SOLVING THE DISAPPEARING TEST PROBLEM

How Splunk Software is Used to Track and Score Millions of Exams

Use Cases

- Troubleshooting Services Delivery
- Improving Operational Efficiencies
Executive Summary

What do you do when 100 “widgets” go into a machine and only 99 come out? And what if those widgets are standardized tests, the “machine” is a $7B exam processing company and a multitude of government and commercial customers—not to mention anxious test-takers—are waiting eagerly on the other end, expecting results?

Executives at one Splunk customer struggled with this exact situation and were considering a variety of costly solutions to address an even costlier problem: missing exams meant broken contracts, lost business opportunities and a tarnished company reputation.

After reviewing the available options, the executives determined that Splunk® Enterprise had the three key components necessary to diagnose and resolve their issue. Splunk software now enables them to leverage their machine data to:

• **Gain visibility across data silos.** They knew the data necessary to find the “lost” exams was somewhere in the enterprise, but they had no way to search across the various data silos efficiently enough to be useful. Splunk software let users search across data silos with one query.

• **Sift data in the ways that mattered.** Log data was naturally organized by system, but technicians needed to filter the data by exam. Because Splunk Enterprise doesn’t force data into pre-determined structures, they were able to filter data from a variety of systems to pinpoint the activity surrounding a specific exam.

• **Automate and simplify manual time-consuming processes.** The visibility enabled by Splunk software allowed the IT staff to access machine data from across the infrastructure and all of its systems. Staff no longer had to write custom scripts, manually search through the machine data to locate a bottleneck or determine a root cause.

• **Clarity into compliance and customer service.** Additionally, this customer found Splunk software to be useful in other ways. Because they could track human graders’ activities by time, they identified employees who were overworking and violating compliance standards. They also used Splunk Enterprise to alert their technicians to problems before customers were affected and the support staff was overwhelmed.

The Disappearing Test Problem

Turning a standardized test into a carefully graded and fully validated document is a significant technical challenge. Over several years, this customer had custom-built a system consisting of five Java applications running on JBoss middleware and an Oracle database. This system performed over 25 separate functions, including matching exams with other registration information, comparing exams to answer keys and even processing results from human-graded portions of exams.

At one time, the system could process as many as five million exams over a single weekend, but unrelated organizational and technical demands led the customer...
to switch from a centralized, mainframe-based system to a distributed architecture.

While this change solved old problems, it created a new problem of exams going missing, forcing technicians to reduce batch sizes from the original five million exams per weekend to as few as 10,000. Even with these smaller batches, the customer found that for every 10,000 exams that went into the system, 9,997 would come out fully processed. The others would get lost in the system, initiating a time-consuming manual investigation to determine the root cause and find the exams.

**The Business Process: Exam In, Grade Out**

The customer’s complex, distributed grading system consisting of the 25+ step exam evaluation process can be grouped into five basic categories (see Figure 1):

1. **Import exam.** Exams in both paper and digital media are imported into the evaluation system.

2. **Match exam to registered student.** Exams are matched to examinee registration data and validated to ensure the right examinee took the right exam.

3. **Machine score the exam.** An answer key is loaded and compared to the answers provided on the test, and then a score is generated.

4. **Human score the exam.** For exams requiring more than machine evaluation, the system queues and presents the test for a grader and processes the results.

5. **Aggregate and publish the scores.** Once an exam is fully reviewed, scores are aggregated and published on a web-based portal for customers and examinees to review.

A small percentage of the exams that passed through this multi-step grading process had issues. Customer test-takers would check the online portal for their test scores only to discover the exam results were “still processing” days after they should have been available.

This led many customers to call the company’s help line. To resolve these issues, a dozen-person, cross-functional team often worked weeks to manually dig through complex, multi-tiered systems. This team represented datacenter infrastructure, hosting, database administration, application development, the software architecture board, customer-facing groups and at least one business executive. This ongoing manual effort cost the company hundreds of thousands of dollars.

