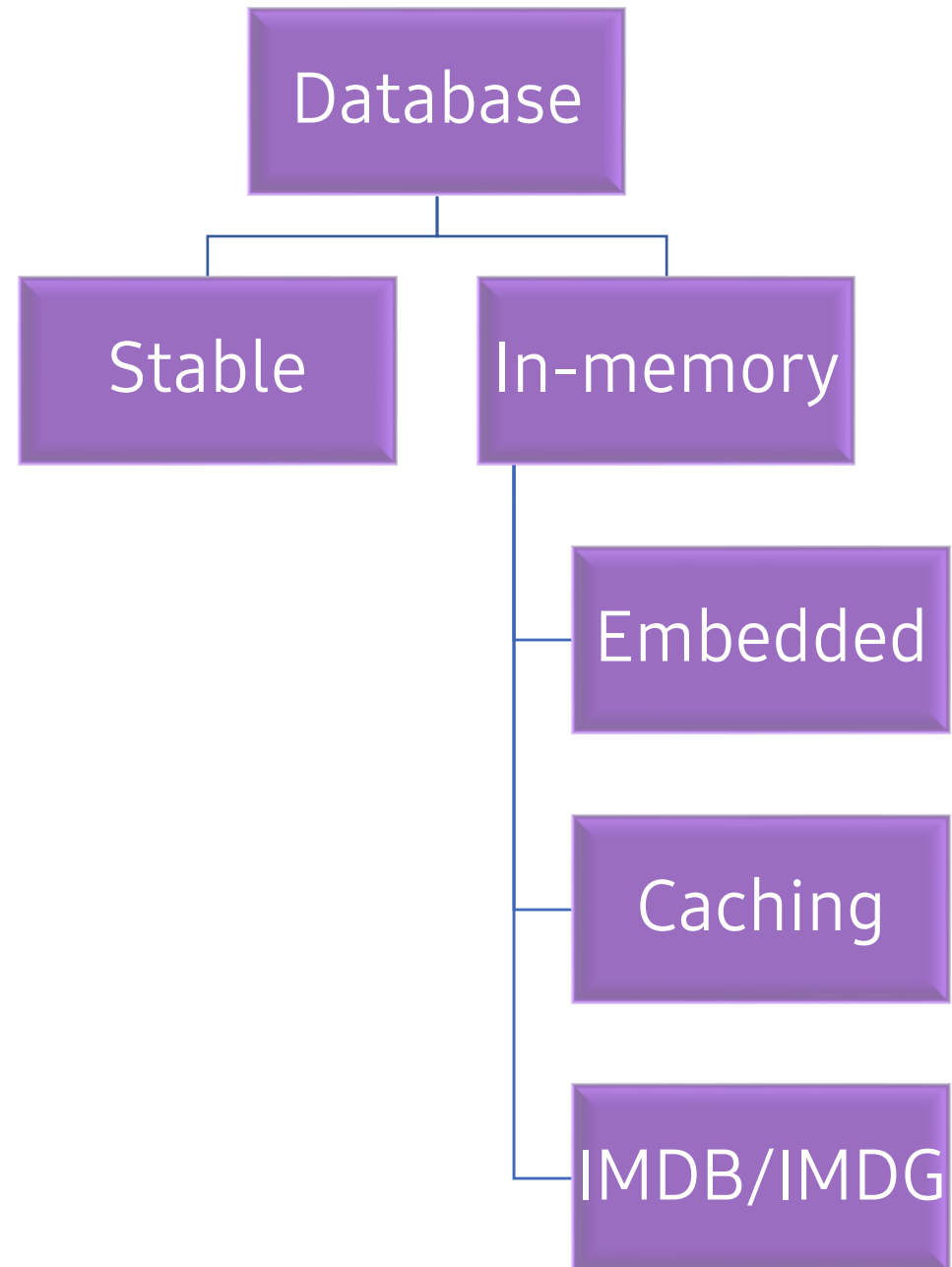


# Day 10

## Topic 2: Time series DB

Lecturer: M.S. Le Minh Tan



```
graph TD; Database[Database] --> SQL[SQL]; Database --> TimeSeries[Time series];
```

Database

SQL

Time  
series

# Time Series

- An ordered sequence of values of a variable (temperature) at equally spaced time intervals (e.g. hourly).

