



AEM 6 OAK: MongoMK and Queries

Marcel Reutegger & Thomas Mueller Graf | Senior Software Engineers



Marcel Reutegger

Software Engineer

At Day/Adobe since 2002

JCR API Specification

Apache member

Apache Jackrabbit / Oak

Marcel Reutegger

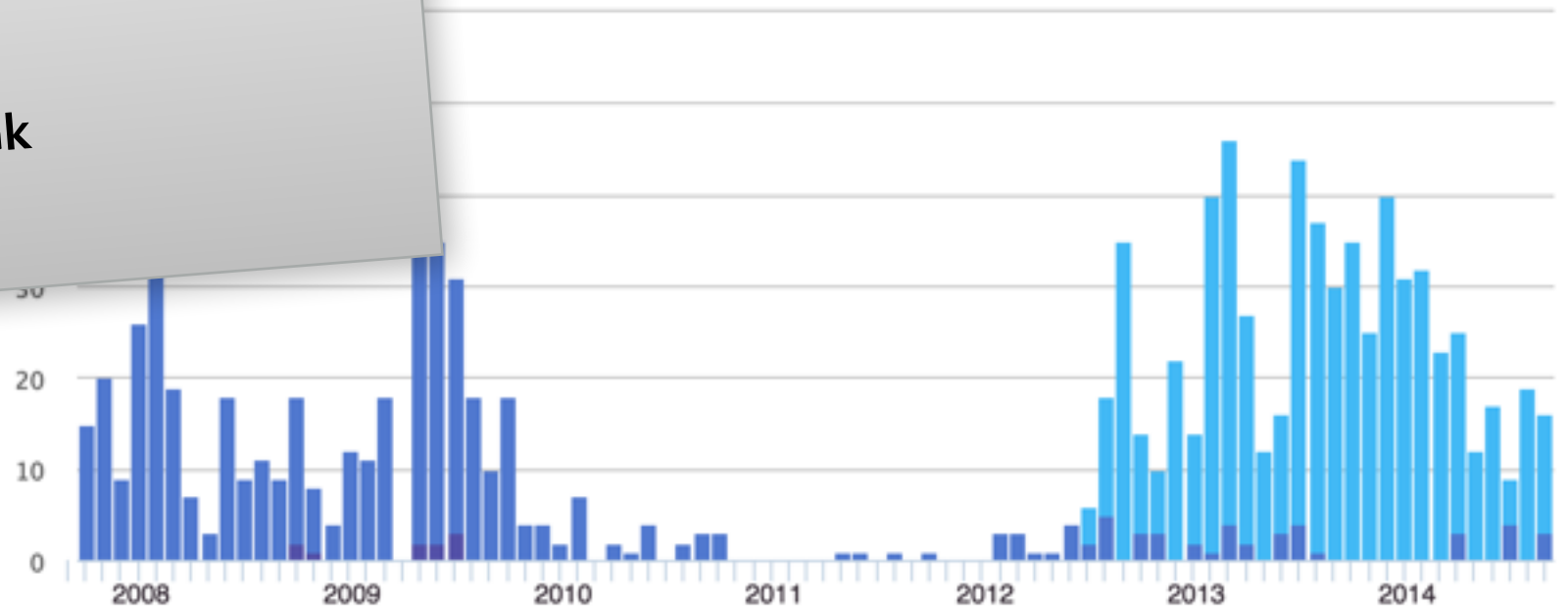
Software Engineer

At Day/Adobe since 2002

JCR API Specification

Apache member

Apache Jackrabbit / Oak



Thomas Mueller

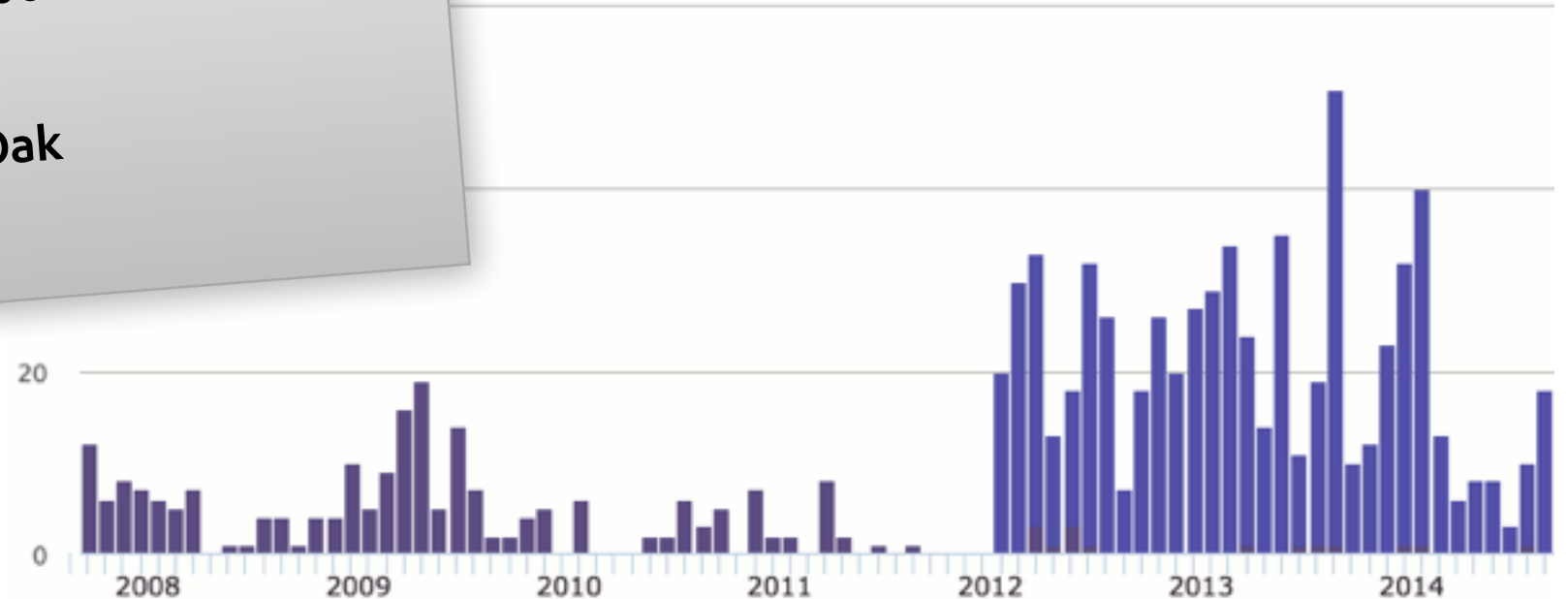
Software Engineer

HypersonicSQL, H2 Database

At Day/Adobe since 2006

Apache member

Apache Jackrabbit / Oak



MongoMK

- Overview & Tools
- Clustering
- Eventual Consistency

Query

- Parsing
- Index Selection
- Execution

NodeStore

MicroKernel

DocumentNodeStore

MongoMK

DocumentMK

NodeStore



MicroKernel

DocumentNodeStore

MongoMK

DocumentMK

NodeStore

MicroKernel

DocumentNodeStore

MongoMK

DocumentMK

Implementation

NodeStore

MicroKernel

Doc

MongoMK

DocumentMK

DocumentNodeStore
or DocumentMK
on MongoDB

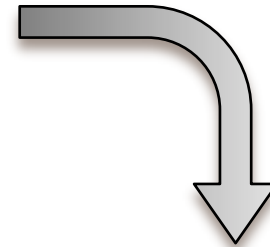
- Prerequisite: MongoDB 2.6.x
- Runmode: -r crx3,crx3mongo
- Optional database name: -Doak.mongo.db=my-author
- Optional URI: -Doak.mongo.uri=mongodb://mongodb.example.org:27017

- nodes: Nodes and Properties
- blobs: DataStore with binary chunks
- clusterNodes: Connected cluster nodes
- settings: Checkpoints

```
{  
  _id : "/home/john",  
  name : "john",  
  email : "john@example.com"  
}
```

MongoMK - Data Model

```
{
  _id : "/home/john",
  name : "john",
  email : "john@example.com"
}
```



```
{
  _id : "2:/home/john",
  name : { "r14979e4b424-0-1" : "john" },
  email : { "r14979e4b424-0-1" : "john@example.com" },
  _deleted : { "r14979e4b424-0-1" : "false" },
  _revisions : { "r14979e4b424-0-1" : "c" }
}
```

```
{  
  _id : "2:/home/john",  
  name : { "r14979e4b424-0-1" : "john" },  
  email : { "r14979e4b424-0-1" : "john@example.com" },  
  _deleted : { "r14979e4b424-0-1" : "false" },  
  _revisions : { "r14979e4b424-0-1" : "c" }  
}
```

```
{
  _id : "2:/home/john",
  name : { "r14979e4b424-0-1" : "john" },
  email : { "r14979e4b424-0-1" : "john@example.com" },
  _deleted : { "r14979e4b424-0-1" : "false" },
  _revisions : { "r14979e4b424-0-1" : "c" }
}
```

Timestamp

Counter

Cluster ID

```
{
  _id : "2:/home/john",
  name : { "r14979e4b424-0-1" : "john" },
  email : {
    "r14979e4b424-0-1" : "john@example.com",
    "r14979e6a941-0-1" : "john.doe@example.com"
  },
  _deleted : { "r14979e4b424-0-1" : "false" },
  _revisions : {
    "r14979e4b424-0-1" : "c",
    "r14979e6a941-0-1" : "c"
  }
}
```

Change email to "john.doe@example.com"

- Mongo shell
- `wget http://s.apache.org/oak-mongo.js`
- `mongo --shell oak-mongo.js`
- `java -jar oak-run-1.1.2.jar console mongodb://localhost/aem-author`

Diagnostic information

```
> oak.indexStats()  
[  
  {  
    "count": 3106,  
    "size": 1031874,  
    "simple": "1007.69 kB",  
    "id": "2:/oak:index/acPrincipalName"  
  },  
  ...  
]
```

```
> oak.systemStats()  
{  
  "nodeStats" : {  
    "ns" : "aem-author.nodes",  
    "count" : 469694,  
    "size" : 418,  
    "avgObjSize" : 933,  
    "storageSize" : 439,  
    "numExtents" : 15,  
    "nindexes" : 5,  
    "lastExtentSize" : 119,  
    ...  
  }  
}
```

```
> oak.formatRevision('r1499901f1f8-0-1')  
r1499901f1f8-0-1 (Mon Nov 10 2014 10:21:21 GMT+0100 (CET))
```

MongoMK - oak-run console

```
> java -jar oak-run-1.1.2.jar console mongodb://localhost/aem-author
Apache Jackrabbit Oak 1.1.2
Jackrabbit Oak Shell (Apache Jackrabbit Oak 1.1.2, JVM: 1.7.0_67)
Type ':help' or ':h' for help.
-----
/> cd home
/home> ls
groups
rep:policy
users
/home> print-node
{ jcr:primaryType = rep:AuthorizableFolder, jcr:mixinTypes =
[rep:AccessControllable], groups = { ... }, rep:policy = { ... }, users = { ... }
}
/home>
```

MongoMK

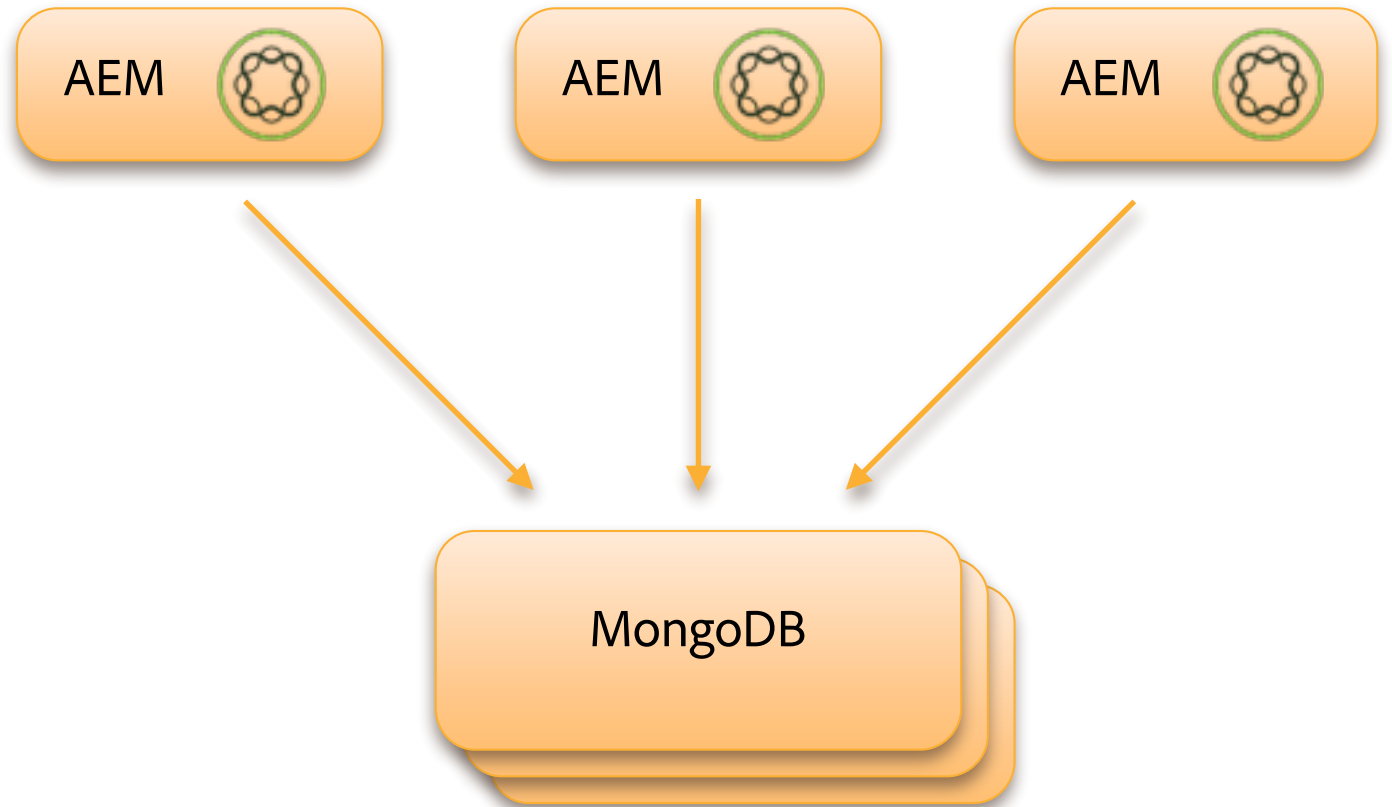
- Overview & Tools
- **Clustering**
- Eventual Consistency

