

Archive Reload Function of the Online Data Management System for Earth Observation Data Exploitation Platforms

Markus Kunze, Stephan Kiemle, Nicolas Weiland, Matthias Hofmann

German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, D-82234 Weßling, Germany
Email: Markus.Kunze@dlr.de, Stephan.Kiemle@dlr.de, Nicolas.Weiland@dlr.de, Matthias.Hofmann@dlr.de

Background and Motivation

The amount of exploitable remote sensing data has increased significantly in recent years. As a result, the requirements of applications for fast processing of large coverages and time series in high resolution have also increased, which confronts Earth Observation (EO) data centers with new challenges.

ESA Data Hub Sentinels user products	2018	2019	2020	2021	2022	2023	2024
Yearly volume [TB]	13503	14111	14245	14433	15423	17085	18047
Average Data Rate [Mbit/s]	3592	3754	3789	3839	4103	4545	4850

Figure 1. Estimated data volume and rates for ingesting all Sentinel user products

Online EO Data Management

In order to cope with these challenges, approaches such as "Bringing users to the data" have emerged. For this purpose, exploitation platforms with large online memory and processing possibilities are being developed, which provide the functionality to process the data on the platform itself instead of transferring them to the user.

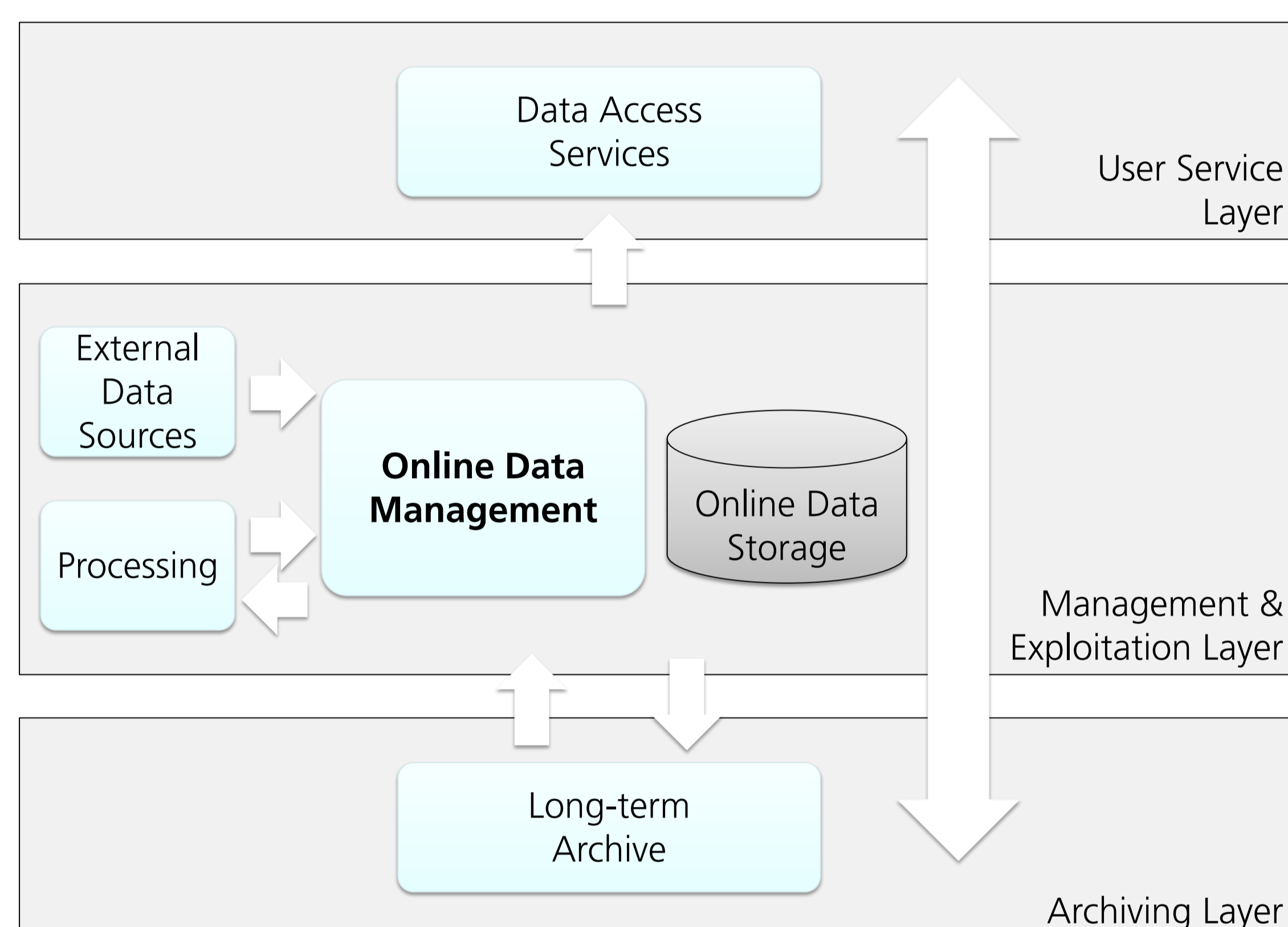


Figure 2. Simplified Example Archive Reload Scenario

The German Aerospace Center DLR is developing an Online Data Management System (ODMS) which provides functionality for managing EO data in exploitation platforms. With the following functions:

