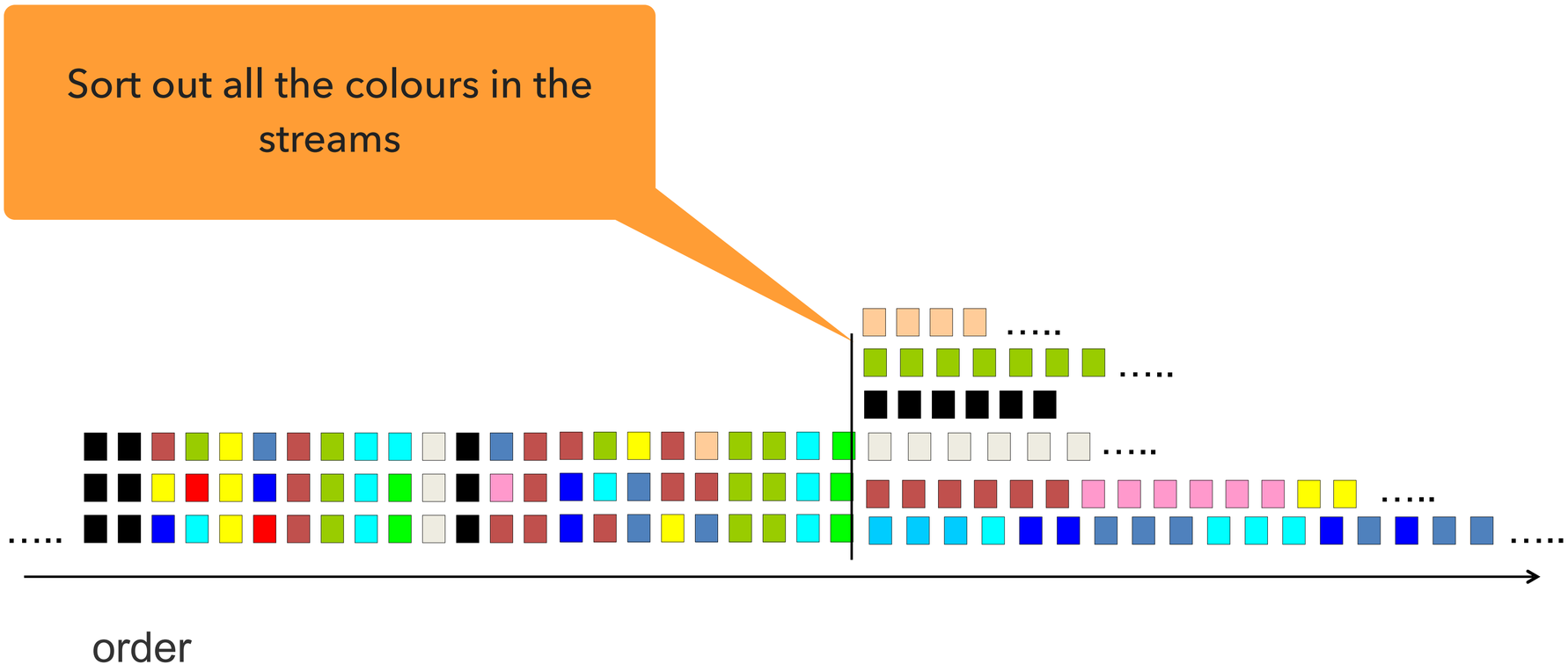
An Outlook on Graph Stream Processing Benchmarks

Riccardo Tommasini - 16/08/2021 - LDBCW@VLDB

Who I Am

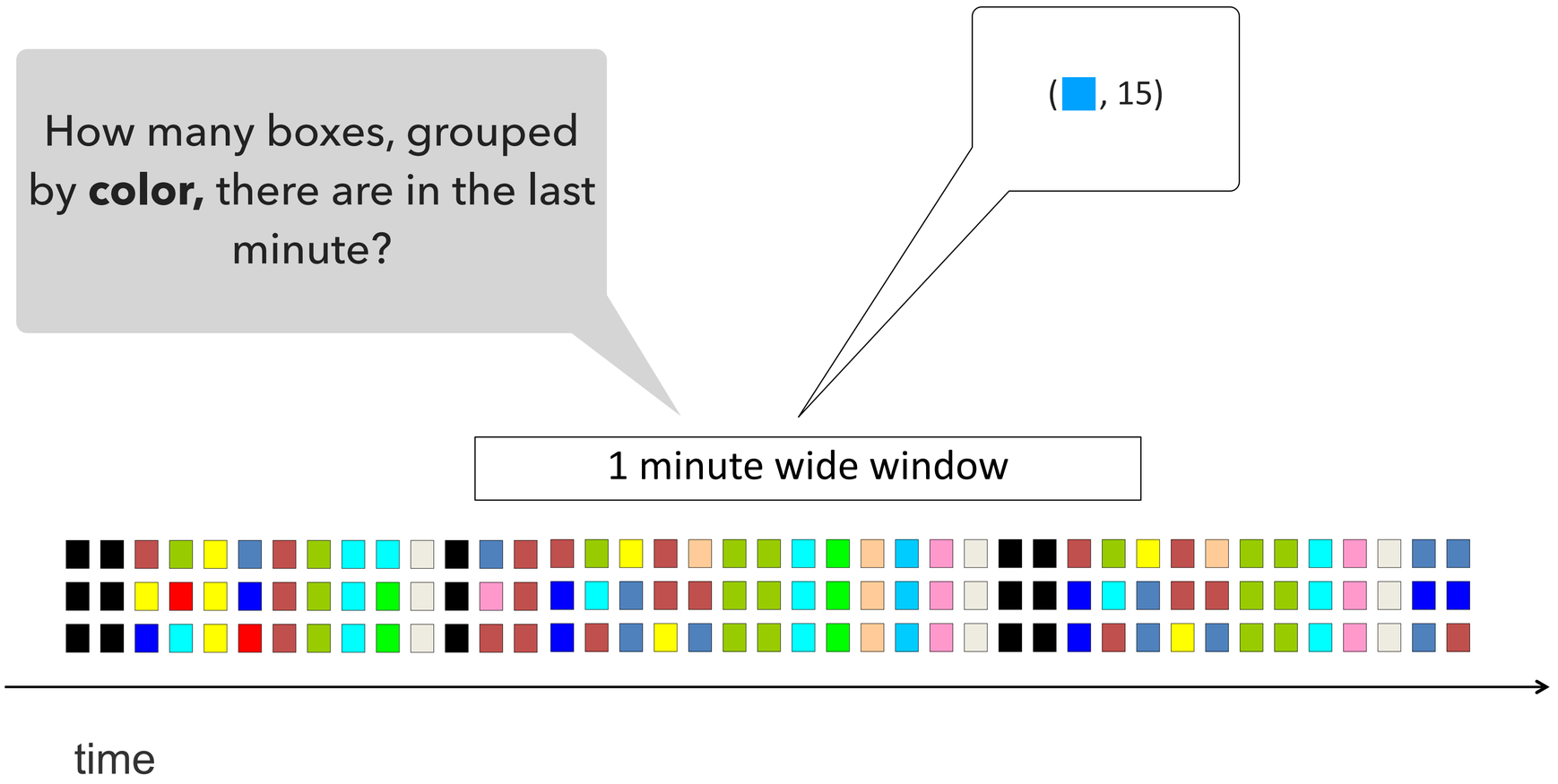
- Assistant Professor @ University of Tartu
- Maitre de Conférence @ INSA Lyon since 01/10/2021
- Maintainer of RSP4J library

