

GSAW 2020
March 2-5, 2020
Renaissance Los Angeles Airport Hotel
Session 5: Data Exploitation

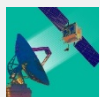
What Can We Learn from One Billion Ground System Log Messages?

Sharon Orsborne

and

Jay Bugenhagen

GMSEC Project NASA Goddard Space Flight Center
Software Engineering Division



GSAW 2020
March 2-5, 2020

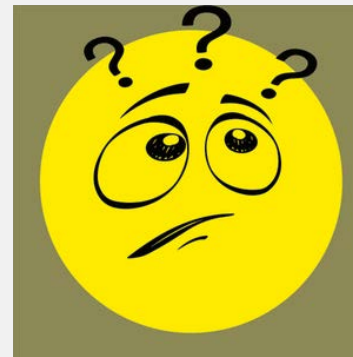
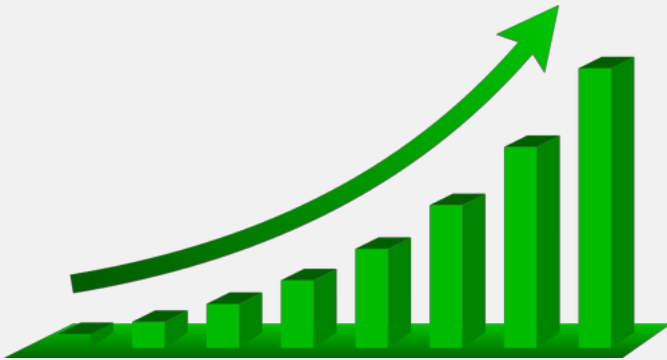
This material is a declared work of the U.S. government and is not subject to copyright protection in the United States.
Published by The Aerospace Corporation with permission.

What Can we Learn from One Billion
Ground System Log Messages?
NASA Goddard Space Flight Center

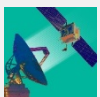
Events/Log Analysis - Current State



- 2016/101:13:34:05.34 ORB Sat XYZ Entering Eclipse. Duration = 12:34.0
- Historically, we have put tremendous effort into processing and displaying spacecraft housekeeping telemetry
- Ground system events/log messages have received much less attention
- With increasing system complexity the challenge is only growing



- “The messages scroll so fast we can’t read them; but if they stop scrolling, we have a big problem!” – ops personnel

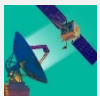


Use Cases



- Real-time dashboards to show:
 - Mission/Enterprise metrics
 - Systems status
 - Message traffic meta-data
- Long-term analysis of non-telemetry information
- Standard (daily) reports of various activities
- Ad-hoc analysis of anomalies
- Integrated real-time and archived data analysis

But really, once the capabilities are available the users will start realizing what can be done with their new powers.



