

Increasing Reliability by Testing



“Perfection is too expensive”

-- source unknown

Testing phases

Evaluation



- inspection
- compliance check
- burn-in test
- install

Acceptance



- Eval == *this?
- burn-in test

Production



- HW self-checks
- scrubbing
- active data verify
- HW “changes”

Background probes (1)

- RAID verify/scrub
 - first run: 492 machines/ 8514 units checked, 298 problems
 - in March: 319 / 5734 / 71 (most due to firmware bug)
- RAID cards' low-level logs
 - correlations with data corruptions (fsprobe)
- SMART (self-tests + counters)
 - 30% of failures predicted; vendors replace these disks
- fsprobe: read-write-compare
 - memory errors as well as other silent corruptions
- memory test
 - in background, using 'memtester'

Background probes (2)

- IPMI log check
 - false positives: fan speed low etc.
 - checking temperatures (calibration!), fan count, logs
 - ~5 vendor calls per week (out of ~5000 machines)
- Hardware scan (inventory)
 - inventory vs. actual
 - CPUs, memory masked by the BIOS
 - Ethernet speed, FW levels
 - undocumented vendor changes (upgrade, replacement)
 - undocumented admin changes (!)
 - some scans (lshw) may trigger hardware actions

Background probes: **foreseen**

- Machine Check Exceptions (mcelog)
 - need to understand exact meaning of messages
- Error Detection and Correction (EDAC)
 - in RHEL4 but needs compliant hardware
 - reports ECC errors, PCI parity errors
- file checksum checking
 - background task on disk servers

