File 20110803.1227 (CDT): Notes from UCDMO conference this morning:

More notes from UCDMO technical sessions:

———————————-

0800 Jeff Jonas, IBM: on 'The Information Sharing Paradox'

He developed the Non-Obvious Relationship Awareness (NORA) software used by Las Vegas gambling establishments to spot collusion between customers and dealers. It started finding insider threats; this prompted funding for the company by In-Q-Tel, CIA's venture capital arm, and was subsequently acquired by IBM.

The concept of 'Enterprise amnesia': for example, Mr Jonas showed a surveillance video of a Las Vegas dealer colluding with a customer who brought his own cold deck to the table. The dealer's telephone number in the employee payroll database was the same as the contact number in the customer's frequent gambler club database entry. The casino already has all that information, yet never made the connection. What causes this, according to Mr Jonas, is the fact that the employee database was never indexed on telephone numbers. If it's not indexed, you'll never find it. Comparison with library card catalogue systems runs throughout his talk. Library card catalogue systems are very non-brittle—they are flexible enough to accommodate any new kind of information acquisition that comes along. Just author/title/subject/call-number indexes, but comprehensively done. Every single card in the catalogue is catalogued all four ways, and none are not included or partly done. That makes information findable no matter what changes later.

'Algorithms at dead end' is the blog post written by Mr Jonas that introduced the concept of 'you can't squeeze data out of a pixel'.

Jigsaw Puzzle metaphor: given a pile of pieces, you don't know how many puzzles are represented. You don't have the box with a picture on it. You don't know how many pieces are missing, or duplicates, or fabricated lies. All you can do is look at one puzzle piece at a time and try to make sense of it in context of everything else you know.

He ran the experiment. Bought a bunch of puzzles, threw away some of the pieces, grabbed a handful from one puzzle and mixed it with just a few from another puzzle, added an overwhelming percentage from an unknown number of other puzzles, included some duplicate pieces on purpose, and included a few pieces that were deliberate lies. Then he gave that bag of pieces to groups of people and timed their interactions, looking for such things as:

1. How long did it take them to find the first connection (two pieces that fit together)?
2. How long until they discovered the first duplicate piece?
3. How long did it take them to deduce that there was more than one puzzle picture here?
4. How long did it take them to identify what the biggest picture probably was?
5. How long to identify that there were pieces that didn't fit anywhere?

Each person was allowed to look at only a single piece per round. They all could see the entire table of pieces already examined. They were to make the best assessment possible for each new piece in the shortest time possible. Periodically, the game suspended for what he called 'data mining rounds' where no new pieces arrived, but everyone just considered what was already on the table.

Results: the first false positive (a piece the players thought fit, but in fact belonged to a different puzzle) went unnoticed. After only 1.3 percent of the pieces had been examined, the players succeeded in guessing the location that the picture represented. After 4 percent of all the pieces had been examined, they found the first connection (two pieces that fit together).

Graphing the time taken by players to examine each puzzle piece, the experimenter found that computing effort declined as the observation space increased; the trend shows that the average computing throughput (disregarding a few outliers) gets faster as more data begins to correlate.

Another discovery that came out of the puzzle experiments—and from application of the idea in software through the NORA project—was that noisy data were faster, overall, to process and find connections in. 'Bad data becomes your friend.' Conclusion: don't clean your data too much. Those variant spellings of names, transposed digits in addresses and phone numbers make it more likely that you will find connections, not less. This was surprising but easily explained: because it is likely there will be multiple pieces of data associated with any particular person in a very large database, those misspellings and transpositions fuzz the data comparison function slightly, enough to spot connections just a little too far apart to be noticed otherwise. The analogy to multiple puzzle pieces being parts of a single larger element in a recognisable puzzle picture (think of a yellow flower comprising ten puzzle pieces) is clear. Two differently shaped puzzle pieces that both contain a spot of yellow on them, but in different places,

are analogous to a phone number appearing twice in the database, but with a transposition error in one of them. Mostly, the phone number is still recognisable at a glance even with the transposition, so if it appears twice in the very large database, that's twice the probability it will be found earlier.

Overall conclusion: the more data you have, the less compute time you need to find correlations.

Example: given a very large database, count the number of unique entities in it. The first time he tried this, he concluded there were three billion people living in the United States. But as connections were found, entities began to collapse together, until the rising curve showing the number of people found began to level off, then descend, and finally converged to the true value.

