Plan your Analytics Evolution

Ready your organization to take advantage of emerging trends and technologies to drive insights from analytics

Gartner expects the democratization of data and analytics to accelerate through 2023[1]

Over the next three years, Gartner predicts[1] that the democratization of analytics expertise will accelerate, giving employees across organizations access to advanced analytics capabilities that will pull valuable, applicable insights from data. While the role of the data scientist remains critical in helping to build and design the enterprise's complex analytics models, new technologies are making it possible for non-specialists to self-serve insight. In this way, intelligence is becoming embedded and ubiquitous at all levels of the organization – from the data scientists' lab to the office, factory floor or retail store.

Enterprises that have embraced these analytics capabilities are already experiencing benefit. From helping retail store associates know which product to offer to a given customer at the right time, to enabling sales reps to identify the most promising accounts across their region, small day-to-day decisions informed by analytics can add up to have a significant impact on the business. A number of these tools also have the advantage of automating some of the more repetitive and mundane tasks that fall to data scientists, enabling them to focus their precious time on more value-adding projects.

Giving employees the ability to control how and where these analytics tools are used is critical in order to stay competitive.

Advanced analytics for the business user

Bringing analytics capabilities to the masses can be done in a number of ways, using a variety of tools. In this paper we will explore a few new and emerging technologies that can be used, and steps to ensure you're equipped to implement them.

The first is augmented analytics, a term coined by Gartner in 2017[2]. This is the use of artificial intelligence (AI) techniques like machine learning and natural language processing (NLP) to automate the process of cleaning and analyzing a company's raw data to convert it into actionable insights. Today, if a business user needs to ask a question of the company's data using analytics, they may often need to work with a data scientist, who understands how to parse the question from the business user into a format the analytics algorithm will understand, then translate the results back out for the user. With augmented analytics, not only is the data scientist freed from the tedious job of cleaning the data and parsing the questions themselves, but business users can also gain more detailed insights, in near-real time. This approach is on the rise, with Gartner predicting that it will be ubiquitous by 2022[3].

Closely tied to augmented analytics is augmented data management, which applies AI and machine learning to all aspects of maintaining data quality as well as identifying new ways to drive value from existing data – for example giving metadata, traditionally mainly used for audit, lineage and reporting, a more active role in uncovering patterns in data usage.

Another trend fast gaining traction is conversational analytics. Indeed, Gartner's prediction here is that by 2021, it will help boost analytics and business intelligence adoption from 35 percent of employees to over half, including new classes of user – especially front-office workers[4]. Using a combination of natural language processing (NLP) and analytics means business users, conversational analytics enables users to ask their questions in a way that's natural to them, just as they would using a search engine online. Then the AI algorithm does the work of parsing the question so the analytics model can be run and the results returned, again in a way the business user can understand. Conversational analytics takes this a step further, enabling the user to ask their questions verbally, as they would with a home assistant device.

Tech trends supporting analytics evolution

These trends all share the objective of delivering richer, more timely insights to the business user. Achieving these goals means enterprises should consider a couple of key developments in IT infrastructure.

The first is the shift towards in-memory processing for larger analytics workloads in the data center, driven by the ongoing need to handle larger data sets with increased speed. As the need for real-time insights grows, IT teams are challenged to enable more data to be processed closer to the CPU. This means holding more active data in-memory, which is becoming more achievable. However, IT teams must consider how to architect for more in-memory processing. An increasingly compelling option is to introduce more persistent memory into the environment.

Advanced analytics workloads are also driving the need for many organizations to rethink their storage media and strategy. Advanced analytics can offer new insights from stored data but if that data is residing on storage sub-systems with slow performance and unpredictable service levels then the potential value cannot be realized. Capacity scaling should also be considered in designing the storage system. Older storage technologies that are not space- and operationally efficient can quickly become a drag on the enterprise, especially as the desire for more and more data accelerates. As a result, IT teams are pressed to allocate increasing amounts of data to the highest tiers of their storage hierarchies so it can benefit from fast, flash-like speed and performance.

Advanced analytics at the edge is also increasingly in demand. Enabling algorithms to run on end user devices makes the edge more intelligent, and so helps avoid backhauling large volumes of data to the cloud or the data center. This means use cases like conversational analytics can be supported with the low latency and high performance that deliver the real-time response users expect. Ensuring the enabling hardware, software and connectivity are in place is an important first step to rolling out this type of analytics.

