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

INSTITUTE OF COMPUTER SCIENCE
Information Systems Laboratory

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

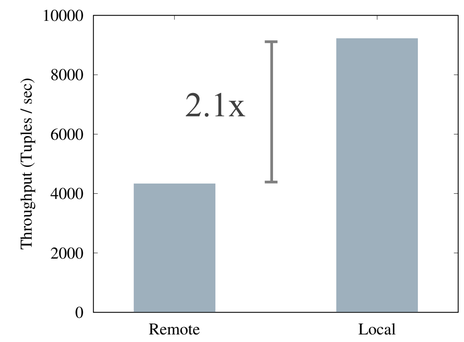
UNIVERSITY OF CRETE

Motivation

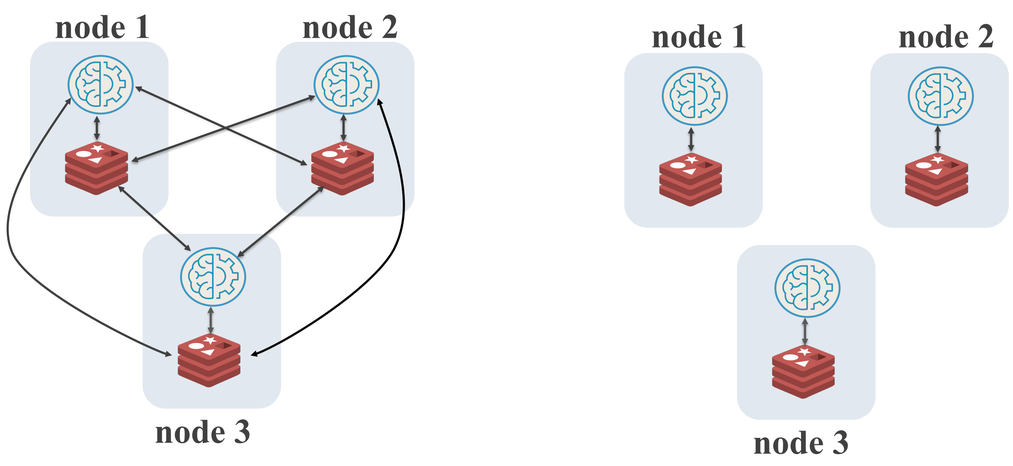
Problem Statement

- 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

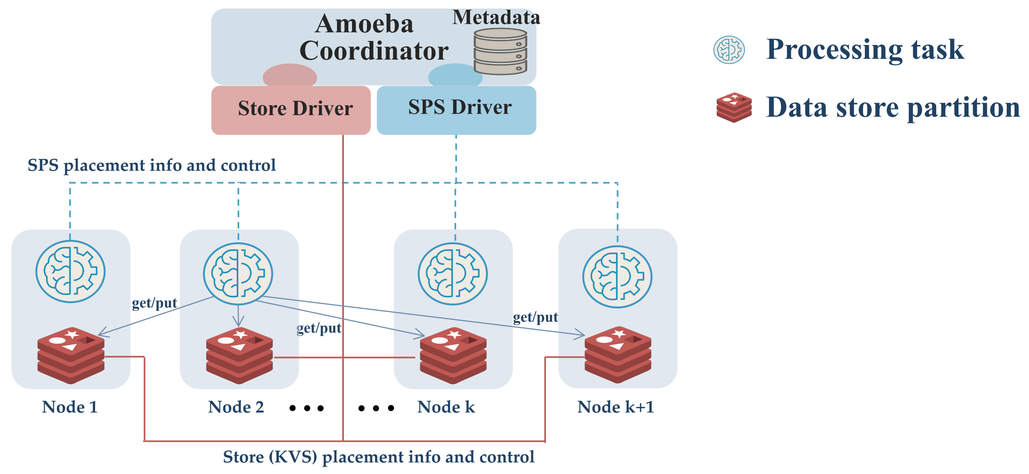
Effect of data locality on performance (Linear Road)



