SIP Timer Engineering

End-to-End Configuration For Optimal Performance

Samuel Rausch, Principle Engineer
Comcast Cable
April 26th, 2010

Overview

- SIP Timers control a wide range of behaviors across all the voice elements in an operator's network
- Proper engineering must take into account the impact of each element's timers on the rest of the network
- The operator's access model also defines specific behaviors
    - Over-the-top versus operator managed
- Timers are also useful in the management of network congestion
- Many thanks to Ruibing Hao, Jiongkuan Hou and Vijey Jenkal for their contributions to Comcast's ongoing network architecture optimizations!

Network Models

- Operators have different network models, each one with unique behaviors and challenges
- Variations include
  - Over-the-top vs operator owned access
  - Redundant vs non-redundant vs non-existent Access SBC's

Critical Timers

Timer	Default value	Section	Meaning
T1	500 ms	17.1.1.1	Round-trip time (RTT) estimate
T2	4 sec.	17.1.2.2	Maximum retransmission interval for non-INVITE requests and INVITE responses
Timer A	initially T1	17.1.1.2	INVITE request retransmission interval, increases exponentially per retransmission
Timer B	64*T1	17.1.1.2	INVITE transaction timeout timer
Timer E	initially T1	17.1.2.2	Non-INVITE request retransmission interval, increases exponentially per retransmission
Timer F	64*T1	17.1.2.2	Non-INVITE transaction timeout timer

- T1 – the base value for all other timers, **absolutely critical!!!**
- T2 – max interval value of exponential backoff timers (Timer A & E)
- Timer A – base interval for INVITE backoff (Timer B)
- Timer B – total INVITE transaction timer
- Timer E – base interval for non-INVITE backoff (Timer F)
- Timer F – total non-INVITE transaction timer

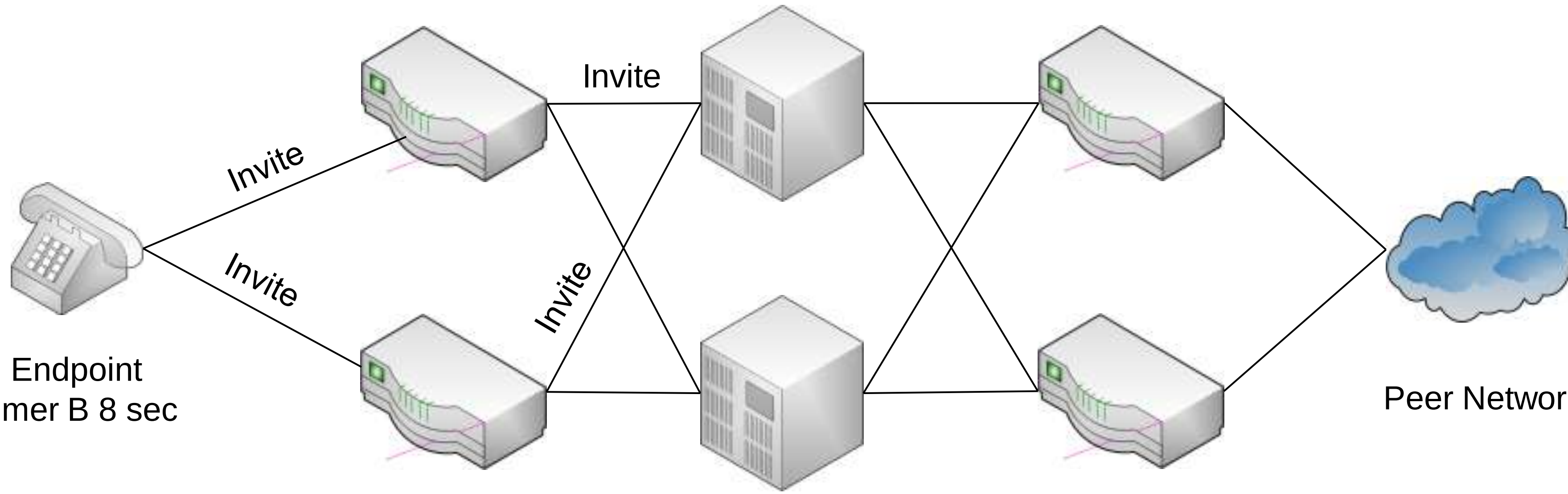
Other Timers

Timer	Default value	Section	Meaning
T4	5 sec.	17.1.2.2	Maximum duration that a message can remain in the network
Timer D	> 32 sec. for UDP	17.1.1.2	Wait time for response retransmissions
Timer G	initially T1	17.2.1	INVITE response retransmission interval
Timer H	64*T1	17.2.1	Wait time for ACK receipt
Timer I	T4 for UDP	17.2.1	Wait time for ACK retransmissions
Timer J	64*T1 for UDP	17.2.2	Wait time for retransmissions of non-INVITE requests
Timer K	T4 for UDP	17.1.2.2	Wait time for response retransmissions

