



FAME2

Le projet FAME2 a pour ambition de doter les serveurs de la gamme FAME2G (développés par ailleurs) de toutes les caractéristiques (matérielles, logicielles, applicatives) nécessaires aux très grands systèmes de calcul intensif et de traitement de bases de données multimédia. Il s'agit de garantir que l'architecture proposée prenne en compte, dans sa conception, les critères d'extensibilité matérielle (nombre de processeurs variables par exemple) et de versatilité des logiciels d'exploitation requis.

Le projet fait partie des initiatives à vocation industrielle de Ter@tec, au sein du pôle de compétitivité francilien System@tic. Il rassemble des acteurs du monde de la recherche et de l'industrie, de jeunes entreprises innovantes et des utilisateurs expérimentés.

Les évolutions technologiques du calcul intensif permettent à de nouveaux acteurs de venir sur ce marché, mais à condition d'apporter des produits basés sur des technologies très compétitives et de constituer un réseau de compétences complémentaires.

OBJECTIF DU PROJET

La durée de vie des grands codes de calcul dépasse celle des systèmes informatiques.

D'où la volonté de s'appuyer sur un environnement pérenne pour le développement des codes de calculs. Cela se traduit par l'utilisation croissante de logiciels libres (open source), ou de produits d'éditeurs indépendants ayant établi un standard de facto, qui peuvent être partagés par de larges communautés d'utilisateurs et de développeurs.

Une telle tendance à la standardisation se retrouve pour le matériel. L'utilisation de composants commerciaux de grande diffusion est en effet de plus en plus répandue, et conduit les constructeurs informatiques à n'effectuer que les développements strictement nécessaires à l'introduction rapide de nouveaux produits sur le marché.

Un deuxième tournant technologique important tient à l'énorme masse de données brassées par les applications du calcul intensif. Un tel besoin s'est accompagné de l'émergence de communautés actives dans le développement d'intergiciels spécifiques de ce domaine.

Enfin, la troisième tendance est l'intégration des caractéristiques multimédias propres aux applications de la simulation haute performance.

Le projet FAME2 tire parti de ces évolutions et vise la réalisation d'un multiprocesseur à mémoire partagée, qui utilisera en 2008 la prochaine génération de processeurs de la société Intel. Comme on cherche à limiter l'augmentation de la fréquence des processeurs au profit de la maîtrise de la consommation d'énergie, un des enjeux du projet est de parvenir à exploiter efficacement un haut degré de parallélisme au sein des serveurs. L'utilisation de processeurs multicœurs, avec plusieurs cœurs d'exécution, conduira ainsi à intégrer plusieurs centaines d'unités de calcul au sein d'un même serveur. Un tel choix exige des compromis et des optimisations tant du point de vue de l'architecture matérielle (hiérarchie et cohérence des informations en mémoire) que de l'architecture logicielle (gestion du parallélisme, environnement d'exploitation et de développement).

Partenaires

Bull
CAPS-entreprise
CEA
DASSAULT AVIATION
ECP – MAS
IFP
ILOG
INRIA-IRISA
INT-ARTEMIS
NewPhenix
RESONATE MP4
Université d'Evry / IBISC
UVSQ-ITACA

Durée du projet : 18 mois

Effort : 100 personnes.an

Contacts :

Claude Camozzi
Bull
claude.camozzi@bull.net

Pierre Leca
CEA
pierre.leca@cea.fr



ENJEUX

Assurer le leadership de l'ensemble des partenaires dans l'utilisation des systèmes Petaflopiques.

Disposer d'une offre de serveurs adaptés à la réalisation de très grands systèmes. Anticiper la mise à disposition des évolutions nécessaires des logiciels de base, des outils d'optimisation et des bibliothèques sur ces nouveaux serveurs.

Mettre progressivement à disposition, sur la gamme NovaScale et ses successeurs, un portefeuille d'applications variées (simulation, grandes bases de données, biologie intégrative) et optimisées.

PROGRAMME DE TRAVAIL

Pour relever de tels défis, le projet FAME2 s'est structuré autour de quatre grands axes.

Le premier axe concerne l'étude de l'architecture logicielle et portera plus particulièrement sur l'évolution du noyau Linux, la gestion des processus de poids légers en exécution ou « threads », et la génération de code pour les processeurs multicœurs.

Le deuxième axe concerne l'architecture matérielle; il prône une attention particulière aux architectures NUMA (Non Uniform Memory Access) régissant l'accès aux différentes régions de mémoire, ainsi qu'aux entrées-sorties.

Le troisième axe tient à la gestion des données : outre les applications bien connues du calcul intensif dans les secteurs de l'énergie et de l'aéronautique, la gestion des grandes bases de données au format XML et des flux multimédias a en effet été sélectionnée comme particulièrement représentative des applications émergentes.

Enfin, un quatrième axe consiste à prendre en compte les contraintes d'intégration du serveur dans un centre de calcul, et la sécurité informatique.

Prévu sur une durée de 18 mois, le projet s'est fixé plusieurs jalons : les points clés et les éléments de base de l'architecture, la démonstration de la faisabilité et de l'efficacité d'une architecture d'interconnexion interne homogène, puis la fourniture d'un démonstrateur pour l'émulation d'un serveur exploitant les microprocesseurs les plus récents avec des niveaux de parallélisme différenciés.

L'étape suivante sera de réaliser un environnement de développement et d'optimisation pour prendre en compte cette nouvelle architecture matérielle, afin de la mettre à disposition d'applications typiques du calcul intensif et de prouver l'efficacité du serveur pour ces applications.

La dernière étape sera alors de proposer des solutions complètes pour l'accès aux très grandes bases de données.

RESULTATS ATTENDUS

- Preuve de la pertinence de l'architecture du nœud de calcul
- Accélération de la disponibilité d'OS et d'outils logiciels efficaces
- Très grandes bases de données multimédias (jusqu'à des centaines de Terabytes), grand nombre d'accès
- Résolutions des grands problèmes combinatoires : cryptologie, affectations quadratiques, ...
- Biologie intégrative (du génome à l'homme, ...)
- Modélisation et simulation numérique : changement d'échelle
- Résolution de grands challenges : simulations à plusieurs dizaines de millions de mailles (champ pétrolière, écoulements d'air autour d'un avion), base de données XML de 10 Terabytes

Communications et événementiels : publications, convention annuelle, séminaires...

SYSTEM@TIC

PARIS-REGION

Le Pôle de Compétitivité se fixe pour objectif de devenir une référence mondiale dans le domaine des systèmes électroniques complexes, c'est-à-dire ceux qui assurent le pilotage, la supervision la régulation ou le contrôle des diverses installations sophistiquées utiles à la quasi-totalité des domaines d'activité du pays : le commerce, la finance, la santé, la sécurité, l'énergie, les transports...

A cette fin, le pôle présidé par Thalès a décidé de regrouper les entreprises spécialisées dans la conception, le fonctionnement et la maintenance de tels systèmes : Alcatel, Bull, Dassault Aviation, EADS, France Télécom, Motorola, Renault et Sagem...

Implantées pour l'essentiel dans la partie ouest de la région Ile-de-France, elles associeront à l'ensemble de leurs travaux plusieurs laboratoires du CEA, du CNRS et de l'Inria ainsi que trois grandes écoles : Polytechnique, Centrale et Supélec. Principaux secteurs d'activité visés : les transports et les télécommunications.

Plus d'informations :

www.systematic-paris-region.org