Query

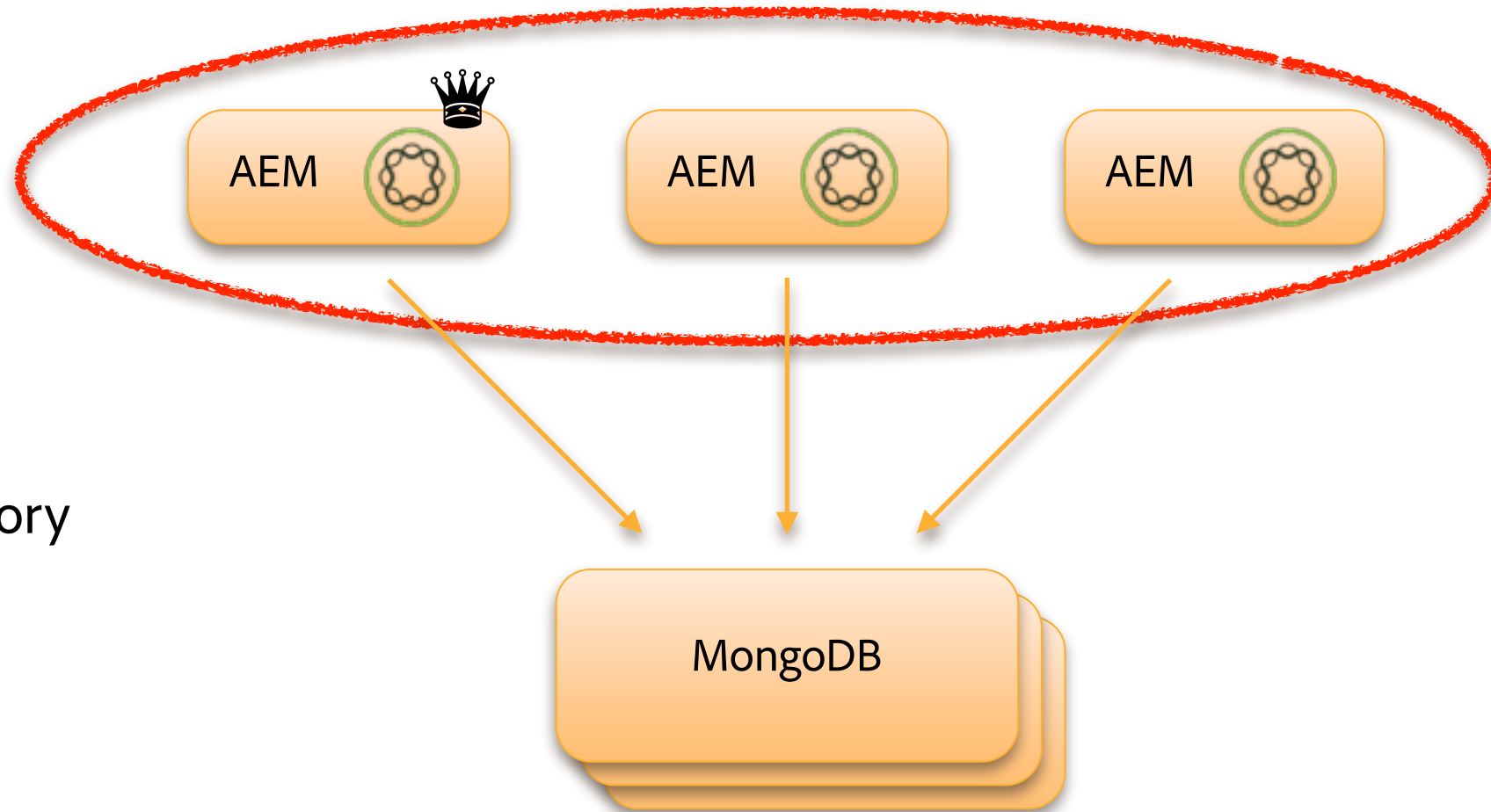
- Parsing
- Index Selection
- Execution

MongoMK - Clustering

- No direct communication
- All cluster nodes are equal
- All cluster nodes can write
- Conflict detection on commit



