swim

WHITE PAPER

Sink or Swim: Transforming Boundless Data into Continuous Intelligence

visit us at **swim.ai**

Contents

Executive Summary	01
So Much Data, so Little Insight	01
Data Growth Fuels Digital Transformation	02
Boundless Data Requires a New Approach	02
Transforming Boundless Data Streams into Continuous Intelligence	04
Swim Enables Continuous Intelligence at Scale	05

swim

Executive Summary

Continuous Intelligence is the next big evolution in software architecture that helps organizations contend with the inundation of information while deriving timely, meaningful insights from boundless data streams. It transcends today's big data and streaming data silos, replacing the limitations of batch-based processing and latency-prone database queries with an "analyze-then-store" model, where computation is stateful, driven by data and enables analysis, learning and predictions from dynamic and static data on the fly.

Whether today's organizations ultimately sink or swim in data depends on their ability to transform boundless streams of data into timely, contextual insights relevant to their business operations. Their applications need to continuously listen to data and always have an answer. Traditional approaches of processing in batch or storing raw, often only ephemerally useful data for later analysis is too slow. Organizations need to continuously and statefully analyze new data, avoiding database limitations, and process, analyze and visualize all data in concert.

So Much Data, so Little Insight

According to IDC's latest Global Datasphere forecast, there will be **175 zettabytes of data** in use by 2025 (see figure 1) – the equivalent of filling about 500 million personal computers or laptops with new data every day. Hundreds of new devices are connected to the Internet every second, further driving up the amount of data generated.

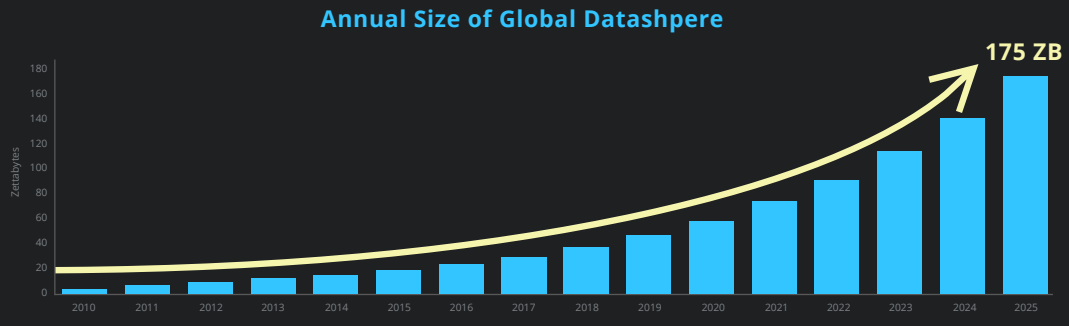


Figure 1: Annual Size of the Global Datasphere (source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere)

For over a decade, British and American military leaders have lamented that they are "swimming in sensors and drowning in data." Credited with coining this phrase, retired General David Deptula (USAF) has spoken out about sensory overload, especially when working with real-time data, large data sets and fast decisions, but the problem applies to a much broader audience. Businesses across every industry - including communications, transportation, logistics, manufacturing and ecommerce - are investing in real-time data analytics and AI capabilities to speed insights, discover trends and be more responsive to dealing with market disruptions while improving customer experience.

Data Growth Fuels Digital Transformation

Organizations are drowning in the huge growth of streaming live data, accelerating the need for "digital transformation" – the process of changing the way we live our personal and professional lives, and how business gets done by using digital technologies.

Digital transformation is profoundly impacting consumers, companies and entire industries worldwide, every day. We live in a data-driven world, which means new data is produced, streamed and consumed everywhere, all the time. Most data streams, e.g. those originating from sensors, clickstreams or other data sources, sometimes referred to as 'the edge', are boundless; they never stop producing or consuming data.