Result: a method for detecting professionally fabricated lies, using only the raw data. They found six examples in this very large database of multiple identities constructed by a person. The correlations between fake IDs were tiny: two fake IDs shared the same library card number in one case, or the telephone number of two apparently unconnected fake IDs in their fishing licence applications was the same. They found six people who all had lots of fake IDs for some reason.

Defence against professionally constructed fake IDs: there are only two possible defences. The first is to collect observations your adversary doesn't know you have. The second is to compute over data your adversary cannot fathom; if your adversary knows what is deducible from the data he knows you have, you are sunk.

Organisations become the victim of the index of the systems they are trying to correlate. A central catalogue or index, the way Google works, is the way to overcome this hurdle.

Top Three Reasons Information Sharing Will Fail:

1. Everybody wants to be an information recipient, but they don't share what they have.
2. 'You can't even share with yourself.'
3. The information sharing paradox.

Solving the problem:

'Discovery must come first'. It must work like a library card catalogue. The policy focus becomes discoverability. If you do not publish your metadata, it is not discoverable.

Best Practices:

1. Solve the discoverability problem first.
2. Data attribution is key; you have to know where you got it from or you won't be able to trust it later.
3. Data tethering; every add/change/delete should be tracked. This is why credit reports display every query that has been made against your credit at the bottom of the report. It was learnt the hard way that that was critical information to remember.
4. Generalised schemas; the library card catalogue system has worked for centuries, accommodating books, tapes, DVDs, and every other kind of media that has come along. The system still works.
5. Information collocation; keep your data and queries in the same index space. When searching for data, it can be illuminating to see what other people have also searched for that was similar to your query.
6. Analytics at ingestion; it is computationally more efficient than batch mode. Getting a slightly wrong answer quickly is better than a right answer that doesn't arrive until tomorrow or next week. Do batch mode searching only when there are no new data to look at.
7. Reduce the risk of unintentional disclosure by centralising your indexes. Fewer copies are easier to protect.

Analytics in the Anonymised Data Space:

This is really interesting. The presenter looked for a way to handle the situation when two entities don't want to share information. He solved it by hashing the personally identifiable information (PII) in the database, then comparing hashes. Sure, that's straightforward. But what he did next was clever: he hashed very spelling variant of every name (or at least the most common spelling variants). He hashed telephone numbers with common transpositions of their digits. This multiplied the amount of data for comparison hugely, but made anonymised comparison practicable. Thinking of a previous argument, he introduced noise to the data to make the 'puzzle solving metaphor' work better. You can still compare the hashes blindly, but you find many more correlations that would be missed by an exact comparison of exact hashes. He reports 'some success' in creating these anonymised indexes.

So, if organisations can share data anonymously, why aren't they doing it? See a blog post titled, 'When Federated Search Bites'.

Conclusion:

'Every time you learn something, ask "how does this relate to what I already know?" '

Backup slides:

More on the puzzle experiments:

They mixed up pieces from 4 puzzles as follows:

10 percent duplicated pieces;

30 percent missing pieces;

40 percent pieces from other puzzles;

6 pieces were outright lies.

The sequence of events observed in puzzle assemblers was familiar; he found good agreement in the times when certain key events happened:

- finding the first connection;
- finding the first duplicate piece;
- identification of the overall picture;
- first realisation that there are at least two puzzles here (and later three, and then four)
- 'this piece doesn't fit anywhere'
- 'let's take this irrelevant part (sky, grass) off the table'
- 'this piece is a lie' (the experimenter was accused by the players of being evil)

Comment from the audience: the puzzle experiments assume that there will be a picture buried in there somewhere.

Comment from John Benner, Booz Allen Hamilton: you said 'more data, less compute'. Could that be better phrased, 'more context, less compute'?

The more general your database schema is, the less brittle it is when adding new data feeds. We have worked with 4000–5000 data feeds, and if adding a new feed makes for a lot of work, you're never going to get anywhere.

Question from the audience: what are three of the most frustrating things you see government agencies doing?

Answer: collocate your data, make it discoverable, share it. The president's Information Sharing Sharing Environment (ISE) is very good.

Question from the audience about massively parallel database search—Hadoop.

Answer: in the puzzle experiments, what was found to be most important was how quickly you can make sense of each new piece as it arrives. This is not a problem that map reduce can solve. But the periodic 'data mining' activities are a problem that Hadoop can solve—no new data are arriving, you just contemplate what's on the table, looking for correlations and connections that you didn't notice before. (The presenter called it 'deep inside reflection'.)