- Sling Discovery
- Topology with one leader
- Heartbeats through repository



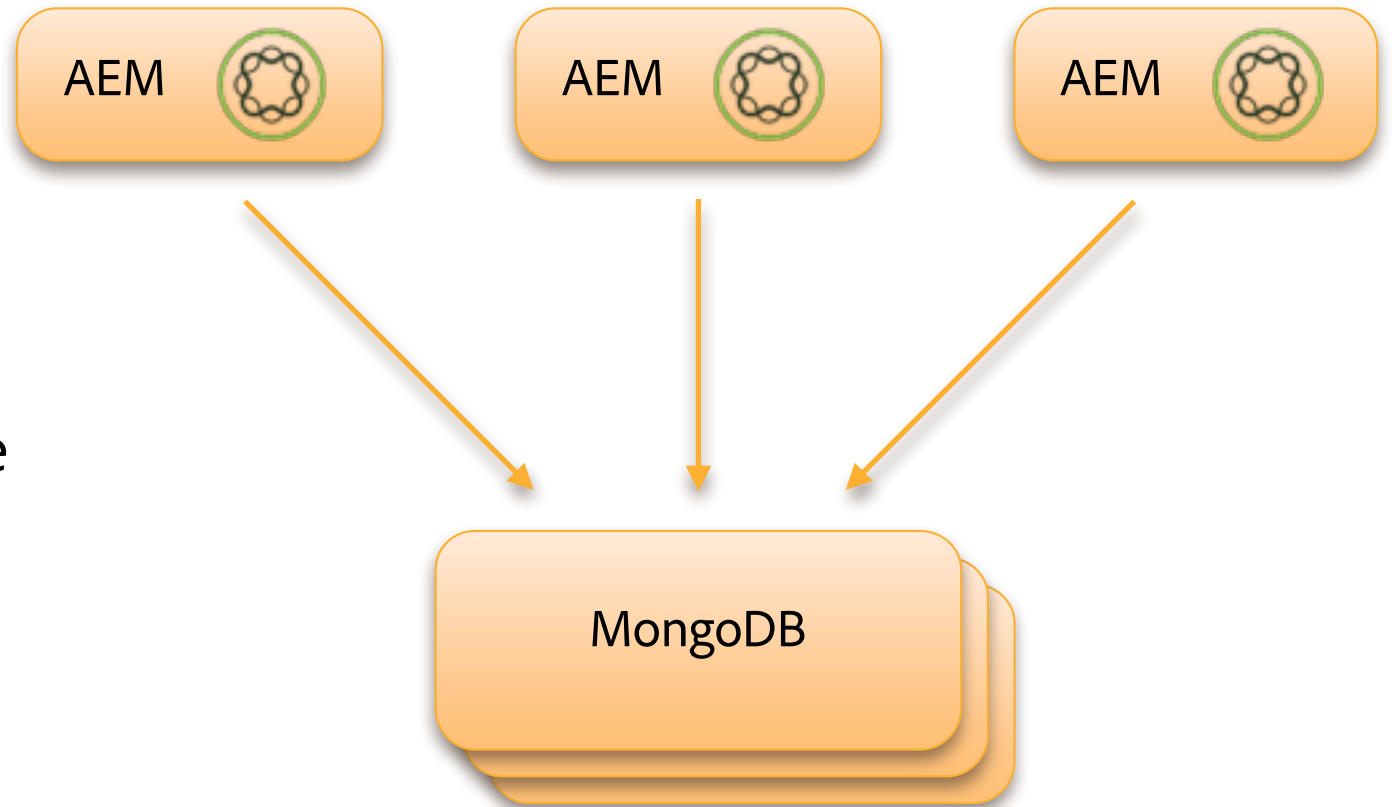
MongoMK

- Overview & Tools
- Clustering
- **Eventual Consistency**

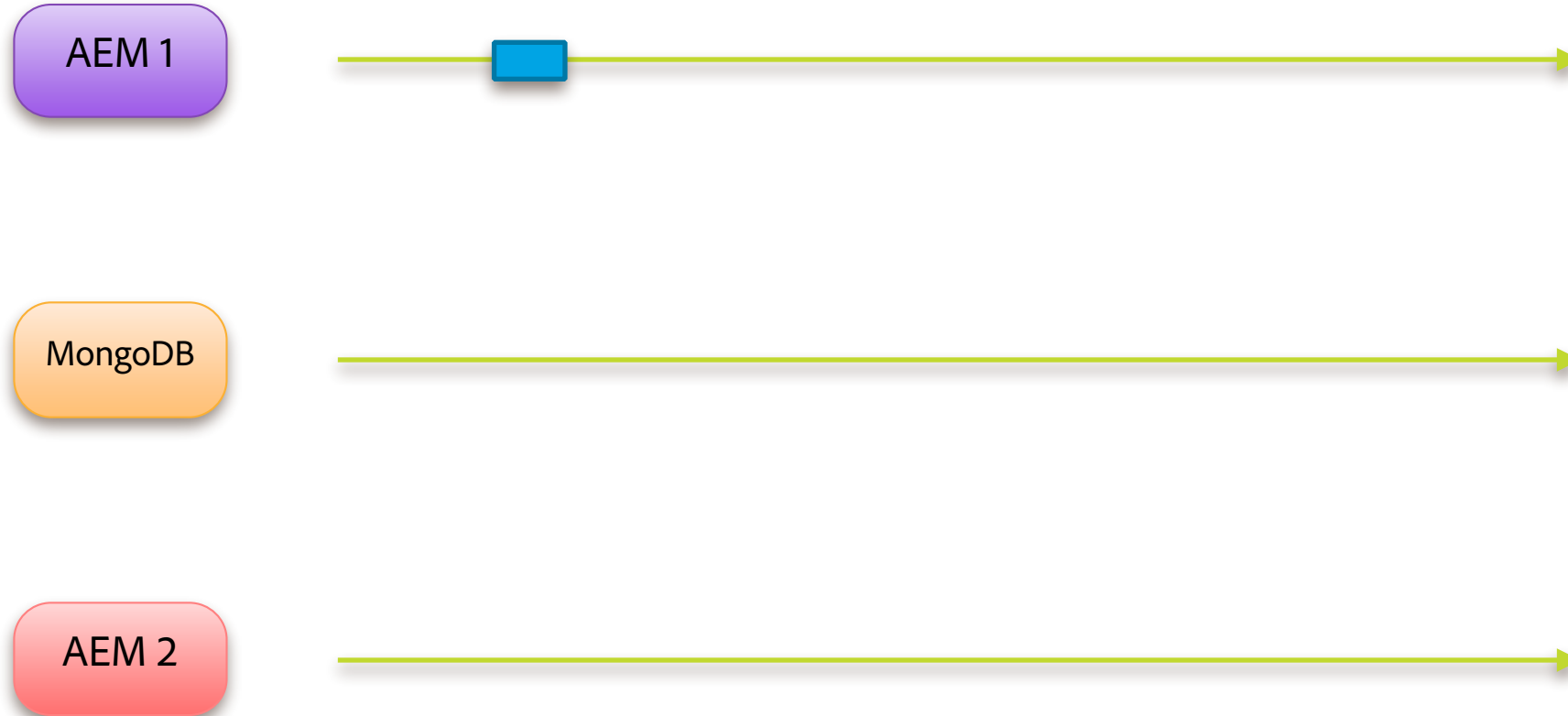
Query

- Parsing
- Index Selection
- Execution

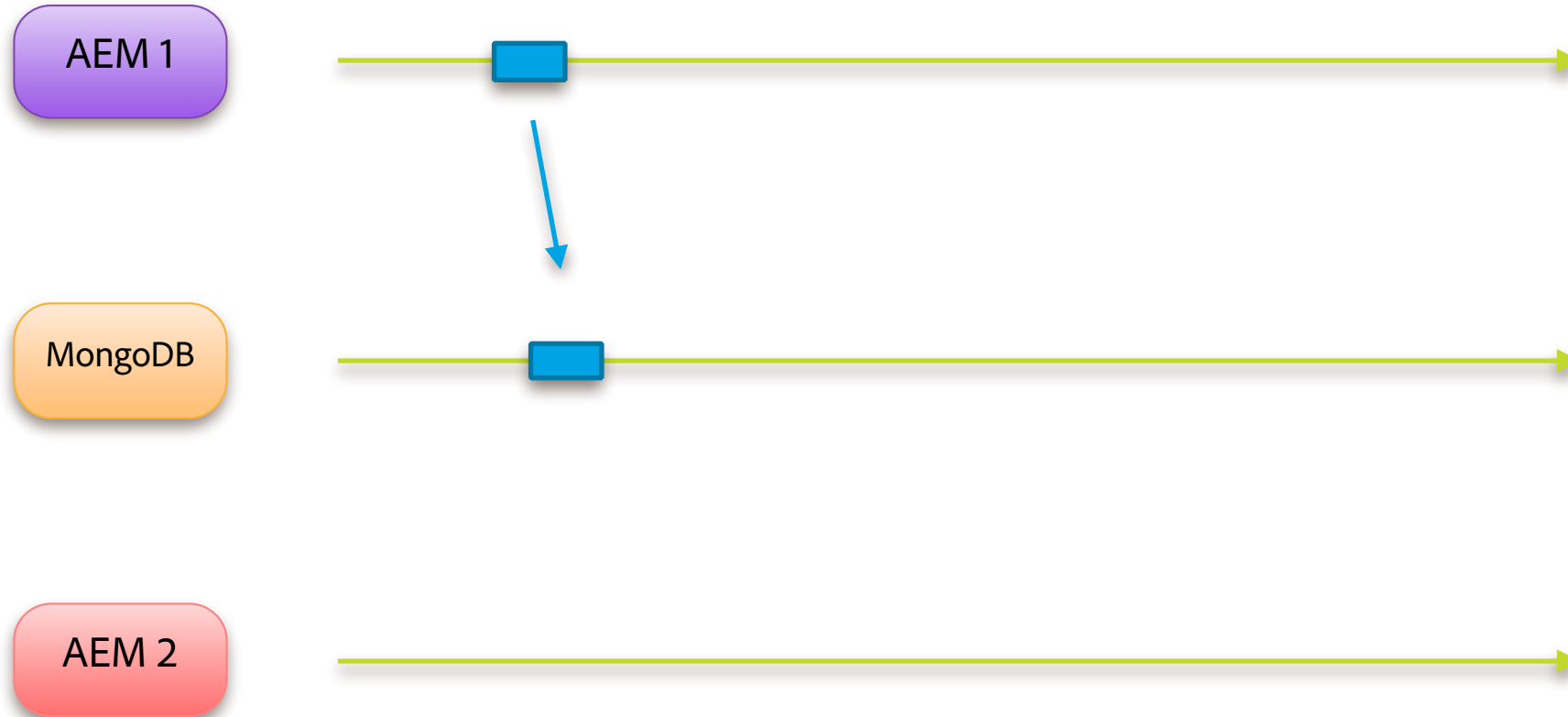
- Changes are not immediately visible on other cluster nodes
- Delay ~ 1 second
- Sequence of events is not the same on all cluster nodes



