



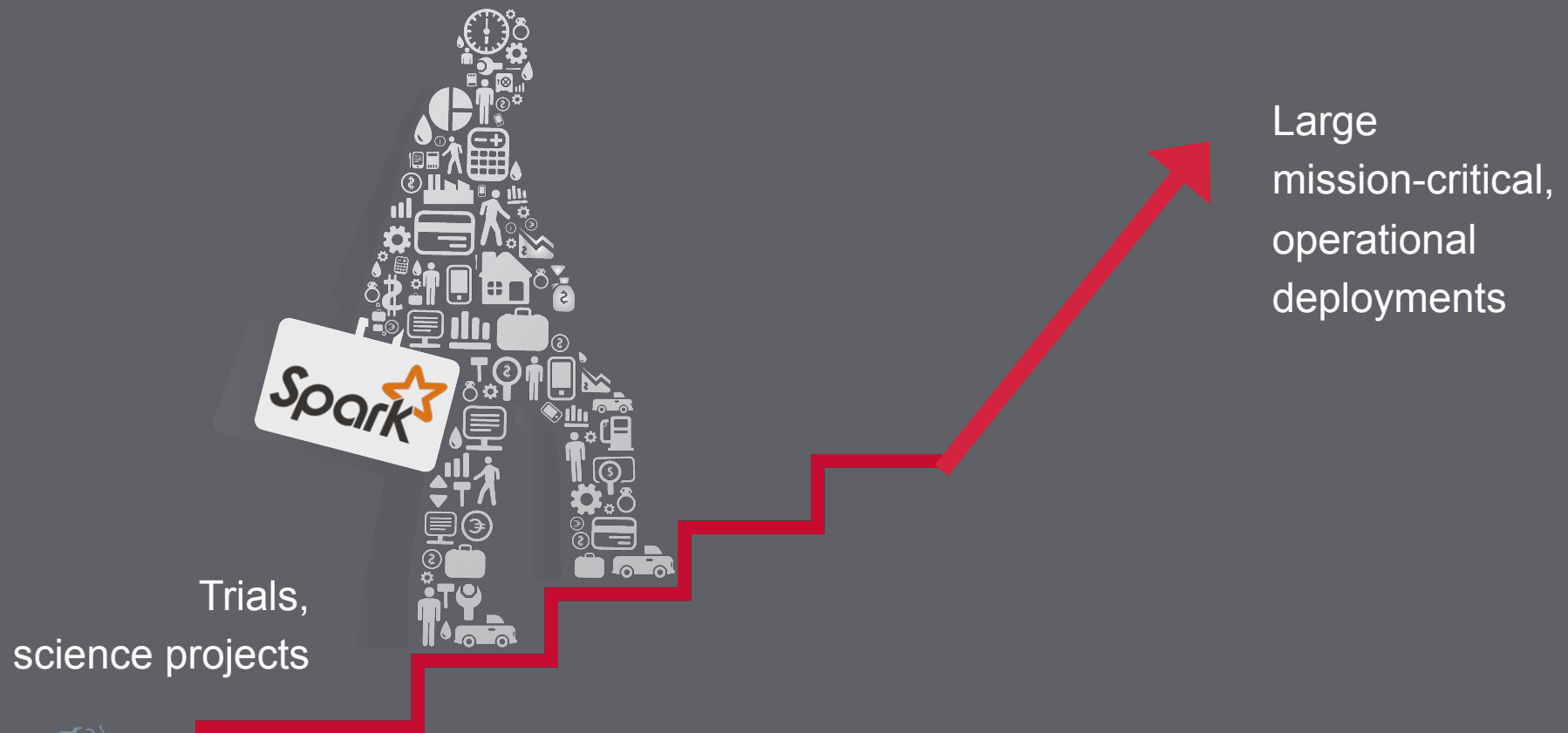
Taking Your Spark To Production Scale

Anil Gadre, SVP Product Management, MapR Technologies

June 15, 2015



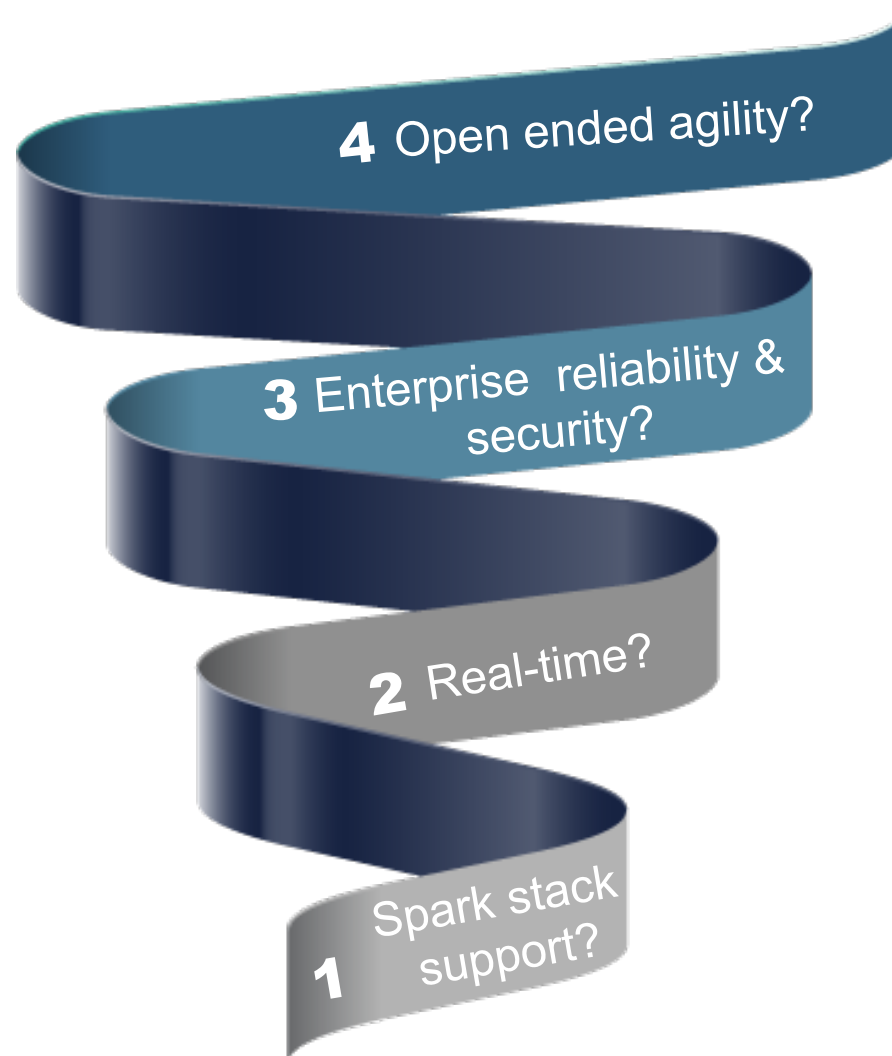
The Journey To Production Scale



Companies with Spark & MapR in Production



Key Issues To Plan For





CISCO™

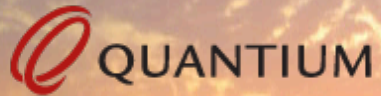
Security Intelligence Operations

Global Managed Security Services
delivered on Hadoop

Spark Stream processing used to
first check for known threats

Data next processed on Hadoop
using MLlib and GraphX

Additional SQL querying done via
Spark SQL



Business transformation
from data to dividend

Delivers Lightning Fast Analytics for Clients

Building largest Hadoop cluster
in Australia

Real-time analytics using Spark on
MapR—reducing data loading time
from hours to minutes

Leverage multi-tenancy,
high-performance and reliability
of MapR

A hand holding a test tube with orange liquid, set against a blurred background of a person in a white lab coat.

NOVARTIS Next-Gen Genomics

Develop flexible platform to keep up with fast changing research techniques

POSIX file access lets bio-informaticians use existing tools with open source tools (Spark)

Graph manipulations can be done reliably and at scale using Spark



Real-Time Customer Analytics

- MapR Data Lake stores both online and archive data
- Spark on MapR reduced ETL processing
- NFS moved data into the cluster seamlessly
- 1/10th Total Cost of Ownership vs. old way
- New customer onboarding cut from months to weeks

Databricks & MapR Strategic Partnership

(since April 2014)



