



Financial Services Data and Analytics Newsletter

July 2022



[Click to launch](#) 

Introduction

Introduction

Topic of the month

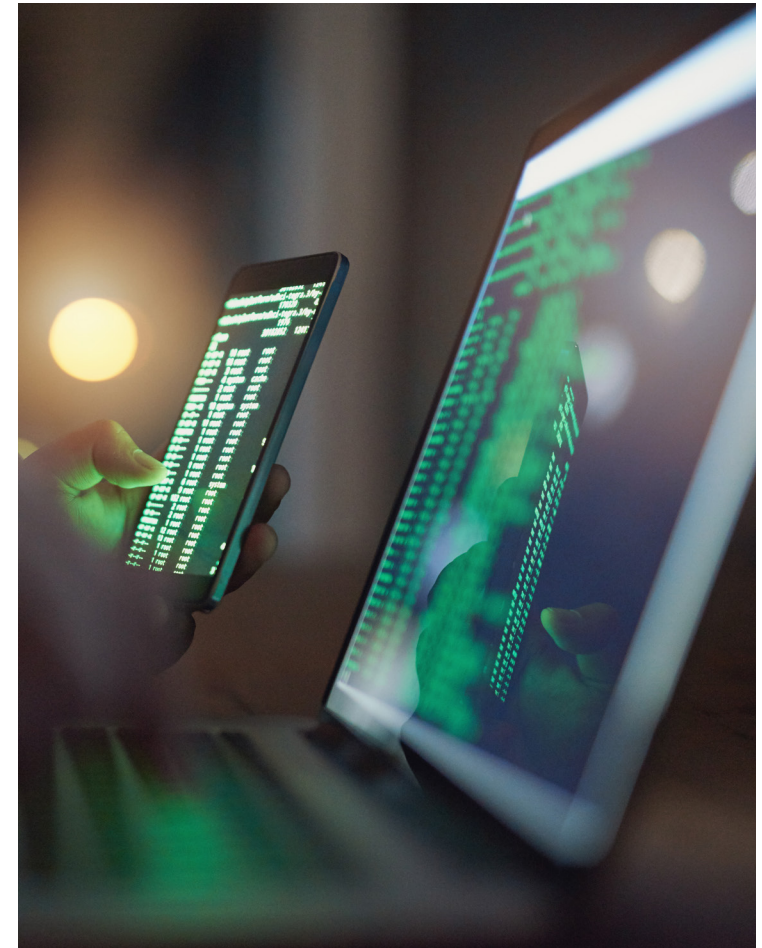
Industry news

Knowledge bytes

Contact us

It is a fairly common need across multiple enterprise functions to display current trends versus historical trends data in reports/dashboards which necessitates the data architects to muse over multiple design patterns. Key questions include computing the historical trends in the database itself vs using an in-memory reporting component (typically cubes) by storing the base level data in cubes and computing trends real time. Both designs have their own advantages and disadvantages with respect to flexibility and cost. What if there is a special engineered database to suit your needs and is optimised to process and retrieve 'time-series' data, measure change over time or aggregate over large periods of data? In this newsletter, we are providing a concept note on such a different type of database called time series databases (TSDBs).

This edition of the newsletter also includes news snippets around new initiatives taken across sectors, with the most interesting one on the different announcements made by the Insurance Regulatory Development Authority of India (IRDAI) in the past couple of months. Execution is going to be key to drive these initiatives to their full potential, but the zeal from the regulator and the industry is definitely commendable.



Topic of the month – Time series databases

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

What are TSDBs?

A TSDB is a database optimised for storing and retrieving time-stamped or time-series data. Temporal ordering, a key characteristic of time-series data, organises events in the order in which they occur and arrive for processing. One can use time-series data to look backward and measure change, or to look forward and predict future change. The TSDB architecture is fully capable of supporting cloud/on-premise/hybrid solutions. The different architecture patterns typically used are – a standalone specialised time series database that is receiving data from one or more data collector endpoints or an extension to the current relational database management system (RDBMS) / NoSQL database, specially designed for time-series data analysis and computation. There are currently many options available for TSDBs – RDBMS options like TimescaleDB and questDB, NoSQL options like Amazon TimestreamDB and Prometheus, and standalone TSDB like InfluxDB. An abstract model of time series data attempts to answer the following question:

Who did what at what time and where?

