www.qconferences.com
www.qconbeijing.com

QCon

QCon北京2014大会 4月25—27日

伦敦 | 北京 | 东京 | 纽约 | 圣保罗 | 上海 | 旧金山

London · Beijing · Tokyo · New York · Sao Paulo · Shanghai · San Francisco

QCon全球软件开发大会

International Software Development Conference

www.infoq.com/cn

InfoQ

@InfoQ

infoqchina

软件
正在改变世界！

Spark: High-Speed Big Data Analysis Framework

Andrew Xia

Weibo: @Andrew-Xia

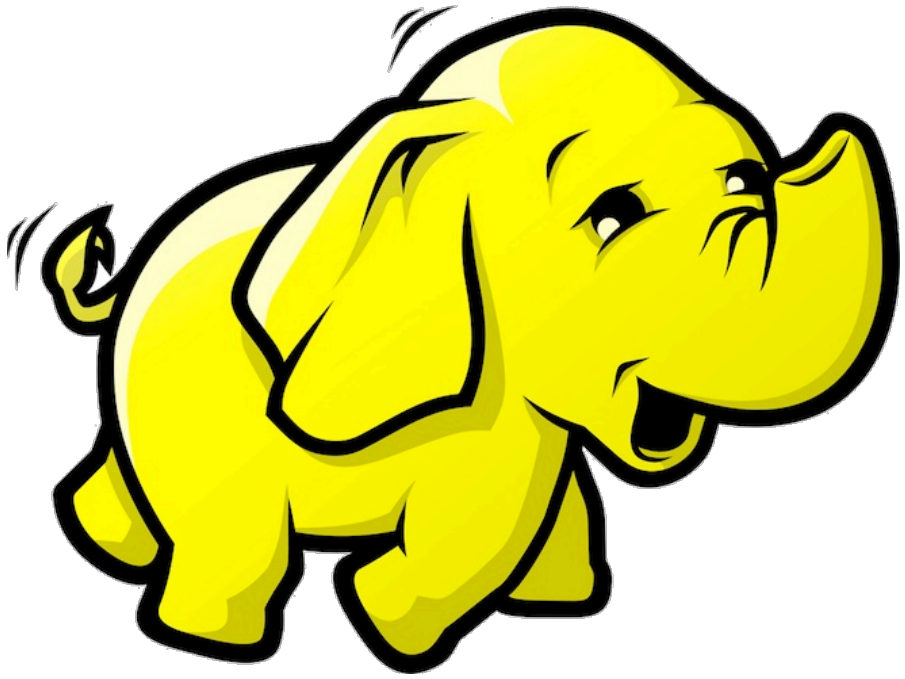
Intel

Agenda

- ***Why need Spark?***
- How to program with spark?
- What is Spark ecosystem?
- Community

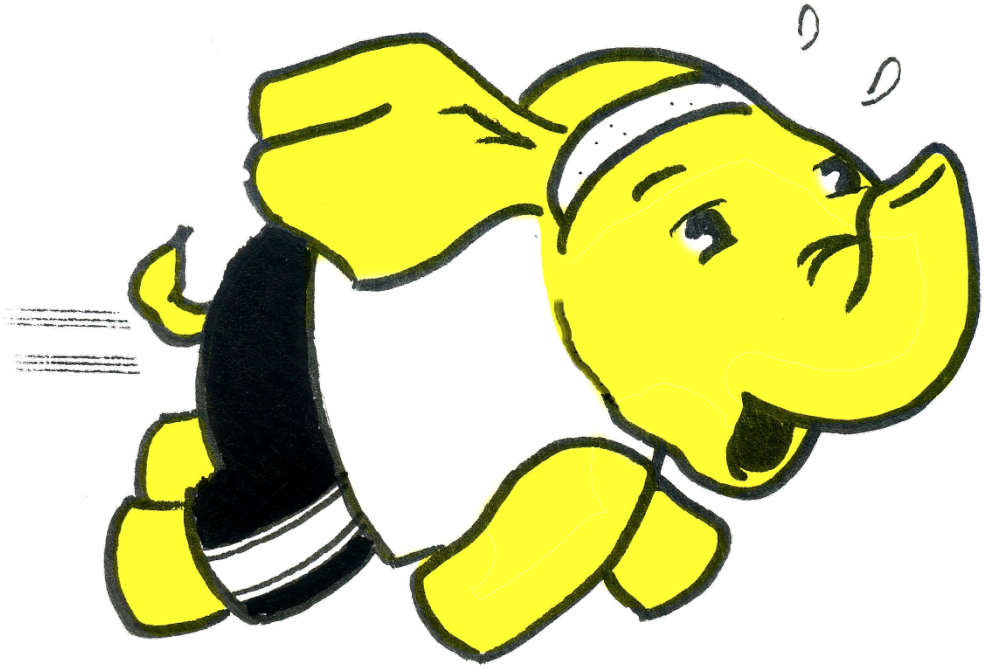
Hadoop

- Simplified big data analysis by giving a reliable programming model for large clusters

