Big NoSQL Data

Michael Carey
Information Systems Group
UC Irvine
mjcarey@ics.uci.edu
Plan for Today’s Talk

• The pre-relational and relational eras
• Moving beyond rows and columns (?)
  1. The object-oriented DB era
  2. The object-relational DB era
  3. The XML DB era
  4. The NoSQL DB era
• AsterixDB as a Big NoSQL Data exemplar
  – User model (DDL, DML, etc.)
  – System architecture and internal highlights
• Conclusions and Q&A
The Birth of Today’s DBMS Field

• In the beginning was the Word, and the Word was with Codd, and the Word was Codd...
  – 1970 CACM paper: “A relational model of data for large shared data banks”

• Many refer to this as the first generation of (real?) database management systems
The First Decade B.C.

- The need for a data management library, or a database management system, had actually been well recognized
  - Hierarchical DB systems (e.g., IMS from IBM)
  - Network DB systems (most notably CODASYL)
- These systems provided navigational APIs
  - Systems provided files, records, pointers, indexes
  - Programmers had to (carefully!) scan or search for records, follow parent/child structures or pointers, and maintain code when anything physical changed
The First Decade B.C. (cont.)

Order (id, custName, custCity, total)

Product (sku, name, listPrice, size, power)

Item (ino, qty, price)

Item-Order

Item-Product

(Product) (Parent child record sets)

123    Fred     LA      25.97

1    2   9.99

2   1   3.99

401    Garfield T-Shirt    9.99    XL    -

544    USB Charger    5.99    -    115V
Enter the Relational DB Era

Order \((id, \text{custName}, \text{custCity}, \text{total})\)

\[
\begin{array}{|c|c|c|c|}
\hline
123 & Fred & LA & 25.97 \\
\hline
\end{array}
\]

Item \((\text{order-id, ino, product-sku, qty, price})\)

\[
\begin{array}{|c|c|c|c|c|}
\hline
123 & 1 & 401 & 2 & 9.99 \\
123 & 2 & 544 & 1 & 3.99 \\
\hline
\end{array}
\]

Product \((\text{sku, name, listPrice, size, power})\)

\[
\begin{array}{|c|c|c|c|c|}
\hline
401 & \text{Garfield T-Shirt} & 9.99 & \text{XL} & \text{null} \\
544 & \text{USB Charger} & 5.99 & \text{null} & 115V \\
\hline
\end{array}
\]

- Be sure to notice that
  - Everything’s now (logical) rows and columns
  - The world is flat; columns are atomic \((1NF)\)
  - Data is now connected via keys (foreign/primary)
As the Relational Era Unfolded

• The Spartan simplicity of the relational data model made it possible to start tackling the opportunities and challenges of a logical data model
  – Declarative queries (Rel Alg/Calc, Quel, QBE, SQL, ...)
  – Transparent indexing (physical data independence)
  – Query optimization and execution
  – Views, constraints, referential integrity, security, ...
  – Scalable (shared-nothing) parallel processing

• Today’s multi-$B industry was slowly born
  – Commercial adoption took ~10-15 years
  – Parallel DB systems took ~5 more years
Enter the Object-Oriented DB Era

Notice that:
- Data model contains objects and pointers (OIDs)
- The world is no longer flat – the Order and Product schemas now have set(Item) and Product in them, respectively
What OODBs Sought to Offer

• Motivated largely by late 1980’s CAx applications (e.g., mechanical CAD, VLSI CAD, software CAD, ...)
  – Rich schemas with inheritance, complex objects, object identity, references, ...
  – Methods (“behavior”) as well as data in the DBMS
  – Tight bindings with (OO) programming languages
  – Fast navigation, some declarative querying

• Ex: Gemstone, Ontos, Objectivity, Versant, Object Design, O2, also DASDBS (sort of)
Why OODBs “Fell Flat”

• Too soon for another (radical) DB technology
  – Also technically immature relative to RDBMSs
• Tight PL bindings were a double-edged sword
  – Data shared, outlives programming languages
  – Bindings led to significant system heterogeneity
  – Also made schema evolution a major challenge
• Systems “overfitted” in some dimensions
  – Inheritance, version management, ...
  – Focused on thick clients (e.g., CAD workstations)
Enter the Object-Relational DB Era

