

# **MYSQL CLUSTER USING DOLPHIN EXPRESS:**

**DELIVERING 'REAL-TIME' RESPONSE TO THE  
DATABASE MARKET**

**Mikael Ronström, Marianne Ronström  
Open Source Consultant  
iClaustron AB**



**Dolphin Interconnect Solutions ASA**

Reg. No. 984 861 060 in the Norwegian Register of Business Enterprises

Revision 1.0, Apr 16<sup>th</sup>, 2007

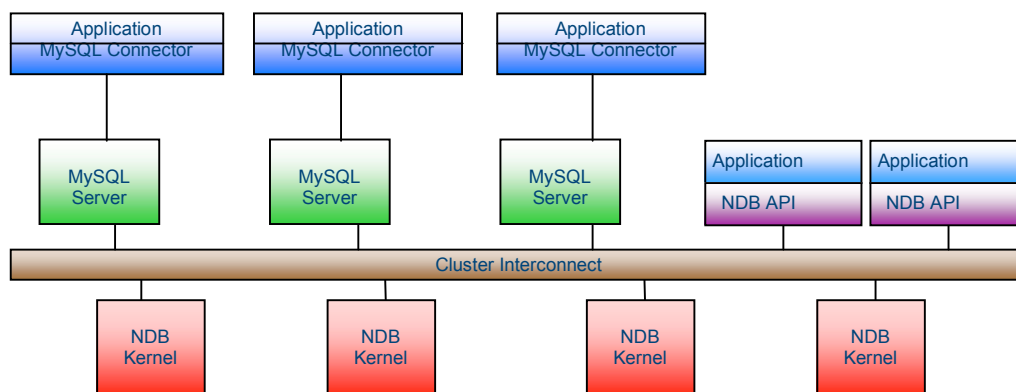
## 1 INTRODUCTION

In this white paper we show how MySQL Cluster in combination with Dolphin Express becomes a `real-time` DBMS suitable for markets with extremely low latency or `real-time` requirements. Low latency is important in delivering fast response time in any database application, but for some markets such as telecom, financial services and many embedded systems, this requirement is a critical requirement. Our results demonstrate that Dolphin Express provides a response time improvement of up to 292% as well as significant performance increases in high-load cases, particularly in larger clusters, making it ideal for any performance sensitive database application and uniquely able to meet the needs of `real-time` database applications.

A companion in-depth technical whitepaper for an audience that's interested in knowing more details on the technical aspects of the solution is available on the Dolphin website at [www.dolphinics.com](http://www.dolphinics.com). Dolphin also provides a set of tools to enable you to perform similar testing to verify the benchmark results.

MySQL Cluster is an Open Source Database technology developed by MySQL AB that brings High Availability, Scalability and High Performance to database applications. In this white paper we show how it can also provide Real-time characteristics. MySQL Cluster uses the standard MySQL technology with a clustered database backend, NDB Cluster. MySQL Cluster can be accessed from the MySQL Connector's which provides access to the full LAMP stack, or it can also be accessed from the NDB API that works directly with NDB Cluster.

## MySQL Cluster



Given that MySQL Cluster is a clustered backend, communication performance is important and therefore the choice of cluster interconnect technology is very important. There are many aspects of the cluster interconnect that are important, but probably the most important is how well integrated the interconnect software is with MySQL Cluster and the latency in the interconnect hardware.

Dolphin's Express cluster interconnects provides low latency on the hardware level in combination with Dolphin Express software, called SuperSockets, which provides a socket API implementation optimized for the cluster interconnect hardware. Using Dolphin Express a standard MySQL Cluster installation can use the Dolphin hardware immediately. The socket implementation is well engineered and makes use of 3 different schemes for performing write operations dependent on size of the writes. There is a very tight integration of interrupt mechanisms. This makes it possible to very quickly wake any process waiting for messages on the socket. It is also possible to locally shut off those interrupts if the software is not currently waiting for messages. Thus in many instances, it is possible to entirely remove the cost of interrupts. This feature is very important in

making Dolphin Express a very efficient solution in high-load cases. This is particularly true in configurations where all CPU's are kept busy. Dolphin and MySQL have developed features in MySQL Cluster that enables this efficiency for Dolphin Express, and is referred in this paper to as the real-time extension of MySQL Cluster.

Furthermore for real-time markets it is not enough to simply be able to process many messages at peak loads. It is also important to have very low latency with very stable response times. As an example in financial markets many companies receive the same information simultaneously and the company that can process the information first has a first-mover advantage and has the best opportunity to make money for its clients. Faster response time of the DBMS is an important part of such information processing for the financial markets.

