NEOSTEL DATA PROCESSING CHAIN

Jiří Doubek[(1)], Agnieszka Sybilska[(2)], Piotr Duźniak[(3)], Piotr Sybilski[(2)], Grzegorz Lech[(2)]

[(1)] Iguassu Software Systems, Evropská 120, Praha 6, 160 00, Czech Rep. Email: jiri.doubek@iguassu.cz
[(2)] Sybilla Technologies, Toruńska 59/004, 85-023 Bydgoszcz, Poland
Email:agnieszka.sybilska@sybillatechnologies.com
[(3)] GMV Innovating Solutions Sp. z O. O., Ul. Hrubieszowska 2, Warsaw, 01-209 Poland Email:pduzniak@gmv.com

ABSTRACT

The NEOSTEL (Fly-Eye) telescope will detect NEOs through wide-field survey observations. The Data Processing Chain (DPC) will simultaneously process images from all cameras, extract detections, build tracklets and check it with the existing catalogue, identify new discoveries and provide SST functionality. The processing is divided into two main components – LightEngine and the Automatic Asteroid and Satellite Processor (ASAP). LightEngine pre-processes raw data and provides plate solution of images in near real time using a variety of catalogues. Identified detections from all cameras are passed to the ASAP, which combines them into a single field and builds tracklets from several images over the same sky region. Tracklets are compared to the Minor Planet Center (MPC) asteroid catalogue and reviewed by an operator. Additional tracking data can be requested through the Scheduling and Command Message (SCM) provided through the Orbit Calculator module, making use of a ranging method for orbit determination.

1 INTRODUCTION

Within the ESA's SSA programme, the system in charge of detecting natural objects such as asteroids, which can potentially impact Earth and cause damage, is the Near-Earth Objects (NEO) segment. Data on NEOs are collected from telescopes and radar systems worldwide. Each of these submits observations to the Minor Planet Center (MPC), which acts as a central clearing house for asteroid and comet observations. The measurements collected are retrieved by the NEO Coordination Centre (NEOCC).

Regular sky surveys are important part of the NEO discoveries. The same part of sky is scanned several times in order to find moving objects. For performing the successful surveys, a telescope with a wide field of view (FoV) shall be available as well as the effective way how to visit the different sky regions. For this purpose the Near Earth Object Survey TELescope (NEOSTEL) is under construction. After it becomes operational it will produce a big amount of data that are needed to be processed in near-real time. For this purpose the NEOSTEL Data Processing Chain is being developed. The NEOSTEL DPC has several components which are responsible for the following parts:

- LightEngine – pre-processing of raw data, astrometric and photometric data reduction.
- ASAP – building tracklets, comparing them with existing asteroid catalogue, reporting tracklets to the MPC.
- Orbit Calculator – providing the NEO scores and other statistics for the observed tracklets.
- Archiving subsystem – storing all data in a redundant system and providing additional services to the other software parts.

All components are further described in the next sections.

2 LIGHT ENGINE

LightEngine (LE) pre-processes frames retrieved from the Archive, performs astrometry and photometry and provides the results to ASAP (as shown in Figure 1, where the protocols used in communication between LE components, ASAP and the Archive are presented). The Figure 1 diagram shows all connections, not a specific data flow.

- The Archive notifies the pre-processing and processing services about new frames via RabbitMQ.
- The pre-processing service uses REST to download and upload frames from/to AstroDrive.
- The processing service uses REST to download and upload frames from/to AstroDrive and notifies ASAP about processing completion via RabbitMQ. It also uses REST to communicate with the catalogue service.
- AstroDrive has its dedicated database manipulated via PostgreSQL engine. AstroDrive forwards frame storage and retrieval operations to the Archive via Minio.
- The thumbnail generator service uses REST to download frames from AstroDrive.

Proc. 1st NEO and Debris Detection Conference, Darmstadt, Germany, 22-24 January 2019, published by the ESA Space Safety Programme Office
Ed. T. Flohrer, R. Jehn, F. Schmitz (http://neo-sst-conference.sdo.esoc.esa.int, January 2019)

- ASAP downloads frame solutions from AstroDrive and thumbnails from the thumbnail generator service via REST.

