Large Scale Data Processing and Applications: *Cloudy DB Research at UC Irvine*

Michael Carey
Vinayak Borkar

Information Systems Group
CS Department
UC Irvine

http://isg.ics.uci.edu

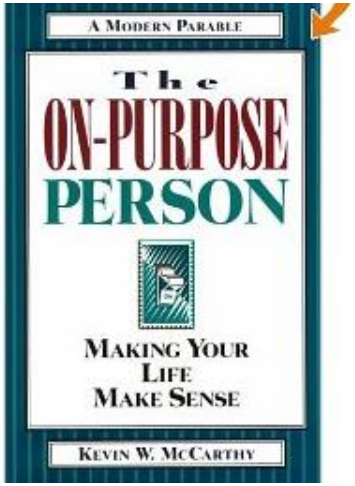
Cloud DB Bandwagons Today

- MapReduce and Hadoop
  - Parallel programming for dummies
  - But now Hive, Pig, Scope, …
  - MapReduce is the new runtime
- DFSs and HDFS (and CSS)
  - Scalable, self-managed, Really Big Files
  - But now BigTable, HBase, …
  - HDFS (or CSS) is the new file storage
- Key-value stores
  - Charter members of the NoSQL movement
  - Includes S3, Dynamo, HBase, Cassandra, …
  - These are the new record managers

Let's Do This Stuff "Right"!

- In my opinion
  - The OS/DS folks out-scaled us (oops)
  - We'd be nuts to build on their foundation
- Identify the lessons, but then do it "right"
  - Cheap open-source S/W on commodity H/W
  - Non-monolithic software components
  - Equal opportunity data access (external sources)
  - Fault-tolerant query execution
  - Little pre-planning or DBA-type work required
  - Tolerant of flexible / nested / absent schemas
  - Types and declarative languages (duh!)

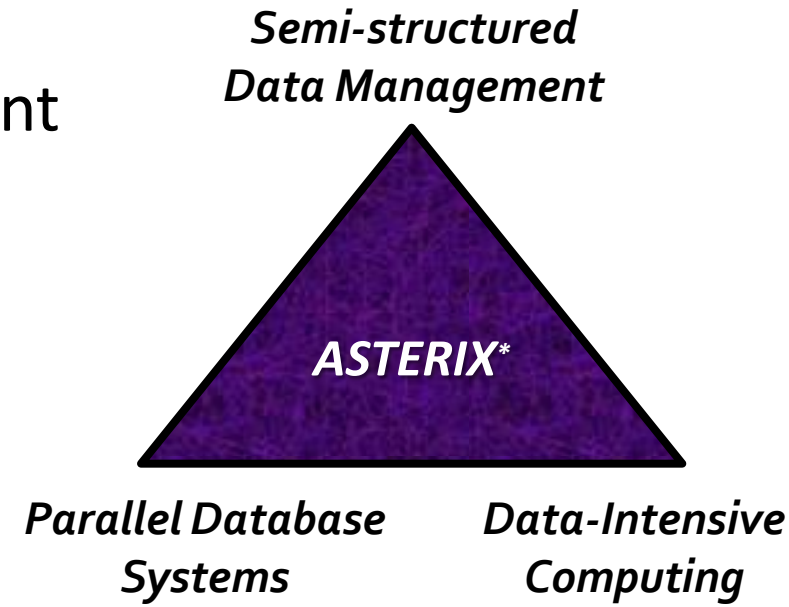
What If We'd *Meant* To Do This?



