Toward a Unified Computing Architecture

For enterprise IT and data administrators, the world is changing fast. For decades, IT organizations have built separate computing infrastructures aimed at addressing specific business problems. Chief among these:

- Traditional transaction processing (OLTP), which was used to operationalize day-to-day functions in business (ordering, inventory, shipping, payment, and so on);
- Analytics processing (OLAP).

Data derived from transactional activity often informed analytics processing; however, the two computing activities were siloed from each other. Not only did transactional and analytical processing require different database infrastructures (relational databases supported transactional processing while data warehouses powered analytical processing), hardware and software for each had to be differentially optimized to achieve acceptable performance. Analytics processing was more resource intensive, and often required reformatting of transactional data for analysis, which led to storage of multiple copies of data in different formats, leading to ever greater expense, complexity and latency.

And OLTP and OLAP are just two of the infrastructure islands in a typical enterprise:

- ERP implementations have grown up to manage major business processes in industry segments and lines of business;
- CRM solutions have taken root to manage customer information;
- Storage systems were established to keep business data safe;
- Enterprise search systems were created to find and retrieve information.

Executive Summary

In an era of high-volume data flows and real-time analytics, an enterprise computing architecture built on a foundation of isolated, fragmented IT infrastructures can slow the generation of critical business insights and prevent visibility into key data sets. A new, converged computing architecture called the data hub provides a powerful, unified infrastructure that can process transactional, analytical, and unstructured data in real time. This new architecture provides immediate visibility into data from across the enterprise, and with fast access to critical insights and constantly shifting patterns as they occur, businesses can react to—and in many cases predict—outcomes to drive better business decisions and processes.

Deriving insights from fragmented data is costly and inefficient. A unified computing architecture offers full and immediate visibility into enterprise data for real-time business insights.

Tim Allen
Senior Alliance Marketing Manager,
Intel Software and Services Group

Michael Demshki
Director, Business Development, Intel Big Data Solutions Group
Each of these capabilities is usually treated as a separate computing activity, often performed in a different computing infrastructure, using proprietary tools, and requiring dedicated IT support teams with focused expertise.

Not only is this incremental and fragmented approach to enterprise computing costly and inefficient, it doesn’t permit a unified view of information. Each infrastructure develops independent silos of data, making it difficult if not impossible to share information for analysis that spans data types or lines of business. And as data grows, it becomes increasingly challenging and costly to manage, transfer, and store.

It’s inefficient to transform and move stand-alone data stores simply to perform business queries on them. But without full visibility into all enterprise data, business insights are highly constrained, making it difficult to arrive at the right answers or recognize important trends.

Today, two developments are breaking down the traditional dichotomies in enterprise computing infrastructures—and bringing big changes to long-standing administrative roles in IT and analytics. First, in addition to transactional and analytical workloads, enterprise data centers today gather, analyze and store huge volumes of diverse unstructured data—the so-called Big Data—from a wide variety of sources, including sensors, web logs, utility monitors, surveillance cameras, social media, video and transportation networks.

Second, technologies such as in-memory databases and NoSQL data stores are fundamentally disrupting the enterprise computing and BI analytics technology models of the past. A new computing paradigm is emerging: with powerful in-memory computing—supported by breakthrough advances in server and networking technologies—the same infrastructure can be used to process transactional, analytical, and unstructured data in real time. This new infrastructure provides immediate visibility into data from across the enterprise to support critical, real-time business decision-making. We call this new converged computing platform the Data Hub: a new computing architecture that integrates an enterprise-class database infrastructure, high-performance data processing capabilities, and advanced business and analytics applications into a single in-memory platform that can process even the largest and most complex workloads in real time.

In the transition to this new computing paradigm, the change is not simply doing what you used to do faster:...
data is compressed and maintained. In-memory computing, all relevant data is analyzed, its inherent value has increased. Therefore, businesses can react to—and in many cases predict—outcomes to drive better business decisions and processes.

