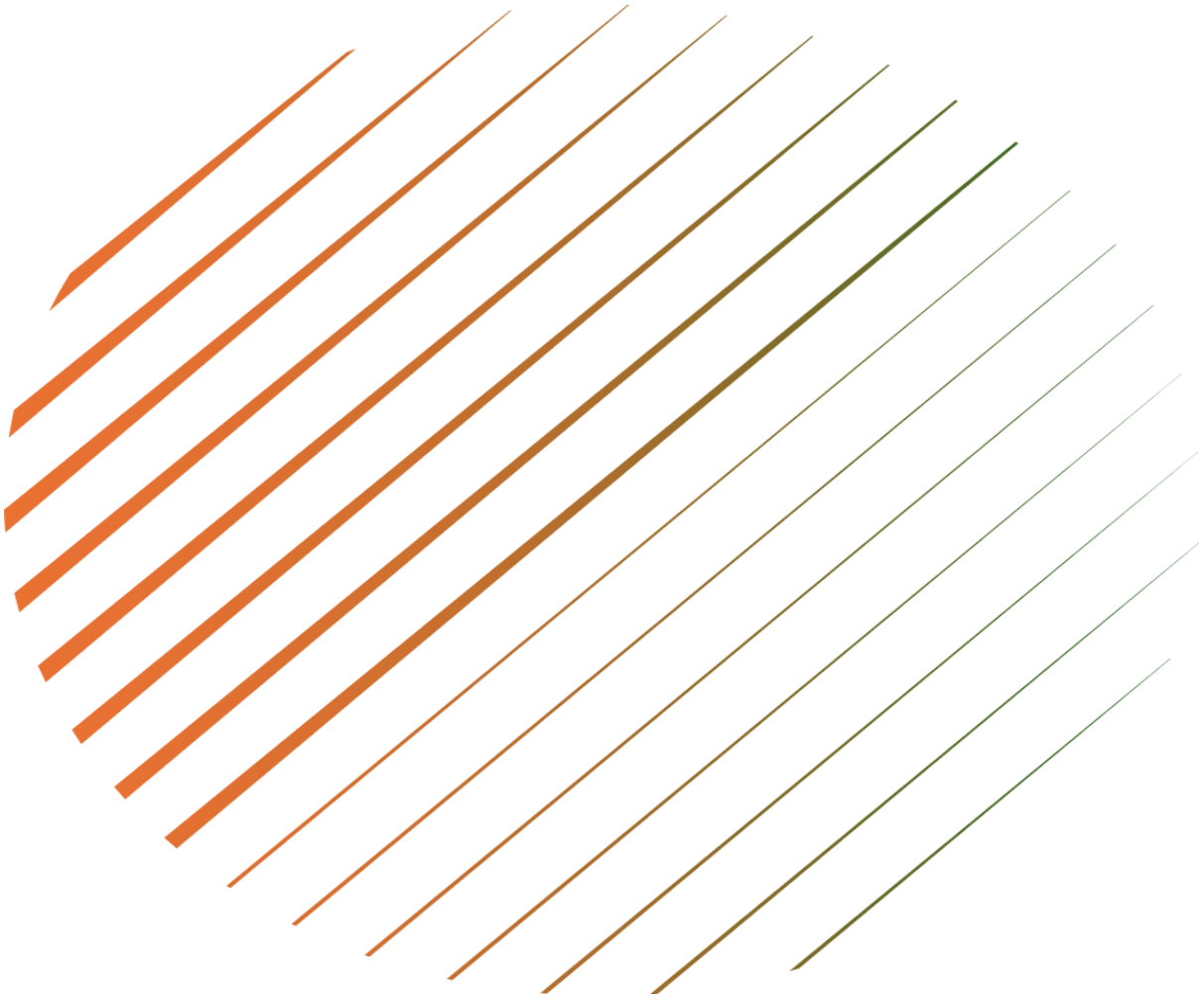
FTP LOG ANALYSIS AND FILE INTEGRITY MONITORING PROJECT



PREPARED BY: PRANIT KALAMBATE

ROLE/PROGRAM: TRAINED IN CEH CURRICULUM / CYBERSECURITY ANALYST TRAINING

DATE COMPLETED: 10-12-2025

1.PROJECT OVERVIEW

1.1 Goal

The primary goal of this project is to perform FTP (File Transfer Protocol) log analysis using Splunk Enterprise to:

- Extract critical FTP-specific fields like username and command.
- Analyze user activity (uploads, downloads, deletes).
- Monitor for security threats such as unauthorized transfers and failed login/transfer attempts.