Evonik shortens maintenance windows and reduces TCO

Leading specialty chemicals company Evonik works with a large and complex supply chain to meet the needs of its global customers. To keep track of all the moving parts and stay ahead of any challenges, it runs SAP HANA, for real-time analytics and reporting. This is based on its SAP HANA in-memory database, and it needed to increase capacity, without sacrificing time or cost.

Evonik ran a proof of concept (PoC) using Intel® Optane™ persistent memory and found that data table reloads after server restarts could be achieved 17x faster[5]. This in turn allows for shorter maintenance windows for SAP HANA patching or configuration as well as lowering overall total cost of ownership (TCO).

17X FASTER
data table reloads after server restarts

Intel® technology enables analytics innovation

Intel has developed a suite of hardware, software, and expertise, supported by a broad solutions ecosystem, to help its customers meet their advanced analytics goals. When it comes to delivering on the promise of democratized insights, building on your existing Intel® architecture can help you optimize performance and accelerate insights without the need to invest heavily in complex new solutions. As it provides support for every stage of the data pipeline, it is possible to scale this capability from existing resources, avoiding creation of data analytics siloes on incompatible systems.

Optimize performance

Advanced, real-time analytics workloads need strong performance, huge memory capacity, and seamless scalability to grow in line with the business and its data. The 2nd generation Intel® Xeon® Scalable processor offers a world-class platform with unparalleled memory footprint for industry standard servers. This enables scalable performance for a wide variety of analytics workloads, including data-intensive in-memory analytics platforms like SAP HANA, Microsoft SQL or Oracle DB. Meanwhile, scale-up and scale-out flexibility support your business evolution.

These capabilities enable IT teams to increasingly deploy larger pools of in-memory analytics to support real-time decision making. Reliability, availability and serviceability (RAS) features also help ensure huge workloads can run with confidence while maintaining data integrity, and data encryption based on hardware-enhanced platform security helps provide efficient data protection at rest and in transit. With strong performance across common in-memory analytics platforms, a platform built on the Intel Xeon Scalable processor also helps accelerate transactions and speed time to results for users.

As part of its commitment to support customers' evolving analytics and AI needs, Intel is launching the 3rd generation Intel® Xeon® Scalable processor in mid-2020, evolving our 4 to 8-socket processor foundation for today's data-intensive digital services. With support for bfloat-16 built-in, the processor will extend the built in AI acceleration of Intel® Deep Learning Boost with unprecedented support for AI training and inference, helping lay a foundation for AI as well as offering enhanced scalability and performance for advanced analytics workloads.

Accelerate actionable insights

Designed to support complex, real-time workloads like augmented and conversational analytics, Intel® Optane™ technology, which includes persistent memory and storage drives, offers a unique combination of affordable large memory footprints and support for data persistence, enabling fast data access and greater workload consolidation. With Intel Optane persistent memory, delivered with 2nd gen Intel Xeon Scalable processors, you can access data faster and optimize CPU utilization. This persistent memory technology provides greater capacity for larger data sets while also addressing the need for real-time analysis at fast speeds.

Delivered with the 3rd gen Intel Xeon Scalable processor, Intel® Optane™ persistent memory 200 series will offer massive capacity and over 225x faster CPU access to persistent data than reading from a conventional SSD[6], helping further accelerate time to insights.

Modern storage systems, such as Intel® Optane™ SSDs deliver high, symmetrical reads/writes with amazingly low latency under load and high quality of service (QoS) to accelerate data into compute so you can analyze data more quickly. Meanwhile, Intel® 3D NAND SSDs store massive amounts of data in a space and operationally efficient manner to scale with growing storage needs.

SIEMENS

Siemens AG enables cost-effective data growth

Siemens AG develops a range of solutions for consumers and industry, including power generation and distribution, intelligent infrastructure and distributed energy systems. To achieve all this, it relies on the ability to process and analyze its data very quickly, and cost effectively.

In order to help support these goals long-term, Siemens AG has implemented 2nd generation Intel® Xeon® Scalable processors with Intel Optane persistent memory to support its digital growth while also minimizing TCO.