<table>
<thead>
<tr>
<th>Order (id, customer, total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item (order-id, ino, product-sku, qty, price)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(123)</td>
</tr>
<tr>
<td>(123)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Product (sku, name, listPrice)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClothingProduct (size) under Product</td>
</tr>
<tr>
<td>ElectricProduct (power) under Product</td>
</tr>
<tr>
<td>401</td>
</tr>
<tr>
<td>544</td>
</tr>
</tbody>
</table>

- Be sure to notice:
  - “One size fits all!” (😊)
  - UDTs/UDFs, table hierarchies, references, ...
  - But the world got flatter again...
    (Timing lagged OODBs by just a few years)
What O-R DBs Sought to Offer

• Motivated by newly emerging application opportunities (multimedia, spatial, text, ...)
  – User-defined functions (UDTs/UDFs) & aggregates
  – Data blades (UDTs/UDFs + indexing support)
  – OO goodies for tables: row types, references, ...
  – Nested tables (well, at least Oracle added these)

• Back to a model where applications were loosely bound to the DBMS (e.g., ODBC/JDBC)

• Ex: ADT-Ingres, Postgres, Starburst, UniSQL, Illustra, DB2, Oracle
Why O-R DBs “Fell Flat”

• Significant differences across DB vendors
  – SQL standardization lagged somewhat
  – Didn’t include *details* of UDT/UDF extensions
  – Tough to extend the innards (for indexing)

• Application issues (and multiple platforms)
  – Least common denominator vs. coolest features
  – Tools (e.g., DB design tools, ORM layers, ...)

• Also still probably a bit too much too soon
  – IT departments still rolling in RDBMSs and creating relational data warehouses
Then Came the XML DB Era

<Order id="123">
  <Customer>
    <custName>Fred</custName>
    <custCity>LA</custCity>
  </Customer>
  <total>25.97</total>
  <Items>
    <Item ino="1">
      <product-sku>401</product-sku>
      <qty>2</qty>
      <price>9.99</price>
    </Item>
    <Item ino="2">
      <product-sku>544</product-sku>
      <qty>1</qty>
      <price>3.99</price>
    </Item>
  </Items>
</Order>

<Product sku="401">
  <name>Garfield T-Shirt</name>
  <listPrice>9.99</listPrice>
  <size>XL</size>
</Product>

<Product sku="544">
  <name>USB Charger</name>
  <listPrice>5.99</listPrice>
  <power>115V</power>
</Product>

Note that
- The world’s less flat again
- We’re now in the 2000’s
What XML DBs Sought to Offer

• One `<flexible/>` data model fits all (XML)
  – Origins in document markup (SGML)
  – Nested data
  – Schema variety/optionality

• New declarative query language (XQuery)
  – Designed both for querying and transformation
  – Early standardization effort (W3C)

• Two different DB-related use cases, in reality
  – *Data storage*: Lore (pre-XML), Natix, Timber, Ipedo, MarkLogic, BaseX; also DB2, Oracle, SQL Server
  – *Data integration*: Nimble Technology, BEA Liquid Data (from Enosys), BEA AquaLogic Data Services Platform
Why XML DBs “Fell Flat” Too

• Document-centric origins (vs. data use cases) of XML Schema and XQuery made a mess of things
  – W3C XPATH legacy (😢)
  – Document identity, document order, ...
  – Attributes vs. elements, nulls, ...
  – Mixed content (overkill for non-document data)

• Two other external trends also played a role
  – SOA and Web services came but then went
  – JSON (and RESTful services) appeared on the scene

• Note: Likely still an important niche market...
Now the **NoSQL DB Era?**

- **Not** from the DB world!
  - Distributed systems folks
  - Also various startups

- From caches → K/V use cases
  - Needed massive scale-out
  - OLTP (vs. parallel DB) apps
  - Simple, low-latency API
  - Need a key K, but want no schema for V
  - Record-level atomicity, replica consistency varies

