Big Data Framework

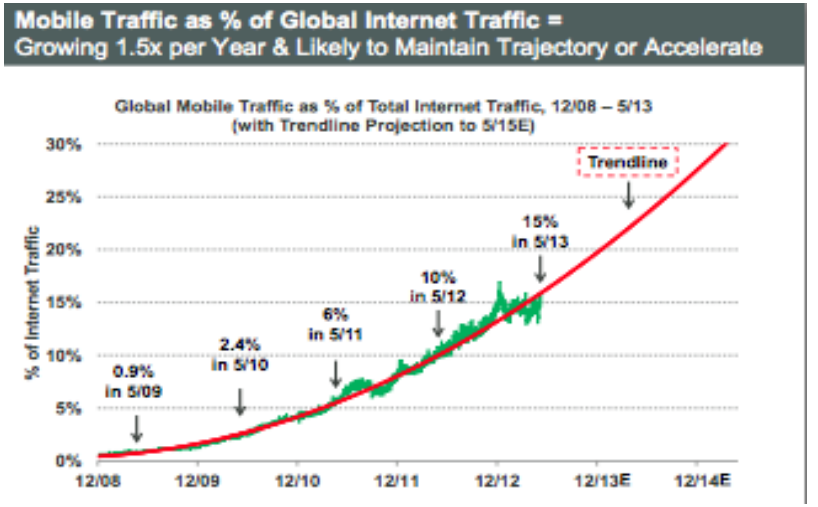
- Ties front end frameworks to cloud infrastructure
- Keeps application code in the apps
- Shortens release cycles and time to market
- Tools and conventions making Big Data easy
- App developers write apps, we take care of the backend

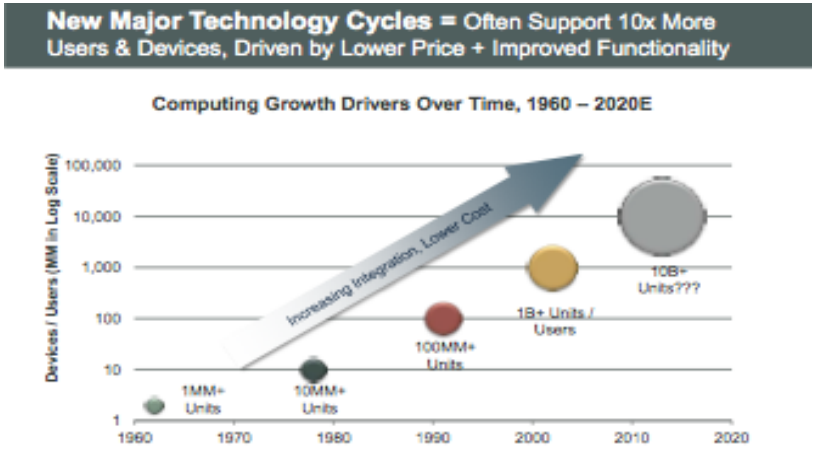
Key Drivers

- Mobile usage continues to grow
- Mobile apps are connected, free or nearly so, and popular
- Even moderately successful apps have huge traffic volumes
- It's rapidly getting worse
- Supporting these volumes is currently a large technical challenge