What is Stream Processing

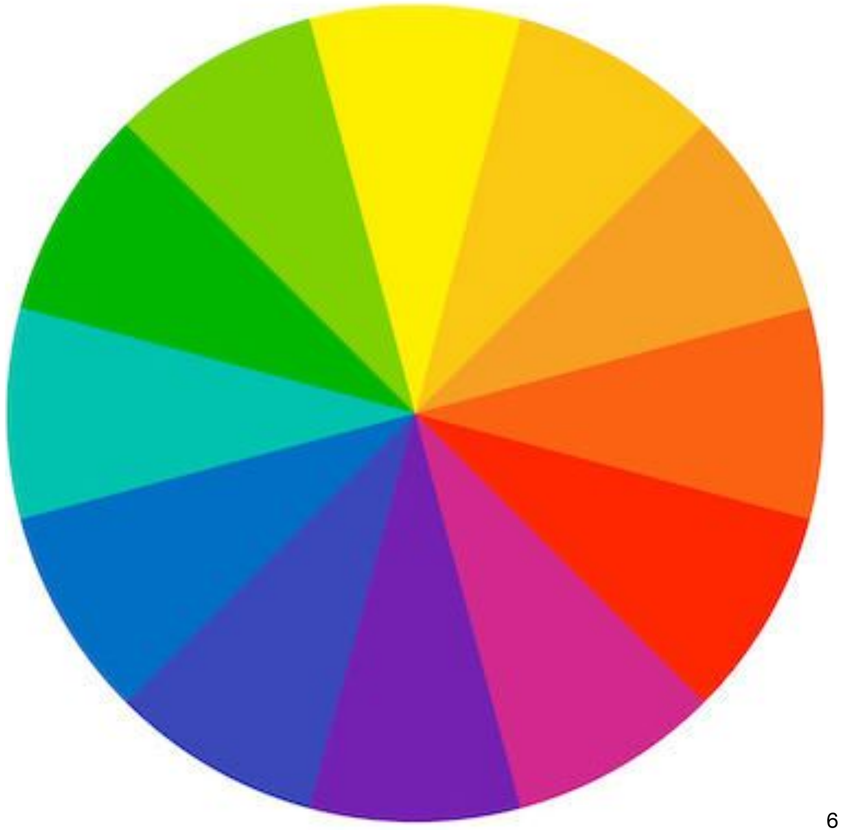


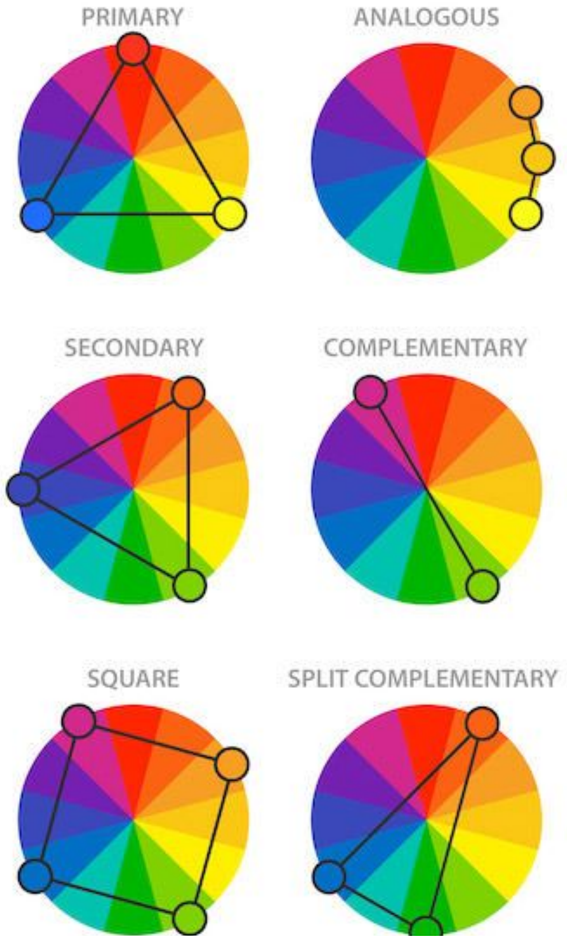
GIF制作

Data Stream Management Systems

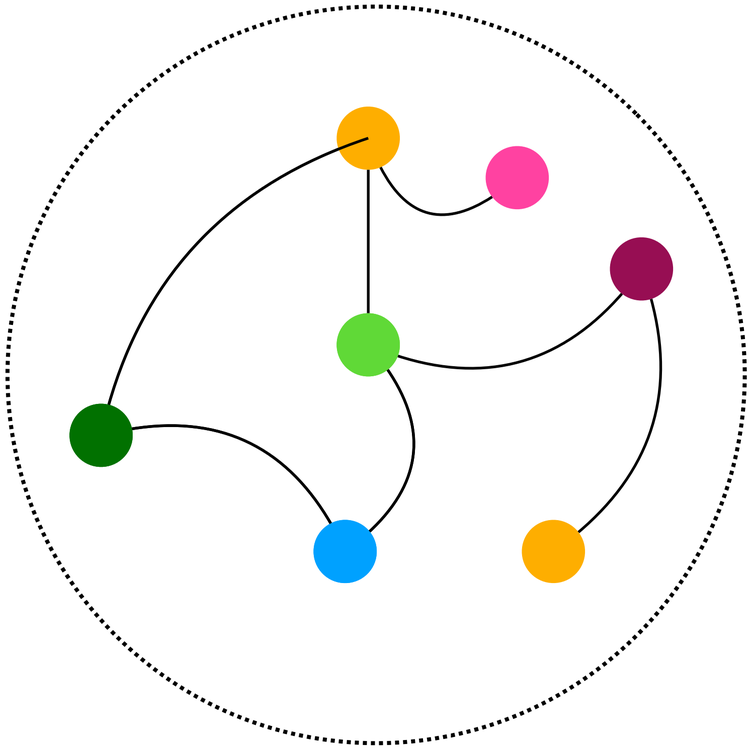


What About Graph?