Events Analytics -

A new NASA initiative that was planned for years



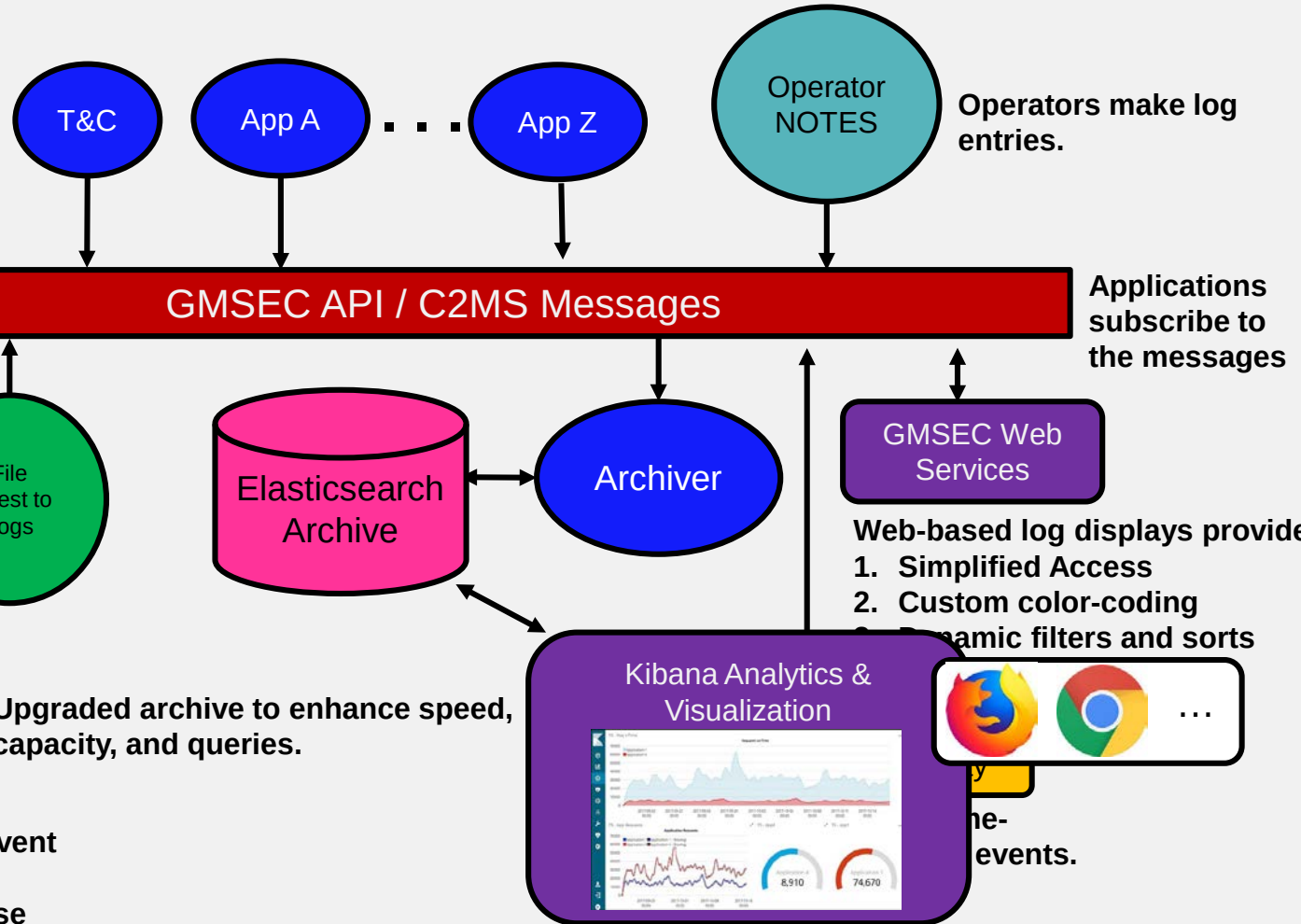
- So, how do we leverage all this event/log data to improve mission ops capabilities?
 - Dan Smith has pushed the idea of smart analysis of log data for many years
- New tools, part of the explosion in big data and analytics, are now available that can quickly be applied to the challenges
- Summer of 2019 – “Let’s see what the smart interns can do!”
 - Quickly decided to look at Elasticsearch and Kibana
 - “50 minutes to learn. No, not an hour.” - NASA intern
 - <https://www.youtube.com/watch?v=gQ1c1uLLyKI>



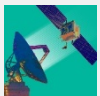
GMSEC Log Msg Approach



Current: Each component writes log messages to the bus.



Converting file contents to event messages creates a single integrated log to help increase situational awareness.



What is Elasticsearch?



Elasticsearch

- “ELK Stack”
 - Open source suite of tools
 - Elastic (the company) – lots of extra functionality and services built for Elasticsearch



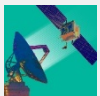
Logstash

Free, easy to learn and use, and... FAST, FAST, FAST!!

