

## Clustered Database Systems for Internet-Scale Applications

During the past decade, Internet applications have experienced huge growth rates. Although this bodes well for the business, it presents enormous complexity, effort, and costs for database architects to maintain large-scale Web applications. As transactions and data volumes explode, there has been no easy and cost-effective way to scale the database platform to accommodate the rising demand.

In order to achieve scale, companies have been forced to use database sharding, requiring extensive architecture and application changes to create and manage isolated database “islands.” Some organizations, desperate to find an alternative solution, have switched to key-value stores, despite its minimalist functionality. Yet, most companies don’t want a stripped-down database; they want a complete solution.

### **Bringing Scale, Efficiency, and Simplicity Where There Was None**

Clustrix, Inc. delivers a unique database solution that provides the full relational, transactional, ACID-compliant functionality of a SQL database, and the flexible data model and scalability of key-value stores. The Clustrix Clustered Database System (CDS) is a fully relational, massively scalable, distributed database that meets the requirements of the most demanding transactional database applications. Designed from the ground up to accommodate Internet-scale OLTP applications, the Clustrix solution is easy to manage and provides massive scalability, rich functionality, high performance, and high availability—at a significantly lower cost than less powerful solutions.

Delivered as a complete and easily-deployed appliance, scaling out is as easy as adding another CLX 4000 Series™ node—even in a live production environment. With each additional node, the database scales incrementally. Whether there are a few or hundreds of nodes in a cluster, the Clustrix cluster appears to applications as a single database—although the data is intelligently spread across the nodes.

### **Fully Distributed and Parallel Architecture**

At the heart of the Clustrix solution is the Sierra Clustered Database Engine, a shared-nothing execution environment that includes the Sierra Parallel Planner and the Sierra Distributed Execution Engine. This unique distributed and parallel architecture enables the Clustrix Clustered Database to execute one query with maximum parallelism and many simultaneous queries with maximum concurrency. It does this by moving the query to the data, not the data to the query. It is through this advanced technology that the CLX 4000 delivers enormous scalability, high read/write workload performance, high availability, load balancing, and self-management.

### **HIGHLIGHTS**

- Incrementally scales to hundreds of nodes
- High performance—even under heavy write loads
- Fully relational database, ACID compliant
- Efficiency and scale of key-value stores
- High availability: automatic failover and self-healing
- Non-disruptive: online deployment, expansion, and schema updates
- Supports all MySQL connectors
- Delivered as a turnkey appliance

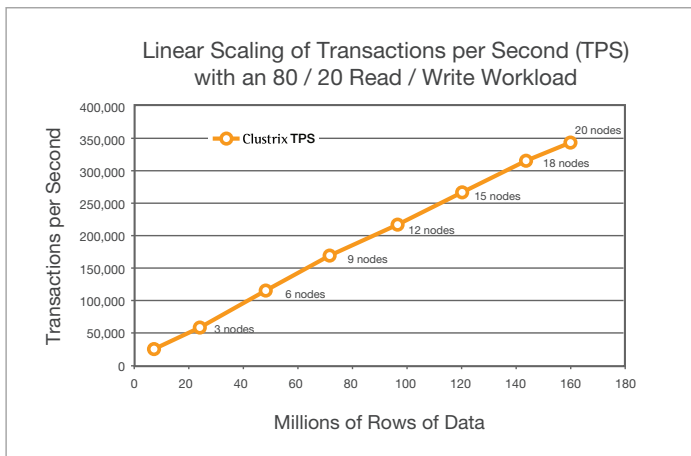


## High Availability

Clustrix's distributed, shared-nothing architecture was designed for high availability. To eliminate unplanned downtime, the system provides automatic failover and recovery to instantly and transparently respond to failures, as well as self-healing to automatically restore the system to its original protection level. Each CLX 4000 Series node contains a battery-backed, non-volatile RAM card that ensures safe transactions in the event of a power failure. To further support continuous operation, the architecture allows online deployment and expansion of nodes as well as online updates to database schema.

## Scaling Performance

One of the most compelling and unique features of Clustrix's database solution is that the distributed, parallel, shared-nothing architecture provides linear growth in transactions per second (TPS), I/O throughput, capacity, and cache. Because the Clustrix solution is fully distributed, eliminating the bottleneck of central coordination common in database clusters, it provides a constant growth rate for all of the above dimensions with each additional node. This means that prior ceilings on scalability have been removed. Test results are impressive: a Clustrix cluster with only 20 nodes processed over 343,000 transactions per second for an application with an 80/20 read/write workload. The same 20-node cluster generated 650,000 TPS with read-only workloads.



## Easy Deployment, Expansion, and Management

CLX 4000 Series appliances deploy in a transparent and non-disruptive manner into a MySQL environment. No changes are required to existing database schemas or applications. The CLX 4000 Series supports common MySQL operations, including MySQL replication, MySQLDumps, and MySQL bin logs. So, you can implement CLX 4000 appliances in your SQL environment without changing your architecture or applications—regardless of whether your database is standalone, replicated, or partitioned. You can use replication to set up your CLX 4000 appliances as a read slave to one or multiple MySQL masters, or as a master to one or multiple MySQL slaves. Or you can choose to take full advantage of the power of a Clustrix solution and replace all your MySQL servers with CLX 4000 appliances, giving you a massively scalable, high-performance, highly available database cluster—managed as a single database. Scaling the database horizontally is as simple as adding one or more CLX 4000 appliances to an existing cluster, and assigning IP addresses—the expansion occurs automatically.

## Sounds Almost too Good to be True

If you would like to increase the scalability, performance, and flexibility of your database system, and reduce operational costs, visit [www.clustrix.com](http://www.clustrix.com) or contact us to learn more about the Clustrix 4000 Series solution.