This is the basis of Big Data

http://www.kpcb.com/insights/2013-internet-trends

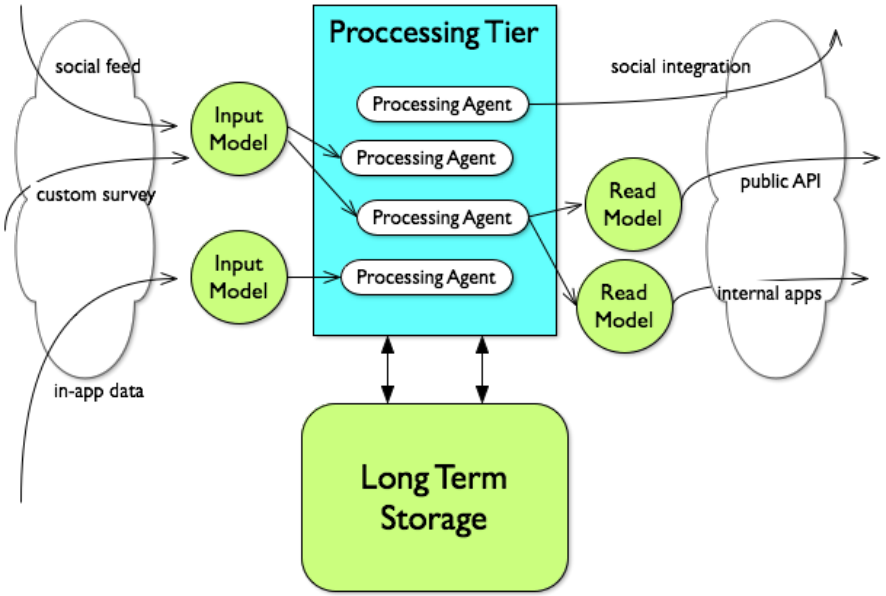




Current Landscape

- Hadoop & R are the two major Big Data solutions
- Both are low level technologies; expert-only systems
  - Requires an advanced degree level of comprehension in either computer science (Hadoop) or statistics (R).
- Last major technical innovation in this space was 10 years ago
- No higher level framework currently exists to lower this barrier

Technical Analysis



Big Data is all about performance

Three ways Big Data drags down performance:

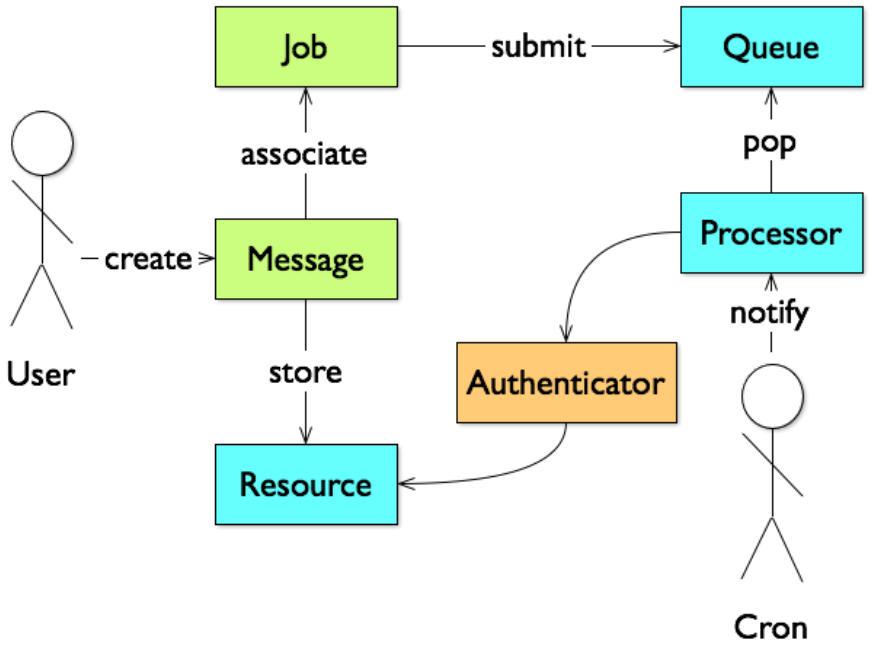
- Concurrency
- Data volume
- Data complexity

All solved using the same pattern

- Split the data into task specific chunks
- Optimize data storage for each task
- Transform data between task specific storage models
- Wash, rinse, repeat until performance is acceptable

Time & resource intensive process

Solution



- Develop a high level Big Data Framework
- Integrate with major mobile & web technologies
- Provide detailed howto documentation
  - Search, shopping cart, video on demand, etc.
- Apps begin life with task specific data models, performant from day one.
- Significantly lowers Big Data's barrier to entry

