November 2019

Detection of end-user device connectivity via active requests

N/A

Detection of end-user device connectivity via active requests

ABSTRACT

Network monitoring services typically do not collect signal or data regarding the state of connectivity between end-user devices on a network segment and their serving access points. Identifying that a network device is up does not guarantee that end users have network connectivity. The practice of only monitoring infrastructure devices results in service outages being identified only through end-user reporting. This disclosure describes a network service that measures end-user connectivity via active requests and responses. Valid end-user addresses are loaded, grouped by network type and device, and sampled for connectivity at periodic intervals. If the fraction of responding end-user devices falls below certain thresholds, network operations personnel are alerted.

KEYWORDS

- Network monitoring
- Service outage
- Service level objective (SLO)
- End-user connectivity
- Network services
- Network operations center (NOC)
- Enterprise network

BACKGROUND

Network monitoring services have a gap in monitoring, specifically, such services do not collect signal or data about the state of connectivity between end-user devices on a network segment and their serving access points. Identifying that a network device is up does not

guarantee that end users have network connectivity. The practice of only monitoring infrastructure devices may result in service outages (which may occur due to network configuration errors or other causes) being identified only through end-user reporting. This is a particularly acute problem for large networks or complex networks, e.g., with a very large number of devices, or serving different operating systems and end-devices (mobile devices, computers, IoT devices, printers, etc.). This gap also limits a network service provider to compute service-level objectives based solely on the state of the network infrastructure.

Although network equipment that simulates virtual user activity through the use of synthetic probes exists, such equipment collects information about real-time network performance, e.g., factors such as response-times, latency, jitter, packet loss, etc., and does not monitor end-user device connectivity. Such equipment has different permissions on the network, e.g., on a different VLAN, or housed in secure locations with more direct connections to edge routers than end-user devices, and can miss instances when only end-user devices are affected. Besides, the data can be collected only as long as the network service is up, which defeats the purpose of detecting end-user service outage. Other network services exist that provide real-time telemetry metrics from end-user access network segments, enabling the characterization of network performance from an end-to-end perspective, e.g., how end-users use and experience the network. However, such network services also only function so long as the network is up.

DESCRIPTION

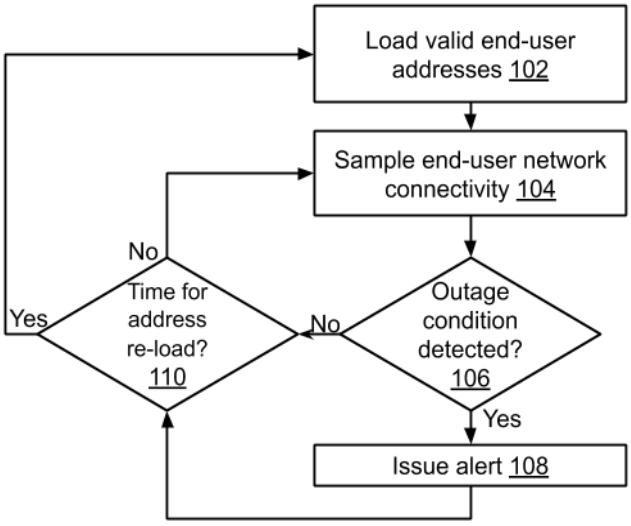

Fig. 1: Network service to measure end-user connectivity

This disclosure describes a network service that measures end-user connectivity via active requests and responses. The operation of the service is illustrated in Fig. 1.

1. *Load valid end-user addresses* (102): At periodic, e.g., hourly, intervals, the service collects a set of IPv4/IPv6 addresses of end-user devices connected to the network. This set can be generated via internet protocols such as address resolution protocol (ARP) or neighbor discovery protocol (NDP). Alternately, control plane network devices can provide a list of connected devices using, e.g., dynamic host connect protocol (DHCP). The addresses are grouped by network type and device, e.g., end-user devices are grouped by logically similar failure modes.
2. *Sample network connectivity through end-user devices* (104): At a period of relatively fine granularity, e.g., every thirty seconds, network connectivity checks are sent out using, e.g., Internet control message protocol (ICMP), to a sample of the collected IP addresses of a group. Sampling end-user devices for network connectivity is explained in greater detail below.

3. *Detect outage condition* (106): Outage conditions are detected based on factors such as the total number of IP addresses connected to the network, the number of responsive IP addresses, the number of unresponsive IP addresses, some form of trend analysis, e.g., long-term averaging or historical comparison, etc. For example, a machine-learning model can detect an outage based on one or more of the preceding factors. As another example, an outage can be defined to occur when the following conditions are met:
    a. The number of responsive end-user devices in a given network segment falls below a first threshold;
    b. The number of responsive end-user devices in a given network segment stays below the first threshold for a period of time, e.g., three minutes; and
    c. The long-term average of the number of responsive devices stays above a second threshold.

