

Sommaire

- Forum ORAP / *SPEEDUP*
- La technologie cluster CGG pour le traitement sismique
- le gouvernement allemand soutient le calcul scientifique
- IDC : un nouveau classement de la performance des ordinateurs
- Actualités BI-ORAP
- Agenda

Forum ORAP / *SPEEDUP*

Pour la troisième année consécutive, le Forum de l'automne sera organisé en collaboration avec l'association suisse *SPEEDUP*. Son thème central sera la bio-informatique et il aura lieu **les 26 et 27 septembre à Lausanne**. Le programme est en cours de finalisation et sera disponible sur les serveurs ORAP et *SPEEDUP* dès que possible.

Inscription : Chantal.letonqueze@irisa.fr

La technologie cluster CGG pour le traitement sismique

Jean-Yves BLANC
Juin 2002

La Compagnie Générale de Géophysique (CGG) est l'un des leaders mondiaux des services géophysiques. L'ensemble des activités du groupe CGG couvre les domaines de l'acquisition de données sismiques, les services et logiciels de traitement et d'interprétation de réservoir pour les compagnies pé-

trolières. CGG est également le premier producteur mondial d'équipements géophysiques, à travers sa filiale Sercel. Basé à Paris, CGG opère mondialement sur terre comme sur mer pour acquérir des données sismiques. Les services de traitement et d'ingénierie de réservoir de CGG offrent tout le spectre de traitement et les activités afférentes de géophysique de réservoir.

Acquisition des données sismiques

L'acquisition d'images du sous-sol est un processus lourd et coûteux, principalement mis en œuvre pour la découverte de gaz et de pétrole.

Cette acquisition nécessite l'emploi de sources spécialement conçues (camions vibrateurs sur terre et *airguns* en mer) pour créer avec précision des trains d'ondes acoustiques qui se propagent dans le sol. Ces ondes se réfléchissent et se réfractent selon les impédances acoustiques des diverses couches du sous-sol qu'elles rencontrent.

De vastes ensembles d'enregistreurs (qui peuvent s'étaler sur plusieurs kilomètres carrés) captureront ces ondes plusieurs centaines de fois par seconde. Une étude typique nécessite plusieurs centaines de milliers de ces tirs. Ces sources et les récepteurs associés sont lentement et précisément déplacés à chaque tir, de façon à couvrir l'ensemble de la zone d'acquisition.

Une étude de grande taille produit plusieurs teraoctets de bruits et d'informations redondantes résultant du parcours des ondes dans les couches du sous-sol. En pratique, ces données brutes sont inutilisables et nécessitent un traitement poussé avant de pouvoir être interprétées avec succès.

Traitement des données sismiques

En appliquant aux données brutes diverses techniques de traitement du signal alliées à une puissance de calcul massive, une image finale du sous-sol peut être élaborée méthodiquement. En effet le traitement sismique est une suite d'étapes, requérant

chacune une intervention humaine pour contrôler la qualité des données et définir avec précision les paramètres à appliquer aux étapes suivantes.

Un logiciel bien conçu, facile à manipuler et au fait des dernières découvertes algorithmiques métier est critique dans ce secteur d'activités. CGG développe son propre logiciel de traitement sismique, Geocluster, utilisé dans plus de 50 sites dans le monde par ses géophysiciens. CGG vend aussi Geocluster comme une solution clé en main pour les compagnies pétrolières, ou encore comme un composant des centres dédiés opérés par CGG pour le compte de ses clients.

Caractéristiques informatiques du traitement sismique

Le traitement sismique a comme principale caractéristique de manipuler de très importantes quantités de données. Informatiquement les conséquences sont doubles : tout d'abord elles classent au niveau calcul cette industrie dans les gros consommateurs de supercalculateurs, et ensuite (ce qui est plus spécifique) fait de l'industrie pétrolière un des plus gros utilisateurs d'entrées/sorties, aussi bien au niveau volumétrie que bande passante I/O.

Par conséquent, l'adéquation des nouveaux hardwares avec les logiciels métiers est essentielle au succès de cette industrie. Celle-ci a donc été traditionnellement parmi les précurseurs au niveau de l'intégration industrielle des nouvelles technologies. Par exemple, les machines vectorielles ont été dans cette industrie presque complètement remplacées par des systèmes Numa dès le milieu des années 90.

