Big Data tools and their applications

Martin Macák

LAB OF SOFTWARE ARCHITECTURES AND
INFORMATION SYSTEMS

FACULTY OF INFORMATICS
MASARYK UNIVERSITY, BRNO

lasaris

Big Data Introduction

What are Big Data?

- **V**olume
- **V**ariety
- **V**elocity
- **V**eracity
- **V**alue

BIG DATA

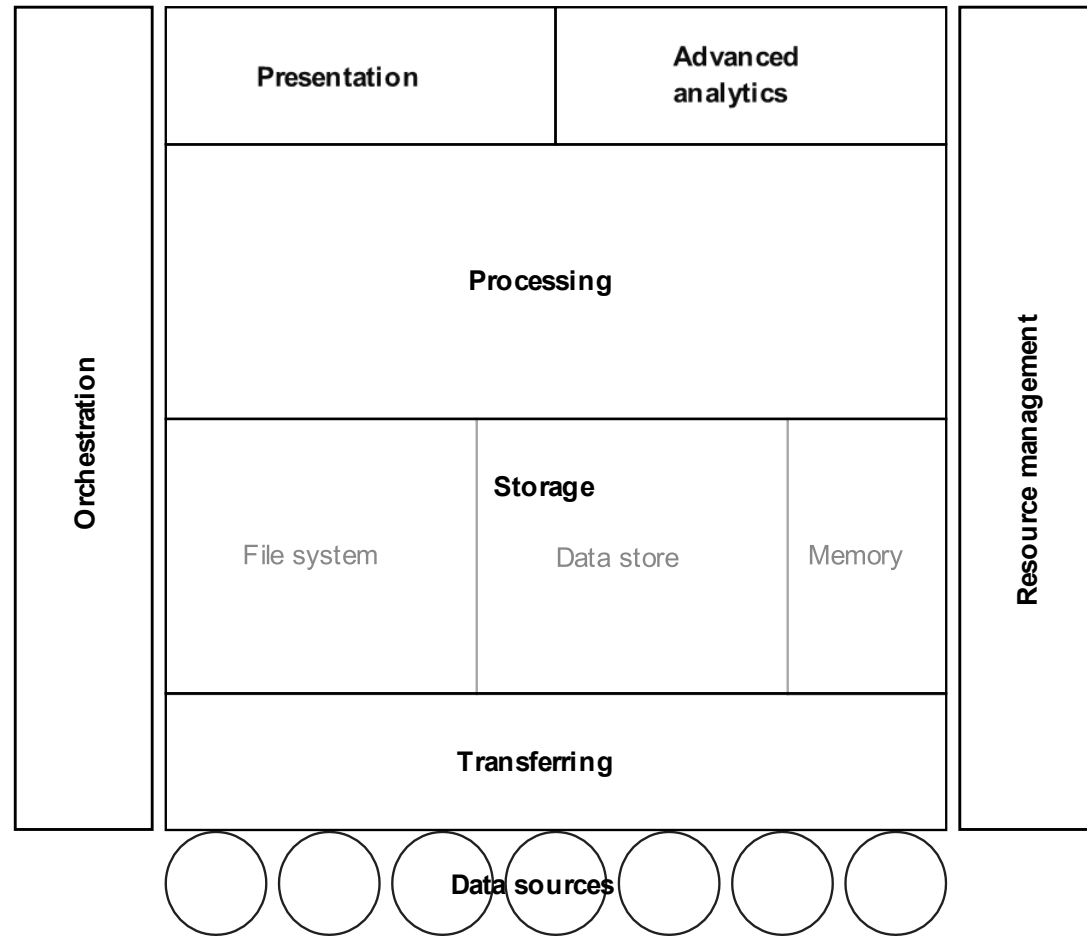lasaris

Where are they used?

