File 20110218.0250: Weekly activity report 0176:

weekly activity report 176 (loughry)
Joe Loughry
Sent: 18 February 2011 02:50
To: Niki Trigoni; Andrew Martin; Joanna Ashbourn
Cc: otaschner@aol.com; anniecruz13@gmail.com; andrea@hpwtdogmom.org;
chip.auten@comcast.net; edloughry@aol.com; diane@dldrncs.com;
Joe Loughry; mmcauliffesl@comcast.net; tom.a.marso@lmco.com

Weekly activity report no. 20110217.1235 (GMT-7) sequence no. 0176, week 5 HT

UCDMO issued a new update of the Cross Domain Inventory baseline (version 3.7.0, dated February 2011) this week that included significant changes in the Transfer category of CDS solutions available for reuse. ISSE was dropped from the list; RM 5.0 is still indicated without a SABI/TSABI annotation footnote, and both ISSE 3.6.x.x and RM 4.x.x and previous versions appear in the second section, as expected, with TTWCS. Operational CDS controlled interfaces on the sunset list are expected to be removed from inventory not later than 31st October of next year. In my meeting with Dr Martin this week, we discussed the interpretation I am beginning to write on the R’’ certification test and evaluation (first case study under the new methodology), beginning with the detailed outline for Chapters 2, 3 and 4 that I am currently working on. In Chapter 4, the case study overview of R’’ begins with transfer CDS solutions on the UCDMO baseline and a necessary amount of contextual narrative. Stakeholders (I have been using the word ‘participants’ until now, but stakeholders is the word used in grounded theory; I need to make sure I understand the definition of that term and that it does not imply unintended meanings) in the case study include the government programme office, COTR, IV&V representatives and liaisons, as well as the certification testers, certification authority---in both its advisory and evaluative capacity, the developer, and field sites and associated project managers and accreditors waiting on a certified CDS. The issues uncovered most obviously include cost, as expressed in time, resources and budget, but especially the time to completion of certification testing and the role of the GPO in pushing the certification testers to minimise their test coverage and the conflict engendered with the government penetration testers vs IV&V contractors. The COTR continually expressed and exhibited clear evidence of the difficulty of balancing this multi-way stress network during CT&E, but apparently could not or dared not bring more pressure to bear on the certifiers or evaluators for fear of slippage of the entire test effort schedule. One thing in particular that kept coming up was the expense of maintaining trained engineers on site during IV&V, regression and penetration testing; an evaluation of whether had this been done it might have resulted in a cost-effective reduction in calendar time needs to be looked into. Another aspect that remains to be analysed is the role of sites and the associated project managers with regard to important accreditations unable to proceed without an approved CDS and precluded by other pre-existing approvals from substituting a different controlled interface in order to make their deadlines. I will have detailed outlines of all three chapters for Dr Martin next week, I hope.

I have been reading this week a new book by Emerson, et al. ‘Writing Ethnographic Fieldnotes’ (The University of Chicago Press, 1995). I am not accumulating new field notes at the present time, but the second half of the book is relevant to coding and determining meaning from field notes, both of which are relevant to grounded theory. I should have that portion of the first case study done in another two weeks. I also finally acquired the out-of-print book by Burton (1993) on

Pentagon procurement policies and procedures; I borrowed a copy from the DU library to read over the weekend.

Dr Gus Hosein of the London School of Economics---and Deputy Director of Privacy International---gave the Information Security and Privacy Programme (ISPP) seminar this Monday on ‘Moving beyond ‘‘impacts’’: cheating for Privacy’. The presenter began by discussing the history of the UK ID card project. Privacy, he said, is not universally agreed to be well defined. It is a complicated mess because human beings are complicated. Surveys attempting to quantify attitudes and perceptions of privacy in the general population consistently return inconsistent results. Legal systems are equally confused, examples of that extending as far back as 1763 in England. Laws are not knowledgeable on technology; politically, in order to be acceptable to lawmakers, the UK ID card had no choice but to promise an hypothetical and theoretically invulnerable system, which they sought to implement with biometrics. It is known by now that biometrics can never be absolutely reliable.

eHealth (electronic health records) is the area where billions in expenditures are going now. But the people selling eHealth systems never talk about data integrity or data security. The tool called analytics (a.k.a. data mining) was the next big attempt to solve a stochastic search problem on privacy-sensitive information. The Department of Homeland Security (DHS) tried it with passenger name record (PNR) screening in the CAPPS2 programme early this century; eventually they gave up because they could not get it to work. The U.S. defence department then seemed to think, ‘where DHS couldn’t succeed, maybe we can’, and proceeded to inaugurate the poorly named Total Information Awareness (TIA) programme. TIA funding was killed by Congress several times; a few years later a panel concluded that data mining would never be able to solve the terrorist identification problem. Didn’t stop other governments from trying, though: Germany, France, and others. Policy-makers keep coming back to the TIA idea because they think these systems can be made perfect. Examples: Matrix (Lexis-Nexis) and Palantir.

One reason these attempts will never work is that technology is changing much more rapidly than policy-makers’ thinking. Referring to a 2008 Comm. ACM article on ambient intelligence, the presenter drew a linkage to metering and the smart grid. ‘The Regulation of Investigatory Powers Act was written for the internet era in 2000, but it ended up being written for a different internet era, one where we all used email through our local ISP.’ In 2008, deep packet inspection was floated in the UK. With the smart grid, ‘What will the government ask our electricity providers, for example, to do? Dr Hosein’s modest proposal: first stipulate that cheating and lying are an important aspect of human nature. We need to maintain the need to protect privacy; to remain human, we need to retain the ability to cheat. But how can we do that when our devices and our infrastructure are telling on us? So, let’s design cheating and failing into our systems.’ The Google Latitude system did it right. Too much information assurance, says the presenter---information that is too accurate---is a drawback. ‘We must be able to change our identifiers with ease’, and ‘this is where the UK ID system failed; they thought there was one person, one identifier, but there are many identifiers’. Eric Schmidt, if I recall correctly, said something in 2010 about teenagers being allowed or encouraged to change their names when it becomes necessary to disconnect from their previous online identities. There was also something in RISKS this week about correlation of visible user names to track on-line identities. The recording of the seminar I got did not contain the question-and-answer period.

