Cooperative Approximate Authorization Recycling

Qiang Wei, Konstantin Beznosov, Matei Ripeanu
Department of Electrical and Computer Engineering, University of British Columbia

1. Introduction

Architectures of modern access control solutions—such as Access Manager, GetAccess, SiteMinder, EJB—are based on the request-response paradigm. In this paradigm, the policy enforcement point (PEP) intercepts application requests, obtains access control decisions (or authorizations) from the policy decision point (PDP), and enforces those decisions.

In large enterprises, PDPs are commonly implemented as dedicated authorization servers, providing such important benefits as consistent policy enforcement across multiple PEPs and reduced administration of authorization policy. The drawbacks are also critical: reduced performance due to communication delays between the PEP and PDP, as well as reduced reliability, since each PEP depends on its PDP and the connecting network.

The state-of-the-practice approach to improving overall system reliability and decreasing processing delays observed by system users is to cache authorization decisions at each PEP—what we refer to as *authorization recycling*. Enterprise authorization solutions commonly provide PEP-side caching. However, these solutions employ a simple form of authorization recycling: a cached decision is reused only if the authorization request in question exactly matches the original request for which the decision was made. We refer to such reuse as *precise authorization* recycling.

To improve authorization system availability and latency, Crampton et al. [1] introduced a *secondary decision point* (SDP). Collocated with the PEP, the SDP resolves authorization requests by inferring *approximate* authorizations. Thus, the SDP provides an alternative source of access control decisions in the event that the PDP is unavailable or slow.

To further improve the availability of access control systems, we propose an approach of *cooperative approximate authorization recycling* (CAAR), where authorization requests are resolved by cooperating SDPs. Our solution is similar to cooperative Web caching, in that system components collaborate to serve the overall request stream. Unlike Web caching, however, CAAR allows computation and sharing of *approximate* access control decisions.

2. CAAR Architecture

CAAR is a collaborative access control system integrating multiple cooperating SDPs. In addition to solving requests received from its PEP, each SDP participates in solving requests from other SDPs. To save bandwidth and reduce load, we use a discovery service which helps find those SDPs that can contribute to resolving an authorization request. These SDPs compute responses locally and send them back to the original SDP.

For the same request, our inference algorithms ensure that different SDPs always compute consistent responses when their caches are fresh. In the case of conflicts caused by stale responses at some SDPs, the SDP in question selects the response that was inferred from the most recent cached data. To manage inconsistencies caused by policy updates at the PDP we plan to use a system based on leases.

CAAR is designed to operate in a large enterprise infrastructure. To limit the scope of assumed trust, CAAR design allows each SDP to trust only those PDP(s) whose responses it enforces; each response issued by a PDP (a.k.a., primary response) is signed by that PDP, whereas all secondary responses are backed by a chain of primary responses by which its validity can be verified.

3. Evaluation

We evaluated the CAAR system through simulations. We used the cache hit rate as an indirect metric for availability improvements. A cache hit means that an authorization request is resolved by the cooperating SDPs without resorting to the PDP. Therefore, a high cache hit rate results in requests being resolved by the local and other cooperating SDPs, even when the PDP is unavailable, thus increasing availability.

We simulated multiple cooperating SDPs that implement inference algorithms for the same Bell-LaPadula access control policy. Our experiments explored the influence of the following four factors on the overall hit rate: (a) cache warmness at each SDP, (b) the number of cooperating SDPs, (c) the overlapping rate between the request spaces of the cooperating SDPs, and (d) the distribution of request popularity (uniform or Zipf).

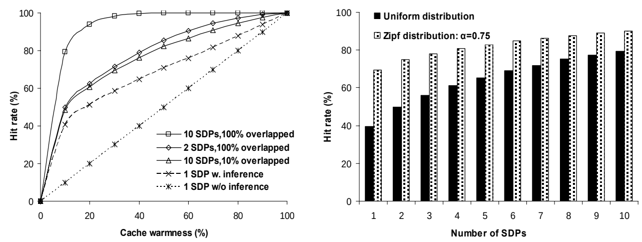

Figure 1: *Left graph*: hit rate as a function of cache warmness with various numbers of SDPs and overlapping rates (for uniform request popularity). *Right histogram*: hit rate as a function of the number of SDPs for both uniform and Zipf-popularity distributions.

Simulation results indicate that the hit rate of the overall system is greater than 50%, even when individual SDPs have a cache warmness of less than 10%. Since small caches can be loaded in memory, this allows cache hits to be served from memory rather than from disk, which could further speed up the computation of approximate authorizations.

Increasing the number of SDPs leads to higher hit rates. However, additional SDPs provide diminishing returns in terms of improving the hit rate, due to increased cache overlapping. For instance, the first SDP brings a 10% improvement in the hit rate, while the 10th SDP contributes only 2%. One can thus limit the number of cooperating SDPs to control the overhead traffic without losing the major benefits of cooperation. The results also imply that in a large system the impact of a single SDP's failure on the overall hit rate is negligible.

Finally, the simulation results indicate that cooperation among multiple SDPs brings higher benefits for uniform request popularity distributions than for non-uniform (Zipf) ones. The reason is that popular requests are already cached locally, while unpopular requests are also unlikely to be solved by other SDPs.

4. Summary

CAAR is a collaborative access control system designed for large enterprises. Our evaluation shows that by combining coordination and inference, multiple cooperating SDPs can significantly improve the availability of an access control infrastructure as observed by application clients.

5. References

[1] J. Crampton, W. Leung, K. Beznosov, "The Secondary and Approximate Authorization Model and its Application to Bell-LaPadula Policies," *ACM Symposium on Access Control Models and Technologies* (SACMAT), Lake Tahoe, CA, USA, June 2006, pp 111-120.