Wrocław University of Science and Technology

Large Scale Data Processing

Lecture 4 – Flink

dr hab. inż. Tomasz Kajdanowicz, Piotr Bielak, Roman Bartusiak, Krzysztof Rajda

November 30, 2021

HR EXCELLENCE IN RESEARCH

Wrocław University of Science and Technology

Overview

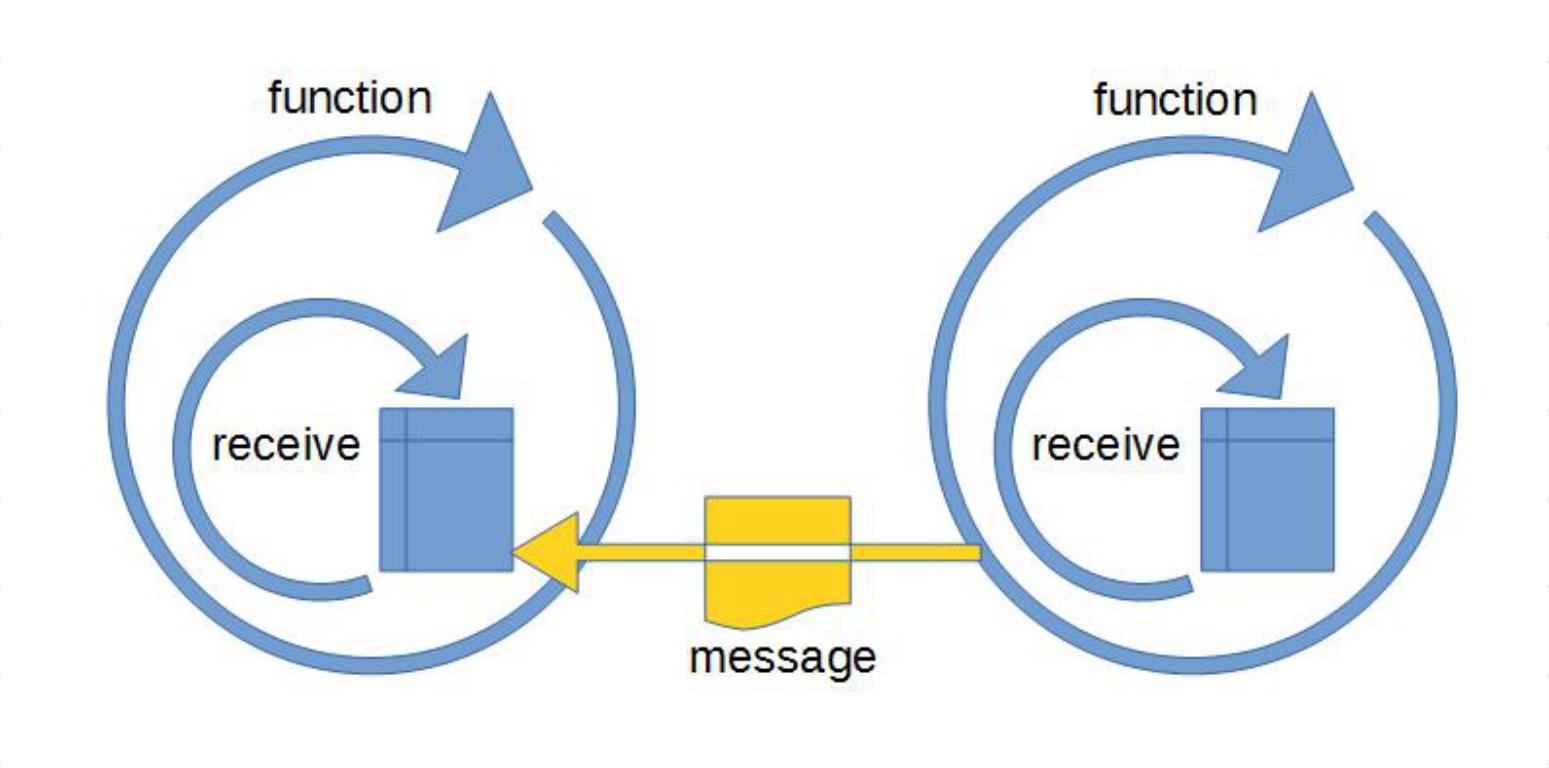
Flink

Flink

Flink

Actor programming

Flink



Actor programming

Flink

- Actors - can send message, modify hist state, spawn new actors
- Message queues
- Actor behaviour