- In the context of this talk, NoSQL does **not** mean
  - Hadoop (or SQL on Hadoop)
  - Graph databases or graph analytics platforms
NoSQL Data (JSON-based)

Collection(Order)

```
{ "id": "123",
  "Customer": {
    "custName": "Fred",
    "custCity": "LA"
  }
  "total": 25.97,
  "Items": [
    { "product-sku": 401,
      "qty": 2,
      "price": 9.99 },
    { "product-sku": 544,
      "qty": 1,
      "price": 3.99 }
  ]
}
```

Collection(Product)

```
{ "sku": 401,
  "name": "Garfield T-Shirt",
  "listPrice": 9.99,
  "size": "XL" }

{ "sku": 544,
  "name": "USB Charger",
  "listPrice": 5.99,
  "power": "115V" }
```

Note that
- The world’s not flat, but it’s less <messy/>
- We’re now in the 2010’s, timing-wise
Current NoSQL Trends

- Popular examples: MongoDB, Couchbase
- Coveting the benefits of many DB goodies
  - Secondary indexing and non-key access
  - Declarative queries
  - Aggregates and now (initially small) joins
- Seem to be heading towards...
  - BDMS (think scalable, OLTP-aimed, parallel DBMS)
  - Declarative queries and query optimization, but applied to schema-less data
  - Return of (some, optional!) schema information
Our Example: **Apache AsterixDB**

Data loads & feeds from external sources (XML, JSON, ...)

AQL queries & scripting requests and programs

Data publishing to external sources and apps

---

**ASTERIX Goal:**
To ingest, digest, persist, index, manage, query, analyze, and publish massive quantities of semistructured information...

---

Just How Big is “Big Data”?

This is big data!
So what went on – and why?

What’s going on right now?
Also: Today’s Big Data Tangle

Spark SQL

mongoDB

(Pig)

Hive

Cassandra

MySQL

HBase

Hadoop
AsterixDB: “One Size Fits a Bunch”

Semistructured Data Management

Parallel Database Systems

1st Generation “Big Data” Systems

BDMS Desiderata:

• Able to manage data
• Flexible data model
• Full query capability
• Continuous data ingestion
• Efficient and robust parallel runtime
• Cost proportional to task at hand
• Support “Big Data data types”
ASTERIX Data Model (ADM)

CREATE DATaverse TinySocial;
USE TinySocial;

CREATE TYPE GleambookUserType AS {
    id: int,
    alias: string,
    name: string,
    userSince: datetime,
    friendIds: { int },
    employment: [EmploymentType]
};

CREATE TYPE EmploymentType AS {
    organizationName: string,
    startDate: date,
    endDate: date?
};

CREATE DATASET GleambookUsers (GleambookUserType)
PRIMARY KEY id;

Highlights include:
• JSON++ based data model
• Rich type support (spatial, temporal, ...)
• Records, lists, bags
• Open vs. closed types
ASTERIX Data Model (ADM)

CREATE DATaverse TinySocial;
USE TinySocial;

CREATE TYPE GleambookUserType AS {
    id: int
};

CREATE TYPE EmploymentType AS {
    organizationName: string,
    startDate: date,
    endDate: date?
};

CREATE DATASET GleambookUsers (GleambookUserType)
PRIMARY KEY id;

Highlights include:
• JSON++ based data model
• Rich type support (spatial, temporal, ...)
• Records, lists, bags
• Open vs. closed types
CREATE DATASURE TinySocial;
USE TinySocial;

CREATE TYPE GleambookUserType AS {
    id: int
};

CREATE TYPE GleambookMessageType AS {
    messageId: int,
    authorId: int,
    inResponseTo: int?,
    senderLocation: point?,
    message: string
};

CREATE DATASET GleambookUsers (GleambookUserType)
PRIMARY KEY id;

CREATE DATASET GleambookMessages (GleambookMessageType)
PRIMARY KEY messageId;

Highlights include:
• JSON++ based data model
• Rich type support (spatial, temporal, ...)
• Records, lists, bags
• Open vs. closed types
Ex: GleambookUsers Data