Too many organisations are trying to do batch processing when they should be doing real-time analytics instead.

Mr Jonas is an extremely good speaker.

0900 Dan Nichols (panel discussion) on 'Challenges in providing cross domain solutions—an industry perspective' with:

- Dan Nichols, UCDMO
- Dr Ron Mraz, Owl Computing
- Bill Ross, General Dynamics C4 Systems
- Ed Hammersla, Raytheon TCS
- Russell Dietz, SafeNet

Dan Nichols: 'From an industry perspective, what are the keys to achieving enterprise interoperability and sustainability?'

Dr Mraz (Owl): 'reliability, reliability, reliability' (which maps to availability in the C-I-A triad).

Mr Ross (G-D): incongruity between C&A schedule and vendor's ability to deliver and field products. Lack of determinism in the process. Uncertain costs in C&A. We need better standards in remote management and cross domain as a service. We need composable security.

Mr Hammersla (Raytheon): you need to have three happy people at the end of the day: the operator, the accreditor and the data owner. If you tweak too much to the security side of the equation, the user won't use it.

Mr Dietz (SafeNet): 'visibility, viability, and variability'. You have huge deployment issues. Viability means the ability to scale and interoperate. Variability refers to the fact that there are an infinite number of trust domains and edges now.

Question from the audience: how to address coalition demands where 'their side is the high side'?

Question from the audience: please address the questions and issues involved in cross domain as a service.

Mr Ross: abstract out the implementation details before you can get to cross domain as a service.

Mr Dietz: manageability makes a lot of sense for a service, but the actual processing of messages is custom and complex.

Comment from the audience regarding Mr Garriss's talk from yesterday about HTML 5. It seems like the number of vulnerabilities is growing exponentially; HTML 5 introduces an order of magnitude more new vunerabilities—we have reached the limit of what hand crafted rules can handle.

Dr Mraz: recall the three V's; from Owl's perspective, a true one-way guard, HTML 5 vulnerabilities are not really a problem for us.

Mr Ross: there will always be a place for filter rules, but there needs to be a paradigm change to handle complex protocols like HTML 5.

Mr Dietz: shift towards contextual analysis to handle complex protocols.

Question from Kevin Dixon, Marine Corps Systems Command: we have a requirement for a hand-held CDS. How can we get industry to adopt our needs as their priorities?

Mr Dietz: look at the devices used in medical institutions right now—eventually the handset and portable device work being done in medical applications will be applicable to the Marines' problem.

Mr Ross: the Trusted Computing Group standards are moving in the right direction to help solve your problem.

Dr Mraz: Owl does a lot of self-funded research in one-way solutions, looking for what customers want. At first, all our customers wanted solutions on Windows NT. We sold that, but we also started working on a Solaris version, and soon after that was done, half our customers immediately switched over to Solaris. So we started working on a Linux version, and that immediately took over. The next thing we expect to be mobile devices, or maybe virtual.

Next question: how does the panel see NIAP as relevant?

Mr Hammersla: NSTISSP No. 11, the requirement that drives US government systems to the Common Criteria, seems less important these days to the intelligence community; they need solutions quicker than the 2–3 years it takes NIAP to evaluate.

Dr Mraz: NIAP works well for certain products where a Protection Profile exists, but more adaptable products are more suited to the accreditation, not NIAP certification.

Mr Hammersla: meeting NIAP requirements has actually reduced the security of some products because of meeting NIAP's specific security controls and requirements.

Question from Dan Nichols: what untapped value can cross domain bring to enterprise Computer Network Defence (CND)?

Dr Mraz: look at our public utility customers. They have to satisfy certain regulatory rules and at the same time make a profit. They are using cross domain to separate SCADA devices from business networks. One company has 20,000 critical infrastructure devices behind a CDS that protects them from new threats long enough that the utility can schedule installation of patches on these SCADA devices at reasonable intervals.

Mr Dietz: the commercial world uses cross domain to create zones of isolation to slow down attack spread and save money that would otherwise be spent chasing problems.

Mr Hammersla: taking significant costs out of the equation at the same time as you're increasing security. For example, replace five PCs on one desk with a single thin client, and you lower the attack surface dramatically.

Mr Ross: you can utilise CDSes as a general sensor for IDS in the network to save money.

Question from the audience: how to reduce costs?

Mr Ross: identify reasonable building blocks.

Mr Hammersla: cost is the highest priority of our development team. (He gave the example of the Slingbox product as being a device that is extremely easy to set up and use.)