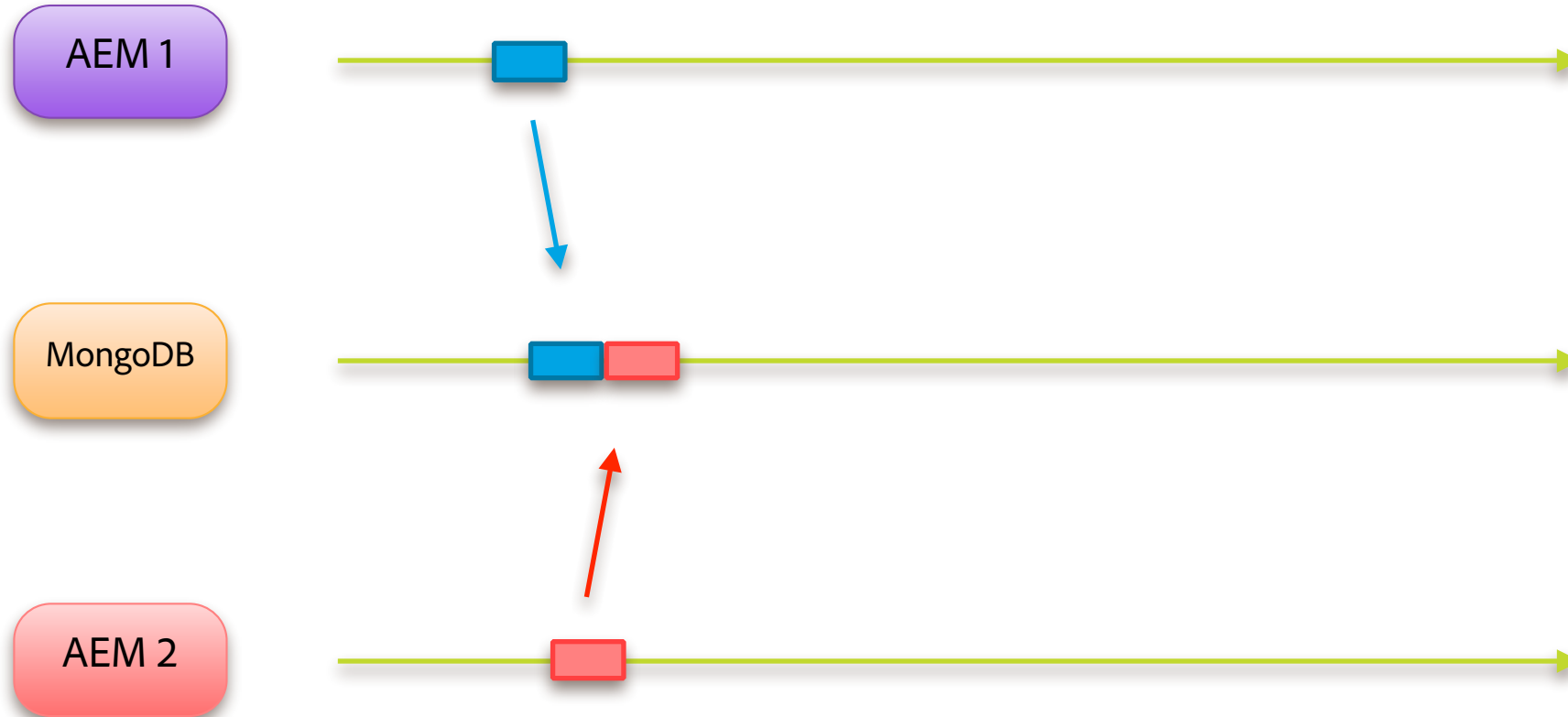
MongoMK - Eventual Consistency



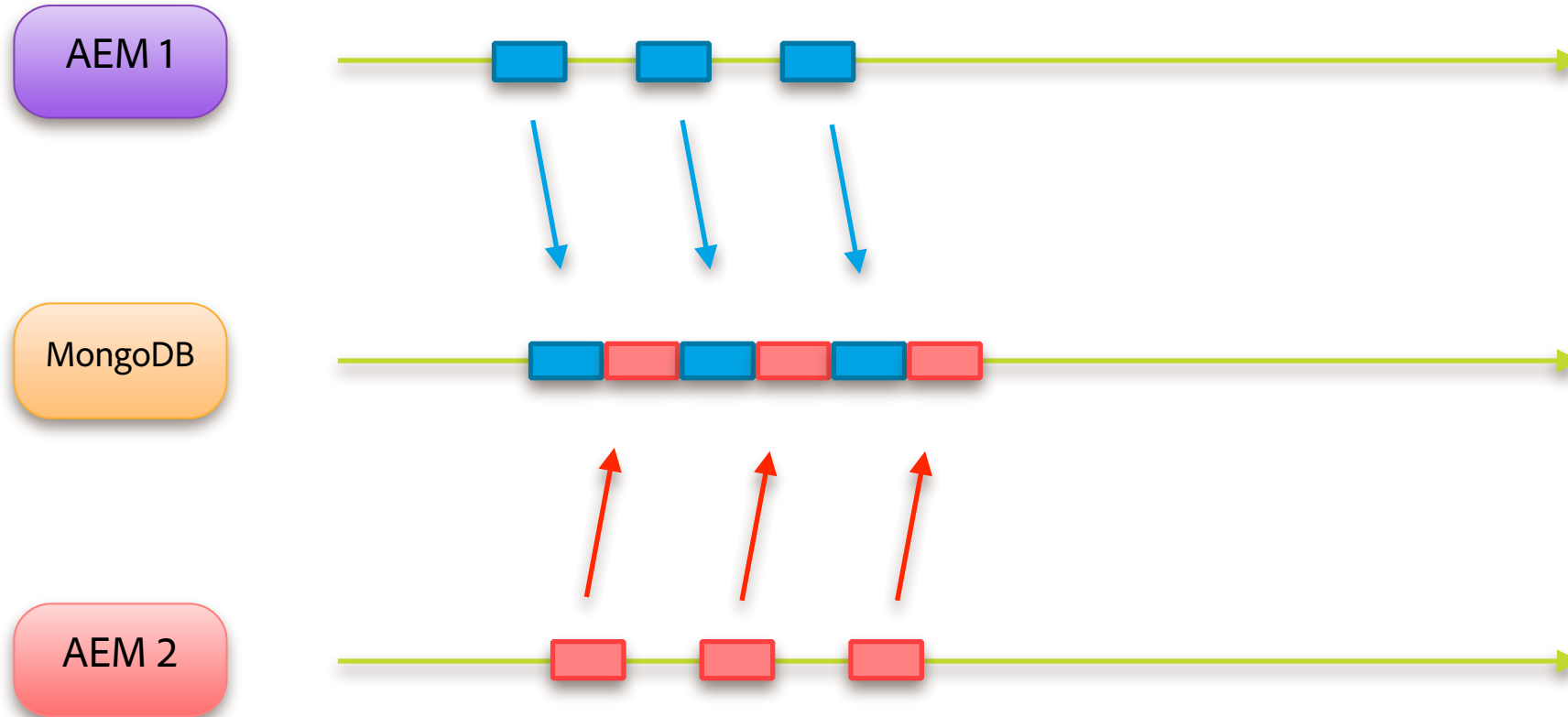
MongoMK - Eventual Consistency



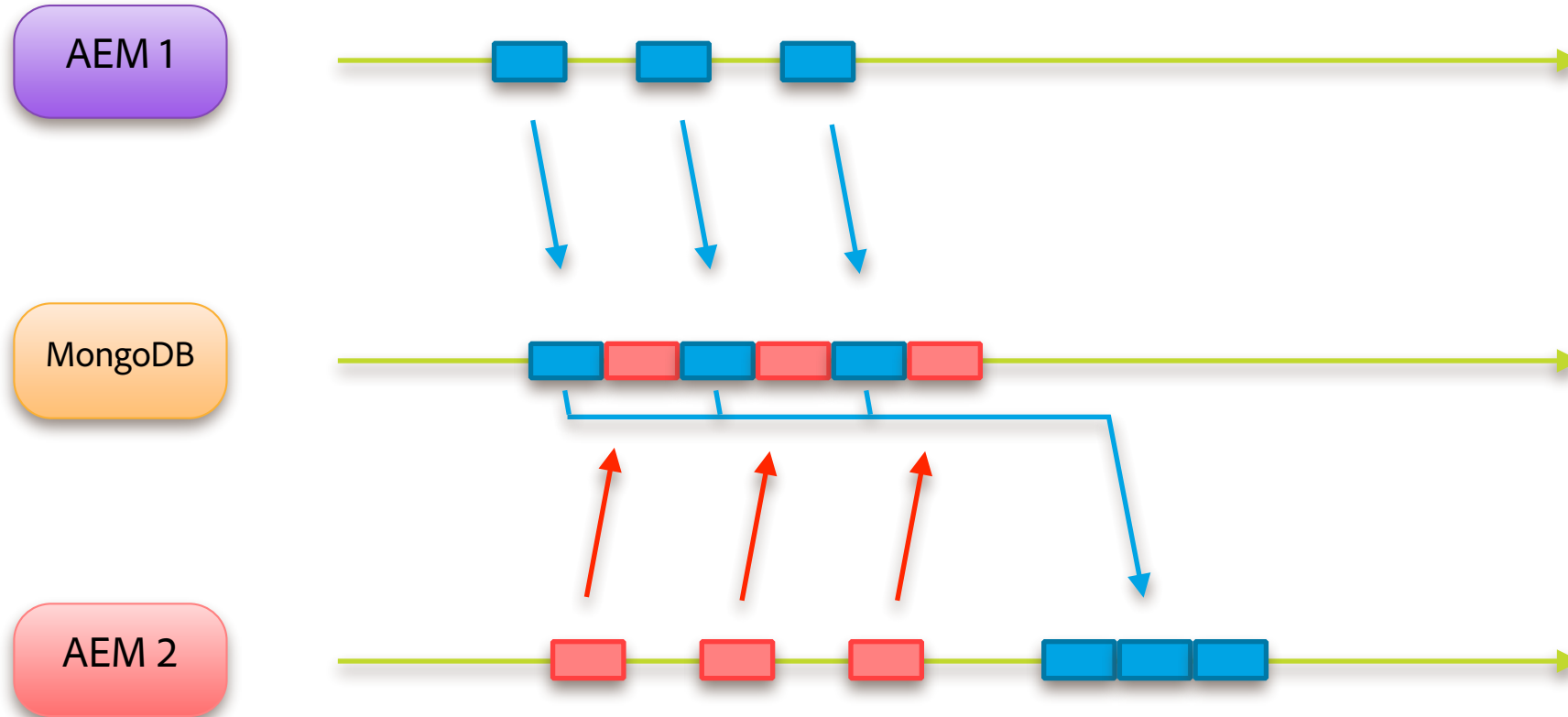
MongoMK - Eventual Consistency



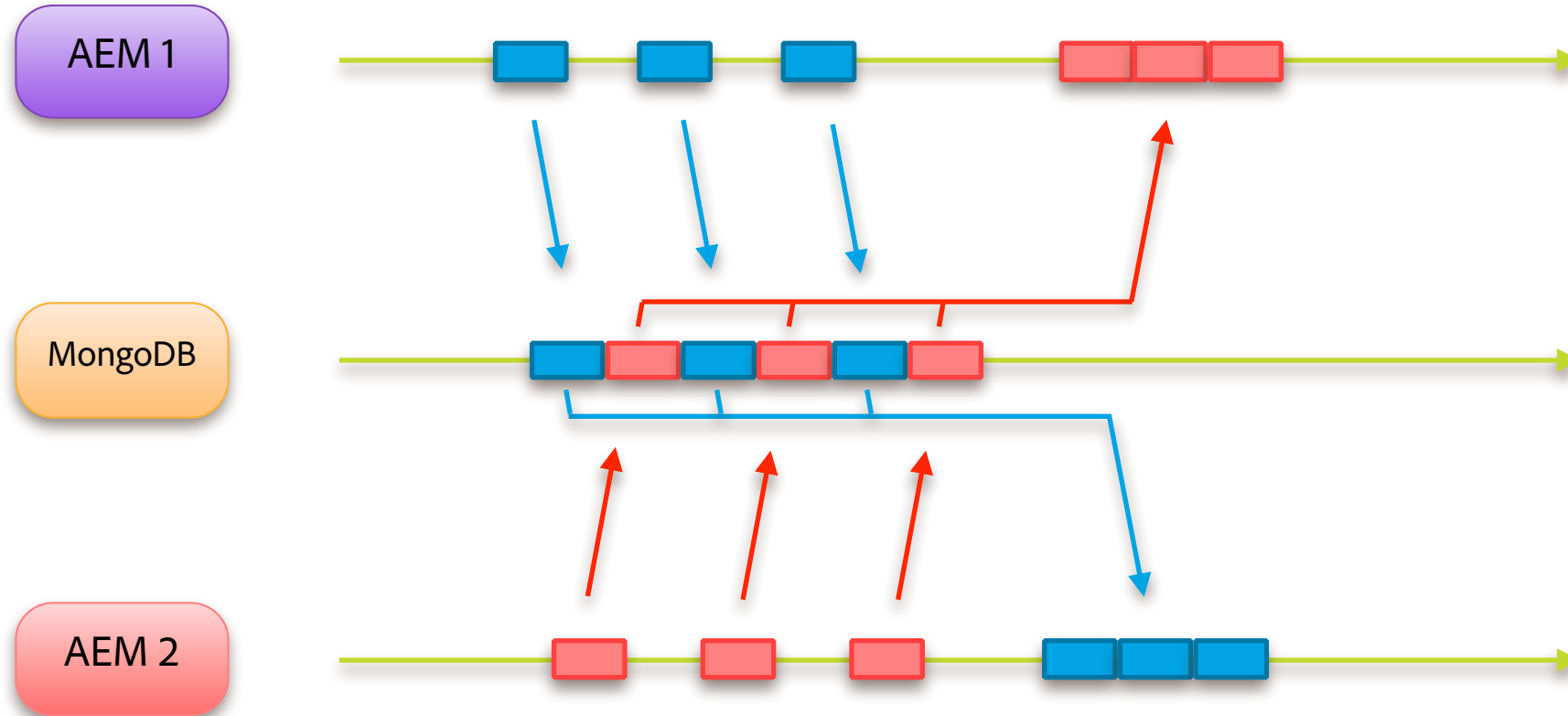
MongoMK - Eventual Consistency



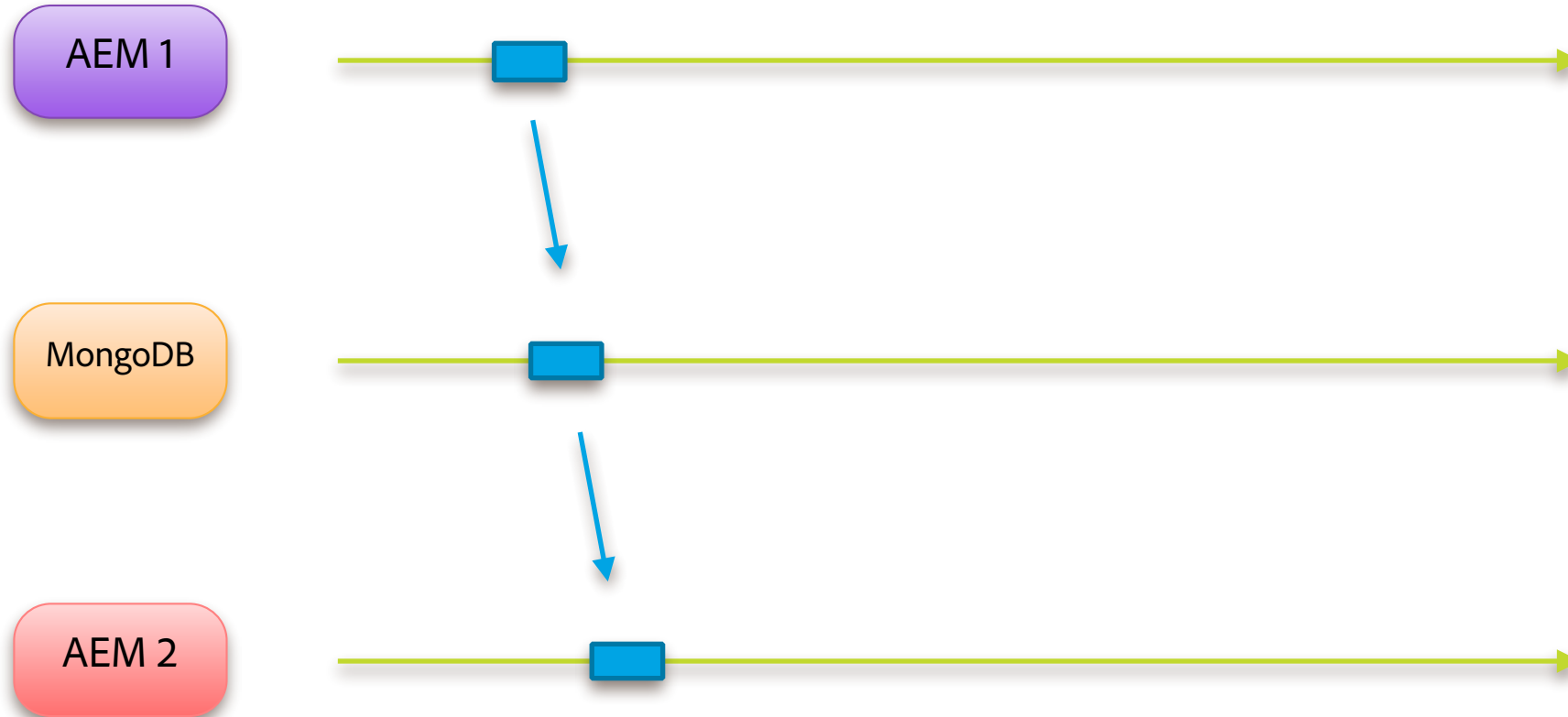
MongoMK - Eventual Consistency



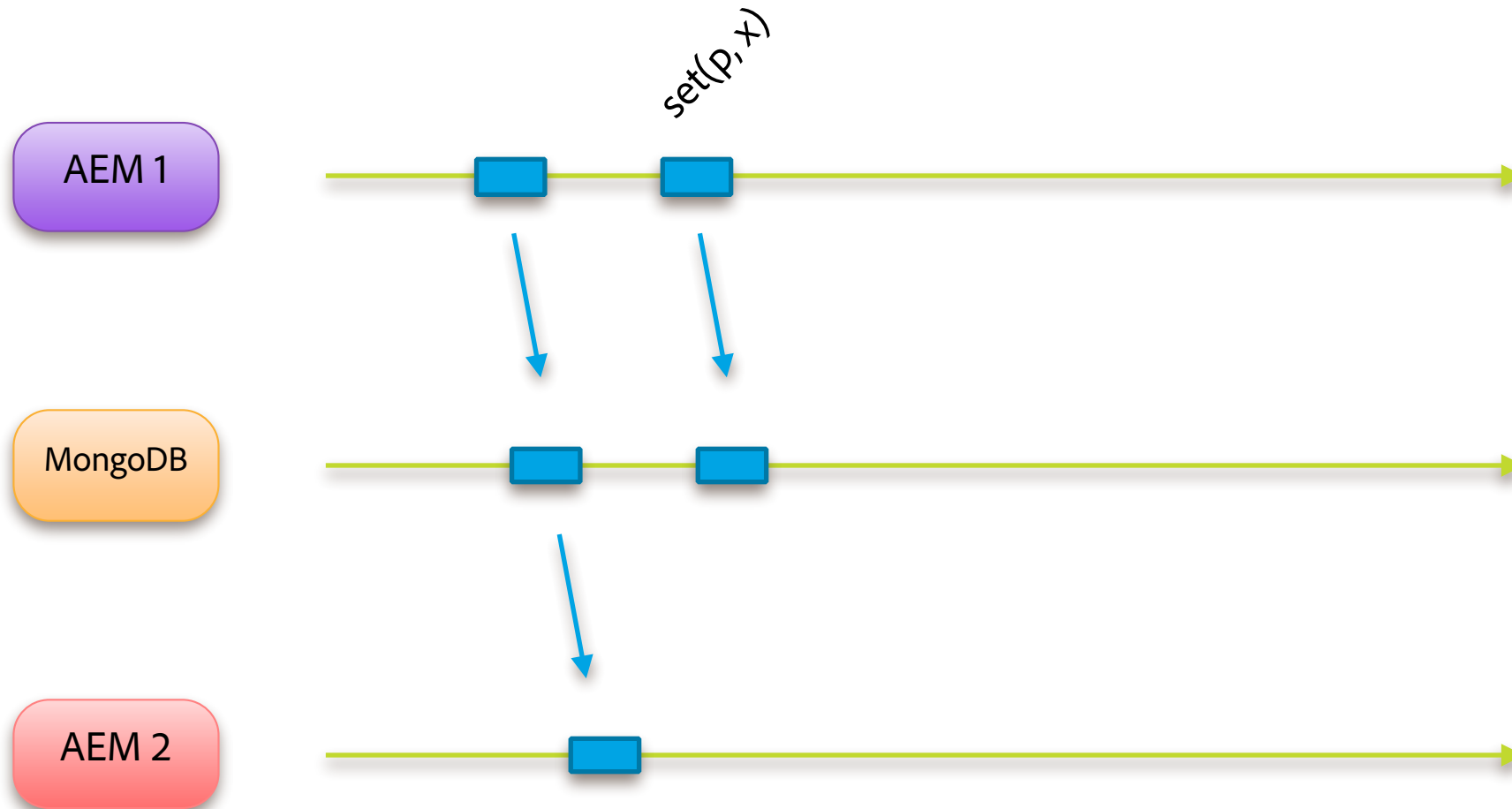
MongoMK - Eventual Consistency



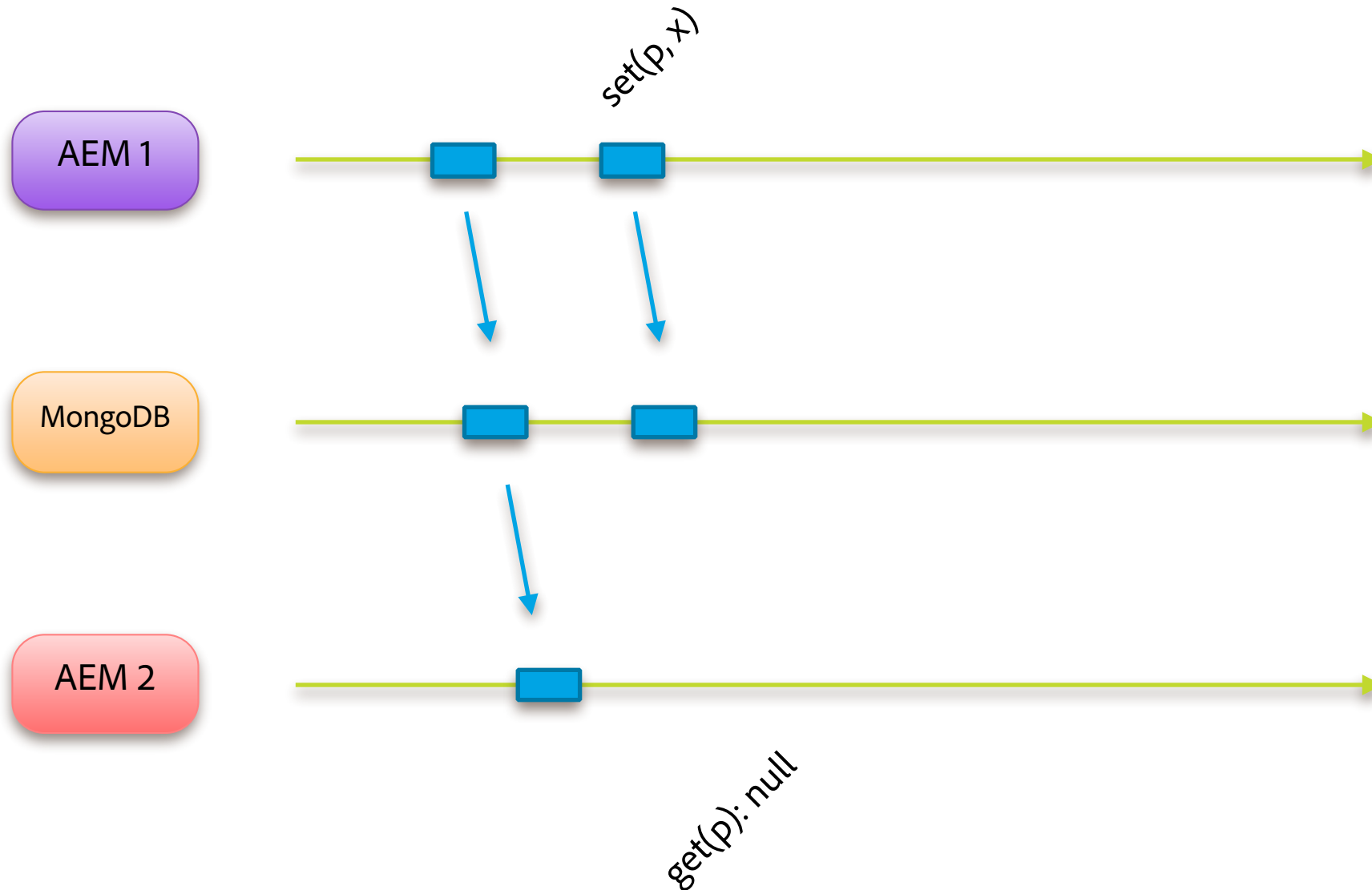
MongoMK - Conflicts



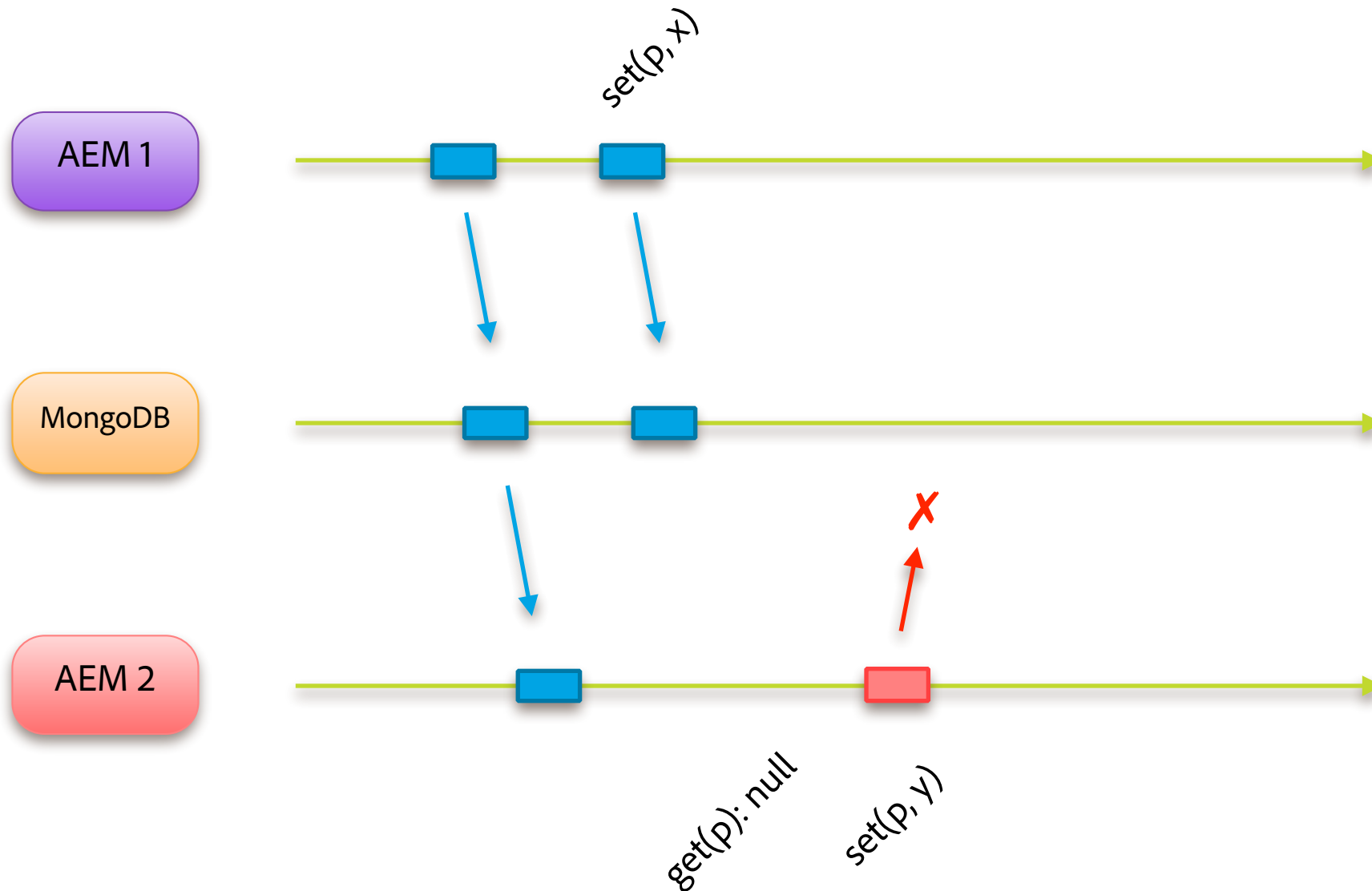
MongoMK - Conflicts



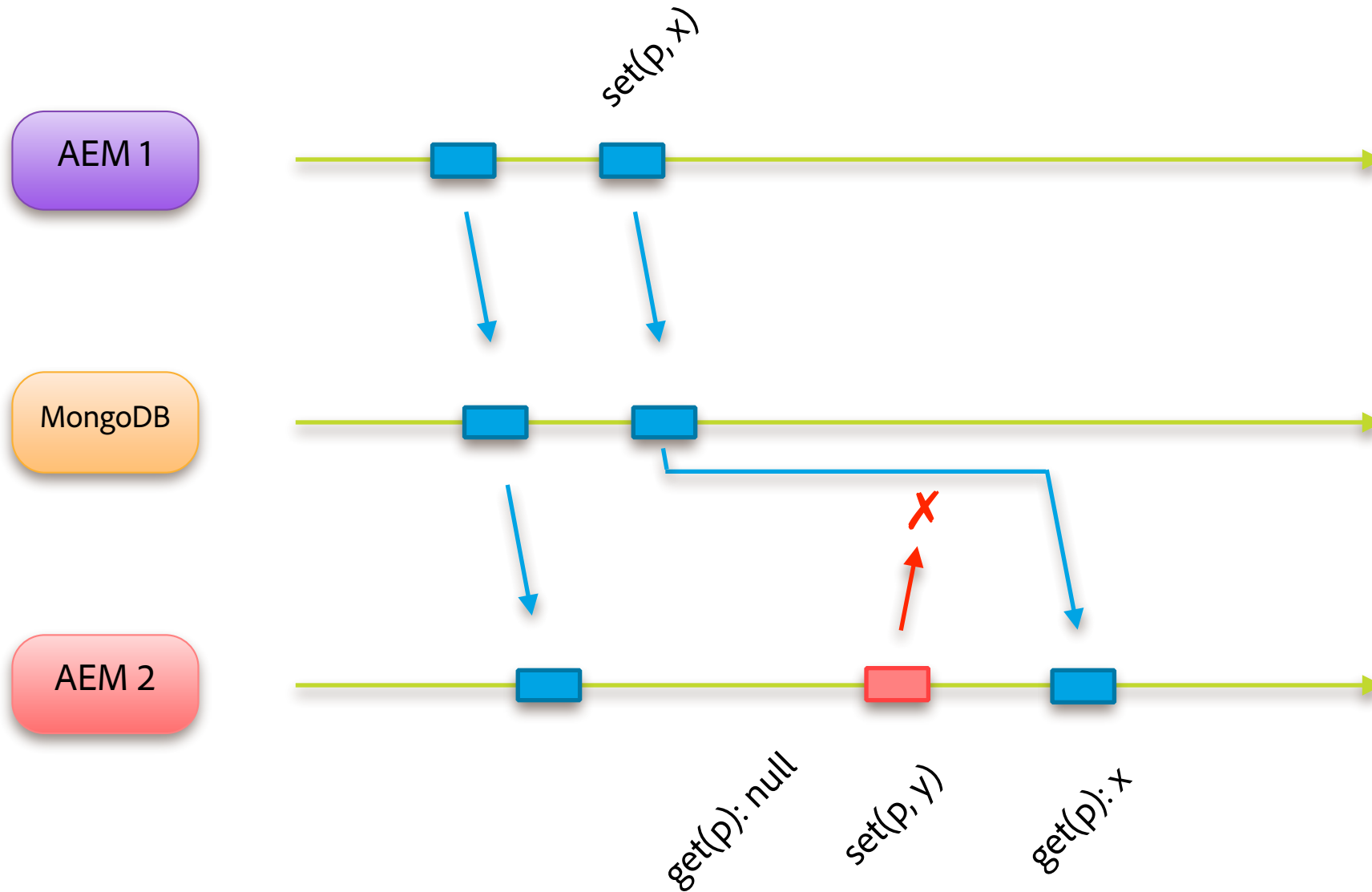
MongoMK - Conflicts



MongoMK - Conflicts



MongoMK - Conflicts



MongoMK

- Overview & Tools
- Clustering
