AT&T Cybersecurity

AT&T Business

# Beginner's guide: Security information and event management (SIEM)

# A rose by any other name
## LMS, SLM/SEM, SIM, SEM, SEC

Although the industry has settled on the term 'SIEM' as the catch-all term for this type of security software, it evolved from several different, but complementary, technologies that came before it.

*continued on next page*



This document is intended to include general information for beginners learning about security information and event management (SIEM). Use of names of third party companies in the document are for informational purposes only and do not constitute any endorsement by AT&T Cybersecurity.

# A rose by any other name
# LMS, SLM/SEM, SIM, SEM, SEC

*continued*

- **LMS** "Log Management System" – A system that collects and stores log files (from operating systems, applications, and more) from multiple hosts and systems into a single location, allowing centralized access to logs instead of accessing them from each system individually.
- **SLM /SEM** "Security Log/Event Management" – An LMS, but marketed towards security analysts instead of system administrators. SEM is about highlighting log entries more significant to security.
- **SIM** "Security Information Management" – An asset management system, but with features to incorporate security information. Hosts may have vulnerability reports listed in their summaries, and intrusion detection and antivirus alerts may be shown mapped to the systems involved.
- **SEC** "Security Event Correlation" – To a particular piece of software, 3 failed login attempts to the same user account from 3 different clients, are just 3 lines in their logfile. To an analyst, that is a peculiar sequence of events worthy of investigation, and log correlation (looking for patterns in log files) is a way to raise alerts when these things happen.
- **SIEM** "Security Information and Event Management" – SIEM is the "all of the above" option. As the above technologies merged into single products, SIEM became the generalized term for managing information generated from security controls and infrastructure. We'll use the term SIEM for the rest of this presentation.

# Q: What's in the logs?!

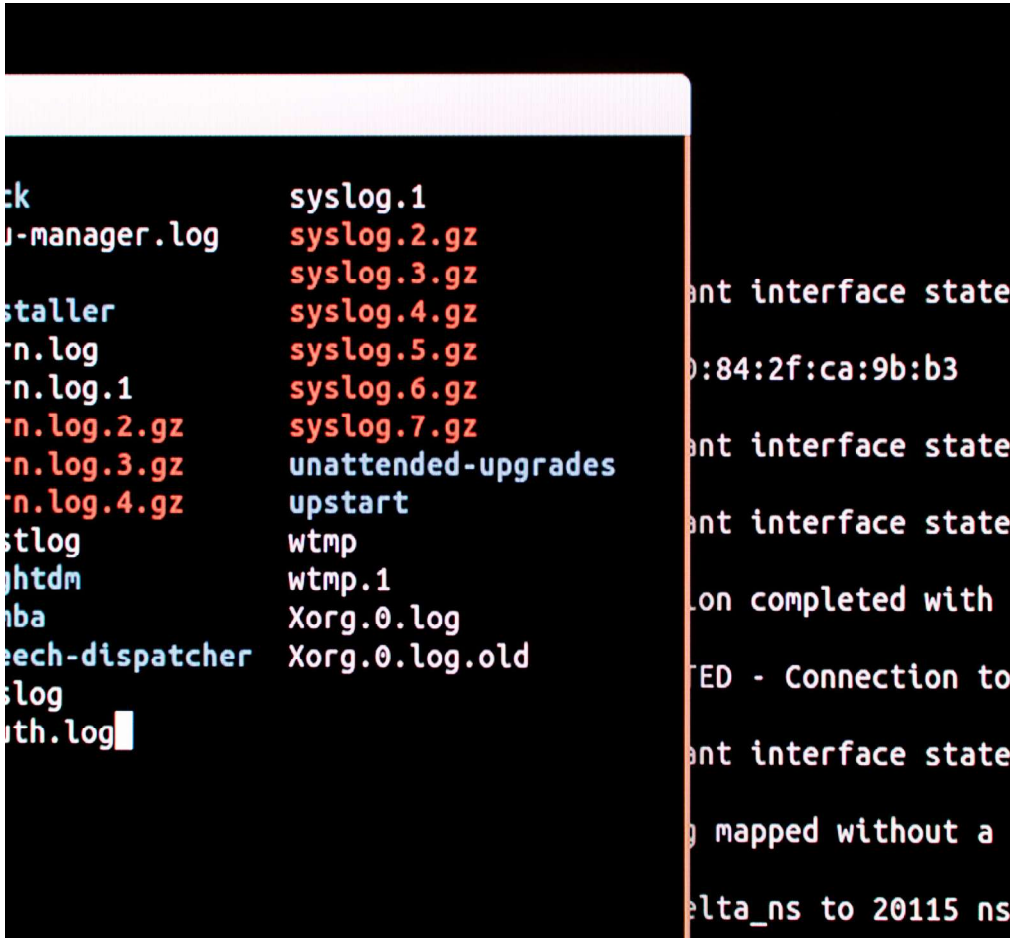
**A:** Logs contain the information you need to answer **"Who's attacking us today?"** and **"How did they get access to our systems?"**