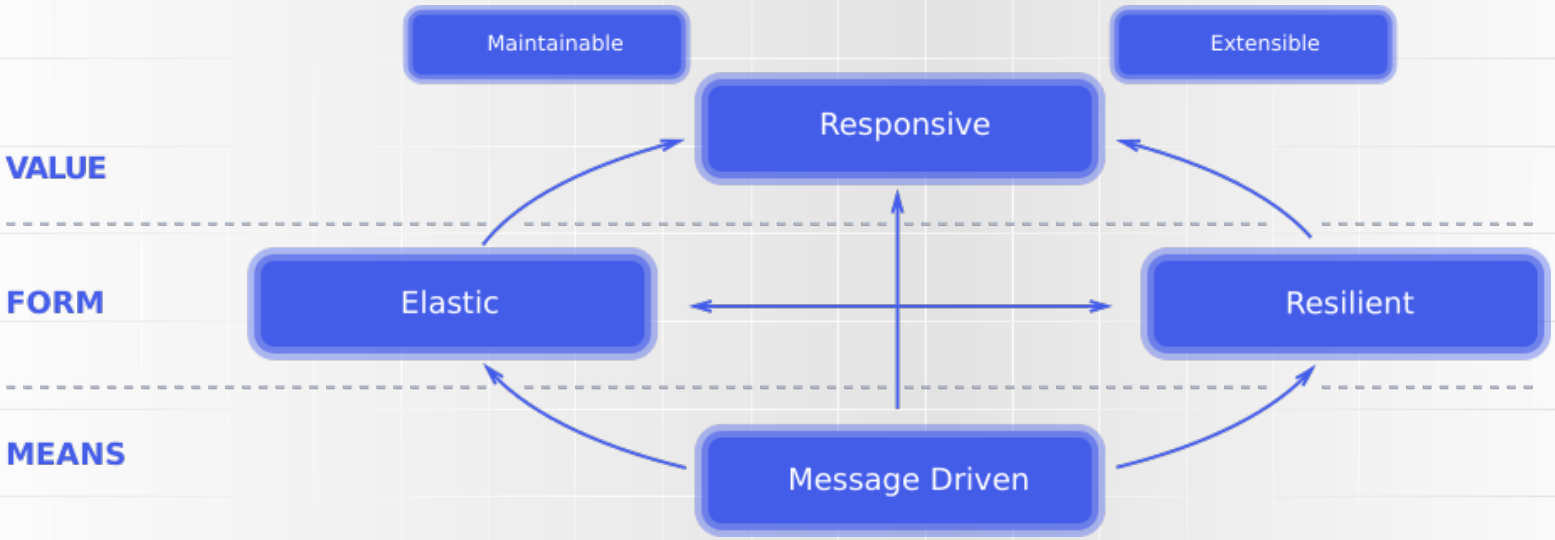
Actor programming

Flink

- Message - envelope with sender and receiver
- Offload response processing to third-actor
- Queues - communication decoupled from processing
- Only local state is mutable

Reactive manifesto

Flink



Akka platform

Flink

Akka platform

Flink

- Akka Actors
- Akka typed
- Akka streams
- Akka persistance
- Alpakka

Akka platform- Akka Actors

Flink

- core
- overhead - 300 bytes per actor
- actor system
- actor primitive
- communication patterns
  - tell
  - ask
  - forward
- coordination
- supervision

Akka platform- Akka typed

Flink

- new actor api
- typed communication
- typed behaviour
- typed state

Akka platform- Akka streams

Flink

- powerfull stream processing API
- processing graph DSL
- backpreasure
- reactive streams
- materializes to actors

Akka platform- Akka cluster

Flink

- cluster formation
- cluster shutdown
- sharding
- singletons
- remote actors
- load balancing

Akka platform- Akka persistence

Flink

- event sourcing - next lecture
- persist actor state - by persisting events
- snapshot actor state - so the recover is not from all events
- recover actors
- latest state behaviour - persist only state not all events

