EAC Voting System Test Laboratory Discussion of TGDC Draft VVSG
March 19, 2008
9:00am – 2:00pm

Hyatt Regency Denver
650 15th Street
Denver, CO 80202

iBeta Written Responses to the Discussion Questions

1. The 2005 VVSG states one of the goals for the next iteration of the VVSG as being to create performance guidelines that promote innovation rather than design orientated guidelines that limits design choices. Do you think this document achieves that goal? Do you view performance guidelines as sufficiently testable?

iBeta response: The VVSG does provide definitions of Performance, Design, and Functional Requirements although the guidelines rely on those definitions throughout all 3 parts. For this discussion, iBeta is defining these as:

> Functional Requirements - Requirements that specify what the delivered system must be able to do; the functions it must perform.
> Design Requirements – Requirements that specify the normal operating environments for voting systems – Non-functional requirements that impose constraints on the design or implementation such as those that specify the normal operating environments for a voting system.
> Performance Requirements – Requirements that specify something about the system itself, and how well it performs its functions. Such requirements are often called 'quality of service requirements.' Examples of such requirements include availability, testability, maintainability, and ease-of-use

If the goal was to provide performance and functional requirements, the large number of design requirements within the guidelines would indicate that this goal was not met. As examples:

Part 1 6.4.1.5-A	Doesn't this limit allowable languages? By this requirement, current voting system applications would be made obsolete. Is this the intention?
Part 1 7.5.1.4-A.5	This design requirement limits innovation by requiring TCPIP and wired connections. Is it the intent of this requirement to impede use of current and emerging technologies?
Part 1 7.5.1.4-B	The discussion implies that this requirement applies to all activation devices even those that do not have the capability of being networked (the preceding requirements do not require that a card activation device be attached to a network). Is this a design requirements forcing a component to be designed to support a capability that is not that system's functional requirement?

Whatever the classification of a requirement, all good requirements are cohesive, complete, consistent, correct, current, and verifiable (testable). As a VSTL, the requirements that force qualitative pass/fail criteria are a major concern. A small set of examples (please see the iBeta VVSG public comments for the full set) is as follows:

Part 1: 3.2.8.1-A "…SHALL be reasonably easy…"
Part 1: 6.4.5-B.c "…SHALL allow for easy access…"
Part 1: 6.4.5-B.d "…SHALL allow for easy adjustment…"
Part 1: 3.3.5-C "…SHALL be easily legible…"
Part 1: 6.4.1.4-B "…SHALL be small and easily identifiable."
Part 1: 6.4.5-B.a "…SHALL allow for a non-technician to easily detect…"
Part 1: 6.4.5-B.b "…SHALL allow for a trained technician to easily diagnose…"
Part 1: 6.4.7-C.1.a "…SHALL provide a means to safely and easily handle, transport, and install…"
Part 3: 5.4.1-C.c …SHALL prioritize testing effort based on the OEVT team's determination of easily exploitable vulnerabilities….
Part 3 4.3-B "obvious inconsistencies"
Part 1: 3.2.6-C …"SHALL be designed to minimize accidental activation."
Part 3 5.2.3-C … "gracefully"
Part 3 2.6.2.1-D"..the archive files SHALL be generated using algorithms and file formats in common usage
Part 3: 3.4.3.3-B The final statement "If the test lab does not possess the required hardware and software to create the build" conflicts with requirement of 2.4.3.1-A.4.

2. How can innovative systems be evaluated for purposes of certification? If the EAC were to undertake creating an innovation class what suggestions would you make regarding the testing of innovative or new technologies?

iBeta response: Innovative systems have been handled in the 1990, 2002, and 2005 standards. The requirements for innovations in the past 18 years have been found in the standards. The issue is interpretation of the standard in terms of the innovation. Technology may change but that will not change the requirements, only the interpretation of the requirement in terms of the innovation. The interpretation identifies if a specific requirements applies to all or some of the innovative voting system. We would suggest that the current EAC interpretation program is the way to address innovations. If an manufacture application is received with an innovation, the manufacturer should submit the requirements for the design of the innovation and their proposal on how the EAC should interpret the current requirements as they pertain to the innovation. The EAC would issue the interpretation of how the VVSG requirements apply to this innovation. If new requirements are needed these can be provided in the interpretation. It is possible that interpretations may require further clarification but in this instance a revision is issued to the interpretation. Perhaps a regular schedule of bi-annual updates should be implemented for incorporating interpretations into the VVSG. Prior to the update the interpretations can be opened for public comment. In this way no one is blind sided by a

new standard or new requirements because everything is already published through the interpretation process.

