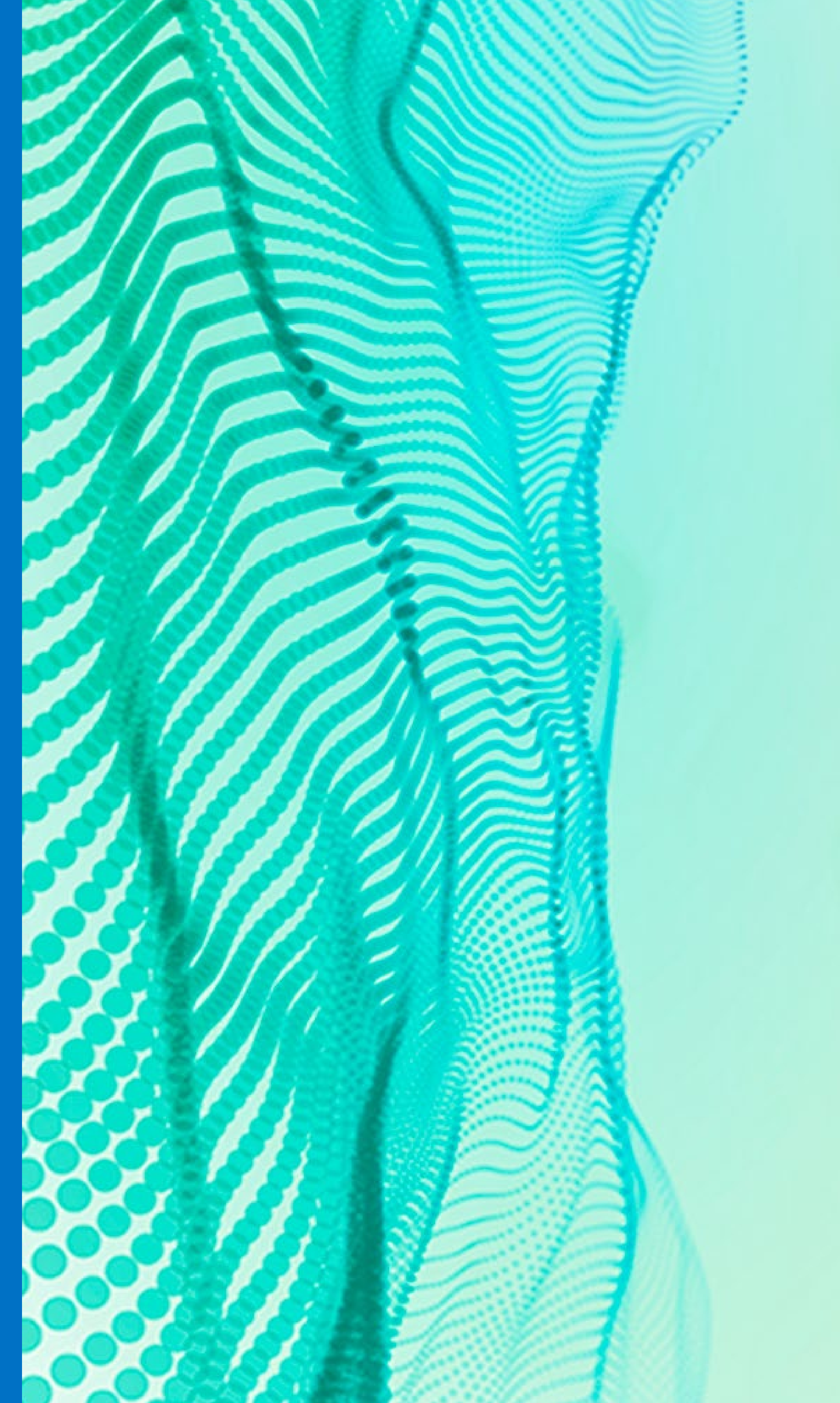
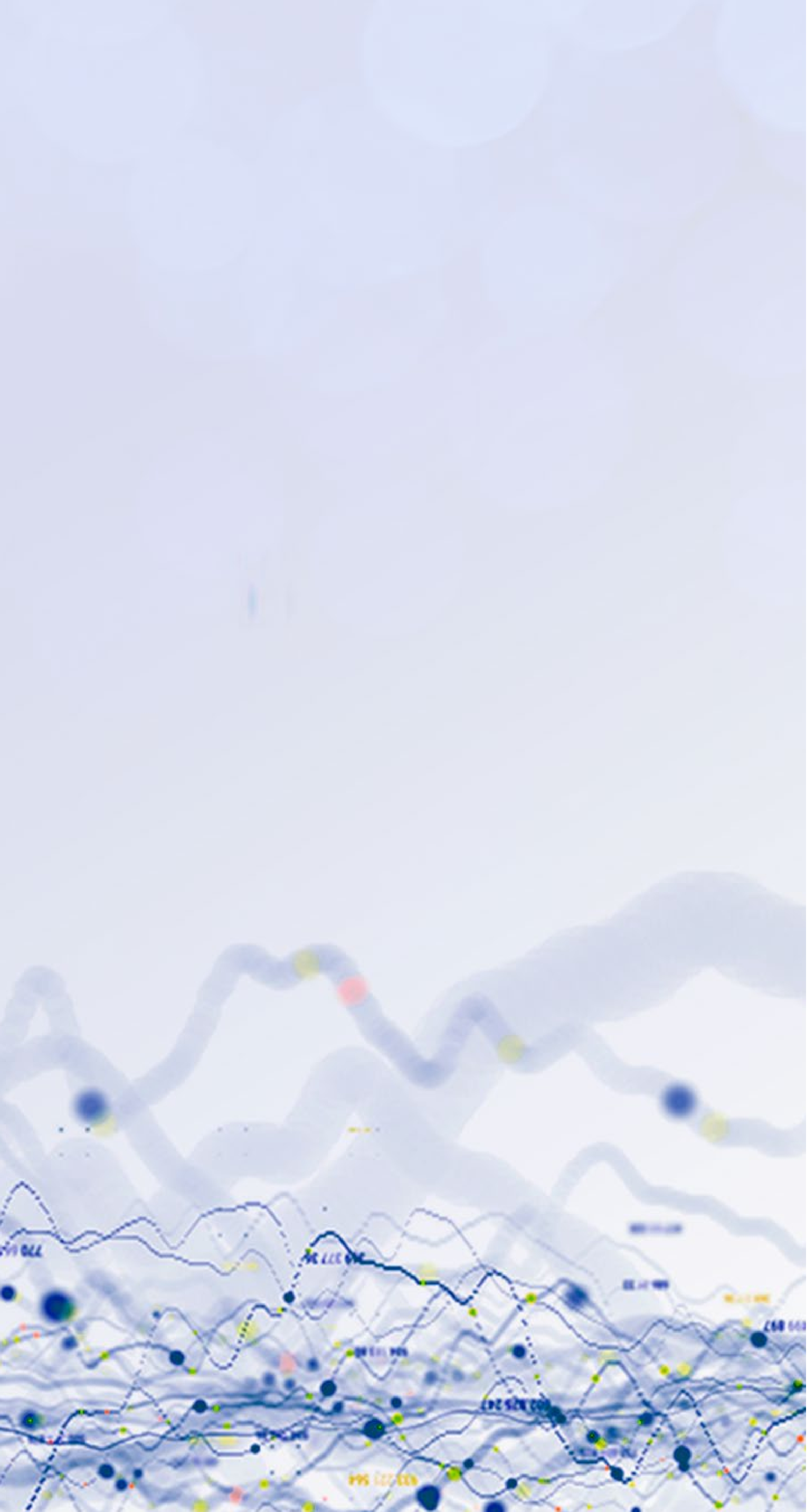
BUILD YOUR IN-MEMORY ANALYTICS STACK

