

3.36pt

# Big Data in Streaming

## elaborazioni e visualizzazione in Real Time

Davide Vergari aka OutOfBounds  
&  
Davide Isoardi aka Abacus

Linux Day 2016

22 Ottobre 2016



# Big Data e il V.V.V.

## Cosa sono?

**Volume:** peta e peta di dati distribuiti

**Velocità:** flusso continuo da tanti (migliaia) device -  
**RealTime**

**Varietà:** dati eterogenei non statici nel tempo



Avercelo grosso, il **DB**, non è sinonimo di BigData

Quindi parallelizziamo le operazioni:  
**Ingestion**, **I/O su file system**, **Elaborazione**

# L'ecosistema Hadoop

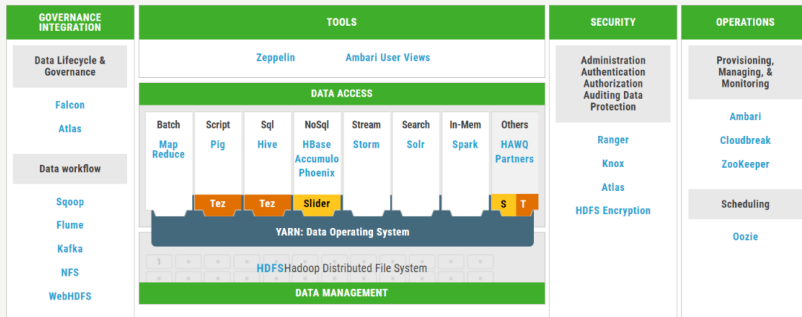
*Non è un tool, ma un insieme di service che completano la soluzione a seconda delle esigenze del business.*



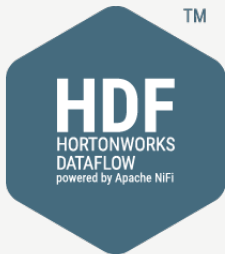
## Hadoop - Da dove nasce



# Hortonworks Data Platform - dove siamo



# Cosa vediamo oggi?



# Riferimenti

- GitHub repo di Davide Isoardi <https://github.com/disoardi>
- GitHub repo di Davide Vergari <https://github.com/dvergari>
- GitHub repo di eCube s.r.l dove ci son le demo complete ed il processors NiFi per la sentiment analysis  
<https://github.com/ecubesrl>

Qui trovate a grandi linee i riferimenti della demo vista nel talk, ovviamente con alcune differenze e, soprattutto, replicabile su una singola macchina. Ad esempio la demo tweetsdemo-service è funzionante anche direttamente su una Sandbox (leggete il README)

## eMail

Pre altre info contattatemi via mail all'indirizzo [abacus@linux.it](mailto:abacus@linux.it)