

Historique

Évolution des systèmes informatiques (à travers les âges)



Jean-Michel Adam
Université Grenoble Alpes – UFR SHS

Historique

L'histoire des systèmes d'exploitation permet de dégager des concepts de base que l'on retrouve dans les systèmes actuels :

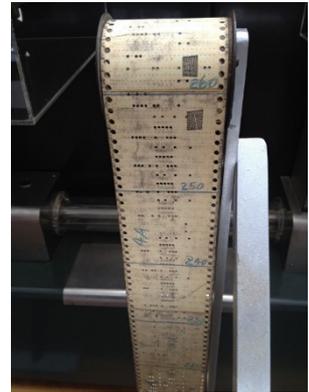
- le traitement par lots
- la multiprogrammation
- le temps partagé

Historique et typologie des OS

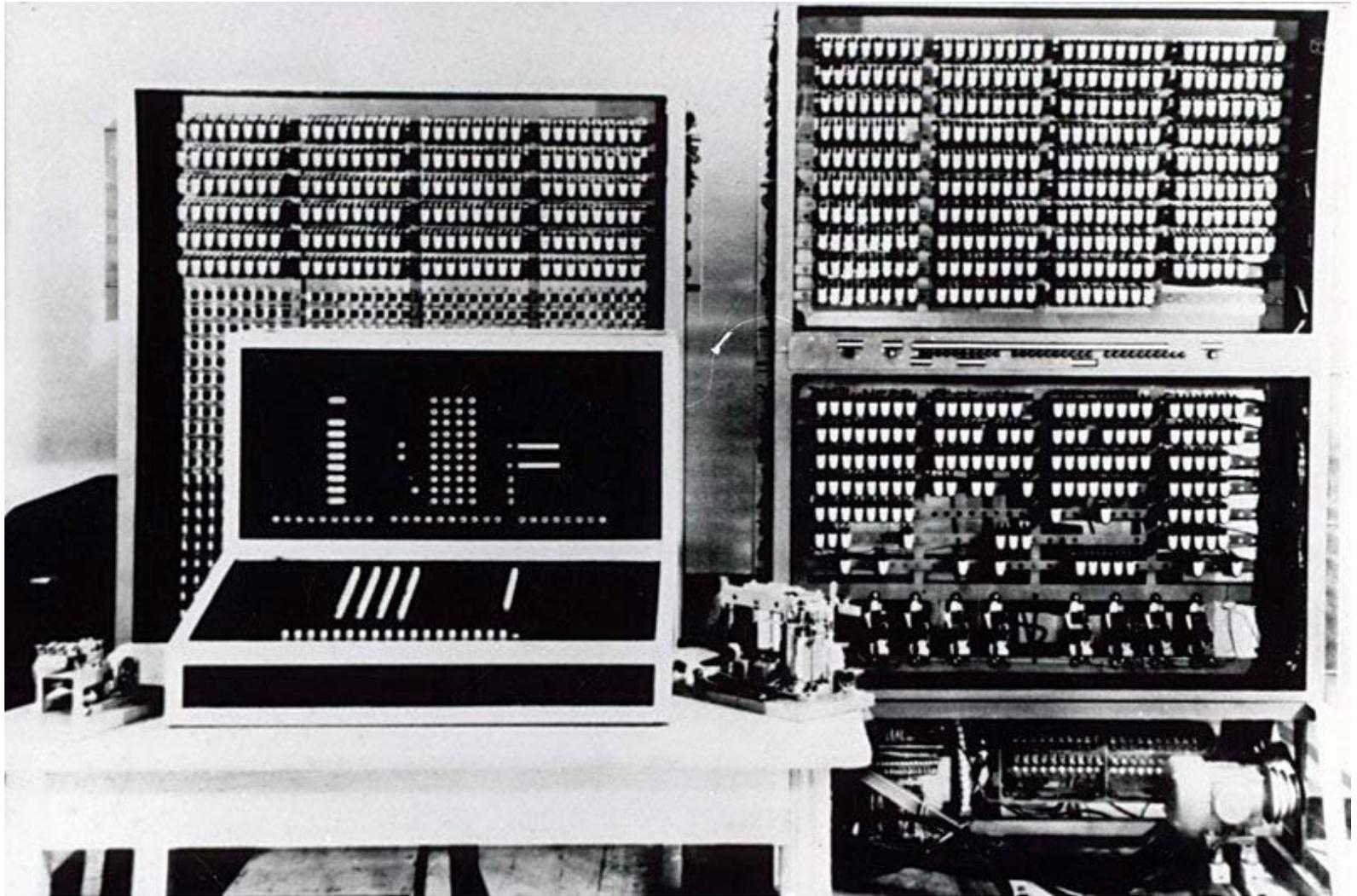
- Il existe 5 générations d'OS
 - Systèmes de traitement par lots
 - Enchaînement d'une suite de programmes sans intervention de l'utilisateur
 - Systèmes multiprogrammés
 - Exécution simultanées de plusieurs programmes pour optimiser l'utilisation du processeur
 - Systèmes en temps partagé
 - Multiprogrammation + temps partagés pour pouvoir gérer plusieurs utilisateurs en même temps sur la machine
 - Systèmes temps réel
 - Garantie d'exécution dans un temps imparti des programmes (systèmes des trains, avions, etc)
 - Systèmes distribués
 - Permet de gérer plusieurs machines sous forme d'une seule machine virtuelle

Premier ordinateur : le Z3

- Le Z3 était un calculateur à relais électromécaniques conçu par l'ingénieur allemand Konrad Zuse.
- Ce calculateur était la première machine programmable pleinement automatique ce qui en ferait le premier ordinateur du monde.
- Il était composé de 2000 relais électromécaniques, fonctionnait à une fréquence d'horloge de 5 à 10 Hz et exploitait des mots d'une longueur de 22 bits.
- Le code et les données étaient stockés sur des rubans perforés en celluloïd.
- Le Z3 fut achevé à Berlin en 1941 et détruit par un bombardement en 1943

