Overview of NIST Big Data Interoperability Framework Volume 8

Dr. Gregor von Laszewski
Assistant Director of Community Grids Lab, Adjunct Associate Professor
Indiana University

NIST Campus
Gaithersburg, Maryland
June 1, 2017
Presentation Overview

- Volume Presentation Outline
- Volume 1, Definitions (Nancy Grady, SAIC)
- Volume 2, BD Taxonomies (Nancy Grady, SAIC)
- Volume 3, Use Cases and General Requirements (Geoffrey Fox, Indiana University)
- Volume 6, Reference Architecture (David Boyd, InCadence Corp.)
- Volume 4, Security and Privacy (Arnab Roy, Fujitsu; Mark Underwood, AVP, Strategic Initiatives, Controls and Countermeasures)
- Volume 8, Reference Architecture Interface (Gregor von Laszewski, Indiana University)
- Reference Architecture Software Implementation Environment and Demonstration (Gregor von Laszewski, Indiana University)
- Volume 7, Standards Roadmap (Russell Reinsch, Center for Government Interoperability)
- Volume 9, Adoption and Modernization (Russell Reinsch, Center for Government Interoperability)
NBDIF Volume Overview

Vol. 1 BD Definitions
Defines common language

Vol. 2 BD Taxonomies
Hierarchy of NBDRA components

Vol. 3 Use Cases & Vol. 5 Arch Survey
Info gathered; requirements extracted

Vol. 4 S&P
Interwoven topics of S&P examined

Vol. 6 NBDRA
Developed NBDRA

Vol. 7 Standards Roadmap
Examine standards wrt NBDRA

Vol. 8 NBDRA Interfaces
Implementation of NBDRA

Vol. 9 Adoption & Modernization
Volume Presentation Outline

- For each volume
  - Scope of the volume
  - Brief recap of version 1
  - Highlights of version 2 accomplishments
  - Summary of version 2 areas needing contributions
  - Topics that could be considered for version 3
Volume 8

Acknowledgement

• NSF for support of Comet (SDSC, Indiana University)
  – Members of the NSF comet project

• NBDRA WG members
  – Tuesday meetings and discussions have considerably contributed to this effort
  – We still need to add your names ... please e-mail me at laszewski@gmail.com ...

• Prior effort conducted as part of cloudmesh at Indiana University
Volume 8 – Opportunities for Contribution

- Volume 8 is in draft stage.
  - We look for additional contributors to the volume 8
  - We want to keep the initial part of the draft as simple as possible and develop contributions based on ASCII text modifications.
  - Contributions can be conducted in multiple ways
    - A) Make modifications to the PDF document with PDF augmentation tools such as Adobe or Skim, send via e-mail.
    - B) Make modifications in ASCII while specifying the section you like to modify (section header, new paragraph, line numbers), send via e-mail.
    - C) Make modifications through git pull requests
Current Working Group Influences

- Working Group Volumes
- Cloudmesh
- Cloudmesh Comet @ SDSC/IU
Volume 6: NIST Big Data Reference Architecture

INFORMATION VALUE CHAIN

Big Data Framework Provider

Processing: Computing and Analytic
- Batch
- Interactive
- Streaming

Platforms: Data Organization and Distribution
- Indexed Storage
- File Systems

Infrastructures: Networking, Computing, Storage
- Virtual Resources
- Physical Resources

Data Provider

System Orchestrator

Big Data Application

Collection

Preparation/Curation

Analytics

Visualization

Access

Data Consumer

IT VALUE CHAIN

Data Preparation/Curation

Collection

Preparation/Curation

Analytics

Visualization

Access

Software Tools and Algorithms Transfer

Service Use

Big Data Information Flow

K E Y :
Volume 8 – Document Scope

• Established operational interfaces
  – for management interactions and dataflow with needed resources between NBDRA components
• Described and defined specific interfaces and the interactions
  – between NBDRA components
• Developed a set of interfaces defined through examples
  – that can be used to create schema based definitions of objects that are manipulated through Big Data design patterns
Volume 8 – Possible Version 3 Topics

• Formalization of the interfaces defined by example
• Validate the interfaces
  – Verifying select use cases
  – Working with domain experts
    • to identify workflow and interactions among the NBDRA components and fabrics
  – Exploring and modeling interactions
    • within a small-scale, manageable, and well-defined confined environment
  – Aggregating the common data workflow and interactions
    • between NBDRA components and fabrics and package them into general interfaces.