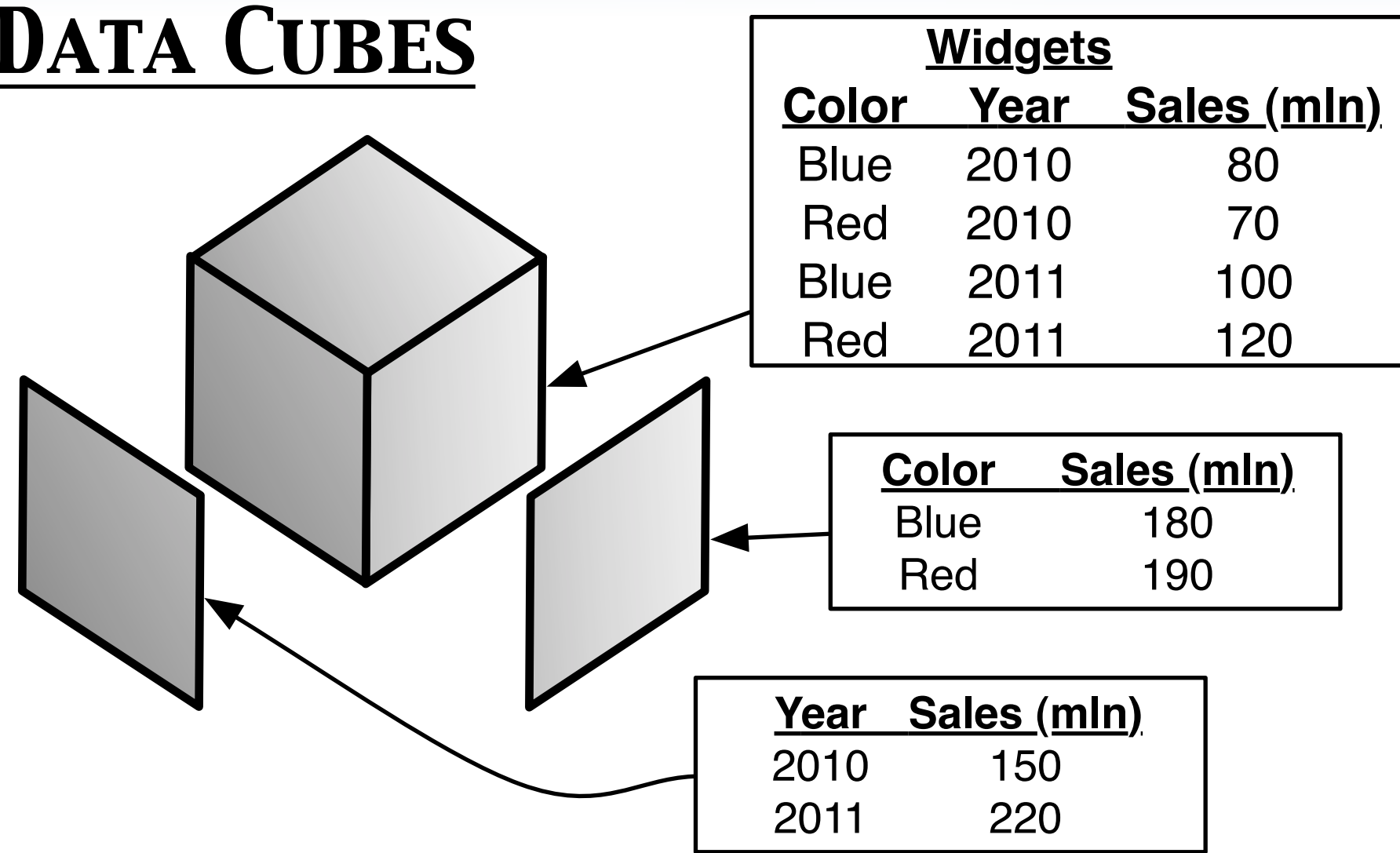
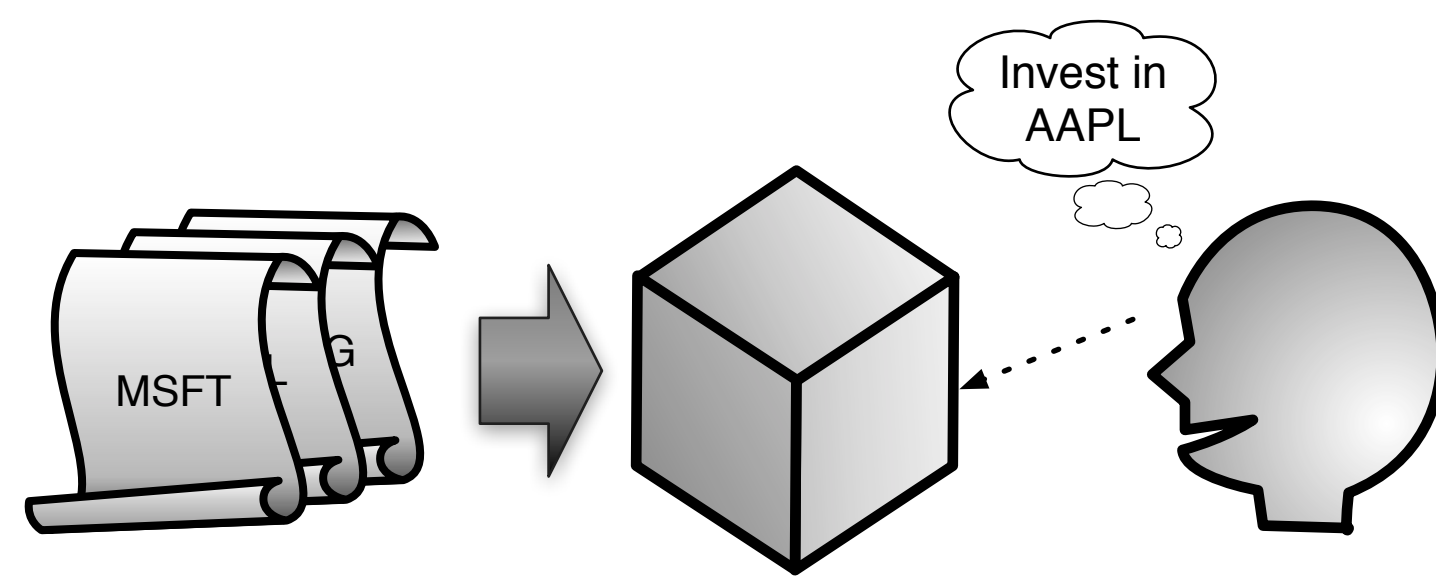


