From Data to Insights: How IT Operations Data Can Boost Quality

By listening to the voice of customers, machines and test data, organizations can advance their digital business objectives and position quality as a key competitive advantage.

Executive Summary

The digital economy has transformed the way organizations do business. To survive in a complex and competitive marketspace, it has become imperative for enterprises to rely on a variety of software systems that drive customer experience and express brand value. The stakes today are higher than ever before, which makes software failure cost an unforgiveable mistake for business.

This has triggered a consequent realignment in business strategies, bringing a renewed focus on technology as a key enabler of digital transformation. Big data and analytics are the engines powering such transformations as they together yield intelligent solutions enabling better decisions, improving customer experience and driving innovation. In fact, by 2018, 67% of the CEOs of global 2000 enterprises will position digital at the center of their corporate strategies.1

This white paper offers an in-depth analysis of the benefits accrued to QA when analytical tools and strategies are applied to data from downstream operations systems. The insights generated by identifying root causes of testing holes and leaked defects, in our view, can be used to improve upstream quality and optimize testing. Error-prone code modules can also be identified, which should eventually lead to a streamlined test strategy before code is released into production. Moreover, by analyzing historical data and developing trends using live data, previously unanticipated defects can be predicted, enabling preemptive quality intervention.

Gaining Actionable Insights from Operational Data

The data residing in systems across the lifecycle can be utilized to map code to tests to defects and to incidents in production systems, making it possible to instantly isolate the root cause of a failure.

Terabytes of data generated every day by production systems can yield results that can be used across the spectrum of testing to enable advanced prediction of failures due to defects or environmental or configuration issues. This predictive ability empowers quality assurance in the digital economy.
How We Got Here

Even with 100% test coverage, applications still end up failing in production due to an assortment of reasons that current test methodologies cannot address. As has been observed, a significant number of production incidents tie back to an oversight in application testing, making it the topmost priority when developing a cost-avoidance strategy.

Systems utilized by IT operations professionals daily generate massive amounts of data. A wealth of information resides in such systems, such as log data, app or infrastructure monitoring data, transaction data, service ticket data, mobile device usage and app crash data, user preferences, app settings, network conditions, etc. While operational personnel, through a variety of advanced analytical tools, regularly leverage this information to enhance operational efficiency, QA teams cannot consume the output with the tools that they use. A significant amount of effort is required to transform this data before it can be leveraged to generate QA insights. In order to enhance quality efficiency, it is imperative that QA teams be able to gather insights from this data, without getting lost when rifling through voluminous operations logs.

Extracting Intelligence from Ops Data

Log files can be classified into three broad categories, based on the data source and SDLC process from which they are generated:

- **Voice of customers**: Incidents logged in traditional incident management systems from production systems.
- **Voice of machines**: System logs, server logs, database logs and application logs generated automatically by machines in production when an application is up and running.
- **Voice of tests**: Application logs generated in the test environment combined with defects, test cases and test execution logs from test management tools and code to test correlations generated by other specialized tools.

Workflow for Capturing Insights from Operations

![Workflow Diagram](image_url)
A simple four-step process can be used to transform and analyze log data:

- **Data massaging:** Input data typically comes in a variety of formats, ranging from unstructured free text to structured XML. This data has to be cleaned; missing data then needs to be represented through data imputation for easy manipulation in the next step.

- **Data analysis:** This involves applying text analytics techniques to identify key predictors and descriptors, which can then be used in a clustering algorithm to identify leading trends.

- **Data transformation:** Analyzed data can be transformed in accordance to business rules to discern correlations and generate a map tying code to tests to defects and to production incidents, thereby identifying probable root causes and possible resolutions.

- **Visualization and representation:** Correlated data can then be depicted graphically or tabulated in an easy-to-understand format that can be used by QA teams for quick analysis and action.

This process can be further extended through supporting utilities to automate failure resolution retroactively or proactively, thereby preempting failure through predictive analytics based on defect and test execution histories.

### The Benefits

By automatically correlating data across the software lifecycle and applying analytics, IT organizations can generate quality insights that traditional testing processes fail to provide. They can effectively collaborate and deliver by maximizing velocity and minimizing risk and also reduce IT spend on incident management. Sample use cases from the point of view of different stakeholders in the software value chain include:

- **QA analyst:** *Can I implement a self-healing solution to create automated test scripts and execute these tests based on defects found in production logs?*
  
  - Identify testing holes through analysis of applications lifecycle management (ALM), operations and incident logs.
  
  - Sample testing holes reports:
    
    - Coverage analyzer based on logs.
    
    - Defect map matrix based on nonstructured data.
    
