



# Syncsort and Vertica Set New World Record

Smash cost/performance barrier by loading 5.4 TB in under 60 minutes on affordable **HP** blade system

## The New Standard in High-Speed Cost-Efficient Computing

Syncsort and Vertica have shattered the previous benchmark held by Microsoft and Unisys by more than doubling the record for ETL performance. Syncsort's data integration product, DMExpress™ v4.8, extracted, transformed, cleansed, and loaded 5.4 TB of raw TPC-H data into the Vertica Analytic Database in 57 minutes 21.51 seconds. The feat was accomplished using highly cost-efficient HP BladeSystem c7000 x86 servers running the RedHat Linux operating system. This achievement underscores claims that the Syncsort-Vertica solution can handle significant volumes of heterogeneous data faster and at a fraction of the cost of any other potential solution.

### Syncsort DMExpress: High-Speed ETL

Syncsort DMExpress is the fastest, most reliable and cost efficient enterprise Data Integration software.\* When dealing with large data volumes, DMExpress replaces long-running processes, such as ETL, so that applications run faster – delivering information to business users when they need it.



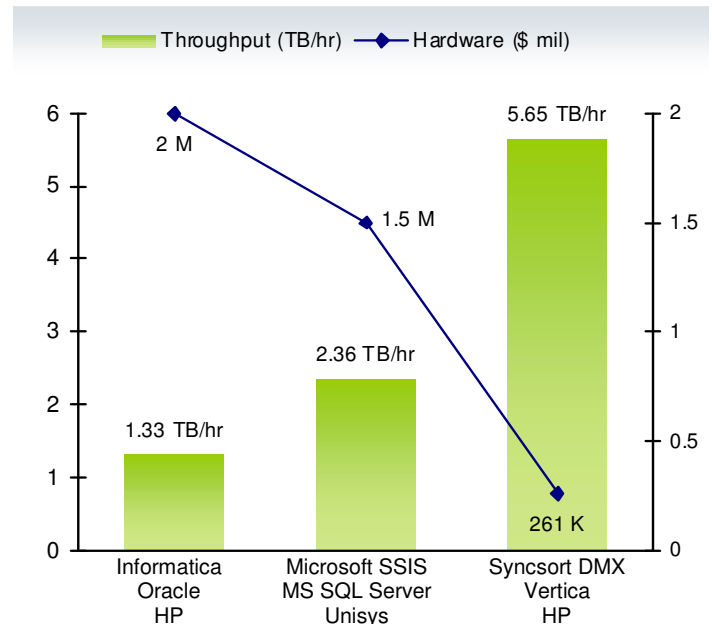
### Vertica: High-Performance Data Management

Vertica is the market innovator for high-performance analytic database management systems. Vertica gives more people the freedom to analyze more data in more ways, faster, so they can out-think and out-innovate the competition. Vertica is the only analytic database built from scratch to handle today's workloads.



### World Record Highlights

- 5.4TB in 57 minutes 21.51 seconds – Syncsort, Vertica, and HP
- Previous record – 1 TB in 25 minutes 20 seconds held by Microsoft and Unisys
- Independently audited by DSS Labs
- Process:
  - ✓ DMExpress reads industry-standard TPC-H data from source disk files into memory.
  - ✓ DMExpress performs data cleansing on the source records to ensure they conform to business rules, and creates a reject file of any records that are non-conforming.
  - ✓ DMExpress passes the data in ASCII form with delimiter separated fields via pipes.
  - ✓ Vertica's copy receives the data via pipes.
  - ✓ Vertica's copy loads the data into the Vertica database.
  - ✓ The process is repeated for all the tables.



\* Based on results of specific tests against leading competitors, including Informatica, IBM DataStage, and others. Results may vary depending on environment.

## Compelling Innovation to Drive ROI

### Syncsort DMExpress

Syncsort has built compelling solutions to performance problems for 40 years. Its latest data integration innovation, DMExpress speeds ETL, data warehousing, BI, and other large volume applications. Running on UNIX, Windows, and Linux, DMExpress integrates disparate data from multiple, heterogeneous sources. It transforms and aggregates the data at high speed to create consolidated views for reporting, analysis, or other downstream applications. The solution integrates a host of compelling innovations, such as:

- Proprietary Algorithms
- Cutting Edge Data Handling Methods
- Dynamic Optimization
- Fast I/O
- Parallel Processing

#### Benefits to Drive ROI

- High Performance
- Breadth of Functionality
- Scalability & Operational Flexibility
- Ease of Use
- World Class Service & Support

### Vertica Systems

Vertica Systems is the market innovator for affordable, high-performance analytic database management systems. Vertica has developed blindingly fast database technology for storing and querying very large databases using industry-standard hardware. The column-oriented, grid-enabled Vertica Analytic Database offers unmatched performance, scalability, and flexibility for real-time, read-intensive analytical applications in financial services, telecommunications, retail, Internet, and government. It's suite of innovation, includes:

- Column store architecture
- Heavy use of compression
- Redundancy & multiple sort orders
- Automatic database design
- Recovery by query
- Continuous Data Loading

#### Benefits to Drive ROI

- High Performance
- Column architecture
- Scalability
- Data Compression
- Reduced Storage and Labor Costs
- Interoperability

### Increase ROI with the Syncsort-Vertica-HP Solution

#### Labor Costs

- ✓ Ease of use increases productivity and saves staff time
- ✓ Avoid hiring cost
- ✓ Fast project implementation reduces project costs

#### Hardware Costs

- ✓ Ability to run on commodity hardware
- ✓ Minimal use of CPU, storage and other expensive computing resources

#### Software and Services Costs

- ✓ Inexpensive license fees
- ✓ Modestly priced and minimally needed training and professional services
- ✓ 24X7 world-class technical support

For more details on the Syncsort-Vertica-HP solution or to download an independent technically audited white paper written by George Spofford of DSS Labs, visit [www.etlworldrecord.com](http://www.etlworldrecord.com).