

# Travaux du laboratoire ID sur les grappes de grande taille

Yves Denneulin

Laboratoire ID-IMAG

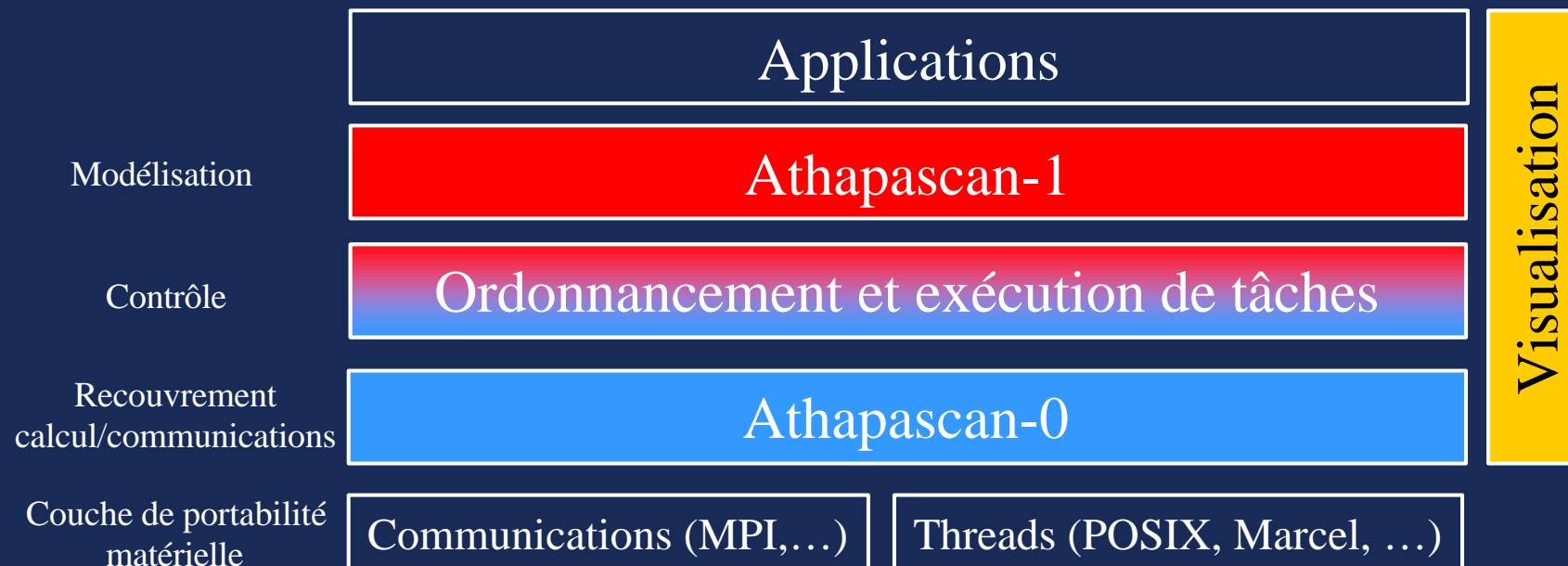


# Introduction

- Laboratoire ID
  - Créé en 2000
  - Anciennement élément du LMC
  - Équipe parallélisme depuis 1992
  - Héberge le projet INRIA APACHE depuis 1996
- Composition
  - 12 permanents
  - 18 doctorants (fort renouvellement cette année), 4 DEAs
  - 5 ingénieurs
- Thèmes de recherche
  - Algorithmique parallèle pour le calcul scientifique
  - Ordonnancement théorique et applications
  - Modèles de programmation parallèle
  - Support d'exécution pour le calcul parallèle et les applications distribuées

# Historique

- Interface de programmation : **Athapascal-1**
  - **Description** implicite du parallélisme (indépendamment de la machine)
  - Ordonnancement **séparé** de l'application (par annotation [Open-MPI])
- Noyau exécutif : **Athapascal-0**
  - Processus légers communicants (*multithreading* distribué  $\Rightarrow$  **recouvrement**)
  - Portable



# Applications

- Dynamique moléculaire
- Simulation océanique
- Décomposition de domaines
- Simulation de trafic routier
- Appariement dense d'images
- Cartographie
- Solutions éprouvées mais...

# Projet grappe 256

- Objectifs
  - Matériel standard
    - Réseau standard (éventuellement rapide)
  - Logiciels libres (GNU/Linux)
- Passage à l'échelle
  - Algorithme
  - Support d'exécution
    - lanceur
  - Ordonnanceur
  - Évaluation de performances et débogage
    - Vision synthétique d'un grand nombre d'éléments
  - Déploiement
    - Grand nombre de machines
    - Administration centralisée
- Projet achat grappe 256 processeurs (128 SMPs)

# Collaborations industrielles

- Microsoft

- Portage et évaluation sur NT de la pile logicielle
- Étude de la scalabilité
- Grappe de 40 machines sous NT/W2000

- Hewlett-Packard

- Administration intranet sous NT
  - Solution à base d'agents sur le réseau
- Problématique
  - Utilisation de machines du commerce
  - Exploitation de jachères de calcul de toutes ressources sur un intranet
  - Contraction/extension d'applications
- Don de 225 machines eVectra
  - PIII 733, 256Mo, 20Go
  - Ethernet 100 Mb + commutateurs Gigabits
  - 81GFlops à Linpack (370ème au Top 500)

# Collaborations industrielles (2)

- Bull
  - Évaluation de performances
    - Outils et théorie, traçage et profiling d'applications parallèles
  - Support d'exécution
    - MPI efficace et thread-aware
    - Couche d'interopérabilité pour implantation de MPICH
    - Déploiement efficace d'applications
      - Clonage de réseau virtuel dynamiquement extensible
- Mandrake
  - Construction d'une distribution orientée grappe
- Polyspace
  - Application de vérification de code

# Projets en cours

- Support d'exécution
  - Impact d'un grand nombre de composants matériels
    - Performances
    - Conception
    - Description (hiérarchie ?)
  - Communication haut débit
    - Exploitation d'architectures matérielles
    - Respect de la norme MPI
  - Système de fichiers distribué
    - Vision globale des systèmes de fichiers + efficacité et facilité d'utilisation
    - Utilisation de l'ensemble des disques locaux d'une grappe
  - Extension/contraction dynamique d'applications
    - Détection d'inactivité
      - Caractérisation de l'inactivité
    - Ajout/retrait d'une machine à une configuration MPI



# Projets en cours (2)

- Administration
  - Déploiement facile et rapide d'une configuration
  - Centralisation de l'administration
  - Prise en compte de l'hétérogénéité matérielle et logicielle
- Surveillance
  - Vision synthétique d'informations pertinentes
  - Prise de logs
- Débugage
  - Prise et exploitation de traces
- Ordonnancement
  - Couplage d'ordonnanceurs
  - Prise en compte de l'architecture physique
    - Latence des communications
    - Hiérarchie mémoire (impact des caches)
  - Ordonnanceur global intra et inter applications

# Projets en cours (3)

- Métacomputing
  - Système de fichiers distribués (cadre DataGrid)
  - Ordonnancement hiérarchique
    - Heuristiques
    - Implantation
- Applications
  - Simulation de tombé de tissus
  - Simulation de mondes en 3D avec Athapascan-1
  - Cartographie