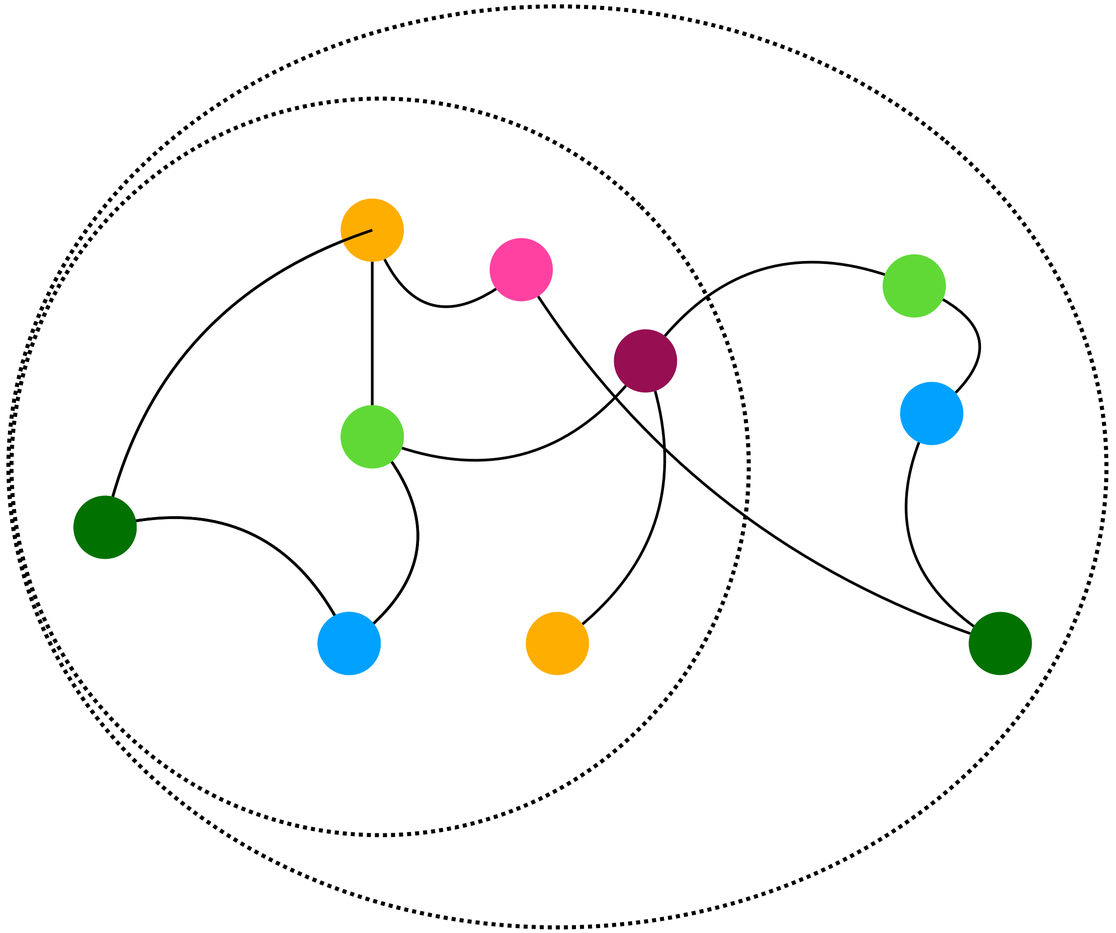




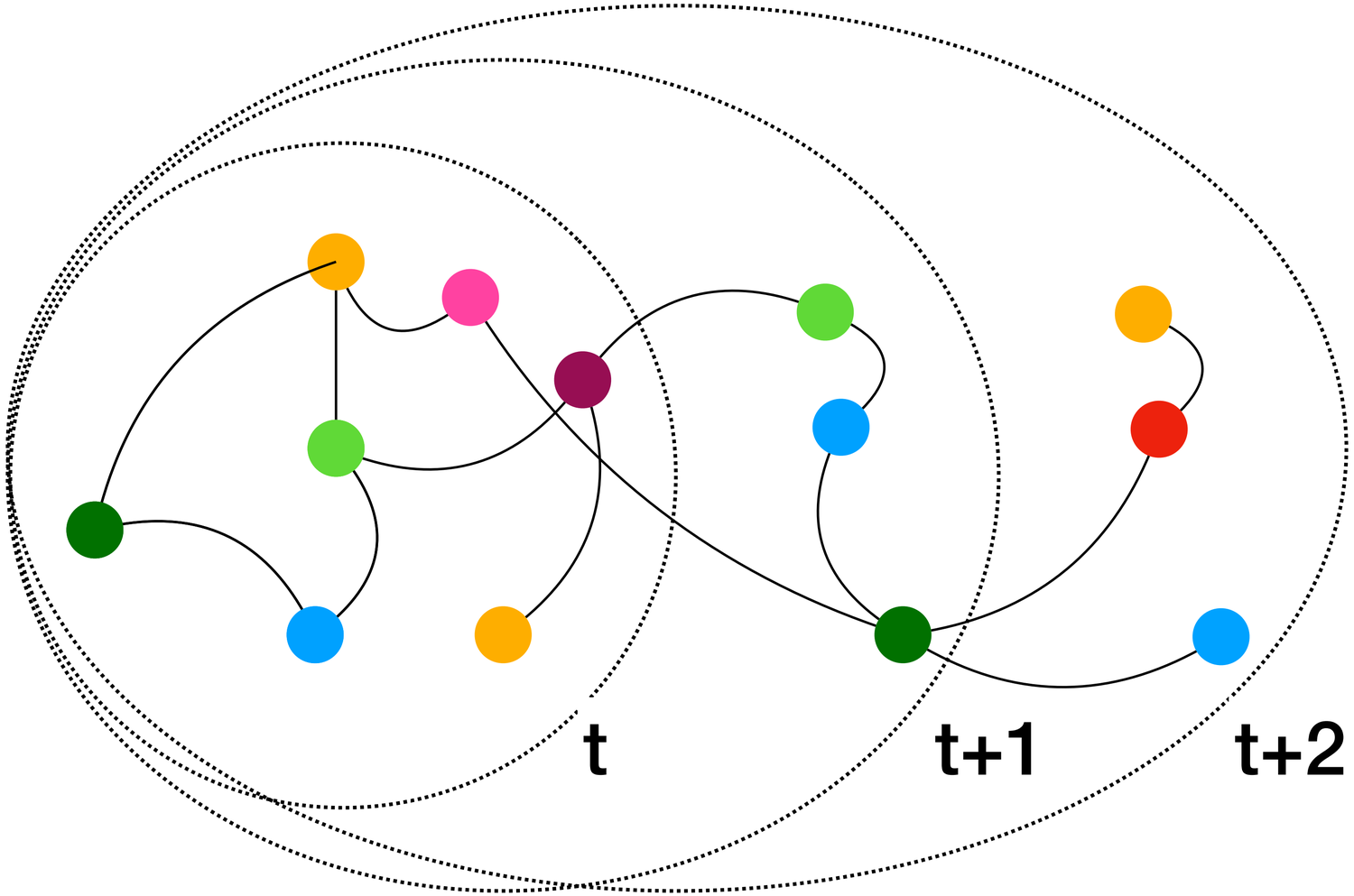
Streaming Graphs



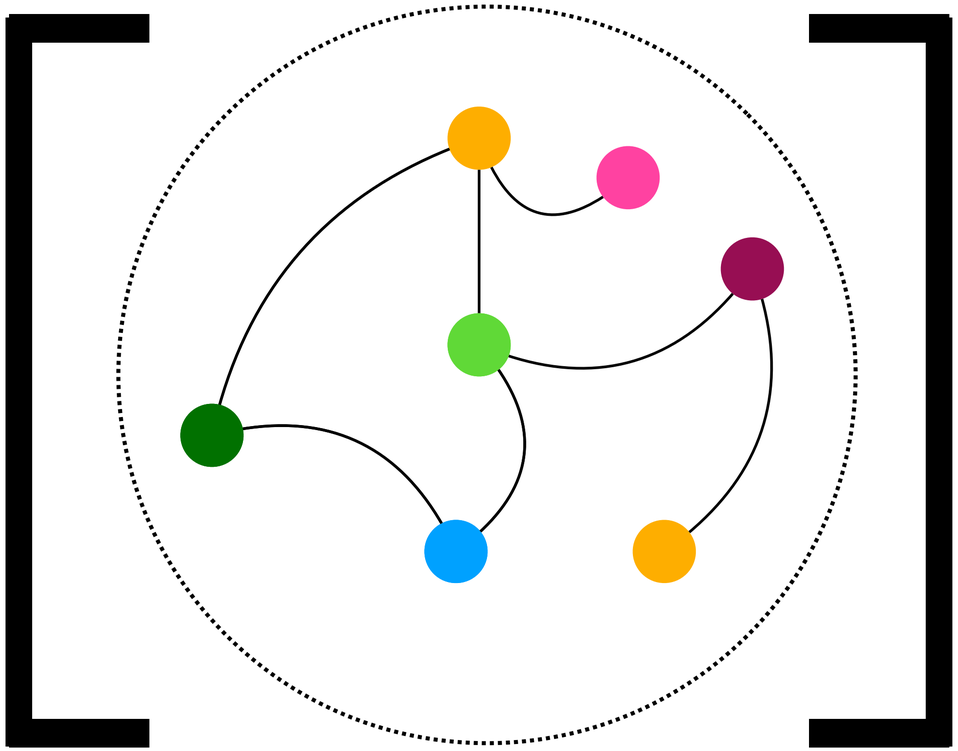
Streaming Graphs



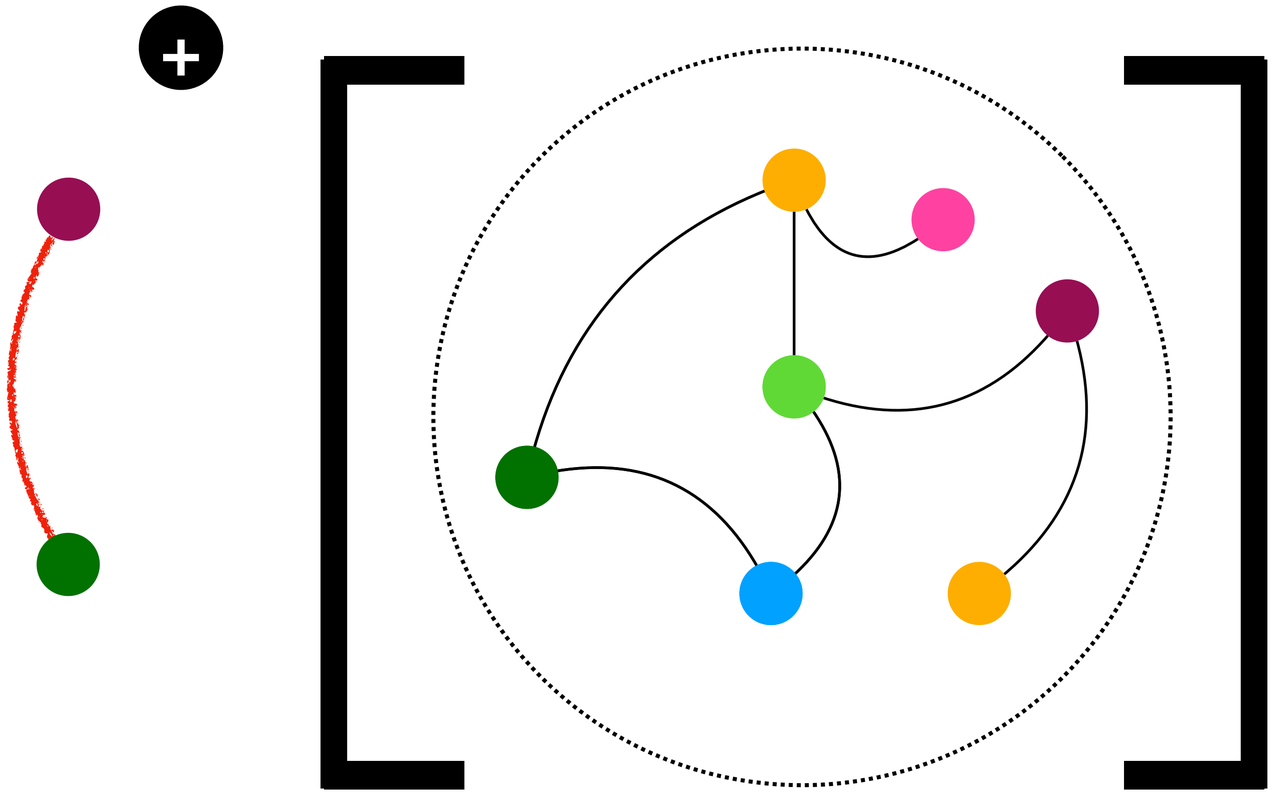
Streaming Graphs



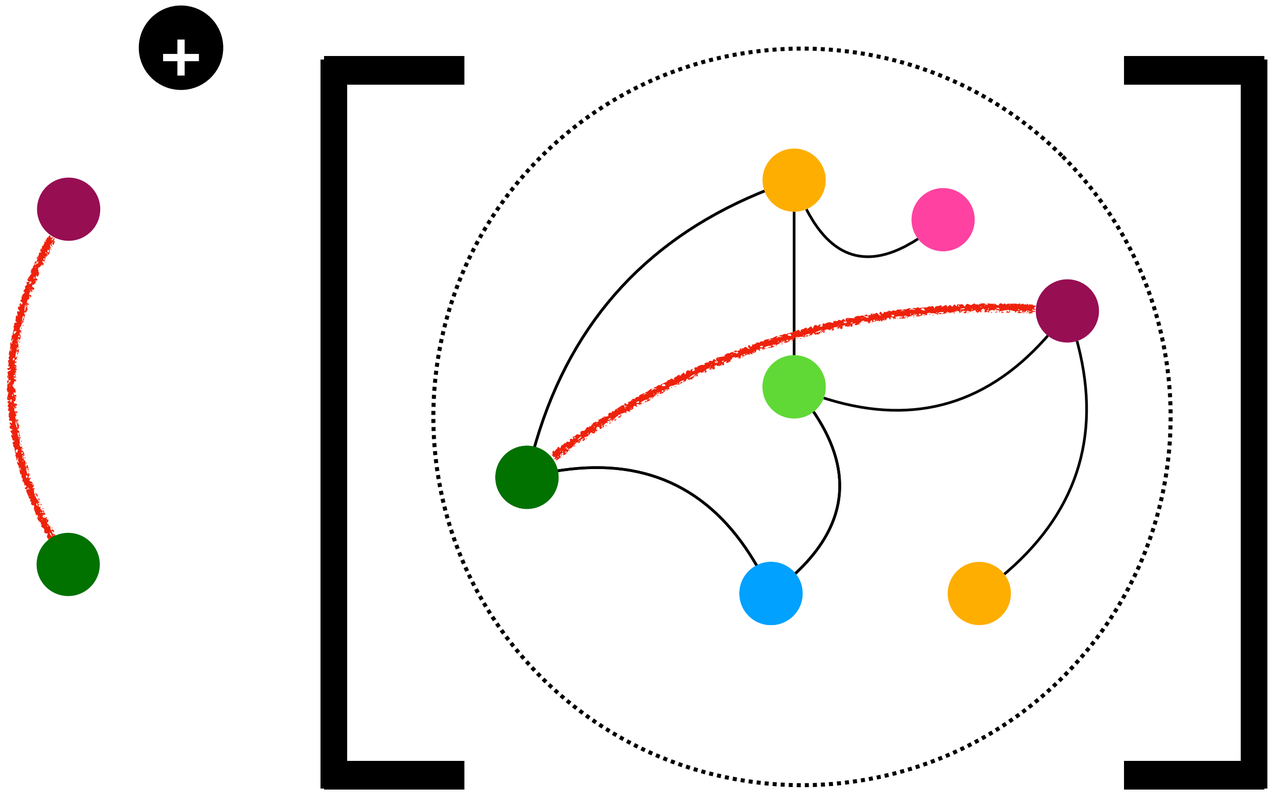
Dynamic Graphs