Path Based Storage

- Leverages highly scalable key/value storage
- Large collections of data are spread over a search space
- The search space is represented by a path
- Paths are hierarchal; query higher up the path for a bigger chunk
- Good for capturing training data from users for machine learning

App records a user's tap metadata at an input path:

```
/in/2014/02/19/13/10/24/N/34/1/16/W/118/23/44
```

Processed input is cascaded through a read path:

```
/suggestions/2014/02/19

/suggestions/2014/02/19/13

/suggestions/2014/02/19/13/10

/suggestions/2014/02/19/13/10/24
```

Another device can see what has happened in the last hour:

```
/suggestions/2014/02/19/13
```

Behavior + Data

- Hadoop and R both process data in batches
  - Resembles a regression to a function/data design
- Can lead to an unresponsive, error prone system
  - Fully testing a batch job running for hours is impossible

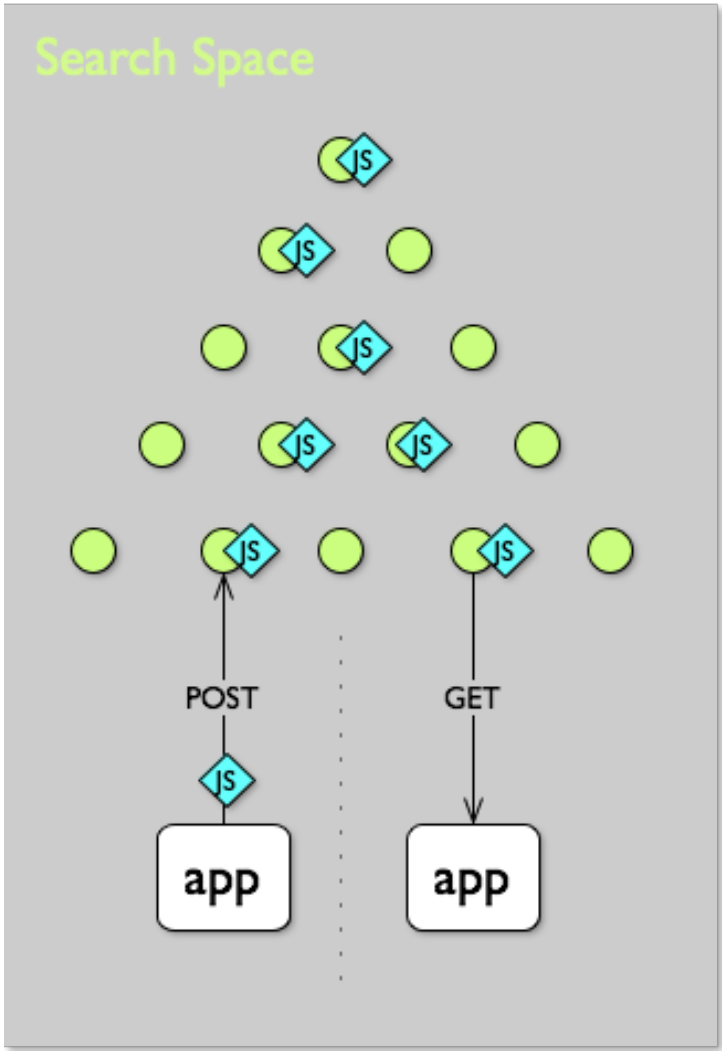
Real Innovation:

Move processing out of the batch

- Data events cascade updates through the search space

Move processing code into the app

- Apps associate functional behavior in JavaScript when updating data



Big Data Wants You!

- Looking for motivated and talented contributors
- Offering highly visible, highly relevant work experience
- Training on GitHub, issue tracking and other current best practices
- Makes for a great resume addition
- Professional references available for successful contributions
- Development advice & support available for academic / freelance / personal projects using the Big Data Framework

Great alternative to larger, more competitive internship programs.