



Département : Informatique

SYSTÈMES DISTRIBUÉS

Enseignante:

Wissal Soltani

Objectifs du cours

- ❑ Connaître les caractéristiques d'un système distribué (SD);
- ❑ Comprendre les concepts et les paradigmes fondamentaux d'un SD;
- ❑ Bien comprendre l'algorithmique distribuée ;
- ❑ Pouvoir raisonner dans un environnement distribué;
- ❑ Manipuler et implémenter des algorithmes distribués.

Plan du cours

- I. Introduction aux systèmes distribués
- II. Définitions et concepts fondamentaux
- III. Motivations pour utiliser les systèmes distribués
- IV. Exemples typiques de systèmes distribués
- V. Avantages et inconvénients des systèmes distribués
- VI. Propriétés clés des systèmes distribués
- VII. Applications réparties
- VIII. Conclusion générale



Contexte Général

Les progrès technologiques

❑ Les réseaux d'ordinateurs sont partout !!!

- ❑ Internet, réseaux de téléphones mobiles, réseaux locaux, etc.
- ❑ Les progrès technologiques des ordinateurs et le haut débit permettent aux différentes applications sur des ordinateurs distincts (et distants) de coopérer pour effectuer des tâches coordonnées;
- ❑ Avant les années 80 : les ordinateurs étaient encombrants et chers (les systèmes centralisés);
- ❑ A partir de la mi-80, deux nouveautés:
 - ❑ Microprocesseurs (moins chers et très rapide)
 - ❑ LANs and WANs

Technologie + besoins

Évolution Technologique

- Standardisation des réseaux de télécommunications;
- Performance des voies de télécommunication (débit et fiabilité)
- Rapport coût/performance des stations;
- Convergence informatique et téléphonie.

Évolution des besoins

- Structure des entreprises et des organisations : communication et partage (ex. Intranet);
- Accès universel à l'information (ex. Web);
- Informatique distribuée "grand public" (ex. vidéo interactive).

```
graph TD; A[Évolution Technologique] --> D[Applications Distribuées]; B[Évolution des besoins] --> D;
```

Applications Distribuées

Système distribué en opposition à système centralisé

- ❑ **Système Centralisé : tout est localisé sur la même machine et accessible par le programme**
 - ❑ Système logiciel s'exécutant sur une seule machine
 - ❑ Accédant localement aux ressources nécessaires (données, code, périphériques, mémoire ...)

- ❑ **Système distribué : une définition parmi d'autres**
 - ❑ Ensemble d'ordinateurs indépendants connectés en réseau et communiquant via ce réseau.

Définition d'un Système distribué (SD)

❑ Définition [Tanenbaum]:

« Un ensemble d'ordinateurs indépendants qui apparaît à un utilisateur comme un système unique et cohérent. »

- ❑ Les machines sont autonomes;
- ❑ Les utilisateurs ont l'impression d'utiliser un seul système.

Définition (pour ce cours)

❑ Définition

Un système distribué est un ensemble d'entités autonomes de calcul (ordinateurs, PDA, processeurs, processus, processus léger etc.) interconnectées et qui peuvent communiquer.

❑ Exemples

- ❑ Réseau physique de machines
- ❑ Un logiciel avec plusieurs processus sur une même machine

❑ Remarque

Généralement, on utilise la même terminologie

Systèmes répartis $\equiv$ Systèmes distribués (distributed systems)

Vision matérielle/ logicielle d'un SD

Vision matérielle : architecture matérielle

- ❑ Machine multi-processeurs avec mémoire partagée ;
- ❑ Cluster d'ordinateurs dédiés au calcul/traitement massif parallèle ;
- ❑ Ordinateurs standards connectés en réseau.

Vision logicielle

- ❑ Système logiciel composé de plusieurs entités s'exécutant indépendamment et en parallèle sur un ensemble d'ordinateurs connectés en réseau

Dans ce cours

- ❑ Conception logicielle des systèmes distribués
- ❑ Par défaut sur une architecture matérielle de type ordinateurs connectés en réseau

Pourquoi des systèmes répartis ?