Co-located systems != aligned systems



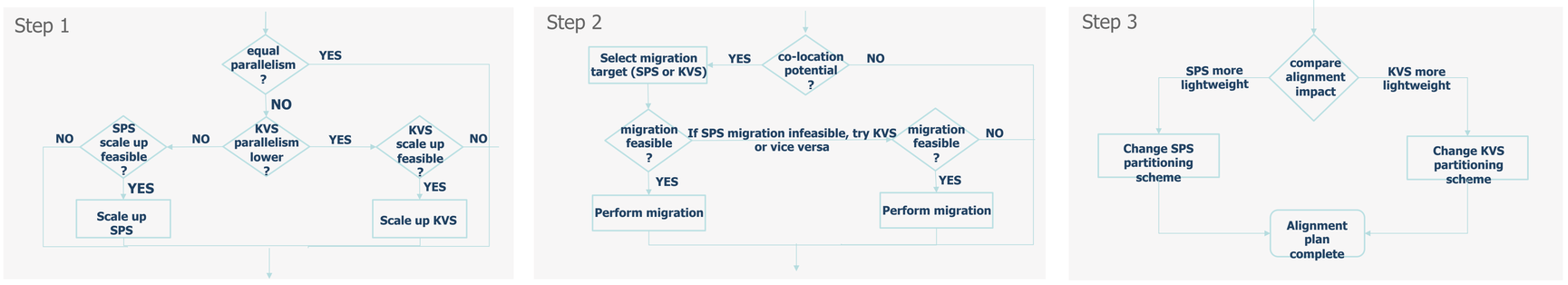
Amoeba: An Approach to Continuous Adaptation



Antonis Papaioannou and Kostas Magoutis, 2021. Amoeba: Aligning Stream Processing Operators with Externally-Managed State, UCC'21, December 6-9, 2021

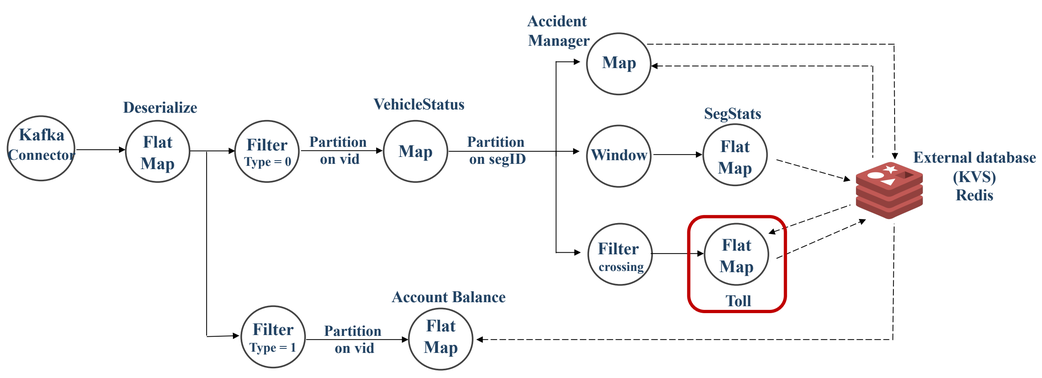
Amoeba planning workflow

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



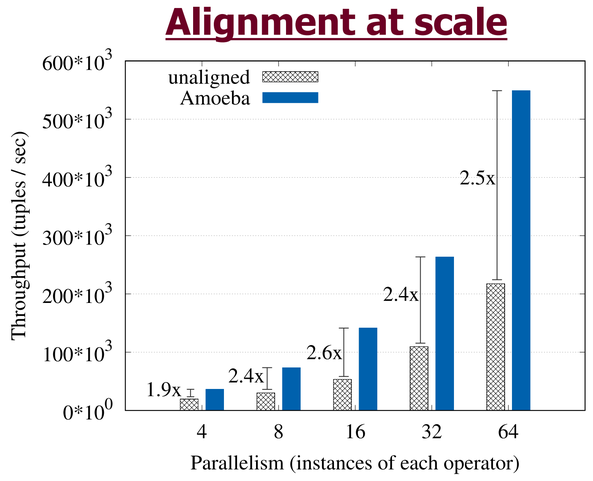
Evaluation

Linear Road benchmark



Redis Table	Updated by	Read by
SegmentStats	AccidentManager, SegStats	Toll
Account Balance	Toll	AccountBalance

