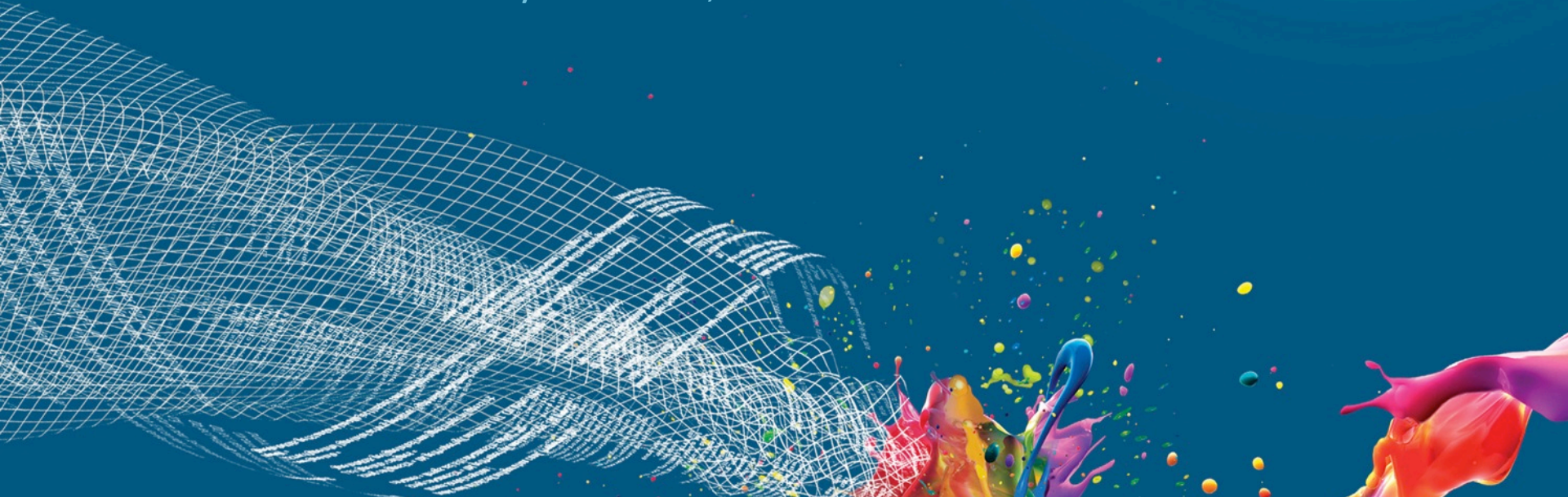


What's Holding Up Your Hadoop?

Eddie Garcia - Security Architect, Office of the CTO



What's Holding Up Your Hadoop?

“It’s too complicated.”

“I don’t have the right resources.”

“Security and compliance are never going to approve this.”