{"id":1, "alias":"Margarita", "name":"MargaritaStoddard", "nickname":"Mags",
"userSince":datetime("2012-08-20T10:10:00"), "friendIds":{{2,3,6,10}},
"employment": [ {"organizationName":"Codetechno", "startDate":date("2006-08-06")},
   {"organizationName":"geomedia", "startDate":date("2010-06-17"),
    "endDate":date("2010-01-26")} ]},

"gender":"F"
}

{"id":2, "alias":"Isbel", "name":"IsbelDull", "nickname":"Izzy",
"userSince":datetime("2011-01-22T10:10:00"), "friendIds":{{1,4}},
"employment": [ {"organizationName":"Hexviafind", "startDate":date("2010-04-27")} ]}

{"id":3, "alias":"Emory", "name":"EmoryUnk",
"userSince":datetime("2012-07-10T10:10:00"), "friendIds":{{1,5,8,9}},
"employment": [ {"organizationName":"geomedia", "startDate":date("2010-06-17"),
   "endDate":date("2010-01-26")} ]
},

......
Other DDL Features

CREATE INDEX gbUserSinceIdx ON GleambookUsers(userSince);
CREATE INDEX gbAuthorIdx ON GleambookMessages(authorId) TYPE BTREE;
CREATE INDEX gbSenderLocIndex ON GleambookMessages(senderLocation) TYPE RTREE;
CREATE INDEX gbMessageIdx ON GleambookMessages(message) TYPE KEYWORD;
// --------------------- and also---------------------------------------------

CREATE TYPE AccessLogType AS CLOSED
{ ip: string, time: string, user: string, verb: string, `path`: string, stat: int32, size: int32 };
CREATE EXTERNAL DATASET AccessLog(AccessLogType) USING localfs
("path"="localhost:///Users/mikejcarey/extdemo/accesses.txt"),
("format"="delimited-text"), ("delimiter"="|"));

CREATE FEED myMsgFeed USING socket_adapter
("sockets"="127.0.0.1:10001"), ("address-type"="IP"),
("type-name"="GleambookMessageType"), ("format"="adm"));
CONNECT FEED myMsgFeed TO DATASET GleambookMessages;
START FEED myMsgFeed;

External data highlights:
• Equal opportunity access
• Feeds to “keep everything!”
• Ingestion, not streams
ASTERIX Queries (SQL++ or AQL)

• **Q1:** List the user names and messages sent by Gleambook social network users with less than 3 friends:

```
SELECT user.name AS uname,
    (SELECT VALUE msg.message
     FROM GleambookMessages msg
     WHERE msg.authorId = user.id) AS messages
FROM GleambookUsers user
WHERE COLL_COUNT(user.friendIds) < 3;
```

```
{ "uname": "NilaMilliron", "messages": [ ] }
{ "uname": "WoodrowNehling", "messages": [ " love acast its 3G is good:)" ] }
{ "uname": "IsbelDull", "messages": [ " like product-y the plan is amazing", " like product-z its platform is mind-blowing" ] }
...
```
Q2: Identify active users (last 30 days) and group and count them by their numbers of friends:

```sql
WITH endTime AS current_datetime(),
    startTime AS endTime - duration("P30D")
SELECT nf AS numFriends, COUNT(user) AS activeUsers
FROM GleambookUsers user
LET nf = COLL_COUNT(user.friendIds)
WHERE SOME logrec IN AccessLog SATISFIES
    user.alias = logrec.user
    AND datetime(logrec.time) >= startTime
    AND datetime(logrec.time) <= endTime
GROUP BY nf;
```

SQL++ highlights:
- Many features (see docs)
- Spatial & text predicates
- Set-similarity matching
Updates and Transactions

- **Q3**: Add a new user to Gleambook.com:
  ```sql
  UPSERT INTO GleambookUsers (  
  "id":667,"alias":"dfrump",  
  "name":"DonaldFrump",  
  "nickname":"Frumpkin",  
  "userSince":datetime("2017-01-01T00:00:00"),  
  "friendIds":{{ }},  
  "employment":{{"organizationName":"USA",  
  "startDate":date("2017-01-20")}},  
  "gender":"M"}  
  );
  ```