Intel recognizes the importance of this transformation, and is not only optimizing its own products and technologies for implementation in data hub architectures, but is also helping the entire ecosystem gain the highest possible benefit from this important transition.

### The Breakthrough of In-Memory Computing

Accessing data from storage for processing has long been a performance bottleneck for advanced computing systems. Storage performance for traditional spinning technologies such as hard disc drives (HDDs) is limited by mechanical processes, leading to long delays between when data is generated, when it is written to storage, and when it is available for analysis.

In these systems, data must be extracted from storage, translated into proprietary analytical formats, and then loaded into the analytics environment. Often, the resulting data models must then be optimized for performance. Despite this optimization and preparatory work, complex queries can still take many hours to complete.

With traditional bulk storage mechanisms—which still form the backbone of storage infrastructures in most enterprises—by the time the data is analyzed, its inherent value has decreased and its cost has increased.

With in-memory computing, all relevant data is compressed and maintained in the main dynamic random access memory (DRAM)—in silicon—rather than stored in slower physical disk storage. This means that data can be accessed so quickly that in-memory databases can run transactional and analytic applications on the same infrastructure. This can eliminate the need for, and cost of, a separate data warehouse. This also means that data is available for analysis as soon as it is generated, and complex analytical queries can be completed in real time and the results funneled back into transactional applications or into further analytics to test out scenarios. With immediate access to troves of relevant data, both truly predictive and robust historical analytics are now greatly improved.

With in-memory computing solutions, a single database can perform both OLTP and OLAP functions—in fact, users can draw on both row and columnar data at the same time to get answers that pull from both processing functions simultaneously, and at blinding speeds.

In-memory computing is the culmination of multiple technical advances. These include the integration of traditional relational row store processing with columnar data processing, and the use of in-memory columnar compression, a set of technologies that speeds the scanning of massive data sets and analytical querying while employing a smaller memory footprint.

Relational databases store data in row-based tables, because it’s the most efficient format for frequent updates required by transactional applications. Columnar data processing is a much faster technology for scanning through very large unstructured data sets and performing analytical querying. What’s unique with in-memory computing solutions is that a single database can perform both functions—in fact, users can draw on both row and columnar data at the same time to get answers that pull from both processing functions simultaneously.

Most in-memory solutions also provide some level of data compression, which greatly increases the amount of data that can be held in memory, making it more immediately available to the server processor for faster scanning. Columnar data is much easier to compress than row-oriented data, and compression levels can be as high as 10-to-1.

Columnar data processing is especially powerful for analytical querying, particularly when supported by Intel® Advanced Vector Extensions (Intel® AVX) and SSE (Streaming SIMD Extensions) instructions built into the Intel® Xeon® processor E7 v2 family. This enables in-memory solutions to tap into the massively increased performance potential of highly parallelized, multicore processing by packing more data elements into the register of a single processor and dividing query processing into multiple threads that work simultaneously. In addition, placing columnar data in SSE registers enables processors to use memory pools much more efficiently than row-based stores, because the in-memory solution can run queries and evaluate data while it is still compressed.

### Big Data, Massive Scalability and In-Memory Integration

Many businesses already generate and store large volumes of data, and integrating big data storage and analytics is a fundamental requirement of data hub architectures.

Big data results from very large flows of unstructured data generated by billions of connected devices and other intelligent data sources in the rapidly expanding Internet of Things. Analysis of this unstructured data can reveal patterns, connections and
relationships that provide valuable business insights, fuel predictive analytics, and improve competitiveness. Big data also results in petabytes of data that must be stored and analyzed.

While it’s possible to build a real-time, in-memory business platform with petabyte scalability, it’s neither practical nor necessary for most businesses.

Apache Hadoop® is the de facto standard for managing huge data flows and analyzing massive, unstructured data sets. Hadoop offers a cost-effective storage solution for integrating very large data volumes with in-memory computing environments. It offers an open source solution for ingesting, preparing, and storing warm data for inclusion in the real-time analytics environment. Users can see the data in Hadoop as an extension of the in-memory database, and queries are automatically federated across both platforms. The combined solution supports real-time analytics acting on petabytes of data.

