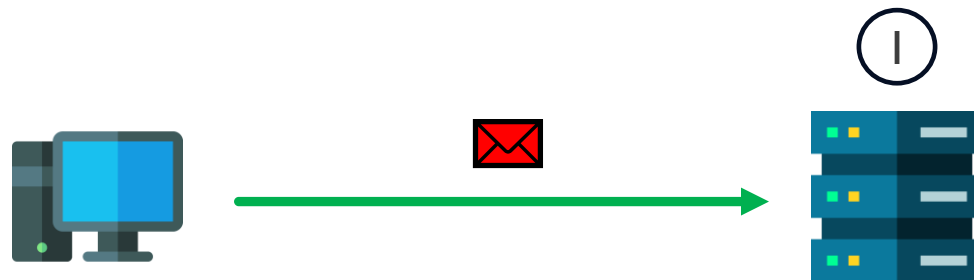
ROUND ROBIN



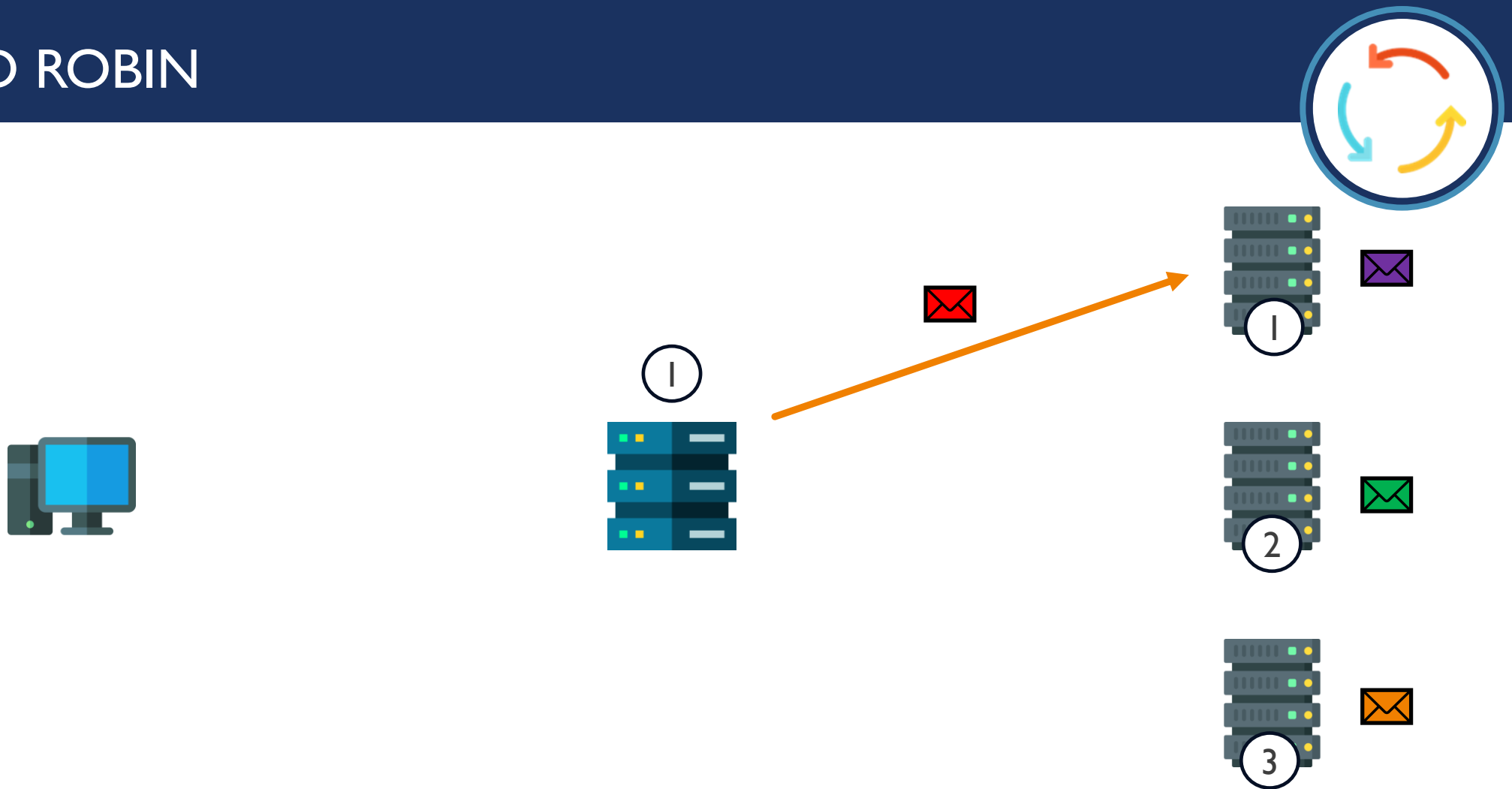
1



ROUND ROBIN



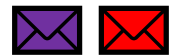
ROUND ROBIN



ROUND ROBIN



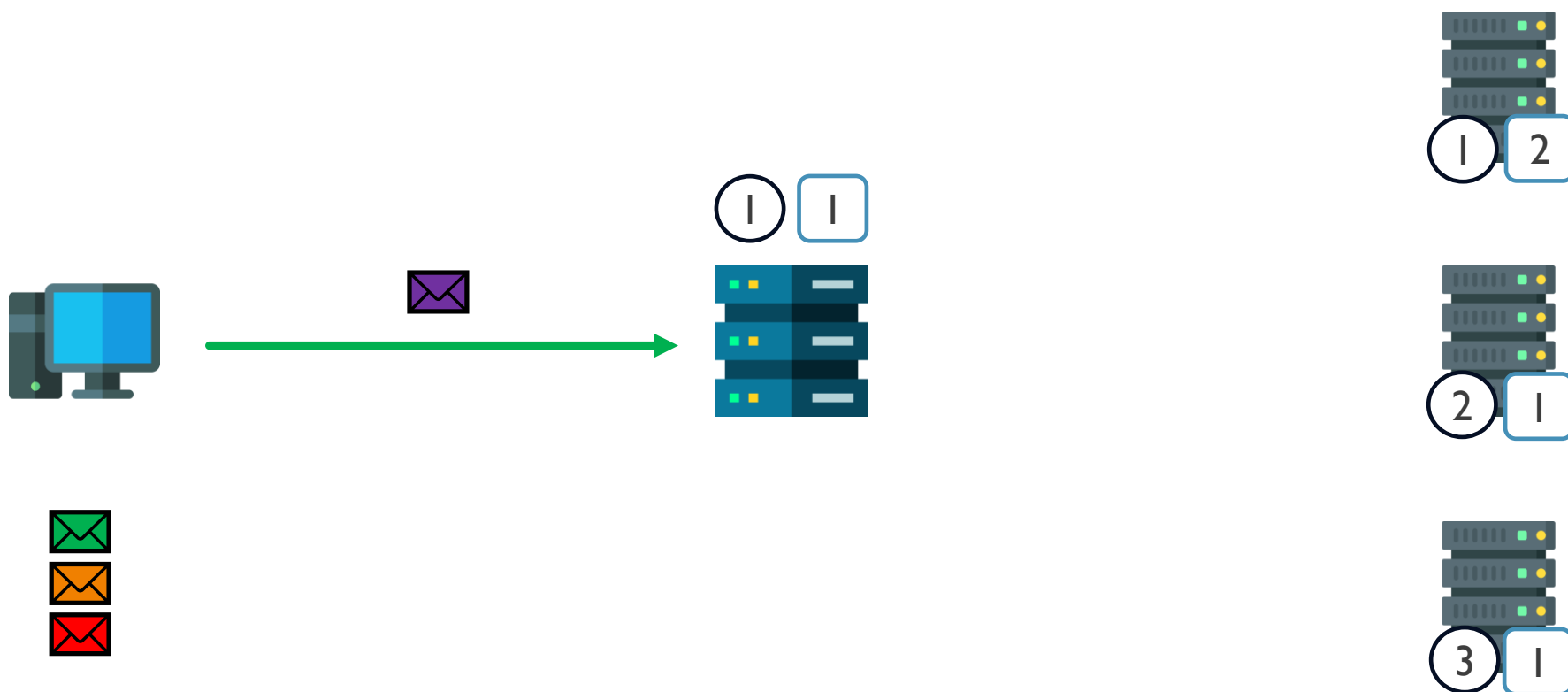
2



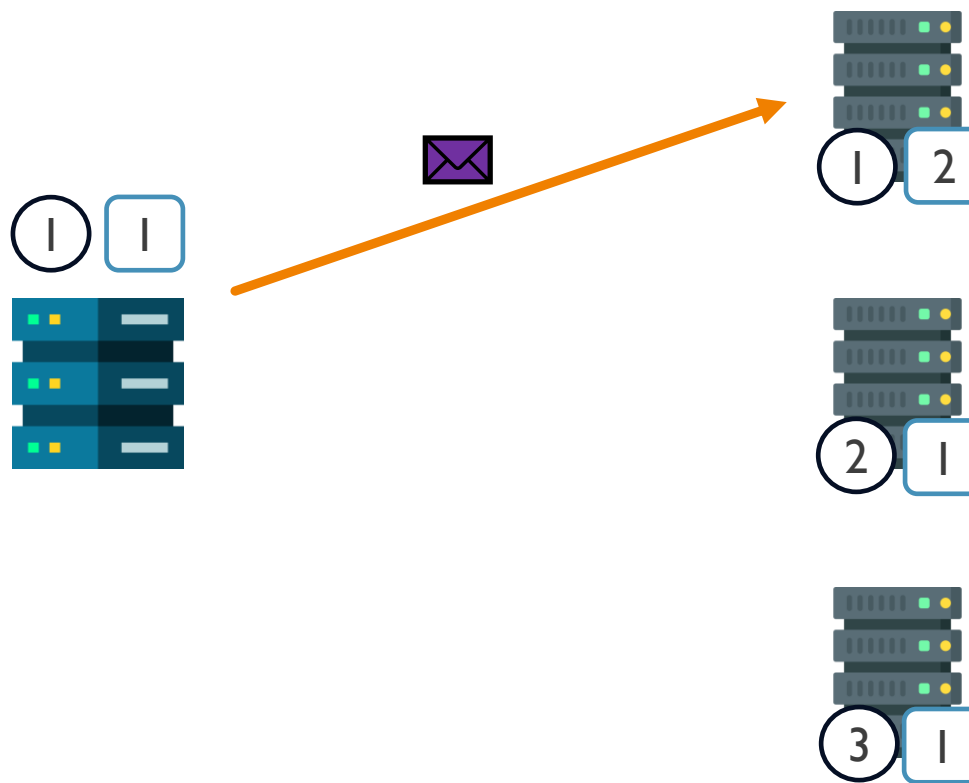
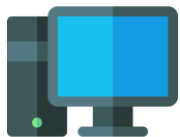
ROUND ROBIN PONDÉRÉ