We may think of security controls as containing all the information we need to be protected, but they often contain only the things they have detected—there is no "before and after the event" context within them.

This context is usually vital to separate the false positive from true detection or the actual attack from a merely misconfigured system.

Successful system attacks rarely look like real attacks, except in hindsight. If this were not the case, we could automate all security defenses without ever needing to employ human analysts.

Attackers will try to remove and falsify log entries to cover their tracks. Having a trusted source of log information is vital to record keeping as to issues that may arise relating to system misuse.






AT&T Business

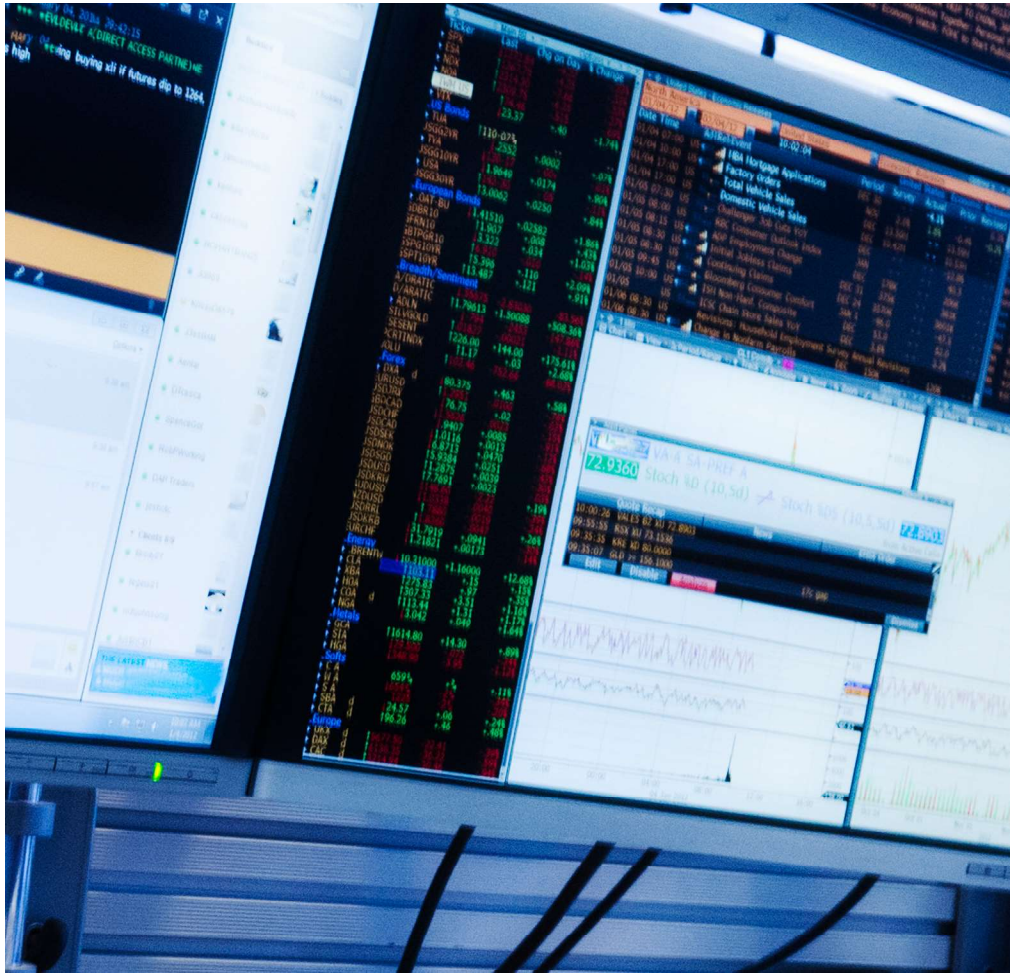
# The blind men and the **security information elephant**

SIEM is about looking at what's happening on your network through a larger lens than can be provided via any one security control or information source.