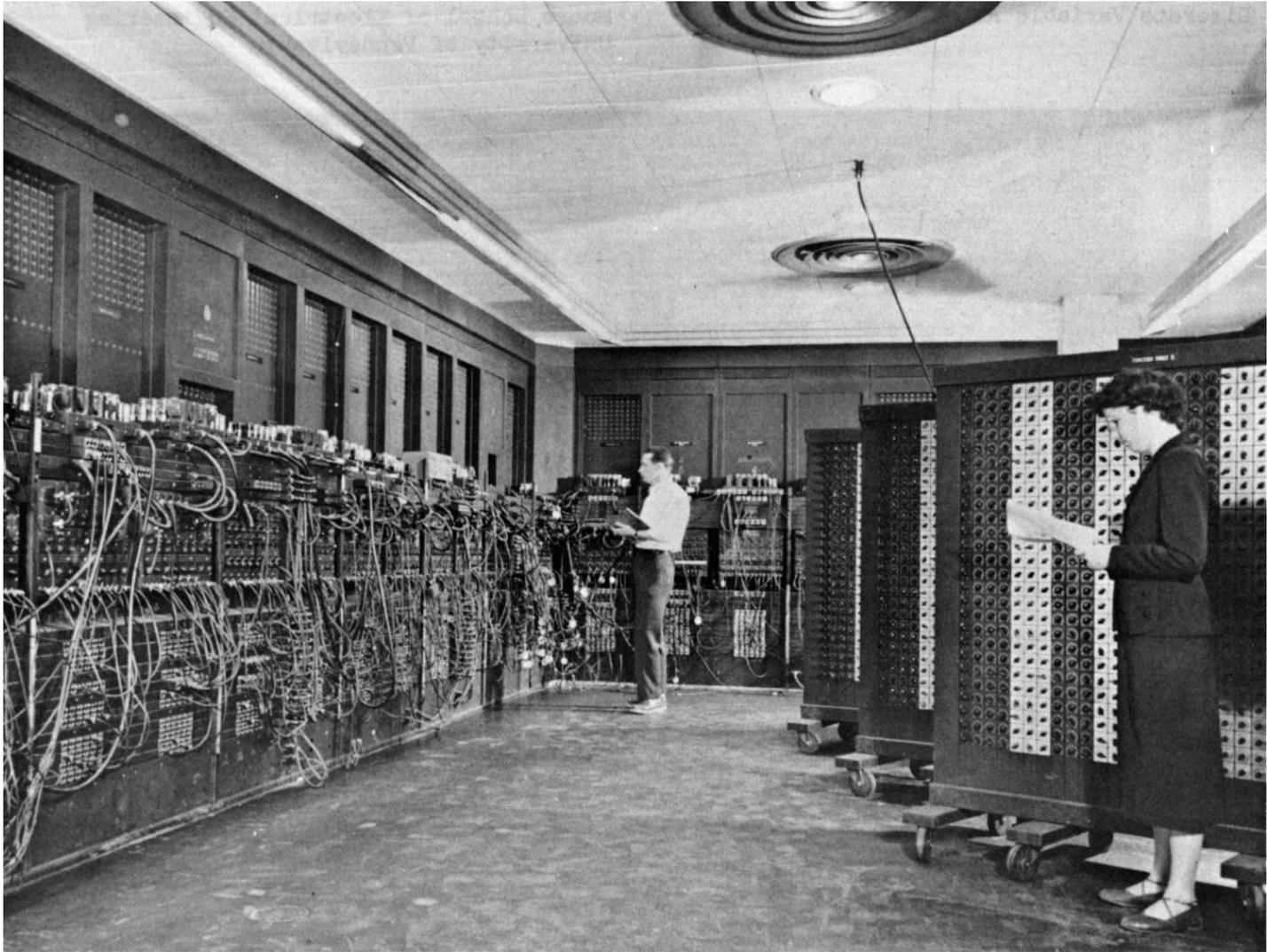


Premier ordinateur : le Z3



Source: computerhistory.org

Le premier aux US : l'ENIAC

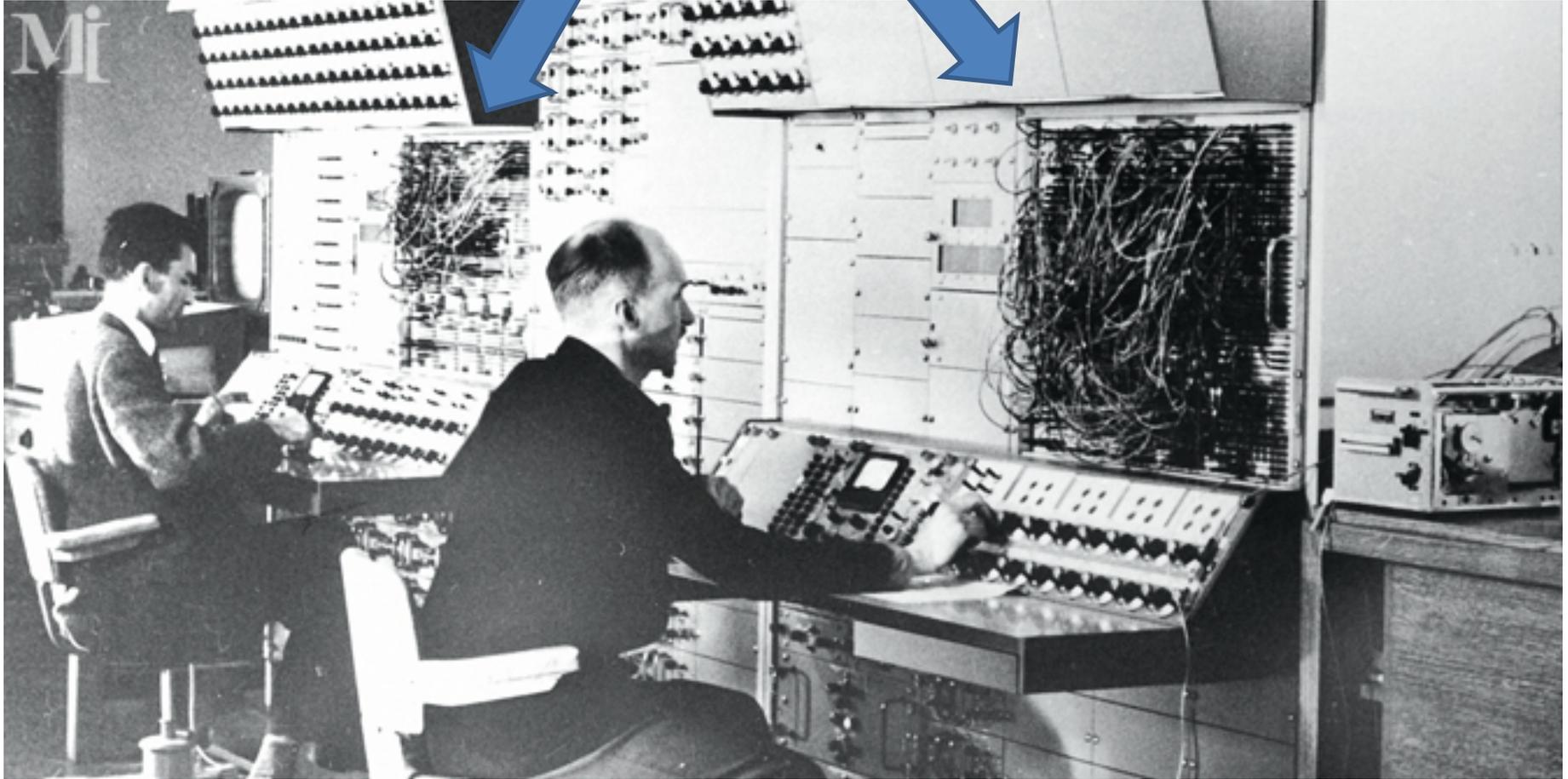


L'ENIAC (photo prise entre 1947 et 1955 – US Army).
Electronic Numerical Integrator Analyser and Computer

Le premier aux US : l'ENIAC

- Sa capacité est de 20 nombres à 10 chiffres signés permettant chacun de réaliser 5 000 additions simples chaque seconde. Il ne peut en revanche gérer que 357 multiplications ou 38 divisions par seconde.
- Physiquement l'ENIAC est une grosse machine, il contient 17468 tubes à vide, 7200 diodes à cristal, 1500 relais, 70000 résistances, 10000 condensateurs et environ 5 millions de soudures faites à la main.
- Son poids est de 30 tonnes pour des dimensions de $2,4 \times 0,9 \times 30,5$ m occupant une surface de 167 m²
- Sa consommation est de 150 kW.

Programmes

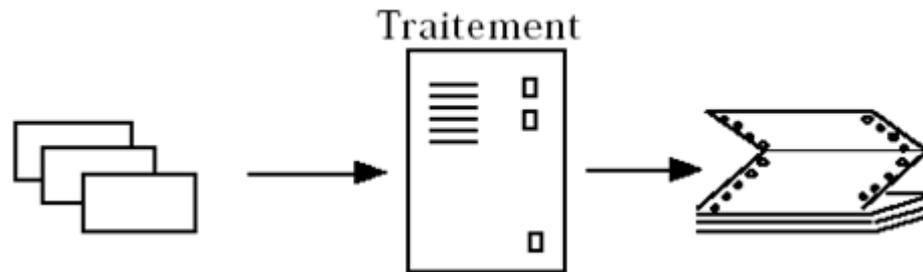


Caractéristiques des premiers ordinateurs

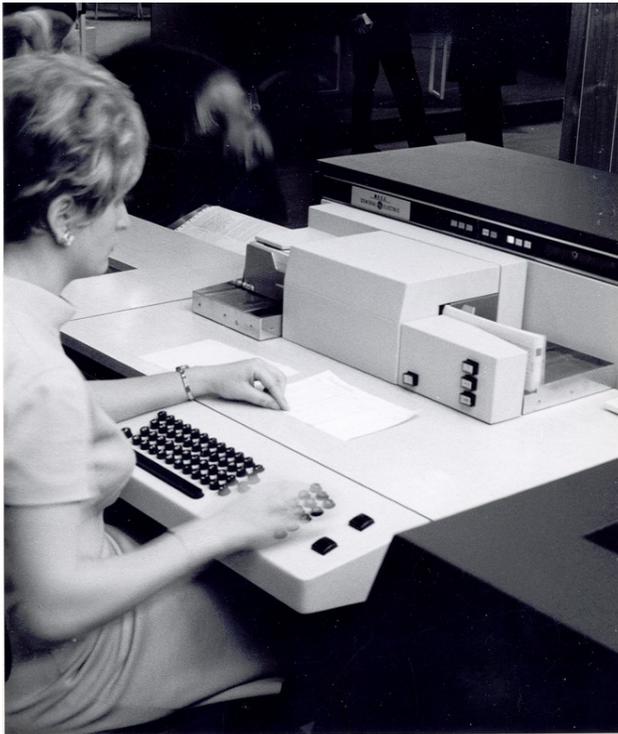
- Très longs à installer.
- Dispositifs d'E/S extrêmement lents.
- Premiers logiciels: assembleurs, bibliothèques de sous-programmes communs (E/S, calcul à virgule flottante), pilotes de périphériques, compilateurs, éditeurs de liens.
- Très faible utilisation du processeur.
- Possibilité de voir ou de modifier directement le contenu de la mémoire
- Ordinateur très sécurisé.

La porte ouverte (fin 50)

- Réserveation de l'ordinateur
- Programmation aux clés, puis par cartes, puis par télétype
- Enchainement de paquets de cartes



