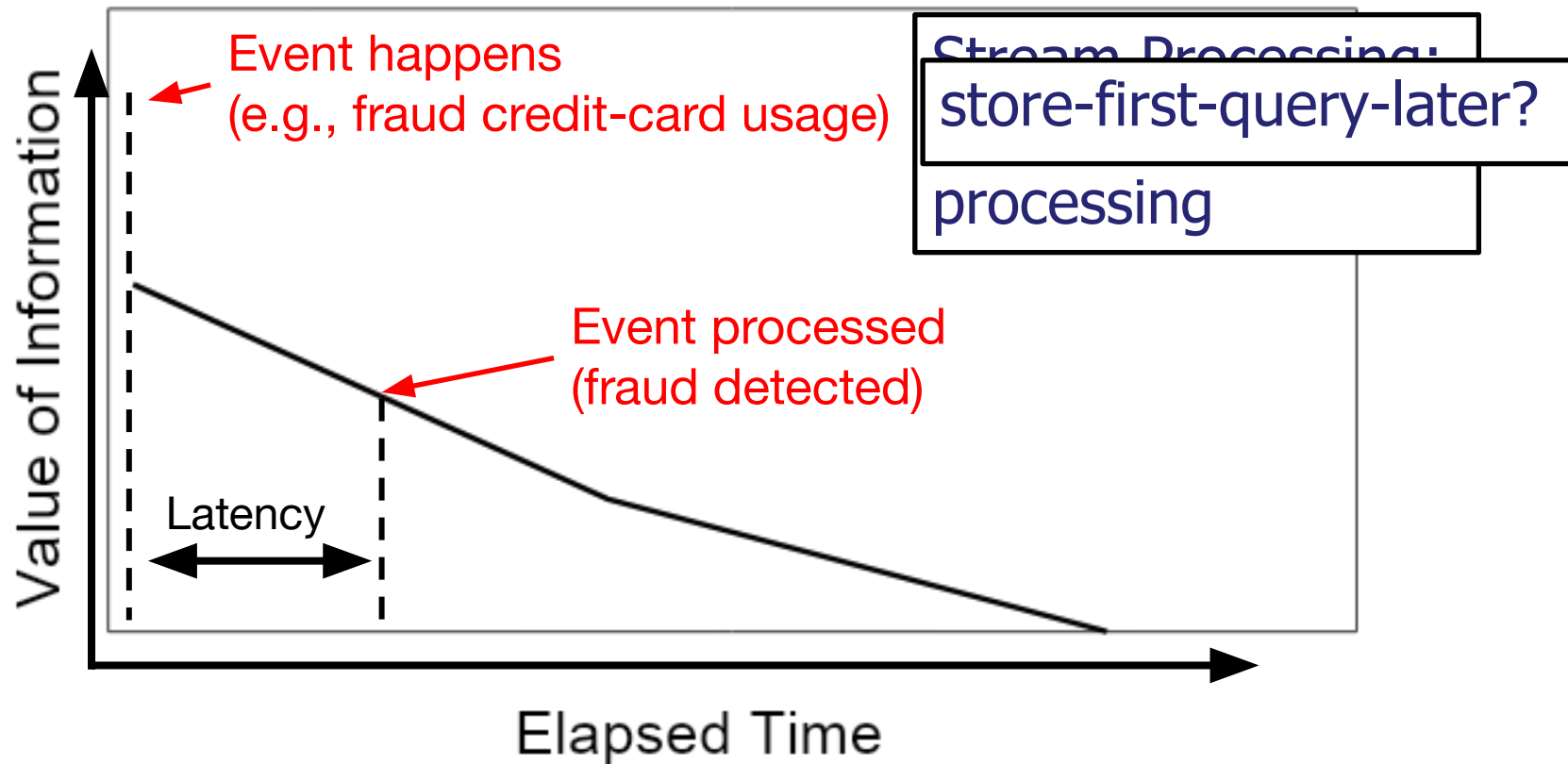# High Performance Stream Processing on Modern Hardware: The Past and The Future

Shuhao Zhang
Postdoc Researcher, DIMA TU Berlin

Outline

- Motivation
  - Why stream processing?
  - Why stream processing on modern HW?
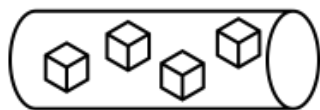- Past Work
- Future Work