- Your intrusion detection only understands packets, protocols, and IP addresses.
- Your endpoint security sees files, usernames, and hosts.
- Your service logs show user logins, service activity, and configuration changes.
- Your asset management system sees apps, business processes, and owners.

None of these by themselves can tell you what is happening in terms of securing the continuity of your business processes and your business …

**But together, they can.**



# SIEM

# A single view of your IT security

SIEM is essentially nothing more than a management layer above your existing systems and security controls.

It connects and unifies the information contained in your existing systems, allowing them to be analyzed and cross-referenced from a single interface.

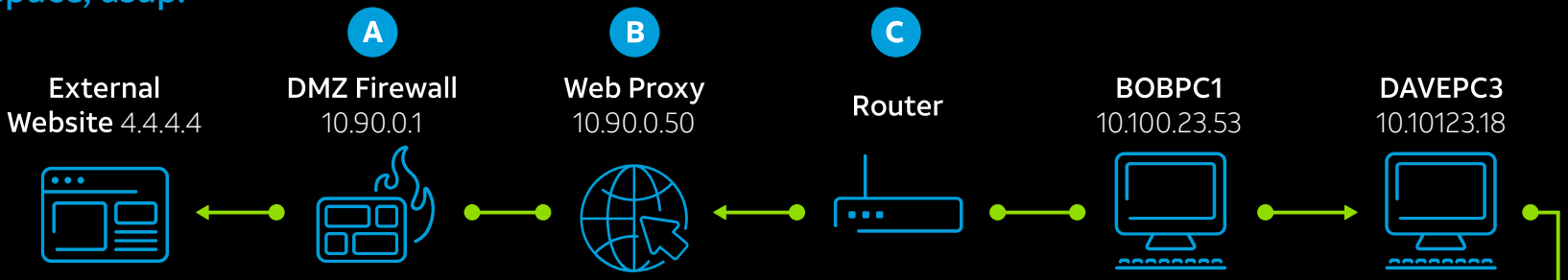
SIEM is a perfect example of the 'garbage in, garbage out' principle of computing:

**SIEM is only as useful as the information you put into it.**

**The more valid the information depicting your network, systems, and behavior the SIEM has, the more proficient it will be helping you make effective detections, analyses, and responses in your security operations.**

# Example of an attack

Bob's machine was compromised by asbss.exe, which originated from a malicious website. This malware then used Bob's account to try and infect another computer, DAVEPC3, and antivirus caught it. But Bob's machine "BOBPC1" is likely still compromised. **We should block the malicious domain and sanitize Bob's workspace, asap.**




**A** `Connection to TCP port 80 - src:10.90.0.50 dst: 4.4.4.4 state: ACCEPTED`

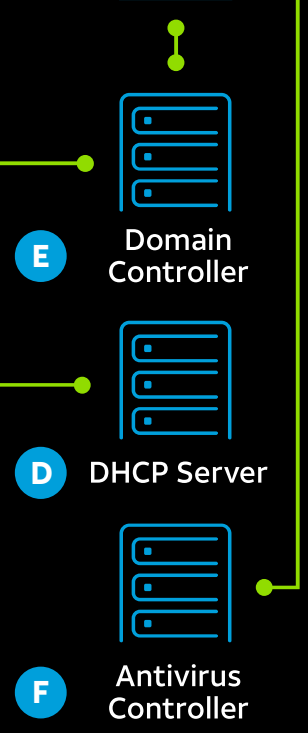
**B** `HTTP Client GET - http://somebadwebsite.org/878732/asbss.exe`