(26.0, 26.5, 28.5, 27.5)

Interval = 30'

# Use cases

- Time series analysis
- Regression analysis
- Time series forecasting

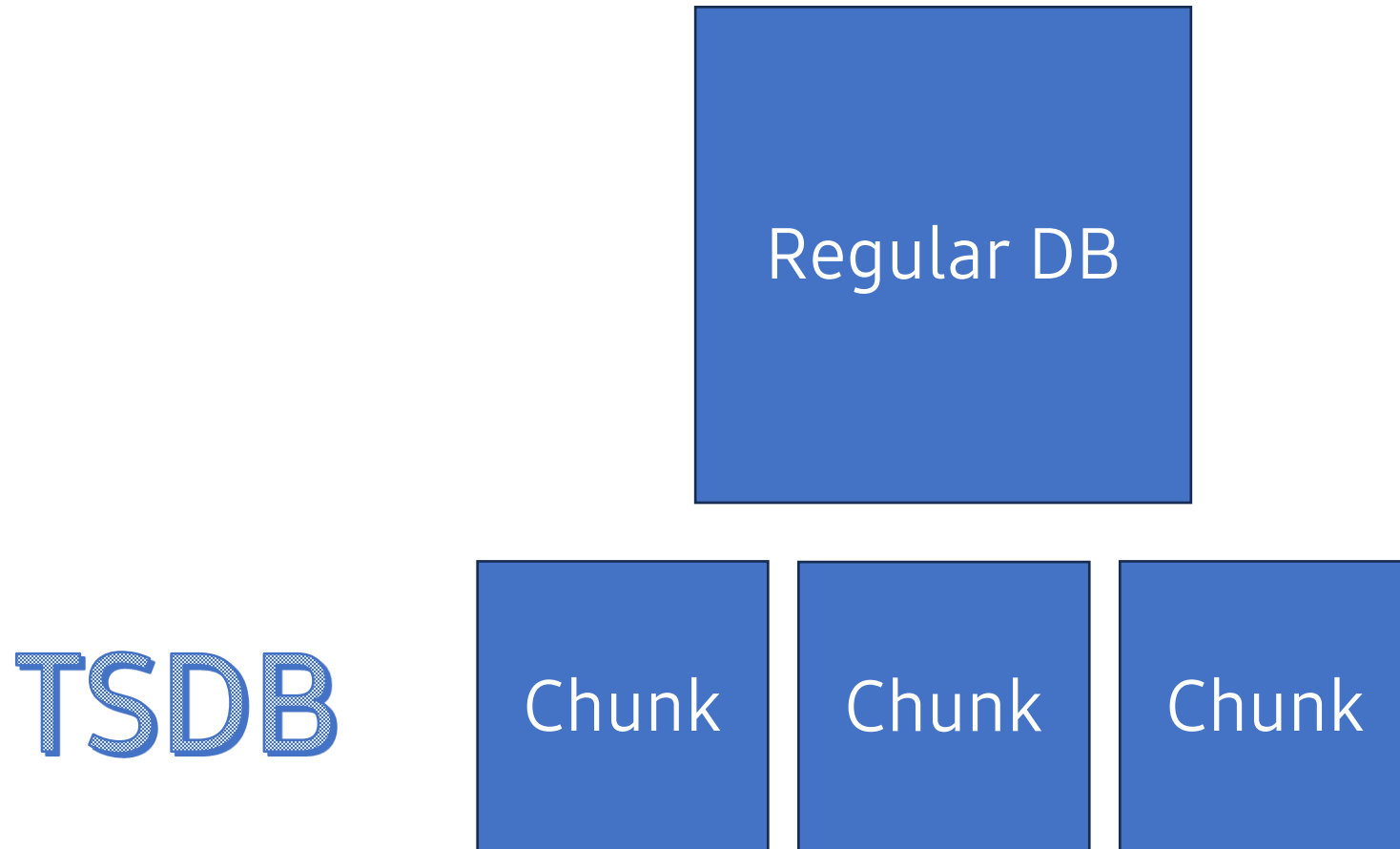
# Outlines

- I. Time Series Database (TSDB)
- II. Benefits & suggestions
- III. InfluxDB
- IV. Exercises

