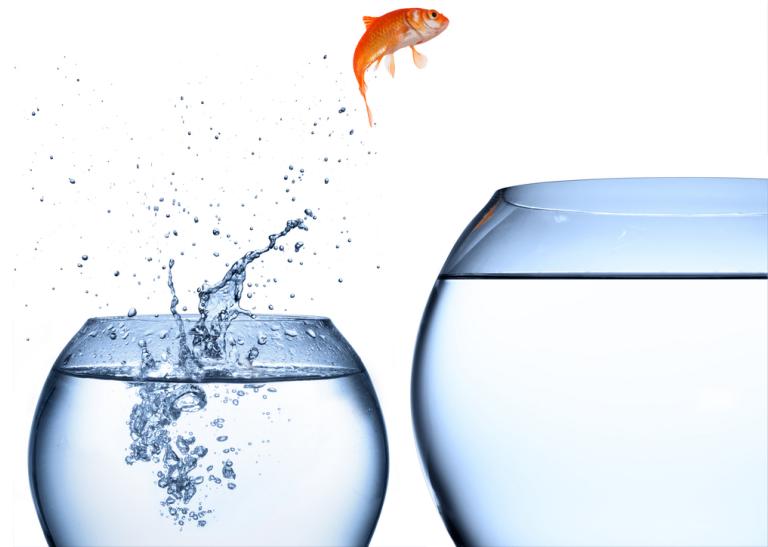


BIG DATA

EMANUELE DELLA VALLE

18 MAGGIO 2018



How Macy's innovates for scale, speed and simplicity

The original presentation as well as the video with the interviews are available online at

<https://it.hortonworks.com/blog/how-macys-big-data-leader-innovates-for-scale-speed-and-simplicity/>

Today's Speakers



Eric

VP, Industry Solutions

Hortonworks



Seetha

Director, Analytic & CRM Systems

Macy's



Dan

Director Product Marketing

Tableau

Bruno

CMO

AtScale



3



Today



- Why BI on Hadoop?
- Where Do I Get Started?
- Lessons Learned



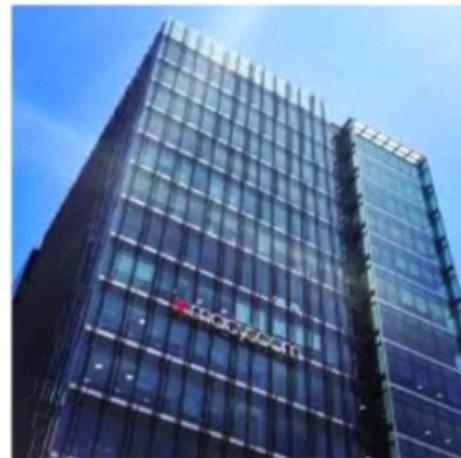
Hortonworks



HOW...do you know you need to change?

Macy's by the Numbers

- ❑ Premier omni-channel retailer
- ❑ 150 year-old Fortune 500 company
- ❑ Macy's and Bloomingdale's brands
 - ❑ ~ 870 stores in 45 states
- ❑ Online division...Macys.com
- ❑ Fiscal 2015 sales of \$27 billion
 - ❑ ~ 167,000 employees

