STREAMSTORE: Aligning Stream-processing Systems with Externally-Managed State

Antonis Papaioannou and Kostas Magoutis
{papaioan, magoutis}@ics.forth.gr

Motivation

Problem Statement
- Scalable stream-processing systems (SPS) often require externally-managed state
- Externally-managed state typically stored in scalable key-value stores (KVS) (such as Redis)
- Remote access to externally-managed state incurs a cost: **Data locality can improve performance**
- Co-locating SPS and their dependent KVS is one step to an optimal SPS-KVS data path
- Alignment is also needed to ensure fully-local access

Co-located systems != aligned systems

Amoeba: An Approach to Continuous Adaptation

Amoeba planning workflow
- Step 1: Adjust parallelism of SPS/KVS
- Step 2: Data or task migration
- Step 3: Align SPS/KVS partitioning scheme

Evaluation

Linear Road benchmark

Amoeba triggers alignment at 60 sec into the run

Alignment w/ data migration (Steps 2-3)

Alignment at scale

Alignment improves performance and per node resource utilization

Ongoing Research

Explore tradeoffs in the decision making process within Amoeba planning workflow