ROUND ROBIN PONDÉRÉ



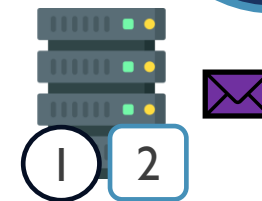
ROUND ROBIN PONDÉRÉ



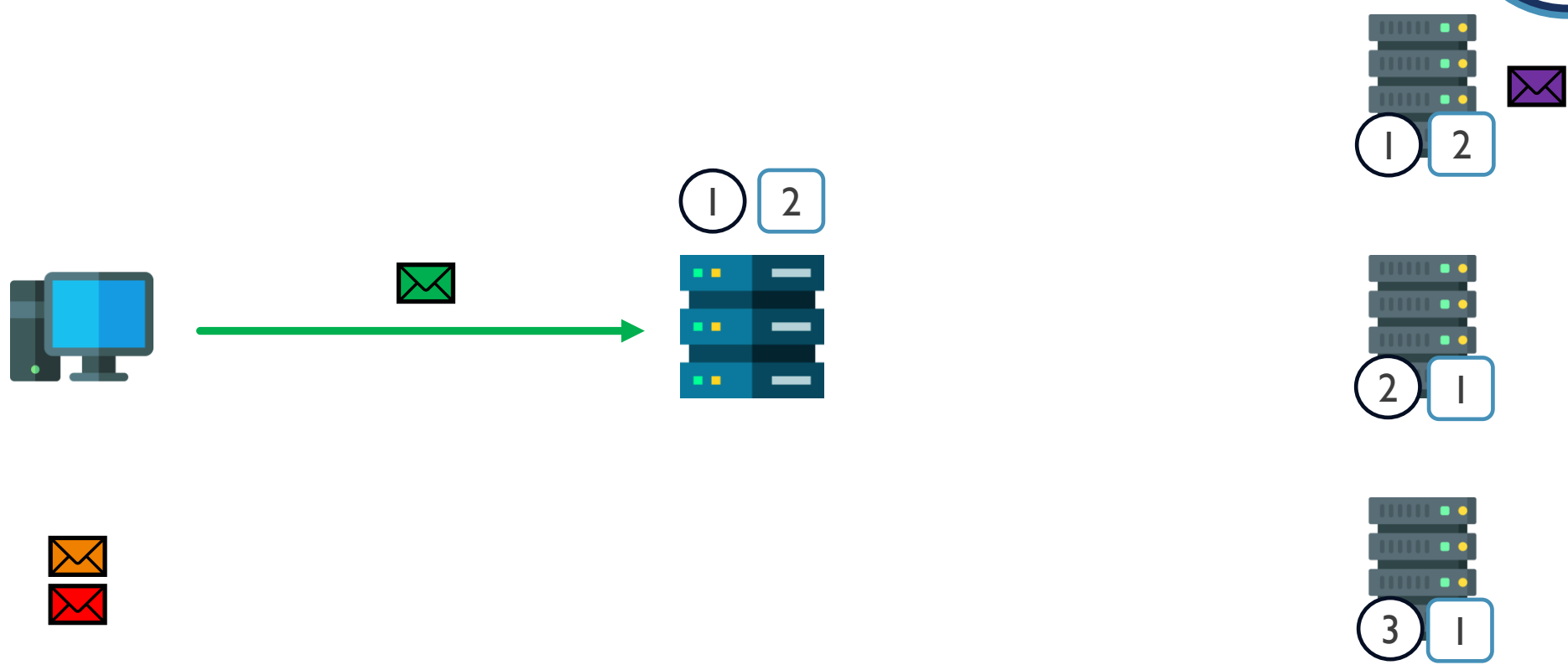
ROUND ROBIN PONDÉRÉ



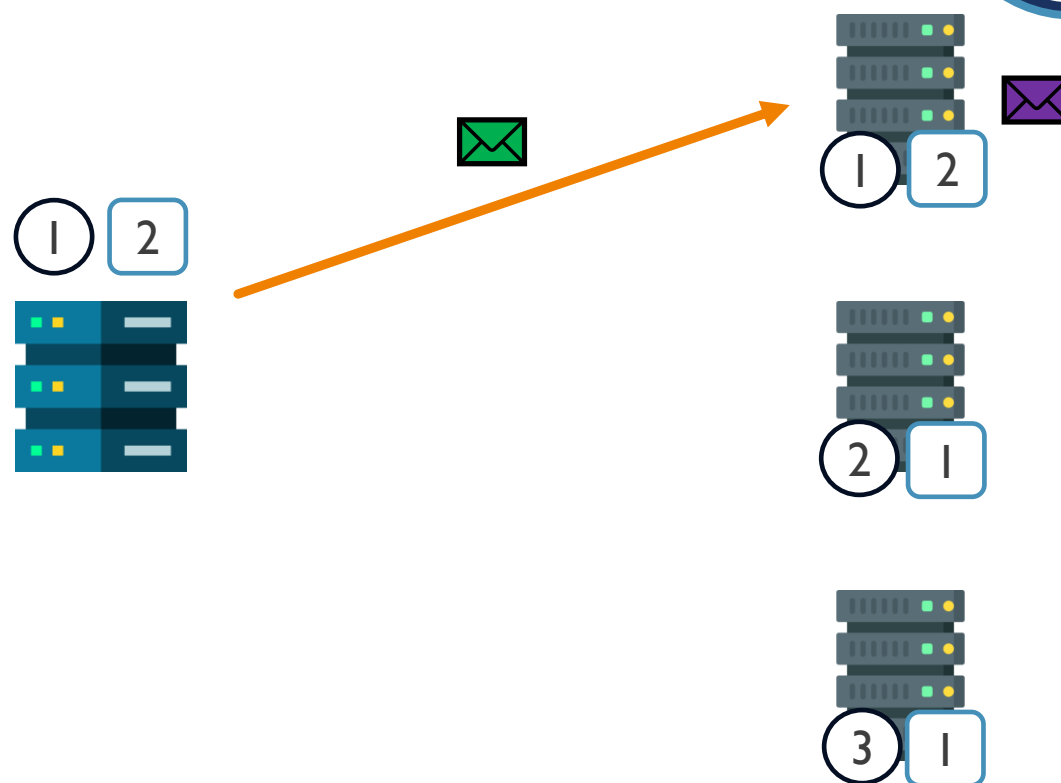
1 2



ROUND ROBIN PONDÉRÉ



ROUND ROBIN PONDÉRÉ



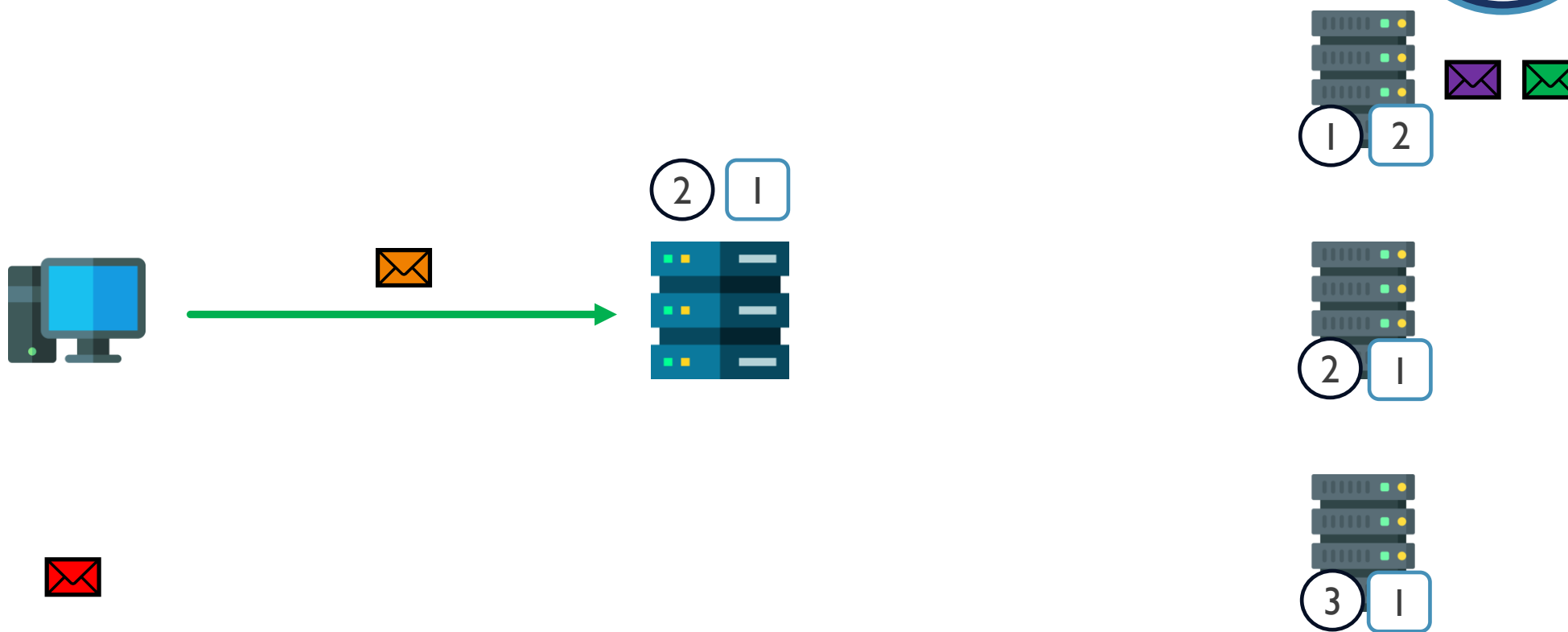
ROUND ROBIN PONDÉRÉ



2 1



ROUND ROBIN PONDÉRÉ



ROUND ROBIN PONDÉRÉ



2 1



1 2



2 1



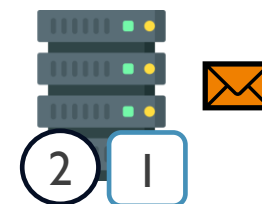
3 1



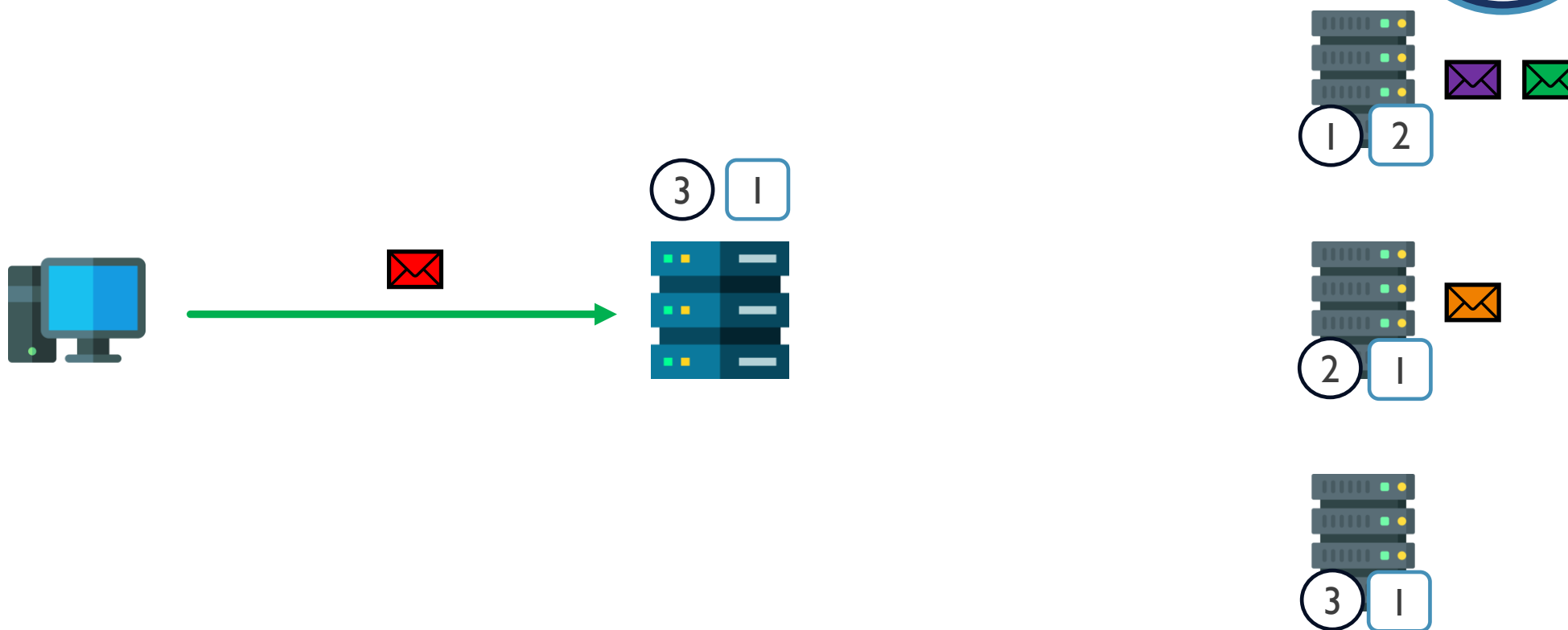
ROUND ROBIN PONDÉRÉ



3 1



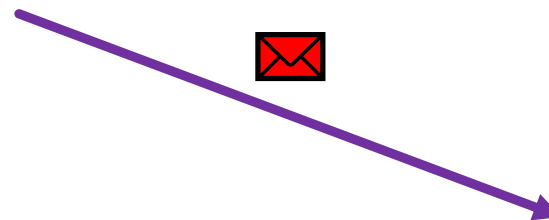
ROUND ROBIN PONDÉRÉ



ROUND ROBIN PONDÉRÉ



3 1



1 2



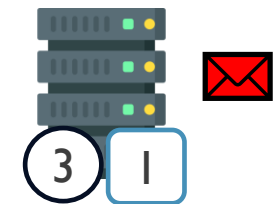
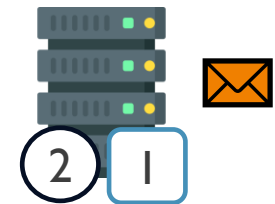
2 1