CONTENTS

IN BRIEF

In-memory databases have moved from being an expensive option for exceptional analytics workloads, to delivering on a far broader set of goals. Using real-world examples and scenarios, in this guide we look at the component parts of in-memory analytics, and consider how to ensure such capabilities can be deployed to maximize business value from both the technology and the insights generated.





THE (R)EVOLUTION OF IN-MEMORY ANALYTICS

The reactive 'rear view mirror' analytics practices that seemed ground-breaking a few years ago can no longer satisfy today's demand for real-time insight. Organizations across multiple industries are adopting more complex analytics use cases (like predictive analytics and AI) that will help them identify and respond to trends and insights in real-time, or even in advance.

One of the drivers for this demand is the fact that our ability to capture and process large volumes of data has increased in recent years, while the cost involved has gone down. IDC predicts that the global datasphere will grow from 33 zettabytes (ZB) in 2018 to 175 ZB by 2025[1], and as our ability to ingest large volumes of data increases, so does the need to extract value from it. Hardware technology advancements, coupled with new analytics and AI algorithms and software help us gain more, complex insights faster and adopt real-time, data-centric decision making.

These developments mean a much broader choice of tools is now available, which can be chosen based on each use case. For example, more traditional approaches can still be fine for batch processing business intelligence (BI) data. However, for real-time workloads like fraud detection, customer service, or patient monitoring an IT infrastructure that supports real-time data processing and analytics is essential. This is where the application and impact of in-memory computing is quickly growing.

In-memory analytics brings a range of advantages to the enterprise, including the ability to reduce IT costs and simplify infrastructure with more uptime; lower latency; greater synchronization of data (compared to older models which kept data in separate containers); reduced data copies, and faster development cycles. Moreover, with its ability to deliver faster, deeper insights at scale and in real time, it helps create new opportunities to drive and enhance service delivery.

As its name suggests, in-memory computing (or in the context of this paper, in-memory analytics) is about moving data up the storage tiers and closer to the CPU to enable faster access for more rapid analytics, and reduced time-to-insight. However, being able to implement and scale this capability often means making architectural changes to your underlying IT infrastructure.

WHAT DOES MEMORY REALLY HAVE TO DO WITH ANALYTICS CAPABILITIES?

As memory technologies have evolved, they have enabled more data to be handled at a lower cost, which in turn drives today's more data-heavy workloads. However, fast memory technologies like DRAM are still more costly than lower storage tiers like NAND, SSDs and hard drives, and tend to increase their cost exponentially as you add more until you hit a 265GB-per-slot ceiling after which no more capacity can be added. This means that when real-time analytics applications need to keep large volumes of data close to the CPU, memory capacity can quickly become a bottleneck. At the same time, retrieving the data from lower storage tiers creates too much latency, creating a poor cost/performance ratio.