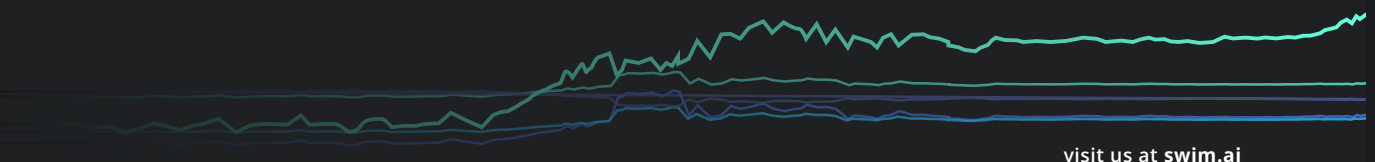
But more importantly, most streaming data is only ephemerally useful, which means its ability to generate insights related to the 'Now' evaporates very quickly, after which it just turns into stale data. One might wonder, given the availability of a large number of modern big data analytics technologies and tools, how companies are dealing with this dilemma. In fact, one of the biggest concerns data leaders have in collecting data is that no one in their enterprise will ever use all of that data, or understand its value, to drive business decisions.

Boundless Data Requires a New Approach

In business, there is a critical need for 'continuous intelligence' – the ability to collect, analyze and act on both real-time (dynamic) and relational (contextual) data as quickly as the data is generated. Global research and advisory firm Gartner has recognized this need and recently noted that by 2022, more than half of all major new business systems will incorporate continuous intelligence using real-time context data to improve decisions.

Yet, there are three significant challenges - time, complexity and efficiency - with existing software architectures and technologies for deriving valuable insights and continuous intelligence from large volumes of data:

- **Time**: Traditional application infrastructure relies on a store-then-analyze approach, which prioritizes data storage over data processing and data analysis delaying critical insights necessary to identify trends and take timely action. To make things worse, many data sources are currently stored or processed in individual siloes, making it difficult to draw useful conclusions across multiple types of data in context.



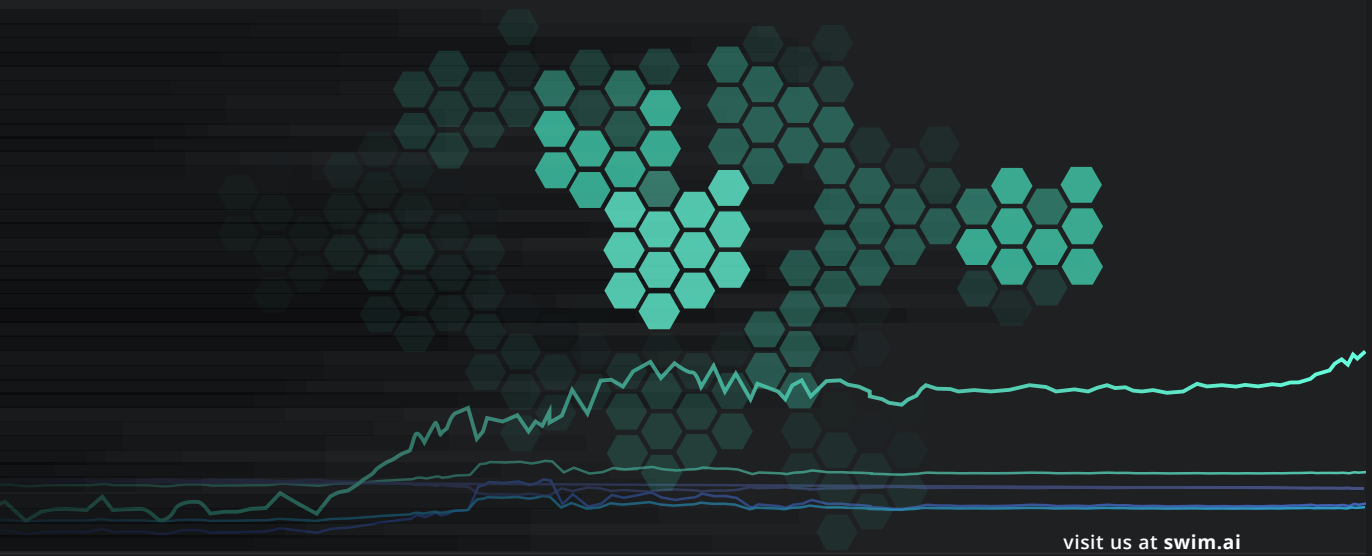
- **Complexity**: Analysts are overwhelmed with massive volumes of data and have little ability to identify critical events, derive immediate insights and make decisions as their data changes, when situational awareness and operational responsiveness matter most.