La porte ouverte

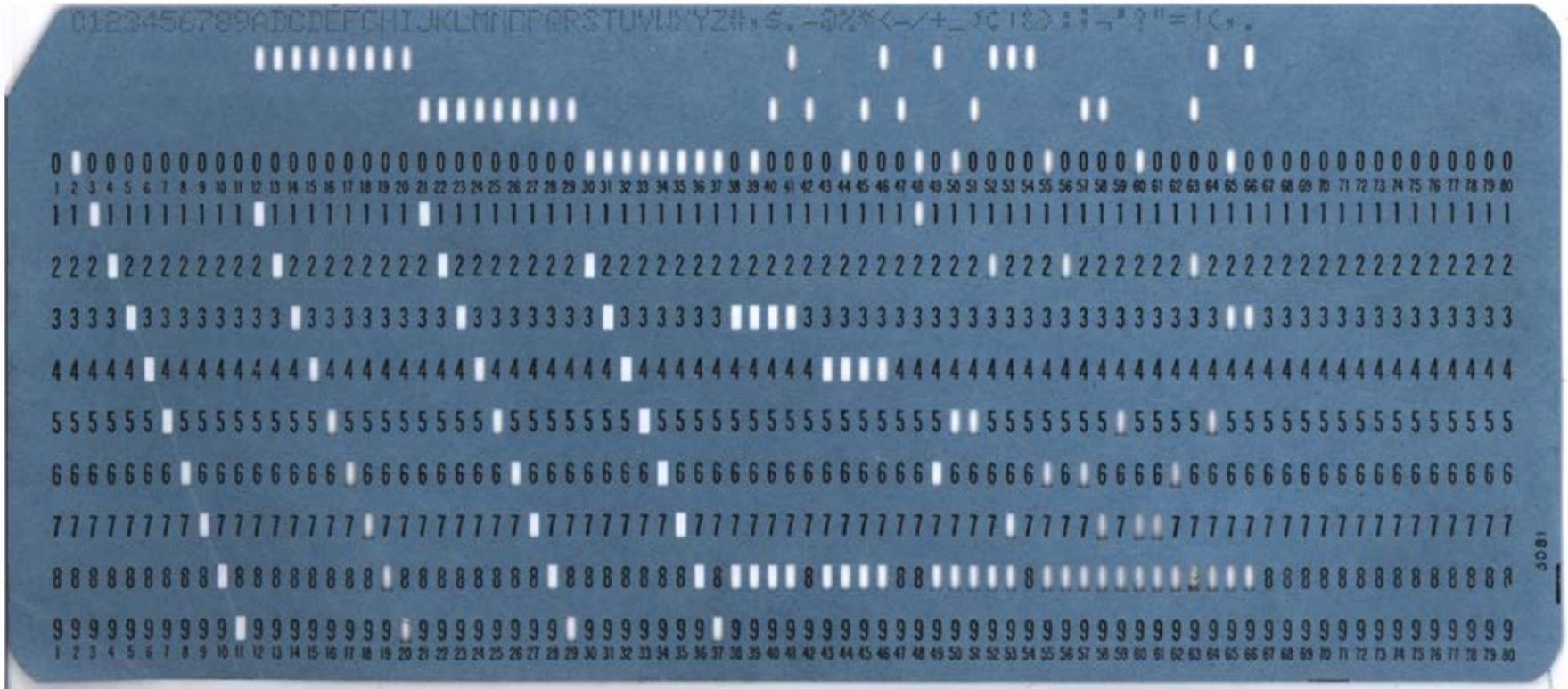


Lecteur de cartes perforées

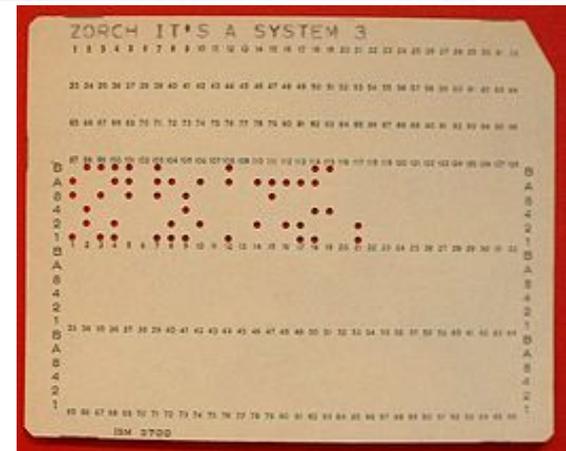
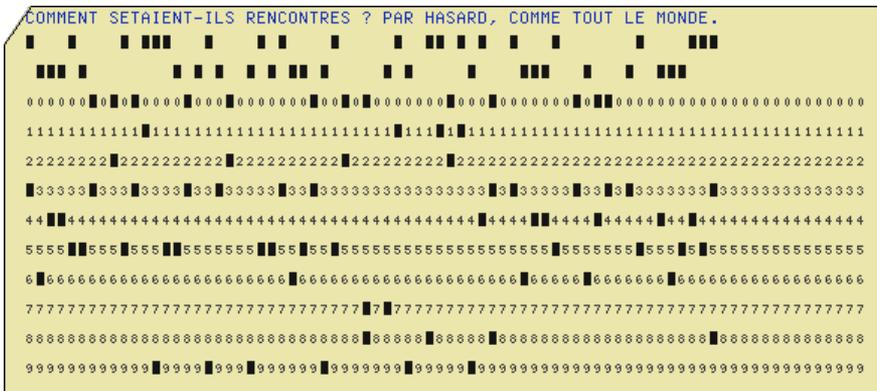


Ordinateur Bull-GE 53

Cartes perforées



Carte perforée à 80 colonnes



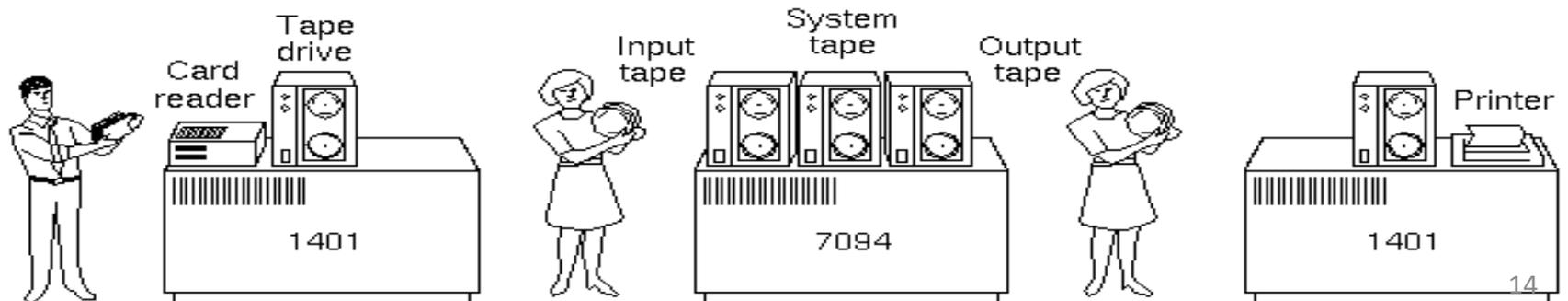
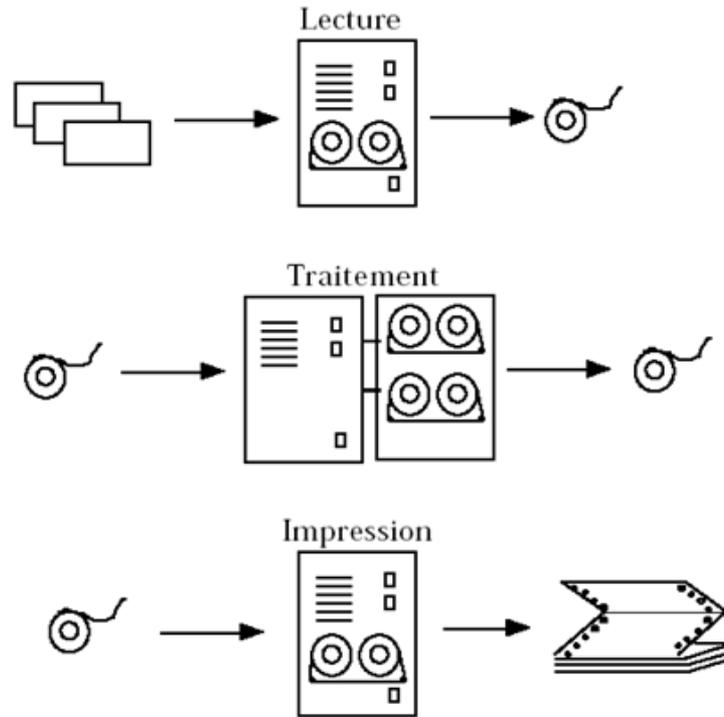
Carte perforée à 96 colonnes

Pour la saisie des programmes et des données



Perforatrice **IBM 029**

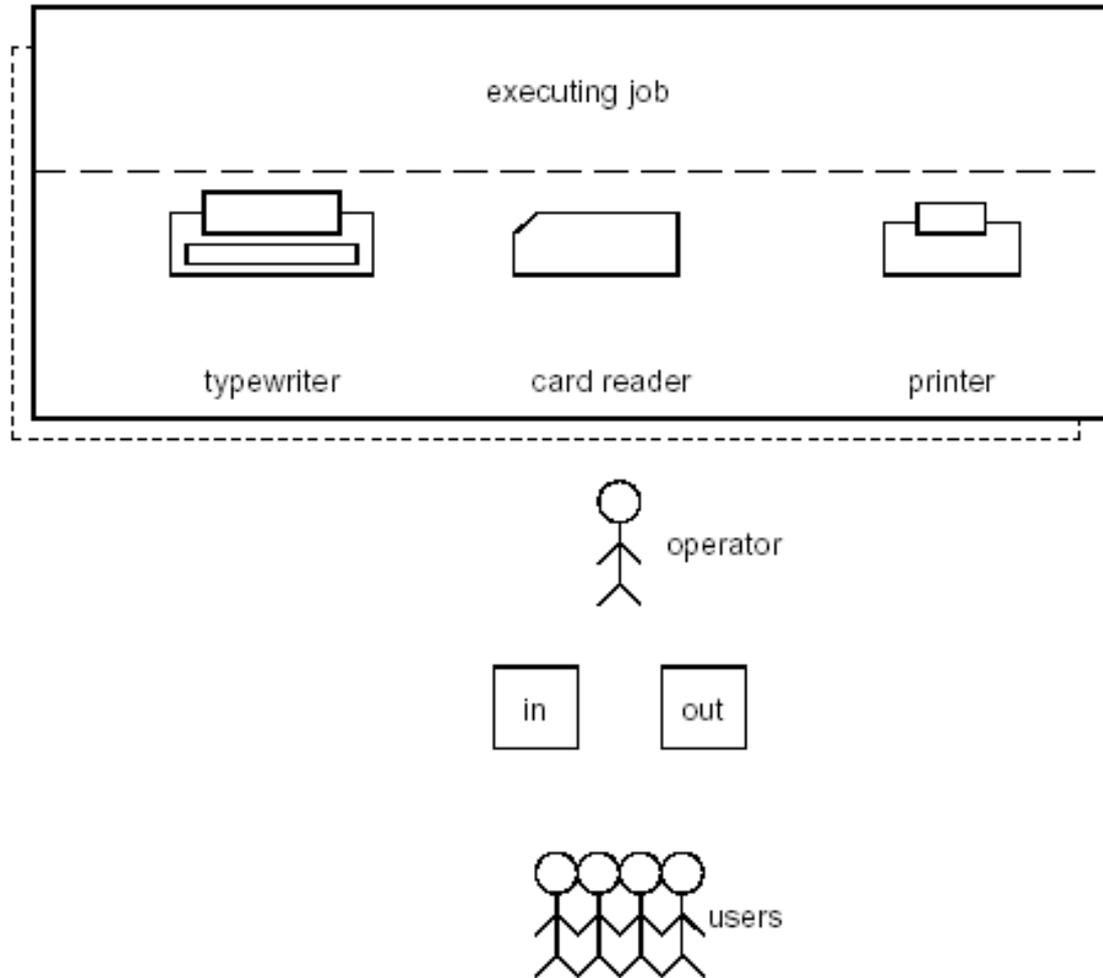
Traitement par lots (années 60)