The challenge can be addressed using recent innovations in storage and memory – for example Intel® Optane™ SSDs for the data center, which include a non-volatile memory technology that offers DRAM-like performance with the persistence of SSD storage. This can help create a smoother continuum between memory and storage, and balance the cost/performance ratio for large-scale in-memory workloads like advanced analytics and AI. It also offers developers creating new AI and analytics applications more choice and support from their hardware platform.

As a next step in this evolution, Intel has just launched Intel® Optane™ DC persistent memory with the latest generation Intel® Xeon® Scalable processor, connecting this advantage directly to the CPU memory bus. It offers the unprecedented combination of high capacity, affordability and persistence. By moving and maintaining larger amounts of data closer to the processor, workloads and services can be optimized to reduce latency and enhance performance.

USE CASES FOR IN-MEMORY ANALYTICS

Clinical Healthcare: In-memory analytics can provide the ability to monitor patient safety in real-time, personalize patient care, assess clinical risk, and to reduce patient readmission – improving organizational efficiency and the patient experience. Process large, memory-hungry files like 3D MRI scans and analyze them in conjunction with patient data from other sources like electronic medical records (EMR).

Retail: In-memory analytics can improve real-time inventory tracking and understanding of how customers act in-store or online, driving operational efficiency, higher sales volumes, better security, and enhanced customer satisfaction. Being able to understand customers on an individual level allows for real-time ad targeting, and personalized in-store and online experiences.

Financial Services: Use in-memory analytics to aggregate transactional data to improve business health and reveal insights about customer behaviors, and to inform real-time decisions about stock trading and fraud detection. Support real-time services like mobile money transfer.

Call Center Analytics and Customer Churn: In-memory analytics can process call center logs, online interactions, interactive voice response, and IT support systems to model customer behaviors and predict churn. Predictive models can also be used to help increase staff retention and reduce recruitment overheads.

Smart Operations Supply Chain: Analytics-driven predictive maintenance enables servicing with minimal disruption and cost, optimizing supply chains and forward planning for spare parts and raw materials. In-memory analytics can be used to power core systems like SAP HANA* to deliver status updates in real time, and to provide customers with real-time order tracking all the way from warehouse to delivery.

REAL-WORLD PERFORMANCE AND TCO IMPROVEMENTS WITH INTEL® OPTANE™ DC PERSISTENT MEMORY

Evonik* strives to be a worldwide leading speciality chemicals company. It produces sophisticated chemicals for a variety of customer needs including 3D printing, tires, and sustainable farming. Understanding the changing needs of customers and the complex supply chain that supports them can only be done with real-time analytics and reporting.

The organization depends on innovative technology to keep up with its infrastructure needs. In order to support real-time insights, those needs included creating greater in-memory database capacity without sacrificing time or cost.

The organization endeavors to achieve the next level in total cost of ownership (TCO) efficiency, and as part of this approach it chose to implement – with support from Accenture® – servers using Intel® Optane™ DC persistent memory.

Previously, Evonik's option to increase memory capacity was limited to investing in larger servers. However, with Intel Optane DC persistent memory, it can now invest in persistent memory modules. This gives Evonik the flexibility to integrate data sets into its SAP HANA* platform more efficiently. Whether data sets grow over time or by acquisition, Evonik is now well positioned to store such massive data sets.

In a recent proof of concept, Evonik found that with Intel® Optane™ DC persistent memory, it saved time during data table reloads after the server was restarted. This allows for shorter maintenance windows for SAP HANA patching or configuration change. In a stable SAP HANA environment with a large memory footprint that supported both Intel Optane DC persistent memory and DRAM, Evonik also achieved a lower TCO. With less server downtime for its SAP HANA systems, there is also more time for productivity.

UNDER THE HOOD OF IN-MEMORY ANALYTICS

As demand for in-memory analytics increases, the IT leader is challenged to extract more value from larger data sets than ever before. It's therefore important to consider the architecture that underpins your in-memory operations and how it meets your needs now and in the future.

Your architectural requirements will depend on the maturity of your organization's capabilities and experience. A traditional architecture may use an operational database to hold and manage current data, which is processed in batches before being directed to the organization's data warehouse for analytics. However, in many enterprises this model is giving way to a single hybrid transactional and analytical processing (HTAP) architecture, which enables you to run analytics on transactional data, eliminating the data hops that come with batch processing and data warehouses. This supports some of the key platforms and applications that benefit from in-memory analytics, such as SAP HANA*.

If considering HTAP, you should also consider which flavor of it will be most appropriate for your business. Point-of-decision HTAP, for example, uses the same underlying infrastructure for both operational and analytics workloads (such as is the case with SAP HANA), enabling you to run analytics reports directly on the operational database. However, in-process HTAP takes things a step further, embedding analytics capabilities into the transaction itself to impact its outcome – for example in use cases such as dynamic ad placement on a website, or real-time payment fraud detection.

ENABLING DYNAMIC AD PLACEMENT WITH INTEL® XEON® SCALABLE PROCESSORS

The digital advertising industry is an increasingly complex and high-stakes environment. As consumers use more devices to access content – from mobile phones and tablets to connected TVs – it's essential for advertisers to ensure they're getting the right ad to the right person, at the right time, and on the right device.

ZypMedia* offers national advertising products to local companies, and wanted to ensure that with around 500,000 transaction requests a second, needing a half-millisecond response time, it had the speed and flexibility to deliver this high quality of service.