1.2 Data Source

The project utilizes FTP server access logs ingested into Splunk.

Source: ftp.log.gz

Sourcetype: ftplog

2.FIELD EXTRACTION PROCESS (ADVANCED CONFIGURATION)

Unlike UI-based extractions, for this project, I used an Inline Regular Expression within the Splunk configuration to define fields permanently. This method demonstrates a deeper understanding of Splunk architecture and log parsing logic.

2.1 Inline Field Extraction Details

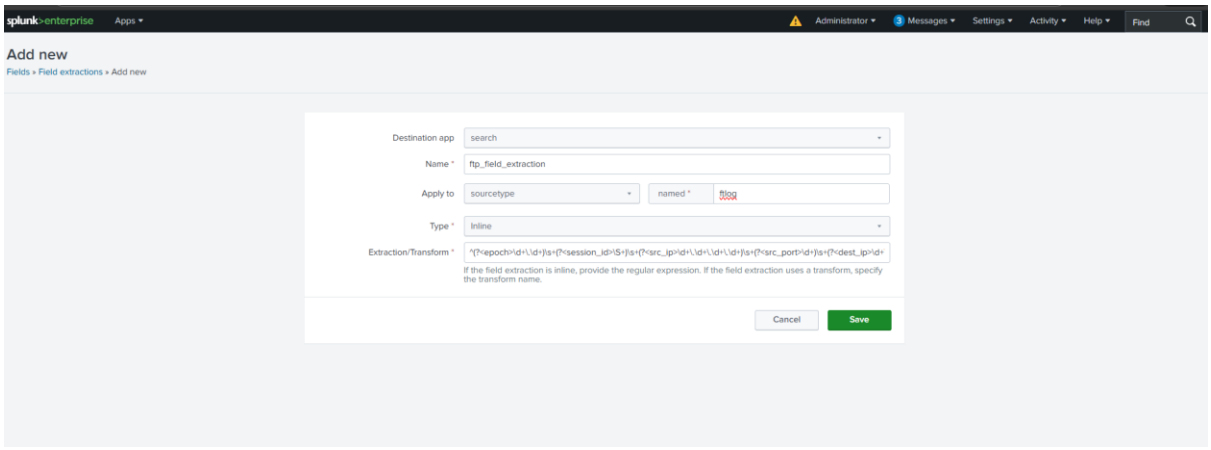
- Configuration Name: ftp_field_extraction
- Applied to Sourcetype: ftplog
- Type: Inline

2.2 Extraction Regular Expression (Regex)

The following specific regular expression was used to parse the raw FTP log events, assigning meaningful names to the extracted data groups ((?<fieldname>...)):

^(?<epoch>\d+\.\d+)\s+(?<session_id>\S+)\s+(?<src_ip>\d+\.\d+\.\d+\.\d+)\s+(?<src_port>\d+)\s+(?<dest_ip>\d+\.\d+\.\d+\.\d+)\s+(?<dest_port>\d+)\s+(?<protocol>\S+)\s+(?<username>\S+)\s+(?<command>\S+)\s+(?<fil

e_path>\S+)\s+\S+\s+\S+\s+(?<response_code>\d+)\s+(?<response_msg>.*)



3. ANALYSIS OF FTP ACTIVITY BASELINES

3.1 Total FTP Activity by Command

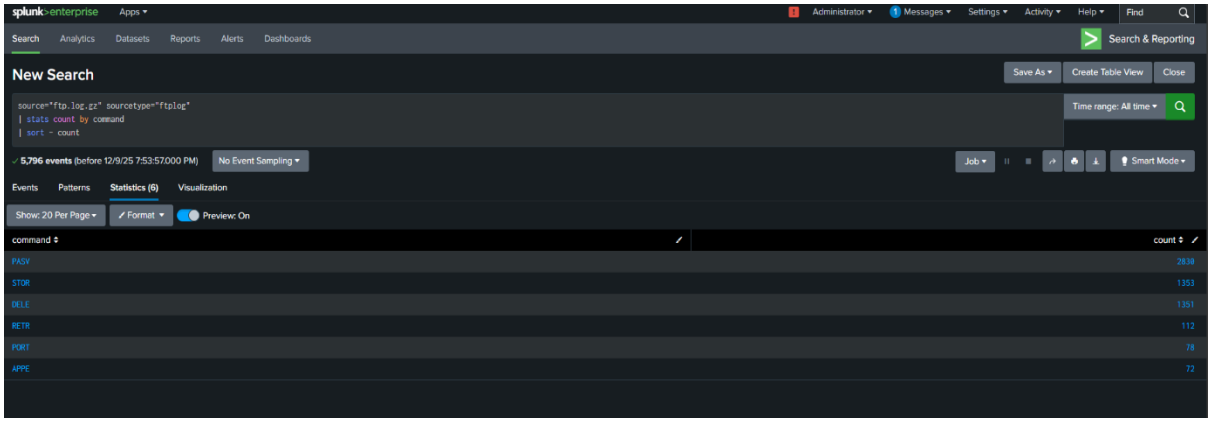
This analysis establishes the baseline of command usage, showing which FTP operations (uploads, downloads, deletions) are most frequent.

