

# Big Data

Internet Research Group

April, 2012

Peter Christy [pchristy@irg-intl.com](mailto:pchristy@irg-intl.com)

John Katsaros [jkatsaros@irg-intl.com](mailto:jkatsaros@irg-intl.com)

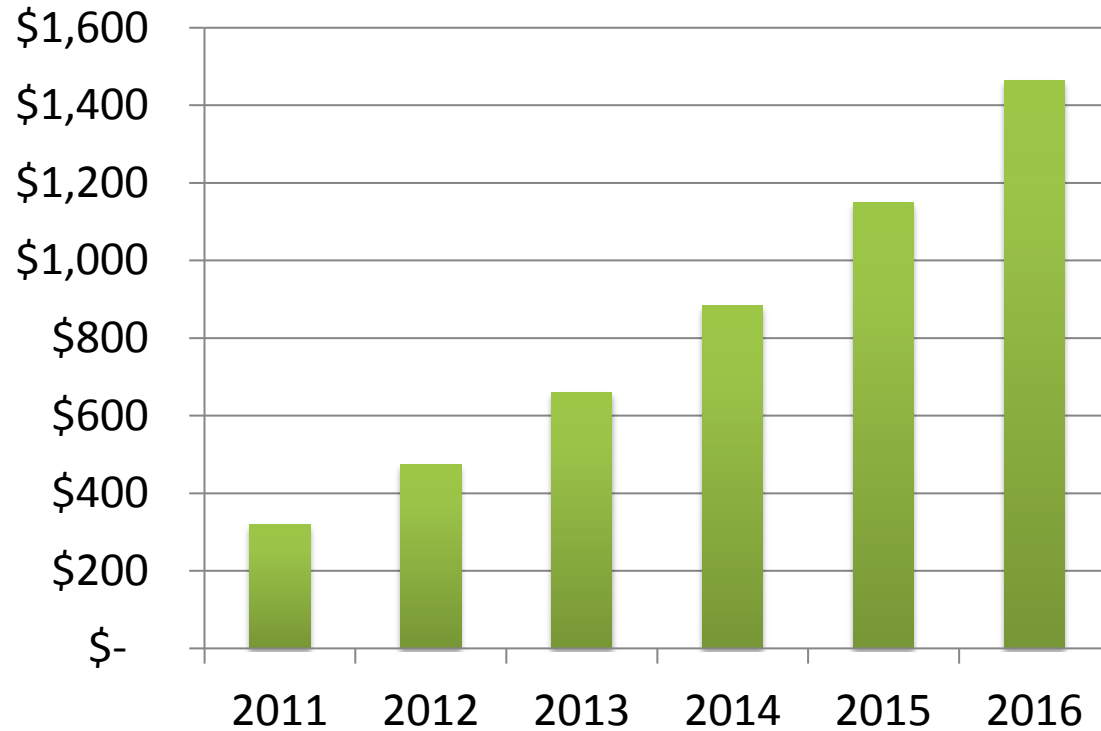
In 2011 organizations realized that they were sitting on an information goldmine. Rather than discarding datasets which seemed too costly to analyze, they have reached the point where a business can affordably analyze vast amounts of information and unlock valuable insights.

A new industry was launched – Big Data

Big Data can trace its roots to Google when in 2004 two Google engineers, Jeffery Dean and Sanjay Ghemawat, published a Usenix paper describing MapReduce – Simplified Data Processing on Large Clusters.

Subsequently Doug Cutting at Yahoo! developed an open source version named Hadoop.

## Big Data Market (M)



Source: IRG Research

**The Big Data Market hit the radar in 2011**

Big Data is often machine generated and includes, but is not limited to, click stream data, log files, (servers, network equipment, apps...) alerts (network, security devices ...) and social media content.



Click Stream



Logs



Alerts



Social Media

At the same time MapReduce was developing, server prices continued to drop and Amazon's AWS service began offering servers priced on a time basis – these events made computing more affordable, especially for Big Data applications.



Server = 8 cents/hour



Server Prices



The early adopters of Big Data ran large Web properties like e-Bay and Facebook – organizations that had a lot of click stream data and large numbers of registered users.