Mr Dietz: as we go through C&A, anticipate hurdles to lower their costs when you eventually get there.

Dr Mraz: modularity and re-use. Each module can be examined as a separate entity. Modularity and layering are key to reducing costs.

Closing comments:

Mr Dietz: need visibility to control and manage CDS. Modularity reduces cost of C&A. Flexibility and variability of use are key to spreading cross domain into commercial applications.

Mr Hammersla: wants to see community expand use of cross domain beyond where it is used today.

Mr Ross: how do we make the process more deterministic? Transfer, access variability in processes is huge. Want a repeatable process independent of interpretation.

Dr Mraz: I am surprised we got no questions on cloud computing. I would have bet a dollar we'd get a question on cloud computing. Enterprise CDS is the way to go in future. Small organisations should concentrate on their mission and outsource CDS to a service.

Move CDS closer to the edge.

Track 2: international CDS

(UK) Mr Hugh Colborn, CDS Policy and Strategy Lead, CESG

'Cross Domain Solutions in the UK: some thoughts from CESG'

Five Eyes partners are in attendance at this session.

CESG is the direct UK equivalent of NSA's IAD. It is pan-government, neither MoD nor civil nor IC only.

'Do our views differ? We all have the same basic problems.'

On the Inbound side: malicious content, appropriate classification, and trusted source.

On the Outbound side: accidental release of information, over- or under-classification, covert channels, malicious or wilful release of information, and malicious content.

The interesting difference is that if you ask a US person, they always say that disclosure of information is the most important problem, followed by malicious content getting in. The UK's answer is exactly the other way around: UK considers malware getting in to the classified system to be the most important problem, and spillage of information out is much less important. That is because 'we don't care who has our information, we don't care about Wikileaks'. The US always feels that accidental disclosure is a disaster. UK care more about malicious content getting in to our systems.

Cross Domain Solutions:

The UK has no good definition of CDS. MoD defined it best as 'exchange of information between trusted networks and those of lesser security or higher risk'. MoD defined the problem; the solution is whatever solves your problem.

Current CDSes tend to be expensive and sometimes don't deliver as much as the user believes they do. Classic example: people believe data diodes do something. They don't, they just pass data in one direction. There is no magic pixie box.

The thought that one puts one magic pixie box at the edge (popular with vendors and beloved by users) is fundamentally flawed.

CDSes are difficult to assess using conventional assurance methods.

The idea that 'high EAL equals high assurance' is wrong. A data diode at EAL7 is irrelevant if someone can take control of the NIC on the end, or of the graphics card, and start running processes there.

'As an attacker, I don't care if the vulnerability I use is in the scope of the TOE or not.'

'High EAL requirements are Nature's way of telling you you're doing something risky.'

CESG is researching problems and investigating proofs of concept for solutions, but not developing GOTS products. We develop architectures, patterns, produce guidance, research papers, and technical notes.

History of the CDS team:

Part of IATP. Started in 2006 as Assured Information Management (AIM). The title was changed from Assured Information Sharing (AIS) to Cross Domain Solutions in January 2011.

But we are not building CDSes.

We have a very strong line of educating HM government and industry on cross domain IA matters. Some building of exemplars, proofs of concept, but mostly we produce architectures and guidance.

A Model for Cross Domain Mitigations: 4 steps.

1. Identify (for example, did it come from where I thought it came from?)
2. Verify (does it look like what I expect?)
3. Transform
4. Render

This model was published in the CESG IA portfolio. It is unclassified, but might not be on the public web site.

Assuring CDS:

We do information assurance and risk management to support the successful delivery of a mission.

End to end solutions.

Building architectures out of components of lesser assurance.

Place assured components into architectures sensibly.

Move with the changing threat.

Patch! Don't defer patching because it gets in the way of your uptime. Patch!

Question from the audience: 'do you believe you can have product assurance if you continue to outsource development of solutions?'

Answer: there is not so much proliferation of solutions in MoD as elsewhere because a lot of what they do is outsourced. There are geographic restrictions on where they can source their equipment. Sometimes there is not a lot you can do about that.

The next talk was by Francois Luneau of CSEC

Canada: 'CSEC Cross Domain Solutions Activities: Past—Present—Future'

All the people working in this department started out doing manual transfer, so they understand the pain. Manual transfer still had to be effective, efficient, and secure.