    Detection of outage conditions is explained in greater detail below.
4. *Issue alert* (108): If outage conditions are detected, network operations personnel are alerted.
5. *Check for time to reload addresses* (110): If a sufficient time (e.g., one hour) has passed since the last loading of valid end-user IP addresses, then the IP addresses are reloaded (102). If not, end-user network connectivity continues to be sampled, at a relatively fine time-granularity, e.g., every thirty seconds (104).

Sampling end-user devices for network connectivity

To prevent the monitoring procedure from loading the network only a fraction of IP addresses are sampled for testing connectivity in any individual sampling event. Depending on the number of responsive end-user devices, the number of IP addresses sampled at the next

sampling event is scaled up or down by a scaling factor. For example, if all end-user devices responded at a sampling event, then at the next sampling event the number of devices to be sampled is scaled down by the scaling factor, thereby reducing network load. If none of the end-user devices responded at a sampling event, then at the next sampling event the number of devices to be sampled is scaled up by the scaling factor, thereby expanding the scope of monitoring. Scaling down continues until a certain non-zero minimum number of end-devices is reached. Scaling up continues until a certain maximum number of end-devices is reached.

To avoid selection bias, every IP address within a group is tested for end-user connectivity before repeating an IP address. In this manner, all IP addresses are eventually sampled, and the probability of frequently resampling the same set of devices is reduced. Further, the set of end-user devices to be sampled for testing end-user connectivity is shuffled, e.g., randomly reordered, when the IP addresses within a group have each been sampled at least once.