- Key-value store-like transactions (w/record-level atomicity)
- Insert, delete, and upsert ops; index-consistent
- 2PL concurrency
- WAL no-steal, no-force with LSM shadowing
AsterixDB System Overview

![AsterixDB System Diagram]
Software Stack

SQL++ or AQL

Apache AsterixDB

XQuery

Apache VXQuery

HiveQL

Hivesterix

Pregel Job

Hadoop M/R Job

Hyracks Job

Algebricks

Pregelix

M/R Layer

Operator Library (join, sort, group-by, etc.)

Storage Library (LSM B-Tree, R-Tree, etc.)

Connector Library (m-to-n, m-to-1, etc.)

HDFS Utilities

Hyracks General-Purpose DAG Execution Engine
Hyracks Dataflow Runtime

- Partitioned-parallel platform for data-intensive computing
- Job = dataflow DAG of operators and connectors
  - Operators consume and produce *partitions* of data
  - Connectors *route* (repartition) data between operators
- Hyracks vs. the “competition”
  - Based on time-tested parallel database principles
  - *vs.* Hadoop MR: More flexible model and less “pessimistic”
  - *vs.* newer SQL-on-Hadoop runtimes: Emphasis on out-of-core execution and adherence to memory budgets
  - Faster job activation, data pipelining, binary format, state-of-the-art DB style operators (hash-based, indexed, ...)

Early test at Yahoo! Labs on 180 nodes (1440 cores, 720 disks)
Query

```sql
use dataverse TinySocial

avg {
  for $m in dataset MugshotMessages
  where $m.timestamp >=
    datetime("2014-01-01T00:00:00")
  and $m.timestamp <
    datetime("2014-04-01T00:00:00")
  return string-length($m.message)
}
```

### Algebricks

- **Partitioned Parallelism!**

### Hyracks (cont.)

```sql
assign $hi := 2014-04-01T00:00:00
assign $lo := 2014-01-01T00:00:00
btree $id := search(msTimestampIdx, $lo, $hi)

sort $id
btree $id := search(MugshotMessages, $id, $id)
assign $l := string-length($m.message)
aggregate $lagg := local-avg($l)
aggregate $agg := global-avg($lagg)
assign $t := $m.timestamp
select $t >= 2014-01-01T00:00:00 and
       $t < 2014-04-01T00:00:00
```
Algebricks Query Compiler Framework

**Metadata Catalog**
- Query String
- Query Parser
- Abstract Syntax Tree
- Translator
- Logical Plan
- Type Inference and Check
- Logical Plan
- Rule-based Logical Optimizer
- Logical Plan
- Rule-based Physical Optimizer
- Physical Plan
- Hyracks Job Generator
- Hyracks Job
- Hyracks Runtime

**Algebricks**
- Logical Operators
- Logical Expressions
- Metadata Interface
- Model-Neutral Logical Rewrite Rules
- Physical Operators
- Model-Neutral Physical Rewrite Rules
- Hyracks Job Generator

**Target Query Language**
- Query Parser (AST)
- AST Translator
- Metadata Catalog
- Expression Type Computer
- Logical Rewrite Rules
- Physical Rewrite Rules
- Language Specifics

**Language Implementations**
- Language Implementations
- Algebricks
- Runtime
Native Storage Management

- **Disk 1**
- **Disk n**
- **IO Scheduler**
- **Working Memory**
- **Buffer Cache**
- **In-Memory Components**
- **Datasets Manager**
- **Transaction Sub-System**
  - **Transaction Manager**
  - **Lock Manager**
  - **Log Manager**
  - **Recovery Manager**

Other components:
- **Memory**
- **Disk**
- **Buffer Cache**
- **In-Memory Components**
- **IO Scheduler**
- **Hadoop HDFS**

UClrvine
LSM-Based Filters

Memory

T16, T17

Disk

[ T12, T15 ]

T12, T13, T14, T15

[ T7, T11 ]

T7, T8, T9, T10, T11

[ T1, T6 ]