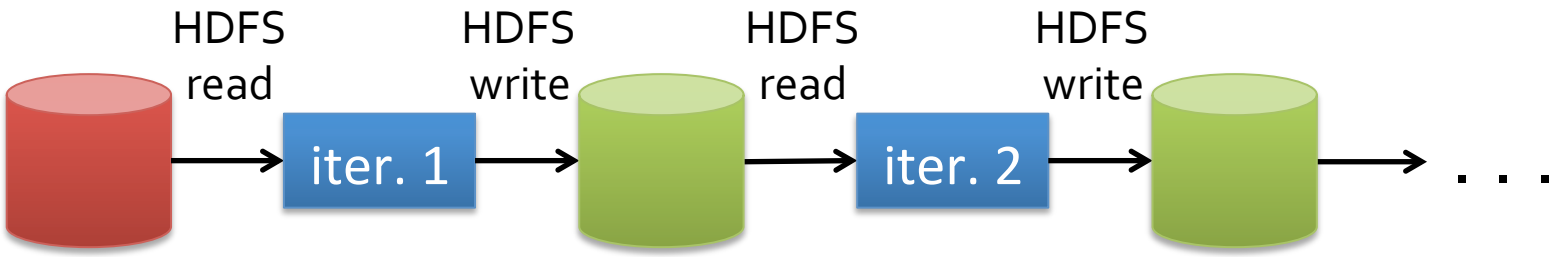


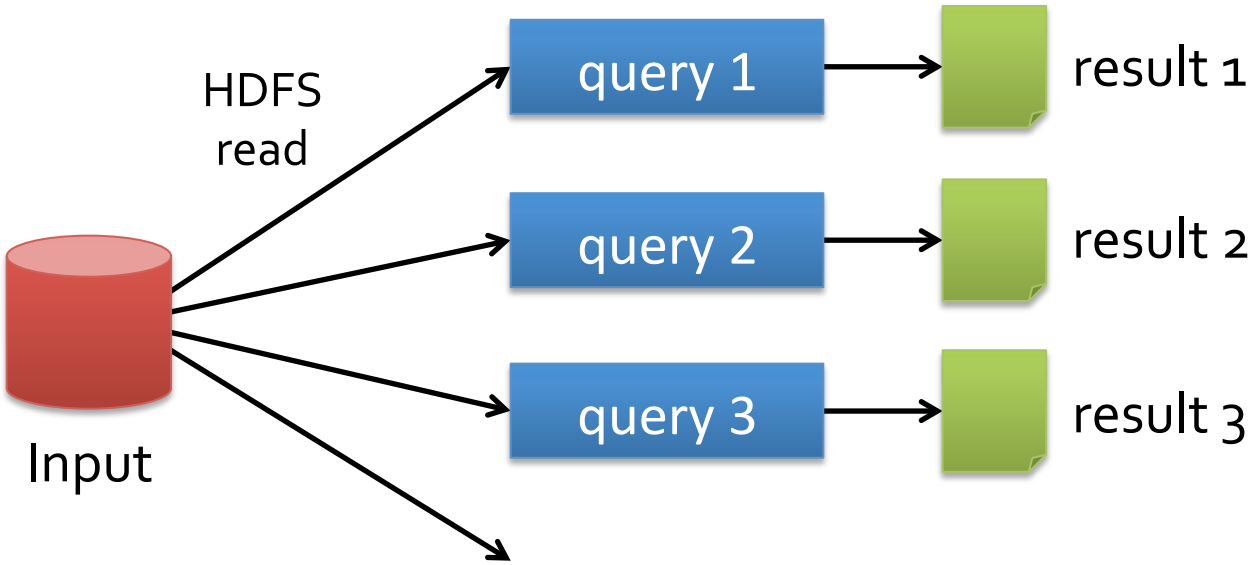
But…

- As soon as it got popular, user want more
  - More *complex*, multi-stage jobs
  - More *interactive*, ad-hoc queries
  - More *real-time* stream processing