CDS SA&A (a.k.a. C&A) advice and guidance: security controls, security assessment, filter security requirements and assessment, CDS prototyping and development, CDS publications.

A lot of C&A people do not have experience with CDS. So CSEC advises and educates the C&A folk.

CDS SA&A (a.k.a. C&A) advice and guidance

Current and Future:

Bodies of evidence required for SA&A being developed by RMF process and System Security Engineering process.

What should be the minimum body of evidence required for SA&A?

- threat scenarios

- T&E report: noncompliant security controls, impact of noncompliant security controls, and recommended compensatory controls

What should we do if we cannot obtain the minimum information required to perform a risk assessment?

Answer: reproduce the evidence.

Security Controls: most T&E reports in Canada are based on DCID 6/3. Some T&E reports are based on NSA SR 1 to SR 9. Some CDS T&E reports are based on NIST SP 800-53. The future is 800-53.

DCID 6/3 had this useful concept in it of Protection Levels.

What should T&E objectives be?

- to verify compliance of security mechanism to the security controls

- to verify effectiveness of security mechanism

- to provide evidence in support of the risk assessment.

What Risk Assessment methodology to use? Various, but RDAC is predominant for SECRET and below in Canada.

RDAC defines a matrix with threats down the left side rows and security controls in the columns along the top. RDAC implicitly creates relationships between threat scenarios and security controls and T&E.

Filter Security Requirements and Assessment:

- not covered in sufficient detail by 800-53.

- what should the filter do?

In the past, we defined filter requirements by High Level Requirements: prevent malware in, prevent spills out.

Currently and in future, we define filter requirements by a detailed requirements document in the 'Filtering Inspection and Sanitisation Guidance Document'.

CDS Prototyping and Development in Canada:

In the past, we developed low to high transfer CDS, XML CDS, and CDS for email.

Currently and in future, focus on integrating CDS solutions and tailoring them using adapter software (adapter design pattern). We are not going to develop any more.

CDS Publications in Canada:

- CDS guidance publications: CDS primer, CDS Business Architecture Document.

- Access: ECDS technical guidance for access solutions, CDS Access Solutions—architecture and risk management framework.

- Transfer: ECDS technical guidance documentation for Portable Media Transfer; data diodes

- FiST: reducing the risk of portable media transfer

- Data filtering: technical basics, security controls document, requirements.

Question from the audience regarding FiST:

Answer: 'Forcing CDS to be more than anti-virus'; understand their filter requirements.

Right now, CSEC is out of the business of building CDS and wants to get out of the C&A business. In future, they want to provide guidance and be the CDSO to point users toward solutions.

Australia: Josh Gill, Cross Domain Evaluations, DSD, on 'Assured Information Sharing: an Australian Perspective'.

DSD is NT authority for Australian government.

Australian government Information Security Manual (the ISM) policy.

Evaluation of CDS system design and products.

C&A of TS systems.

Advice to government.

AISEP Common Criteria

Crypto and High Assurance ('High Grade' is their cryptographic High Assurance term)

Cross Domain.

CD Evaluations:

DSD provide planning advice to system owners, architectural assessments in support of C&A, and T&E.

'Our customers are the whole of the Australian government': defence agencies, intelligence agencies, government ministries, foreign affairs and Trade, law enforcement; plus a high number of small agencies with limited budget for infrastructure and staff.

Examples of requests for service we have received:

- 'what firewall does agency A need to connect to agency B?'

- 'which firewall can I use to connect unclassified to SECRET?' (this was a real question!)

- 'DSD policy saus I need a CDS. What is it?'

- 'why can't I use virtualisation to separate information with different classifications or caveats?'

Thoughts on these common questions:

DSD does not certify solutions. The accreditation authority takes advice from DSD. We don't have a CDMO organisation yet.

DSD needs to make their policy more clear.

Need to improve policy and guidance documents.

The topic of virtualisation comes up very, very often. Pressure for DSD to revise policy to allow separation by virtualisation. Lots of pressure from CIOs throughout Australia.

What do we do well? (According to our customers)

- guidance early in projects (greatly reduces pain of C&A)

- guidance on specific technical questions

- receiving information shared by our generous international partners, i.e., the U.S.

- testing and breaking systems (fun, but really only a small part of our job, due to ITAR and intellectual property restrictions on what we can get our hands on to break)

What have we not done well?

- providing enough information for customers to make their own decisions.

- enough information to systems designers and builders.

- availability to customers and accreditation authorities.

- feedback our successes and challenges to those who've shared information and given time to help.