5 steps to a successful Hadoop project from POC to production

Solving business problems in a repeatable
and measurable fashion

Step 1: Identify a real Hadoop business use case

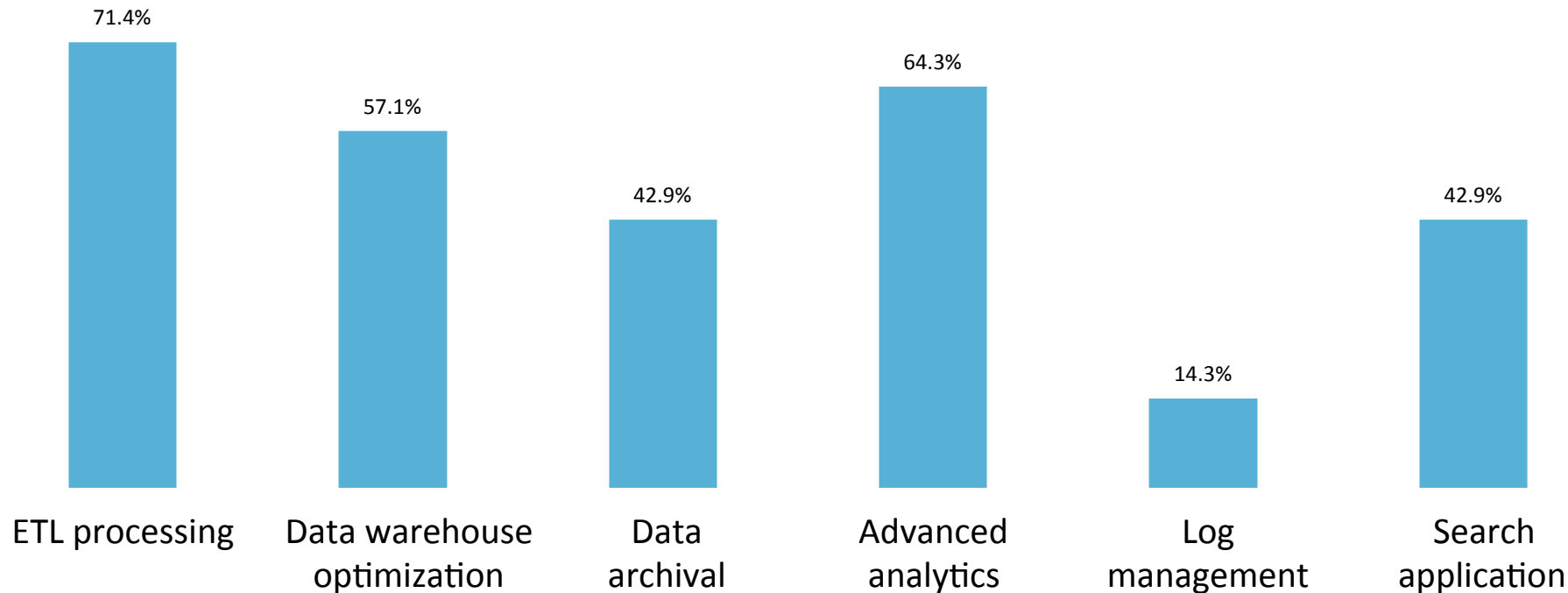


Hey... no one's using
Big Data/NoSQL stuff
around here, right?



