TRANSACTIONAL SEMANTICS IN AIR

Course: CS 876 Streaming Data Systems
Name: Shivankar Pilligundla

Links

- Demo video: https://youtu.be/QPyl8dl_ymQ
- Project Report: https://docs.google.com/document/d/1OCbM_DFrQJRowh6358Ryyl0iNV9isGNY/edit?usp=sharing&ouid=106949801848675640556&rtpof=true&sd=true
- Github: https://github.com/shivankar-p/TransactionSemanticsAir

Problem Statement

- Emerging stream applications rely on the support of shared mutable states, where application states may be concurrently read and modified by multiple threads during stream processing.

- Transactional SPEs (TSPEs) have been recently proposed with built-in support of shared mutable states.

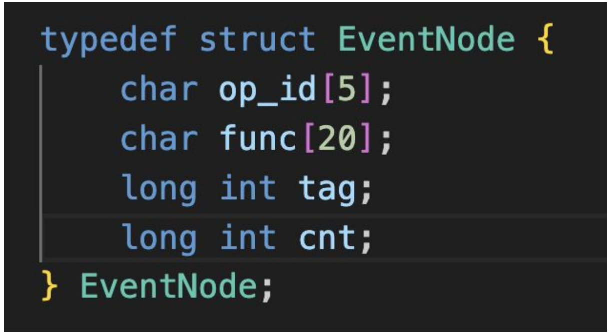
- TSPEs adopt transactional semantics during the processing of continuous data streams, where accesses to shared mutable states are modelled as state transactions

- AIR can be used as a TSPE due to its efficient Message passing interface and low communication overheads.

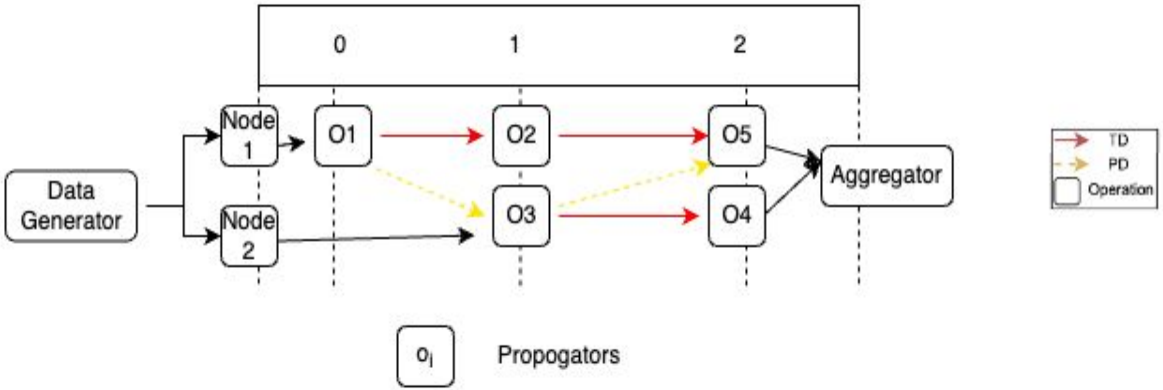
Approach

- Divide the processing into multiple phases namely data collection, mapping, execution and result generation phase.
- Data collection happens outside of AIR where transactions are split into operations and dependencies among them like LD, PD, TDs are identified.
- Mapping phase maps these operations efficiently to AIR dataflows, vertices in an efficient way to minimize coordination overhead.
- Execution phase starts with processing the operation logic in each vertex and sending and receiving signals from its dependencies or dependants.
- Result generation phase writes to the final state of the system.

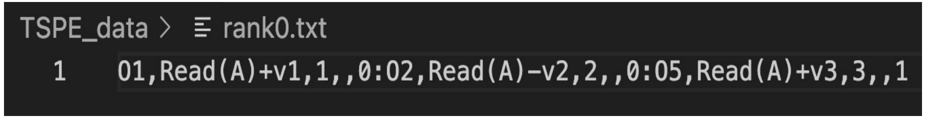
Implementation



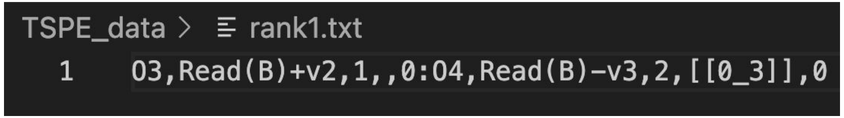
```
typedef struct EventNode {
    char op_id[5];
    char func[20];
    long int tag;
    long int cnt;
} EventNode;
```



Results & Evaluation


TSPE_data > rank0.txt

```
O1,Read(A)+v1,1,,0:O2,Read(A)-v2,2,,0:O5,Read(A)+v3,3,,1
```


TSPE_data > rank1.txt

```
O3,Read(B)+v2,1,,0:O4,Read(B)-v3,2,[[0_3]],0
```

- The implementation is ran on a basic TPG with 5 operations.
- O1, O2, O5 in rank 1 and O3, O4 rank 2.
- There is a dependency from O4 in rank 2 to O5 in rank1.

Conclusion

- It is pretty much feasible to extend AIR to function as a TSPE.

- I have successfully addressed the lack of transactional semantics in AIR for streaming data processing.

- Provided a foundation for further development and fine-tuning, making AIR a competitive TSPE in the streaming data domain.

Project Category

- I would nominate my work into **Excellent** Category.
- I have done adequate study about state of the art Transactional Stream Processing and its increasing popularity in the industry
- I have tried to implement a basic foundation for AIR to operate as a TSPE. By implementing the initial phases.
- Mapping and coordination has been tested for different TPGs to understand if the mapping is efficient.