

Smartindustry

ACCELERATING DIGITAL TRANSFORMATION

MENU

Smart Industry Forum

Share your perspective. Click here to contact the Smart Industry Forum editor.

The new rise of time-series databases

Submitted by [Michael Risse](#) on Mon, 02/26/2018 - 15:09

Time-series data storage and management has long been an interesting—if quiet—market category. It's been a multibillion-dollar business for years and a mainstay in process-manufacturing plants since the 1980s. But recently, the category has been getting another look from investors and companies large and small.

Why?

For starters, time-series data volumes are huge: way back in 2010 manufacturing companies were generating 1,800 petabytes of data per year (twice as many as the next closest vertical). Much of that was time-series data [Figure 1]. And manufacturing data volumes have only continued to grow in recent years thanks to new Internet of Things (IoT) and Industrial Internet of Things (IIoT) deployments.

These vast data volumes attract attention because “data has gravity”—meaning that whoever stores the data will attract high-value add-ons such as management, security, analytics and consulting services. The result? To attract these other business opportunities, time-series data-storage platforms can be licensed for less than they cost.



Seeq's Michael Risse



Figure 1: Manufacturing data, most of it stored in time series databases, dwarfs all other segments, including government. The source data for this slide is from McKinsey Global Institute's seminal June 2011 report: "Big data: the next frontier for innovation, competition and productivity."

Venture interest and open source options

To see what's happening, just follow the dollars. On January 24 of this year, [Timescale](#), an open-source time-series database (OSTSDB) company, secured \$12.4 M Series A funding led by Benchmark Capital. This was soon followed by InfluxData, which [scored \\$35M](#) in a Series C funding on February 12, led by Sapphire Ventures, bringing their total funding to \$60M.

If the names don't ring a bell, Sapphire Ventures is the venture arm of SAP, and Benchmark Capital and Battery Ventures are both very successful venture funds. (Benchmark has nearly \$3B under management and was an early-stage investor in companies ranging from Twitter to Dropbox to Instagram. Battery Ventures has nearly \$7B in assets.) The investors are likely looking at graphs [Figure 2] showing that time-series databases have recently been the fastest growing segment in the database market. InfluxData, for example, claims 115,000 active sites using their product.

Time Series – the Fastest Growing Database

DB-Engines also ranks time series database management systems (Time Series DBMS) according to their popularity. Time Series Databases are the [fastest growing segment](#) of the database industry over the past year.

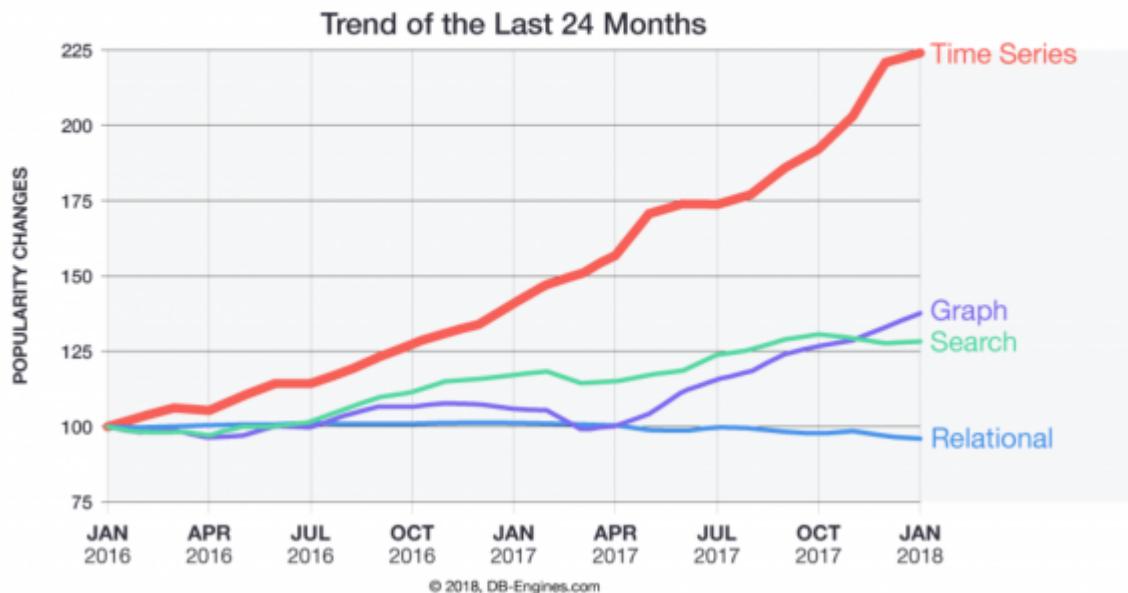


Figure 2: Time series databases are experiencing explosive growth due to their ability to efficiently store and provide access to large volumes of data. Source: see website: https://db-engines.com/en/ranking_categories

That said, Hortonworks (NASDAQ: HDP), a leader in Hadoop and big-data implementations with process-manufacturing companies, has itself been adding features and patterns to address time-series database opportunities. Their added value is enabling manufacturers to analyze any type of data for batch, interactive, or real-time applications by unlocking siloed data sets from both operational technology and information technology systems. By centralizing customer-process data into a single open-source platform, Hortonworks is able to democratize industrial data analysis by providing a single view of operations for their customers. By virtue of their funding, Timescale and InfluxData are now separated from a pack of OSTSDB companies or open-source efforts, including OpenTSDB, Prometheus, Druid, KairosDB and others. Net of another funding event, it seems Timescale and InfluxData may be staging a repeat of the recent CloudEra/Hortonworks battle among big-data startups.

