Best Practices: Measuring the Success of Enterprise Testing

To determine whether an enterprise solution is ready to deploy in production, enterprise IT organizations must perform substantial testing before deployment. Testing requires extensive planning and numerous resources, but enterprises should also determine whether the investment in testing is paying off.

BY CYNTHIA LOVIN AND TONY YAPTANGCO

Related Categories:

Enterprise management

Enterprise testing

Performance

Planning

Scalable enterprise

System deployment

Visit www.dell.com/powersolutions for the complete category index.

Determining the success of a product development and test effort is a matter of ensuring the organization has delivered a product that meets its customers' quality expectations.[1] Of course, an organization cannot wait for its customers to provide feedback as to whether it has achieved this goal, especially if it has not. Rather, an organization must measure several dimensions of the development and testing to determine whether it has been successful in meeting quality goals.

Developing a measurement strategy

A measurement strategy needs to address four key aspects:

- **Process:** Metrics for the effectiveness and efficiency of the test process, which address the consistency of test efforts either by comparing current efforts with past performance or by measuring the consistency of test activities for different teams within an organization
- **Quality:** Metrics that address the quality of the entity being tested, typically addressing both the volume and the severity of defects
- **Schedule:** Metrics that assess the overall schedule adherence to the test plan or that compare the duration per test phase against past programs
- **Cost:** Metrics that measure investments in a particular test project

A good approach to measurement is provided by the Six Sigma framework, which describes four families of measures. These families are known as the PFQT measures, which detail the productivity (process), financials (cost), quality, and timeliness (schedule) of a project.[2] An effective measurement strategy includes metrics that address these four areas.

Building a measurement infrastructure

Prior to implementing a measurement strategy, organizations must create a measurement infrastructure to track the number and type of defects reported during the development and test process as well as the number of resources allocated to test activities. This infrastructure also should set the goals and standards for the development and test organizations.

[1] This article is part of a series of articles examining best practices in enterprise testing. For previous articles, visit www.dell.com/powersolutions.

[2] For more information, see "Data Management Plans Can Reduce Project Cycle Times" by David Wetzel, *iSixSigma,* March 9, 2005, www.isixsigma.com/library/content/c050309b.asp.

Defect-tracking system

A defect-tracking system can be as simple as a local database or as sophisticated as one of the many defect-tracking tools available today. A key characteristic of a defect-tracking system is the ability to categorize defects by problem area, state, and severity. The problem area typically can be associated with the development team responsible for the component. The state of a defect is its current phase in the defect life cycle. Typically defect phases include *created*, *fixed*, and *closed*, but they can also include other states such as *assigned* (assigned to a specific developer for resolution) or *rejected* (determined not to be a valid defect). For severity, defect-tracking systems may use a scheme of three to five defect types. For example:

- **High:** These are referred to as severity 1 defects in many defect-tracking tools. These critical defects cause serious problems at customer sites and must be fixed prior to the product being shipped.
- **Medium:** Known as severity 2 defects in many defect-tracking tools, these defects can cause some part of a product feature set to deviate from requirements. In some instances, these defects are specific to a configuration or to a specific set of user actions.
- **Low:** These are referred to as severity 3 defects in many defect-tracking tools. These defects generally do not affect the functionality of the product but may give a poor perception of the rigor of the development and test process—for example, misspelled words or inconsistencies between the user interface and the user guide. Many organizations choose not to fix all of these defects, although a high volume of these types of defects can cause the perception of a poor-quality product.

Metric	PFQT family	Description
Defect management	Quality	Trend of defects found and fixed per week; indicator of defect volume and required capacity to resolve outstanding defects
Defect arrival rate	Quality	Pareto distribution of incoming defects; may indicate complexity of defects to be fixed and possible feature hot spots
Defect incoming/ closure gap	Quality	Indicates how much change is going to occur on the product; may be used to support decisions about increasing or decreasing test coverage
Defect aging	Quality	Pareto distribution of outstanding defects by severity classified by < 7 days, < 14 days, < 30 days, and > 30 days; another measure of organizational capacity and defect complexity
Test execution progress	Timeliness	Test cases executed compared to plan; pass rate can be correlated to number of defects to be found and focus areas for regression testing

Figure 1. Common operational metrics

Resource-tracking system

A resource-tracking system is necessary to measure the number of hours spent on test activities. Useful metrics include the number of total hours spent on the project (for planning, execution, and reporting), the number of manual or automated test hours, and the number of hours spent during each test phase.

Goals and standards

Goals and standards are the benchmarks in a measurement system. These may be defined as a percentage improvement over historical programs, or they may be based on an industry source or a best practice from another organization doing similar work.