3 1



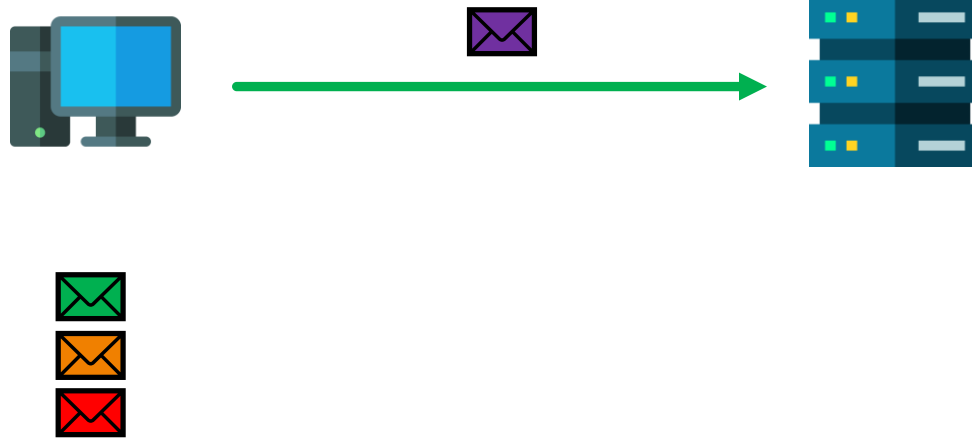
ROUND ROBIN PONDÉRÉ



HACHAGE D'IP



HACHAGE D'IP



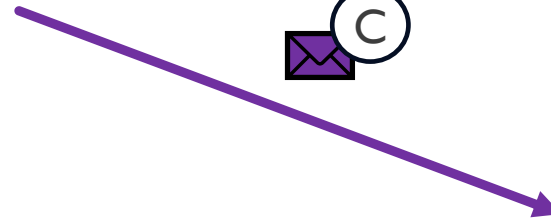
HACHAGE D'IP



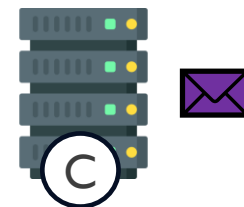
HACHAGE D'IP



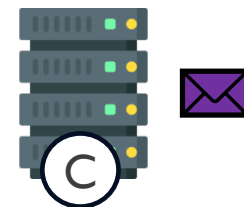
HACHAGE D'IP



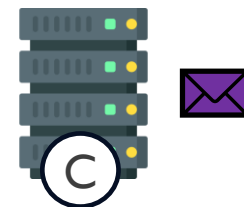
HACHAGE D'IP



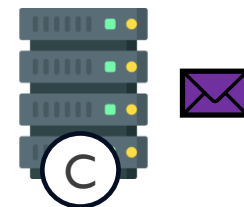
HACHAGE D'IP



HACHAGE D'IP



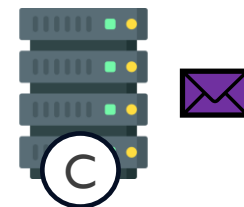
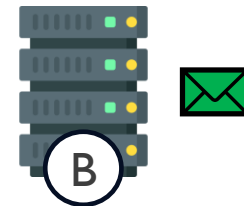
HACHAGE D'IP



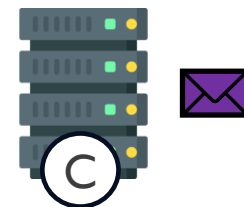
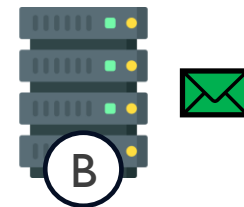
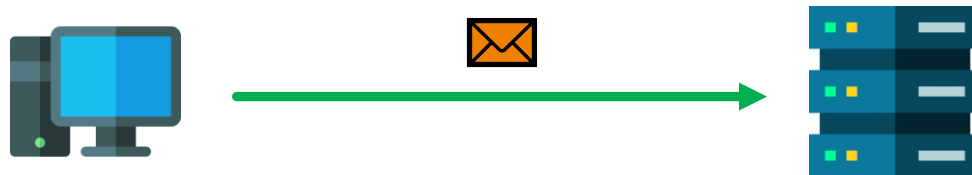
HACHAGE D'IP



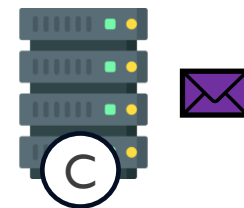
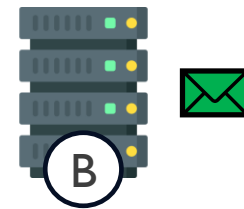
HACHAGE D'IP



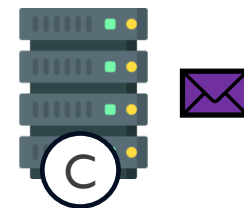
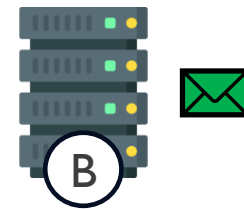
HACHAGE D'IP



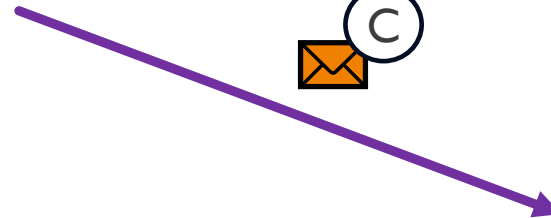
HACHAGE D'IP



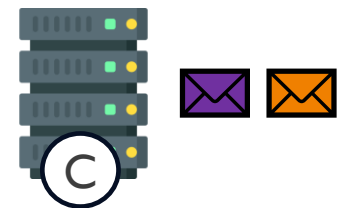
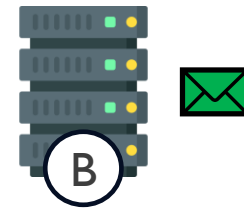
HACHAGE D'IP



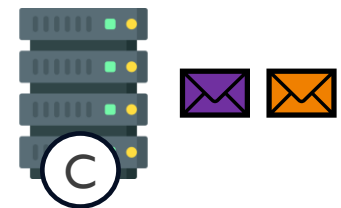
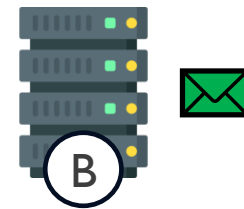
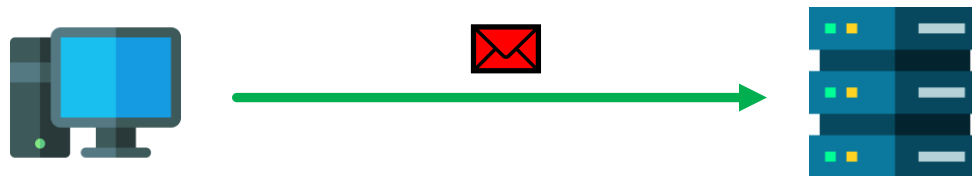
HACHAGE D'IP



HACHAGE D'IP



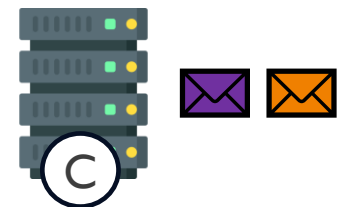
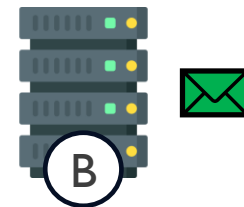
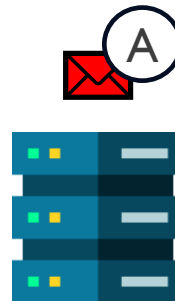
HACHAGE D'IP



