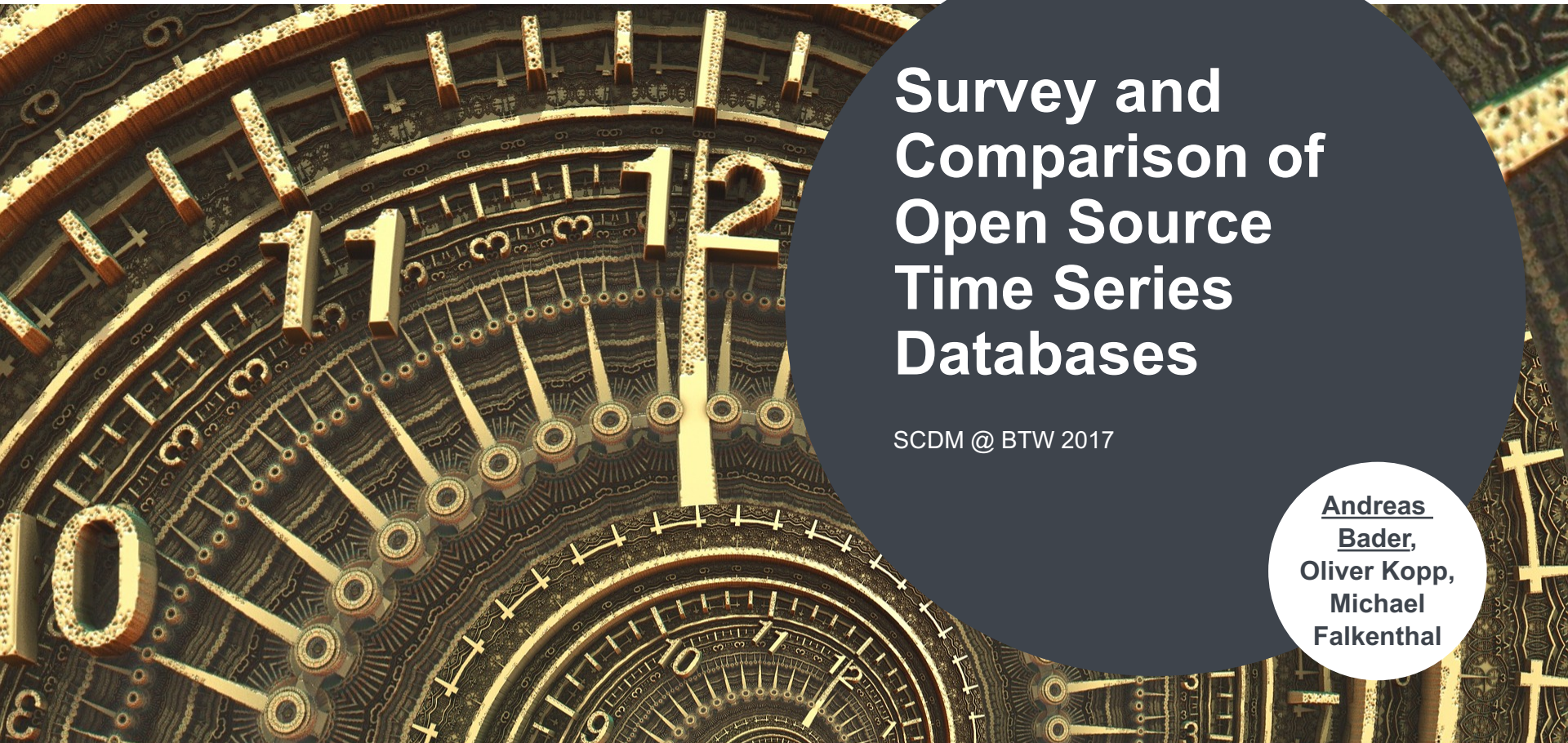
University of Stuttgart
Germany

IPVS



Survey and Comparison of Open Source Time Series Databases

SCDM @ BTW 2017

Andreas Bader,
Oliver Kopp,
Michael Falkenthal

Comparison of Open Source TSDBs

What is a time series data?

- A row of data that consists of a timestamp, a value, optional tags

ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Comparison of Open Source TSDBs

What is a time series data?

- A row of data that consists of a timestamp, a value, optional tags

timestamp

ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Comparison of Open Source TSDBs

What is a time series data?

- A row of data that consists of a timestamp, a value, optional tags

timestamp

value

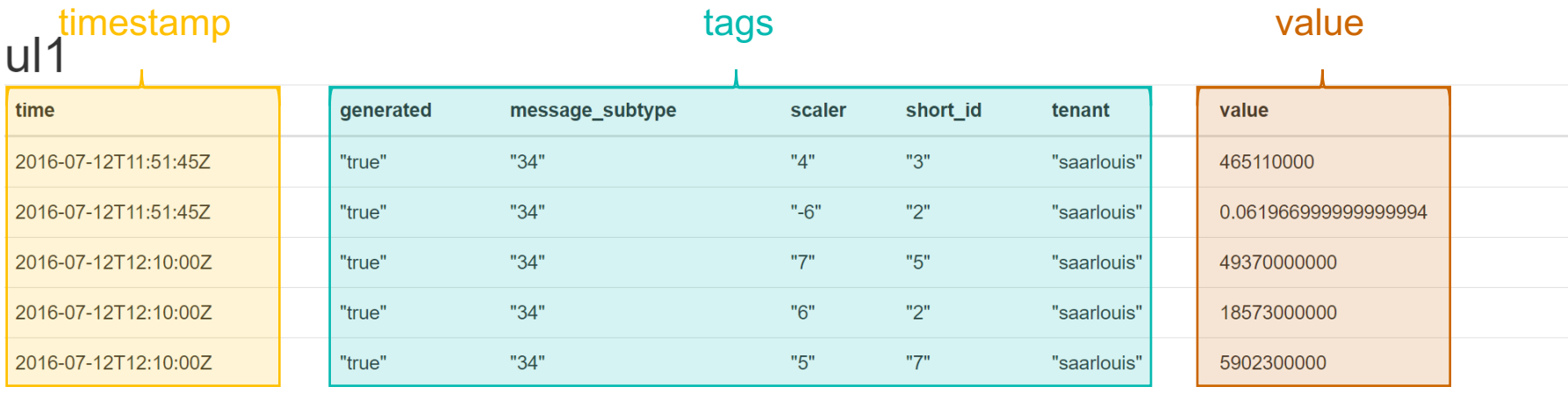
ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Comparison of Open Source TSDBs

What is a time series data?

- A row of data that consists of a timestamp, a value, optional tags



Comparison of Open Source TSDBs

What is a Time Series Database (TSDB)?

- A DBMS is called TSDB if it can
  - store a row of data that consists of timestamp, value, and optional tags
  - store multiple rows of time series data grouped together (e. g., in a time series)
  - can query for rows of data
  - can contain a timestamp or a time range in a query

ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Comparison of Open Source TSDBs

What is a Time Series Database (TSDB)?

- A DBMS is called TSDB if it can
  - **store a row of data that consists of timestamp, value, and optional tags**
  - store multiple rows of time series data grouped together (e. g., in a time series)
  - can query for rows of data
  - can contain a timestamp or a time range in a query

ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Comparison of Open Source TSDBs

What is a Time Series Database (TSDB)?

- A DBMS is called TSDB if it can
  - **store a row of data that consists of timestamp, value, and optional tags**
  - **store multiple rows of time series data grouped together (e. g., in a time series)**
  - can query for rows of data
  - can contain a timestamp or a time range in a query

ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Comparison of Open Source TSDBs

What is a Time Series Database (TSDB)?

- A DBMS is called TSDB if it can
  - **store a row of data that consists of timestamp, value, and optional tags**
  - **store multiple rows of time series data grouped together (e. g., in a time series)**
  - **can query for rows of data „SELECT * FROM ul1“**
  - can contain a timestamp or a time range in a query

ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Comparison of Open Source TSDBs

What is a Time Series Database (TSDB)?

- A DBMS is called TSDB if it can
  - **store a row of data that consists of timestamp, value, and optional tags**
  - **store multiple rows of time series data grouped together (e. g., in a time series)**
  - **can query for rows of data „SELECT * FROM ul1“**
  - **can contain a timestamp or a time range in a query**

“SELECT * FROM ul1 WHERE time >= '2016-07-12T12:10:00Z‘”

ul1

time	generated	message_subtype	scaler	short_id	tenant	value
2016-07-12T11:51:45Z	"true"	"34"	"4"	"3"	"saarlouis"	465110000
2016-07-12T11:51:45Z	"true"	"34"	"-6"	"2"	"saarlouis"	0.061966999999999994
2016-07-12T12:10:00Z	"true"	"34"	"7"	"5"	"saarlouis"	49370000000
2016-07-12T12:10:00Z	"true"	"34"	"6"	"2"	"saarlouis"	18573000000
2016-07-12T12:10:00Z	"true"	"34"	"5"	"7"	"saarlouis"	5902300000

Outline

- Motivation
- Comparison of open source TSDBs
- Live Demo
- Conclusion and Outlook

Motivation

Why comparing Open Source TSDBs?