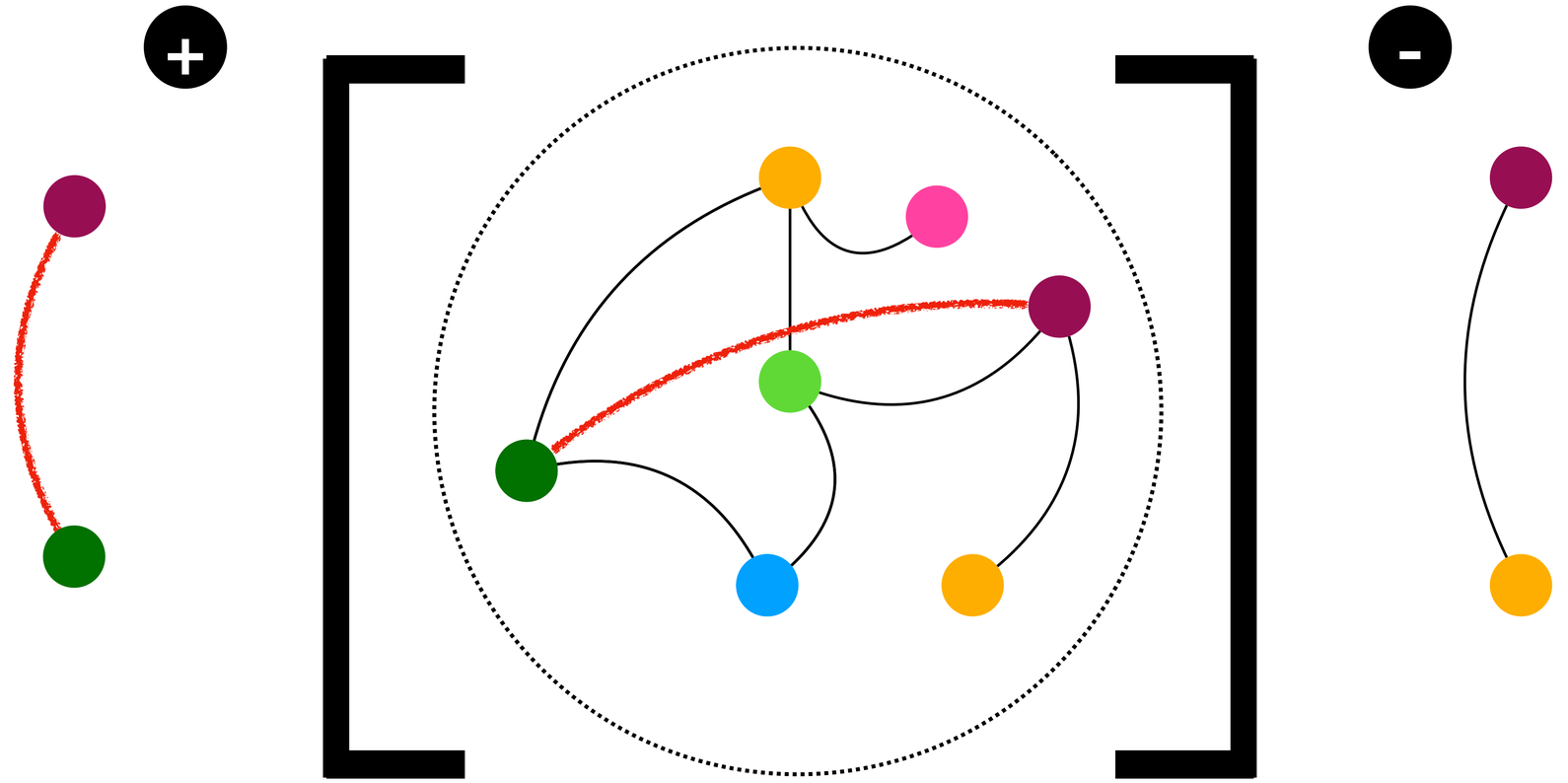
Dynamic Graphs



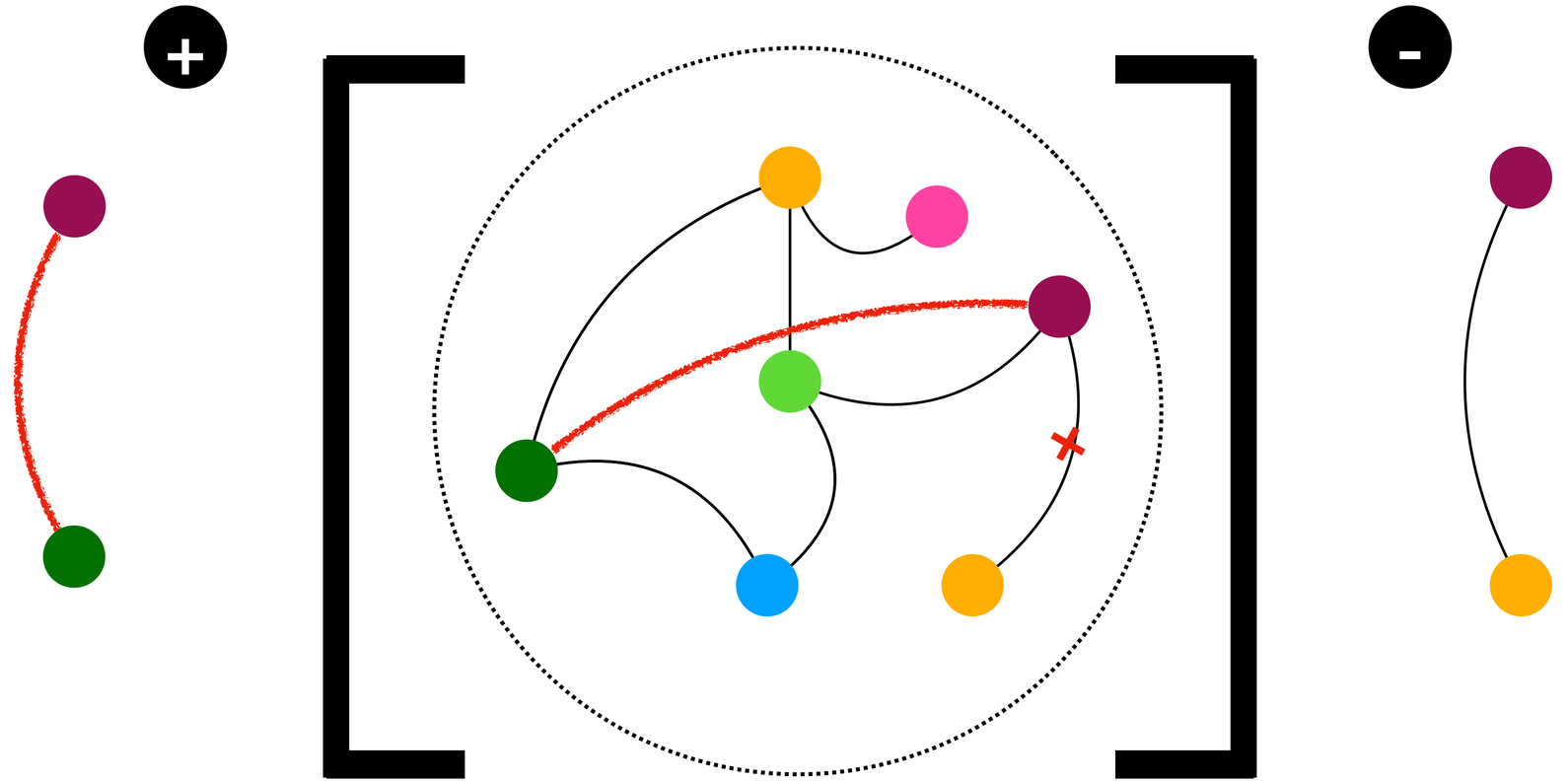
Dynamic Graphs



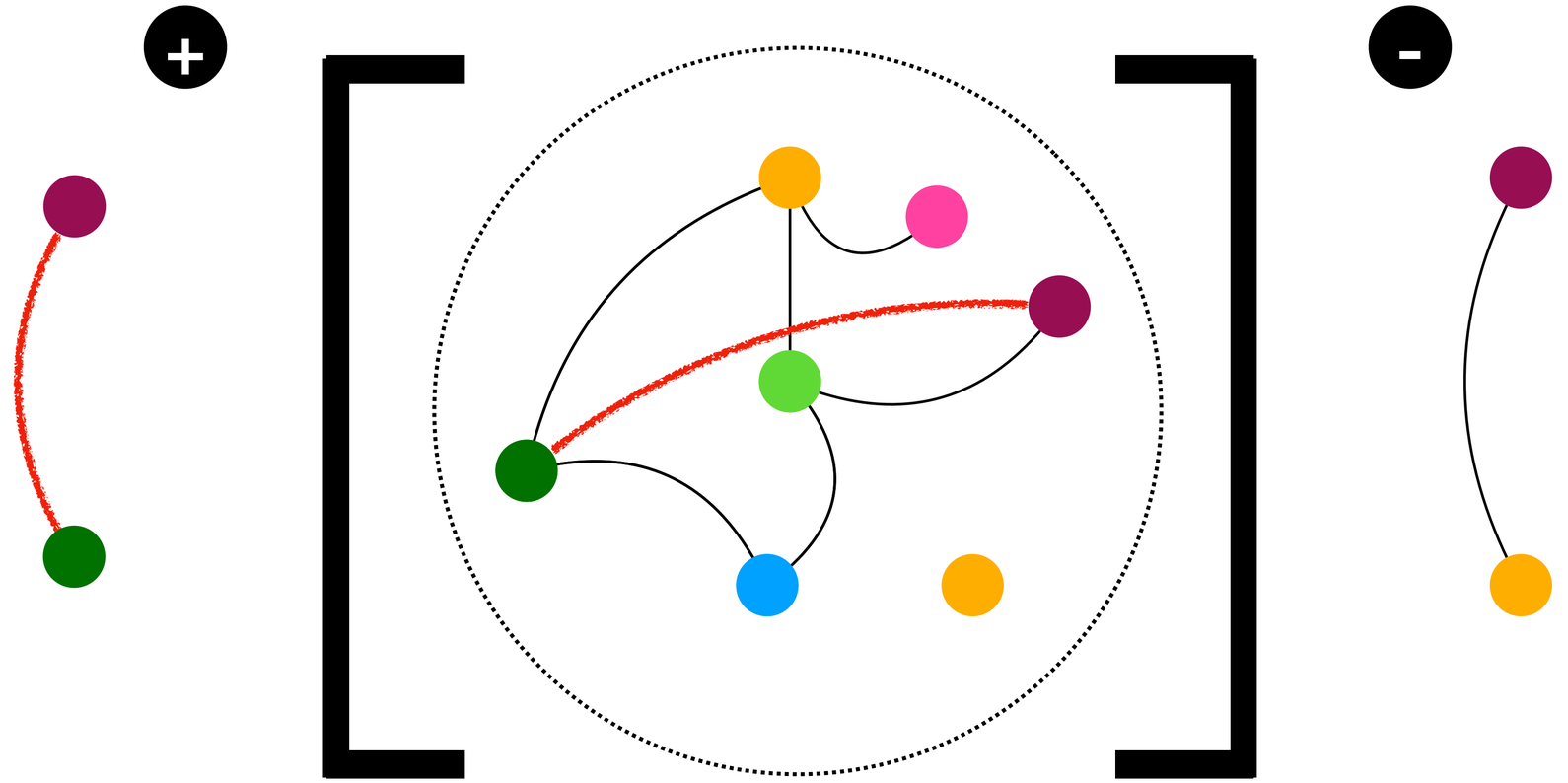
Dynamic Graphs



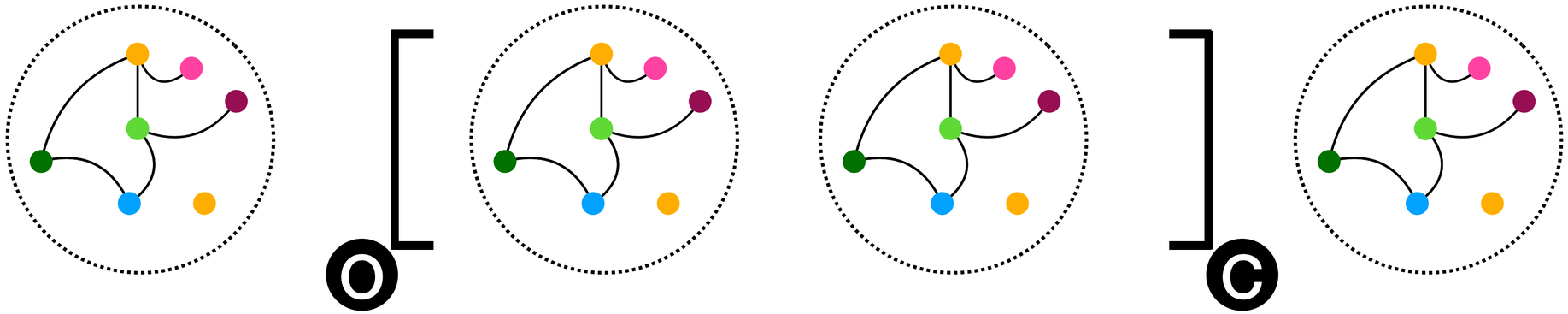
Dynamic Graphs



Dynamic Graphs



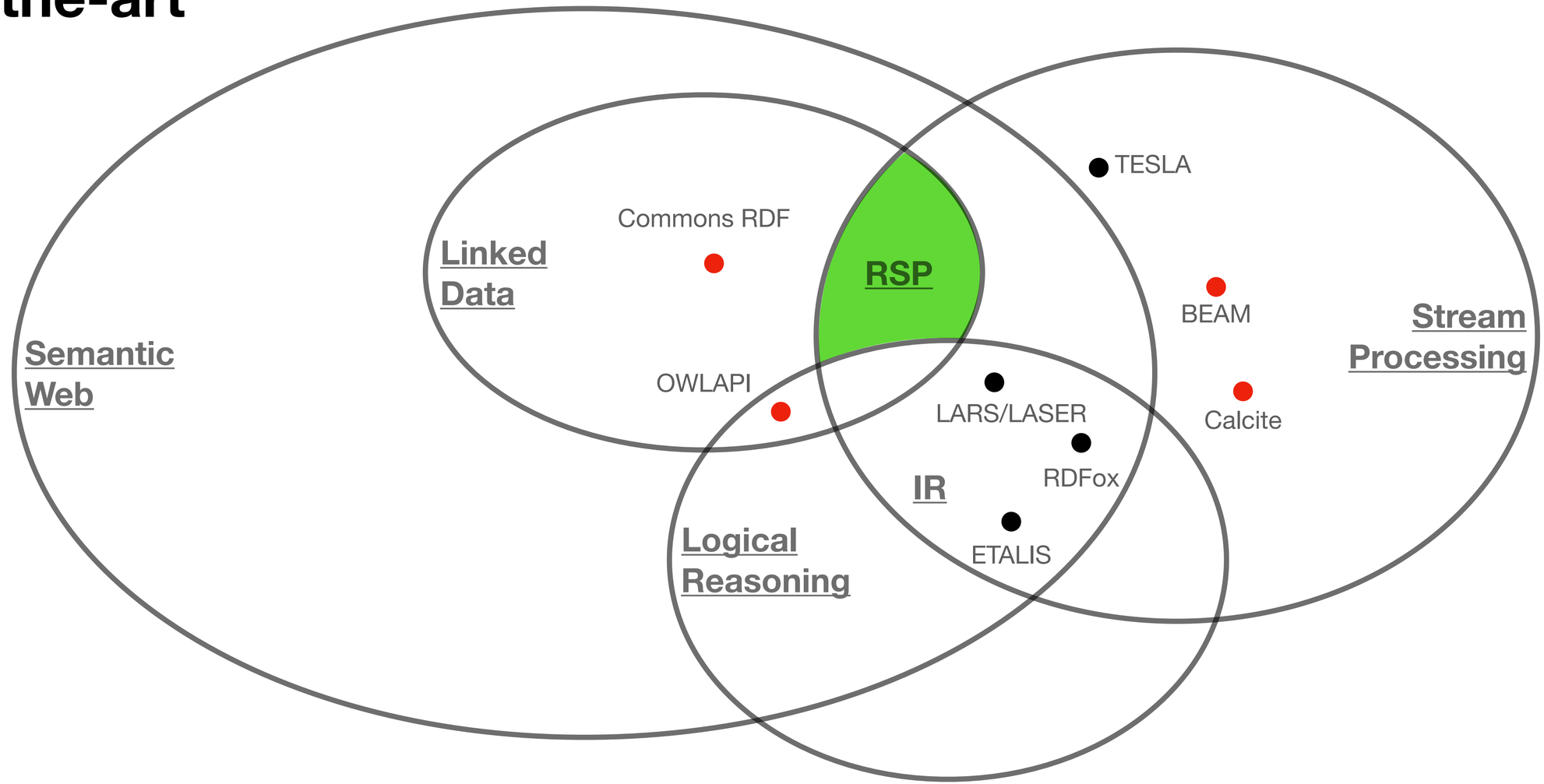