- Eventual Consistency

Query

- Parsing
- Index Selection
- Execution

Jackrabbit 2.x: Apache Lucene

- one index
- hard to extend or change
- limited scalability



Jackrabbit Oak: Pluggable Indexes

- 5 index types,
including Apache Solr
- 50+ index definitions with AEM
- cost based optimizer



MongoMK

- Overview & Tools
- Clustering
- Eventual Consistency

Query

- **Parsing**
- Index Selection
- Execution

XPath Query



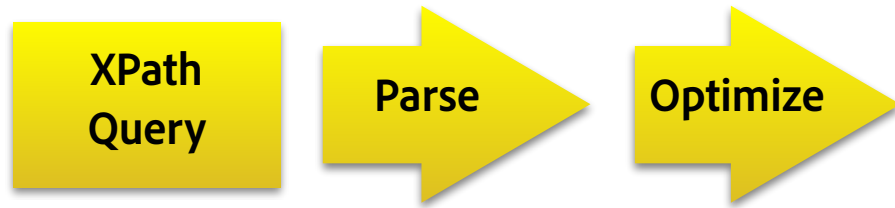
Oak Query: Parsing

**XPath
Query**

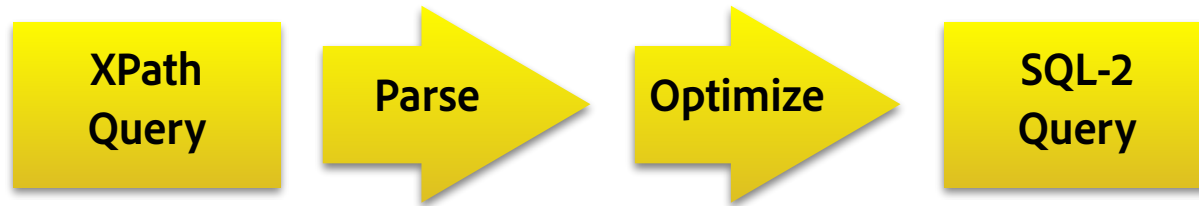
Parse



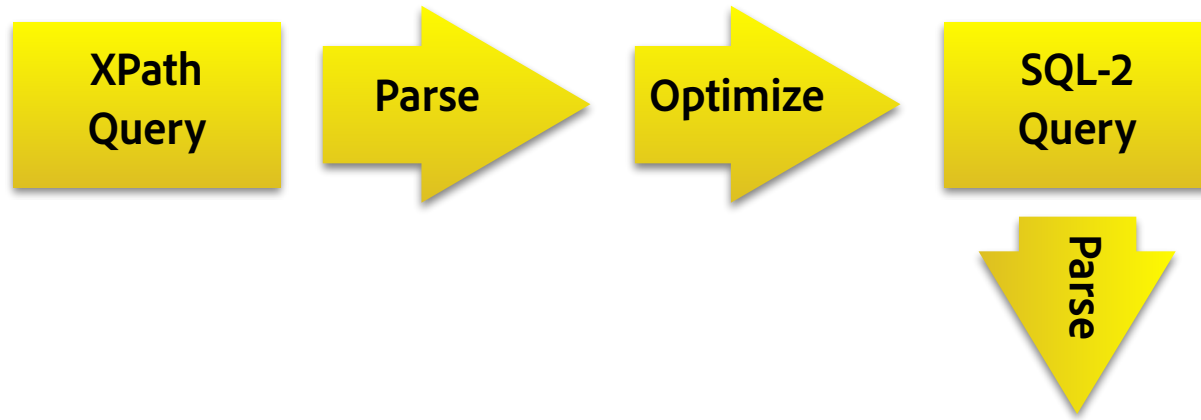
Oak Query: Parsing



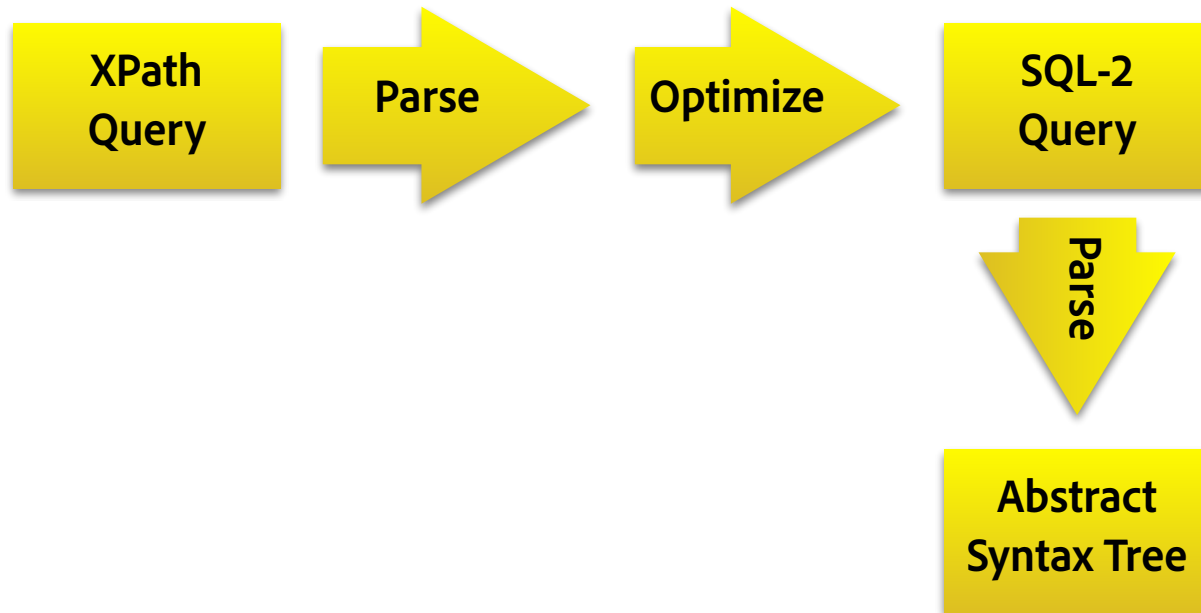
Oak Query: Parsing



Oak Query: Parsing



Oak Query: Parsing



Query Languages:

XPath

**XML Path Language
Defined by W3C**

SQL

**For Databases
Oak: Low Level**

Query Languages:

XPath

```
/jcr:root//*  
[@size='L']
```

SQL

```
select * from  
[nt:base]  
where size = 'L'
```

Query Languages:

XPath

```
/jcr:root/content//images/*  
[@size='L']
```

SQL

```
select b.*  
from [nt:base] as a  
inner join [nt:base] as b  
on ischildnode(b, a)  
where name(a) = 'images'  
and isdescendantnode(a, '/  
content')  
and b.[size] = 'L'
```


Query Languages:

XPath

```
/jcr:root/content//images/*  
[@size='L']
```



preferred

SQL

```
select b.*  
from [nt:base] as a  
inner join [nt:base] as b  
on ischildnode(b, a)  
where name(a) = 'images'  
and isdescendantnode(a, '/  
content')  
and b.[size] = 'L'
```

XPath

```
//*[@a = 'x' or @a = 'y']
```

SQL

```
select [jcr:path],  
[jcr:score], *  
from [nt:base] as a  
where [a] in('x', 'y')
```

XPath

```
//*[@a = 'x' or @b = 'y']
```

SQL

```
select [jcr:path],  
[jcr:score], *  
from [nt:base] as a  
where [a] = 'x'  
union select [jcr:path],  
[jcr:score], *  
from [nt:base] as a  
where [b] = 'y'
```

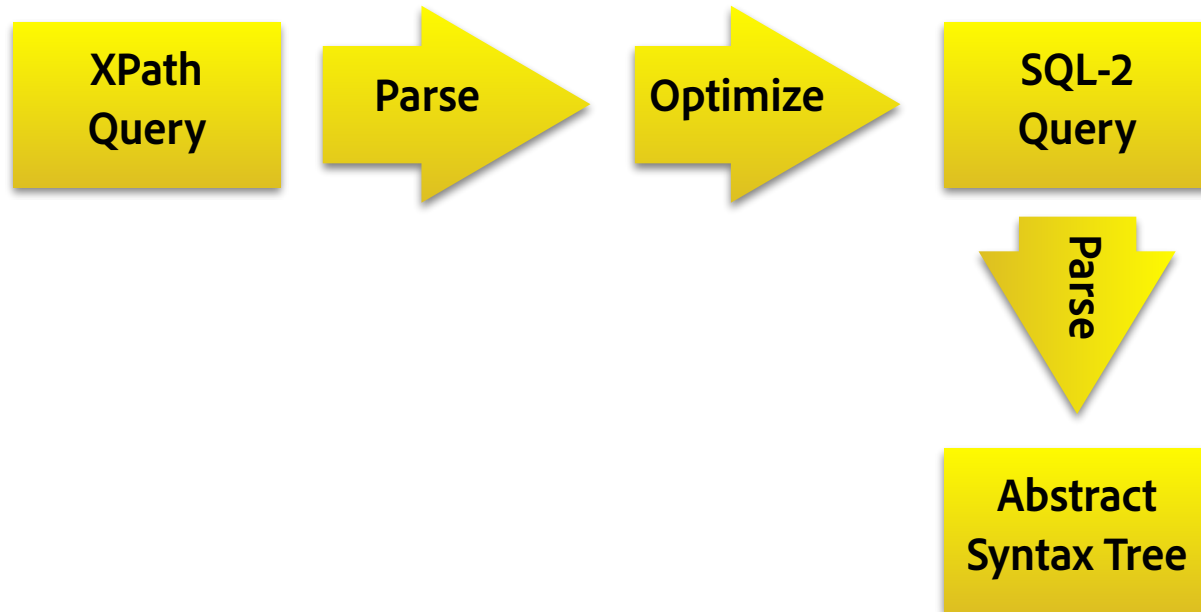
MongoMK

- Overview & Tools
- Clustering
- Eventual Consistency

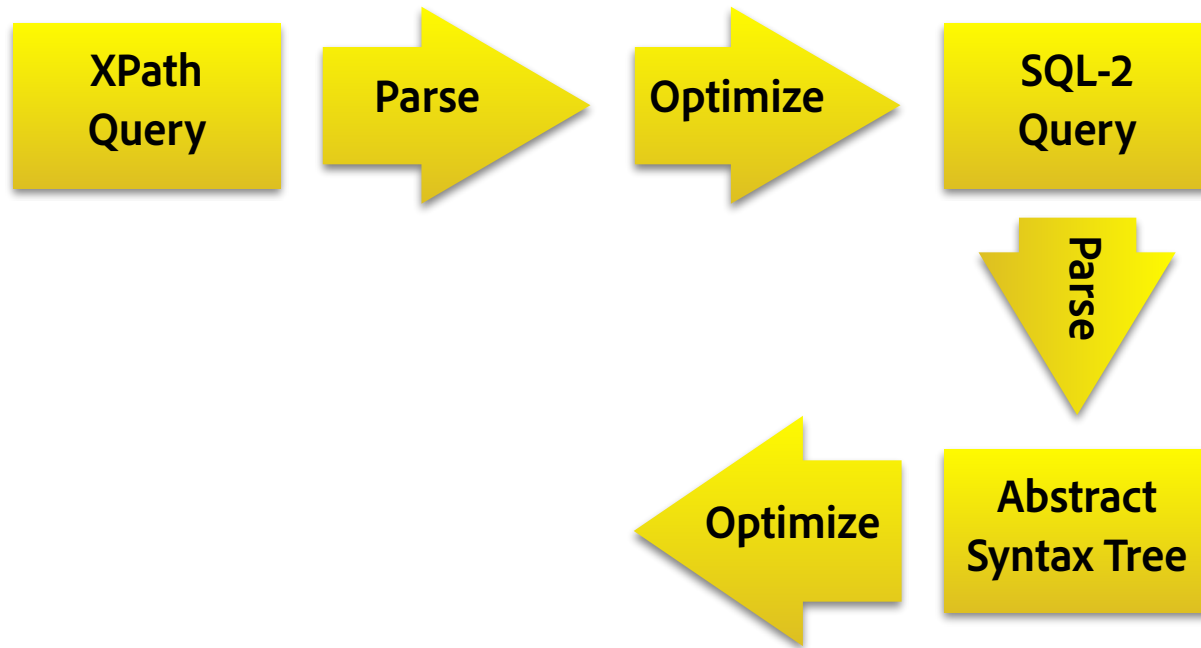
Query

- Parsing
- **Index Selection**
- Execution

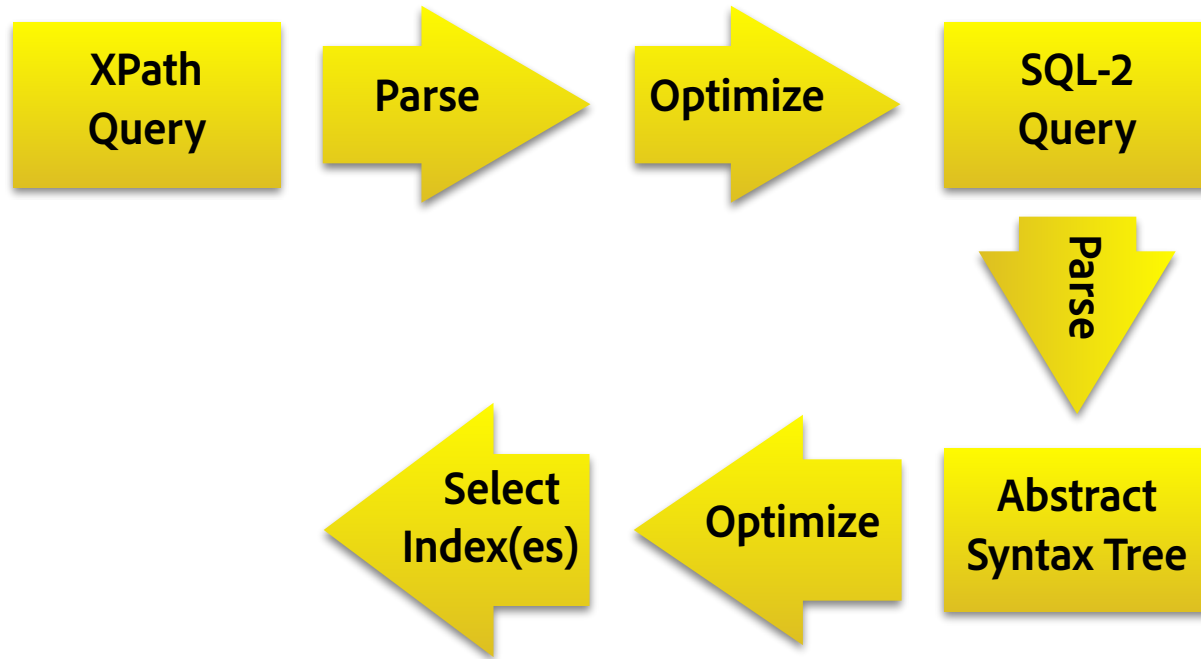
Oak Query: Index Selection



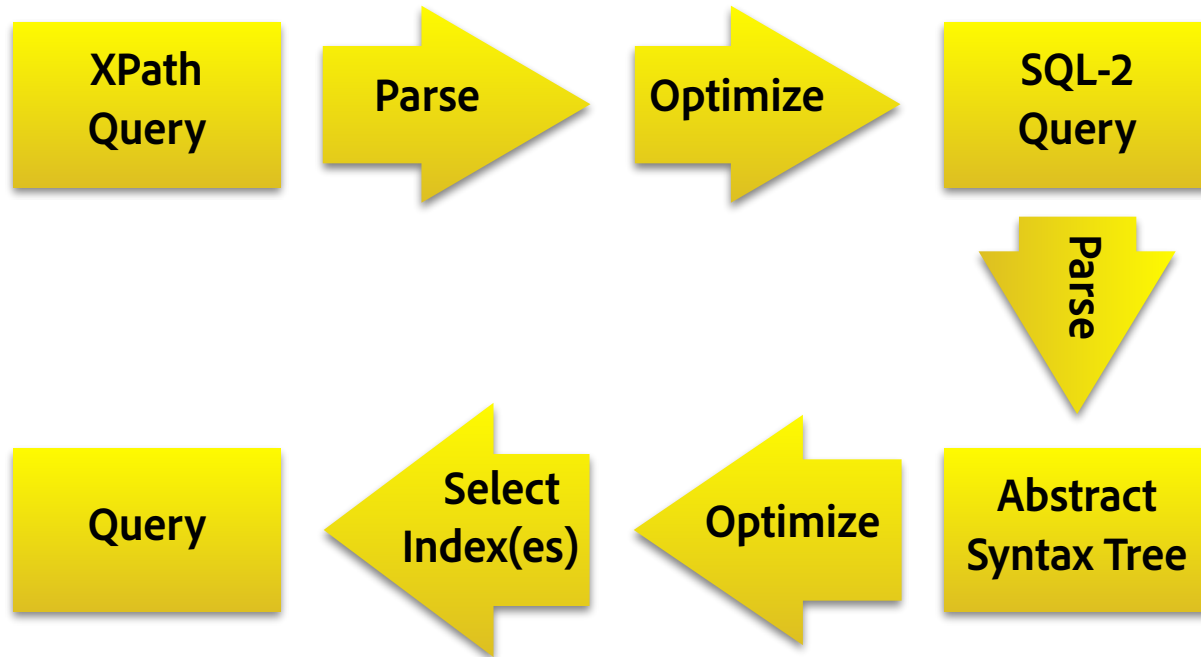
Oak Query: Index Selection



Oak Query: Index Selection



Oak Query: Index Selection



Index Types:

Property Index

- Synchronous
- Unique or Non-Unique

Nodetype Index

- Uses Property Index

Traversing "Index"

- Traverse the repo...
- or just one/few node(s)

Ordered Index

- Async



Full-Text Index Types:

Aggregate Index

- query-time aggregation

Lucene Index

Solr Index



New Index Type:

Lucene Property Index

- **async**
- **supports properties, later also full-text**
- **index-time aggregation**
- **supports ordering**



XPath

```
/jcr:root  
/content/dam/outdoors//  
element(*, dam:Asset)  
[@tag = 'x' or @tag = 'y']  
order by @jcr:lastModified
```

SQL

```
select [jcr:path],  
[jcr:score], *  
from [dam:Asset] as a  
where [tag] in('x', 'y')  
and isdescendantnode(a,  
'/content/dam/outdoors')  
order by [jcr:lastModified]
```