T1, T2, T3, T4, T5, T6

Oldest Component

Intuition: Do NOT touch unneeded records

Idea: Utilize LSM partitioning to prune disk components

Q: Get all tweets > T14
Transaction Support

• Key-value store-like transaction semantics
  – Entity-level transactions (by key) within “transactors”
  – Atomic insert + delete operations (including indexing)
  – Concurrency control (based on entity-level locking)
  – Crash recovery (based on no-steal logging + shadowing)
  – Backup and restore support (just in case... 😊)

• Expected use of AsterixDB is to model, capture, and track the “state of the world” (not to be it)...

SELECT ... FROM Weather W...
// return current conditions by city
(Long serializable reads)
A Peek at Performance

<table>
<thead>
<tr>
<th></th>
<th>Users</th>
<th>Messages</th>
<th>Tweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asterix (Schema)</td>
<td>192</td>
<td>120</td>
<td>330</td>
</tr>
<tr>
<td>Asterix (KeyOnly)</td>
<td>360</td>
<td>240</td>
<td>600</td>
</tr>
<tr>
<td>Syst-X</td>
<td>290</td>
<td>100</td>
<td>495</td>
</tr>
<tr>
<td>Hive</td>
<td>38</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>Mongo</td>
<td>240</td>
<td>215</td>
<td>478</td>
</tr>
</tbody>
</table>

Table 2: Dataset sizes (in GB)

<table>
<thead>
<tr>
<th>Batch Size</th>
<th>Asterix Schema</th>
<th>Asterix KeyOnly</th>
<th>Syst-X</th>
<th>Mongo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.091</td>
<td>0.093</td>
<td>0.040</td>
<td>0.035</td>
</tr>
<tr>
<td>20</td>
<td>0.010</td>
<td>0.011</td>
<td>0.026</td>
<td>0.024</td>
</tr>
</tbody>
</table>

Table 4: Average insert time per record (in sec)

10 Node IBM Cluster
- 40 cores
- 40 disks
- GB Ethernet switch
Similar schema/queries to earlier examples.
# A Peek at Performance (cont.)

<table>
<thead>
<tr>
<th></th>
<th>Asterix</th>
<th>Asterix</th>
<th>Syst-X</th>
<th>Hive</th>
<th>Mongo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Schema</td>
<td>KeyOnly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rec Lookup</strong></td>
<td>0.03</td>
<td>0.03</td>
<td>0.12</td>
<td>(379.11)</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Range Scan</strong></td>
<td>79.47</td>
<td>148.15</td>
<td>148.33</td>
<td>11717.18</td>
<td>175.84</td>
</tr>
<tr>
<td>— with IX</td>
<td>0.10</td>
<td>0.10</td>
<td>4.90</td>
<td>(11717.18)</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Sel-Join (Sm)</strong></td>
<td>78.03</td>
<td>96.76</td>
<td>55.01</td>
<td>333.56</td>
<td>66.46</td>
</tr>
<tr>
<td>— with IX</td>
<td>0.51</td>
<td>0.55</td>
<td>2.13</td>
<td>(333.56)</td>
<td>0.62</td>
</tr>
<tr>
<td><strong>Sel-Join (Lg)</strong></td>
<td>79.62</td>
<td>99.73</td>
<td>56.65</td>
<td>350.92</td>
<td>273.52</td>
</tr>
<tr>
<td>— with IX</td>
<td>2.24</td>
<td>2.32</td>
<td>10.59</td>
<td>(350.92)</td>
<td>14.97</td>
</tr>
<tr>
<td><strong>Sel2-Join (Sm)</strong></td>
<td>79.06</td>
<td>97.82</td>
<td>55.81</td>
<td>340.02</td>
<td>66.45</td>
</tr>
<tr>
<td>— with IX</td>
<td>0.50</td>
<td>0.52</td>
<td>2.62</td>
<td>(340.02)</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>Sel2-Join (Lg)</strong></td>
<td>80.18</td>
<td>101.24</td>
<td>56.10</td>
<td>394.11</td>
<td>313.17</td>
</tr>
<tr>
<td>— with IX</td>
<td>2.32</td>
<td>2.32</td>
<td>10.70</td>
<td>(394.11)</td>
<td>15.28</td>
</tr>
<tr>
<td><strong>Agg (Sm)</strong></td>
<td>128.66</td>
<td>232.30</td>
<td>130.64</td>
<td>83.18</td>
<td>400.97</td>
</tr>
<tr>
<td>— with IX</td>
<td>0.16</td>
<td>0.17</td>
<td>0.14</td>
<td>(83.18)</td>
<td>0.19</td>
</tr>
<tr>
<td><strong>Agg (Lg)</strong></td>
<td>128.71</td>
<td>232.41</td>
<td>132.19</td>
<td>94.11</td>
<td>401</td>
</tr>
<tr>
<td>— with IX</td>
<td>5.53</td>
<td>5.55</td>
<td>4.67</td>
<td>(94.11)</td>
<td>8.34</td>
</tr>
<tr>
<td><strong>Grp-Aggr (Sm)</strong></td>
<td>130.20</td>
<td>232.77</td>
<td>131.18</td>
<td>127.85</td>
<td>398.27</td>
</tr>
<tr>
<td>— with IX</td>
<td>0.45</td>
<td>0.46</td>
<td>0.17</td>
<td>(127.85)</td>
<td>0.20</td>
</tr>
<tr>
<td><strong>Grp-Aggr (Lg)</strong></td>
<td>130.62</td>
<td>234.10</td>
<td>133.02</td>
<td>140.21</td>
<td>400.10</td>
</tr>
<tr>
<td>— with IX</td>
<td>5.96</td>
<td>5.91</td>
<td>4.72</td>
<td>(140.21)</td>
<td>9.03</td>
</tr>
</tbody>
</table>