Une équipe d'experts informatiques est chargée à CGG de la veille technologique et de la mise en place des nouvelles technologies au sein de Geocluster et dans les centres de traitement de la Compagnie. Cette équipe a décidé en 1999 d'évaluer le potentiel des clusters basés sur des machines bon marché.

La problématique cluster à CGG

Il s'agissait pour CGG de mesurer le ratio prix/performance de clusters appliqués aux divers types de travaux que l'on rencontre en traitement sismique. Compte tenu de nos spécificités et des enjeux économiques, ces tests devaient se faire dans des conditions industrielles aussi proches que possibles de nos environnements de production, et de s'efforcer de mesurer le coût de possession total (TCO, *total cost of ownership*) des clusters, ie. en prenant en compte tous les éléments : matériels, logiciels, structurels et opérationnels.

Les premières études réalisées nous ont rapidement montré que le choix le plus prometteur était, pour une utilisation industrielle lourde, de se limiter tout d'abord aux types de traitement a priori les plus favorables. En effet adapter l'architecture et l'environnement des clusters à un usage générique menaçait de se révéler un processus long et coûteux. Or les contraintes économiques de l'industrie pétrolière faisaient que nous étions plutôt à la recherche de technologies nouvelles nous permettant de diminuer rapidement nos coûts de traitement, ou même de pouvoir industrialiser de nouvelles découvertes algorithmiques.

La définition des cahiers des charges et la planification des projets a été effectuée en 2000. Avec l'aide de la R&D et des équipes applicatives, un projet en cinq phases a été établi : choix des traitements à cibler, portage des codes afférents, sélection du hardware, benchmarks en vraie grandeur et tests de production. Dès l'origine, le projet a également anticipé un déploiement pour 2001 (à valider bien sûr par la dernière phase du projet).

Avantages et inconvénients des clusters

Si nos premières expérimentations ont rapidement montré le potentiel des clusters, basés sur des processeurs Intel ou Alpha, sur certains de nos algorithmes ; elles nous ont également permis de réaliser que ces architectures sont complexes, ont une fiabilité limitée, peuvent nécessiter de profondes modifications logicielles et requièrent des moyens d'administration importants. Tous ces facteurs augmentent de façon conséquente le TCO, initialement alléchant du fait du prix modéré du hardware entrant dans la réalisation des clusters. Par conséquent, notre approche a consisté à minimiser autant que possible chacun des ces inconvénients.

Le premier élément clé réside dans le choix judicieux des algorithmes qui vont être utilisés sur les clusters. Etant donné l'écart entre la puissance de calcul des nœuds et la plupart des réseaux d'intercommunication existants, des modules de calculs intensifs ne requièrent que peu de communications inter-processus se sont donc révélés comme de très bons candidats. Les algorithmes d'imagerie avant-sommation présents dans Geocluster possèdent de telles caractéristiques. Le pré-positionnement des données sur les nœuds réduit encore ces communications. En effet ces algorithmes appliquent un nombre relativement petit de paramètres variables (les champs de vitesses modélisés) à un nombre très élevé de données fixes (les traces sismiques enregistrées sur le terrain), pouvant par conséquent être pré-positionnées sur les

nœuds, et nécessitent d'être synchronisées seulement après plusieurs heures de calculs, lorsqu'un sous-ensemble des résultats est devenu disponible.

En outre utiliser des clusters pour un usage spécialisé permet aussi de réduire la complexité de l'environnement industriel (administration, soumission de jobs, tolérance aux pannes ...) à mettre en œuvre. Cet aspect, très important dans la phase de déploiement industriel, a en effet nécessité un travail significatif, d'intégration de composants logiciels domaine public avec des éléments propriétaires nous permettant de disposer au final d'un environnement d'exploitation ad hoc.

Ensuite le choix du hardware composant ces clusters est un autre élément déterminant. Dès le début, outre l'aspect économique qui restait primordial dans ce projet, nous avons apporté un soin particulier à la simplicité, la résilience aux pannes et à l'extensibilité du design. Nos benchmarks nous ont révélé sur nos applications un gros écart aussi bien de prix/performance que de prix sec entre les nœuds bi-processeurs et les nœuds SMP avec plus de deux CPUs (pour lesquels on paie quasi-systématiquement un surcoût lié à du hardware haute disponibilité dont on peut a priori se passer dans un usage scientifique et technique). Au niveau des réseaux d'interconnection, les réseaux rapides spécialisés (tels Myrinet et SCI) se sont rapidement révélés être de piètres candidats en raison de leur coût élevé, de leur non extensibilité à des tailles de clusters très importantes (plusieurs milliers de nœuds) et de leur non-tolérance aux pannes. Par conséquent nos choix se sont portés vers des combinaisons d'ethernet 100 base T et gigabit.