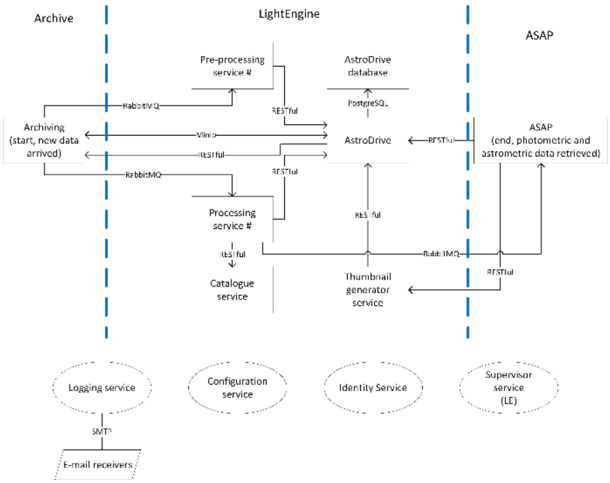


Figure 1 Interactions between LightEngine components and other services

LightEngine is divided in the following software components:

Component	**Description**
Pre-processing service (PPS)	The pre-processing service is responsible for: • Basic correction of the frames. • Combining darks, biases, flats into master darks, master flats, master biases. • Applying master calibration frames on the scientific frames. Multiple instances of the pre-processing service can be running simultaneously. It is expected that there will be one instance running for each camera.

Processing service (PS)	The processing service is responsible for:
	• Astrometry.
	  ○ Feature detection.
	  ○ Plate solving.
	  ○ Catalogue matching.
	• Photometry.
	Multiple instance of the processing service can be running simultaneously.

AstroDrive (AD)	AstroDrive serves as the Archiving Service's overlay for the most of the LE. It provides a way to obtain and store frames and frame solutions.
Thumbnail generator service (TGS)	The thumbnail generator provides ASAP with frame thumbnails based on requested parameters.
Catalogue service (CatS)	The catalogue service provides the processing service with requested catalogue data.
Generally available services developed by Sybilla Technologies for LightEngine and other services used in NEO-DPC	
Supervisor service (SS)	The supervisor monitors other services' health and is able to restart them in case of failure.
Logging service (LS)	The logging service serves as a universal sink for logs emitted from the other services. The service can send e-mail notifications to configured receivers.
Configuration service (ConS)	The config service is responsible for providing other services with configuration data.
Identity service (IS)	All clients connecting to the services in the intranet from outside of the intranet (e.g. a CLI app, a Web app) must be authenticated and authorized. The service provides authentication and authorization services for both the outside and inner entities.

All LightEngine services with user interface will provide the current status of the services via SignalR updates and allow for retrieval of the historical data via RESTful API, requesting the data. Commands will be executed via RESTful API and monitored via SignalR. Each service will provide minimal status information BasicData and minimal set of commands listed in Table 1.

Table 1 BasicData for each service.

Property name	**Type**	**Description**
Id	String (UUID)	Unique ID of the service.
Name	String	The name of the service.
ServiceState	String (enum): unknown, starting, started, normal operation, malfunction, permanent failure, stopping, stopped	High level state description of the service. Initial state unknown is just after the creation of the service, starting is usually automatically initiated.
StartTime	String (date and time)	Service start time

ProductionTime	String (date and time)	Information provided in state creation time.
Decay	String (duration)	Time during which the information provided in status may be trusted (usually three times the typical update loop).
IsOperational	Boolean	Simple Boolean flag showing if the service is ready to accept new request, fulfil its role in the system.

3 ASAP

The Asteroid and Satellite Automatic Processor (ASAP) is software that was developed for processing space debris and asteroids images from optical sensors in the context of ESA Test-Bed Telescopes. It is being redesigned for the NEOSTEL DPC to work effectively for the large FoV and provide results in near-real time.

ASAP contains several services which are responsible for specific tasks. The first is the Detection Collector Service (DCS) which collects astrometric data coming from LightEngine and stores them in a local database for further processing. After all images from the same survey field are ready, the DCS inform the Loners and Tracklets Service (LTS). It combines data from all images, builds tracklets and store them in the database. Then the MPC Service (MPC) compares the found tracklets with the asteroid catalogue and identifies potential new discoveries. The human operator can go through the discoveries on a web page where the tracklet information and thumbnails are shown. Confirmed asteroids are automatically sent to the MPC.

