

EVALUATION OF DIFFERENT CLOUD COMPUTING RESOURCES FOR LARGE SCALE SATELLITE DATA PROCESSING

POTENTIAL AND LIMITATION

DR. TUONG-THUY VU

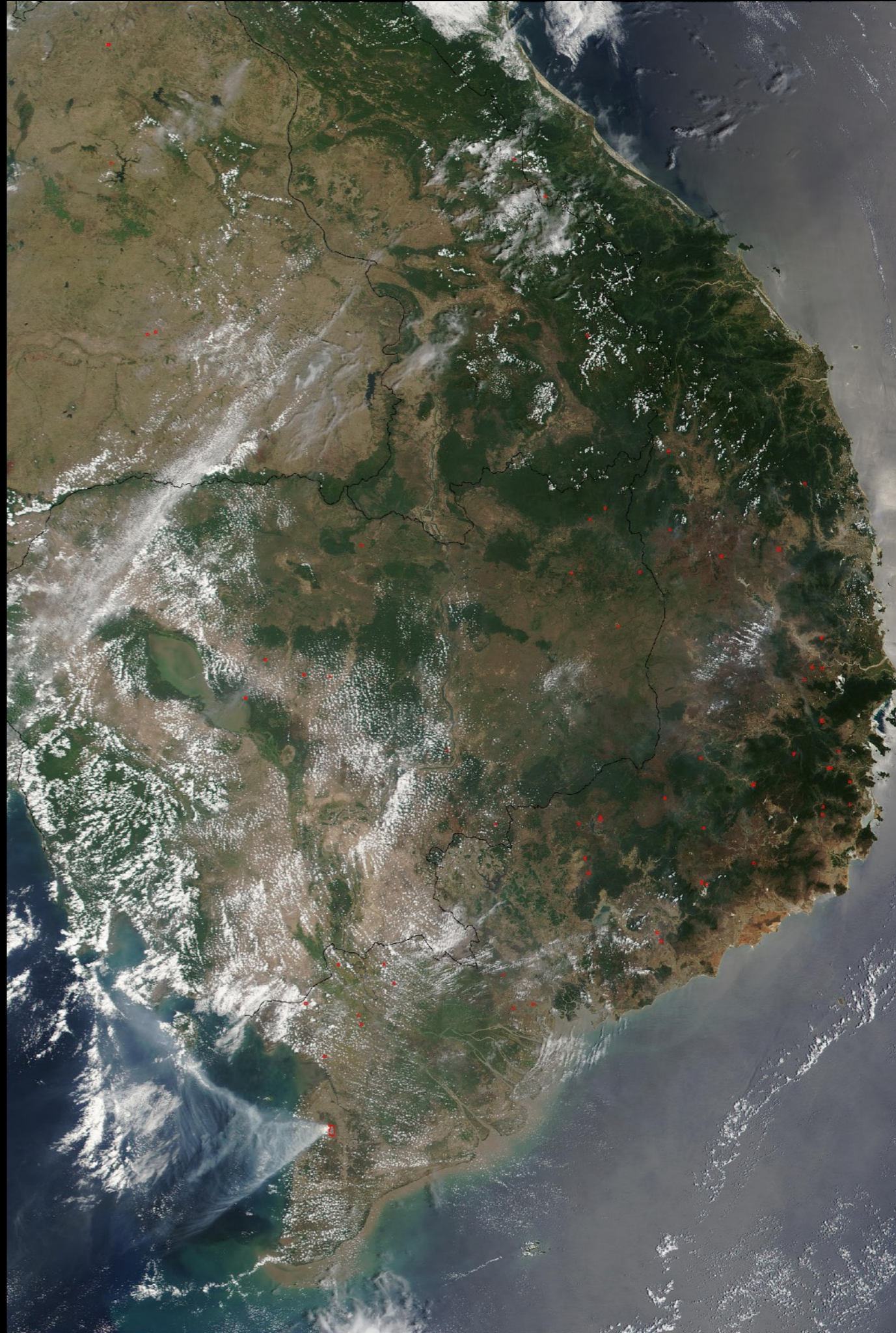


BIG DATA

EARTH
OBSERVATION

SATELLITE
