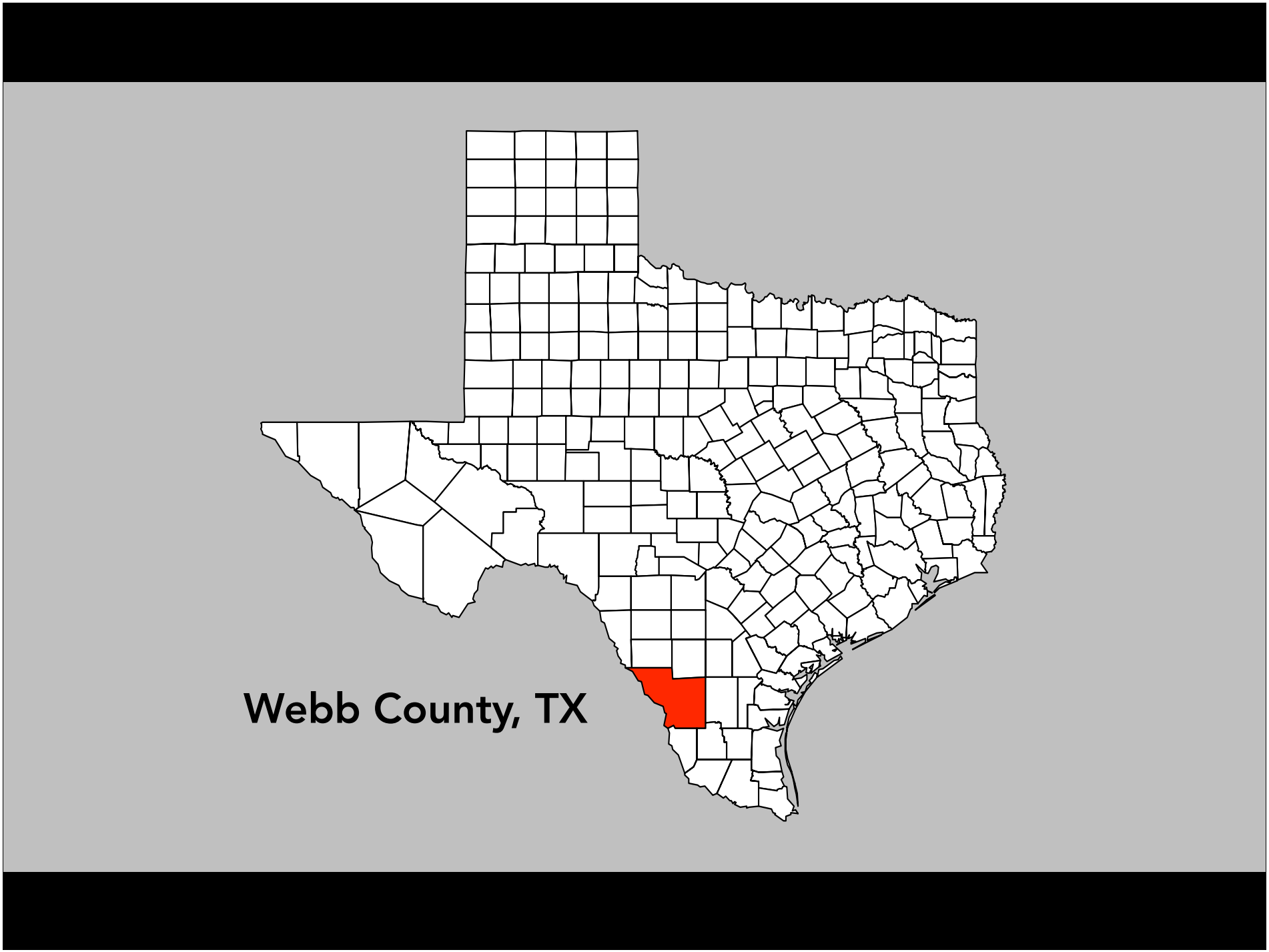
Adventures in e-voting research

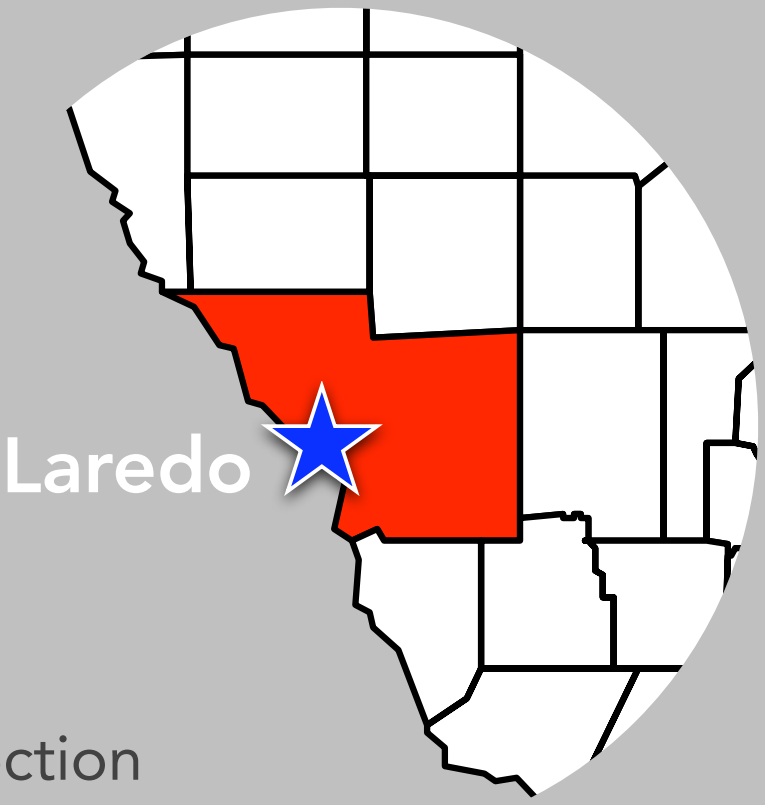
Dan S. Wallach
Department of Computer Science
Rice University
(Joint work with Dan Sandler)

What I Did Over Spring Break

(March 2006)

by Dan S. + Dan W.


Webb County, TX



March 7, 2006:

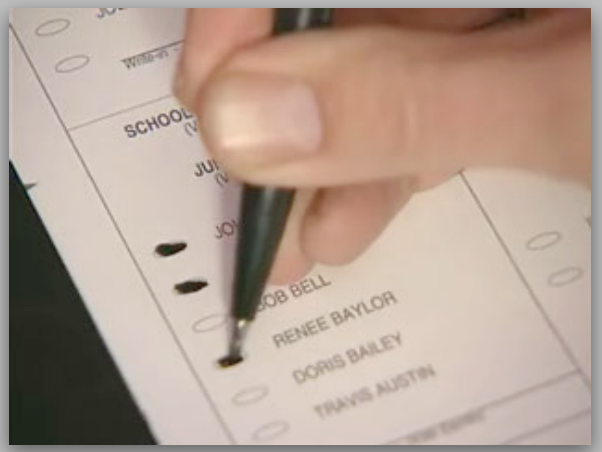
2006 Democratic primary election

(County's first use of DREs)

An unusual situation

Voters given a choice:

OR



DRE

(ES&S iVotronic)

Paper

(central ES&S op-scan)

Flores v. Lopez

~50,000 votes cast

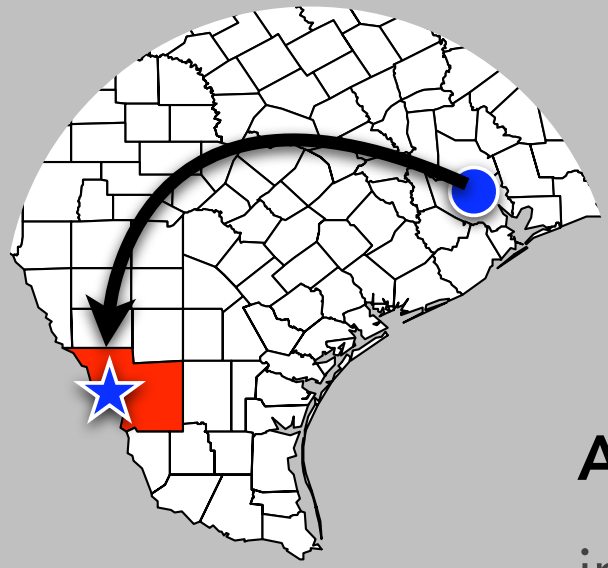
Margin of victory: ~100 votes

The loser suspected the DREs

…because he looked better on paper

Lawsuit

Bring in the experts!