The Detection Collector Service (DCS) receives the astrometric and photometric data from LightEngine. For each individual camera it includes the plate solution and details for the stellar and non-stellar detections. The DCS stores all information into the ASAP database and checks when images from one field are completed.

The Loners and Tracklets Service collects astrometric data for a specific field and combines information from all cameras to a single frame (Figure 2).

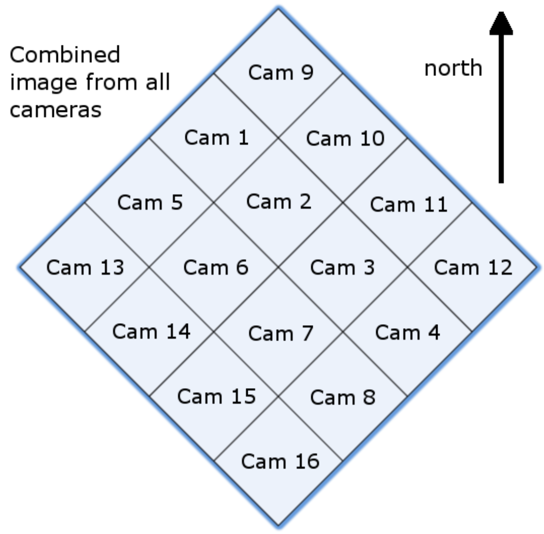


Figure 2. Combined image from 16 NEOSTEL cameras.

Several combined images from a same tessellation field are part of a sequence. Such sequence is used for building potential asteroid tracklets.

Tracklets are formed from loners. Loners are detections that are present on one combined image but not on the others. It shall be noted that loners can be points or streaks. Point loners shall change its position from one image to another. The set of all images has to be compared to find the point loners. Streak loners have a noticeable speed and position angle with respect to the telescope tracking.

Tracklets represent moving objects – potential asteroid or satellites. They are composed of the combination of loners from different images. A loner from the first image is combined with a loner from a second, third, etc. Such a combination is selected on specific criteria. First tracklet position can be any loner from the first image. Second tracklet position shall be selected from the following image within the specified distance from the first position. The other tracklet positions shall reflect the speed and position angle of the previous detections. The situation is illustrated on Figure 3.

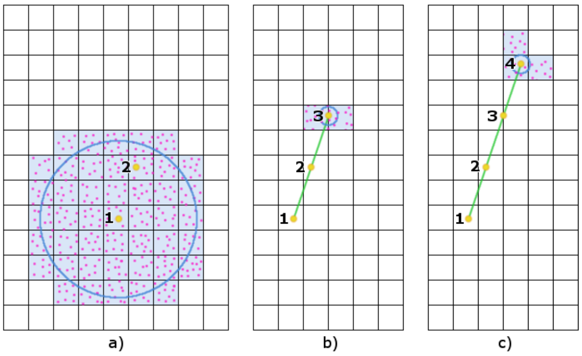


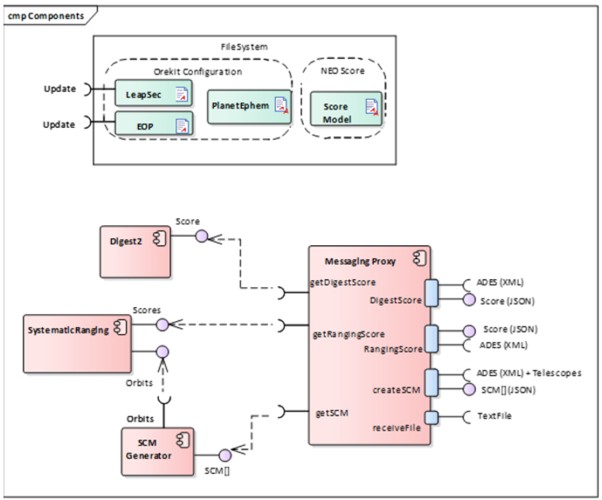
Figure 3. Building tracklets from loners. Figure a) shows a loner from the first image marked by "1". The blue circle is the maximum distance to a possible second loner from the second image. The loner marked by "2" shows a possible candidate for a tracklet. Figure b) contains point loners "1" and "2" from the previous step. The possible third position "3" shall be in line with the previous ones. Figure c) shows the tracklet with new position "4". It was found using the same methods as in the previous case.