IMAGE DATA

CLOUD
COMPUTING

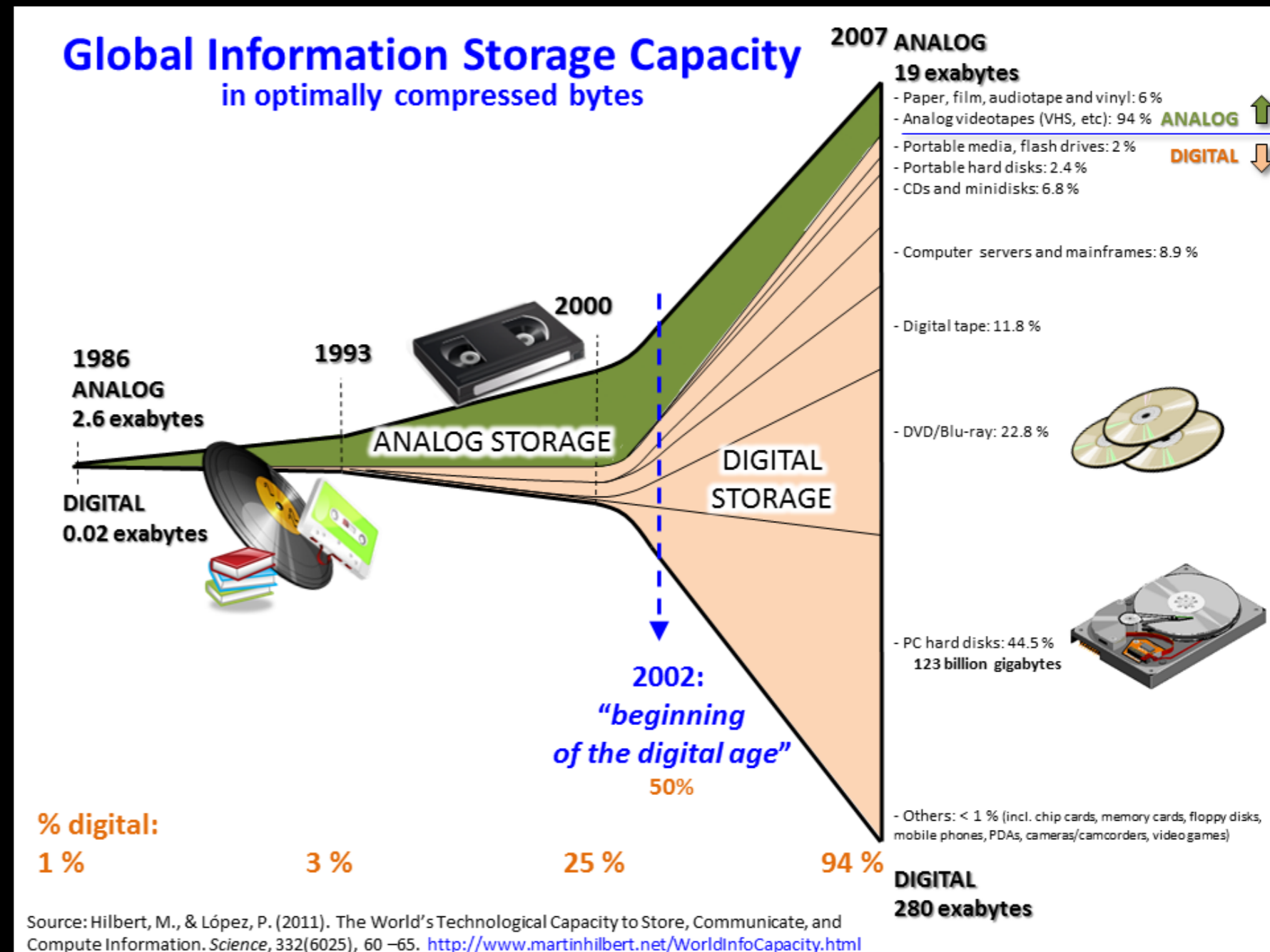


CHALLENGE

analysis, capture, data curation, search, sharing, storage, transfer, visualisation, querying, updating and information privacy.