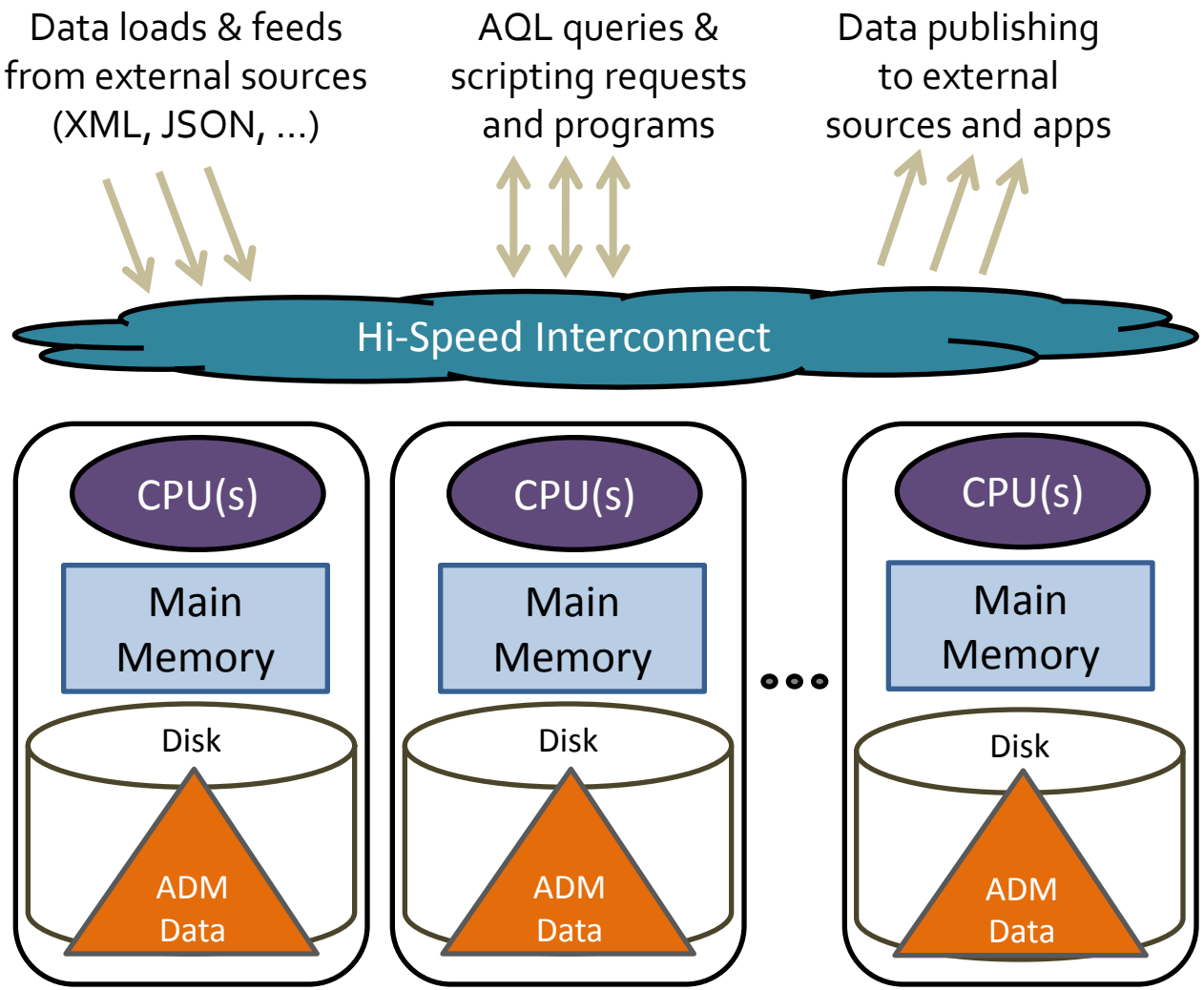
- What is the "right" basis for analyzing and managing the data of the future?
  - Runtime layer (and division of labor)?
  - Storage and data distribution layers?
- Explore how to build new information management systems for the cloud that…
  - Seamlessly support external data access
  - Execute queries in the face of partial failures
  - Scale to thousands of nodes (and beyond)
  - Don't require five-star wizard administrators
  - ….

The ASTERIX Project

- Semistructured data management
  - Core work exists
  - XML & XQuery, JSON, …
  - Time to parallelize and scale out
- Parallel database systems
  - Research quiesced in mid-1990's
  - Renewed industrial interest
  - Time to scale up & de-schema-tize
- Data-intensive computing
  - MapReduce and Hadoop popular today
  - Various language efforts (Pig, Hive, Jaql, …)
  - Ripe for parallel DB query processing ideas and support for stored, indexed data sets



ASTERIX Project Overview



ASTERIX Goal:
To *ingest*, *digest*, *persist*, *index*, *manage*, *query*, *analyze*, and *publish* massive quantities of semi-structured information…

(ADM = ASTERIX Data Model)