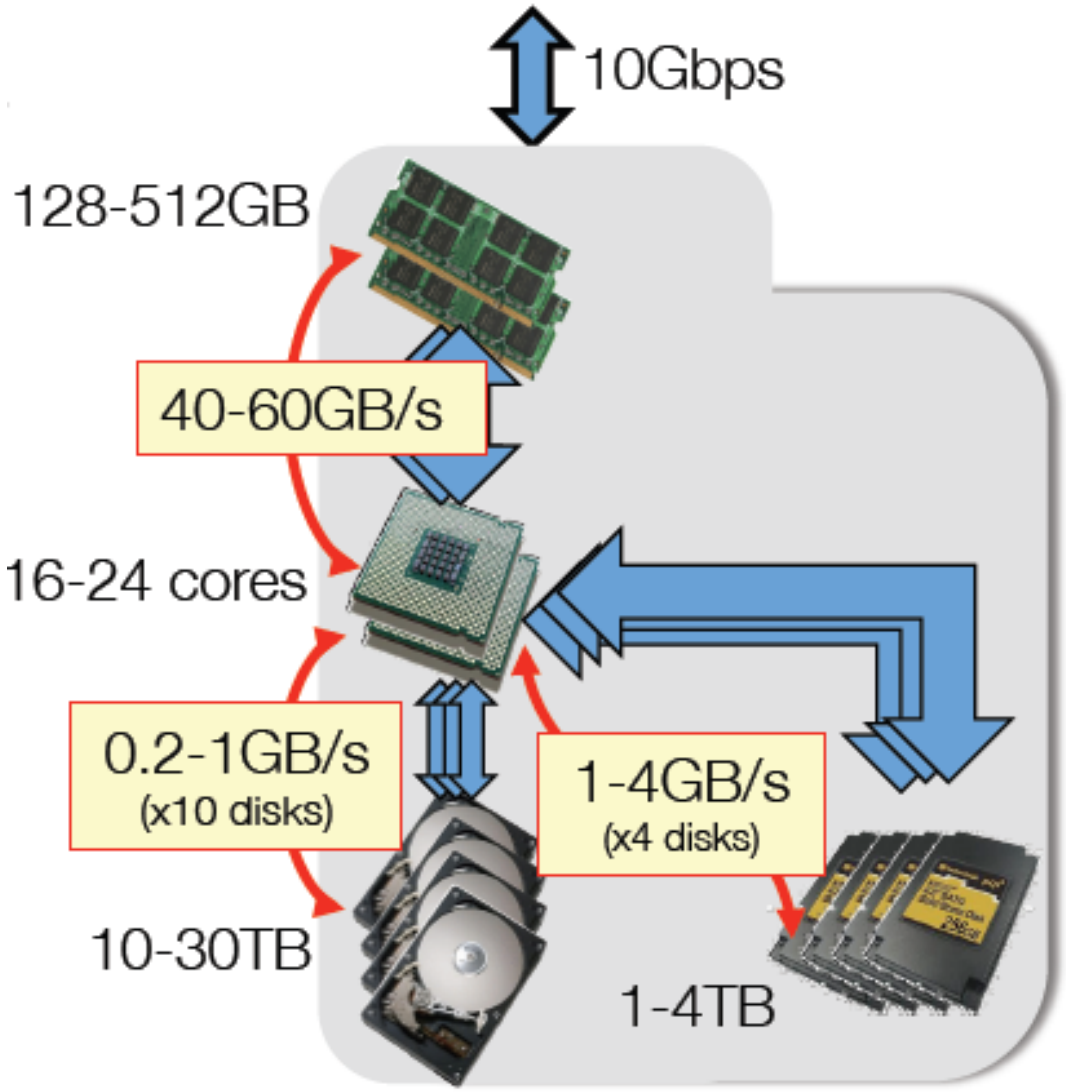
Hadoop data sharing





Slow due to replication, serialization, and disk IO

High-end Data Center Node

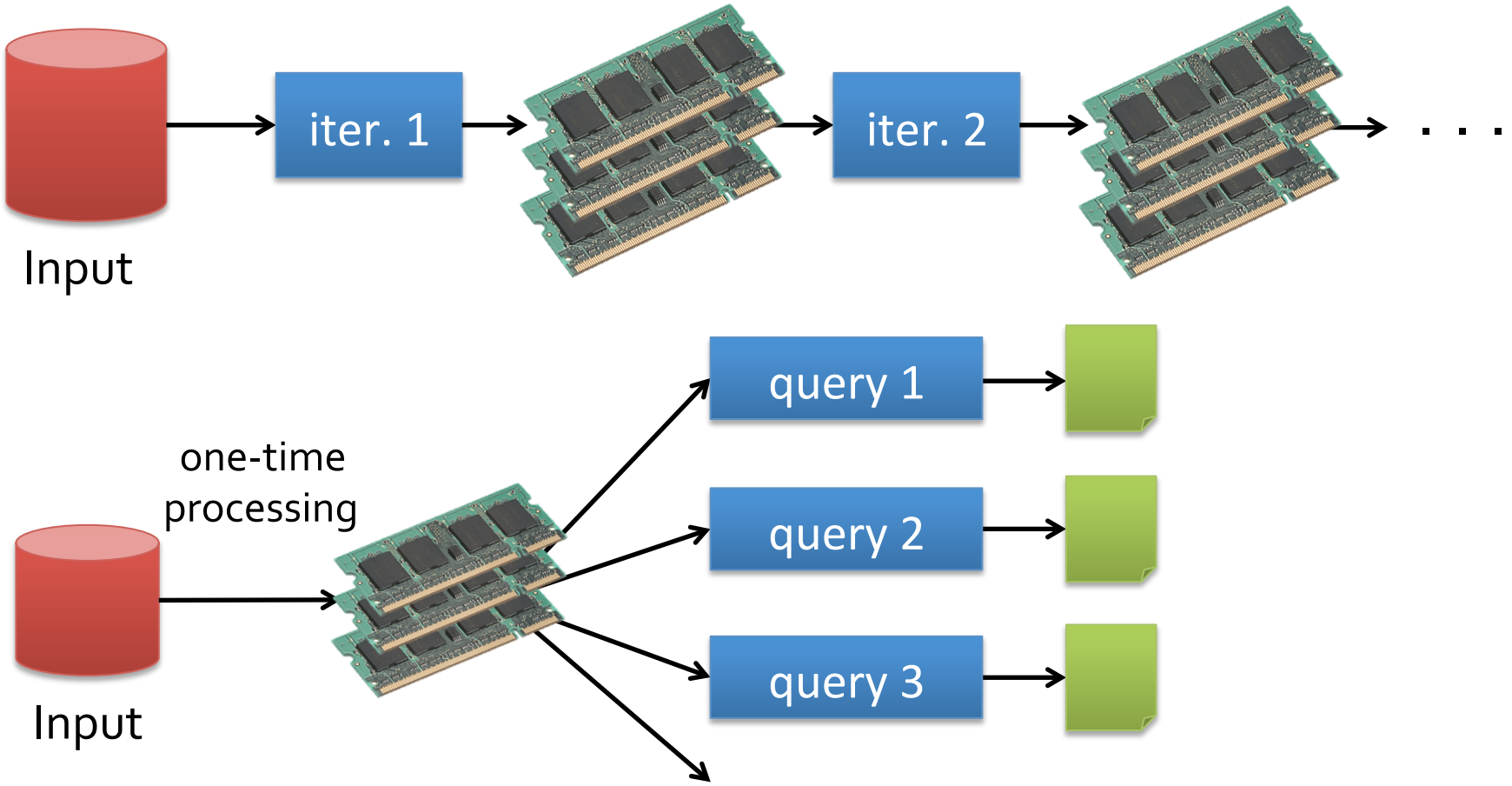


High-end Data Center Node(cont.)

- Memory bus >> disk & SSDs
- Memory density (still) grows with Moore's law
- Many datasets fit into memory
  - The inputs of over 80% of jobs in Facebook, Yahoo!, and Bing clusters fit into memory
  - 1TB = 1 billion records @ 1 KB

Memory is **King**

Goal : In-Memory Data Sharing



10-100× faster than network and disk

- Matei started Spark from 2009
- PhD from UC Berkeley, now assistant professor in MIT
- As CTO and co-founder, databricks raised $14M

Agenda

- Why need Spark?
- ***How to program with spark?***
- What is Spark ecosystem?
- Community

RDD: Resilient Distributed Datasets

- Restricted form of distributed shared storage
  - Immutable, partitioned collections of records
  - Can only be built through coarse-grained deterministic transformations (map, filter, join, …)
- Unify many current programming models
  - Data flow : MapReduce, Dryad, SQL, …
  - Specialized for iterative apps: BSP (Pregel), iterative MapReduce (Haloop) …

Spark Programming Interface

- Input & output
  - Local file, HDFS, HBase, etc
- Operations on RDD
  - Transformation: build new RDD
  - Actions: compute and output results
- Controller on RDD
  - Partitions: layout across nodes
  - Persist: storage RDD in RAM, on disk, etc

Example: Log Mining

Load error messages from HDFS into memory, then interactively search for various patterns

```
lines = spark.textFile("hdfs://...")
errors = lines.filter(_.startsWith("ERROR"))

errors.persist()

Mysql_errors = errors.filter(_.contains("MySQL")).count
http_errors = errors.filter(_.contains("Http")).count
```