"Intel® Optane™ persistent memory offers new infrastructure sizing, security, scalability….with lower TCO. Digitalization and innovation across Siemens requires faster processing of the huge and ever-growing data volumes," said Mamun Natour, responsible for technology and security for application and digitalization platforms at Siemens AG.
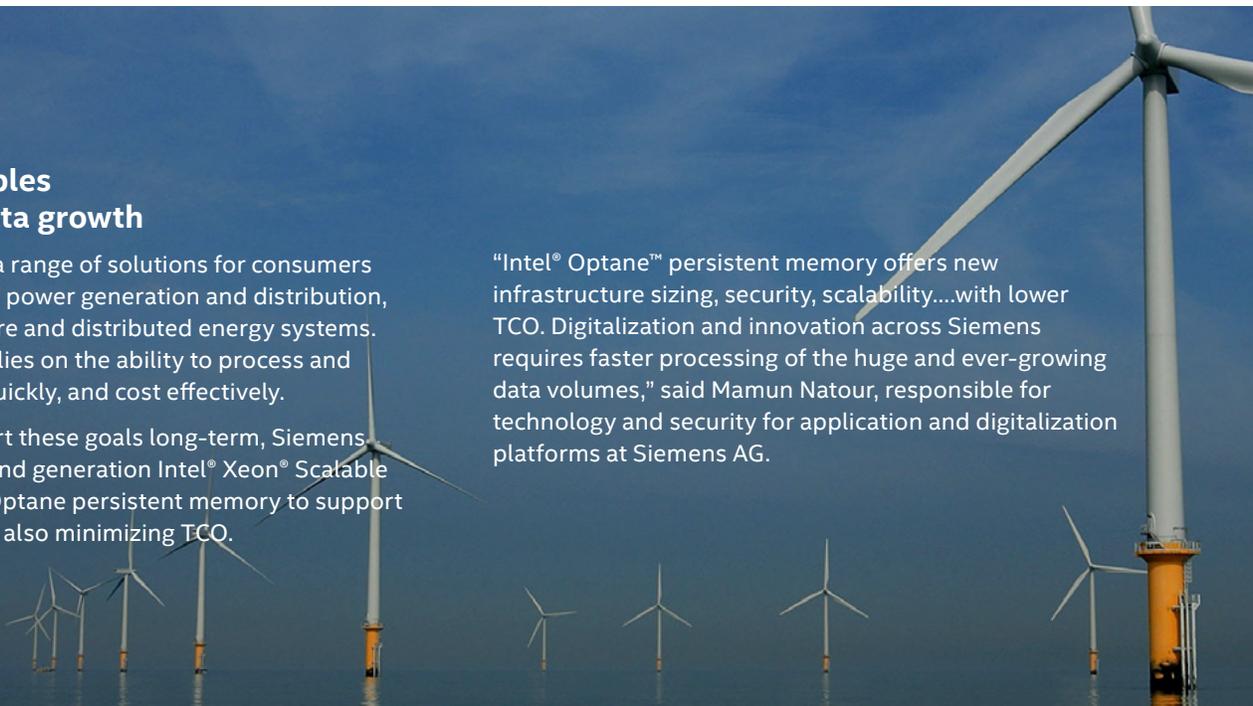
By 2025, 75 percent of data will be created at the edge[7]. The faster this data is ingested, analyzed and moved, the sooner it can deliver insights, so enabling analytics at the edge is increasingly important. Intel offers an expansive portfolio of processing, software tools, and networking technologies to give organizations the choice and flexibility they need to scale seamlessly from edge to cloud. Intel's ecosystem of collaborators and ready-to-implement solutions enables companies to deploy more nodes of intelligence at the edge of their operations, helping to amplify productivity, increase revenue, take smarter risks, and launch new business models quickly.

Simplified solutions

Introducing advanced analytics capabilities for the first time can be a challenge for already stretched IT teams. Intel works closely with the analytics solution provider ecosystem to deliver pre-defined and verified infrastructure solution stacks that help alleviate the process. With a choice of hardware and software elements, you can accelerate and simplify implementation while building a solution that's right for you.

Intel's ecosystem includes hundreds of solutions and support for a wide range of open source analytics, AI and digital transformation initiatives, as well as popular platforms like Oracle Exadata, SAP and Splunk. This ecosystem offers innovative in-memory analytics and AI capabilities, such as Anaconda for open source data science and machine learning based on Python and Data Robot, which enables automated AI for the enterprise. Intel has also worked with many organizations in this space to optimize their applications for Intel Optane persistent memory to deliver enhanced consolidation and performance improvements. This range of options provides enterprises looking to take advantage of analytics capabilities with the choice and flexibility to find the most appropriate solution for them.

Meanwhile, Intel® Select Solutions for advance analytics are workload-optimized to deliver high performance, price performance and improved security offered through Intel's solution provider partners. These solutions currently include:

- **Intel® Select Solutions for Microsoft SQL Server:** Scale out multitenant database-as-a-service solutions for a wide variety of analytical and operational workloads.

- **Intel® Select Solutions for SAP HANA:** A validated and SAP-certified solution with 2nd gen Intel Xeon Scalable processors and Intel Optane persistent memory.

