Anomaly detection with ExtraHop Addy

Published: 2018-01-11

The ExtraHop Addy™ service is a cloud-based service that applies machine learning techniques to automatically determine what is normal versus unusual behavior in your IT environment. Unlike other machine learning solutions that rely on logs or agent data, the Addy service applies machine learning technology to your wire data without requiring you to configure anything.

Addy learns about normal network behavior by analyzing the data stored on your Discover appliance. After the Addy service is activated, you can then browse detected performance and security anomalies in the ExtraHop Web UI and investigate root causes for issues on your network.

Overall, Addy offers the following types of help:

- Uncover hidden issues before they create problems for your users
- Collect high-quality, actionable data to identify root causes of anomalies
- Find unknown performance issues, security issues, or infrastructure quirks
- Gain deeper insight into your network behavior

Important: The Addy service does not analyze sensitive information and data types.

Here are important considerations about anomaly detection with the Addy service:

- You must have an Addy service license.
- You must have full system privileges, access to the Admin UI, and access through any firewalls to connect a Discover appliance to the Addy service through ExtraHop Cloud Services. For more information, see Connect to the ExtraHop Addy service.
- You must have at least four weeks of wire data metrics stored on your Discover appliance before Addy can detect anomalies.
- On a Command appliance, you can access anomalies on a connected Discover appliance if that Discover appliance is connected to the Addy service.

Navigating anomalies

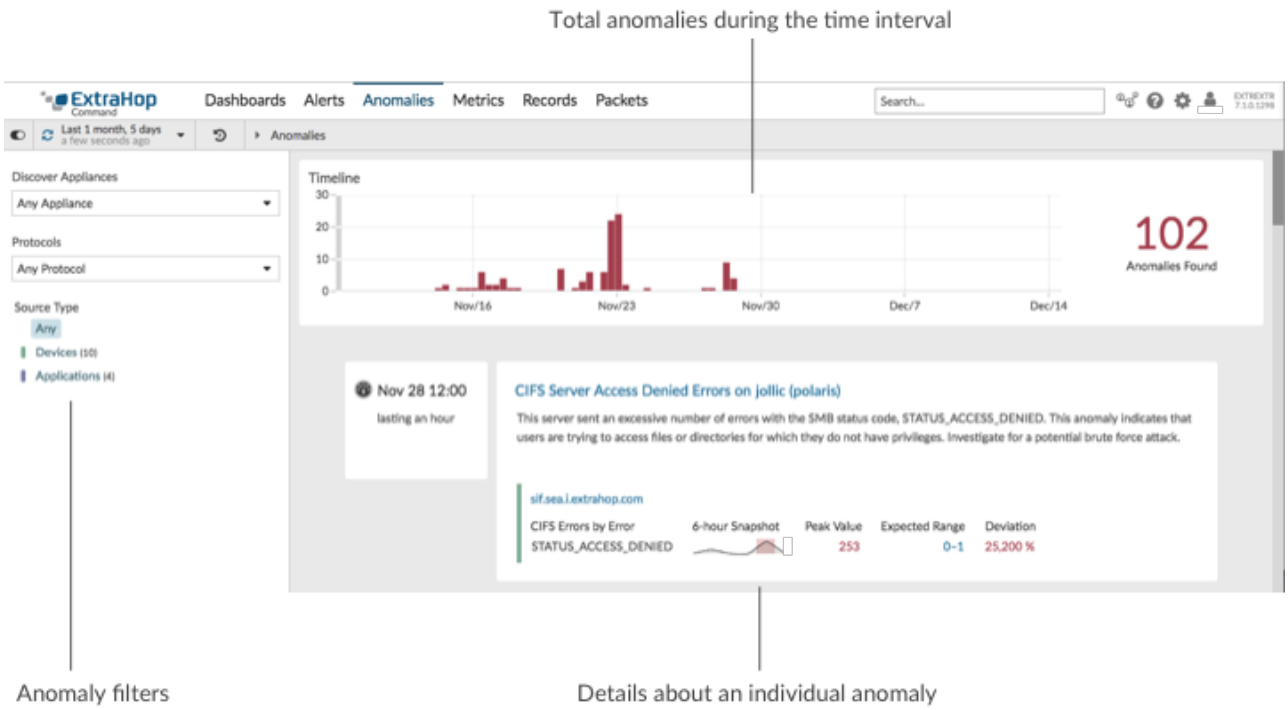
After connecting to ExtraHop Cloud Services, the Addy service automatically begins to calculate the expected range of normal metric values from four weeks of stored Discover appliance metrics, and then detects anomalies.

To browse anomalies, log into the Web UI on the Discover or Command appliance and click **Anomalies** at the top of the page.

