FAST AI SOLUTIONS TO TECHNICAL CHALLENGES USING LOGS

Solutions from BIG DATA SCIENCE RESEARCH:

BDSR makes it possible, to overcome the real-life big data business problems, by providing swift, high performance, accurate and leading-edge Analytical solutions to the companies.

Importance of Log Management:

Logs contain information related to events, which can include device states, monitor readings, errors, warnings and a variety of other information. Logs are often further characterized as **Data logs, Audit logs, Application logs, Events** and a variance of similar terms.

Problem Space where Log Management is necessary:

1. There is a huge challenge to incorporate data sets into analytics, as they are becoming bigger and more divergent. If this is neglected, it will formulate gaps and lead to unseemly insights.

2. We frequently come across the issues, while we are working with IT equipment, the issues may be corresponding to the server faults, malware attacks, DDoS Attacks (it will create a flood of traffic into server), configuration problems and hardware failure.

3. Precisely quantifying the robustness of IT assets across applications requires unerring enterprise-level application logging. This type of logging enables the enterprise to meticulously compute the health of its IT assets across the applications and congregate relevant criteria to reinforce those findings.

4. In order for an organization to maintain concurrence and high information security, each log needs to be analyzed and audited in the proper manner. Managing all of this information, especially in an environment with a more expansive network, produces a massive amount of convexity and creates a lot of haul on IT resources.

5. To know from where we are getting different kind of concerns, our system is generating logs that is very mammoth in volume. Logs coming from various sources, may not use the same format while creating or reporting them. Many solutions implement a common log format to address this problem, but not all logs can be forced to adhere to said format.

   There is no assurance that an incoming log will match what is being collated

and analyzed already. This requires more time and effort to discover key information and interpret that information accordingly.

6. These logs will help in understanding the troubleshooting problems, optimizing system, network performance, recording the actions of users and providing data useful for investigating malicious activity.

7. Apart from these, we can also track down the computer security logs that will mainly contain the user authentication attempts and security device logs that record possible attacks.

8. We can't understand directly anything after looking into the large volume of logs alone until we have the special assistance of log management tools.

9. These log management tools will parse all the sheer volume of the system generated logs and give the classification of different kinds of events happened.

10. This, indeed is another big problem. Despite of a huge demand for big data scientists and Big Data analysts that has been created in the market, there is a severe scarcity of finding skillful, versatile, maestro data scientists and analysts for the capacious amount of data being produced every moment.

11. Additional issue with standard log management is speed. Determining the correct balance and educating users requires a substantial investment of time and resources, making log management an analytical and procedural nightmare.

LogMiner approach :

LogMiner discovers the list of sources available in a Splunk server along with additional details that are available about each of the sources. LogMiner presents those sources and details in an intuitive way (such as a table or any other alternative representation). This view should guide the user to understand the details of available sources in the Splunk server. Log Miner unearths patterns at the source level. Pattern mining is accomplished at the source level. The below are the features of LogMiner.

1. **Root cause Mining:**
   Avails in implementing Root Cause Analysis for identifying component failure causal factors, which in turn enables tracking investigation findings. Diagonising and resolving major root causes as determined in Root Cause mining investigations prevents recurrence of the trigger incident as well as potential related incidents.

2. **Workflow mining:**

Performs **workflow mining**, the goal of which is to extract information about processes from transaction logs.

3. **Problem Solving:**
   Enables justification of any kind of riddles based on such log data.

4. **Detects security complications:**
   It is used to detect security incidents, operational problems, policy violations.

5. **Assists in audits**:
   Useful in auditing and forensics situations like employee internet abuse, computer misuse, fraud involved while using computers, accidental company data disclosure, data theft, deliberate disclosure of company data.

6. **Reducing space consumed by Logs and improve search performance**:
   Vigilance on the transaction log size, diminish the transaction log, affix to or expand a transaction log file, optimize the rate of increase of transaction log, and administer the growth of a transaction log file.

   Enables refined search and yields relevant results based on the (correct or incorrect) queries.

Few use cases of LogMiner:

1. **Use case of Log analysis for Web App Visitor behavior:**

   Log analysis is one of the superlative ways to comprehend your web application visitors' behavior.
   It exposes the information about:
   1. The number of users that have visited the application.
   2. On which products they had spent most of the time.
   3. What are they expecting from our products?
   4. What is their navigation path, from the moment they are entering into our application to the moment they are leaving out of our web application?
   5. At what time, we have a greater number of users visiting our application, etc.

   All these different kinds of user actions are parsed and their behavior is plotted graphically; these representations will help us to apprehend the correlation between the different variables.
   For example, at peak time, where more users will be visiting our application, we can advertise our products and send newsletter. We can optimize our application, based on the issues that will encounter during that peak time. We

can upgrade our application when user traffic is less. We can make an analysis to create an impact of our brand.We can note the repeated user visiting our application and target him to become a customer.

