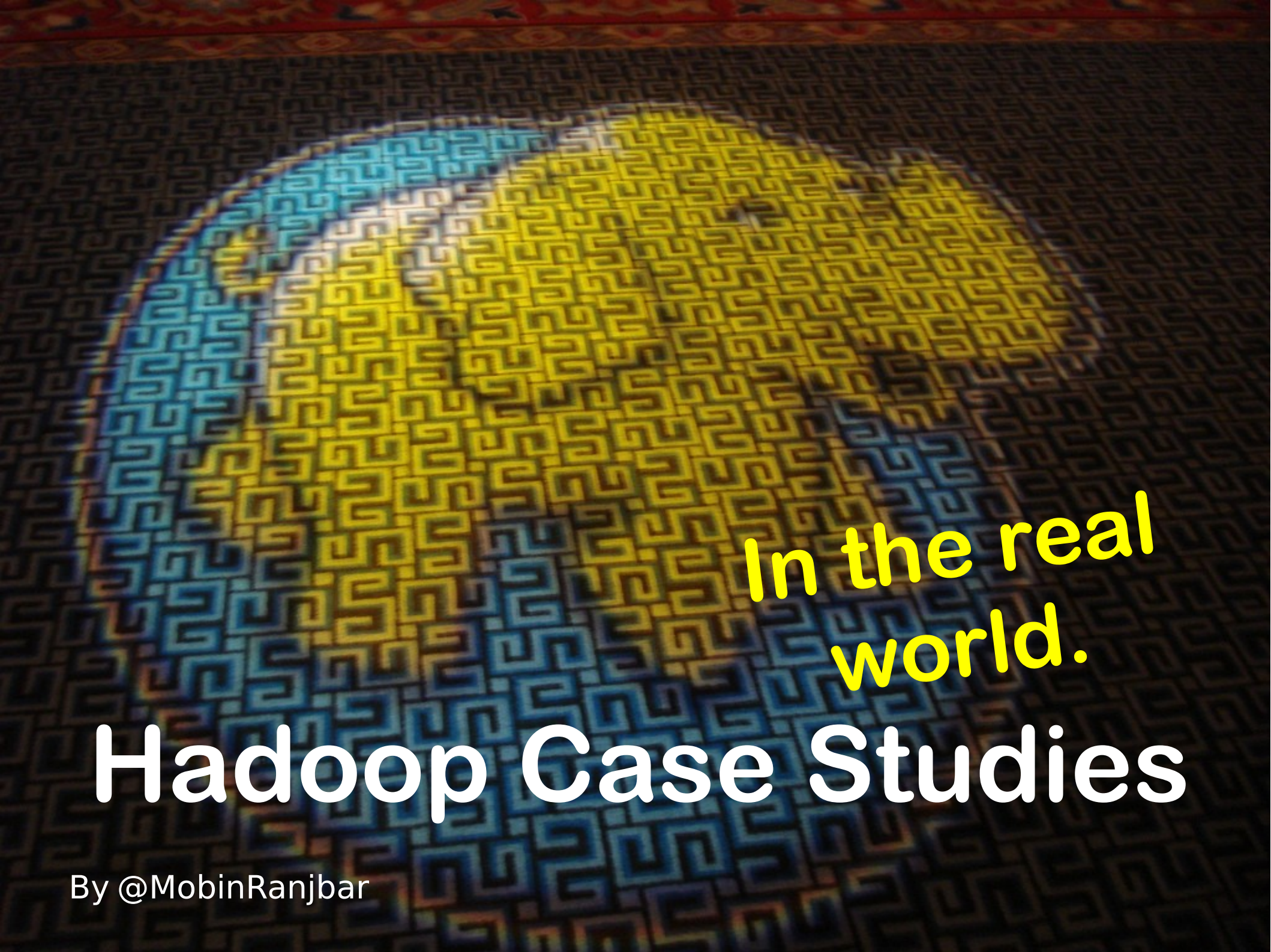
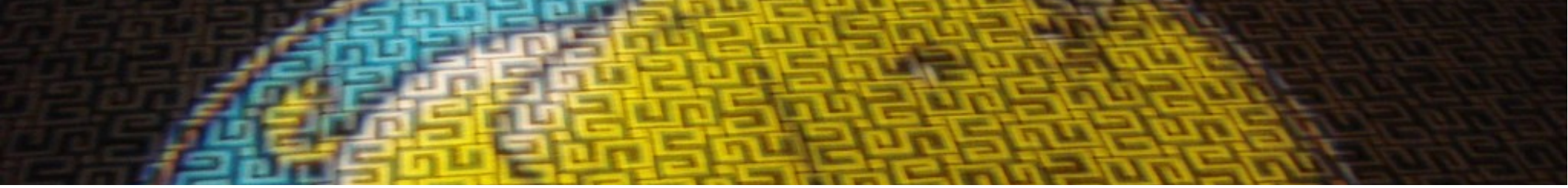

In the real world.
Hadoop Case Studies
By @MobinRanjbar

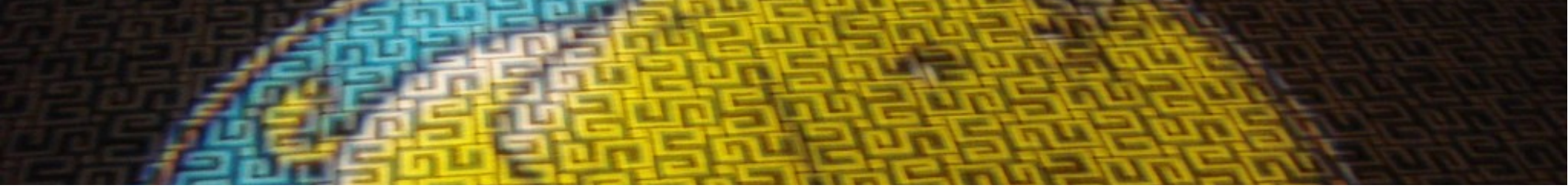


What is Apache Hadoop?