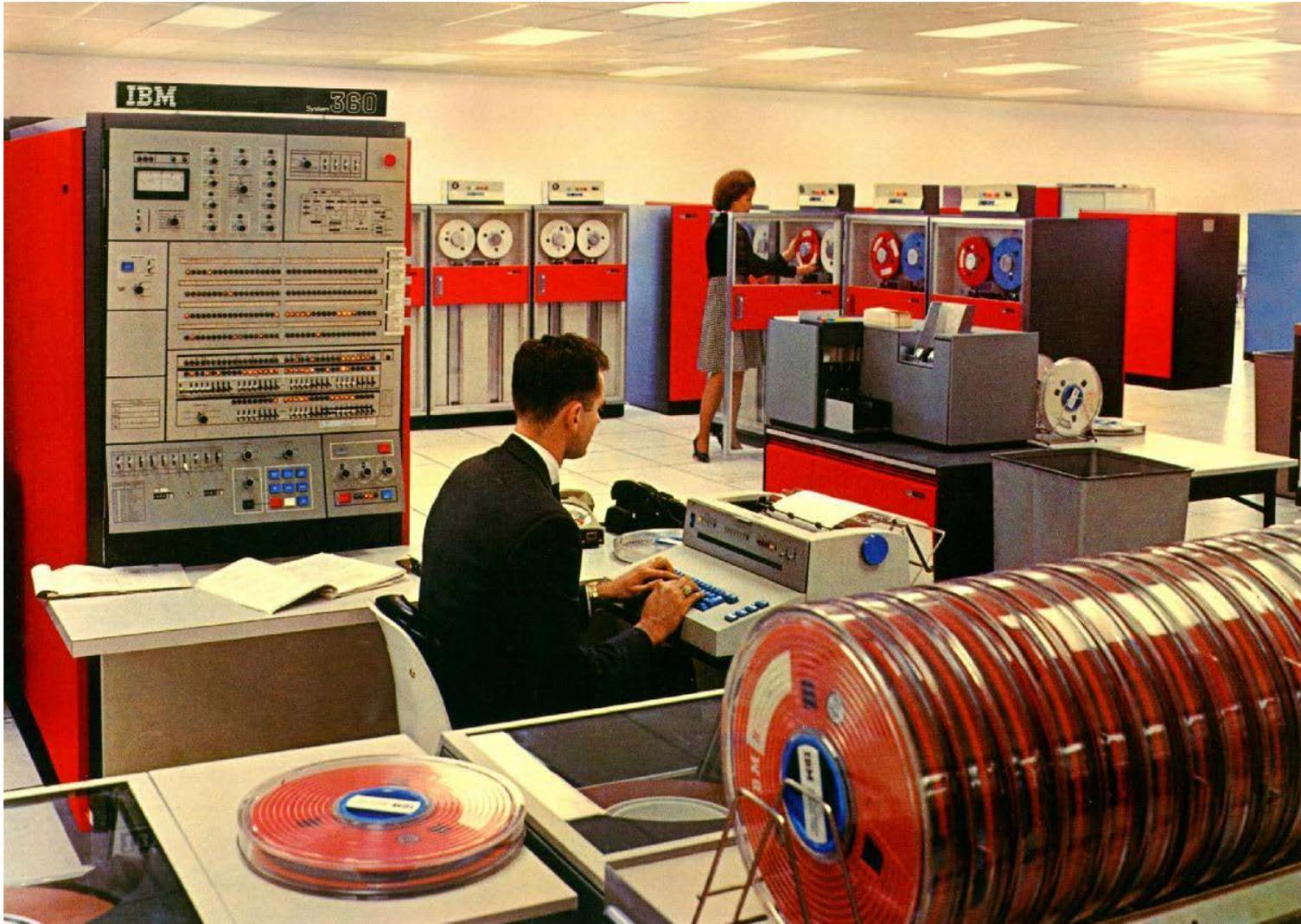
Traitement par lots (E/S tamponnées) ou « Batch processing system »

- Début de la programmation en langage de plus haut niveau.
- Les travaux (lots) sont préparés, et sont exécutés les uns à la suite des autres.
- Un operateur était chargé d'effectuer les tâches répétitives : démarrage de l'ordinateur, lancement des travaux, collecte des résultats (impression ou enregistrement sur bande magnétique)
- Il n'était plus possible à l'utilisateur d'inspecter la mémoire ou de la modifier.
- Un **moniteur d'enchaînement** permet l'exécution en séquence de la série de lots préparés à l'avance. Ce moniteur est présent en mémoire en permanence.

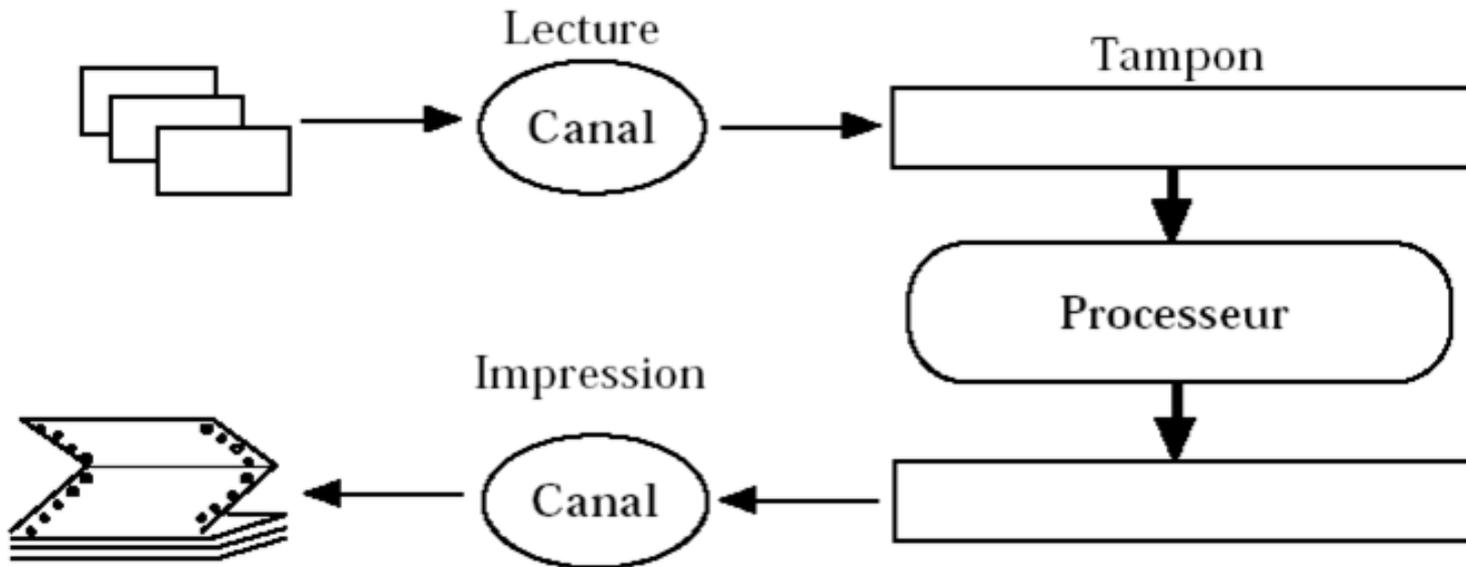
Traitement par lots (E/S tamponnées) ou « Batch processing system »



Salle machine



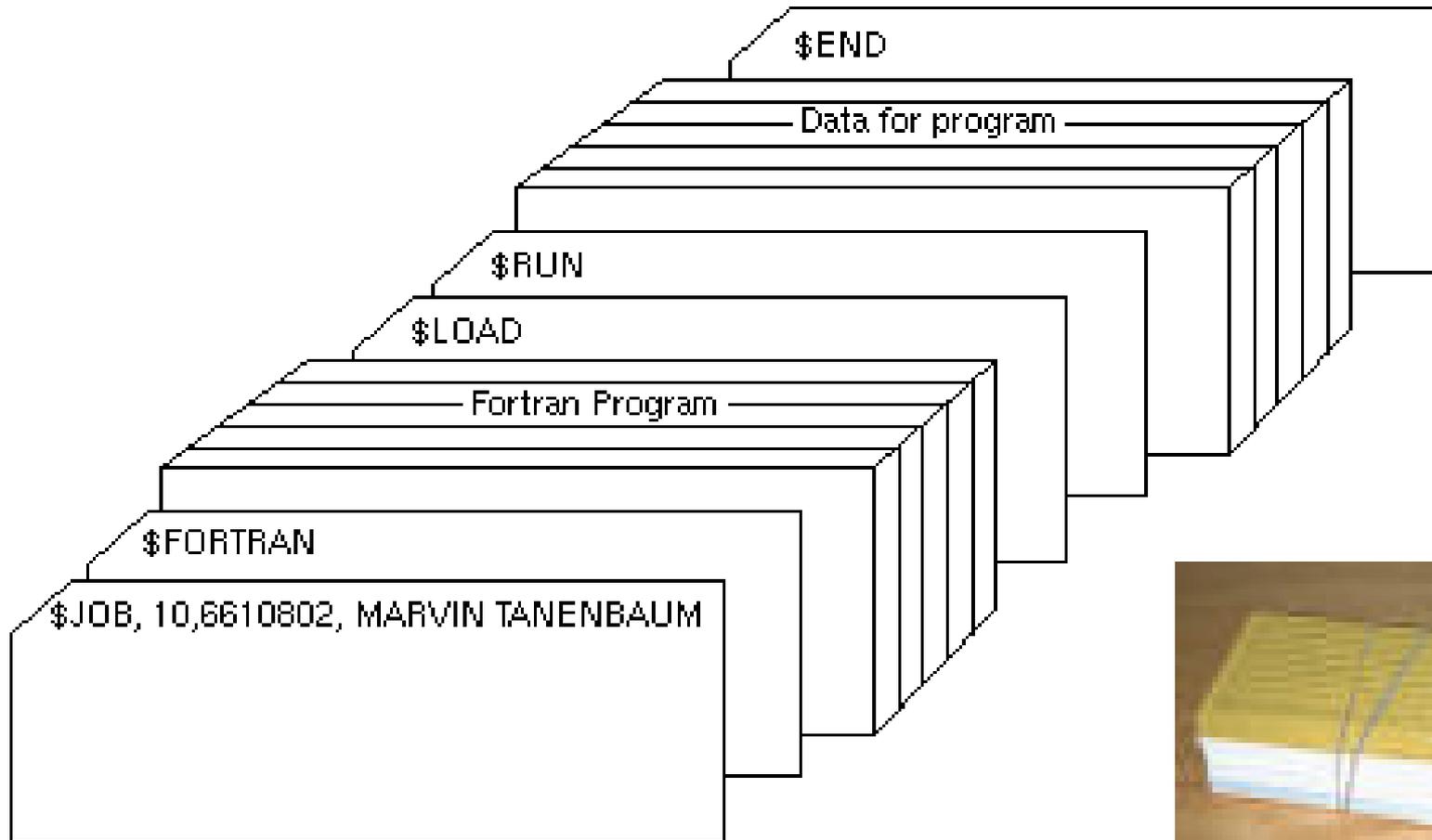
Entrées/Sorties tamponnées (65-70)



