

Assessing End-User Reliability Prior To Product Ship

Mario Garzia,
Mujtaba Khambatti,
Mingtian Ni
Microsoft Corporation

Assessing Windows Vista Reliability

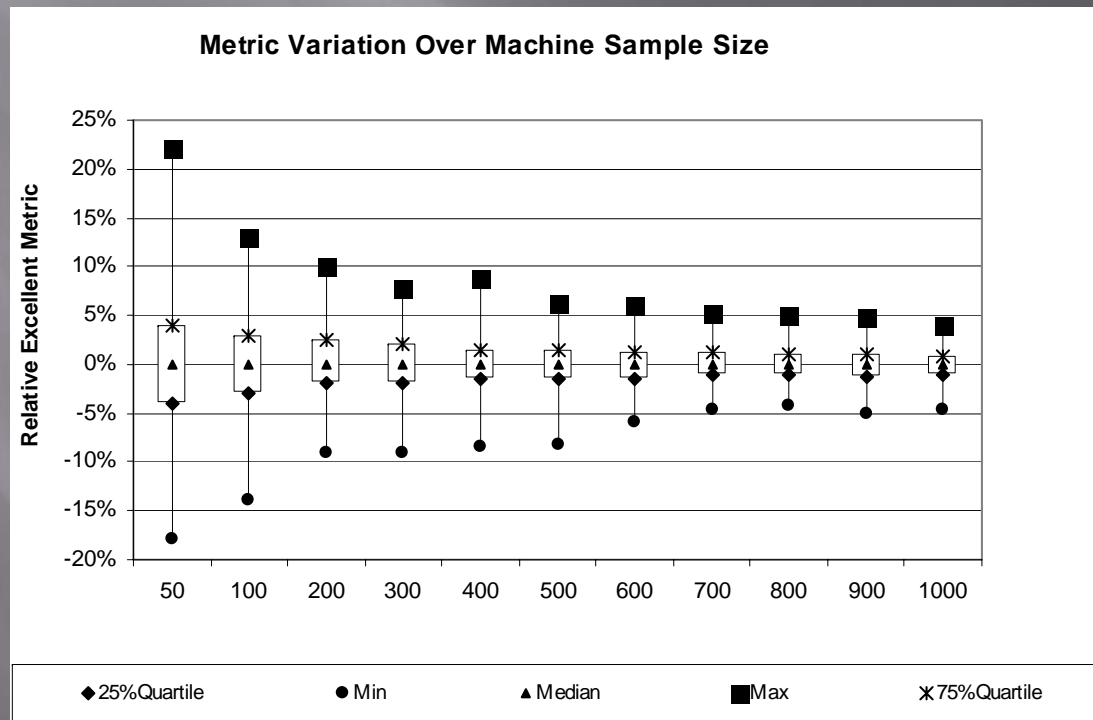
- ▣ The Windows operating system supports a very broad set of platforms and users
- ▣ Assuring reliability requires going beyond lab tests to cover the broad set of usage scenarios
- ▣ Each major Windows Vista build was deployed to users to assess reliability against objectives
 - Windows Vista reliability objectives were based on expected improvement over XP SP2 baseline
 - Reliability of escrow bits assessed using data collected from internal Microsoft deployments as criteria for build release
 - Post build release, data was collected from extensive internal and external deployments
 - ▣ For comparable population scenarios, the data from broad deployments was consistent with escrow results
- ▣ Data collection based on Windows Vista instrumentation tracking system activity
 - System data automatically sent back to MS (with user consent)
 - Rich information on type & cause of disruption
 - Prioritized reliability bugs based on user experience impact

Client Reliability Metrics

- ▣ For client systems reliability means providing the user with disruption free operation
 - ▣ Disruptions include failures as well as planned events
- ▣ Windows reliability assessment tracked user experience
 - ▣ Excellent, Good or Poor experience groupings
 - ▣ Each computer categorized as excellent/good/poor based on the number of disruptions experienced per unit of time
- ▣ Reliability metric used is % of systems with an Excellent, Good or Poor experience
 - ▣ This metrics provides good reliability estimates within a short period of time and with a reasonable number of systems

Results Variability Based On Sample Size

- User experience groupings into percentages results in metrics that provide good results with short runtime and few systems
 - Variability is within 10% for samples greater 200 systems, and within 5% for samples above 700 systems
 - When using 200+ systems, we were able to obtain good results with two weeks of runtime



Reliability Improvement Across Milestones

- Windows Vista reliability results exceeded objectives at RTM
 - The excellent experience group improved across product milestones
 - The poor experience group was 1% smaller than target by RTM
- While release criteria was based on a two week runtime, results remained consistent with data collected post release with thousands of systems running for 2+ months



Graph shows the % improvement for each Experience group relative to it's target value.

Concluding Remarks

- ▣ Assessing the reliability of mass market products requires going beyond laboratory tests and evaluating the product in actual use
- ▣ To be practical, release criteria need to provide good results within a short period of time and with a reasonable number of systems
- ▣ Results are greatly influenced by the observed usage scenarios, requiring broad external beta releases to assure a broad operational profile coverage
- ▣ The process described has been used successfully during the Windows Vista development cycle to assure reliability objectives were met before product ship

Microsoft[®]

© 2002 Microsoft Corporation. All rights reserved.

This presentation is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS SUMMARY.