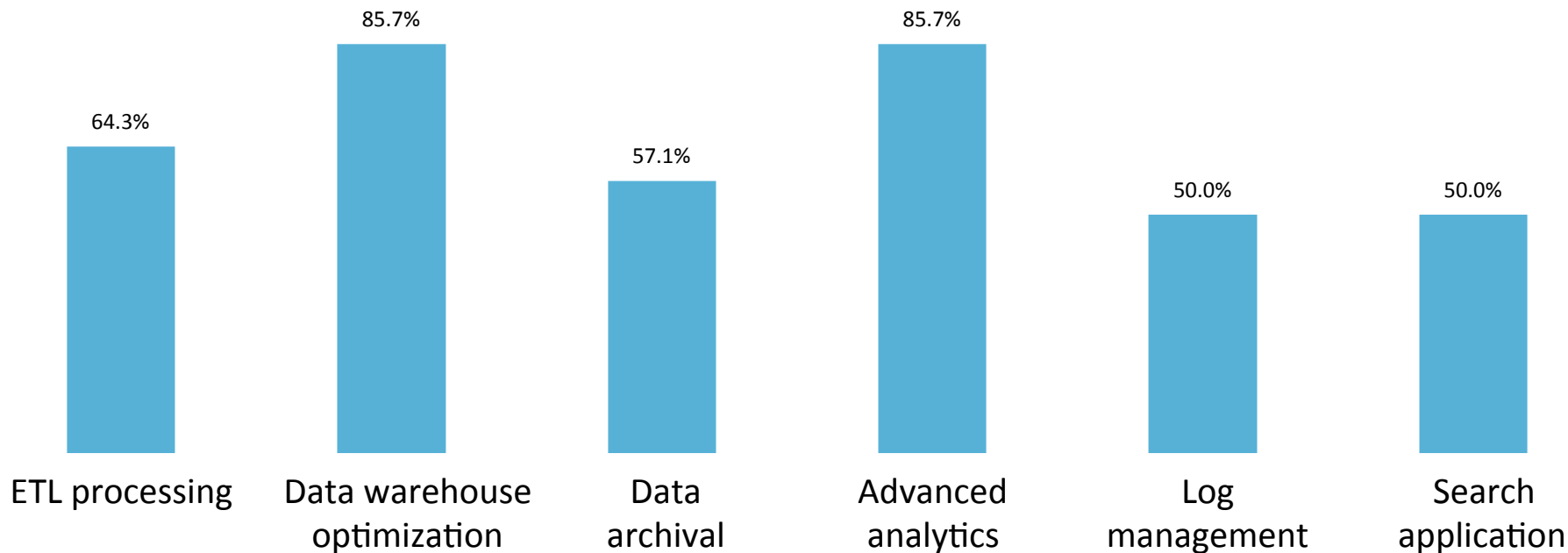
Usage of Hadoop – today

What are you currently using your Hadoop infrastructure for?



Usage of Hadoop – next 6-12 months

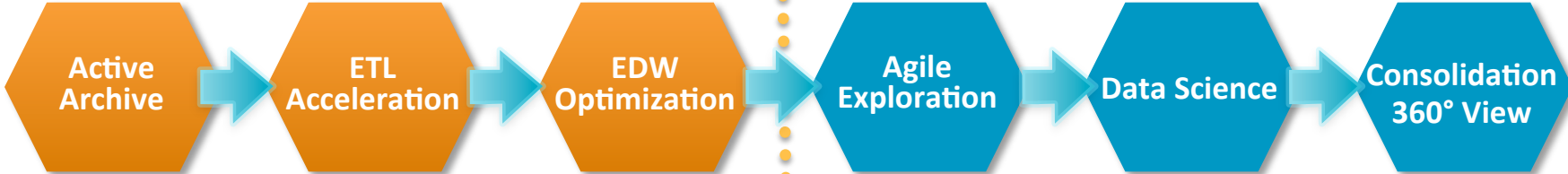
How do you anticipate using your Hadoop infrastructure for in the next 6-12 months?



Journey to Achieve Full Potential

Operational Efficiency

Information Advantage



IT



Business

Step 2: Get ahead of the potential challenges



Top Challenges Hadoop Operations

1. Integration with other IT management tools
2. Diagnostics and troubleshooting
3. Configuration Management

Top Challenges Organization Issues

1. Too entrenched in existing technology
2. Lack of skills - operational
3. Lack of understanding of Hadoop and its business benefits

Top Challenges

Data Warehouse Optimization

1. Lack of SQL compliance to ANSI standards
2. Lack of supported data types
3. Lack of workload resource management

Top Challenges Advanced Analytics

1. Lack of robust analytic tools integration
2. Lack of workload resource management
3. Lack of data analytics skilled resources


Top Challenges

Security and Compliance

1. Lack of unified security model for Hadoop
2. Native Hadoop audit and access management
3. Kerberos setup and management

Step 3: Secure your data assets





SECURITY ASSERTION

Protect
ALL the data.

Data Breaches Happen

SECURITY

Home Depot Data Breach Could Be the Largest

By NICOLE PERLROTH SEPTEMBER 8, 2014 6:58 PM 130 Comments



Ebay faces a class action lawsuit

John Ribeiro @JohnRibeiro

Ebay faces a class action lawsuit.

The consumer privacy attorney says the state of Louisiana has inadequate security measures for its customers.

The e-commerce giant is continuing to cause controversy according to the court.

Ebay informed users that it had no evidence that it was no evidence that it advised users to contact their credit card companies to contain the breach.

