AMD Boosts Data Warehouse Performance with Parallel Processing Appliance

“Before, we had a very tough time maintaining our service level agreement in delivering data to ... analysts. By delivering data that they want on time, we relieve them of worry about data volumes and query structure and help them focus on analysis.”

Renjye Yu, IT Senior Manager, AMD

In testing semiconductor wafers, AMD uses a data warehouse to process and analyze one terabyte of data each week. When its data warehouse began foundering under the load, AMD switched to Microsoft SQL Server 2008 R2 Parallel Data Warehouse powered by AMD Opteron™ processors. With the new system, AMD has increased performance and enabled a sustainable and scalable solution.

Solution Spotlight
- Improves performance of data warehouse used to analyze wafer electrical tests, helping analysts uncover microprocessor performance trends sooner.
- Processes 10,000–13,000 reporting queries a day, meeting the needs of sophisticated analysts.
- Estimated to reduce data warehouse support time by up to 90 percent.
- Cuts data storage costs by an average of three to six times thanks to data compression.

Business Needs
AMD is a global semiconductor firm with a presence in the graphics, personal computing, and server markets. At different points in the semiconductor wafer fabrication process, AMD performs a variety of wafer electrical tests and loads the data into a data warehouse for analysis. Because of the large number of tests run, AMD adds nearly a terabyte of data each week to this data warehouse.

The data warehouse team struggled to accommodate this rapid expansion and to deliver the necessary query and data-loading performance required by about 50 analysts. “These are sophisticated business analysts who understand how to mine the data and run thousands of ad hoc queries,” says Salim Saba, IT Director for Data Warehouse at AMD. “Their work generates data that others in the AMD engineering community use to create reports, so it’s important that their work not be held up.”

However, analysts were not happy with the solution’s data latency. They and the data warehouse team spent a great deal of time deleting data sets, loading the latest data first, and chopping large queries into many smaller queries to speed up performance time. “We had to retain just six months of data rather than the year’s worth that analysts wanted,” says Rajaro Chitturi,
Database and Applications Manager at AMD. “We were reacting on a daily basis to keep the data warehouse performing to the level that our users required.”

**Solution**

AMD wanted a massively parallel processing (MPP) data warehouse platform that would support its larger data sets and provide better performance. When AMD learned about Microsoft SQL Server 2008 R2 Parallel Data Warehouse, which combines Microsoft SQL Server 2008 Enterprise software with industry-standard appliances, the company knew that it had found its new wafer-test data warehouse solution. Simple to deploy, SQL Server 2008 R2 Parallel Data Warehouse is delivered as a prebuilt appliance with software, hardware, and networking components already installed. It provides high-speed parallel query processing, highly scalable data storage, and high-speed data transfer.

“With traditional databases, storage plays a key role, because to do any reporting, you need indices,” says Saba. “However, storing multiple indices in a database to attain adequate query performance took up to 75 percent more space than the data itself. With Parallel Data Warehouse, we don’t need indices to this extent but rather use hash keys to tell the database where and how to store the data.”

AMD uses a custom reporting tool and Microsoft SQL Server 2008 R2 Reporting Services to pull data from Parallel Data Warehouse and create reports.

AMD and Microsoft worked together to create a custom version of Parallel Data Warehouse powered by AMD Opteron™ 6100 Series processors. An HP ProLiant DL385 G7 server computer houses the processors.

**Benefits**

By migrating its data warehouse to SQL Server 2008 R2 Parallel Data Warehouse, AMD has significantly improved loading and reporting performance. It can deliver data faster for analysis and has lowered storage costs by using data compression.

**Improves Data-Loading Performance**

Since migrating its electrical test data warehouse to Parallel Data Warehouse, AMD has seen data-loading times improve dramatically. “We used to worry about backlogs, but no more,” says Chitturi. AMD runs an average of 1,500 loads per day, and data loads to a given table range from four-minute to four-hour intervals. AMD averages about 500,000 file loads a day, with file sizes ranging from bytes to gigabytes. “We are really proud of the latest AMD Opteron™ processor inside Parallel Data Warehouse, which contributes to the great performance of this system,” Chitturi says.

AMD is also processing more reporting queries than it previously could—between 10,000 and 13,000 a day—with an average runtime of a few seconds and virtually no performance issues. Backup is also faster: it used to take the data warehouse team a week to back up its data, but with Parallel Data Warehouse it takes up to two hours at most.

**Delivers Data to Analysts Faster**

With better data warehouse performance, analysts can uncover microprocessor performance trends sooner and generate data needed by engineers across AMD. This ultimately helps AMD get faster processors to market sooner. “Our job on the data warehouse team is to provide the company with a single version of the truth so that all employees can make informed decisions and help the company grow,” Saba says. “Parallel Data Warehouse is helping us to deliver a sustainable and scalable data warehouse solution.”

Renjye Yu, IT Senior Manager at AMD, adds, “Before, we had a very tough time maintaining our service level agreement in delivering data to these analysts. By delivering data that they want on time, we relieve them of worry about data volumes and query structure and help them focus on analysis.”

**Reduces Support Work Needed by an Estimated 90 Percent**

The data warehouse team has more time to load data and enhance the system rather than spending time chopping up queries to fit the limits of the data warehouse. Additionally, because of the user complaints about the previous system, the data warehouse team had one employee devoted full time to addressing performance-related support tickets. With Parallel Data Warehouse, AMD has reduced support work to just a few hours a week.

**Cuts Storage Costs Through Data Compression**

AMD has also reduced its storage costs, because of the built-in compression in Parallel Data Warehouse. “We have seen data compression increase on average by three to six times depending on the data search,” Saba says. “This helps to lower storage costs and improve query performance.”

---

**Software and Services**

- Microsoft Server Product Portfolio
  - Microsoft SQL Server 2008 Enterprise
  - Microsoft SQL Server 2008 R2 Parallel Data Warehouse
- Technologies
  - Microsoft SQL Server 2008 R2 Reporting Services

**Hardware**

- AMD Opteron 6100 Series processors
- HP ProLiant DL385 G7 server computers

---

This case study is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS SUMMARY.