Identifying key measurements

A measurement system needs to provide metrics in three categories: operational, project assessment, and business.

Operational metrics

Operational metrics are designed to measure the project progress during the test execution phase—whether that phase is unit, product, or system testing. Organizations may use these metrics to assess the stability of the product at a given point in time, release readiness, and adherence to test entrance and exit criteria. Operational metrics are typically analyzed weekly during the test cycle to enable the team to make adjustments to the plan as required. Figure 1 describes common operational metrics.

Project assessment metrics

A key objective of any organization is to execute the test cycle for a given project according to the test plan. Project assessment metrics are designed to provide the organization with feedback on the efficiency of test planning, execution, and reporting efforts. These metrics can help increase the accuracy of planning efforts as well as ensure that comparable resources are being utilized for similar programs across the organization. Project assessment metrics should be produced shortly after the end of the test project, and may be useful in helping test and development teams drive improvements. Figure 2 describes common project assessment metrics.

Business metrics

Ongoing improvement is a requirement for any organization. Business metrics address whether an organization is executing tests

Metric	PFQT family	Description
Mean time to close	Productivity	Average time required to close a defect after it has been discovered
Test plan effectiveness	Productivity	Number of defects discovered per 100 hours of test case execution
Invalid defect rate	Productivity	Number of invalid defects expressed as a percentage of all defects; typical reasons a defect would be considered invalid are "working as designed," "already reported," "invalid configuration," and "cannot duplicate"
Adherence to resource plan	Financials	Variance by percentage in overall labor costs as compared with the test plan
Defect incoming rate by test phase	Quality	Percentage of defects found by test phase (unit test, product test, and system test); a best practice is discovery of 60% of defects during unit test, 35% during product test, and 5% during system test
Defect closure rate by test phase	Quality	Percentage of defects found by test phase (unit test, product test, and system test); defect closure should not lag significantly behind defect discovery
Volatility index	Timeliness	Number of scope changes introduced during the test cycle, including an impact analysis
Schedule adherence	Timeliness	Variance of calendar days from beginning of unit testing to end of system testing as compared with the test plan

Figure 2. Common project assessment metrics

Metric	PFQT family	Description
Defect discovery accuracy	Productivity	Number of defects discovered per month for all projects by percentage of valid versus invalid
Test plan effectiveness	Productivity	Number of defects discovered per 100 hours of test case execution as compared to historically similar programs
Test case execution profile	Productivity	Number of overall tests executed, percentage of manual versus automated tests, and defect discovery by each type
Adherence to resource plan	Financials	Variance by percentage in overall labor costs as compared with historically similar programs
Adherence to capital plan	Financials	Variance by percentage in overall capital costs as compared with historically similar programs
Field escapes	Quality	Number of customer support escalations after field release, including pareto distribution of reasons for escape
Defect incoming rate by test phase	Quality	Comparison of the incoming rate of defects by test phase as compared with historically similar programs
Defect closure rate by test phase	Quality	Comparison of the closure rate of defects by test phase as compared with historically similar programs
Schedule adherence	Timeliness	Variance of calendar days from beginning of unit testing to end of system testing as compared with historically similar programs

Figure 3. Common business metrics

with consistency and improving in its test efforts. Business metrics compare test planning, execution, and reporting efforts with historical efforts of a similar type. They are also used to assess consistency of test activities within the same team, among different teams with the same mission, and among various regions of a global organization. Business metrics are typically reviewed monthly or quarterly. Figure 3 describes common business metrics.

Gauging the success of test efforts

The primary goal of measuring the development and test process is to help ensure that enterprises are meeting the quality expectations of customers. Quality measures indicate whether a product is performing to the specifications defined by the marketing team and implemented by the product development group.

No one measure can be used to determine the quality level of a product. It may be tempting to continually introduce more and more metrics to assist in the management team's analysis, but a succinct list of measures can be used effectively. Of course, customers will always provide the final verdict as to whether an enterprise has met its quality goals and whether its measurement system needs tuning.

Measurement systems address key aspects of product development and testing, including consistency, effectiveness, and efficiency. A well-defined measurement system can help organizations deliver high-quality products and provide a basis for continual process improvement.

Cynthia Lovin is a senior consultant test engineer in the Dell Product Group Global Test Department. She has Six Sigma Green Belt certification and a B.A. in Business Administration from The University of Texas at Austin.

Tony Yaptangco is the director of the system test group within the Dell Product Group Global Test Department. Tony has a B.S. in Computer Science from San Diego State University and an M.S. in Engineering Management from National Technological University.