    - Tailored reports to solve specific client problems.

  Once these testing holes or gaps are identified by correlating ops and test data, it can be further integrated with external test case generation utilities to auto-generate automated test scripts. These test scripts can then be executed using continuous integration solutions, and the results recorded to verify defect resolution, making a fully automated, self-healing process from a QA perspective.

- **Operations analyst:** *Can I build system intelligence to automate defect triage?*
  
  - Identify the root cause of a production incident from the voice of customers and associate it with testing holes or gaps.
  
  - Use advanced machine learning algorithms to learn and predict “triage root cause” values, thereby eliminating the manual component of root cause identification and enabling quicker defect triage.

- **Developer:** *Can I automatically identify functional gaps to feed into unit tests based on application logs?*
  
  - Identify and highlight functional gaps in code using baseline application binary scans.
  
  - Trace back testing holes or gaps to code elements to identify functional gaps, which can then be added to unit tests. Improved test coverage enhances the quality of the deliverable.
Insights from Operations: A Working Model

Quick Take

A U.S. discount footwear retailer wanted to make its QA process more intelligent. The retailer was looking at ways to test more intelligently and effectively to ensure minimal leakage of defects to a mission-critical production system.

By analyzing the operations logs, actionable insights were generated that helped improve future testing iterations through the processes outlined in this white paper. The client was able to test coverage and reduce cycle time by focusing on high-value tests. Moreover, minimized defect leakage helped the client deliver an improved customer experience. Figure 4 offers a snapshot of the log files that were taken as input and the inferences that were generated.

From Data to Insights

<table>
<thead>
<tr>
<th>INPUTS</th>
<th>INFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPLICATION SERVER LOGS</td>
<td>IDENTIFIED POTENTIAL DEFECT RECURRENTITY FROM THE TESTING PHASE.</td>
</tr>
<tr>
<td>DEFECT SUMMARY LOGS FROM THE TEST MANAGEMENT TOOL, HP ALM</td>
<td>IDENTIFIED NEW ERRORS WHICH WERE MISSED IN THE TESTING PHASE, READ 13,250 ERROR MESSAGES.</td>
</tr>
<tr>
<td>TEST CASES</td>
<td>FOUND FUNCTIONAL GAPS IN CODE WHICH WAS SUBSEQUENTLY USED BY THE DEVELOPMENT TEAM FOR ANALYSIS AND REMEDIATION.</td>
</tr>
<tr>
<td>LOG ANALYSIS CLASSIFIED THE ERROR MESSAGES AS 90% TESTING HOLES AND 10% DEFECT LEAKAGE.</td>
<td></td>
</tr>
<tr>
<td>CODE ELEMENTS TO DERIVE E2E MAPPING FROM DEFECTS TO CODE</td>
<td>MAPPED PRODUCTION ERRORS TO RESPECTIVE TEST CASES.</td>
</tr>
</tbody>
</table>

Defect Description ~ 63% Traceable to Test Case Description. Test Case Description ~ 70% Traceable to Defect Description.
Moving Forward: Shifting Gears

As enterprises balance between legacy and digital IT, addressing new paradigms of QA will require an ecosystem thinking for assurance solutions that are open, community-based and consumed in an as-a-model manner. Building on next-gen automation and intelligence capabilities, these solutions should be able to drive business and technology assurance outcomes for the enterprise.

Quality at speed requires enterprises to look beyond traditional realms of testing to improve test cycle times, while delivering predictable quality. The utilization of operations data has in our view resulted in improved quality of applications and has created positive user experiences.

The evolving maturity of analytics engines built on artificial intelligence, natural language processing and machine learning techniques empowers enterprises to use their operations data to create a quality-led competitive advantage. This not only translates into lower software development costs but also has widespread implications on customer experience and resulting brand value that are exceptionally valued in the digital world.

Footnote
1 Culled from IDC’s IT Industry, Digital Transformation and CIO Agenda webcasts.

About the Author

Vikul Gupta is Director for Digital Assurance and DevOps with Cognizant’s Quality Engineering & Assurance business unit. He has more than 16 years of experience in strategy, management, delivery and development on multiple technologies, and deep domain experience in DevOps, cloud and infrastructure automation, with roles ranging from product developer to key strategist. Currently, Vikul is responsible for creating differentiated DevOps service offerings and intellectual property that defines the go-to-market strategy for these offerings. He is a graduate of the National Institute of Technology, Surat. Vikul can be reached at Vikul.Gupta@cognizant.com.

About Cognizant

Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process services, dedicated to helping the world’s leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 100 development and delivery centers worldwide and approximately 255,800 employees as of September 30, 2016, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.