ecocloud

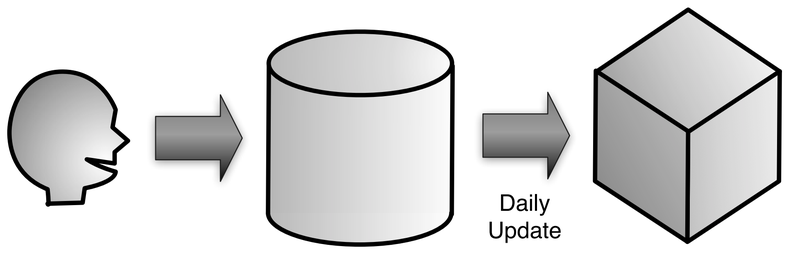
Data Cubes

Widgets

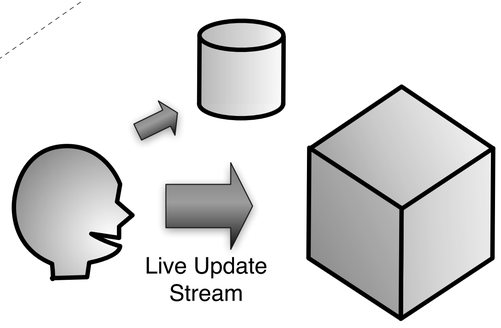
Color	Year	Sales (mln)
Blue	2010	80
Red	2010	70
Blue	2011	100
Red	2011	120

Color	Sales (mln)
Blue	180
Red	190

Year	Sales (mln)
2010	150
2011	220



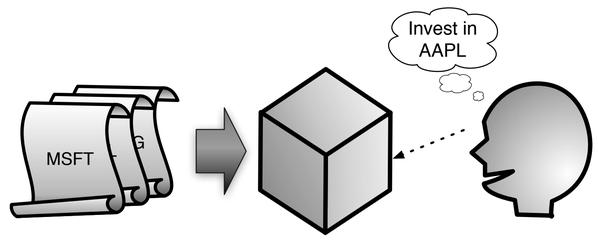
Traditional Data Cube Construction



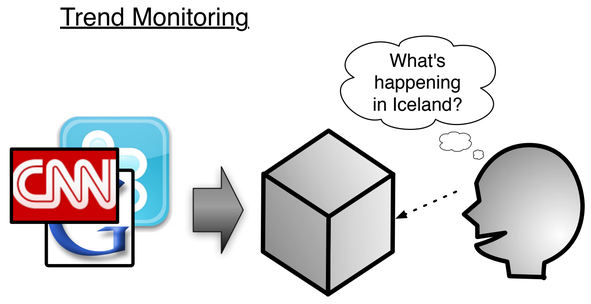
Cumulus Data Cube Construction

Applictions

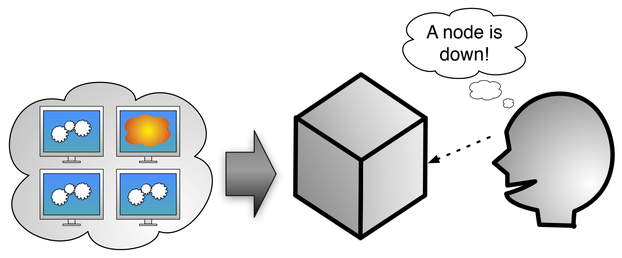
Stock Market Predictions



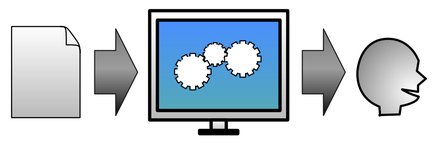
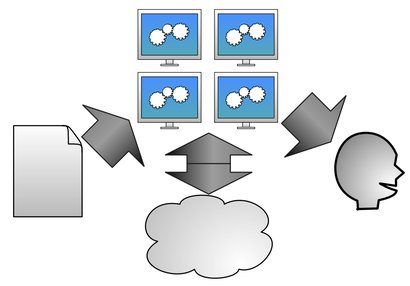
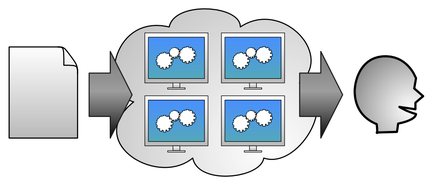
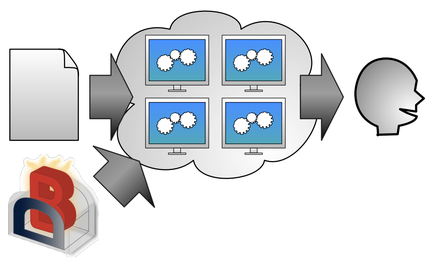
Trend Monitoring