**The Investigation: So Much Data, So Little Time**

Technicians troubleshooting the missing exams had no alternative but to manually sift through the various system events. This was problematic for a variety of reasons:

**The data was not centralized.** They had volumes of data from each of their five Java applications, their Oracle database and their JBoss middleware system. They also had logs tracking each exam as it went through the system.
through the system and logs tracking the overall enterprise service bus. The logs were of different formats and were located in different files on different servers. When exams went missing, technicians were forced to comb through each set of logs manually, a process that was tedious at best and futile at worst.

**Error messages did not help solve the problem.**
In some cases, technicians would find helpful error messages. For example, if a database error occurred, they might find a message like “record locked for update.” That helped them to pinpoint the culprit—the database—but they could not easily connect the error message to a specific exam. They were still preventing from finding the lost exam and resolving the customer’s issue.

**They could not correlate the data that would be most useful.** Logs were naturally organized by application or system, but the technicians were tracking down lost exams, not lost systems. Their system logs included test IDs, but isolating those test IDs and viewing the progress of exams across all systems was essential to their process and yet impossible with their existing tools.

The Splunk Solution: Operational Timesheets
Using Splunk Enterprise, the customer’s team focused on the three critical pieces of information necessary to find a lost test:

- The test ID
- The processes the test had completed
- The time when the test completed each process.

All of these data elements were present in the logs, but the technicians could not easily gather and correlate them. Splunk software enabled them to capture and visualize a test’s journey through the system (see Figure 2).

**Step #1—Gathering and correlating the data.** As indicated in Figure 2, each of the five Java applications wrote the test ID and its own workflow ID to an “authoritative transaction log” when it began work on a given test. For example, when the first application (known as “A1”) began work on test “jdoe1,” it wrote the following line to a log:

```
2011-10-02 09:01:17.12 Start attempt for import WFID=A1 TestID=jdoe1
```

**Figure 2.** Gathering and correlating the data.
Once those logs were indexed into the Splunk platform, the technicians had enough information to track a test all the way through the system. To view all the data related to test ID “jdoe1,” for example, the Splunk query is simply:

\[ \text{TestID}=\text{jdoe1} \]

If they wanted to see how far test ID “jdoe1” had progressed through the system, they could use the statistical analysis command “last,” as in:

\[ \text{TestID}=\text{jdoe1} | \text{stats last(WFID)} \]

and Splunk would display the most recent workflow ID (WFID) reached by test “jdoe1.”

If they wanted to identify those tests that hadn’t completed the entire cycle (as evidenced by reaching workflow ID “J5”), the command would be:

\[ * | \text{stats last(WFID) as LastWorkFlow by TestID | search LastWorkFlow!=J5} \]

Finally, if they wanted to investigate an issue, they could select one of the test IDs from the above search and do another simple search for all data associated with that test ID. The search would reveal all data related to that test and would identify the precise time when that test got lost in the system. After isolating the search to that second or minute, the technicians could search across all their data for issues reported during that narrow timeframe, reducing the technicians’ analysis time from hours and even days to seconds.

Once the process was automated, the customer no longer needed the standing cross-functional team to search for lost tests, reducing resolution time from weeks to minutes and saving an estimated $360,000 per year.

**Step #2—Using scheduled searches and lookup tables to expedite resolution.** Once technicians discovered that even the most basic of Splunk searches could dramatically improve their ability to locate lost exams and identify issues, they began to discover Splunk software’s other benefits.

**Scheduled Search.** Scheduled Search is a Splunk feature that allows the user to create and save a query and then ask Splunk software to trigger that query on a scheduled basis. In this case, technicians created a query to isolate and correlate the three key pieces of data (test ID, workflow ID, and timestamp) and then ran this query against their data early each morning, when other users were unlikely to be making demands on the system.