Challenges to how we do business:

- Technology convergence:

+ non-high-grade encryption

+ 'COTS for classified'

+ virtualisation

+ SOA, ESB

+ mobile computing
+ bring your own device to work
+ sophisticated content inspection
- Process, Policy, and Strategy:
+ rationalisation of national security classifications (see below)
+ demand for flexible business processes: 'just make it happen'
+ soverignty
+ information sharing

Rationalisation of National Security Classifications in Australia:

Trying to go from an 8-level system (TS, S, C, 'highly protected', restricted, protected, 'X-in confidence', unclassified) to seven slightly different levels:

TS maps straight to TS.

SECRET and 'highly protected' both map to S.

C maps to C.

'Restricted' and 'Protected' both map to 'Protected'.

'X-in confidence' and unclassified both map to a group of three new levels: 'dissemination limiting marker', 'unclassified', and 'government'.

Where to from here for Australia?

- provide enough information for customers to make their own decisions;
- provide enough information for system designers and builders;
- further development of ISM policy for gateway, CDS, firewall, content filtering, and data inspection;
- publication of DSD Guide to Secure Configuration of CDS;
- development of a CDS assessment guide;
- continue to engage with SME's—the 5eCDSWG, Australian government, IT security advisors, industry experts—all have provided constructive feedback.
- 'Data diodes provide excellent isolation but they are not a CDS';
- improve throughput of evaluations, availability to customers and accreditation authorities (continual education and discussion—we have a forum called 'onsecure' at `onsecure.gov.au`);
- feedback our successes and challenges.

Email: `josh.gill@defence.gov.au`.

1400 Phyllis Lee, NSA on 'Virtual Server Separation'

Why virtualisation?

- Separation can be accomplished via two separate computers. Low risk for accreditation, more expensive in cost, space, weight, and power, and it doesn't scale.
- Process separation on the same piece of hardware by an OS. The separation is only as good as the OS provides.
- Virtualisation. The second and third pictures look very similar. Why do we consider them different? Because virtualisation allows us to pull security relevant entities into different containers where they can be controlled (communications into and out of each virtualisation container can be controlled).

Assumptions:

1. Hypervisor is less complex than an OS. It presents a smaller attack surface.
2. Communications into and out of a container can be constrained and controlled.

NSA kept getting asked whether virtualisation was sufficient for separation. So they did this:

- writing requirements for allowing multiple Communities of Interest (COI) to coexist on the same virtualised server.
- participate in DMTF (Distributed Management Task Force) with all the major virtualisation vendors in the membership.
- NSA working group of virtualisation experts.

NSA's assumption is that testing laboratories do not have virtualisation experts. So NSA needs to provide guidance to labs for performing evaluation of vendor products.

NSA asked itself two questions:

1. What would be the idealised architecture for virtualisation?
2. How would we defend it?

More NSA assumptions:

- virtualisation is configured correctly.

- hardware is not buggy.

Finally, NSA wants to engage vendors.

The output of this process is:

- DMTF: virtualisation protection profile (VPP) at EAL1 or EAL2. (Note: NIAP has been revamped; see below.)

- white paper: what it is, what it takes to do their VPP; learning curve.

- government input into VPP. 'Will there be an Appendix C?' ('Appendix C' is NSA's term for the place where they put all their nice-to-have, or want-to-have features.)

They are hoping to have this done ('assurance activities') in the next 2–3 months.

Continuing Work:

- agreeing and writing down requirements, how to measure if a requirement was met, 'Appendix C'

- publishing a VPP.

'It is important every year or so to go back and look at previous guidance and ask yourself if it still makes sense.'

Question from the audience: will you be pulling the TCG's TPM into the VPP?

Answer: Yes. TCG are working on things relevant to the DMTF.

Question from the audience about NSA's assumptions that labs have no virtualisation experts, that hardware is not buggy, and that virtualisation is always configured correctly. The questioner (from NRO) basically called out Phyllis Lee on all these assumptions. She looked ready to kill, but ignored the question and went on. The questioner said things like 'we cannot achieve our goals without virtualisation, but we know that containers contain covert channels, even though NSA cares not about covert channels any more. Some of us still do care.' [uneasy laughter from audience]

Another question from the audience: 'given some of the concerns expressed, do you expect to see an actual SABI accreditation any time soon?'

Answer: 'I am not working on a problem I think is hopeless.' Have to look at the assumptions, e.g., that the box is configured correctly. The box must be configurable, else it is useless. These are things we need to figure out.

