

# Piria : le calcul intensif au service de l'image

Le nouveau moteur permet d'effectuer une recherche de **3,7 millions d'images par seconde**, ce qui est 5 fois plus rapide que précédemment.

Cette performance record a été obtenue [sur un supercalculateur conçu et fourni par Bull](#), en utilisant le logiciel de recherche multimédia spécialement développé par le CEA LIST dans le cadre du projet FAME2 .

Il ouvre la voie à un vaste champ applicatif allant de la veille stratégique à la comparaison d'images médicales, des « fouilles » de données sur Internet au commerce électronique ou à la gestion de contenu. Aujourd'hui, les moteurs sur Internet effectuent des recherches d'images uniquement à partir de leur description textuelle (noms, légende).

En effectuant les recherches à partir de l'analyse du contenu des images, le moteur de recherche Piria développé par le CEA, apporte une solution beaucoup plus puissante, ouvrant la voie à un vaste champ applicatif : de la veille stratégique à la comparaison d'images médicales, des « fouilles » de données sur Internet au commerce électronique ou à la gestion de contenu.

Le principe de la recherche d'images par le contenu **est de calculer pour chaque image de la base une signature visuelle ou codée et de regrouper ces signatures dans un index**. La requête, matérialisée par une image, donne lieu à une réponse sous forme d'images similaires. Ces techniques de recherche par le contenu, qui analysent en premier lieu la valeur des pixels, sont intrinsèquement très consommatrices en calcul.

## **Un besoin impératif de la puissance de calcul**

Dans le cadre du projet FAME2 auquel participe le CEA, les chercheurs ont pu accéder à de puissants moyens de calcul haute performance pour tester l'application de recherche d'images Piria dans une base de données de grande dimension.

Le test a nécessité l'adaptation du code du moteur Piria à l'architecture parallèle du supercalculateur développé par Bull (88 cœurs de processeurs Intel Itanium et 50 téra-octets de disque), permettant d'intégrer la base de **22 millions d'images d'un volume de 2,9 téra-octets**.

Les résultats de ce développement ont été présentés durant l'été 2007 : les 22 millions d'images ont été indexées en moins d'une semaine de calcul, en exploitant 48 cœurs de processeurs Intel Itanium du supercalculateur.

Le moteur Piria permet d'effectuer en 6 secondes une recherche parmi 22 millions d'images, au lieu de 15 secondes pour une recherche d'image parmi 11 millions avec le système Cortina, système de recherche d'images par le contenu accessible par Internet et développé par l'Université de Californie de Santa Barbara.

Cette épreuve était l'un des grands défis que le projet FAME2 se proposait de relever.