- **Efficiency**: More data are produced or consumed than can be efficiently stored or processed in real-time. To compensate, users often constrict data outputs by periodic sampling or processing only select data sets for historical analysis – often hours, days, weeks, or even months later. The resulting insight delay reduces or eliminates the principal value of real-time and rich, contextual data.

With today's big data analytics approaches, most business-critical data (such as ERP data, CRM data, etc.) exist and are generated in one central location – the enterprise datacenter or the cloud (see figure 2). All data is first stored there in databases, data warehouses or data lakes, and then scheduled for later processing and analysis. Big Data analytics are extremely sophisticated. You can aggregate, reduce, transform, group, correlate, which means you can have a rich set of capabilities for understanding your static business data. The challenge is that the static data reflects only past events and struggles to account for what's happening in the business right now.

As data sources started to proliferate and more and more data was generated at the 'edge', event streaming allowed applications to consume this data by subscribing to topics that were relevant to them. According to Gartner, by 2022 more than 50% of all enterprise data will be generated at the edge. But for analytical purposes, most of the streaming data is 'shipped' to a central location – typically the cloud -, where it is stored and then scheduled again for later processing and analysis. Streaming analytics approaches will then look at events coming in, apply filters or perform some simple transformations, but streaming analytics really are designed to understand solely that streaming picture. They aren't necessarily designed to incorporate the static, contextual data that represents what the business is trying to achieve. Streaming analytics don't account for that business context.

Both with big data analytics and streaming analytics, which tend to be siloed in most IT departments, the analysis of data and generation of insights happens on a 'store first, then analyze' basis. But neither on-premises nor cloud-hosted "store-then-analyze" architectures can deliver fast responses for continuous, real-time situational intelligence.



Transforming Data Streams into Continuous Intelligence

Continuous Intelligence represents the next step in the Digital Transformation journey. With Continuous Intelligence, data processing and analysis can happen anywhere – at the edge, in the cloud, or anywhere in between; the key is to perform any analysis, learning and predictions closest to the data. More importantly, with Continuous Intelligence both static and dynamic data are processed together for contextual analysis. And since most streaming data is of ephemeral value only, Continuous Intelligence applies an "analyze-then-store" paradigm, processing incoming live data first and then storing it – or part of it - only if required (e.g. for compliance reasons) or useful for future purposes. Continuous Intelligence does not mean that persistence becomes an afterthought, but it is no longer on the performance path, either; Continuous Intelligence does not involve latency-prone database queries typical for (mini-)batch processing to drive contextual, actionable insights at every moment.

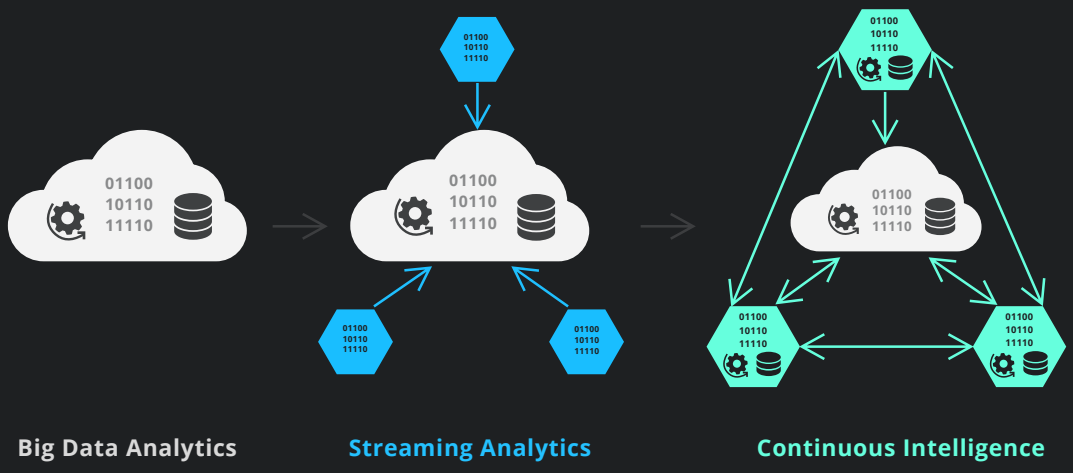


Figure 2: Continuous Intelligence - Overcoming the Store-then-analyze Paradigm (source: Swim)

It is important to remember that even today's most performant databases are updated by clients and queried by users. They don't drive (re-)computation of insights when the state of individual data sources or of the entire system model changes. To transform boundless data streams into Continuous Intelligence, we need to ensure that applications are always ready to respond with analysis or predictions based on the newest data, so performing analysis only when applications are polled by user actions is inadequate. The arrival of new data should be used to trigger all re-evaluations that might be needed, so that applications can deliver a continuous stream of insights or predictions at any time and to all users.