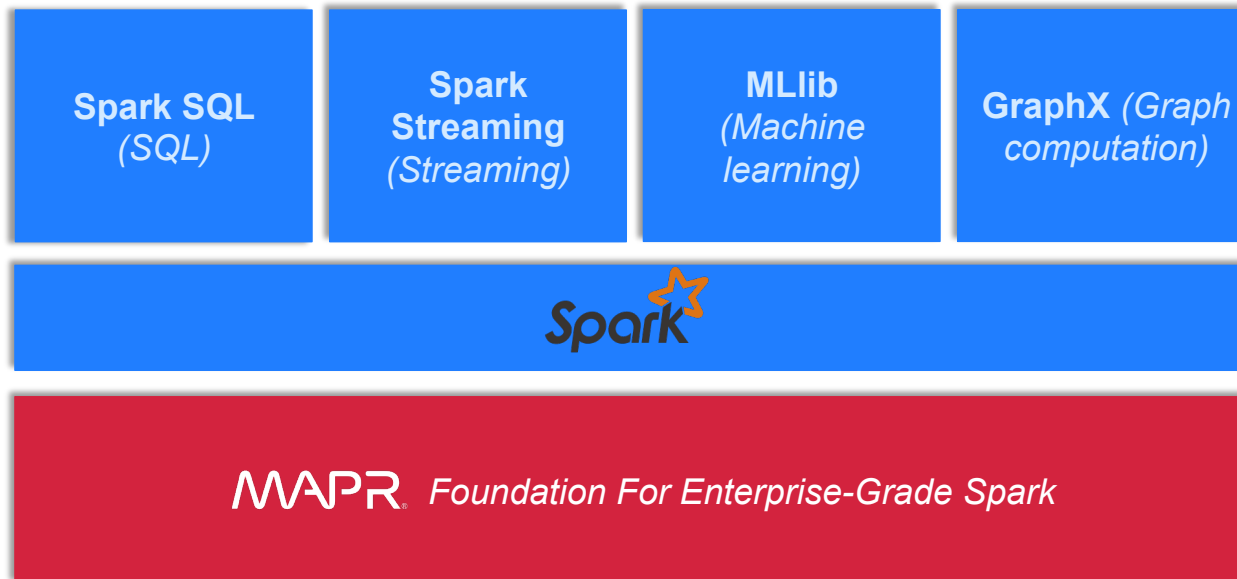
Support for the
complete Spark stack

Engineering & roadmap
collaboration

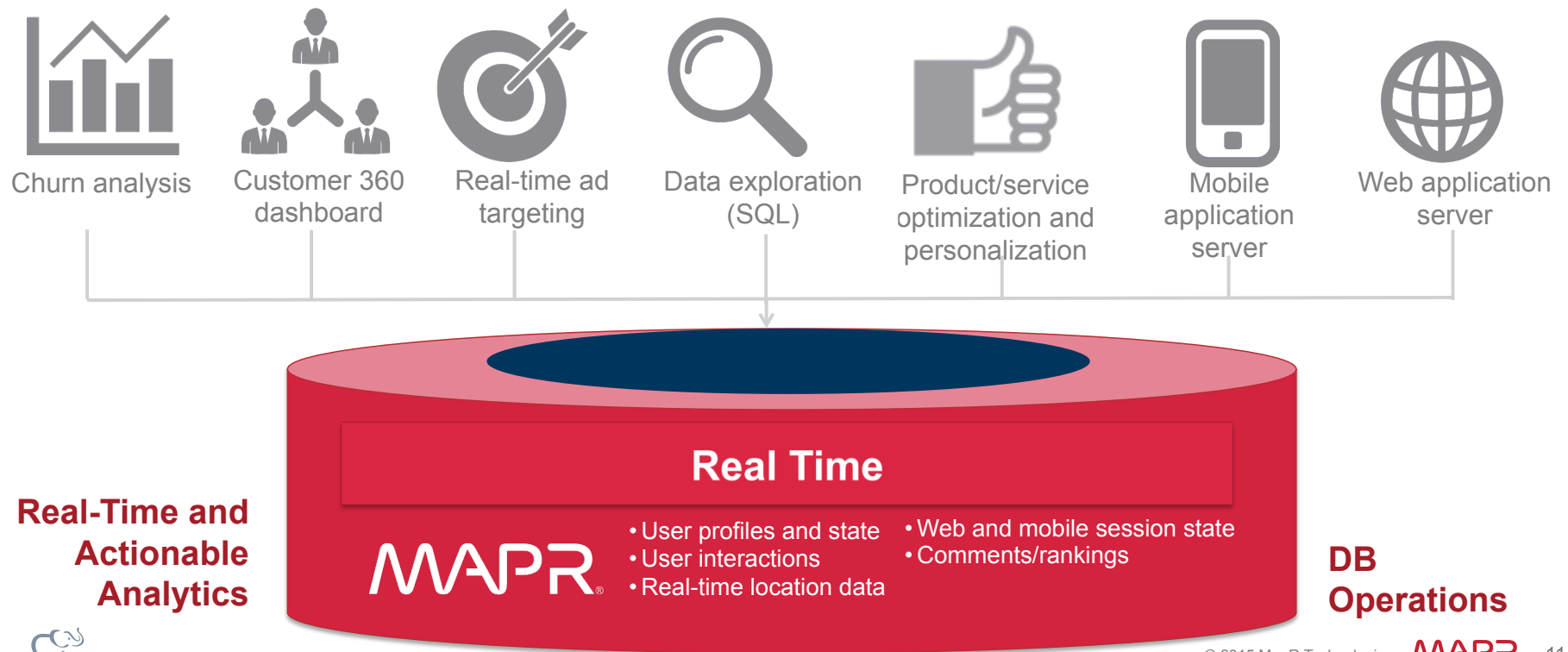
Back-end support



The Most Complete Spark Environment



Operations + Analytics on One Hadoop Platform with SQL Access



Spark + MapR = Ready For Production Success

High Performance

World-record performance on disk

24/7 Best-in-class Global Support

Strategic partnership with Databricks to ensure enterprise support for the entire stack

Operational Data Store

MapR-DB + Spark = real-time analytics

Reliability for Production

SLA-Driven Applications

- High availability
- Data protection
