IPSec-HA Recap
Yoav Nir
Where are we?

- draft-ietf-ipsecme-ipsec-ha accepted as a WG item early this year
- Now at version -09
- Draft was approved by the IESG on 15-Jul
- Now in RFC Editor's queue
- Let's go over the issues and the terminology
Terminology

This is a single gateway

This is a cluster
Terminology

This is a single gateway

This is a cluster member

This is a cluster

This is a cluster
Terminology

• Availability – portion of the time a system can do its work. Expressed as percentage or “nines”
• High Availability – the property of a system where the down time is low.
• Fault Tolerance – a property of a system where functionality is maintained even following a specified set of fault condition.
Terminology

This is a hot standby cluster

This is a load-sharing cluster
Terminology

This is a hot standby cluster following a failover
Terminology

• Failover is when a part of the load goes from one cluster member to another.
• In HS cluster a standby member becomes the active member, and the formerly active member either becomes a standby member, or is out of commission.
• In LS the decision function changes.
  – So the handling of a certain peer, SA, or selector migrates.
  – Or one of the members is out of commission.
Terminology

• Tight Cluster – a cluster where all the members share an IP address.
• Loose Cluster – a cluster where the members don't share an IP address.
  – They may share a DNS name
  – They may use RFC 5685 redirect to send traffic to the correct gateway.
Terminology

• Tight Cluster – a cluster where all the members share an IP address.
• Loose Cluster – a cluster where the members don't share an IP address.
  – They may share a DNS name
  – They may use RFC 5685 redirect to send traffic to the correct gateway.
• We don't care about loose clusters.
  – They're out of scope.
Terminology

• Synch channel is the means by which cluster members communicate in order to share state.
Problems

• Out of scope:
  – How the synch channel work
  – Synchronizing policy

• In scope:
  – Any behavior of a cluster, that may appear different from that of a single gateway.
  – Any altered behavior following a failover.
Problems

• Lots of state (section 3.2):
  – IKE SAs
    • Keys
    • Authentication Information
  – IPsec SAs
    • With replay counters
  – SPD Cache entries
Problems

• IKE Counters (section 3.3)
  – An implementation MUST keep careful track of Message Ids, both inbound and outgoing.
• Synch after every IKE exchange?
Problems

• Outbound SA counter (section 3.4):
  – MUST NOT reuse a replay counter value.
• Synch after every IPSec packet?
  – Not feasible!
• Synch occasionally?
  – State will mismatch with peer after failover.
• Does the peer actually enforce this?
Problems

• Reminder:
  – IKE Message Counters
    • MUST NOT repeat
    • MUST NOT skip
    • MUST process in order
  – IPSec Replay Counter
    • MUST NOT repeat
    • May skip as much as you want
    • Enforcement is OPTIONAL.
      – If you enforce, MUST NOT process outside window
Problems

• Inbound SA Counter (section 3.5):
  – Like the previous problem, only causes a security vulnerability
  – Should not accept a packet with an old replay number.

• Synch after every packet?
  – Not practical and you might still miss.

• Live with it, assuming an attacker can't both replay and cause/detect a failover?
  – After all, enforcement is OPTIONAL.
Problems

• Missing Synch Messages (section 3.6):
  – No transport is 100% reliable.
  – If failover happened, there's a good chance some synch messages are missing.

• We have to assume that our state is mismatched with the peer's.
  – Maybe there's an SA we don't know about.
  – Maybe an SA was deleted.
Problems

• Simultaneous use of IKE or IPSec SA by more than one member (section 3.7):
  – Relevant for LS cluster
  – Replay counters cannot synch.

• Solutions fall into two broad categories:
  – “Sticky” - only one member handles a particular class of traffic, so no shared SA.
  – “Duplicate” - Similar SAs, one for each member with the same peer.

• Also a problem choosing distinct IVs.
Problems

• Overloading the load balancer (section 3.8) – We'd like the IPsec SA to directly to the member, bypassing the load balancer.

• draft-arora-ipsecme-ikev2-alt-tunnel-addresses addresses addresses this.

• Later on, we'll talk about whether this is interesting for the WG.
Problems

• Allocation of SPIs (section 3.9):
  – SPIs for inbound SAs MUST be distinct.
  – Members MUST NOT create two SAs with the same SPI, at least not with the same peer.
  – Do we really need a protocol extension to solve this?
    • We think not
• That's it for the problems.
• We'll come back to these slides when we discuss the solutions later.