Workload vs. Ingest-Based Pricing: Which Is Right for Your SOC?

DEVO

The SIEM marketplace is full of choices. There are almost as many SIEM pricing models as there are SIEM vendors. But most pricing models fall into one of two groups: workload-based (aka CPU) pricing models and ingest-based pricing models. Each main group has many variations, but they typically break down into these two categories. Workload-based pricing is calculated by the compute power you need to power your searches, dashboards, alerts, etc. Ingest-based pricing is based on the total amount of data ingested into the SIEM platform.

Choosing the pricing model that's right for your organization's SOC is almost as hard as choosing the right vendor. This guide will help you decide which of the two major pricing models is the best fit for your Security Operations Center (SOC). If you need additional resources to help with your decision-making process, please read our Buyer's Guide to Next-Gen SIEM.

Let's begin with a brief overview of the two models and outline some of the key factors for consideration. Then we'll identify key decision criteria for selecting the pricing model that best fits your needs.

As with choosing a SIEM, there isn't a single best choice; it depends on the needs and circumstances of your SOC. Let's start by examining the strengths and limitations of both options.

STRENGTHS AND LIMITATIONS OF WORKLOAD-BASED PRICING

Strengths of Workload-based Pricing:

1. Often has a lower TCO per TB than ingest-based pricing
2. You don't need to know data volume to predict future pricing
3. Good for SOCs with few analysts running a small number of queries

Limitations of Workload-based Pricing:

1. Requires very accurate prediction of required CPU
2. Often results in slow performance at the busiest, most critical times
3. Poor choice for larger SOCs with many analysts running a high volume of queries

Workload-Based Pricing

Workload-based pricing is calculated on the compute power needed to search your data, rather than the total amount of data you ingest. This can benefit large organizations that ingest huge volumes of data or organizations with difficulty predicting data volumes. The obvious benefit to workload-based pricing is that it can result in a "lower cost per TB." Proponents of workload-based pricing say it delivers higher value because not all data is created equal; some data is searched much more often than other data and you only pay for the searches, not the ingest.

The tradeoff with workload-based pricing is it requires accuracy to provision the compute power necessary to support everything you need your SIEM to do. If you under-provision your CPU, the SIEM will perform poorly across the board. That means slower searches, slower performance for dashboards, and slower performance for alerts. As a result, it also means delays in detections, investigations and resolutions of incidents and threats. The point of having a SIEM is to detect and respond to threats quickly — a slow SIEM is almost as bad as no SIEM at all.

KEY QUESTIONS TO ANSWER

The first thing to ask yourself when planning CPU capacity is: How many users will there be and how many searches will they run? A general rule is to provision one CPU per user per query in a given interval.

But the CPU load isn't just about the searches analysts initiate. Alerts also count as searches since an alert is simply a query with a notification policy based on the result of the query. The CPU load must also include the total number of alerts that will be running. Otherwise, you might not have sufficient CPU capacity for the queries your team runs due to all of the configured alerts.

But alerts aren't the only types of queries running in the background that consume CPU. Dashboards also consume a considerable amount of CPU. A dashboard is just a query that displays its results in a graph. Just like alerts, dashboards can cause CPU contention with user-run queries and result in slow response times for both. You also need to consider the total number of dashboards you'll have. Finally, enrichments (also called lookups) are queries that also consume CPU. They populate tables with data from other tables to add context to the data and enable SOC analysts to make accurate decisions. Take these different types of "background queries" into account when scoping out your total CPU needs.

CRUNCHING THE NUMBERS

If you aren't using workload-based pricing in your current SIEM, it's critical to accurately scope out the amount of CPU you'll need. Let's say you have 10 users. Multiply 10 by the total queries they run in a typical 5-to-10-minute interval. If 10 people each run 10 queries in a 10-minute interval, then 100 CPU is a good place to start. This assumes the number of queries is relatively constant — that's a big assumption — but more on that later.

Next, add up the dashboards used daily — don't count those used infrequently. Let's call it 20 dashboards. Add that to your previous number and you need 120 CPUs for user queries and dashboards. And don't forget to add the total number of alerts as well. This is an often-overlooked number, and it results in either slow query performance or alerts with extremely long lag times between the event and the notification. Alerts must be as real time as possible, so don't under-provision the CPU they require. It's important to add up all the alerts in a typical schedule (i.e., business hours on weekdays). If it's 100 alerts, for instance, include it in your subtotal, which now has reached 220 CPU.

One more thing. Remember to add enrichments as well. These can be queries related to threat intelligence, 3rd-party API integrations, etc. Let's call it another 50 queries that run *every hour*. That's a total of 270 CPUs, which is a good estimate for everyday usage.