Note: Receive email notifications for anomalies by configuring an anomaly alert. For more information about alerts, see the following topics:

- Configure Addy anomaly alert settings
- Add a notification to an alert configuration to receive emails when an anomaly is generated
- Alert History

The following figure shows how anomalies are displayed on the Anomalies page:



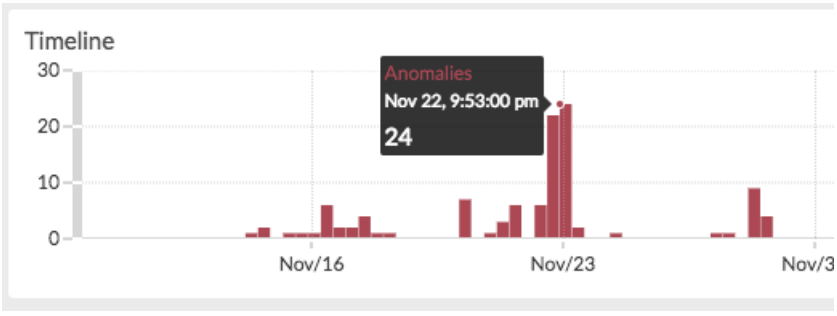
Interpret anomalies

The Anomalies page displays the total number of anomalies for the selected time interval and details about each detected anomaly. The following sections show you what information you can learn from anomalies.

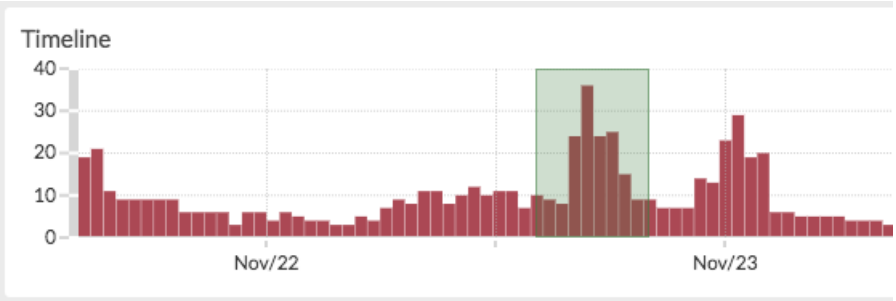
View total anomalies over time

The Timeline chart provides a summary of detected anomalies (y-axis) over time (x-axis) for the selected time interval. Each bar in the chart represents the total number of concurrent, active anomalies that were detected during a specific time period. Look for the tallest bar to determine when the most anomalies occurred in a time period.

Hover over a bar to view information, such as date, time, and the number of detected anomalies for a specific time period.



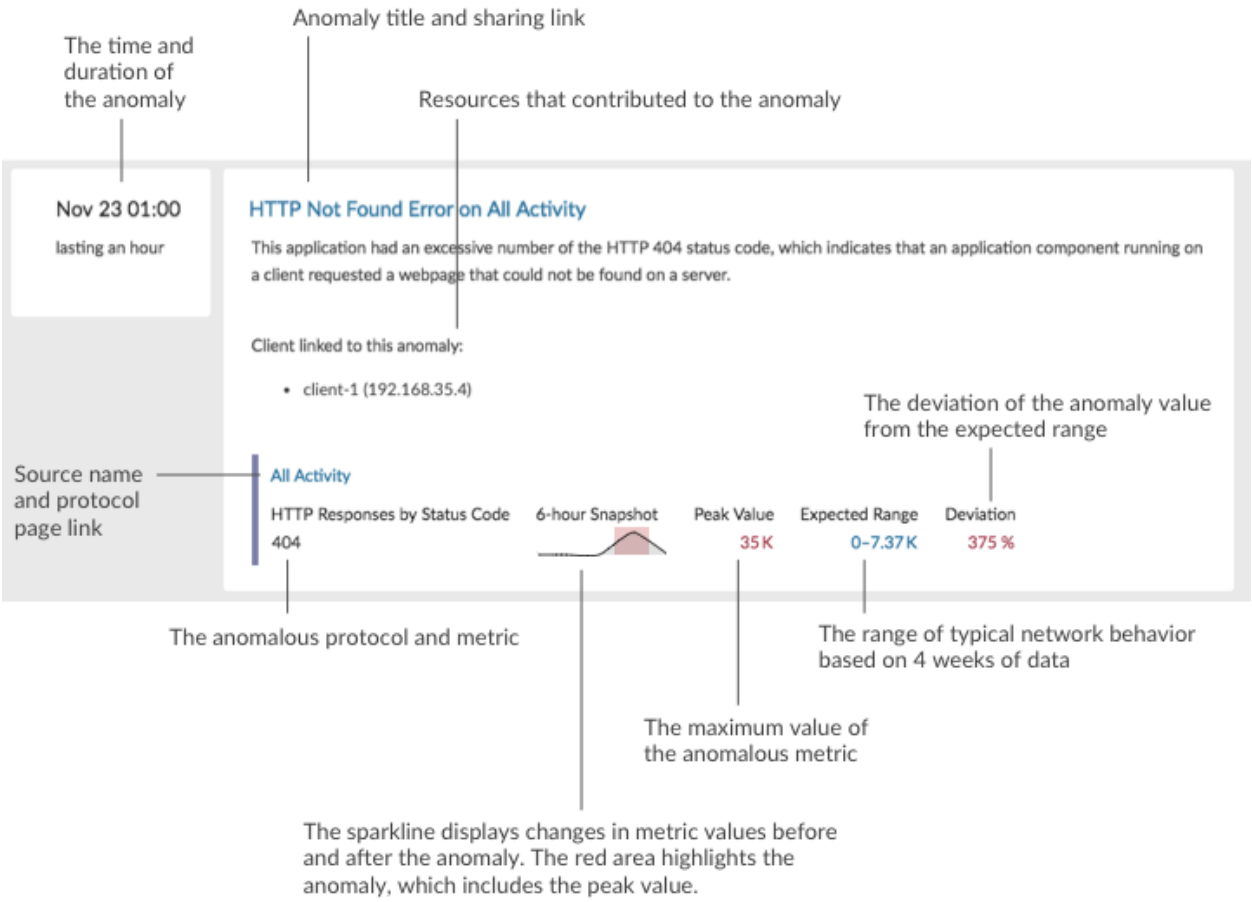
Click and drag across an area on the chart (which will become highlighted in green) to zoom in on a specific time range. The time interval in the Discover or Command appliance dynamically updates to match the new time range in the chart, and details about each anomaly that was detected in that time range are displayed below the chart.