Motivation

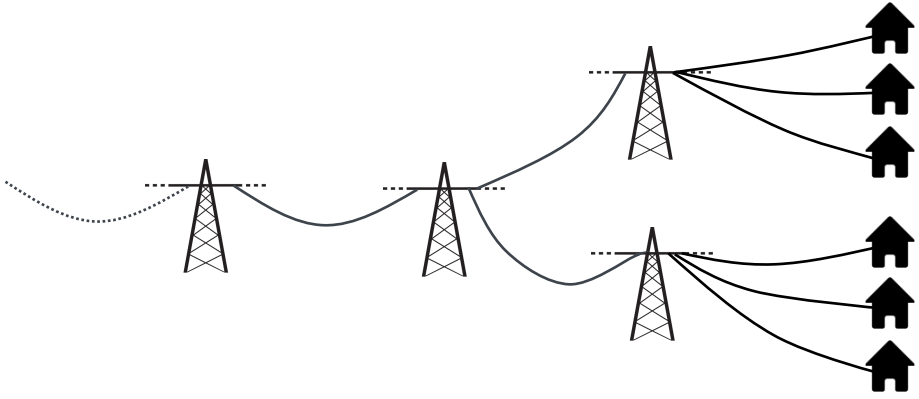
NEMAR Project

- New market role
- Sensor data from smart grids
- Smartly acting on energy markets
- Smart help for operational management & decision support

Motivation

NEMAR Project

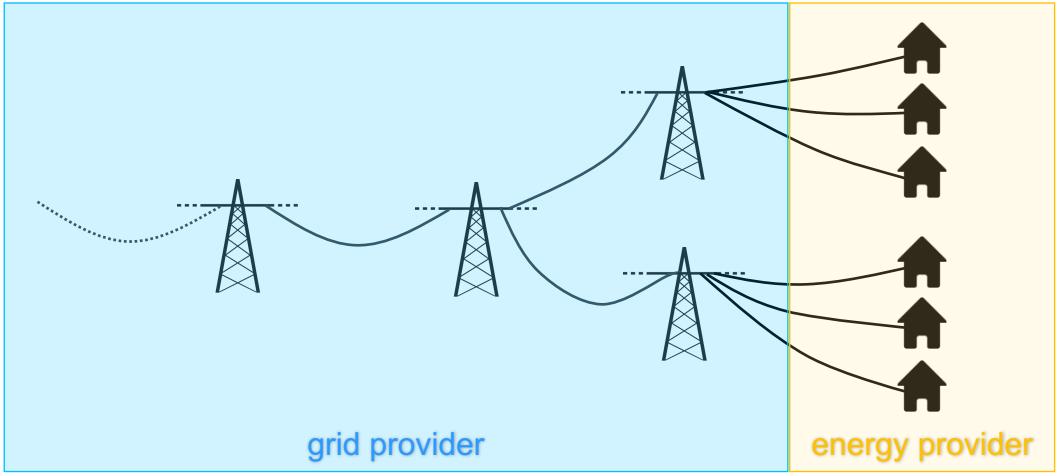
- New market role
- Sensor data from smart grids
- Smartly acting on energy markets
- Smart help for operational management & decision support



Motivation

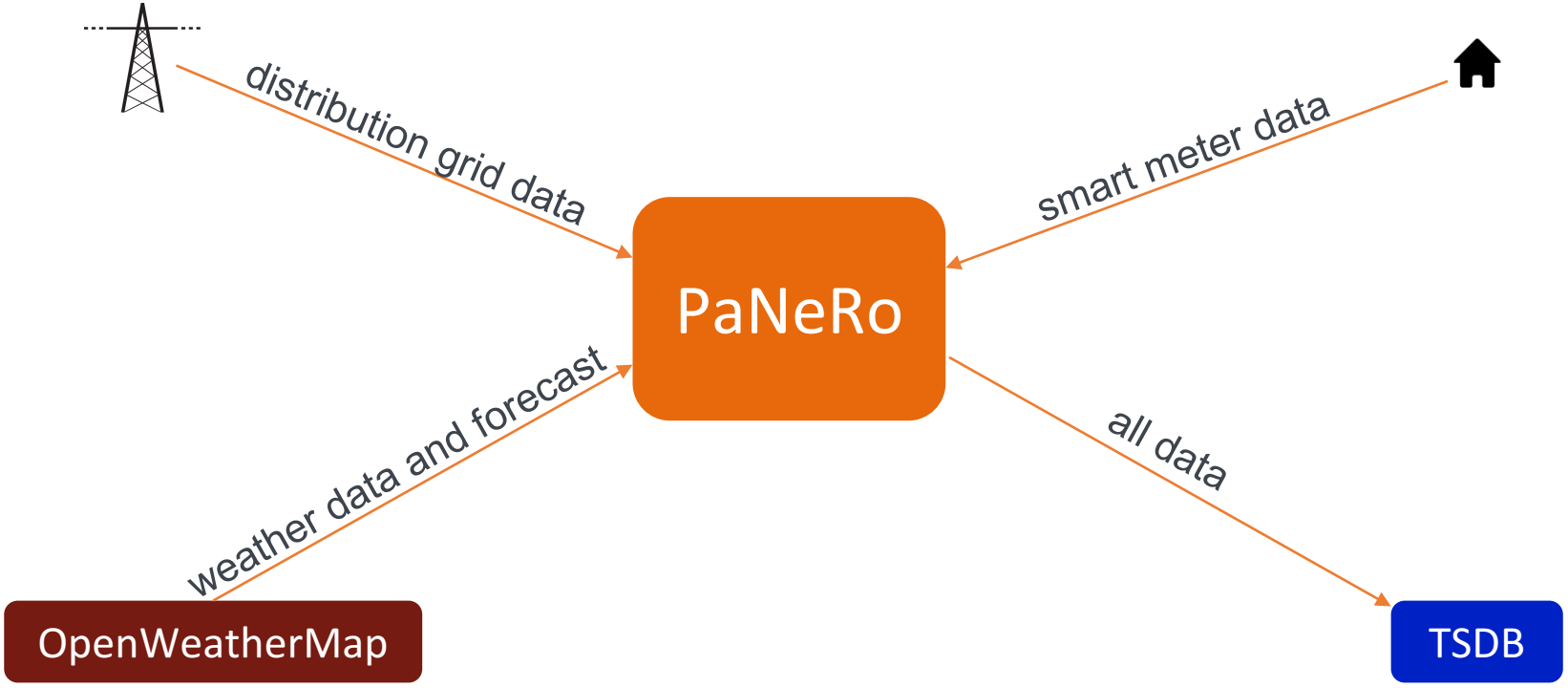
NEMAR Project

- New market role
- Sensor data from smart grids
- Smartly acting on energy markets
- Smart help for operational management & decision support