**C** `%SEC-6-IPACCESSLOGP: list ACL-IPv4-E0/0-IN permitted tcp 10.100.23.53(38231) > 10.90.0.50(3129), 1 packet`

**D** `Lease for 10.100.23.53 Assigned to BOBPC1 - MAC:AE:00:AE:10:F8:D6`

**E** `Authentication Package: Microsoft_Authentication_Package_V1_0 Logon Account: BRoberts Source Workstation: BOBPC1 Error Code: 0x00000064`

**F** `Client: DAVEPC3 - Successfully Removed - C:\Windows\Temp\asbss.dll - Reason: Win32/RatProxyDLL18 105`






AT&T Business

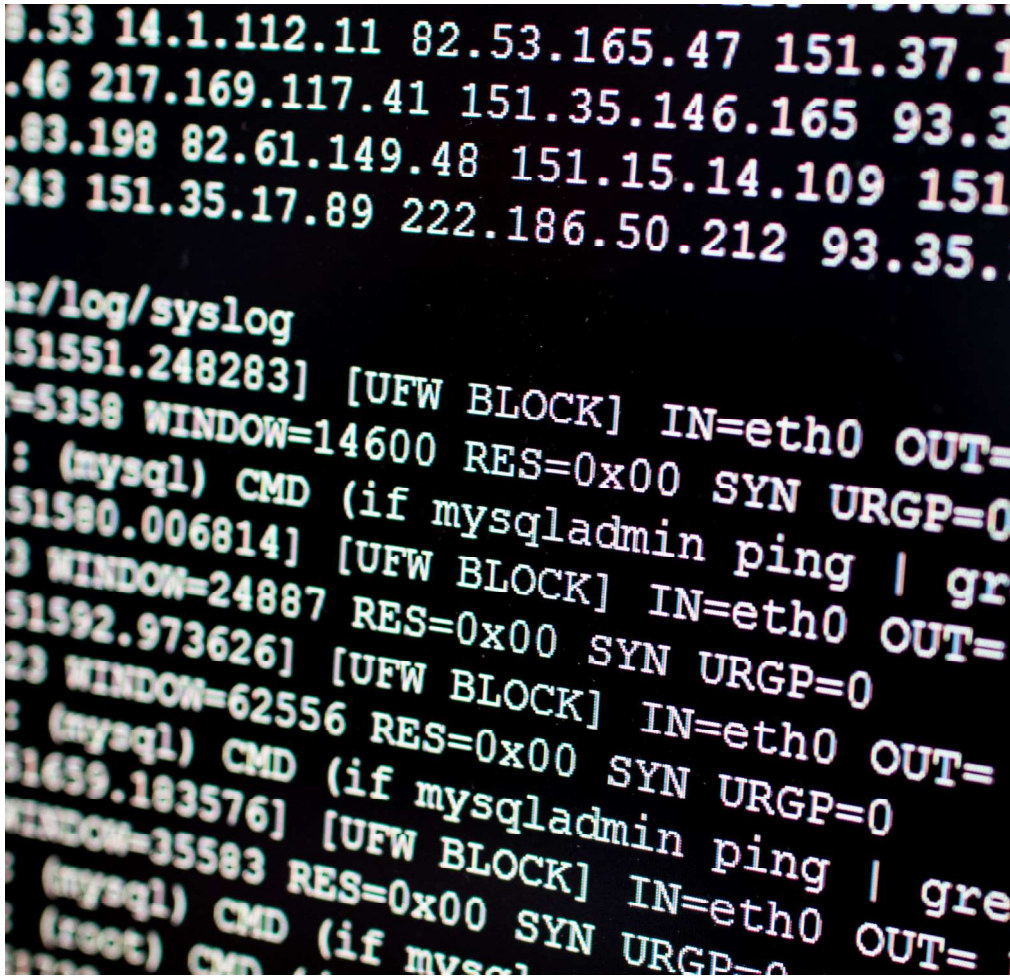
# The importance of context

Log collection is the heart and soul of a SIEM. The more log sources that send logs to the SIEM, the more that can be accomplished with the SIEM.

But logs on their own rarely contain the information needed to understand their contents within the context of your business. For example, with only the logs, all an analyst sees is “connection from host A to host B.”

Yet to the system administrator, this becomes “daily activity transfer from point of sales to accounts receivable.” The security analyst needs this information to make reasoned assessment of any security alert involving this connection.