BIG DATA

Volume
Variety
Velocity



EARTH OBSERVATION SATELLITE DATA



LANDSAT 5

160 MB (compressed)
366 MB (uncompressed)

LANDSAT 7

900 MB (compressed)
1.6 GB (uncompressed)

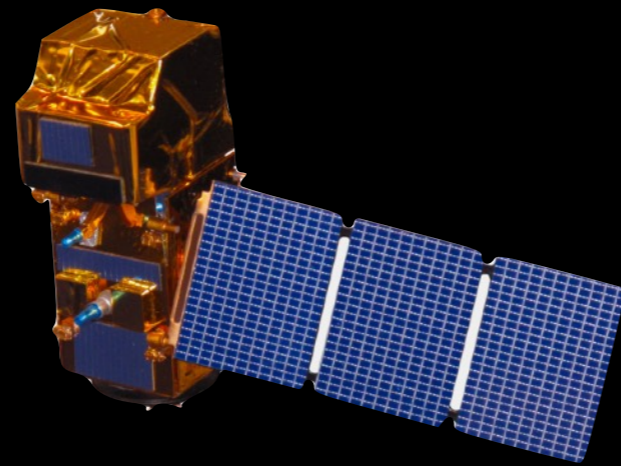
LANDSAT 8

1 GB (compressed)
2 GB (uncompressed)

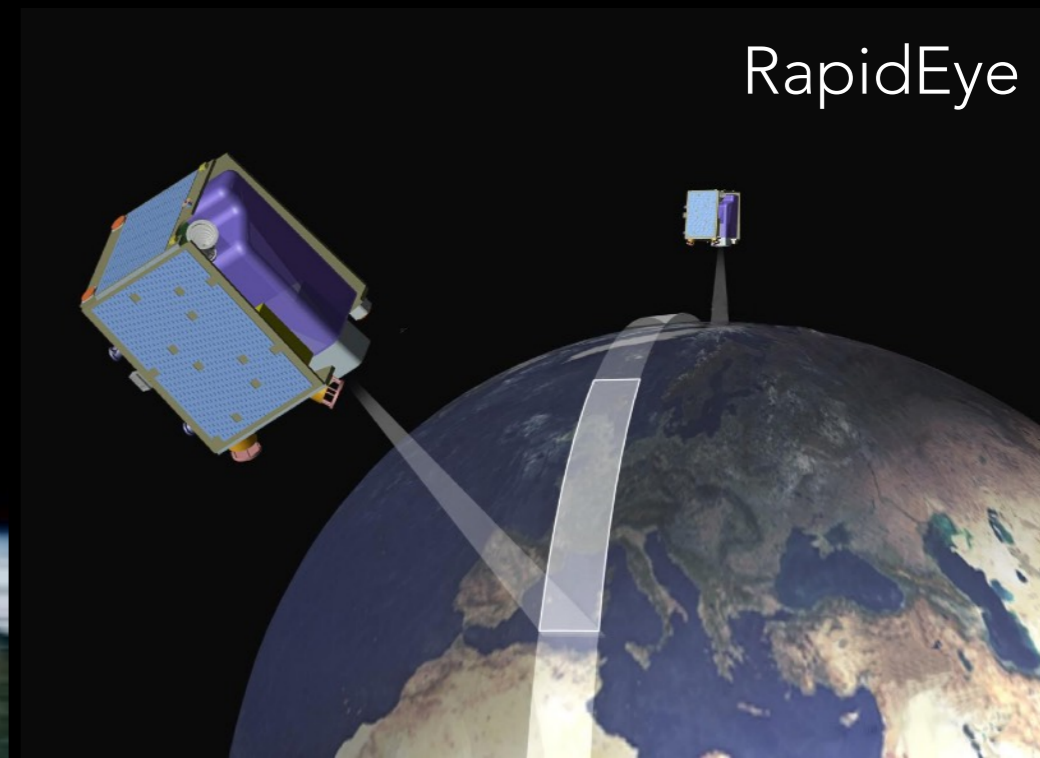
1 scene

EARTH OBSERVATION SATELLITE DATA

Multi-sensors