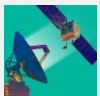


Kibana

But also: scalable, full-featured, large community support, plug-in architecture, web-based (Kibana), resilient, etc.



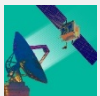
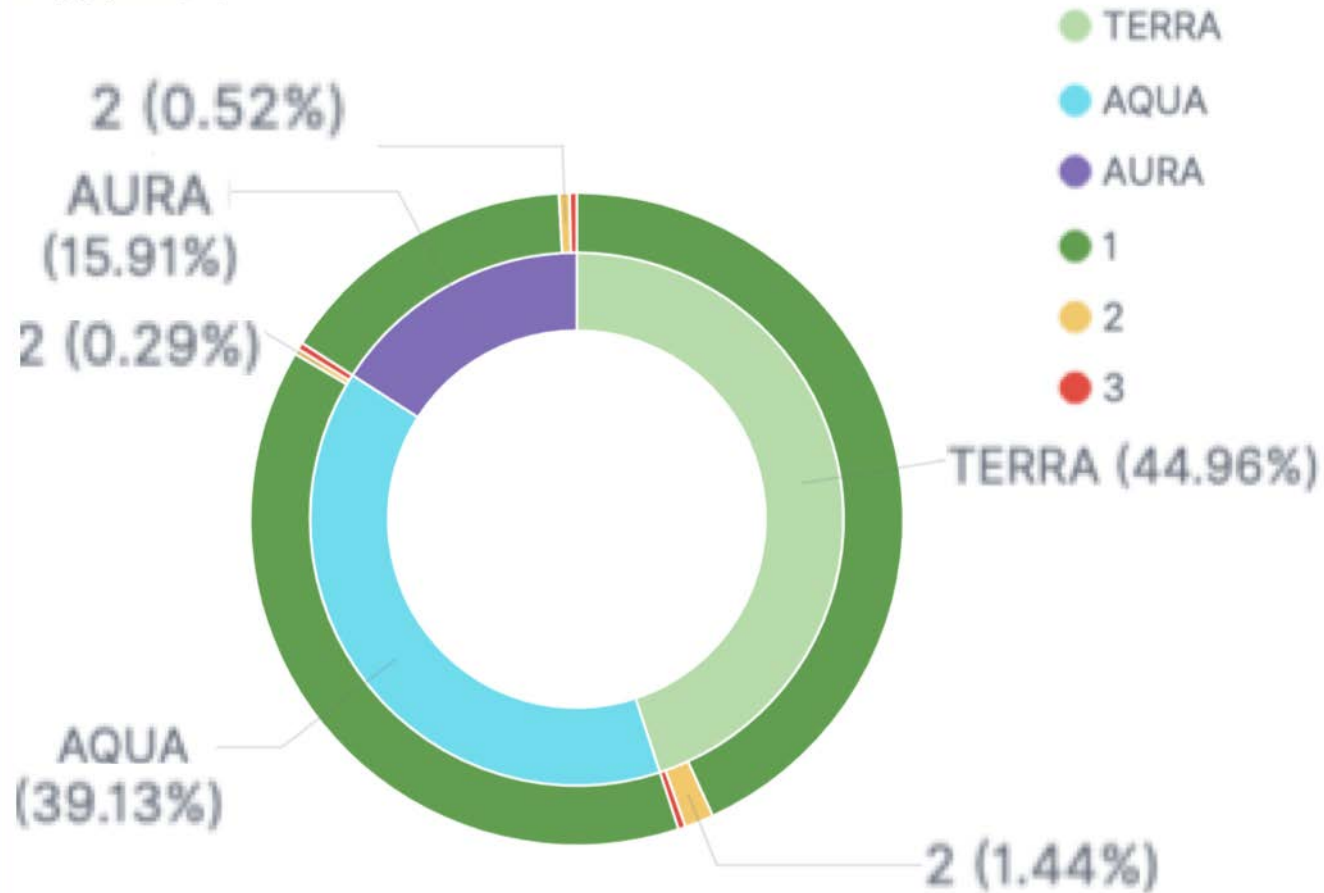
Sample Output - Dashboard



