

Problems to store, transfer and process the Big Data

COURSE: COMPUTING CLUSTERS, GRIDS, AND CLOUDS

KRISTIAN KOUROS

LECTURER: ANDREY SHEVEL

ITMO UNIVERSITY

05/06/2017

Outline

- Introduction
- Big Data Characteristics
- Importance of Big Data
- Storage Problems
- Transfer Problems
- Processing Problems
- Conclusion

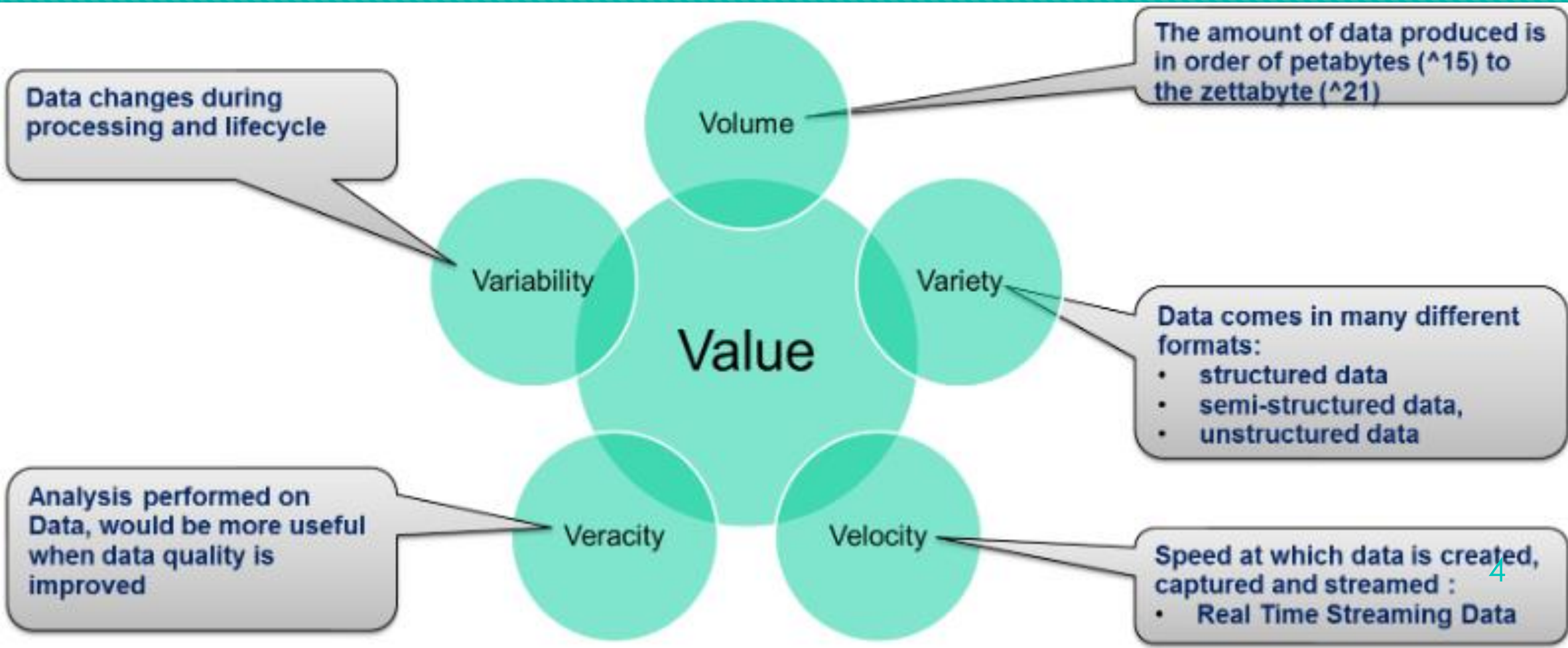
Introduction

- What is Big Data?

“Data sets that are so *large* and *complex* that traditional data processing application software is inadequate to deal with them.”

- Generated by IoT devices such as mobile devices, remote sensing, software logs, cameras, microphones, RFID readers, Social Media and WSN.
- Complex type of data: Combination of structured, semi-structured, unstructured , homogeneous or heterogeneous data.