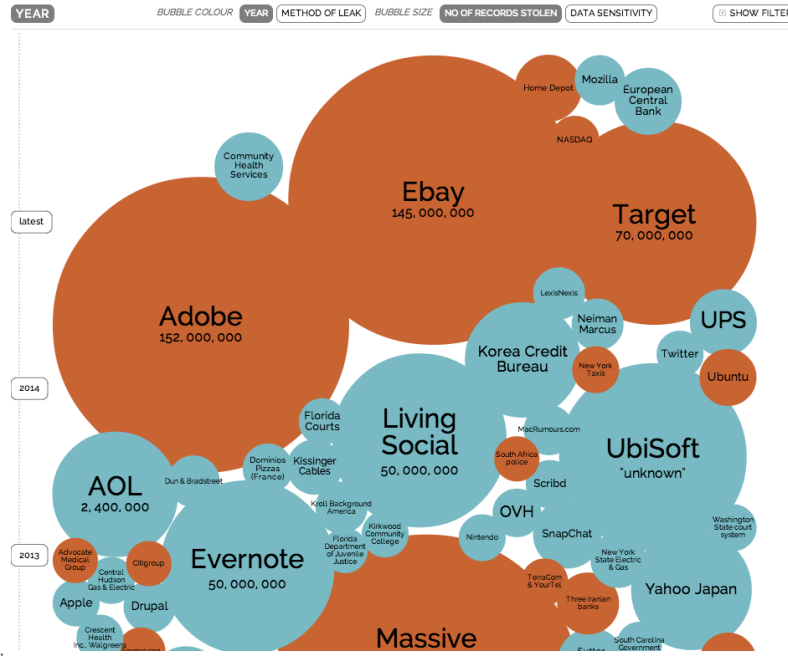
"The thieves had access to passwords, email addresses and credit card numbers at a minimum," according to a lawyer in the District of Louisiana.

- E-MAIL
- FACEBOOK
- TWITTER
- SAVE
- MORE

IN THEATERS OCT 3

World's Biggest Data Breaches

Selected losses greater than 30,000 records



Each Bulletin: Credit Card Breach at 33 P.F. Chang's Locations

The breach more extensive than disclosed

Share 30 Facebook 7 Email Print



Picture illustration taken in Vienna July 9, 2013. Picture taken July 9, 2013.

(Reuters) - Adobe Systems Inc said on Tuesday that the scope of a cyber-security breach disclosed nearly a month ago was far bigger than initially reported, with attackers obtaining data on more than 38 million customer accounts.

of July

he s China 7 were me 11— breach elease. d eir t card an 18 states, with eight

Regulations Dictate Data

A set of laws, guidelines, and specifications that dictate how a business must operate



HIPAA
Compliance



Common Pain Points of Hadoop Security

Perimeter

- Manual setup of separate Kerberos cluster
- Manual, error-prone configuration
- Ongoing maintenance

Access

- Redundant security policies for every access path
- Manual mirroring
- Threat of rogue insiders