Security analysts’ limited bandwidth can make it difficult to be familiar with every system your IT operation depends on, but the true value of logs is in correlation to get actionable information. More on correlation later.






AT&T Business

# SIEM recipes
# Ingredients for a proper SIEM deployment

## Logs and alerts

### Security controls

- Intrusion detection
- Endpoint security (antivirus, etc.)
- Data loss prevention
- VPN concentrators
- Web filters
- Honeypots
- Firewalls

### Infrastructure

- Routers
- Switches
- Domain controllers
- Wireless access points
- Application servers
- Databases
- Intranet applications

## Knowledge

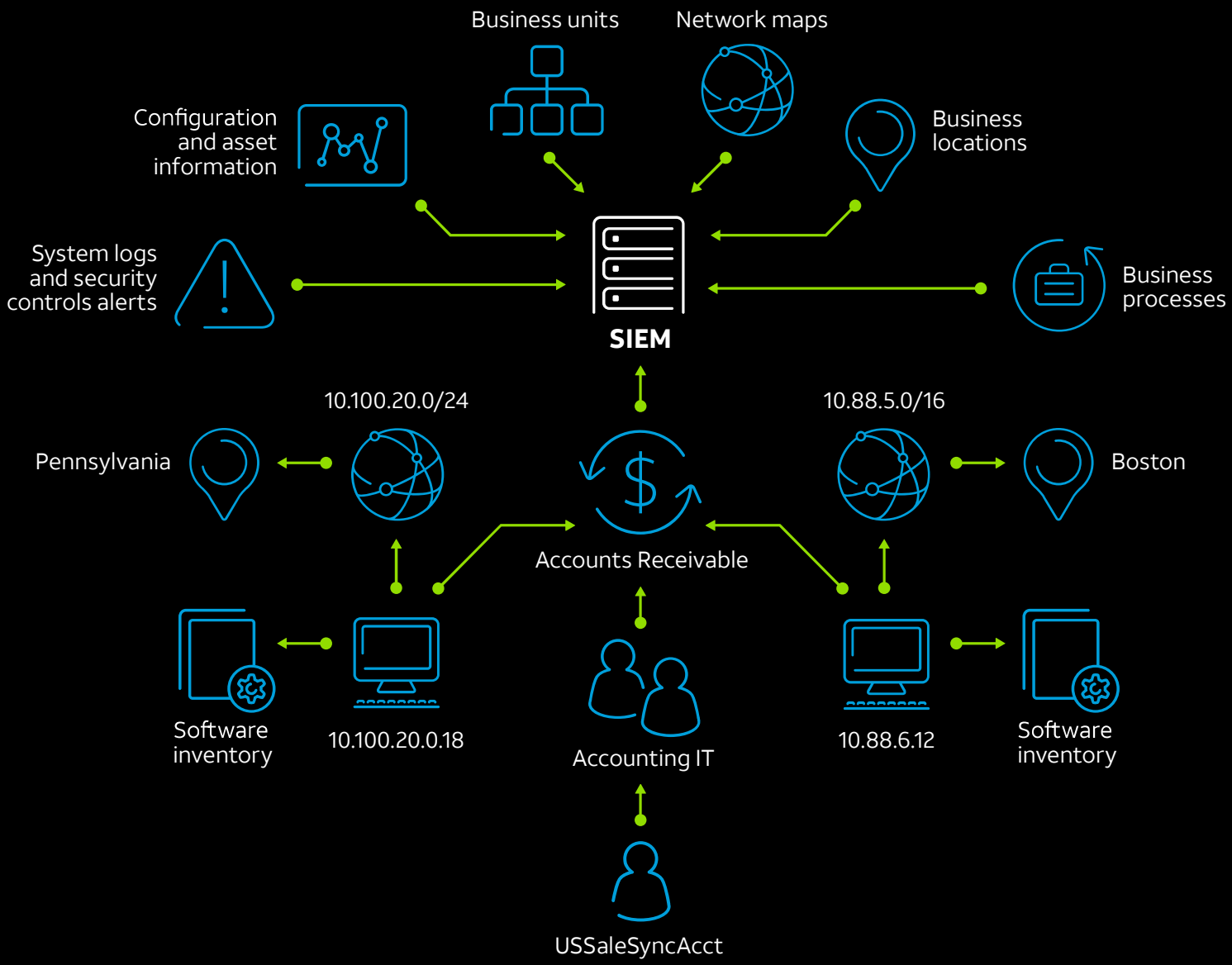
### Infrastructure information

- Configuration
- Locations
- Owners
- Network maps
- Vulnerability reports
- Software inventory