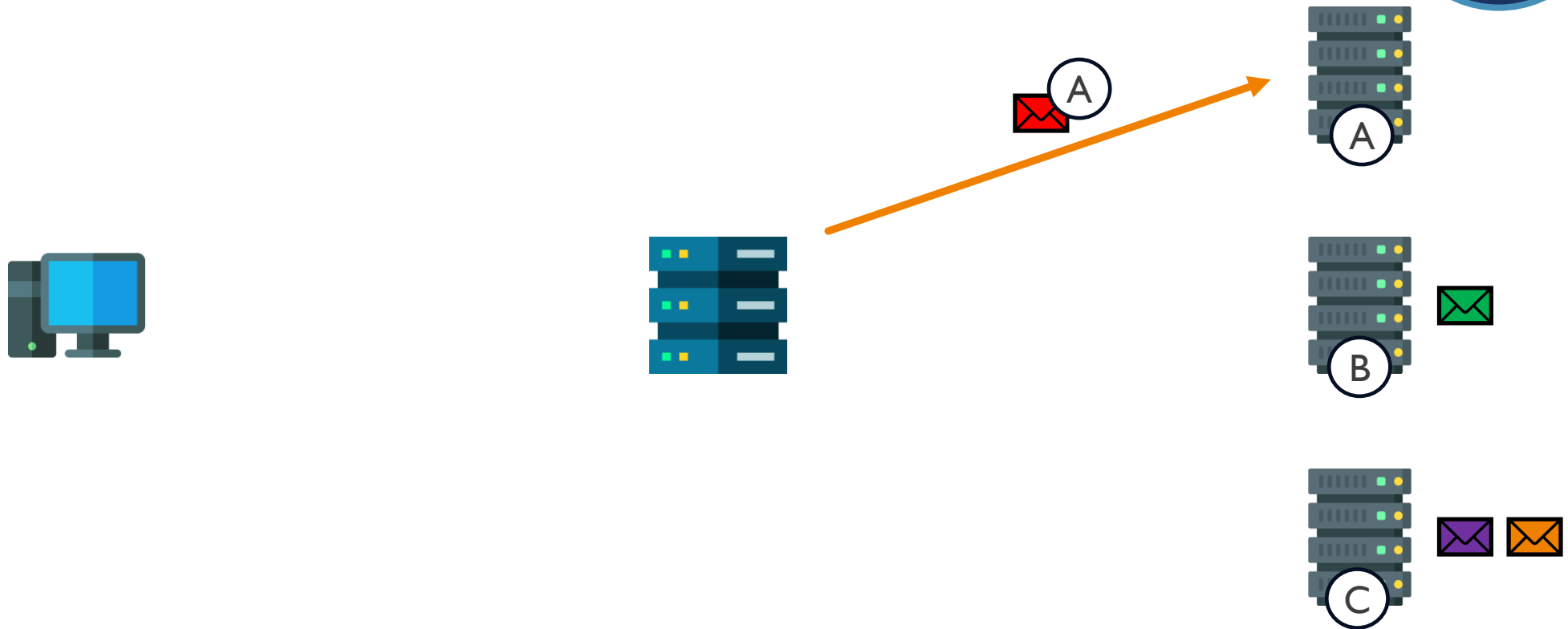
HACHAGE D'IP



HACHAGE D'IP



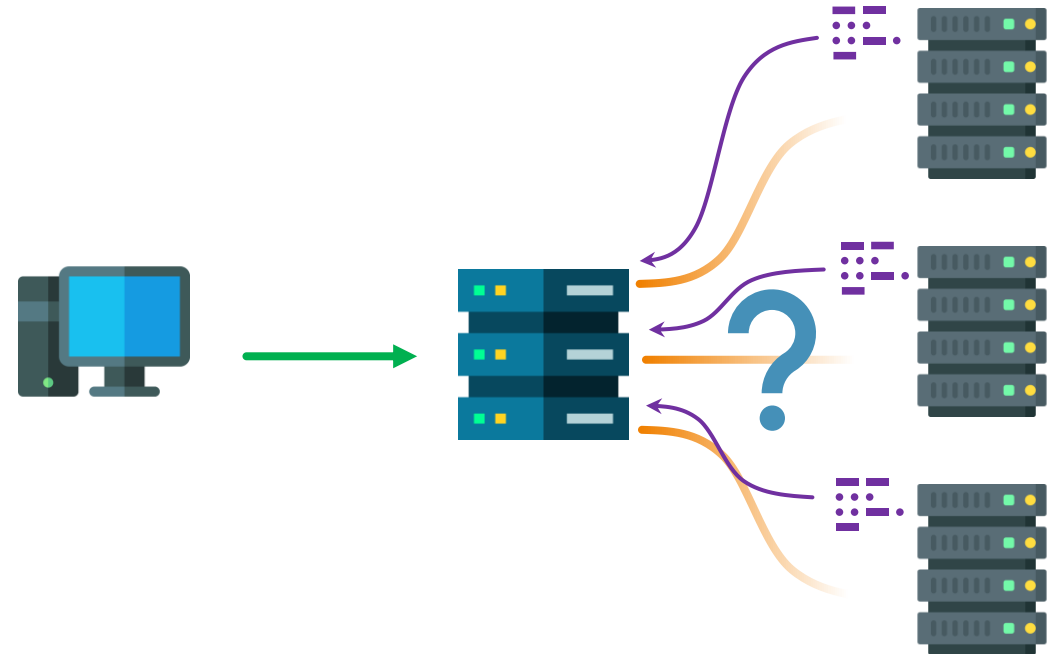
HACHAGE D'IP



HACHAGE D'IP



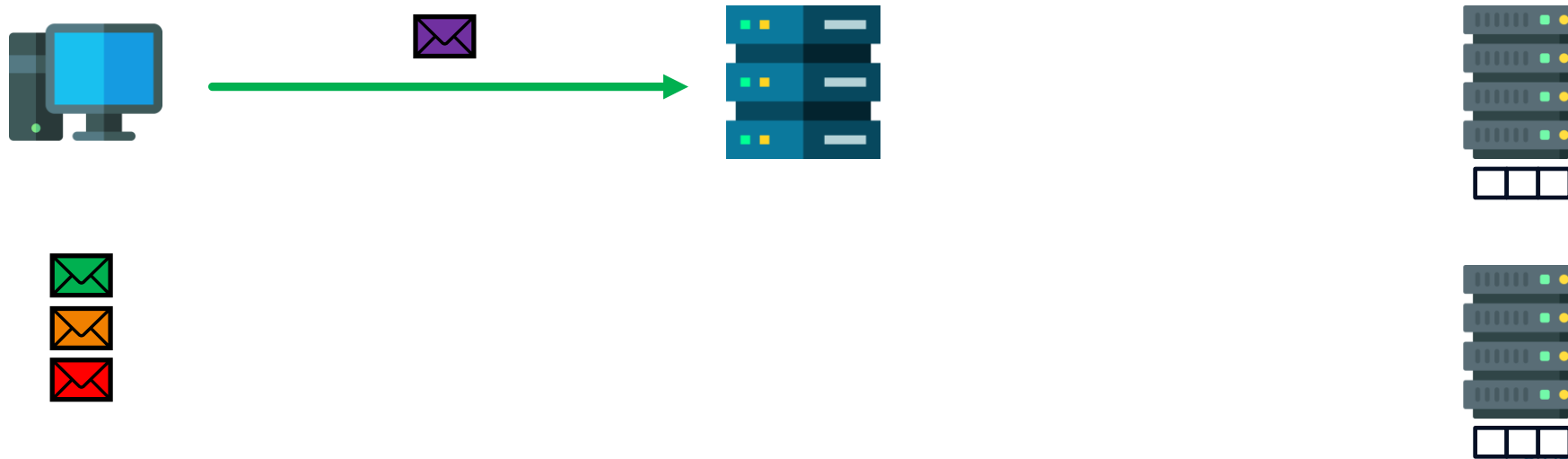
APPROCHES DYNAMIQUES



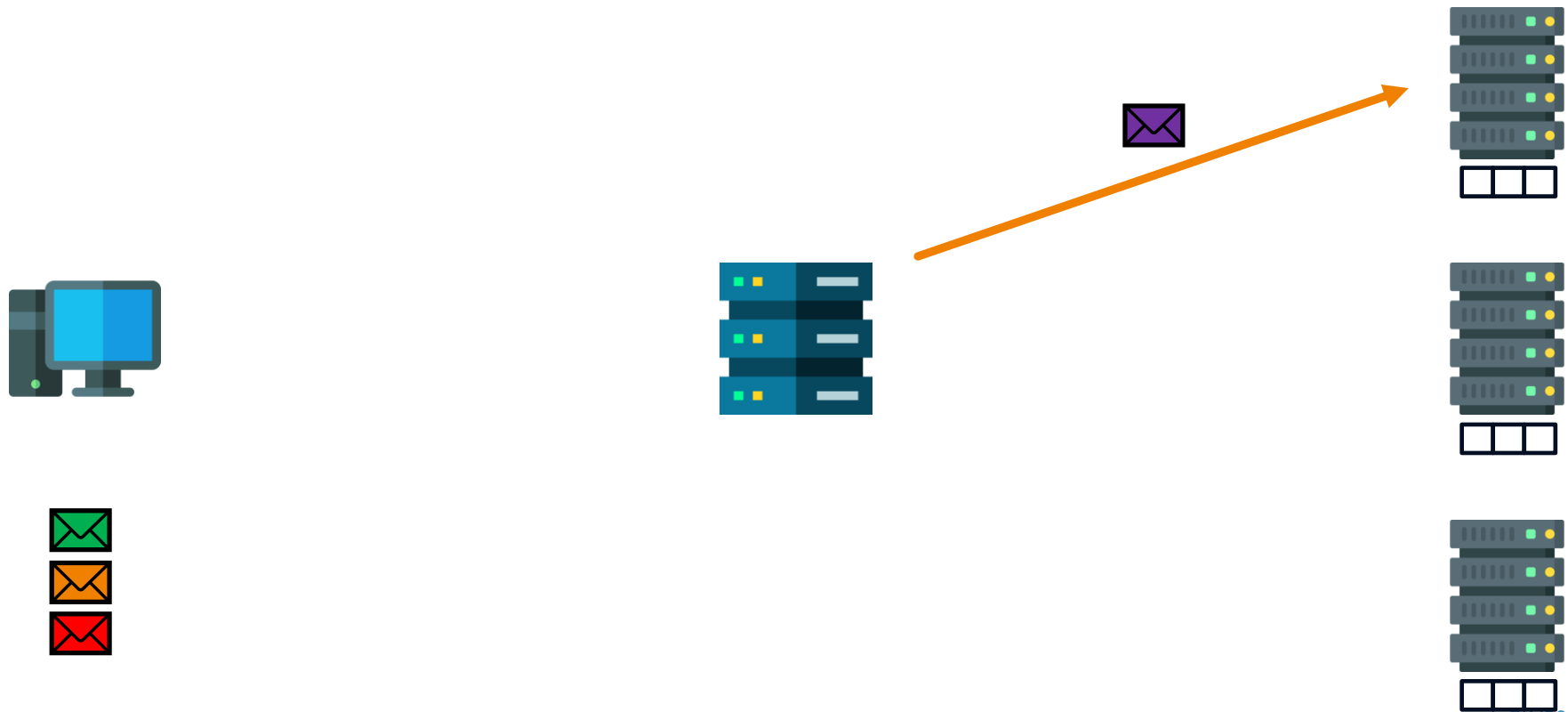
LEAST CONNECTION



LEAST CONNECTION



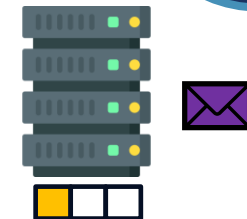
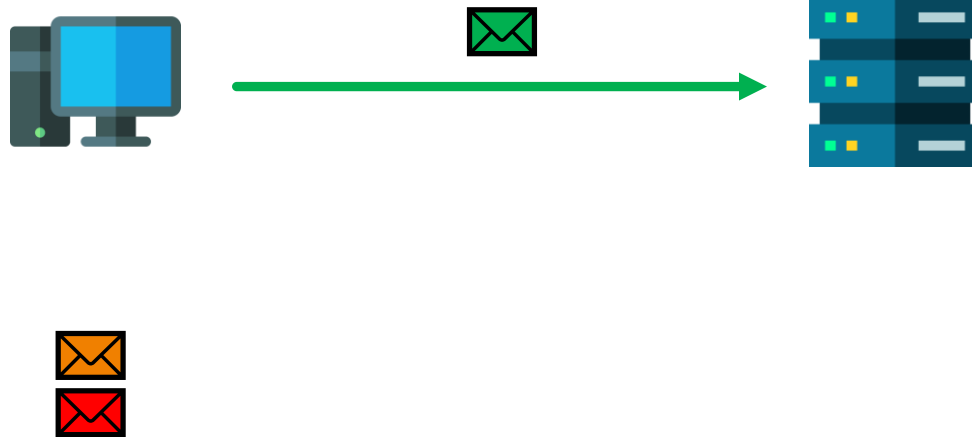
LEAST CONNECTION



LEAST CONNECTION



LEAST CONNECTION



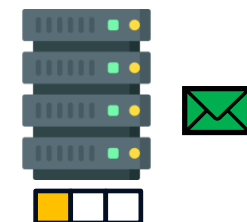
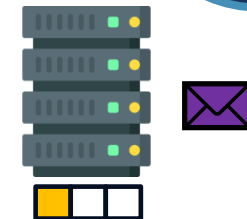
LEAST CONNECTION



LEAST CONNECTION



LEAST CONNECTION



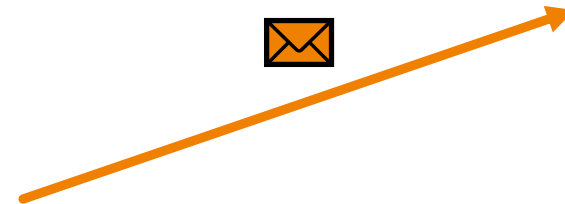
LEAST CONNECTION



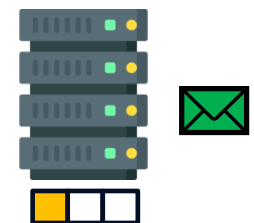
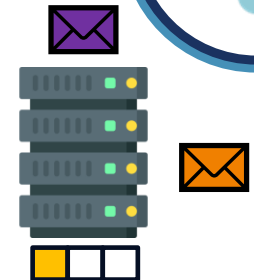
LEAST CONNECTION



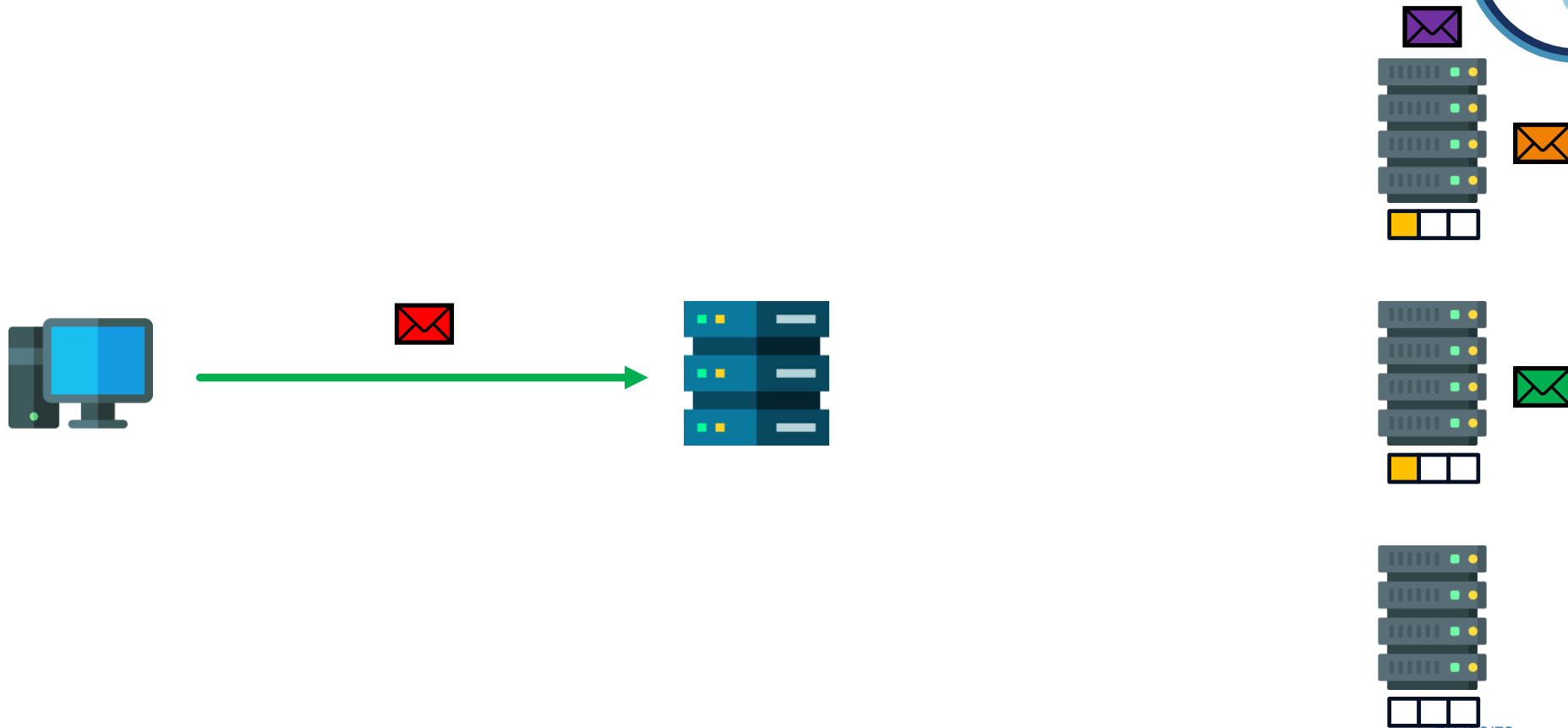
LEAST CONNECTION



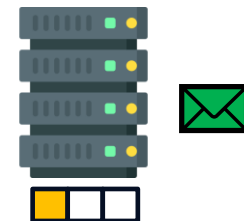
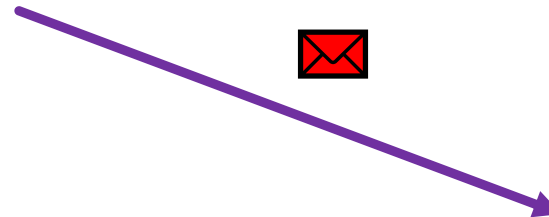
LEAST CONNECTION