'For this use case, we know what the security requirements are.'

Question from the audience about High Assurance. Answer: virtualisation is not HA. HA is probably not achievable on commodity hardware. Is virtualisation as good for separation assurance as two separate boxes? No.

Question from the audience about NIAP: 'did they change the EAL levels of NIAP?'

Answer: go to the NIAP web page. They will only recognise international evaluations up to EAL2 now. NIAP EAL2 does not mean what EAL2 used to mean. They added a lot of new assurance requirements to EAL2.

DoD policy will be to only accept international evaluations to the EAL2 level, even if it says on the box higher than EAL2.

Question from the audience: does the VPP contain assumptions for communicatiom between containers?

Answer: at present, they are only looking at separation, and they are constraining communications.

1430 Scott Loftin, HS-CDS Senior Technical Director, Concurrent Technologies Corporation, on an NRO project to virtualise servers.

'Virtualisation is the execution of a guest OS atop another OS, called the host, where the guest is abstracted from the underlying hardware.'

Three kinds of virtualisation: full, partial, and paravirtualisation.

Why virtualise CDS? Scalability, availability, security through improved monitoring and separation. Improve deployment speed. Better hardware capacity utilisation. Abstract away periodic hardware product life cycle changes.

Concerns: lack of formal testing criteria, potential increase in cost, possible performance impact.

Different types of hypervisors: bare metal (type 1) or hypervisor hosted on an OS (type 2).

HS-CDS: for NRO, a cloud of virtual, multi-vendor guards with remote monitoring and management of that cloud.

All three of Radiant Mercury, ISSE, and AFT have been virtualised successfully by them. They have run 128 virtual RMs and 32 virtual ISSEs on the same box at the same time.

They can create more virtual guards instantly. They can choose the appropriate mixture of guards, e.g., a bunch of AFT to handle an influx of Microsoft Office documents, on the fly.

They have initiated C&A at one IC agency to create a normalised image metadata repository. They have just started the C&A effort; haven't even had a PDR yet.

HS-CDS architectural diagram. The current C&A will be PL-3, so only one set of remote monitoring and remote maintenance workstations will be allowed to exist and must be in the same same security domain.

Question from the audience: why are you doing C&A to PL-3 if you have PL-4 guards?

Answer: because the system we are building is a PL-3 system, or maybe PL-3 with DCID 6/3 Appendix E because of foreign persons. It only requires PL-3.

The virtualisation platform is VMware ESXi 4.1 host OS, 64-bit, 96 cores, 512 GB RAM, 2 TB storage, Solaris 10TX guest OS. RM 5.01 will be the only guard being taken through accreditation, not ISSE.

Question from the audience: why did they choose VMware over Xen or Red Hat's KVM?

Answer: Xen's Dom0 (if not broken up) is a big technical risk.

Lessons Learnt:

- use dedicated physical management LAN and operate all VMs at the same security level. It makes accreditors more comfortable.

- make use of stateless VMs. Makes it easy to move a transfer in progress from a failed or compromised VM to a new clean one. Just extract any forensic information (log files, whatever) from the old VM before you destroy the old one.

- take advantage of VM cloning for creating VMs, testing, and disaster recovery. You can create standby VMs for failover purposes easily. Plus, you only have to patch once.

- improve uptime by patching a newly created set of VMs before bringing them on-line, then instantly cut-over from old to new VMs; this is possible because they are stateless.

- take advantage of dynamic and remote VM cloud configuration. You can view real-time hardware resource utilisation, change number of CPUs, memory, and storage on the fly.

- use hardware routers and switches as opposed to virtual networking.

- create or utilise products and tools for all system and application logging. I.e., centralise, centralise, centralise. 'Don't be like the Raytheon High Speed Guard where you have to shut down every few hours to pull off the logs so the guard doesn't stop working.'

- oversize your physical computing hardware. But don't create a giant single point of failure. Consider using a small number of smaller virtualisation servers instead of one big one.

- understand software licencing agreements and update policies. Some vendors charge more for virtualisation on very large hardware.

- don't underestimate security and testing.

Takeaways:

- Virtualisation, while complex, allows for increased flexibility;

- but risks exist; pay as much attention to the security of the management console as to the hypervisor itself;

- virtual and non-virtual machines (and guards) can be used together in a cloud environment.

1545 Dan Bradley: 'Don't Use a CDS and Avoid This Conference'

Of the NSA Secure Information Sharing and Validation Division: `djbradl@missi.ncsc.mil`.