Motivation

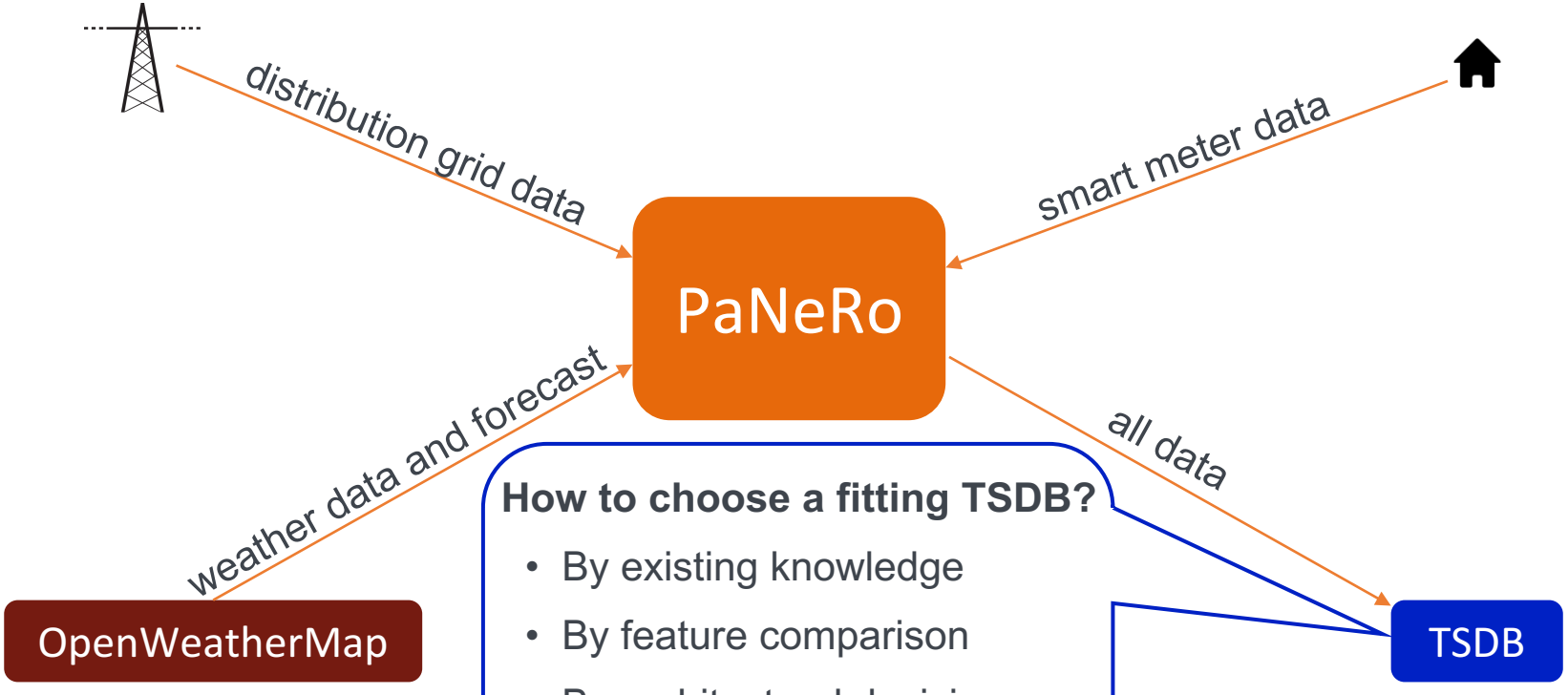
PaNeRo Platform for NEMAR



From: Oliver Kopp, Michael Falkenthal, Niklas Hartmann, Frank Leymann, Holger Schwarz, Jessica Thomsen: **Towards a Cloud-based Platform Architecture for a Decentralized Market Agent**. In: INFORMATIK 2015, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2015

Motivation

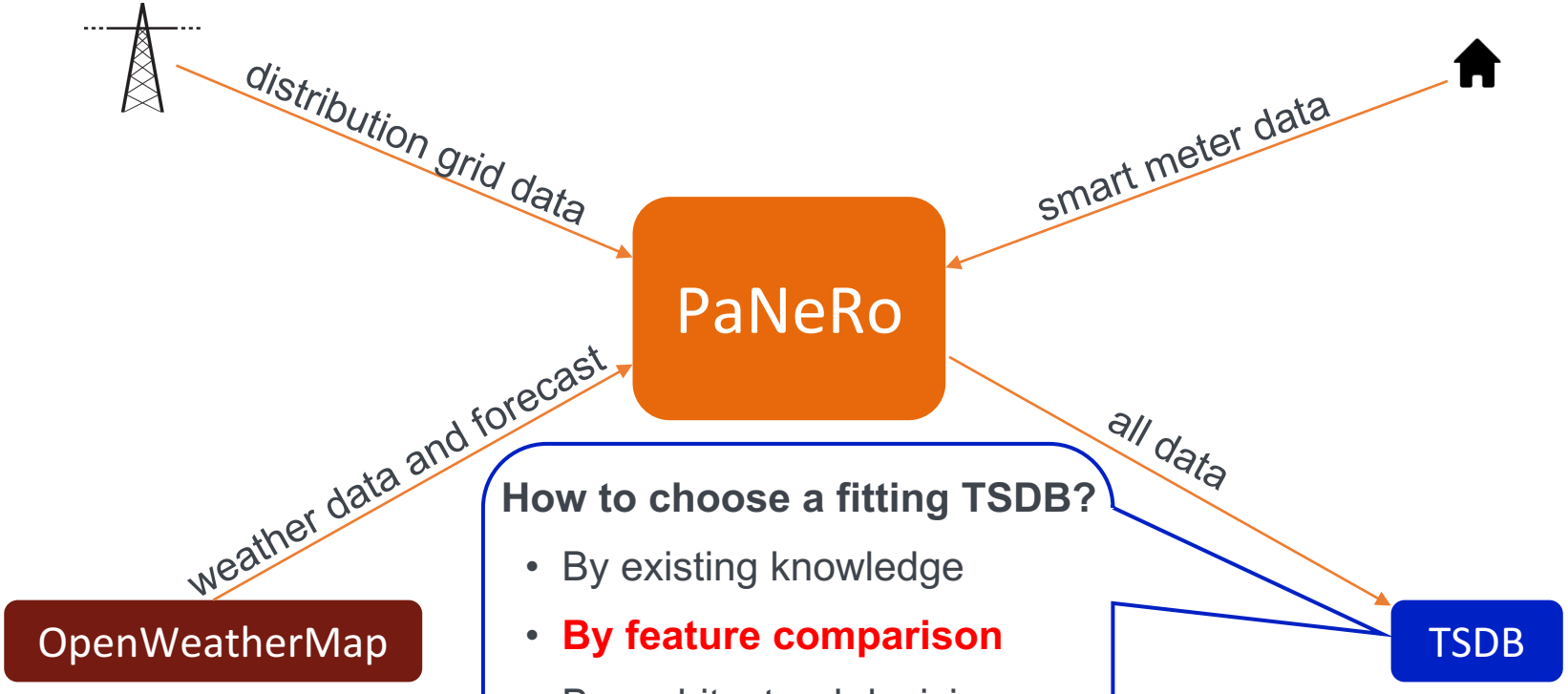
PaNeRo Platform for NEMAR



From: Oliver Kopp, Michael Falkenthal, N msen: **Towards a Cloud-based Platform Architecture for a Decentralized Market** (LNI), Gesellschaft für Informatik, Bonn 2015

Motivation

PaNeRo Platform for NEMAR



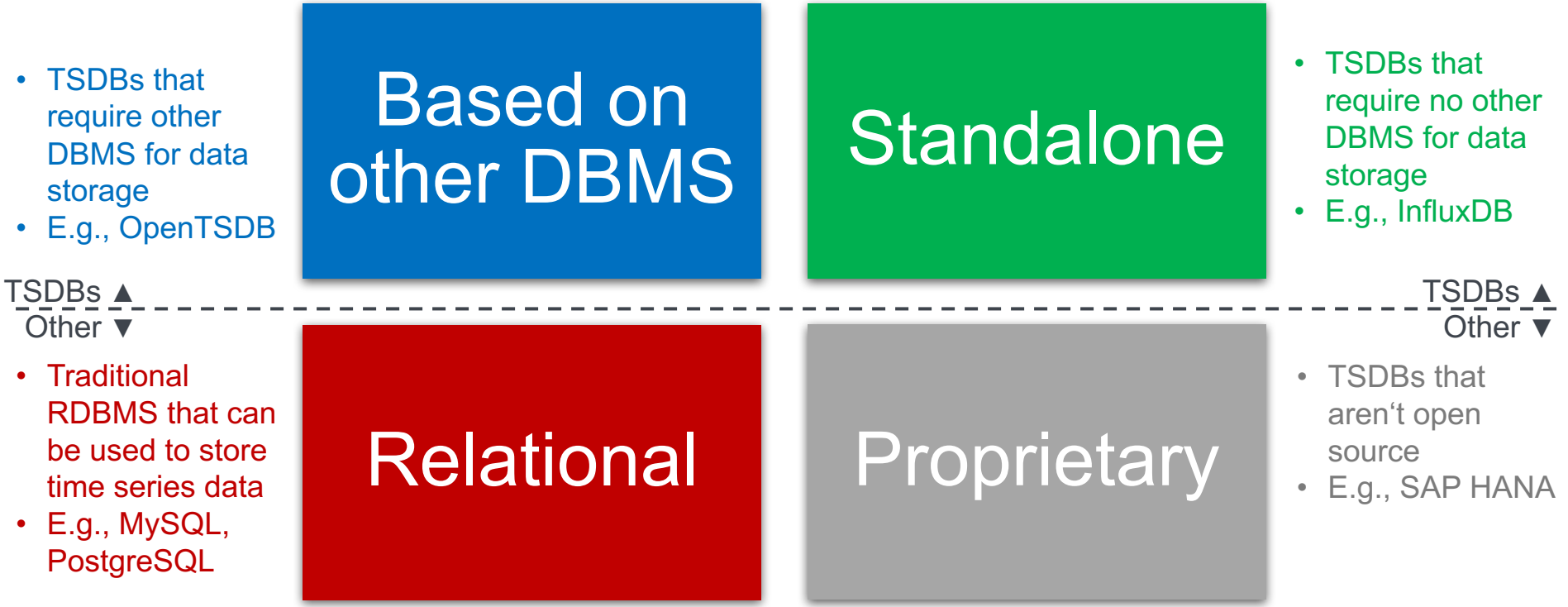
From: Oliver Kopp, Michael Falkenthal, N msen: **Towards a Cloud-based Platform Architecture for a Decentralized Market** (LNI), Gesellschaft für Informatik, Bonn 2015

Comparison of Open Source TSDBs

How to compare Open Source TSDBs?

