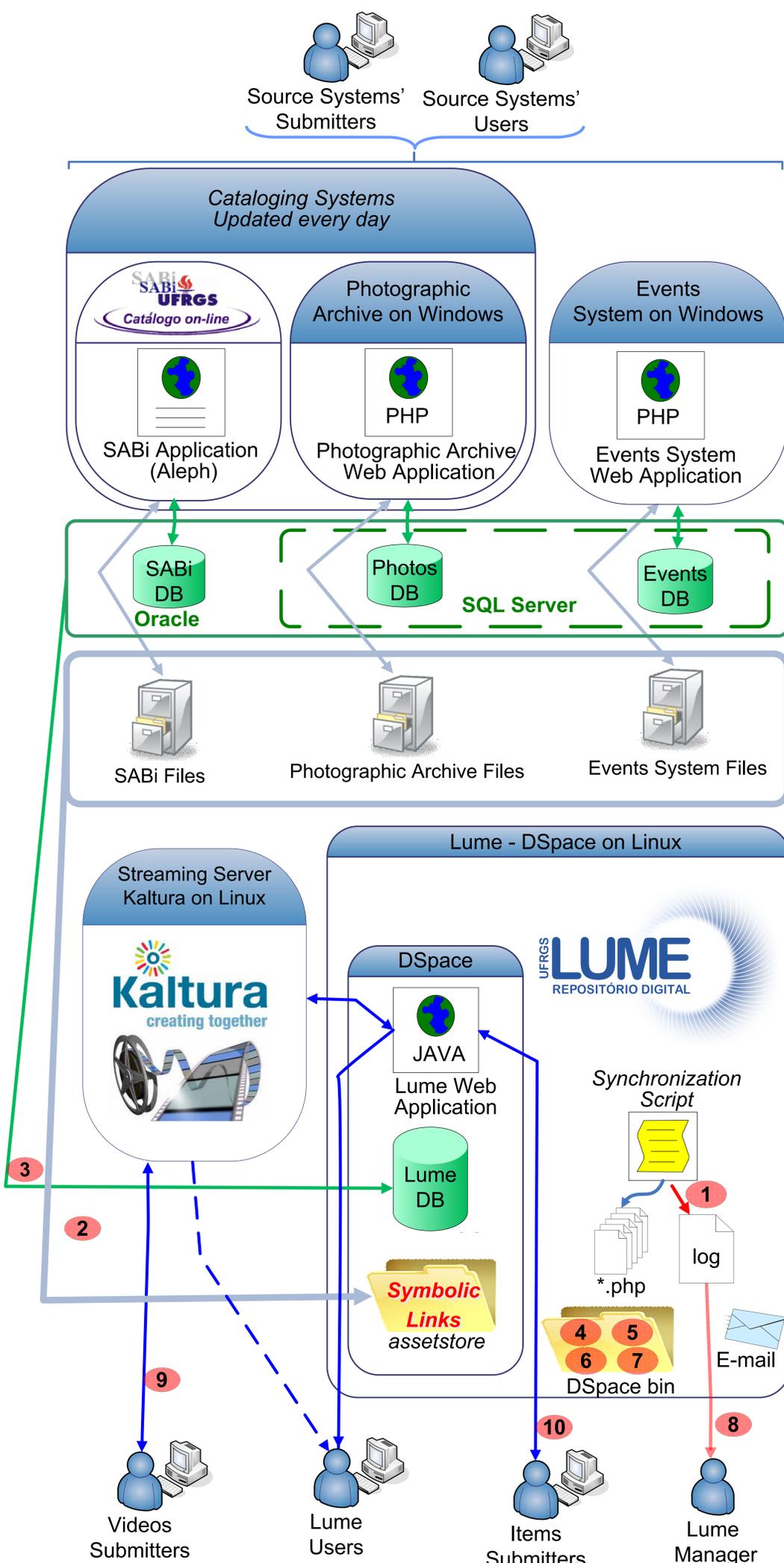


Manuela Klanovicz Ferreira, Afonso Comba de Araujo Neto, Raquel Almeida, Zaida Horowitz, Janise Silva Borges da Costa, Caterina Groposo Pavão
Centro de Processamento de Dados - Universidade Federal do Rio Grande do Sul (UFRGS)
E-mail: comissao@cpd.ufrgs.br

Abstract

Lume is the Digital Repository of Federal University of Rio Grande do Sul (UFRGS - Universidade Federal do Rio Grande do Sul) and aims at fostering the dissemination of digital content produced in the context of this University. The data sources for the repository can be classified into three groups: institutional cataloging systems, institutional events management system, and special archives which items are submitted manually in an individual basis. A detailed diagram showing Lume's customized usage of DSpace is presented, highlighting the developed extensions that allow the synchronization of Lume with the several data sources with distinct updating patterns, and how the provision of embedded videos was enabled.



Lume - Digital Repository of UFRGS

Lume is the Digital Repository of UFRGS and was implemented in 2008 using DSpace - Institutional Repository System. Its purpose is to preserve, disseminate, and increase the use and visibility of documents produced by the University that are preserved due their scope and/or historical characteristics. UFRGS is more than 100 years old and covers several fields of knowledge, with 90 undergraduate courses, 71 master programs, and 68 doctoral programs distributed on four centers in Rio Grande do Sul, Brazil.

Synchronizing DSpace with other systems

- 1) Synchronization process progress is saved into a log file, allowing the identification and spotting of eventual errors;
- 2) Symbolic links referring to the original files in the source systems are created in an asset store folder, making it possible for DSpace to access them without having to create copies of the files;
- 3) Metadata are translated to Dublin Core format and imported from the source system to Lume. To insert an item and corresponding metadata and symbolic links into Lume, the following DSpace database tables are accessed:

- | | |
|-----------------------------|------------------------------|
| I. public.item | V. public.bundle |
| II. public.handle | VI. public.item2bundle |
| III. public.collection2item | VII. public.bundle2bitstream |
| IV. public.bitstream | VIII. public.metadatavalue |
| | IX. public.resourcepolicy |

Steps 4 to 7 result from running four distinct commands provided by DSpace (available at <install-dir>/dspace/bin):

- 4) A **cleanup** command is run to clean deleted bitstreams;
- 5) The **filter-media** command creates thumbnails from images, and extracts plain text from PDF's so that it can be indexed by Lucene;
- 6) Search indexes are updated with data from new items by the **index-update** command;
- 7) The **sub-daily** command sends e-mails notifying collection subscribers of recent additions to the repository;
- 8) Finally, Lume managers are informed via e-mail if any errors were detected during the synchronization process.

Submission on DSpace and Kaltura video streaming

- 9) Authenticated members (LDAP) of Lume are offered the possibility to personally submit new items directly into DSpace.
- 10) The availability of video submissions and streaming is being planned using Kaltura. Users will be able to submit videos, having them deposited in Kaltura and their metadata associated with DSpace items. The Kaltura player will then be used to show the embedded video.

Conclusions

Integrating distinct systems is not an easy task. Even when managing common data, each system is designed for a specific administrative purpose, so they have distinct architectures and data models. In this sense, the adoption of patterns and standard communication interfaces in the development of information systems is essential to simplify their interoperability.

