
CORBA haute performance

« CORBA à 730Mb/s ! »

Alexandre DENIS

PARIS/IRISA, Rennes

Alexandre.Denis@irisa.fr

Plan

- ✓ Motivations : concept de grille de calcul
- ✓ CORBA : concepts fondamentaux
- ✓ Vers un ORB haute performance
- ✓ Conclusion et perspectives

Ressources de la grille

✓ Réseau hiérarchique

- ◆ WAN : Internet, VTHD, etc.
- ◆ LAN : Ethernet
- ◆ SAN : Myrinet, SCI

✓ Machines

- ◆ Grappes de PCs
- ◆ Calculateurs parallèles



Applications de la grille

✓ Applications parallèles

- ◆ grille vue comme un supercalculateur

✓ Applications de couplage de codes

- ◆ exécution de plusieurs codes
 - écrits dans des langages différents
 - qui utilisent des paradigmes de communication différents (MPI, DSM, ...)
- ◆ contraintes de localisation (visualisation, confidentialité, ...)

Structural Mechanics

Optics



Thermal

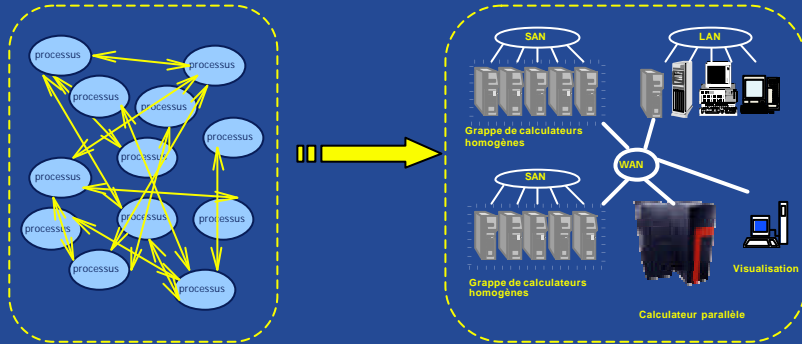
Dynamics

Satellite design

Approche à la « Globus »

✓ Modèle à plat « flat »

- ♦ la structure n'apparaît pas au niveau applicatif
- ♦ grille vue comme une seule grappe MPI
- ♦ tous les nœuds sont égaux



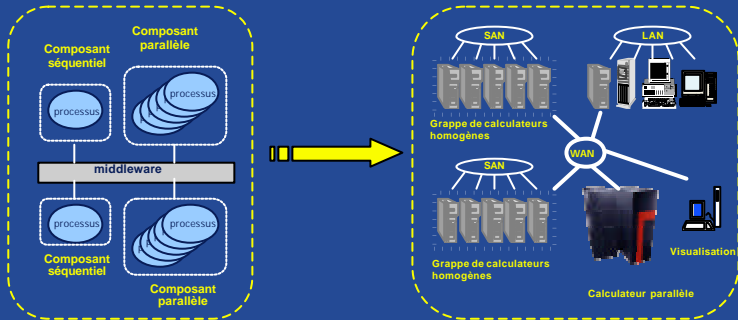
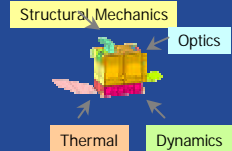
Approche du projet PARIS

✓ Composants

- ◆ structuration de l'application
- ◆ encapsulation du parallélisme

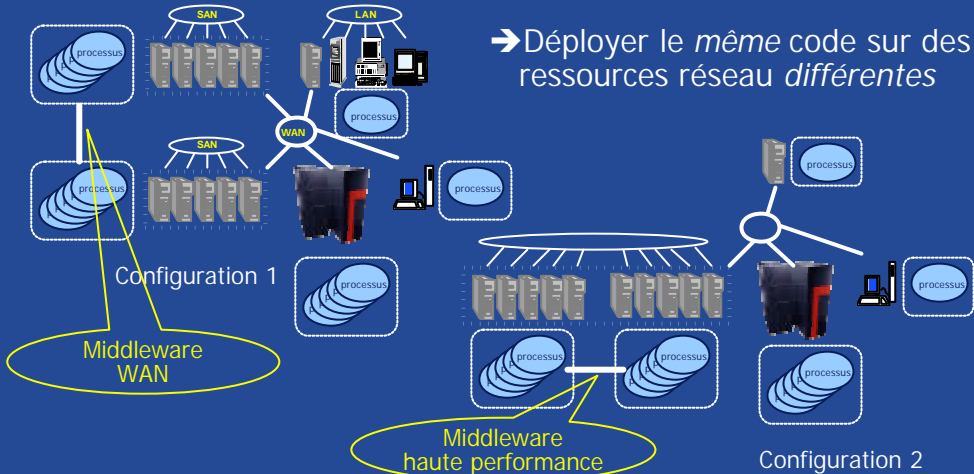
✓ Couplage

- ◆ composants reliés par un middleware



Placement composants-ressources

→ Déployer le *même* code sur des ressources réseau *différentes*



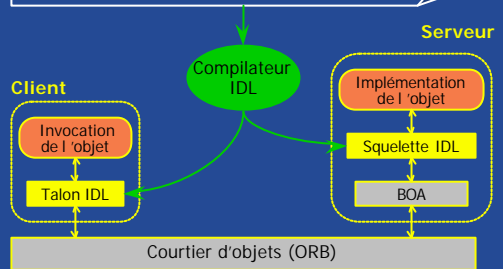
→ Le middleware *s'adapte* au réseau disponible