Name (who): Describes the subject that produces the data, which can be a person, a monitoring metric, or an object.

Tags (who): Additional information to describe/classify the 'name'.

Timestamp (when): Time is the most important and fixed axis feature of time-series data and is a key attribute to what distinguishes it from other data.

Location (where): Used for locating the monitored object which is described by one or more tags.

Values (what): The value or status corresponding to the data. Multiple values or statuses can be provided, which are not necessarily numeric.

The TSDB architecture treats the above five tags as 'indexes' on which data gets loaded, organised, compressed, and retrieved from database for performing different kinds of operations like aggregation or down sampling. Back in time, TSDBs supported only four indexes, namely – metric, tags, timestamp and values. But with advent in the field of geospatial temporal analysis, another parameter of 'location' was added to further boost the utilities of TSDB in various fields such as analysis of customer growth in certain geographical areas over a period.

Below is a snapshot of data getting stored in a TSDB.

```
{
  "name": "device.status",
  "tags": {
    "deviceId": "aliyun-tablestore",
    "version": "2.0",
    "model": "unify"
  },
  "location": {
    "lat": 30.1276799186,
    "lon": 120.0848161441
  },
  "timestamp": 1531918377012,
  "values": {
    "status": "active",
    "temperature": 30,
    "cpu": 50
  }
}
```

Who

Where

When

What

Topic of the month – Time series databases

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

Features of TSDB

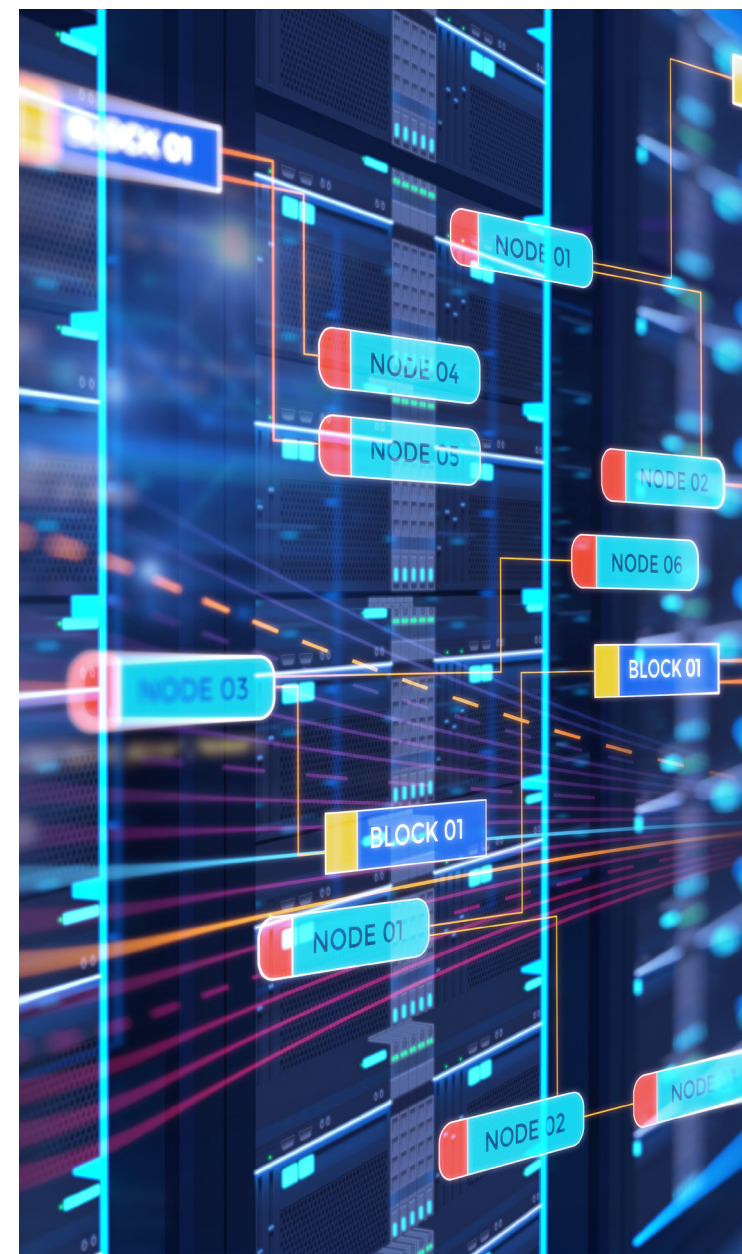
The most significant characteristic of TSDBs (along with supported SQL queries) is to provide time-series data specific functions and features to further exploit the benefits of temporal indexing. These features can majorly be divided into four categories:

Aggregates: A query can be for a single time series or multiple time series. For range queries or multiple time series, the results are down sampled, grouped and aggregated to give the user a holistic view of data to gain insights, e.g. **aggregate window, aggregate rate.**

Selectors: These functions are used for selection of specific datasets and records from a single or multiple time series, according to desired requirements. Results can be returned based on highest/lowest average, count, rate, etc., for a selected window, e.g. **quantile, highest average/lowest average.**

Transformations: These set of queries support varies time-dependent values' extraction and transformations that are not possible in any other type of DB (without significant coding efforts). For example, calculating running derivatives and exponential moving time weighted averages. Functions involve various kinds of calculations, moving averages, timeframe moving average, etc., e.g. **double EMA, timed moving average.**

Geotemporal: This is one of the most unique features of TSDBs. This is achieved by categorising the received location input in S2 cells (mathematical mechanism that helps computers translate earth's spherical 3D shape into 2D geometry), e.g. **geo grid filter, geo shape data.**



Topic of the month – Time series databases

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

Query example 1:

We shall now explore how TSDB's inbuilt function helps in solving the following problem statement.

For instance, consider the following problem statement and simplicity of generating output from TSDB vs RDBMS:

Which dynamic six-hour periods saw the most log ins from users on tablet devices, for last one week starting 1 January 2020?

TSDB (TimeScale database)	RDBMS (MS SQL database)
Select time bucket ('6 hours', login timestamp, timestamp '2020-01-01 08:00:00') as device bucket, device type, count(*) as logins by device, From user logins Where login timestamp > now() - interval '1 week And device type = 'tablet' group by device bucket, Device type order by logins by device desc;	Pseudo query/logic: <ol style="list-style-type: none">1. Filtering the login timestamp only for device type = 'tablet' for last one week starting 1st Jan 2020 and ordering the data by login timestamp2. Aggregating the count of logins for every hour3. Running a loop against dataset generated from step 2 against each hour reference and sum up counts for 6-hour window4. Picking up the six-hour window that has maximum logins from data set generated in step no 3
Time bucket () and interval are inbuilt functions in TSDB, to help in solving problem statement	No special function available and hence the code is lengthy, complex and needs developer support
Power of solving simple to complex problem statements at hands of end user without any special technical support	End user will have to take help from special technical resource for designing query to solve problem statements

Topic of the month – Time series databases

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

Query example 2:

We shall explore another such problem statement solved by TSDB.

For instance, consider the following problem statement and simplicity of generating output from TSDB vs RDBMS:

How to trace vehicle movement as to how many trucks exit from Los Angeles on a daily basis and what is the weight change during the same?

TSDB (TimeScale database)	RDBMS (MS SQL database)
<pre>Select time bucket('1 day', time) as day, count(*) as trucks exiting, sum(weight) as tonnage FROM vehicle movements Where ST within(last location, ST Polygon ((select geom from cities where name='los angeles'),4326)) and not ST within(current location, ST Polygon((select geom from cities where name='los angeles'),4326)) group by day Order by day DESC LIMIT 30;</pre>	No such feature available



Topic of the month – Time series databases

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

Use cases for TSDBs

In the last three years, the trend in adoption of TSDB for business use cases has grown exponentially to harness the evolving nature of data and explore the potential of driving more real-time use cases to solve complex problems. Following are some of the major use cases where TSDB has been very promising:

IoT applications and use cases

Internet of things (IoT) devices, wearables generate enormous time-series data which can be increasingly used for data analysis and uncovering the hidden patterns related to health, safety, financial activity, etc., which can generate revenue for client offering subscriptions to real-time alerts.

Analysing and predicting customer characteristics

Data generated from customers' online activities can be subjected to time-series database analytics to understand the spend pattern, predict the customer characteristics like shopping behaviour and accordingly set up campaigns and alerts to attract the customer to a client's ecosystem network.



Anomaly detection

Temporal ordering allows the user to analyse time-series information to compare current data to historical data, detect anomalies and generate real-time alerts, or visualise historical trends. These anomalies and historical trends can further be used to reinforce current mechanism and train artificial intelligence (AI) models for more resilient and seamless automation. For example, in ATMs and banking apps, temporal analysis can help learn user activity and detect any malicious anomaly, allowing the bank to better protect the consumers from frauds and cyberattacks.