View details for individual anomalies

Each anomaly provides detailed information about the type of issue that occurred, when the issue occurred, and the source of the issue. Individual anomalies are listed below the Timeline chart, and they are sorted by their start time. The most recent anomaly is listed first.

The following figure shows you what type of information is provided within an individual anomaly:



Title

The title includes the anomalous metric and the device or application name that is the cause of the anomaly. Click the title to share an anomaly.

Description

The description provides information about what the anomaly means. For most anomalies, Addy automatically surfaces detail metrics identified with Addy's machine learning capabilities, so you can immediately begin your investigation.

For more information, see Investigate anomalies with Addy.

Duration

The duration of the anomaly indicates how long the anomalous value was detected by the Addy service.

The minimum duration of an anomaly is one hour, because Addy detects anomalies by analyzing metric data with 1-hour granularity. If the duration value is displayed as ONGOING, the anomalous metric is in the process of being detected.

Sparkline

Sparklines are simple line charts that show you the metric behavior that led up to the anomaly. The sparkline charts display a snapshot of metric data from the time frame around the duration of the detected anomaly (such as 6 hours), and not the overall time interval from the top of the page (such as the last 7 days).

Peak Value

The peak value is the maximum value from observed data that deviated from expected ranged for the duration of the anomaly.

Expected Range

The expected range includes values that represent a normal background level of activity, which is calculated based on 4 weeks of data. The expected range is the basis for comparison with observed values to detect changes in metric activity.

Deviation

A deviation is the quantity calculated by the Addy machine learning engine to indicate the extent of change from an expected range.

How ExtraHop Addy works

This section provides some background information on how the ExtraHop Addy service identifies anomalies.

Anomalies are unexpected deviations from normal patterns in device or application behavior. By detecting an anomaly as soon as it happens, you can identify and resolve a potential issue before it becomes a larger problem. You can also review historical anomaly data to investigate issues related to known security or network outage events.

In most network monitoring tools, anomalies are detected through manually-configured alerts and trend models for individual devices. However, as your network changes—because of hardware reconfigurations or the addition of applications to your network—these types of alerts and models can become quickly outdated and potentially inaccurate. Addy automatically delivers consistent and accurate results about anomalous metrics and protocols without requiring manual configuration for individual devices. The Addy machine learning engine analyzes the historical behavior of individual devices, and automatically adapts to each device across time when there are changes to the expected range of data in your network.

Here is how Addy anomaly detection generally works: the metrics that the Addy machine learning engine analyzes come from wire data that is collected by your Discover appliance. The Discover appliance processes this data, generates metrics, and associates the metric data with protocols, devices, and applications. Addy retrieves a subset of protocol metrics from the Discover appliance to analyze and report results about detected anomalies.

The algorithm that drives the machine learning engine in Addy calculates the expected range of normal network behavior and adapts to changing variations in protocols and metric data. Outliers, or anomalies, are then detected based on three variables:

- Observed data, collected in real-time by the Discover appliance
- Expected range data, calculated from four weeks of historical data collected by the Discover appliance
- Threshold values, which are automatically adjusted by the algorithm based on historical metric data and heuristics defined by the IT networking domain experts at ExtraHop

Note: If you need to define a specific threshold value for an anomaly, which might be associated with a service level agreement (SLA) for example, we recommend manually configuring an alert in the Discover appliance.

Essentially, an anomaly is detected when observed data deviates from the expected range of data by a significant amount. You can then view analysis results about anomalies on the Anomalies page in the Web UI of the Discover appliance. For each anomaly, Addy provides the measured deviation (which is the difference between the observed value and the expected range), the anomaly value, and the expected range of normal metric values at the time of the anomaly.

Addy also provides anomalous 50th percentile or 75th percentile values for a subset of metrics that account for server processing time.

Related topics

Check out the following resources that are designed to familiarize new users with the Addy service.

- Connect to the ExtraHop Addy service
- Find and filter anomalies with the Addy service
- Investigate anomalies with Addy