DATA CUBES

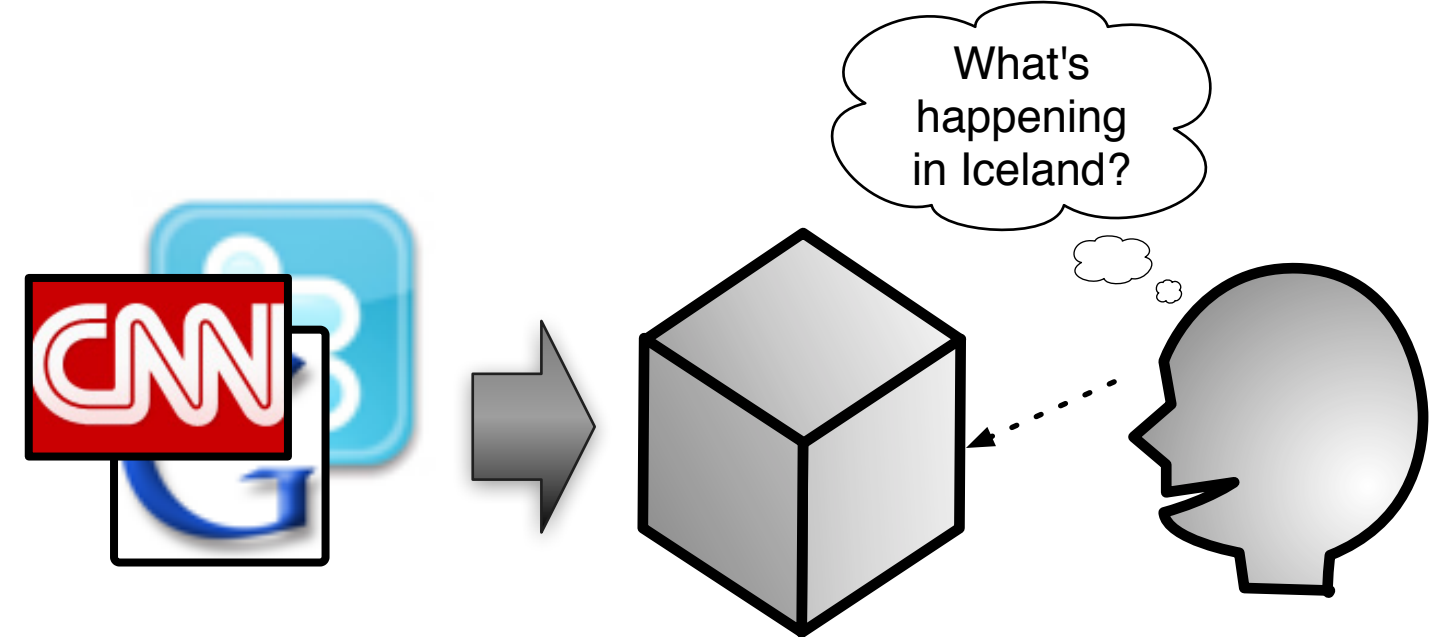


APPLICATIONS

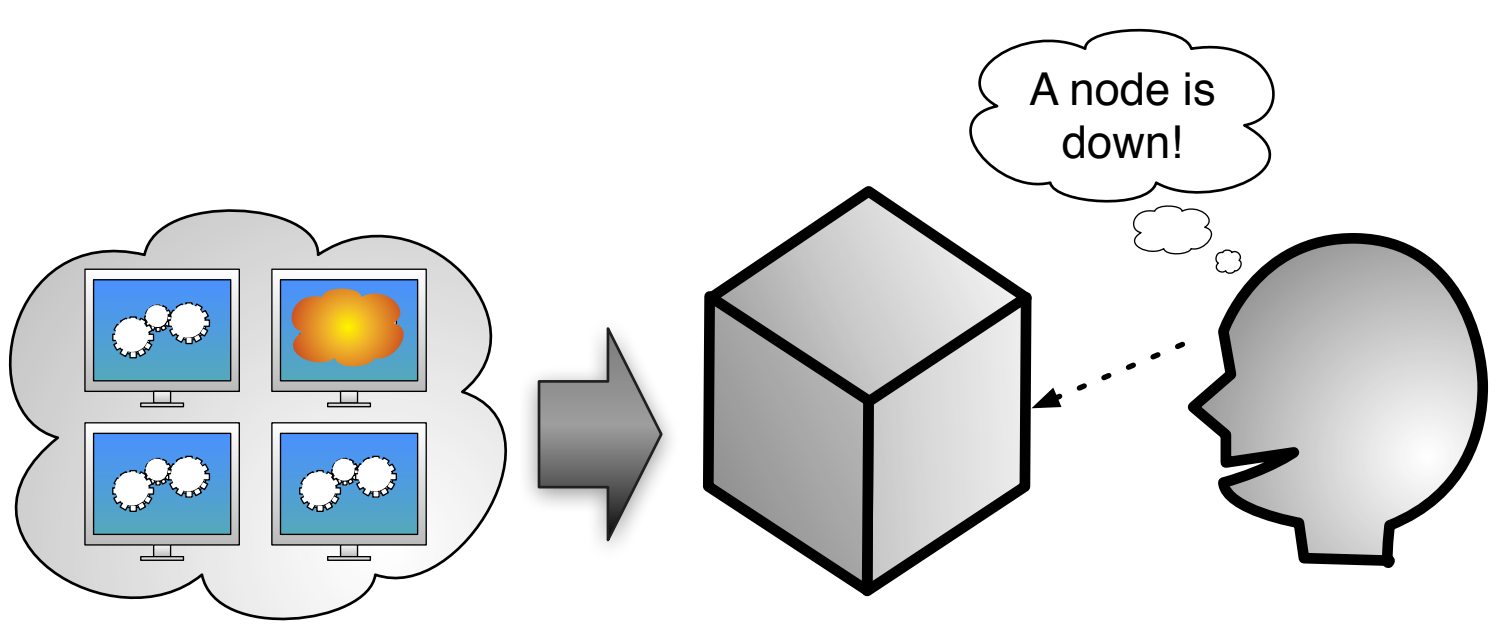
Stock Market Predictions



Trend Monitoring



Network Management