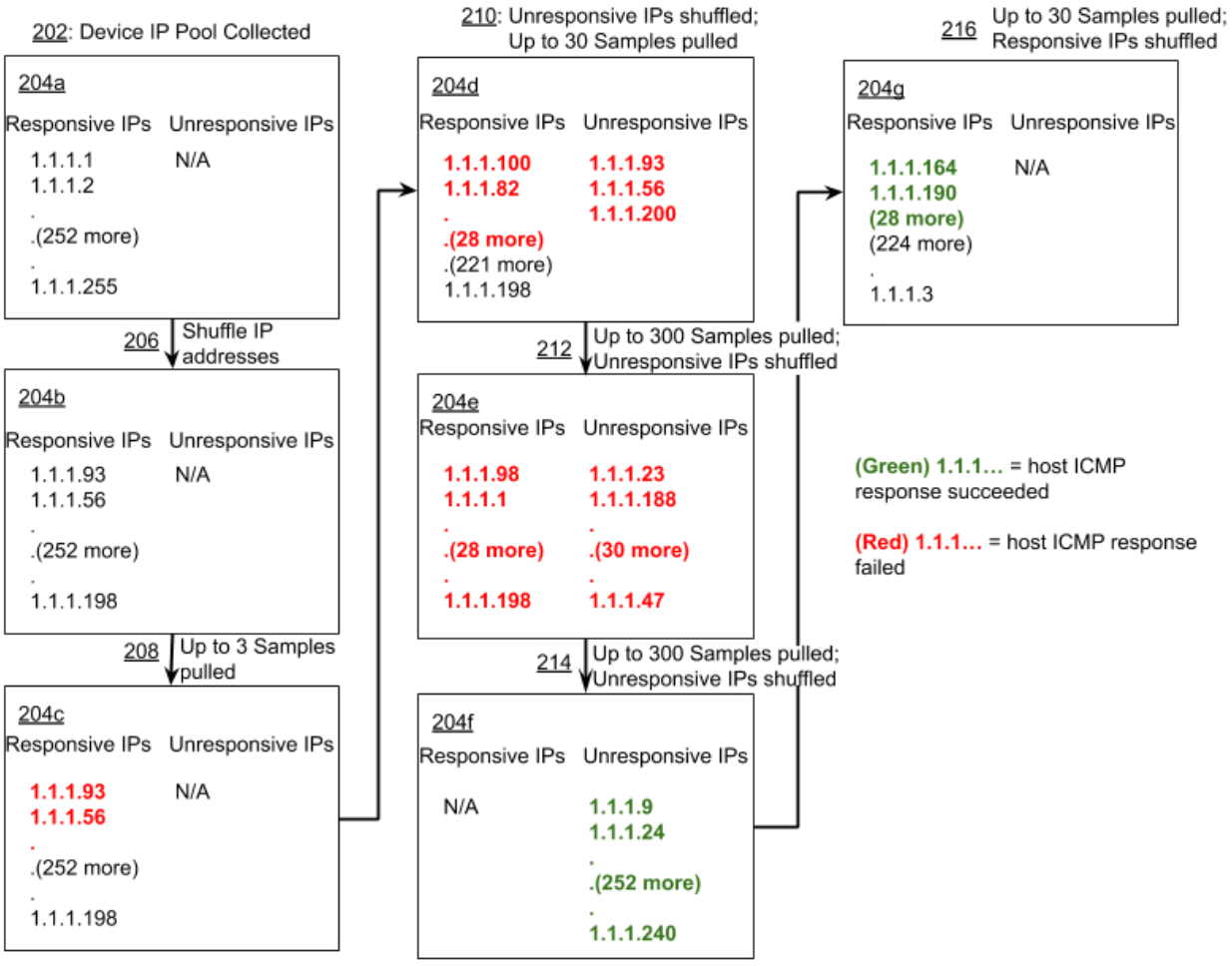

Fig. 2: Sampling end-user devices for network connectivity

Fig. 2 illustrates in greater detail sampling end-user devices for network connectivity, including the features of shuffling end-user devices and changing between sampling events the number of end-user devices being sampled. The network service maintains a list of responsive IPs, e.g., end-user devices that responded to ICMP messages, shown generally in green; unresponsive IPs, e.g., end-user devices that did not respond to ICMP messages, shown generally in red; and end-user devices that are yet to be tested for connectivity in the most recent sampling cycle, shown in black.

A pool of valid end-user device IP addresses is collected (202). These are represented (204a) as untested IP addresses 1.1.1.1 through 1.1.1.255. The IP addresses are shuffled (206), to

get a shuffled address list (204b). Three end-user device addresses (1.1.1.93, 1.1.1.56, 1.1.1.200) are selected for purposes of testing connectivity (208). None of the three are found responsive (204c). This results in these three end-user addresses being moved to the unresponsive IP list (204d) and a scaling up of the number of end-user devices to be tested to thirty (210).

None of the thirty end-user devices are found responsive. These unresponsive devices are moved to the unresponsive IP list (204e), and the number of end-user devices to be tested is further increased (212) to three hundred.

None of the three hundred end-user devices are found responsive. These unresponsive devices are moved to the unresponsive IP list (204f), and the number of end-user devices to be tested remains at a maximum of three hundred (214).

All two hundred and fifty-five end-user devices are found responsive (204f). These responsive devices are moved to the responsive list (204g), and the number of end-user devices to be tested is decreased to thirty (216). The responsive IP list is shuffled, and the next cycle of testing end-user devices for connectivity commences.

Detecting outage conditions

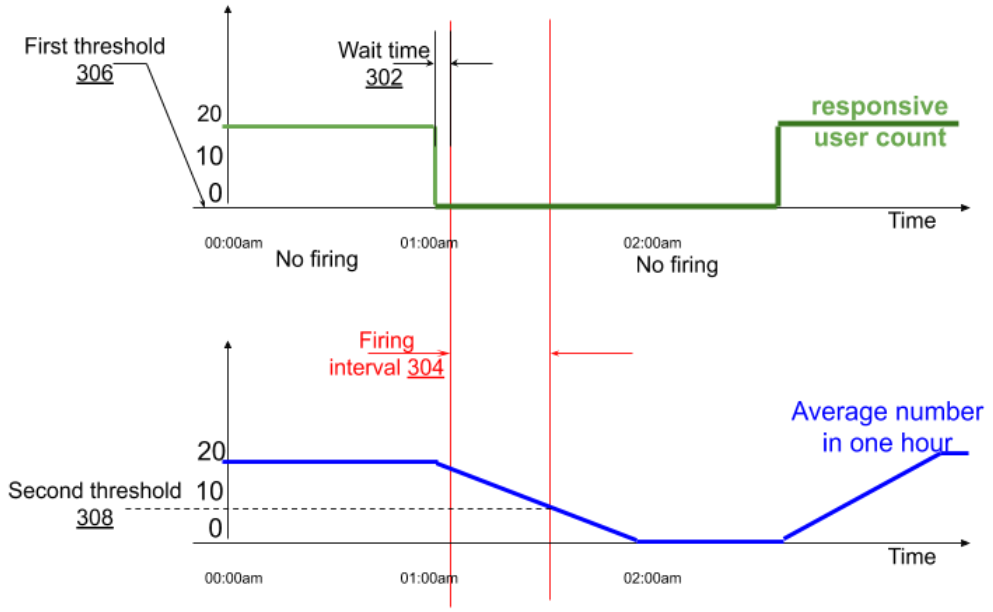

Fig. 3: Detecting outage conditions

Fig. 3 illustrates an example of detecting outage conditions based on the number of responsive users (green curve) and the long-term average of the number of responsive users (blue curve). A first threshold (306), operative on the number of responsive users, is set to zero. A second threshold (308), operative on the average of the number of responsive users, is set to eight. As explained before, an outage is detected, e.g., alerts are fired (304) during an interval in which