Middleware : CORBA



- ✓ Accès à des objets à distance
 - ♦ localisation transparente
 - ♦ prise en charge de l'hétérogénéité
- ✓ Implémentations courantes :
 - ♦ presque exclusivement TCP...

```
interface MatrixOperations {  
    const long SIZE = 100;  
    typedef double Vector[ SIZE ];  
    typedef double Matrix[ SIZE ][ SIZE ];  
    void multiply(in Matrix A,  
                in Vector B, out Vector C );  
};
```



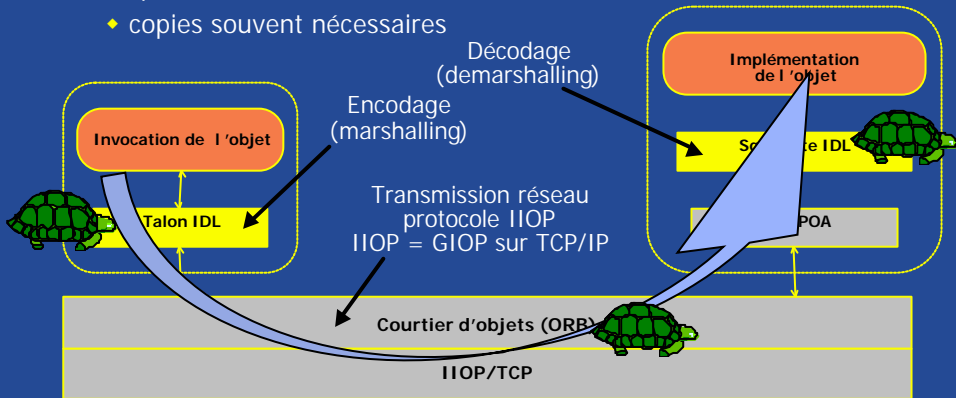
Anatomie d'une requête CORBA

✓ Encodage et décodage

- ♦ format GIOP
- ♦ alignement des données spécifié
- ♦ copies souvent nécessaires

✓ Transmission réseau

- ♦ limitée à TCP la plupart du temps...



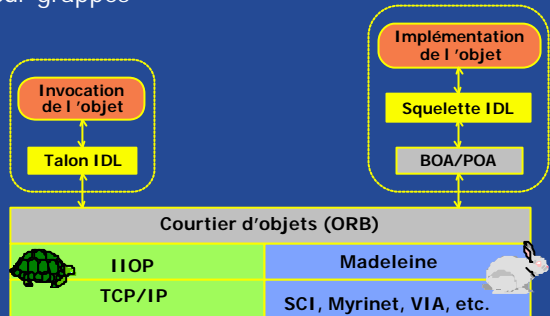
ORB sur Madeleine

✓ Madeleine : bibliothèque de communication pour grappes

- ◆ module de PM² (<http://www.pm2.org>) - LIP, ENS Lyon
- ◆ vise la haute performance sur grappes
- ◆ approche « zéro copie »

✓ Portabilité

- ◆ Myrinet, SCI, VIA, etc.
- ◆ interface uniforme pour les différents réseaux



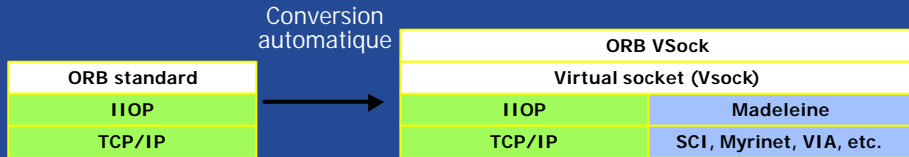
Interopérabilité

✓ Compatibilité avec le « monde extérieur »

- ◆ Madeleine pour contacter les nœuds de la grappe
- ◆ IIOP pour dialoguer avec les autres ORB

✓ *Virtual sockets*

- ◆ modification minimale de l'ORB
 - pour ne pas dépendre des évolution de l'ORB
- ◆ possibilité de porter d'autre middlewares sur VSocket
 - DSM, JVM, etc.
- ◆ implémente un sous-ensemble des fonctionnalités « socket »



Transparence

✓ Services fournis par VSocket

- ♦ sélection automatique du protocole
 - en fonction du nœud distant
 - en fonction de l'objet distant
- ♦ résolution d'adresses
 - en fonction des adresses IP
- ♦ multiplexage au-dessus de Madeleine

