Make a Data Driven Investment Decision About Big Data

Matt Farrellee, Director, Emerging Technologies, Red Hat
Irshad Raihan, Sr. Principal, Big Data Product Marketing, Red Hat
So What Is Big Data, Anyway?
What Makes Big Data Different?

“Big” is the easy problem to solve

- Analytics: Single Pass ➔ Iterative
- Data Formats: Consistent ➔ Inconsistent
- Queries: Standard ➔ Ad hoc
The Old Data Lifecycle

Multiple Silos. Multiple Views. Multiple Goals.

- Data Center Operator: Manage
- Architect: Build
- App Developer: Code
- Business User: Query
The New Data Lifecycle

One Language. One View. One Goal.

Data Center Operator

Business User

Ingest

Integrate

Act

Discover

Architect

App Developer
Barriers to Big Data Success

Lack of agile, open, and cost effective enterprise-grade solutions

I am locked into a vendor stack
Data Center Operator

I want more than canned BI queries
Business User

I want to use my favorite dev framework
App Developer

I need to integrate data across silos
Architect
Big Data Innovation Cannot Happen in a Bubble
## The Red Hat Way Applied to Big Data

<table>
<thead>
<tr>
<th>Category</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Open</strong></td>
<td>- Workload optimized, modular blocks</td>
</tr>
<tr>
<td></td>
<td>- Completely interoperable with partner ecosystem</td>
</tr>
<tr>
<td><strong>Agile</strong></td>
<td>- Agile data integration and application development</td>
</tr>
<tr>
<td></td>
<td>- IaaS and PaaS options for development and deployment</td>
</tr>
<tr>
<td><strong>Enterprise Ready</strong></td>
<td>- Resiliency - scalability, high availability, disaster recovery</td>
</tr>
<tr>
<td></td>
<td>- Security - comprehensive, normalized role-based access</td>
</tr>
<tr>
<td><strong>Cloud Ready</strong></td>
<td>- Decoupled application, data integration and infrastructure</td>
</tr>
<tr>
<td></td>
<td>- Seamless transition from physical &amp; virtual to cloud</td>
</tr>
<tr>
<td><strong>Cost Effective</strong></td>
<td>- Open Source software and communities</td>
</tr>
<tr>
<td></td>
<td>- Skills and tools reuse</td>
</tr>
</tbody>
</table>
Choosing The Right Big Data Technology

- Monolithic, Scale up
- Distributed, Scale Out

- Schema on Write
- Schema on Read

- Consistency, Integrity
- Scale, Data Access

- Offline, Batch
- Online, Real Time