We do not support the idea of the EAC creating an innovative class for two reasons.

1. The position paper "Voting Systems Innovations Class" prepared at the direction of the STS Subcommittee of the TGDC on June 26, 2007 conflicts with the direction to the VSTLs provided in EAC NOC 07-002. The position paper states, in Section 5 - Steps in Reviewing Submissions; that the steps involve an initial submission of a proposal followed by one or more rounds of review by the test lab. Per the EAC Draft Lab Manual and the EAC NOC 07-002: VSTL Work with Manufacturers Outside of Voting System Certification Engagements, VSTLs are NOT to participate in work outside of a federal certification test effort with the manufacturer.
2. Further we believe the class structure currently identified in the VVSG is confusing and adds unnecessary repetition of requirements. An example is the classification of optical scanners. The VVSG identifies separate classes for a scanner reading a manually marked ballot and machine marked ballot. The scanner isn't different, it's the ballot that is different. The class structure is adding false complexity and as such we believe using it is impractical. We would suggest Part 1 and 2 of the VVSG should be structured as a compliance matrix, with identifiers that make it visually straightforward to recognize if a functional requirement applies to a particular type of voting system. In the instance of optional functionality there is also a visual clue, such as shading.

   By using a compliance matrix in the functional requirements guideline, it would be feasible to add a new innovation by either inserting a row with a new requirement or a column with a new type of voting system. The actual implementation in the standard might not use a matrix per se but symbols next to each voting system type for required, optional, and not applicable requirements. Each part would have an appendix with the matrix that goes into the report. As currently written, the burden on the test labs is to create this matrix (each VSTL must use its own judgment/interpretation), go through the NVLAP audit process with individual matrices for evaluation (burden on NVLAP to assure consistency of VSTL test process), and then train their test staff. This is inefficient and adds unnecessary cost to the process.

VVSG #	Req. Description	DRE	Ballot Marker	Pct Scan	Central Scan	Add New
1.0	Ballot Preparation					
2.0	Precinct Functions					
3.0	Accessibility					
4.0	VVPAT Functions					
5.0	Reporting Functions					

3. Is Open Ended Vulnerability Testing (OEVT), as presented in the proposed guidelines, feasible in a conformance assessment process? What advantages or

disadvantages do you see with OEVT? If the EAC were to require OEVT how could it best be included into the EAC's Testing and Certification Program?

iBeta response: iBeta applauds that the EAC is issuing an RFI and then an RFP for generating the threat model.

The OEVT, as presented in the proposed guidelines, is not a repeatable test (Part 3 Section 5.4: "Open-ended vulnerability testing (OEVT) is conducted without the confines of a pre-determined test suite. It instead relies heavily on the experience and expertise of the OEVT Team Members, their knowledge of the system, its component devices and associated vulnerabilities, and their ability to exploit those vulnerabilities.")

As currently defined, the OEVT requirements do not allow the VSTL to comply with requirement 4.13.3 of the NIST Handbook 150-22 "The Certification Test Report plus the laboratory's records of the certification test shall contain sufficient information to allow repeating, reproducing and/or auditing the entire certification test." In order to allow the VSTL's to accomplish an OEVT, the handbook would need to be modified.

A modification to the handbook may not be prudent as the policy currently in-place has merit. The OEVT, being ad hoc, relying on the skills of individual test team members, and not repeatable, goes against the CMMI Maturity Level 2 (Process Repeatable) and drives the maturity level to 1 – the starting point for use of a new process which depends on ad hoc processes and individual effort.

The cost/benefit of an ad hoc test method based on the non-repeatable knowledge and focus of a selected staff and varying staff needs to be determined. If OEVT is incorporated, perhaps a consortium of VSTL security experts could be deployed to provide consistency of test process to all voting systems submitted for federal certification.

4. How could the processes of the VVSG be modified to incorporate minor revisions without incurring the costs (time and money) of a total system test, and still maintain the integrity of the standard?

iBeta response: The VVSG provides the requirements for what every voting system must do. We believe that incorporation of minor revisions should not be addressed in the VVSG. The methodology of how changing systems will be verified and validated to meet the requirements of the VVSG is a policy decision that should be contained in the EAC lab manuals, where these policies are currently identified.

