

Accelerating Analytics with Intel® Solid-State Drives

The State of Michigan helps speed educational data analysis by up to 60 percent while achieving an eight-month return on investment



“In our testing, the Intel SSD 910 Series delivered significantly better performance than the SAN for both short and long queries....With one long query, we saved two hours and 45 minutes—a reduction of more than 60 percent. The improved performance will help reduce user frustration and increase productivity.”

*Steve St.Laurent,
Database Administrator,
Michigan Department of Technology,
Management & Budget (DTMB)*

The Michigan Department of Education works with the Center for Educational Performance and Information (CEPI) to analyze a wide range of data from the state’s K-12 schools. The goal is to generate new insights that can help boost academic performance, improve planning and budgeting, optimize the utilization of assets, and more. To accommodate the growing use of analytics from a variety of users, CEPI needed a way to provide large-capacity temporary disk space for workload processing while conserving its storage area network (SAN) resources. A pilot program showed that installing Intel® Solid-State Drives (Intel® SSDs) in analytics servers could significantly improve query performance and deliver a rapid return on investment (ROI).

CHALLENGES

- **Accommodate more analytics and larger queries.** Support the increasing use of analytics at the Department of Education by providing sufficient temporary disk storage space for more, larger queries and analytics workloads.
- **Reduce the time to new insights.** Accelerate query results and generate new insights rapidly.
- **Control IT costs.** Meet the storage requirements of the analytics application without having to use costly SAN resources.

SOLUTION

- **Intel SSDs.** The CEPI team ran a pilot program that used the Intel SSD 910 Series within Intel® Xeon® processor-based servers as temporary disk storage.

TECHNOLOGY RESULTS

- **Improved performance up to 60 percent.** Tests showed that Intel SSDs significantly reduced processing time for some analytics workloads compared with the using the SAN. The larger the query, the better the performance increase.

BUSINESS VALUE

- **Anticipating an eight-month return on investment.** CEPI can achieve an ROI in just eight months by using Intel SSDs while conserving SAN capacity for other functions.
- **Gained greater capacity for analytics.** CEPI can better accommodate the growing use of analytics by a wide range of users.

To accommodate the rising use of data analytics by a broad array of users, the IT professionals at CEPI had to rethink a key component of its analytics infrastructure—the temporary (or “scratch”) disk space required by the SAS* analytics application. The SAS servers pull in data from Microsoft SQL Server* databases and then keep the data in a scratch disk space during processing.

“In some cases, ad hoc queries are not well formed,” explains Steve St.Laurent, a database administrator for the DTMB, which oversees CEPI. “A user might need information for just one or two school districts, but the query will pull in information for every school district in the state. We need a large scratch space to accommodate those types of queries.”



Intel SSDs speed query results

Currently, CEPI uses a SAN to provide that disk space, but this approach is too expensive, especially since the capacity needed for scratch disk space fluctuates over time. "We have 500 GB of SAN capacity reserved for scratch disk space, and we pay for that capacity every month, whether it is fully utilized or not," says John Carey, IT section manager with DTMB. "We need a more cost-effective way to provide large-capacity temporary storage."

In considering new approaches, performance is a key consideration. "Using a disk within the analytics server might help improve performance compared with using the SAN, but we realized that traditional hard disk drives wouldn't be able to keep up with the high number of simultaneous transactions," says Carey. "We needed to explore other options."

Building a Better Scratch Disk with Intel SSDs

The CEPI team decided to evaluate Intel SSDs. "We considered other SSDs," says St.Laurent. "But to deliver the performance that some of our data architects require, we decided to test the Intel SSDs—we knew they were among the fastest drives available."

The CEPI team was also impressed with the long-term dependability of Intel SSDs. "Engineers from Intel IT provided real-world usage information demonstrating these were enterprise-class drives that could deliver the reliability and longevity we need for analytics," says St.Laurent.

To gain approval for this solution from the state, the CEPI team needed to conduct

a pilot that could confirm the benefits of using SSDs as part of the analytics environment. With help from Intel, the CEPI team installed two 800 GB drives from the Intel SSD 910 Series in an HP ProLiant* DL580 server, which was equipped with the Intel Xeon processor 7500 series. Using a second drive helps to ensure high availability by providing backup capacity in the event of a problem with the primary drive. The CEPI team ran typical short and long queries using the SAN as the scratch disk and then using the Intel SSDs.

Boosting Query Performance by Up to 60 Percent

The testing showed a clear advantage for using the Intel SSDs and helped CEPI gain approval for implementing the new drives. "In our testing, the Intel SSD 910 Series delivered significantly better performance than the SAN for both short and long queries," says St.Laurent. "In fact, the larger the query, the better the performance. With one long query, we saved two hours and 45 minutes—a reduction of more than 60 percent. The improved performance will help reduce user frustration and increase productivity."

Using internal Intel SSDs instead of the SAN for scratch disk space will also help conserve SAN resources. "Today, a poorly formed query could severely affect the performance of the SAN," says St.Laurent. "By using SSDs inside the servers, we can avoid those performance hits and reserve the SAN for other functions, including the storage of our tremendous volumes of source data."

LESSONS LEARNED

The CEPI team envisions how the value of using the Intel SSD 910 Series will increase in the future. "As queries grow in size and complexity, it will become even more critical to provide high-capacity, high-performance scratch disk space," says John Carey, IT section manager with the Michigan DTMB. "Using the SAN would be too costly and could inhibit certain queries. By integrating Intel SSDs into our analytics servers, we have the right model going forward."

Anticipating a Rapid ROI

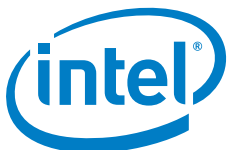
Adopting Intel SSDs could also help CEPI save money. "By choosing Intel SSDs, we can make a one-time purchase and avoid the ongoing costs we incur by using the SAN for scratch disk space," says St.Laurent. "We can afford to set aside more capacity for scratch disk space without having to pay ongoing SAN charges."

"Even by purchasing two drives per server to help ensure availability, we expect an eight-month return on investment," says Carey. "Everything past that would be 100 percent savings."

Helping to Revitalize Education

"By helping to accommodate greater use of analytics by more users across the Department of Education, we are doing our part to revitalize education in Michigan, which is one of the top priorities of our governor," says Carey. "At the same time, the cost savings will contribute to our goal of enhancing operational efficiency. We can provide better services while saving taxpayer money."

Find a solution that is right for your organization. Contact your Intel representative, visit [Business Success Stories for IT Managers](#), or explore the [Intel IT Center](#).



This document and the information given are for the convenience of Intel's customer base and are provided "AS IS" WITH NO WARRANTIES WHATSOEVER, EXPRESS OR IMPLIED, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NON-INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS. Receipt or possession of this document does not grant any license to any of the intellectual property described, displayed, or contained herein. Intel® products are not intended for use in medical, lifesaving, life-sustaining, critical control, or safety systems, or in nuclear facility applications.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. **For more information go to <http://www.intel.com/performance>**

Intel does not control or audit the design or implementation of third-party benchmark data or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmark data are reported and confirm whether the referenced benchmark data are accurate and reflect performance of systems available for purchase.

© 2013, Intel Corporation. All rights reserved. Intel, the Intel logo, Intel Inside and Xeon are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

0113/LJ/TDA/XX/PDF

328536-001US