Regional storage support for LSST related science

SAŠA SIMIĆ

FACULTY OF SCIENCES, DEPARTMENT OF PHYSICS,
UNIVERSITY OF KRAGUJEVAC, SERBIA

Regional storage support for LSST related science

- LSST collaboration in order to encourage and support preparations of teams and participants has initiated this kickstarter project.

- Kickstarter projects are founded by Heising-Simons foundation and organized by LSST collaboration.

- As it is stated on their site, the goals of this project are:
  - to promote excellence in Rubin-related research by enabling science collaboration members to dedicate time to preparatory work and
  - to ensure that research opportunities and facilities are accessible on a fair and equitable basis to all science collaboration members and to overcome common barriers to entry and participation

- Who can participate: Individuals/Team or Partnership.

- Who is eligible: LSST Science collaborations TVS, SMWLV and SSSC.

Regional storage support for LSST related science

- Scope of activities:
  - Science Collaboration task force activities
  - Student training
  - Faculty summer salary
  - Institutional research partnerships
  - Supporting service contributions to the Science Collaborations
  - Enabling research under unusual circumstances, e.g. caregiving in a pandemic
  - <u>Equipment, software and service costs</u>
 - Support for junior/soft-money researchers
  - Collaboration-based Training
  - Visiting Scientist Program

Regional storage support for LSST related science

- Our partnership include scientist from two groups related to LSST science.
  - Serbian team – SER-SAG with involved institutions:
    - Faculty of Sciences, University of Kragujevac (UniKg – PI);
    - Astronomical Observatory Belgrade (AOB);
    - Faculty of Mathematics, University of Belgrade (UniBg);
  - Croatian team – CPG
    - Department of Physics, University of Rijeka (UniRi);
    - Ruđer Bošković Institute, Zagreb (RBI)
    - Hvar Observatory of the University of Zagreb's Faculty of Geodesy (HO).
- The main goal of our proposal is to improve the readiness of teams involved in LSST operations.
- Our intention was to improve our research capacities in the areas of data storage, processing, and administration for LSST-related science.

Regional storage support for LSST related science

- Both teams have their in-kind contributions to LSST project.

- The SER-SAG group work on frequency-time domain periodicity mining of Active Galactic Nuclei light curves.

- The SER-SAG team has access to the supercomputer in state owned data center in Kragujevac.

- The in-kind of CPG team will be explained in next talk.

- The CPG hosts HPC facility Bura which will be one of the LSST Scientific Processing Centre.

- So, realisation of this project allow us to share resources between two teams (SER-SAG and CPG) in process of preparation for beginning of work of LSST telescope.
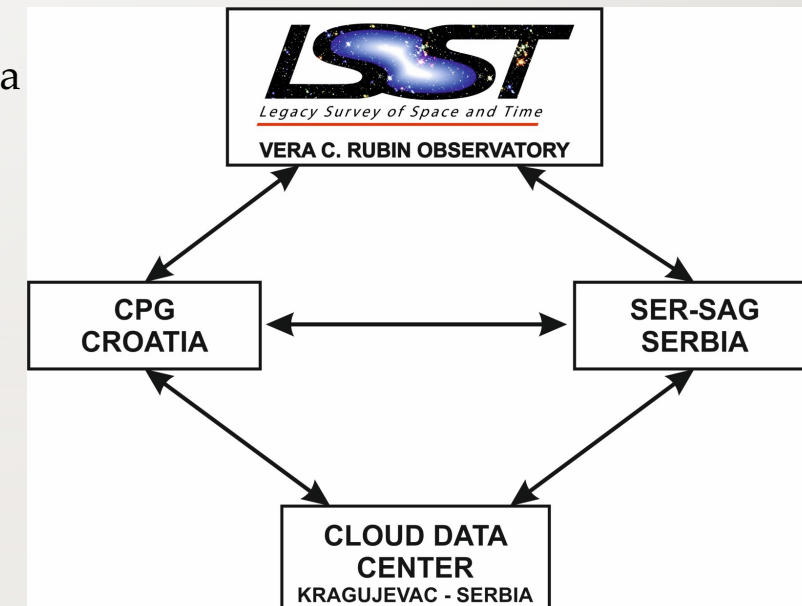
Regional storage support for LSST related science – results

- Since the project is officially finished, let me present achieved results:

- We stated that major aims of the project were:
  - (i) installation of data storage for the purpose of supporting the intermediate and final products of LSST SER-SAG and CPG in-kind contributions,
  - (ii) development of regional scientific cooperation between SER-SAG and CPG in-kind teams and
  - (iii) development of astronomy at the University of Kragujevac in Serbia.

- The first and main task is accomplished with the building of the 250TB cloud data storage at the Faculty of Sciences at the University of Kragujevac, for storing the large amount of intermediate and final outcomes of scientific analysis of LSST data. The cloud data storage is fully operational, and resources are available to be shared among the partnerships involved for the LSST related scientific research. Currently, all necessary services on the cloud server are ready and can be distributed among the scientists from participating institutions.

Regional storage support for LSST related science – results

- The project and cloud data storage were presented to both in-kind teams as well to the wider scientific community during:

  - V Conference on Active Galactic Nuclei and Gravitational Lenses (V AGN & GL).

  - University of Rijeka and University of Belgrade, during mutual exchange of visits.

  - TVS science colloquia which is agreed with the TVS group leader to be held on November 7. at 22:00h CEST.

- The kickstarter project has been promoted to students who expressed interest to help project implementation, and thus learn about the LSST and its related science. This resulted with one student summer internship, organized during the kickstarter project at the University of Kragujevac.

Regional storage support for LSST related science - hardware component

- For this purpose, we purchase following hardware:

  - QNAP TS-h1677XU-RP - NAS 16 bay storage

  - 16 x Seagate 18TB Exon HDD

  - 2 x Western Digital 2TB SSD

  - APC smart 1500 UPS

  - Additional support high speed (10GB/s) LAN cables, extension card and support monitor

Regional storage support for LSST related science - hardware component

Inside view of QNAP NAS storage

Regional storage support for LSST related science - hardware component

One of 16 Seagate 18TB HDD and mount

Regional storage support for LSST related science - hardware component

QNAP NAS
Storage mounted
in rack

Regional storage support for LSST related science - hardware component

APC smart UPS mounted in the rack

Regional storage support for LSST related science - hardware component

All equipment rack mounted
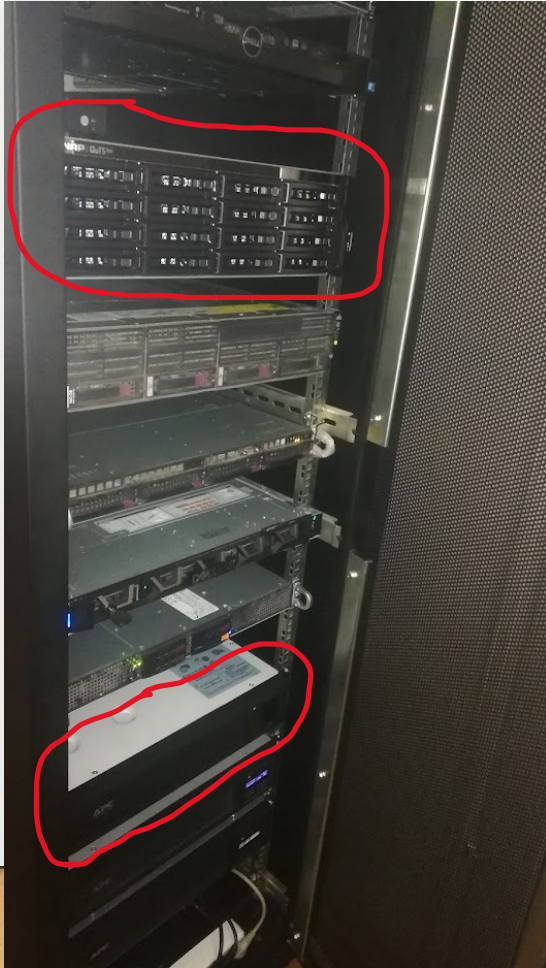
Regional storage support for LSST related science – software component - outline

- Individual and group account policy

- Accessing the data storage
  - Login to server

- Keeping the data on the server
  - Synchronizing the data
  - Backup the data
  - Mobile devices connection

- Sharing data with others, shared directory

- Services
  - ftp, web, ssh, VPN, webDAV

- Data security
  - Antivirus, IP blocking, SSL certificate, RAID

Regional storage support for LSST related science – user accounts

- Users place their data in the Home directory, or any other directory for which they have access.

- The data can be accessed in different ways. Logging directly to the server and using **FileStation5** file manager.

- All standard functions for file mange are implemented and available.

- Need to have in mind possible restrictions on some files/dirs, regarding account limitations.
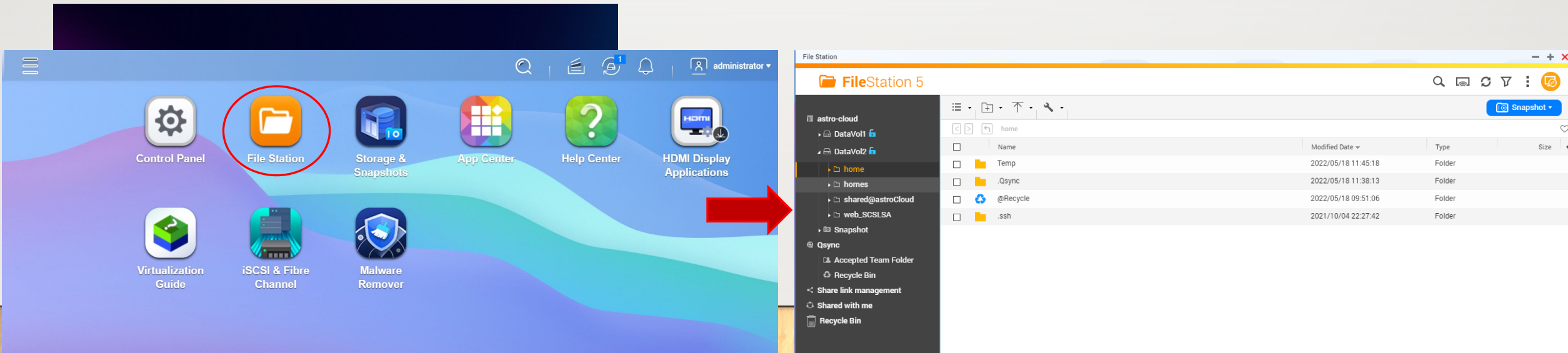
Regional storage support for LSST related science – automatic data management

- File synchronization using Qsync - means that …

- Allows real time synchronization of data on your computer and data on ser-sag.storage

- It is necessary to have client program installed on your computer.

- It is necessary to pair directories on your local drive and directory on the server.

- Qsync client automatically start with OS. There are versions for Win, Mac and Linux.

- All files are then recorded in its latest version and on both places (server and local PC).

Regional storage support for LSST related science – automatic data management

Regional storage support for LSST related science – automatic data management

- Managing data with mobile devices.

- Qfile – remote access to files on NAS

- Qsync – sync/backup data with NAS
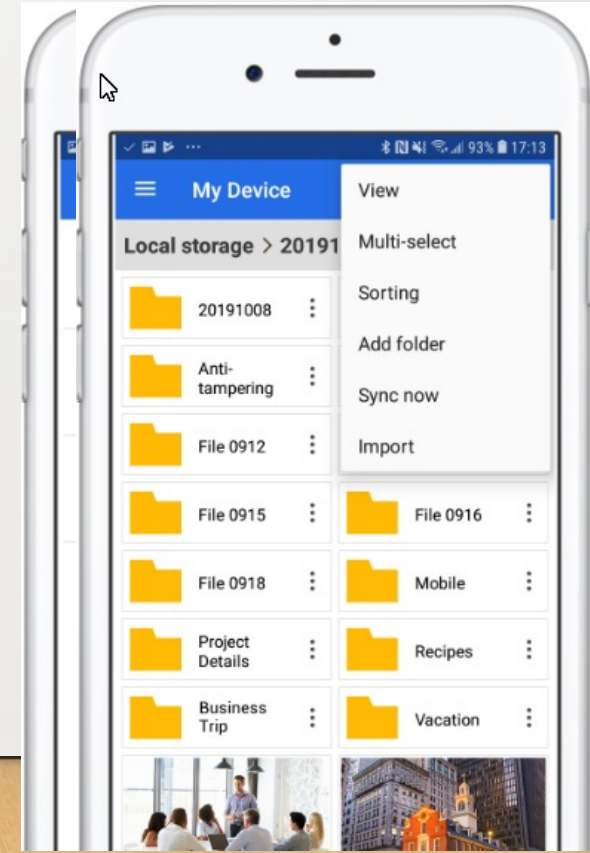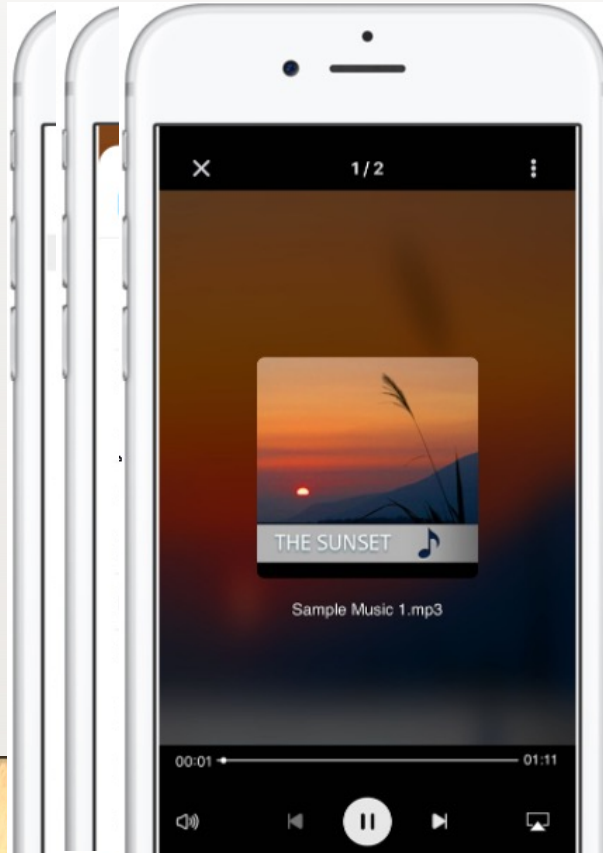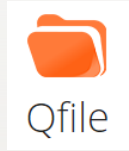
Regional storage support for LSST related science – shared directoriums

- In case of collaborative work maybe it is best to organize data within shared directoriums.

- Shared directoriums are created outside of users Home directoriums.

- In this case, multiple persons can access to directorium contents in same time.

- Shared directoriums can be seen from outside QNAP environment.

- Rules of access requlate what user can do.

- Can be used in different third party softwares

- Can be accessed by programing languages (python, C++, …)

- Security is not disturbed by multiple access.
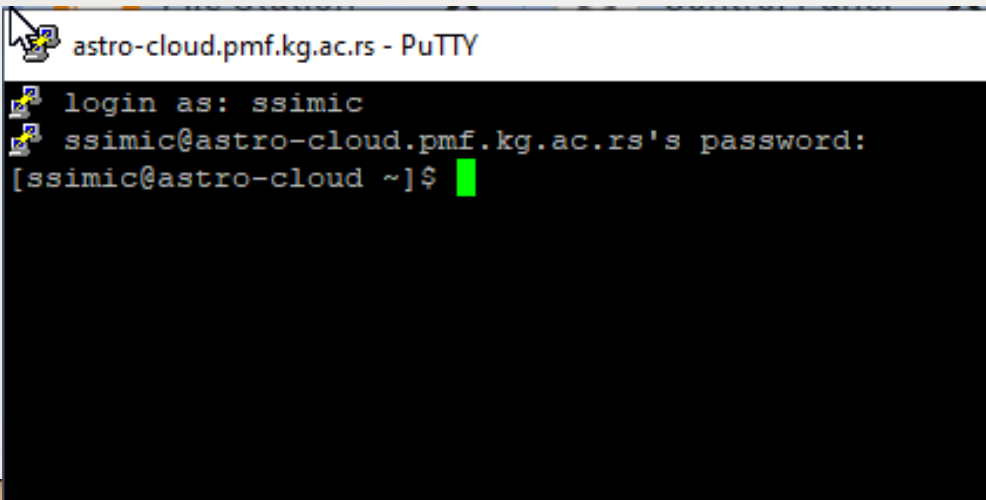
Regional storage support for LSST related science – shared d...

- Accessing it by typing server and directorium na...
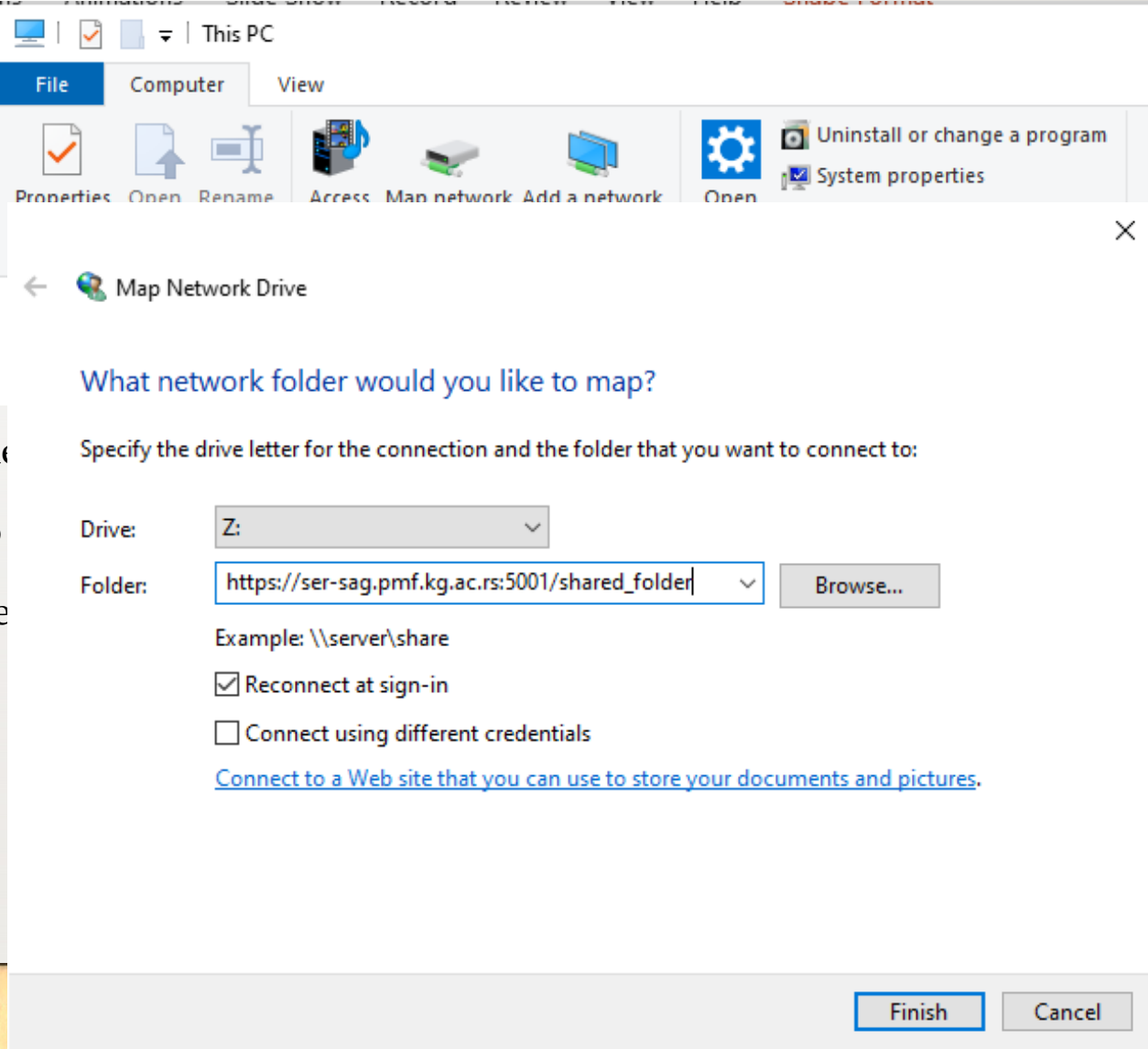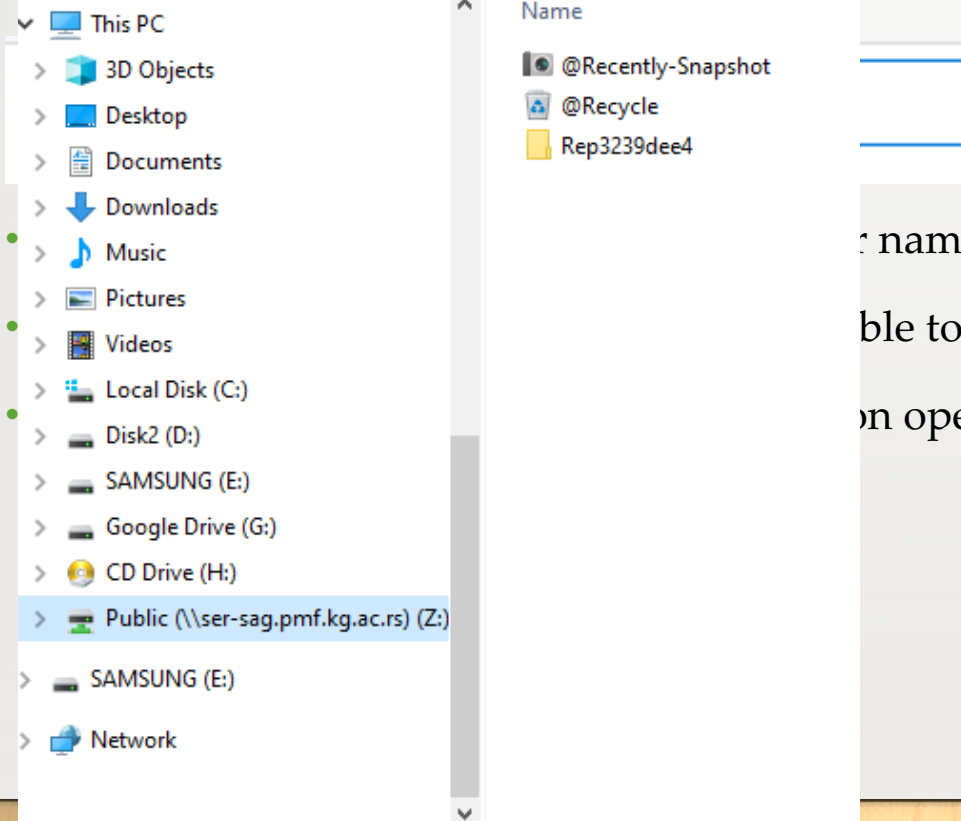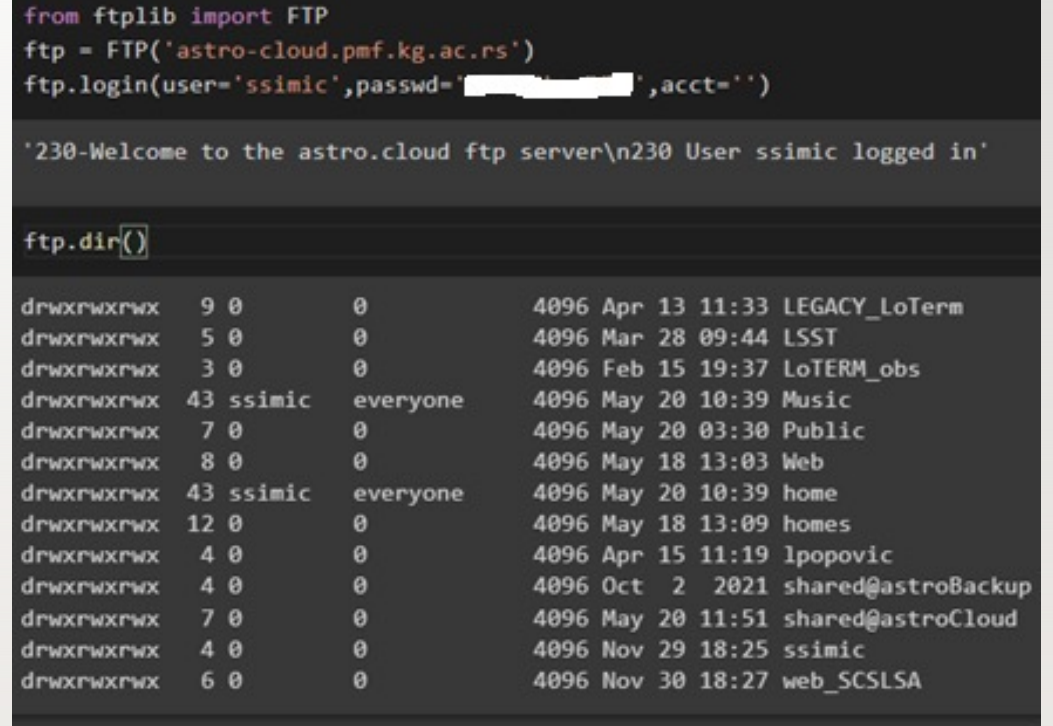
Regional storage support for LSST related science – shared directoriums

- Accessing programmatically shared directoriums.

- Most common way is by using ftp.

- Data on the server can be accessed via ssh protocol.

Regional storage support for LSST related science – web server

- Data storage features Web Server.

- Allows implementation of Internet site.

- This is convenient and allows access to customized browser user interface. (presentation of computation or data in browser, or using the software via web browser)

Web Server

After enabling this function, you can upload the webpage files to "Web" network share to publish your website.

☑ Enable Web Server ⓘ

Port number: 80

☐ Enable HTTP compression

☐ Enable secure connection (HTTPS)

TLS version compatibility: 1.2 and later ▼ ⓘ

☐ Enable strong cipher suites ⓘ

Port number: 8081

☐ Force secure connection (HTTPS) only

Maximum number of clients: 64

☑ Do not allow Web Server embedding in IFrames

Allow specific websites to embed Web Server in IFrames   Allowed Websites
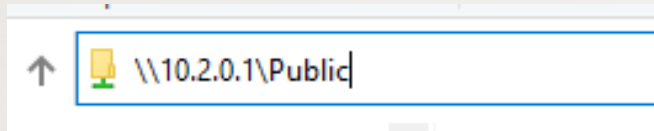
☐ Enable X-Content-Type-Options HTTP header ⓘ

☐ Enable Content-Security-Policy HTTP header ⓘ

After enabling this service, click the following link to enter to Web Server.

Regional storage support for LSST related science – VPN

- For more security and efficiency use VPN (Virtual Private

- Simulate like you are user in the local netwok.

- Allows secure, encrypted connection.

- Need to create VPN connection on your local computer.

↑ 🖳 \\10.2.0.1\Public

Add a VPN connection

VPN provider
| Windows (built-in) | ⌄ |

Connection name
| ser-sagVPN |

Server name or address
| ser-sag.pmf.kg.ac.rs |

VPN type
| L2TP/IPsec with pre-shared key | ⌄ |

Pre-shared key
| •••••• | ⊚ |

Type of sign-in info
| User name and pa |

User name (optional
| |

Password (optional)
| |

🖳 Network 3
Connected

⠲ QNAP_ser-sag
Connected

Save   Cancel

Regional storage support for LSST related science – security

Regional storage support for LSST related science – security

- Account protection.

- An SSL certificate is a digital certificate

Regional storage support for LSST related science – security

- RAID on 16 disks system.

- Redundant Array of Independent Disks (RAID) is a virtual disk technology that combines multiple physical drives into one unit. RAID can create redundancy, improve performance, or do both.

- We need to have small discussion on how to organize server disks in most efficient way to our demands.

- The main trade is do we need data server of large capacity, higher security or greater speed.

RAID mode	Description	Operation	Advantages	Disadvantages	Recovery
RAID 0	Striped disks	Data is split evenly between two or more disks.	Large size and the fastest speed.	No redundancy.	If one or more drives fails, this results in array failure.
RAID 1	Mirrored disks	Two or more drives have identical data on them.	A single drive failure will not result in data loss.	Speed and size is limited by the slowest and smallest disk.	Only one drive is needed for recovery.
RAID 3	Striped set with dedicated parity	Data is split evenly between two or more disks, plus a dedicated drive for parity storage.	High speeds for sequential read/write operations.	Poor performance for multiple simultaneous instructions.	A single drive failure will rebuild.
RAID 5	Striped disks with distributed parity	Data is split evenly between three or more disks. Parity is split between disks.	Large size, fast speed, and redundancy.	The total array size is reduced by parity.	A single drive failure will rebuild.
RAID 10	1+0; Striped set of Mirrored Subset	Four or more drives are made into two mirrors that are striped.	Larger size and higher speed than RAID-1, and more redundancy than RAID-0.	No parity.	Only one drive in a mirrored set can fail.

Thank you for attention