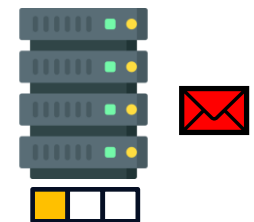
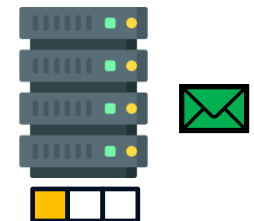
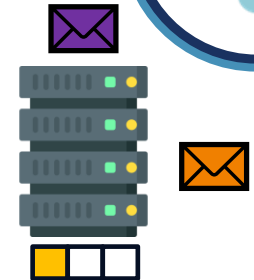
LEAST CONNECTION



LEAST CONNECTION



LEAST CONNECTION

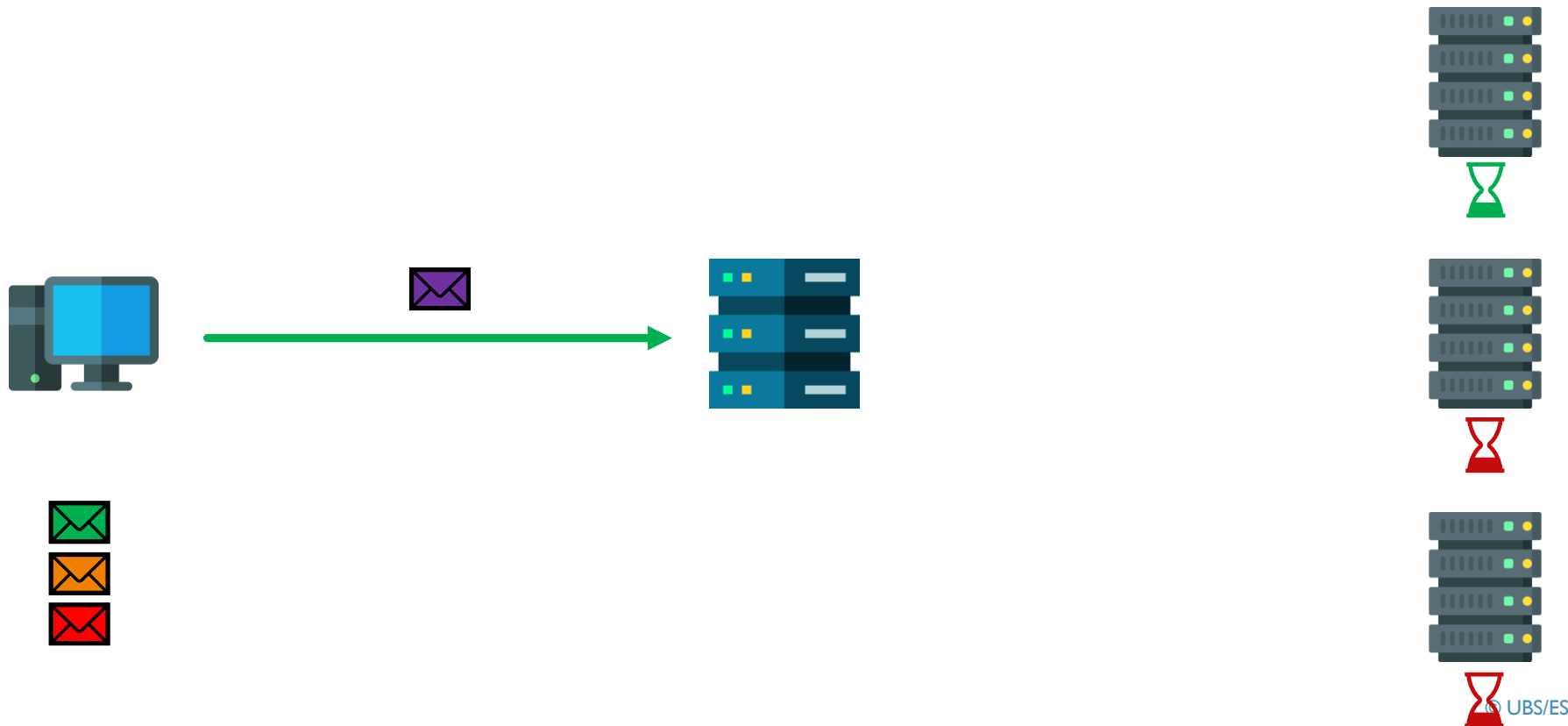


TEMPS DE RÉPONSE

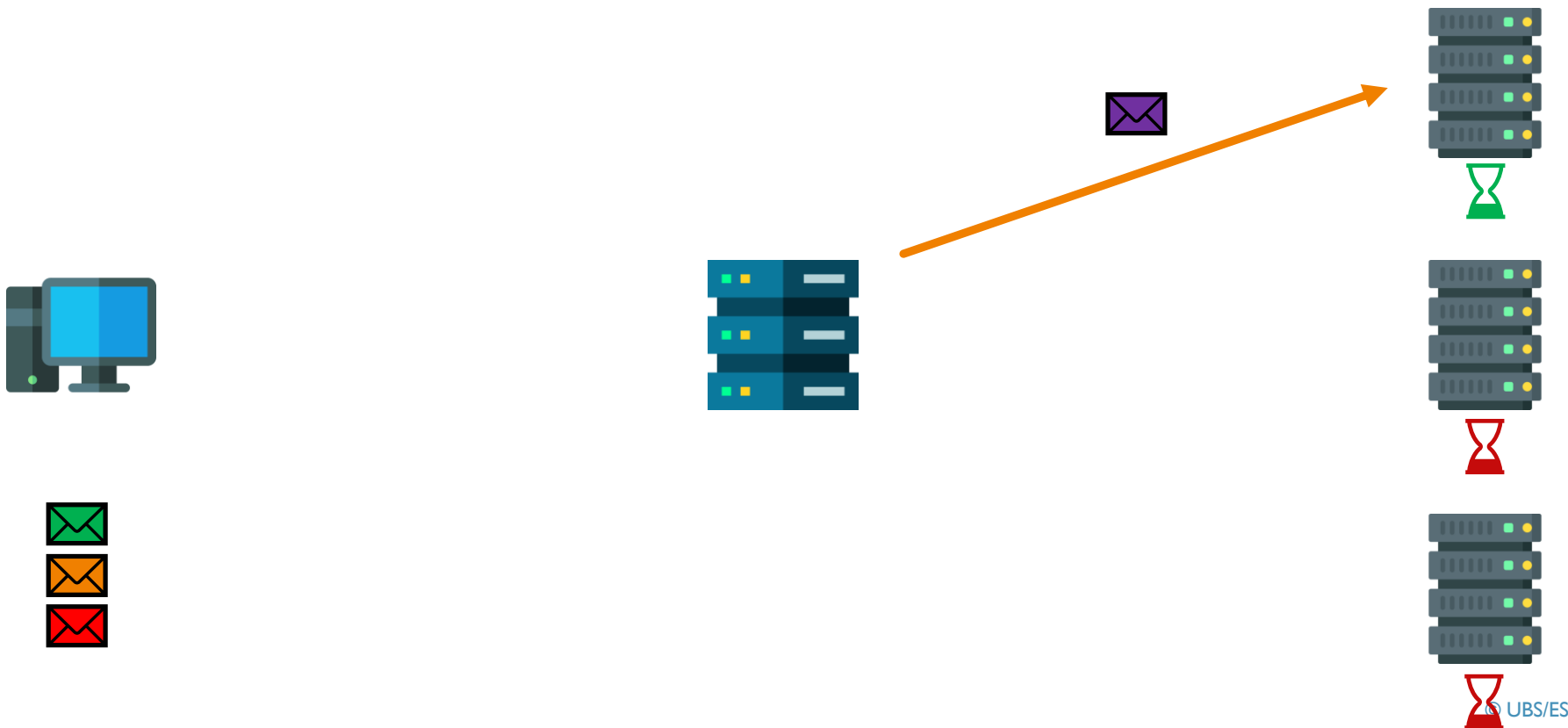


UBS/ES

TEMPS DE RÉPONSE



TEMPS DE RÉPONSE

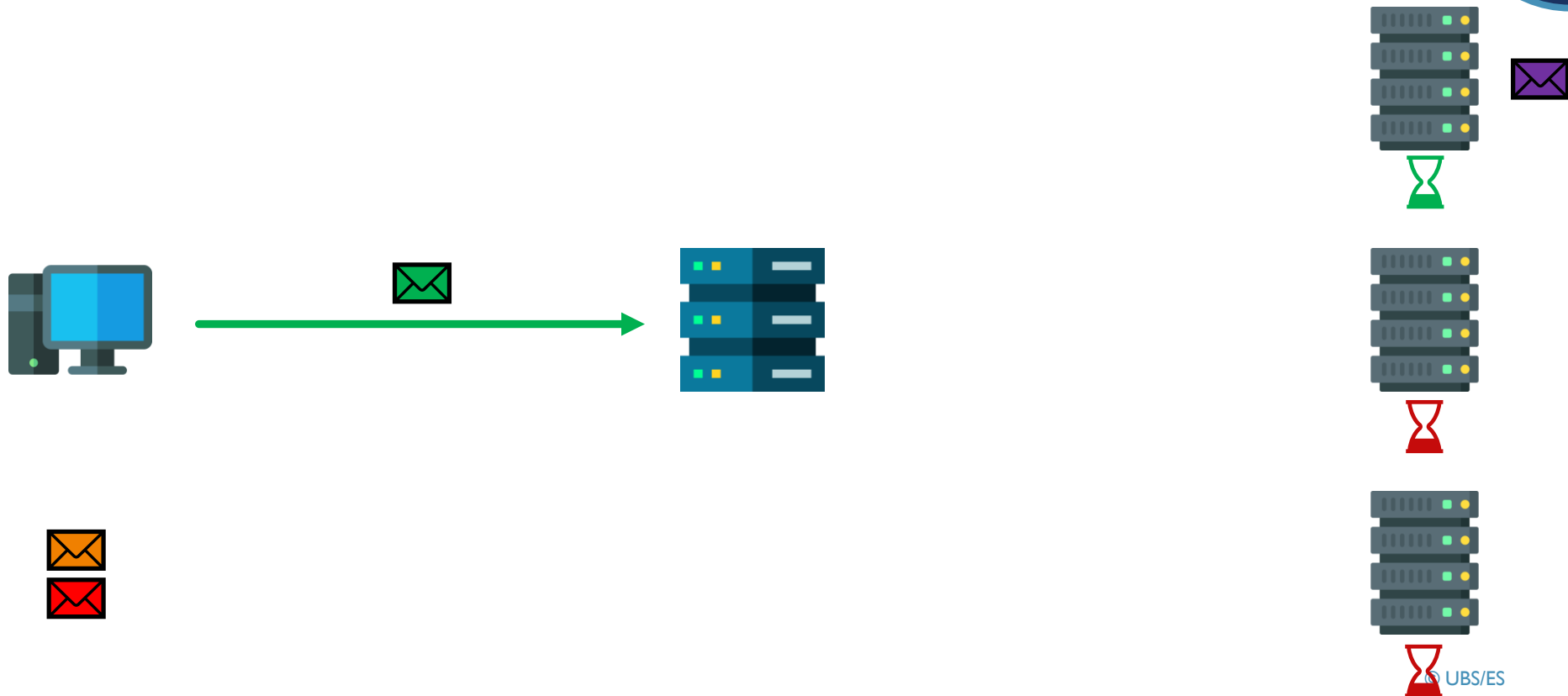


TEMPS DE RÉPONSE

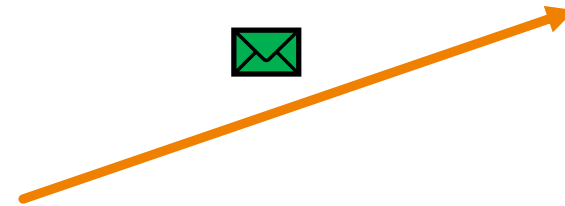


UBS/ES

TEMPS DE RÉPONSE



TEMPS DE RÉPONSE