Keep in mind that when an incident occurs, it's an all-hands-on-deck situation with users running many more queries than usual. It's wise to over-provision by 10% to 20% to account for those "fire drill" high activity scenarios.

Most SIEMs will report on the number of queries run per user as *well* as the time it takes each query to execute. Since the CPU won't be available for a new query until it finishes its current query, you also need to account for the number of long-running queries in your SOC.

Counting the total number of queries executed in an hour during your busiest times is a good way to determine if your scoping exercise was accurate or requires adjusting.

Your total compute power can grow quickly. Again, the danger of under-scoping your compute power will result in slow SIEM performance across the board. That slow performance will make it difficult to get prompt answers.

In an all-hands-on-deck scenario where your analysts are all banging away in response to an incident, slow performance will adversely impact the SOC's ability to respond quickly and decisively.

STRENGTHS AND LIMITATIONS OF INGEST-BASED PRICING

Strengths of ingest-based pricing:

1. Unlimited users and unlimited queries
2. Fast query performance
3. Good for larger SOCs with many use cases

Limitations of Ingest-based Pricing:

1. Hard to predict data volume over time
2. Unexpected spikes in data volume can increase costs
3. Not good for environments with unpredictable data volumes

Ingest-Based Pricing

Ingest-based pricing has been the de-facto SIEM model for a long time. Some vendors base ingest on the number of events per second, the number of devices monitored, or just total data volume. But these are variations on the total amount of data ingested.

INSIDE THE NUMBERS OF INGEST-BASED PRICING

As the attack surface has grown due to the proliferation of mobile devices, IoT, and hybrid cloud environments, there has been a parallel explosion in the volume of activity a SOC must watch and a corresponding increase in total data volume. This has driven up the cost of ingest-based models dramatically.

Predicting long-term data volume can be a challenge. Some SOC teams simply don't know how much data they expect to bring in over the next three weeks, let alone the next year. They certainly don't want to be surprised with a large bill due to a sudden increase in data volume. Although most ingest-based SIEMs have ways to mitigate a flood of data using agents or collectors that filter out and/or truncate unwanted data, the unpredictability of data volume can be a big concern.

On the other hand, most ingest-based SIEMs automatically scale up their CPU based on the data volume ingested, which is a big benefit. This ensures users always have adequate search capabilities regardless of the volume of data. Maintaining fast search performance during large spikes in data volume should be a fundamental asset of ingest-based SIEM vendors. Ingest-based SOCs also never need to worry about how many users are running searches or if a few long-running searches are tying up resources. For larger SOCs with many analysts, the availability of "unlimited users/unlimited queries" can be very attractive.

PICKING THE RIGHT MODEL FOR YOUR ORGANIZATION

Now that we've looked at both models, let's review some qualifying questions to help you determine the right model for your SOC. Just as signs help you navigate an unfamiliar road, here are some signposts you can use to arrive at the best pricing model for your needs.

Data Volume Consistency

Data volume consistency is a major consideration. Most organizations have growing data volumes, which is not surprising since we are living in the data age! For some organizations, data growth is fairly consistent. For others, it's explosive and unpredictable. If you have a SIEM, you should be able to go back and look at data volume over the last several months to spot a consistent pattern. If it's easy to identify the trend of data volume over time, then ingest-based pricing should be easy to predict for your organization. Even if volume is increasing, it should be increasing at a predictable rate. If, however, you can't spot that volume trend ingest-based pricing may be risky because you won't be able to accurately predict the cost.

If you don't currently use a SIEM and don't know how much data you have, look at the total attack surface you need to secure. A good way to quantify this is to determine the total number of devices that will be sending data to the SIEM. Include both on-premises and cloud-based sources. Is the number of devices expected to remain relatively constant over time? Even if you expect the number of devices to grow, will it grow predictably? If the answer is yes to both, then ingest-based costs will be predictable. If not, then ingest-based pricing could be risky and workload-based pricing may be the better option.

Number of Users

The number of SIEM users is another important consideration. More users will generate more SIEM activity, and you'll need more CPU to support it. There isn't a hard-and-fast number that will point you to workload-based or ingest-based pricing, but a general rule of thumb is 10 concurrent users. If your SOC has fewer than 10 people concurrently using the SIEM, workload-based pricing might be a good fit because you don't have a large group simultaneously competing for CPU resources. If you have more than 10 users or expect to exceed 10 within six months, ingest-based pricing may be the better option for your SOC.