5. What are the implications of the proposed usability benchmarks to you as a Voting System Test Laboratory? What are your current capabilities to test using human subjects?

iBeta response: It is our understanding that the usability benchmarks are a means of assessing compliance with section 301 of HAVA. In reading this section of HAVA we

don't see this type of testing actually being required. From a test perspective our interpretation is the requirements of HAVA are functional.

Regardless, it is difficult to assess the implications of the proposed usability benchmarks because so little information is provided in the guidelines.

1) Are these benchmarks valid? They are based on "preliminary" research and called "tentative" values. The caveats expressed in the guidelines lead one to be concerned
    a. How can there be one number for paper and DREs when studies of voters casting ballots as intended show different rates?
    b. Was the same voter and ballot used on various systems? Were there adjustments for voter familiarity with the ballot?
2) Assuming these are valid benchmarks, in order for the test to be valid the test conditions need to be controlled so that they are essentially the same as those that were run for the benchmark test. If you are running a performance benchmark for a software upgrade you would run the test in the old configuration to set the benchmark. After you upgraded the software you would then run the exact same test to judge the performance of the upgrade. Usually the upgrade will pass if it achieves equal or better performance. The guideline provided no information on the test conditions necessary to implement such an approach.
    a. Ballot design:
        i. What type of election is this? If it was a primary was it closed or open?
        ii. How many ballot designs were there?
        iii. What voting variations were on the ballot? Were there complex variations? Was there a long proposition?
        iv. How many contests were contained on the ballot? What were the "vote for" options?
    b. Ballot type:
        i. Were these visual ballots and/or audio ballots?
        ii. Were these in English only or other languages? If there were other languages, what were they?
        iii. What part of the input was manual and non-manual?
    c. Voter profile:
        i. What was the size of the test group?
        ii. What factors need to be taken into account when making up the test group: age, gender, education, literacy, or English proficiency? Are there other factors?
        iii. What part of the group was made up of voters with disabilities? What were the levels of disability?
        iv. What was the voter's level of experience with the voting system?
    d. Jurisdiction profile:
        i. In assessing a system as better or poorer what factors of the jurisdiction are being taken into account? Per the Voting Rights Act, multi-lingual ballot capabilities are required in 60% of the states. Only 10-12% require support of the iconographic languages

3) The benchmark is predicated on the voter making their intended choices, but there is no information on how you capture the voter's intent in order to assess if this was met.

While we question that this is even appropriate testing to meet the requirements of HAVA, as a test lab we have the capabilities to run this type of testing but in the absence of sufficient test requirements we cannot assess our current readiness.

6. Are there any changes to the VVSG, in either scope or depth, which would significantly reduce the cost (time and/or expense) of compliance without adversely affecting the integrity of the VVSG or the systems that are derived from its implementation?

iBeta response: It is difficult to argue against excess in the testing realm as you can appear to be the enemy of quality when, in fact, the guidelines should demand more stringent manufacturer development and test processes focusing on defect prevention rather than end of life-cycle (certification testing) detection.

Having said that, the workmanship requirements (Part 1 Section 6.4) are forcing the manual source code review of millions of lines of code through static analysis but not providing for the development of reliable and maintainable code. This labor intense, manual review cost needs to be assessed against the little or no appreciable technical advantage gained during that static test.

The use of software static analysis tools to eliminate dead code, flag vulnerable constructs, calculate complexity, and depth of inheritance structure could, with the corresponding functional requirements in the guidelines, provide for improved quality of the code source at a relatively small cost.

The large number of design requirements within the guidelines also limits the manufacturer's ability to use development practices. With IBM's purchase of Rational, Agile development has gone mainstream but one of the major principles of Agile development is that working software is delivered frequently (weeks rather than months) which is not realistic given the certification process. In ten years, if the methods of Agile were to be dominate family of development processes, this VVSG, as currently written, would not allow for voting manufacturers and VSTLs to tap into that labor pool or keep our employees current. Although not the intent, the design requirements may be driving this voting industry to antiquity.

We would suggest some efficiencies could be gained by incorporating standard test methods, standard test plan format, standard test report format and a standard conformance matrix in the VVSG. This would provide an improvement in usability of the VVSG and facilitate testing, tester training, reporting and EAC/state/jurisdiction review of report. Standardization of all of these elements could in their way contribute to reducing the cost of testing.