Graph Streams



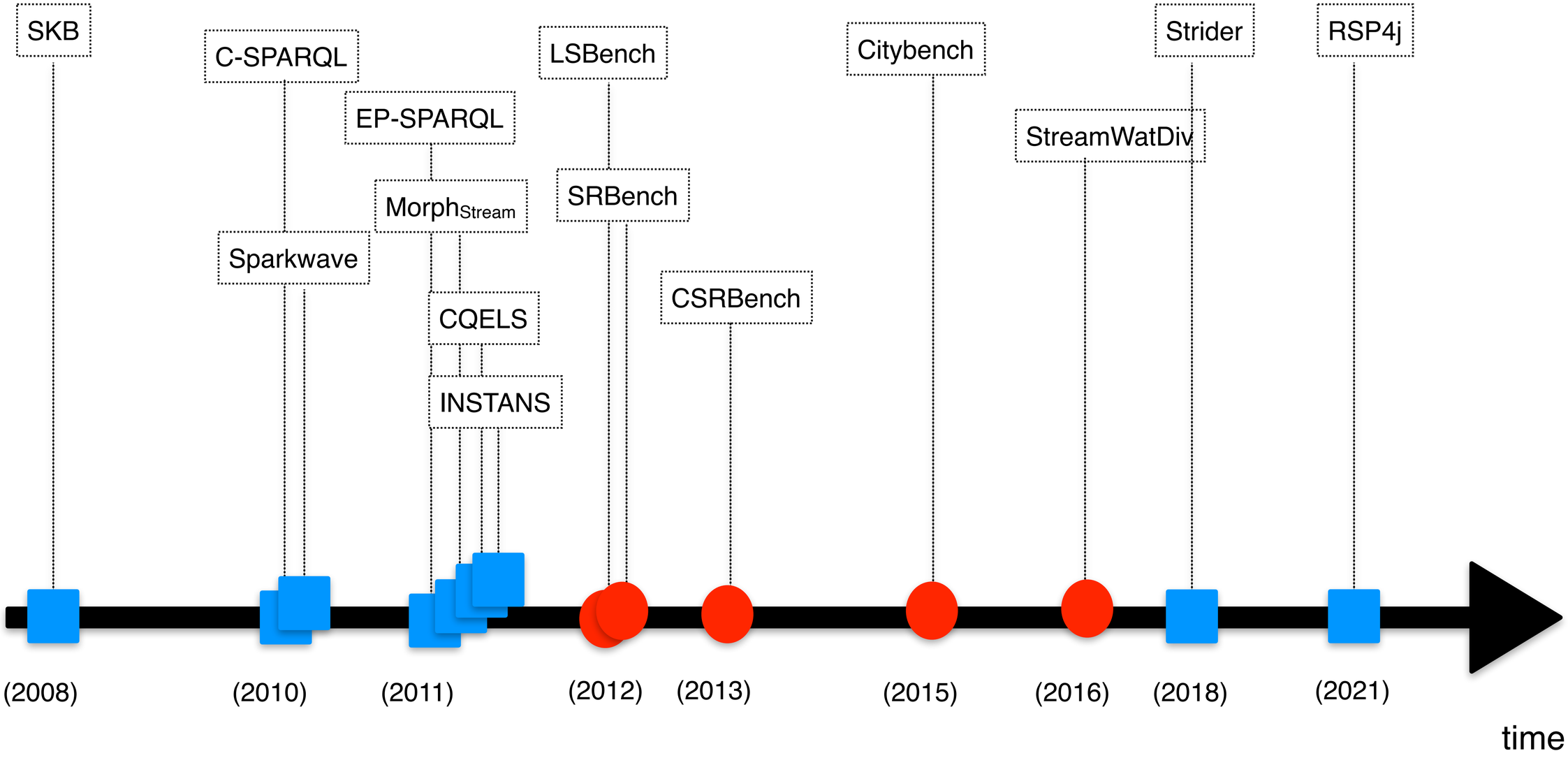
RDF Streams

RDF Stream Processing

State-of-the-art



RSP State-of-the-art: Timeline



LSBench
SRBench
CityBench
StreamWatDiv
DEBS Gran Challenge

LSBench

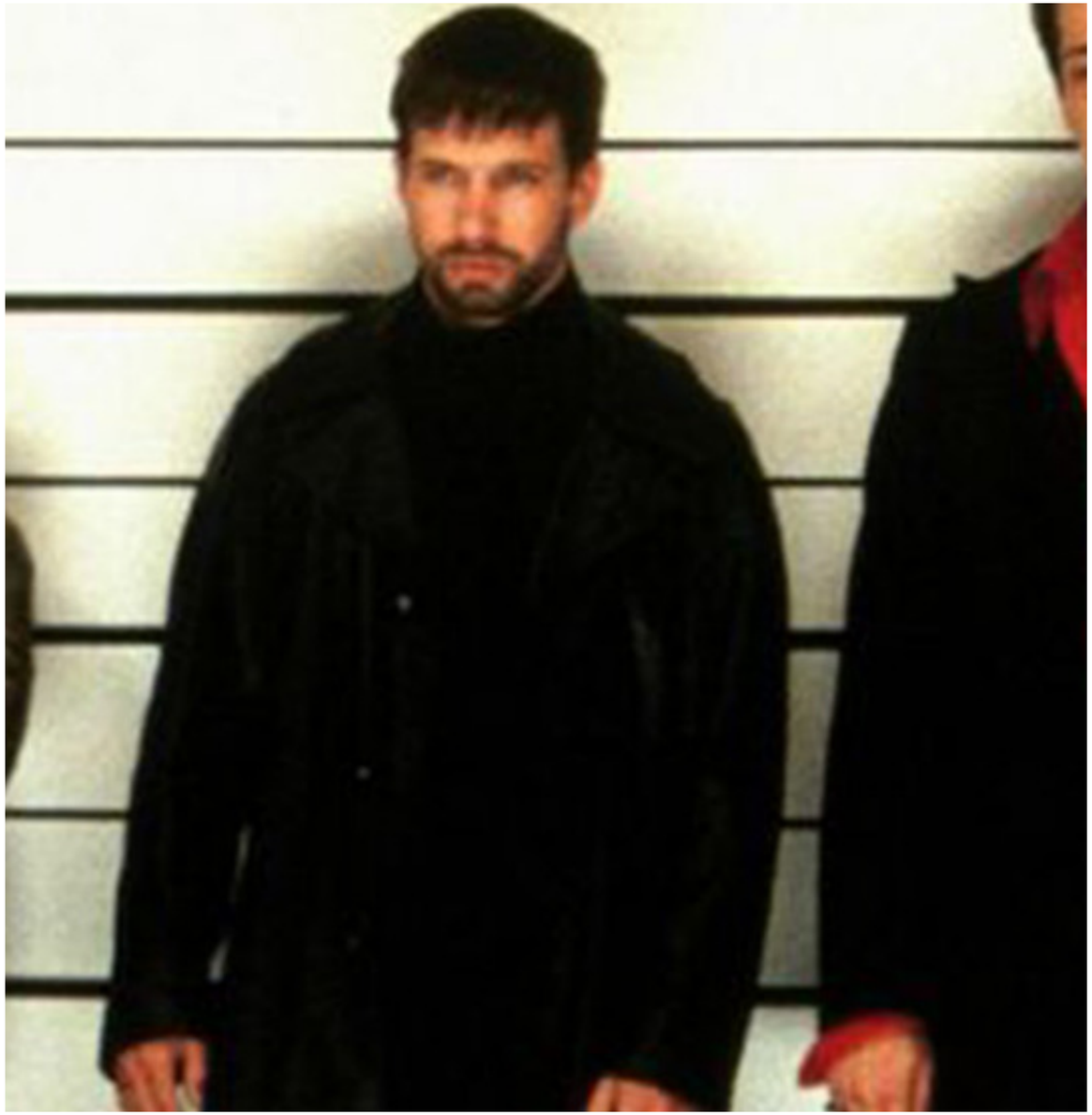
Le Phuoc et al.