April 12–13

initial investigation (Dan & Dan)

Webb Co. data

Raw binary data from Compact Flash cards

Opaque, undocumented format

Text output from DRE tabulator

For each machine:

- **"IMAGELOG.TXT"** (cast ballots)
- **"EVENTLOG.TXT"** (more on that later)

A smoking gun?

Evidence of evil DREs?

HACKS???

(how could we?)

What we (really) found

Anomalies in the event logs

Per-machine records

List of important election events

e.g. “terminal open,” “ballot cast,” …

Example event log

Votronic	PEB#	Type	Date	Time	Event
5140052	161061	SUP	03/07/2006	15:29:03	01 Terminal clear and test
	160980	SUP	03/07/2006	15:31:15	09 Terminal open
			03/07/2006	15:34:47	13 Print zero tape
			03/07/2006	15:36:36	13 Print zero tape
	160999	SUP	03/07/2006	15:56:50	20 Normal ballot cast
			03/07/2006	16:47:12	20 Normal ballot cast
			03/07/2006	18:07:29	20 Normal ballot cast
			03/07/2006	18:17:03	20 Normal ballot cast
			03/07/2006	18:37:24	22 Super ballot cancel
			03/07/2006	18:41:18	20 Normal ballot cast
			03/07/2006	18:46:23	20 Normal ballot cast
	160980	SUP	03/07/2006	19:07:14	10 Terminal close

Problem #1
Logs starting mid-day

```
03/07/2006 15:29:03  Terminal clear and test
03/07/2006 15:31:15  Terminal open
```

Polls opened around **7 AM** across Webb Co.

What happened between 7 and 3:30?
Lost votes?

(10 total machines)

Problem #2

Election events on wrong day

Votronic	PEB#	Type	Date	Time	Event
5142523	161061	SUP	02/26/2006	19:07:05	01 Terminal clear and test
	161115	SUP	03/06/2006	06:57:23	09 Terminal open
			03/06/2006	07:01:47	13 Print zero tape
			03/06/2006	07:03:41	13 Print zero tape
	161109	SUP	03/06/2006	10:08:26	20 Normal ballot cast
					[... 9 more ballots cast ...]
	161115	SUP	03/06/2006	19:29:00	27 Override
			03/06/2006	19:29:00	10 Terminal close

The election was held on 03/07!

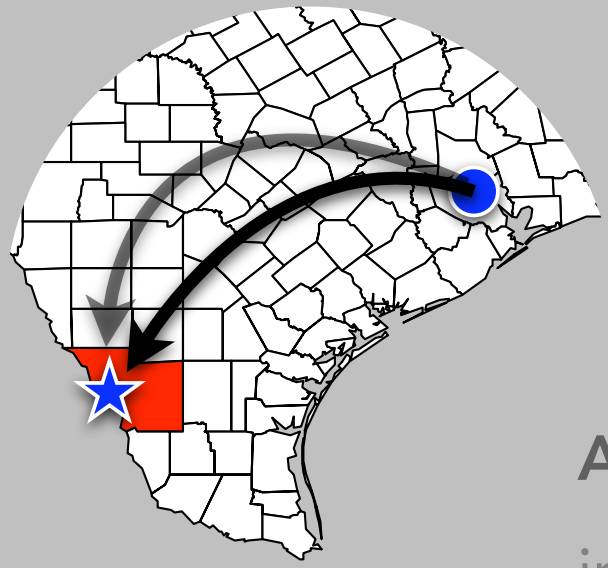
(4 machines / 41 votes)

Votronic	PEB#	Type	Date	Time	Event
5145172	161061	SUP	03/06/2006	15:04:09	01 Terminal clear and test
	161126	SUP	03/06/2006	15:19:34	09 Terminal open
	160973	SUP	03/06/2006	15:26:59	20 Normal ballot cast
			03/06/2006	15:30:39	20 Normal ballot cast
	161126	SUP	03/06/2006	15:38:37	27 Override
			03/06/2006	15:38:37	10 Terminal close

26 machines with exactly two ballots cast the day before

We learned that these were probably L&A test votes, erroneously **included in the tally**

(52 votes)



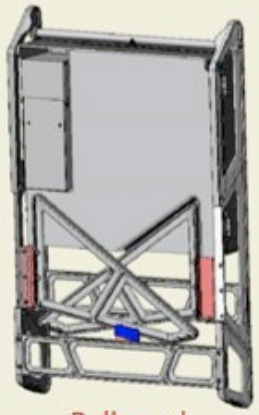
April 12–13

initial investigation (Dan & Dan)

April 24–25

follow-up trip (just Dan)

Booth Setup Sequence

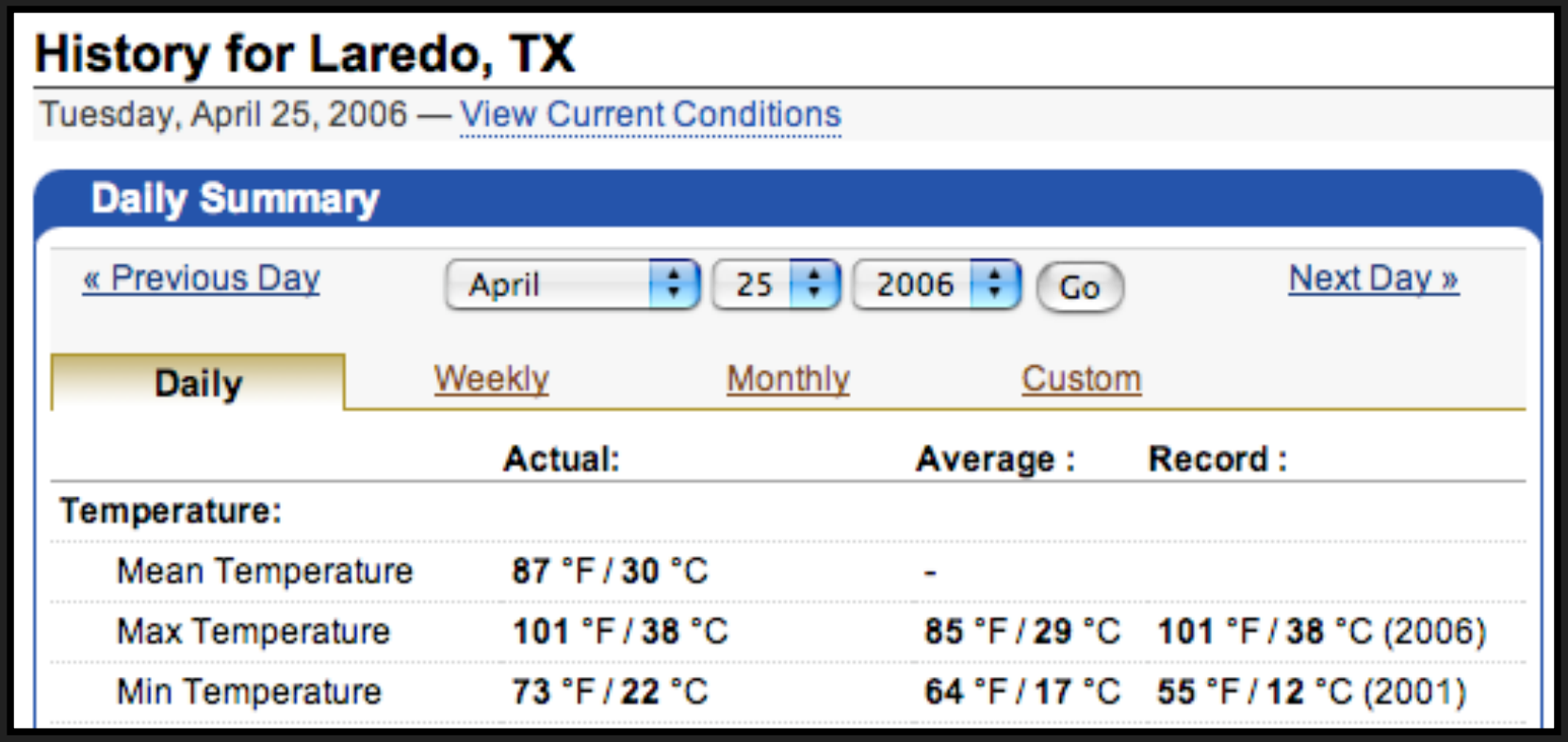


Delivered

Fold out legs

Pivot up platform
and lock upright

Unlock privacy
screens and add iVotronic



History for Laredo, TX

Tuesday, April 25, 2006 — View Current Conditions

Daily Summary

« Previous Day | April | 25 | 2006 | Go | Next Day »