Landsat

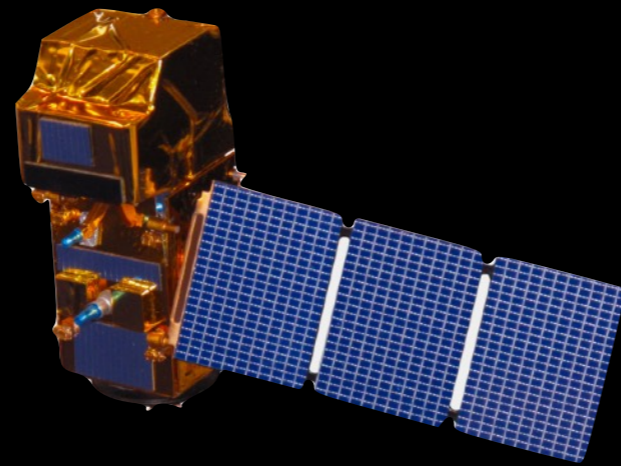


EARTH OBSERVATION SATELLITE DATA

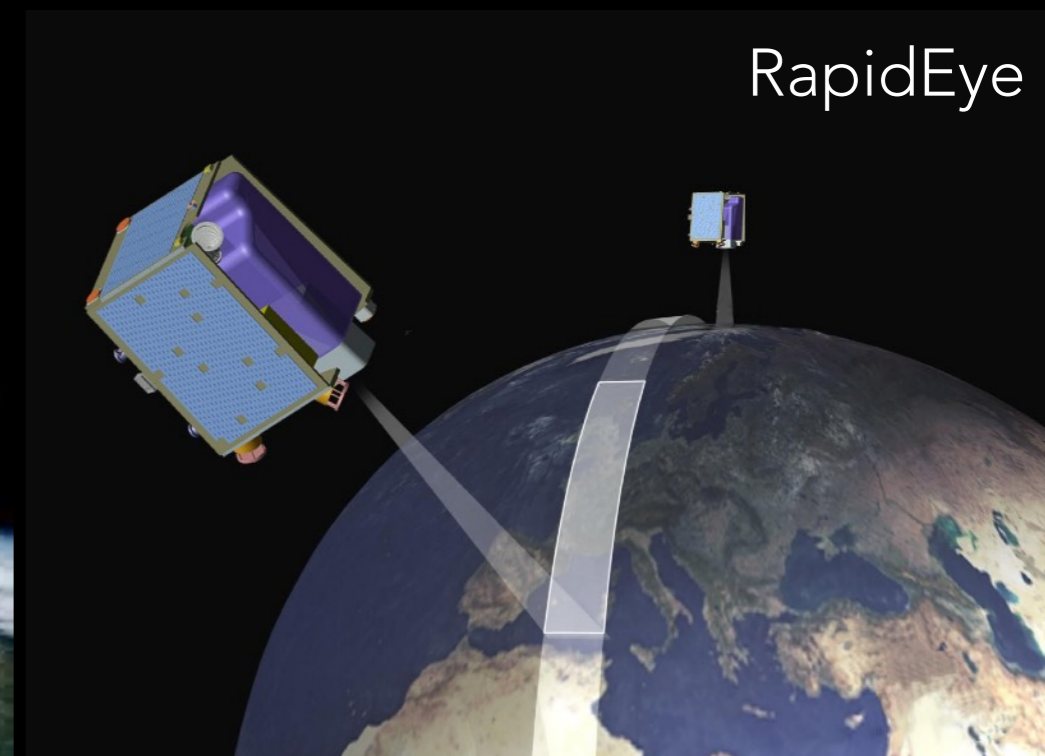
Multi-sensors



Ground-truth



Landsat



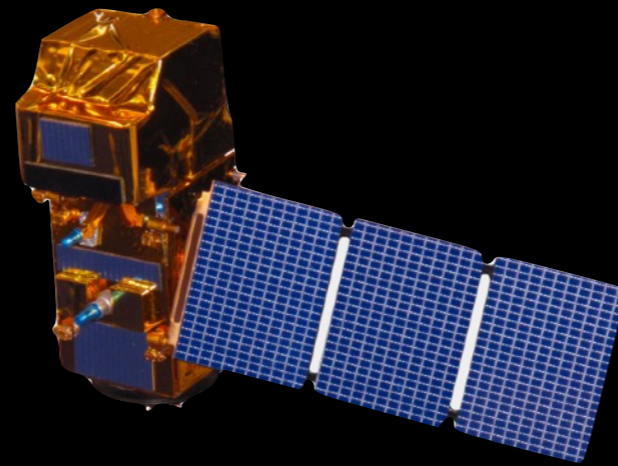
EARTH OBSERVATION SATELLITE DATA

Multi-sensors

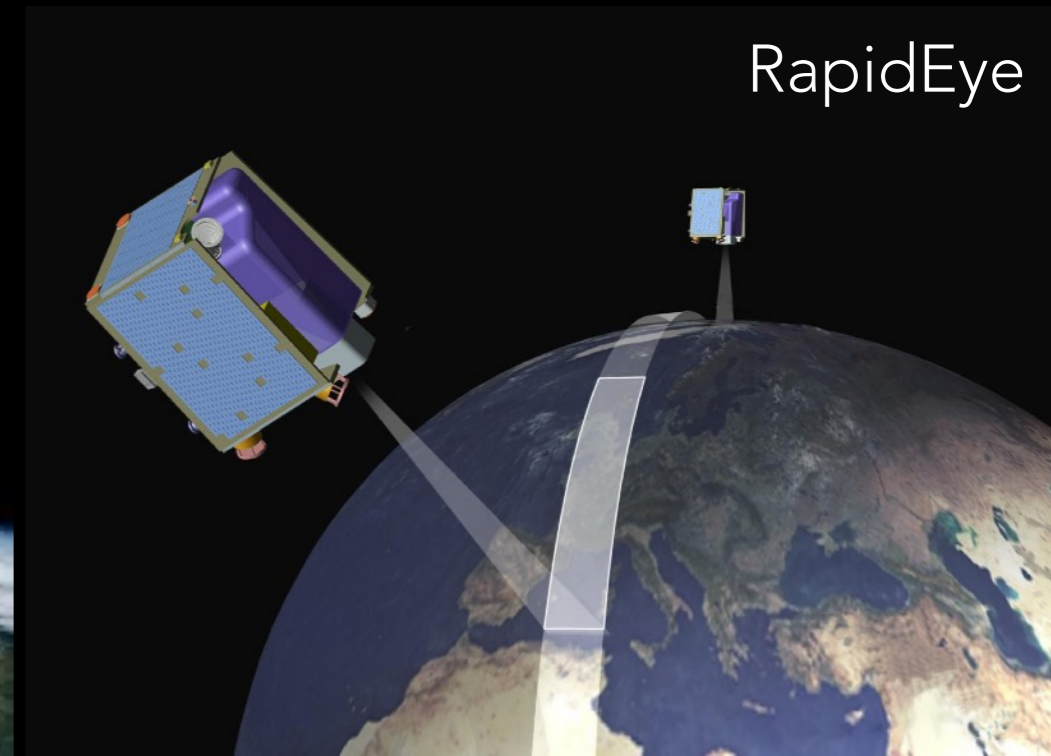


Ground-truth

Modelling



Landsat



EARTH OBSERVATION SATELLITE DATA

