High KDC load from user analysis jobs

Description

During the period from May 26th to July 3rd the KDCs at CERN got several times close to their load limits due to fast kinit request loop from individual experiment users from ATLAS.

Impact

- Both the Windows and Heimdal KDC services experienced over several weeks several periods of high load from individual clients due to high frequency "kinit" requests. The services were getting close to their load limit and there was a significant risk of affecting the availability of the CERN single sign-on service for all users.
- Protective measures had to be put in place in order to be able to react on high-load events in a fast way. Additional monitoring had to be put in place to quickly identify the originating users and to disable their access to avoid more global impact.

Time line of the incident

May 26th – A large number of token requests from individual ATLAS users observed (several thousand token request per second). First suspicion is a problem with the pool_extractIdentifier utility.

June 5th - Further analysis by the POOL, ROOT and XROOT teams showed a more general problem for user jobs connecting to CASTOR from POOL, ROOT or XROOT via the XROOT Kerberos authentication plugin.

June 7th – ATLAS analysis users have been contacted directly about ongoing high load and the suspected problem with XROOT security plugin in ROOT 5.26. Users are starting to migrate to ROOT 5.28 where possible. Additional teams using 5.26 are being identified from statistics obtained from the KDC logs and informed. Unfortunately not all users can easily migrate quickly to a new ROOT version.

June 9th – The problem in the XROOT security plugin has been identified and it has been confirmed to be already fixed in ROOT 5.28: a high frequency attempt to re-authenticate was triggered when the XROOT security plugin received an error from the lower-level Kerberos client library. The cause for this originating error was not clear yet – several candidate sources

were investigated. A back-port bug fix to remove the fast re-authentication loop for the 5.26 release series is being prepared by the ROOT team.

June 17 – The root cause triggering the XROOT authentication loop has been identified as an erroneous replay attack detection by the Kerberos client in SL5. A truncation of a timestamp exchanged leads to this problem for users trying contacting the server several times per second: (http://anonsvn.mit.edu/viewvc/krb5?view=revision&sortby=log&revision=20413).

This has allowed defining a workaround to the problem, which involves temporary disabling the replay cache by setting

$ export KRB5RCACHETYPE=none

until the experiments have fully moved to the bug-fixed ROOT versions or the Kerberos client library problem has been fixed.

June 30th – the bug-fixed version 5.26f is in production for ATLAS and users are migrating to this release.

July 3rd – The last CASTOR instance has been upgraded to include the workaround of disabling the replay cache on the XROOT server. The load on the KDC instances at CERN is normal again.

Analysis

After some initial suspicion of a pool command line utility or a service configuration change as the root cause, the xroot security plugin for Kerberos has been identified as the origin of the frequent kinit requests. The code problem had already been fixed back in ROOT release 5.28a, but this version was not yet in production by ATLAS and several other experiments. The analysis of the problem took significant time as the repeated kinit requests were only triggered when the xroot plugin received an authentication error from the Kerberos client library, which takes place only under certain conditions. One of these conditions was found to be a skew in the clock synchronization of the service machines, but this has been quickly ruled out as major contribution by a synchronization check of all service machines.

A further and dominating contribution was found to be a bug in the Kerberos client library itself, which truncates the timestamp exchanged between server and client during the authentication process. This timestamp truncation leads during periods of high request rate from a single client to a false detection of replay attacks by the

Kerberos client library – which then in turn triggered the problem in the xroot security plugin resulting in a kinit loop. A temporary workaround has been introduced on all CASTOR XROOT servers to avoid the impact of experiment users who cannot quickly upgrade

We would like to point out that the risk mitigation, analysis and resolution this particular incident was only possible with the contribution from many different areas including: Experiment responsible, POOL, ROOT, XROOT, CASTOR and the Kerberos teams.

Follow ups

- The ROOT team has very quickly reacted and back-ported the security plugin bug-fix from 5.28a to the ROOT versions used by the experiments in production (eg 5.26 for ATLAS, 5.27/06 for CMS). This change successfully prevented the re-occurrence of the kinit loop, but would have required a full experiment software re-deployment on the experiment side for to reach all affected users. The latter was not considered feasible on a short time scale.
- Since the root cause for triggering the kinit loop was identified as a false replay attack detection a workaround on the xroot server side was introduced to disable this detection until either the Kerberos client problem has been fixed in SL5 or the experiment builds with affected root versions have been fully replaced by corrected root versions. With this workaround in place, both KDC services run stable also with experiment clients running older, still faulty builds.