**Lookup Tables.** Lookup tables are flat, comma-separated files that enable the user to provide extra information about a piece of indexed data. For example, the raw indexed data might only contain “employee=Sondra.” The user could instruct Splunk to link that raw indexed data to a lookup table with email address and other information. If the final lookup table looked like this:

```
employee,email,role,location
Sondra,srussell@splunk.com,Sales Engineer,Washington D.C.
```

the following search would find the raw data containing “employee=Sondra”

\[ \text{location}="\text{Washington D.C.}" \]

Lookout tables were used in an unusual way to efficiently track the state of any given exam at any given time (see Figure 3). Using the following scheduled search (query).

\[ * | \text{table WFID over TestID | outputcsv} \]
the customer created and continually updated the lookup table (see Figure 4).

![Figure 4. Workflow ID lookup table.](image)

The final lookup table read like a timesheet for each exam and gave the customer a powerful, easily searched snapshot of the entire system status that it could link with raw data when needed. This combination of high-level system state summary and low-level system data allowed technicians to both quickly resolve tactical, one-off issues and answer strategic, organizational-level questions, including:

- Where is each exam in the system at this moment?
- What, if any, exams are “stuck” in the system?
- Which workflow stages take longer to process than expected?
- How long does the average exam take to fully process?

Continually analyzing the data from this perspective enabled the customer to move from a reactive organization—where users were forced to reduce batch sizes and devote extra hours to researching technical issues—to a proactive organization with the information at hand to improve service and accurately forecast capacity.

**Soft ROI: HR Compliance and Proactive Customer Service**

As with many Splunk customers, this company soon found other uses for its data, including reducing the demand for technical support and identifying hourly workers who are overworking.

**Reducing costly escalations and improving the customer experience.** The company provides exam takers with access to a web portal to check the status of their exams. Before its Splunk deployment, the company had difficulty identifying which exams were lost and therefore mostly relied on test-takers to call technical support and self-identify as having a problem.

Once Splunk Enterprise was in place, the customer used the alert feature to trigger a Python script. The alert identifies tests that needed to be resubmitted and the Python script sends an email to the examinee requesting the resubmission. This process yields the triple benefits of speeding up the resubmission process, reducing the inbound call volume to technical support and improving the overall customer experience.

**Preventing HR labor violations.** The customer employs hourly workers to hand-grade exams that require a high level of evaluation. Since these graders typically work from home, the customer’s human resources department found it difficult to ensure that employees were complying with standard regulations, including the requirement that workers take periodic breaks and that they not work between 1 a.m. and 6 a.m.

By querying the data, the operations team was easily able to identify employees who weren’t in compliance. The team reported a dramatic improvement—its investigation identified and counseled over 30 out-of-compliance employees and today it reports only rare instances of non-compliance.
Using Splunk Software, They Made the Grade

In this use case, we explored how Splunk software can enable operational intelligence for a classic workflow problem: how do you find stray objects in a system with a single expected path to successful completion? This use case demonstrated:

**Elimination of data silos.** Because Splunk software indexes all kinds of data, the customer’s siloed data no longer hinders its staff from quickly identifying lost exams and rectifying the issues.

**Correlations drive analytics.** Because Splunk software correlates different types of data, the customer could link error messages in one system’s logs to useful evidence in other logs and finally resolve issues.

**Flexible analytics powered by a read-time schema.** Because Splunk Enterprise keeps the data unstructured, the customer could view the data through the lens that mattered at the time it queried the data, not the lens that mattered to the system at the time it generated the log.

**Significantly reduce manual intervention and labor costs.** Because the Splunk platform provides real-time visibility of end-to-end data flow through the customer’s business process, when an error or bottleneck occurs, staff can diagnose and resolve the issue immediately, often in an automated way.

---

**About Splunk**

Splunk Inc. (NASDAQ: SPLK) is the pioneer in analyzing machine data to deliver Operational Intelligence for security, IT and the business. Splunk provides the enterprise machine data fabric that drives digital transformation. More than 12,000 customers in over 110 countries use Splunk in the cloud and on-premises. Join millions of passionate users by trying a [free trial of Splunk](#).