Burn-In Test: Purpose



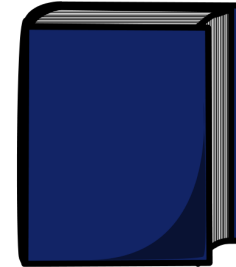
Burn-In Test: Properties



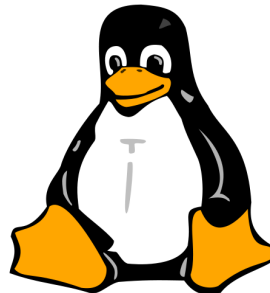
Decisive



Automatic

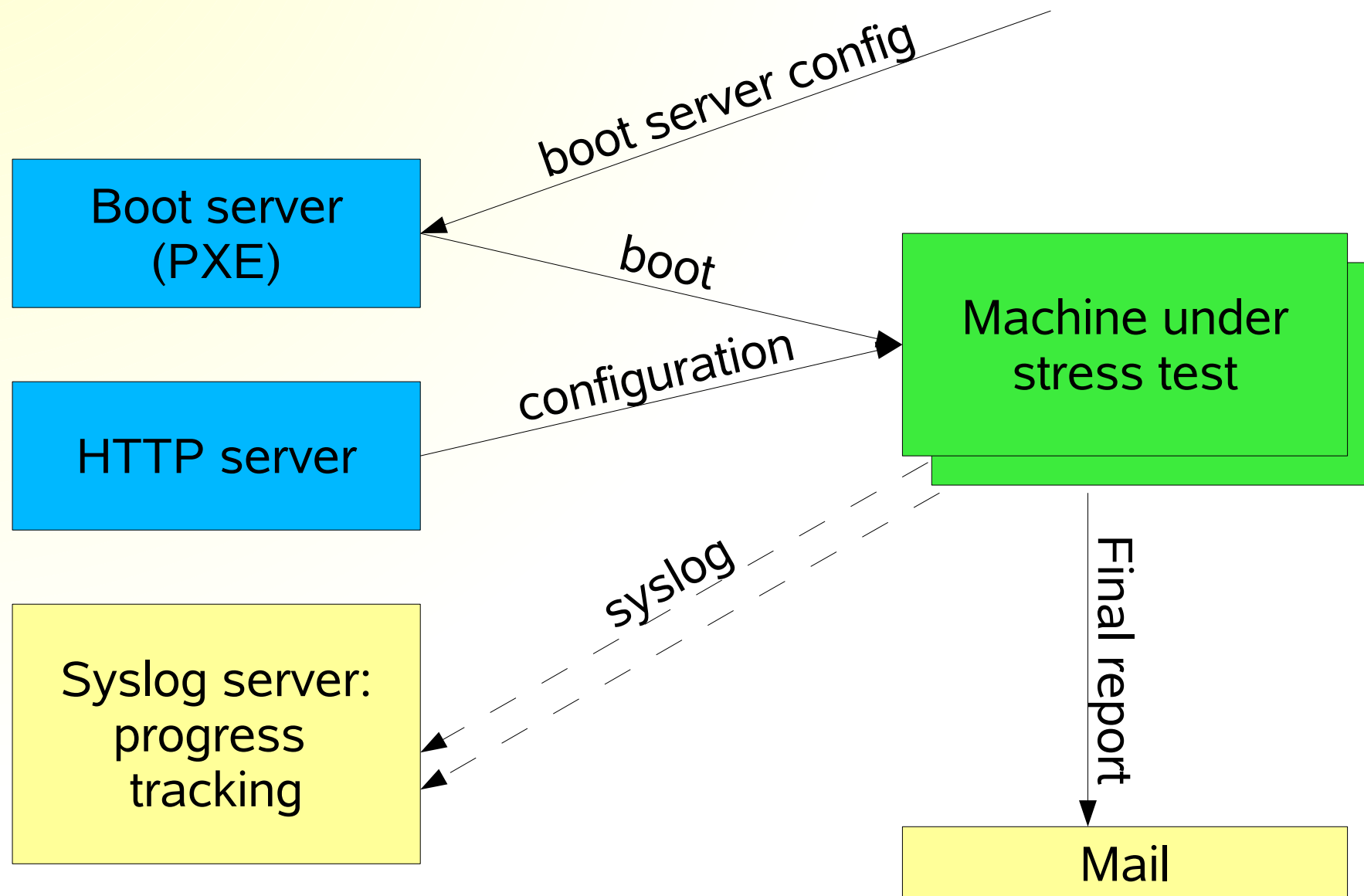


Descriptive

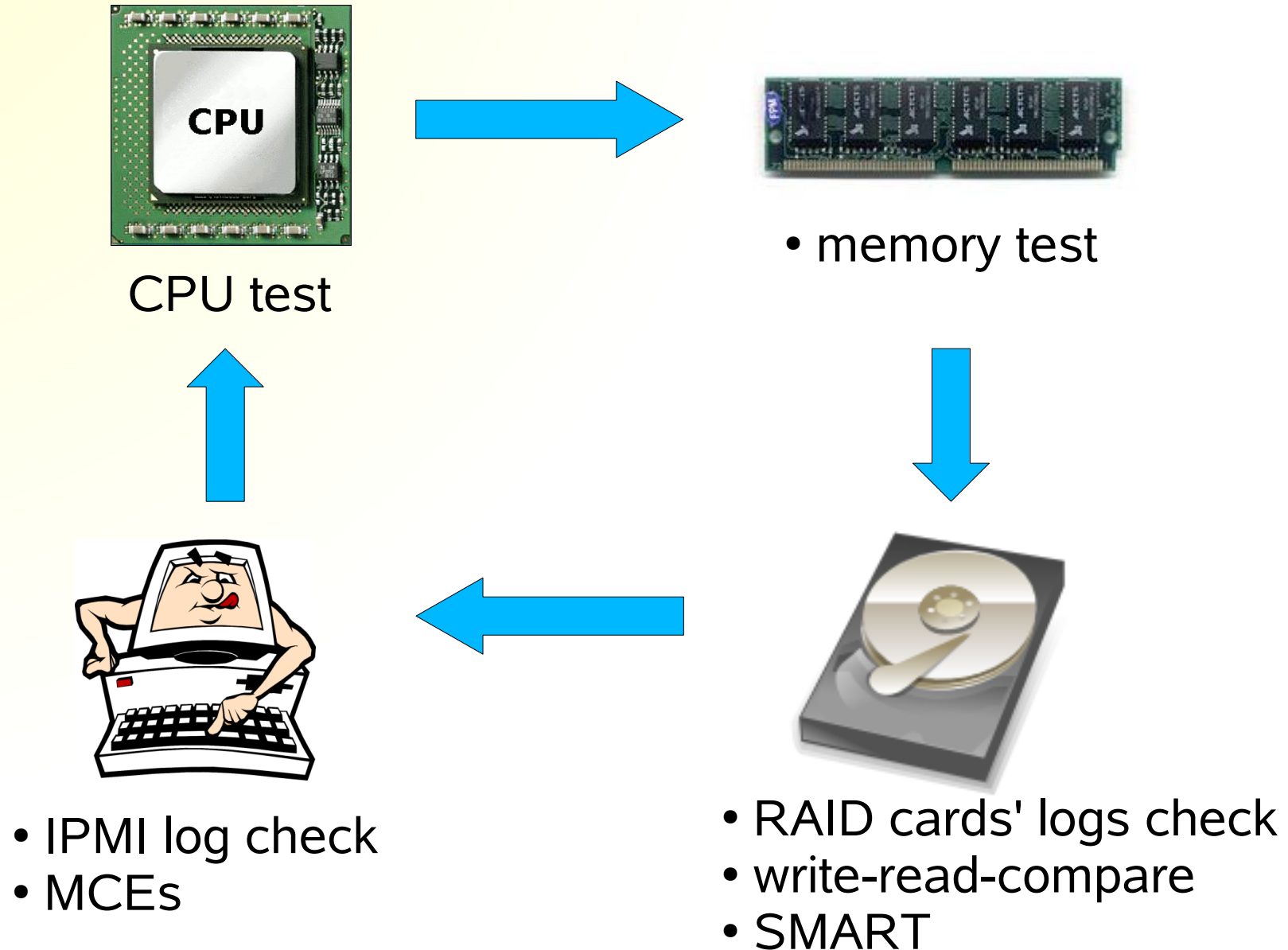


Maintainable

Burn-In Test: Architecture



Burn-In Test: One week



Burn-In Test: Tracking Tool

Firefox

File Edit View Go Bookmarks Tools Help

http://syslog-bit.cern.ch/bit.cgi?filter=lx&F=Ok

http://syslog...lter=lx&F=Ok

Burn-in test status

[See all hosts](#) (click on the "x" to limit to a given host)

Filter on the beginning of hostname (e.g. 'lx') here:

x	Syslog:	in Lemon:	Started:	Currently:	Status:	Last version:
x	lxb8021	Lemon	?	Logfile size > 10240k !	failed	?