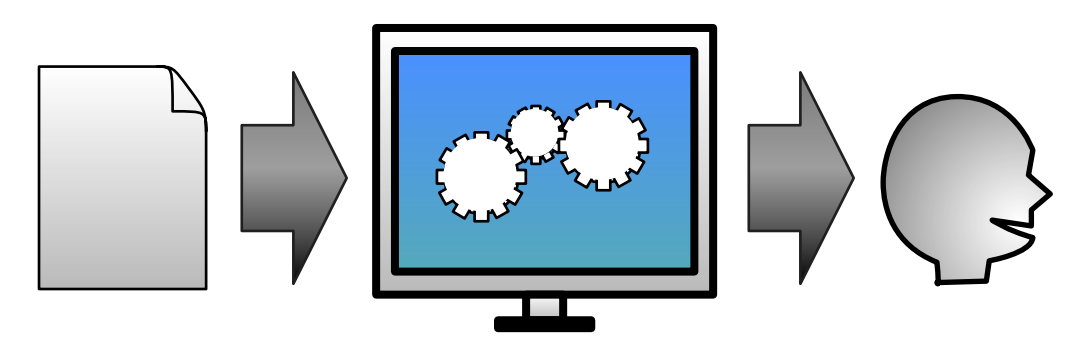
STREAM MONITORING

Approach

Design

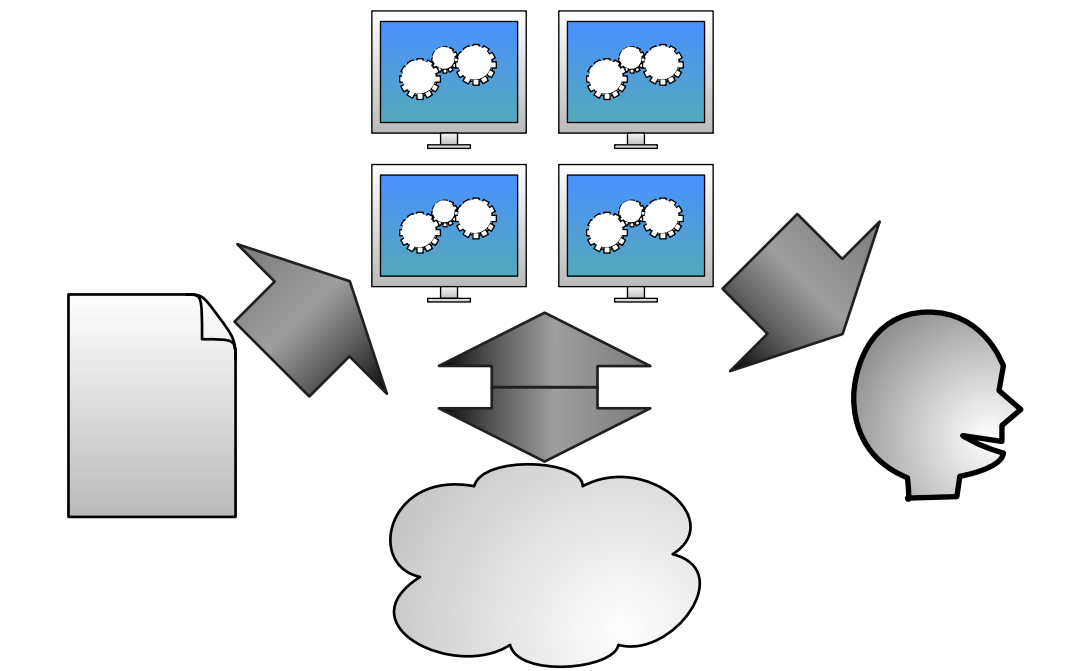
Limitations

Stream Processors
(Streambase, Cayuga)