Wednesday, the Oxford Security Reading Group met to discuss the paper

'Safe to the Last Instruction: Automated Verification of a Type-Safe Operating System' by J. Yang and C. Hawblitzel (Toronto: PLDI'10, June 5--10, 2010, pp. 99--110), suggested this time by Anbang. I enjoyed this paper. The tiny OS introduced by the authors is interesting in its own right (it has a wonderfully draconian approach to exception handling and thread life-cycle), but the combination of Typed Assembly Language (TAL) with fully automated verification of an entire, if minimal, OS is what was new in this paper. The authors trimmed their TCB down to a couple of assemblers, a linker and a boot loader (and bizarrely, an ISO image generator) but managed to get all the compilers into the untrusted set and they demonstrated convincingly that the boot loader, if it had to be trusted, was at least unable to propagate unsafe code past initialisation of the kernel.

John said something good during the discussion: 'shame it doesn't work on multiprocessors. It seems like everything in the verified world is stuck in 1970'. There followed some discussion of the Singularity OS, written entirely in managed code and with Spec#; managed code handled a lot of the work for the authors of Singularity and enabled them to include the amount of functionality that they did. Verve was developed by two people in less than a year. I suspect that Verve could be self-hosting (the authors don't say) but Bartok at least could be hosted on it---and certainly the assemblers; possibly also the Beat compiler but not C#, because the .Net libraries contain unsafe code. Shamal complained about how little functionality the OS has; I disagree. I think their method of killing threads on exception in certain instances is sort of an elegant little hack. Consider threads in Verve OS to be an opportunistic resource; they may be spent speculatively on things like querying the keyboard interface for activity---because they are very low cost (see the cycle counts in Section 7) when compared to the performance of seL4---and the simple IPC mechanism in Verve is a nice trade-off against the requirement for MMU hardware in seL4. I like the design of Verve.

There was some discussion of TPM hardware and where it ought to be integrated into the Verve OS. I argued for the boot loader; it is already trusted, although it is not clearly described in the paper and likely kept as simple as possible for that reason, but the TPM measurement could be done there and afterwards dropped out of memory entirely. I would even argue for bank switching to give an additional layer of assurance---it would be inexpensive in terms of circuitry and could be integrated directly into a verified MMU. If the advantage of Verve is that no unsafe code ever executes, then to maintain that discipline and simultaneously retain the primary advantage of the design, preserves the attribute of transparent simplicity. Not everyone in Reading Group agreed with that.

I like TAL because it feels like programming on the bare JVM. In reading this paper, I found it useful to look up another paper by Morrisett, et al. (ACM SIGPLAN, 1999) that was related to one of the references; TALx86 is another typed assembly language inspired by limits in the safety of the JVM. Clearly, there must be a second paper from Yang and Hawblitzel to come. Some of the unanswered questions in this paper suggest a direction in which the authors might be going: to extend the Nucleus to incorporate MMU protection for devices, which would probably result in a smaller Nucleus, but at the same time offer the additional layer of protection of discrete memory spaces for applications---a win, so long as the MMU is already on the board. The kernel would would grow to include some MMU control functionality; at the same time, a TPM could be incorporated, perhaps treated like any other device on the other side of the MMU and kept in user space, yet with a guarantee of safety and memory protection (protect from inside threats and outside) around the TPM measurement.

I think their TCB could be made smaller. The anecdotes at the end about the problematic kernel debugger struck a chord with John; he related similar advice deriving from Webinos. I was puzzled by the postconditions in the example in Section 4.1; the postcondition ‘ensures a < b && a == x + old(a);’ seemed incomplete to me; I wondered why it did not contain an additional clause: ‘&& b == y + old(b)’. If the additional postcondition expression were included, the function seems to be tautological, but when it is left out, I wonder if the authors believe that ‘a < b && a == x + old(a)’ somehow logically implies ‘b == y + old(b)’. It makes me want to look up the paper by Barnett, et al. (2006) for more on postconditions in Boogie, but I did not have time.

GSS reports for Hilary term are due in two weeks. I have been worried about finances for a few months; this week a potential alternative means has come about and, combined with refinancing and moving some funds around, it may alleviate that particular stress factor soon. My next meeting with Dr Martin is set for Tuesday, 22nd February at 5pm UK time; Dr Martin will be on travel but we will do it over the network.

My current tasks, in priority order, are:

1. Detailed outline of Ch. 2, 3, 4.

2. Update the literature survey with references on grounded theory and Qual. research.

3. Import remaining R’’ case study source material into ATLAS.ti.

4. Figure out a way to make event traces in the H. unit linkable.

5. [Waiting] Figure out if I can use the chronological record in my lab notebook as a source for the ‘memo writing’ activity that occurs later [not done yet]

6. Plot tasks on a new Google Calendar as blocks in a 168-hour week. Establish limits on non-thesis work times. [Still not done; I forgot to do it last week.]

7. Survey article needed for summer [not started yet].

Joe Loughry
Doctoral student in the Computing Laboratory,
St Cross College, Oxford

End of WAR 0176.

References