- **Ingestion** of large volumes of data by ensuring integrity (not only for the data, but also for the system)
- Establish an **inventory** of all managed data, supporting EO specific queries
- Provide **access** to data via specific data access services
- **Evict** data according to rules which may contain EO specific conditions
- Provide transparent **storage** for the data
- **Monitor** and **report** all tasks and provide workflow **controlling** capabilities
- **Reload** data via LTA interface

Archive Reload

From time to time, data will have to be evicted from the Online Storage, based on specific rules. However, in many exploitation use cases large data set series of historic data need to be held online to be processed into consistent time series products. Therefore the ODMS shall provide a functionality to efficiently reload evicted data from the backend archives. This introduces additional complexity in the overall concept and results in particular challenges for the development of the system.

- **Performance** optimization for bulk data reloading
- Handling **priorities**
- Ensuring **consistency** of data holdings the online platform and the backend archives
- Respecting **quotas**
- Compliance with **IT security** constraints

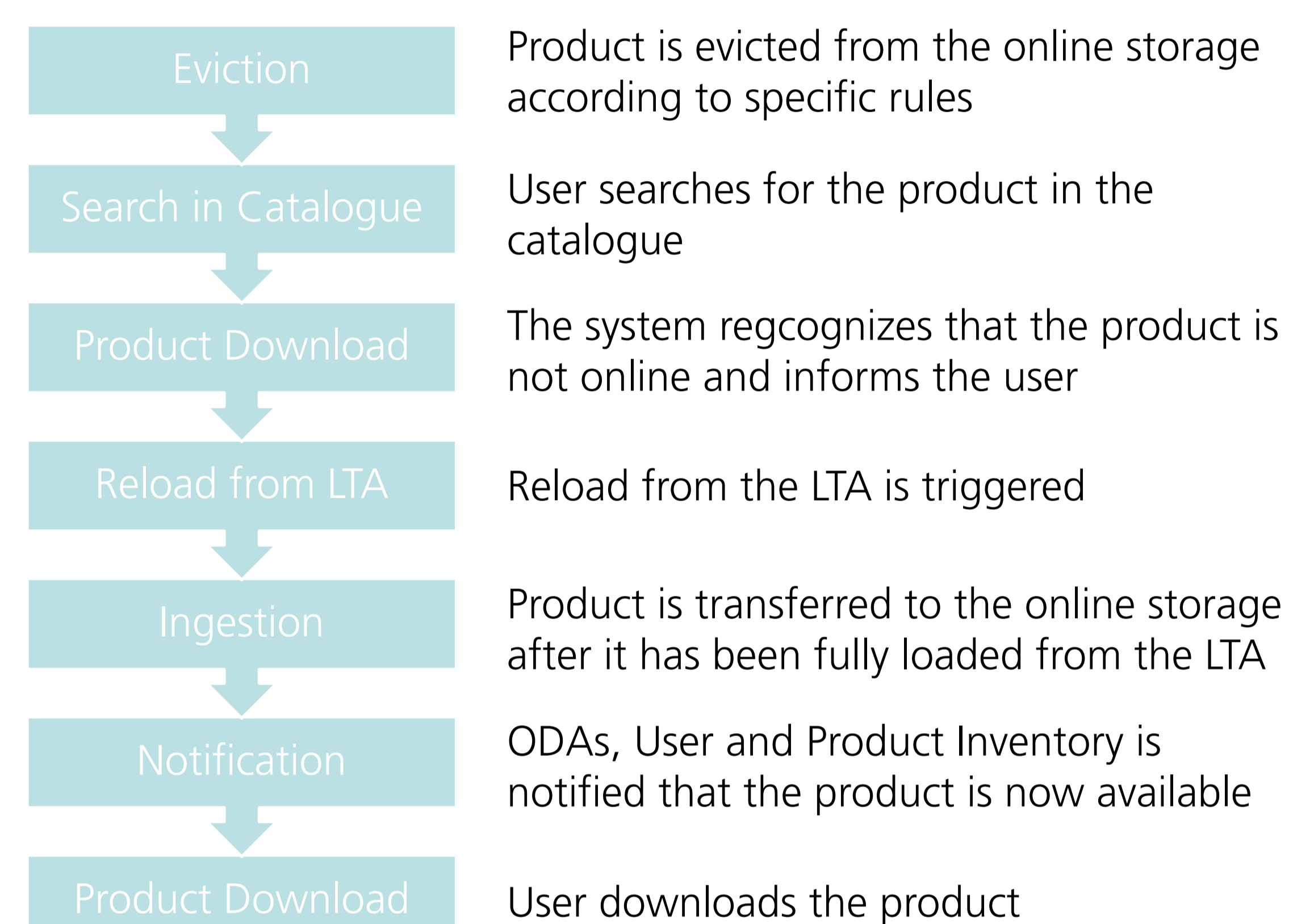


Figure 3. Simplified Example Archive Reload Scenario

Conclusion

In order to fulfill all requirements and to be able to map the processes presented, the architecture shown in Figure 4 was created.