SPL Query:

source="ftp.log.gz" sourcetype="ftplog" | stats count by command | sort - count

Results (Partial): The results show a clear hierarchy of activity:

- PASV (Passive Mode) is the most frequent command (2838 counts), followed by STOR (Upload) (1353 counts).
- DELE (Delete) is also very common (1351 counts), indicating high file management activity.
- RETR (Download) is relatively low (112 counts), suggesting the server is used more for uploads/storage than downloads.



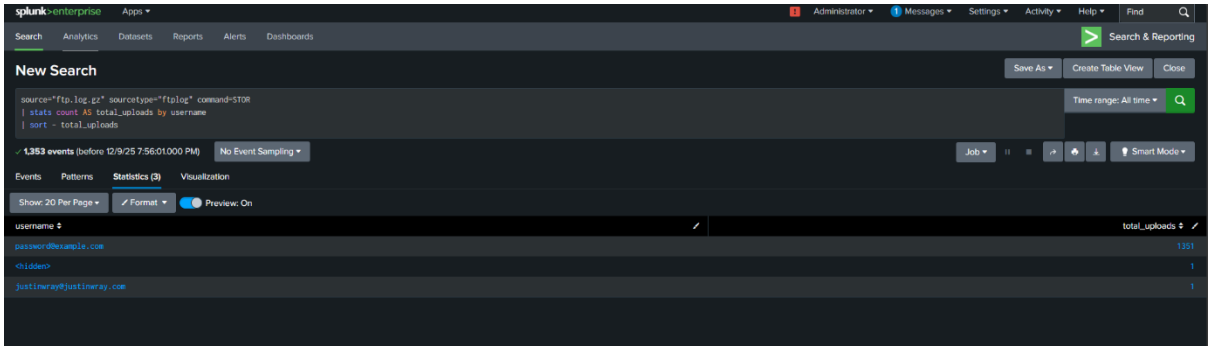
3.2 Most Frequent FTP Users

This analysis shows which usernames are generating the most overall activity. I focused specifically on upload activity (STOR command), as uploads are crucial for file integrity monitoring and assessing file storage usage.

SPL Query (Uploads):

source="ftp.log.gz" sourcetype="ftplog" command="STOR" | stats count AS total_uploads by username | sort - total_uploads

Results (Partial): The user password@example.com accounts for the vast majority of upload activity (1351 uploads). Other users like justinwray@justinwray.com and <hidden> show minimal activity.


splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
Search
Analytics
Datasets
Reports
Alerts
Dashboards
Search & Reporting
New Search
Save As
Create Table View
Close
source="ftp.log.gz" sourcetype="ftplog" command=STOR
| stats count AS total_uploads by username
| sort - total_uploads
Time range: All time
1,353 events (before 12/9/25 7:56:01.000 PM)
No Event Sampling
Job
Smart Mode
Events
Patterns
Statistics (3)
Visualization
Show: 20 Per Page
Format
Preview: On
username
total_uploads
password@example.com
1351
<hidden>
1
justinwray@justinwray.com
1

