

Data in IT Management

Serge Mankovski
CA Labs, CA Technologies
May 2012

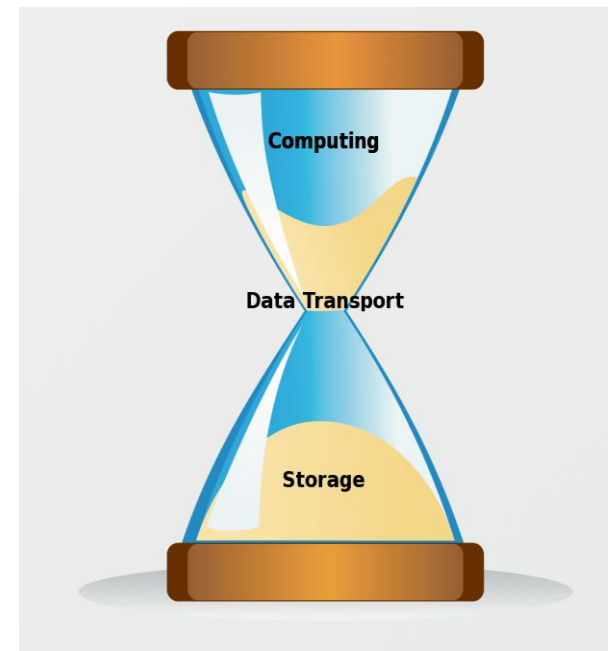
Disclaimer: Does not represent view of CA, just my view as a researcher

Enterprise IT Management (EITM)

- Makes all software and hardware in Data Center operate
 - Compute
 - Store
 - Move

Operational definition of Big Data

If it takes more than one week to make a copy of data set, it is Big Data



Data in EITM

- Front-end
 - SaaS
 - PaaS
 - Cloud
- Back-end
 - Log
 - Event stream
 - Alarm
 - Automation
 - Problem resolution
- Front-end/Back-end correlation
 - Multi-perspective

Big Data once

Big Data twice

Big Data trice

Data Processing Requirement in IT

— Mining

- Trend detection
- KPI mining
- Change Management

— Explorative

- Root Cause/Troubleshooting
- Impact Analysis

— Escalation Management

- Alarming
- Failure Detection

Off-line

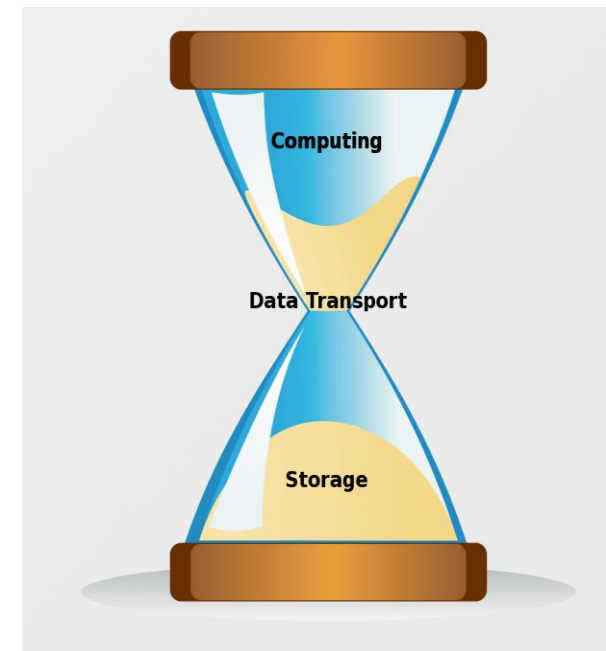
Near-line

Online

360 degrees view of data (think Cloud)

Will be more common as data becomes really large

- Big Data as Data Store
- Big Data as Data Warehouse
- Big Data as Analytic Platform
- Big Data as Event Stream
- Big Data as Not-only-SQL
- Big Data as Non-structured data
- Big Data as Distributed Computation System



Problem of benchmarking Big Data Systems (think Service)



Data Representation, Data Processing, and Measuring Scale are
“mutually dependent variables”

Need finer grain “étalon du mètre” for measuring measuring scale

Place data on a “etalon bench with marks” and let vendors build processing around it

Reality check

- Data physically co-located
logically separated
- Important to know how one
data “view” system compares to
another

Needs to happen

- Discussion about the benchmark
should take into the account “**the
bench**”
- “The bench” should provide a
neutral **measuring service** free
of vendor bias
- “The bench” should provide
elastic facility for “plugging in”
and measuring data processing
“views”

Establish Etalon “Benchmark Data Cloud Service”