- ❑ Besoin de communication et de partage d'information;
- ❑ Réalisation de systèmes à haute disponibilité;
- ❑ Partage de ressources (programmes, données, services);
- ❑ Réalisation de systèmes à grande capacité d'évolution;
- ❑ Aspects économiques (rapport prix/performance);
- ❑ Adaptation de la structure d'un système à celle des applications (géographique ou fonctionnelle);
- ❑ Besoin d'intégration (applications existantes).

Exemples

- ☐ WWW ;
- ☐ Contrôle du trafic aérien ;
- ☐ Système de courtage ;
- ☐ Banques ;
- ☐ Super calcul distribué ;
- ☐ Système de fichier distribué ;
- ☐ DNS (Domain Name System) ;
- ☐ Systèmes Pair-à-pair (P2P).

Exemple 1 : Serveur de fichiers

❑ **Serveur de fichiers**

- ❑ Accès aux fichiers de l'utilisateur quelque soit la machine utilisée
- ❑ Machines du département informatique
 - ❑ Clients : Anis

❑ **Un serveur de fichiers**

- ❑ Sur toutes les machines : /home/anis est le « home directory » de l'utilisateur anis
- ❑ Physiquement : fichiers se trouvent uniquement sur le serveur
- ❑ Virtuellement : accès à ces fichiers à partir de n'importe quelle machine cliente en faisant « croire » que ces fichiers sont stockés localement

❑ **Serveur de fichier : Intérêts**

- ❑ Accès aux fichiers à partir de n'importe quelle machine
- ❑ Système de sauvegarde associé à ce serveur
- ❑ Transparent pour l'utilisateur

❑ **Inconvénients**

- ❑ Si le réseau ou le serveur plante : plus d'accès aux fichiers

Exemple 2 : Web

- ❑ Un serveur web auquel se connecte un nombre quelconque de navigateurs web (clients)
- ❑ Accès à distance à l'information
 - ❑ **Accès simple**
 - ❑ Serveur renvoie une page HTML statique qu'il stocke localement
 - ❑ **Traitement plus complexe**
 - ❑ Serveur interroge une base de données pour générer dynamiquement le contenu de la page
 - ❑ **Transparent pour l'utilisateur** : les informations s'affichent dans son navigateur quelque soit la façon dont le serveur les génère.

Exemple 3 : Calculs scientifiques

☐ Plusieurs architectures matérielles généralement utilisées

- ☐ Ensemble de machines identiques reliées entre elles par un réseau dédié et très rapide (cluster)
- ☐ Ensemble de machines hétérogènes connectées dans un réseau local ou bien encore par Internet (grille)

☐ Principe général

- ☐ Un (ou des) serveur distribue des calculs aux machines clients
- ☐ Un client exécute son calcul puis renvoie le résultat au serveur

☐ Avantage

- ☐ Utilisation d'un maximum de ressources de calcul

☐ Inconvénient

- ☐ Si réseau ou serveur plante, alors arrêt du système

Intérêts des systèmes distribués

❑ Utiliser et partager des ressources distantes

- ❑ Un même service peut être utilisé par plusieurs acteurs, situés à des endroits différents;
- ❑ Système de fichiers : utiliser ses fichiers à partir de n'importe quelle machine;
- ❑ Imprimante : partagée entre toutes les machines.

❑ Optimiser l'utilisation des ressources disponibles

- ❑ Calculs scientifiques distribués sur un ensemble de machines.

❑ Système plus robuste

- ❑ Duplication pour fiabilité : deux serveurs de fichiers dupliqués, avec sauvegarde;
- ❑ Plusieurs éléments identiques pour résister à la montée en charge ...