Daily | Weekly | Monthly | Custom

	Actual:	Average :	Record :
Temperature:			
Mean Temperature	**87** °F / **30** °C	-	
Max Temperature	**101** °F / **38** °C	**85** °F / **29** °C	**101** °F / **38** °C (2006)
Min Temperature	**73** °F / **22** °C	**64** °F / **17** °C	**55** °F / **12** °C (2001)

source: wunderground.com

Machines containing only two votes

Everything appeared normal

Most likely L&A test votes

Others

Hardware clock set incorrectly

Just enough to account for anomaly

This is not proof of correct behavior!

Problem #3
Insufficient audit data

We couldn't collect data from every machine

Many were cleared after the election!

(Only the CF card "dumps" remain.)

Paper records missing

Zero tapes

Cancelled ballot logs

Observations

"Mistakes were made."

Violations of election procedures

Counting test votes in final results

Loss of zero tapes and other paper logs

Erasure of some machines

Local (mis)configuration

Hardware clocks set wrong

These things cast doubt on the results

Honest mistakes
or illegitimate votes?

No way to be sure.

Believable audits impossible.

These things happen in real elections.

Detailed report:

http://accurate-voting.org/wp-content/uploads/2006/09/webb-report2.pdf

Research goals

Make it easier to audit results after election day

Make it harder to make mistakes on election day

Prove

every vote tallied is valid

every valid vote is present

Tolerate

accidental loss/deletion of records

election-day machine failure

How?

Connect the machines together.

Benefits of the network

Store everything everywhere

Massive redundancy

Stop trusting DREs to keep their own audit data

Link all votes, events together

Create a secure timeline of election events

Tamper-evident proof of each vote's legitimacy

Auditorium

Ingredient: **hash chains**

```
                "Machine turned on" (HASH = 0x1234)
"Cast a vote after event 0x1234" (HASH = 0xABCD)
"Cast a vote after event 0xABCD" (HASH = 0xBEEF)
 "Turned off after event 0xBEEF" (HASH = 0x4242)
```

Every event includes the hash of a previous event ("hash chaining")

Result: **precedence** — "X must have happened after Y"

To alter or delete a single record,

you must alter every subsequent event as well

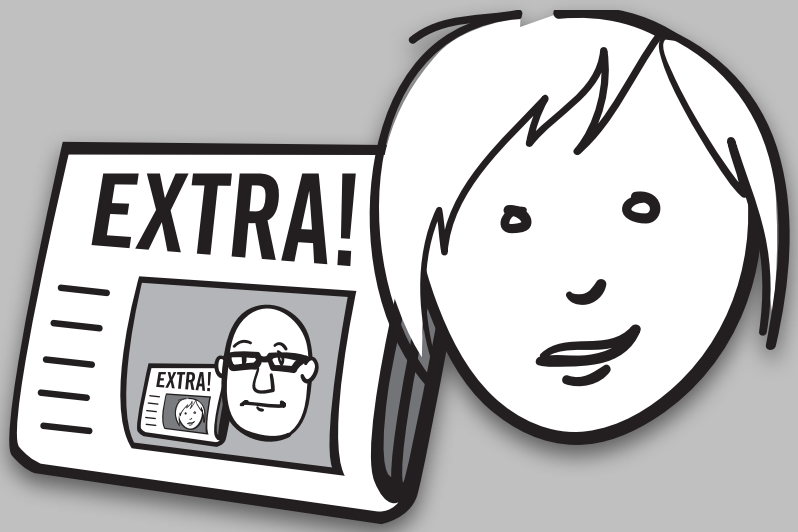
Ingredient #2: **timeline entanglement**

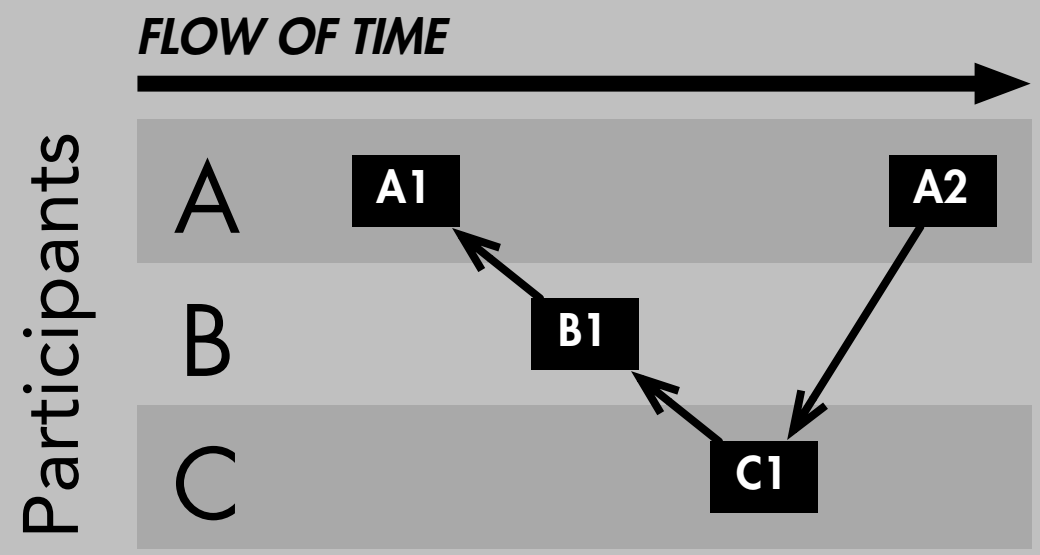
Entanglement = "chain with hashes from others"

Result: event precedence **between participants**

Malicious machines can't retroactively alter their logs

This would upset the global timeline!





B1 incorporates HASH(**A1**)
C1 incorporates HASH(**B1**)
A2 incorporates HASH(**C1**)

