

# DAAD

Techniques de gestion de données distribuées à grande échelle qui tiennent compte de l'autonomie, de la dynamique et de l'hétérogénéité des applications et des bases de données

Técnicas de computação de dados distribuídos à grande escala, que respeitam a autonomia, dinâmica e heterogeneidade das aplicações e dos bancos de dados

Patrick Valduriez (INRIA et LINA, Nantes) [www.sciences.univ-nantes.fr/lina/ATLAS/Members/Valduriez.html](http://www.sciences.univ-nantes.fr/lina/ATLAS/Members/Valduriez.html)

Vidal Martins (PUCPR et LINA, Nantes)

Esther Pacitti (INRIA et LINA, Nantes)

Marta Lima de Queirós Mattoso (UFRJ-COPPE, Rio de Janeiro) [www.cos.ufrj.br/~marta/](http://www.cos.ufrj.br/~marta/)

Alexandre de Assis Bento Lima (UFRJ-COPPE, Rio de Janeiro)

Eduardo Laber (PUC-Rio, Rio de Janeiro)

Projet supporté par CAPES, CNPQ, COFECUB et INRIA (2003-2005)

## Descriptif

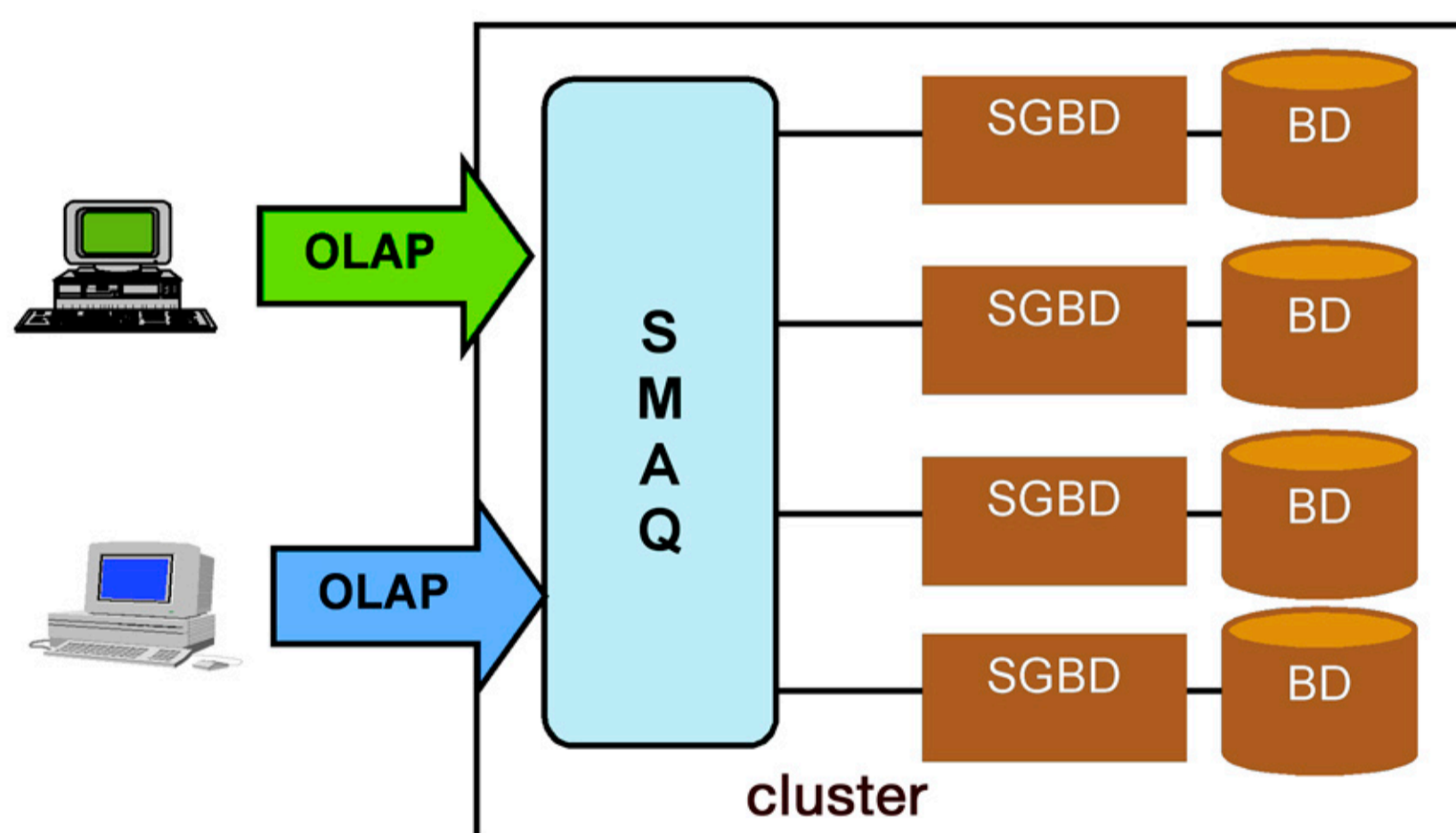
Dans un système de gestion de données distribué à large échelle (cluster, grid), le problème principal concerne l'efficacité du traitement des requêtes qui doit supporter l'autonomie, la dynamique et de l'hétérogénéité des applications et des bases de données. Pour ce faire, nous nous appuyons sur une architecture pair-à-pair (P2P) qui respecte l'autonomie des nœuds et nous exploitons le parallélisme et l'équilibrage de charge pour obtenir de hautes performances.