流计算(stream processing)

Event happens
(e.g., fraud credit-card usage)

Event processed
(fraud detected)

Value of Information

Latency

Elapsed Time

Stream Processing:
store-first-query-later?

processing

流计算在阿里巴巴的应用

source: <u>Single's Day Global Shopping Festival on 11/11 2019, Alibaba with Apache Flink</u>

Real-time Events — ingest → **Stream Processing Engine (Flink)** — append → **Live Report / Dashboard**
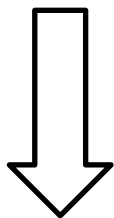
Continuous analytics

- 2 billion events/sec
- 3 TB/sec

- Online training
- Real Time Fraud Detection
- Dynamic Pricing
- …

(sub)second latency

4

流计算的广泛应用

Application
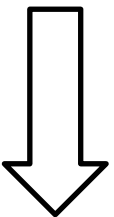
Smart Manufacturing　　Auto-Vehicles　　Digital Banks

Network

large volume of data streams

System
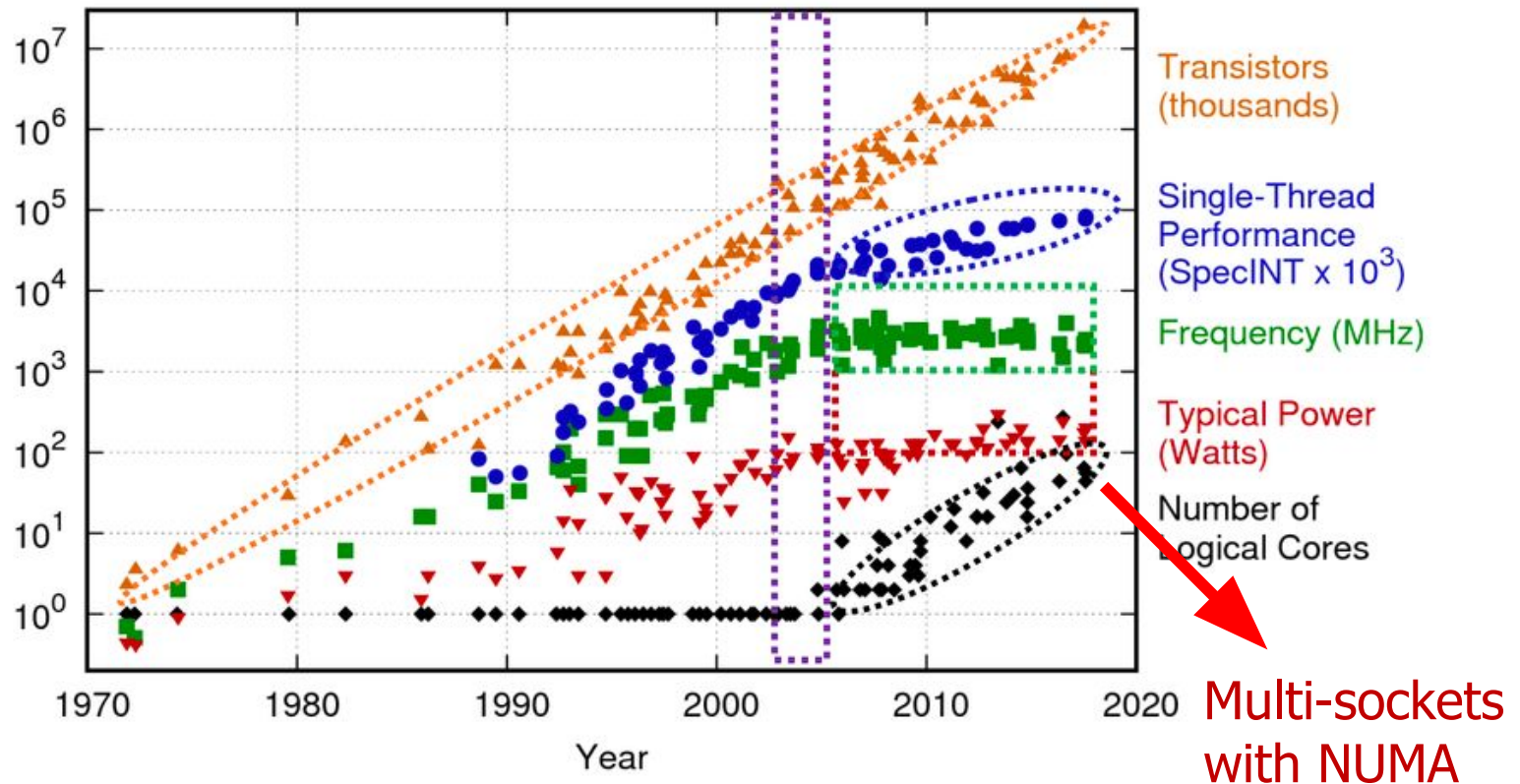
Flink　　STORM　　Spark Streaming　　…

