Data Value – Some thoughts and Experiences

July, 2016
Barry A Rudolph
My background

- **IBM**
  - Broad background in most divisional functional areas
  - WW responsibility for disk storage system P&L
  - WW responsibility for tape and archive storage system P&L
  - WW responsibility for networking hardware system P&L
  - Retired in Nov 2010

- **Board of director roles**
  - VelociData, Inc – private startup in stream analytics
  - Spectra Logic – Private – deep storage experts

- **CEO - VelociData**
  - Extreme performance - stream data transformation and analytics
  - FinTech, Enterprise ETL, large network systems management
Why should we care

Why does it Matter?

- Data Value may be important to:
  - Make fundamental decisions on capital and expense
    - Value of availability, value of performance, value of data access etc.
  - Define a backup strategy and make capital and period expense decisions
  - Define a Disaster Recovery strategy and make capital and period expense decisions
  - Define an archive strategy and make capital and period expense decisions
  - Define a strategy around future analytic value of current period data
  - Define and implement a risk management strategy – legal discovery, data compromise impact
  - Define and implement an insurance strategy
  - ETC ........
Simple conceptually but hard in reality

So, it sounds easy? Data Value Assessment Options:

- **Define business critical environments/Applications:**
  - Systems or environments that represent the primary assets of the enterprise, or represent significant risk... e.g.
    - Mainframe transaction systems
    - CRM systems
    - Product Development Environments
  - Define the system linkages (what systems feed other systems)
  - Associate the data with the systems – to define value

- **Define value heuristically**
  - Assess usage patterns
  - Assess data flows and linkages
  - e.g. This is done inherently in data storage devices and distributed systems to make decisions on caching and locality

- **These may be quite synergistic**
Approaches in the field

What do companies do:

- Very sophisticated analysis of critical systems and data associations
- Data classification—manual and with existing tools
- Operational decisions aligned to specific environments viewed as critical
- Backup everything
- Use systems with inherent redundancy (HDFS)
- Hope for the best
- Hope no one notices
Revenue and Profit Drivers

- Critical applications that are needed for business continuity – Disaster recovery
- Critical applications that represent the tactical revenue and profit generation
- Critical applications that are aligned with strategic initiatives
- Critical data that may be valuable to non-competitive businesses – define potential buyers

Additional Risk Drivers

- Define the data that if lost or stolen has the largest impact on:
  - Brand equity
  - Direct expense to recover adjacent costs
  - Value to competitors
  - Legal and regulatory impacts