It worked with Aerospike*, a database company committed to helping its customers tackle Internet-scale problems while retaining the flexibility of an on-premise solution. Running its software on Aerospike's database and Intel® Xeon® Scalable processors, ZypMedia achieved excellent performance results.

CONSIDERATIONS WHEN BUILDING YOUR IN-MEMORY ANALYTICS STRATEGY

Keep in mind that wherever you're using your data (from Hadoop* software, to column-based open source like Cassandra*, Oracle* databases, or your SAP HANA platform), the way you consume the data should determine the technology you use to hold and manage it. The type of approach you choose should depend on the business need, and it's important to bear in mind that more traditional approaches remain an option when your analytics requirements are not so time- or memory-intensive.

When embarking on your strategy development, follow these rules of thumb to ensure all your bases are covered:

1. Avoid the temptation to be swayed by tactical initiatives or passion projects by prioritizing the development of a **holistic and long-term business data strategy**. Make sure all those involved understand the business objectives and success criteria, as well as the data needed to support the decisions and analytics that will help meet them.

2. Carefully evaluate the **IT variables, requirements and criteria** that will impact how you ingest, store and consume your data to meet your business goals. This may include considerations like:

a. Legacy applications – which business-critical software do your users rely on, and how will it be integrated into your in-memory analytics strategy without impacting performance, availability or security?

b. Workload placement – most organizations today operate a hybrid multi-cloud model, with different workloads and applications running across a mix of on-premise data center, local or hosted private cloud and public cloud. Make sure you have clear and repeatable criteria for deciding which workloads should be run in which environment, and that your in-memory analytics workloads are vetted as part of this process.

c. Agility and time-to-market – use cases for in-memory analytics focus on delivering real-time insight and action, so it can be a useful approach when agility and quick delivery to market are important. Less time-sensitive operations could more cost effectively be handled using batch processing or other more traditional methods.

d. Capacity – as with any advanced analytics workload, in-memory requires huge memory capacity in order to deliver on the promise of real-time insight and action. Make sure your storage and memory architecture is designed with these requirements in mind, and consider technologies like Intel Optane DC persistent memory to cost effectively address the capacity, performance and persistency needs of your hot data.

e. Cadence – look at each data element that you're bringing in and treat it according to how it will be used – each has its own characteristics. For example, some data elements may need to be processed at a high frequency, on an ongoing basis, but may be relatively low volume. Others may come in huge volumes but can be processed in batches to make them manageable. These different models will have different hardware and software requirements though, so you need to think holistically and make sure you're covering all bases.

3. Ensure you have the required **skills within your workforce** to implement, execute and monetize your in-memory analytics objectives. You may choose to do this by training up existing team members or hiring in new skills.

4. Follow some core **guiding principles for in-memory database architectures:**

a. Create a distributed environment that allows for separate components such as CPU, memory and storage to be added as needed over time.

b. Ensure your environment is relational and multi-modal, able to support SQL databases and semi-structured data such as JSON*.

c. Enable mixed media support to ensure that as data ages it can be moved down from memory to storage tiers to help optimize data management costs without compromising latency or analytics performance.

APPLYING THESE PRINCIPLES TO REAL-WORLD SCENARIOS

1 Every in-memory analytics deployment will differ depending on the way the organization responds to the points outlined above. However, it can be instructive to consider some concrete examples.

Let us take a financial services organization looking to do fraud detection and prevention using in-memory analytics. A real-time use case like this needs high-performing systems that enable the organization to ingest the data about a particular transaction as it is going on and then run analytics on it against the customer's historic transactions and behavior patterns.

The historical data may be stored on a warm storage layer such as Intel Optane SSDs or NAND-based SSDs that offer high performance on processing. Meanwhile, the live transaction data could be processed initially through the point of sale. The edge infrastructure can be arranged to perform some data cleansing and filtering (for example removing any cancelled transactions that contain no useful insights), and even analysis, before the critical data is passed to a centralized repository. The organization may choose to use CPU-based edge devices to handle this process, or – if the data needs to be compressed and/or encrypted, or if power availability at the edge is low – a field programmable gate array (FPGA) may be a better fit.

Once the live transaction hits the centralized repository (cloud or on-premise), the organization can compare the live and historical data to identify points of similarity or discrepancy, with your in-memory analytics application helping identify these patterns.

This is where high-performance, low-latency hardware is critical to enable near instantaneous results. Make sure your server platform offers high-capacity, scalable and cost-effective memory to support this need.

2 As a second example, let's consider a healthcare organization diagnosing and treating cancer patients. In this scenario, regulatory obligations around data privacy and security are likely to out-rank performance and latency requirements (although these will of course also be important). From an IT infrastructure perspective, this means focusing more on storage and memory than compute performance and network speed.

This challenge is compounded by the sheer size of much of the data the organization holds, with a single MRI scan being as large as multiple gigabytes in size. Being able to quickly and significantly scale your capability to handle and analyze data in real-time therefore is crucial in order to deliver instantaneous insights that underpin diagnosis and treatment decisions.

The system must also be able to support processing peaks when it is ingesting huge volumes of new files, meaning burst capacity must be worked into the architectural plan. This is generally difficult with a GPU-based system, which tends not to offer the memory capacity headroom needed to run these large-scale in-memory analytics workloads. In this case, a CPU-based platform with enhanced memory capabilities – such as the latest generation Intel® Xeon® Scalable processor with Intel Optane DC persistent memory – can help by offering increased capacity to scale during times of peak demand.