Systemes Distribués : Inconvénients

☐ Si problème au niveau du réseau

- ☐ Le système marche mal ou plus du tout

☐ Bien souvent, un élément est central au fonctionnement du système : serveur

- ☐ Si serveur plante : plus rien ne fonctionne
- ☐ Goulot potentiel d'étranglement si débit d'information très important

☐ Sans élément central

- ☐ Gestion du système totalement décentralisée et distribuée
- ☐ Nécessite la mise en place d'algorithmes +/- complexes

Propriétés des systèmes distribués

- 1) Hétérogénéité
- 2) Fiabilité des systèmes distribués
- 3) Sécurité des systèmes distribués
- 4) Transparence

Propriétés des systèmes distribués (1/10)

- ❑ Un **système réparti** (ou distribué de «distributed system») est un système comprenant un ensemble de **processus** et un **système de communication**
 - ❑ Ensemble composé **d'éléments reliés** par un **système de communication**
 - ❑ les éléments ont des fonctions de :
 - ❑ **traitement** (processeurs);
 - ❑ **stockage** (mémoire);
 - ❑ **relation avec le monde extérieur** (capteurs, actionneurs).
 - ❑ Les différents éléments du système ne fonctionnent pas indépendamment mais collaborent à une ou plusieurs tâches communes.
- ⇒ **Conséquence** : une partie au moins de l'état global du système est partagée entre plusieurs éléments (sinon, on aurait un fonctionnement indépendant)

Propriétés des systèmes distribués (2/10)

☐ **Système distribué = éclaté**

- ☐ Connaissance des éléments formant le système : besoin d'identification et de localisation
- ☐ Gestion du déploiement et de la présence d'éléments essentiels

☐ **Communication à distance est centrale**

- ☐ Techniques et protocoles de communication
- ☐ Contraintes du réseau : fiabilité (perte de données) et temps de propagation (dépendant du type de réseau et de sa charge)

☐ **Naturellement concurrent et parallèle**

- ☐ Chaque élément sur chaque machine est autonome
- ☐ Besoin de synchronisation, coordination entre éléments distants et pour l'accès aux ressources (exclusion mutuelle ...)

Propriétés des systèmes distribués (3/10)

Hétérogénéité

- ❑ Des machines utilisées (puissance, architecture matérielle...);
- ❑ Des systèmes d'exploitation tournant sur ces machines ;
- ❑ Des langages de programmation des éléments logiciels formant le système ;
- ❑ Des réseaux utilisés : impact sur performances, débit, disponibilité ... ;
- ❑ Réseau local rapide ;
- ❑ Internet ;
- ❑ Réseaux sans fil.

Propriétés des systèmes distribués (4/10)

Fiabilité

Nombreux points de pannes ou de problèmes potentiels:

❑ Réseau

- ❑ Une partie du réseau peut-être inaccessible
- ❑ Les temps de communication peuvent varier considérablement selon la charge du réseau
- ❑ Le réseau peut perdre des données transmises

❑ Machine

- ❑ Une ou plusieurs machines peut planter, engendrant une paralysie partielle ou totale du système

❑ Peut augmenter la fiabilité par redondance, duplication de certains éléments

- ❑ Mais rend plus complexe la gestion du système

Propriétés des systèmes distribués (5/10)

Fiabilité

❑ Tolérance aux fautes

[Capacité d'un système à gérer et résister à un ensemble de problèmes]

- ❑ Le système doit pouvoir fonctionner (au moins de façon dégradée) même en cas de défaillance de certains de ses éléments ;
- ❑ Le système doit pouvoir résister à des perturbations du système de communication (perte de messages, déconnexion temporaire, performances dégradées) ;
- ❑ Le système doit pouvoir facilement s'adapter pour réagir à des changements d'environnement ou de conditions d'utilisation.

Propriétés des systèmes distribués (6/10)

Sécurité

- ❑ Nature d'un système distribué fait qu'il est beaucoup plus sujet à des attaques
 - ❑ Communications à travers le réseau peuvent être interceptées
 - ❑ On ne connaît pas toujours bien un élément distant avec qui on communique
- ❑ Le système doit pouvoir résister à des attaques contre sa sécurité (violation de la confidentialité, de l'intégrité, usage indu de ressources, déni de service)

Solutions

- ❑ Connexion sécurisée par authentification avec les éléments distants
- ❑ Cryptage des messages circulant sur le réseau

Propriétés des systèmes distribués (7/10)

