

Hypertable

*CS227 Li Jin, Yang Lu
Brown University*

The background features several sets of concentric circles in a lighter shade of blue, resembling ripples in water. These circles are positioned in the lower right and bottom center of the slide.

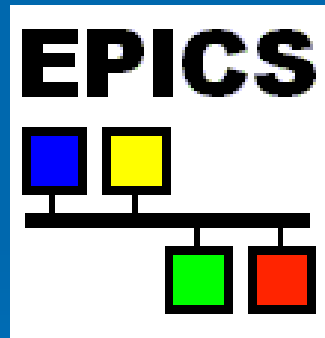
Schedule

- Overview
- Data Model
- Architecture
- Case Study

High Performance, Open Source Scalable Database

- Modeled after **Bigtable**
 - Implemented in C++
 - Project Started in March 2007
 - Runs on top of HDFS
 - Thrift Interface for all popular languages
 - Java
 - PHP
 - Ruby
 - Python
 - Perl, etc.
- 

Hypertable Deployments



Data Model



Table: Visual Representation

crawlDb Table

<i>row key</i>	title	content	anchor
com.facebook.www	Facebook Home	<!DOCTYPE html ...	anchor:com.apple.www/ Facebook
com.yahoo.www	Yahoo!	<html><head>...	anchor:com.redherring.www/ Facebook
com.zvents.www	Discover Things To Do - Zvents	<html xmlns="http...	anchor:org.slashdot.www/ Zvents
org.hypertable.www	Hypertable: An Open Source, High Performance, ...	<!DOCTYPE html ...	