Traitement par lots (E/S tamponnées)

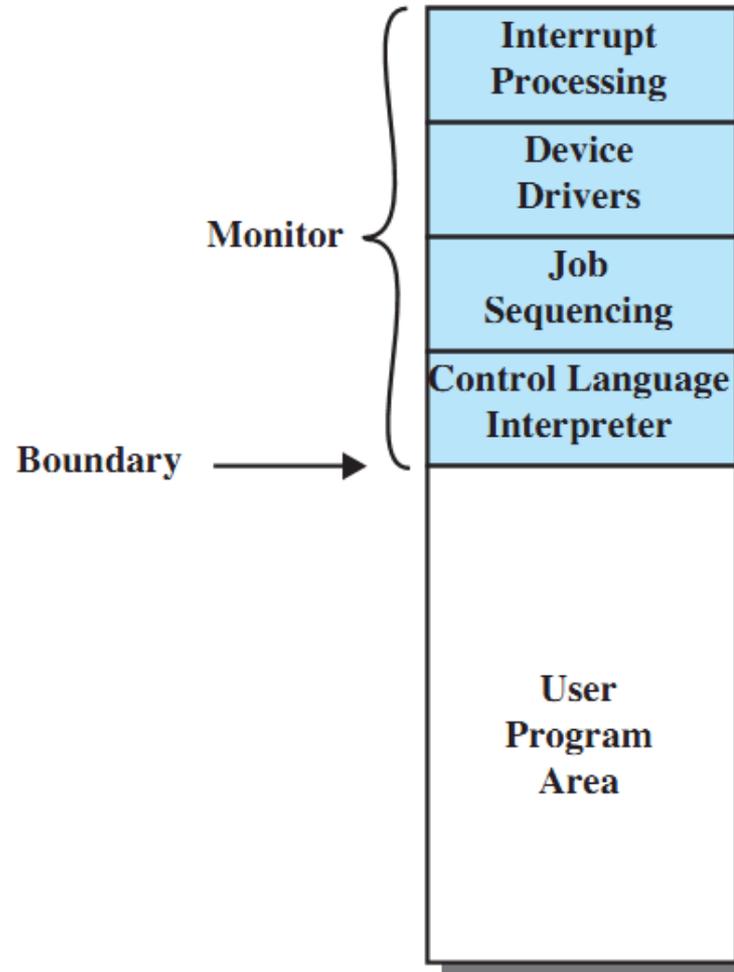
- Protection de l'ensemble des travaux de perturbations en cas d'erreur :
 - limitation du temps d'exécution,
 - supervision des entrées/sorties
 - protection de la mémoire occupée par le moniteur d'enchaînement
- Usage d'une horloge et d'instructions privilégiées
- Utilisation d'un langage de contrôle des lots (JCL = Job Control language) et de cartes de contrôle