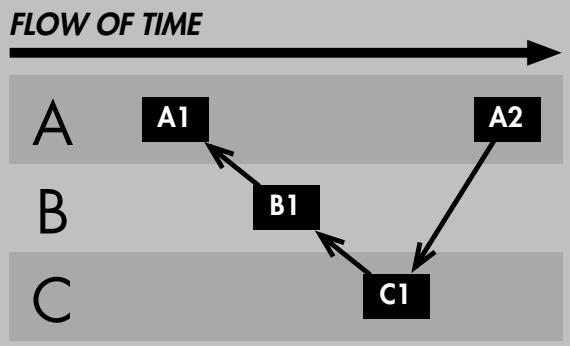
Ingredient #3: **Broadcast**

All-to-all communication

Allows entanglement

Widespread replication

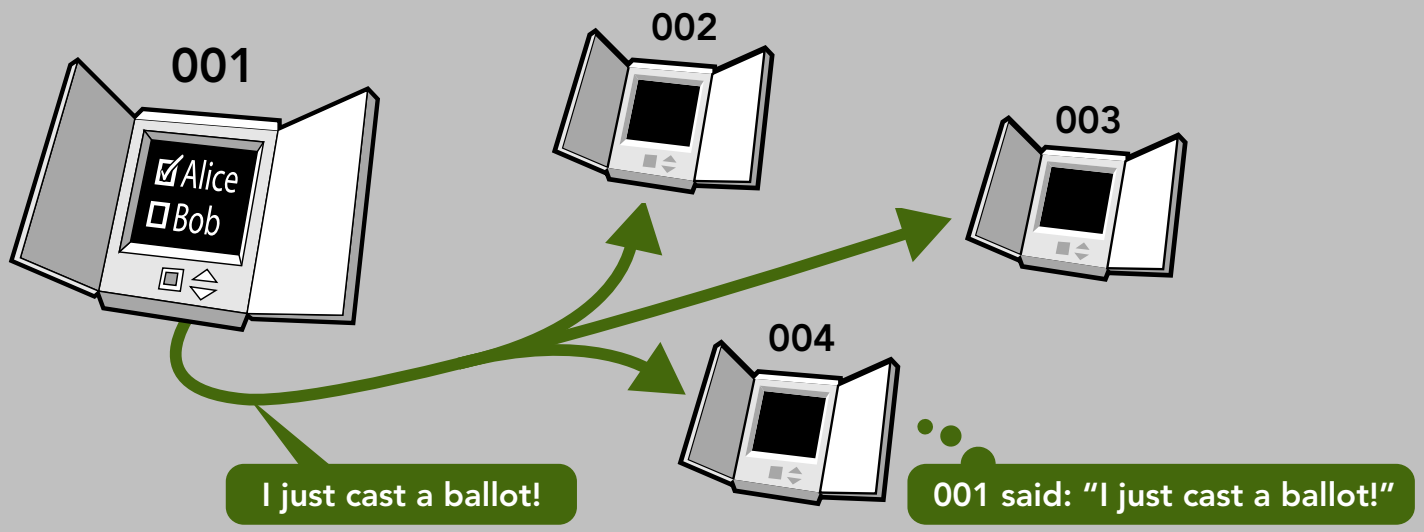
Broadcast + entanglement detail



Abstract view
(B1 succeeds A1)

Broadcast entanglement =

Auditorium



Everyone hears everything in the Auditorium.

The Papal Conclave

Proceedings **closed to outsiders**

All ballots cast **in plain view**

All ballots **secret**

The supervisor console

Shows status of all machines

Votes cast, battery running low, etc.

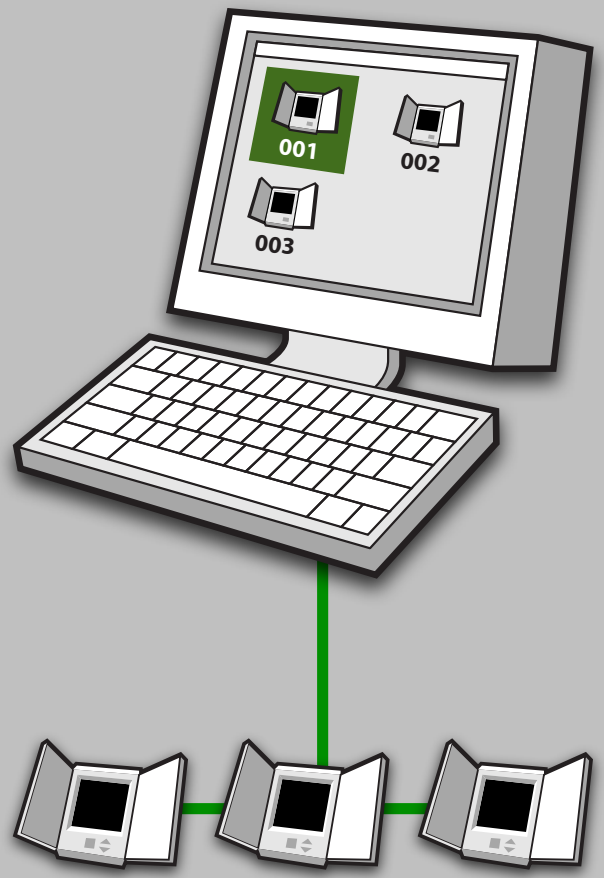
Helps conduct the election

Less opportunity for poll-worker error

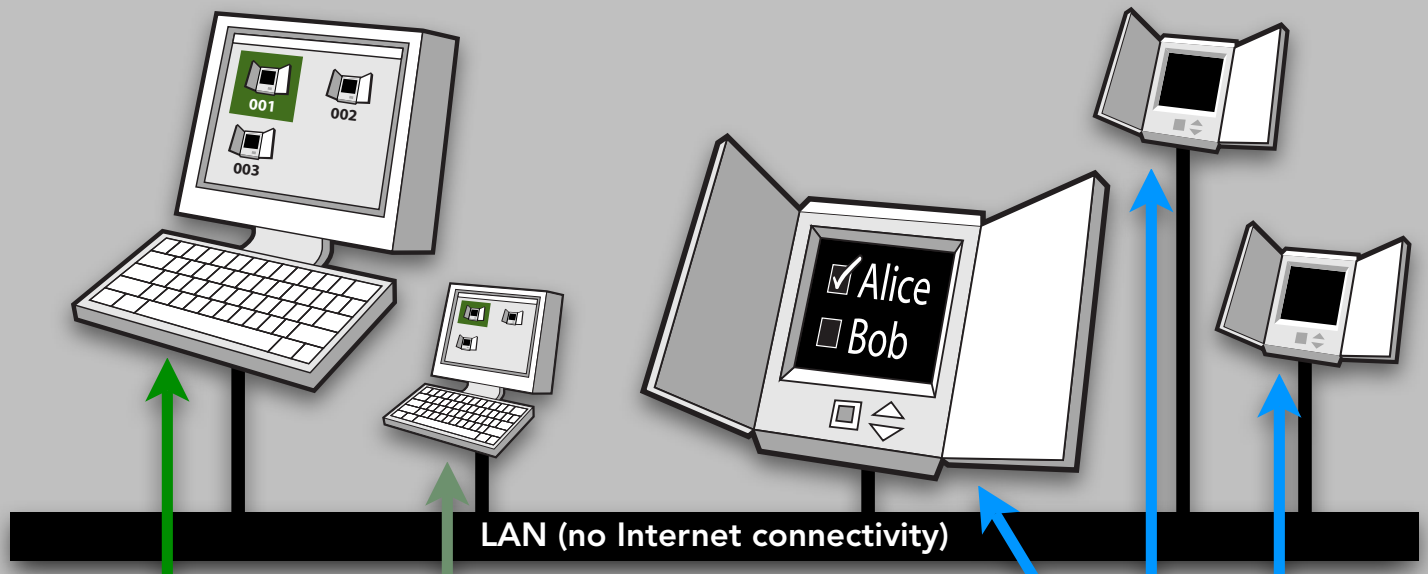
Ballots distributed over the network

Booths are stateless, interchangeable

Supervisor has a hot spare, too



Voting in the Auditorium



SUPERVISOR
Monitors, displays booth status
Broadcasts vote authorizations
Records all broadcast messages

SUPERVISOR (BACKUP)
Ready to assume supervisor's responsibilities at any time
Records all broadcast messages

BOOTHS
Listen for vote authorizations (ballots)
Capture voter selections
Broadcast encrypted votes
Record all broadcast messages

Failure scenarios

"Attacks"

Attacks (1)

Early machine exit (*e.g.* equipment failure)

Votes safely replicated on other machines

Votes provably legit (authorized by supervisor, etc)

Late machine entry

Cleared? *See above.*

Hot spare? *Logs prove the machine hasn't been used.*

Attacks (2)

Ballots cast on the wrong day

Clock set wrong? *Hash chain OK; votes legit*

Test votes? *No hash chain connection to poll opening.*

Intentional subversion

Stuffed ballots? *Like test votes: invalid.*

Removed ballots? *Provably missing from hash chain.*

Mega attacks

Switched results

Scenario

Malicious parties in control of precinct

Day before election: attackers conduct a **secret election**

Swap those results for the election day results

Secret election could also be *post facto*

Countermeasures?

Election start nonce (“launch code”) — added to `(polls-open)`

Quickly publish hash of final `(polls-closed)` event

Concurrent shadow election

Scenario

Malicious parties create duplicate precinct

On election day, conduct secret election using appropriate start nonce

Countermeasures?

TPM to resist duplication of booth key material (signed by high-ranking election officials)

Booth capture

Scenario

Armed attackers take control of the polling place by force and stuff ballots—or destroy them—until the police arrive

Detection

Trivial

Countermeasures?