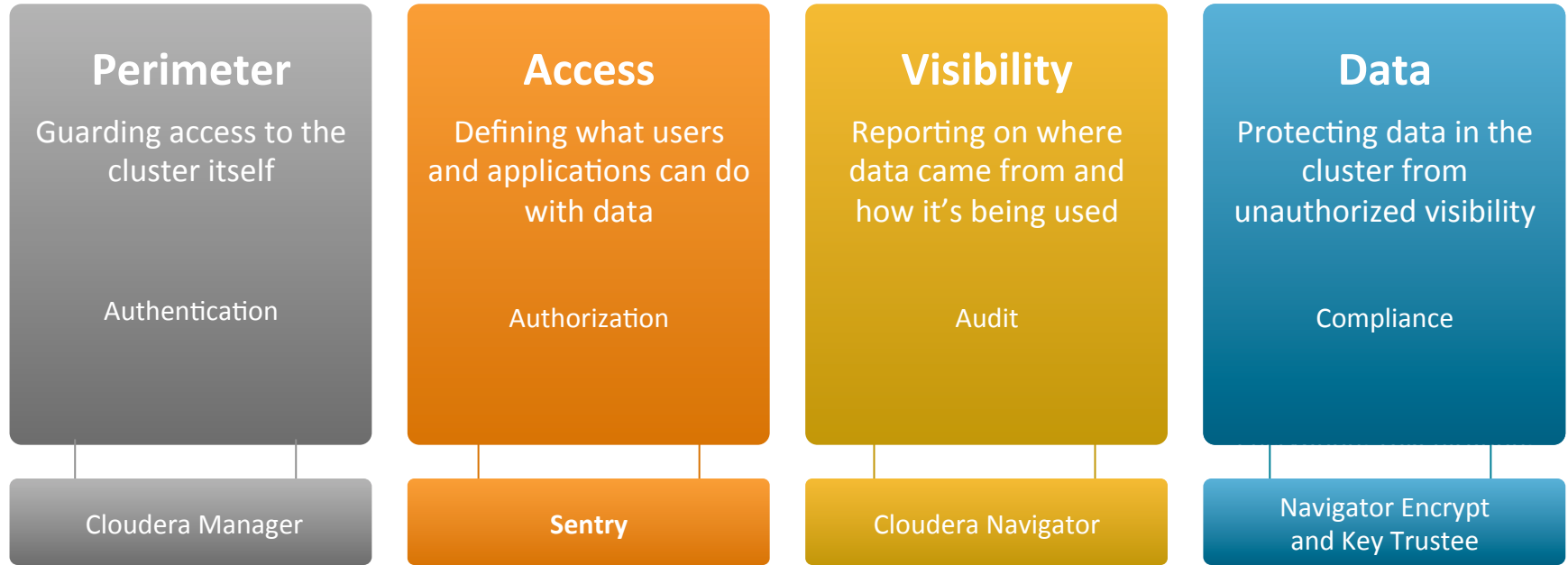
Visibility

- No unified audit trail with manual point-in-time user access history
- Cannot meet core data governance requirements with limited lineage

Data

- Risk of data breach
- Limited integration with existing systems
- Cost of encryption and key management

Cloudera - Comprehensive Hadoop Security



Step 4: A Big Data Center of Excellence



Why Build a Center of Excellence?