✓ Dynamicité

- ♦ CORBA est client/serveur avec connexion dynamique
- ♦ Madeleine a une topologie réseau fixe
- ➔ projection d'une topologie client/serveur sur un graphe complet d'interconnexion
- ➔ application chargée dynamiquement par des processus résidants sur la grappe

Performances

✓ Débit

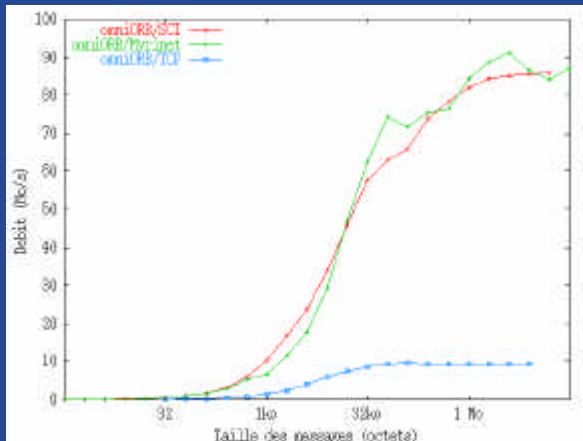
- ◆ SCI : 86Mo/s
- ◆ Myrinet : 91Mo/s
- ◆ 92 à 99% du débit Madeleine exploité

✓ Latence

- ◆ CORBA/IIOP (standard) 150 μ s
- ◆ CORBA/SCI ou Myrinet 50 μ s

✓ Amélioration nette !

- ◆ En latence **et** en débit

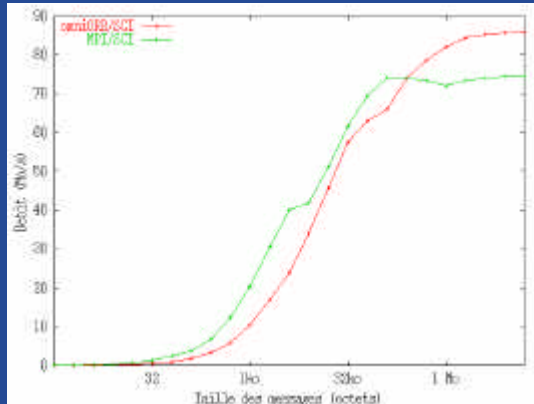


CORBA sur SISI/SCI, BIP/Myrinet et TCP/Ethernet-100

CORBA et MPI

✓ Comparaison entre CORBA et MPI

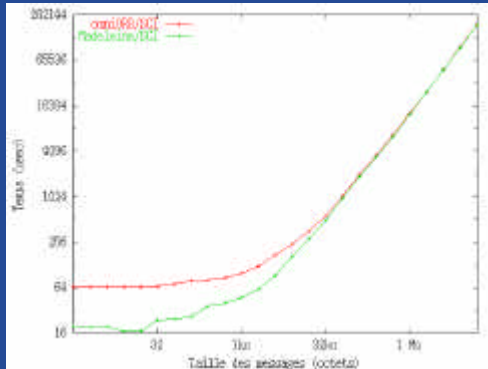
- ♦ omniORB/Makeleine - MPI/Makeleine
- ♦ performances du même ordre de grandeur
- ➔ structurer l'application ne pénalise pas les performances



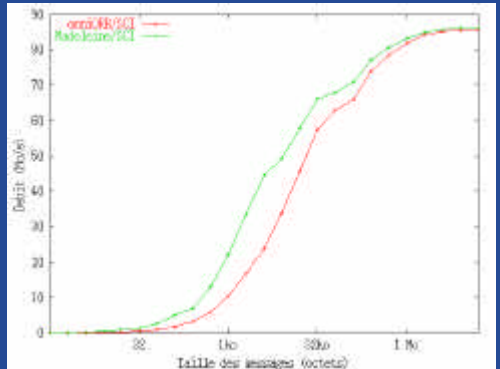
Latence

✓ Quel protocole pour CORBA/Madeleine ?

- ♦ GIOP pas optimal
 - surcoût logiciel $\approx 40 \mu\text{s}$
 - entêtes longs superflus à l'intérieur d'une grappe



Temps de CORBA et Madeleine sur SCI



Débit de CORBA et Madeleine sur SCI

Conclusion

*Implémentation CORBA
complète opérationnelle !*

✓ CORBA

- ◆ favorise la structuration des applications
- ◆ sa réputation de lenteur est une légende !

✓ Structuration des applications en composants

- ◆ simplifie l'allocation des ressources
- ◆ ne pénalise pas les performances

→ Validation de notre approche de structuration des applications

→ Intégration dans Padico, notre plate-forme d'objets distribués haute performance

Perspectives

✓ Optimisation de la latence

- ♦ utilisation d'un ESIOP

✓ Padico : intégration CORBA + MPI

- ♦ multiplexage de l'accès au réseau
- ♦ factorisation des threads de communication

*Coopération plutôt que concurrence
des accès aux ressources*