State of the Art

Issues

- the overwhelming number of tools (more than 6000 open source tools [1])

- each tool targets a specific purpose(s)

- no single best tool for a given purpose

- immaturity of tools (limitations, high change rate, lack of proper support)

- new tools are being created, many old tools die

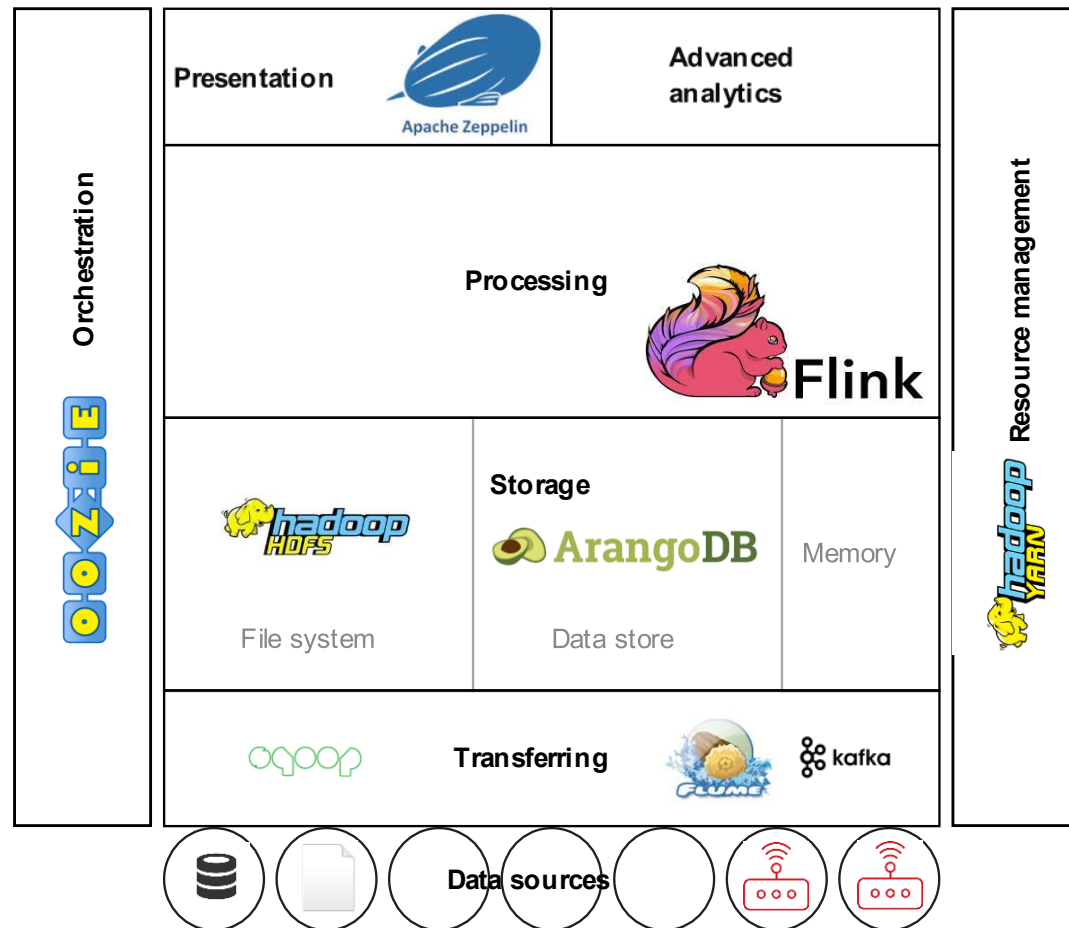- in each domain may be necessary to use a different tool for the "same" problem