ASTERIX User Model

- Data model (ADM) with open and closed types

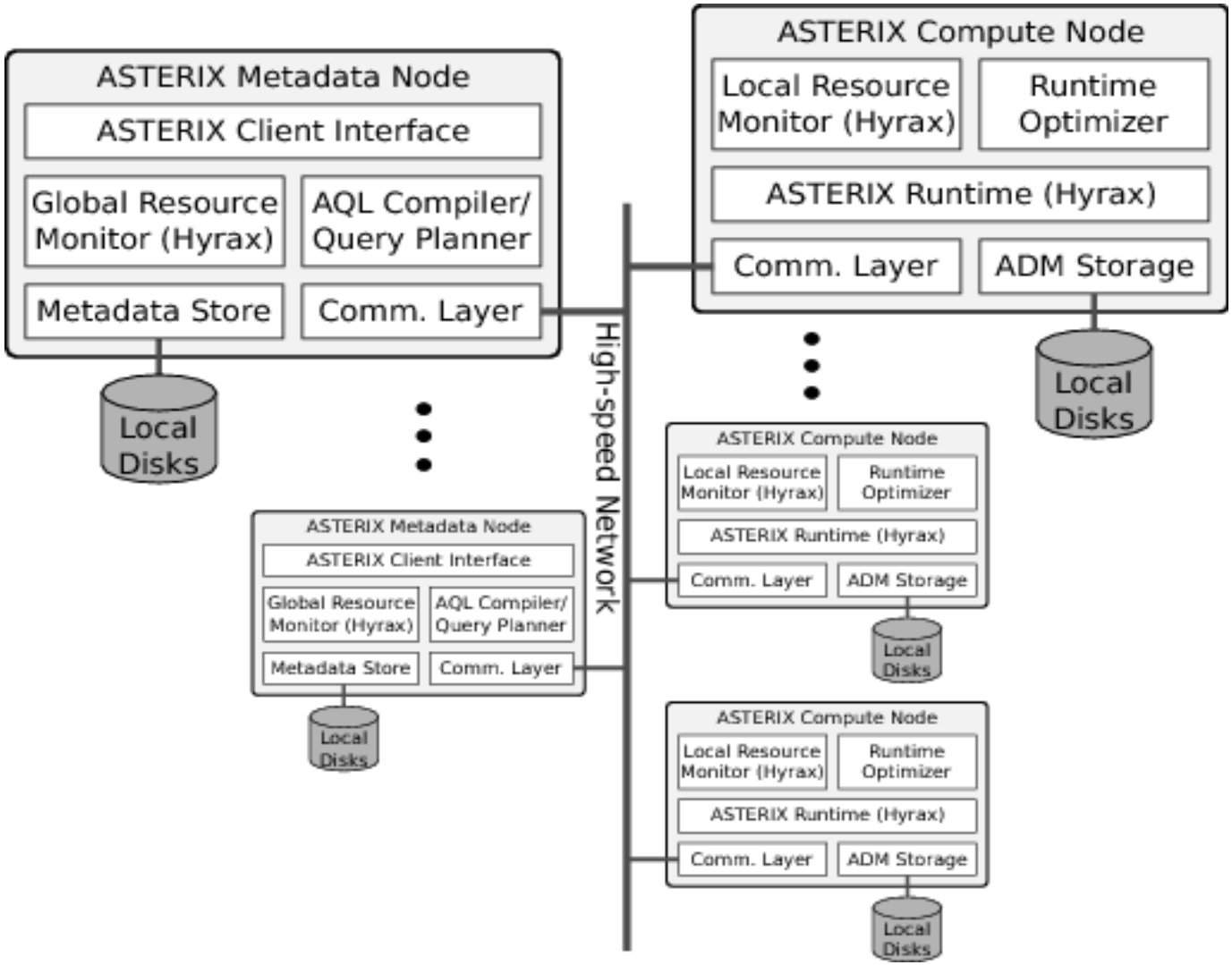
```
declare open type UserType as {
  name: string,
  email: string,
  interests: <string>,
  address: AddressType,
  member_of: <
    {
      sig_name: string,
      chapter_name: string,
      member_since: date
    }
  >
}
declare closed type AddressType as {
  street: string,
  city: string,
  zip: string,
  latlong: point2d
}
```

- Query language (AQL) for nested and semi-structured data queries

```
for $user in dataset('User')
where some $i in $user.interests
        satisfies $i ="movies"
return {"name": $user.name };
```

