September, 2024

# Understanding Entity-Based Asset Counting

The 5.3.0 release introduces a new *entity-based asset counting model*. This model has already been in use on managed deployments (those with an "s" in their version number; for example, 5.2.0s), but is now the default model for new Stellar Cyber platforms in 5.3.0. Existing platforms can also convert to the new entity-based model.

## Summary

Entity-based asset counting introduces several key improvements over the classic asset counting model by simplifying counting and leveraging only high-confidence sources for attributing entities against the subscribed license count.

This simplified model produces more consistent results and also reduces the system overhead required to determine entity counts, allowing the entity-based asset counting model to scale beyond what could previously be support with the classic model.

## What Gets Counted?

At a high level, the entity-based model counts *devices* and *users* as entities:

- **Devices** – any private IP address observed in the sample window (either from a trusted data source or inferred from observed traffic; see below for details).
- **Users** – any logical account with a valid email address as an identifier observed in the sample window, including accounts assigned to non-human actors such as service accounts.

## How Are Devices Identified?

Devices can be identified either from trusted data sources or inferred from certain types of network data. The total device entity count is the sum of all unique IP addresses from both sets of data.

### Trusted Endpoint Data Sources

Some data sources, such as EDR solutions, provide an inventory of their assets. In this model, the discovered hosts' private IP addresses are reported as assets. Data from a managed list of sources that are known to produce less accurate results is excluded from consideration.

### Inferred Devices

For data originating from sensors, firewalls, and other network devices, asset IP addresses are inferred from network flow data, subject to the following rules:

- The IP address must belong to one of the following groups:
  - Private address spaces defined in RFC-1918
  - CGNAT address space defined in RFC-6598
  - Customer-defined IP ranges
- The IP address must be observed to send traffic twice within a sample period.

## How Are Users Identified?

Users are identified either from trusted user data sources or inferred from observed activities from specific productivity platforms. The total count is the sum of all unique email addresses from these sources.

### Inferred Users

For data originating from Microsoft 365 or Google Workspaces, user entities are extracted from activity logs in connected Microsoft 365 and Google Workspaces environments.

#### Microsoft 365

For user entities discovered using email addresses from Microsoft 365, either  the **UserId** or **MailboxOwnerUPN** field must be a valid email address.

#### Google Workspaces

For user entities discovered using email addresses from Google Workspaces, the following must all be true:

- The **email** field is a valid email address
- The **profileId** field has a value
- The **owner** and **calendar_id** fields are not populated (this omits certain accounts created by third party applications that have integrated with Google Directory or Google SSO)

# Frequently Asked Questions

## How are devices with multiple IP addresses handled?

If you have a device with multiple IP addresses, each unique IP address is considered to be a unique entity for asset counting purposes. Each unique IP address often has a unique set of applications listening on the address, even when it is assigned to a host with other IP addresses. Each unique IP address has a unique fingerprint and security profile and is treated as a unique entity. The logical association at the host level is only used contextually.

## Many devices on my network use DHCP. Is each device counted multiple times for each IP lease it receives?

Yes, each observed IP that meets the criteria listed above is counted as a unique entity. Stellar Cyber recommends the following approaches to limit this:

- Eliminate double-counting by using IP reservations with DHCP where possible.
- Set DHCP lease periods to at least 24 hours to limit this to only counting devices twice

## My network consists of many mobile devices. When a device switches wireless access points, it may receive a new IP address. Are these counted as unique entities?

Yes, each observed IP address that meets the criteria listed above is counted as a unique entity. Stellar Cyber recommends configuring 802.1x session persistence to reduce the number of duplicate entities. This can also be configured to eliminate double-counting when a device changes connectivity mediums (for example, a devices switches from a wired to a wireless connection).

## I have multiple sites with overlapping IP spaces. Are each site’s IPs counted independently?

No. If two separate sites with overlapping IP spaces are sent to the same Data Processor (DP), two devices with the same IP address are counted as a single entity. Because this can cause confusion in correlation models, Stellar Cyber recommends using globally unique IP spaces where possible. In cases where this is unavoidable, we recommend deploying a separate Data Processor on each site to ensure the address spaces do not collide in correlation models.

## I use auto-scaling groups. Is each ephemeral instance’s IP counted uniquely?

Yes. Each unique observed IP is counted as a unique entity. Stellar Cyber recommends configuring the auto-scaling group to use sequential IP addressing in specific ranges to ensure double-counting does not occur. If you are using DHCP, choose a scheme that reissues the lowest available address in the pool.

## I use eBPF to monitor network traffic from Kubernetes. Are the IP addresses of my containers counted as entities?

Yes. Each container’s IP address (including shared addresses on virtual interfaces) is considered a unique entity. Stellar Cyber recommends using predictable addressing in Kubernetes to avoid double-counting.

### I have multiple domains registered in Microsoft 365 and users have email aliases in each domain. Are these counted multiple times?

Yes, if activities are observed for the same user but with duplicate email addresses, the number of unique email addresses is used to determine the entity count. Stellar Cyber recommends encouraging your users to use only the primary domain for authentication – all subsequent actions will be associated with a single identifier.

### Are service accounts counted as entities?

If the service account has a valid email address, it is counted as a user entity. In most cases, monitoring these non-human user accounts is important (for example the AWS root account). The presence of a valid email address is the simplest determining factor and this is used by the asset counting model.