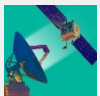
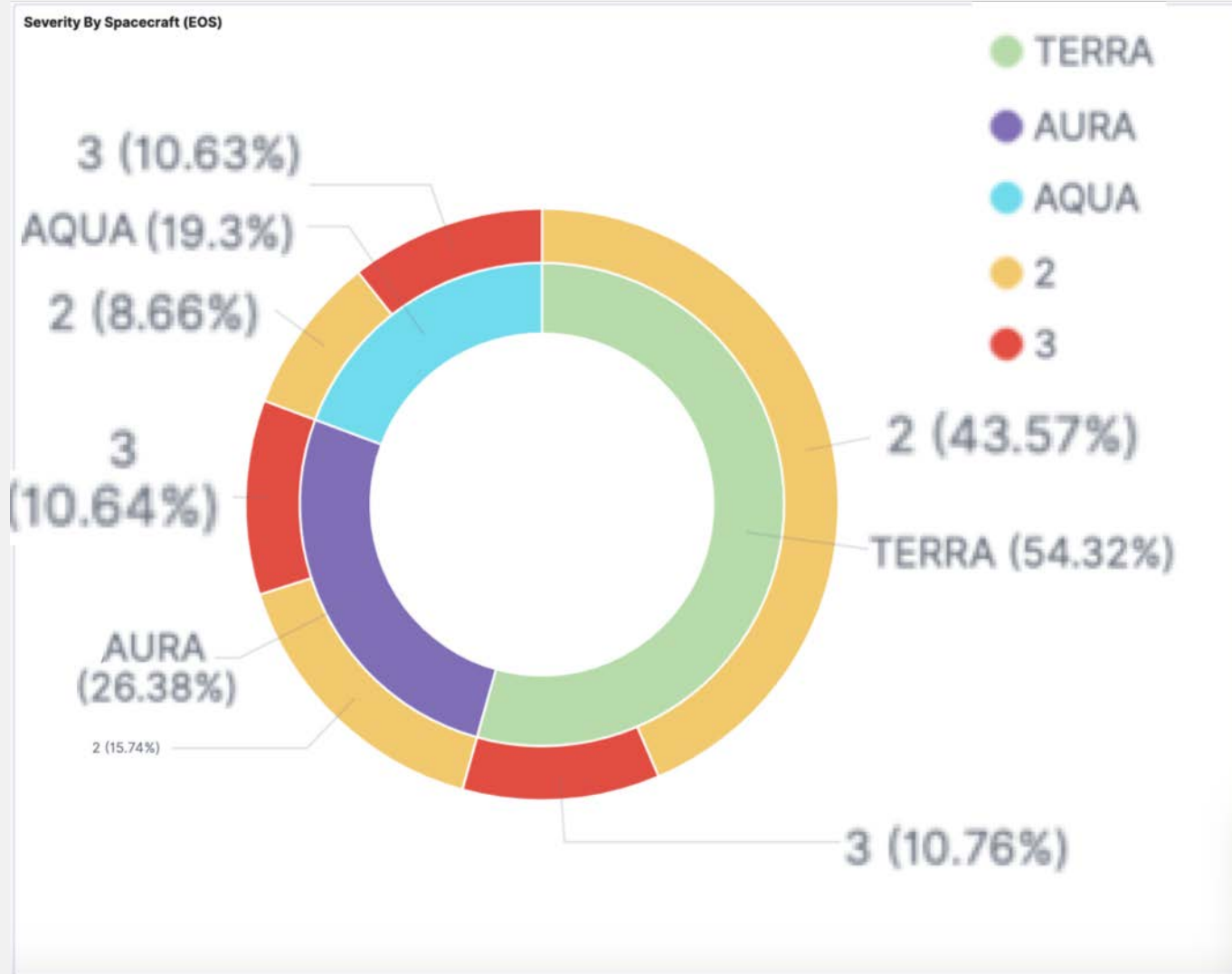
Messages Severity by Spacecraft