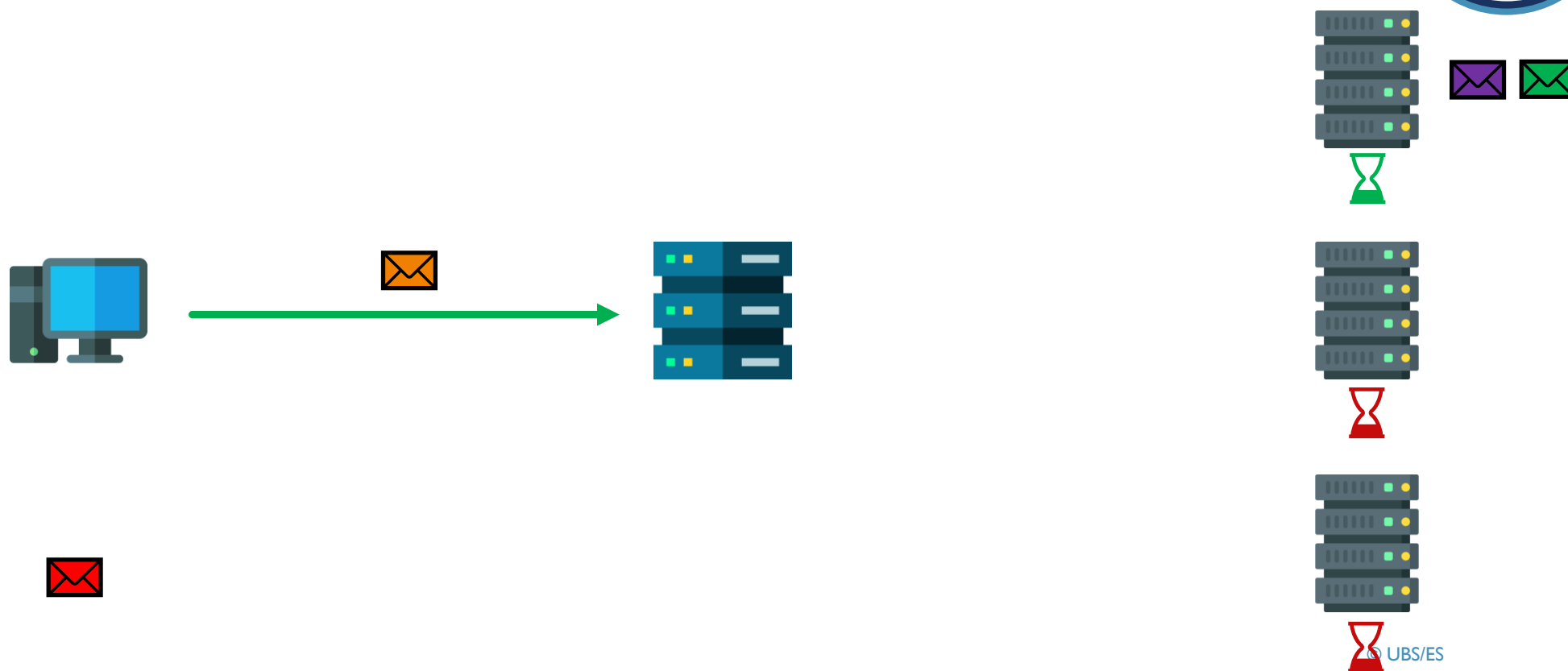
UBS/ES

TEMPS DE RÉPONSE

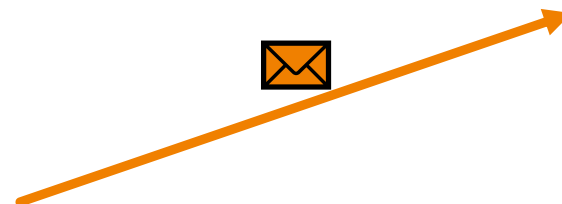


UBS/ES

TEMPS DE RÉPONSE



TEMPS DE RÉPONSE

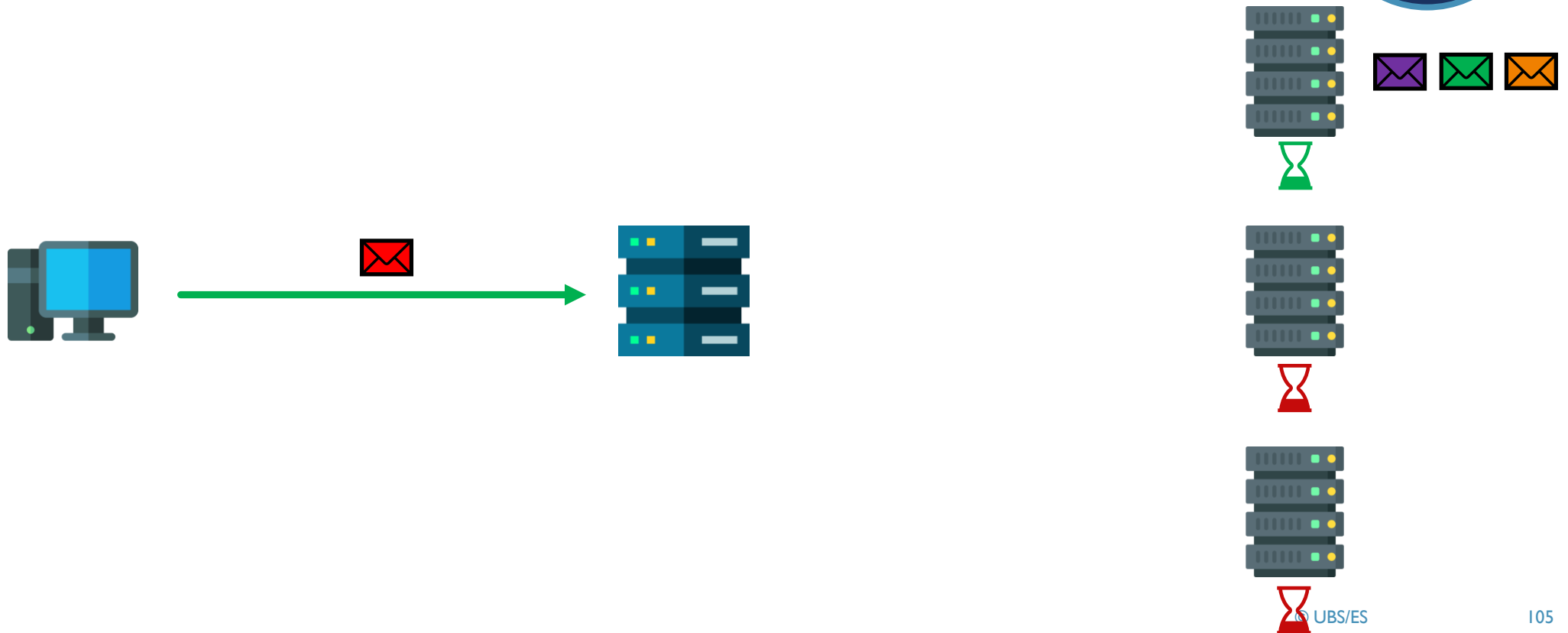


UBS/ES

TEMPS DE RÉPONSE



TEMPS DE RÉPONSE



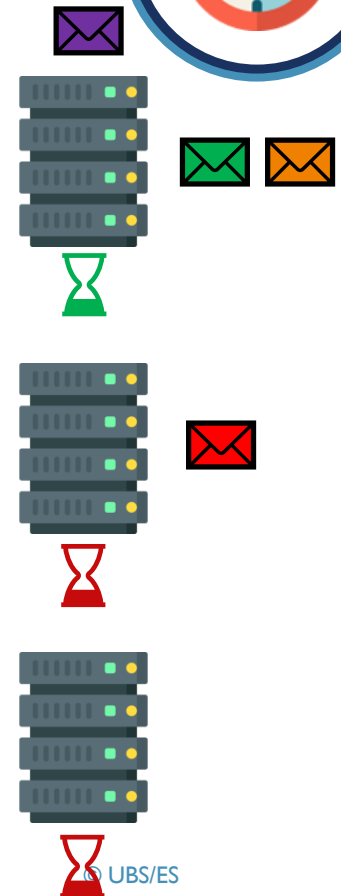
TEMPS DE RÉPONSE



UBS/ES



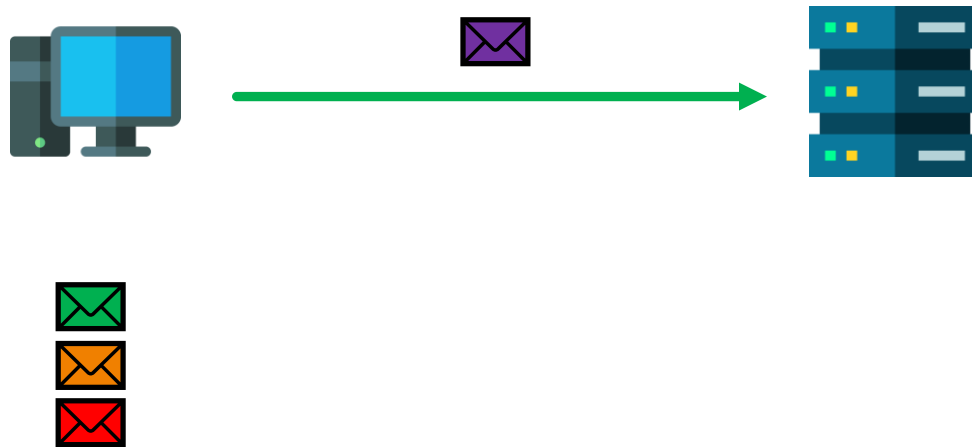
TEMPS DE RÉPONSE



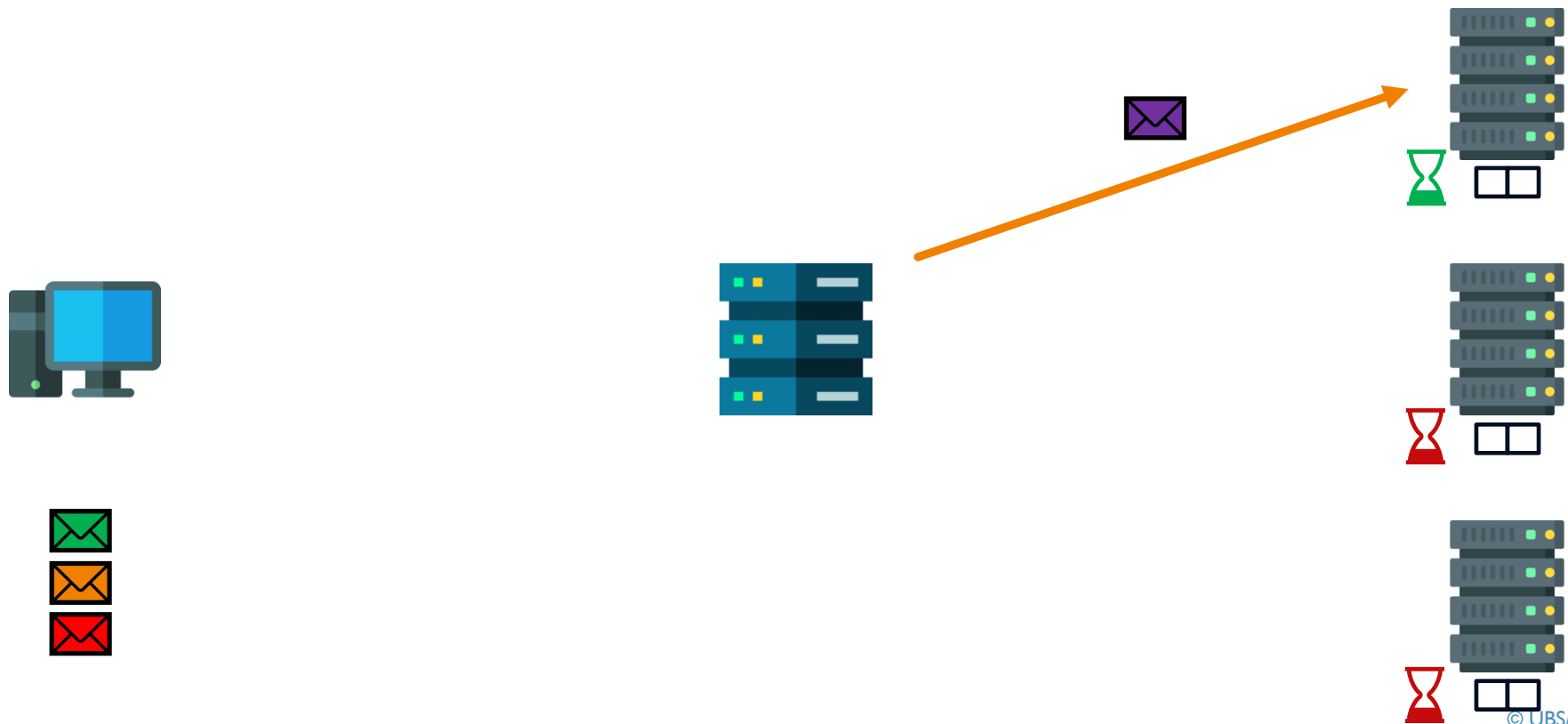
TEMPS DE RÉPONSE PONDÉRÉ



TEMPS DE RÉPONSE PONDÉRÉ