# I. Time Series Database (TSDB)

- A DB that is optimized for time series/time-stamped data.
- You can ... time series:
  - Create
  - Enumerate
  - Update
  - Destroy
  - Organize

# What's the benefits? Data location





## 2. Fast, easy range queries

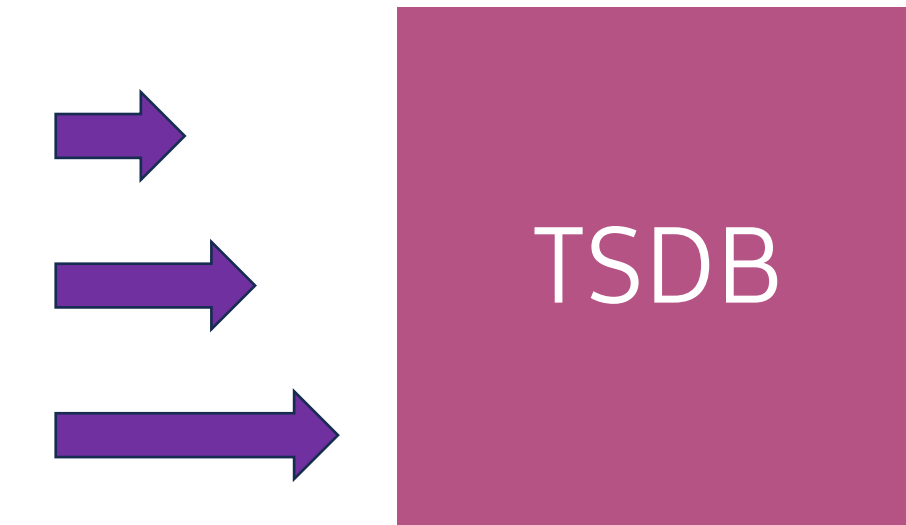
Id	Name	Faculty
15110124	Tam	KTDL
15110125	Toan	In & TT
15110123	Tan	CNTT
15110126	Mai	ATTT

Timestamped	Temp	Humid
9:00	26.5	0.50
9:30	27.0	0.70
10:00	28.0	0.80
10:30	29.0	0.85



Pre-sorted

### 3. High write performance



**FAST!**

FAST!

## 4. Data compression

StudyYear	Course	AvgScore
22	NMLT	8.8
21	NMLT	9.0
20	NMLT	8.7
20	KTLT	8.5

...

...

...

StudyYear	AvgScore_NMLT	AvgScore_KTLT
22	8.8	8.7
21	9.0	9.2
20	8.7	8.3
20	7.5	8.5

...

...

...

# Data roll-ups

StudyYear	AvgScore_NMLT	AvgScore_KTLT
22	8.8	8.7
21	9.0	9.2
20	8.7	8.3
20	7.5	8.5

...

...

