

Toward a More Effective and Efficient Boulder County LAT

In the round of Logic and Accuracy Tests testing associated with the Nov 2, 2004 general election, the three LAT executions required 20, 9 and 10 hours respectively to execute. Working through the LAT was inefficient and unsatisfying. We could have had better, more complete results earlier.

These notes are intended to begin a dialog aimed at improving the LAT. Having been through the process several times during the 2004 primary and general election, it seems like the following areas, if improved, would significantly increase the effectiveness and efficiency of the LAT:

- Little information was made available before the LAT about how it would be conducted and as a result testers disagreed about its purpose and scope, leading to several hours of clarification and delay. Prior release of the detailed LAT process¹ (expanding on a memo like this) should help.
- Resolution boards disagreed about what constituted a vote or voter intent, leading to several more hours of clarification. The actions of human resolution boards need to be efficiently accounted for or standardized during the LAT.
- The test strategy required the scan and resolution of all ballots on all stations. By the end of the third LAT and the 24th scan of every ballot, the test deck was worn out and some ballots were torn. Besides requiring an enormous amount of time, this amount of scanning and resolving fatigued the test deck, introducing another variable of uncertainty in the LAT results.
- When discrepancies between hand tallies and machine counts arose, it was difficult to find out what happened. A per-ballot report is necessary to understand discrepancies.
- Important vote-counting functions were not included in the LAT. These functions primarily reside on the Tally Station and include duplicate ballot detection and accumulation across vote-counting stations.

A restructuring of the LAT can address major parts of these issues. It can be shortened, be made more efficient and cover more functionality. What follows are some ideas that might help to get us there.

Proposed Changes to the LAT

The large number of scans and resolves for each ballot is the primary reason that the test takes so long. The current LAT scans/resolves each ballot on all eight vote-counting stations for each of the LAT1, LAT2 and LAT3 executions². We feel that we can prudently eliminate, say, 75% of these redundant scans and still have confidence in the integrity of the LAT – i.e., that it confirms that the vote-counting equipment is ready to accurately count votes.

We propose that the test deck be scanned/resolved just twice³ for each execution of the LAT – once on each of two groups of vote-counting stations. The CVRs from each group are then tallied and the tallies from the

¹ In this memo the LAT is a test that confirms that the Boulder County vote-counting system is ready and that it will accurately count votes. We recognize that there are differences of opinion on what should be included in the LAT. Here we focus on a test that is limited in scope to the vote-counting hardware, its associated application software and the Election Database as constructed by Boulder County elections managers. The surround system (i.e., ballot construction, ballot security, absentee vote processing, damaged ballot processing, resolution team training and operation, etc.) is not considered part of the test.

² The test is conducted three times – before the counting of votes (LAT1), on election day (LAT2) and after vote counting (LAT3).

³ There are, of course, many choices that would reduce the number of scans. For example, the test deck could be scanned and resolved just once during each of the LATs – that is, divided up into eight sub-decks and processed individually on the eight machines. This would cut even more time off of the LAT and have little impact on test quality. We think either approach, or a number of other approaches, would be appropriate. However, in this memo the procedures we describe use the two groups of machines as the example.

two groups compared. Use of Tally Station functionality in this way is not included in the current test. Our proposal produces a shorter, more complete test of the equipment.

There are a few other improvements that can be inserted. They are described in the detailed process section below. But the main flow could look something like the following:

- If a per-ballot report isn't available, the test deck is constrained to contain at most one ballot from each precinct⁴. Exceptions could be made when there is more than one ballot style for a given precinct, in which case testers should be informed that the per-precinct report won't be able to give per-ballot results for that precinct.
- Partisan teams of voters prepare a test deck of ballots and create a per-ballot hand tally. The teams divide their test ballots into four groups⁵ – Early Voting, Absentee and two groups of Election Day ballots. A duplicate ballot is artificially generated and inserted into one of the ballot groups.
- The eight vote-counting stations are divided into two 4-station groups and the stations within a group are assigned a role that matches one of the ballot groups – EV, AB or ED. The test ballot groups are then scanned and resolved on the respective machines in group 1. The same is then done for the machines in group 2. A specific ballot will be scanned twice in each of the three LAT executions, for a total of six scans. By changing the roles of the various vote-counting stations in LAT1, LAT2 and LAT3, we can ensure that these six scans are performed on six different stations.
- Discrepancies in tallies between groups of machines or between executions of the LAT need to be understood. That is, testers need to be able to look at a tally and understand how the equipment handled auto-resolves and how the human resolution board resolved the rest. A per-ballot report that breaks out the tally by these two components is essential for this level of analysis. In addition, it would help us evaluate the training and consistency of the Resolution Teams. Until such a report is available we purpose to remove human resolution inconsistencies from the tests. That is, when the ballots are resolved, races that require human resolution are manually recorded so that they can be consistently and unambiguously scored the same way throughout the test. The purpose is to remove differences in human judgment so that discrepancies found in the tallies can be understood using the existing reports and audit logs.
- The CVRs from the four machines in each group are tallied and a standard Canvass Report and a per-ballot report⁶ generated. The duplicate ballot detection utility is run and a report generated.
- The per-ballot scoring can then be compared to the test deck hand tally to verify correct scoring. The Canvass Reports from the two groups of machines can be compared to verify the scoring across machines.