Multi-sensors

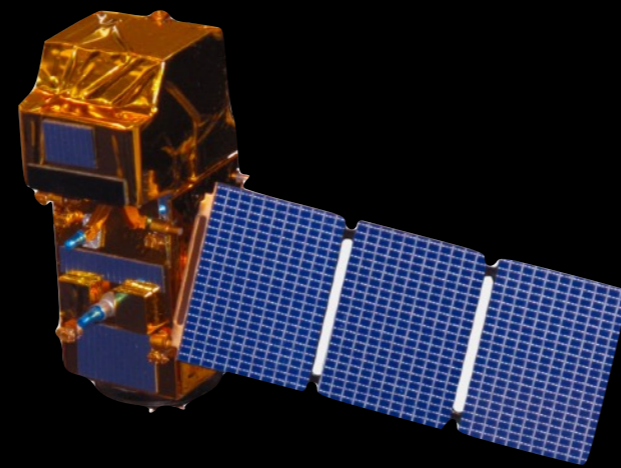


Ground-truth

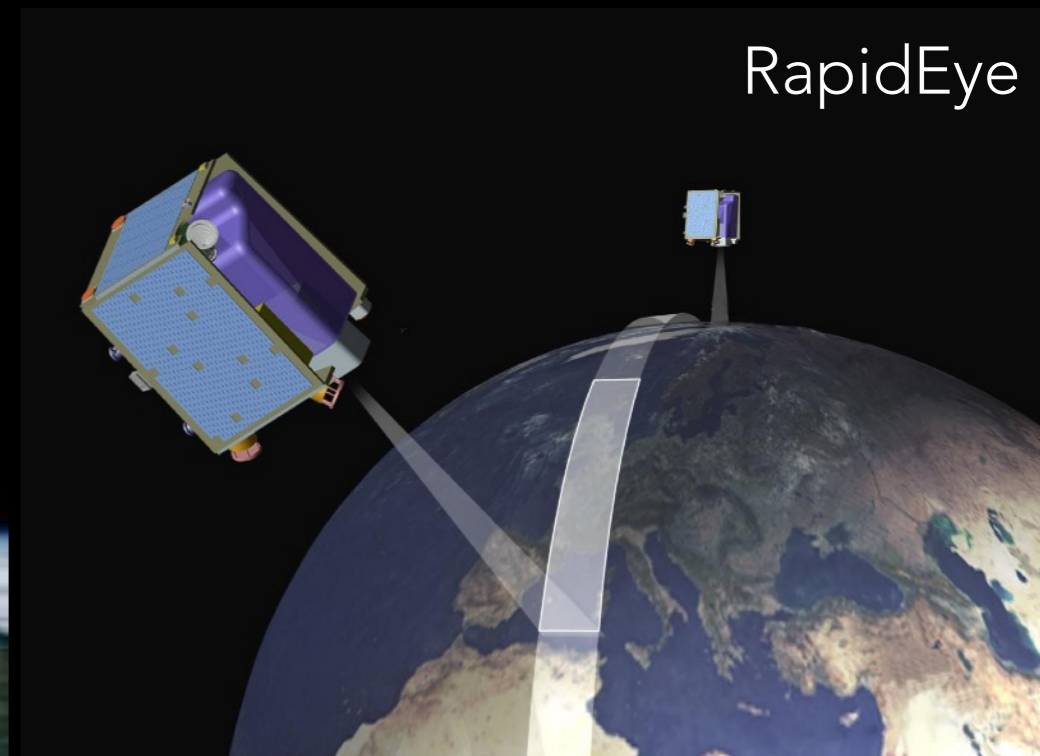
Modelling

Spatial analysis

.....



Landsat



- Production pipeline for Landsat and NASA Global Imagery and Browse Service – 37 steps, 7 petabytes of data, and 670,000 processor hours.
- NEX Downscaled Climate Projections (NEX-DCP30) and NEX Global Daily Downscaled Projections (NEX-GDDP) pipeline – 40 terabytes (TB) of data and 320,000 processor-hours.
- Estimating biomass by counting trees across the United States by using machine learning and computer vision – 40 TB of data/100,000 compute hours.

Building Petabyte Data Production Systems with the NASA Earth Exchange

(<http://www.nas.nasa.gov/SC15/demos/demo22.html>)