Big Data characteristics



Importance of Big Data

With the proper advanced analytics and visualization techniques we can uncover hidden patterns and find unknown correlations

- Cost reduction
 - Direct (Ex. Traditional database vs Hadoop)
 - Indirect (Ex. UPS)
- Data-driven decision making (Social Media field research)
- Improvement in products and services

Big Data Problems



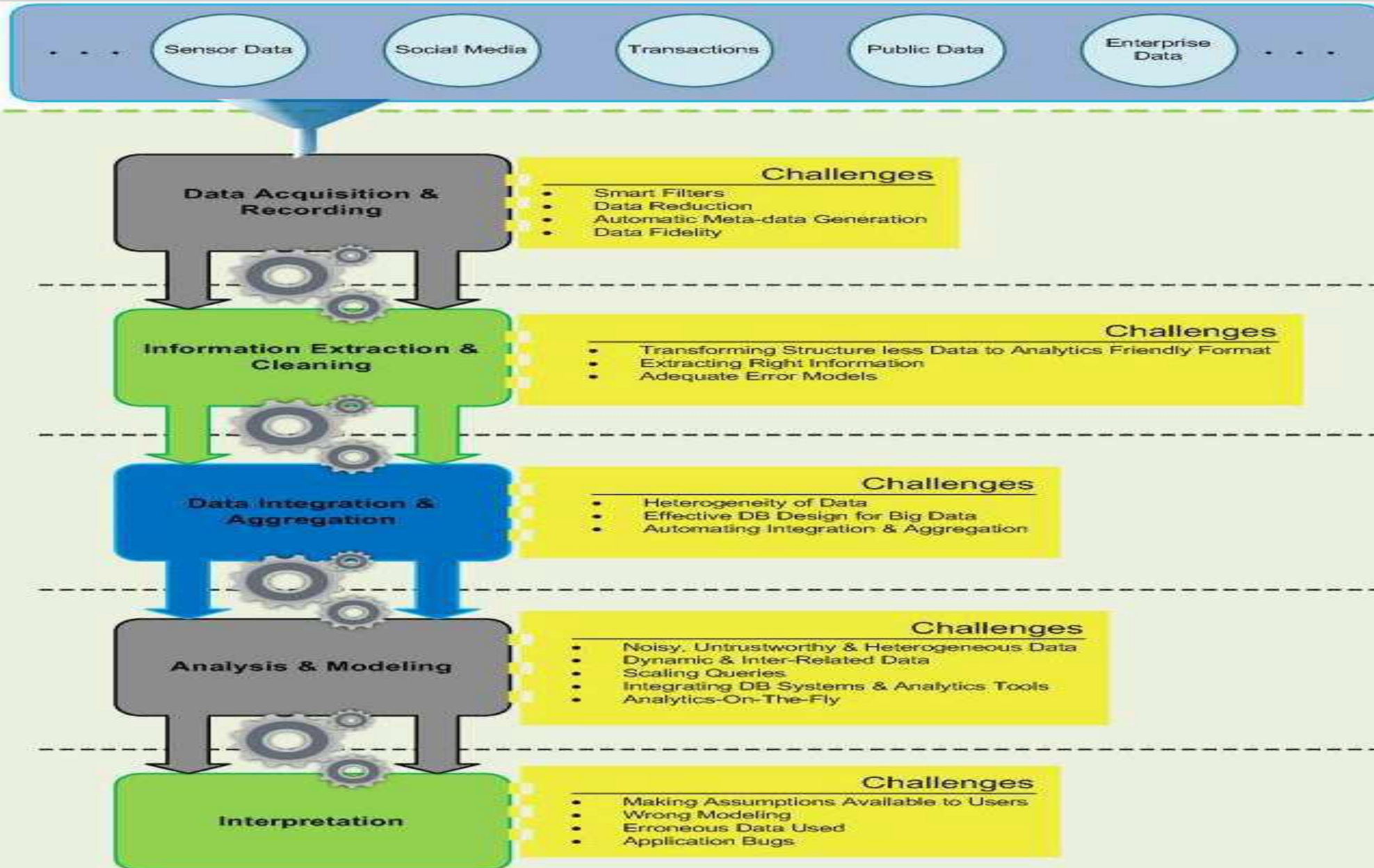
Storage

- Current storing technologies cannot handle the huge, heterogeneous amount of data.
- New organizing techniques are required.
- Hyper scaling computing environments

Transfer

- Overwhelm of traditional communication networks.
- It would take longer to transfer than to process !!
- Possible solutions:
 1. Process the data “in place” and transmit only the resulting information (Transition to Fog architecture).
 2. Perform triage on the data and transmit only that data which is critical to downstream analysis.

Processing



Conclusion

- Big Data , the inevitable future
- Current situation is manageable
- The future demands new algorithms, techniques and technologies

End of presentation

References

- [1] “Big Data Challenges” Nasser and Tariq, J Computer Engineering Information Technology 2015, 4:3
- [2] “Big Data in Cloud Computing: features and issues “, Pedro Caldeira Neves, Bradley Schmerl , Jorge Bernardino and Javier Cámara
- [3] “Big Data: Issues and Challenges Moving Forward,” S. Kaisler, F. Armour, J. a Espinosa, and W. Money, 46th Hawaii Int. Conf. Syst. Sci., pp. 995–1004, 2013.
- [4] "Big data", Wikipedia, 2017. [Online]. Available: https://en.wikipedia.org/wiki/Big_data