

Incident Portmortem

UCL Moodle Production is Down

by

Hugo Martinez

Catalyst IT Europe Limited

Version: 1.0

March 2023

Internal use only

Table of Contents

Introduction.....	2
Overview.....	2
Related WRs.....	2
Team allocation.....	2
Timeline.....	2
Root cause analysis.....	2
Conclusion.....	3
Lessons learned.....	3
Problem Resolution.....	3
Next steps.....	3
References.....	3
Relevant Matrix chat log.....	3
WRMS Tickets.....	3
Gitlab commits.....	3

Revision History			
Modified By	Date	Version	Change
Hugo Martinez	27/03/23	1.0	Initial Version

Introduction

Overview

The KMS key's grants were not available for ASG auto-scaling role. Because the role couldn't decrypt the AMI, new instances couldn't be spun up causing the cluster to run out of capacity.

Related WRs

- [WR#404019](#)
- [WR#403945](#)

Team allocation

Incident commander	Michiel Dethmers
Communications Lead	Michiel Dethmers
Operations Lead	Hugo Martinez
Reporting Lead	Hugo Martinez

Timeline

- Time of alert/issue being noticed 9:16.
- Time of identification 9:45.
- Time of issue being declared an incident 10:34.
- Incident response team allocated 9: 30.
- Sequence of findings and primary decision:
 - Confirmed that AMIs were not being decrypted.
 - Re-enabled KMS grants for ASG auto-scaling role.

Root cause analysis

It seems like the KMS grants for the ASG auto-scaling role were removed as part of the code update done on Friday. The KMS grants are moving from the ECS master Terraform modules to the AWS accounts module. The KMS grants and the ASG auto-scaling role are now considered prerequisites for the ECS master, so they need to be managed by a different part of the code. The KMS grants were set in the AWS account module before the staging update, but evidence indicates that when the code from staging was applied, it removed all the KMS grants, even though they were set twice by two different code bases. It's the first time we have seen this behaviour. We will need to update the code accordingly to avoid similar situations in the future.

For reference the code updating I'm referring too was done on the staging cluster on Friday in preparation for the code upgrade plan for production. As part of our recent infrastructure as code development we have made a number of changes and improvements that are now in the last testing phase where we get the upgrade process ready for production. During the upgrade we found that the Varnish module wasn't applying the new network configuration correctly which caused the staging site to be briefly unavailable on Friday. It was this same code upgrade what removed the correct grants from the ASG auto-scaling role. The first issues has been resolved and a code patch is being review. The second issue is now being investigated.

Conclusion

We need to update the Infra-code so it handles KMS grants for AMI in a different way.

Lessons learned

We need to update the Infra-code so it handles KMS grants for AMI in a different way.

Problem Resolution

Resetting the KMS grants for the ASG auto-scaling role.

Next steps

We need to update the Infra-code so it handles KMS grants for AMI in a different way.

References

Relevant Matrix chat log

WRMS Tickets

- WR#404019
- WR#403945

Gitlab commits