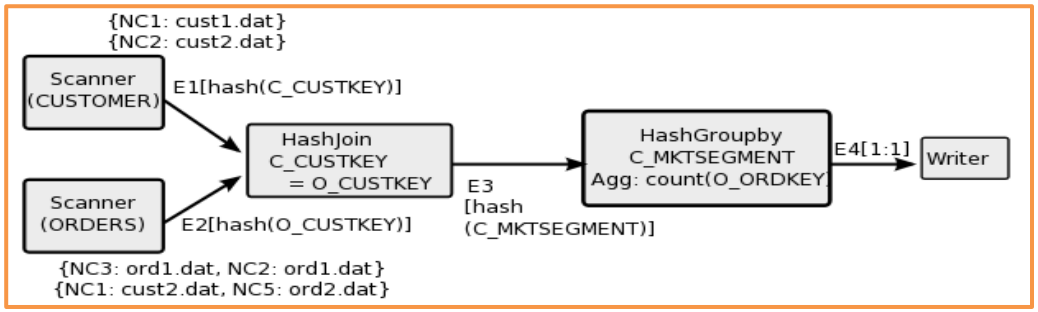
- Support for both stored and external datasets

ASTERIX and Hyracks



Hyracks In a Nutshell

- Partitioned-parallel platform for data-intensive computing
- Job = dataflow DAG of operators and connectors
  - Operators consume/produce partitions of data
  - Connectors repartition/route data between operators

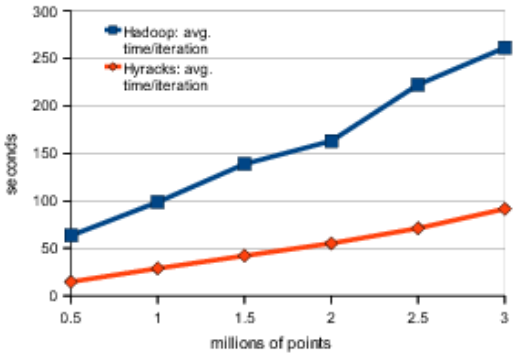


- Hyracks *vs.* the "competition"
  - Based on time-tested parallel database principles
  - *vs.* Hadoop: More flexible model and less "pessimistic"
  - *vs.* Dryad: Supports data as a first-class citizen

Hyracks Performance

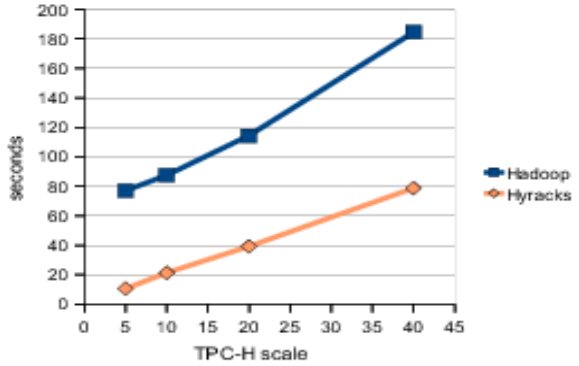
(40 cores / 40 disks)

- K-means (on Hadoop compatibility layer)

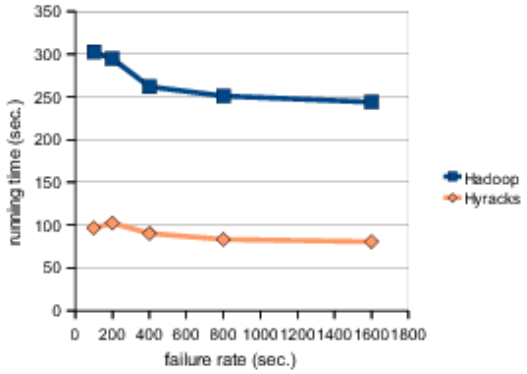


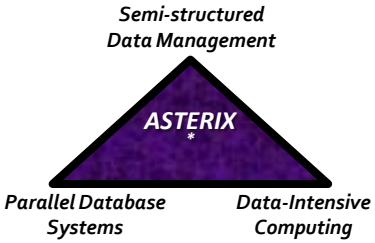
(Faster ☺)

- DSS-style query execution (TPC-H-based example)

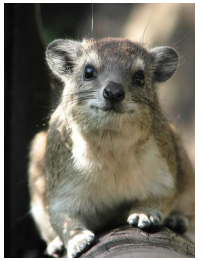


- Fault-tolerant query execution (TPC-H-based example)





In Summary



- Ask not what cloud software can do for you, but what you can do for cloud software…!
- We're asking this very question at UCI
  - ASTERIX: Parallel semistructured DBMS
  - Hyracks: Partitioned-parallel data runtime
- If you're interested, the first Hyracks release (in open source at Google Code) is coming in ~1-2 weeks