Partial destruction is recoverable from intact machines

Software tampering

Scenario

Malicious software

Introduced by poll workers, voters, “field upgrades”

Countermeasures

TPM / “trusted computing” technologies

End-to-end verifiable cryptography (more on this later)

Conclusion

In real elections…

Mistakes are made

Data is lost

Auditing is…challenging

Auditorium

an *auditable* record of election day

All election events linked in a secure timeline

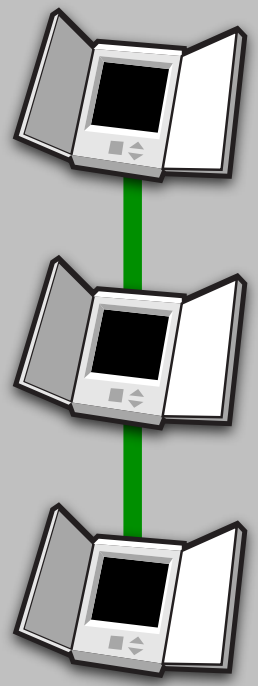
No ambiguity about when a vote was cast

Entanglement + broadcast = recoverability

A lost machine's votes are safe and believable

Composable with other secure e-voting ideas

software independence, secure vote storage, trusted computing



Don't fear the (air-gapped) **network!**

Future work (VoteBox++)

Homomorphic encryption of ballots

Safe for election observers to see in real time

Broadcast to the Internet!

Voters cryptographically challenge the voting machine

Adaptation of Benaloh technique [EVT '07]

Real-time Auditorium message validation at large scale

General-purpose first-order logical predicate evaluation

thanks

VoteBox team

Kyle Derr, Corey Shaw, Ted Torous

Rice Computer-Human Interaction Lab

Mike Byrne, Sarah Everett, Kristen Greene

NSF/ACCURATE

fin

See our Electronic Voting Technology '07 paper:

http://accurate-voting.org/wp-content/uploads/2007/08/evt07-sandler.pdf

VoteBox

The VoteBox platform

Testbed for e-voting research & experimentation

Auditorium drives all election events

distribute ballots, collect votes, open/close polls

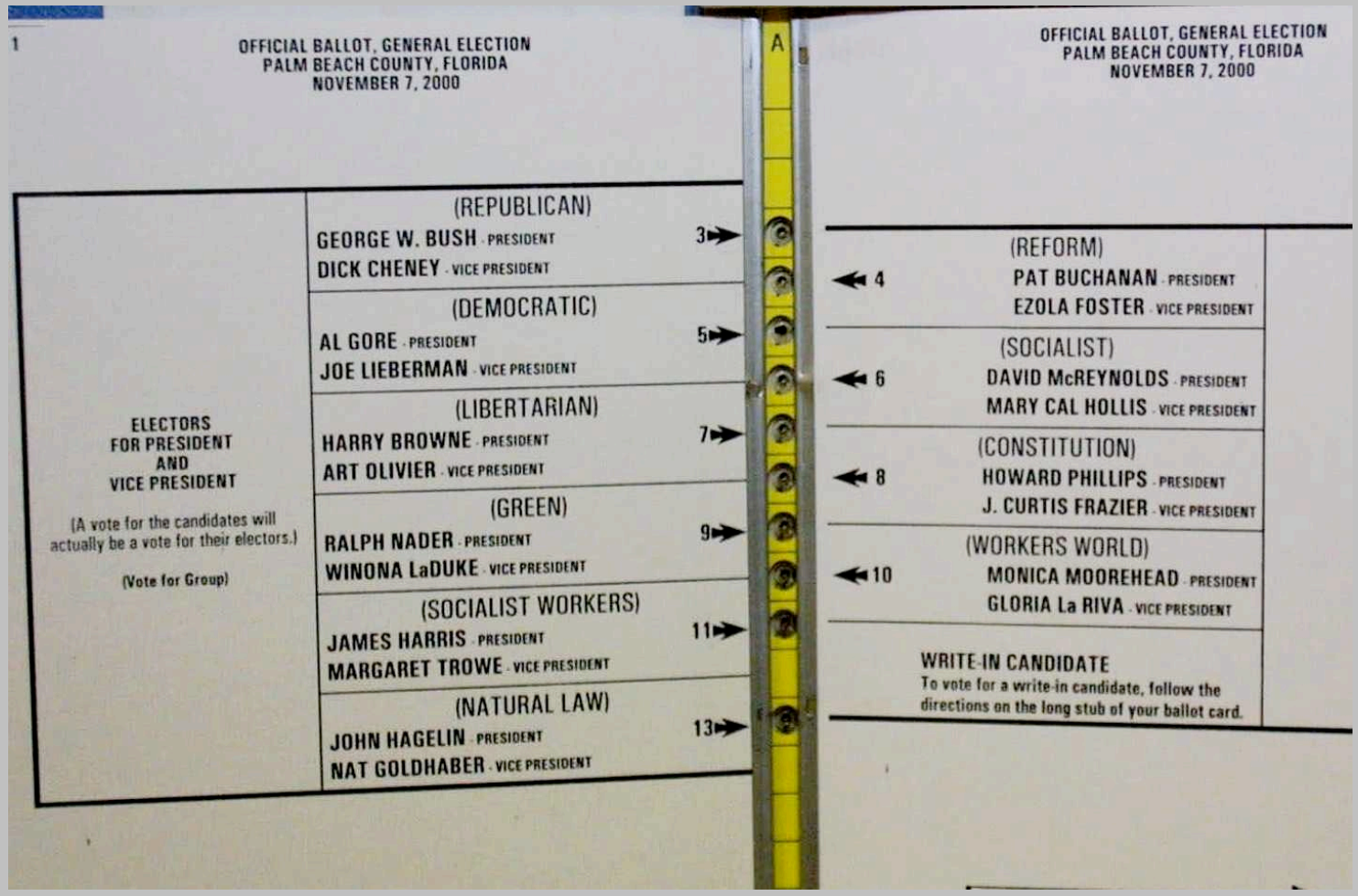
Pre-rendered user interfaces [Yee06]