Le dernier aspect à prendre en compte était l'évolution rapide du hardware composant les clusters, le rendant obsolète en quelques mois. Un leasing d'équipements sur de courtes périodes est très coûteux, et l'obsolescence rapide est un phénomène très disruptif pour de grands centres de traitement. Par conséquent un travail pointu d'ingénierie financière sur l'acquisition, la détention et l'upgrade des clusters a été nécessaire pour aboutir à un TCO aussi bas que possible.

L'évolution du projet clusters à CGG

Le choix des algorithmes à mettre en œuvre a été dicté par les contraintes économiques de notre industrie, qui nous ont poussé à réduire les prix de certaines étapes du traitement sismique et à pouvoir industrialiser de nouvelles techniques d'imagerie. Par conséquent, nous nous sommes concentrés dans un premier temps sur le portage d'un faible nombre de

modules très gourmands en calculs et en I/O (quatre en fait), avant d'envisager le portage complet de Geocluster sous Linux, préambule à sa disponibilité sur clusters.

Le portage des quatre codes initialement ciblés s'est étalé sur un peu plus d'un an, échelonné en plusieurs versions, de façon à pouvoir apprendre au fur et à mesure de leur utilisation lors des benchmarks, des tests de production, puis enfin de la mise en production lourde. Il s'agissait de codes massivement parallèles qui étaient initialement designés pour nos architectures Numa habituelles. Ces modules Geocluster ont été entièrement ré-écrits pour un maximum d'efficacité sur l'architecture cluster que nous avons choisie.

La sélection du hardware s'est effectuée en deux étapes. Tous d'abord en demandant à un large panel fournisseurs (une quinzaine avait été consultés) de répondre de façon chiffrée à notre cahier des charges. A l'issue de cette étape, trois systèmes avaient été retenus en *short list*, et les constructeurs sélectionnés nous ont alors chacun prêtés un petit cluster pour une semaine d'évaluation sur les codes ciblés avec des problèmes de tailles adaptées. A l'issue de ce choix, nous avons retenu des clusters basés sur des nœuds IA-32 fournis par Dell et IBM.

Les benchmarks en vraie grandeur avaient des buts multiples. Tout d'abord identifier clairement les points à améliorer au niveau performance : si les processeurs IA-32 peuvent fournir des performances en calculs très intéressantes, elles dépendent beaucoup de la nature de ces calculs et du degré des optimisations réalisées. Ensuite nous avons cherché à déterminer quels sont les aspects applicatifs et systèmes qui diffèrent notablement des systèmes Numa que nous utilisons. Puis nous nous sommes assurés que le degré d'intégration offert à nos géophysiciens était suffisant pour assurer une production efficace. Et enfin nous avons été à même de nous faire une première idée du degré de tolérance aux pannes à mettre en œuvre. Ces benchmarks se sont étalés sur près de six mois, échelonnés sur la période de portage des codes et du début des tests en production.