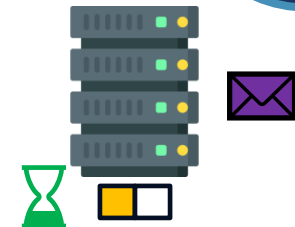
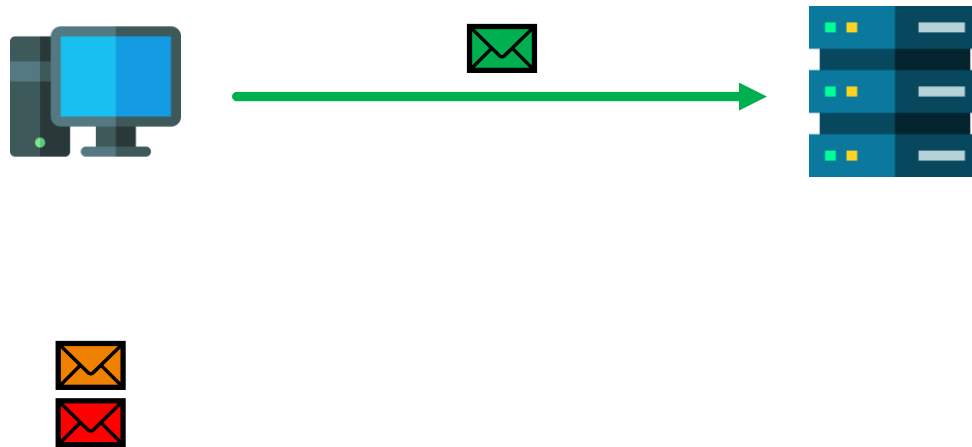
TEMPS DE RÉPONSE PONDÉRÉ



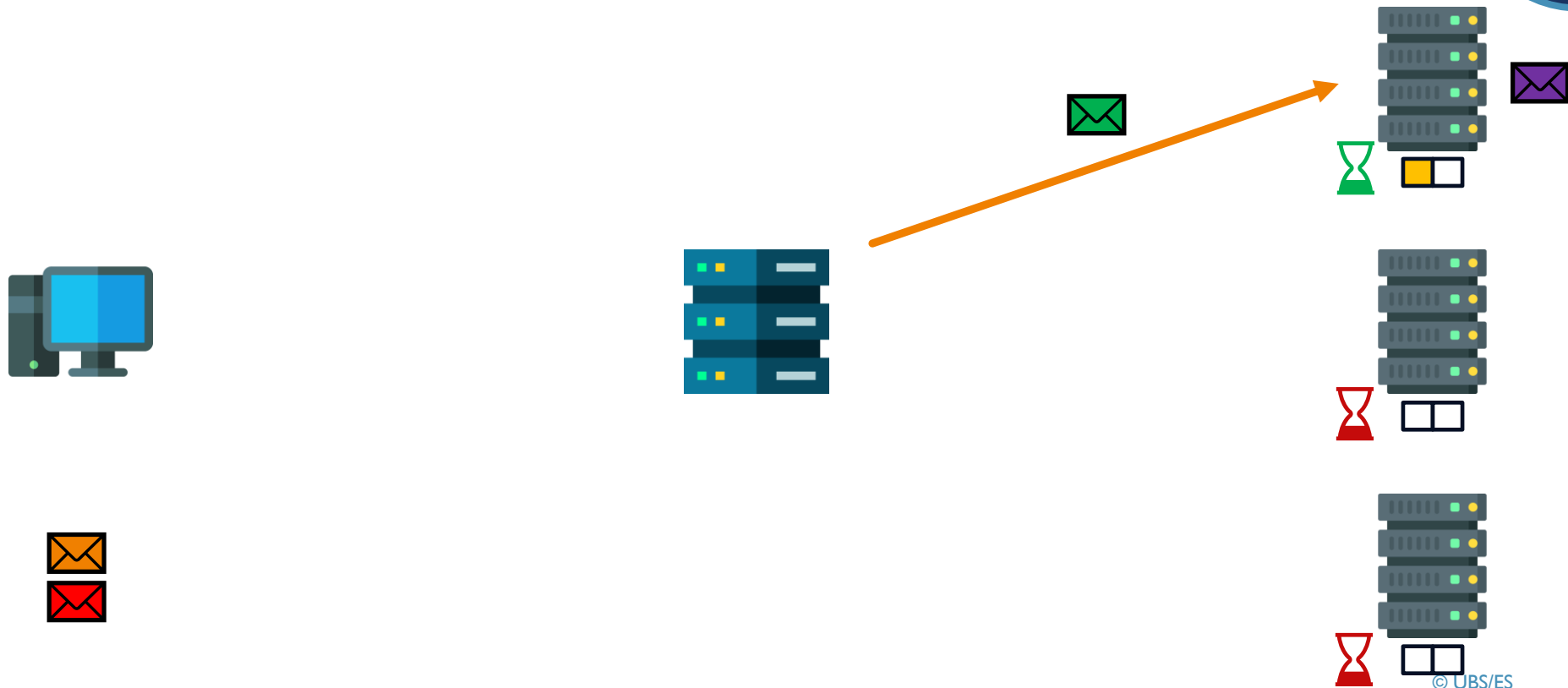
TEMPS DE RÉPONSE PONDÉRÉ



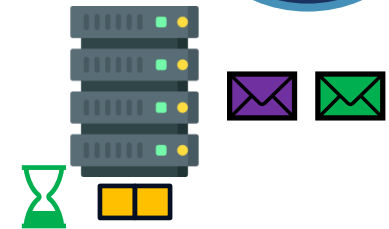
TEMPS DE RÉPONSE PONDÉRÉ



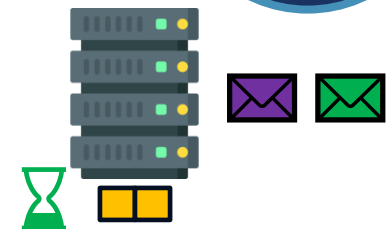
TEMPS DE RÉPONSE PONDÉRÉ



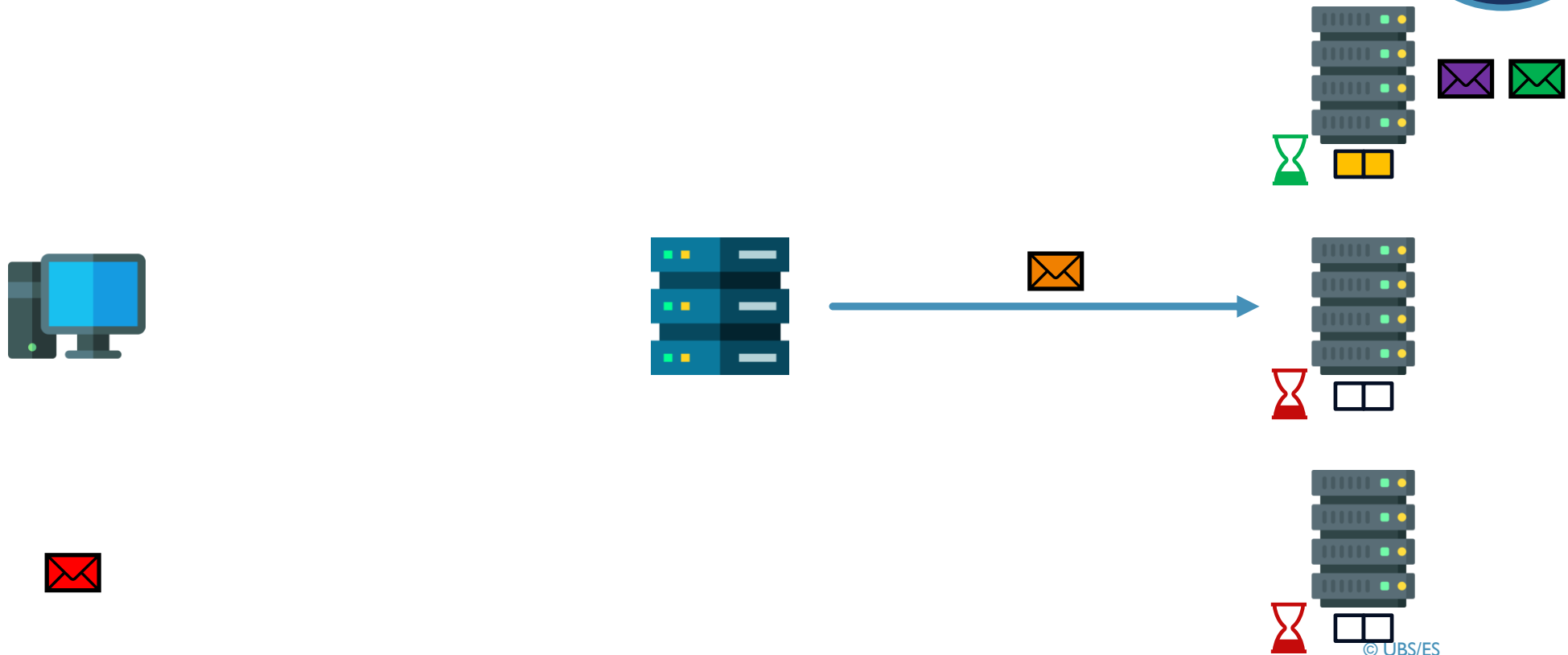
TEMPS DE RÉPONSE PONDÉRÉ



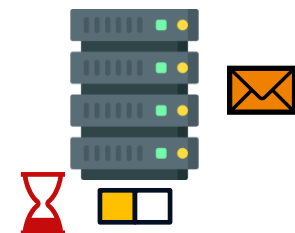
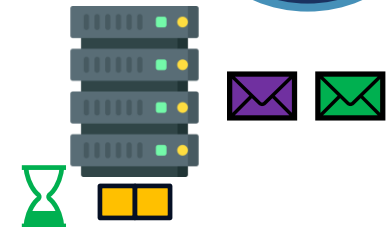
TEMPS DE RÉPONSE PONDÉRÉ