Table: Key

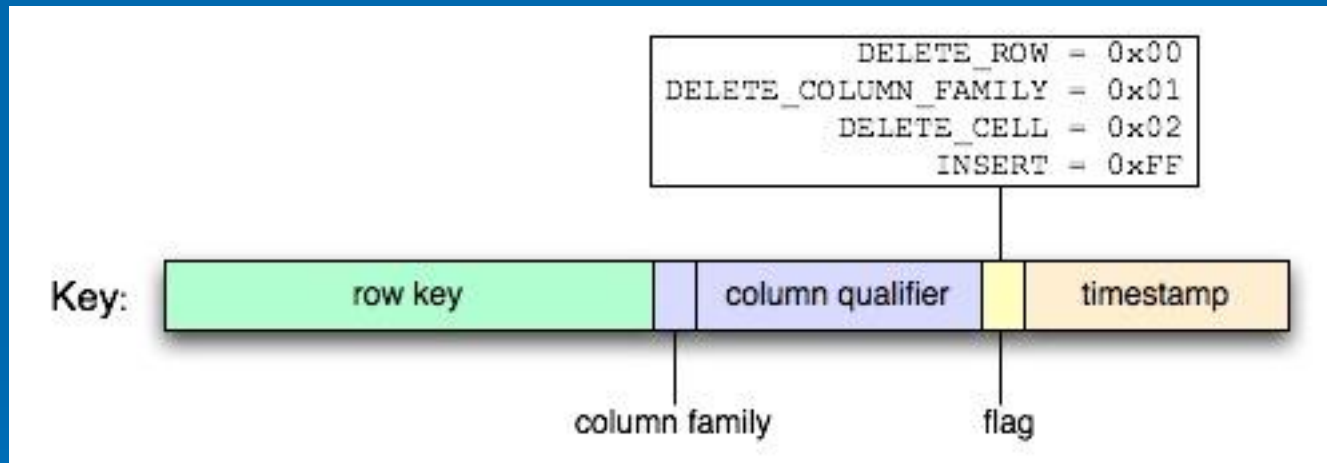


Table: Actual Representation

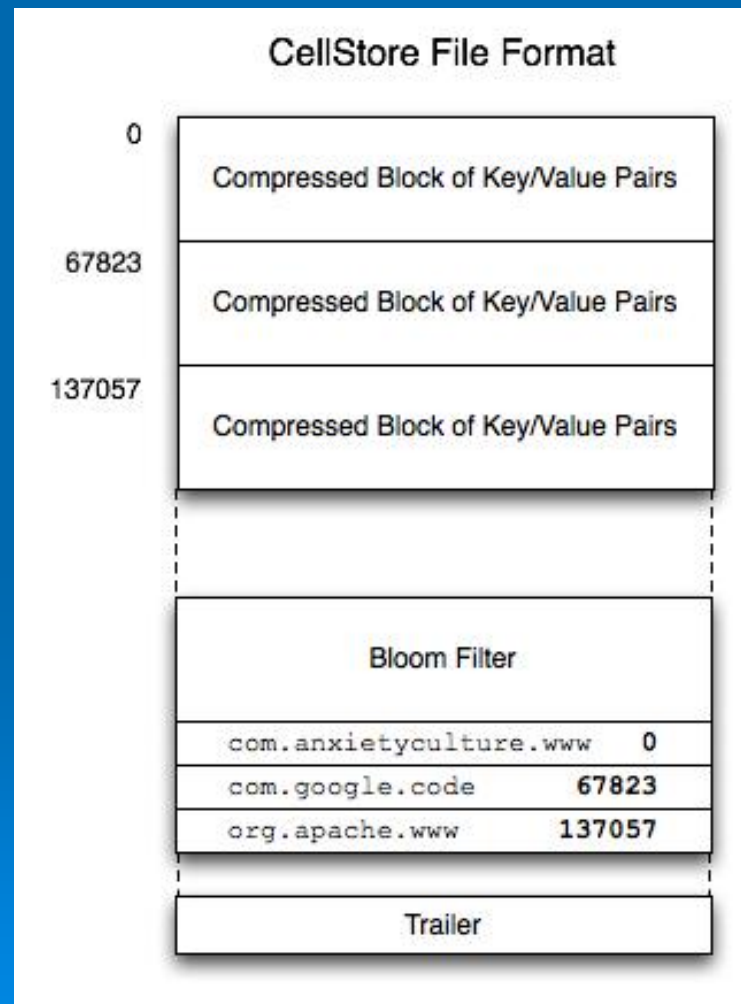
crawldb Table

key

value

com.facebook.www title 2008-02-11 15:14:01	Facebook Home
com.facebook.www title 2008-02-03 19:27:57	Facebook Home
com.facebook.www title 2008-01-22 08:46:28	Facebook Home
com.facebook.www content 2008-02-11 15:14:01	<!DOCTYPE html PUBLIC "-//W3C//DTD...
com.facebook.www content 2008-02-03 19:27:57	<!DOCTYPE html PUBLIC "-//W3C//DTD...
com.facebook.www content 2008-01-22 08:46:28	<!DOCTYPE html PUBLIC "-//W3C//DTD...
com.facebook.www anchor:com.apple.www/ 2008-02-11 15:14:01	Facebook
com.facebook.www anchor:com.apple.www/ 2008-02-03 19:27:57	Facebook
com.facebook.www anchor:com.apple.www/ 2008-01-22 08:46:28	Facebook
com.facebook.www anchor:com.redherring.www/ 2008-02-11 15:14:01	Facebook
com.facebook.www anchor:com.redherring.www/ 2008-02-03 19:27:57	Facebook
com.yahoo.www title 2008-02-10 21:12:09	Yahoo!
com.yahoo.www title 2008-02-04 03:46:22	Yahoo!
com.yahoo.www title 2008-01-22 08:46:28	Yahoo!
com.yahoo.www content 2008-02-10 21:12:09	<html><head><meta http-equiv="Content-...
com.yahoo.www content 2008-02-04 03:46:22	<html><head><meta http-equiv="Content-...
...	...