- Simple: Social Media (Synthetic)
- Relevant: Static + Streaming
- ~~Portable: not for the code~~
- Scalable: Generator
- KPI: Max Throughput & Memory

SRBench

Y Zhang et al.

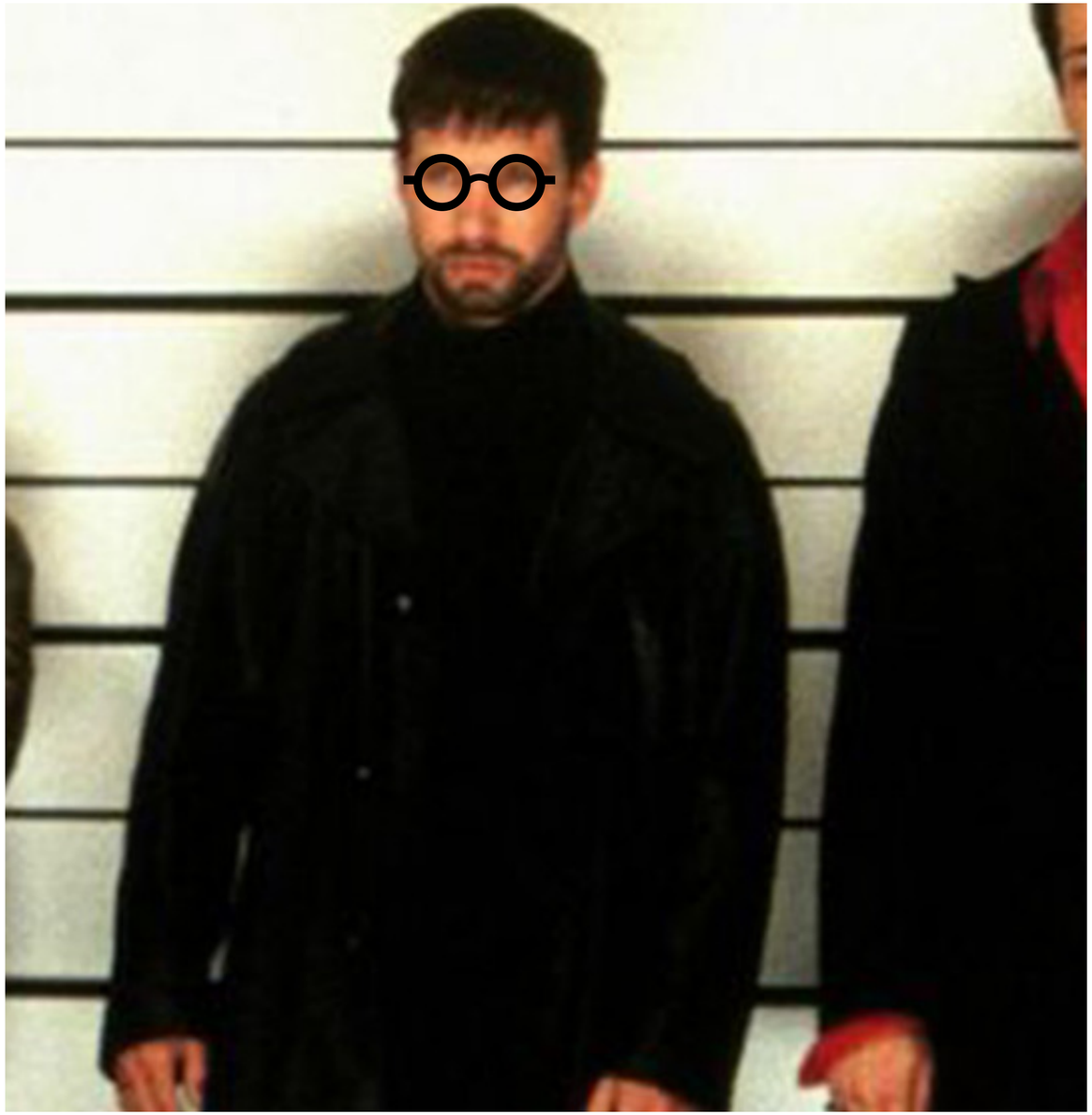
- Simple: Weather/Sensor
- Relevant: Real Data
- Portable: data only + query
- ~~Scalable~~
- KPI: Query Coverage



(C)SRBench

Daniele Dell'Aglio et al.

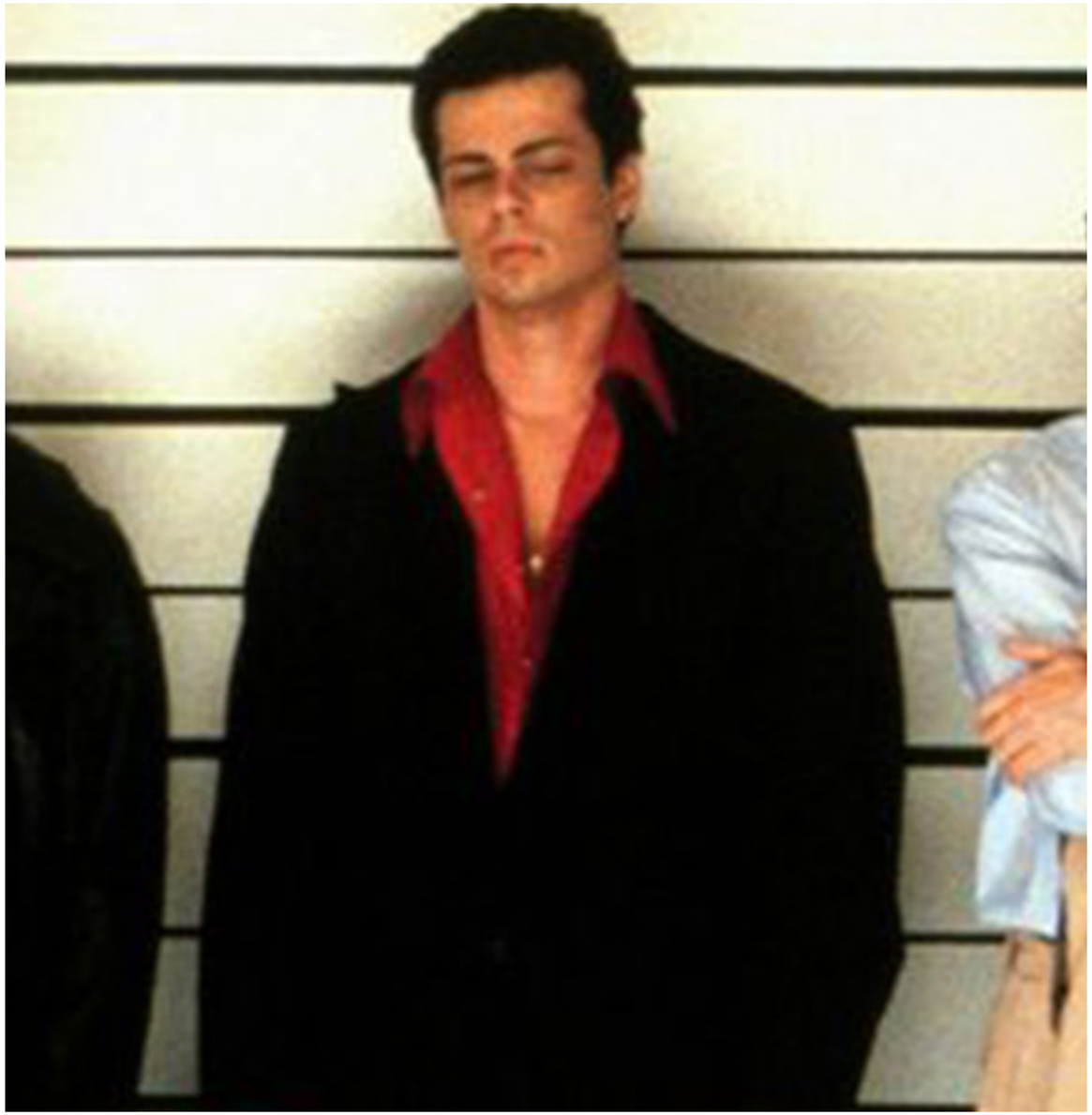
- Simple: Weather/Sensor
- Relevant: Real Data
- Portable: data only + query
- ~~Scalable (a bit): parametric query~~
- KPI: Query Coverage & Correctness



Citybench

MI Ali et al.