Transparence

- ❑ Fait pour une fonctionnalité, un élément d'être invisible ou caché à l'utilisateur ou un autre élément formant le système distribué
 - ❑ Devrait plutôt parler d'opacité dans certains cas ...
- ❑ **But**: cacher l'architecture, le fonctionnement de l'application ou du système distribué pour apparaître à l'utilisateur comme une application unique cohérente
- ❑ L'ISO définit plusieurs transparences (norme RM-ODP)
 - ❑ **RM-ODP** = *Reference Model of Open Distributed Processing*
 - ❑ Accès, localisation, concurrence, réplication, mobilité, panne, performance, échelle

Propriétés des systèmes distribués (8/10)

Transparence

☐ Transparence d'accès

- ☐ Accès à des ressources distantes aussi facilement que localement
- ☐ Accès aux données indépendamment de leur format de représentation

☐ Transparence de localisation

- ☐ Accès aux éléments/ressources indépendamment de leur localisation

☐ Transparence de concurrence

- ☐ Exécution possible de plusieurs processus en parallèle avec utilisation de ressources partagées

☐ Transparence de réplication

- ☐ Possibilité de dupliquer certains éléments/ressources pour augmenter la fiabilité

Propriétés des systèmes distribués (9/10)

Transparence

❑ Transparence de mobilité

- ❑ Possibilité de déplacer des éléments/ressources

❑ Transparence de panne

- ❑ Doit supporter qu'un ou plusieurs éléments tombe en panne

❑ Transparence de performance

- ❑ Possibilité de reconfigurer le système pour en augmenter les performances

❑ Transparence d'échelle

- ❑ Doit supporter l'augmentation de la taille du système (nombre d'éléments, de ressources ...)
- ❑ Le système doit préserver ses performances lorsque sa taille croît (nombre d'éléments, nombre d'utilisateurs, étendue géographique)

Propriétés des systèmes distribués (10/10)

Quelques difficultés

Propriété	Difficulté engendrée
 Asynchronisme du système de communication (pas de borne supérieure stricte pour le temps de transmission d'un message)	⇒ Difficulté de détecter les défaillances
 Dynamisme (la composition du système change en permanence)	⇒ difficulté de définir un état global du système ⇒ Difficulté d'administrer le système
 Grande taille (nombre de composants, d'utilisateurs, dispersion géographique)	⇒ la capacité de croissance (scalability) est une propriété importante, mais difficile à réaliser

Applications réparties

- 1) Distinction entre un système et une application répartie
- 2) Exemples d'applications réparties
- 3) Programmation classique versus programmation répartie

Applications réparties (1/7)

❑ Distinction entre “système” et “application”

- ❑ **Système** : gestion des ressources communes et de l'infrastructure, lié de manière étroite au matériel sous-jacent
 - ❑ **Système d'exploitation** : gestion de chaque élément
 - ❑ **Système de communication** : échange d'information entre les éléments
 - ❑ **Caractéristiques communes** : cachent la complexité du matériel et des communications, fournissent des services communs de plus haut niveau d'abstraction
- ❑ **Application** : réponse à un problème spécifique, fourniture de services à ses utilisateurs (qui peuvent être d'autres applications)
 - ❑ Utilise les services généraux fournis par le système
 - ❑ La distinction n'est pas toujours évidente, car certaines applications peuvent directement travailler à bas niveau (au contact du matériel). Exemple : systèmes embarqués, réseaux de capteurs

Applications réparties (2/7)

Système réparti

Un ensemble de systèmes calculatoires autonomes (ex.: ordinateur, serveur, terminal, etc.) sans mémoire physique commune qui communiquent à travers un réseau quelconque



Application répartie

Un ensemble de processus qui tournent sur un système réparti afin de fournir ou utiliser un service déterminé

Applications réparties (3/7)

❑ **Application répartie** = traitements **coopérants** sur des **données réparties**

❑ **Traitements** :

- ❑ Description : programme
- ❑ Exécution : flot d'exécution (processus)

❑ **coopération** = communication + synchronisation

- ❑ modèle d'exécution
- ❑ interface de programmation (et/ou langage)
- ❑ modèle de programmation
- ❑ outils de développement
- ❑ environnement d'exécution : services systèmes (pour différents types d'infrastructures)

Applications réparties (4/7)