x	lxb7373	Lemon	Mar 21 11:23:53	00-bit-cpu.sh loop 2 since Mar 25 08:44:12	PASSED	0.25-i386-bit-0.11-2
x	lxb7317	Lemon	Mar 27 16:30:04	finished Apr 4 13:16:09	PASSED	0.25-x86_64-bit-0.11-2
x	lxb7374	Lemon	Mar 21 11:24:04	00-bit-cpu.sh loop 2 since Mar 25 15:57:00	PASSED	0.25-i386-bit-0.11-2
x	lxb0035	Lemon	Mar 21 15:11:02	05-bit-memtester.sh loop 2 since Mar 26 13:55:48	PASSED	0.25-i386-bit-0.11-2
x	lxbladehp06	Lemon	Mar 27 15:03:56	finished Mar 31 21:01:00	failed	0.30-x86_64-bit-0.14-3

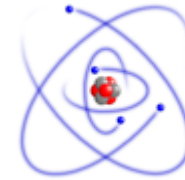
Summary

PASSED	lxb7373 , lxb7317 , lxb7374 , lxb0035
failed	lxbladehp06

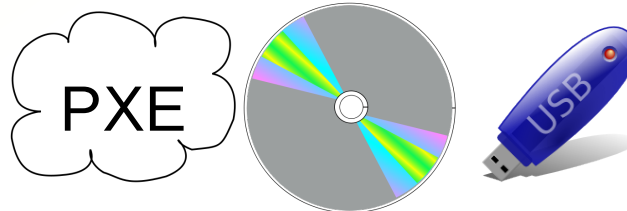
U-Boot Linux environment



Memory-based



Standard SL(C)4



Boot: network, CD, USB stick..