Base RDD

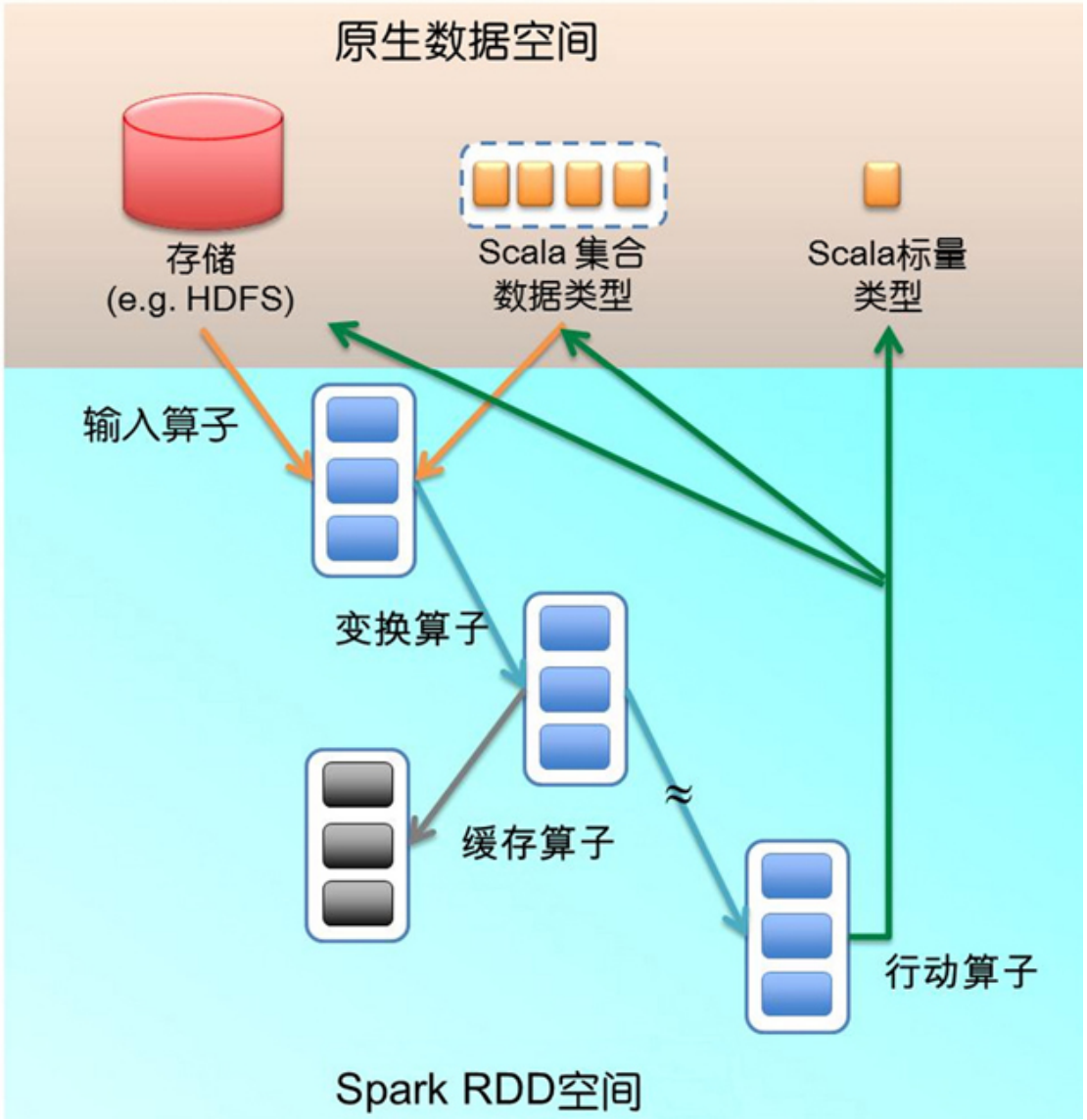
Transform operation

Controller of RDD

Action operation

Result: scaled to 1 TB data in 5-7 sec (vs 170 sec for on-disk data)

