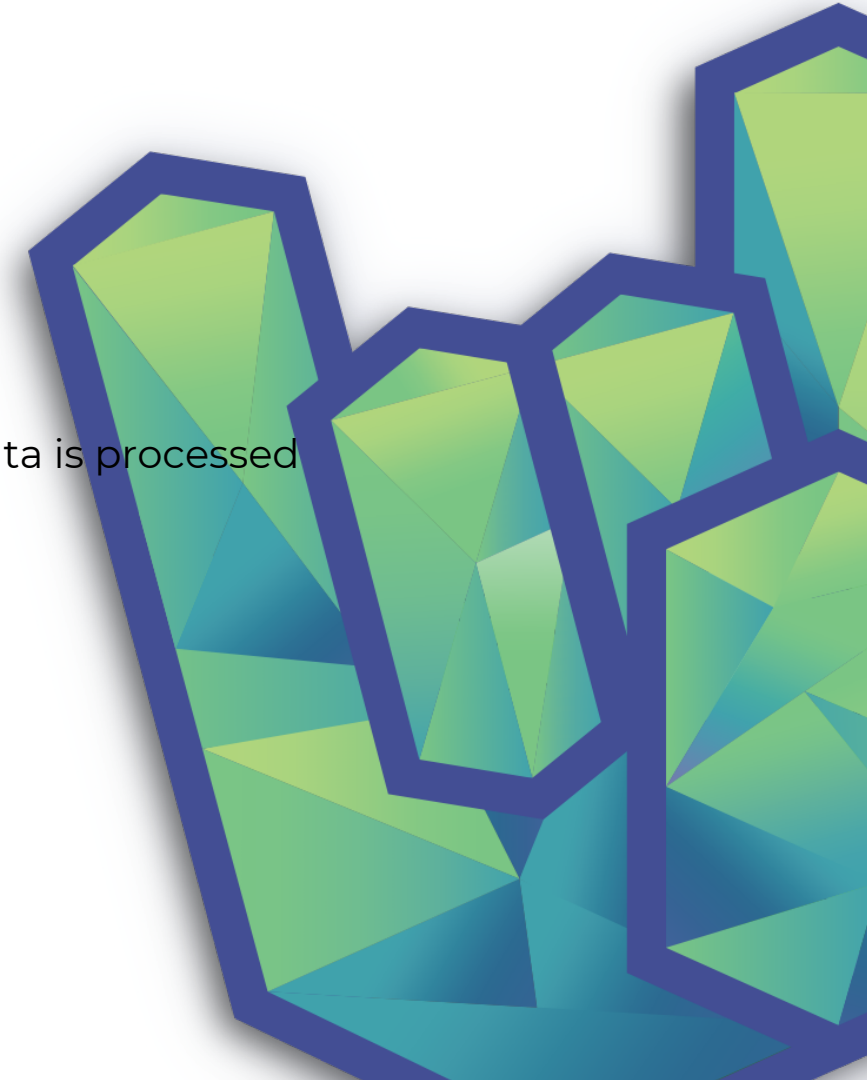
State

State

Flink can retain & update information while data is processed



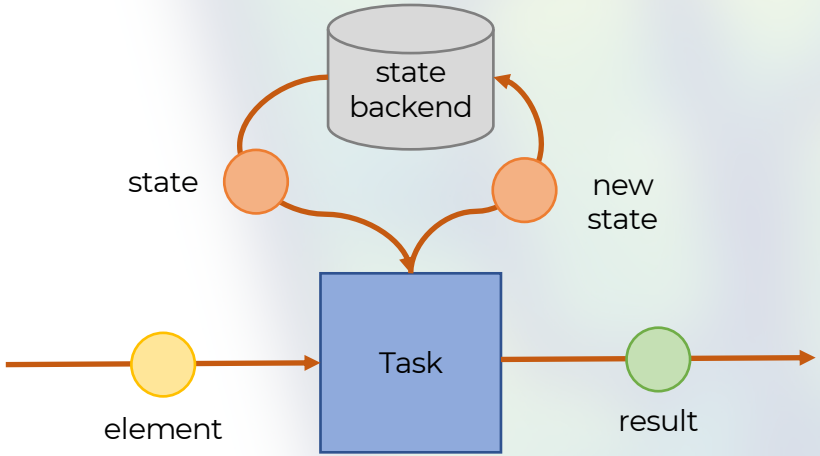
State Mechanics

Flink needs a state backend to retain it

- in-memory backend (HashMapStateBackend)
- file-based backend
- RocksDB backend
- (advanced) custom backend

State management process

1. event arrives
2. fetch state
3. update state
4. emit result



State Types

Operator state

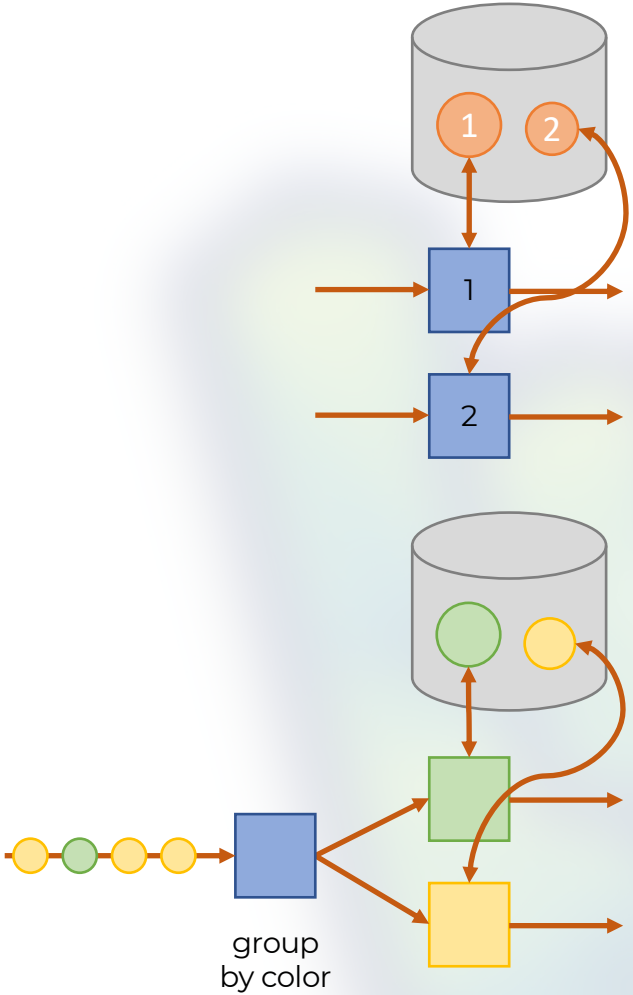
- private to an operator task
- cannot be read/written by another operator task

Keyed state

- one instance per key

About keys

- a piece of data that is used to group events
- multiple events can be assigned to the same key
- an event can be assigned to a *single* key
- examples of keys
  - hash code of a String field of the event
  - a number field of the event, mod 10
  - a numeric event type (enum)



Flink rocks