INTEL® OPTANE™ DC PERSISTENT MEMORY FOR IN-MEMORY ANALYTICS

- **Powerful computing:** Delivered with the next generation Intel® Xeon® Scalable processors, helping transform system performance for large, memory-bound data challenges like real-time analytics and decision making. For example, 2nd Generation Intel Xeon Scalable processors featuring Intel Optane DC persistent memory deliver 2.4x better performance running SAP* HANA* 2 compared to a five-year-old system[2].
- **Fast storage that performs like fast memory:** Persistence and large capacity accelerate the entire analytics lifecycle – data ingestion, preparation, storage, discovery, development, and deployment – to make scalable in-memory analytics a reality.
- **Move fast, store more, process everything:** Keep significantly larger data sets closer to the processor to drive down latency and extract more value for less. Run in-memory databases faster, for example, with over six times the transaction throughput in Oracle TimesTen*[3].
- **Restart sooner and reduce wait times:** Rapid restart times for in-memory databases compared to DRAM-only cold restarts and reduced wait time help boost efficiency so critical business decisions can be made and actioned faster.
- **Reduce system downtime:** Persistent memory removes the need to read data from disk into RAM in a system restart, so applying security patches and other administrative tasks can be done faster. This helps reduce system downtime (and its resultant delays or missed SLAs) by accelerating time to recovery to maximize system uptime and service delivery.
- **Lower total cost of ownership (TCO) for analytics workloads:** Achieve 36 percent[4] more virtualized machines (VMs) for multi-tenant virtualized online transaction processing (OLTP) databases and 30 percent[5] lower estimated hardware cost per VM using 2nd Generation Intel Xeon Scalable processors with Intel Optane DC persistent memory.

DELIVERING IN-MEMORY ANALYTICS

Whether you're embarking on your first in-memory analytics initiative, or have done it before, this section is designed to help you think through some of the steps that will help make your next project a success.

1. Prepare Your People

- Start with the business objective and let that be your guide. Ensure you're only bringing in in-memory analytics if it makes sense for the organization. You'll be generating insights so make sure someone is there to benefit from them.
- Involve stakeholders from across the business to get ideas and perspectives that will help you identify the best project to start with.
- Pick business and IT collaborators who are willing to collaborate and drive change. The best laid plans can be stymied by culture or resistance to change.

2. Prepare Your Project

- Identify low-hanging fruit to prove business value quickly, and keep it simple. Identify a high return on investment (ROI) business use case that can benefit from real-time analytics and start with that. You can scale and/or add complexity later, once you have demonstrated value and built support for the initiative.
- Explore around your chosen use case to identify opportunities to create a ripple effect. Ideally use cases may be related, for example appearing in the same line of business, or by having shared data needs. Try to pick a project that could be easily scaled or replicated if it goes well.

3. Prepare Your Data

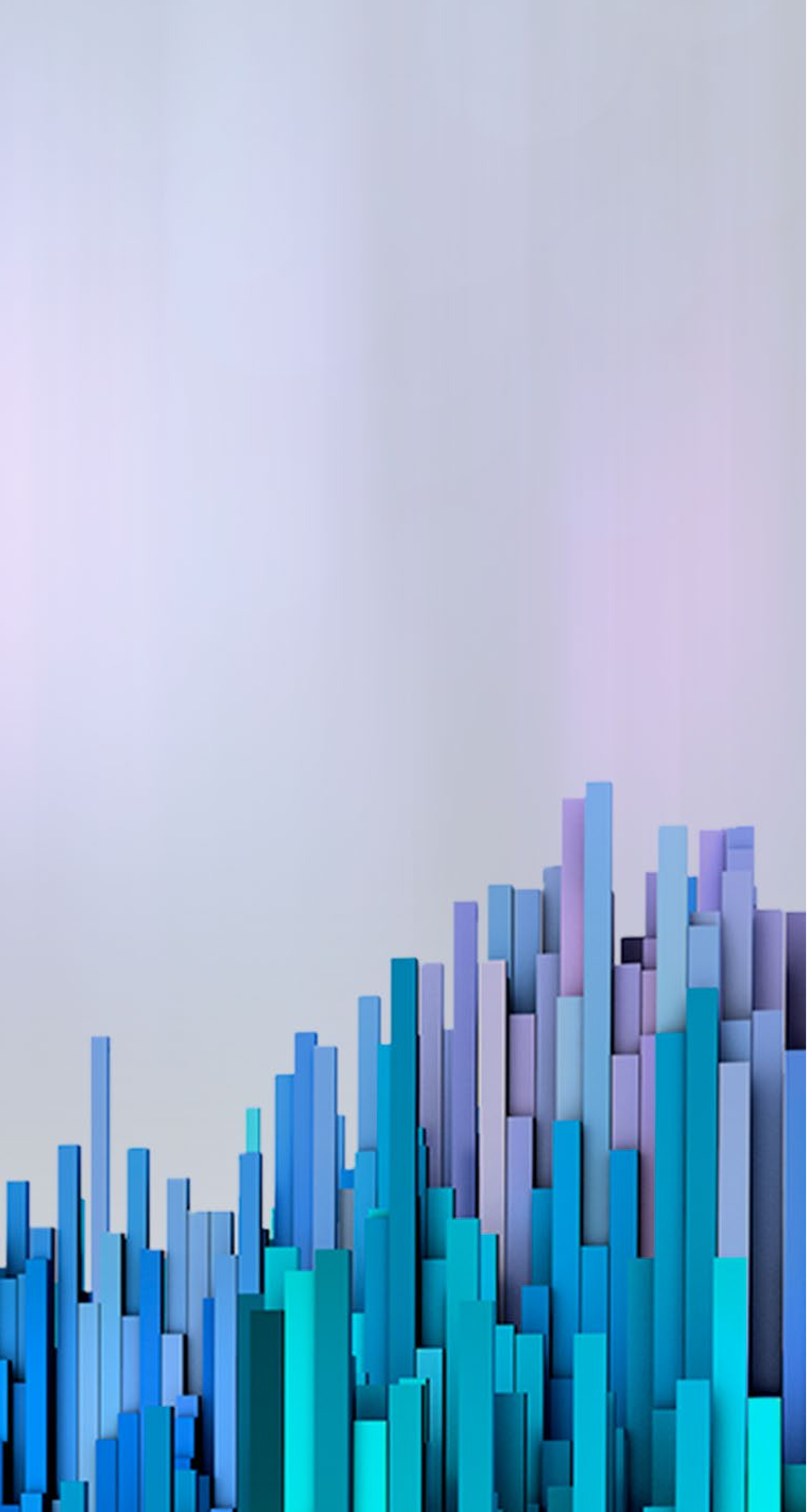
- Identify a data consumption model. Will your users need self-service capabilities to access insights on-demand, and splice them multiple ways? Or will automated reports on a regular cadence do the job?
- Understand the data elements that will be needed to meet the business requirement. What data types are you dealing with? How often is the data generated, ingested, processed, and analyzed? What business processes and applications must it intersect with?
- Invest in building your foundational data model, ideally by getting everyone together to whiteboard and sketch things out. Think about scalability – will your proposed model work tomorrow just as well as today? What other features might you need to add as business needs evolve?
- Consider your data governance and management strategies – who will own it? How and when will it be cleaned and updated? How will it be secured, and can you demonstrate compliance with regulations and laws?