Spark workflow

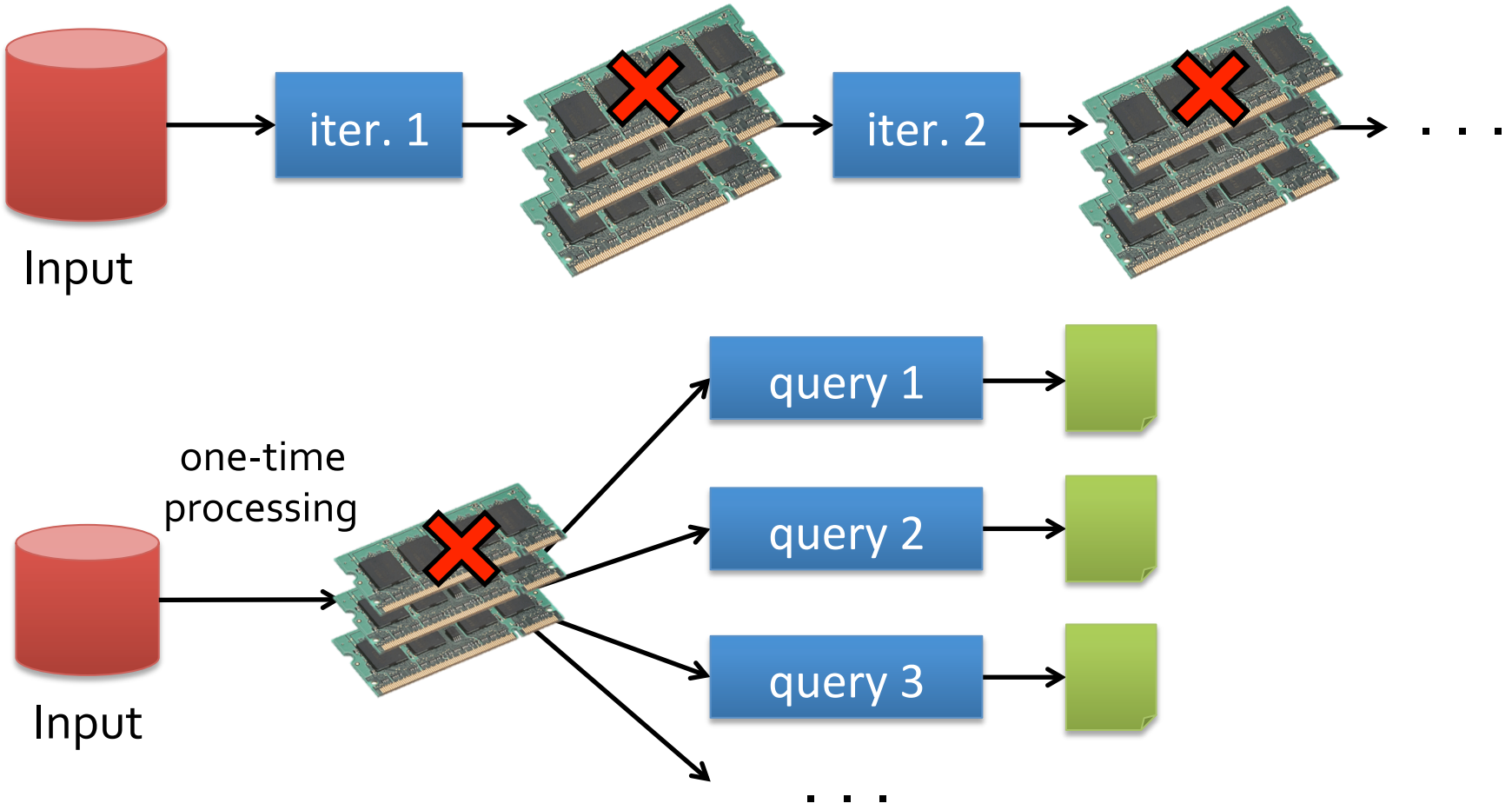


Not enough memory to cache?

FAQ1: Not enough memory?

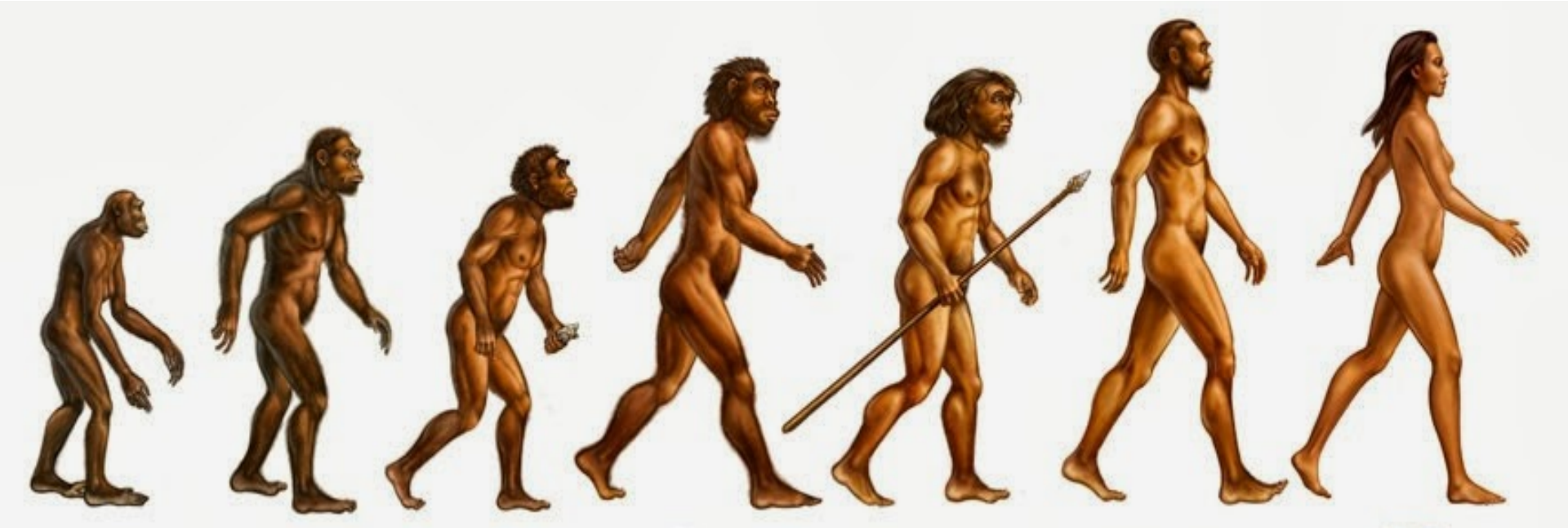
- Graceful degradation
- Scheduler takes care of this
- Other options
  - MEMORY_ONLY
  - MEMORY_ONLY_SER
  - MEMORY_AND_DISK
  - MEMORY_AND_DISK_SER
  - DISK_ONLY