Comparison of Open Source TSDBs

Categories



- TSDBs that require other DBMS for data storage
- E.g., OpenTSDB

- TSDBs that require no other DBMS for data storage
- E.g., InfluxDB

- Traditional RDBMS that can be used to store time series data
- E.g., MySQL, PostgreSQL

- TSDBs that aren‘t open source
- E.g., SAP HANA

Search for terms like „TSDB“, „Time series“, … on Google, ACM, IEEE

➢ 83 found TSDBs, 50 of them open source

Comparison of Open Source TSDBs

How to store time series data in RDBMS?

- First approach: Timestamp as primary key
  - One value per timestamp per table

Timestamp	Value	Host
2016-07-12	1.22	example.org
2016-07-12	5.33	

- Second approach: Tags and date as combined primary key
  - Tags are optional → same issue as above

Timestamp	Value	Host
2016-07-12	1.22	example.org
2016-07-12	5.33	

- Third approach: Use an auto-incrementing primary key

ID	Timestamp	Value	Host
1	2016-07-12	1.22	example.org
2	2016-07-12	5.33	

Comparison of Open Source TSDBs

Real-world example: VividCortex (I)

- SaaS platform
- MySQL Community Server + InnoDB (storage subsystem)
- Ingesting 332,000 values/s
- 3 AWS EC2 Servers (8 vCPUs, 26 GB Ram → ~ t2.2xlarge)
- Basic queries like Insert or SUM
- Trade-Offs:
  - Batch-wise ingestion into Vectors
  - Vectors consist of delta values
  - Ad-hoc queries are not possible → using a service instead
  - Grouping/Sharding must be manually decided when cluster is built

From: VividCortex: Building a Time-Series Database in MySQL, 2014, url: http://de.slideshare.net/vividcortex/vividcortex-building-a-timeseriesdatabase-in-mysql

Comparison of Open Source TSDBs

Real-world example: VividCortex (II)

- MySQL

ID	Sec1	Sec2	Sec3	Sec4	Sec5	Sec6	Sec7	Sec8	…
1	5.5	+0.7	-0.8	+0.3	-2.33	+1.0	-3.2	+0.0	…
2	3.7	+1.2	-3.4	+2.3	-0.55	+0.3	-5.0	+2.0	…
…	…	…	…	…	…	…	…	…	…

- InnoDB

Host	Time series	Timestamp	ID
db.example.org	CPU Temperature	2016-07-12T00:00:00Z	1
db2.example2.org	RAM Utilization	2016-07-11T00:00:01Z	2
…	…	…	…

From: VividCortex: Building a Time-Series Database in MySQL, 2014, url: http://de.slideshare.net/vividcortex/vividcortex-building-a-timeseriesdatabase-in-mysql

Comparison of Open Source TSDBs

Real-world example: VividCortex (II)

- MySQL

ID	Sec1	Sec2	Sec3	Sec4	Sec5	Sec6	Sec7	Sec8	…
1	5.5	+0.7	-0.8	+0.3	-2.33	+1.0	-3.2	+0.0	…
2	3.7	+1.2	-3.4	+2.3	-0.55	+0.3	-5.0	+2.0	…
…	…	…	…	…	…	…	…	…	…

- InnoDB

Host	Time series	Timestamp	ID
db.example.org	CPU Temperature	2016-07-12T00:00:00Z	1
db2.example2.org	RAM Utilization	2016-07-11T00:00:01Z	2
…	…	…	…

From: VividCortex: Building a Time-Series Database in MySQL, 2014, url: http://de.slideshare.net/vividcortex/vividcortex-building-a-timeseriesdatabase-in-mysql

Comparison of Open Source TSDBs

Real-world example: VividCortex (II)

- MySQL

ID	Sec1	Sec2	Sec3	Sec4	Sec5	Sec6	Sec7	Sec8	…
1	5.5	+0.7	-0.8	+0.3	-2.33	+1.0	-3.2	+0.0	…
2	3.7	+1.2	-3.4	+2.3	-0.55	+0.3	-5.0	+2.0	…
…	…	…	…	…	…	…	…	…	…

- InnoDB

Host	Time series	Timestamp	ID
db.example.org	CPU Temperature	2016-07-12T00:00:00Z	1
db2.example2.org	RAM Utilization	2016-07-11T00:00:01Z	2
…	…	…	…

From: VividCortex: Building a Time-Series Database in MySQL, 2014, url: http://de.slideshare.net/vividcortex/vividcortex-building-a-timeseriesdatabase-in-mysql

Comparison of Open Source TSDBs

Real-world example: VividCortex (II)

- MySQL

ID	Sec1	Sec2	Sec3	Sec4	Sec5	Sec6	Sec7	Sec8	…
1	5.5	+0.7	-0.8	+0.3	-2.33	+1.0	-3.2	+0.0	…
2	3.7	+1.2	-3.4	+2.3	-0.55	+0.3	-5.0	+2.0	…
…	…	…	…	…	…	…	…	…	…

- InnoDB

Host	Time series	Timestamp	ID
db.example.org	CPU Temperature	2016-07-12T00:00:00Z	1
db2.example2.org	RAM Utilization	2016-07-11T00:00:01Z	2
…	…	…	…

From: VividCortex: Building a Time-Series Database in MySQL, 2014, url: http://de.slideshare.net/vividcortex/vividcortex-building-a-timeseriesdatabase-in-mysql

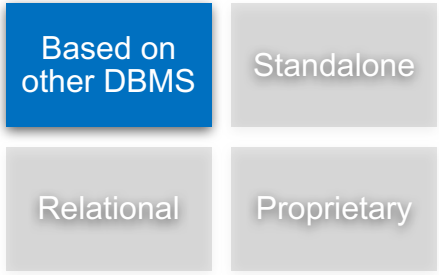
Comparison of Open Source TSDBs

How to rank by popularity?

By amount of citations/papers	DB-Engines	Google ranking by amount of search hits
• Many open source TSDBs do not have any papers or are not mentioned	• 17 of 85 TSDBs only	• Fitting search terms required • Partially represents amount of discussion

Comparison of Open Source TSDBs

Group 1: Based on other DBMS



TSDB	Search Hits
OpenTSDB	**12900**
Rhombus	**11700**
Newts	**6610**
KairosDB	**3130**
BlueFlood	**2010**
Gorilla	1520
Heroic	1490
Arctic	1330
Hawkular	1220
Apache Chukwa	858
BtrDB	637
tsdb: A Compressed Database for Time Series	634
Energy Databus	605
Tgres	445
SiteWhere	436
Kairos	380
Cube	266
…	…

TSDB	Search Hits
…	…
SkyDB	190
Chronix Server	148
MetricTank	<100

Compared TSDBs are written in bold.

Comparison of Open Source TSDBs

Group 2: Standalone

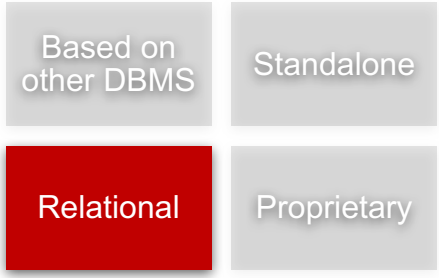
TSDB	Search Hits
Elasticsearch	**38000**
MonetDB	**37200**
Prometheus	**33700**
Druid	**28900**
InfluxDB	**28900**
RRDtool	22600
Atlas	7960
Gnocchi	5320
Whisper	5210
SciDB	4140
BlinkDB	2250
TSDB	1640
Seriesly	1330
TsTables	1100
Warp 10	1020
Akumuli	741
DalmatinerDB	527
…	…

TSDB	Search Hits
…	…
TimeStore	443
BEMOSS	391
YAWNDB	385
Vaultaire	339
Bolt	176
GridMW	<100
Node-tsdb	<100
NilmDB	<100

Compared TSDBs are bold.

Comparison of Open Source TSDBs

Group 3: Relational



TSDB	Search Hits
MySQL Community Edition	**309000**
PostgreSQL	**131000**
MySQL Cluster	23800
TimeTravel	743
PostgreSQL TS	<100

Compared TSDBs are written in bold.

