Contribution ID: **33** Type: **Oral Presentation**

Multi-domain anycasted high availability for stateful services in RCauth.eu, now made simple

Friday, March 24, 2023 11:50 AM (30 minutes)

Use of 'anycasting' internet addresses ('IP anycast') in load balancing and high availability, and for traffic engineering reasons, is a widely deployed technique for content delivery networks to lower latency for access to frequently accessed content such as web pages and video. Using the properties of the Border Gateway Protocol (BGP) as a variable-length path-vector protocol for routing internet packets, distinct hosts in multiple places in the internet announce the same network address to serve the same content. This provides redundancy of service provisioning, and at the same time offers the possibilities for traffic engineering by varying the perceived path length in the 'default-free zone' of the global routing table.

The most common deployment of anycast is a single organization managing all underlying hosts, and then announcing their hosts either using their own autonomous system, or from a range of autonomous systems all under single administrative control. The provisioning hosts themselves are also usually 'stateless' – they either service static content or obtain any state required from upstream sources that are not publicly exposed.

The RCauth.eu federated token translator is a service that issues end-user 'PKIX' certificates with globally unique, persistent, and non-reassigned identifiers based on eduGAIN-federated authentication. However, the uniqueness and non-reassignment must be guaranteed by the service itself, and hence it maintains state in a back-end database that is consulted and updated on issuance of each certificate.

The initial deployment of RCauth.eu consisted of simple hardware security module and security controls at a single site, Nikhef in Amsterdam, which could sustain a very low issuance volume. For deployment in more communities and infrastructures, and in the European Open Science Cloud, a most robust solution was required. A collaboration of Nikhef (Amsterdam, The Netherlands), GRNET (Athens, Greece), and STFC (Didcot, Oxford, UK) therefore initiated a more robust setup using a distributed RCauth.eu service, where each site hosts a fully replicated instance. Since the user experience must be consistent (a persistent, unique, and mostly unchanging credential based the user's federated identity), the service has to be supported by a distributed database that retains near-synchronous state across all instances. However, since the expected total issuance volume for RCauth.eu is unlikely to exceed the capacity of one instance, the primary purpose of the distributed setup is to provide redundancy and rapid fail-over, rather than load balancing.

In building the distributed RCauth.eu, we reviewed several distributed high availability techniques that aim to remove single points of failure, and work without operator intervention. Since the transaction flow in token translation can take several minutes (due to the user authentication interaction with the home organisation), failures occurring during that period must be absorbed, and be independent of the settings on the client devices. It should also work across administrative domains and across countries and regions. Based on those requirements, we selected BGP Anycast as the most appropriate technology, but engineered the system in such a way that it minimally affects existing systems and network operations. We demonstrate that we can build a stateful anycasted service across three countries and two autonomous systems, achieve rapid (seconds-scale) failover, can synchronise databases over transport-protected L4 virtual circuits while maintaining a consistent database state. And by considering an integrated approach of service and host management, internal routing, and eBGP engineering, we show how to build a highly available multi-domain and multi-national service without requiring additional autonomous system resources.

Primary author: GROEP, David (Nikhef and Maastricht University)

Co-authors: SALLÉ, Mischa (Nikhef); GKINIS, Kyriakos (GRNET); LIAMPOTIS, Nicolas (GRNET); FURNELL, Will (UKRI-STFC); JENSEN, Jens (UKRI-STFC)

Presenter: GROEP, David (Nikhef and Maastricht University)

Session Classification: Network, Security, Infrastructure & Operations

Track Classification: Track 7: Network, Security, Infrastructure & Operations