How to recovery RDD?



FAQ2:How to recovery?

- Lineage*:* track the graph of transformations that built RDD
- Checkpoint: lineage graphs get large



Agenda

- Why need Spark?
- How to program with spark?
- ***What is Spark ecosystem?***
- Community

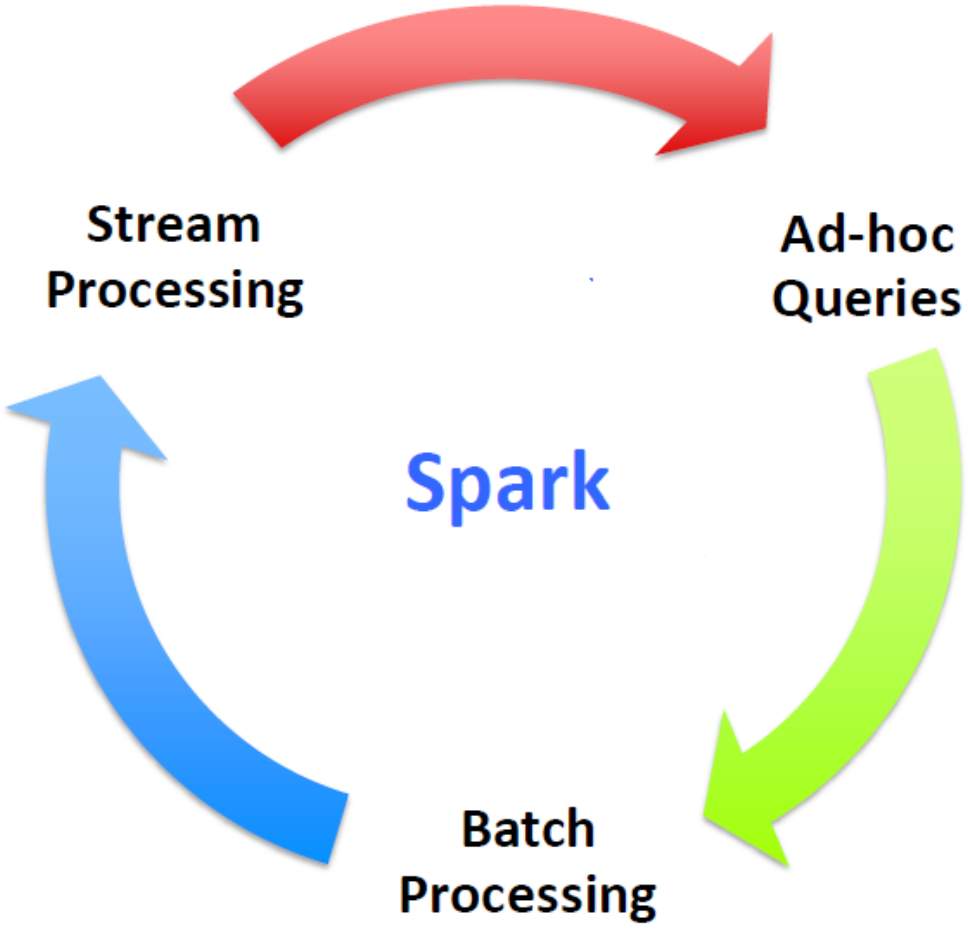
Data Process Goals

- Low latency queries on live data
  - Streaming process
  - Enable decisions on real-time data
  - E.g., detect & block spam in real-time
- Low latency queries on historical data
  - Interactive/ Ad-hoc query
  - Enable faster decisions
  - E.g., identify why a site is slow and fix it
- Sophisticated data processing
  - Batch processing
  - Enable better decisions
  - E.g., trend analysis