Silos Obstruct Scale, Prevent Flexibility, and Drive Up Costs



**Novice
Skill Level**



**Redundant
Infrastructure**



**Incompatible
Systems**



**Disparate
Data**

Big Data COE Program Roles

Staff Centrally and Train to Scale

Management & Leadership



Program Manager



Executive Sponsor



Big Data Visionary

Business & Data



Lead Data Scientist



Data Wranglers



Lead Business Analyst



LOB Rep



LOB Rep



LOB Rep

Technology & Ops



Data Warehouse Specialist



Developers



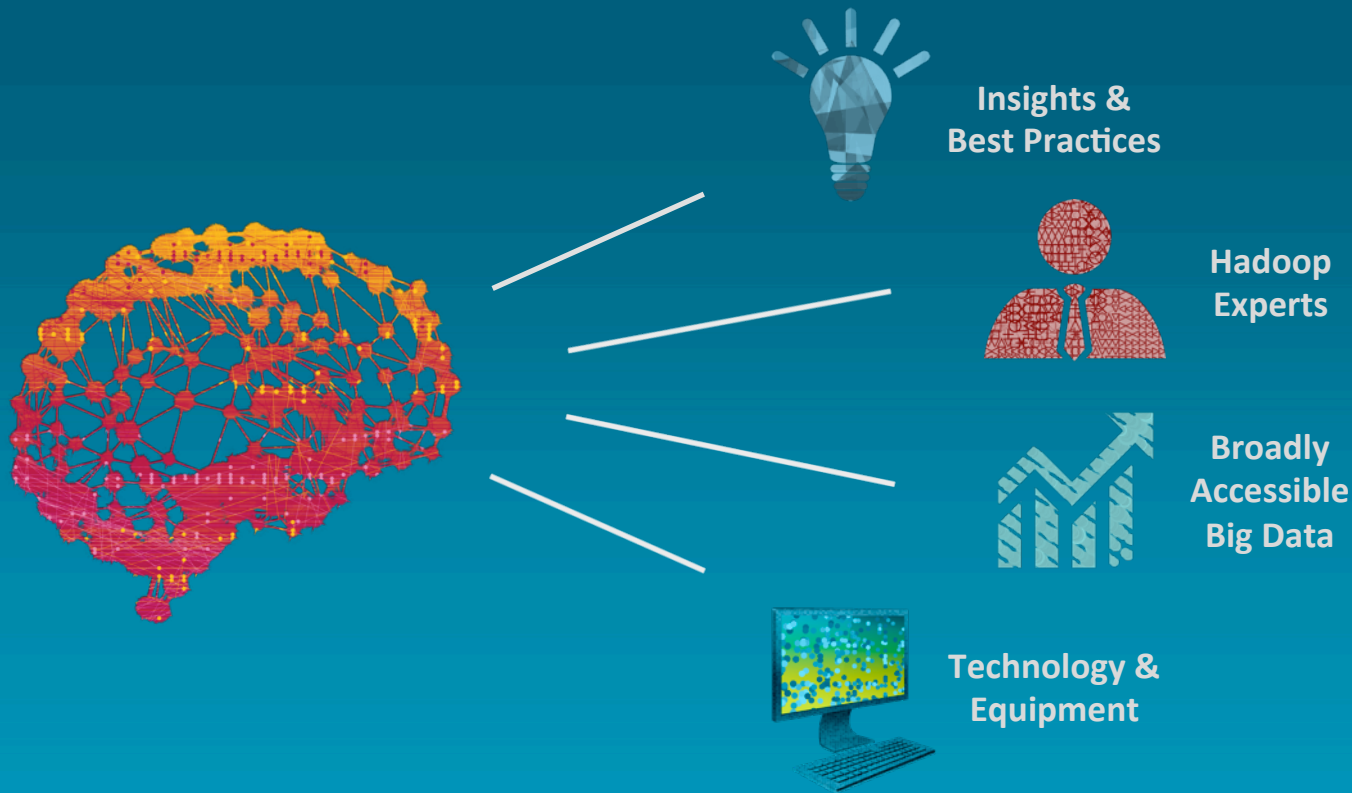
Architects



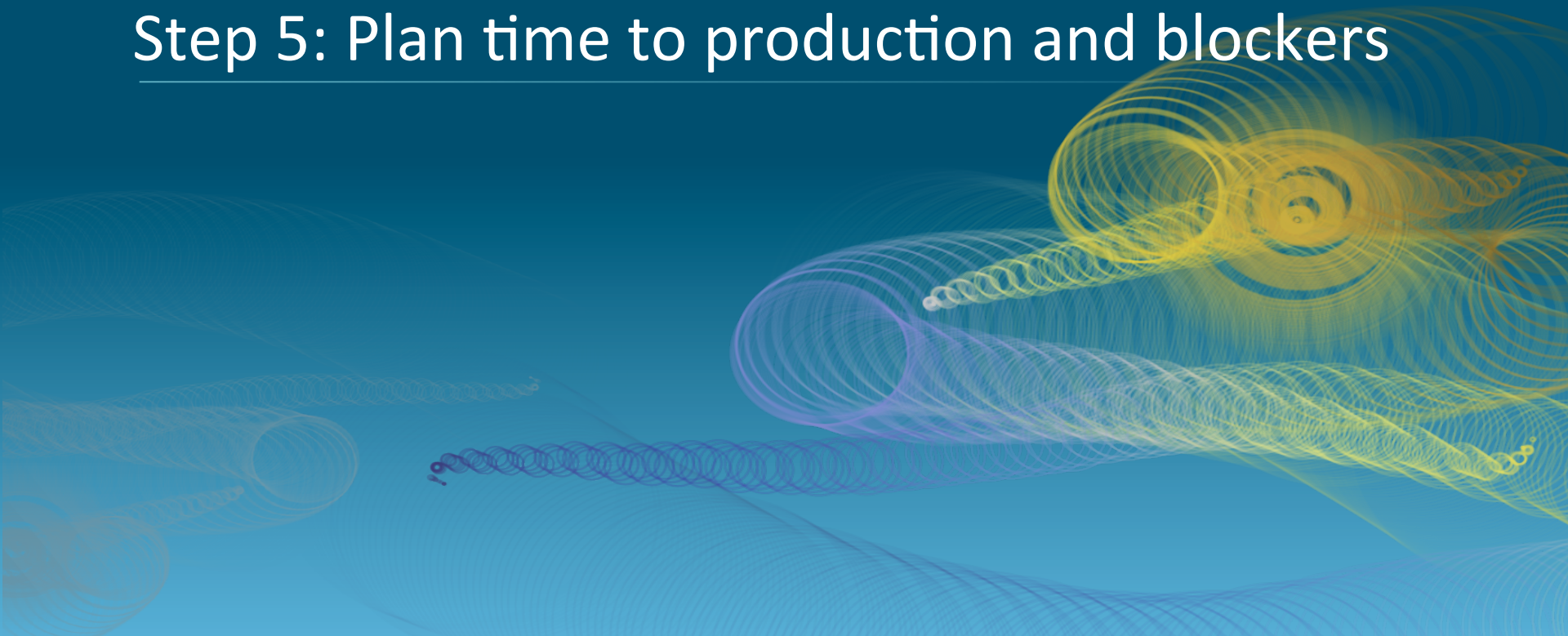
IT & InfoSec

Big Data as a Service Drives Success

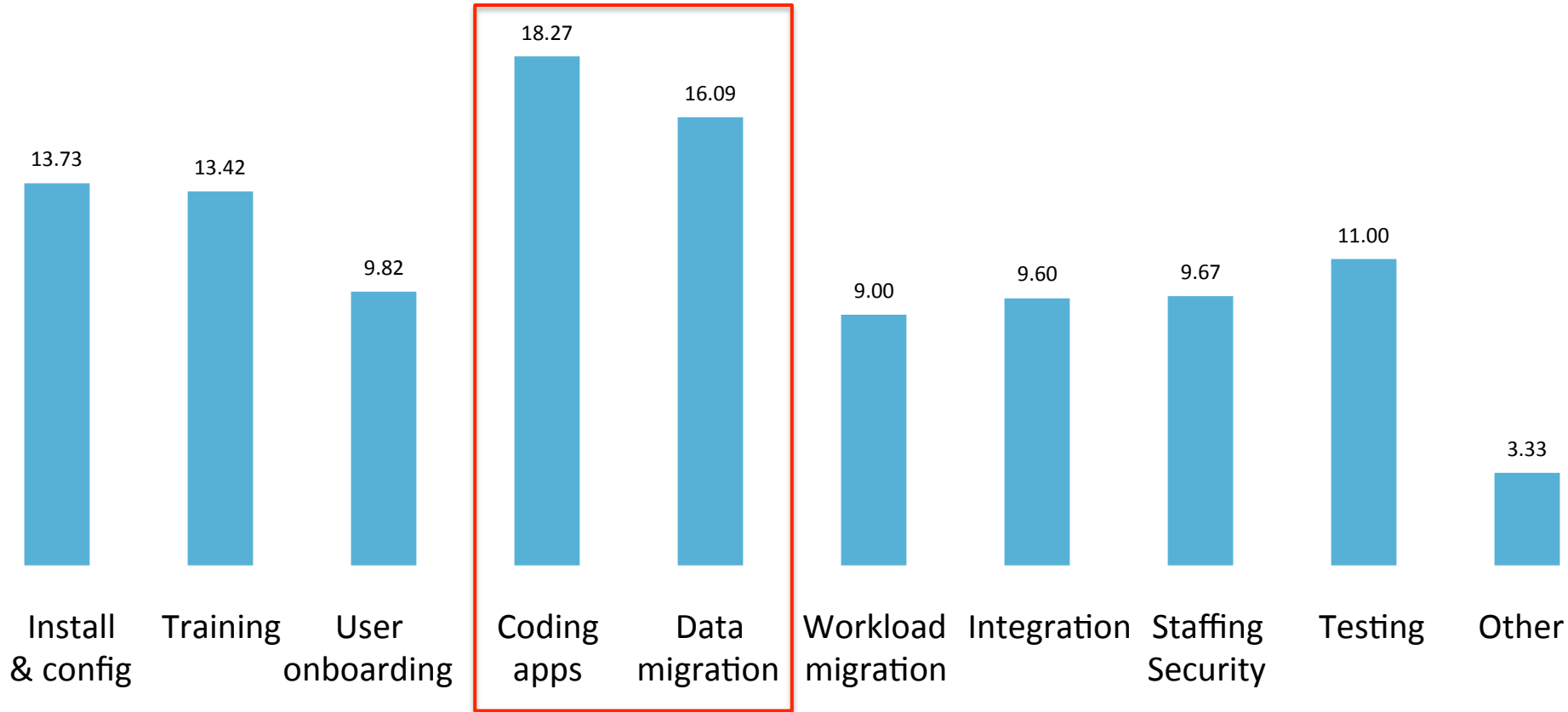
Coordinate Workloads to Distribute Advantages Widely and Seamlessly



