## Open Science Grid Bulk Host Certificate Use Case

June 27th, 2012

Von Welch (vwelch@indiana.edu)

## Use Case: Requesting Multiple Host Certificates in a Single Action

A number of sites on the Open Science Grid (OSG) operate clusters comprised of a large number (hundreds-to-thousands) of computer systems. For various technical reasons each system in these clusters often needs its own host certificate (e.g., the use of the GLExec system[[1]](#footnote-1)).

Since manual methods of obtaining individual host certificates would be cumbersome at this scale, the OSG with its current PKI provider, the DOE Grids CA, has developed the technique of *bulk host certificate requests*. Simply put this allows an appropriately authorized registration authority agent (a “Grid Admin”) to request multiple certificates in a single administrative operation instead of having to request each certificate individually.

The number of certificates that a Grid Admin can request in a single request is limited to 50[[2]](#footnote-2). This is to limit the impact of accidents and help prevent overloading of resources. This means 50 certificates is a common number of certificates to be requested in a bulk certificate request and that number is used as an example throughout this document.

## DOE Grids PKI Implementation

With the current OSG PKI provider, the Grid Admin runs a script[[3]](#footnote-3) that accepts a list of hostnames for which certificates are desired as input and iterates through that list, creating a CSR for each, contacting the DOE Grids CA with the request, retrieving the resulting certificate in a synchronous manner, and writing that certificate along with its associated private key on the local filesystem.

Since Grid Admins are authorized agents for the DNS domain that their host certificates fall in, meaning no manual vetting is required, and the DOE Grids CA responses fairly quickly (a few seconds) to a certificate signing request, the process of requesting 50 certificates usually takes a couple of minutes. Grid Admins are, roughly speaking, used to starting the bulk request scripts, going and getting a cop of coffee, and having it complete by they time they get back to their desk.

## Challenge with Implementation with the DigiCert API

Our goal is to continue providing our use community with the same experience they are used to today with the DigiCert API, that is in the course of a couple of minutes they can request and retrieve 50 certificates.

First, we note the following differences between the DigiCert API and our current DOE Grids API:

1. The DigiCert API requires two calls instead of one to issue a host certificate. With the DOE Grids CA a single call takes a CSR as input and returns a certificate. With the DigiCert API are two calls, the first of which provides the CSR and then second of which retrieves the resulting certificate.
2. The time between requesting a host certificate and its issuance it longer and in practice seems to be unbounded. That is after making the call to request a certificate, there is a period after that when calls to retrieve the certificate will fail. We have been informed by DigiCert this period should be a minute, but we’ve routinely seen times approaching ten minutes.
3. We have no method other than attempting to retrieve a certificate to determine if it has been issued. However, DigiCert has stated that frequent attempts to retrieve a certificate overload their system and can slow down the certificate issuance process described in the previous bullet.

These factors put us in a difficult situation: we’d like to complete a bulk certificate request as quickly as possible (with a couple of minutes), but to obtain all the requested certificates we are required to poll DigiCert to see if they have been issued, which in turns has the potential to slow down their issuance.

## OSG’s Current DigiCert API Implementation

To help understand the situation, we briefly describe our current approach with the DigiCert API in a simplified form, omitting unnecessary detail.

As described previously, with the DOE Grids CA, bulk certificate requests were simply a matter of iterating through the desired certificates, requesting each in turn. With the DigiCert API since there is a minimum of one minute from when we request a certificate until when it is ready, continuing with this serial approach would mean that it would take on the order of an hour for 50 certificates to be issued. Hence we have adopted a parallel approach as follows.

Currently OSG logic in performing a bulk host certificate request with the DigiCert API is to request all 50 certificates in rapid succession. We then wait for 60 seconds (indicated to us as the approximate time it will take for certificates to become available for retrieval) and then attempt retrieval of all certificates.

If all 50 certificates are obtained, the operation is successfully completed.

If some certificates cannot be retrieved yet, we wait for 5 seconds (a period indicated to use by DigiCert as a reasonable period) and then attempt retrieval of all certificates not yet retrieved.

After that, if all certificates have been obtained, the operation is successfully completed. If not, we again wait 5 seconds and attempt retrieval of all missing certificates, and so forth.

## Suggested Improvements

Since DigiCert understands their infrastructure well, it is our hope that with the understanding of the problem this document provides, they can best architect and implement an appropriate solution.

We do however offer the following suggestions:

1. Given the current implementation, have a well-defined process we should follow for requesting and retrieving multiple certificates. We’ve derived the current “wait 60 seconds and then attempt repeated with a 5 second delay” process based on some back and forth, but hopefully this document will provide DigiCert with understanding of the big picture and they can provide a preferred process.
2. Provide a firm bound on the time period from request to availability for retrieval. One minute has been stated as a goal, but the firmer that is, the fewer requests we have to make of DigiCert infrastructure during polling.
3. Having to be concerned about overloading DigiCert API services by issuing too many requests makes us concerned about their fragility. Ideally they would be robust enough to handle a higher number of requests without concern.
4. Implement API calls that allow for operation on multiple certificates as a single action. Our implementation would be simplified if we could, with one call, request or retrieve N certificates.
5. Implement a mechanism to notify us when a certificate is ready to be retrieved. Currently, we must poll (repeated attempt retrieval) until a certificate is ready, which as discussed previous puts undue load on DigiCert systems. Allowing us to register a callback or similar mechanism for DigiCert to inform us when a certificate(s) were ready for retrieval would allow us to avoid this.

1. https://wiki.nikhef.nl/grid/GLExec [↑](#footnote-ref-1)
2. They are also limited to one such request per day. [↑](#footnote-ref-2)
3. https://vdt.cs.wisc.edu/svn/software/osg-cert-scripts/bin/multi-cert-gridadmin [↑](#footnote-ref-3)