Longest Word - Hadoop

● ● ●

Eias Kaminski, Florian Korch

htw. Hochschule für Technik
und Wirtschaft Berlin

University of Applied Sciences

Gliederung

Unsere Aufgabe

- 6 Texte in verschiedenen Sprachen vorgegeben
- aus jeder Sprache das längste Wort finden

Hadooop

"The Apache Hadoop software library is a framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. Rather than rely on hardware to deliver high-availability, the library itself is designed to detect and handle failures at the application layer, so delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures." - https://hadoop.apache.org/

Hadoop Setup (Mac OS)

- Installation von Hadoop über homebrew
  - homebrew ~ chocolaty ~ apt-get
- Konfiguration über Anpassung von 4 Dateien
  - hadoop-env.sh
  - core-site.xml
  - mapred-site.xml
  - hdfs-site.xml
- Erstellen eines neuen SSH-Keys
- Festlegen von Start- und Stopskripten für Hadoop
  - hstart
  - hstop

Implementation ->

Quellen

- "Hadoop – Apache Hadoop 3.3.4." https://hadoop.apache.org/docs/stable/ (accessed Jan. 23, 2023).

- Diwakar, "Installing Hadoop on Mac," *Beer&Diapers.ai*, Nov. 19, 2019.

  https://medium.com/beeranddiapers/installing-hadoop-on-mac-a9a3649dbc4d (accessed Jan. 23, 2023).

- "hdfs_user_guide.pdf." https://hadoop.apache.org/docs/r1.2.1/hdfs_user_guide.pdf (accessed Jan. 23, 2023).