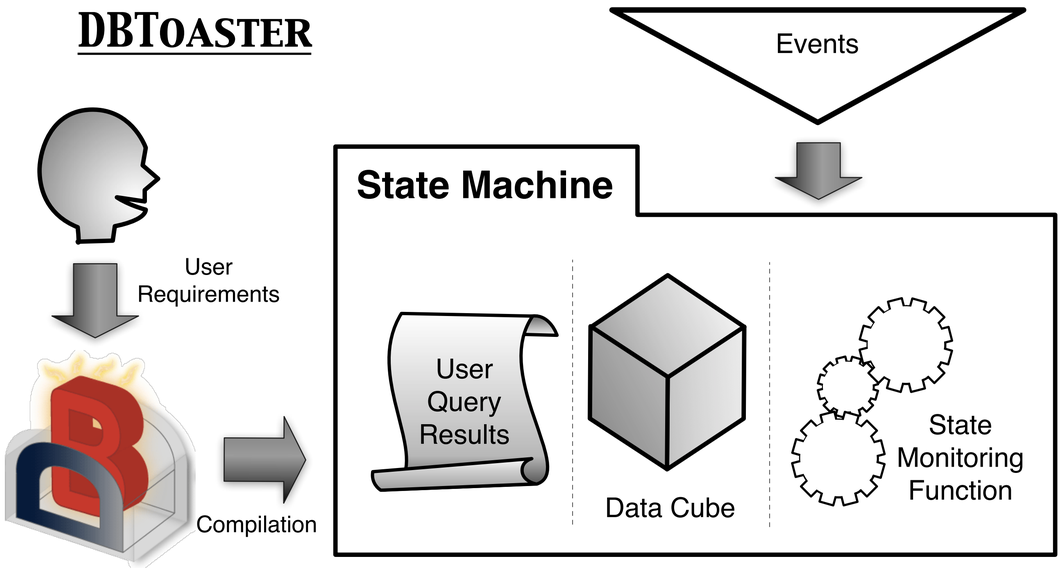
Network Management



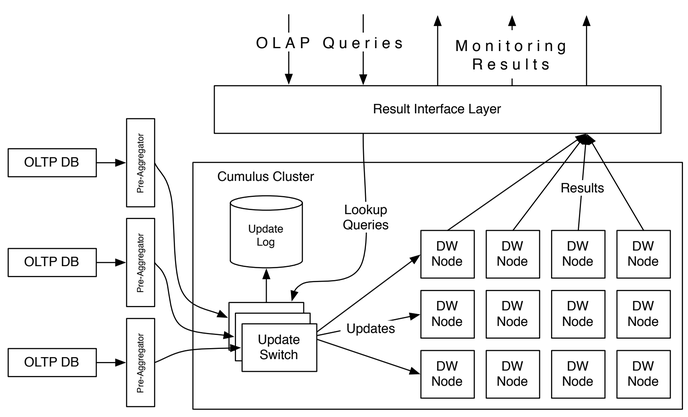
Stream Monitoring

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		

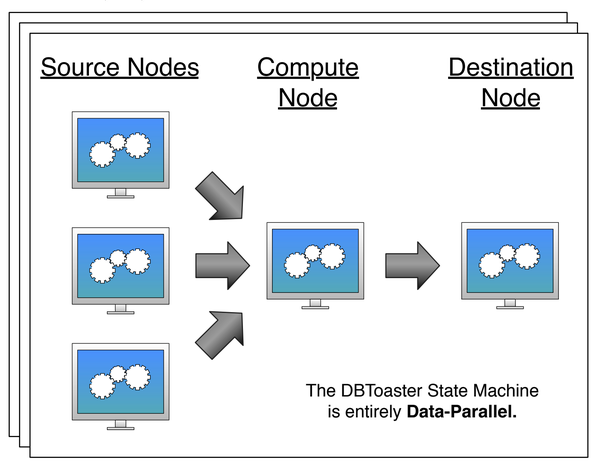
DBToaster



Cumulus Architecture



For All k_1, k_2, …



The DBT State Machine

State: Multikey Maps

Map_1

$< k_1, k_2, k_3, \ldots > \rightarrow$ [value]

