Request for Comments: 7322 Mohammed Siddiq

Category: Architecture May 2024

Enhancing File storage

Implementing Needle in Haystack

Abstract:

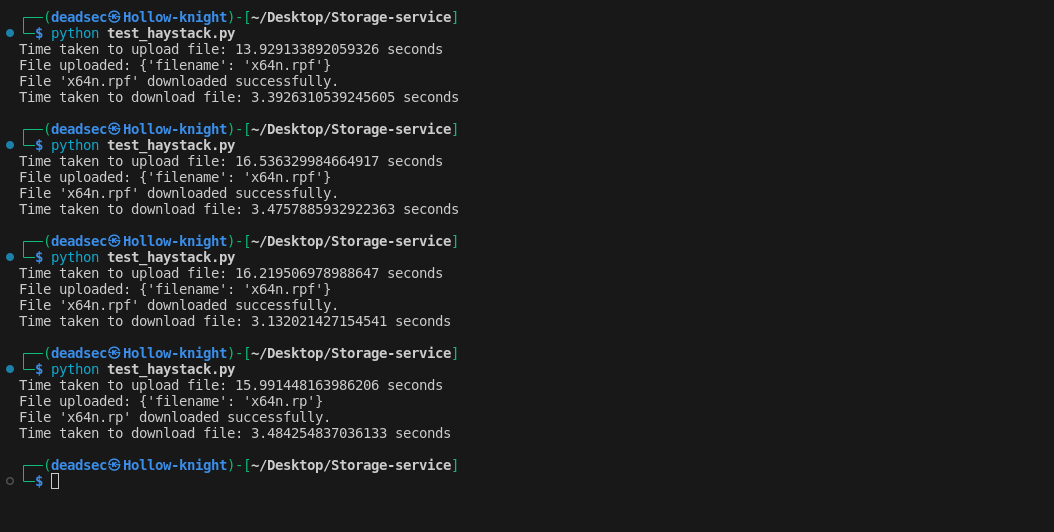
This document describes the how our current implementation of Object Storage Service lacks compared to Sample haystack implementation. The documentation contains testing of current storage implementation as well as haystack. It also highlights missing applications in haystack which needs to be implemented, as well as other necessary modifications.

The Great Testing:

The test was run on Lenovo Legion Y740-I5RH

System Specifications: i7-9750H [cores:6, hreads12, max: 4.50 GHz, base: 2.60 GHz], 16gb ram, GTX 1660 ti, 512gb SSD,1TB HHD.

The dummy file used for testing purpose is “**x64n.rpf**” which is 1.6gb in size the below image shows the testing time taken to upload as well as download the given file on a haystack storage service model.



The same test couldn’t be performed on Current Storage service as it crashed the system due to high CPU usage while base64 encoding of dummy file

Switching from JSON to SQLite:

To store keys,offset and size of upload it would be more efficient to switch to SQLite db than to use a json, where in casse of million of keys, offset , and other pair values , it would be inefficient to store them in json and open it than read its content and than search for the key.

The methodology in which keys, offset and size of files can be stored in DB is that, a key can represent each table in db [which is ofcourse unique ] and the offset and size of multiple files can be stored in table columns.

Or we can can have a Table in Database whose one column has keys and the other column has the size and offset.

| key1 | offset,size |

| key1 | offset,size |

| key2 | offset,size |

Handling Update endpoint:

Just like the current implementation a key to existing object in haystack is provided with the content that needs to be appended/updated , the haystack connects to the table with the key name and at the end of table appens the new values of uploaded content.

Logging

Logging of all updates in storage service