Given the significance of low response times, Dolphin and MySQL have developed a number of tweaks to the MySQL Cluster software that makes its real-time behaviour improve considerably. These are currently in the process of being integrated with the MySQL Cluster Carrier Grade Edition. Without using those tweaks, the Dolphin Express solution decreases the standard deviation of response times considerably. With these optimizations implemented and properly used, the standard deviation decreases significantly and the application behaviour becomes very predictable. Thus, the risk of losing economical advantage simply because of random effects in the software is decreased.

**2 PERFORMANCE ASPECTS OF INTERCONNECT HARDWARE**

The cluster interconnect can improve a number of performance parameters. These are:

- 1) Latency Improvements
- 2) Efficiency Improvement in CPU-limited scenarios
- 3) Efficiency Improvement where there are unused CPU resources
- 4) Bandwidth Improvements

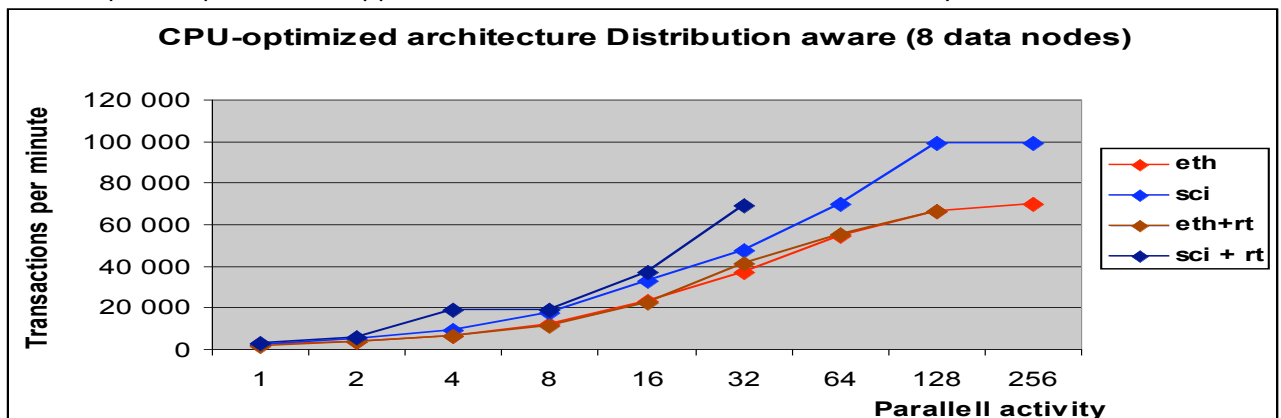
The reason we differ between efficiency improvements in CPU-limited scenarios and scenarios where there is abundant CPU resources are that they are affected by different factors.

**2.1 Latency Improvements**

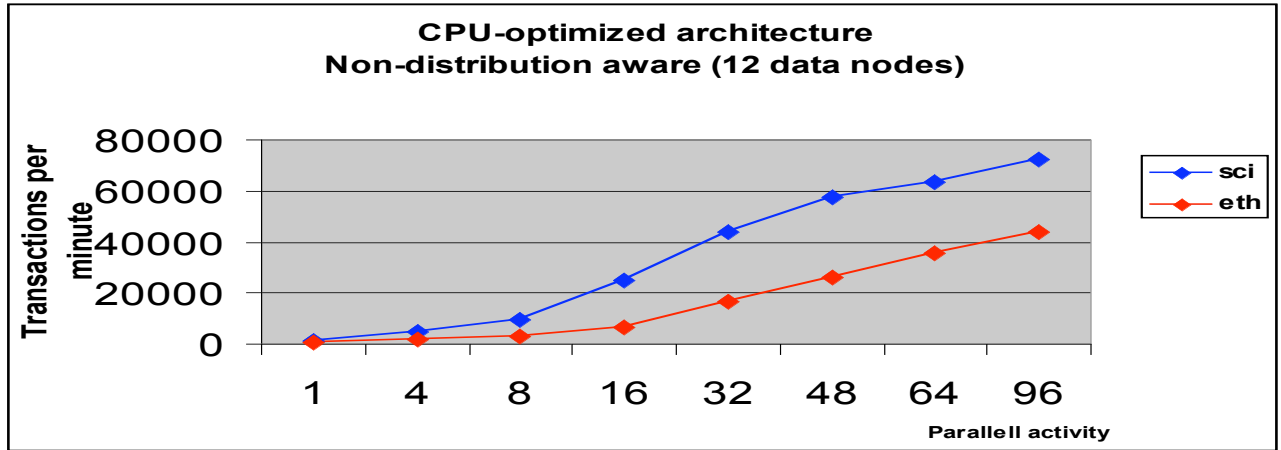
The improvement of latency has a great effect on response time and Dolphin Express achieves this through the HW and its integrated software solution. The improvement can be as great as a factor of four although a factor of two is more normal. In our benchmarks, we demonstrate this factor as very dominant when the number of parallel threads of activity is low.