Num sistema de gerência de dados distribuídos à larga escala (cluster, grid), o problema principal é a eficiência em processamento de consultas que devem respeitar a autonomia, dinâmica e heterogeneidade das aplicações e dos bancos de dados. Para isso, utilizamos uma arquitetura par-a-par (P2P) que respeita a autonomia de cada nó e explora o paralelismo e o balanceamento de carga da arquitetura para obter alta performance.

## Résultats obtenus

Définition du middleware SMAQ (SMashing Queries) pour clusters de bases de données autonomes et répliquées ; techniques de traitement de requêtes décisionnelles (OLAP) distribuées ; validation avec un cluster de 64 nœuds et le système PostgreSQL.

Definição do middleware SMAQ (SMashing Queries) para clusters de bancos de dados autônomos e replicados ; técnicas de processamento de consultas (OLAP) distribuídas ; validação com um cluster de 64 nodos e o sistema PostgreSQL.



## Publications communes

C. Furtado, A. Lima, E. Pacitti, P. Valduriez, M. Mattoso.

**Physical and Virtual Partitioning in OLAP Database Clusters.** Int. Symp. on Computer Architecture and High Performance Computing, Rio de Janeiro, Brazil, 2005

R. Akbarinia, V. Martins, E. Pacitti, P. Valduriez.

**Replication and Query Processing in the APPA Data Management System.** Distributed Data & Structures 6 (WDAS): Records of the 6th International Meeting (Lausanne, Switzerland), Waterloo: Carleton Scientific, 2004

A. Lima, M. Mattoso, P. Valduriez.

**Adaptive Virtual Partitioning for OLAP Query Processing in a Database Cluster.** Brazilian Symposium on Databases (SBB), Brasilia, Brazil, 2004

A. Lima, M. Mattoso, P. Valduriez.

**OLAP Query Processing in a Database Cluster.** Int. Conf. on Parallel and Distributed Computing (Euro-Par 2004), Pisa, Italy, 2004

F. Porto, E. Laber, P. Valduriez.

**Cherry Picking: a Semantic Query Processing Strategy for the Evaluation of Expensive Predicates.** Brazilian Symposium on Databases (SBB), Manaus, Brazil, 2003