- Simple: Smart-City/sensor
- Relevant: Real Data, static and streaming
- ~~Portable: generation is hard-coded~~
- Scalable: query generator



StreamWatDiv

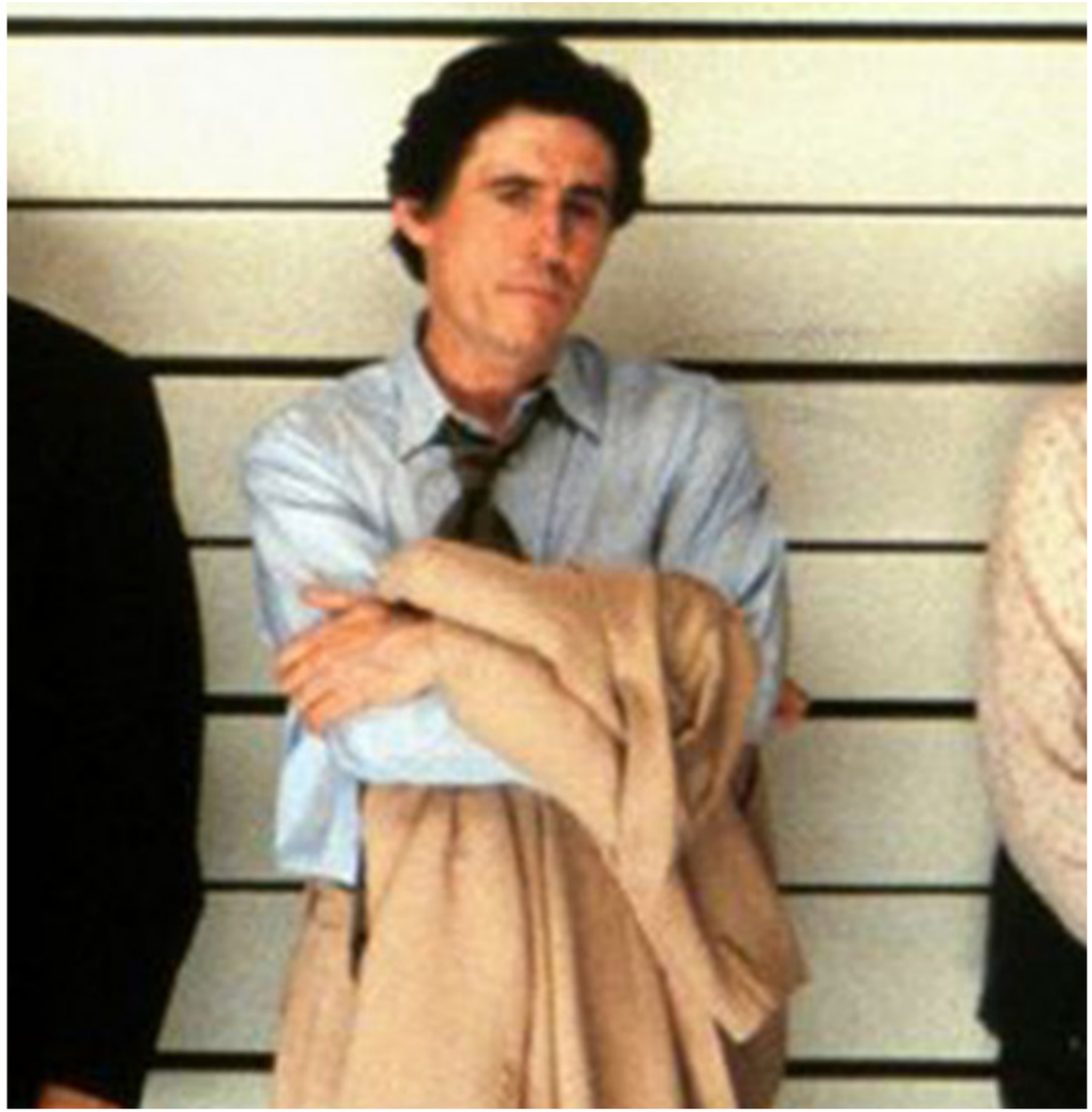
L Gao et al.

- Simple: User Activities
- Relevant: static + streaming
- Portable: test driver
- Scalable: Generator
- KPI: throughput & latency
- https://dsg.uwaterloo.ca/watdiv/stream-watdiv

DEBS Gran Challenge 2017

V Gulisano et al.

- Simple: User Activities
- Relevant: real data
- Portable: RDF
- ~~Scalable~~
- KPI: throughput



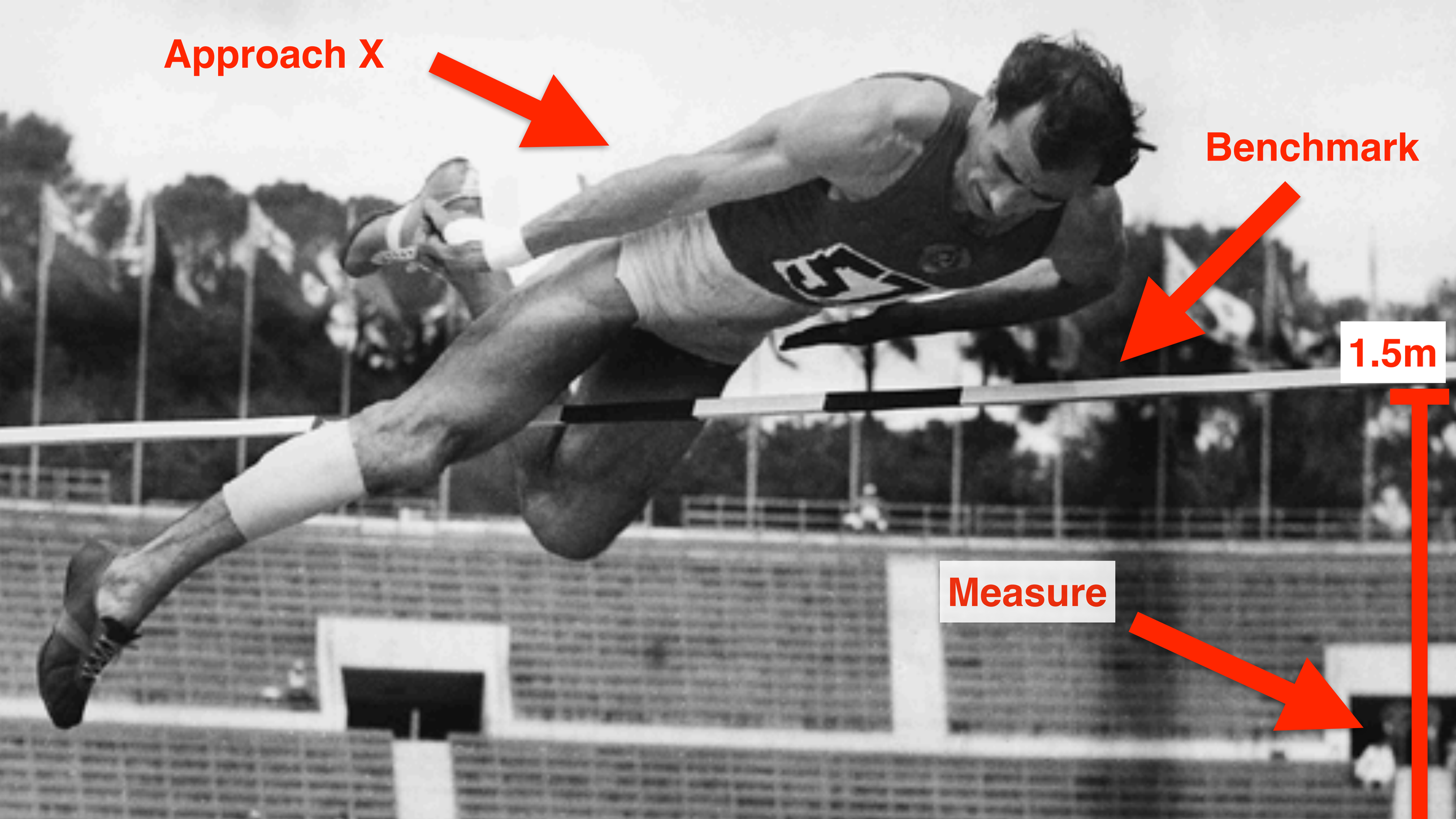
“

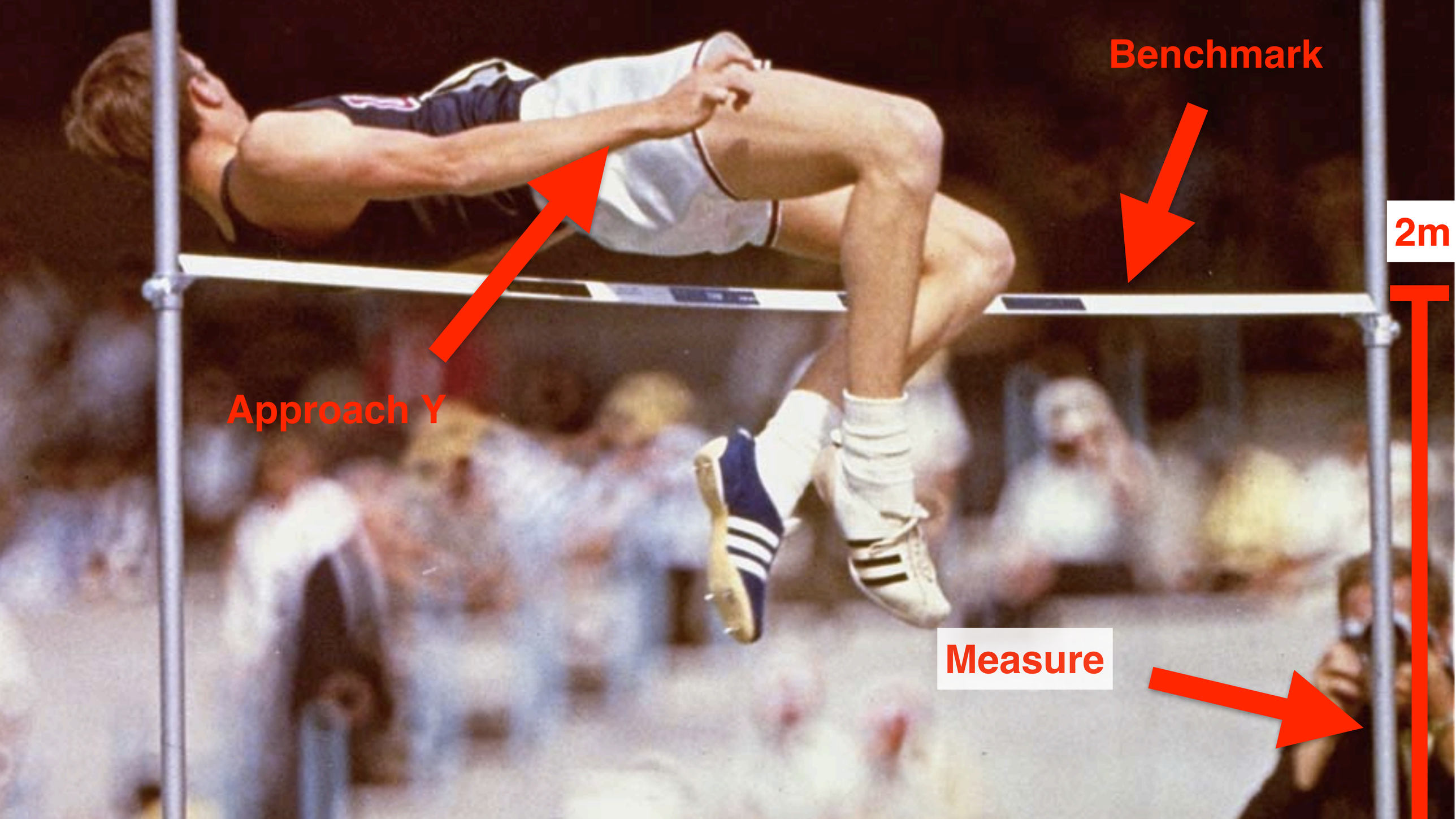
The goal of a domain specific benchmark
is to foster technological progress
by guaranteeing a fair assessment.