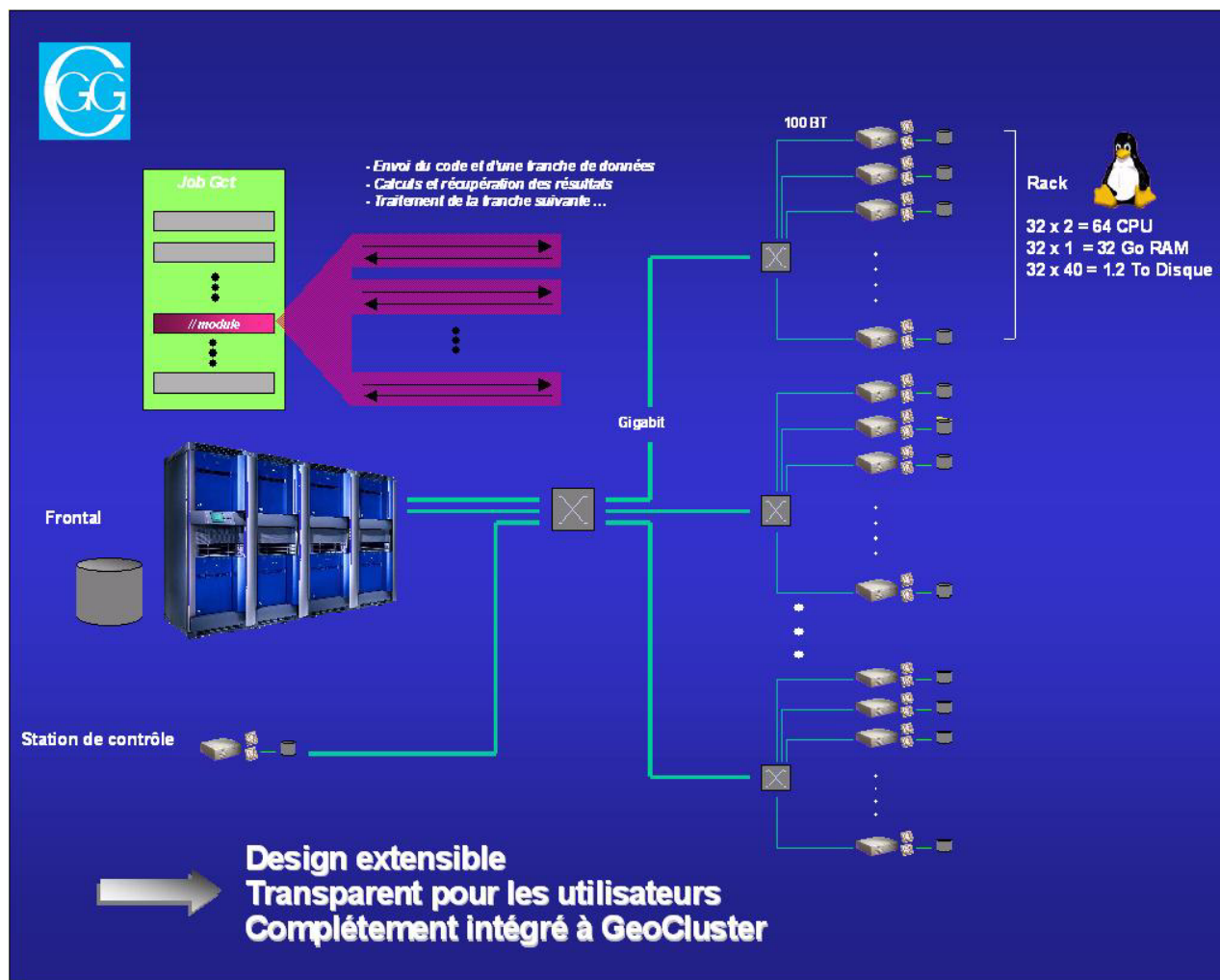
Les tests de production se sont révélés également une étape très importante qui a duré une dizaine de mois (seconde moitié 2000 et début 2001). Tout d'abord c'était l'occasion d'assembler les premiers clusters selon l'architecture que nous avons choisie. Ensuite cela nous a permis de mesurer dans un environnement industriel les performances que l'on était à même de soutenir dans la durée. Et ainsi d'affiner le ratio prix/performance des clusters sur ce type d'applications, condition préalable à la décision de

déployer – ou non – ce type d'architecture à grande échelle. Enfin il s'est agi de la première expérience en vraie grandeur en matière d'environnement d'exploitation et de tolérance aux pannes ; et cela fait une vraie différence à cette échelle.

A l'issue de ces tests de production, le management de CGG a décidé de déployer notre technologie clusters en opérations. Les premiers grands clusters ont été installés au printemps 2001 à Massy (256 CPUs), Houston (512 CPUs) et Londres (1024

CPUs). Ces chiffres ont ensuite doublé tous les six mois et CGG met actuellement en œuvre plus de 6000 CPUs, uniquement dans nos propres centres de traitement. Nous en avons aussi installé un nombre certain chez nos clients et dans les centres dédiés que nous opérons pour le compte des grandes compagnies pétrolières.