Outline

- **Motivation**
  - Why stream processing?
  - **Why stream processing on modern HW?**
- Past Work
- Future Work

The End of Moore's Law

Source: BACKUS: Comprehensive High-Performance Research Software
Engineering Approach for Simulations in Supercomputing Systems, 2019
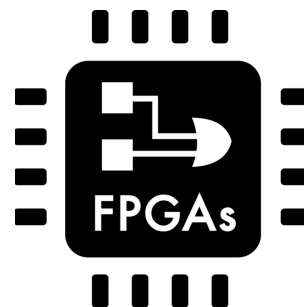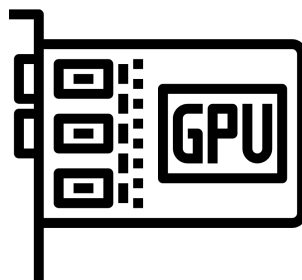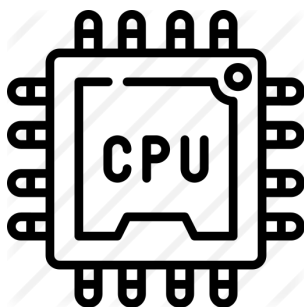
在新硬件上部署流计算

Network

large volume of data streams

my research focus

System

Hardware

研发目标

Trend – –

(i) Stream processing is becoming more critical for big data applications;

(ii) Modern architectures are developing rapidly (CPUs, GPUs)
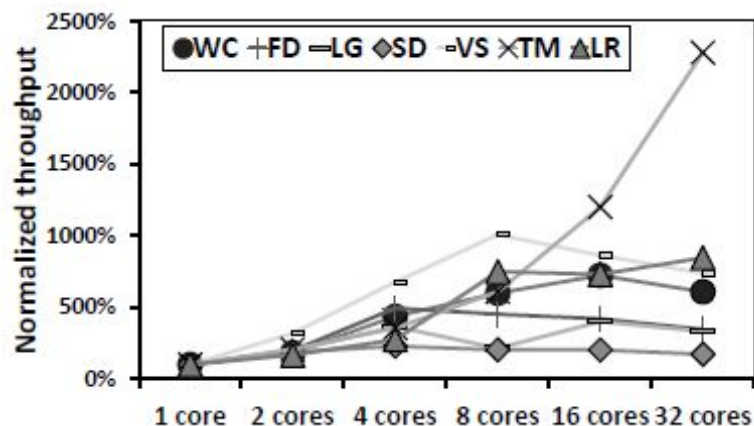
Vision – –

We need a high-performance stream processing engine (SPE) on modern architectures.
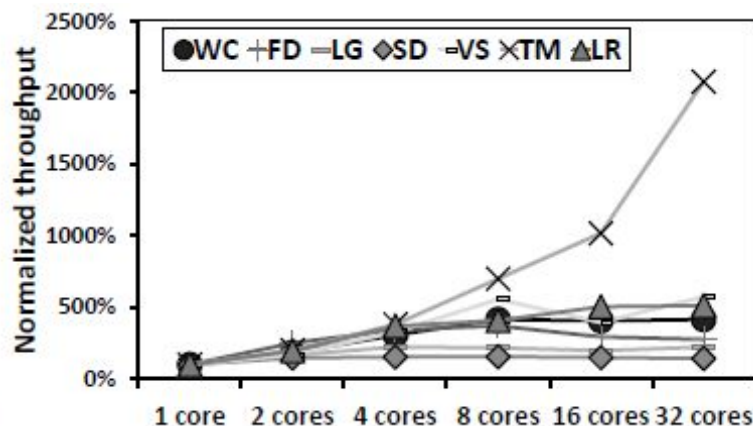
Outline