XPath

```
/jcr:root  
/content/dam/outdoors//  
element(*, dam:Asset)  
[@tag = 'x' or @tag = 'y']  
order by @jcr:lastModified
```

SQL

```
select [jcr:path],  
[jcr:score], *  
from [dam:Asset] as a  
where [tag] in('x', 'y')  
and isdescendantnode(a,  
'/content/dam/outdoors')  
order by [jcr:lastModified]
```



best
index wins

Query Debug Log

- See jackrabbit.apache.org/oak/docs
- Enable debug level logging for `org.apache.jackrabbit.oak.query`

```
xpath: /jcr:root/content/dam/outdoors//element...
```

```
sql2: select [jcr:path],...
```

```
cost for aggregate lucene is Infinity
```

```
cost for ordered is Infinity
```

```
cost for nodeType is 10005.0
```

```
cost for property is 4.0
```

```
cost for traverse is 199997.0
```



XPath

```
/jcr:root/*  
[jcr:contains(., 'marley')]  
order by @jcr:score
```

SQL

```
select [jcr:path],  
[jcr:score], *  
from [nt:base] as a  
where contains(*, 'marley')  
and isdescendantnode(  
a, '/')  
order by [jcr:score]
```



full-text
index wins

Index Management

- Currently, under /oak:index/
- Plan: install as package
- Re-index is usually not needed



