DIRECTV® and Oracle Rdb: The Challenge of VLDB Transaction Processing

William L. Gettys, Principal Engineer
Oracle Rdb Engineering
Weaverville, North Carolina
gettys@nova.enet.dec.com

System Description

The end of the Cold War has brought significant changes for GM Hughes Electronics, one of the world's leading satellite and defense electronics companies. Their response to the loss of defense revenue was to marry their satellite communications expertise with the rapidly expanding entertainment industry to produce DIRECTV®, the first all-digital direct broadcast satellite (DBS) service in the United States.

For years, customers in rural areas have used large, unsightly satellite dishes to receive television programming. The cost, size and complexity of these systems has limited their appeal.

Hughes has now launched two geosynchronous satellites that use higher powered transmitters to send streams of compressed digital data to 18 inch antennas that can be mounted inauspiciously. A specialized video processor decompresses the signal and displays it on the consumer's television with better-than-broadcast quality audio and video. The programming offered includes a number of cable-like television broadcast services, music and scores of offerings of pay-per-view movies, sports and special events. The low cost and very high quality of the system have created a demand for receivers that significantly outstrips the current supply. Consumer frustration with cable television companies has added to the demand.

Business Arrangements

At the outset, DIRECTV® faced a number of daunting tasks: building and launching the satellites, developing advanced digital compression technology, working with vendors to develop the receivers, contracting with suppliers of programming and learning to market satellite television on an unprecedented scale to a massive new audience. On top of all this, a system for billing, collections and customer service had to be developed.

Prudently, DIRECTV® chose to partner with others for much of this project. Digital Equipment Corporation was named the prime contractor for the billing system and DBS Systems Corporation of Charlotte, North Carolina wrote the billing and customer service software. They chose Oracle Rdb as the database engine for the billing system and other systems used in the project.

In an arrangement that resembles an information utility, Digital owns and manages the hardware and software for the billing system. Digital is compensated only for enrolling subscribers, sending them bills and maintaining their records in the database.

SIGMOD ' 95,San Jose , CA USA

Other contractors in various locations were chosen for other roles in the project. An Rdb based data warehouse is used to consolidate information from the various business partners and to support the marketing of this unique new service.

Requirements and Challenges

The complexity of the business arrangements and the desire to differentiate DIRECTV® through excellent customer service made for stringent requirements of the database system.

Database Size

Consumer research and market analysis pointed to very significant interest in DIRECTV® among television and electronics aficionados. Early estimates indicated an ultimate database size of at least 500 Gbytes for the billing system.

Availability

Hughes offers 24 hour customer service for DIRECTV®. Since most customer service activities require read/write access to the billing system, Hughes requires 99.9% availability with no planned outages for maintenance. This means less than 9 hours of total down-time per year. Individual service interruptions cannot exceed 15 minutes.

Performance

To minimize customer frustration, DIRECTV® requires that most on-line transactions complete in less than one second as measured at the customer support representative's workstation. Some of these workstations are located thousands of miles from the billing system data center.

Data Integrity

Obviously, the database must be secured from hardware and software failure by backing it up to tape. The backup operation must proceed without interrupting read or write access or severely reducing response time.

Disaster Recovery

In the event of a disaster that destroys the billing system computers, operations must resume within 24 hours.

Reporting

A large number of pre-defined reports are generated each night to summarize the previous day's activities. These are distributed to appropriate managers by 8 am each day. In addition, the contract provides for ad hoc access to the entire collection of information stored in the billing system.

Data Warehouse

A data warehouse combining detailed information from all systems is used to study trends and effects of marketing programs.

Database Features

A number of design techniques and features of Rdb were used to meet these requirements.

Shadow database

The requirement that the database never be unavailable for more than 15 minutes is a particularly difficult challenge. Digital's OpenVMS clusters provide for rapid failover so that if a single node in the cluster fails, processing can continue on the remaining nodes. Digital's transaction processing monitor, ACMS, makes application failover particularly easy within a cluster. Recovering from the failure of an entire cluster, however, cannot be accomplished within the 15 minute requirement. To provide for this eventuality and to meet the disaster recovery requirement, a shadow cluster and databases are maintained in a separate location. A fiber optic network link connects the two clusters.