To be clear, Continuous Intelligence is not meant to solve every data analytics problem an organization may have. It is about high-frequency analytics, enabling always-on 'listening' to the data, monitoring, analyzing, learning and projecting outcomes, not about explorative adhoc analytics. Its focus is on continuously monitoring changes and driving responses for known conditions, rather than looking for the needle in the haystack, i.e. defining what questions to ask the data, and only then formulate queries for later reports.

swim

Swim Enables Continuous Intelligence at Scale

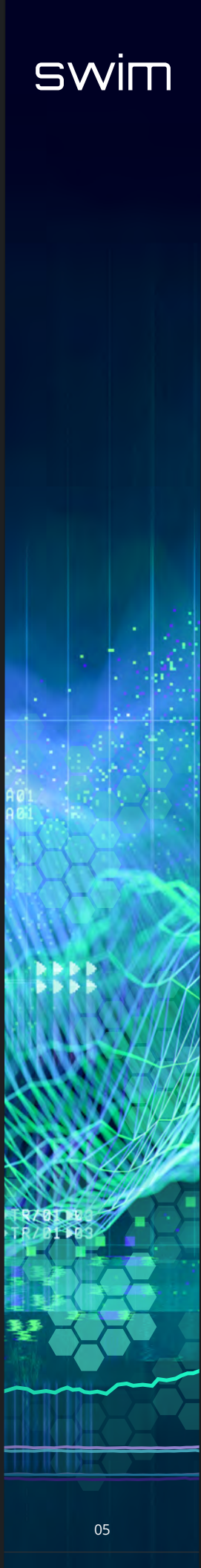
Swim is software in motion. Just as continuous motion of objects doesn't leave any gaps, Continuous Intelligence from Swim never misses any insights. Processing real-time and contextual, historical data in concert, Continuous Intelligence from Swim provides always-on, continuous data integration, analytics, visualization and automated real-time responses, integrated into business operations and providing ongoing decision support.

Swim's unique architecture enables dynamically distributed applications that are fast and efficient at massive scale. Today, a Fortune 100 company is using Swim to monitor and manage the health of their nationwide network in absolute real time. The application spans millions of real-world data sources - processing over 4 petabytes of data per day with results in under a second. For comparison, that's the same amount of data as watching over 10,000 4K ultra-high definition videos every day. By using Swim Continuum, the network service provider has quickly achieved enterprise-grade, production-ready management of **Continuous Intelligence** applications at scale.

Swim makes it easy to develop, deploy and run applications that statefully process streaming and batch data on-the-fly, using an "analyze-then-store" architecture; thus, Swim applications are constantly able to drive operational decisions that demand high resolution, contextual computation. So when your business needs to know, it already does. No need to query databases first to generate relevant and timely insights.

About Swim

Swim offers Swim Continuum, the first open core, enterprise-grade platform for building, managing and operating continuous intelligence applications on-premises, in the cloud or at the edge. It provides businesses with complete situational awareness and operational decision support at every moment. Built upon the open source SwimOS core, Swim Continuum provides unprecedented performance and efficiency for operationalizing high-frequency, contextual data analytics and real-time visualizations of massive amounts of streaming and batch data. Its single, production-ready platform monitors and manages all Swim operations, creates engaging, connected user experiences and seamlessly interoperates with existing enterprise systems. For more information, visit www.swim.ai and follow us on Twitter @swim.



Copyright © 2020 by Swim.ai, Inc. All Rights Reserved. Swim.ai, SwimOS, Swim Continuum and the Swim logo are trademarks or registered trademarks of Swim.ai, Inc. and/or its affiliates in the U.S. and worldwide. All other product names, logos, and brands are the property of their respective owners.