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.

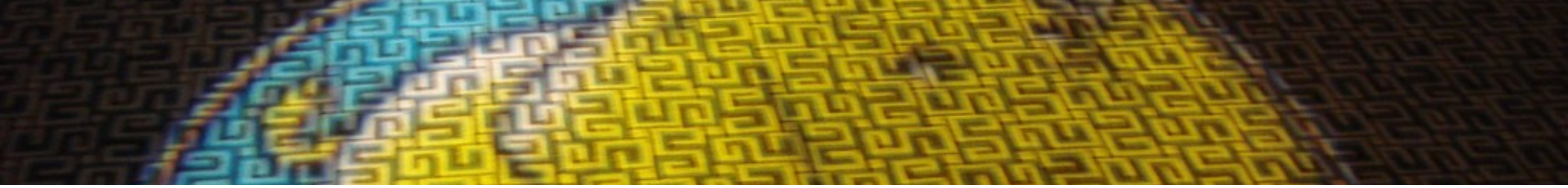
Rather than rely on hardware to deliver high-availability, the library itself is designed to detect and handle failures at the application layer, so delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures.

Source: http://hadoop.apache.com



Why Apache Hadoop?

- It is distributed.
- It can handle structured/semi-structured/unstructured data.
- It does not need expensive hardwares.
- It is open source.
- It has a variety of tools for machine learning and etc.
- And etc.



Case Studies

Economy and Financial Systems

- Fraud Detection in bank accounts.
- Analysing customers data to know how are they spending money.
- Who is more eligible to get a loan based on their history.
- Modern currencies like Bitcoin can take advantage of it in mining process.

Communications

- Content–centric Networking / Named Data Networking
- Analysing internet traffic in miliseconds.

News and Media

- News classification and recommendation systems.
- What do people like to watch and read without asking them what they like(based on their behaviour).

Medical and Pharmacy

- Accurate medical prescription based on patients' lifestyle, nutrition and genetic disorders.
- Analyzing and decreasing prescription drug side effects.
- Discovering cancers.
- Brain genome mapping and sequencing.

Sales and Marketing

- Analyzing customers behaviour. Where do they look the most in a market store.
- Finding the related products based on their previous purchases.

Oil and Gas Exploration and Discovery

- Weather, soil, and equipment data can be analyzed to predict the success of drilling operations and make more intelligent decisions about drilling sites.
- Oilfield managers need to analyze well data, seismic data, industry news and potentially social media to evaluate potential oilfields.
- Processing seismic data that can be used to discover seismic trace signatures that were previously overlooked.

Weather Forecasting

- More accurate weather forecasting.
- Analysing and forecasting the risk of flood.
- Discovering climate changes in the future.

Geology

- Forecasting the risk of earthquick based on aftershocks data.
- Analysing volcanoes' behaviour.

Astronomy and Space

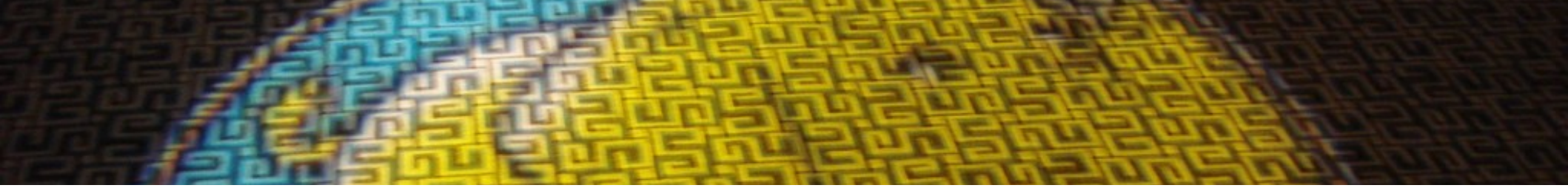
- Astronomical Hi–Res Image Processing.
- Exploring new galaxies and planets.
- Analysing Meteorite impact models.

Green and Renewable Energy

- GreenHadoop: Leveraging Green Energy in Data-Processing Frameworks.
- Vestas utilized big data to maximize power generation and reduce energy costs.
- Predicting the amount of solar energy that will be available in the near future.

Crime and Urban Management

- Accurate election result prediction.
- Preventing crime before it happens.



And anywhere you can find high volume of unstructured data with high rate of growth, Apache Hadoop will change the game.

Presented by Mobin Ranjbar

Twitter: @MobinRanjbar

Linkedin: linkedin.com/in/mobinranjbar