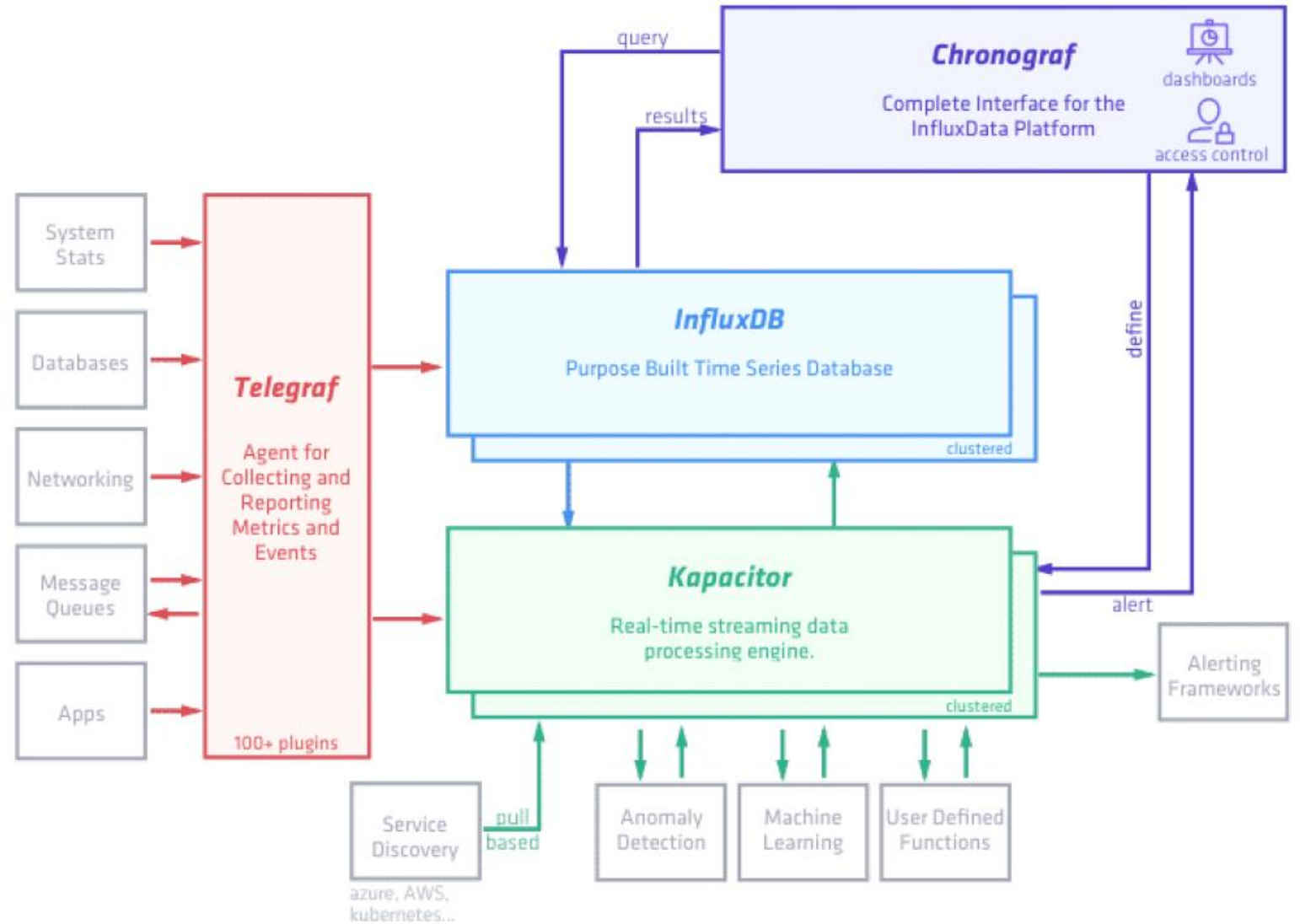
...

5. Scalability: Functionalities that work best with time-series.
6. Usability: Features for time series data analysis.

# Suggestions (2017)

1. InfluxDB
2. RRDtool
3. Graphite
4. OpenTSDB
5. Druid
6. Prometheus

# III. InfluxDB



# TICK Stack



# InfluxDB

- Time (1st column)
- Fields (next  $n$  columns): Non-indexed
  - Key – value
- Tags (next  $m$  columns): Indexed
  - Key – value

# Measurement

name=passengers				
time	minors	adults	location	driver
2015-08-18T00:00:00Z	1	2	1	doe
2015-08-18T00:00:00Z	2	2	1	jones
2015-08-18T00:06:00Z	1	1	1	doe
2015-08-18T00:06:00Z	0	1	1	jones
2015-08-18T05:54:00Z	0	2	2	doe
2015-08-18T06:30:00Z	2	2	2	doe
2015-08-18T06:06:00Z	3	1	2	jones
2015-08-18T06:30:00Z	0	4	2	jones

Point →



Time

Fields

Tags

Series Number	Retention Policy	Measurement	Tag set
Series 1	autogen	passengers	location=1, driver=doe
Series 2	autogen	passengers	location=1, driver=jones
Series 3	autogen	passengers	location=2, driver=doe
Series 4	autogen	passengers	location=2, driver=jones

# IV. Exercises

1. Install InfluxDB v2 (1 pt)

- <https://docs.influxdata.com/influxdb/v2/install/>
- Test: “influxd” command

2. Set up (2 pts)

- <https://docs.influxdata.com/influxdb/v2/get-started/setup/>
- Test: Bucket named “nhom\_n”.