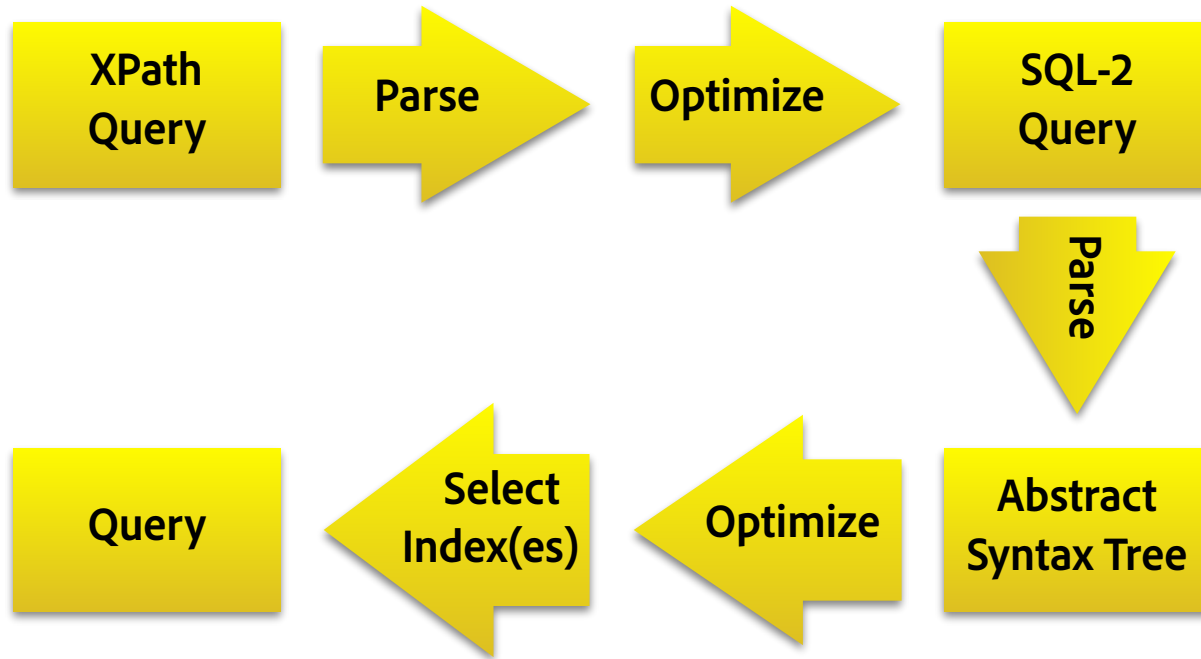
MongoMK

- Overview & Tools
- Clustering
- Eventual Consistency

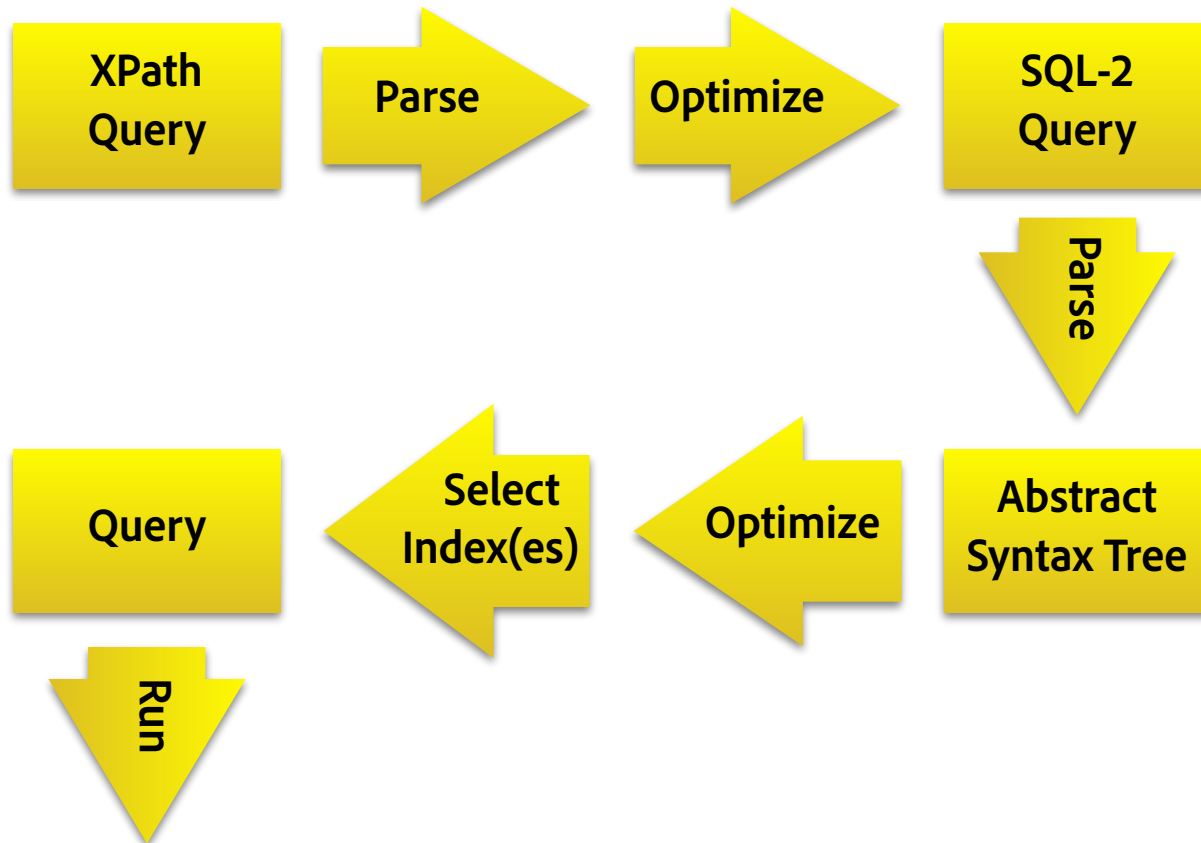
Query

- Parsing
- Index Selection
- **Execution**

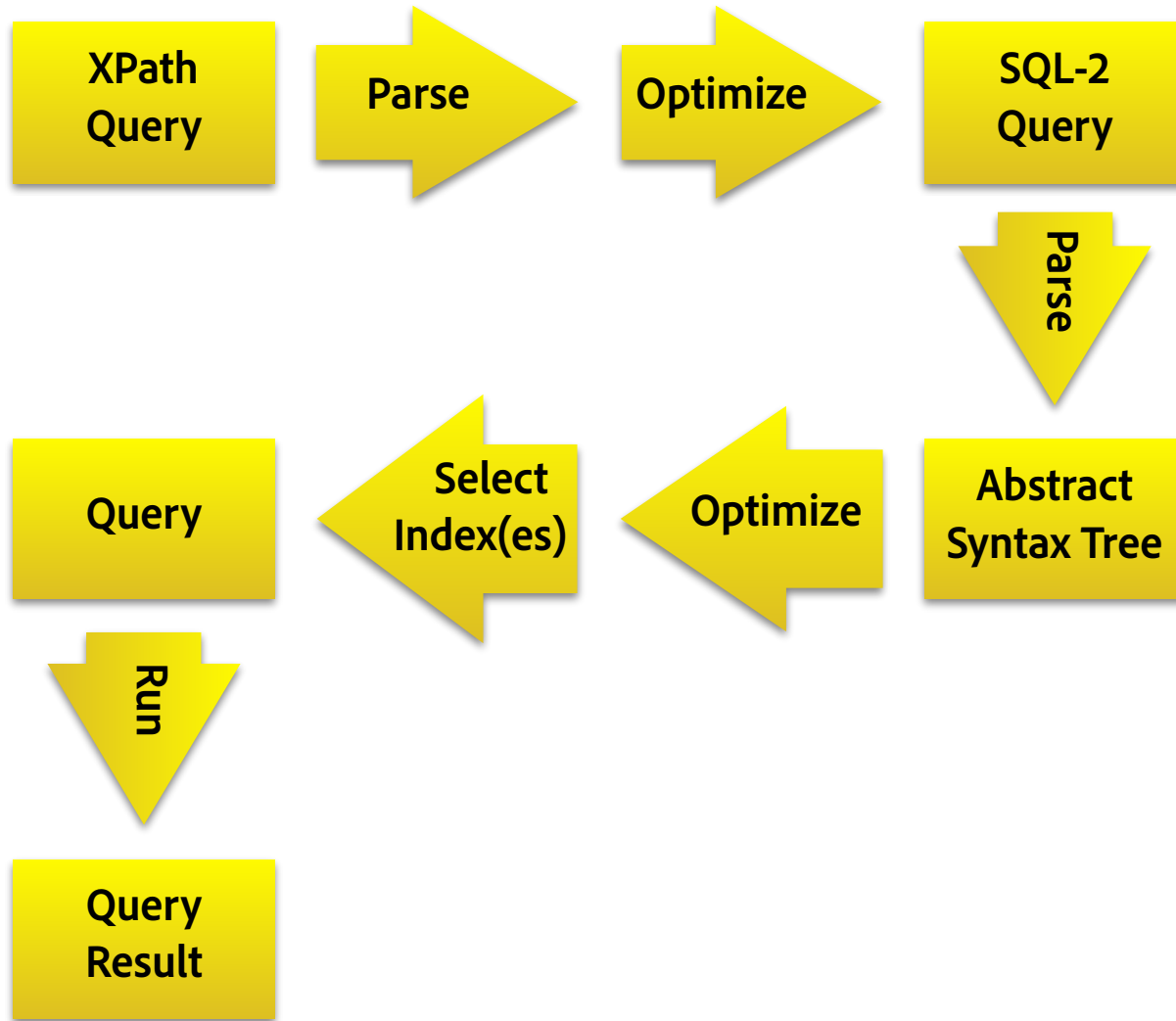
Oak Query: Execution



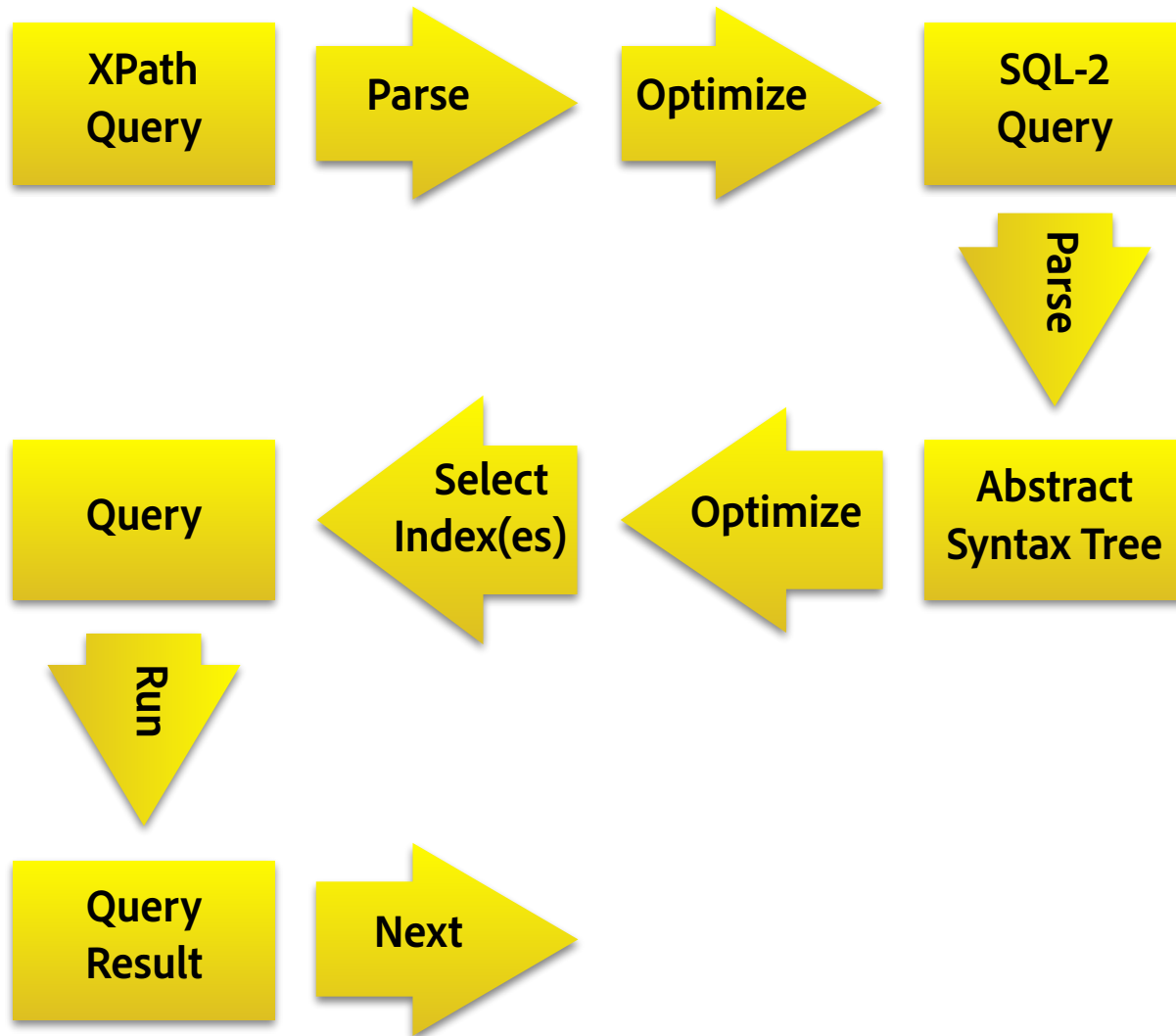
Oak Query: Execution



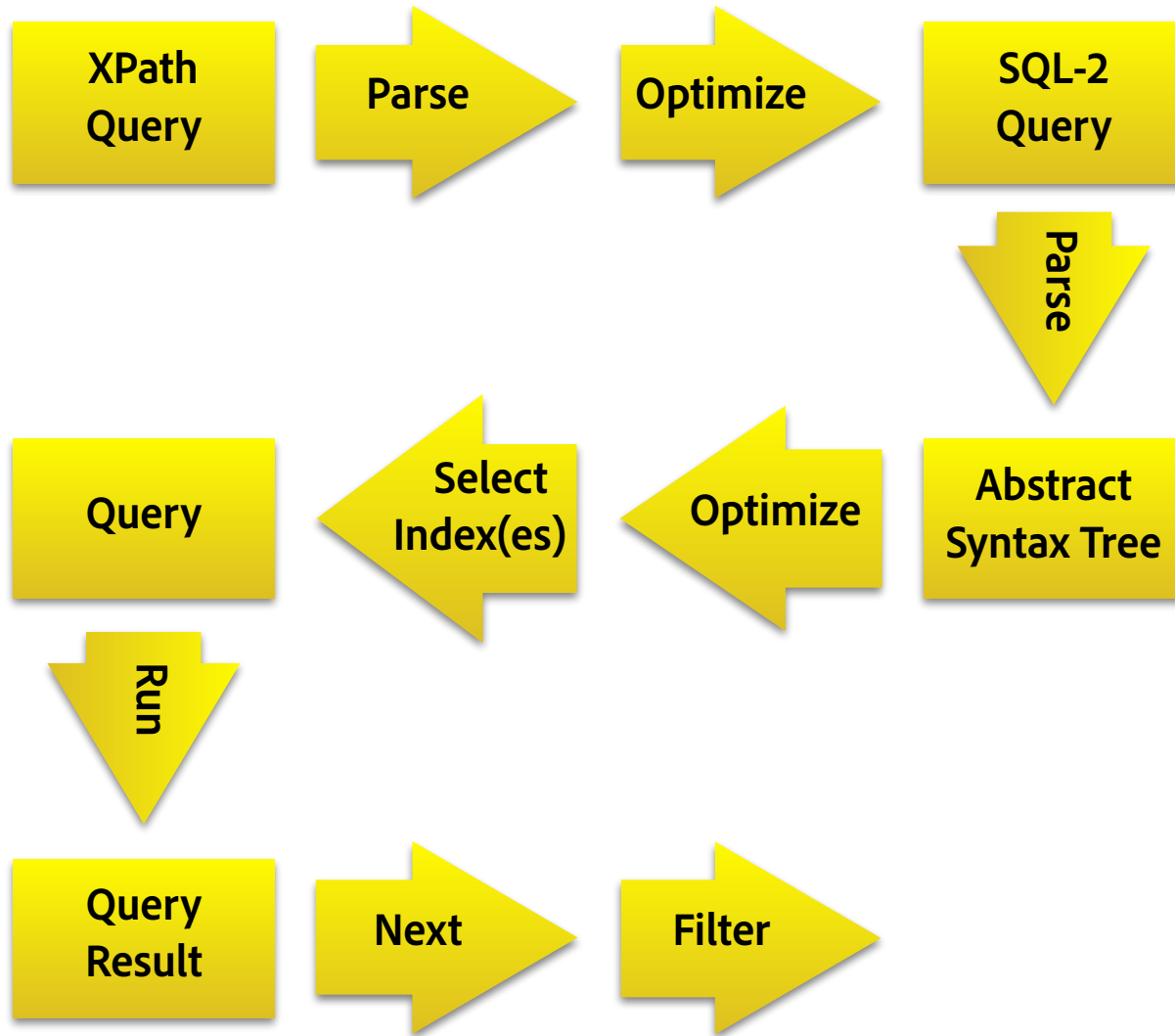
Oak Query: Execution



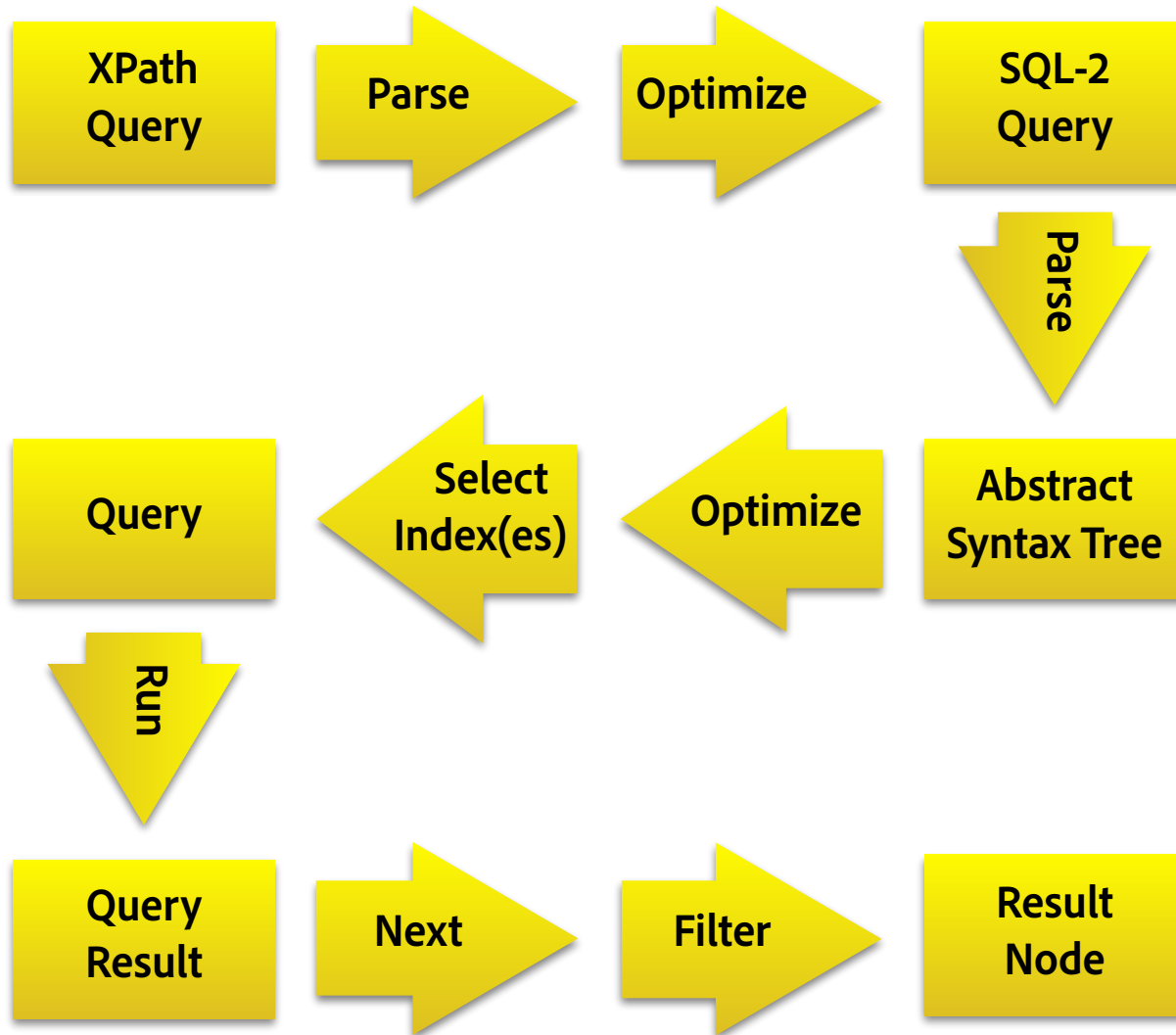
Oak Query: Execution



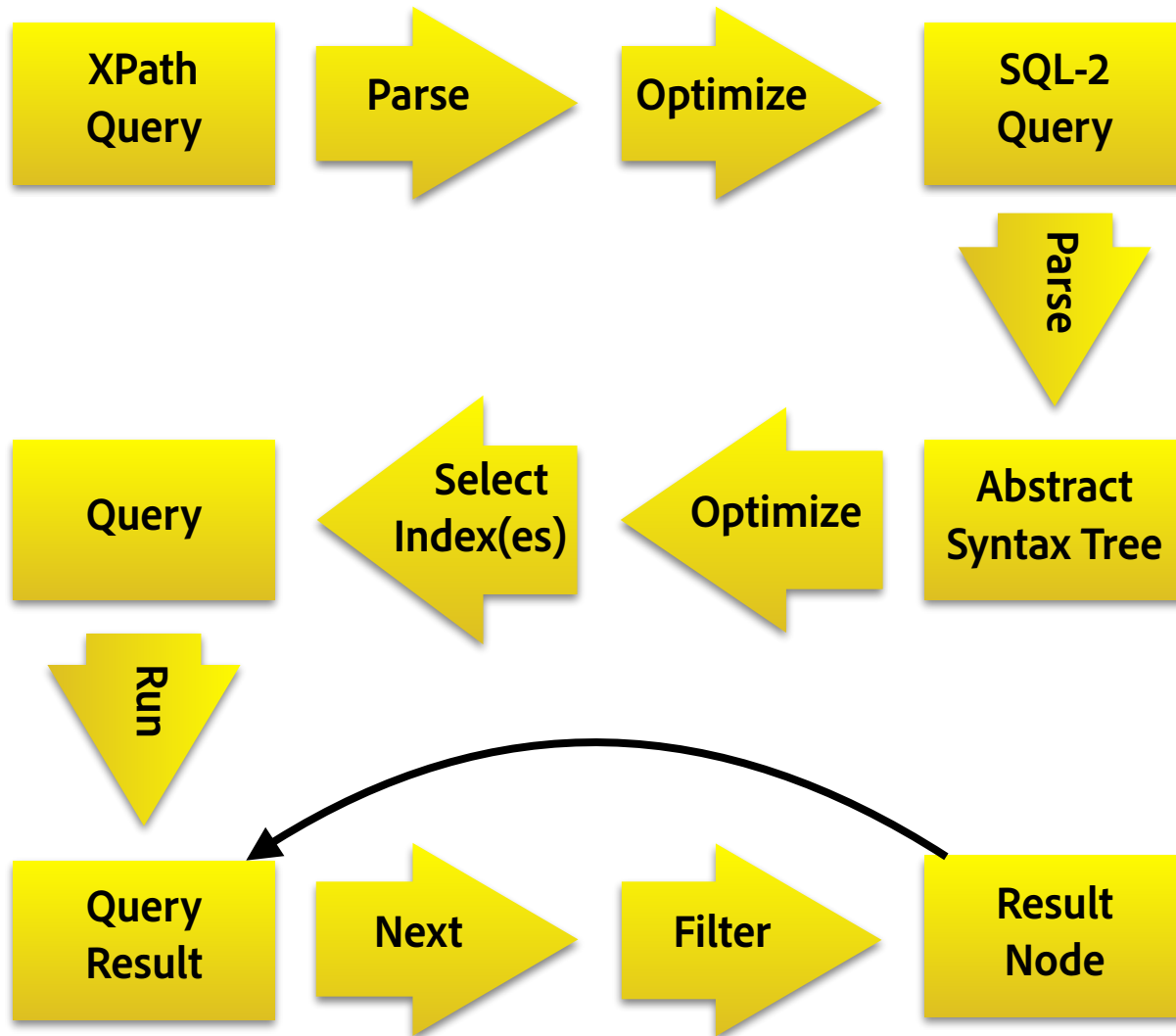
Oak Query: Execution



Oak Query: Execution



Oak Query: Execution



Warnings in Error Log

- Queries are logged if many nodes are read

```
/jcr:root/content//*[@tag = 'x']
```

```
*WARN* ... Traversed 20000 nodes
```



Read Limits

- Slow queries don't stop by default
- To make them stop, use JMX:
QueryEngineSettings
LimitInMemory, LimitReads
(or a system property)



Tools

- CRX DE, Log files: error.log, query.log
- Tread Dumps, Heap Histogram, Profiler
- Felix Script Console:
<http://felix.apache.org/documentation/subprojects/apache-felix-script-console-plugin.html>
- Index Manager:
<http://adobe-consulting-services.github.io/acs-aem-commons/features/oak-index-manager.html>
- Oak-Run (console)



Images: Creative Commons Licensed

Binoculars

<https://www.flickr.com/photos/67611651@N03/13610540345>

from Robert Körner (modified)



Adobe