SIEM

What is SIEM

- SIEM stands for **Security Information and Event Management.**
- It is security management solution that helps in collecting, parsing and correlating events from various log sources.

Components of Car

Wheels

Engine

Battery

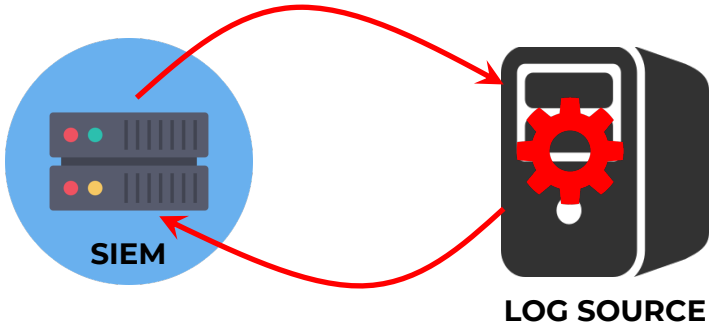
Steering

Components of SIEM - COLLECTOR

- Collect logs from various Log Sources
- 2 Categories of Collection

Log Sources = A server, application or an appliance from where the logs are collected

Pull



Connect to **10.10.10.5**

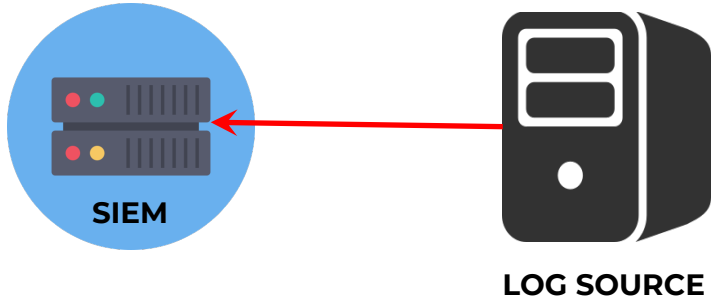
Login Using **siemadmin/p@ssword**

Go to **C:\System32\dns**

Read the file **dns.log** and get the logs

Every **10 minutes**

Push

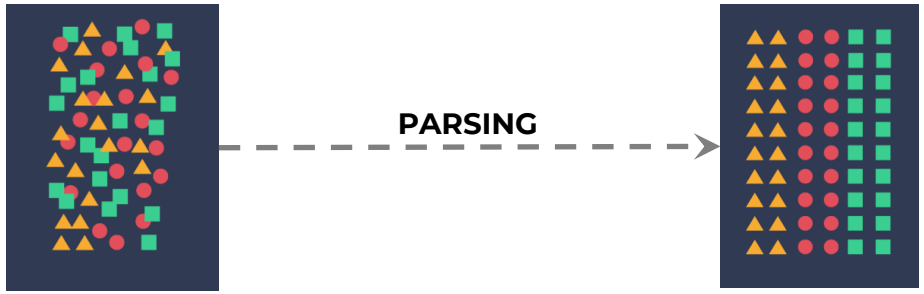


Wait for the logs to come. **Listening for logs**

When a log is generated, send a copy of it to **10.20.20.5** (SIEM IP) On Port **514**

Components of SIEM - PARSER

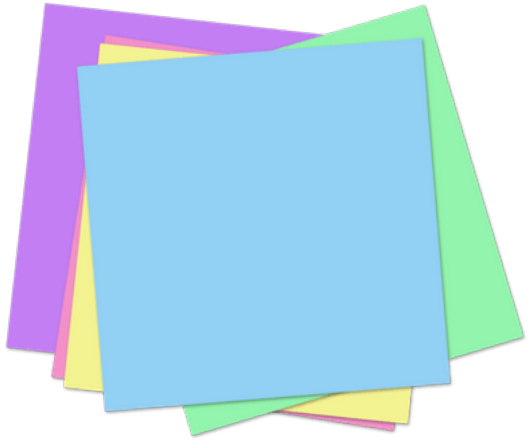
- Process of Converting unstructured data into structured format



- Extract Meta-data like Source IP, Destination IP, Port No., Username, Host name, etc.

Components of SIEM – Why Parse?

- Deliver a seminar
- 500 students attended it
- You have given a small sheet of paper to collect their information



Name
Mobile
Email
College
City
Gender
Age etc.

Not a form, just a blank page

Can we expect everyone to write all the details in the same order?

Can we expect everyone to write all the information?

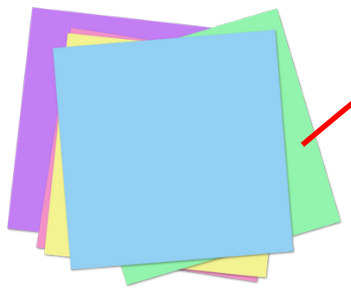
- Data of 500 students
- Try to derive information out of the data.