4. Plan Your In-Memory Analytics Environment

- Study the software ecosystem available to you. A wide range of open source and proprietary applications are available to support in-memory analytics use cases, so you may be able to adopt something that already exists. If not, make sure you have access to the development skills that will help you refine or build your ideal app.
- Select your vendors for key infrastructure pieces. Run pilots or proofs of concept to ensure your chosen solutions will work in your real world.
- Consider how you will implement your in-memory analytics capabilities, depending on your organizational needs and the maturity of your IT infrastructure.
- You may consider options such as bare metal servers; virtualized machines and hyper-converged infrastructure; containers with orchestration, or the cloud.

5. Deliver

- Work in increments, and don't expect big bang results overnight. Things will not always go to plan and there will be hiccups along the way, but treat these as opportunities to learn and optimize for the next cycle.
- Agree and put in place measures for reporting on the performance of your in-memory analytics initiatives, and maintenance of the systems and algorithms supporting them. It's likely you will need to revise and adjust your prediction models over time in order to maintain the accuracy and quality of the insights you derive.



IN-MEMORY ANALYTICS AND INTEL

Intel's range of data center technologies are designed and optimized to simplify the complexity of in-memory analytics. The latest generation Intel Xeon Scalable processors, with Intel Optane DC persistent memory deliver the performance, persistence and capacity required to improve service delivery by accelerating data-driven insights. These technologies transform system performance for I/O intensive queries, for example by delivering up to 8x speedup improvement with IO-intensive queries using Intel® Optane™ DC persistent memory and hard drives, compared to DRAM and SSDs[6], while helping lower data replication latency and enabling the activation of larger 'working' data sets in memory. The introduction of persistence in memory on the CPU also helps accelerate in-memory database start times.

By using Intel Optane SSDs as part of your storage architecture as well, you can expand the size of your data sets beyond DRAM's limits while breaking through performance bottlenecks, boosting throughput and cost-effectively optimizing compute resource. This can help deliver the rapid scalability and flexibility often demanded by in-memory analytics workloads.

In addition to our continued commitment to hardware innovation, we work closely with our ecosystem to develop fully integrated, benchmarked software and hardware stacks designed to support specific workloads for faster service delivery and time-to-value. Intel® Select Solutions for Microsoft SQL Server* are optimized for in-memory advanced analytics use cases like online transaction processing (OLTP).

We will always generate more data than we can process, so it's important to focus any data strategy on maximizing the value of our data resources. Increasingly, this means enabling real-time analytics use cases through innovation in algorithms, software and the hardware that underpin them. Organizations planning to make the most of these developments will benefit from prioritizing and maximizing their investments in in-memory analytics.

Benchmarks show you can accelerate time-to-insight and enhance your customers' experience with the latest Intel® technology for real-time analytics applications.

- Gain up to 3.7x higher online transaction processing (OLTP) performance (HammerDB*) with 2nd Generation Intel® Xeon® Platinum 8280 vs a five-year-old system[7].
- Achieve up to 2.3x faster time-to-insights (BigBench*) with 2nd Generation Intel® Xeon® Gold 6248 vs a five-year-old system[8].
- Recent HiBench* benchmarks found it delivered up to 4.3x higher performance than a five-year-old system[9].