Alignment at scale

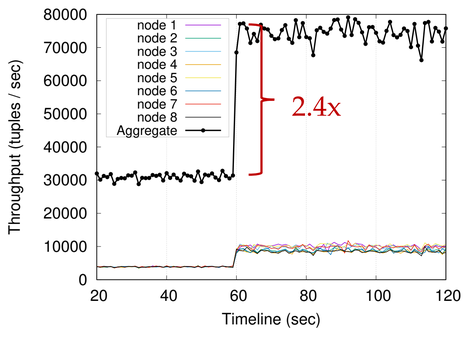


Evaluation testbed: Amazon EC2

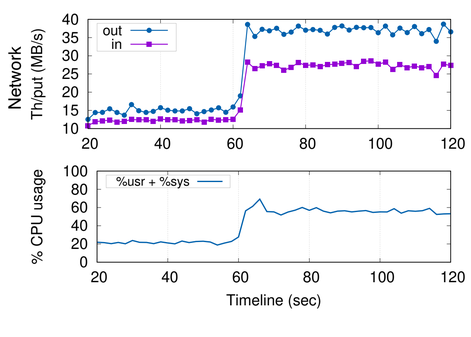
	#VMs	VM type	cores	DRAM	SSD*
Flink JM	1	c5d.large	2	4GB	50GB
Flink TM / Redis	1-64	c5d.xlarge	4	8GB	100GB
Kafka cluster	3	r5ad.large	2	16GB	75GB

*All instances are equiped with dedicated SSD instance store

Dynamic partition-scheme alignment (Step 3)

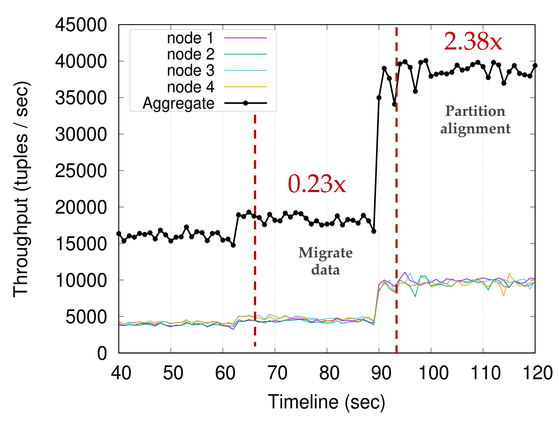


Amoeba triggers alignment at 60 sec into the run

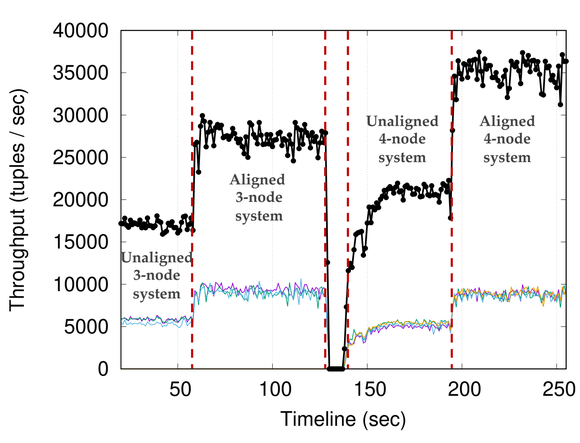


Alignment improves performance and per node resource utilization

Alignment w/ data migration (Steps 2-3)



Coordinating elasticity actions (Steps 1-3)



Ongoing Research

Explore tradeoffs in the decision making process within Amoeba planning workflow

H.F.R.I. Hellenic Foundation for Research & Innovation

17th ACM European Conference on Computer Systems (EuroSys 2022), April 5-8, 2022, Rennes, France

We thankfully acknowledge funding by the Hellenic Foundation for Research and Innovation (H.F.R.I.) through the STREAMSTORE faculty grant (Grant ID HFRI-FM17-1998)