smaller trusted software stack

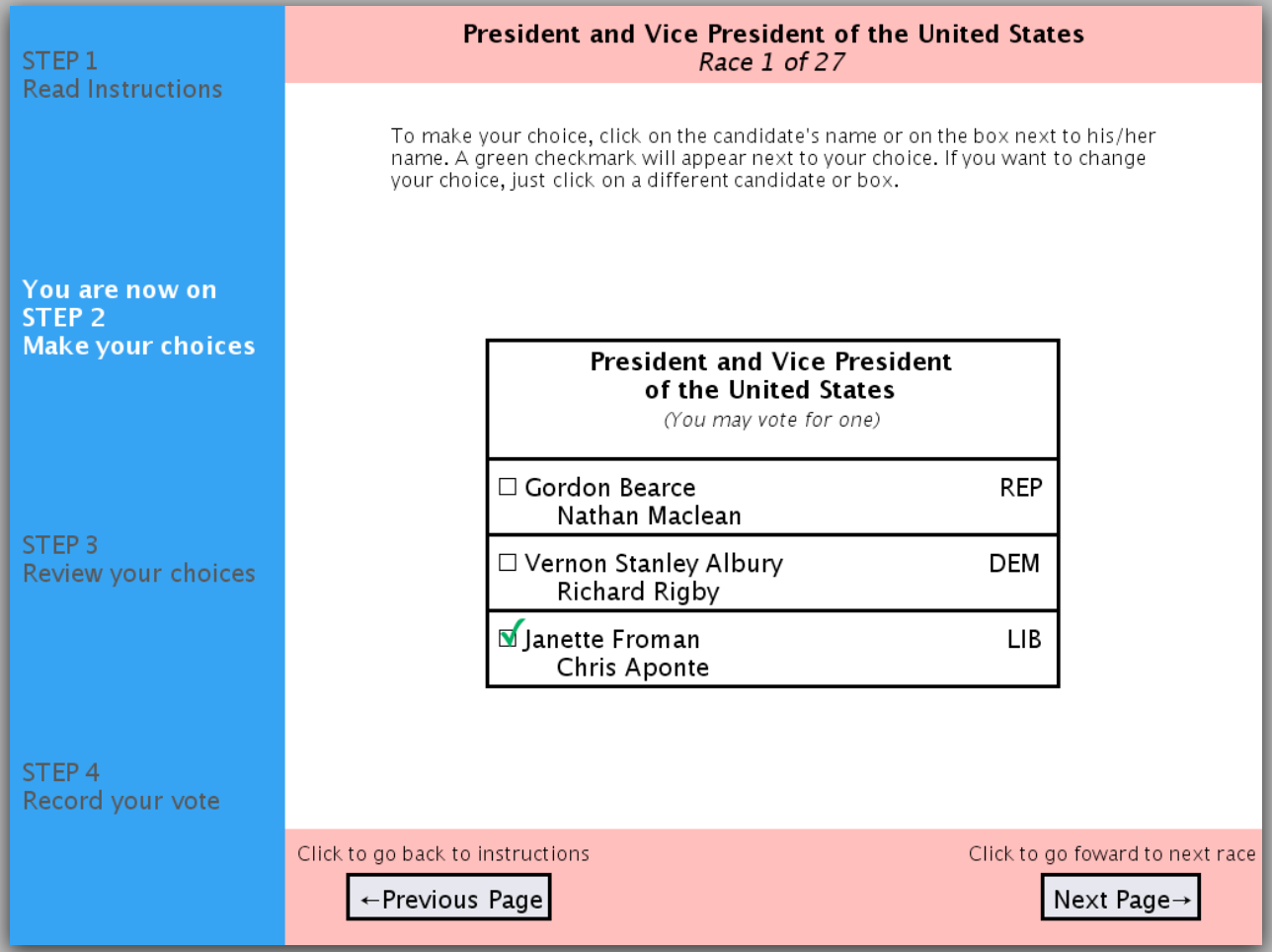
Flexible UI + ballot preparation tools

human factors research [Everett07]