Existing big data projects

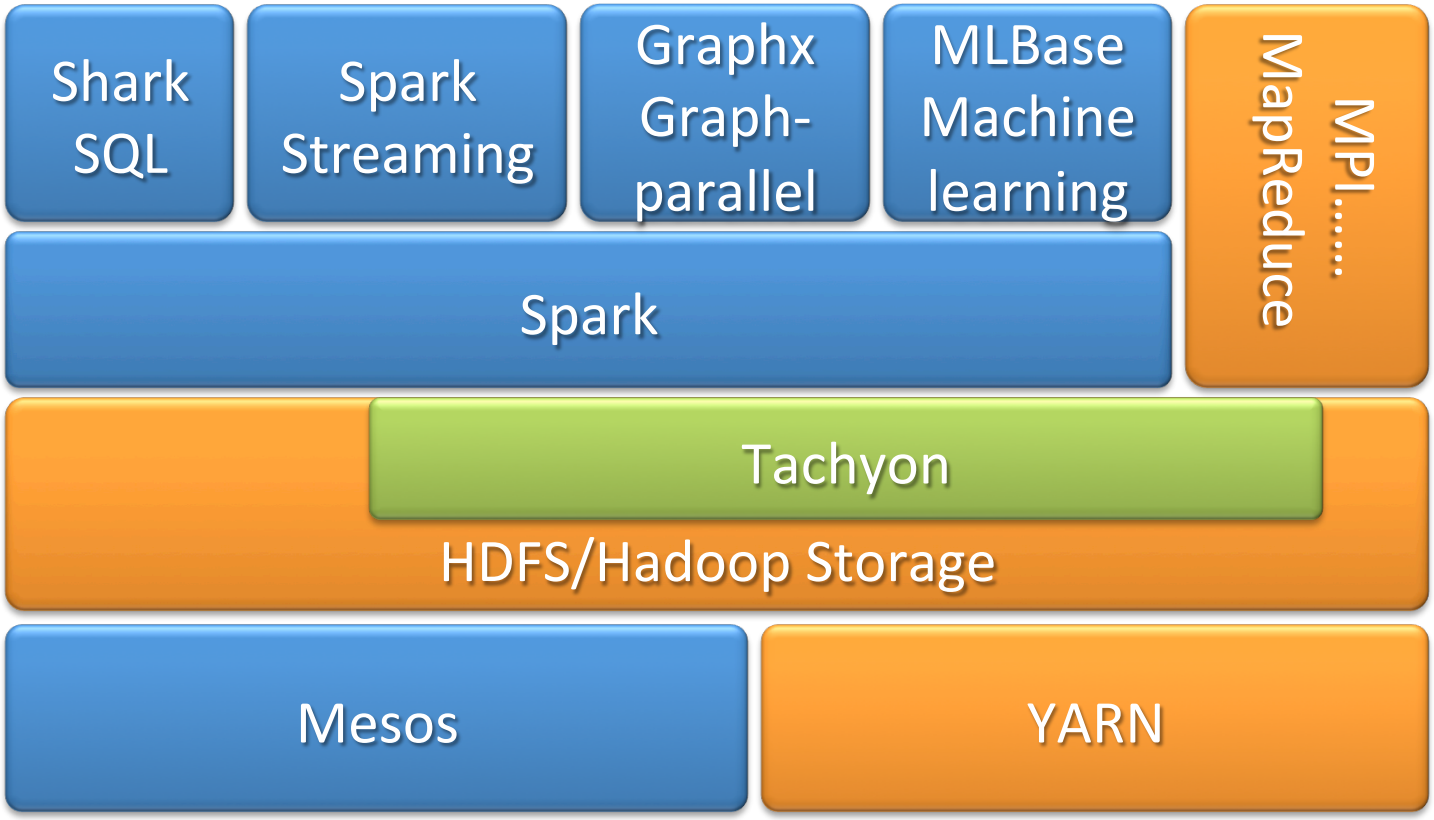
- Streaming process
  - Twitter Storm
  - Apache S4
  - LinkedIn Samza
- Ad-hoc query
  - Google Dremel
  - Cloudera Impala
  - Apache Drill
- Batch process
  - Apache Hadoop
  - Apache Hive

Vision of Spark ecosystem



One stack to rule them all!

Spark ecosystem



Agenda

- Why need Spark?
- How to program with spark?
- What is Spark ecosystem?
- ***Community***

Community

- Apache incubator
- Global
  - 3000 people attended online training
  - 1100+ meetup members
  - 24 companies
  - 23 committers
  - 90+ contributors
- Intel China
  - 3 committers
  - 7 contributors
  - 70+ patches

Collaboration

- Machine learning algorithm optimization
- Real-time ranking system
- Recommendation system

Summary

- ***Memory is King!***
- ***One stack to rule them all!***
- ***Contribute to community!***


intel®
Leap ahead™

特别感谢
QCon上海合作伙伴