Exemple de lot



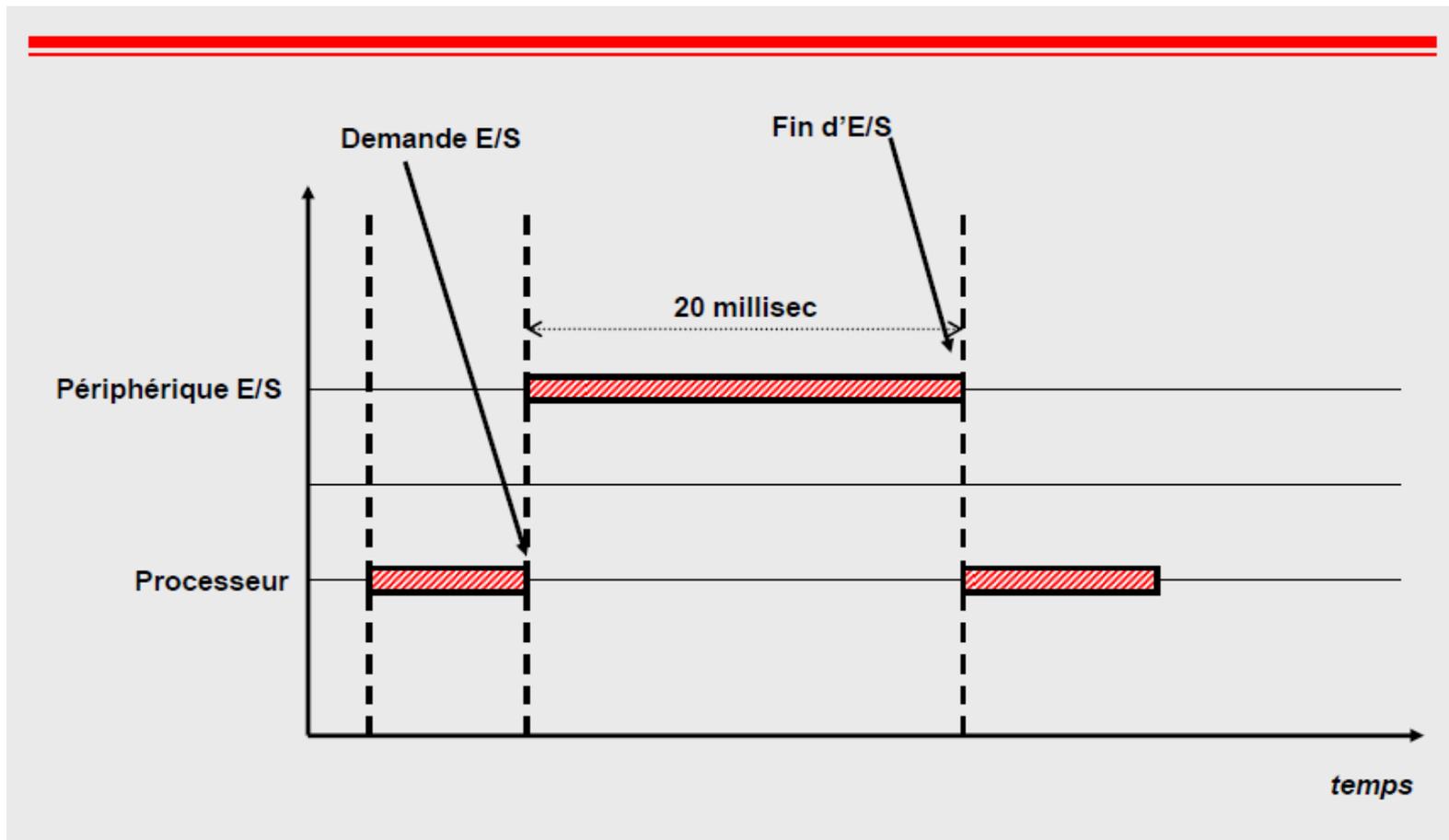
Traitement par Lots

Organisation de la mémoire



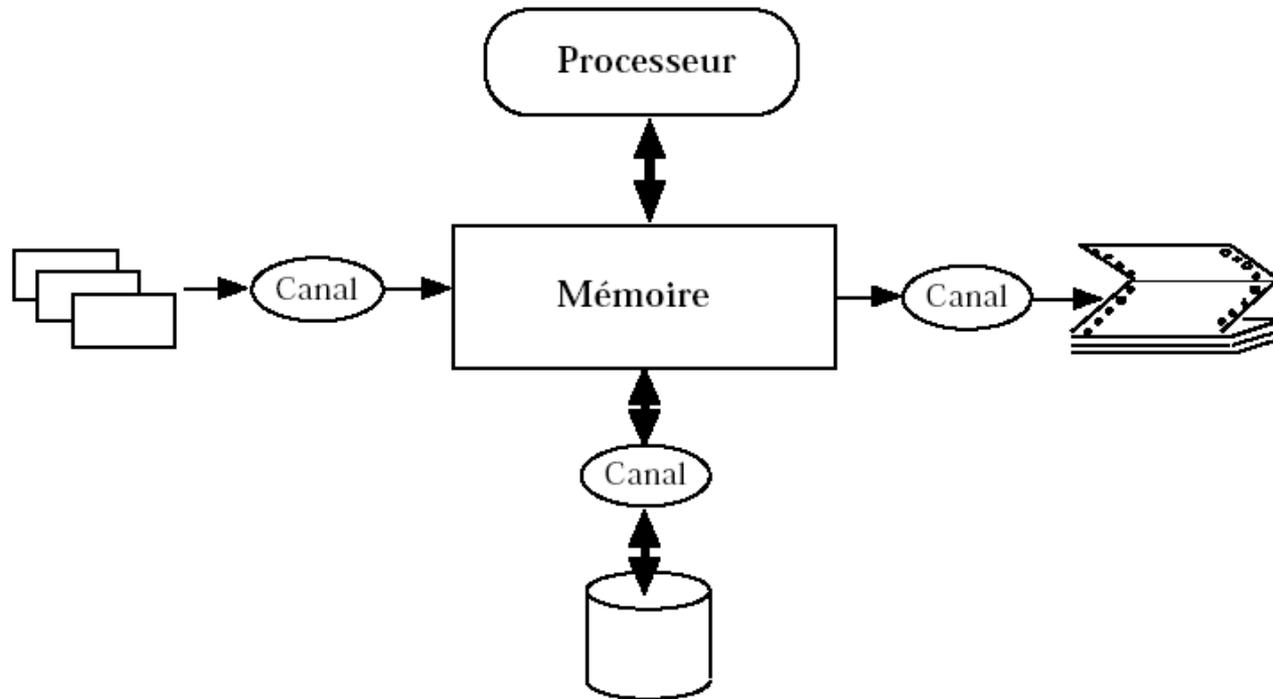
Traitement par Lots

Faible utilisation du processeur



La multiprogrammation

Multi-programmation...



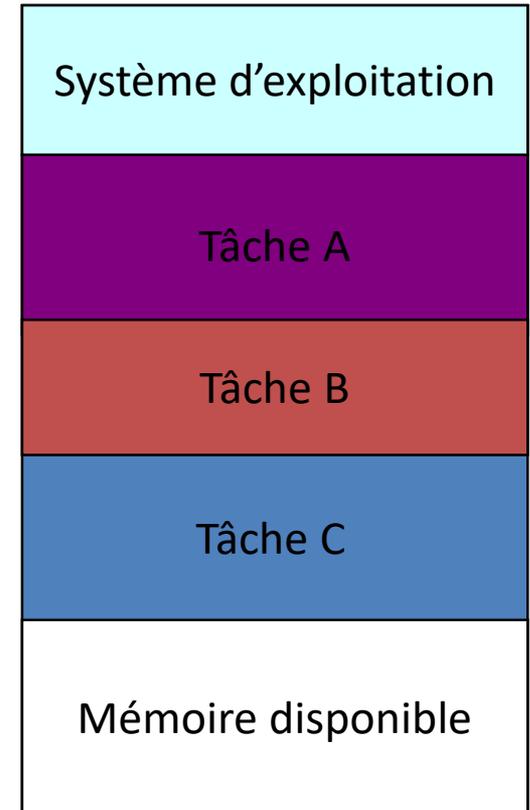
...et temps partagé (70-)

amélioration du taux d'occupation du processeur
multi-utilisateurs

La multiprogrammation

- C'est le partage de la mémoire entre plusieurs travaux en cours simultanément.
- Permet de mieux utiliser le processeur pendant les entrées/sorties (très lentes et gérées par les unités d'échange)

Organisation de la mémoire



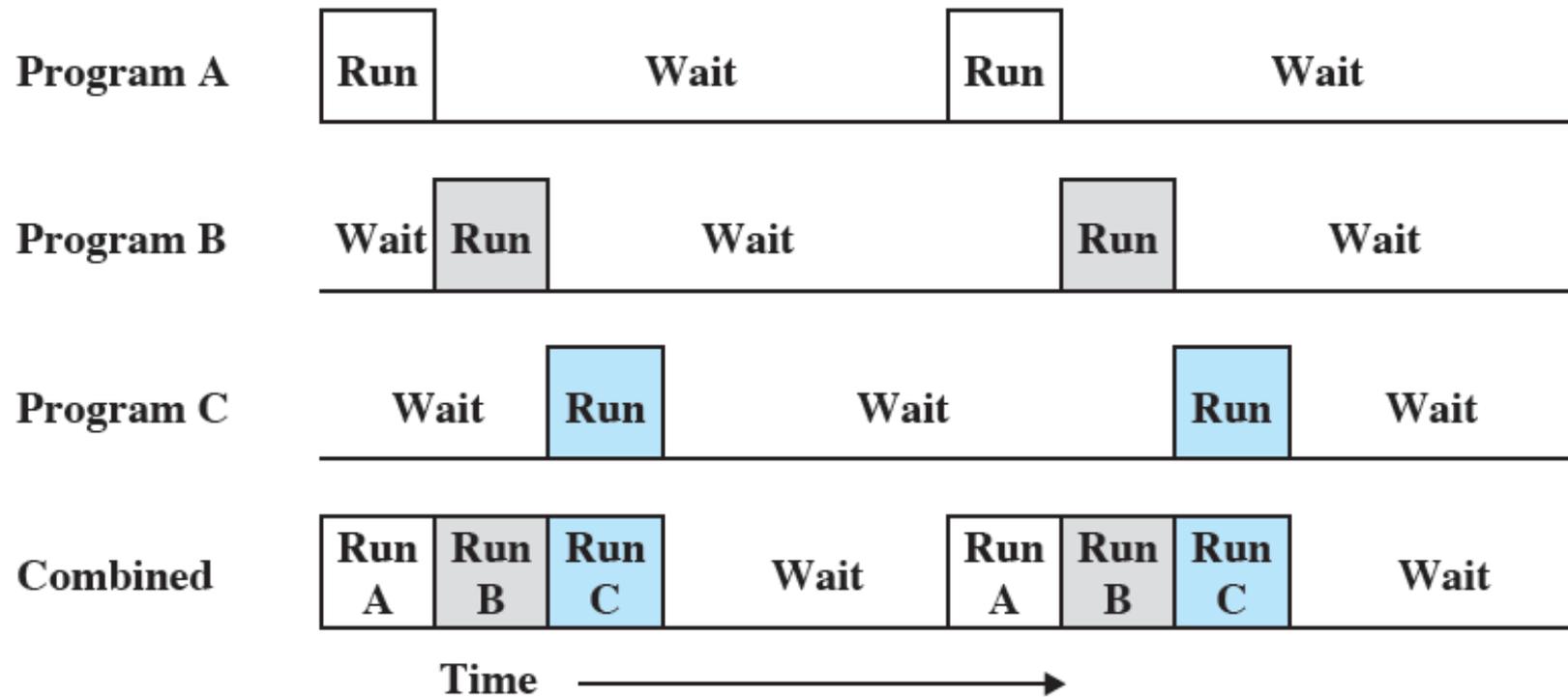
La multiprogrammation

Principe :

- Un travail en attente peut utiliser le processeur libéré par un travail qui effectue une opération d'entrée/sortie
- Le processeur peut changer d'affectation avant la fin d'un travail pour satisfaire des contraintes de temps de réponse.

Multiprogrammation

Exemple avec 3 programmes



Multiprogrammation

- Pour que cela fonctionne :
 - le temps de réaffectation du processeur à une tâche doit être très bref
 - il y a présence simultanée en mémoire de plusieurs programmes ou morceaux de programme.
- Rôle central de la mémoire
- Importance des flux entre mémoire centrale et mémoire secondaire

Multiprogrammation

- Des dispositifs spéciaux de réimplantation des programmes et de protection de la mémoire sont nécessaires:
 - Eviter qu'une tâche puisse lire / écrire dans la zone mémoire affectée à une autre tâche
 - Eviter qu'une tâche puisse manipuler la zone réservée au système autrement que par les appels système
 - Eviter qu'une tâche puisse lire / écrire des données d'E/S d'une autre tâche
- Apports de la multiprogrammation :
 - meilleure utilisation des ressources : meilleur équilibre de charge
 - réduction du temps de réponse pour les travaux courts

Temps partagé

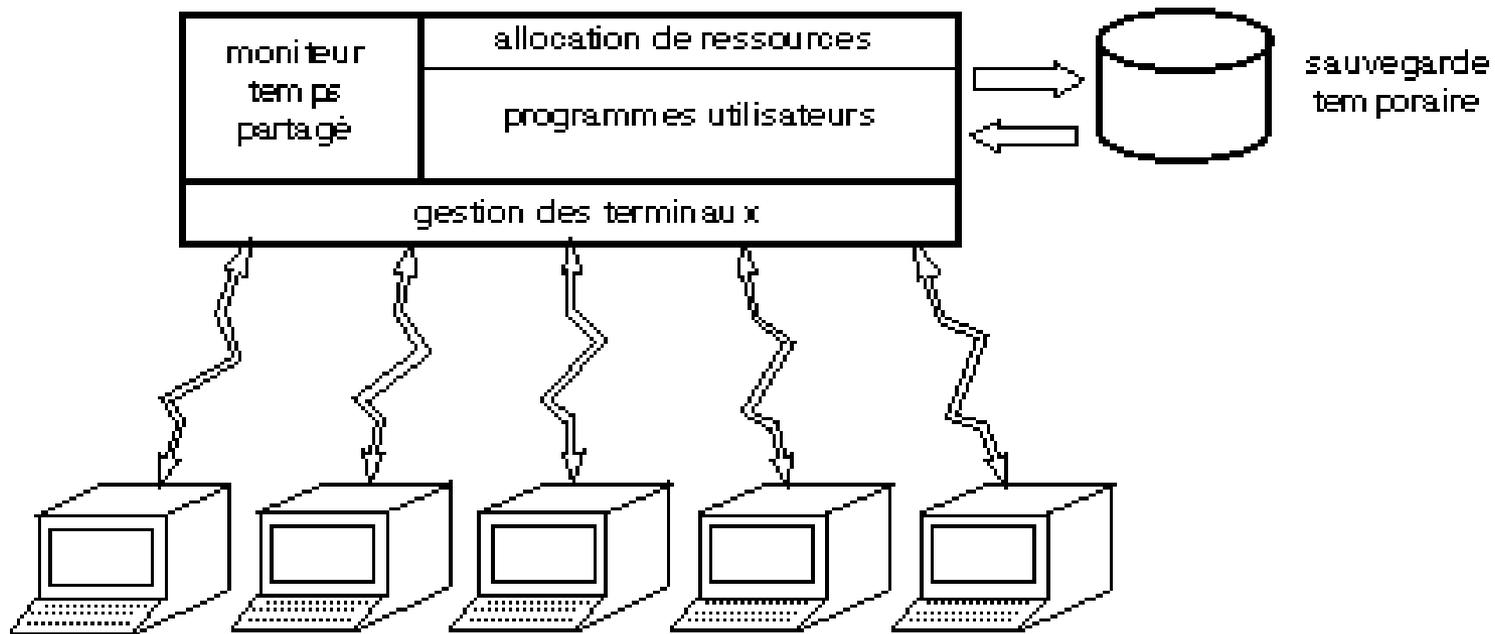
- Fonction : Partager le temps du processeur et les autres services ou ressources commun(e)s, afin d'offrir à chaque usager l'équivalent d'une machine individuelle.
- Accès au système par des terminaux,
- Utilisation interactive du système.