Table: Physical Data Layout



Access Groups

- Provides control over physical layout
 - Row oriented
 - Column oriented
 - Hybrid
- Reduces I/O

```
CREATE TABLE MyTable (  
    a, b, c, d,  
    ACCESS GROUP first(a),  
    ACCESS GROUP second (b, c, d)  
);
```

Scaling (part I)

session table

0000-020310 ...
0000-100101 ...
0000-121200 ...
0000-145121 ...
0000-230039 ...
0000-347987 ...
0000-493602 ...
0000-512100 ...
0000-972981 ...
0000-988135 ...
0000-992873 ...
0001-039266 ...
0001-080220 ...
0001-085830 ...
0001-096853 ...
0001-165562 ...
0001-200238 ...
0001-272091 ...
0001-297291 ...
0001-871523 ...
0001-897217 ...
0001-933990 ...
0001-986956 ...
0002-049398 ...
0002-230212 ...
0002-283487 ...
0002-302340 ...
0002-345222 ...
0002-383049 ...
0002-451287 ...
0002-512621 ...
0002-727212 ...
...

crawldb table


com.anxietyculture.com ...
com.burningbuilding.www ...
com.cnn.www ...
com.cplusplus.www ...
com.danga.www ...
com.davidshrigley.www ...
com.ehow.www ...
com.google.code ...
com.google.www ...
com.highscalability ...
com.ifcfilms.www ...
com.lifehacker ...
com.myspace.www ...
com.nyelabs ...
com.nytimes.www ...
com.omnigroup.www ...
com.readymech ...
com.wolfram.integrals ...
com.yahoo.www ...
org.apache.wiki ...
org.apache.www ...
org.cmake.www ...
org.hypertable.www ...
org.sftw.www ...

Scaling (part II)




Scaling (part III)

Server-1



0000-230039 ...
0000-347997 ...
0000-493602 ...
0000-512100 ...
0000-972981 ...
0000-988135 ...
0000-992873 ...
0001-039266 ...
com.anxietyculture.w ...
com.burningbuilding ...
com.cnn.www ...
com.cplusplus.www ...

Server-2



com.danga.www ...
com.davidshrigley.w ...
com.chow.www ...
com.google.code ...
com.google.www ...
com.highscalability ...
com.ilcfilms.www ...
com.lifehacker ...
0001-897217 ...
0001-933990 ...
0001-986956 ...
0002-049398 ...

Server-3




0002-383049 ...
0002-451287 ...
0002-519821 ...
0002-727212 ...
com.readymech ...
com.wolfram.integral ...
com.yahoo.www ...
org.apache.wiki ...
0002-861825 ...
0002-892000 ...
0003-038532 ...
0003-049908 ...