4.SECURITY ANALYSIS AND FILE INTEGRITY MONITORING

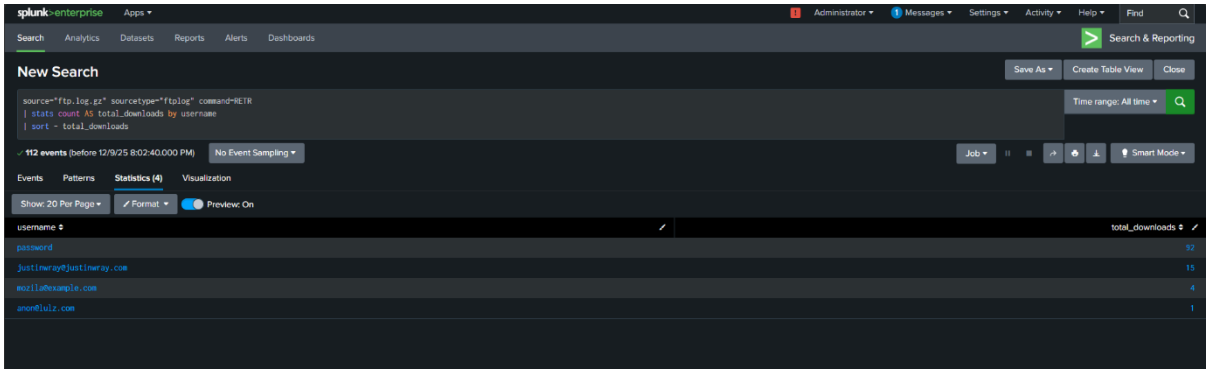
4.1 Top Downloads by User

Monitoring download activity is crucial to detect large-scale data retrieval, which could indicate data theft or unauthorized access.

SPL Query:

source="ftp.log.gz" sourcetype="ftplog" command="RETR" | stats count AS total_downloads by username | sort - total_downloads

Results (Partial): The user password (92 downloads) is the primary downloader, followed by justinwray@justinwray.com (15 downloads). Monitoring this list helps establish a baseline for who should be accessing data.



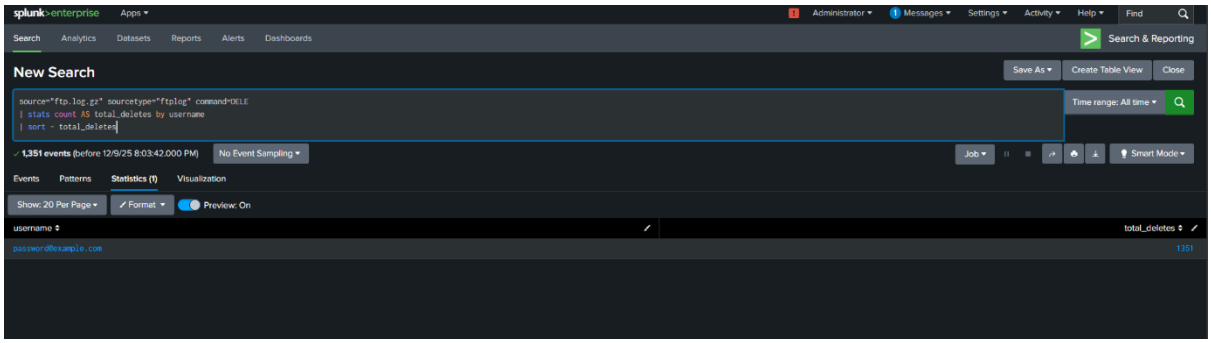
4.2 File Deletion Monitoring (High-Risk Activity)

File deletion is a high-risk activity that must be closely tracked to ensure file integrity.

SPL Query:

source="ftp.log.gz" sourcetype="ftplog" command="DELE" | stats count AS total_deletes by username | sort - total_deletes

Results (Partial): The user password@example.com performed all the detected deletion activities (1351 deletions), making them the sole focus of file integrity checks.



4.3 Monitoring for Failed Commands

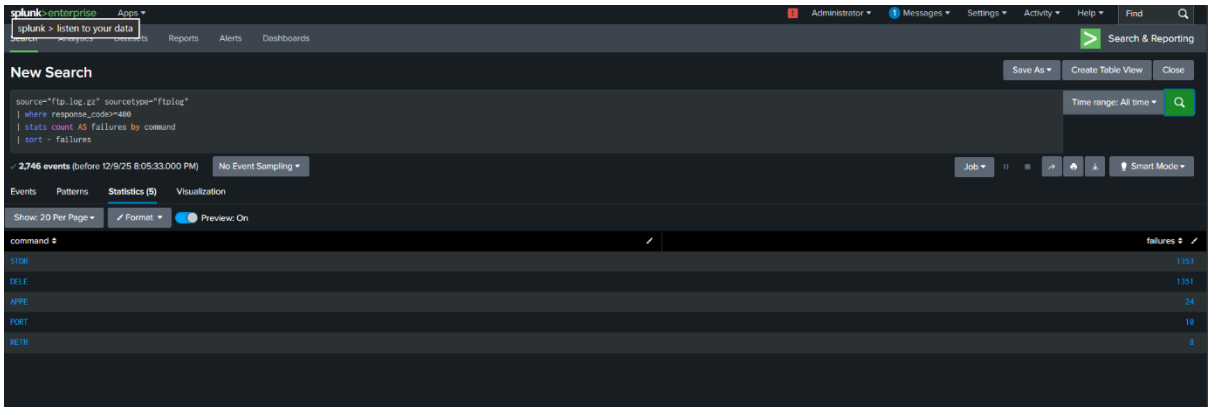
Failed commands (response code 4xx/5xx) often indicate unauthorized access attempts, failed transfers, or permission issues.

