

ROOT CAUSE ANALYSIS



Zuul (authentication) outage

The following is a detailed accounting of the service interruption that Rally users experienced on June 24-25, 2015 beginning at 11:01 pm.

Root Cause Analysis Summary:

Event Date	June 24-25, 2015
Event Start	Multiple incidents: June 24 11:01 pm June 24 11:14 pm June 25 9:15 am
Downtime Start	same as above
Time Detected	same as above
Time Resolved	Multiple: June 24 - 11:12 pm June 24 - 11:20 pm June 25 - 9:18 am
Downtime End	same as above
Event End	same as above
Root Cause	Enabled haproxy load balancers that healthcheck against zuul every 5 seconds. The load caused by the increase in request rate went above what zuul could tolerate and all ALM services were disabled Total Downtime: 11 + 6 + 3 = 20 minutes Total Impaired Availability: 20 minutes Time to Detect: 0 minutes Time to Resolve: Avg 6.6 mins

Issues and Remediations:

Issues	Remediations
Temporary Zuul change	Zuul changed to respond immediately with a 200 so that it can handle the load from HAproxy. This is what is running now and allowed us to re-enable Zuul authentication.

Future Preventative Measures:

Actions	Description
Re-enable Zuul Healthcheck	Idea: Memoize the result from a healthcheck for some amount of time and respond with that.
HA Proxy Enhancement	Move HAproxy to a service based configuration. This would reduce the number of HAproxy requests.
Determine Zuul request rate capacity	Need to understand this. Loadtest Zuul to find out the answer. (Use the off DC to reduce risk of bringing down main prod instance)
Determine future of Zuul	EOL Zuul w/ a new auth solution or move to a new datastore