Intel® Technology Differentiators

Optimized performance: World-class, highly integrated 2nd generation Intel® Xeon® Scalable processors deliver scalable performance for a range of advanced analytics applications, including in-memory platforms like SAP HANA or Microsoft SQL, to support real-time decision making and enhanced reliability. With the upcoming 3rd generation Intel Xeon Scalable processor, these capabilities are extended with additional speed frequencies to enhance performance.

Accelerated insights: Intel® Optane™ technology delivers a unique combination of affordable large capacity and support for data persistence, enabling fast data access and greater workload consolidation. This enables fast handling of large data sets, and quick data analysis while helping reduce storage costs.

Simplified solutions: A rich ecosystem of solution providers and support for open source platforms combined with –workload-optimized and verified Intel® Select Solutions, enable you to get up and running quickly.

Learn more

- **eGuide:**
  From Data to Insights – Optimize the Four Stages of Your Data Pipeline to Make Your Business Analytics-Driven

- **eGuide:**
  Build Your In-memory Analytics Stack

- **Reference Architecture:**
  High Performance Data Analytics with Splunk

- **Solution Brief:**
  Intel® Select Solutions for Microsoft SQL Server

- **Solution Brief:**
  Intel® Select Solutions for SAP HANA

- **White Paper:**
  Transforming Intel's Security Posture with Innovations in Data Intelligence

- **Webpage:**
  Intel® Select Solutions for Advanced Analytics

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

[1] https://www.gartner.com/en/newsroom/press-releases/2019-10-21-gartner-identifies-the-top-10-strategic-technology-trends-for-2020

[2] Hype Cycle for Emerging Technologies, 2017. https://www.gartner.com/document/3768572?ref=solrAll&refval=239846505

[3] Predicts 2020: Analytics and Business Intelligence Strategy, https://www.gartner.com/document/3978987?ref=solrAll&refval=239843443

[4] Gartner Top 10 Data and Analytics Trends, https://www.gartner.com/smarterwithgartner/gartner-top-10-data-analytics-trends/

[5] SAP HANA* simulated workload for SAP* BW edition for SAP HANA* Standard Application Benchmark Version 2 as of 30 May 2018. Baseline configuration with traditional DRAM: Lenovo ThinkSystem SR950* server with 8x Intel® Xeon® Platinum 8176M processors (28 cores, 165 watt, 2.1 GHz). Total memory consists of 48x 16GB TruDDR4* 2,666 MHz RDIMMs and 5x ThinkSystem* 2.5" PM1633a 3.84 TB capacity SAS 12Gb hot-swap solid-state drives (SSDs) for SAP HANA* storage. The operating system is SUSE* Linux* Enterprise Server 12 SP3 and uses SAP HANA* 2.0 SPS 03 with a 6TB dataset. Average start time for all data finished after table preload for 10 iterations: 50 minutes.

New configuration with a combination of DRAM and Intel® Optane™ persistent memory: Intel Lightning Ridge SDP with 4x CXL QQ89 AO processor (24 cores, 165W, 2.20 GHz). Total memory consists of 24x 32GB DDR4* 2666 MHz and 24x 128GB AEP ES2, and 1x Intel® SSD S3710 Series 800GB, 3x Intel® SSD P4600 Series 2.0TB, 3x Intel® SSD Series S4600 1.9TB capacity. BIOS version WW33'18. The operating system is SUSE* Linux* Enterprise Server 15 and uses SAP HANA* 2.0 SPS 03 (a specific PTF Kernel from SUSE was applied) with a 1.3TB dataset. Average start time for optimized tables preload (17x improvement). Testing conducted by Evonik and Accenture on May 30, 2018.

[6] Intel® Optane™ persistent memory: Results based on Intel testing on February 20, 2019. Configuration: Intel® C620 series chipset, 28-core Intel® Xeon® Scalable processor (QDF QQYZ), 2,666 megatransfers per second (MT/s), 256 GB, 18 W, 32 GB DDR4 DRAM (per socket), 128 GB Intel® Optane™ persistent memory (per socket), firmware: 5336, BIOS: 573.D10, WW08 BKC, running Linux* OS4.20.4-200.fc29*. Performance tuning quality of service (QoS) disabled, IODC=5(AD).

[7] https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders/

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.

Intel does not control or audit third-party data. You should review this content, consult other sources, and confirm whether referenced data are accurate.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation.

Your costs and results may vary.

Intel, the Intel logo, and other Intel Marks 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 · 0420/IP/CAT/PDF · 342828-001EN