Managing a DB2 Parallel Edition Database

Gilles Fecteau

IBM CANADA Laboratory
1150 Eglinton Avenue East
North York, Ontario, Canada

Introduction

In this presentation, the author will describe how to manage a DB2 Parallel Edition (PE) Database. DB2 PE belongs to the IBM family of open DB2 client/server database products including DB2/6000, DB2/2, DB2 for HP-UX, and DB2 for the Solaris Operating Environment. DB2 PE employs a **shared nothing** architecture in which the database system consists of a set of independent **logical database nodes**. Each logical node represents a collection of system resources including, processes, main memory, disk storage, and communications, managed by an independent database manager. The logical nodes use message passing to exchange data with each other. Tables are partitioned across nodes using a hash partitioning strategy. The cost-based parallel query optimizer takes table partitioning information into account when generating parallel plans for execution by the run time system.

In other **share nothing** databases, the issue of skew imbalance and the management of large number of nodes is perceived as significantly more complex than in a *Share Disk* or a *Symetric Multi-processor (SMP)* environment.

This presentation will focus on the tools and utilities provided by DB2 PE to alleviate this problem and make it easier to manage a large database in the **share nothing** environment.

Storage Model

A DB2 PE system can be configured to contain one or more physical nodes. Each node need not be of the same power, some can be uni-processors, some SMP.

That DB2 PE system can house one or more databases, each database containing multiple tables (of different sizes).

In order to provide an appropriate management of the database schema, facilities are provided to divide a database in logical group of tables. Each group (called NODEGROUP) can in turn be housed on a subset of logical nodes. Logical nodes are themselves located on physical nodes with a many to one relationship (one or more logical node per physical node).

The presentation will explain how the physical database design can take into consideration the plurality of NODEGROUPs to achieve the correct level of de-clustering for each tables.

Skew correction

Each NODEGROUP has associated with it a partition map used for skew correction. Tables are horizontally partitioned using a hash algorithm applied against the partition key (user defined or defaulted by the system). The hashing is done over a large (4096) number of partitions. The partition map is then used to regroup partitions on logical nodes.

The presentation will show how an ANALYZE function is used to derive the best partition map to correct for skew in either data volume or transaction volume.

Dataload

The presentation will describe the high speed load techniques available with DB2 Parallel Edition.

Redistribution

The presentation will describe the Redistribution command.

Populating de-normalized tables

In Decision support, user frequently create de-normalized tables and populate them from complex SQL queries. This presentation will cover how this can be done in parallel from either SQL statements or interfaces to the fast load utility.

Parallel backup and recovery

Finally the presentation will cover the parallel backup and recovery mechanisms of DB2 Parallel Edition.