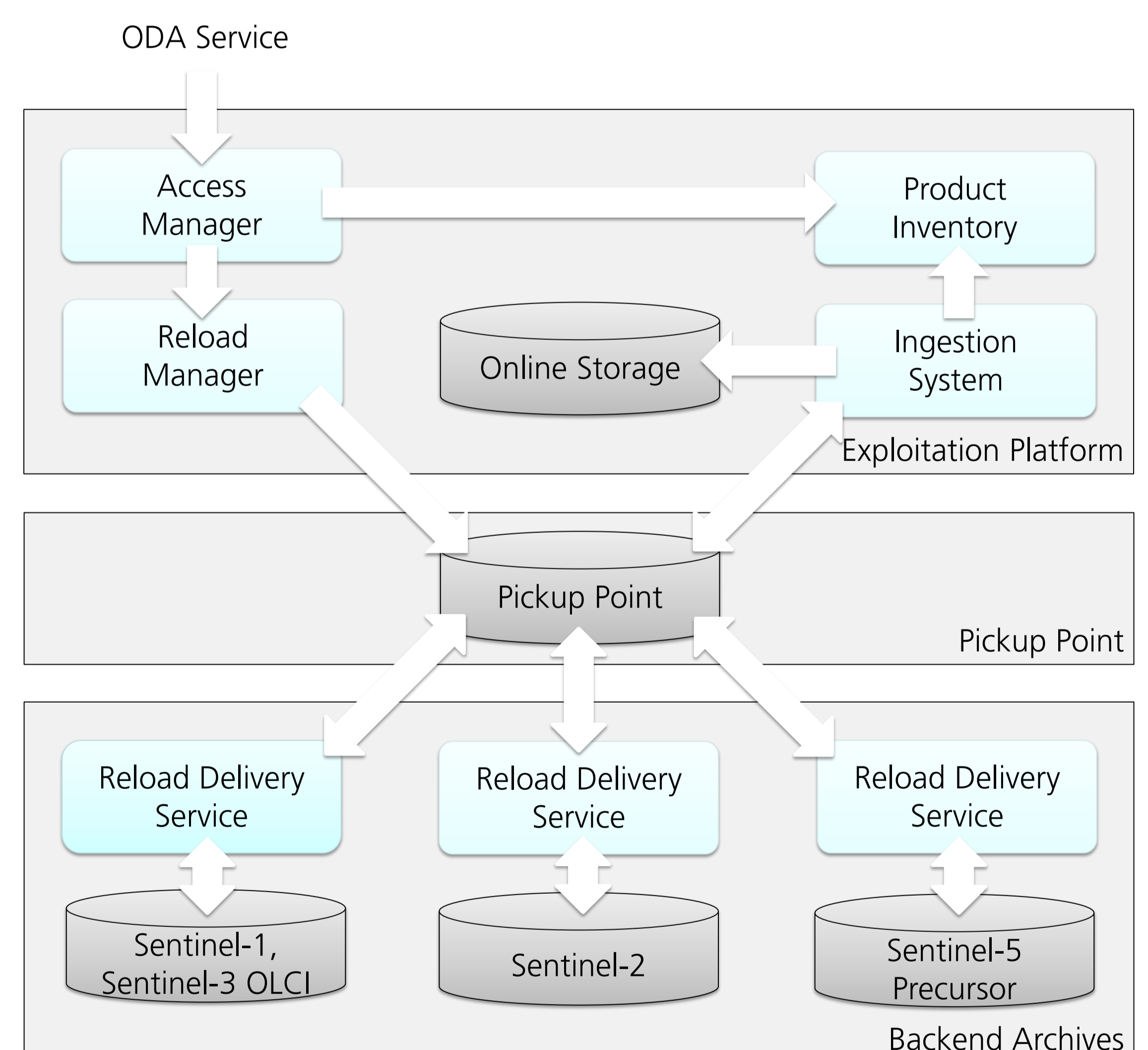


Figure 4. Example implementation of Archive Reload in projects like CODE-DE and partly DIAS

The Archive Reload Function of the Online Data Management System for Earth Observation Data Exploitation Platforms is used in projects such as CODE-DE and DIAS Mundi. Most of the system components are re-using existing open-source software, components developed for the DLR multi-mission German Remote Sensing Data Center. To further optimize the performance of the Archive Reload, we will focus on supporting bulk data reloads in the future.

Selected References

- [1] C. Reck, G. Campuzano, K. Dengler, T. Heinen, M. Winkler (2016) German Copernicus Data Access And Exploitation Infrastructure, Proceedings of BIDS'2016
- [2] N. Weiland, S. Kiemle, M. Kunze, T. Keßler (2017) Online Earth Observation Data Management Proceedings of BIDS'2017