The MPC Service compares all tracklets with the existing asteroid catalogue. For objects that are not matching with known asteroids, the PNG thumbnails and NEO scores are requested and stored. The thumbnails are provided by the Thumbnail generator Service of LightEngine. Thumbnails fixed on object and fixed on sky are provided. The NEO scores come from the Orbit Calculator. The MPC Service regularly checks the ASAP database in order to find confirmed tracklets that shall be sent to the MPC. Only confirmed object that are marked for automatic submission are sent to the MPC. Tracklets can be confirmed automatically by ASAP or by a specific user. ASAP can submit the asteroid astrometry by email or by ADES pipeline. The email submission uses the MPC 80 column format and is sent them through the provided SMTP server. In the case of the ADES, the XML file is produced and validated with the ADES XSD schema. The file is then sent through the ADES HTTP pipeline.

ASAP provides a web interface for the NEOSTEL processing results. It visualizes data stored in the ASAP database. For new discoveries, it shows to user its properties, thumbnails and orbit data. User can mark the object as real/noise or not sure. The decision is stored in the ASAP database. The asteroid web page makes possible to see several objects in one page. In such a way the user can quickly make the visual confirmation of multiple new discoveries. The number of objects visible per web page is configurable. The objects that can are inspected by one user are locked to the others.

4 ORBIT CALCULATOR

Orbit Calculator is one of NEOSTEL project's subsystems. It is responsible for calculating NEO score and systematic ranging scores from received observation data. It is also generating universal follow-up SCM messages from received orbital information.



Orbit Calculator is receiving from ASAP tracklets of uncorrelated objects via REST interface. Using Minor Planet Center's library Digest2, "neo scores" are calculated. They are representing a semi-probability of an object being a NEO. In case of user request, additional Systematic Ranging scores might be computed.

Systematic Ranging is an Orbit Calculator's component developed by ESA. It is using a novel orbit determination approach for short-track observations. It is performing a raster scan in the poorly-constrained space of topocentric range and range rate, while the plane of sky position and motion are directly tied to the recorded observations. Systematic Ranging library is a sophisticated method that outputs thousands of probable orbits, that match observations and are within previously specified bounds (not orbiting Earth, part of Solar System) assuming asteroid may be a NEO. This, so called, "cloud of orbits" is handled by Orbit Calculator's ScmGenerator component.

Scm Generator is responsible for computing ephemerides, Ra Dec topocentric coordinates, using telescope's position information. Orbit calculator is using space dynamics library Orekit for its computations. Only after operator's confirmation asteroid is scheduled for follow-up, because statistical approach of Systematic Ranging method is computationally complex.

Ephemerides have to cover large part of a night, so there are several pointing required, spaced by approximately 15 minutes - it can be changed by system's operator. For each pointing epoch, each orbit from a cloud is being propagated. This operation crates several groups of state vectors, separated by pointing time span (~15 minutes).

Each state vector is then processed to obtain pointing from Telescope's site. Sun-centric state vectors are transformed to geocentric EME2000, and then these are used to calculate object – telescope [ra,dec] angular direction. This operation creates several groups of [ra,dec] points in sky's coordinates.

For each group one point is chosen that minimizes the sum of distances to other points. This is a "center of a cloud", for which telescope should be directed to and have the highest probability of reobservation.

The list of points and epochs is then embedded in SCM message and sent back to ASAP for further propagation to target telescopes.

SCM is a CEN-CENELEC standard format for observing system commanding and scheduling. This standard aims to ease the planning and operation processes and to reduce the efforts from researchers that use several different observing systems and/or simulation software products.

5 ARCHIVING SYSTEM

The Archiving Service is a subsystem of the NEOSTEL DPC responsible for cataloguing and storing FITS images managed by the DPC during their whole lifecycle, receiving them from the telescope and storing the evolution of their data and metadata, and making them to the rest of the DPC subsystems. It also manages the export of stored data.

The Archiving System functionalities are exposed to other systems in the DPC through HTTP REST interfaces and RabbitMQ message queues. The components of the Archiving Service are Ingestor, Data Service, Catalogue, Object Storage, UUID Generator, Exporter, Thumbnail service, Preview image service and Packager

The Ingestor is in charge of receiving and orchestrating the archival of new data in the archive. It receives data, generates a unique id for it, extracts its metadata, and coordinates the archival of the data and cataloguing of the metadata in the Object Storage and the Catalogue database.