Akka platform- Alpakka

Flink

- akka stream connectors
- Cassandra
- Kafka
- AWS

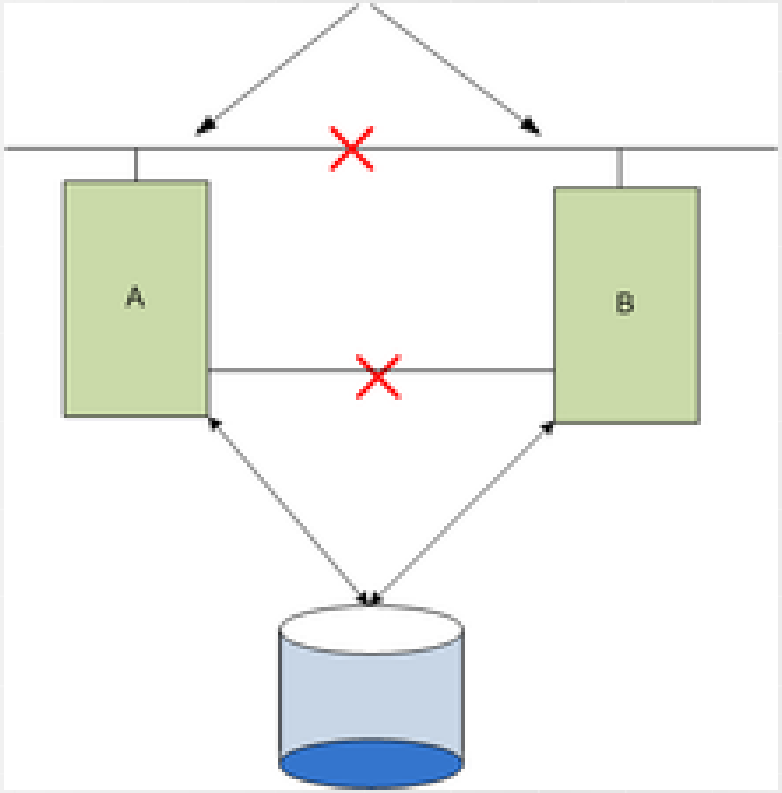
Akka cluster problems

Flink

- Split-brains
- Cluster formation
- Failure detection

Akka cluster - SB

Flink



Akka cluster - SBR

Flink

- Detect that some nodes are not responding
- Apply configured policy to handle split

Akka cluster - SBR - Keep majority

Flink

- cluster know it size
- on split, each part also knows size, and the size of new "cluster"
- down smaller part

Akka cluster - SBR - Static quorum

Flink

- configure quorum-size
- compare new "cluster" size to quorum
- down if size is smaller

Akka cluster - SBR - Keep oldest

Flink

- singletons are kept on oldest node
- on failure, if oldest node is not part of the new "cluster" - down