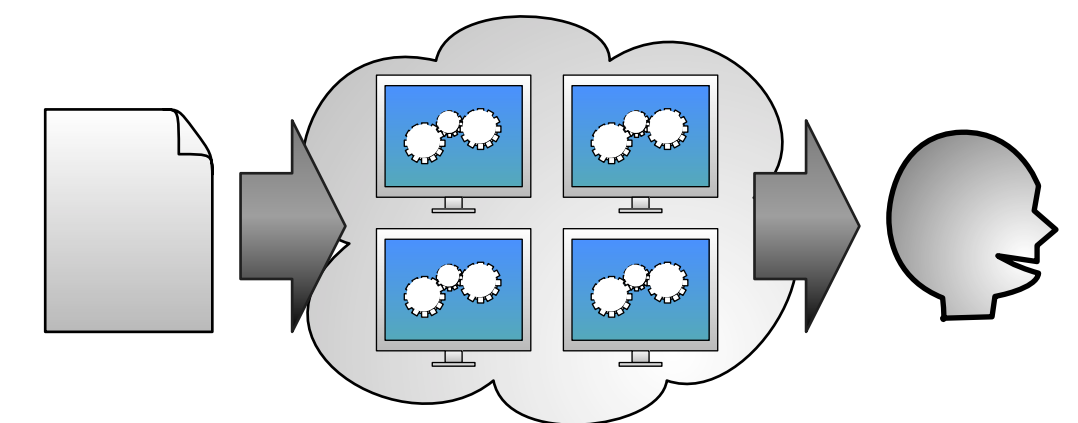
- Limited Storage
- Limited Processing Power
- Limited Query Complexity*

Cloud Storage
(PNuts, Cassandra, BigTable)



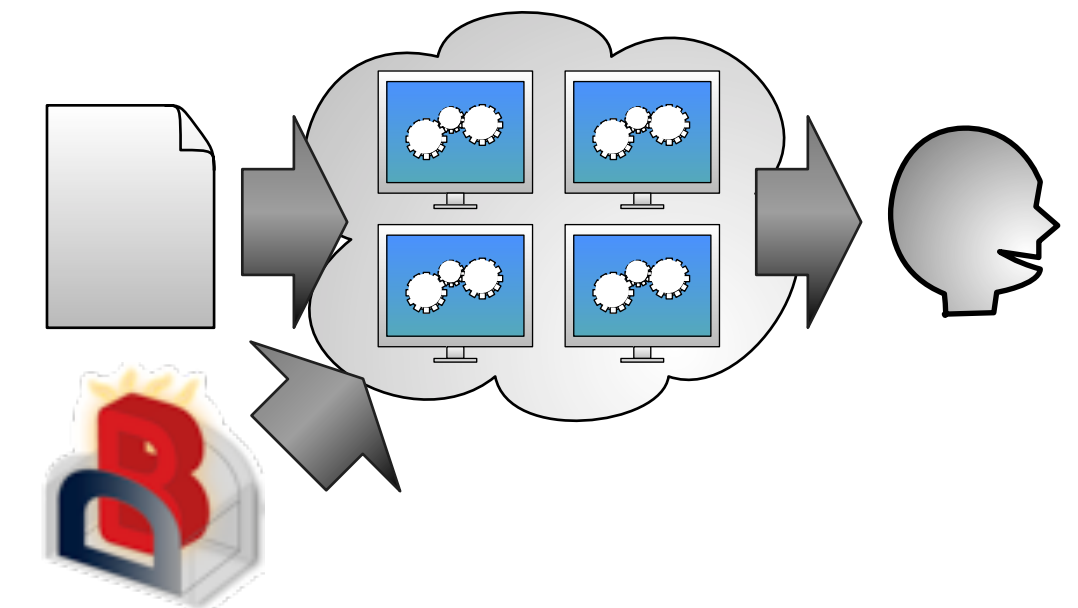
- High Development Costs
- Consistency Limitations
- High Network Overheads

Cloud Batch Processing
(Hadoop, Map/Reduce)