CLOUD COMPUTING

... a service over the Internet

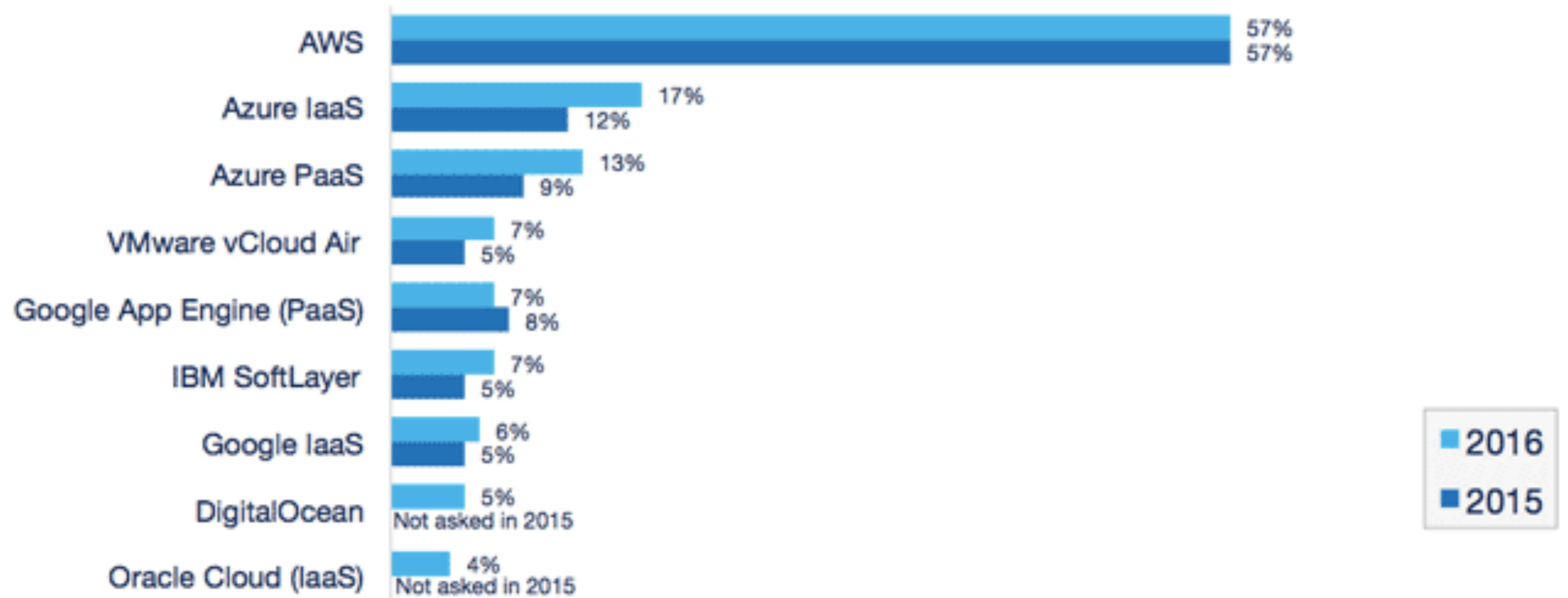
SOFTWARE (SAAS)

PLATFORM (PAAS)

INFRASTRUCTURE (IAAS)

Public Cloud Adoption 2016 vs. 2015

% of Respondents Running Applications



Source: RightScale 2016 State of the Cloud Report



Source: Gartner (May 2014)

■ 2016
■ 2015

CURRENT TREND

LANDSAT ON
AWS

DIGITALGLOBE

REMOTE SENSING DATA STORED ON THE CLOUD-COMPUTING PLATFORM, PROVIDED ACCESS VIA WEB PORTAL, DOWNLOADABLE

ESRI CLOUD

ENVI CLOUD

SOFTWARE COMPANIES PROVIDED PROCESSING SERVICES ON THE CLOUD

ESA CLOUD
TOOLBOX

NASA EARTH
EXCHANGE

AGENCIES PROVIDE CLOUD-BASED PROCESSING TOOLS

CURRENT TREND

LANDSAT ON
AWS

DIGITALGLOBE

REMOTE SENSING DATA STORED ON THE CLOUD-COMPUTING PLATFORM, PROVIDED ACCESS VIA WEB PORTAL, DOWNLOADABLE

ESRI CLOUD

ENVI CLOUD

SOFTWARE COMPANIES PROVIDED PROCESSING SERVICES ON THE CLOUD

ESA CLOUD
TOOLBOX

NASA EARTH
EXCHANGE

AGENCIES PROVIDE CLOUD-BASED PROCESSING TOOLS

INCREASING RESEARCH ON SPATIAL DATA PROCESSING SERVICES ON THE CLOUD

The most open platform so far ...

