**NIST Big Data Public Working Group (NBD-PWD)**

**NBD-PWD-2014/M0306**

**Source: NBD-PWG**

**Status: Draft**

**Title: Ref. Architecture + Tech. Roadmap Meeting Minutes for April 22, 2014**

**Author: NBD-PWG Subgroup Co-Chairs**

**Participants reviewed the AIs from the previous call.**

1. RA Phase I : Orit prepared a condensed version of David’s M0303, included it in the RA document, and emailed it on Apr 19th for the editors’ review.
2. RA Phase II: The three approaches identified at the previous call and listed below have been discussed:
   1. Pattern or “Top Down”: Choose/define a representative generic, but simple use case (a.k.a., a pattern). Map it to the RA blocks with the emphasis on interfaces, i.e. show the API flow (list existing technologies as illustrations) between the RA components in order to implement the pattern. **Bob Marcus** presented his example proposal for the use case of using SQL interfaces for Hadoop DB. Bob also shared a 3rd party paper presenting a broader discussion of this approach and stated that standardizing such a popular approach could be an important candidate for standardization. David B. commented that this particular use case would be potentially limiting for implementing highly scalable BD systems. Orit L. commented that it would be more useful to show the sequence diagrams using the NIST BD RA diagram and the agreed RA terminology as the baseline. Wo commented that specific technologies will not be mentioned in the NIST output documents, but a White Paper showing how a specific Use Case using specific technologies and APIs can be mapped to the RA would be mostly welcome.
   2. Interfaces or “Bottom Up”: **David Boyd** presented his initial text describing the functionality of the interface between the Data Provider and the Application Provider (see M0304). The text was well accepted. Wo noted that the interface needs to capture both scenarios: Data Provider- initiated and Application Provider- initiated. David suggested to provide a number of “sequence diagrams” to illustrate how the already described “types of interactions” are implemented/used for different push and pull scenarios. Orit suggested that the description of this particular interface would be structured in terms of the Big Data Vs’ characteristics, i.e., Volume, Velocity, Variety, etc. Along the same lines, the description of other interfaces can/should be organized according to the taxonomy characteristics relevant to each interface.
   3. Taxonomy also “Bottom Up”: Start with the graphics prepared by Geoffrey Fox and shared with the group in Jan-Mar this year. Organize the technologies and products listed in the graphics in a Taxonomy tree/graph. The resultant taxonomy can be used for validation of the approach (b) for completeness, etc. **Nancy** couldn’t attend the call, so this discussion was postponed to the next meeting.
3. Action items:
   1. Expand the interface (or Bottom Up) to include sequence diagram with focus on Big Data Vs scenarios – David
   2. Explore other “Top Down” examples other than query operation – Bob
   3. Start with the graphics prepared by Geoffrey Fox and shared with the group in Jan-Mar this year. Organize the technologies and products listed in the graphics in a Taxonomy tree/graph. The resultant taxonomy can be used for validation of the approach (b) for completeness, etc.

**Participants (on the Web call) included:**

* Geoffrey Fox
* Orit Levin (Microsoft)
* Vivek Navale
* John Rogers (HP)
* Ovace A. Mamnoon (HP)
* David Boyd (L-3 Data Tactics)
* Mark Underwood (Krypton Bros)
* Ashok Malhotra (Oracle)
* Keith Hare (JCC Consulting, Inc.)
* William Miller
* Pavithra Kenjige
* Wo Chang
* Andrey Shevel
* Alicia Zuniga-alvarado
* Quyen Nguyen
* Dan (NIST)
* Arnab Roy (Fujitsu)
* Bob Marcus
* Ray Wulff
* Brent Comstock (Cox Communications)
* Jon Hudson (Brocade)
* Manoj (CyberIQ)
* Tim Zimmerlin
* Pw Carey (Compliance Partners, LLC).

**Chat log**

(1:01 PM) Ashok Malhotra joined.

(1:03 PM) Geoffrey Fox joined.

(1:03 PM) Tim Zimmerlin joined.

(1:03 PM) Mark Underwood (Krypton) joined.

(1:04 PM) Orit Levin (Microsoft) joined.

(1:05 PM) Manoj (CyberIQ): i'm getting invalid password code for the bridge. Is 272-30-504 correct code?

(1:06 PM) Ashok Malhotra: yes, that's correct

(1:07 PM) Manoj (CyberIQ): @ashok. thanks. finally got in.

(1:07 PM) David Boyd(L-3 Data Tactics) disconnected.

(1:07 PM) David Boyd(L-3 Data Tactics) joined.

(1:08 PM) Orit Levin (Microsoft) disconnected.

(1:09 PM) David Boyd(L-3 Data Tactics) disconnected.

(1:09 PM) David Boyd(L-3 Data Tactics) joined.

(1:09 PM) Orit Levin (Microsoft) joined.

(1:11 PM) Pw Carey, Compliance Partners, LLC joined.

(1:14 PM) Pw Carey, Compliance Partners, LLC: My biased thinking is that every other organization will be more than happy to let this group take the lead.....for various reasons....

(1:15 PM) Geoffrey Fox disconnected.

(1:17 PM) Geoffrey Fox joined.

(1:18 PM) \_JON\_HUDSON\_(\_BROCADE\_) joined.

(1:20 PM) \_JON\_HUDSON\_(\_BROCADE\_) disconnected.

(1:20 PM) Pw Carey, Compliance Partners, LLC: You earlier mentioned Use Cases is this a task area worthy of our efforts.....?

(1:21 PM) Mark Underwood (Krypton): Here is the link to thje Ontology for Big Data meeting / group

(1:21 PM) Mark Underwood (Krypton): <http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2014_Symposium>

(1:22 PM) Geoffrey Fox disconnected.

(1:23 PM) Geoffrey Fox joined.

(1:24 PM) Brent Comstock, Cox Communications joined.

(1:28 PM) PavithraKenjige joined.

(1:29 PM) Pw Carey, Compliance Partners, LLC: After Snowdon, how can we best address standards for; security, privacy, confidentiality....?

(1:30 PM) Pw Carey, Compliance Partners, LLC: Sorry, we meant Snowden...

(1:33 PM) Pw Carey, Compliance Partners, LLC: We'll forward Nancy's minutes to you.....

(1:34 PM) Pw Carey, Compliance Partners, LLC: We raised the issue of aligning both the Big Data Frameworks with the Cloud Eco-System Frameworks.....as we recall...

(1:38 PM) David Boyd(L-3 Data Tactics) disconnected.

(1:38 PM) Pw Carey, Compliance Partners, LLC: We'll be happy to support Orit in these efforts.....

(1:38 PM) David Boyd(L-3 Data Tactics) joined.

(1:38 PM) Pw Carey, Compliance Partners, LLC: Rather, in her efforts....

(1:39 PM) Geoffrey Fox disconnected.

(1:39 PM) Arnab Roy (Fujitsu) joined.

(1:40 PM) Geoffrey Fox joined.

(1:45 PM) Pw Carey, Compliance Partners, LLC: And this means you'll be keeping Orit in the loop via document/status updates regarding Security & Privacy, yes?

(1:47 PM) PavithraKenjige disconnected.

(1:48 PM) PavithraKenjige joined.

(1:48 PM) Mark Underwood (Krypton): Pw - right. I see it working in parallel starting fromt he current Ref Architecture diagram (that's the current spreadsheet we've begun) but sync up later to address S&P.

(1:49 PM) Mark Underwood (Krypton): As long as the walkthroughs include use cases with S&P concerns, i.e., not pure sicence use cases, the interface S&P issues are likely to be surfaced anyway

(1:49 PM) David Boyd(L-3 Data Tactics) disconnected.

(1:49 PM) David Boyd(L-3 Data Tactics) joined.

(1:52 PM) Mark Underwood (Krypton): Agree that both approaches Orit mentions could uncover design patterns

(1:55 PM) Pw Carey, Compliance Partners, LLC: ITU Big Data Standards & EU Big Data Standards are two organizations working within these areas....http://www.itu.int/en/ITU-T/techwatch/Pages/big-data-standards.aspxhttp://ec.europa.eu/digital-agenda/en/news/information-and-networking-days-h2020-work-programme-2014-2015-connecting-europe-facilityhttp://ec.europa.eu/digital-agenda/en/news/information-and-networking-days-h2020-work-programme-2014-2015-connecting-europe-facility

(1:56 PM) Mark Underwood (Krypton): thx, Pw

(1:56 PM) PavithraKenjige disconnected.

(1:58 PM) David Boyd(L-3 Data Tactics) disconnected.

(1:58 PM) David Boyd(L-3 Data Tactics) joined.

(1:59 PM) Pw Carey, Compliance Partners, LLC: What if we all sign a Technological Celibacy Agreement .... ? Respectfully submitted: Pw Carey

(2:00 PM) David Boyd(L-3 Data Tactics) disconnected.

(2:00 PM) David Boyd(L-3 Data Tactics) joined.

(2:02 PM) Geoffrey Fox disconnected.

(2:03 PM) Pw Carey, Compliance Partners, LLC: A White Paper might pass the Non-Endorsement smell test.....

(2:08 PM) Mark Underwood (Krypton): Pw: True, has advantages. A wiki could be more of a "live" document

(2:09 PM) Pw Carey, Compliance Partners, LLC: We think you suggestion is good...the hard part is not adding a subliminal NIST endorsement....hard not to do.....thanks....Pw

(2:09 PM) Pw Carey, Compliance Partners, LLC: Rather, we think your suggestion is good.....

(2:10 PM) Ray Wulff joined.

(2:11 PM) Pw Carey, Compliance Partners, LLC: We can't....

(2:12 PM) Pw Carey, Compliance Partners, LLC: But will it be available over the Internet....?

(2:12 PM) David Boyd(L-3 Data Tactics): This is dave.  I am trying to go but fighting my new corporate structure.

(2:12 PM) Ashok Malhotra: Jim Melton and I will attend from Oracle

(2:12 PM) David Boyd(L-3 Data Tactics): Will know for sure next week.

(2:13 PM) \_JON\_HUDSON\_(\_BROCADE\_) joined.

(2:13 PM) Arnab Roy (Fujitsu): I am planning to attend.

(2:13 PM) Pw Carey, Compliance Partners, LLC: We would think, the more going would be a very good thing for multiple reasons....

(2:14 PM) David Boyd(L-3 Data Tactics): If I can go I hope to present a workshop item on a capability maturity model for big data.

(2:15 PM) Pw Carey, Compliance Partners, LLC: Mr. Boyd, if you need any support in preparation, please let us know and we'll be glad to assist you in your efforts, regardless of the task....

(2:15 PM) \_JON\_HUDSON\_(\_BROCADE\_) disconnected.

(2:16 PM) Pw Carey, Compliance Partners, LLC: How to address NSAs efforts on behalf of privacy....?

(2:18 PM) Geoffrey Fox joined.

(2:19 PM) Orit Levin (Microsoft): My participation in Amsterdam and China is under internal approval process. Just FYI for now.

(2:19 PM) \_JON\_HUDSON\_(\_BROCADE\_) joined.

(2:19 PM) Mark Underwood (Krypton): I have Jon, Arnab as probable for Amsterdam, Orit and David requesting $ :)

(2:24 PM) PavithraKenjige joined.

(2:24 PM) Pw Carey, Compliance Partners, LLC: You may want to concentrate via Mobile Technologies & Nono-Technologies....just a thought....or two...

(2:26 PM) Pw Carey, Compliance Partners, LLC: Oops....rather Nano-Technologies....

(2:26 PM) Pw Carey, Compliance Partners, LLC: Nono-Technology hasn't been invented yet, perhaps next week....

(2:29 PM) David Boyd(L-3 Data Tactics) disconnected.

(2:29 PM) David Boyd(L-3 Data Tactics) joined.

(2:30 PM) David Boyd(L-3 Data Tactics): Orit sorry I go bumped off. I could not hear all you said.

(2:30 PM) Tim Zimmerlin disconnected.

(2:30 PM) Brent Comstock, Cox Communications disconnected.

(2:30 PM) Mark Underwood (Krypton) disconnected.

(2:30 PM) David Boyd(L-3 Data Tactics): Send me an email asking your question again.

(2:30 PM) Geoffrey Fox disconnected.

(2:30 PM) Orit Levin (Microsoft): ok!

(2:30 PM) David Boyd(L-3 Data Tactics) disconnected.

(2:30 PM) Orit Levin (Microsoft) disconnected.

(2:30 PM) Manoj (CyberIQ) disconnected.

(2:31 PM) Pw Carey, Compliance Partners, LLC: Dear Sir: Our best wishes...