

Title

Storage Cluster outage Postmortem

Date

19/9 2018

Authors

Peter Kondrup

Status

The current status is complete, action items in progress

Summary

An outage of all three nodes in one of our minor storage clusters caused the whole system to fail.

Because of previous undetected outages of nodes in the Cluster it became unavailable when the last node stopped working.

No data has been lost during the outage.

Impact

All customers and system integrations was unable to use the system while the Cluster was unavailable from 10:29 to 10:48 (19 minutes).

Root Causes

Undetected outages of nodes previously made the system exposed to down time when the last node failed.

Missing monitoring of nodes to be aware of outages.

We don't know exactly why the nodes stop working.

Trigger

Outage of the last of three nodes in the storage cluster.

Resolution

A reboot and a re-sync of the cluster solved the issues

Detection

We got an alarm from our overall system monitoring 10:29

Action items

Ensure that we have monitoring on node failure in storage cluster (done)

Ensure that critical part of the system keeps working if possible (planned)

Lessons Learned

It was good, that our overall monitoring told us that something was wrong. Based on that we could start investigating what has happened and inform our customers.

It was very bad, that we was not aware of the outages in the cluster. We definitely need monitoring of all part of the system to detect if for example a node stop working in a storage cluster.

Timeline

10:29 Alarm from overall monitoring

10:36 Investigation started

10:40 Information to customers available online

10:48 Problem solved and everything was back to normal operation

Supporting information

None