June 2023

TECHNIQUES TO PROVIDE EDGE RELAYS WITH PRIVACY-PRESERVING CACHES

Giovanna Carofiglio

Luca Muscariello

Michele Papalini

Mauro Sardara

TECHNIQUES TO PROVIDE EDGE RELAYS WITH PRIVACY-PRESERVING CACHES

AUTHORS:
Giovanna Carofiglio
Luca Muscariello
Michele Papalini
Mauro Sardara

ABSTRACT

Real-Time Communication (RTC) traffic typically leverages one or more media bridges (often located in the cloud) and, in order to reduce latency and/or offload cloud resources to the edge, one or more real-time edge relays can be utilized in order to optimize such traffic. However, edge relays may also be exploited by malicious entities and, thus, certain protective mechanisms are typically utilized at edge relays that, while reducing the probability of being exploited, can reduce the throughput of such edge relays. Techniques presented herein may help to preserve data privacy at edge relays through the use of a time-local caching mechanism.

DETAILED DESCRIPTION

Real-Time Communication (RTC) traffic typically leverages one or more media bridges that are often located in the cloud in order to orchestrate a group communication and switch media across participants involved in the group communication. To reduce latency and offload cloud resources to the edge, one or more real-time edge relays can be deployed in which such edge relays may be considered opportunistic, intermediate nodes located in-path that seek to optimize the traffic (e.g., by sending retransmissions at lower latency, providing cost saving by offloading from the cloud in the case of 'edge-local' participants).

However, moving RTC traffic to an edge relay is not without challenges. For instance, cache partitions of an edge relay should be protected from malicious entities that may seek to perform replay attacks, leverage cache pollution techniques, and/or the like. Simply encoding variable parts of the cached data will also reduce cache locality at the edge relay, thereby leading to reduced cache throughput.

It is important that the performance of the caching system at a media relay node is insensitive to the key encoding used by a requestor.

In order to address such issues, techniques are presented herein that provide for first partitioning the cache of an edge relay and then performing lookups on these partitions based on a time locality. More specifically, a data sender may encode the variable part of the data name with randomized values, with the objective of reducing replay-attacks and cache pollution mechanisms by anonymizing the less significant bits of the name field (e.g., "/AppId/MeetingId/mediaId/SequenceNumber"). However, this can reduce cache locality at the edge relay node up to making cache throughput very low.

Figure 1, below, illustrates various operations that can be utilized to facilitate such cache partitioning and lookups.

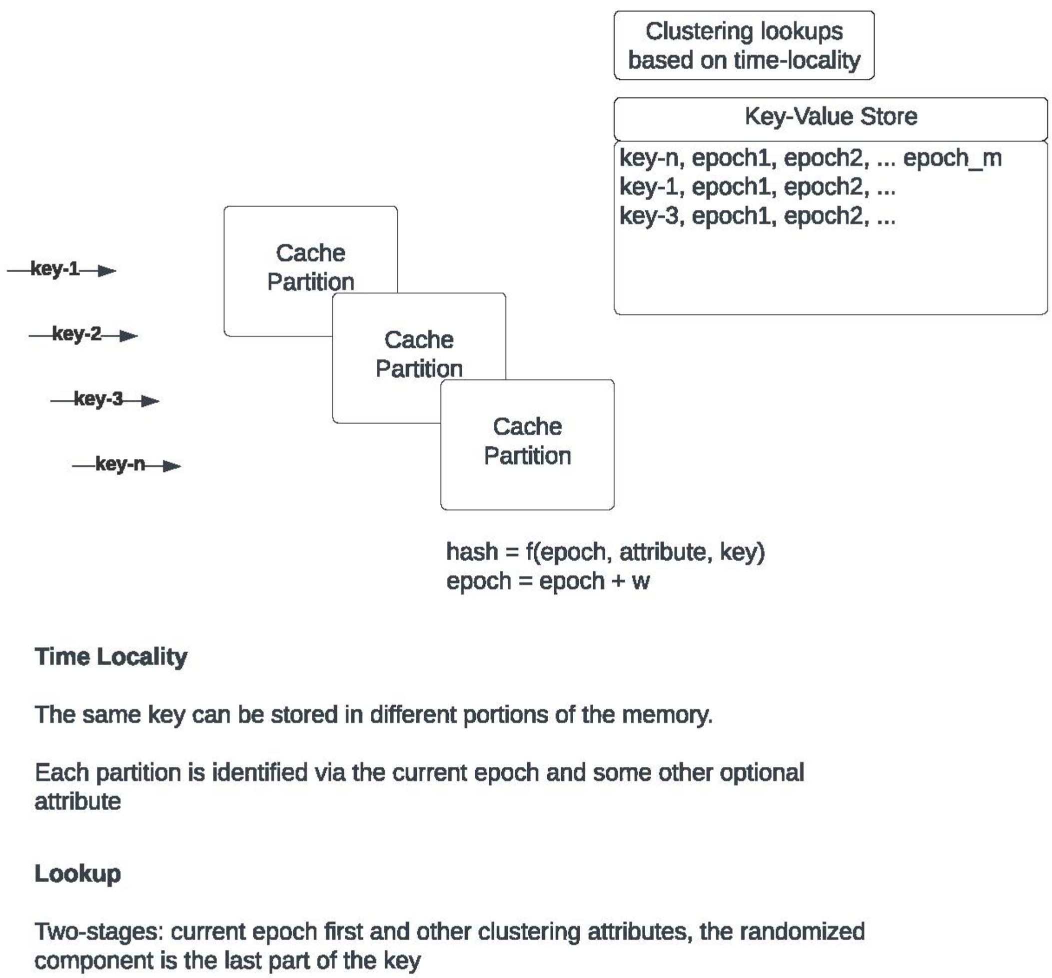

Figure 1: Example Cache Partitioning and Lookup Operations

As shown above in Figure 1, a packet cache for the edge relay node can be sharded and partitioned by time-based-epochs. For the partitions, lookups may be performed in two stages: using the name and the epoch via a key-value store. For a real-time application, the epoch update determines when a key in the cache becomes obsolete or addressable by searching backwards in time (e.g., by searching the previous epoch).

Accordingly, the time-local caching mechanism as presented herein may help to preserve data privacy at edge relays via a time-local caching mechanism.