Integrating Hadoop with in-memory computing platforms is key to optimizing data throughput capability versus cost across all enterprise data types and all business requirements.

**The Data Hub Architecture**

The data hub model builds on the innovations of in-memory computing. It combines massively parallel processing power and a highly performance-optimized platform that enables entire databases to operate in memory. Combined with unstructured data access via Hadoop, the data hub is a vastly scalable platform with the power and flexibility to run a wide variety of workloads, including transactional batch processing, unstructured data analytics, and business intelligence applications, with all of an enterprise’s data immediately available for processing (although this white paper advocates structuring a data hub architecture upon an in-memory database, it is not an absolute requirement; traditional databases optimized with advanced multi-threaded and multicore processors can also support the data hub model).

The data hub also offers robust security, with encryption and data protections, plus data governance controls and management.

**Data hub architectures based on Intel® technologies**

represent a paradigm shift in how enterprises manage and use data, enabling continuous real-time analytics that deliver improved insights into business dynamics and customer sentiment and behaviors.

An in-memory data hub represents an enormous shift in the way businesses can manage and use data. The speed and scale of in-memory technologies allows businesses to host transactional and analytical applications on the same database and process data with near zero latency. Operational data is available for analysis as soon as it is generated, and complex queries can be completed in seconds, enabling real-time analytics.

Accessing an organization’s data through a single, converged data hub offers abundant and measureable business benefits. Maintaining data in an in-memory database does away with the need to store multiple copies of data sets in multiple formats for processing in stand-alone computing environments. Having a single copy of data available for both transactional and analytical processing saves operational and capital expenses, reduces infrastructure complexity and IT management overhead, which can lead to dramatic savings in total cost of ownership.

Having immediate access to accurate data and information as it occurs fundamentally changes how organizations do business, allowing organizations to make faster, smarter decisions and act in real time to help ensure optimal outcomes.

**High Performance Foundation for a Real-Time Data Hub**

Intel offers a cost-effective, high performance server platform with the memory capacity, parallel execution resources, system bandwidth and advanced reliability features required to address the demands of powerful, mission-critical data hub environments. Implemented with optimized in-memory database software solutions, these technologies open the benefits of the data hub architecture to an expanding range of industries.

Data hub technologies and products from Intel and its partners include:

**Intel® Xeon® Processor E7 v2 Family.** The Intel Xeon processor E7 v2 was designed with in-memory computing in mind. These processors provide up to double the performance over the preceding generation, not only speeding time to insight but also enabling businesses to process twice as many queries in a given time frame. Intel Xeon processor E7 v2 also supports major increases in memory capacity, up to three times as much as the previous generation. These processors deliver a range of high-performance capabilities for business processing, analysis, and technical computing.
Intel® Run Sure Technology.
Available on the Intel Xeon processor E7 v2 family, Intel Run Sure Technology maximizes server uptime by boosting reliability, availability, and serviceability (RAS). Resilient System Technology is included, which integrates processor, firmware, and software layers to help diagnose fatal errors, contain faults, and automatically recover to keep the server operating. Another feature group, Resilient Memory Technologies, helps ensure data integrity within the memory subsystem.

Intel® Xeon Processor E5 v3 family.
The Intel Xeon processor E5 v3 family provides exceptional compute power and data throughput for rapid processing of large or complex data sets. These new processors add 50 percent more cores and cache over the previous generation and provide performance improvements up to 3X. They help address the growing demands placed on computing infrastructure, from supporting business growth, enabling new services, delivering new applications in the enterprise, to extending workloads into the cloud. This processor is a cost-effective solution for the Hadoop ingestion portion of the data hub architecture.

