



Contribution ID: 5

Type: **not specified**

Handling Big Data - an Overview of Mass Storage Technologies

Monday, August 26, 2013 4:00 PM (40 minutes)

According to IDC forecasts, Big Data-related IT spending is to rise 40% each year between 2012 and 2020, and the total amount of information stored world-wide will about double every two years. It means that the, so called, digital universe will explode from 2.8 zettabytes in 2012 to 40ZB, or 40 trillion GB, in 2020. This is more than 5200 gigabytes for every man, woman, and child alive in 2020. It is therefore crucial to build storage systems robust and scalable enough to not only safely hold fast-growing amount of data but also capable of making all the information they hold available for access in an efficient manner.

This presentation will introduce basic principles, characteristics and fundamental technologies used in today's storage systems. It will compare distributed POSIX-like filesystem approach to the web-scale systems implementing only a reduced set of operations, like GET, PUT & DELETE. There will be a discussion of advantages and disadvantages of both technologies and their applicability to various requirements. This will lead to a characterization of efficient access protocols, in particular the proposed HTTP 2.0 standard based on Google's SPDY protocol.

Author: JANYST, Lukasz (CERN)

Presenter: JANYST, Lukasz (CERN)

Session Classification: Plenary talks

Track Classification: Big Data and Large Storage Systems