### Business information

- Business process mappings
- Points of contact
- Partner information

# How a log file is generated in your network




10.100.20.18 Initiated Database Copy using credentials USSalesSyncAcct to remote Host 10.88.6.12 - Status Code 0x44F8

# Behold:

## The power of correlation

Correlation is the process of matching events from different systems (hosts, network devices, security controls…anything that sends logs to the SIEM).

Events from different sources can be combined and compared against each other to identify patterns of behavior invisible to individual devices.

They can also be matched against the information specific to your business.

Correlation allows you to automate detection for the things that **should not** occur on your network.

# The beauty of **log correlation**

Log correlation is the difference between:

```
"14:10 7/4/20110 User BRoberts
Successful Auth to 10.100.52.105 from
10.10.8.22"
```

and ...

```
"An account belonging to marketing
connected to an engineering system
from an office desktop, on a day when
nobody should be in the office"
```

# Slow cook for 8 hours
# **Serve to hungry analysts…**

Your network generates vast amounts of log data–a Fortune 500 enterprise's infrastructure can generate 10 terabytes of plain-text log data per month without breaking a sweat.

You can't hire enough people to read every line of those logs looking for bad stuff. We're serious, don't even try this. Even if you succeeded, you'd be so bored you'd never actually spot anything even if it was right in front of your face. Which it would be.

Log correlation lets you locate the interesting places in your logs, and that's where analysts start investigating. And they're going to find pieces of information that lead to other pieces of information as the trail of evidence warms up.

Being able to search through the rest of those logs for that one thing they suspect resides there is one of the other key functions of a SIEM.

It's a good thing a SIEM is fundamentally a…

# . . . giant database of logs.

It would be amazingly useful if every operating system and every application in the world recorded their log events in the same format. They don't. Most logs are written to be readable by humans, not computers.

This makes using regular search tools over logs from different source more difficult.

Want to see? These two logs say the same thing to a human being, but are very different from the machine's point of view.

```
"User Broberts successfully authenticated to 10.100.52.105 from client 10.10.8.22"
"100.100.52.105 New client connection 10.10.8.22 on account: Broberts: Success"
```

Long story short: we need to break down every known log message out there into a normalized format.

```
"User [USERNAME] [STATUS] authenticated to [DESTIP] from client [SOURCEIP]"
"100.100.52.105 New client connection 10.10.8.22 on account: Broberts: Success"
```

So when you see a SIEM product that talks about "how many devices it supports" it's talking about how many devices it can parse the logs from.

# Searches, pivoting, and cross-correlation

Breaking log entries down into their components, normalizing them, is what allows us to search across logs from multiple devices and correlate events between them.

Once we've normalized logs into a database table, we can do database-style searches, such as:

```
Show [all logs] from [all devices] from the [last two weeks], where the
[username] is [Broberts]
```

This is what allows us to do automated correlation as well as: matching fields between log events, across time periods, and across device types.

```
If a single host fails to log in to three separate servers using the same
credentials, within a 6-second time window, raise an alert
```

Just as with any database, event normalization allows the creation of report summarizations of our log information.

```
What user accounts have accessed the highest number of distinct hosts in
the last month?
```

```
What subnet generates the highest number of failed login attempts per day,
averaged out over 6 months?
```

# But wait, **there's more!**

Let's review:

- A SIEM is a recording device for the systems that form your information infrastructure.
- A SIEM allows you to give analysts access to information from these systems without giving them access to the systems themselves.
- Event correlation allows you to encode security knowledge into automated searches across events and asset information, to alert on things happening within your infrastructure. This creates a starting point for human analysis into the sea of log data.

But to keep up with today's threat landscape, you need more that just a SIEM. **You need relevant data, a unified approach, and integrated threat intelligence to truly get a holistic view of your security posture.**