How many students belong to Bengaluru?	**30 minutes**
How many students are between age 18 & 20	**30 minutes**

It takes lot of time to pull information from unstructured data

Components of SIEM - Why Parse?

- Take 30 minutes and enter the details of all the students in a spreadsheet.



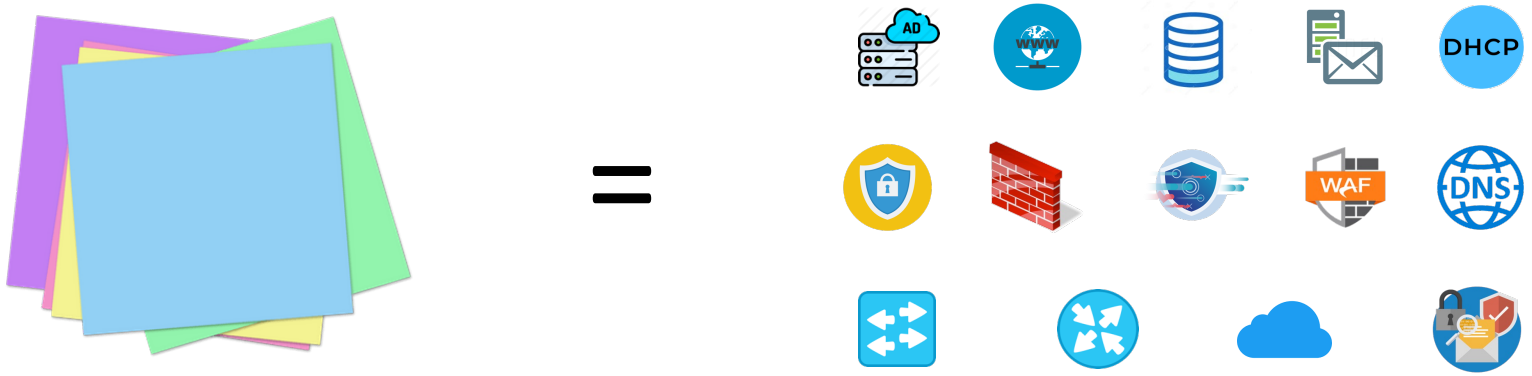
Name	Mobile	Email	College	City	Gender	.....	Age

How many students belong to Bengaluru? **10 seconds**

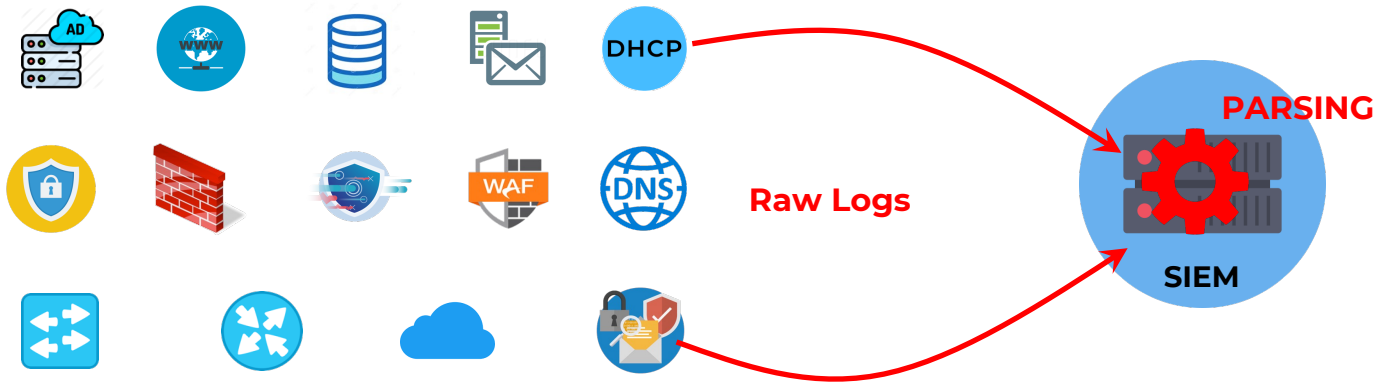
How many students are between age 18 & 20 **10 seconds**

PARSING

Components of SIEM - Why Parse?



No standardization



Components of SIEM - PARSER

- Extract Meta-data like Source IP, Destination IP, Port No., Username, Host name, etc.