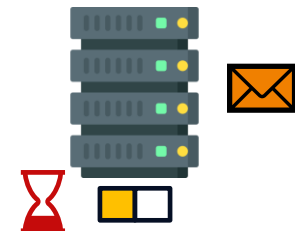
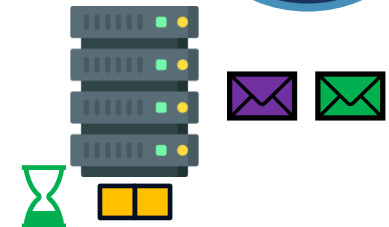
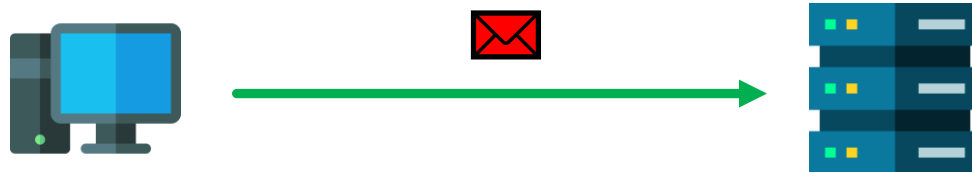
TEMPS DE RÉPONSE PONDÉRÉ



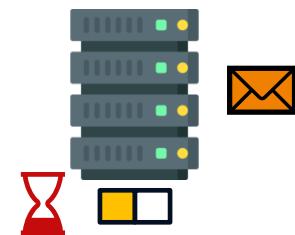
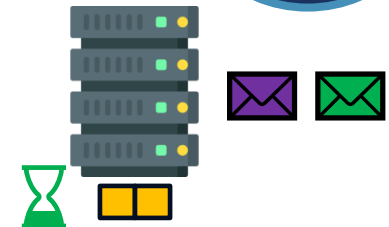
TEMPS DE RÉPONSE PONDÉRÉ



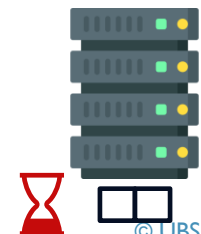
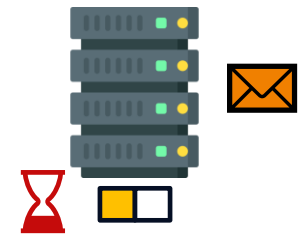
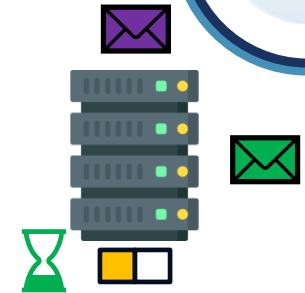
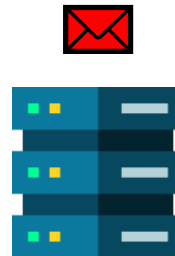
TEMPS DE RÉPONSE PONDÉRÉ



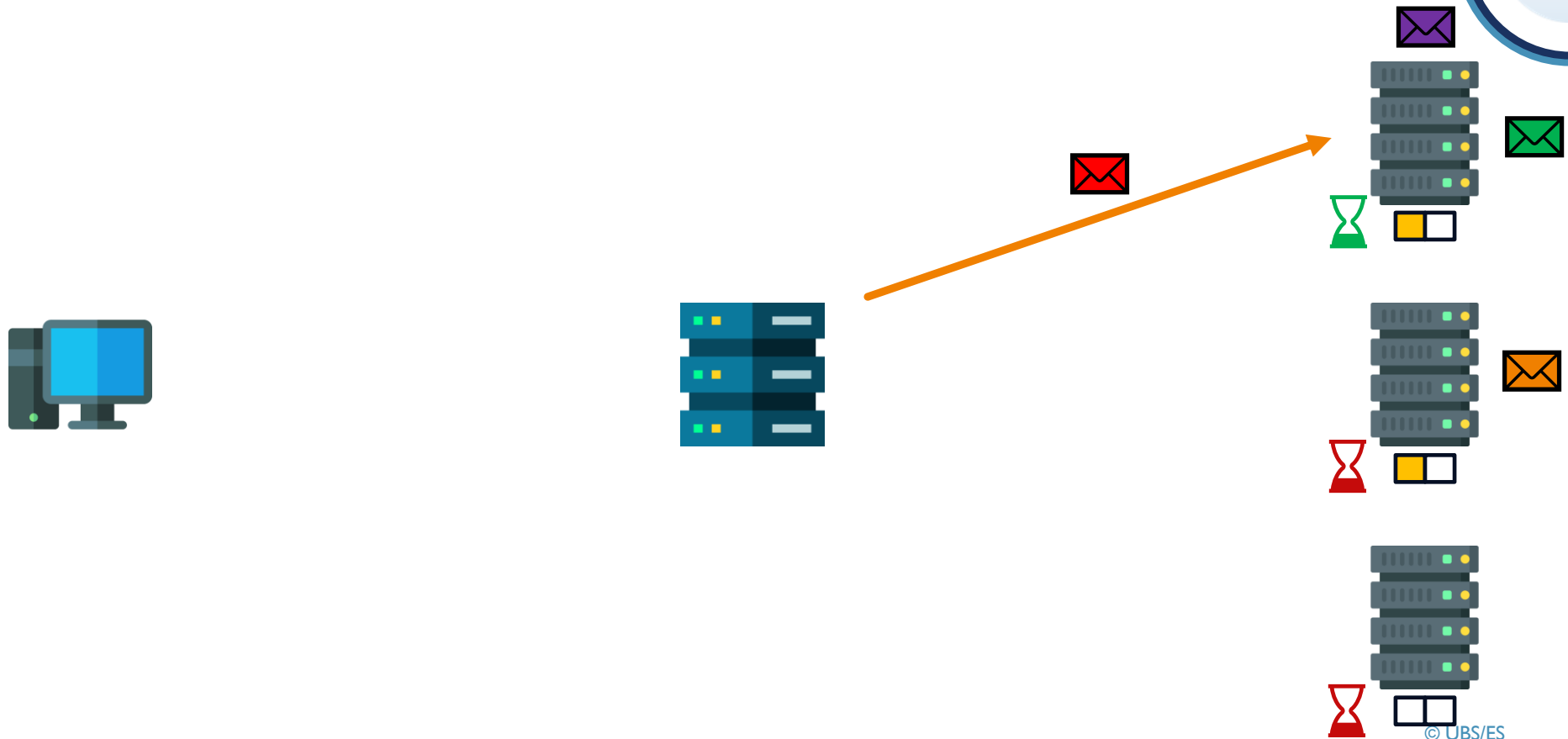
TEMPS DE RÉPONSE PONDÉRÉ



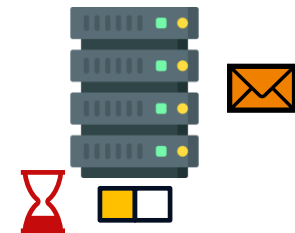
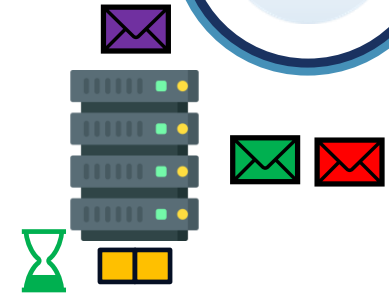
TEMPS DE RÉPONSE PONDÉRÉ



TEMPS DE RÉPONSE PONDÉRÉ



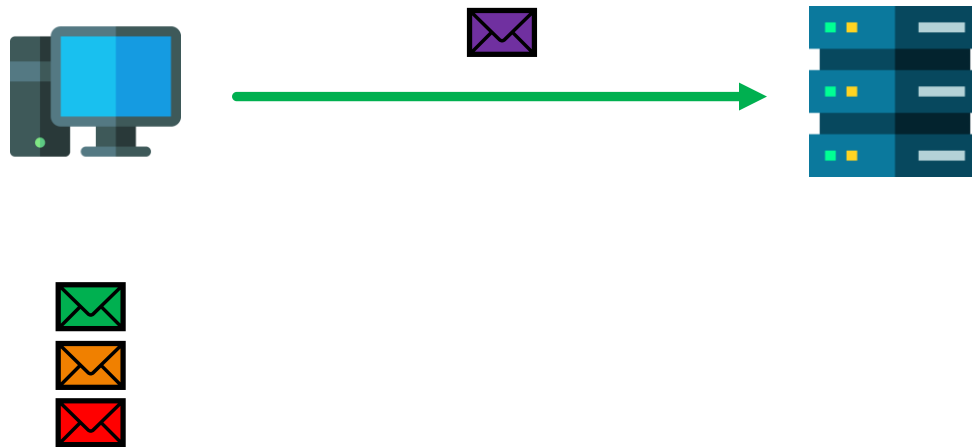
TEMPS DE RÉPONSE PONDÉRÉ



BASÉ RESSOURCES



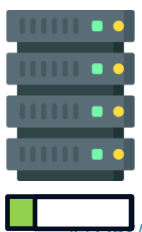
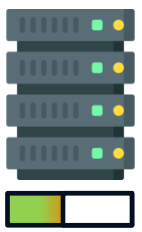
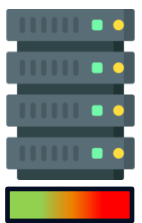
BASÉ RESSOURCES



BASÉ RESSOURCES



BASÉ RESSOURCES



BASÉ RESSOURCES



BASÉ RESSOURCES



BASÉ RESSOURCES



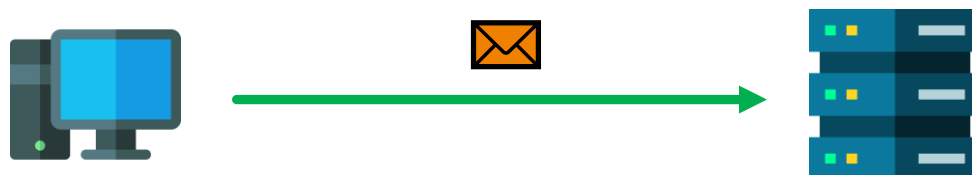
BASÉ RESSOURCES



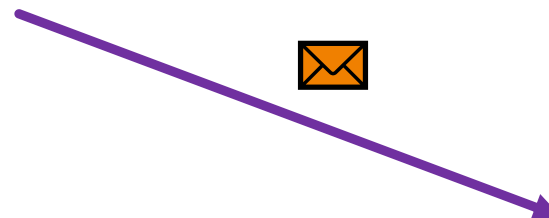
BASÉ RESSOURCES



BASÉ RESSOURCES



BASÉ RESSOURCES



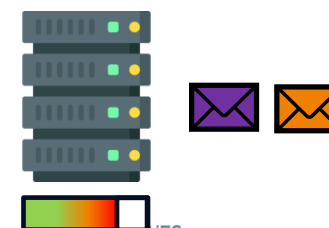
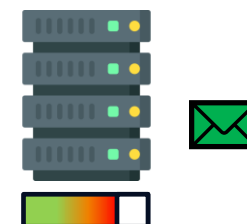
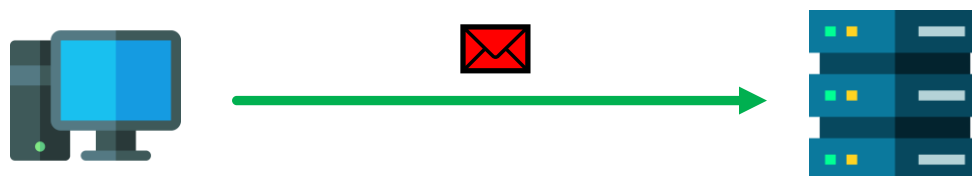
BASÉ RESSOURCES



BASÉ RESSOURCES



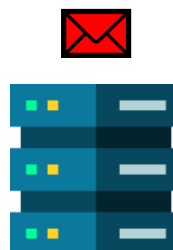
BASÉ RESSOURCES



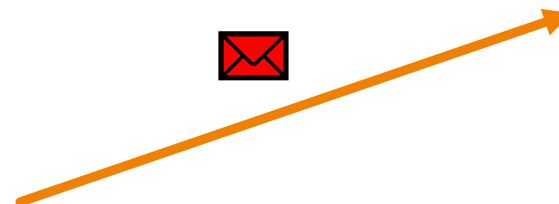
BASÉ RESSOURCES



BASÉ RESSOURCES



BASÉ RESSOURCES



BASÉ RESSOURCES