Intel® Advanced Vector Extensions (Intel® AVX). Intel AVX is a 256-bit instruction set extension that improves processor performance due to wider vectors, new extensible syntax, and rich functionality. This results in faster, improved data management for applications such as analytics, imaging, audio/video processing, scientific simulations, financial analytics and 3D modeling.

Intel® Networking Technologies.
Intel® Ethernet Gigabit networking technologies provide the scalable, high-throughput features required to meet the I/O demands of the massively data-intensive environment of a converged data hub. With Intel Ethernet Controllers achieving up to 40 gigabits per second, Intel networking technologies offer significant benefits over traditional storage architectures, including a 45 percent reduction in power per rack, a significant reduction in infrastructure costs, and twice the server I/O bandwidth.1

Intel® Solid State Drives Data Center Family (Intel® SSDs). Intel SSDs for data center applications complement the in-memory capabilities of the data hub server platform by supporting low-latency, high-bandwidth persistent storage. They offer full end-to-end data protection, consistent performance with low latencies, AES 256-bit encryption for excellent data security, and high capacities for growing storage needs. Intel SSDs can withstand large amounts of data writes; they produce less heat and noise than HDDs; and they typically have about one-half of the power requirements of hard disc drives.6 The new Intel Solid State Drive Data Center Family for PCIe® incorporates NVM Express® (NVMe) technology for substantial performance gains and improved processor utilization, with up to six times faster data transfer speed than 6 Gbps SAS/SATA SSDs.7

Intel® Cache Acceleration Software (Intel® CAS). Intel CAS maximizes performance of enterprise applications and eliminates I/O bottlenecks with very low effort and upfront costs. Intel CAS employs a data caching policy that allows the user to select which data type should be accelerated, providing better control and performance optimization. This “hot data” is placed on an SSD cache, sized appropriately to fully utilize its capacity, thus reducing the storage system latency and significantly improving performance—up to three times more performance on transactional database processing and up to 20 times faster processing of read-intensive business analytics.8 Intel is also working with Cloudera to optimize the Cloudera Enterprise Data Hub® for integration with Intel Xeon processors and in-memory databases to seamlessly extend Hadoop’s highly scalable storage capacity to data hub architectures.

An in-memory data hub architecture offers businesses opportunities for revenue growth by providing the insights to create valuable new services and by delivering predictive analytics to anticipate customer needs.

Empowering Real-Time Business
Data hub architectures based on Intel® technologies represent a paradigm shift in how enterprises manage and use data, enabling continuous real-time analytics that deliver improved insights into business dynamics and customer sentiment and behaviors. Supported by in-memory database technologies, the data hub helps ensure that transactional, analytical and unstructured data is integrated into the same processing structure for maximum intelligence gathering and reduced complexity.

This provides businesses with opportunities for revenue growth by providing the insights to create valuable new services and improved products that demonstrably address the needs of customers. Armed with real-time market insights and predictive analytics, businesses can anticipate customer needs with more personalized products and spur greater innovation due to faster, more comprehensive insights into market demands.

Data integration is not the only critical issue—with in-memory data hubs handling more and heavier workloads, data integrity and system uptime is more important than ever.
Servers based on the Intel Xeon processor E7 v2 family provide mission critical reliability to rival RISC architectures, but with greater flexibility, and reduced total cost of ownership. Intel Xeon processor E7 v2 family-based servers offer up to an 80-percent performance advantage at up to 80 percent lower four-year total cost of ownership over servers based on IBM POWER7+. 

Intel® processor-based data hub architectures also deliver cost savings and efficiencies to drive rapid ROI. The modern server platforms, networking technologies and storage solutions that drive data hubs are much more efficient than previous generations of data center equipment, requiring less power to operate and to cool.

These modern technologies also offer reduced operational costs through simplified, centralized IT management, and reduced capital expense through more efficient, less redundant data storage. Data hub architectures based on Intel products and technologies also offer improved insights into operational efficiencies through management interfaces that monitor and optimize performance.

For more information on Intel processor-based data hub architectures, go to www.intel.com/xeon.