1. Personalize the visit – lead to longer visits
2. Spam mitigation – reduce annoyance
3. Suggesting friends you may know – increases member interconnections

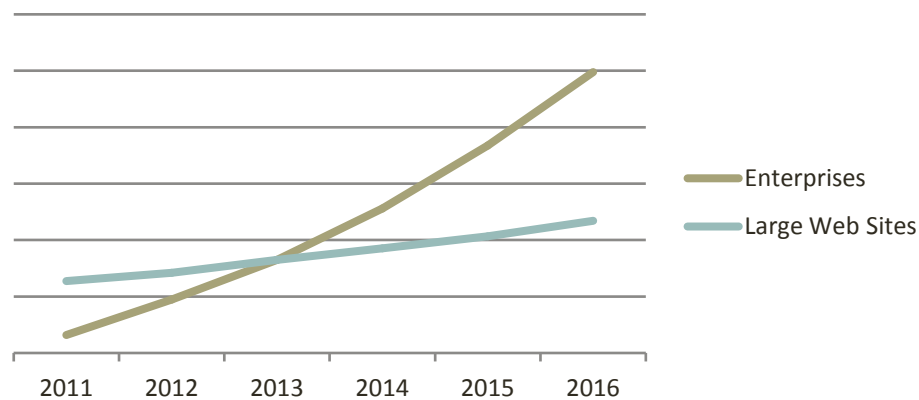


Recently enterprises which operate large Web sites have begun working with Hadoop and Big Data

1. Customize the visit – increase brand awareness
2. Suggesting products (e.g., financial services)
3. Cross selling/multi-channel (e.g., Disney) CRM

Enterprise adoption of Big Data will grow quickly and in 18 months spending will exceed the amount spent by large Websites.

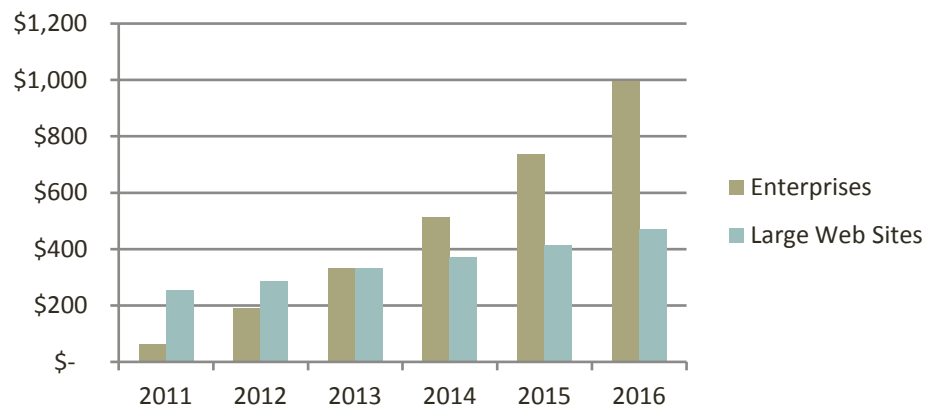
### Big Data Segment Spending Comparison



