

Cluster de PC : l'expérience de l'équipe d'Analyse Numérique et EDP d'Orsay.

V. Louvet

*Université d'Orsay, Département de Mathématique,
Equipe EDP et Analyse Numérique*

Rencontres MATHRICE, le vendredi 29 mars 2002

PLAN

1. Besoins de l'équipe,
2. Configuration matérielle du cluster,
3. Administration,
4. Sous-clusters,
5. Compilateurs, bibliothèques et utilitaires,
6. Exploitation,
7. Perspectives et projets.

Besoins de l'équipe

- Développement (recherche et enseignement) :
 - Programmation,
 - Parallélisation,
 - Validation et mise au point.
- Analyse :
 - Analyse de performances,
 - Exploitation des résultats.
- Exploitation :
 - Calculs de tailles intermédiaires, avant l'exploitation massive sur les machines des centres de calculs.

Besoins de l'équipe

Machine expérimentale et de développement

Calculs scalaires et parallèles

Configuration matérielle du cluster

- 8 PC homogènes :
 - Biprocesseurs Pentium III 550 MHz,
 - 512 Mo RAM,
 - Cartes GigaEthernet Intel Pro,
 - 6,4 Go DD.
- Réseau rapide :
 - Switch GigaEthernet Cisco à 8 ports.
- Frontal :
 - Biprocesseur Pentium III 800 MHz,
 - 512 Mo RAM,
 - 40 Go DD
- Switch Ecran/clavier/souris

Administration

- – Systèmes d'exploitation :
 - RedHat 6.2,
 - Frontal : maître NIS et serveur de fichiers.
- – Outils de gestion testés :
 - Installation et clonage : Kickstart, ghost.
 - Monitoring : bview, bWatch, jmon.
 - Management : mat.

Pas de solution globale satisfaisante

Beaucoup de temps perdu

Administration : Alinka Raisin

- Avantages :
 - Installation par leurs soins,
 - Administration par serveur web pour :
 - le clonage,
 - l'arrêt, le redémarrage et la reconfiguration des noeuds,
 - la définition de sous-cluster,
 - PBS, Mosix,
 - la visualisation de la charge et des jobs en cours.
 - Maintenance assurée : pas de temps perdu si il y a problème (allo Alinka !)

Administration : Alinka Raisin

- Inconvénients :
 - Coût.
 - Pas de comptabilité (CPU/User, ...) et outils d'analyse faibles.
 - Pas de multi-frontal (dans notre version actuelle) : un peu léger pour Mosix.
 - Même version de noyau pour toutes les machines.
 - Sécurité ?

Sous-clusters

1. Sous-cluster séquentiel

- MOSIX : Migration de process et répartiteur de charge.
- 4 machines dédiées aux calculs séquentiels, OPEN MP, logiciels commerciaux (matlab).

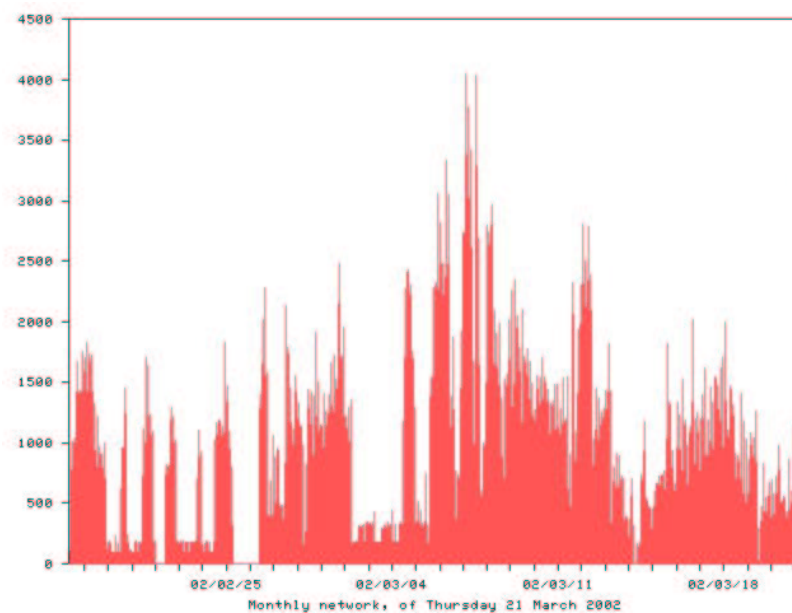
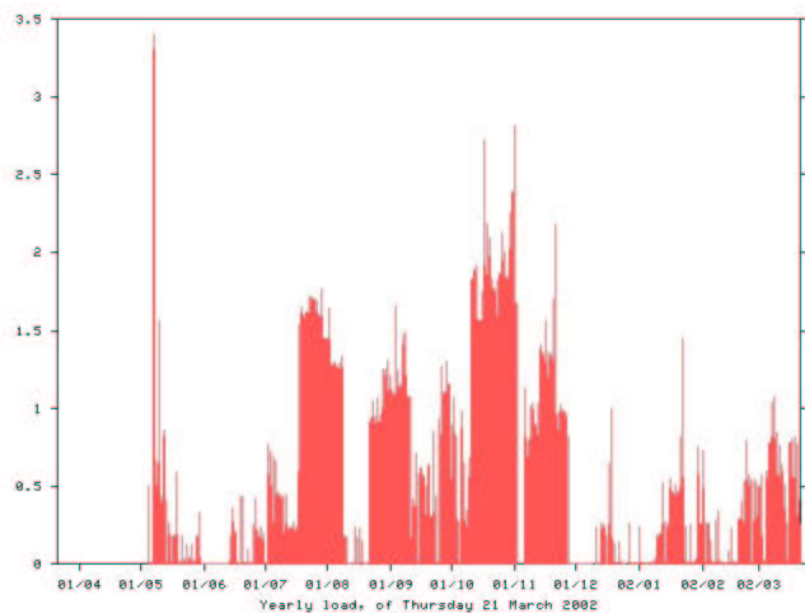
2. Sous-cluster parallèle