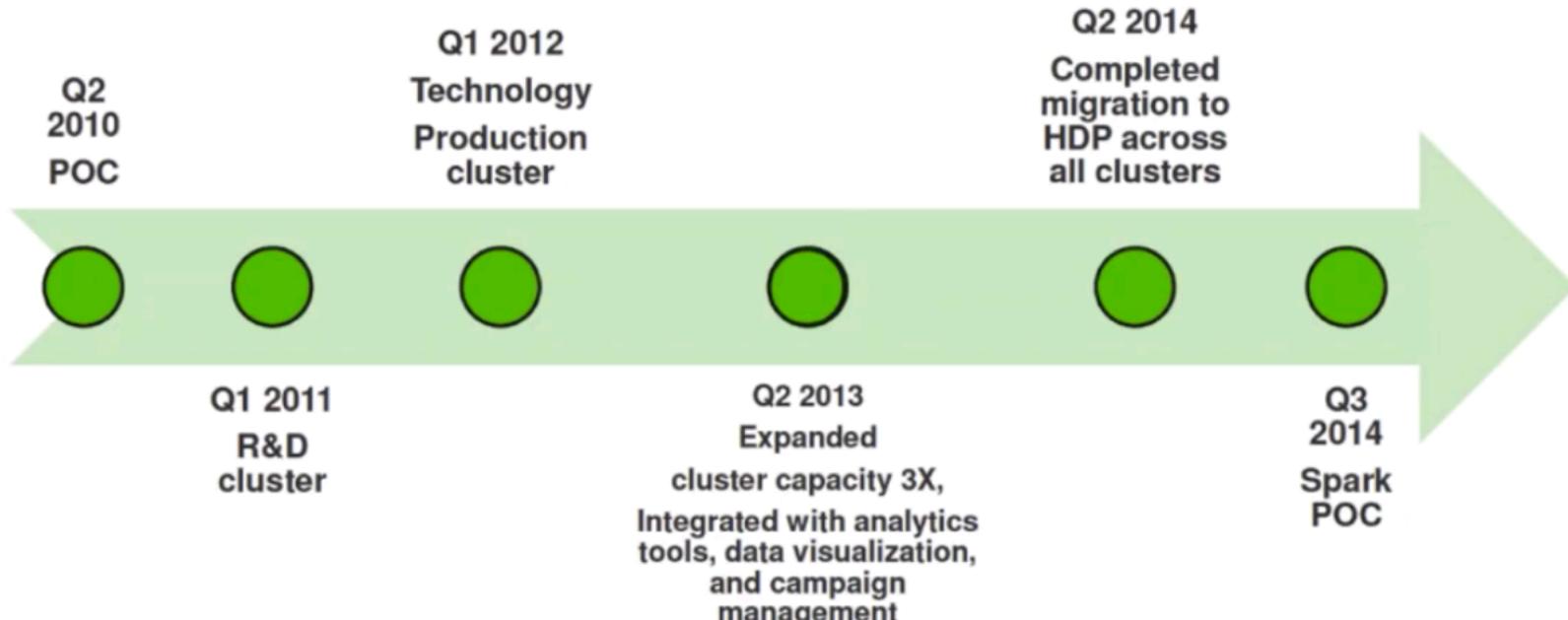


Data & Platform Challenges



- ❑ Big volume, unstructured data processing
not supported by RDBMS
- ❑ Archived data inaccessible → can't analyze
- ❑ Large analytic queries run off-hours, by DBAs
- ❑ Large analytic queries / model builds limited several per week
 - But 100s queries queued per day
- ❑ Predictive analytics taking off
 - But unable capitalize on it with traditional system limitations

Hadoop Journey @ Macy's



Typical BI on Hadoop Journey



Operational Efficiency



Cheap Storage
ETL Acceleration

Rationalization & Agility



EDW Optimization
Data Science

Information Advantage



Exploration
Self-Service BI
Traditional Reporting Analytics

Complimentary Maturity Assessment @ atscale.com/maturity

WHEN...do you know you need BI on Hadoop

WHAT...do you do to get started with BI on Hadoop?

Signs You Need BI on Hadoop



- Data Discovery and Optimization Needs**
 - Understand what data lives in Hadoop
 - Early detection technical issues + failures after new releases
- Marketing, Sales and Product Insight Needs**
 - Segmentation Analysis for better targeting, in real-time
 - Customer Journey analysis across devices, properties, brands
- Users Needs Outgrown Ability to Deliver Insights**
 - Current BI tools and process don't scale
 - Users demand integration for multiple BI tools (Tableau, Excel, Custom...)

Key Considerations

- ❑ Data Discovery and Optimization Needs
 - Avoid Data Movement
 - “Freedom within a framework”
- ❑ Marketing, Sales and Product Insight Needs
 - Optimize for One Semantic Layer
 - Provide Re-usable data assets
- ❑ Users Needs Outgrown Ability to Deliver Insights
 - Preserve existing BI Investment (tools and skillset)
 - Reduce one-off and processes that require IT prepare data for business-users

Use Cases to Prioritize (Business)



- ❑ Getting right message to the right customer at the right time
- ❑ High-performance analytics with near real-time data processing
- ❑ Enhance website with real-time product recommendations
- ❑ Expand analytics to larger org w/o constraints and limitations
- ❑ Support advanced analytics with additional tools and processes

Use Cases to Prioritize (IT)



- Store, process, analyze large volumes data
- Ad-hoc data mining and complex analytics
- Unstructured and big volume data processing
- Model building and optimization at scale
- Support for ad-hoc analysis, reporting and data visualization

WHAT...questions do you need to ask?

What is most important from Business standpoint?

- Ad-hoc analytics platform scalable for big data
- Operationalize insight
- Plug-in to business execution
- Analyze without limitations
- Leverage all forms enterprise data
- Analytical insights through mining, reporting, visualization

BI on Hadoop...for Technology



What is important from Technology standpoint?

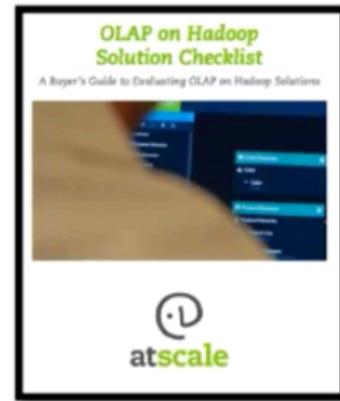
- High volume processing/throughput beyond RDBMS capacity
- Linear scalability
- Large storage capacity
- Ease of use (SQL support)
- Cost effectiveness or Maintainability
- Power of open-source community tools

BI on Hadoop Checklist



Your Guide to Key Considerations Along the Way

1. Installation & Configuration
2. Compatibility
3. Maintenance
4. Performance & Scalability
5. Functionality



WHAT...did you learn?



Key Takeaways



1. Data Scientists Beware

- Machine Learning '1' Big Data use-case...there are more...

2. Resource Flexibility

- Don't go it alone
- Tap into BI & Hadoop community experts to help ramp



Key Takeaways – cont'd



3. Perspective and Action

- Data explosion can be competitive advantage & innovation
- Get ahead - Manage - Operationalize - Optimize

4. Focus

- No need to boil the ocean - crawl, walk, run.

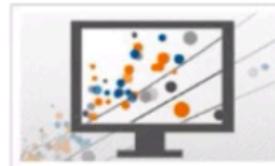
Next Steps



BI on Hadoop Cheat Sheet: 5 Things You Should Know
<http://atscale.com/5-things>



Do's and Don'ts of “BI on Hadoop”
<http://info.atscale.com/do-bi-on-hadoop>



Request a demo
www.atscale.com/try

Bonus contents :-P