Raw Log Jan 1 20:28:02 knight sshd[20336]: Failed password for root from 218.49.183.17 port 49869 ssh2

Parsed Event

Date	Time	Event	Username	Source IP	Source Port	Protocol	Hostname
Jan 1	20:28:02	Failed Password	root	218.49183.17	49869	ssh	knight

Each of these columns are called fields. **Fields of a log** **Metadata**

During parsing, SIEM also does **NORMALIZATION** = Bringing all type of logs in one standard format

Components of SIEM - STORAGE

- SIEM collects millions of logs from thousands of logs sources.
- Needs to store the collected logs.

Huge space
Usually in Terabytes (TBs)

Required for Forensics and Compliance purposes

- Few SIEM will store the Raw Logs and the Parsed Events separately
- Few SIEM store the Raw Log and the associated Metadata (extracted fields) together

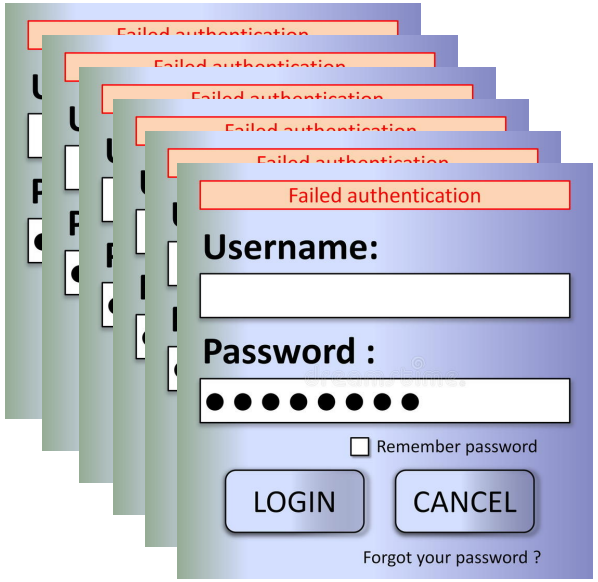
- Before the logs are written on to the database, they are **indexed**.
- Every vendor have their own way (algorithm) of indexing. This is their secrete sauce. This is what makes one SIEM faster than the other

Components of SIEM - CORRELATION

- This is the intelligent part of the SIEM
  - Without Correlation, SIEM will just be a log collection tool.
  - The correlation component makes the SIEM a solution capable of identifying attacks

- Correlation = Set of conditions that indicates suspicious activity

Example of Suspicious Activity

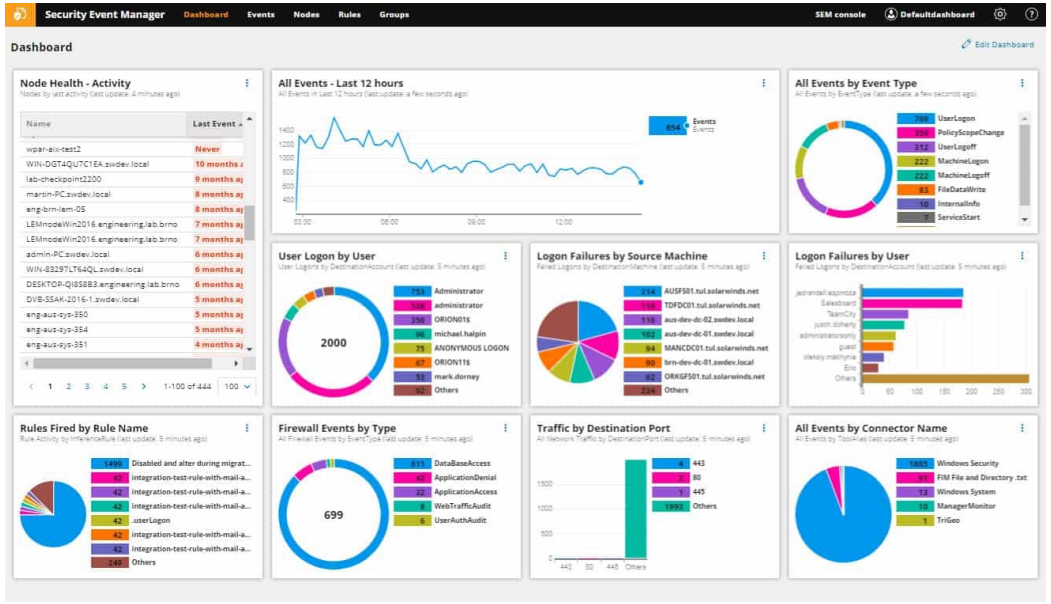


- In the above example, the conditions are:
  - Failed login event
  - Same user
  - 20 times
  - In 1 minute

Possible Brute-force Attack

Components of SIEM - MANAGER

- Manage the overall configuration of the SIEM
- Usually the frontend of SIEM (User Interface)



- Perform search and create reports and dashboards

Components of SIEM - AGGREGATION

- Combing similar logs that are generated over a period of time
  - Similar = Same attributes like Source IP, Username, Destination IP, etc.
  - Time period could be 10 seconds or 1 minutes or 10 minutes.

