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

**NBD-PWD-2016/M0565**

**Source: NBD-PWG**

**Status: Draft**

**Title: Web chat from Meeting of September 27, 2016**

**Chat Log D:\\_wo\1DMG\2015\\_BigDataWG\Docs\ChatLog 2016\_09\_27 14\_56.rtf**

**Geoffrey Fox (to Everyone)**: 1:02 PM: better version

**Geoffrey Fox (to Everyone)**: 1:02 PM: better version

**Geoffrey Fox (to Everyone)**: 1:02 PM: https://docs.google.com/presentation/d/15P4idfiujyKB85emaN3-1--YmlVIsbViyWXzRllLTcs

**Wo Chang (to Everyone)**: 1:03 PM: Geoffrey: do you want me to pass the presenter to you?

**Geoffrey Fox (to Everyone)**: 1:04 PM: yes but mike is not working

**Mark Underwood (to Everyone)**: 1:11 PM: Yes, role proliferation is a key observation

**Mark Underwood (to Everyone)**: 1:17 PM: We''re trying to worry about some of that (SDN sec) in IEEE P1915.1

**Tim Zimmerlin (to Everyone)**: 1:17 PM: Ansible ecosystem supports extensible "design patterns" and "instantiation templates". I.e., approximately called Playbooks & Roles

**Mark Underwood (to Everyone)**: 1:22 PM: Tim, ack'd

**Cavan Capps (to Everyone)**: 1:35 PM: Geoffrey, I have a question

**Mark Underwood (to Everyone)**: 1:45 PM: Can't talk now, but I think we will have to help readers of anything we put into v2 standards crosswalk from this notion of "roles" to the narrower (or broader!) InfoSec meaning

**Mark Underwood (to Everyone)**: 1:47 PM: "Playbook" plays well [sic] for security scenarios

**Tim Zimmerlin (to Everyone)**: 1:48 PM: Mark, if you assign user permissions at the granularity of "playbooks" and "roles", then what?

**Tim Zimmerlin (to Everyone)**: 1:49 PM: ...could UNIX file or directory permissions work?

**Mark Underwood (to Everyone)**: 1:51 PM: Solutions are all over the map

**Tim Zimmerlin (to Everyone)**: 1:51 PM: AFAIK, Cloudmesh is a very effective (clever) software shim API that hides the custom cloud APIs from Ansible and every other open source solution.

**Mark Underwood (to Everyone)**: 1:52 PM: Geoffrey, I'd like to see a few sentences on this ambiguity about orchestration. Might help Laurie as well as my drafts

**Tim Zimmerlin (to Everyone)**: 1:53 PM: ...this allows playbooks to run by interfacing directly to Cloudmesh everywhere.

**Russell C Reinsch (to Organizer(s) Only)**: 1:53 PM: separate standard for orch and deployment. very similar. vpel is running systems, workflow standard. language for system devops is tosca. specifies deployment systems.

**Tim Zimmerlin (to Everyone)**: 2:03 PM: Live at https://github.com/cloudmesh/portal

**Tim Zimmerlin (to Everyone)**: 2:08 PM: Bookmarkable link: http://cloudmesh.github.io/client/

**Tim Zimmerlin (to Everyone)**: 2:10 PM: Geoffrey, sounds to me like you have a lot of interest...time for another MOOC.

**Tim Zimmerlin (to Everyone)**: 2:11 PM: Your work grounds much of the NBD-PWG volumes.

**Mark Underwood (to Everyone)**: 2:12 PM: nice work, and thanks

**Mark Underwood (to Everyone)**: 2:12 PM: RA doesn't have a clear orchestration message

**Ann Racuya-Robbins (to Everyone)**: 2:15 PM: Very helpful. Thank you!

**Mark Underwood (to Everyone)**: 2:16 PM: Laurie- I'm taking a stab at this in the V2 Vol 4 draft - domain-specific roles loosely connected to Ansible or Kupernetes roles

**Frank Farance (to Everyone)**: 2:18 PM: frank@farance.com

**Tim Zimmerlin (to Everyone)**: 2:23 PM: Gregor, sincere thanks for ALL your hard work...looking forward to many others using Cloudmesh, etc.

**Tim Zimmerlin (to Everyone)**: 2:24 PM: Yep! Big data is a team sport when real world datasets are involved.

**Gregor (to Everyone)**: 2:29 PM: Thank you very much, we will work hard to make improvements.

**David W. Boyd (to Everyone)**: 2:31 PM: Have to sign off now. Good discussion today.

**Mark Underwood (to Everyone)**: 2:32 PM: Gregor - What would it take to clone-and-create a sample dataset that incorporates a simplistic PII-like scenario - e.g., we add demographics to the fingerprint set and protect it using in-app as well as SDN features - high level discussion here: http://kbros.co/2czm5zK

**Mark Underwood (to Everyone)**: 2:36 PM: clone and create a synthetic PII set

**Tim Zimmerlin (to Everyone)**: 2:36 PM: Mark, you are recommending that provenance metadata be added (using pseudo data values). Where is such metadata? sql? RDF? XML docs?

**Tim Zimmerlin (to Everyone)**: 2:37 PM: ...are URL links used? URIs?

**William Miller (to Everyone)**: 2:38 PM: Wo it is now possible to combine the fingerprint image with metadata information.

**Mark Underwood (to Everyone)**: 2:39 PM: http://searchsdn.techtarget.com/tip/Software-defined-networking-security-enables-granular-policy-control

**Tim Zimmerlin (to Everyone)**: 2:39 PM: SDN is almost layer 2 today. SDN can block access to subnets...for metadata...to only senior LEOs.

**Tim Zimmerlin (to Everyone)**: 2:40 PM: Mark, is falling back into the traditional 3 tier server architecture: client, Web frontend, and databases.

**Tim Zimmerlin (to Everyone)**: 2:41 PM: Cloudmesh might be able to construct 3 independent virtual clusters...then we could overlay them.

**Gregor (to Everyone)**: 2:42 PM: yes tim that is what one of our future usecases actually is

**Gregor (to Everyone)**: 2:43 PM: the underlaying IaaS woudl follow a policy vetting.

**Gregor (to Everyone)**: 2:43 PM: e.g. the data with security restrictions is tahn executed on the secure IaaS

**Mark Underwood (to Everyone)**: 2:43 PM: I think we're onto something interesting here, but I'm not sure what the critical resources are to nurse the baby

**Tim Zimmerlin (to Everyone)**: 2:45 PM: First, build and vet the script API. Second, build one or more REST UIs.

**Robert Reyling (to Everyone)**: 2:46 PM: How many of Geoffrey's students were/are involved with the Cloudmesh exercise(s)?

**Tim Zimmerlin (to Everyone)**: 2:46 PM: ...experts need scripts...to run and share and integrate with apps.

**William Miller (to Everyone)**: 2:48 PM: It should be understood that SDN and SD-WAN are differnt and how they would be used with Big Data systems.

**Tim Zimmerlin (to Everyone)**: 2:49 PM: Geoffrey, Lancaster and Cambridge are starting a new big data analytics statistics API.

**Tim Zimmerlin (to Everyone)**: 2:51 PM: ...MIT Institute for Data, Systems, and Society is making big data (statistics) the common domain for all 5 MIT schools (opened July 2015).

**Tim Zimmerlin (to Everyone)**: 2:53 PM: ...sell them on Cloudmesh...

**Gregor (to Everyone)**: 2:55 PM: coudl you send me the slides

**Gregor (to Everyone)**: 2:56 PM: laszewski@gmail.com