Reasons Why You Might Not Want to Use a CDS:

Why is cross domain so difficult?

Why does it sometimes not meet expectations?

The Laws of IA Physics. Sometimes we want to do things that might not be a good risk.

Threat needs to drive security requirements.

The Laws of IA Physics vs stated technical requirements.

'Maybe we don't have a good Javascript guard or a good executable guard because those are just intractably hard technical problems.'

Don't impeded information sharing by imposing technical requirements not backed up by threats.

Both threats and user expectations are going up. Our choices are, therefore:

1. Increase our tolerance for risk ['increase risk']
2. Talk the users into accepting less capability ['lower CDS expectations']
3. Improve CDS risk mitigation capabilities.

'Have you noticed that CDS capabilities today are not significally different from 10 years ago?'

The Laws of IA Physics:

- complexity is bad.

- trust and trustworthiness: if there is a disconnect between the level of trust and the level of trustworthiness in your solution, then risk is heightened.

- risk is end-to-end.

What is a 'Good' CDS? One that is sufficiently trustworthy for the amount of trust we need to place in it.

Don't use a CDS:

- is your data already on the high side?

- is it worth the risk to move it up?

- can you move it via a lower risk manual process? Don't write off air-gapping, but also don't do air-gapping poorly.

- is the benefit worth the cost and time?

- can you scan or retype the information?

- can you 'swivel chair'?

Contact your service or agency Cross Domain Service Element (CDSE) for guidance on the C&A process.

Examples of not using a CDS:

- KVM, commercial remote computing, thin client or zero client solutions: operate natively on the target network.

- Digital senders: these devices are like a scanner that creates a PDF and emails it automatically to the high side. Nearly foolproof and low-risk filters that work on any data format.

- Authorised 'air gap' transfer process: air-gap done properly.

'Have You Seen this AOA?'

1. the CDS we cannot remotely afford.
2. the CDS we have structured our requirements to justify.
3. (something completely random)
4. copy and paste from another AOA.

Why is it so difficult to build a CDS?

- mitigation of content risk is highly dependent on the data type.

- file complexity

- structured data permits use of Explicit Allow policy.

- unstructured data only supports an Explicit Deny policy.

- tight data specifications produce better quality filters.

Alternatives:

1. Use an access solution.
2. Add a channel to an existing CDS. Is there a Radiant Mercury already nearby?
3. Use a unidirectional solution.
4. Use highly structured fixed format content.
5. Isolate your CDS from community networks. Local risk vs community risk.
6. Use a 'mission network' to support collaboration within an enclave.

Example of a Mission Network:

When the mission is split across networks, risk is high and you need lots of CDSes. If you truly have one mission, put it on a common network. This is where the Afghanistan Mission Network (AMN) was two years ago. Now they are on one network.

Three Types of CDS and How to Use Them:

1. MILS/access solutions. There is no risk if you don't mix the content. Cryptographic tunnelling across a common transport is an example. This is the lowest-risk solution.

2. Structured data. Reduce the need for unstructured transfers. Prefer to use well structured and highly structured XML, for example, instead of PDF.

3. Unstructured data. Use a mission net. Use data transformation to neuter the content.

'Want less risk? Don't cross boundaries.'

Data Transformation Example:

Afghanistan Theatre Video Bridge (ATVBv1.0) uses H.323, a very complex video protocol. If there is a worse format for cross domain, it would be hard to find one.

NSA proposed an H.323 to analogue to H.323 bridge. User feedback has been positive. Reliability is excellent. Conclusion: data transformation can help solve the unstructured data problem.

Bottom Line: is a CDS worth the time, cost, and life cycle support requirements? Is it worth the risk? Did your AOA include a realistic analysis of non-CDS requirements? Can you use a mission network instead? Can you use data transformation to reduce unstructured data?

Question from the audience about the ATVB.

Answer: Yes, that was a CDS. Data transformation is a type of CDS.

Question from the audience about the VLR (Very Low Risk) programme. Is it an alternative to C&A?

Answer: VLR looks at the environment as what needs to be evaluated for risk, rather than what is inside the dotted line around the CDS. If you think VLR is a cheaper alternative to CDS C&A, you are misled. People who have tried using VLR have found it to be much, much more expensive.

Instead of assessing the controls on the CDS, the VLR programme assesses the security controls on the entire environment.

Dan Bradley is a very good and entertaining speaker.

References