- Disaster recovery



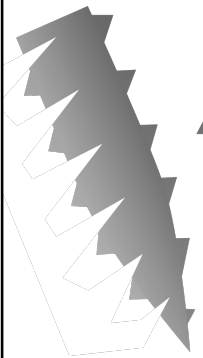
MapR Introduces 3 New Spark-Based Quick Start Solutions

Real-Time Security Log Analytics

Time Series Analytics

Genome Sequencing





APACHE
DRILL

Self-Service Data Exploration

Single SQL Interface for Structured
and Semi-Structured Data

Data Agility with Less IT Required



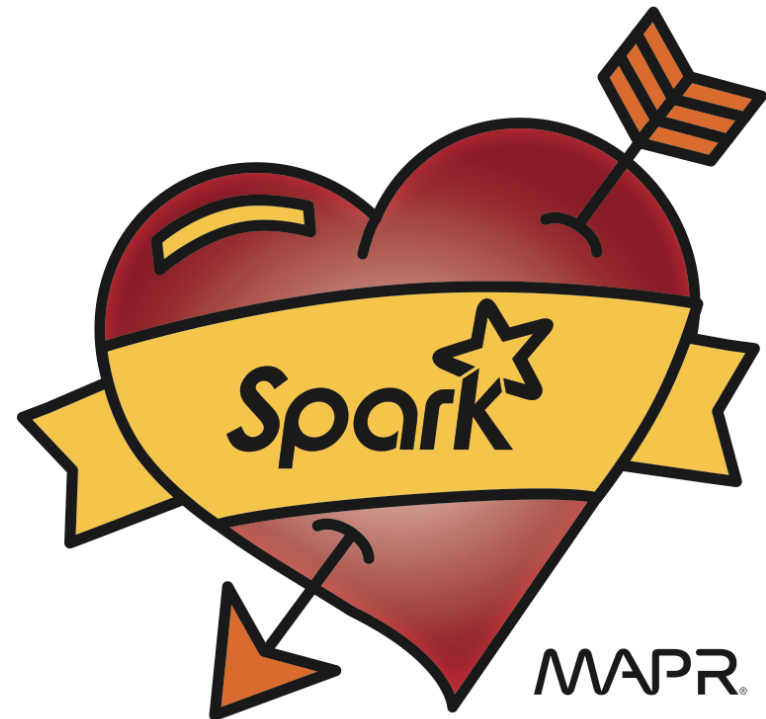


Free
On-Demand Training
www.mapr.com/training



Get Your Tattoo In The MapR Booth!

Show off your
Kickstart My Heart skills
and enter to win
Xbox 360 & Guitar Hero



MAPR[®]

Top-Ranked **Hadoop**
Distribution

Top-Ranked **NoSQL**

Top-Ranked **SQL-on Hadoop**
Solution