FURTHER READING

- Solution Brief: Intel® Select Solutions for Microsoft SQL Server* Business Operations >
- Video: Intel & SAP Executive Interview (Lisa Davis & Chris Hallenbeck): SAP HANA* Revolutionizes Data Management with Intel® Optane™ DC Persistent Memory >
- Blog: Access More Data with Intel® Optane™ DC Persistent Memory and SAP HANA* 2.3 Extension Nodes >

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 complete information visit www.intel.com/benchmarks

Performance results are based on testing as of the date set forth in the configurations and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure.

[1] The Digitization of the World From Edge to Core, November 2018, IDC and Seagate, https://www.seagate.com/files/www-content/our-story trends/files/idc-seagate-dataage-whitepaper.pdf

[2] Performance results are based on testing by Intel IT as of March 12, 2019 and may not reflect the publicly available security updates. No system can be absolutely secure.

Baseline: three-node (1-master + 2-slave) SAP HANA 2 scale-out configuration. Per Node: 4x Intel® Xeon® processor E7-8880 v3 (2.3 GHz, 150 W, 18 cores), CPU sockets: 4; Microcode: 0x400001c, RAM capacity: 64 x 32GB DIMM, RAM model: DDR4 2133 Mbps; storage: GPFS, approximately 21.8TB of formatted local storage per node, SAN storage for backup space only; network: redundant 10GbE network for storage and access, redundant 10G network for node-to-node; OS: SUSE 12 SP2, SAP HANA: 2.00.035, GPFS: 4.2.3.10. **Average time of 50 individual test queries executed 30-50 times each, for a total of approximately 25,000 steps: 2.81 seconds.**

New configuration, one-node (1-master) SAP HANA 2 scale-up configuration: CPU: 4 x 2nd Generation Intel® Xeon® Platinum 8260 processor (2.2 GHz, 165 W, 24 cores), CPU sockets: 4; Microcode: 0x400001c, RAM capacity: 24 x 64GB DIMM, RAM model: DDR4 2133 Mbps; Intel Optane DC persistent memory: 24 x 126GB PMM; storage: XFS, 21TB; network: redundant 10GbE network; OS: SUSE 15, SAP HANA: 2.00.035, Intel BKC: WW06. **Average time of 50 individual test queries executed 30-50 times each, for a total of approximately 25,000 steps: 1.13 seconds.**

[3] 6.49x improvement performance results are based on Oracle TimesTen IMDB, running TPTPM benchmark. Hardware configuration for baseline and persistent memory systems is 2x Intel® Xeon® processors with 28 cores, 2 threads per core. BIOS version1.0134. Storage is 2x6TB SSD DC P4608. Operating system is Red Hat Enterprise Linux 7.5, Kernel 4.18. Baseline system has 1,536Gb DDR4 quad rank @2666MHz. Persistent memory system has 192GB DDR4 dual rank @2666MHz and 6TB of Intel® Optane™ DC Persistent memory. Date of test = March 15th, 2019.

[4] Configurations: 1-node, 2x 26-core 2nd Gen Intel Xeon Scalable Procesor, HT on, Turbo on, 768GB, 0(24 slots /32GB / 2666 DDR)1x Samsung PM963 M.2 960GB, 7 x Samsung PM963 M.2 960GB, 4x Intel SSDs S4600 (1.92 TB), 1xIntel X520 SR2 (10Gb), Windows Server 2019 RS5-17763, OLTP Cloud Benchmark, test by Intel as of 1/31/2019 vs. 1- node, 2x 26-core 2nd Gen Intel Xeon Scalable Procesor, HT on, Turbo on, 192GB, 1TB(12 slots / 16 GB / 2666 DDR + 8 slots /128GB / 2666 Intel Optane DCPMM), 1x Samsung PM963 M.2 960GB, 7x Samsung PM963 M.2 960GB, 4x Intel SSDs S4600 (1.92 TB), 1x Intel X520 SR2 (10Gb), Windows Server 2019 RS5-17763, OLTP Cloud Benchmark, test by Intel as of 1/31/2019.

[5] Baseline: 1-node, 2x Intel Xeon Platinum 8276 cpu, DRAM - 384GB (12x32GB); Total System Cost=$34931[cpu cost=$17438, Memory subsystem cost – Total Cap: 768GB (384GB /Socket)=$8993, Storage cost=$7200, Chassis, PSUs, Bootdrive etc. =$1300]; Total VMs per system=22, Cost per VM=$1587 vs. DCPMM Config: 1-node, 2x Intel Xeon Platinum 8276 cpu, DCPMM- 512GB (4x128GB AEP + 6x16GB DRAM, 2-2-1, Memory Mode, 5.3:1); Total System Cost=$33243 [cpu cost=$17438, Memory subsystem cost - Total Cap: 1024GB (512GB/ Socket)=$7305, Storage cost=$7200, Chassis; PSUs; Bootdrive etc.=$1300]; Total VMs per system=30, Cost per VM=$1108.