Comparison of Open Source TSDBs

Group 4: Proprietary

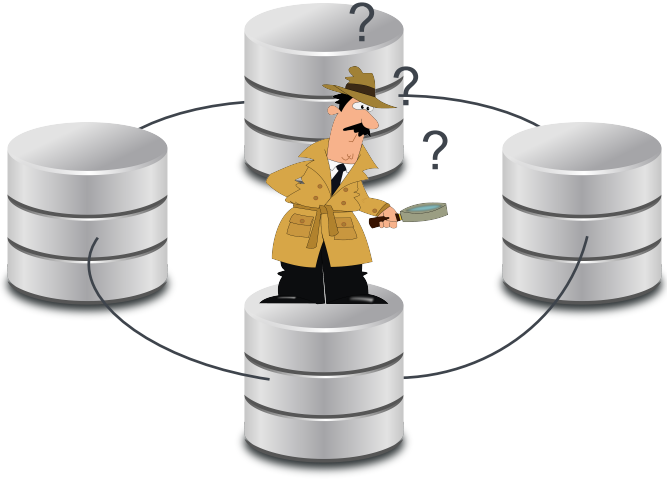
Based on other DBMS | Standalone | Relational | **Proprietary**

TSDB	Search Hits
Microsoft SQL Server	94000
Oracle Database	71500
Splunk	30600
SAP HANA	22100
Treasure Data	15000
DataStax Enterprise	12500
FoundationDB	11300
Riak TS	9720
TempoIQ	8810
kdb+	8220
IBM Informix	7580
Cityzen Data	6400
Sqrrl	5460
Databus	5100
Kerf	3850
Aerospike	3740
OSIsoft PI	3200
…	…

TSDB	Search Hits
…	…
Geras	3030
Axibase Time Series Database	2420
eXtremeDB Financial Edition	1660
Prognoz Platform	1440
Acunu	1360
SkySpark	1240
ParStream	1140
Mesap	741
ONETick Time-Series Tick Database	503
TimeSeries.Guru	464
New Relic Insights	233
Squwk	191
Polyhedra IMDB	149
TimeScape EDM+	<100
PulsarTSDB	<100
Uniformance Process History Database (PHD)	<100