Akka cluster - SBR - Down all

Flink

- kill the cluster

Akka cluster - SBR - Lease

Flink

- utilize external "lock"
- K8S
- time based
- lose lock if not updated in time

Partitioning

Flink

- Horizontal
- Vertical
- Functional

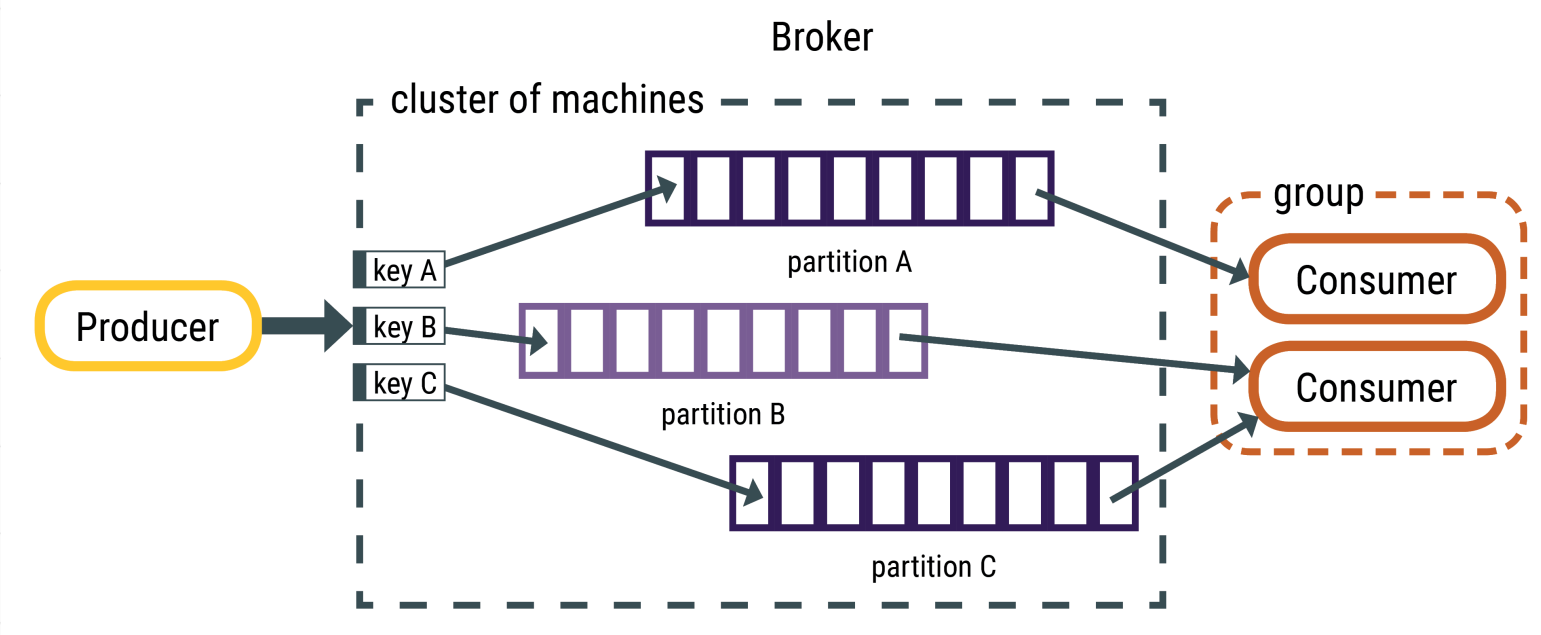
Partitioning - Horizontal

Flink

- Range
- Hash

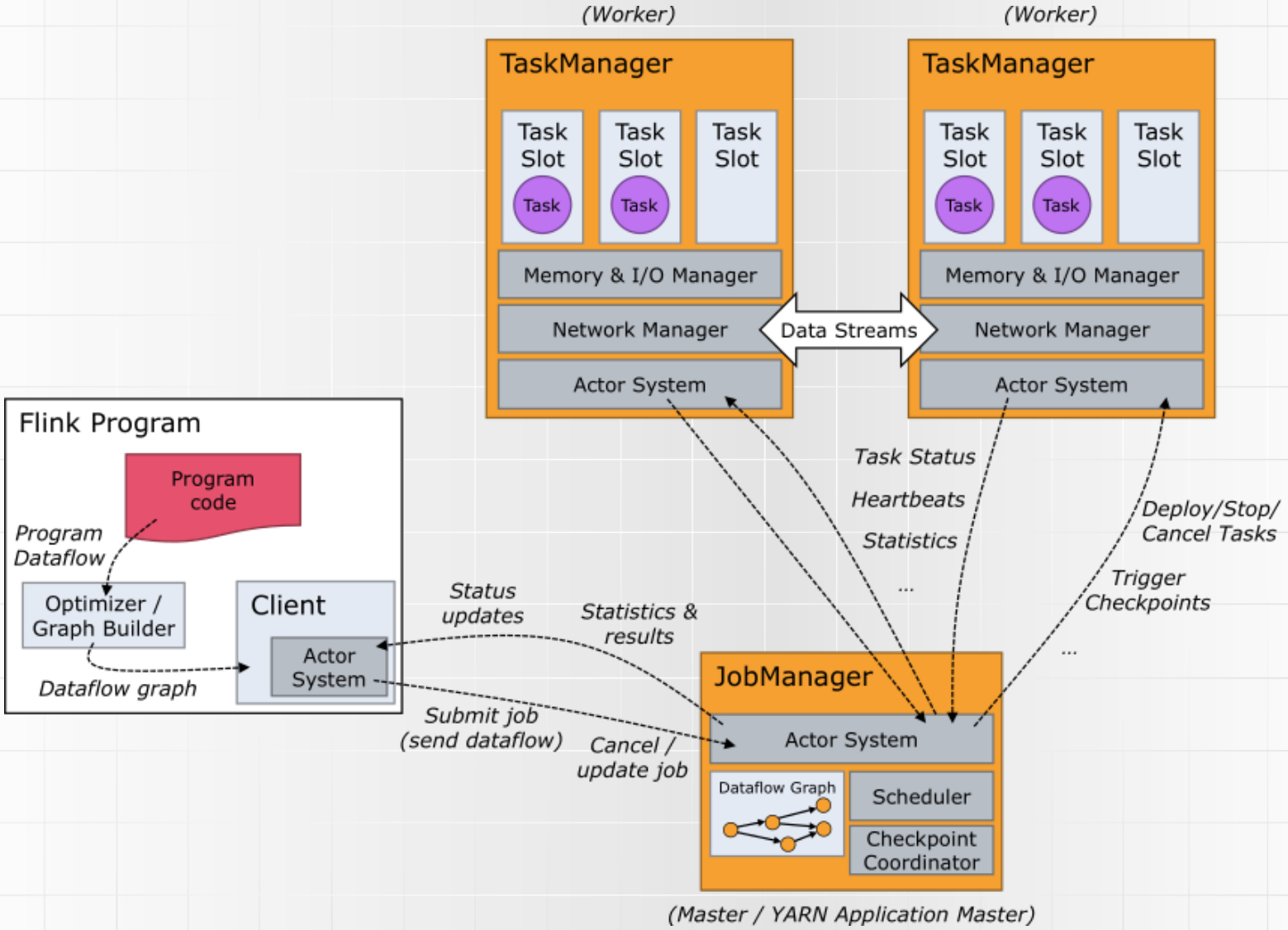
Partitioning

Flink



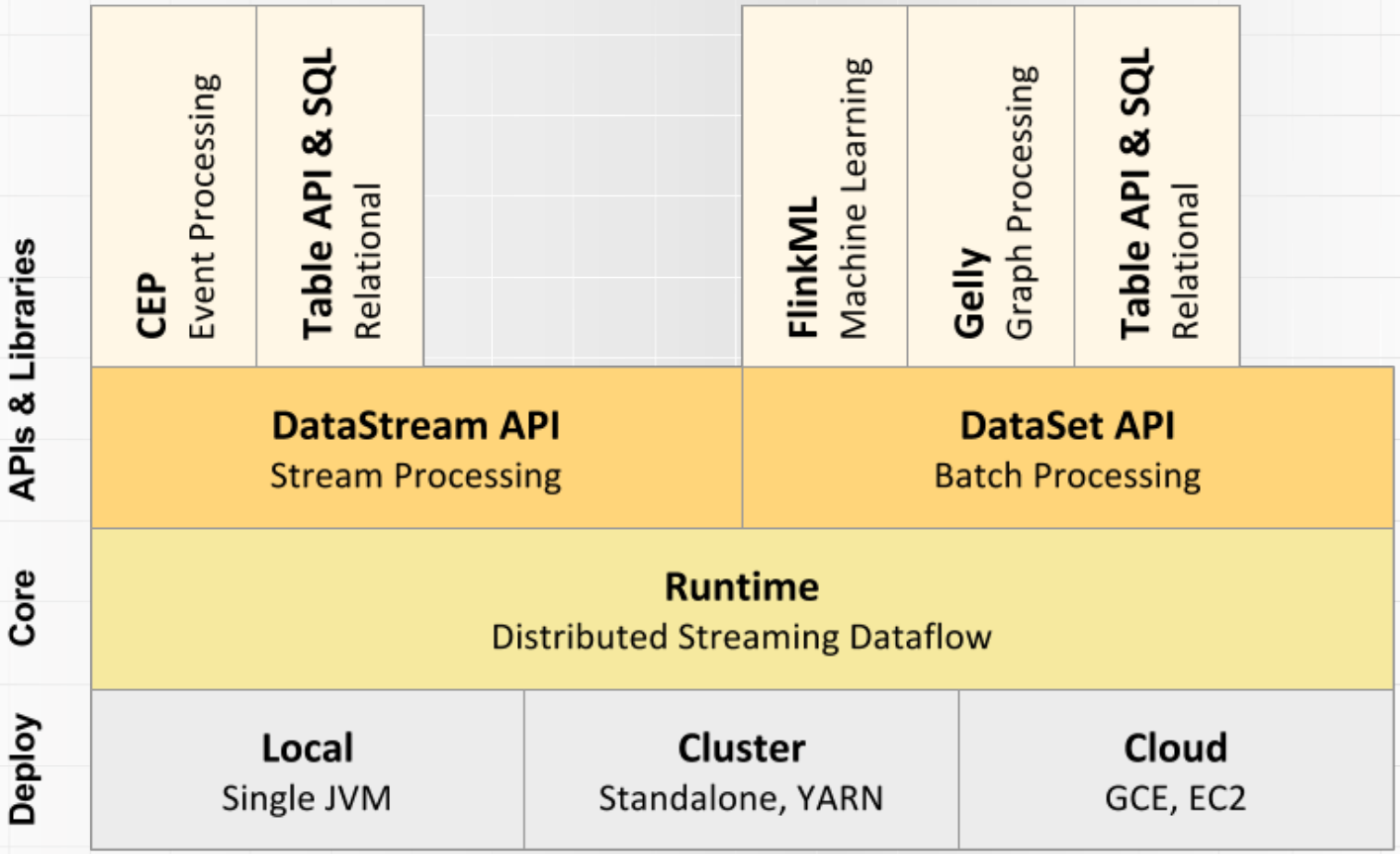
Architecture

Flink



Architecture - components

Flink



Concepts - Programs and DataFlows

Flink

- never ending flow of data
- transformations - operators
- "DAG"
- streaming dataflow - 1-n sources -> 1-n sinks

Concepts - Parallel DataFlows

Flink

- parallel by partitioning
- each operator - one or more subtask
- parallelism of stream = parallelism of upstream
- one-to-one - preserve partitioning
- redistributing - repartition

Concepts - Windows

Flink

- impossible to count all elements in a stream
- scope aggregation operations
- time drive
- data driven
- tumbling
- sliding