**2.2 Efficiency Improvement in CPU-limited scenarios**

When there aren't enough CPU resources and the system is heavily loaded the CPU resources used to handle communication is the decisive factor in the efficiency. As shown below Dolphin Express brings a factor of 40-70% efficiency improvement to the system. Thus Dolphin Express can support at least 40% more transactions in heavy load scenarios.

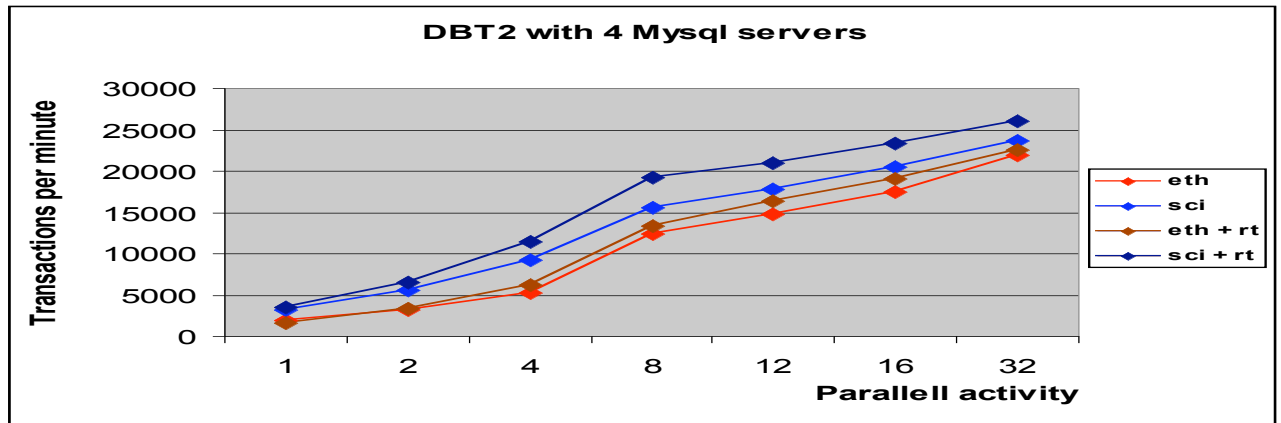


The graph above was in a scenario where the transaction coordinator is placed where data of the transaction is residing. In the graph below we show a similar case where transaction coordinator is randomly chosen. In the graph below we can see that the improvement by Dolphin Express can be as much as 292% in this case. The ability to choose where the transaction coordinator is located is a new feature currently being integrated in the MySQL Cluster Carrier Grade Edition.



**2.3 Efficiency Improvement where CPU resources are abundant**

When there are abundant CPU resources available, the decisive factor of the efficiency of a cluster interconnect, is dependent on how it manages interrupts to the processor. In cases with high load and abundant CPU resources Dolphin Express still supersedes the Ethernet interconnect by at least 15% higher transaction rate, even in small clusters.



What we see in the above graph is that up to 8 threads we can handle more transactions at a very good maintained response time. We improve throughput by 650% by only increasing response time by 50%. Whereas from 8 to 32 parallel activities we only increase throughput by 30% at the cost of a 200% increase in response time. Thus optimum for real-time applications is here at 8 parallel activities where we can see that Dolphin Express using the real-time extension have approximately a 50% performance advantage compared to Ethernet.

**2.4 Bandwidth Improvements**

The Dolphin Express hardware can handle 3-5 times higher bandwidth compared to Gigabit Ethernet. This is relevant in applications storing large objects in MySQL Cluster. In this white paper and the companion technical white paper we have not made any benchmarks testing these improvements.

**2.5 Conclusion**

MySQL Cluster using Dolphin Express can be used to build a `real-time` DBMS product with very high performance and very high availability. Our results demonstrate that Dolphin Express provides response improvements of up to 292% as well as significant performance increases in high-load cases, particularly in larger clusters.