So, whether public companies or startup venture, OSTSDB and big-data vendors are now significant players in the time-series storage market.

The public cloud arrives

Storing large volumes of data in the cloud is increasingly, if not already, a “when” not an “if” question for many companies. Consequently, the big public-cloud platforms are paying more attention to the largest sources of data.

For example, Microsoft [recently introduced](#) a Cassandra interface to Azure CosmosDB, their NoSQL cloud-data service, which brings them into the market for time-series storage. (For context, Cassandra is an open-source database and a popular choice for storing time-series data, so a Cassandra interface to CosmosDB is an obvious fit for time-series data storage.) What’s more, CosmosDB has a graph-database interface, which means it has both of the services required for modern historian functionality: a Cassandra interface for time-series storage, and a graph-database interface for defining and accessing asset models and hierarchies.

Of course, interfaces by themselves don’t make a historian or a time-series database product successful. There are many other factors involved, and it remains to be seen how Microsoft prices their service and differentiates it from open-source offerings, and how they work with partners offering historians on top of Azure. These and other decisions will go a long way to determining the success or failure of Microsoft’s foray into the market.

Honeywell's recently announced Uniformance Cloud Historian, for example, runs on Azure and leverages Microsoft data services as its platform for distributed storage and management. These types of industry partnerships will be crucial for success within process-manufacturing verticals.

Finally, Microsoft won't be making their decisions on CosmosDB and time-series data in a vacuum. Amazon with DynamoDB and Google with BigTable are both making their own arguments for using their NoSQL offerings for time-series data storage. This list could go on and on: PTC/Thingworks has established partnerships to support their IIoT platform with time-series storage options, plus there are time-series storage services in GE Predix and Siemens Mindsphere. Beyond these offerings, 2nd tier IIoT-platform offerings supporting time-series data could fill a dictionary.

The incumbents

As mentioned earlier, data historians (also called process historians) have been used by process-manufacturing companies for decades. Every process-automation vendor offers at least one historian, like DeltaV Continuous Historian from Emerson Process Automation. And others have multiple historians due to a history of acquisitions, like Schneider Electric. Some historians are sold separately by dedicated historian firms like Canary Labs, and others are offered in the context of the software company's principal offering, as with Inductive Automation's Ignition SCADA system.

But for all the historians available for sale, one vendor stands apart in market share among high-value oil & gas, chemical, power generation and other process-industry customers: OSIsoft and its PI infrastructure platform*.

If the new entrants—startups and clouds—are affecting OSIsoft's business, it's hard to see from the outside. As a private company, they don't release earnings, but an investment in OSIsoft last year by Softbank suggests expectations of further growth. There are also public examples of OSIsoft's momentum. Their upcoming user conference, rebranded PI World, is expected to be their largest ever, with a doubling of space for partners and sponsors. With new investors, a growing partner ecosystem, and new efforts in edge and IIoT deployments, it would seem OSIsoft sees opportunity for growth, despite challenges from new participants.

Certainly, OSIsoft's established position is a point of confidence for customers, as is its support for existing investments and IT requirements—and that position is validated by industry observers, including ARC: "ARC research indicates OSIsoft has been the market leader in process historians for many years," commented Janice Abel, principal analyst at ARC Advisory. "The company has a well-established and loyal customer base, a large partner ecosystem, and the OSIsoft PI historian connects to data from more than 450 different sources, which to the best of our knowledge far exceeds any competitors' products."

Perhaps the actual impact on OSIsoft of the open source and cloud entrants to the time-series database market is an increase in the awareness of and need for a proven, enterprise-ready solution delivered out of the box.

Conclusion

With incumbents and challengers using both open-source and cloud services, the market for time-series storage in recent months has taken a strong turn to the interesting. This market, even with its strong incumbents, is attracting both top-tier venture-capital firms and the largest public cloud platforms. For now, it would seem all boats are rising on a tide of interest in the market segment, as IoT and IIoT interest and deployments continue to grow.

While it's been interesting, it's likely only the beginning of what promises to be a wild race, with multi-billion-dollar prizes at stake.

*As a point of disclosure, Seeq is an OSIsoft ISV partner and a Gold Sponsor of their upcoming user conference, and many Seeq customers are OSIsoft's best customers using their "all you can eat" Enterprise Agreements.

Michael Risse is vice president at Seeq Corporation.