



Baidu BigSQL Delivers Faster Spark Interactive Queries with Lower TCO

Baidu's BigSQL data processing platform is based on Spark SQL* and has many features and performance enhancements that improve on it. As Baidu's business expands, the scale of hot data grows rapidly. Memory scaling is needed to deliver the same level of performance that users demand. However, the high cost of Dynamic Random-Access Memory (DRAM) adds increasing pressure to the total cost of ownership (TCO). To lower TCO while ensuring satisfactory performance, Baidu deployed Intel® Optane™ persistent memory and used it to optimize its ad hoc query service – Tuling. Supported by Intel Optane PMem, the cluster offloaded more than 30% of the workload from Tuling². Additionally, the average query latency reduced by 20%².

“In order for Baidu BigSQL* to provide users with high-performance ad hoc query services, large memory is needed to cache hot data locally on compute nodes to avoid DFS I/O slowing performance down. With Intel Optane persistent memory, we managed to ensure outstanding cache performance, while at the same time greatly improving cluster processing and achieving significant TCO benefits.”

LI Shiyong, Senior System Engineer, Baidu

Products and Solutions

[2nd Gen Intel® Xeon® Scalable processors](#)
[Intel® Optane™ persistent memory](#)

Industry

Internet Software,
IT Services

Organization Size

10,001+

Country

People's Republic
of China

Learn more

[Case Study](#)