- 4 machines dédiées aux programmes parallèles (MPI, PVM).

Compilateurs, bibliothèques et utilitaires

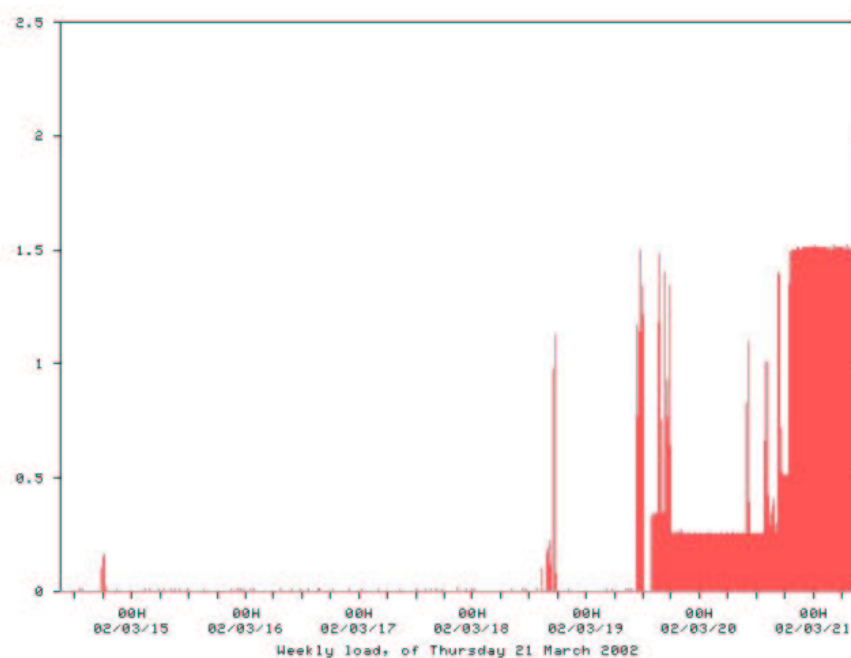
- Compilateurs :
 - GNU, NAG f95,
 - PGI (fortran, C, C++) pour OPEN MP.
- Bibliothèques :
 - MPICH,
 - PVM.
- Système de batch :
 - OPBS fourni avec Alinka Raisin

Exploitation : sous-cluster séquentiel



Charge annuelle et échanges réseau mensuels du sous-cluster séquentiel.

Exploitation : sous-cluster parallèle



Charge hebdomadaire du sous-cluster parallèle.

Exploitation

- Utilisateurs :
 - 21 users, dont 15 réguliers : pas en charge complète.

Perspectives et projets

- **Matériel** :
 - PC : de 8 x 2 processeurs vers 8 x 4 processeurs (OPEN MP)
 - Réseau : de l'ethernet vers myrinet.
- **Projet GTOOCS** :
 - Programmation Orienté Objet pour le calcul scientifique,
 - Environnement de programmation : couplage de codes / exploitation des résultats / surveillance des codes (CORBA).
- **Divers** :
 - Nombreux projets en décomposition de domaines,
 - Micromagnétisme ...
- **XtremWeb (Global Computing)** :
 - Environnement d'exécution massivement distribué sur Internet.