**Table 3: Average query response time (in sec)**
Example AsterixDB Use Cases

• Potential use case areas include
  – Social data analytics
  – Cell phone event analytics
  – Behavioral science
  – Education
  – Public health
  – Power usage monitoring
  – Cluster management log analytics
  – ....
Current Status

• 4 year initial NSF project (250+ KLOC), started 2009
• Now officially *Apache AsterixDB*!
  – Semistructured “NoSQL” style data model
  – Declarative queries, inserts, deletes, upserts (SQL++)
  – Scalable parallel query execution
  – Data storage/indexing (primary & secondary, LSM-based)
  – Internal and external datasets both supported
  – Rich set of data types (including text, time, location)
  – Fuzzy and spatial query processing
  – NoSQL-like transactions (for inserts/deletes)
  – Data feeds and indexes for external datasets
  – ....
Research Roadmap: Big Active Data

- Flexible NoSQL Data Model
- Scalable Storage & Indexing
- Full Declarative Query Capability
- Web data types & Search I
- Fast Continuous Data Ingestion
- Web data types & Search II
- Windowed Aggregation
- Continuous Query Support
Research Roadmap: Big Data Viz
Commercial Use: Big Data Analytics

Couchbase Data Platform

- Service-Centric Clustered Data System
- Multi-process Architecture
- Dynamic Distribution of Facilities
- Cluster Map Distribution
- Automatic Failover
- Enterprise Monitoring/Management
- Security
- Offline Mobile Data Integration
- Streaming REST API
- SQL-like Query Engine for JSON
- Clustered* Global Indexes
- Lowest Latency Key-Value API
- Active-Active Inter-DC Replication
- Local Aggregate Indexes
- Full-Text Search*

Operational Analytics*
Big Data Analytics (cont.)

- Separate Couchbase services, separate nodes
  - Multi-Dimensional Scaling
  - Workload isolation
  - Same flexible data model
- Parallel shadowing of data(sets) via DCP
  - Low impact on data nodes
  - Low latency
For More Information

• Asterix project UCI/UCR research home
  – http://asterix.ics.uci.edu/

• Apache AsterixDB home
  – http://asterixdb.apache.org/

• SQL++ Primer
  – http://asterixdb.apache.org/docs/0.9.1/index.html

QUESTIONS...?