- Motivation
  - Why stream processing?
  - Why stream processing on modern HW?
- Past Work
  - What is the current limitation?
- Future Work

观察现象

Apache Storm

Apache Flink

6/7 apps perform worse with more CPU cores on a 32-core machine

*Source: Shuhao Zhang, et al., "Revisiting the Design of Data Stream Processing Systems on Multi-Core Processors," *2017 ICDE*

深入理解

■ Computation
Useful time

▨ Front-end stalls
Instruction fetch
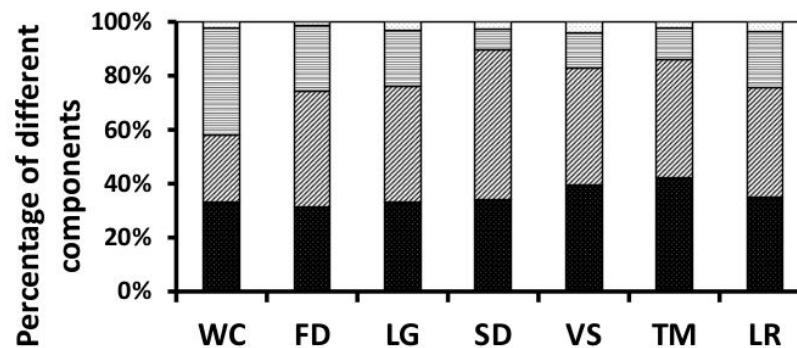
☰ Back-end stalls
Data fetch

▦ Bad speculation
Branch misprediction
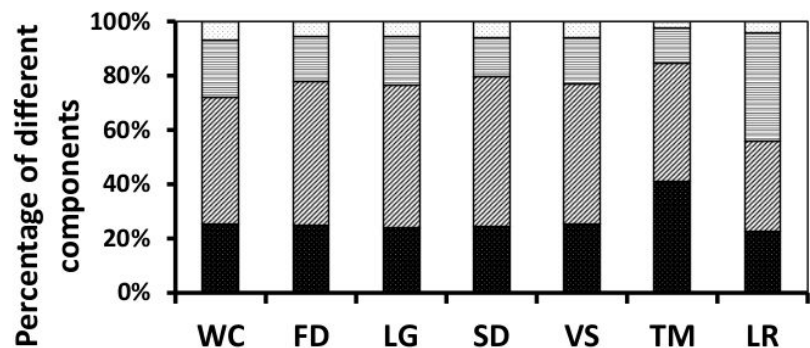
(a) Storm