Severity By Spacecraft (EOS)



Messages Severity ≥ 2 by Spacecraft

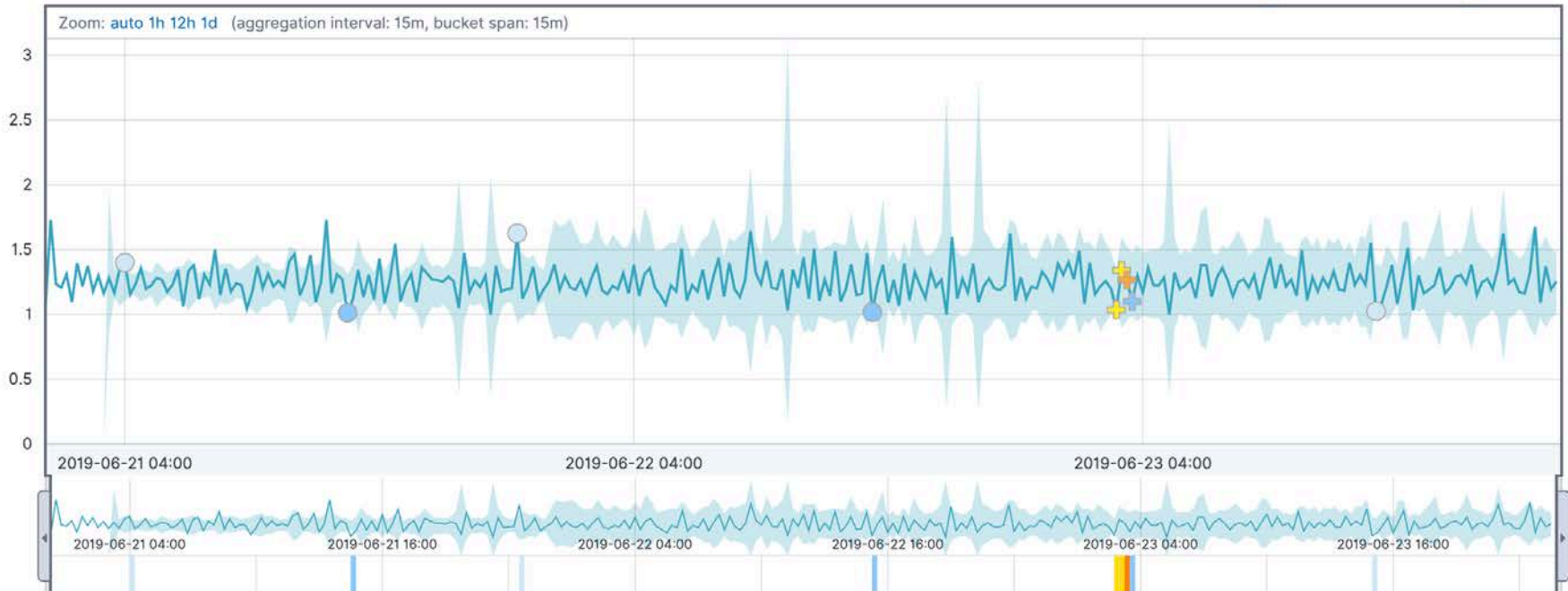


Machine Learning



Single time series analysis of avg severity

☒ show model bounds ☒ annotations



Anomalies

Severity threshold

warning

Interval

Auto

time	max severity ↓	detector	actual	typical	description
> June 23rd 2019	62	mean(severity)	1.267	1.219	↑ Unusually high
> June 21st 2019	18	mean(severity)	1.014	1.224	↓ 1.2x lower
> June 22nd 2019	7	mean(severity)	1.019	1.218	↓ 1.2x lower



GSAW 2020
March 2-5, 2020

What Can we Learn from One Billion
Ground System Log Messages?
NASA Goddard Space Flight Center

Mission Feedback



"All of our many servers' logs are distributed and in different formats – anomaly investigation is a pain in the you know what."