L'architecture mise en œuvre



Dans le cas des clusters, comme dans le reste du traitement sismique, le cœur du design est centré autour de l'accès aux données très volumineuses à traiter. La structure actuelle de nos centres de traitement assumant que les données à traiter sur cluster sont stockées sur de vastes disques Raid 5 (utilisés comme cache pour des HSM), les clusters sont donc attachés à des machines frontales qui leur permettent d'accéder aux données. Suivant la taille du cluster et de la frontale, un ou plusieurs liens gigabit, trunkés ou non, sont mis en œuvre. Cette frontale est aussi un élément important du mécanisme permettant de ren-

dre transparent aux utilisateurs l'usage des clusters (ils tournent en fait leurs jobs Geocluster sur la frontale).

La brique de base de notre architecture étant le rack, les clusters de grandes tailles mettent en œuvre des switchs Ethernet entièrement gigabit permettant la connexion avec les racks de calculs, ces switchs sont onéreux mais performants et tolérants aux pannes. Cet aspect est particulièrement important, la perte d'un tel switch impliquant la perte de tous les racks qui y sont attachés. Ces switchs sont placés dans un rack maître, qui contient en outre la station de contrô-

le chargée de fournir les services systèmes Linux (DHCP, NIS, DNS, *system imager* et monitoring), ainsi que les nœuds de spare et le système de concentration clavier/écran/souris.

Les racks de calculs, brique de base de notre design, sont composés de 32 nœuds bi-processeurs, d'un switch 100 baseT / double lien gigabit, et d'un système de concentration clavier/écran/souris. Les nœuds sont typiquement des bi-Pentium III à 1.46 GHz, avec 1 Go de mémoire et 73 Go de disque local. Le système de concentration clavier/écran/souris est dépendant du fournisseur, même si leurs fonctionnalités sont similaires.

L'installation logicielle du cluster est complètement automatisée à l'aide de divers composants logiciels du domaine public. A titre d'exemple, on upgrade couramment 3000 CPUs depuis la station de contrôle du cluster en 45 mn ... L'administration et le monitoring des clusters est réalisé à l'aide de plusieurs outils CGG. La soumission et le suivi des jobs est faite via le système intégré de Geocluster, le Job Manager. Le monitoring système du cluster est réalisé grâce à plusieurs outils maison permettant d'avoir une compréhension plus ou moins globale et détaillée des informations recherchées ; aucun outil domaine public ou du marché ne s'est révélé à l'usage assez pratique ou efficace à l'échelle de nos clusters. L'administration elle-même est opérée à la fois par nos outils d'installation et d'upgrade pour les opérations lourdes, et par des jeux de commandes globales plus ou moins sophistiquées pour les opérations courantes.

Considérations environnementales

L'installation de clusters de grandes tailles (plusieurs centaines, voire milliers de processeurs) impose un certain nombre de contraintes environnementales qu'il est impératif de prendre sérieusement en compte afin d'obtenir opérationnellement un système fiable.

Chaque rack de calcul dans la configuration CGG requière 4.5 kW d'alimentation électrique, presque entièrement dissipée sous forme de chaleur. Cela nécessite donc une capacité de refroidissement et de circulation d'air (de façon à éviter les points chauds) en relation avec le nombre de racks installés. Il faut également prévoir de l'espace devant et derrière les racks, en sus de leur place au sol : les racks sont désignés pour aspirer l'air froid devant (prévoir donc environ 1 m de libre) et rejeter l'air chaud sur l'arrière (1.5 m de libre avec dalles percées).

Par conséquent de nombreux aménagements

sont régulièrement menés dans les salles machines devant recevoir des clusters : ajout d'UPS (*uniterruptible power supply*), prises électriques et câblage ethernet dans les faux planchers, climatisation.

Amélioration des performances

Basé sur la charge existante, CGG utilise plus de 6000 CPUs d'une puissance pic de plus de 7 Tflops. Cette énorme puissance de calculs, fournie par Dell et IBM, combinée avec Geocluster permet à CGG de produire des travaux de grande qualité pour ses clients de l'industrie pétrolière, 24 heures par jour, 7 jours sur 7.

Les technologies clusters évoluent rapidement et sont en compétition permanente contre les gros systèmes Unix traditionnels, eux aussi objets de farouches améliorations. Actuellement le facteur limitant à une utilisation encore plus large des clusters dans notre industrie réside dans le coût de positionnement des données sur les nœuds. L'arrivée des systèmes de fichiers distribués hétérogènes reposant sur la technologie SAN (*storage area network*) va changer la donne dans un futur proche.