Number of Queries, Alerts and Dashboards

This metric closely aligns with the number of users. Even a small number of concurrent users — if they are "power users" who generate a large number of queries, alerts, and dashboards — may mean workload-based pricing is best for your organization. Remember, every query, every alert, and every dashboard requires compute power to run. Your compute needs grow exponentially for each of these uses. If most users have at least an intermediate skill level and are capable of creating several queries, alerts and dashboards, then ingest-based pricing will deliver the highest query performance.

Conversely, if the majority of your team are "lightweight users" who only use the dashboards created for them, and repeatedly use a small number of queries, then workload-based pricing will work well. If you have just a few power users who run the complex queries, create efficient dashboards for others to view, and generate a small number of alerts, then it's unlikely the less sophisticated users will tax your CPU.

Use Cases

How many use cases will you have for your SIEM? Is its primary use going to be for alerting and incident response, or will you also use it heavily for threat hunting and investigations? The more use cases you have, the more CPU you'll need to support them.

Workload-based pricing is good for SIEMs with a small, tightly focused number of use cases. If you will primarily use the SIEM for identifying known threats and incident response, then workload-based pricing works well. But if you also plan to use the SIEM for threat hunting and investigations, ingest-based pricing will work better for you.

Threat hunting typically requires more advanced queries, correlation of results from multiple data sources and tables, and going back further in time. Advanced queries usually take longer to execute. Do you want the SIEM to perform advanced, CPU-intensive tasks such as memory dump analysis and packet trace analysis and correlate the results with log data? In an environment where CPU is limited, these tasks can cause slow query performance for all users and frustrate SOC analysts. For this reason, ingest-based pricing is better for SOCs with more sophisticated and varied use cases.

Dedicated Administrators

Another indicator is if you have skilled and experienced SIEM administrators. If veteran admins are keeping an eye on your SIEM, they can help mitigate the risks associated with workload-based pricing. One of the major problems with workload-based pricing is having a user create queries and dashboards that consume large amounts of precious CPU. This has a ripple effect on query performance for all other users. But dedicated, experienced administrators can look for resource-intensive queries, stop them, or help tune them to run more efficiently.

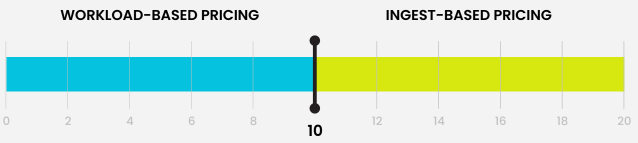
Conversely, if you don't have at least a few dedicated administrators to police your users and their queries, then the unlimited CPU capabilities of an ingest-based pricing model will be better for your SOC. In that scenario, users are free to run as many queries as they need and the SIEM will scale up CPU as required.

Reading the Signs

The following infographics will help you understand if your particular situation and needs point to ingest-based or workload-based pricing.

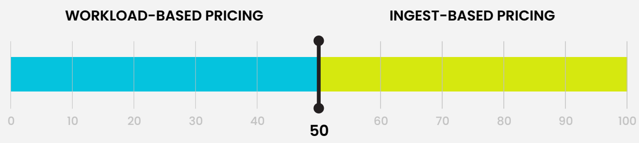
SCORE YOUR DATA VOLUME PREDICTABILITY

Score Your Number of Users



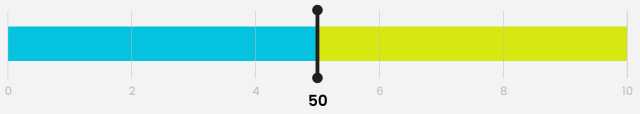
Fewer than 10 users is best for workload-based pricing. A larger team is better served by ingest-based pricing.

Score Your Total Queries, Alerts, Dashboards



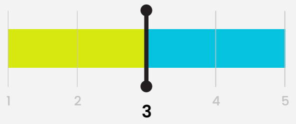
A total of less than 50 is good for workload-based pricing. Higher than 50 works better with ingest-based pricing.

Score Your Total Use Cases



Fewer than 50 use cases is best for workload based pricing. Ingest based pricing is better for larger number of use cases.

Score Your Dedicated Administrators



Ingest based pricing is better where there are fewer than 3 full time SIEM administrators. Workload base pricing works better with 3 or more FTE's to police CPU contention.

Want to learn more?
Contact Devo to learn how ingestion-based pricing will strengthen your organization's security posture — without breaking the bank.

Devo
255 Main Street
Suite 702
Cambridge, MA 02142

Devo is the only cloud-native logging and security analytics platform that releases the full potential of your data to empower bold, confident action. With unrivaled scale to collect all of your data without compromise, speed to give you immediate access and answers, and clarity to focus on the signals that matter most, Devo is your ally in protecting your organization today and tomorrow. Devo is headquartered in Cambridge, Mass. Learn more at **www.devo.com**.