

# An Ultra Highly Available DBMS

Svein-Olaf Hvasshovd, Svein Erik Bratsberg, Øystein Torbjørnsen  
Clustra AS, 7485 Trondheim, Norway  
svein-olaf.hvasshovd@clustra.com

## 1 Introduction

Mainstream database management systems are designed for general use. Various compromises have been done to satisfy the most common users and the largest markets. One application which has been mostly ignored, is the network equipment made for the telco operators.

The equipment used in the telco industry has requirements differing from traditional database applications with respect to availability and real-time performance. This has caused the telco manufacturers to develop their own hardware, operating systems and programming languages.

The Internet market is starting to demand the same reliability as the telco industry has provided for decades, thus the market for such equipment is growing rapidly. The telco industry and Internet world are melting together as they are fighting each other for market shares. The growth of the Internet increases the market for database management solutions based on standard computer equipment, but designed for the *requirements* of the telco market:

**Availability** should be at least 99.999%, corresponding to average 5 minutes unavailability per year including all maintenance work, e.g. SW and HW upgrades.

**Real-time response** times should be in range of one to ten milliseconds for a specified percent of the transactions (say 98%).

**Throughput rates** In average thousand subscribers generate one call per second. The large network components today serve one million subscribers,

---

*Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the VLDB copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Very Large Data Base Endowment. To copy otherwise, or to republish, requires a fee and/or special permission from the Endowment.*

Proceedings of the 26th VLDB Conference,  
Cairo, Egypt, 2000.

thus handling one thousand calls per second. In average a call can lead to from one to ten transactions.

**Scalability** Typical for services today is that they start out as small trial services. If it becomes a success, it grows rapidly to huge dimensions. The service provider does not want to buy big and expensive equipment before he knows the service will be successful. On the other hand, he does not want to throw out the initial investment when upgrading to a system handling the growing load.

**Open interfaces** The DBMS must use open interfaces as SQL, ODBC and Java to ensure interoperability with other data sources and applications.

### Run on commodity HW/SW

For today's mainstream database management systems several of these requirements are not satisfied. Although claiming high availability, the systems can only achieve an availability which is one or two orders of magnitude worse than the requirements. The main obstacles for this are system maintenance and long takeover times in the case of failures.

Another requirement not satisfied by mainstream systems is the response times for update transactions. They all synchronous write the log to disk before committing, and combined with a group commit strategy, transaction response times are highly variable and in the range of 50 milliseconds or higher for update transactions.

## 2 The Clustra Parallel Data Server

This presentation covers basic software and hardware architecture for a high availability DBMS, including interconnect and disk solutions. Basic transaction execution strategies for parallel systems will be discussed.

Special focus will be made on fault-tolerance mechanisms, including replication, take-over, recovery and repair. We will also go through methods for on-line system maintenance, including on-line backup and restore, software and hardware upgrades and schema changes.