CGG dispose en effet en interne de toute la technologie propriétaire pour assurer sa production sur clusters avec le portage de Geocluster sous Linux et sur clusters. Le démarrage effectif de la production générique lourde sur cluster dépend maintenant, pour une part de la disponibilité technique effective de technologies SAN supportant plusieurs centaines, voire milliers, de clients, et d'autre part de l'analyse financière et des risques une fois la faisabilité technique avérée.

Jean-Yves Blanc (jyblanc@cgg.com) est architecte informatique pour la Ligne de Produits Traitement et Réservoir de CGG ; il est actuellement basé à Massy, près de Paris.

Pour plus d'informations : www.cgg.com.

Le gouvernement allemand soutient le calcul scientifique

Le Ministre Fédéral de l'Education et de la Recherche (BMBF) a donné, en juin, un aperçu des financements que son Ministère apporterait à l'informatique scientifique dans le nouveau programme qui couvre la période 2002 - 2006. Les domaines

de recherche prioritaires seront les suivants :

- Les nanotechnologies
- Les technologies de base pour les télécommunications
- Technologies et services pour l'Internet
- Systèmes logiciels (dont HPC)

Dans le domaine du calcul de haute performance, les sujets de recherche suivants ont été mis en avant : gestion des données, nouveaux procédés parallèles performants pour le data-mining dans des grandes bases de données, méthodes parallèles pour le non-numérique, bibliothèques parallèles réutilisables, ... Le financement du HPC en 2002 sera de 4 millions d'euros sur un budget total 2002 de 644 millions d'euros pour l'informatique scientifique.

(Cet article reprend un article paru dans Pri-meur, signé de Uwe Harms)

IDC : un nouveau classement de la performance des ordinateurs

Alors que la 19^{ème} édition de la liste TOP500 a été annoncée en juin dans le cadre de la conférence ISC2002 à Heidelberg (nous en ferons un compte-rendu dans le prochain Bi-ORAP), IDC (International Data Corporation) a présenté un système de classement qui prend en compte trois critères :

- la performance du processeur : SPECfp, SPECint et Linpack Rmax ;
- la mémoire : bande passante et le benchmark Stream Triad ;
- le nombre de processeurs et la mémoire totale du système ainsi que la bande passante du réseau d'interconnexion.

IDC distingue ensuite quatre segments de marché :

- les ordinateurs de très grande performance pour le calcul scientifique et technique ;
- les serveurs de calcul pour l'entreprise (prix de vente supérieur à 1 million de dollars)
- les serveurs de calcul pour une division (prix de vente entre 250.000 et 999.000 \$)
- les serveurs départementaux (moins de 250.000 \$)

Les dix systèmes les plus performants, selon les critères IDC sont :

1. NEC (5120 proc.) : Earth Simulator (Japon)
2. IBM Power3 (8192 proc.) : ASCI White (USA)

3. HP/Compaq (4096 proc.) : LANL (USA)
4. HP/Compaq (3016 proc.) : Pittsburgh SC (USA)
5. HP/Compaq (2560 proc.) : CEA/DAM (F)
6. SGI (6144 proc.) : ASCI Blue Mountain (USA)
7. IBM Power3 (3328 proc.) : NERSC (USA)
8. IBM eServer (928 proc.) : ORNL (USA)
9. NEC SX-6 (192 proc.) : ?
10. HP/Compaq (1536 proc.) : LANL (USA)

Il semble que les données de base de la liste TOP500 et celles de IDC diffèrent un peu. Mais les différences de classement, pour des configurations identiques, existent et devraient provoquer des réflexions intéressantes sur les critères permettant de comparer les performances "réelles" des ordinateurs.

<http://www.idc.com/hpc>

Actualités Bi-Orap

➔ ERRATUM

Nous avons indiqué, dans le dernier numéro de Bi-ORAP, que le noeud de calcul du futur Cray SV2 devrait avoir une performance crête supérieure à 10 GFlops. C'était largement sous-estimé car Cray nous a fait savoir que cette performance serait supérieure à 50 GFlops, la performance d'un processeur devant être très largement supérieure à 3 GFlops. Nous espérons pouvoir présenter l'architecture du SV2 dans un prochain Bi-ORAP.