(b) Flink

Processor stalls

70% of the execution times are wasted

解释现象

L1-ICache size: 32 KB; L2-Cache size: 256KB; L3 Cache size: 20MB
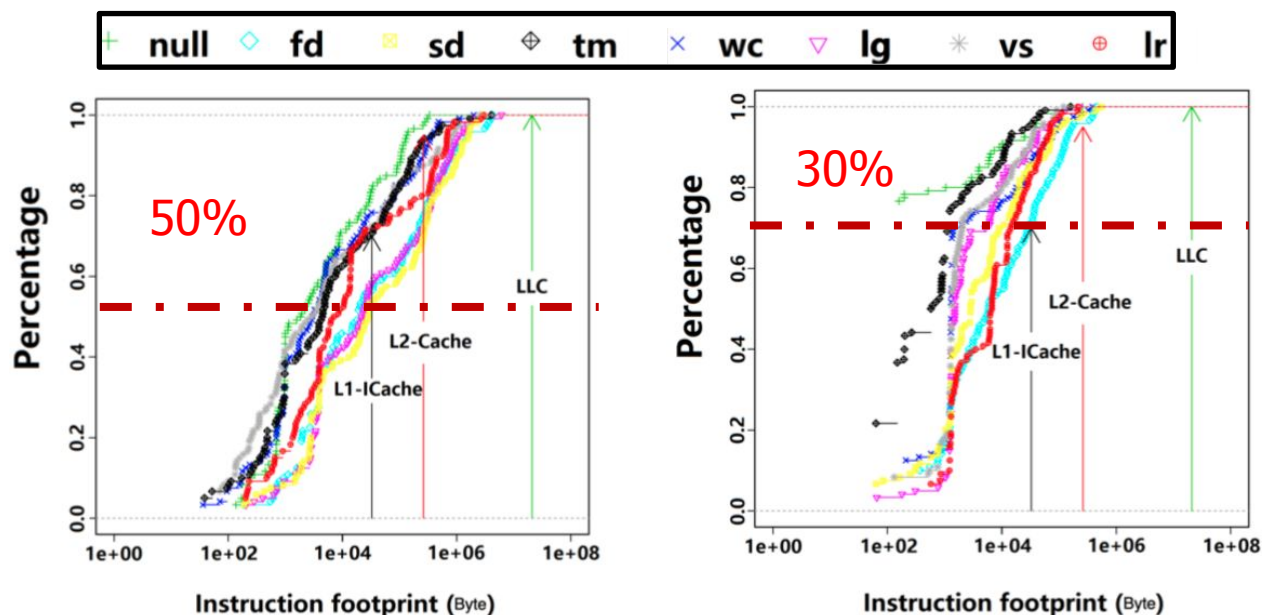
(a) Storm　　　　　(b) Flink

50% instruction footprints ≥ instruction cache

Cache trashing

Poor performance

Point (x,y): y% of instruction footprints ≤ x