Financial trend analysis

Storing and sorting data in a time-series format allows stockbrokers and traders to analyse the previous trends of the stock market and use the same for predictive modelling and result forecasting. These autonomous trading algorithms continuously collect data on how the markets are changing to optimise returns, both in the short and long term.

Medical insurance tracking

Tracking each aspect of patient data (age, admitted or discharged, days to recovery, etc.) during the pandemic helped us understand how we arrive at the daily counts, allowing us to better analyse trends, accurately report totals, and act. Such details and analysis during the pandemic impacts public policy in cities and towns, and insurance premiums, allowing for the Government and organisations to adapt and make informed decisions.

Topic of the month – Time series databases

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

TSDB vs RDBMS vs NoSQL databases

TSDBs are comparatively new as compared to NoSQL or more conventional RDBMS. Thus, it becomes imperative to analyse the differences and offerings of each database engine to further our understanding.

Factors	TSDB	RDBMS	NOSQL
Key features	Columnar data storage, time variable indexed and specialised design for time-series data management	Known for highly robust, structured data storage framework	Highly adaptive to dynamic data generation and highly scalable
Data storage method	Append only	Insert/update	Insert/update
Data purging	Auto deletion of records post aggregation, available as out of the box	Custom routines to be built for data purging	Custom routines to be built for data purging
Scalability and durability	Designed to achieve highest scalability and durability	Limited scalability	Highly scalable
Compression and storage	Highly optimised for storage due to efficient compression ratios	Limited compression support	Better compression support than RDBMS
Analytical function support	Comes with inbuilt analytical functions/aggregates for unleashing analytical insights	Use of stored procedures/functions/extensions can be used for analytical use cases	No inbuilt support for analytics; extensions like Python/R/Scala can be used for generating analytical outcomes
Nature of data	Data stored in structured/unstructured format	Data stored in structured format	Data stored in unstructured format
Querying support	Basic/advanced querying levels depend on the product and underlying architecture	Highly optimised for simple to complex querying	No inbuilt support for data querying
Use cases	High time-precision applications, e.g. IoT, time-based application monitoring and optimisation, real-time data streaming, etc.	Design of robust enterprise-level solutions and applications	Analytics and reporting for unstructured data generated from sources like social media

Industry news

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

1. NBFCs and associated FinTechs are leveraging big data to fulfil credit requirements of smaller businesses

Transerve, a new-age FinTech, is helping NBFCs to assess the credit risk of customers based on their addresses/home location. By classifying various areas/regions, borrowers belonging to certain regions will be classified as low-/high-risk debtors. Additional checks would be implemented for high-risk debtors.

2. Union Bank of India debuts on metaverse

Union Bank of India has debuted on metaverse with its Uni-verse virtual lounge which will showcase the bank's products and services revolving around current account savings account (CASA), loans, Government welfare schemes and digital initiatives. It has also created an open banking sandbox to collaborate with FinTechs and innovate banking products.

3. SCB Asset Management outperforms market with big data, AI and FinTech

Thailand's top asset managers are being enabled by technology and big data to study the non-linear and dynamic aspects that provide new opportunities. They are now capable of reclassifying funds in ways that are less susceptible to market capitalisation and provide do-it-yourself (DIY) products to the market.

4. Kotak General Insurance to digitalise vehicle inspection to avoid fraud

With the help of AI, Kotak General Insurance has partnered with Inspektlabs to detect damages in vehicles. Now, customers will upload a video of their vehicle for policy renewal, the complete process will be automated and inspection reports of the damage will be generated post uploading the video. According to the company, this process will help in the underwriting process, prevent fraud, save time and cost, and lead to customer satisfaction.

Industry news

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

5. Max Life Insurance saves INR 800 crore in claims payout using their in-house predictive underwriting decision system – Shield

Max Life insurance is leveraging technology and data to improve its operational efficiency and reduce human intervention. Its in-house product, Shield, has helped in automating and digitalising 75% of non-term business and saving almost INR 800 crore by minimising fraud. Shield helps in catching fraudulent policies at an issuance stage using an AI bases as a model instead of rejecting them at the claims stage.