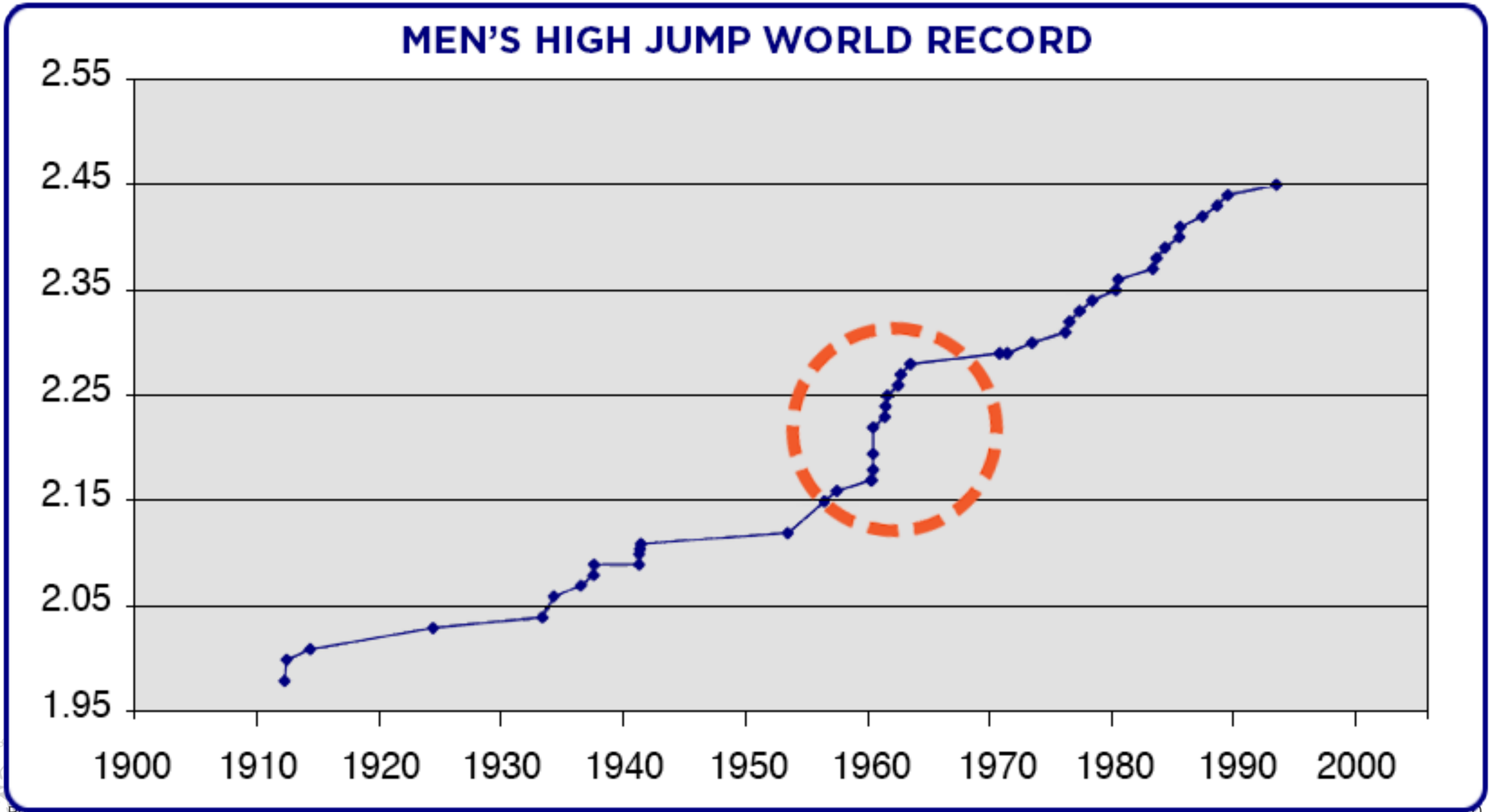
”

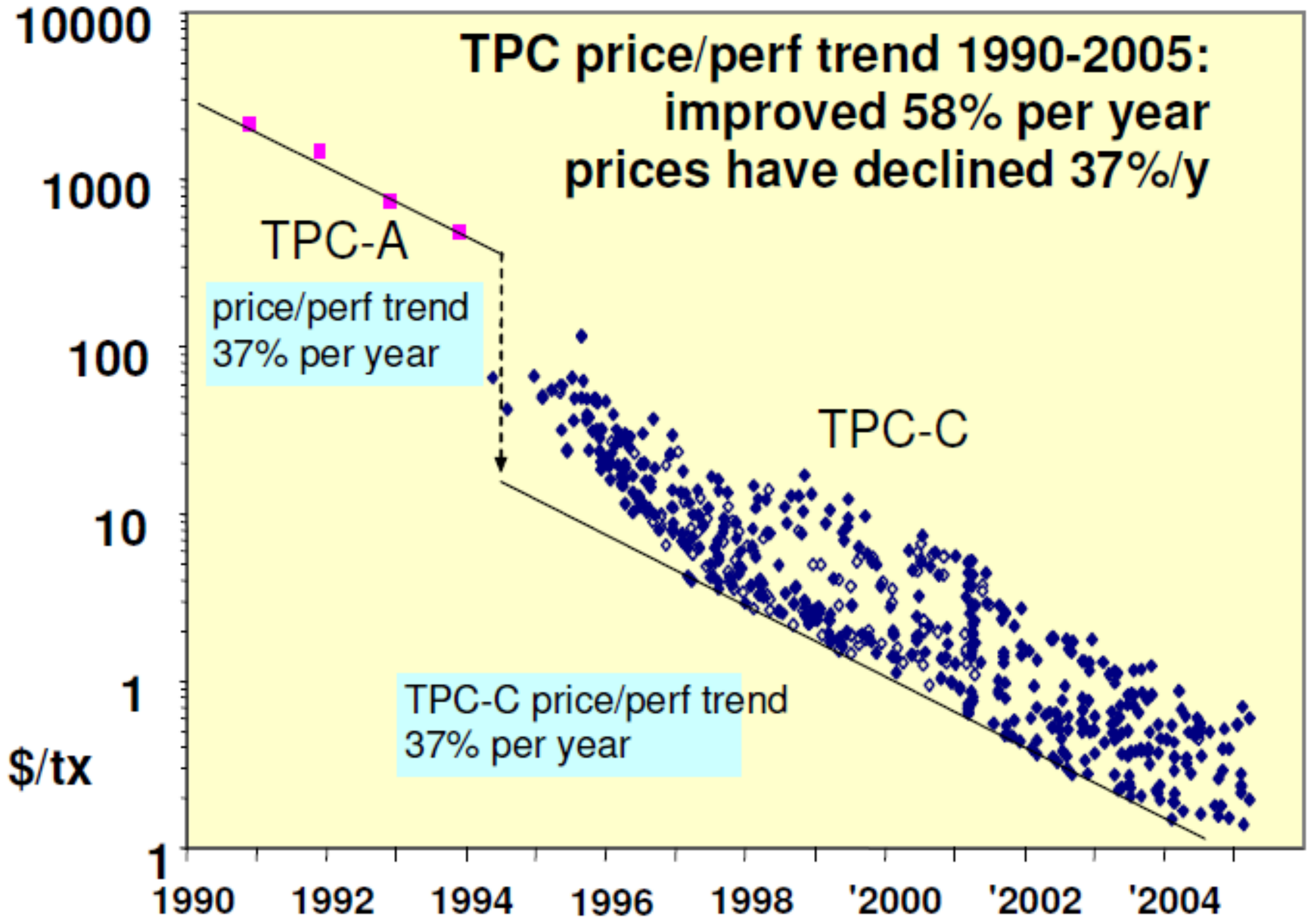
- Jim Gray, The Benchmark Handbook

for Database and Transaction Systems, 1993


Approach X
Benchmark
1.5m
Measure


Benchmark
2m
Approach Y
Measure


MEN'S HIGH JUMP WORLD RECORD
2.55
2.45
2.35
2.25
2.15
2.05
1.95
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000

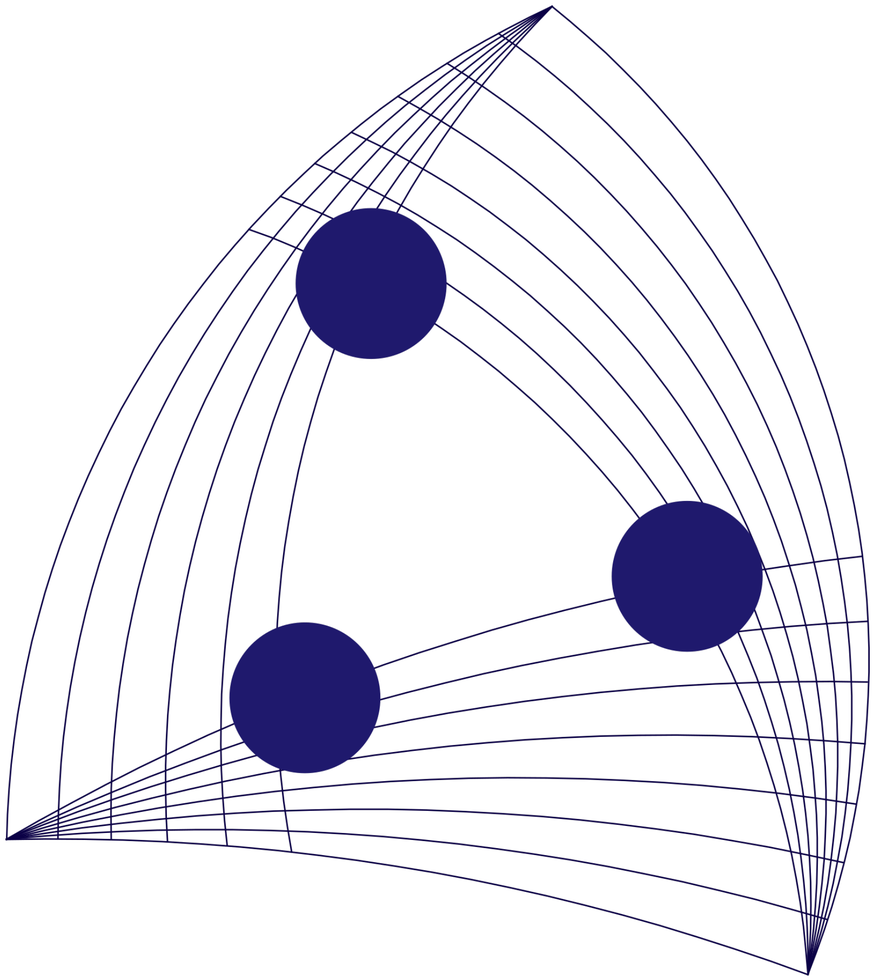


© Jim Gray, 2005

Conclusion

Edges through time

- Fast prototyping of Graph Stream Processing engines
- Potentially a test driver for benchmarks
- ESWC 2021 Best Resource Paper award!



RSPY