When the systems were initially installed, a tape backup of the primary database was restored on the secondary cluster. Every hour or so, the after image journal of the primary database is backed up and applied to the secondary database. In the event of a complete failure, the active journal from the primary database would be applied to the secondary database. This process allows the applications to resume on the failover database within the 15 minute time window without losing any committed transactions.

Even a disaster that results in the complete loss of the primary data center allows the secondary cluster to take over processing within 15 minutes and only a small number of transactions would be lost. (Unlike funds transfer systems, the cash value of each transaction is small, but there are very many of them.)

Reporting Database

Hughes requires that reports be consistent with each other and reflect the complete set of transactions for the previous day. This could be accomplished by adding date predicates to the queries, but to reduce the impact of reports and ad hoc queries on the production database, a second reporting database is used to isolate the reporting workload.

Originally, the reporting database was produced by backing up the primary database each midnight and restoring a copy in a separate location. Building additional indexes makes creating the reports feasable.

To speed the process even more, the applications have now been modified to log all changes to a series of daily change tables. Many reports can now be generated directly from the change tables. At midnight each night, the collection of changes is applied to the reporting database by multiple, parallel loading processes. The reporting database can now have a completely different physical design from the on-line database to facilitate faster report generation and ad hoc query performance.

Table Partitioning

One of the most significant challenges of the DIRECTV® requirements is caused by the newness of the business. At the time that the service was first offered to the public, there were only a few hundred customers in the database. This same system has to be capable of growing to accommodate millions of customers and two years of their history without an interruption of service. The cost of deploying a database pre-sized to accommodate the anticipated growth would have been prohibitive.

Rdb's ability to partition tables and their indexes across multiple disk drives allows the designers to spread the IO activity evenly and to allow for gradual growth. Version 6.0 allows on-line addition of storage areas and open ended modification of storage maps and indexes to meet the needs of DIRECTV®. With less than five minutes down time, we can add additional areas on new disk drives and revise the storage maps to accommodate months of anticipated growth. The benefit of timing capital purchases of hardware to the revenue generated by additional business is obvious.

History Database

Coincidental record clustering is used to place all transaction rows for a single customer on the same database page even though they exist in multiple tables. The problem with this technique is that the page size must be quite large to accommodate the transactions of the most active customers. Pages assigned by the hashing algorithm to less active customers contain large amounts of empty space.

A compromise was developed: all transaction information for the current and prior billing periods are kept in one database with relatively small buffer and page sizes. Older transactions are moved to a history database which has a mix of page sizes. Customers who show little activity have their transactions moved to storage areas with small page sizes. The most active customers' transactions are stored in areas with large page sizes. This partitioning of tables is done automatically by Rdb based on an activity level code supplied by the application.

Ordered hashing is used to distribute the transaction rows more uniformly than would occur with random hashing.

Multiple Billing Centers

It is possible that the database size and processing load can be accommodated in a single large OpenVMS cluster with new hardware now being developed. To be sure that even exceedingly rapid growth could be accommodated, the database and application were designed to be partitioned across multiple databases on multiple clusters.

Enhancements

Some enhancements to Rdb version 6.0 were added specifically for DIRECTV®. A few more will be added to future releases.

After Image Log Shipping

The next major release of Oracle Rdb will automate the process of keeping the shadow databases current. Rdb currently uses an AIJ logging server process to write after image journal records for all users. This process will be enhanced to automatically transfer journal records at high speed to a receiving process that will apply the changes to the shadow database. There is a trade off between execution speed and transaction integrity. Users of this feature will be able to choose from three integrity options.

Data Loading Improvements

One of the enhancements planned for the next major release of Oracle Rdb is automatic parallelization of bulk data loads. Prototypes indicates that order of magnitude performance improvements can be expected. This effort will benefit both the reporting database and the data warehouse.

Data Distribution

Future engineering work on replication will likely simplify the process of populating the reporting database. This effort is a cooperative venture between Oracle's New England Development Center and the engineers at Company headquarters at Redwood Shores, California.