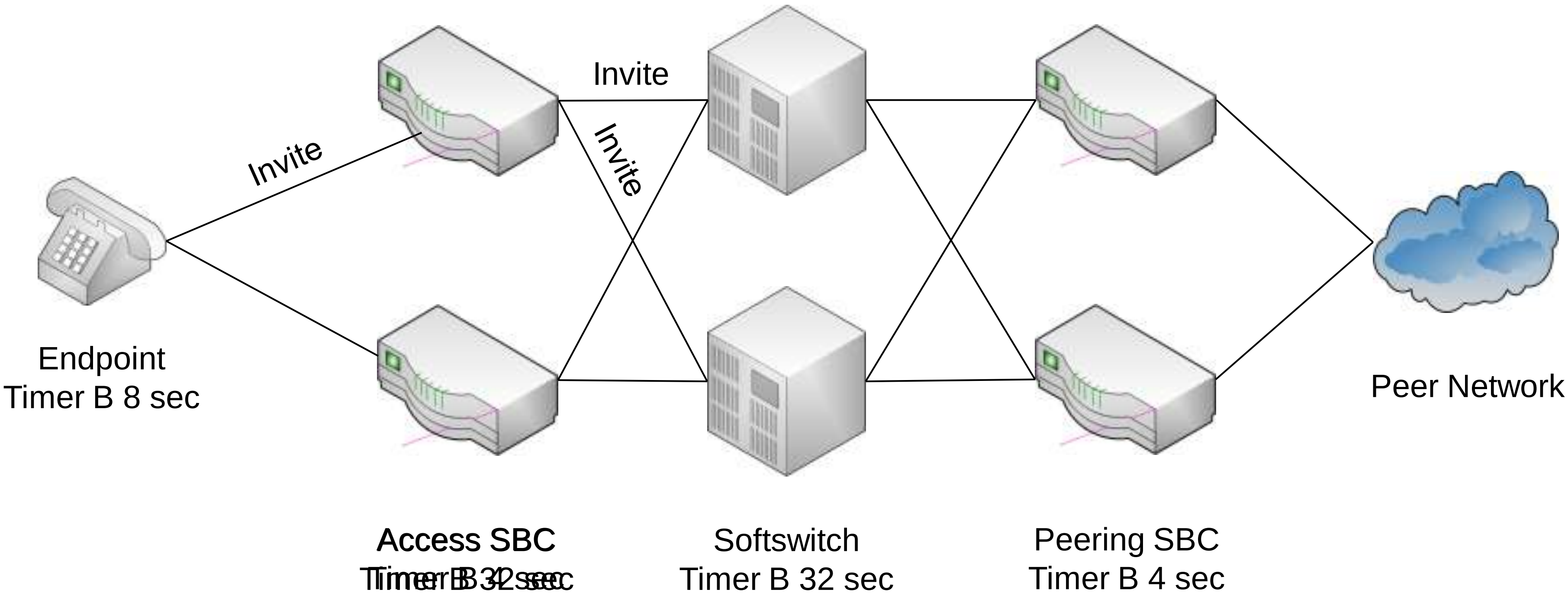
Challenges

- Not all elements will offer configurable timers
  - Many endpoints and softswitches have fixed timers
- Many vendors implement non-standard timers or counters in place of those defined by RFC 3261 (tisk-tisk)
- Over-the-top ITSP's must accurately budget round trip packet timers to correctly configure the T1 timer
- Retry timers must be calculated cumulatively from hop to hop so that a device doesn't route advance before its next-hop has a chance to resolve any upstream issues
  - The following slide will illustrate a "stranded" Invite transaction due to improperly configured Timer B values

Cumulative timer evaluations – Timer B (Invite Timeout)



Cumulative Timer Evaluations – Timer B (Invite Timeout)



- Because of poorly configured Timer B, call fails
  - If Endpoint fails to Cancel the failed Invites, call may progress through the network after the Endpoint abandons it

First A-SBC fails the Invite after 4 sec and the call advances to Secondary Softswitch

- This can be fixed by shortening Timer B in the A-SBC so it will advance before the Endpoint's Timer B expires

Cumulative Timer Evaluations - Conclusion

- By altering the timers in just one device, end-to-end call flows will be affected but not necessarily produce the desired outcome
- Timer values should be higher at the edge and decrease as the message progresses towards the network core
  - This ensures each element's timer is long enough to allow the next-hop to route-advance in case of a failure or impairment
- Over-the-top providers must be particularly sensitive to access-side timers to allow for a reasonable number of retries without unduly extending post-dial delay
- Timers can also be used to control congestion
  - Shorter timers mean fewer retransmitted messages, less congestion during failures
  - Timers set too short actually create excess traffic, negative call results

Example One

	Endpoint	A-SBC	Softswitch	Peering SBC	Peer Gateway
Timer T1	400ms	500ms	500ms	500ms	250ms
Timer T2	4s	4s	4s	4s	n/a
Timer B	4s	4s	6s	1s	1s
Timer F	8s	4s	n/a	n/a	n/a
Retrans for REGISTER	4	3	n/s	n/s	n/a
Retrans for INVITE	3	3	3	3	3

- In this example, the operator owns the end-to-end network
- There are two Access SBC’s and two Peering SBC’s with a geographically redundant softswitch in the middle

Example Two

	Endpoint	A-SBC	Softswitch	Peering SBC	Peer Gateway
Timer T1	650ms	500ms	500ms	500ms	250ms
Timer T2	4s	4s	4s	4s	n/a
Timer B	3s	2s	6s	1s	1s
Timer F	9s	4s	n/a	n/a	n/a
Retrans for REGISTER	4	3	n/s	n/s	n/a
Retrans for INVITE	2	3	3	3	3

- In this example, the operator is providing services over-the-top of a third party's access network
- There are three Access SBC's and two Peering SBC's with a single softswitch in the middle

Summary

- Accurately predicting a customer's experience requires end-to-end SIP timer engineering
  - Start by setting T1 to account for round trip packet times
  - Calculate Timer B and Timer F on each element to allow it's upstream element time to route advance
  - Make sure elements send Cancel before route advancing
- Timers can be used to "influence" congestion
  - During an outage/impairment, how many retransmits are useful? How many are superfluous?
- Test, test, then test some more!

Thank you for your time today…

QUESTIONS???