Temps partagé

Nécessité de garantir un temps de réponse acceptable pour l'exécution de tâches élémentaires.

Principe : allocation du processeur aux programmes des usagers par tranches de temps très brèves ou quanta.

Temps partagé

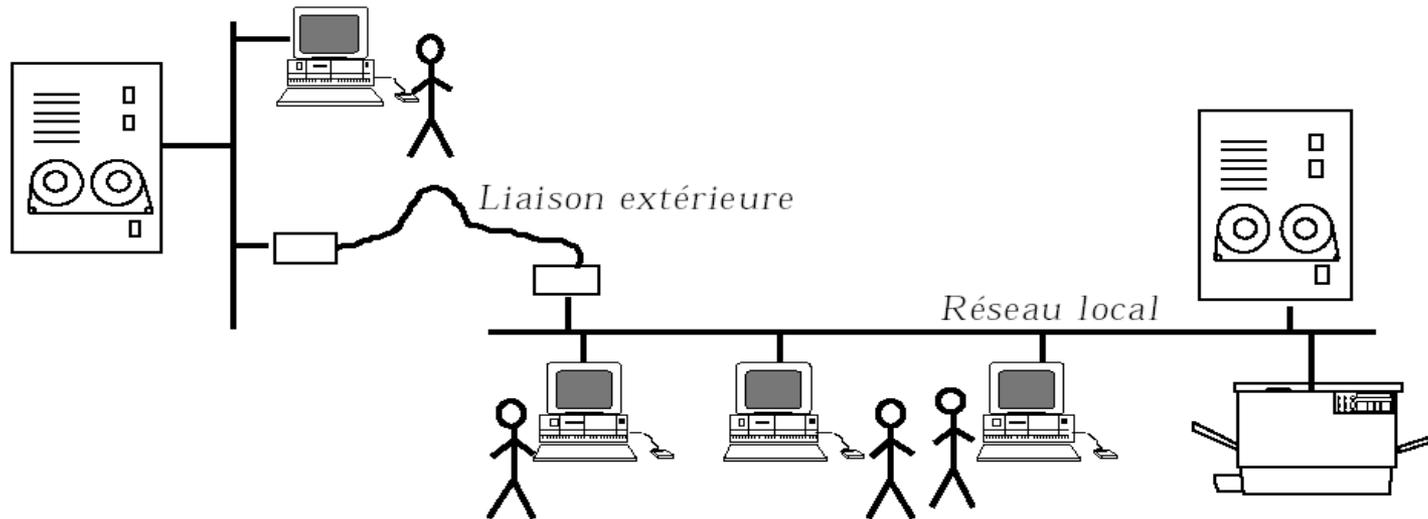


Temps partagé

- Problème de la taille de la mémoire centrale : celle-ci ne peut héberger tous les programmes en cours de tous les usagers : utilisation de mémoire virtuelle.
- Les machines sont équipées de terminaux graphiques sur lesquels on peut suivre l'évolution de plusieurs activités dans des fenêtres distinctes.

Historique et évolution actuelle

Réseaux, machines individuelles (75-80)



Explosion avec la micro-informatique (mi 80)

Ordinateurs personnels

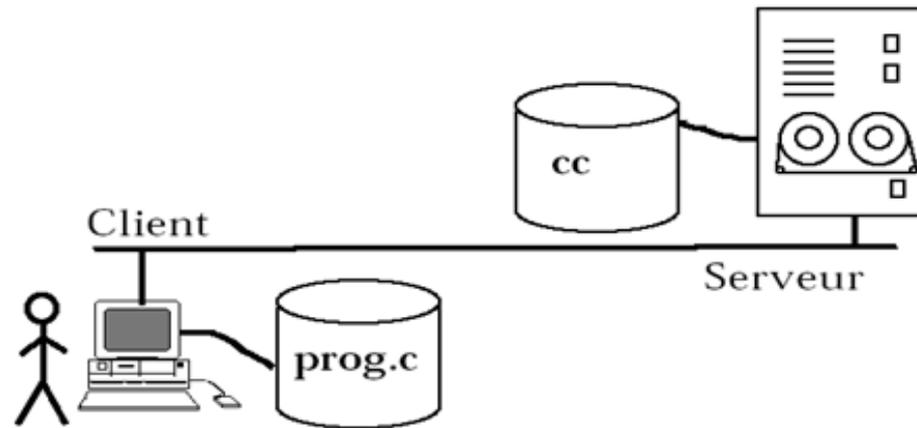
- Au départ (1980-1990)
 - Mono-tâches (MS/DOS, Windows3/11)
 - Dédiés à un utilisateur
 - Pas de sécurité
 - Très interactifs et facile d'utilisation
- Evolution actuelle
 - Multi-tâches
 - Multi-utilisateurs
 - Multiprocesseurs
 - Connexion aux réseaux

Systemes mobiles

- Au départ (fin du XXème siècle)
 - Assistants personnels (PDAs)
 - Téléphones cellulaires
 - Large gamme d'applications
 - Systemes d'exploitation spécifiques
 - Contraintes
 - Mémoire limitée
 - Processeurs lents
 - Petits écrans
 - Faible consommation d'énergie
- Evolution actuelle
 - Smartphones, Tablettes, Phablettes
 - Très large gamme d'applications
 - Systemes d'exploitations se standardisent
 - Les contraintes mémoire, processeur, écran disparaissent
 - La consommation d'énergie augmente