- session

Concepts - Time

Flink

- event time
- ingestion time
- processing time

Concepts - Statefull operations

Flink

- Simple processing - look only at current event
- remember information in between events - statefull operations
- partition by key
- only on "keyed" streams

Concepts - Checkpointing

Flink

- fault tolerance - reply + checkpointing
- checkpoint - might be expensive, do only in some intervals
- on failure - load checkpoint and reply events after the checkpoint

Concepts - Batch processing

Flink

- under the hood - streams
- no checkpoints
- in-memory state
- iterations

Concepts - Tasks and operators

Flink

- operator - for example map()
- task - chained operators, executed on single thread
- chaining - optimization
- configurable - programmable API

Concepts - Job managers, task managers, clients

Flink

- Job manager - master
- Task manager - worker
- Client - starts processing

Concepts - Task slots

Flink

- Task manager -Single JVM
- Task slot - part (or whole) part of task manager
- control on task isolation granularity
- task slot sharing
- needs at least as many task slots as max-parallelism

Concepts - State backend

Flink

- key/value store
- in-memory map
- RocksDB
- snapshotting as part of checkpoint

CEP

Flink

- Complex Event Processing
- on-top
- detect event patterns
- "regexp" for event streams

Serialization

Flink

- by default internal efficient serialization for some primitive types, POJOs, case classes
- fallback - Kryo - schema in every event
- fallback - custom

Object reuse

Flink

- Flink - Java
- mutability
- user program can mutate objects, that is why Flink needs to copy them to be safe
- instruct Flink programatically to reuse
- be immutable (copy where needed manually)

Wrocław University of Science and Technology

Next week

Flink

- Cassandra, CQRS, ES

Wrocław University of Science and Technology

Large Scale Data Processing

Lecture 4 – Flink

dr hab. inż. Tomasz Kajdanowicz, Piotr Bielak, Roman Bartusiak, Krzysztof Rajda

November 30, 2021