6. Convin's AI-based agent assist platform increases BFSI's customer satisfaction (CSAT) score by 30%

The banking, financial services and insurance (BFSI) industry engages in almost 70% of all customer interactions via calls. Convin has launched a new platform which supports multichannel conversation for customers and automates call reports to give an analysis on call performance, outcome from the calls, sentiment analysis and an alert mechanism to avoid inappropriate calls for its BFSI clients. It is expected to improve customer experience, increase productivity of the agents and develop better sales strategies.



Knowledge bytes

Introduction

Topic of the month

Industry news

Knowledge bytes

Contact us

1. Assets Under Management (AUM) of NBFCs anticipated to grow at a higher rate in FY 23.

As per a note published by one of the leading rating agencies, AUM of NBFCs is expected to grow in FY23.¹ AUM growth rate of NBFCs in FY22 was 9.5%, for FY23 and this growth rate will fall between 9–11 % as per this report.

2. To ease technical glitches within banks and NBFCs, RBI issues guidelines for outsourcing IT services

As per the draft, regulated entities would not be required to take prior approval from the RBI before availing IT/ITeS services. However, banks and non-banking financial companies (NBFCs) would have to ensure that usage of these services does not diminish its ability to fulfil its obligations to customers or impede supervision by the supervisory authorities.

3. Government keen on public sector banks exploring FinTech partnership

For the revival of the economy, the finance ministry has asked public sector banks to explore co-lending opportunities with FinTechs. It has also asked to focus on technology and data analytics to accelerate their lending while keeping a check on frauds through IT security and cyber security systems.

4. IRDAI taking steps to increase insurance penetration and improve customer experience

IRDAI intends to increase insurance penetration in India by improving customer experience and enlarging the targets of insurers. Towards this, IRDAI has also allowed both life and non-life insurers to launch products without any prior approvals and given them state-wise targets, thus expecting innovation and customisation in the industry. Some of these guidelines include creating a national health exchange to settle claims, standardisation and benchmarking for hospitals, and regulation on the high vehicle insurance rates and formulation of guidelines towards single coverage for multiple vehicles of a single customer.

1 | <https://www.icra.in/Media/OpenMedia?Key=10695e5a-5faa-428c-9cad-6a38a732f19e>

Introduction

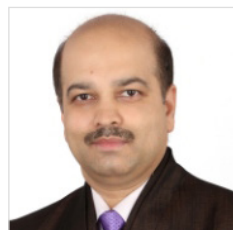
Topic of the month

Industry news

Knowledge bytes

Contact us

Contact us



Mukesh Deshpande

Partner, Technology Consulting
mukesh.deshpande@pwc.com
+91 98450 95391



Hetal Shah

Partner, Technology Consulting
hetal.d.shah@pwc.com
+91 9820025902

Acknowledgements

This newsletter has been researched and authored by Aman Mann, Aniket Borse, Anuj Jain, Arpita Shrivastava, Dhananjay Goel, Fenil Thakkar, Harshit Singh, Krunal Sampat, Mamta Kumawat, Shyam Mishra.

About PwC

At PwC, our purpose is to build trust in society and solve important problems. We're a network of firms in 156 countries with over 295,000 people who are committed to delivering quality in assurance, advisory and tax services. Find out more and tell us what matters to you by visiting us at www.pwc.com.

PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

© 2022 PwC. All rights reserved.

pwc.in

Data Classification: DC0 (Public)

In this document, PwC refers to PricewaterhouseCoopers Private Limited (a limited liability company in India having Corporate Identity Number or CIN : U74140WB1983PTC036093), which is a member firm of PricewaterhouseCoopers International Limited (PwCIL), each member firm of which is a separate legal entity.

This document does not constitute professional advice. The information in this document has been obtained or derived from sources believed by PricewaterhouseCoopers Private Limited (PwCPL) to be reliable but PwCPL does not represent that this information is accurate or complete. Any opinions or estimates contained in this document represent the judgment of PwCPL at this time and are subject to change without notice. Readers of this publication are advised to seek their own professional advice before taking any course of action or decision, for which they are entirely responsible, based on the contents of this publication. PwCPL neither accepts or assumes any responsibility or liability to any reader of this publication in respect of the information contained within it or for any decisions readers may take or decide not to or fail to take.

© 2022 PricewaterhouseCoopers Private Limited. All rights reserved.

AW/July 2022 - M&C 21205