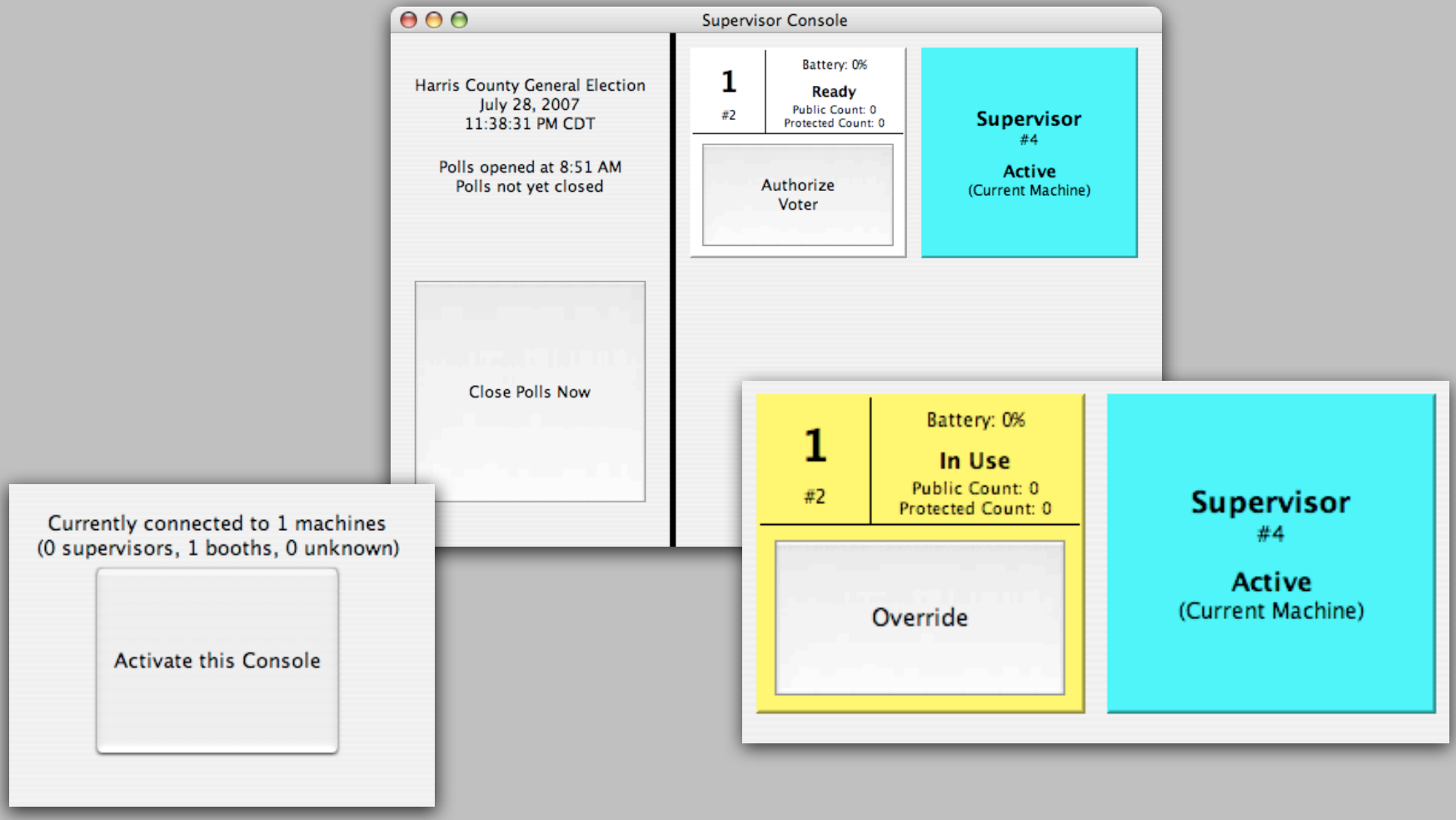
Human factors matters



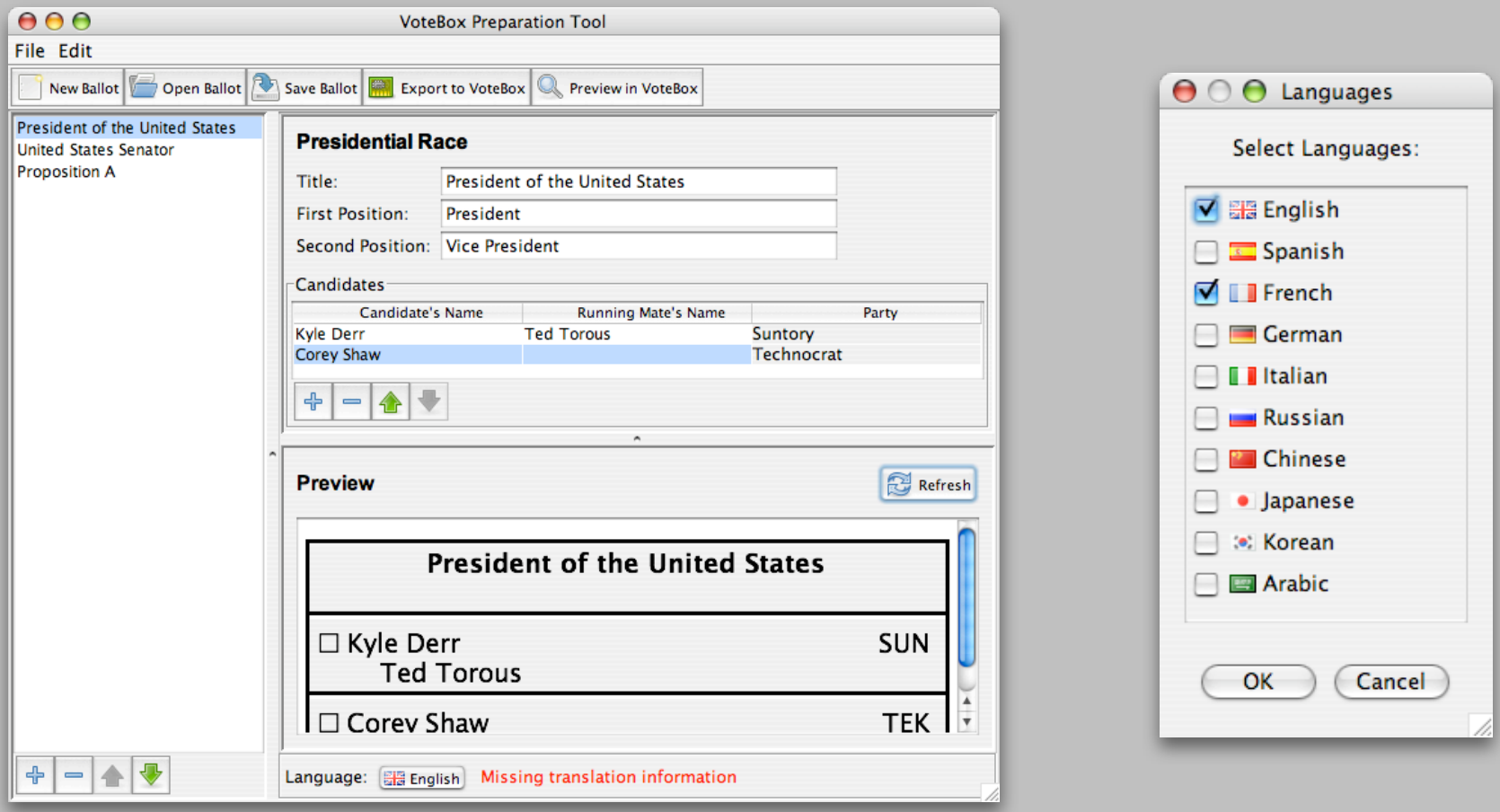
VoteBox booth UI



VoteBox supervisor UI



VoteBox ballot preparation tool



An example election

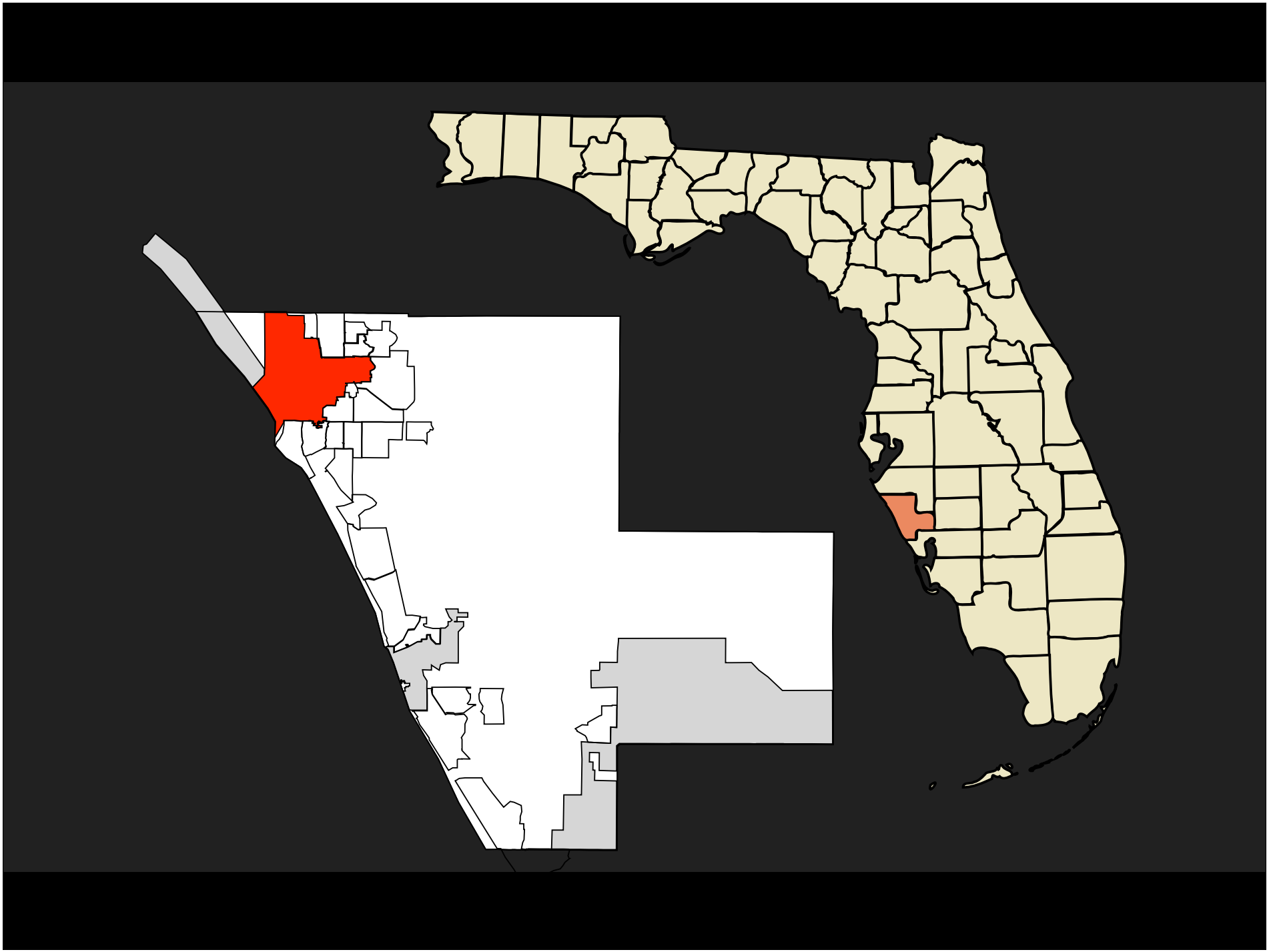
```
supervisor: (polls-open)
supervisor: (authorized-to-cast booth1 nonce1)
    booth1: (cast-ballot nonce1 <ciphertext>)
supervisor: (authorized-to-cast booth2 nonce2)
supervisor: (authorized-to-cast booth3 nonce3)
    booth3: (cast-ballot nonce3 <ciphertext>)
supervisor: (override-cancel-ballot nonce2)
    booth2: (override-cancel-confirm nonce2)
supervisor: (polls-closed)
```

Not shown: lower-level Auditorium data
(including hash chains)

Sarasota, Florida

CD-13 Race, November 2006

Christine Jennings v. Vern Buchanan



In a nutshell...

Did voting machines steal a Democratic victory?

In Katherine Harris' old Florida district, more than 18,000 votes went missing -- and a Republican won a House seat by 369 votes.

By Katharine Mieszkowski

The recount is over in the 13th Congressional District in Florida. The lawyers have won -- and the Democrat has lost. As in the presidential election of 2000, that loss appears to have been caused by a glitch in the voting process. But this time, the controversy centers on the very electronic voting machines many counties around the country purchased after the 2000 election in hopes of avoiding the sort of debacle that produced Bush v. Gore.

On Monday, Florida election officials named Republican Vern Buchanan the victor in the race for the House seat that Katherine Harris -- the Katherine Harris who was Florida's secretary of state during the 2000 recount -- vacated to run for the Senate. The Florida Elections Canvassing Commission, which is made up of Gov. Jeb Bush and two other elected Republican officials, said that the results of the recount showed Buchanan had beaten Democrat Christine Jennings by 369 votes in a race where nearly 240,000 votes were cast. The commission awarded the victory to Buchanan despite the fact that the mystery of more than 18,000 missing votes has not been resolved.

Photo: AP/J. Scott Applewhite

Christine Jennings, the Democratic candidate in Florida's unresolved 13th Congressional District, second from left, after posing with freshman members of the House for a group photo on the steps of the Capitol in Washington on Nov. 14, 2006.

Results (Sarasota County)

	Total votes	%	Election Day	Early Voting	Absentee	Provisional
Vern Buchanan	58,632	47.24	36,619	10,890	11,065	58
Christine Jennings	65,487	52.76	39,930	14,509	10,981	67
Overvotes	1		0	0	1	0
Undervotes	18,412		12,378	5,433	566	35

Undervote rates by race

U.S. Senate	1.14%
Congress	12.90%
Governor	1.28%
Atty General	4.36%
C.F.O.	4.43%

Absentee	2.5%
ES&S iVotronic	14.9%

Theory #1: Rational abstention

Nobody seriously believes this.

Theory #2: Human factors

Were voters confused by the ballot design?

OFFICIAL GENERAL ELECTION BALLOT
SARASOTA COUNTY, FLORIDA
NOVEMBER 7, 2006

CONGRESSIONAL

UNITED STATES SENATOR
(Vote for One)

Katherine Harris	REP	☐
Bill Nelson	DEM	☐
Floyd Ray Frazier	NPA	☐
Belinda Noah	NPA	☐
Brian Moore	NPA	☐
Roy Tanner	NPA	☐
Write-In		☐

U.S. REPRESENTATIVE IN CONGRESS
13TH CONGRESSIONAL DISTRICT
(Vote for One)

Vern Buchanan	REP	☐
Christine Jennings	DEM	☐

STATE

GOVERNOR AND LIEUTENANT GOVERNOR
(Vote for One)

Charlie Crist Jeff Kottkamp	REP	☐
Jim Davis Daryl L. Jones	DEM	☐
Max Linn Tom Macklin	REF	☐
Richard Paul Dembinsky Dr. Joe Smith	NPA	☐
John Wayne Smith James J. Kearney	NPA	☐
Karl C.C. Behm Carol Castagnero	NPA	☐
Write-In		☐

Theory #3: Machine malfunction

Did engineering failures of the machines *induce* the underotes?