By reducing the number of scans and resolves we are slightly lowering our confidence that the test will find a problem that only appears on a specific machine with a specific ballot. But, our intuition is that our confidence isn't lowered much, especially when the full LAT1, LAT2 and LAT3 sequence is considered. What we gain, however, is a substantial reduction in test execution time and substantial improvement in test coverage. With these changes, the test would include full Tally functionality – the ability to accumulate EV, AB and ED votes, the ability to accumulate across machines and the ability to detect duplicate ballots.

⁴ The restriction '*at most one ballot from each precinct*' is not desirable and should be removed once an adequate and timely per-ballot report is available. Without a per-ballot report, this restriction allows the standard Precinct Report or the standard Canvass Report or the standard Election Abstract to substitute for the per-ballot report. It appears that the HART SERVO® archive and reporting application might produce just the report that's needed, and if so it should be made a part of the procedure. If not, another per-ballot report should be provided by HART.

⁵ Other possible groups – Provisional, Emergency, Damaged.

⁶ As above, if a per-ballot report isn't available, the test deck must be restricted to have at most one ballot per precinct. Then the standard Canvass Report (or a similar standard report) is equivalent to per-ballot report.

Process Details

The LAT process cannot begin until the Election Database is completed. Once that is in place, the LAT can proceed as follows:

MBB Preparation

This procedure should be scheduled to run back-to-back with the Test Deck Preparation, below, to provide participants the option of verifying MBB preparation.

- Report out the version information for all election software so that testers can verify that certified software is being used.
- Clone the Election Day election database into two test databases, one for each of the two groups of vote-counting stations. This will ensure that the test will not contaminate the election tally and that the two test groups can be tallied separately.
- Modify the two test databases as necessary to ensure that ballots printed from them have the words “TEST BALLOT” printed on them.
- Create MBBs for each of the 4 machines in group 1 from one of the test databases. Then create MBBs for the remaining machines from the remaining test database.

Test Deck Preparation

As is the current procedure, partisan teams prepare portions of the test deck. The teams vote their ballots and prepare a per-ballot tally. The ballots and tallies are private and secured until use in LAT1.

- Using one of the MBBs created above, print the pool of test ballots – 25 per party or jurisdiction.
 - Ballot selection is constrained as follows:
 - o One ballot selected from each major ballot style.
 - o Additional ballots selected from random precincts.
 - o If a per-ballot report isn’t available, have at most one ballot per precinct except as necessary for ballot style coverage.
- Each party or jurisdiction is given a portion of the test ballot pool. The ballots are hand voted and a per-ballot tally created.
- The party or jurisdiction divides their ballots into four groups – EV, AB and two groups of ED ballots. The group that each ballot ends up in is recorded on the per-ballot hand tally.
- The ballots and hand tally are private and secured until the results of the first LAT are provided to the testers.
- When it’s time to start the LAT1 test execution, assemble all party and jurisdiction ballots into EV, AB and two ED groups. To allow for duplicate ballot detection, photo-copy one of the AB ballots and insert it in the EV group. Record which ballot this is.

Test Execution

In the current LAT, the ballots from each party or jurisdiction are first scanned, resolved and reported individually so that the hand tallies can be verified before a combined test deck is made. That step is no longer needed – the per-ballot report will allow the hand tallies to be viewed and justified at the end of the test.

The eight vote-counting stations are divided into two 4-station groups. By using two operator/resolution teams, the two groups of machines can be scanned and resolved simultaneously. This would require some coordination – ballots and resolution board answers would have to be exchanged – but it would speed up the overall test.

For each of the two groups:

- Verify that certified versions of the election software are being used.
- Run the zero report.
- Determine the role (EV, AB, ED, ED) for the four vote-counting stations in the group. Record the roles so that the roles used during LAT1, LAT2 and LAT3 differ.
- Scan the test deck groups into the respective machines.
- Resolve them, using a real operator and a Resolution Team. On the first resolve of a sub-group of ballots, the team will record all races that are resolved. Subsequent resolves⁷ will use those notes to ensure that human resolutions are consistently resolved throughout LAT1, LAT2 and LAT3.
- Move the CVRs via the MBBs to the Tally Station and tally them. Run the duplicate ballot detection utility and deal with the known duplicate. Accumulate the votes from all machines in the group and produce a single Canvass Report (or election abstract which shows the accumulated tally for each precinct) and a single per-ballot report for the collection of four machines.

Test Verification

For a faster, more accurate verification, soft copies of the reports and audit logs need to be available to the testers.

- Verify that the duplication report found the known duplicate ballot.
- Verify that the per-ballot report from one group of machines matches the per-ballot hand tallies created by the test deck teams. In some cases, the original ballots may need to be inspected to verify that the per-ballot hand tallies were done correctly.
- Verify that the Canvass Reports across the two groups of machines agree.
- Investigate discrepancies using the audit logs on the pertinent scanning machines.
- Provide all these reports to the testing teams on a CD-ROM or other computer media to allow them to prepare for the data processing needed during the canvas and to allow for for more efficient and inclusive testing.

Expected Cost

Some of the above steps are more costly than in the current process. For example, the creation of reports at test conclusion (the duplication report and the canvass report) is new. But these steps are not brand new, just new to the LAT. Nonetheless, these steps will require more time to execute.

But we expect the test execution savings and the analysis of discrepancies savings to be significant. Once the test is prepared and the test deck created, we see no reason that the test cannot be executed and analyzed in 3 hours.

Don Hayden
Neal McBurnett
November 30, 2004

⁷ The record and playback of the Resolution Team actions isn't needed if a per-ballot report that breaks out machine auto-resolves and human resolutions is available.