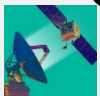
"Now we'll be able to detect and understand ground system anomalies much quicker."

"We're planning to consolidate our logs with GMSEC and Logstash/Beats and use Elasticsearch tools to ease analysis."

"We can combine ground system, spacecraft, and trending info to give better real-time situational awareness."

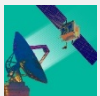
"Now I won't have to write perl scripts to manipulate and grep in the log files."

"Determining an unusual severity scheme was made apparent with just a few queries."





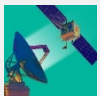
- Getting Started is Quick & Easy
 - Day 1: Downloaded ELK stack, created dummy database and explored queries.
 - Day 2: Wrote a bare bones Java application to feed messages from the bus into Elasticsearch db.
 - Day 3-5: Put actual mission messages into Elasticsearch db and started generating visualizations and characterizing the data - YouTube videos!
- The challenges:
 - Learning how to optimize Elasticsearch configuration
 - Exploring machine learning



Future Work



- Auto-generation of shift/status reports based on event message analysis
- Text-Based Data Analytics
 - Discover/learn trends or sequence “signatures”
- Move to natural language interface
 - Allow broader search ability
 - Support voice recognition
- Create an Ops Assistant
 - Imagine an Amazon Echo for each ops position
 - What would the conversation be?
- Access to telemetry values and event messages could be seamless, allowing queries, reports, and displays to easily intermix the two
- Time frame on plots could be based on query to event log (ie. Plot parameters x, y, and z for the last orbit, or for when John was logged on yesterday).



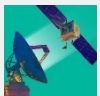


- Elasticsearch is the current path forward for several NASA projects due to cost, robustness, and ease of implementation
- We think that advanced analysis of events/log message will provide powerful capabilities for the mission user regardless of Agency or type of mission
- This is a powerful approach, especially using the GMSEC open architecture— the tools will work regardless of what other products are part of the configuration

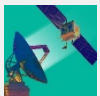
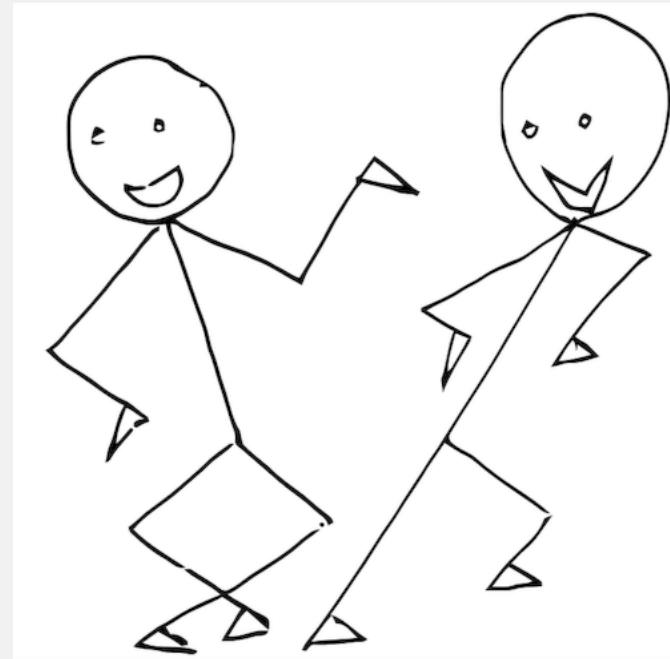
Sharon's
Picture

Jay's
Picture

Theresa's
Picture



THE END





Questions?