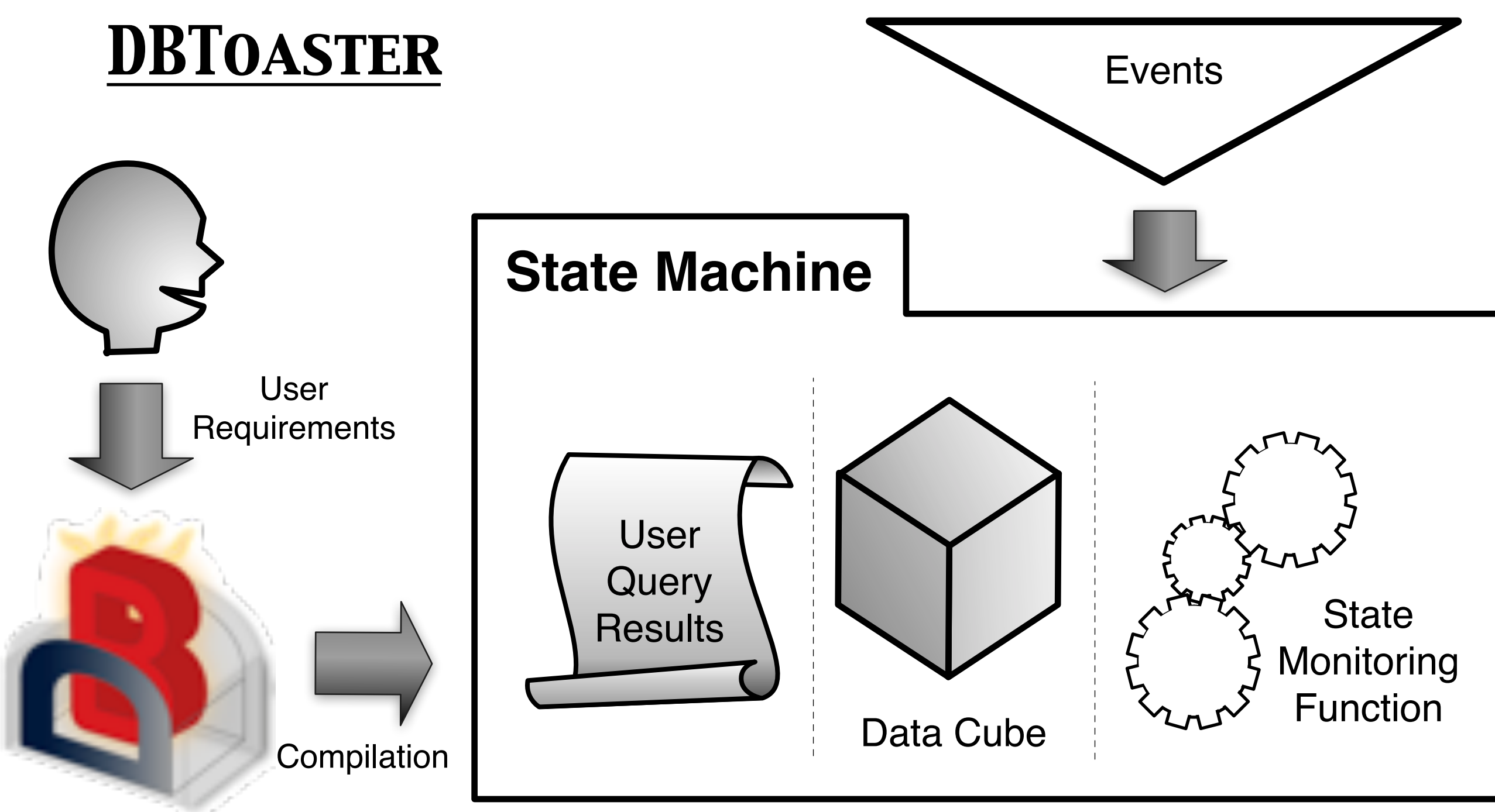


- High Latency
- High Network Overheads

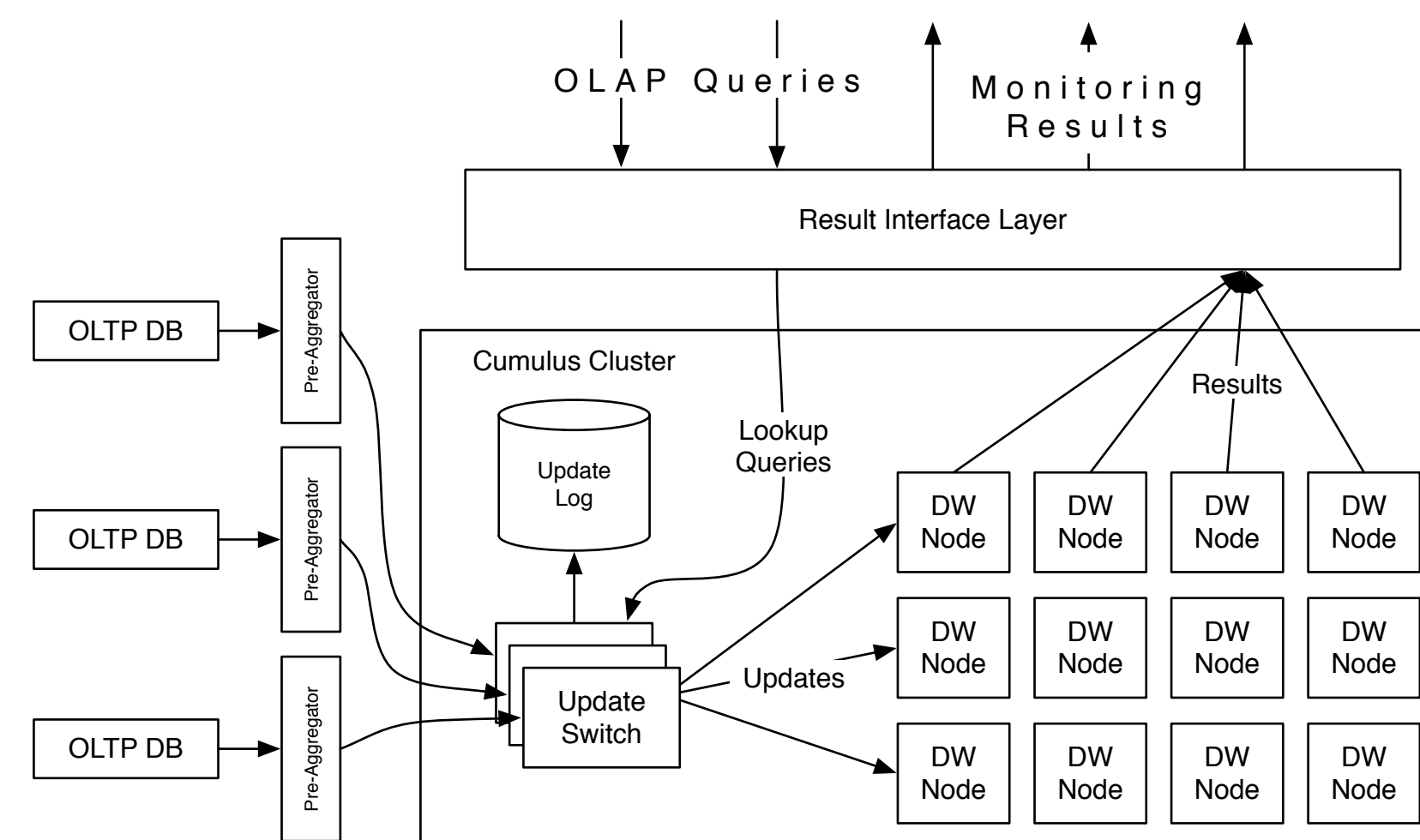
Cumulus



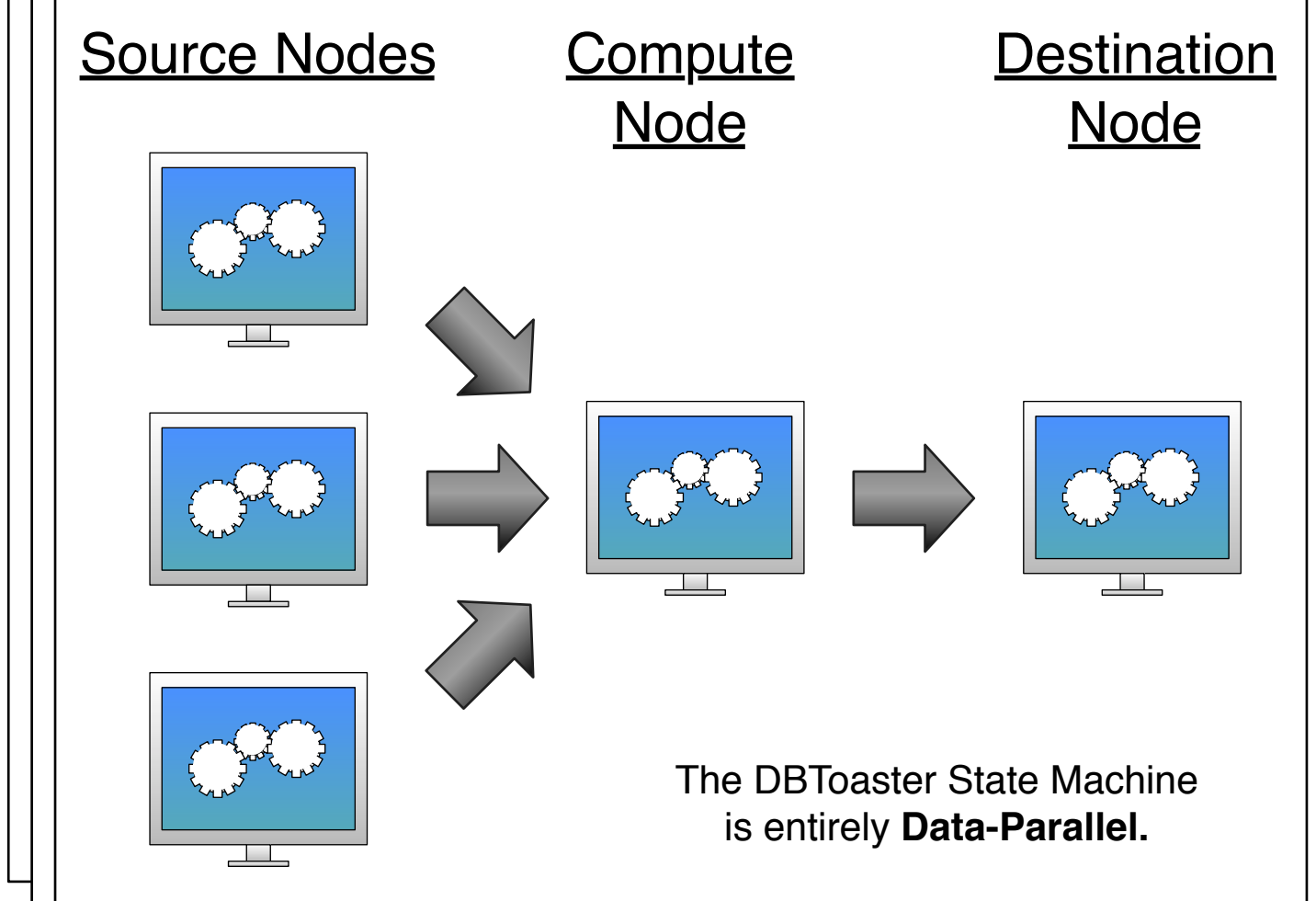
DBTOASTER



CUMULUS ARCHITECTURE

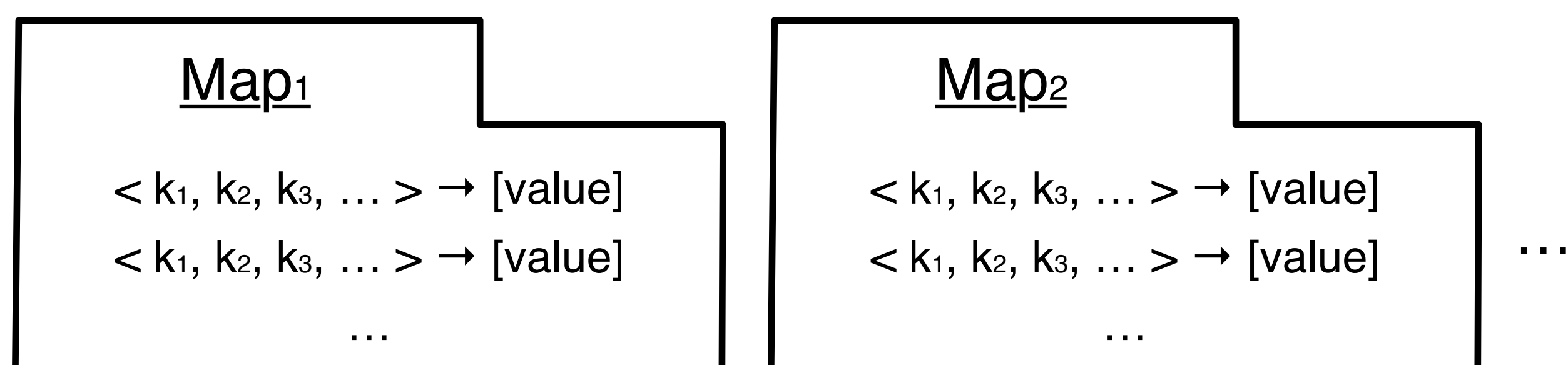


For All k_1, k_2, \dots



THE DBT STATE MACHINE

State: Multikey Maps



Events: Triggers

ON Event(param₁, param₂, ...){ (e.g., ON TABLE INSERT)
statement 1
statement 2
...
}

Statements

for all (k_1, k_2, \dots) {
update Map₁[k_1 , param₁, ...] += f(k , param, Map₂, ...)
}

Distribution Challenge:

Ensure triggers run on consistent view of the state.