A. the number of responsive end-user devices falls to or below the first threshold;
B. the number of responsive end-user devices in a given network segment stays at or below the first threshold for a wait time (302); and
C. the long-term average of the number of responsive devices stays above the second threshold.

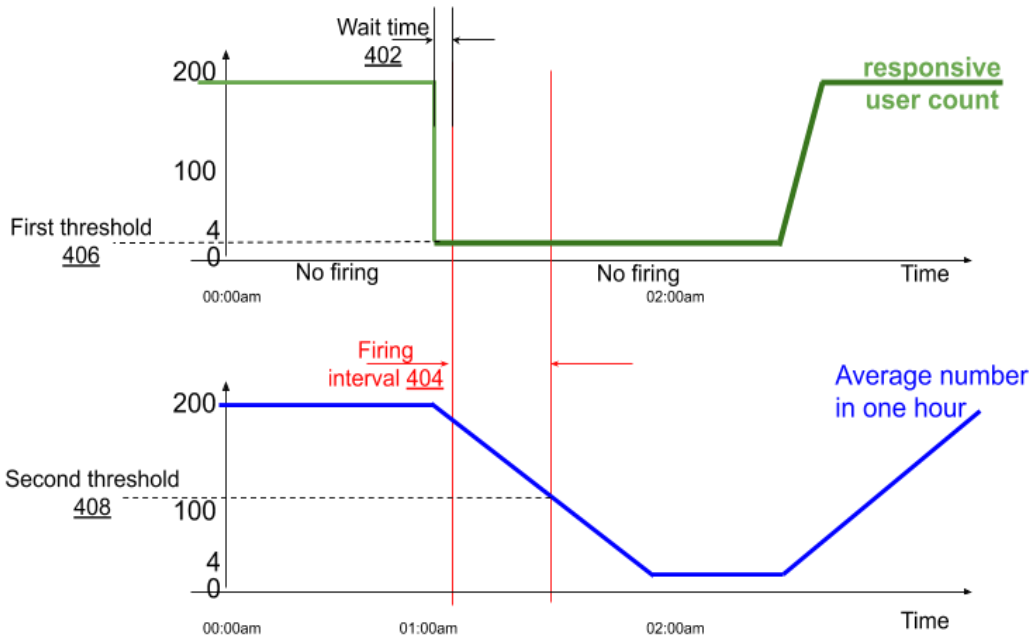

Fig. 4: Detecting outage conditions

Fig. 4 illustrates another example of detecting outage conditions based on the number of responsive users (green curve) and the long-term average of the number of responsive users (blue curve). A first threshold (406), operative on the number of responsive users, is set to four. A second threshold (408), operative on the average of the number of responsive users, is set to one hundred. As explained before, an outage is detected, e.g., alerts are fired (404) during an interval in which

A. the number of responsive end-user devices falls below the first threshold;

B. the number of responsive end-user devices in a given network segment stays below the first threshold for a wait time (402); and

C. the long-term average of the number of responsive devices stays above the second threshold.

The purpose of the second threshold, which ends the alert-firing interval, is to allow for the possibility that a lack of responsive devices may not be indicative of outage. For example,

end-user devices may not respond simply because the business day has come to an end and users have gone home with their devices.

Per the techniques of this disclosure, service outages are almost immediately detected without explicit reporting by users. The described network service that detects outages is a black-box monitor, e.g., it does not require equipment or software installation, and it does not rely on the state or reachability of a network device (or gateway). The network service is agnostic to operating systems (at both the user and the host) and can monitor all network segments across an IT infrastructure. The feature of sampling network connectivity ensures that monitoring traffic is a small fraction of total network load, e.g., the network service avoids exerting excessive load on the network. The described service has limited network impact and is scalable based on address population sizes.

CONCLUSION

This disclosure describes a network service that measures end-user connectivity via active requests and responses. Valid end-user addresses are loaded, grouped by network type and device, and sampled for connectivity at periodic intervals. If the fraction of responding end-user devices falls below certain thresholds, then network operations personnel are alerted.