Comparison of Open Source TSDBs

Criteria Group 1: Distribution & Clusterability



Comparison of Open Source TSDBs

Criteria Group 2: Functions

Distribution & Clusterability

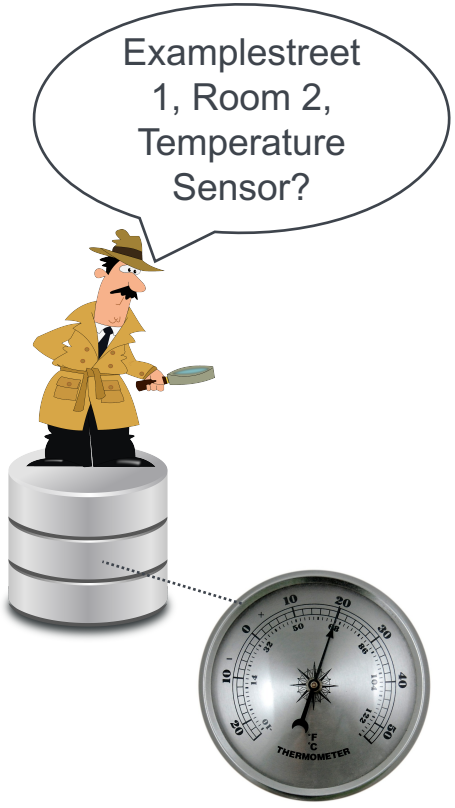
Functions

Comparison of Open Source TSDBs

Criteria Group 3: Advanced Functions

Distribution & Clusterability

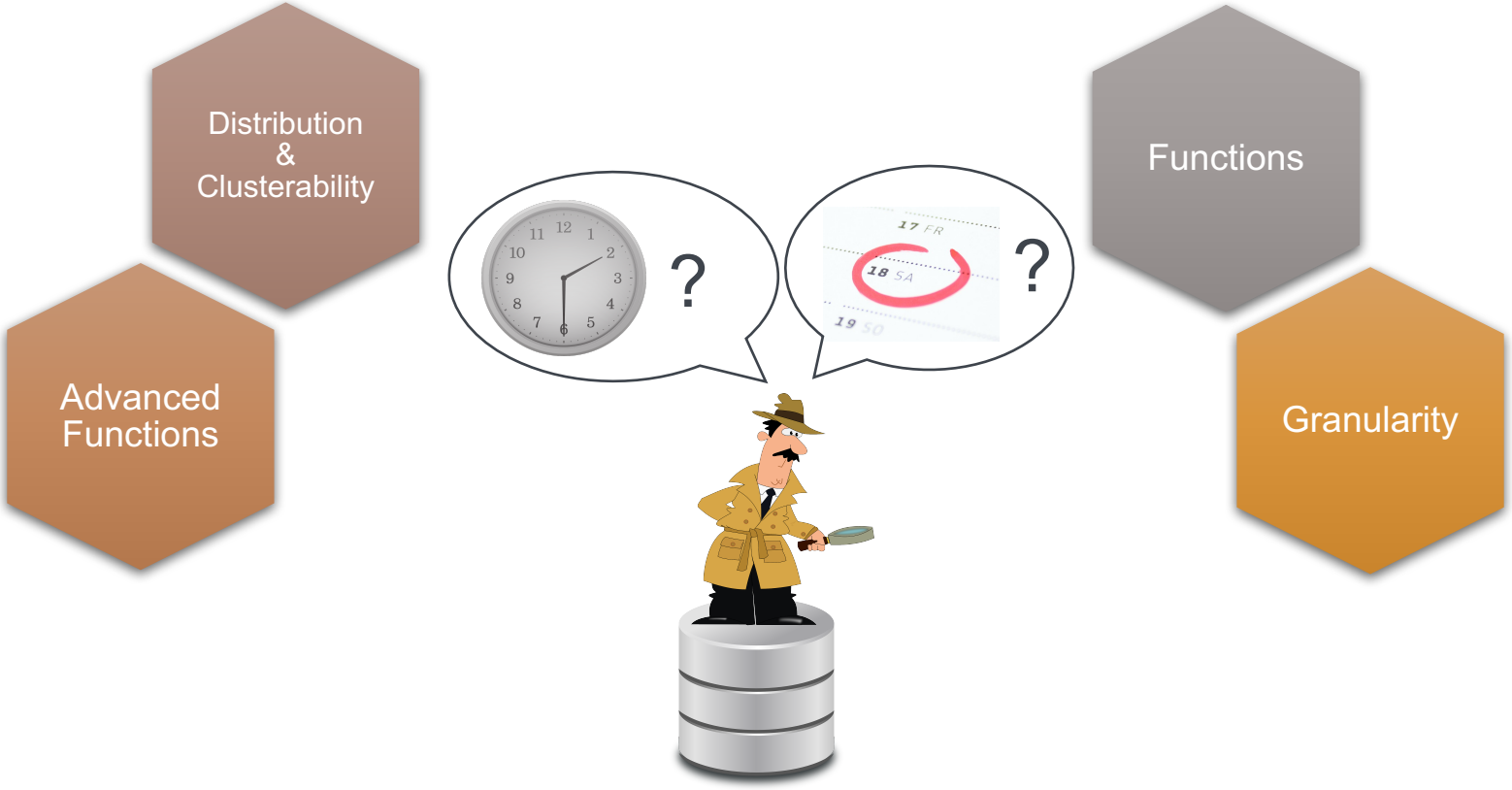
Advanced Functions



Functions

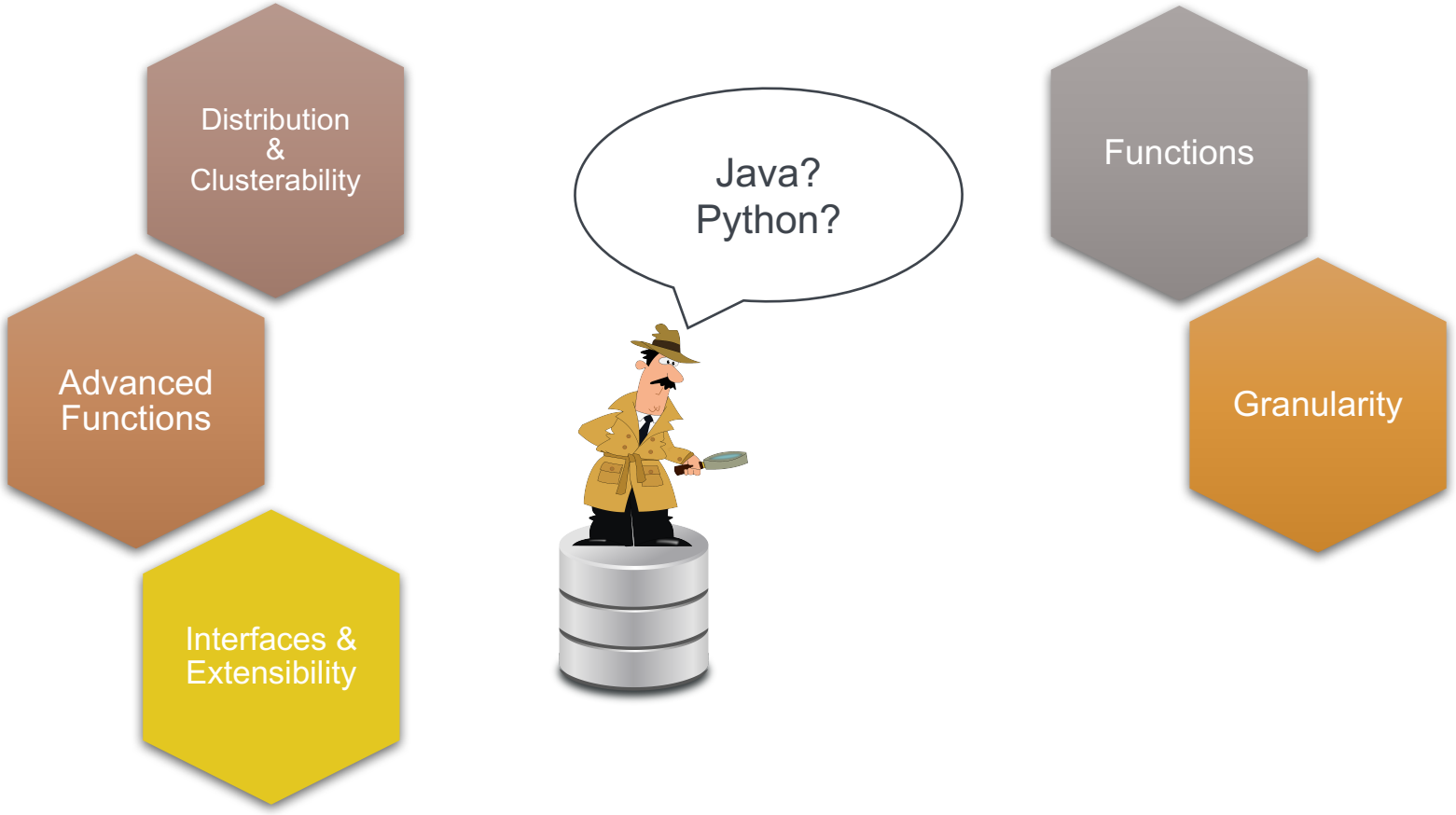
Comparison of Open Source TSDBs

Criteria Group 4: Granularity



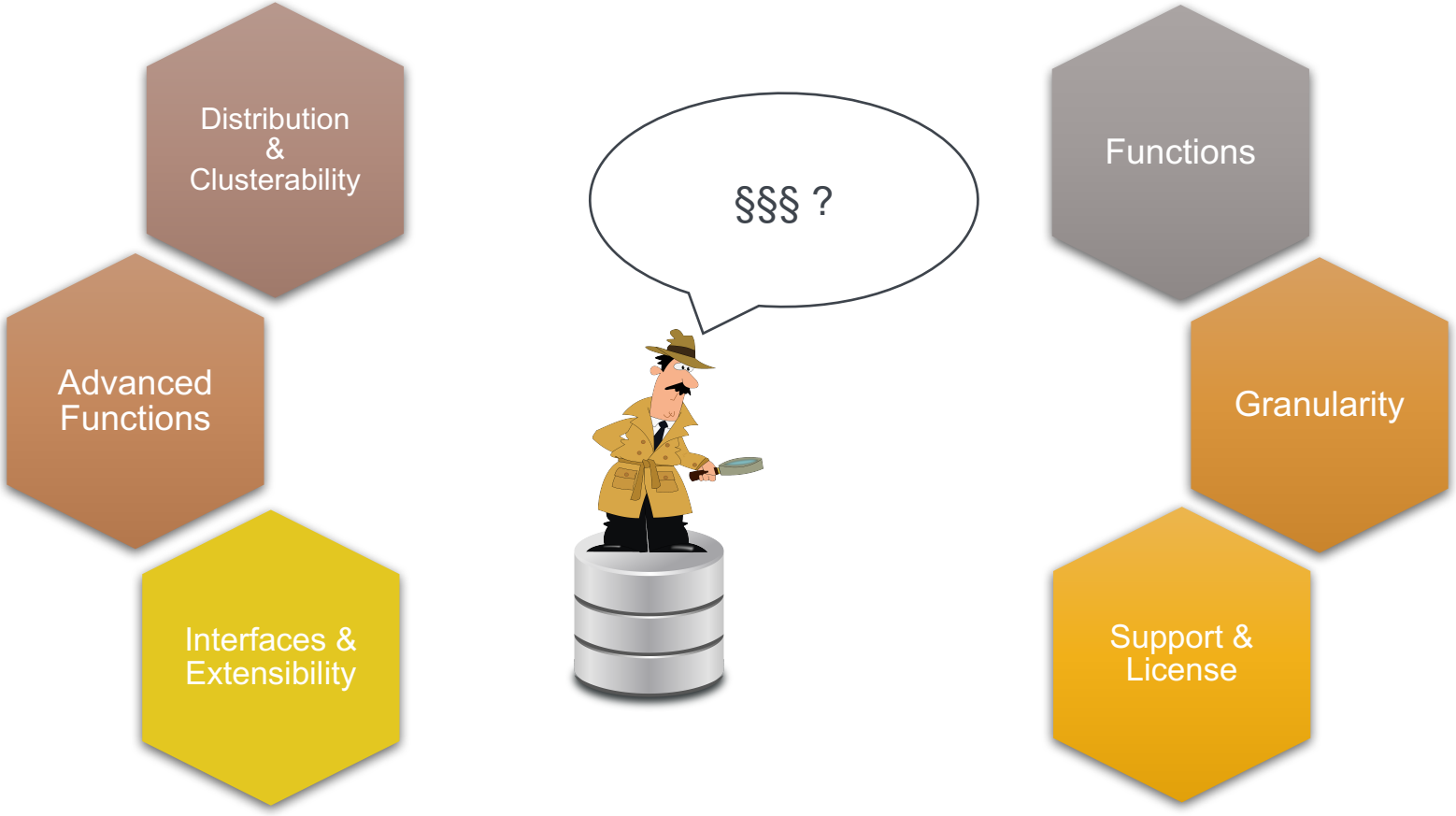
Comparison of Open Source TSDBs

Criteria Group 5: Interfaces & Extensibility



Comparison of Open Source TSDBs

Criteria Group 1: Support & License

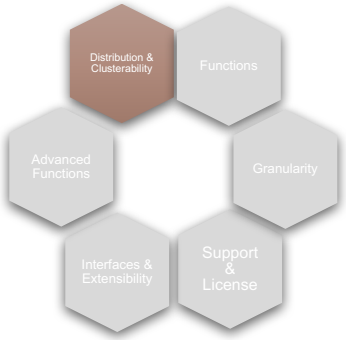


Comparison of Open Source TSDBs

Six Criteria Groups

Comparison of Open Source TSDBs

Criteria Group 1: Distribution & Clusterability

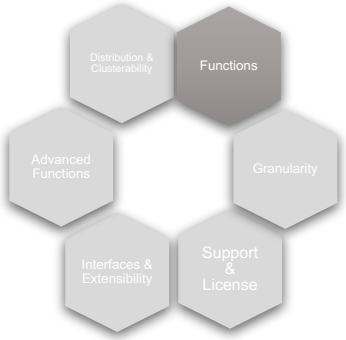


TSDB	High Availability	Scalability	Load Balancing
Group 1: Based on other DBMS			
Blueflood	✔	(✔)	(✔)
KairosDB	(✔)	(✔)	(✔)
NewTS	(✔)	(✔)	(✔)
OpenTSDB	(✔)	(✔)	(✔)
Rhombus	(✔)	(✔)	(✔)
Group 2: Standalone			
Druid	✔	✔	✔
Elasticsearch	✔	✔	✔
InfluxDB	✔	✘	✔
MonetDB	✔	(✔)	(✔)
Prometheus	✘	(✔)	(✔)
Group 3: Relational			
MySQL	✘	✘	✘
PostgreSQL	✘	✘	✘