CONSISTENCY

State Machine Requirements

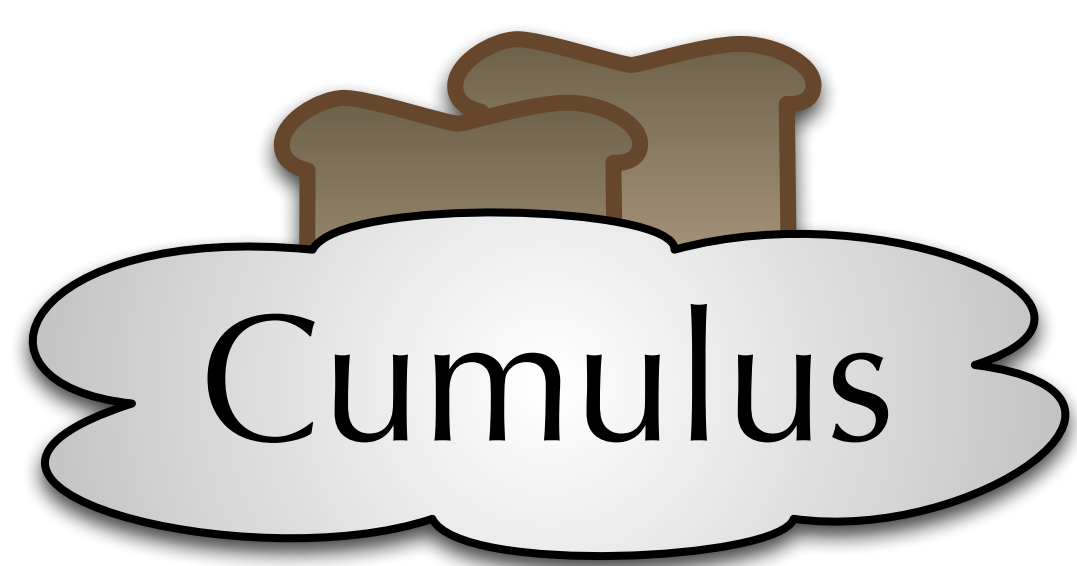
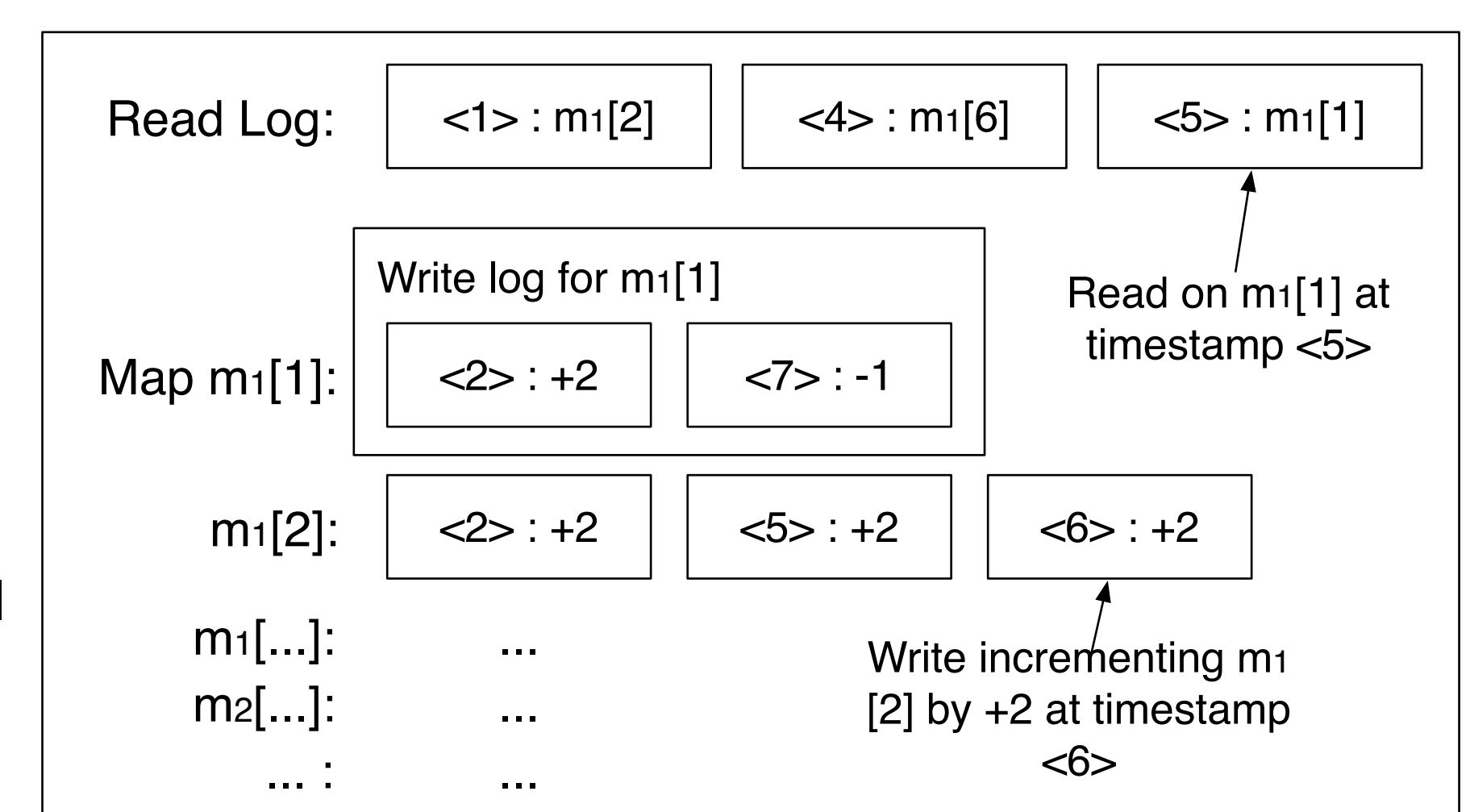
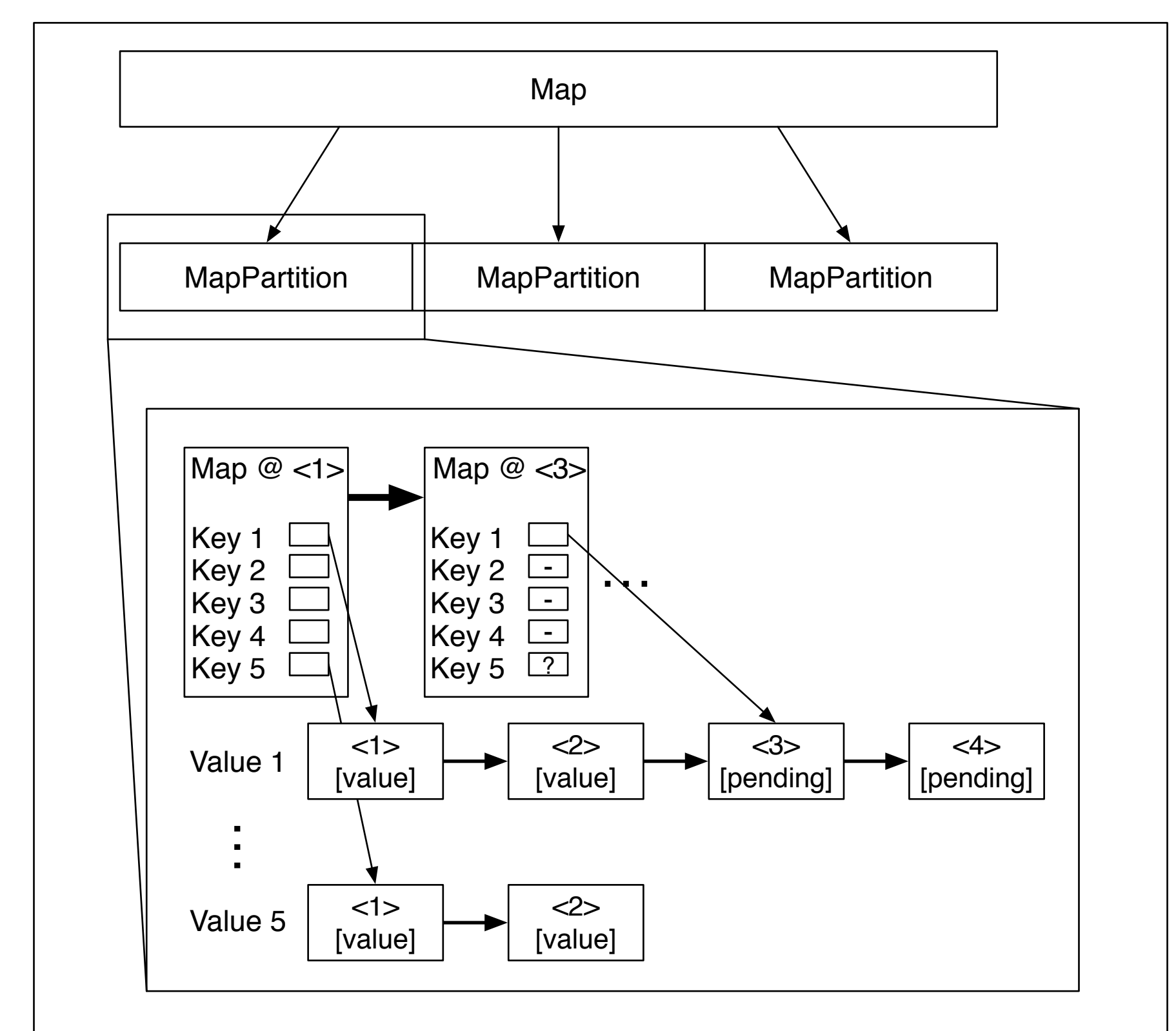
- #1 : Events are atomic.
- #2 : Events must be serialized (in any arbitrary, but consistent order)

Serialized execution requires nodes to know about all events in the system.

Problem: Serialized execution is not scalable!

Solution: Allow out of order execution
(Separate Chronological and Logical clocks)

Order Conflict		Fix
Chronological Order	Logical Order	
Read	Write	Corrective Updates (Read Log)
Write	Read	Versioned Maps (Write Log)
Read	Read	No Correction Needed
Write	Write	Updates are Deltas



Oliver Kennedy
Aleksandar Vitorovic
Christoph Koch
The D.A.T.A. Lab (EPFL)