Systemes répartis et client-serveur



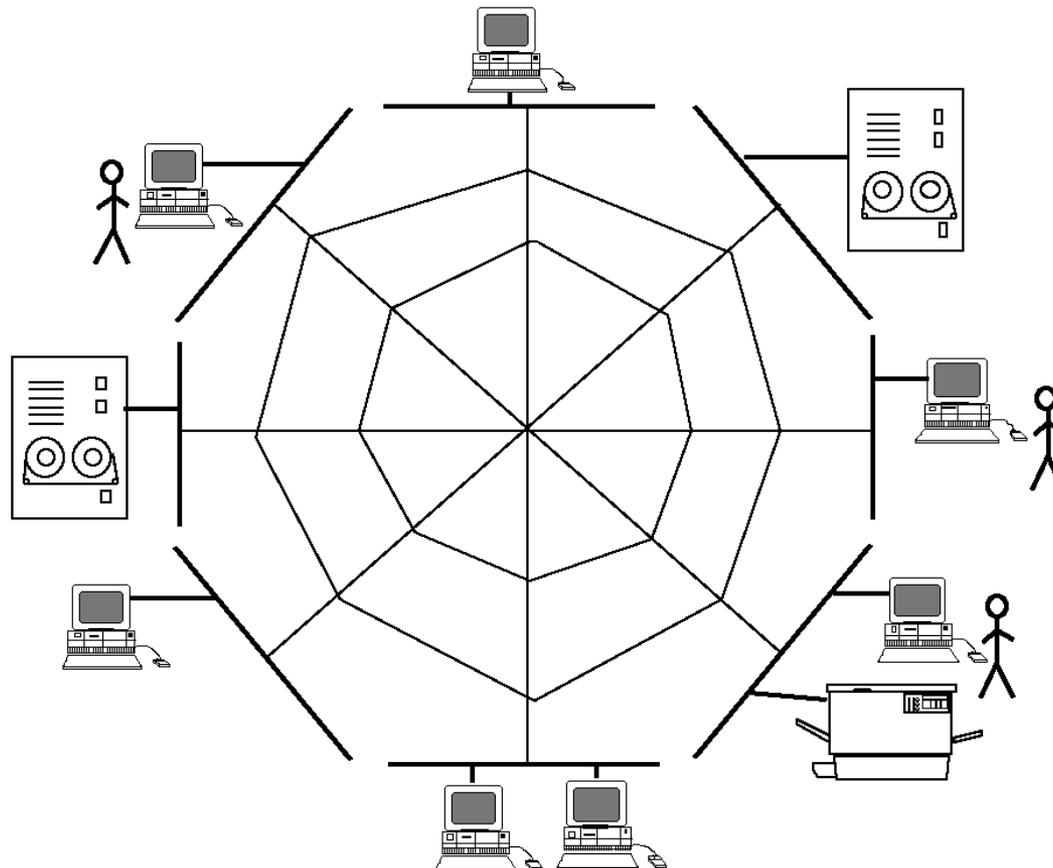
Accès transparent

Partage de ressources coûteuses

Maintenance logicielle centralisée

Situation actuelle

Le Client-Serveur généralisé et mondialisé



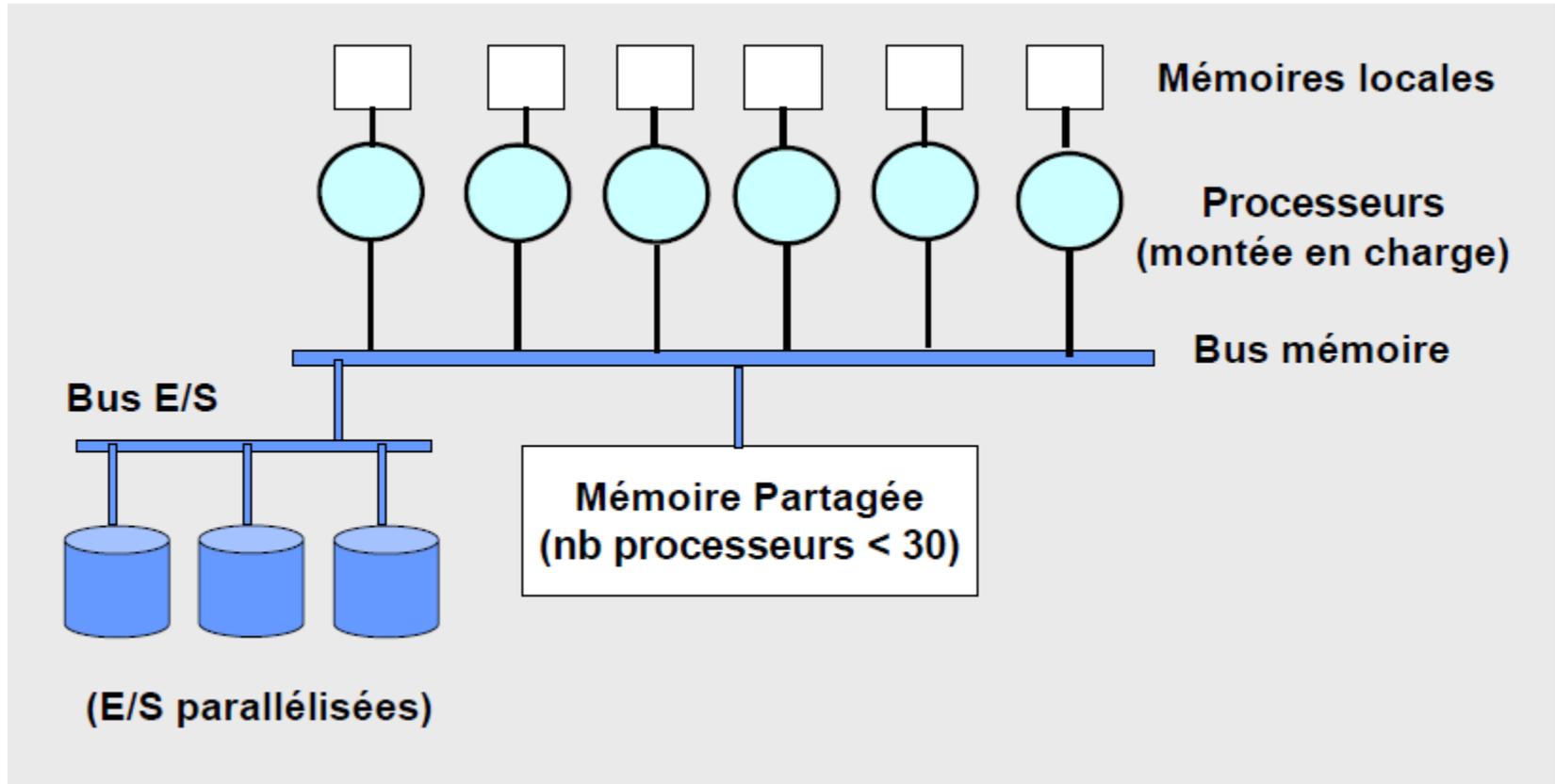
- Services Web
- Accès à des bases de données
- Accès à des plateformes fournissant des ressources et des services (bureau virtuel)
- Ressources disponibles de n'importe où sur Internet

= cloud computing

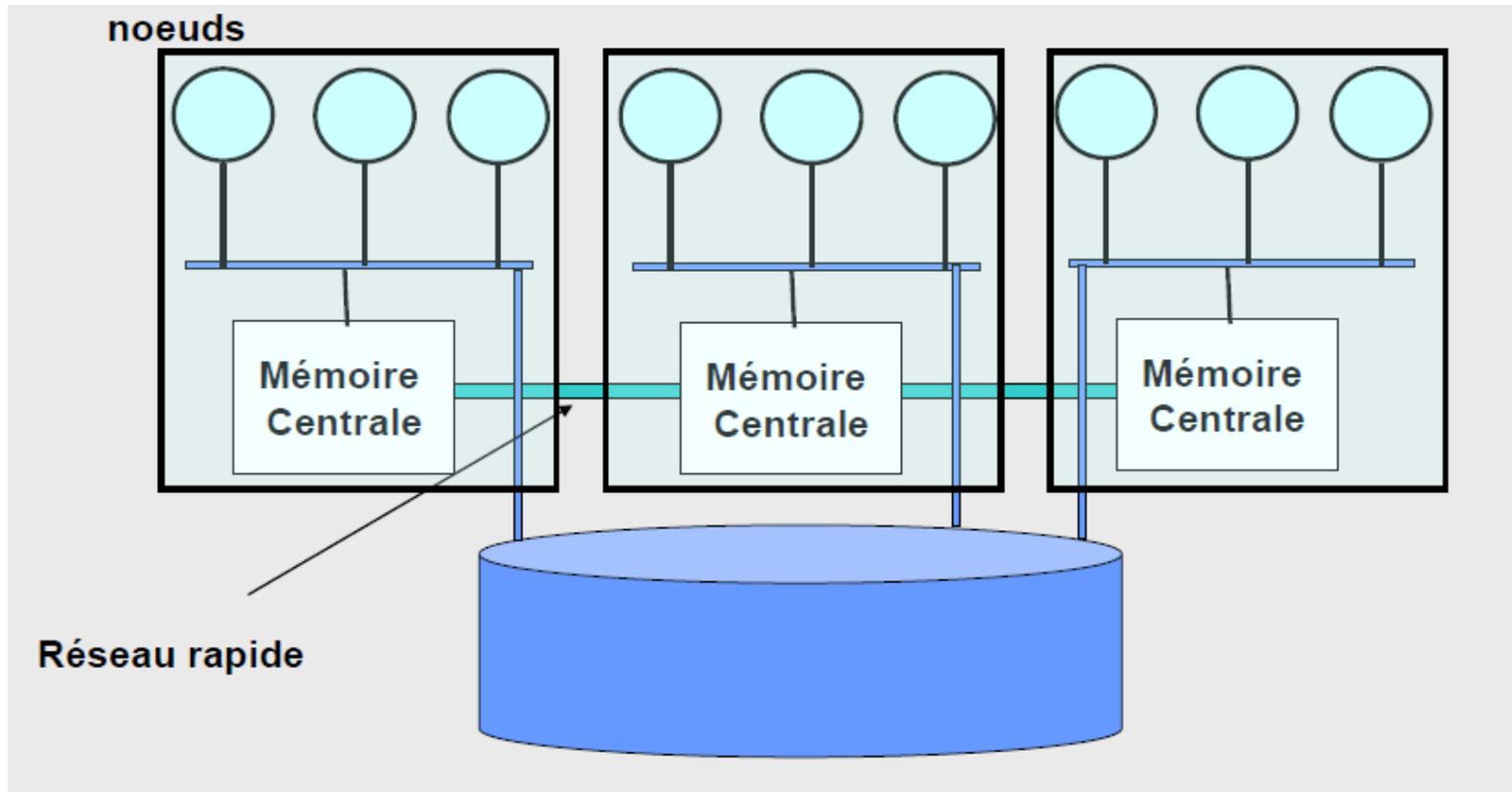
Systemes parallèles multiprocesseurs

- SMP (Symetric Multi Processeurs)
 - Systemes généralistes utilisés pour des serveurs (BD, Web, SI, Fichiers, etc.)
 - Processeurs standards
 - Partage total de la mémoire
- Machines de calcul parallèles
 - Architectures spécialisées
 - Processeurs spécifiques pour le calcul vectoriel
 - Réseaux spécialisés
 - Partage total ou partiel de la mémoire

Systemes parallèles multiprocesseurs



Systemes clusterisés



- Plusieurs nœuds connectés par un réseau spécialisé : échange rapide de données de grande taille
- Permettent : la montée en charge, le passage à l'échelle, une haute disponibilité

Grilles de calcul

- Infrastructures réparties composées d'un ensemble de machines hétérogènes et potentiellement distantes
- Milliers de machines
- Calcul intensif
- Garanties de disponibilité
- Système de réservation des ressources matérielles

Grilles de calcul



Les serveurs du Centre de données du CERN forment le niveau 0 de la Grille de calcul mondiale pour le LHC (Grand Collisionneur de Hadrons / Large Hadron Collider)
(Image : CERN)

Le Cloud Computing



- Système réparti dans lequel les clients utilisent des services déportés sur un ou plusieurs serveurs distants, sans avoir à administrer ces services ni les machines qui les supportent
- Les services proposés par le Cloud peuvent s'exécuter au sein d'architectures cluster
- 3 modèles
 - IaaS (Infrastructure as a Service)
 - PaaS (Platform as a Service)
 - SaaS (Software as a Service)

Systemes répartis

Englobent maintenant

- Architectures distribuées fortement couplées
 - Systèmes parallèles multiprocesseurs
 - Architectures clusters
- Architectures distribuées faiblement couplées
 - Grilles de calcul
 - Cloud computing
 - Applications s'exécutant sur Internet

Le Système d'Exploitation réparti est mis en œuvre par une couche logicielle fonctionnant au dessus des SE des serveurs du système réparti.

Un site à visiter

<http://www.computerhistory.org/timeline>