The Data Service is a façade to the internal catalogue and object storage for external data consumers. It allows performing queries over the catalogue and retrieving data from the object store. It also allows image data and metadata updates.

The Catalogue is a database that contains metadata of the data archived in the object storage. Everything that has been archived by the overall Archive has an entry in this database. FITS images stored in the archive can evolve through different versions. The database has special indexes and will have a specific design to optimise for the most important queries performed by the data service as received from LightEngine.

The Object Storage is in charge of storing the FITS images and allowing their retrieval. FITS images stored in the archive can evolve through different versions.

UUID Generator is responsible for generating unique identifiers for FITS images ingested by the system.

The Exporter is able to export data to external storage and keeping track of what has already been exported and when. This is used to replicate the contents of the archive in a remote location by sending physical media.

The thumbnail service manages insertion, retrieval and deletion of thumbnails. They have a given time-to-live period, after which they get deleted from the system.

The Preview image service manages insertion, retrieval and deletion of preview image objects. Each preview image is associated to an image version.

The Packager prepares packages of files as per user request. It receives a list of image versions and a folder structure to employ, and generates a package (uncompressed zip) for the user to download.

6 PROPOSED HARDWARE CONCEPT

The remote site where the DPC hardware will be located has following limitations:

1. Site will be unmanned. There will be no human operator present during normal operation.
2. Limited access during the winter, site may be inaccessible by up to one week (7 days).
3. Remote support for the observatory will be available only during the working hours (according to German public holidays calendar, 8:00-16:00).

Taking into account access and support limitations as well as strict performance and reliability requirements the server infrastructure has been designed to accommodate multiple latent faults and show a reasonable level of resilience. Redundant power supplies, cooling and communication has been envisioned together with multiple fall-back options for the server infrastructure failures allowing the processing to continue even with the partial hardware malfunctions. The layout of the server rack cabinet for the proposed hardware is presented in Figure 4.

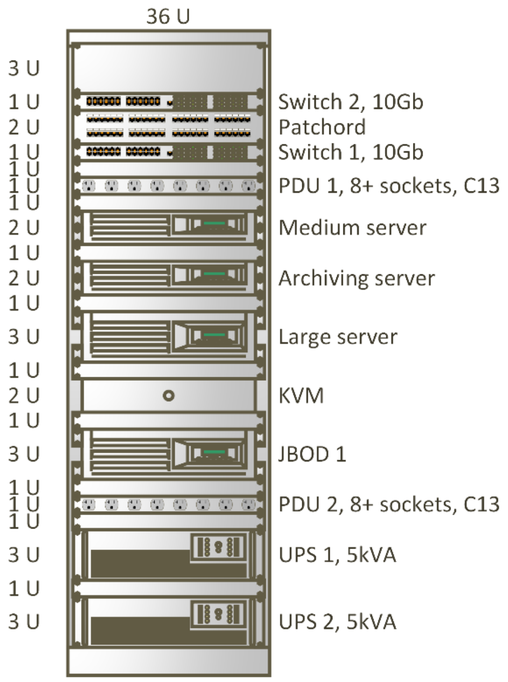


Figure 4: Example of the equipment layout in the rack cabinet hosting the NEO DPC servers.

Following mini-cluster of servers is proposed to solve the redundancy and performance requirements:

- Large server (2-3U Server, 2 x Intel Xeon 2GHz 35M Cache 14 Core, 512 GB RAM, RAID controller, 2x SSD 240GB)
- Medium server (2-3U Server, 2 x Intel Xeon 2.1GHz 20M Cache 8 Core, 256 GB RAM, RAID controller, 2x SSD 240GB)
- Archiving server (2-3U Server, 2 x Intel Xeon 2.1GHz 20M Cache 8 Core, 128 GB RAM, RAID controller, 2x SSD 240GB)
- HDD enclosure, JBOD (3-4U, redundant power supply, 19 x 10 TB)
- 2 switches, necessary cabling, etc. supporting redundant 10Gbps connections
- Transportation disk set (5x 10 TB)
- 2 Power Distribution Units
- 2 UPSes