Exemples d'applications réparties

- ❑ Navigation web, transfert de fichiers,
- ❑ Guichets de banque (GAB (Guichet Automatique de Banque), DAB (Distributeur Automatique de Banque),
- ❑ Commerce électronique,
- ❑ Jeux en réseaux,
- ❑ Télévision interactive

Applications réparties (5/7)

❑ Programmation classique versus programmation répartie

- ❑ La plupart des applications réparties sont de type client/serveur: le client demande des services à un serveur
- ❑ En programmation classique, lorsque un programme a besoin d'un service, il appelle localement une fonction/procédure/méthode d'une librairie, d'un objet, etc.
- ❑ En programmation répartie, l'appel de fonction/procédure/méthode peut se faire à distance
 - ⇒ Proposer des méthodes/concepts/outils permettant de simplifier le développement d'application réseau client/serveur, en essayant de s'abstraire de l'aspect "distant"

Applications réparties (6/7)

❑ Programmation classique versus programmation répartie

Programmation classique

- ❑ L'utilisateur du service et le fournisseur de service se trouvent sur la même machine :
 - ❑ Même OS
 - ❑ Même espace mémoire
 - ❑ Même capacité de calcul CPU
 - ❑ Pas de problème de transport
 - ❑ Disponibilité du service assuré (tant que l'on a accès à la librairie)
- ❑ **Un même langage de programmation**

Programmation répartie

- ❑ L'utilisateur et le fournisseur de service ne se trouvent pas sur la même machine: deux machines différentes (sans compter celles traversées)
 - ❑ OS différents
 - ❑ Espace mémoire non unitaire : "passer un pointeur comme argument" ?
 - ❑ Problème de transport : pare-feu (firewall), réseau, etc.
 - ❑ Retrouver le service ? où se trouve-t-il ? qui le propose ?
- ❑ **Deux langages différents**
 - ❑ Paradigmes de programmation différents : qu'est ce qu'un objet pour un langage procédural ? comment gérer les erreurs ?

Conclusion générale

- Les systèmes distribués sont aujourd'hui au cœur de l'infrastructure technologique moderne. Ils permettent de partager les ressources, d'améliorer la disponibilité des services et de répondre à des besoins de performance et d'évolutivité.
- Cependant, ils introduisent aussi des défis majeurs liés à la gestion de la concurrence, à la sécurité, à la tolérance aux pannes et à la complexité de la synchronisation.
- La bonne compréhension de leurs concepts fondamentaux, de leurs architectures, ainsi que des modèles de programmation associés, est indispensable pour concevoir, déployer et maintenir des systèmes distribués efficaces et sûrs.
- Le futur des technologies de l'information repose largement sur la maîtrise de ces environnements complexes mais extrêmement puissants.



QCM

- **1. Quel est l'objectif principal d'un système distribué ?**

A) Exécuter toutes les tâches sur une seule machine

B) Permettre à plusieurs ordinateurs de coopérer comme un système unique

C) Supprimer toute interaction entre les machines du réseau

D) Limiter l'accès aux ressources partagées

- **2. Quelle est une caractéristique clé d'un système distribué ?**

A) L'indépendance totale des machines

B) L'illusion d'un système unique pour l'utilisateur

C) L'absence de réseau de communication

D) La présence d'un seul serveur central

- **3. Quel est l'avantage principal d'un système distribué par rapport à un système centralisé ?**

A) Moins de consommation énergétique

B) Fiabilité accrue grâce à la redondance

C) Moins de complexité logicielle

D) Aucune interconnexion entre les machines

- **4. Quel élément n'est pas une application d'un système distribué ?**

A) Le Web

B) Un ordinateur personnel isolé

C) Un système de fichiers distribué

D) Le DNS

- **5. Quels sont les principaux défis d'un système distribué ?**

A) Facilité de gestion et de maintenance

B) Coordination et synchronisation des processus

C) Absence totale de pannes

D) Exécution locale uniquement

- **6. Quel est l'un des inconvénients majeurs d'un système distribué ?**

A) L'accès limité aux ressources

B) La dépendance au réseau

C) L'absence de fiabilité

D) Le manque d'applications

- **7. Quelle propriété des systèmes distribués garantit l'accès aux ressources distantes comme si elles étaient locales ?**