Source: IRG Research  
Source: IRG Research

## In 2013 Enterprise Big Data Spending will surpass Large Website Spending

### Big Data Segment Spending Forecast (M)

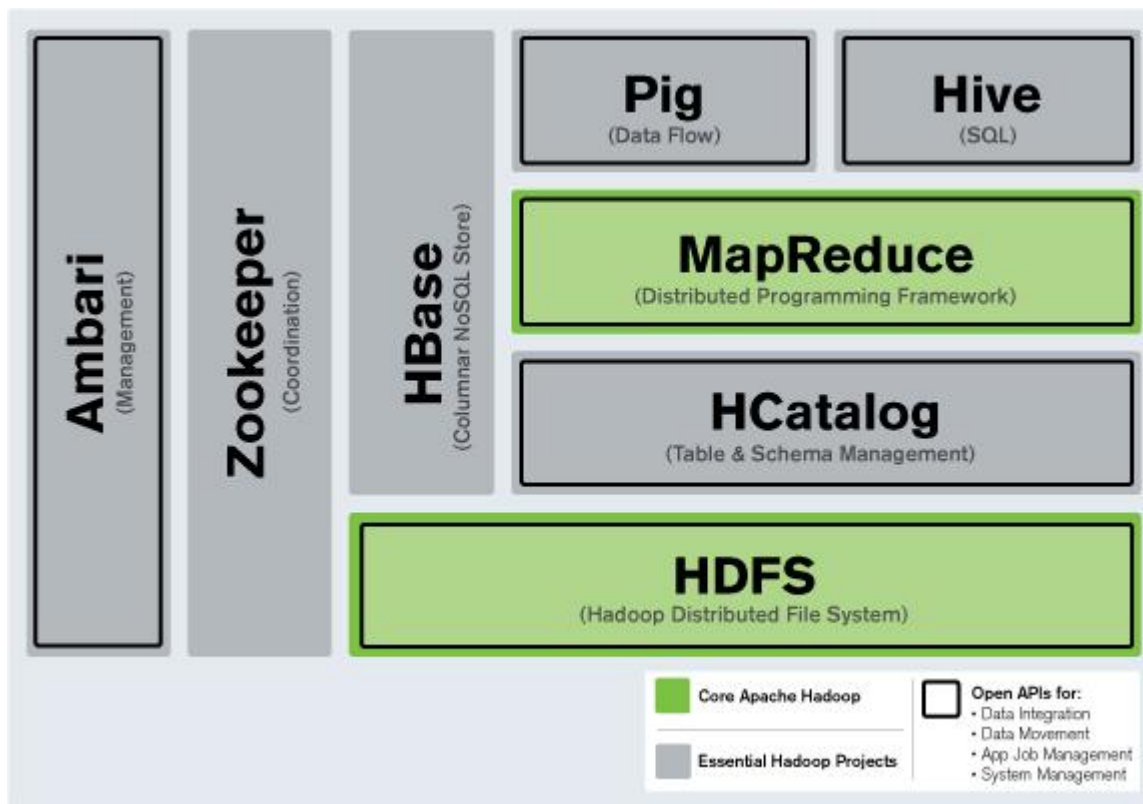


Source: IRG Research

Finally, when (and how) will SMB's and Enterprises not running large web properties adopt Big Data Technology?

1. When Business Intelligence products become useable and widely available
2. Through platforms like Splunk which already has 3,500 IT user organizations
3. When SMBs are presented with compelling value propositions

Meanwhile Big Data developers are extending functionality ranging from better easier to manage systems to higher performance systems to simplified Business Intelligence platforms.



Source: Hortonworks

There seems to be plenty of money available to fund Big Data companies

**Accel Partners Launches \$100mm Big Data Fund**

Accel's Big Data Fund aims to fund transformative early stage and growth companies throughout the Big Data ecosystem, from next generation storage and data management platforms to a wide range of revolutionary software applications and services – i.e. data analytics, business intelligence, collaboration, mobile, vertical applications and many more. We believe the future multi-billion software companies will be emerge from the Big Data ecosystem.