2. **Use case of Root cause mining:**

Root cause mining is a method of problem solving used for identifying the root causes of faults or problems.

Root Cause Mining can be accomplished in the following steps:

1. Identifying and extracting the root problem clearly.
2. Finding the available values for the root problem.
3. Getting the sample entity for each type of value.
4. Obtaining the sequence of events for each entity.
5. Distinguish between the root cause and other causal factors.
6. Establish a timeline from the normal situation up to the time the problem occurred.
7. Establish a causal graph between the root cause and the problem.

The below figure (Fig.1) shows LogMiner UI:

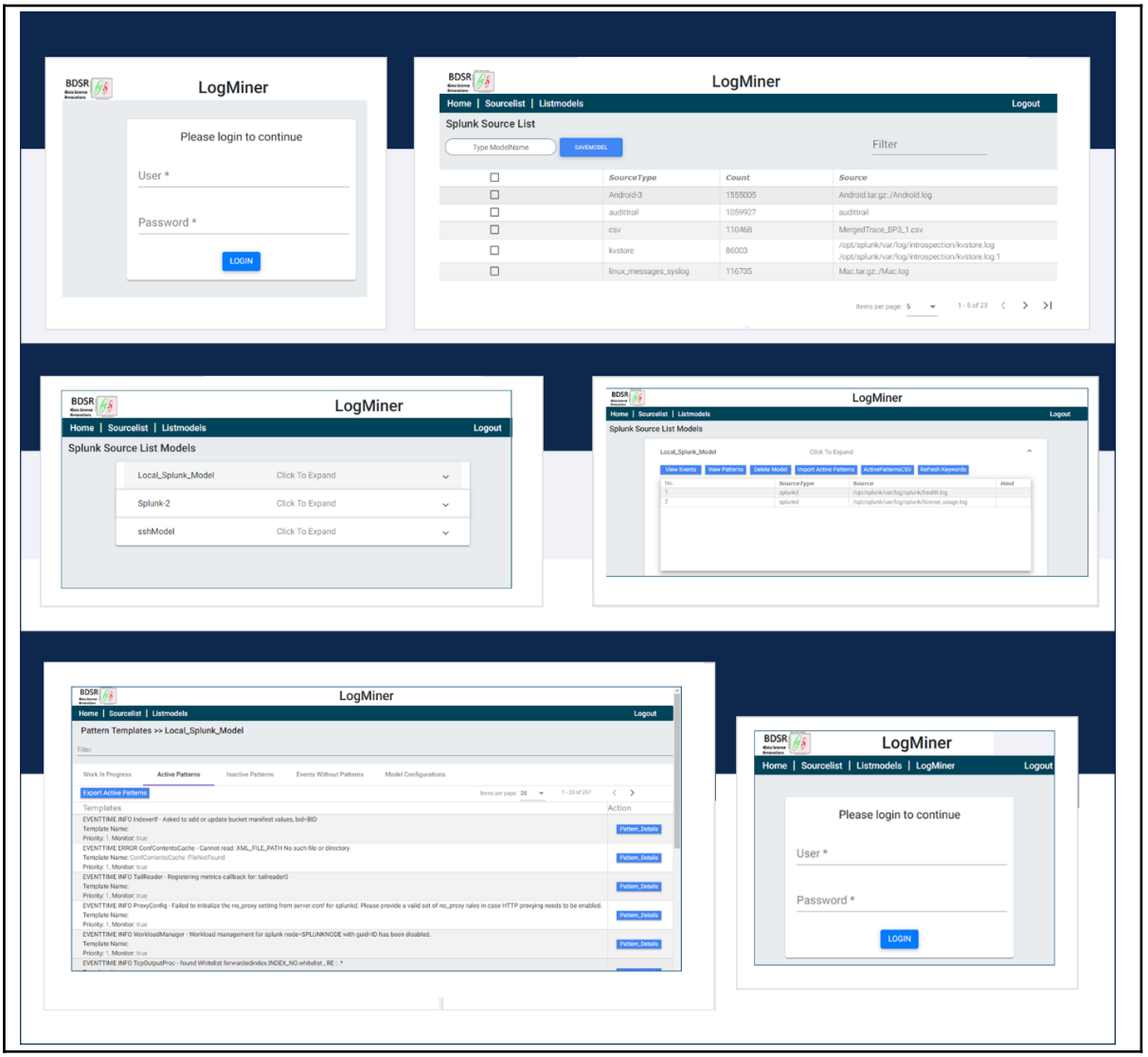

Fig 1

The below figure (Fig.2) shows handling events:

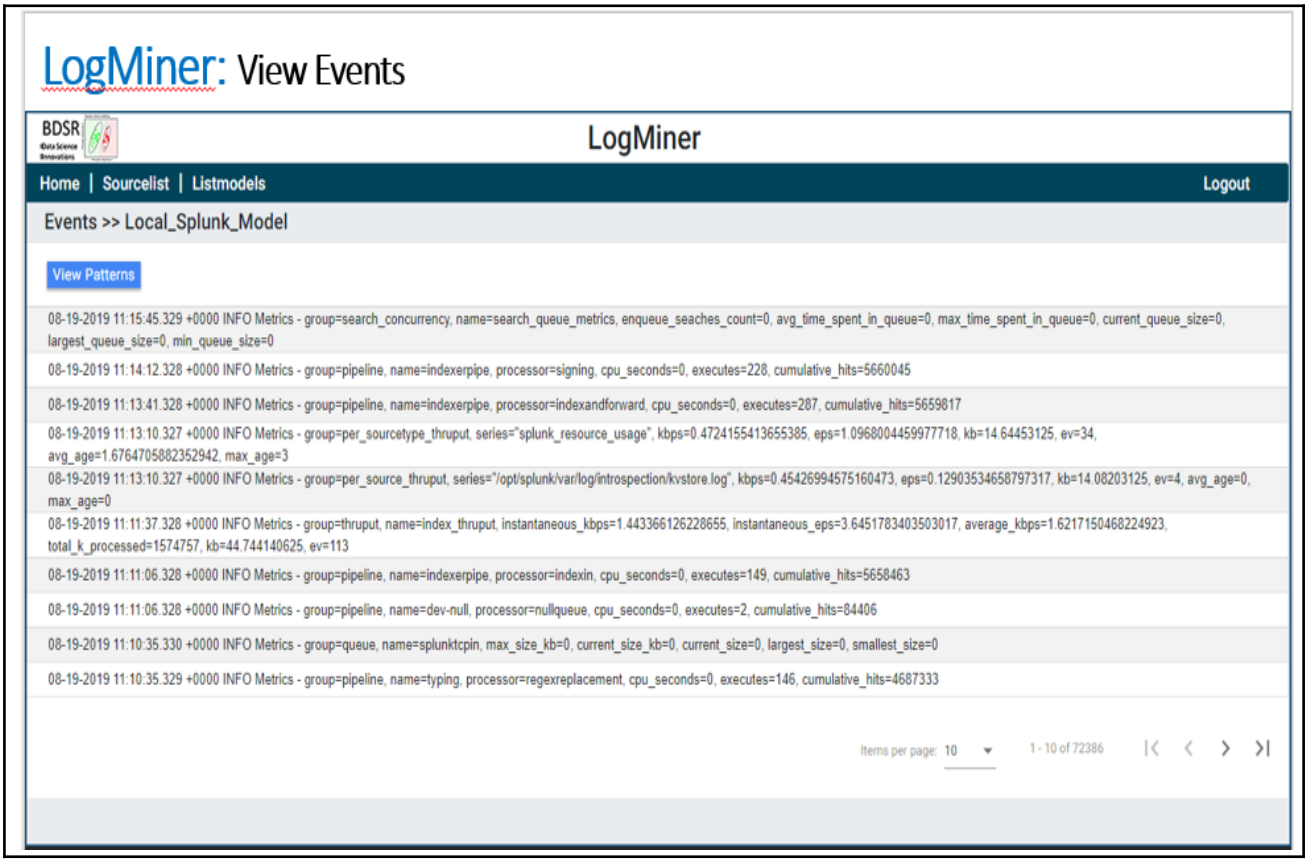

Fig.2

Prospects:

The companies who would have the following Requirements can seek solutions from us:

1. Use Splunk environment.
2. Require real time insights, using log data from application or system software.
3. Root cause mining of Splunk related environment.
4. Who have issues in search performance, installation, indexing, cluster-based, reporting.
5. Identification of patterns of log events.
6. Health monitoring of error patterns.
7. Field extraction of the event patterns.
8. Who require solutions development for big data and text mining business problems.

Solution Space LogMiner provides, to the problems in analyzing logs

To overcome all these hurdles, we have developed the product called “Log Miner” to parse the data into events and transactions where anyone can get understanding over the generated logs by any system or any other IT equipment.

The following are the solutions to the hectic problems where Log Miner will help the IT Companies to track down their large volumes of log data.

1. Event patterns that are happening in Splunk environment.
2. Drilldown to events related to an event pattern.
3. Fields extraction from event patterns.
4. Health monitoring of Error or warning patterns.
5. Events without Patterns.