A) Transparence d'échelle
B) **Transparence d'accès**
C) Sécurité renforcée
D) Fragmentation des données

- **8. Pourquoi la tolérance aux pannes est-elle importante dans un système distribué ?**

A) Pour permettre une exécution plus lente
B) **Pour assurer la continuité du service malgré les défaillances**
C) Pour éviter toute communication entre les nœuds
D) Pour réduire la charge du réseau

- **9. Quelle architecture matérielle est typiquement utilisée dans les systèmes distribués ?**

A) **Ordinateurs multi-processeurs avec mémoire partagée**
B) Un unique serveur avec plusieurs clients connectés
C) Une architecture où chaque machine fonctionne de manière autonome sans réseau
D) Un seul ordinateur avec plusieurs utilisateurs

- **10. Quel est l'impact principal de la transparence de localisation ?**
 - A) L'utilisateur doit connaître l'emplacement exact des ressources
 - B) Les ressources peuvent être accessibles indépendamment de leur localisation physique
 - C) Le système devient plus difficile à gérer
 - D) La communication entre machines devient impossible
- **11. Quelles sont les principales causes de pannes dans un système distribué ?**
 - A) Un mauvais choix d'algorithmes
 - B) La perte de messages, la défaillance d'une machine ou la surcharge du réseau
 - C) L'absence de logiciels de gestion
 - D) Un trop grand nombre d'utilisateurs
- **12. Quelle est la différence entre un système centralisé et un système distribué ?**
 - A) Un système centralisé repose sur un seul ordinateur, alors qu'un système distribué utilise plusieurs machines connectées
 - B) Un système distribué n'a pas besoin de communication entre machines
 - C) Un système centralisé est toujours plus fiable qu'un système distribué
 - D) Un système distribué n'a pas de réseau

- **13. Quelle est l'utilité d'un système de fichiers distribué ?**

- A) Permettre aux utilisateurs d'accéder aux fichiers uniquement depuis leur machine locale
- B) Centraliser toutes les ressources sur un serveur unique
- C) Assurer l'accès aux fichiers depuis n'importe quelle machine connectée au système
- D) Réduire la vitesse d'accès aux fichiers

- **14. Quel est un exemple typique d'un système distribué ?**

- A) Un ordinateur personnel
- B) Une application mobile hors ligne
- C) Un système de contrôle du trafic aérien
- D) Une calculatrice scientifique

- **15. Qu'est-ce que la transparence de réplication ?**

- A) Empêcher la duplication des données
- B) Augmenter la fiabilité en dupliquant certaines ressources
- C) Rendre les données inaccessibles
- D) Éviter toute communication entre les nœuds

- **16. Pourquoi la sécurité est-elle un enjeu majeur dans les systèmes distribués ?**

A) Car les communications entre machines peuvent être interceptées

B) Car les systèmes distribués ne sont jamais attaqués

C) Car chaque machine est totalement isolée

D) Car aucun utilisateur ne peut accéder aux ressources

- **17. Quel est l'objectif principal d'une application répartie ?**

A) Réduire le coût du matériel

B) Permettre à plusieurs processus de coopérer sur des données réparties

C) Remplacer les systèmes d'exploitation classiques

D) Fonctionner sans interaction entre les composants

- **18. Quelle est une des difficultés de la programmation répartie par rapport à la programmation classique ?**

A) Nécessité d'un seul langage de programmation

B) Gestion de la communication et synchronisation entre machines

C) Simplicité accrue par rapport à un programme local

D) Exécution exclusivement en local

- **19. Pourquoi la transparence d'échelle est-elle importante ?**

A) Pour garantir que le système fonctionne indépendamment du nombre d'éléments connectés

B) Pour empêcher la croissance du système

C) Pour s'assurer que le système reste limité à un petit nombre d'utilisateurs

D) Pour éviter l'optimisation des performances

- **20. Qu'est-ce qu'un système pair-à-pair (P2P) ?**

A) Un système centralisé avec un serveur unique

B) Un réseau où tous les nœuds sont clients et serveurs en même temps

C) Un système où un seul ordinateur contrôle l'ensemble du réseau

D) Un système distribué sans aucune communication entre machines



**Merci pour votre
attention**