$< k_1, k_2, k_3, \ldots > \rightarrow$ [value]

…

Map_2

$< k_1, k_2, k_3, \ldots > \rightarrow$ [value]

$< k_1, k_2, k_3, \ldots > \rightarrow$ [value]

…

…

Events: Triggers

ON Event($param_1$, $param_2$, …){ *(e.g., ON TABLE INSERT)*

statement 1

statement 2

…

}

Distribution Challenge:
Ensure triggers run on consistent view of the state.

Statements

for all (k_1, k_2, …) {

update Map_1[k_1, $param_1$, …] += f(k, param, Map_2, …)

}

Consistency

State Machine Requirements

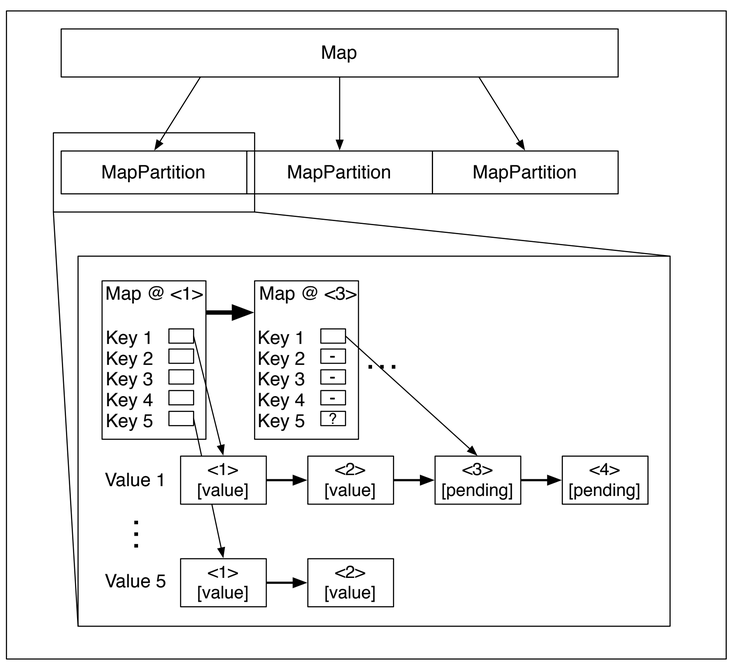
#1 : Events are atomic.

#2 : Events must be serialized (in any arbitrary, but consistent order)

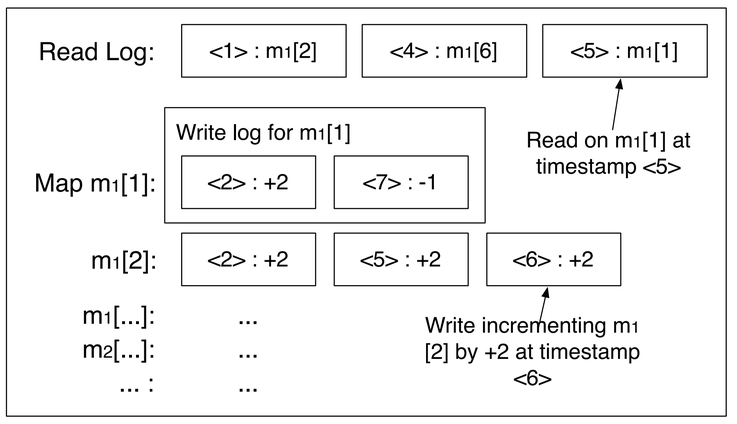
Serialized execution requires nodes to know about all events in the system.

Problem: Serialized execution is not scalable!

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



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



Cumulus

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

TOASTER

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE