RefArch 1 Aug 2013 1100-1300

Meeting Notes

Orit Levin began the call at 1100 and noted that this group and Security/Privacy should be working together, recommending the schedule of a joint meeting.

Wo Chang stated that NIST values volume of quality feedback/inputs. NBD WG and Subgroups Co-Chairs’ technical inputs/submissions are welcome as long as stated clearly as personal opinion. Reminder of goal: vendor-neutral environ that can be used by standards organizations as input. He reinforced the observation that there were four areas (Data Source, Transformation, Infrastructure, and Data Usage) in common across all submitted RAs. Wo made a last call for any additional changes and was approved due to no additional inputs.

Orit Levin noted that document M0059 page 9 shows the common 4 areas seen across all RA submissions. This document is a summary of rough agreement to date and not representative of a final RA. This subgroup’s first round goal was to gain a rough outline of terms for cross-pollination to definitions/taxonomy group. Document 59 is the present state of understanding toward this goal.

Gary Mazzaferro asked if we should discuss the RA communication plan in this forum noting that communication outputs maybe not clear to other working group consumers. Wo Chang agreed communication outside the NBD is very important and noted that we also require deeper understanding of the roles in our RA first.

Gary Mazzaferro added that roles, structure, and framework messages to consumers at present is confusing. In the final product we should consider multiple RA diagrams/viewpoints. Gary noted that is was an issue hit in the Cloud working group and that Big Data adds more complexity to the RA than cloud. ***This facet was parked for line item discussion: number of diagrams and structure of the final Reference Architecture.*** Orit Levin agreed that the inputs and perspectives vary but commonalities exist. An all-view will be built but more detailed viewpoints should be built as well.

Gary Mazzaferro observed that the RA submissions should be viewed as different perspectives of a common architecture rather than multiple disconnected architectures. He noted that further reinforcing the benefits of this strategy we do not have to select elements from each but can coalesce them on commonalities. Gary provided a DODAF link at <http://dodcio.defense.gov/dodaf20/dodaf20_viewpoints.aspx>

Gary indicated he believes this is in line with Orit Levin’s strategy to come up with common lexicon

Bob Marcus noted our 14 August RA has to give list of planned deliverables and that we need to get this ASAP and set writing assignments. He believes we are at risk of missing that date because we don’t have any depth to our output and that we must drill down now.

Gary Mazzaferro noted again that there exists no cross-working group communication plan and that we need a top-level communication plan.

Orit Levin pointed out that with correction we can use document 59 as a seed for the draft RA.

Wo Chang indicated that the Networking and Information Technology Research and Development (NITRD, worked closely with the OSTP on Big Data Initiative) Big Data steering group would be interested in our effort. We need to drill down our RA in order to make progress.

Without detail we have no message to give. Gary Mazzaferro agreed but expressed concern that the Sept 27 deadline for even high-level RA is unrealistic noting the multi-year timelines for the Cloud working groups. He further requested advice on outputs and execution strategy.

Bob Marcus offered to write up a description of the RA he submitted and suggested Gray Mazzaferro write up the messaging aspect. Orit Levin asked that everyone with strong opinions should write their recommendations up for structured dialog and analysis.

Gary Mazzaferro asked for clarification of the outputs stating he understood we were to have an approach identified by 27 Sept and not the deliverables themselves. He recommended we build a plan, and communicate the plan to industry for comment and expressed concern with the requirement for deliverables without a final plan and strategy to build them. Wo Chang reinforced that the 27 Sept deliverable is the Big Data roadmap—not plan. On 14 Aug we meet with all of the co-chairs to establish common goals and start writing. The 27 Sept Roadmap should contain a survey of RAs. Wo asserted that there is no intent in the charter to seek community feedback before drafting the Roadmap.

Gary Mazzaferro asserted execution can proceed by one of two approaches—scientific method or marketing methods. He explained that the scientific method takes longer but is more agnostic. The marketing method is less agnostic but fast. He noted that we should look for a balance and asked how we will be tollgating after the 27 Sept initial deadline. He also indicated that he does not perceive that the scientific method approach is possible on this schedule. He explained that we don’t know the White House expectations and risks explicitly and asked for a list of expectations to guide initiative goals.

Wo Chang asserted that there is no mandate from the White House and the NBD is NIST’s initiative to explore Big Data from the standard and interoperability perspective and we should focus on what we set in the charter to build a good Roadmap. He requested that any ideas for messaging and communication to the industry be written up and submitted. In the meantime we must proceed.

Wo Chang addressed concerns over the prioritization and order of execution of the project restating that the external stakeholders want to see BD roadmap. We may not have anything out by 27 sept but at minimum we hope to generate consensus on BD language. Wo Chang stated that the charter guides the outputs and we need to drill into the RA content. If we miss the 27 Sept deadline we will continue to push forward and report any significant progress. The discussion of order of execution and strategy was tabled.

Wo noted that we need to drill immediately into actors, roles, and responsibilities.

Gary Mazzaferro referring to the document he will deliver said he will map stakeholder positions to 4 block common components as identified in 59.

Orit Levin noted that we also need to dive into each block and that the Transform block will likely be the most challenging. The Infrastructure block will require a significantly deeper dive as well. Orit asked for ideas to seed deep dive. Orit’s idea is that we could proceed in parallel with the other subgroups using this diagram (slide 9 of document 59) as a communication medium with the other groups and asked if this is a good idea. Pratik Thakkar voted for parallel execution as suggested.

Wo Chang noted that the Requirements subgroup has published five Use Cases and asked how we intend to map these to the RA. One use case is the NASA using Terabyte removable disks of data that they send to the lab. Wo asked how Big Data can help in that Use Case and noted that our RA needs to accommodate legacy architecture. It was asked that since Use Cases can be very different and is it our responsibility to tailor the abstract RA to Use Cases? Wo responded that that may be the case and asked in response how we can help customer to get to a concrete architecture with our abstraction. Orit Levin: Added a question: how will we show mapping generic RA to specialized use case?

Pratik submitted that since there are other subgroup dependencies on our output we should ask them that they need and ensure we answer those requests in our deep dive.

Wo Chang stated that all subgroups should study the submitted Use Cases and added that a web search Use Case is in need of analysis. ***Orit Levin accepted the task to start joint meetings with all subgroups. (OL).***

Bob Marcus identified that all groups’ activities are still a work in progress and that this group needs to get into the details of our deliverables ahead of the join meetings. Gary Mazzaferro countered that the RA is a response to other sub group work and that we may be recreating the work of the other groups. Since we are not focusing on RA but on Use Cases and models we may need to reconsider the 27 Sept commitments.

Orit Levin stated the goals for this next week: gather write-ups from contributors with a strong opinion on the outputs and to gather from the Security Subgroup a list of what they need from us.

Wo Chang reviewed the deliverables due this week: ***All submitters write up their RA proposals for next meeting*** ***(proposers)***. Pritik noted that we will compare and coalesce in the 9 Sept meeting.

Gary Mazzaferro accepted the task ***write outline for white paper*** ***(GM)***. Has submitted one to the Roadmap Subgroup and they are proceeding with it.

Orit Levin asked that he includes details of his RA blocks in his outline.

Pratik identified a task to write down agreements on what categories we will use in the whitepaper.

William Miller noted that members of the Use Case Subgroup will submit more Use Cases next week. One will be based on cargo shipping, which represents many areas of logistics. ***William will submit inputs for reference.***

Orit Levin’s concluding comment:. We will get the summary, action items, and minutes out in next 1-2 days due to travel and asked that we advise if anything missing before next week.

Orit Levin concluded the meeting at 13:06.

Notes submitted by Don Krapohl, 6 Aug 2013.

Chat Transcript:

(11:28 AM) Orit Levin (Microsoft): yes. sure people can go and fetch and look at the document(s) independently from the web sharing.

(11:28 AM) PavithraKenjige26 joined.

(11:28 AM) Orit Levin (Microsoft): http://bigdatawg.nist.gov/\_uploadfiles/M0059\_v1\_4187065247.pptx

(11:28 AM) PavithraKenjige disconnected.

(11:28 AM) Wo Chang: Chang: I can see it now.

(11:29 AM) \_Cherry Tom\_(\_IEEE-SA\_): slide 9

(11:31 AM) PavithraKenjige26 disconnected.

(11:31 AM) PavithraKenjige joined.

(11:31 AM) Wo Chang: Chang: fully agreed with Bob: the slide 9 is a top level and we need to drill down more detail

(11:37 AM) Wo Chang: Chang: slide 9 may not be the final diagram but at least contain the four basic key components

(11:37 AM) Wo Chang: Chang: Once we can drill down more, the picture may look differently but the key components are still there

(11:38 AM) Wo Chang: Chang: I think once start wo Chang:rking on the actors, roles, and responsibilites, then we may have a better picture

(11:38 AM) Tom Wo Chang:teki (Acentia): I think the diagram is being proposed as to what is in common among various indvidual submissions

(11:38 AM) Wo Chang: Chang: this is correct Tom.

(11:39 AM) tbeyene disconnected.

(11:39 AM) Tom Wo Chang:teki (Acentia): if we can agree on the common framewo Chang:rk we can then expand each of the high level blocks in a top down manner

(11:39 AM) Wo Chang: Chang: agree again Tom.

(11:41 AM) tbeyene joined.

(11:42 AM) Bob Marcus (ET-Strategies): On August 14, you must present a "description of planned deliverables". I think that you must accelerate your discussion to achieve this goal.

(11:43 AM) Tom Wo Chang:teki (Acentia): the proposal is not limited to a single diagram

(11:43 AM) Wo Chang: Chang: Gary, can you clarify which three diagrams? I think this is helpful if you have them in mind.

(11:43 AM) Pratik Thakkar (Philips) joined.

(11:43 AM) William Miller (MaCT USA): i propose having a structed, unstructed, and hybrid diagrams

(11:44 AM) Bob Marcus (ET-Strategies): The diagram with 4 block is an openinh high-level slide for a deliverable. It has no "Big Data" content and needs to be followed by much more information.

(11:44 AM) Pw Carey (Compliance Partners, LLC): So....what are the Planned Deliverables...and their drop-dead delivery date...and who will be on the Teams helping to create them.....Yes?

(11:44 AM) Wo Chang: Chang: We definitely need multiple diagams to convey different concepts.

(11:45 AM) Pw Carey (Compliance Partners, LLC): We volunteer Pw to help out in this effort....

(11:45 AM) Wo Chang: Chang: PW: the planned deliverables should be a consensus based of collective ideas and outline

(11:45 AM) Pw Carey (Compliance Partners, LLC): The effort being the Final Deliverables....effort...

(11:46 AM) Tom Wo Chang:teki (Acentia): what if we were to take the several digrams that have been submitted (and possibly others) and consolidate from them the areas identified in this diagram, for example come up with an expansion of Transformation based on what has been submitted?

(11:46 AM) Wo Chang: Chang: The outline should be consistent with other subgroups; that's why the joint meetings are important

(11:46 AM) Pw Carey (Compliance Partners, LLC): As a bonifide team player we agree with whatever Team RA commits to....

(11:47 AM) Bob Marcus (ET-Strategies): To PW Carey: Need to decide on planned deliverables and begin giving wtite assignments. The time for long teleconference discussions at high level should end by August 14

(11:47 AM) Pw Carey (Compliance Partners, LLC): Sounds good....Aug. 14th....

(11:47 AM) Brian Beecher, Joules Technology Inc. : Excellent!

(11:48 AM) tbeyene disconnected.

(11:49 AM) Bob Marcus (ET-Strategies) disconnected.

(11:49 AM) Bob Marcus joined.

(11:50 AM) Tom Wo Chang:teki (Acentia): the point gary makes about persepctives is interesting - if we thake that tack we need to agree on what the "perspectves" should be and what information they convey

(11:50 AM) Wo Chang: Chang: Gary: can you upload the referred document so we can review and discuss? thanks!

(11:51 AM) tbeyene joined.

(11:51 AM) gary mazzaferro: http://dodcio.defense.gov/dodaf20/dodaf20\_viewpoints.aspx

(11:53 AM) Pw Carey (Compliance Partners, LLC): We agree with Bob's position....

(11:53 AM) Tom Wo Chang:teki (Acentia): +1

(11:54 AM) Brian Beecher, Joules Technology Inc. : Let get on track

(11:55 AM) Tom Wo Chang:teki (Acentia): I have to go. I wo Chang:uld be willing to take a writing assignment

(11:55 AM) Tom Wo Chang:teki (Acentia) disconnected.

(11:58 AM) Pw Carey (Compliance Partners, LLC): Pw will volunteer Pw to take a writing assignment, too....

(12:00 PM) Pw Carey (Compliance Partners, LLC): Please look at the DoDAF Arch Reference the link was posted within the Public Chat....quite a good start, actually....

(12:00 PM) tbeyene disconnected.

(12:02 PM) Bob Marcus: I volunteer to write up a description of the contents of my Reference Architecture diagram.

(12:02 PM) tbeyene joined.

(12:04 PM) Pw Carey (Compliance Partners, LLC): Bob...we'll be happy to review your efforts.....Respectfully yours, Pw

(12:05 PM) William Miller (MaCT USA): evalute structure vs unstructed vs hycrid apporach and evaluate the gaps

(12:05 PM) Pw Carey (Compliance Partners, LLC): pwc.pwcarey@gmail.com is a good way to reach us....

(12:05 PM) Pw Carey (Compliance Partners, LLC): What's a 'HYCRID APPORACH'....?

(12:11 PM) Bob Marcus: We must have initial draft deliverables by September 27 in order to be taken seriously

(12:13 PM) Pw Carey (Compliance Partners, LLC): We believe they don't know....they're relying upon us to take that responsibility....just our opinion...backed up by millions of years of experience.....Respectfully yours, Pw

(12:18 PM) PavithraKenjige disconnected.

(12:19 PM) PavithraKenjige joined.

(12:20 PM) Bob Marcus: What are the new roles needed for Big Data?

(12:22 PM) gary mazzaferro: I started a drill down document

(12:22 PM) tbeyene disconnected.

(12:22 PM) tbeyene joined.

(12:22 PM) gary mazzaferro: It was posted to the file share

(12:23 PM) gary mazzaferro: http://bigdatawg.nist.gov/\_uploadfiles/M0071\_v1\_3322037625.pptx

(12:24 PM) Pw Carey (Compliance Partners, LLC): Thanks.....

(12:24 PM) Bob Marcus: I think the key new Role at this time is someone who knows how to input and output data from Big Data data stores. This is new because traditional data people have never used tools like Map-Reduce.

(12:27 PM) Brian Beecher, Joules Technology Inc. : This is good bigdata should be flexable to handle large disk clusters and cloud

(12:27 PM) Brian Beecher, Joules Technology Inc. : data disk clusters and cloud data

(12:28 PM) Brian Beecher, Joules Technology Inc. : That is good we need to get an idea of what those requirements

(12:29 PM) Pw Carey (Compliance Partners, LLC): We're currently researching MapReduce....

(12:30 PM) Pratik Thakkar (Philips): @Notes/Scribe: My name is Pratik Thakkar not Prakit :)

(12:30 PM) Brian Beecher, Joules Technology Inc. disconnected.

(12:30 PM) Brian Beecher, Joules Technology Inc. joined.