Server-4



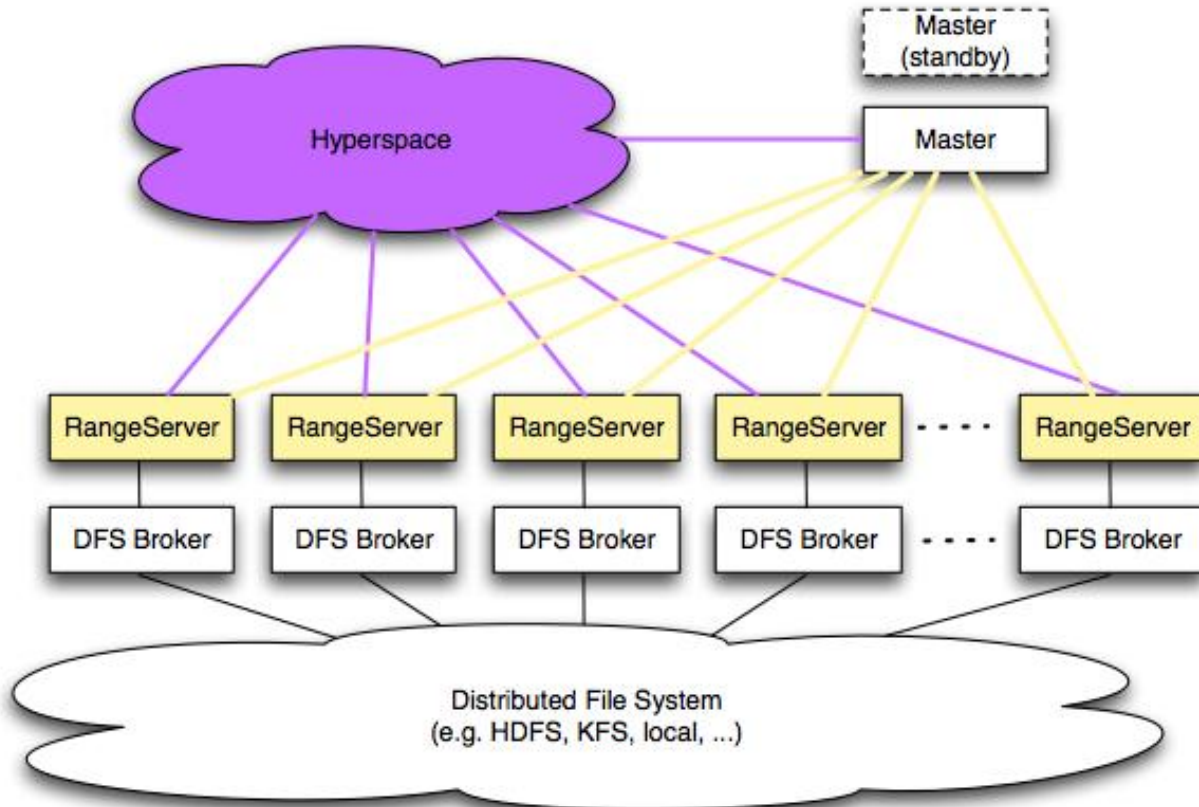
0001-200238 ...
0001-272091 ...
0001-297291 ...
0001-871523 ...
0000-020310 ...
0000-100101 ...
0000-121200 ...
0000-145121 ...
org.apache.www ...
org.cmake.www ...
org.hypertable.www ...
org.sfw.www ...