SPL Query (Failed Commands):

source="ftp.log.gz" sourcetype="ftplog" | where response_code=400 | stats count AS failures by command | sort – failures

The query filters for a 400-series response_code, which typically signifies client-side errors or authentication/permission failures.

Results and Interpretation: The top failed commands are STOR (1353 failures) and DELE (1351 failures). This is a critical finding:



Massive Failure Rate: The total number of failed uploads (1353) and deletes (1351) exactly matches the total number of attempts by the primary user (password@example.com) from Section 3.2 and 4.2.

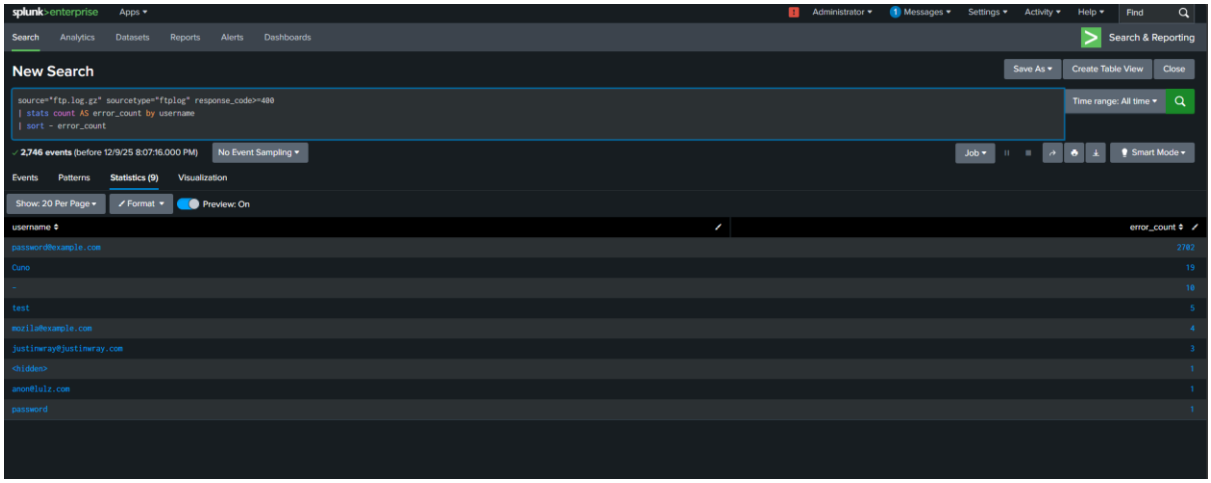
Conclusion: The primary user failed every single upload and delete attempt, likely due to incorrect permissions or a session error. This suggests a major file management problem or a persistent security policy enforcement issue.

4.4 Top Error Generating Users:

Splunk SPL

source="ftp.log.gz" sourcetyp="ftplog" response_code=400 | stats count AS error_count by username | sort - error_count

The user password@example.com accounts for 2702 error events, confirming they are the source of all these failures.


splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
Search
Analytics
Datasets
Reports
Alerts
Dashboards
Search & Reporting
New Search
Save As
Create Table View
Close
source="ftp.log.gz" sourcetype="ftplog" response_code>=400
| stats count AS error_count by username
| sort - error_count
Time range: All time
2,746 events (before 12/9/25 8:07:16.000 PM)
No Event Sampling
Job
Smart Mode
Events
Patterns
Statistics (9)
Visualization
Show: 20 Per Page
Format
Preview: On
username
error_count
password@example.com 2702
Cuno 19
- 10
test 5
mozilla@example.com 4
justinwray@justinwray.com 3
<hidden> 1
anon@lulz.com 1
password 1

5.CONCLUSION

This project successfully analyzed FTP logs to establish operational baselines and uncover a critical security/integrity issue. By using the stats command to track activity by username and command, the project demonstrated effective File Integrity Monitoring. The analysis of failed commands revealed a pervasive issue where the primary user failed every upload and delete attempt, an incident requiring immediate administrative action to correct permissions and prevent data loss or service disruption. This work highlights my ability to use Splunk to manage file integrity and diagnose security incidents across critical network protocols.