[6] Baseline: 1-node, 2x Intel® Xeon® Platinum 8280L CPU @ 2.70GHz processor on S2600WF (Wolf Pass) with 768GB DDRGB (DDR Mem: 24 slots / 32GB / 2666 MT/s) total memory, ucode 0x0400000a running Fedora release 29 kernel Linux-4.18.8-100.fc27. x86_64-x86_64-withfedora-27-Twenty_Seven, and 9 decision support I/O intensive queries, storage is HDD (ST1000NX0313) * 8, 0- Gigabit SFI/SFP+ network connection. Source: score: geomean baseline 1. Tested by Intel, on 24 Feb 2019, vs. New configuration: 1- node, 2x Intel® Xeon® Platinum 8280L CPU @ 2.70GHz processor on S2600WF (Wolf Pass) with 192GB DDR + 1TB Intel® Optane™ DC persistent memory GB (DDR Mem: 2 slots / 16GB / 2666 MT/s + 8 slots / 128GB / 2666 MT /s) total memory, ucode 0x0400000a running Fedora release 29 kernel Linux-4.18.8-100.fc27. x86_64-x86_64-withfedora- 27-Twenty_Seven, and 9 decision support I/O intensive queries, storage is 8x HDD (ST1000NX0313), 10- Gigabit SFI/SFP+ Network connection. score: geomean speedup 8x. Tested by Intel, on 24 Feb 2019.

[7] 1-node, 2x Intel(R) Xeon (R) CPU E5-2697 v2 on Canoe Pass with 256 GB (16 slots / 16 GB / 1866) total memory, ucode 0x42d on RHEL7.6, 3.10.0-957.el7.x86_64, 2 x Intel DC P3700 PCI-E SSD for DATA, 2 x Intel DC P3700 PCI-E SSD for REDO, HammerDB 3.1, HT on, Turbo on, result: OLTP Warehouse workload=2.24M, test by Intel on 2/1/2019 vs. 1-node, 2x Intel® Xeon® Platinum 8280 CPU on Wolf Pass with 384 GB (12 slots / 32 GB / 2933) total memory, ucode 0x4000013 on RHEL7.6, 3.10.0-957.el7. x86_64, 2 x Intel SSD DC P4610 for DATA, 2 x Intel SSD DC P4610 for REDO, HammerDB 3.1 on commercial database, HT on, Turbo on, result: OLTP Warehouse workload=8.45M, test by Intel on 2/1/2019.

[8] BigBench* claim of 2.3X: 1+4-node, 2x Intel® Xeon® processor E5-2697 v2 on S2600JF with 128 GB (8 slots / 16GB / 1866) total memory, ucode 0x42d on CentOS-7.6.1810, 4.20.0- 1.el7.x86_64, 1x 180GB SATA3 SSD, 3 x Seagate ST4000NM0033 (4TB), 1 x Intel I350, TPCx-BB v1.2 (not for publication) / 3TB/ 2 Streams, Mllib, Oracle Hot-Spot 1.8.0 _191, python-2.7.5, Apache Hadoop-2.9.2, Apache Spark-2.0.2, Hive 2.2 + CustomCommit, HT on, Turbo on, result: queries per min=265, test by Intel on 1/24/2019 vs. 1+4-node, 2x Intel® Xeon® Gold 6148 processor on S2600WF with 768 GB (384 GB used) (12 slots* / 64 GB / 2400 (384GB used)) total memory, ucode 0x400000A on CentOS-7.6.1810, 4.20.0-1.el7.x86_64, Intel SSD DC S3710, 6 x Seagate ST2000NX0253 (2TB), 1x Intel X722, TPCx-BB v1.2 (not for publication) / 3TB/ 2 Streams, Mllib, Oracle Hot-Spot 1.8.0 _191, python-2.7.5, Apache Hadoop-2.9.2, Apache Spark-2.0.2, Hive 2.2 + CustomCommit, HT on, Turbo on, result: queries per min=622, test by Intel on 1/12/2019.

[9] Geomean of SparkKmeans, SparkSort, SparkTerasort. 1+4-node, 2x Intel® Xeon® processor E5-2697 v2 on S2600JF with 128 GB (8 slots / 16GB / 1866) total memory, ucode 0x42d on CentOS-7.6.1810, 4.20.0-1.el7.x86_64, 1x 180GB SATA3 SSD, 3 x Seagate ST4000NM0033 (4TB), 1 x Intel I350, HiBench v7.1 / bigdata, Mllib, OpenJDK- 1.8.0_191, python-2.7.5, Apache Hadoop-2.9.1, Apache Spark-2.2.2, , HT on, Turbo on, result: SparkKmeans=119.5M, SparkSort=121.4M, SparkTerasort=107.4M, test by Intel on 1/23/2019 vs. 1+4-node, 2x Intel® Xeon® Gold 6248 processor on S2600WF with 768 GB (384 GB used) (12 slots* / 64 GB / 2400 (384GB used)) total memory, ucode 0x400000A on CentOS-7.6.1810, 4.20.0-1.el7.x86_64, Intel SSD DC S3710, 6 x Seagate ST2000NX0253 (2TB), 1 x Intel X722, HiBench v7. 1 / bigdata, Mllib, OpenJDK-1.8.0_191, python-2.7.5, Apache Hadoop-2.9.1, Apache Spark-2.2.2, HT on, Turbo on, result: SparkKmeans=1235.8M, SparkSort=518.4M, SparkTerasort=589.3M, test by Intel on 1/23/2019.

Optimization Notice: Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice Revision #20110804

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com

Intel, the Intel logo, Xeon, and Optane are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

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

© Intel Corporation 0519/MMK/CAT/PDF 338680-002EN