How reliable is Windows anyway?

Kostas Tourlas
kt@adelard.com
Does it matter?

- Why do we care?
  - safety related apps running on Windows
    - Windows as off-the-shelf component (COTS)
  - the way of things to be?
  - COTS issues central to our work

- Why would anyone else care?
  - business-critical applications
  - embedded Windows (CE for automotive etc.)
Our approach

- Be up-to-date
  - public-domain studies, liaise with Microsoft

- Investigate
  - own XP reliability study
  - look “inside” Windows

- Specialise
  - industry requirements (e.g. SW01, ARINC)

- Advise
  - best practice
  - application-OS integration testing
Not much out there

- Anecdotes abound, facts are few
- Technical studies
  - most commissioned by Microsoft
  - some independent
  - compare 2000/XP against NT/98
  - disparate operating environments
  - results hardly comparable
  - yet some synthesis possible
Public domain data

- National Software Testing Labs
  - highest MTBF quoted: >2800 hrs for Win2K
  - >200 hrs for Win98
  - real-user environment

- e-testing Labs
  - uptime in “user days” before crashing:
    - 90 days for Win2K, 124 days for XP
    - WinNT managed 5 days
  - 31 “user days” ~ 7 calendar days
  - script-based, all Microsoft apps
Public-domain data (contd)

- Dependable Systems of Systems project
  - large academic network of PCs
  - large number of reboots
  - Win2K reboot rate 1.3 times that for NT
    - Win2K less reliable
    - may reflect user practices
  - Win2K availability of 96.3% vs. 93.7% for NT
Microsoft’s data on crashes

**NT4**

- **Device drivers 16%**
- **Anti-virus 12%**
- **Other third-party drivers 16%**
- **Core NT 43%**
- **Hardware Failure 13%**

**Windows 2000**

- **Hardware 22%**
- **3rd party Kernel code 11%**
- **Anti-Virus 4%**
- **system Config 34%**
- **Non Signed Drivers 20%**
- **Signed Drivers 7%**
- **Core NT 2%**

Improvements minimise downtime, recovery time
Adelard’s own study

- Academic network of ~600 XP PCs
  - observed Jan to Mar 2003
- Logged shutdown, blue-screen, re-boot events
- Highly unpredictable, changing environment
- Our goals
  - estimate reliability of WinXP
  - determine range of feasible analyses
    - supported by data from XP’s event logging
Win2K/XP event log

Uptime Report for: \STAFFA
Time Zone: GMT Daylight Time

System Events as of 07/04/2003 14:40:45:

Date:       Time:        Event:               Comment:
---------- ----- ------  ----------- --------  ---------- -------------------
20/01/2003  14:44:03  Shutdown
20/01/2003  14:44:53  Boot                  Prior downtime: 0d 0h:0m:50s
20/01/2003  15:44:54  Shutdown             Prior uptime: 0d 1h:0m:1s
[.. .]       
09/03/2003  06:42:42  Bluescreen            STOP 0x0000007a
[.. .]       
07/04/2003  14:29:34  Abnormal Shutdown     Prior uptime: 0d 3h:10m:9s
07/04/2003  14:38:03  Boot                  Prior downtime: 0d 0h:8m:29s

Current System Uptime: 0 day(s), 0 hour(s), 4 minute(s), 55 second(s)
Time-to-Bluescreen analysis

- Linear vs. logarithmic distribution fit
- Mean of 1.8, or TTBS of 600~1200hrs approx
The picture over time

Time to bluescreen (moving average)

- Start of term (semester)
- Weekly variation

Elapsed time (hrs)
How XP fared overall

- 300 bluescreens
  - MTTBS range of 600-1200 hrs
  - 130 PCs blue-screened (21%), over 500 didn’t
  - users rebooted every 15hrs on average
  - potentially multiple failure rates

- Over 25,000 application crashes
  - 6% were WinExplorer crashes
  - exact amount of use unknown

- Suggests XP resilient to application failure
Windows failure analysis

- What causes Windows to fail?
  - hardware & kernel software (drivers)
  - vulnerable system processes (logon, explorer)
  - may not crash but “hang”

- What causes an application to fail?
  - termination by the OS
  - disruption by the OS
  - depleted resources, covert channels
Best practice

- **Windows users**
  - certified hardware and drivers
  - keep it single-app, minimum hardware
  - configure-out features (e.g. some services)
  - monitor logs, review regularly

- **Application developers**
  - Win2K/XP native compatibility
  - Test rigorously!
Integration testing

- “How-to” advice
- Confidence in Windows
  - verify all drivers
  - kernel memory leaks?
  - investigate blue-screens
- Confidence in your application
  - trap subtle page-faults
  - monitor API calls, open handles, registry use
  - Windows 2000/XP compatibility tool
- Performance monitoring
Conclusions

- Win2K/XP more reliable than NT
  - factor of 5-10?
  - wide variability across environments & between versions
  - need to characterise application environment
  - data hard to generalise from

- Much more work still needed
  - Adelard to start new XP study
  - failure modes
  - generic dependability case

- Stick to best practices, test thoroughly