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

Criteria Group 2: (Basic) Functions

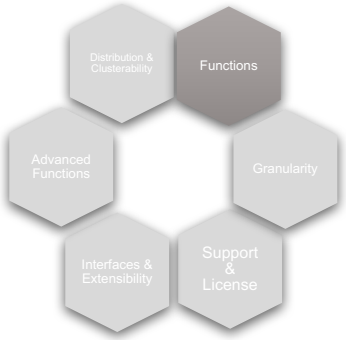


TSDB	INS	READ	SCAN	UPD	DEL
Group 1: Based on other DBMS					
Blueflood	✔	✔	✔	✘	✘
KairosDB	✔	✔	✔	✘	✔
NewTS	✔	✔	✔	✘	✘
OpenTSDB	✔	✔	✔	✘	✘
Rhombus	✔	✔	✔	✔	✔
Group 2: Standalone					
Druid	✔	✔	✔	✘	✘
Elasticsearch	✔	✔	✔	✔	✔
InfluxDB	✔	✔	✔	✔	✔
MonetDB	✔	✔	✔	✔	✔
Prometheus	✔	✔	✔	✘	✘
Group 3: Relational					
MySQL	✔	✔	✔	✔	✔
PostgreSQL	✔	✔	✔	✔	✔

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

Criteria Group 2: (Aggregating) Functions

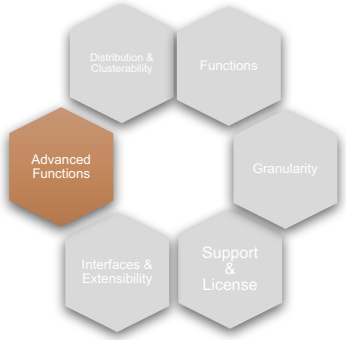


TSDB	AVG	SUM	CNT	MAX	MIN
Group 1: Based on other DBMS					
Blueflood	✔	✘	✔	✔	✔
KairosDB	✔	✔	✔	✔	✔
NewTS	✔	✘	✘	✔	✔
OpenTSDB	✔	✔	✔	✔	✔
Rhombus	✘	✘	✔	✘	✘
Group 2: Standalone					
Druid	✔	✔	✔	✔	✔
Elasticsearch	✔	✔	✔	✔	✔
InfluxDB	✔	✔	✔	✔	✔
MonetDB	✔	✔	✔	✔	✔
Prometheus	✔	✔	✔	✔	✔
Group 3: Relational					
MySQL	✔	✔	✔	✔	✔
PostgreSQL	✔	✔	✔	✔	✔

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

Criteria group 3: Advanced Functions

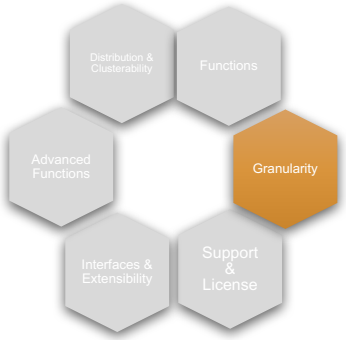


TSDB	Continuous Calculation	Tags	Long-Term Storage	Matrix Time Series
Group 1: Based on other DBMS				
Blueflood	✘	✘	✔	✘
KairosDB	✘	✔	✘	✘
NewTS	✔	(✔)	✘	✘
OpenTSDB	✘	✔	✘	✘
Rhombus	✘	(✔)	✘	✘
Group 2: Standalone				
Druid	✔	✔	✔	✘
Elasticsearch	✔	✔	✘	✔
InfluxDB	✔	✔	✔	✘
MonetDB	✔	✔	✘	✔
Prometheus	✔	✔	✘	✘
Group 3: Relational				
MySQL	✔	✔	✘	✔
PostgreSQL	✔	✔	✘	✔

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

Criteria Group 4: Granularity



TSDB	Downsampling	Smallest Sample Interval	Smallest Granularity for Storage	Smallest Guaranteed Granularity for Storage
Group 1: Based on other DBMS				
Blueflood	✘	✘	1 ms	1 ms
KairosDB	✔	1 ms	1 ms	1 ms
NewTS	✔	2 ms	1 ms	1 ms
OpenTSDB	✔	1 ms	1 ms	> 1 ms
Rhombus	✘	✘	1 ms	1 ms
Group 2: Standalone				
Druid	✔	1 ms	1 ms	1 ms
Elasticsearch	✘	1 ms	1 ms	1 ms
InfluxDB	✔	1 ms	1 ms	1 ms
MonetDB	✔	1 ms	1 ms	1 ms
Prometheus	✔	1 ms	1 ms	1 ms
Group 3: Relational				
MySQL	✔	1 ms	1 ms	1 ms
PostgreSQL	✔	1 ms	1 ms	1 ms

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

Criteria Group 5: Interfaces and Extensibility (I)

TSDB	APIs & Interfaces	Client Libraries	Plugins
Group 1: Based on other DBMS			
Blueflood	CLI, Graphite, HTTP (JSON), Kafka, statsD, UDP	✘	✘
KairosDB	CLI, Graphite, HTTP (REST+JSON, GUI), telnet	Java	✔
NewTS	HTTP (REST+JSON, GUI)	Java	✘
OpenTSDB	Azure, CLI, HTTP (REST+JSON, GUI), Kafka, RabbitMQ, S3, Spritzer	✘	✔
Rhombus	✘	Java	✘
Group 2: Standalone			
Druid	CLI, HTTP (REST+JSON, GUI), Samza, Spark, Storm	Java, Python, R	✔
Elasticsearch	HTTP (REST+JSON)	Groovy, Java, .NET, Perl, PHP, Python, Ruby	✔
InfluxDB	Collect, CLI, Graphite, HTTP (InfluxQL, GUI), OpenTSDB, UDP	✘	✔

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

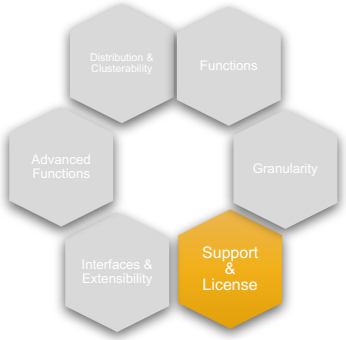
Criteria Group 5: Interfaces and Extensibility (II)

TSDB	APIs & Interfaces	Client Libraries	Plugins
Group 2: Standalone			
MonetDB	CLI	Java (JDBC), Mapi (C Binding), Node.js, ODBC, Perl, PHP, Python, Ruby	✘
Prometheus	CLI, HTTP (JSON, GUI)	Go, Java, Python, Ruby	✘
Group 3: Relational			
MySQL	CLI	J, Java (JDBC), ODBC, Python, …	✔
PostgreSQL	CLI	C, C++, Java (JDBC), ODBC, Python, Tcl, …	✔

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

Criteria Group 6: Support and License



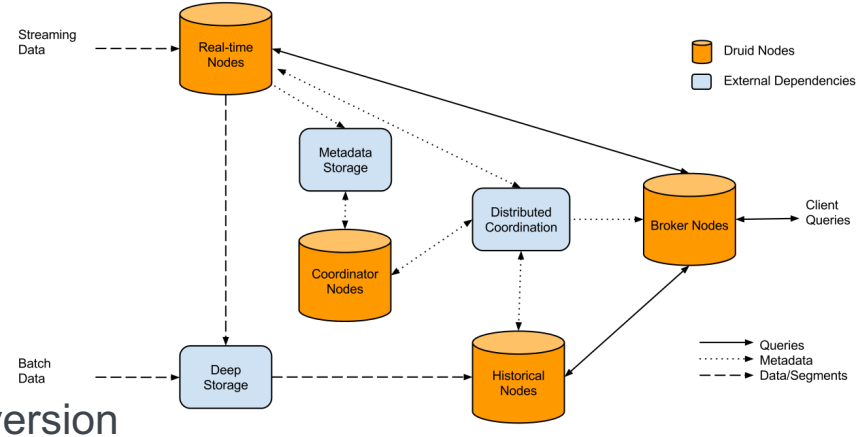
TSDB	LTS & Stable Version	Commercial Support	License
Group 1: Based on other DBMS			
Blueflood	✘	✘	Apache 2.0
KairosDB	✘	✘	Apache 2.0
NewTS	✘	✘	Apache 2.0
OpenTSDB	✘	✘	LGPLv2.1+, GPLv3+
Rhombus	✘	✘	MIT
Group 2: Standalone			
Druid	✘	✘	Apache 2.0
Elasticsearch	(✔)	✔	Apache 2.0
InfluxDB	✘	✔	MIT
MonetDB	✘	✔	MonetDB Public License
Prometheus	✘	✘	Apache 2.0
Group 3: Relational			
MySQL	✔	✔	GPLv2
PostgreSQL	✔	✘	The PostgreSQL License

✔ fullfilled
(✔) partially fullfilled
✘ not fullfilled

Comparison of Open Source TSDBs

Which TSDB would you choose?