在新硬件上实现高性能流计算

users

Stream Analytics	Transaction Stream	. . .

(Multi-)Query Optimization	Query Deployment	Transactional State Management

Execution Optimization

Modern Hardware (CPU, GPU...)

stream processing engines
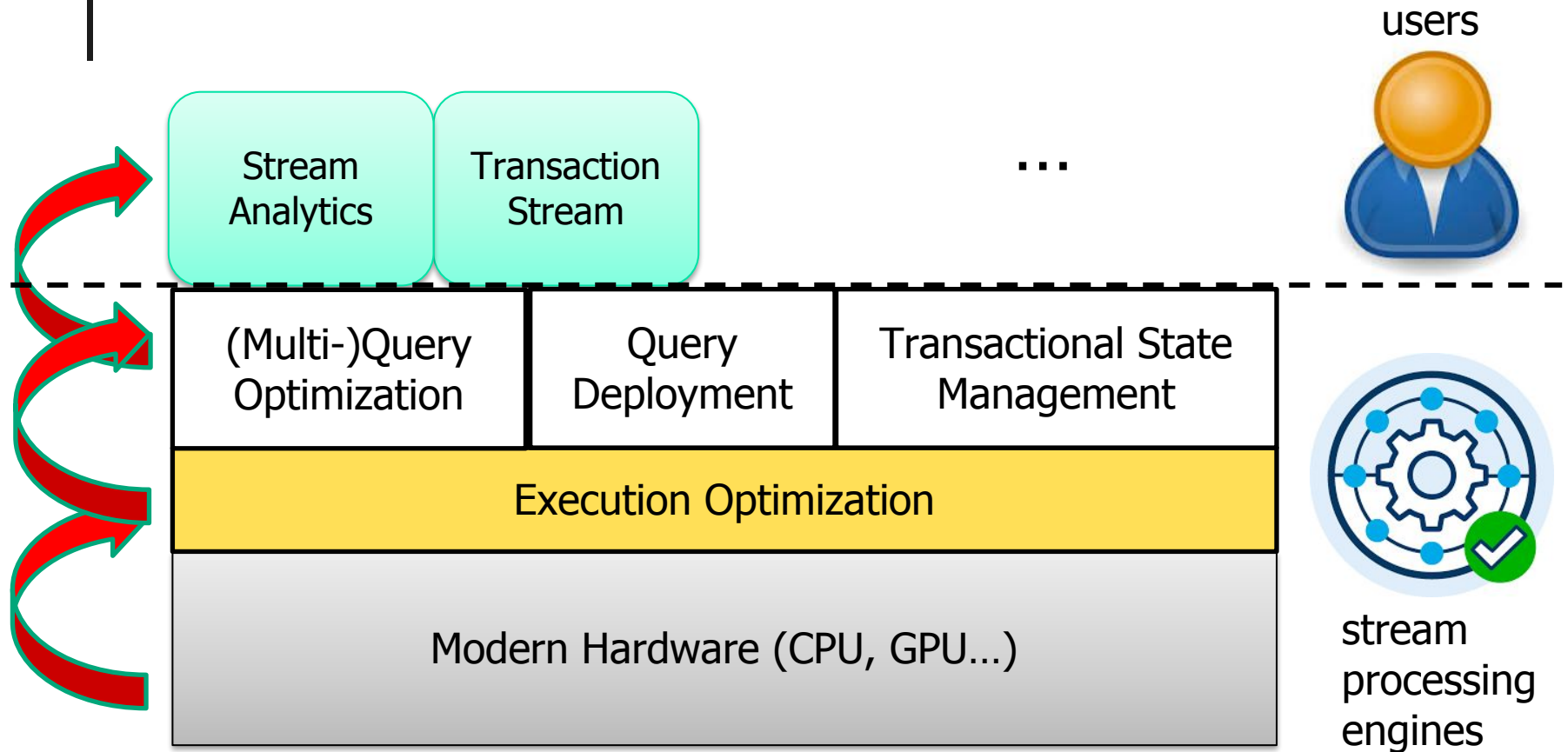
See our ICDE 17 x 2, SIGMOD 19, ICDE 20, ATC 20...