➔ Nouveau cluster Linux au HPC2N

Le "High-Performance Computing Center North", situé à l'Université d'Umea (Suède), a construit un nouveau cluster Linux ayant une performance crête de 800 GFlops. Ce cluster est constitué de 240 processeurs Athlon MP2000+ (120 noeuds) interconnectés avec le système WulfKit3 de Dolphin Interconnect Solutions.

➔ Bull : FAME et Itanium2

Le 28 mai, Bull a présenté un prototype à 16 processeurs de son futur serveur SMP "FAME" ("*Flexible Architecture for Multiple Environments*") dans le cadre du "Intel Development Forum" à Munich. Cette famille de serveurs (de 2 à 16 processeurs) utilisera le futur processeur Itanium 2 (connu sous le nom de code McKinley) qui apportera un gain de performance de 50% par rapport au premier Itanium, en particulier grâce à un doublement de la taille du cache de niveau 3.

Cette famille FAME fonctionne sous Linux et GCOS et est destinée aux marchés des applications scientifiques et des "entrepôts de données". Elle correspond à la stratégie de Bull selon laquelle les systèmes d'entreprise utiliseront de plus en plus de technologies standard. Les premiers systèmes devraient être commercialisés avant la fin de cette année 2002.

➔ Cray

- Cray a signé avec la DARPA (Defense Advanced Research Projects Agency) un accord pour la première étape du programme "*High Productivity Computing Systems*". Ce programme de la DARPA a pour objectif de fournir au Département de la Défense des sauts technologiques pour la sécurité nationale et les industriels. Ce contrat de 3,6 millions de dollars doit permettre de mettre en place un programme de recherche destiné à réaliser des systèmes "trans-petaflops", ordinateurs capables d'effectuer plus de 10^5 calculs par seconde, plus faciles à programmer et plus tolérants aux pannes.
- Cray a reçu pour 19 millions de dollars de commandes du gouvernement américain (dont le Département de la Défense) ; elles portent sur les futurs systèmes SV2 qui devraient être livrés en 2003.
- ARSC (The Arctic Region Supercomputing Center) a commandé un système SX-6 de 8 processeurs. Rappelons que l'accord entre Cray et NEC prévoit que Cray a les droits exclusifs de distribuer les ordinateurs des séries SX en Amérique du nord.

➔ IBM

- Le NOAA (National Oceanic & Atmospheric Administration) a signé un contrat avec IBM selon lequel ce dernier lui fournira un système comprenant 2752 processeurs qui renforcera ses moyens de prévision du temps, des inondations et du climat.

<http://www.cpc.ncep.noaa.gov/>

- Le consortium universitaire italien CINECA (Milan) a inauguré, fin mai, son nouveau superordinateur IBM SP Power 4. Avec 512 processeurs, ce système a une performance théorique de 3 GFlops ; il se situe en 32^{ème} position dans le dernier TOP500, et en 4^{ème} place en Europe.

<http://www.cineca.it>

➔ NEC

- Le UK Met Office a choisi un SX-6 comme future machine de production. Ce système, comprenant 30 noeuds, sera installé en mars 2004 et délivrera une puissance 6 fois supérieure à la puissance combinée des deux Cray T3E en exploitation actuellement. Un an plus tard, l'ajout de nouveaux noeuds permettra de doubler encore la puissance disponible.

<http://www.met-office.gov.uk/corporate/pressoffice/pr20020617.html>

- L'Institut Météorologique Danois procède à une augmentation de la configuration de son système SX-6. En phase finale, la configuration sera de 48 processeurs, 224 Go de mémoire, 4 To sur disques, et une performance crête de 460 GFlops.

➔ SGI

La NASA augmente la configuration de son principal superordinateur : un SGI Origin 3800 à 1024 processeurs. Les processeurs ont été remplacés par des MIPS R14000A à 600 MHz (augmentation de puissance de 37%). La puissance de la nouvelle configuration est maintenant de 1,23 TFlops. La "*NASA Advanced Supercomputing Division*" pilote un projet de construction du "*NASA's Information Power Grid*", un réseau de superordinateurs distribués géographiquement, de grandes bases de données et d'instruments scientifiques accessibles de toute la communauté scientifique de la NASA.