Our current results
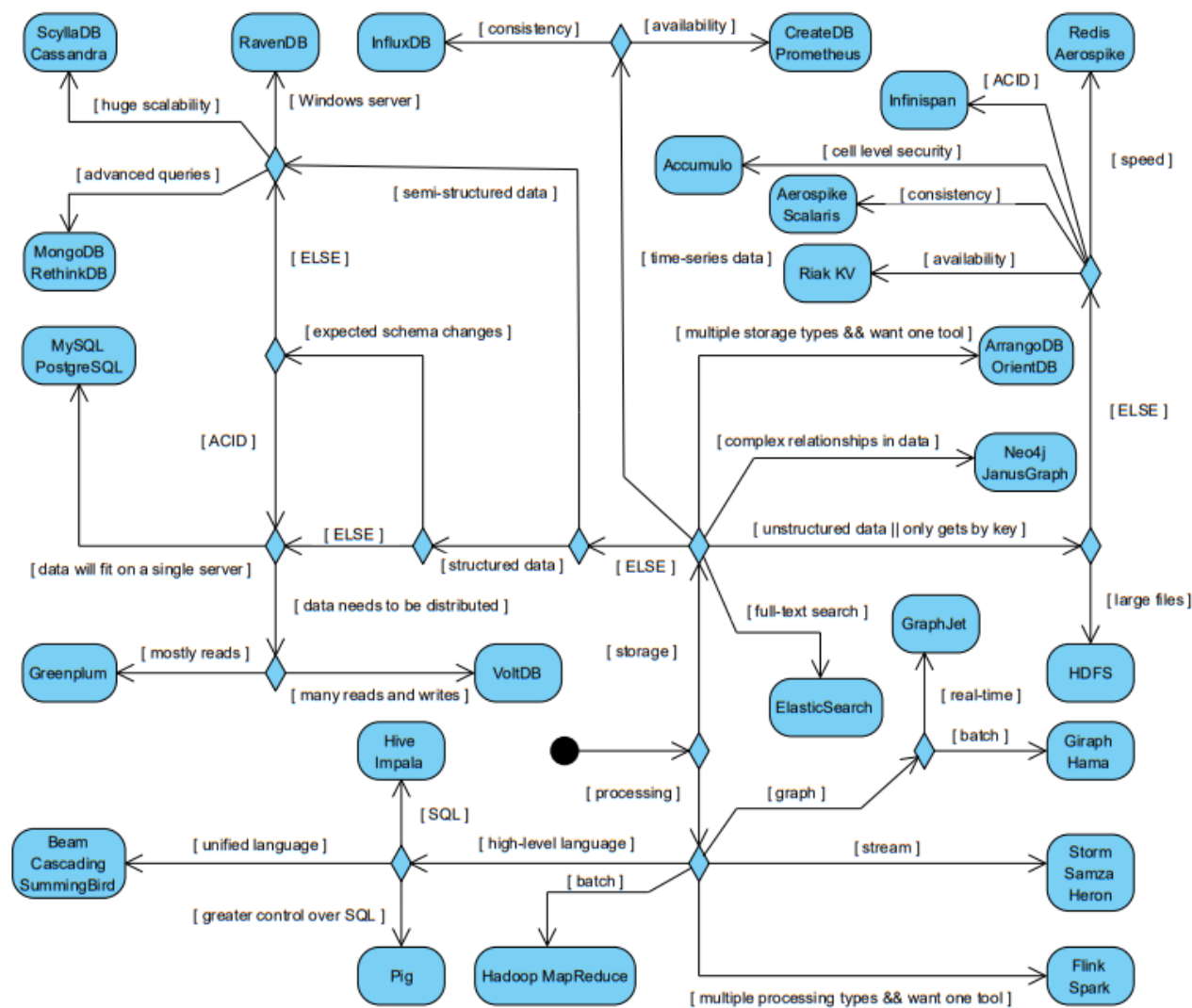
Proposed Architecture

Proposed Architecture – real example

Proposed Tool Selection Diagram

[2]

Our future plans

Recommender platform for Big Data analysis techniques

1. Examination of existing Big Data techniques

2. Examination their applicability in different domains

3. Identification of criteria driving the decision of the technique selection

4. Evaluation of results using benchmarks and real case studies

5. Implementation and validation of the recommender platform

lasaris

Resources

- 1. GÖKALP, Mert; KAYABAY, Kerem; ZAKI, Mohamed; KOÇYIĞIT, Altan; ERHAN EREN, P; NEELY, Andy. *Big-Data Analytics Architecture for Businesses: a comprehensive review on new opensource big-data tools*. 2017. Available from: https://cambridgeservicealliance.eng.cam.ac.uk

- 2. MACÁK, Martin. *Tools for big data analysis* [online]. Brno, 2018. Available from: https://is.muni.cz/th/po5g7/. Master's thesis. Masaryk University, Faculty of Informatics. Thesis supervisor Barbora Bühnová.

lasaris

Summary

- Big Data Introduction
  - 5 Vs

- State of the Art of Big Data Tools
  - many, immature, no single best tool

- Our current work
  - architecture, tool selection diagram

- Our future plans
  - recommender platform for Big Data techniques

Martin Macák, FI MU Brno
macak@mail.muni.cz