# The Hadoop Ecology

## Hadoop Releases



## Hadoop Infrastructure Management



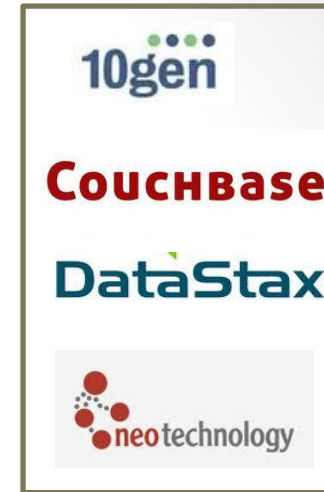
## Data Integration



## Big Data Analytics



## No SQL



## Data Presentation



## Our list of interesting Big Data startups

| Company                 | Investment    | Investors   | Location           |
|-------------------------|---------------|---|--------------------|
| 10gen                   | \$31M         | Sequoia, Flybridge and Union Square Ventures                      | Redwood Shores, CA |
| BackType                | \$1.32M       | YCombinator, True Ventures, lowercase, Freestyle                  | San Francisco, CA  |
| Cloudera                | \$76M         | Greylock, Meritech Capital, Accel, Ignition Partners              | Palo Alto, CA      |
| Couchbase               | \$30M         | Accel Partners, Ignition Partners, Mayfield, North Bridge, Docomo | Mountain View, CA  |
| DataStax                | \$13.7M       | Lightspeed, Crosslink Capital                                     | Burlingame, CA     |
| Datameer                | \$11.75M      | Kleiner Perkins, Redpoint   | San Mateo, CA      |
| Hadapt                  | \$9.5M        | Norwest Venture Partners, Bessemer Venture Partners               | Cambridge, MA      |
| HortonWorks             |               | Benchmark Capital   | Sunnyvale, CA      |
| HStreaming              |               |   | Chicago, IL        |
| Karmasphere             | \$5M          | Hummer, Winblad, USVP   | Cupertino, CA      |
| Kitenga                 |               |   | Santa Clara, CA    |
| MapR Technologies       | \$9M          | Lightspeed, NEA   | San Jose, CA       |
| Mintigo                 | \$9M          | Sequoia, Giza   | Menlo Park, CA     |
| Neo Technologies        | \$10.6M       | Fidelity, Sunstone, Conor   | Menlo Park, CA     |
| Pentaho                 | \$32M         | Benchmark, Index, NEA   | Orlando, FL        |
| StackIQ                 | \$3M          | Antham, Avalon  | La Jolla, CA       |
| Talend                  | \$28M         | Balderton Capital, AGF Private Equity, Galileo Partners           | Los Altos, CA      |
| Tableau Software        | \$15M         | NEA, Meritech Capital Partners                                    | Seattle, WA        |
| <b>Total Investment</b> | <b>\$284+</b> |   |                    |



# And the Venture Companies that Funded these startups

| Company           | Selyo | Urbano | Ycombinator | Truora | Forrester | Greylock | Accel | Ignition | Northbridge | DLPI | MLP | CRS | Redpoint | NEA | Bessemer | Hummer | USVP | NEA | AVP | BAI | AGF | Index | Fidelity | Sunstone | CGI |
|-------------------|-------|--------|-------------|--------|-----------|----------|-------|----------|-------------|------|-----|-----|----------|-----|----------|--------|------|-----|-----|-----|-----|-------|----------|----------|-----|
| 10gen             | X     | X      | X           |        |           |          |       |          |             |      |     |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| BackType          |       |        | X           | X      | X         | X        |       |          |             |      |     |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| Cloudera          |       |        |             |        | X         | X        | X     | X        |             |      |     |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| Couchbase         |       |        |             |        | X         | X        | X     | X        | X           |      |     |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| DataStax          |       |        |             |        |           |          |       |          | X           | X    |     |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| Datameer          |       |        |             |        |           |          |       |          |             | X    | X   |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| Hadapt            |       |        |             |        |           |          |       |          |             |      | X   | X   |          |     |          |        |      |     |     |     |     |       |          |          |     |
| HortonWorks       |       |        |             |        |           |          |       |          |             |      |     |     | X        |     |          |        |      |     |     |     |     |       |          |          |     |
| HStreaming        |       |        |             |        |           |          |       |          |             |      |     |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| Karmasphere       |       |        |             |        |           |          |       |          |             |      |     |     |          |     | X        | X      |      |     |     |     |     |       |          |          |     |
| Kitenga           |       |        |             |        |           |          |       |          |             |      |     |     |          |     |          |        |      |     |     |     |     |       |          |          |     |
| MapR Technologies |       |        |             |        |           |          |       |          | X           |      |     |     |          |     |          |        | X    |     |     |     |     |       |          |          |     |
| Mintigo           | X     |        |             |        |           |          |       |          |             |      |     |     |          |     |          |        |      |     |     |     |     |       |          |          | X   |
| Neo Technologies  |       |        |             |        |           |          |       |          |             |      |     |     |          |     |          |        |      |     |     |     |     |       | X        | X        | X   |
| Pentaho           |       |        |             |        |           |          |       |          |             |      |     |     |          | X   |          | X      |      |     |     |     |     | X     |          |          |     |
| StackIQ           |       |        |             |        |           |          |       |          |             |      |     |     |          |     |          |        |      |     | X   | X   |     |       |          |          |     |
| Talend            |       |        |             |        |           |          |       |          |             |      |     |     |          |     |          |        |      |     |     | X   | X   | X     |          |          |     |
| Tableau Software  |       |        |             |        |           | X        |       |          |             |      |     |     |          |     |          |        | X    |     |     |     |     |       |          |          |     |

Thank You

John Katsaros [jkatsaros@irg-intl.com](mailto:jkatsaros@irg-intl.com)