Outline

- Motivation
  - Why stream processing?
- Past Work
  - What is the current limitation?
- **Future Work**

Stream Processing Market Is Fast Growing

Source: "Event Stream Processing Market Research Report – Global Forecast till 2023" (May. 2019)

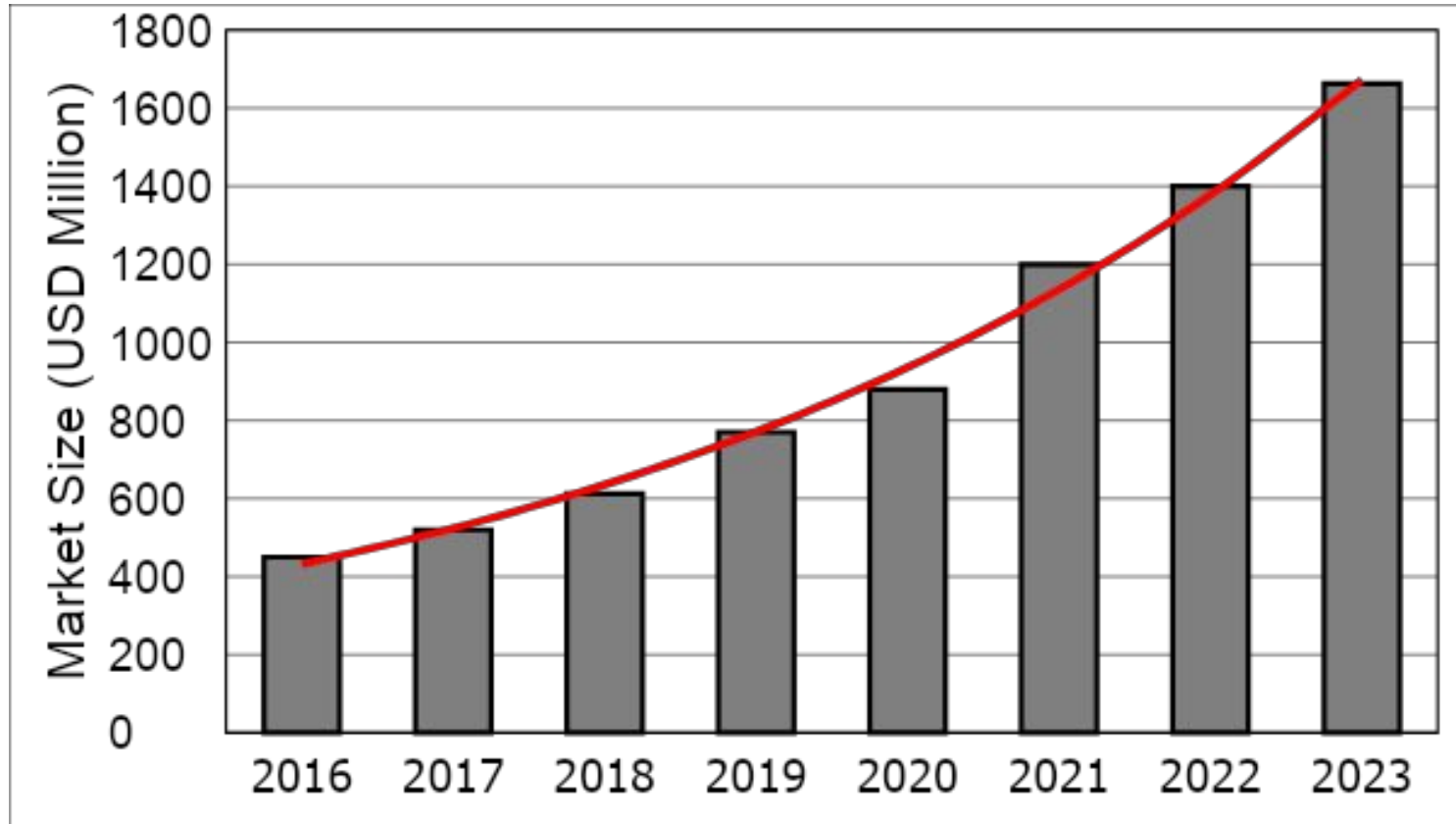# The Trend of Stream Processing (1)：应用复杂化