- If stable version, matrix time series, and commercial support do not matter: Druid
  - Five different service types
  - Different ingestion methods
  - Support for Real-time Analysis
- If commercial support matters: InfluxDB
  - SQL-like language InfluxQL
  - Some cluster features only available in paid version
  - Only load-balancing is still possible
  - Grafana integration (for visualization)
- Otherwise: ElasticSearch, MySQL, or PostgreSQL



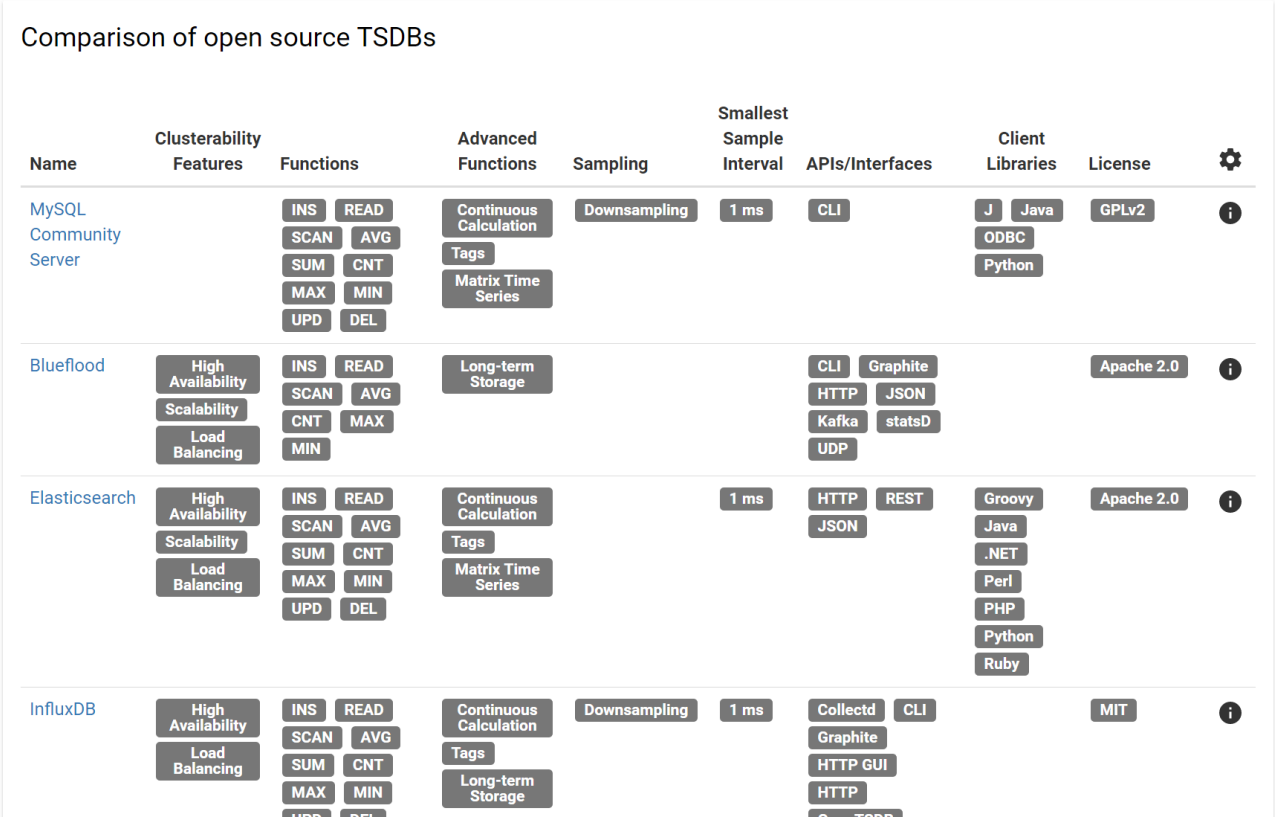
Comparison of Open Source TSDBs

Live Demo

Comparison of Open Source TSDBs

Interactive Version: Ultimate Comparison of Open Source TSDBs

https://tsdbbench.github.io/Ultimate-TSDB-Comparison/



Comparison of open source TSDBs

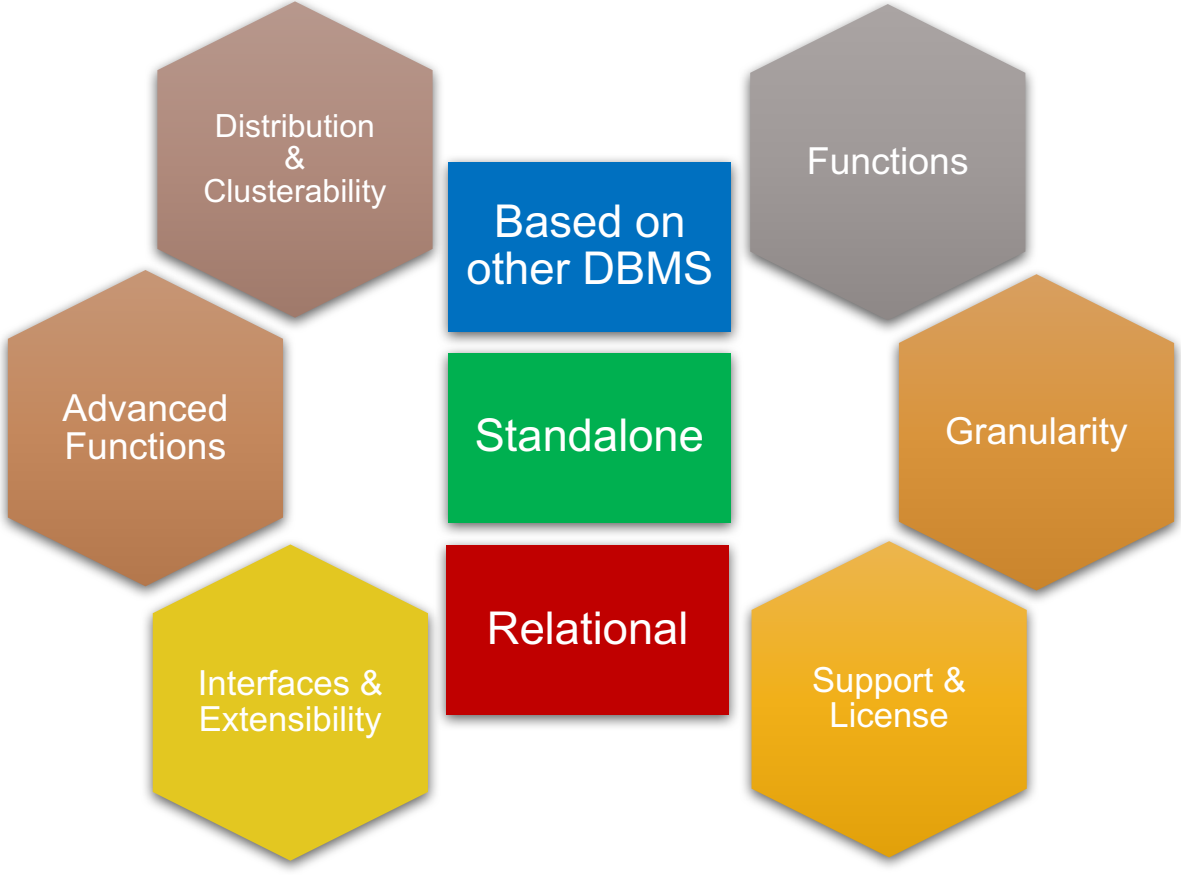
Name	Clusterability Features	Functions	Advanced Functions	Sampling	Smallest Sample Interval	APIs/Interfaces	Client Libraries	License
MySQL Community Server		INS READ SCAN AVG SUM CNT MAX MIN UPD DEL	Continuous Calculation Tags Matrix Time Series	Downsampling	1 ms	CLI	J Java ODBC Python	GPLv2
Blueflood	High Availability Scalability Load Balancing	INS READ SCAN AVG CNT MAX MIN	Long-term Storage			CLI Graphite HTTP JSON Kafka statsD UDP		Apache 2.0
Elasticsearch	High Availability Scalability Load Balancing	INS READ SCAN AVG SUM CNT MAX MIN UPD DEL	Continuous Calculation Tags Matrix Time Series		1 ms	HTTP REST JSON	Groovy Java .NET Perl PHP Python Ruby	Apache 2.0
InfluxDB	High Availability Load Balancing	INS READ SCAN AVG SUM CNT MAX MIN	Continuous Calculation Tags Long-term Storage	Downsampling	1 ms	Collectd CLI Graphite HTTP GUI HTTP		MIT

Comparison of Open Source TSDBs

Conclusion and Outlook

Comparison of Open Source TSDBs

Conclusion and Outlook



- Systematic Search
- 83 TSDBs
  - 50 of them open source
  - 12 of them compared
- Grouped in four groups
- Ranked by popularity
- Compared with 26 criterias
- No compared TSDB supports all features
- Most supported: load balancing, high availability, and scalability
- Most lacking: stable versions, commercial support, support for all query types
- Choice highly depends on the use case
- Next step: Performance Comparison (https://github.com/TSDBBench/)