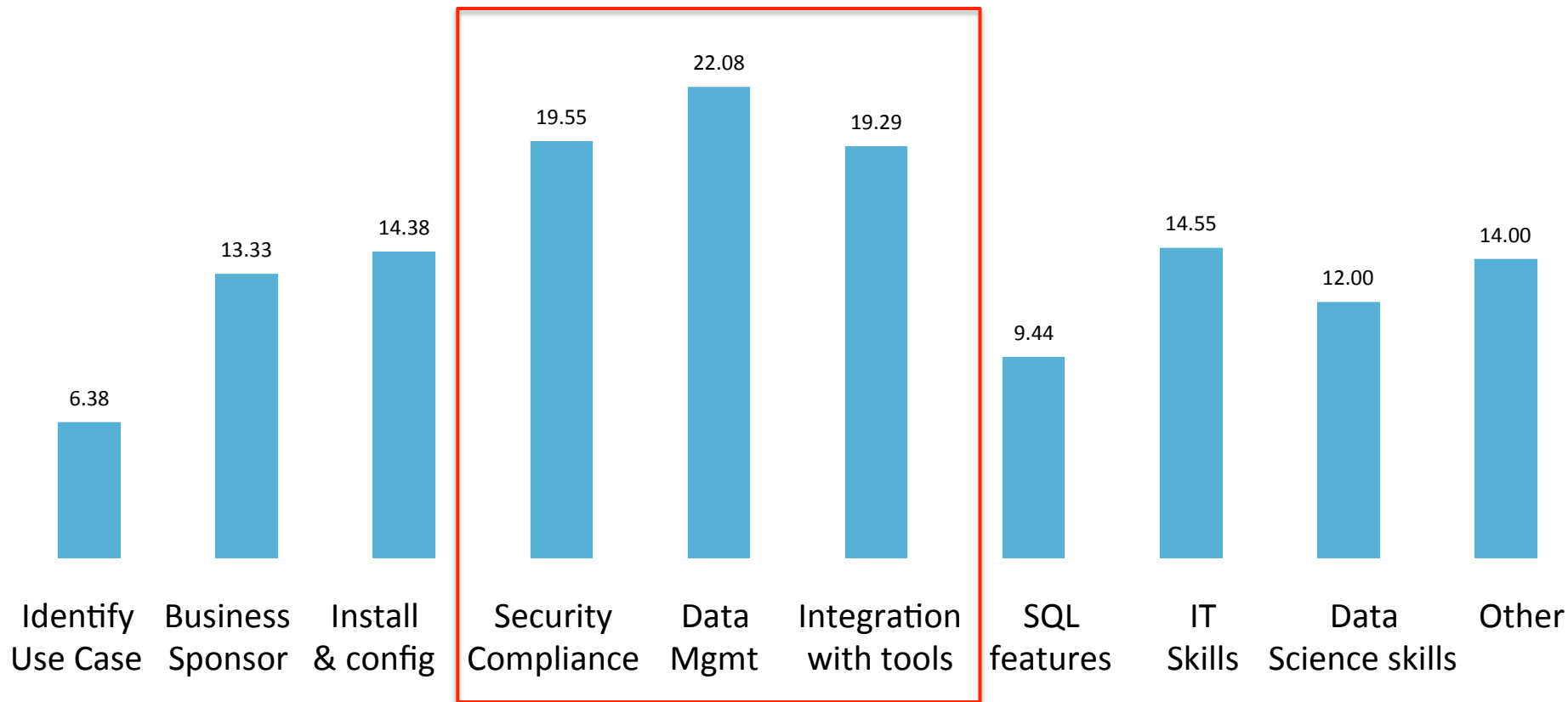
Step 5: Plan time to production and blockers



Percentage of time spent to Production-Ready



Blockers to Production-Ready



Recap



5 Steps for a Production-Ready Hadoop

1. Identify a real Hadoop business use case
2. Get ahead of the potential challenges
3. Secure your data assets
4. A Big Data Center of Excellence
5. Plan time to production and blockers

Step 6: Ask Bigger Questions





cloudera[®]

Ask Bigger Questions

Thank You!

Eddie Garcia

eddie@cloudera.com