- Simple Map-Reduce style stream processing is getting outdated.
- Complex stream processing is getting widely deployed.
  - 实时的机器学习 i.e., Online Learning
  - 流数据挖掘 (e.g., Twitter@KDD19)
  - Live Audio Streaming Services (e.g., Prof.He)
  - Transactional Stream Processing (e.g., S-Store sigmod'15, TStream icde'20)
 - Large Scale Stream Graph Processing (e.g., WuKong+S sosp'17)

The Trend of Stream Processing (2)：需求差异化

- Not just performance.
- But also…
  - Batch + Stream Processing (e.g., Neptune socc'19)
  - Privacy & Security (TimeCrypt, NSDI'20)
  - Unified Computing (e.g., DeltaLake vldb'20)
  - Reproducibility/Deterministic
  - Serverless/Stateful Function
  - …

The Trend of Stream Processing (3)：硬件多元化

- Specialized stream processing engine is getting more and more important.
  - HW accelerator (e.g., GPU, FPGA, APU etc.)
  - IoT环境 (e.g., NebulaStream)
  - 安全计算硬件 (e.g., StreamBox-TZ)
  - 持久内存 (e.g., Stream Processing with state storing in NVRAM)
  - …

机遇与挑战共存

应用复杂化

系统优化

需求差异化

硬件多元化
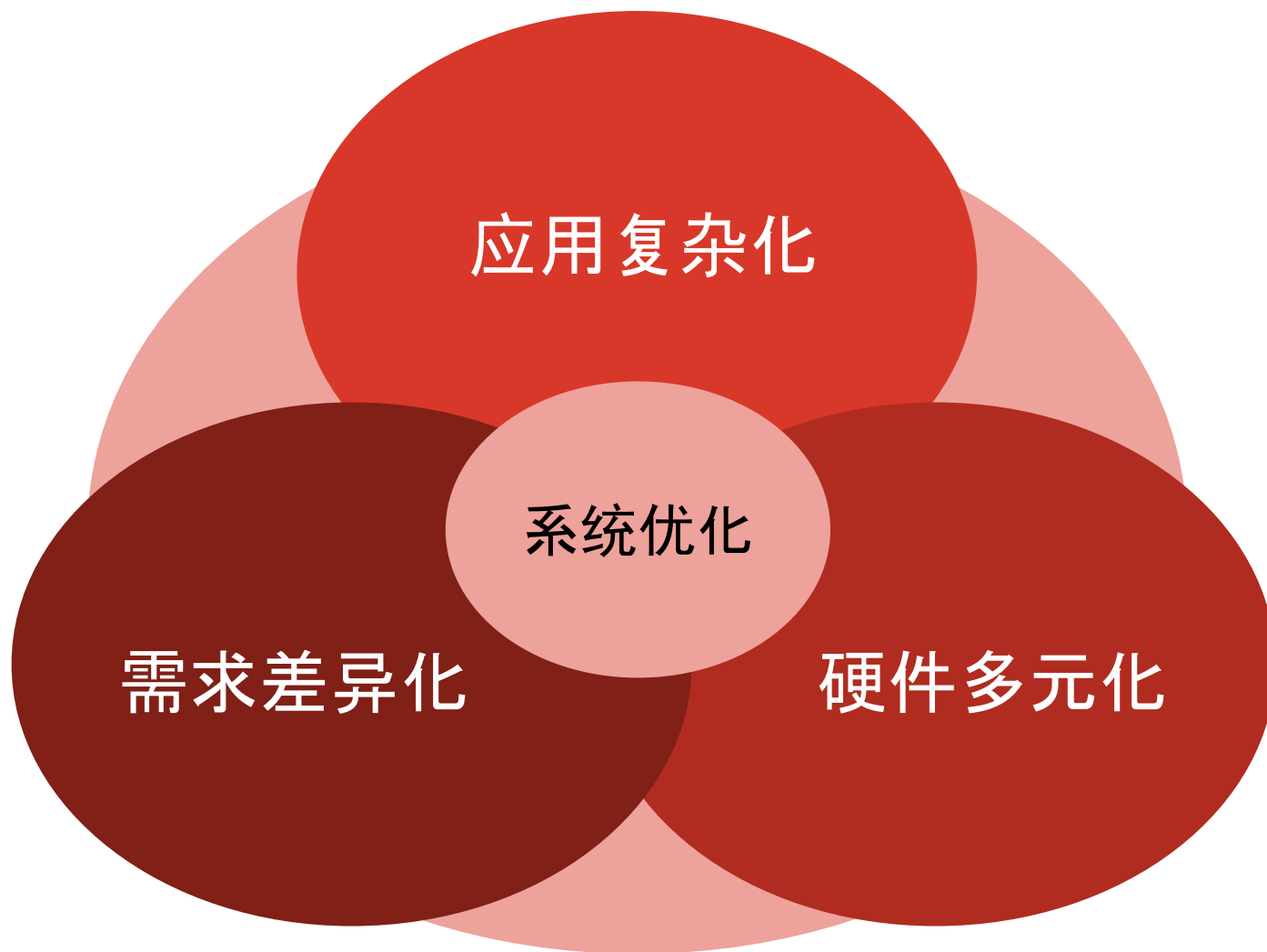
Publication Summary

- Patens

  First authored patents registered in the USA x2

- Journal

  IEEE TPDS 16 x2, ACM SIGMOD Rec. 19 ...

- Conference (only top-tier is shown)

  VLDB 13, SC 16, ICDE 17 x2, SIGMOD 19, ICDE 20, IJCAI 20, USENIX ATC 20, SIGMOD 21 (in revision) ...

- System demo and Workshop

  VLDB 13, IEEE BIGMM 19, VLIoT 20

- Code distribution
  - OmniDB
  - BriskStream
  - TStream

21

My Profile Page

Shuhao Zhang

Assistant Professor

Website: https://shuhaozhangtony.github.io/
Pillar / Cluster: Information Systems Technology and Design
Research Areas:: Secure Computing, Artificial and Augmented Intelligence, Data Science, Networked and Autonomous Systems

I'm looking for phd students.

| BIOGRAPHY | EDUCATION AND WORK EXPERIENCE | PROFESSIONAL ACTIVITIES | SELECTED KEY PUBLICATIONS |

Biography

Shuhao Zhang will be joining ISTD Pillar of SUTD as a faculty in early 2021. He is currently a **postdoctoral researcher** a Technische Universität Berlin working with Prof.Volker Markl on the topic of data stream processing and management. He obtained his Ph.D. degree from the school of computing at National University of Singapore in 2019, under the supervision of Prof. Bingsheng He. He had previously worked as Research Scientist/Assistant with Prof. Bingsheng He,

https://istd.sutd.edu.sg/people/faculty/shuhao-zhang

22