(12:31 PM) Wo Chang: Chang: combined use cases: http://bigdatawg.nist.gov/\_uploadfiles/M0050\_v3\_9138947655.docx

(12:31 PM) Pw Carey (Compliance Partners, LLC): Introduction to Hadoop & MapReduce...Hadoop MapReduce Fundamentals 1 of 5 via YouTube....

(12:31 PM) Wo Chang: Chang: Within the document, there are also mentions what big data challenges to them

(12:32 PM) Wo Chang: Chang: The main point of use cases is to help us draw the big data requirements

(12:33 PM) \_Cherry Tom\_(\_IEEE-SA\_): wo Chang:uld it help to have written list of questions exchanged between subgroups? like liaison statements?

(12:34 PM) Bob Marcus: To Tom" A written exchange will be useful.

(12:35 PM) tbeyene disconnected.

(12:38 PM) gary mazzaferro: I wo Chang:uld suggest creating an outline for the RA white papaer

(12:38 PM) William Miller (MaCT USA): can i show u a use case that i will be submitting to the use case wo Chang:rking group

(12:39 PM) Pw Carey (Compliance Partners, LLC): Please....

(12:40 PM) William Miller (MaCT USA): hard to get in a wo Chang:rd

(12:41 PM) Brian Beecher, Joules Technology Inc. disconnected.

(12:42 PM) Brian Beecher, Joules Technology Inc. joined.

(12:43 PM) Pw Carey (Compliance Partners, LLC): In the upper right-hand corner there is a wheel icon which you can use to reduce the 'echo' effect.....good luck...

(12:44 PM) Pw Carey (Compliance Partners, LLC): More like a 'sprocket'....

(12:46 PM) Pratik Thakkar (Philips): @Gentleman/Lady: Taking notes - My name is "Pratik" :)

(12:47 PM) Pw Carey (Compliance Partners, LLC): Pronounced...."Prah....Teak"....no?

(12:47 PM) gary mazzaferro: Thank you and applogies for mispronouncing your name

(12:49 PM) Don Krapohl: don@donkrapohl.com

(12:49 PM) PavithraKenjige disconnected.

(12:50 PM) PavithraKenjige joined.

(12:55 PM) Pw Carey (Compliance Partners, LLC): We don't see ISO/IEC 27001 Re: Security, et cetera.....do we have rotten eyes....?

(12:55 PM) Bob Marcus: What Big Data technologies are used or needed for the Logistics Use Case?

(12:56 PM) Orit Levin (Microsoft): We (the co-chairs) will summarize the AIs for the next call and send them to the mailing list.

(12:57 PM) Pratik Thakkar (Philips): Sorry: Is AI = Agenda Items?

(12:58 PM) Pw Carey (Compliance Partners, LLC): Sounds like it....

(12:58 PM) Pw Carey (Compliance Partners, LLC): If not....it should be....

(12:59 PM) Wo Chang: Chang: I think AI=Action Items

(1:00 PM) Orit Levin (Microsoft): Action Items, it is ;-)

(1:01 PM) Pw Carey (Compliance Partners, LLC): Well...how about that....AI translates to ACTION ITEMS....I...like it...

(1:02 PM) Pw Carey (Compliance Partners, LLC): Thank you for the presentation, too...

(1:02 PM) Wo Chang: Chang: It will be helpful to include what are the big data requirements. Very good use case.

(1:02 PM) Bob Marcus: Need to drill down from Use Cases to Big Data requiremensts. There probably is not enough time to do that to use in RA deliverables

(1:03 PM) Pratik Thakkar (Philips): Good presentation!

(1:03 PM) Pw Carey (Compliance Partners, LLC): Bob....please send me any items you want us to wo Chang:rk on.....Respectfully yours, Pw

(1:05 PM) Bob Marcus: To PW: Any written conent on specific Reference Architectues will be valuable,

(1:06 PM) Pw Carey (Compliance Partners, LLC): Ok....we'll look into specifics regarding RA (Reference Architectures)....