Server-5

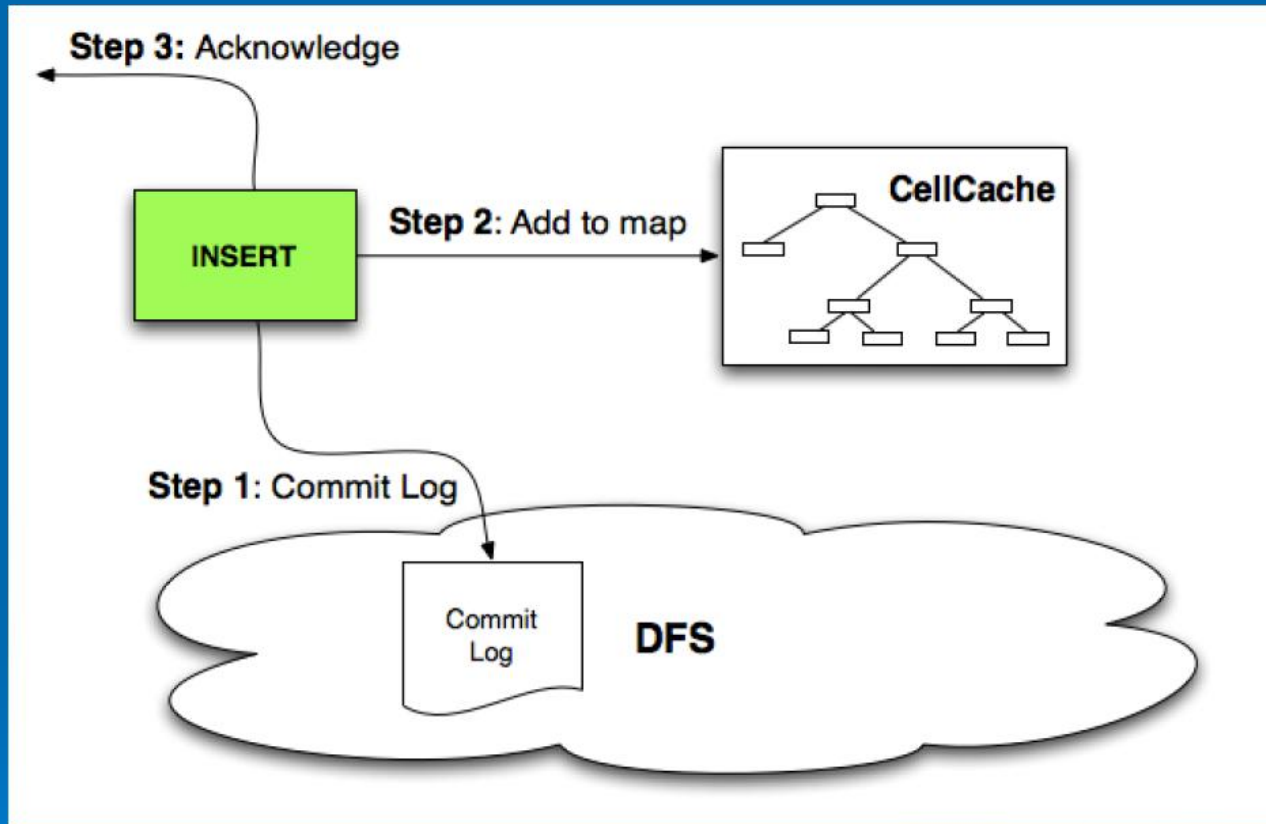


0002-230212 ...
0002-283487 ...
0002-302340 ...
0002-345222 ...
com.myspace.www ...
com.nyelabs ...
com.nytimes.www ...
com.omnigroup.www ...
0001-080220 ...
0001-085830 ...
0001-098853 ...
0001-165582 ...

