Accelerate Business Analytics
Performance Cost Effectively with Intel® Cache Acceleration Software and Intel® Solid-State Drives (SSDs)

Intel® Cache Acceleration Software (Intel® CAS) with Intel® Solid-State Drives increases application performance with intelligent caching to enable up-to-date business intelligence while decreasing time and cost.

SAS, the global leader in high-performance analytics, enables customers to measure what matters most today, reveal best actions, expose threats, and gain predictive insights that compel the right actions. In a world where “big data” grows exponentially every day, SAS provides a suite of intuitive interfaces and the ability to deploy real-time analytics directly to mobile devices, so every decision maker—everywhere—can monitor key metrics and make informed decisions.

One challenge many companies face when monitoring this glut of structured and unstructured data is analyzing it and getting results in a timely fashion. Cheap, abundant hard disk drive (HDD) storage has made it easy to save all kinds of data, yet these HDDs are also a bottleneck when users try to retrieve the data for analysis.

All organizations with access to large data collections should try to harness the most relevant data and use it for better and faster decision-making.

Challenges

• Reduced productivity. Existing hard disk drives experience significant lag when querying data sets larger than five million records. This forces developers to spend additional time waiting for queries to load.

• Cost of improving performance. Lengthy query load times impact the quantity and timeliness of customer-ready data analysis. A full migration to SSD-based storage improves performance dramatically, but the hardware and time investment required may be prohibitive for some customers.

Solution

• Intel® CAS and Intel® Solid-State Drives. We suggest a seamless upgrade using Intel® CAS for SAS input and migrating the SAS WORK directory to an Intel® SSD. This approach improves performance cost-effectively—nearly twice as fast as HDD only...for about one-tenth the cost of a full SSD implementation.

Benefits

• Nearly double performance with a minimum investment. Intel® CAS paired with Intel® SSDs for both caching and SAS operations enables developers to perform queries in significantly less time, allowing them to complete their analyses faster and reduce costly load waiting time...all while using existing storage. This solution is significantly faster than HDD-only systems, yet costs about 10% of a total solid-state solution.

“SAS understands that customers need fast access to big data for making critical business decisions. Quickly getting the right information impacts the bottom line. Investment in the infrastructure that allows faster access to their data is necessary, but it doesn’t always have to break the bank. In joint testing with Intel, SAS discovered that Intel® Cache Acceleration Software improved data retrieval significantly at a reasonable cost. Intel® CAS will allow many of our cost-conscious customers to improve their analytics performance economically.”

— Cheryl Doninger, Senior Director Research and Development, SAS
What is Intel® CAS and how does it work?

Intel® CAS is a server-side caching solution that supports directly attached SSDs and SSDs accessed through SAN storage protocols (such as iSCSI, Fibre Channel, FCoE, etc.), and provides a write-through cache for HDD-based storage solutions. Available on Windows® and Linux® platforms, Intel® CAS integrates seamlessly, with no changes to your application or backend storage.

As an application reads data from storage, Intel® CAS copies the data to the cache, which serves subsequent reads at higher speeds. Meanwhile, all data is written synchronously to backend storage and the cache.

Why Intel® CAS with Intel® Solid-State Drives?

Our High Endurance Technology (HET) provides 30 times the endurance of standard multilevel cell (MLC) technology. Intel® SSDs with HET provide considerable write endurance: Ten full drive writes per day over five years.

- Intel® SSD 910 series. Delivers extremely high performance storage with 180 K random reads and 75 K random writes.¹
- Intel® SSD DC S3700 series. Delivers data at a breakneck pace, with consistently low latencies² and tight IOPS distribution (75 K random read IOPS³).

Intel® CAS and Intel® SSD performance benefits

Intel and SAS jointly tested a typical 40-user SAS mixed analytics environment operating on an HDD-only system. With this as our baseline, we compared performance against the same system with three changes:

- Intel® CAS and Intel® SSD for caching input.⁶
- Intel® CAS and Intel® SSD for caching input, plus Intel® SSD for SAS WORK.⁷
- Intel® SSD for caching input, SAS work, and output.⁸

As the graph above shows, adding Intel® CAS and an Intel® SSD for caching improves performance by 36%, yet costs less than a nickel on the dollar compared to a full SSD solution. Adding Intel® CAS plus Intel® SSDs at the cache and SAS operation layers nearly doubles performance, at one-tenth the cost of a full SSD solution.

¹. Read and write IOPS (Full LBA Range, Iometer* Queue Depth 32 per NAND module). Default/max performance mode: Random 4 KB reads of up to 180,000 IOPS; random 4 KB writes of up to 75,000 IOPS.
². 50 μs typ.; <500 μs max.: 100 GB 4 KB random writes QD=1 at 99.9% of the time across 100% span of the drive.
³. Random 4 K reads.
⁴. System configuration: Generic “white box” server with Sandy Bridge EP 4S (2.9 GHz) processor and 256 GB of RAM running RHEL 6.3.
⁵. Mixed workload: Traditional SAS v9.3 mixed analytics workload in Linux OS environment.
⁶. Same as the HDD setup, with one 100 GB Intel® SSD DC S3700 and Intel® CAS for caching input.
⁷. Same as the HDD setup, with one 100 GB Intel® SSD DC S3700 and Intel® CAS for caching input, plus one 800 GB Intel® SSD DC S3700 for SAS WORK.
⁸. Eleven 800 GB Intel® SSD DC S3700, with an onboard controller.