Agenda

- 10 juillet : **CMPP'2002** : 3rd International Workshop on "Constructive Methods for Parallel Programming" (Dagstuhl, Allemagne)
- 20 juillet : **NeSC** Workshop on Applications and Testbeds on the Grid (Glasgow, Ecosse)
- 21 au 24 juillet : **GGF5** : Global Grid Forum (Edinburgh, Ecosse)
- 24 au 26 juillet : **HPDC-11** : International Symposium on Grid Computing (Edinburgh, Ecosse)
- 5 au 7 août : **WOMPAT 2002** : Workshop on OpenMP Applications and Tools (Fairbanks, Alaska)
- 18 au 21 août : **ICPP 2002** : International Conference on Parallel Processing (Vancouver, Canada)

- 18 au 21 août : **MSA 2002** : International Workshop on Metacomputing Systems and Applications (Vancouver, Canada)
- 25 au 30 août : **TCS 2002** : International Conference on Theoretical Computer Science (Montréal, Canada)
- 26 au 28 août : **Pervasive 2002** : International Conference on Pervasive Computing (Zurich, Suisse)
- 27 au 30 août : **Euro-Par 2002** : (Paderborn, Allemagne)
- 9 au 10 septembre : **EGPGV02** : Fourth Eurographics Workshop on Parallel Graphics Visualization (Blaubeuren, Allemagne)
- 22 au 25 septembre : **PACT 2002** : 11th International Conference on Parallel Architectures and Compilation Techniques (Charlottesville, Va, Etats-Unis)
- 23 au 26 septembre : **Cluster 2002** : IEEE fourth International Conference on Cluster Computing (Chicago, Il, Etats-Unis)
- 24 au 26 septembre : **iGrid 2002** Conference (Amsterdam, Pays-Bas)
- 29 septembre au 2 octobre : **EuroPVMMPI 2002** : 9th European PVMMPI User's Group Meeting (Linz, Autriche)
- 29 septembre au 2 octobre : **DAPSYS 2002** : 4th Austrian-Hungarian Workshop on Distributed and Parallel Systems (Linz, Autriche)
- 7 au 9 octobre : **EMSOFT'02** : Workshop on Embedded Software Development (Grenoble)
- 13 octobre : International Workshop on reliable peer-to-peer distributed systems (Osaka, Japon)
- 21 au 22 octobre : **EuroPAM 2002** : European Conference and Exhibition on Numerical Simulation for Virtual Engineering (Antibes)
- 21 au 23 octobre : **HeteroPar'02** : International Workshop on Algorithms and Tools for Parallel Computing on Heterogeneous Clusters (Moscou, Russie)
- 23 au 25 octobre : 3rd International Conference on **Linux Clusters** (St Petersburg, Fl, Etats-Unis)
- 27 octobre au 1er novembre : **IEEE Visualization 2000** (Boston, Etats-Unis)
- 28 au 29 octobre : **VolVis'02** : Symposium on Volume Visualization and Graphics (Boston, Etats-Unis)
- 28 au 30 octobre : **DISC 2002** : 16th International Symposium on Distributed Computing (Toulouse)
- 3 au 5 novembre : ACM 2002 Java Grande **ISCOPE** : International Symposium on Computing in Object-Oriented Parallel Environments (Seattle, Etats-Unis)
- 16 au 22 novembre : **SC2002** : High Performance Computing and Networking (Baltimore, Etats-Unis)
- 18 novembre : **Gridcomputing** : 3rd International Workshop on Grid Computing (Baltimore, Etats-Unis)

Des informations complémentaires, en particulier les adresses http de ces manifestations, sont disponibles sur le serveur Web d'ORAP.

Appel à informations

Le contenu de BI-ORAP dépend, pour partie, de ses lecteurs ! N'hésitez pas à nous communiquer toute information concernant vos activités dans le domaine du calcul de haute performance : installations de matériel, expérimentations de nouvelles technologies, applications, organisation de manifestations, formations, etc.

Merci d'adresser ces informations au secrétariat d'ORAP ou directement à Delhaye@irisa.fr



HOISE - Europe On-line Information Service

PRIMEUR ! - *Advancing European Technology Frontiers*

<http://www.hoise.com/primeur/>

Organisation Associative du Parallélisme
Structure de collaboration créée par
le CEA, le CNRS et l'INRIA.

Secrétariat : chantal.letonqueze@irisa.fr
 IRISA, campus de Beaulieu, 35042 Rennes cedex
 Tél : 02.99.84.75.33, Fax : 02.99.84.74.99
<http://www.irisa.fr/orap>