System Components



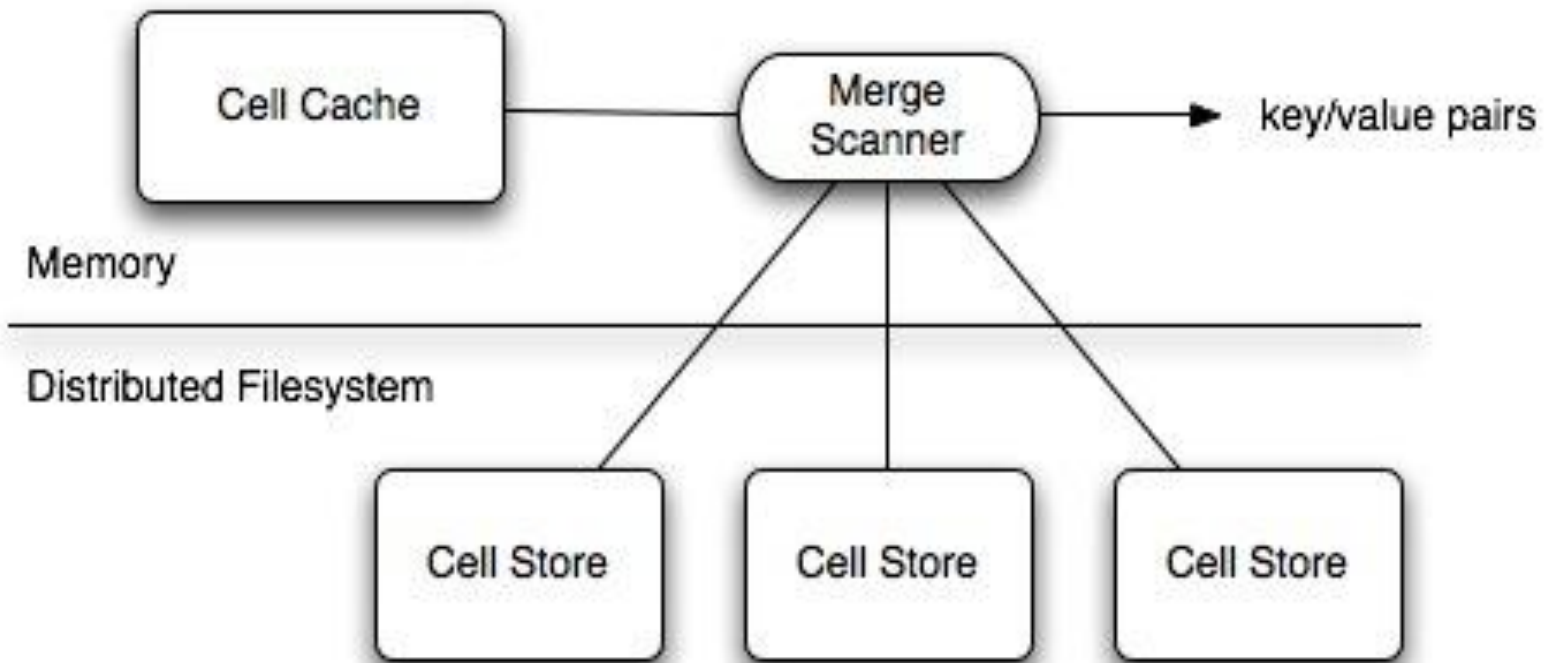
Query Handling: Write




Hypertable	Bigtable
CellStore	SSTable
CellCache	Memtable

Heap	
Insert	$\log(n)$
lookup	$\log(n)$
merge	$\log(n)$

Query Handling: Read



Transaction Support

- Single-row transactions
 - Entire row is guaranteed to be in a single range.
- Transactions with the data scattered over multiple RangeServers is not yet implemented.
- Maybe not necessary for a OLAP system. 

Indexing

Hyperspace

Access Group	Files
Content; Title	/hypertable/data/CellStore1
Anchor	/hypertable/data/CellStore1

Metadata 0 Root

Row key	Location
1:com.yahoo.www	192.168.79.5
2: us-ri-02912.www	192.168.79.6

Metadata 1 Tablet

Row key	Location	Files
1:com.facebook.www	192.168.79.5	
1:com.yahoo.www	192.168.79.6	

Metadata 1 Tablet

Row key	Location	Files
2:us-ri-02906	192.168.79.5	
2:us-ri-02912	192.168.79.6	



Client Access

- Thrift API:
 - supports C++, Java, PHP, Python, Perl, Ruby
 - `get_cells()`, `set_cells()`...
- HQL:
 - `select`, `insert`, `delete`..

Case Study: Tribalytic

- A market research tool using Twitter data.
- Example: Given a keyword “coffee machine”, it tells you people usually tweets about coffee machine around 9:00 14:00 and 21:00.

Case Study: Tribalytic

- One single table called “hits”.
 - Use keywords as row key.
- All queries search on row key.



Case Study: Tribalytic

- Process 3.5 million tweets a day.
- Push 50 million records to “hits” table a day.
- Less than 250 ms to locate and load 40,000 record into memory. (In the coffee machine example).

Case Study: Tribalytic

Row Key	Tweets: Li	Tweets: Yang	Tweets: ...	Tweets: ...
Coffee machine

Case Study: Tribalytic

Tweet Id	User	Timestamp	Content
----------	------	-----------	---------

Keyword	List(Tweet Id)
---------	----------------

OR

Keyword	List(User)	List(Timestamp)	List(Content)
---------	------------	-----------------	---------------

OR

Keyword	List(Tweet Object)
---------	--------------------

Best Fit Use Cases

- Serve massive data sets to live applications.
- Good for applications that need to scan over data ranges.



Q&A