Did voters *see* their undervotes on the summary screen?

Poor touchscreen calibration

Poor touch sensitivity

Hardware and software failures

Manufacturing defects

Dan Rather Reports had a long piece on this issue

Angle of view to the screen

Theory #4:
Fraud!

No evidence to support this.

Exceptionally difficult to prove.

Never ascribe malice to what can adequately be explained by incompetence. – Napoleon

Machine vs. human error

Critical concept relative to Florida law

If the summary screen showed “Jennings” and the machine recorded “none”, then Jennings would win

Regardless, the machines failed to capture voter intent

Experts on both sides agree Jennings would have won

State investigations

"Recount"

Same results as before (largely meaningless)

"Parallel" election tests

Poorly conducted, inconclusive results

Software examination

Found nothing (except unrelated security holes)

Never looked at the hardware (big part of the CA effort)

What happened?

State lawsuits

Judge denied plaintiff's discovery motion

Congressional Committee on House Administration

Investigation ongoing (GAO)

Florida banned electronic voting systems

What's next?

One year later, we still don't know what happened

We need better recount / challenge procedures

Discovery is more important than vendor trade secrets

Jennings running for (re)election in 2008

More details in my report (with David Dill):

http://www.cs.rice.edu/~dwallach/pub/sarasota07.html

The California Top-To-Bottom Study

Summer 2007

Biggest study of its kind, ever

40+ researchers (source code, “red team,” documentation, accessibility)

three vendors (Diebold, Sequoia, Hart InterCivic)

http://www.sos.ca.gov/elections/elections_vsr.htm

Significant flaws found with each vendor

Viral attacks possible!

Diebold and Sequoia "conditionally recertified"

Only one machine per precinct for accessibility

Other votes on paper

Hart InterCivic has comparable sanctions

Revised conditions newly announced

(e.g., reboot SERVO from CDROM after every eSlate)

Hart eSlate architecture

Local network in the polling place

Controller sees all machines, collects all votes together

Cryptography?

HMAC-SHA1 for integrity checking of cast ballots

Single shared key for the entire election

OpenSSL in some places, but incorrect cert checking

No crypto on voting-machine local network

Network protocol?

Messages that directly read and write to memory

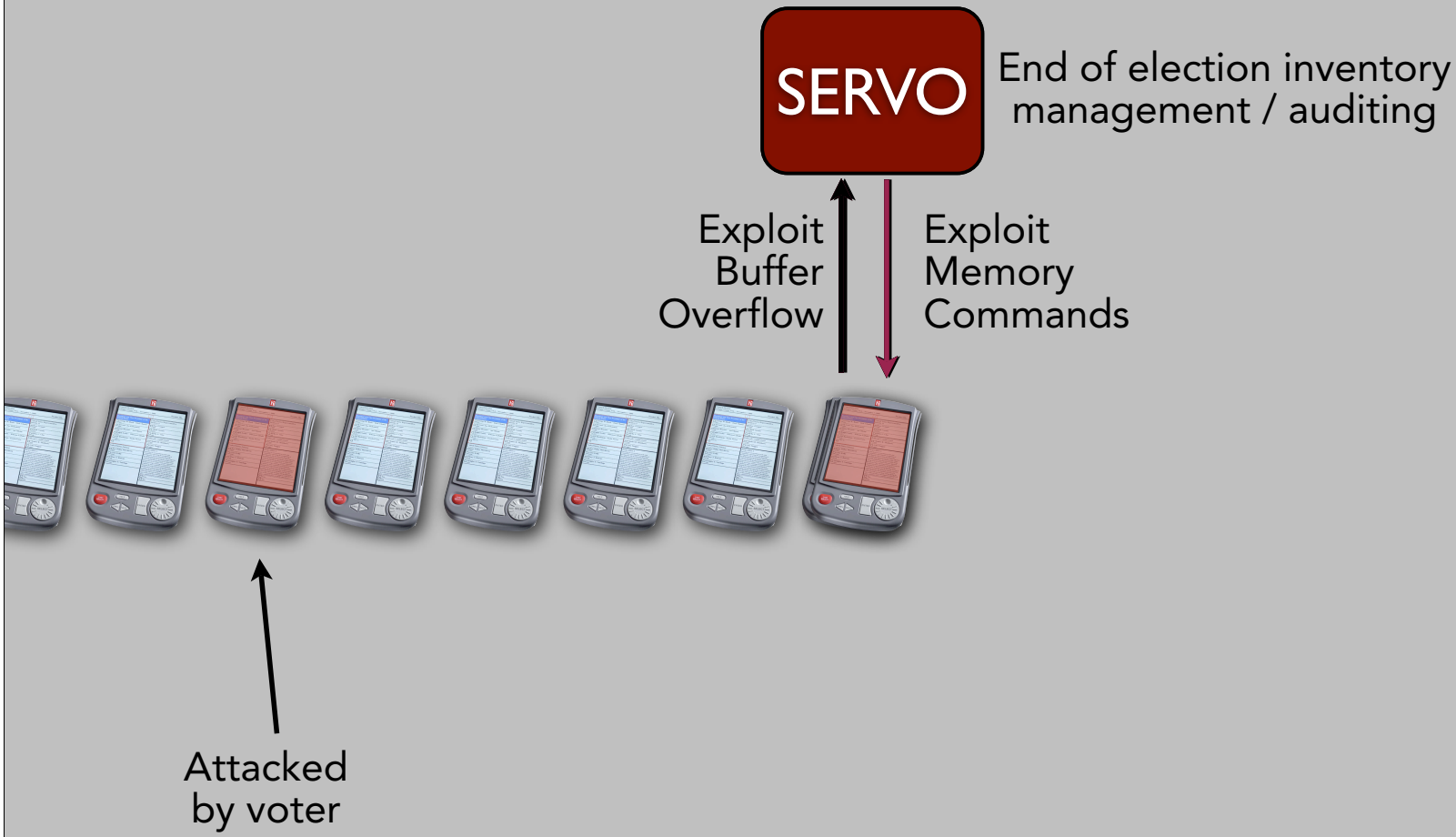
Officially used to test whether code is authentic

Also allows votes to be extracted or changed

Enables virus injection

Regular voters have access to the network port

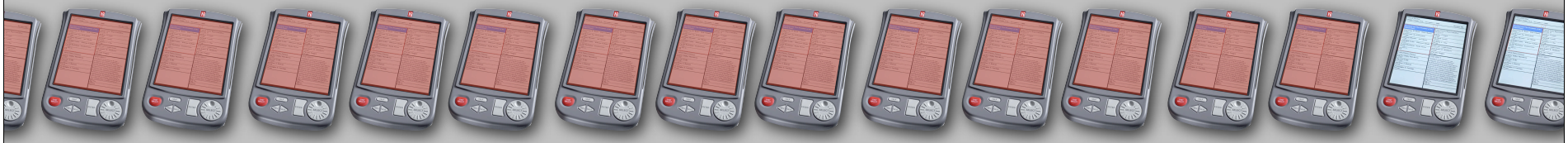
Viral attacks?



Viral attacks?

SERVO

End of election inventory management / auditing



All subsequent machines compromised.

Attacked by voter

No easy way to clean a compromised machine

Must replace internal chips by hand

No easy way to detect compromised machines

Hacked machine can correctly answer network queries

Other Hart problems

Audio unit can be overheard with a short-wave radio

“Adjust votes” feature in tabulation system

Premier (née Diebold) and Sequoia had similar problems

What about ES&S iVotronic?
(Not considered in CA study.)

What's next?

Other states may follow California's lead

Limit use of DREs to one per precinct

Mandatory audits to compare paper to electronic records

Vendors will (hopefully) engineer better products

Optical scan paper ballots growing in popularity