Main page
Recent changes
Server admin log (Prod)
Server admin log (RelEng)
Deployments
SRE/Operations Help
Incident status

Cloud VPS & Toolforge
Cloud VPS documentation
Toolforge documentation
Request Cloud VPS project
Server admin log (Cloud VPS)

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page

Print/export
Create a book
Download as PDF
Printable version

Incident documentation/20170419-ToolsProxy

< Incident documentation

Contents [hide]

Summary

The active Tools proxy filled its local disk and we were unable to compensate as it kept growing without an obvious reason. Andrew B and Chase P failed over to the secondary node in the cluster with a brief period of outage due to Nova returning errors and bad documentation. The secondary proxy displayed a similar growth on disk and a rogue kubelet process was found to be creating tons of requests to an errant local api-server that did not exist, but Nginx was there to continually log the errors creating issues. Stopping kubelet and an Nginx reload (with some manual stuck worker cleanup) seemed to stabilize growth.

Timeline

- Chase saw tools-proxy-01 alerts on lack of disk space in the wikimedia-labs IRC channel

```
[12:20:06] <chasemp> !log tools clean up disk space on tools-proxy-01
```

- Chase did some cleanup for logs and various ephemeral data but the disk space was still growing and outpacing cleanup
- Chase asked Andrew B if he had any ideas on why the disk usage on tools-proxy-01 was still increasing as `df -Th` and `du -sh` did not line up with usage. Andrew B surmised open file handles were eating space and could not release
- We decide to fail over to tools-proxy-02 to reboot tools-proxy-01 to see if we recover disk space
- Reviewing https://wikitech.wikimedia.org/wiki/Nova_Resource:Tools/Admin#WebProxy for fail over procedure several errors were noted and are to be addressed in bug T163390). Chase determines redis replication is indeed active and things on the proxy seem normal for failover.
- tools-proxy-01 is now at 100% capacity and can no longer take new webservice registrations or perform normal functions. There are no more logs or ephemeral items to clean up to compensate.
- tools-proxy-02 is at 14% disk usage for comparison
- Andrew issues the nova command to move the designated floating IP to tools-proxy-02
- Nova outputs errors and for a time registers the floating IP as incorrectly associated with both tools-proxy-01 and tools-proxy-02 simultaneously. tools.wmflabs.org goes down and alerts in the #wikimedia-operations channel.

```
[12:47:07] <icinga-wm> PROBLEM - Host tools.wmflabs.org is DOWN: CRITICAL -
Time to live exceeded (tools.wmflabs.org)
```

- nova finally takes the appropriate command to associate the floating IP
- Chase changes https://wikitech.wikimedia.org/wiki/Hiera:Tools to set 'active_proxy_host: tools-proxy-02'
- Chase runs clush commands to run puppet on exec and webgrid nodes
- Andrew runs puppet on labservices100[1|2] and tools-services-[01|02]

```
[12:53:57] <icinga-wm> RECOVERY - Host tools.wmflabs.org is UP: PING OK -
Packet loss = 0%, RTA = 2.08 ms
```

- tools.wmflabs.org is back and things stabilize
- Chase runs clush to ensure puppet is run across the fleet for good measure
- Chase notices tools-proxy-02 is up to 32% disk usage and climbing. Chase reduces the usage by cleaning up ephemeral logs and waits for outcome. Disk usage reduces but is still growing more than shrinking.
- Chase asks Bryan, Andrew and Madhu to hop on a hangout to debug
- Chase looks back at tools-proxy-01 which is now 85% usage. This seems like delayed release of resources as no other actions were taken on this instance since failover. Theory: As the instance was no longer continually growing the logs outpacing its ability to release space it finally was able to normalize.
- tools-proxy-02 shows constant errors logs in /var/log/nginx/error.log (many per second)

```
2017/04/19 15:17:55 [error] 19896#19896: *674120 open()
"/usr/share/nginx/html/api/v1/nodes" failed (2: No such file or directory),
client: 127.0.0.1, server: localhost, request: "GET /api/v1/nodes?
fieldSelector=metadata.name%3D127.0.0.1&resourceVersion=0 HTTP/1.1", host:
"127.0.0.1:8080"
```

- Chase recalls we had a similar issue of kube-proxy being misconfigured to contact the api-server at localhost on 2017-04-01 and yuvipanda was able to compensate with packaging changes (which caused this issue in that case as well)
- Bryan investigates where the kubelet config is coming from and it seems the answer is nowhere Theory: kubelet is installed as a dependency of kube-proxy but where we apply a good config for kube-proxy via Puppet for tools-proxy* hosts we do not do the same for the implicit kubelet in this case. Kubelet using the default config and the default api-server is continually spamming the local webserver.
- Madhuvishy stops kubelet and runs Puppet to see if Puppet attempts to actively manage the service and it does not. Kubelet stays down and the error log barrage ceases.
- On tools-proxy-01, Andrew issues a reload for Nginx as it seems unable to release workers effectively, and the workers in pending shutdown are holding open fd's on the nginx error log. Nginx workers remain in a pending state for >15 minutes.

```
20671 nginx: worker process is shutting down
```

```
root@tools-proxy-01:/# ls -l /proc/20670/fd
l-wx------ 1 www-data www-data 64 Apr 19 15:14 55 -> /var/log/nginx/error.log
```

- The group theorizes that the proxy_read_timeout in the Nginx configuration is not allowing the workers to gracefully exit in a timely manner
- Bryan `kill -9` 's the hung Nginx workers on tools-proxy-01
- Chase creates a Phabricator paste to track the storage growth on tools-proxy-02 to ensure it remains stable at (after >6h it has)
- Chase looks back at tools-proxy-01 which is now 54% usage with no action taken other than failing off of it (6h later it is at 22% with no other intervention)

Conclusions

We need to be careful with components of k8s getting included by default as it is not always benign. We currently are getting docker on the tools-bastion-03 hosts as well as an incidental. Our documentation for HA components within Tools is lacking and incorrect. We have settings in our reverse Nginx proxy that are meant to be one size fits all or to allow for edge cases that may not be sustainable.

Actionables

- Fix Tools Proxy failover documentation (bug T163390)
- Stop kubelet from running on the Tools Proxies (bug T163391)
- Investigate reducing proxy_read_timeout for nginx reload and worker sanity (bug T163393)

Category: Incident documentation