Increase the Performance

Decreasing the No. of Logs Indexed

Saves Disk Space

- SIEM will still maintain a record of how many times the event occurred (Event Count).
- Time of First Event and Time of Last Event

Date	First Time	Last Time	Event Count	Event	Username	Source IP	Source Port	Protocol	Hostname
Jan 1	20:28:02	20:32:16	36	Failed Password	root	218.49183.17	49869	ssh	knight

Components of SIEM

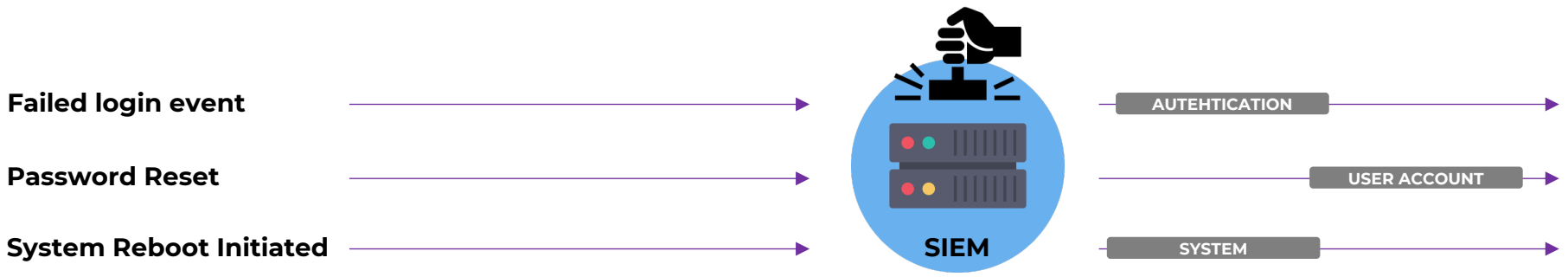
Aggregation – Alternate definition

Log aggregation is the process of collecting logs from multiple computing systems, parsing them and extracting structured data, and putting them together in a format that is easily searchable and explorable by modern data tools.

Components of SIEM - CATEGORIZATION

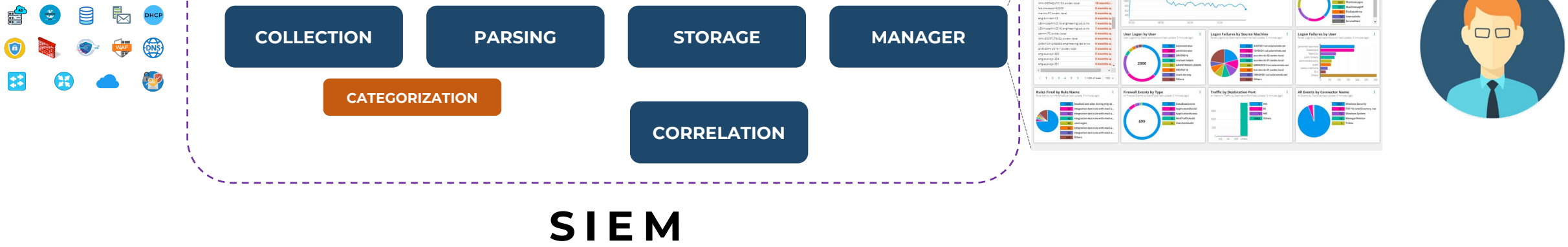
- Grouping of logs based on the type of event.

AUTEHTICATION | SYSTEM | USER ACCOUNT | MALWARE | CONFIG



Date	First Time	Last Time	Event Count	Categorization	Event	Username	Source IP	Source Port	Protocol	Hostname
Jan 1	20:28:02	20:32:16	36	Authentication	Failed Password	root	218.49183.17	49869	ssh	knight

Components of SIEM



Capabilities of a Next Generation SIEM

Machine learning	• Detecting threats based on machine learning algorithm. • To reduce dependency on Correlation Rules, which are typically written by admins
UEBA	• User and Entity Behavior Analytics • User and Asset context • Machine Learning to detect any anomalies in behavior of a machine or a user
Support Threat Hunting	• Detection of threats cannot be left to Correlation rules and Machine Learning • Most SOC teams now employ Threat Hunting practice to proactively detect threats • A next-gen SIEM should support Threat Hunting
SOAR	• Security Orchestration, Automation and Response • Automatically respond to threats (in near real-time)