hwraidtools

hwraidtools

tw_cli

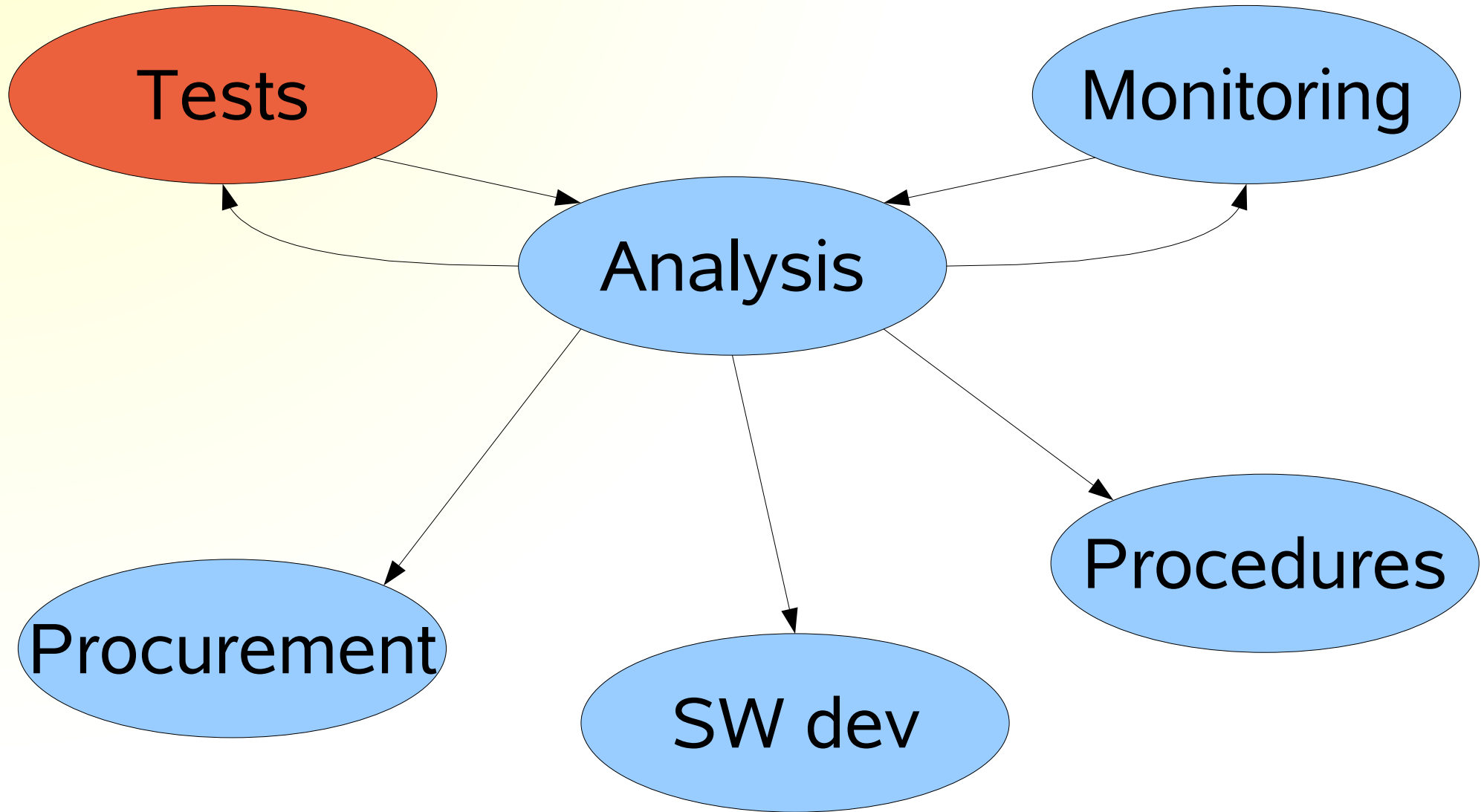
areca_cli

3ware RAID

Areca RAID

?

Feedbacks



However...

Computers are unreliable, but
humans are even more unreliable.

-- Gilb's Laws of Unreliability

More...

Burn-In Test:

<https://twiki.cern.ch/twiki/bin/view/FIOgroup/TsiProcNewBurn>

Monitoring sensors and tests (algorithms):

<http://cern.ch/elfms>

Contact:

Andras.Horvath@cern.ch,
Tim.Bell@cern.ch