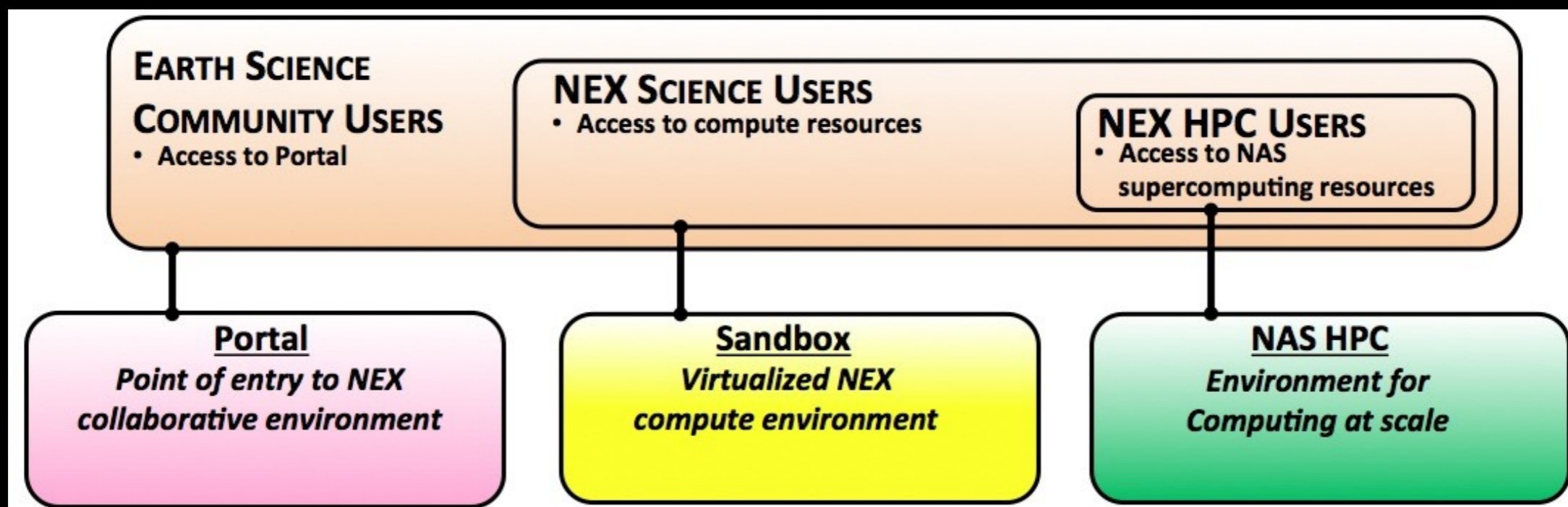


Google Earth Engine



NASA EARTH EXCHANGE

- A collaborative compute platform
- Data
- Tools & Libraries
- Models
- Disk storage
 - Pleiades 1.3 PB total
 - Sandboxes: Inxsrv105: 163 TB total and externally accessible sandbox 305TB





DATA SET

- Landcover
- Climate Datasets
- MODIS Land
- MODIS Atmosphere
- Landsat 5
- Landsat 7
- Landsat 8
- Landsat GLS
- Web-Enabled Landsat
- ASTER
- TRMM
- AVHRR
- Ecosystem Modeling Datasets
- DEM
- Soils
- Aerial



Ecosystem Models

- Terrestrial Observation and Prediction System (TOPS) North America
- Biome-BGC USA
- Lund-Potsdam-Jena Dynamic Global Model (LPJ) Global

Climate Models

- The Weather Research and Forecasting (WRF) Model Regional



TOOLS

Analytical Tools

MATLAB, IDL/ENVI,

R for statistical computing and graphics

Numpy and SciPy for Python

Matplotlib for plotting and graphics

GNU Octave

Database Management Systems

MySQL, SQLite, MongoDB

Data Manipulation Tools and Libraries

MODIS Reprojection Tool (MRT)

MODIS Reprojection Tool for Swath (MRTSwath)

HDF-EOS To GeoTIFF Conversion Tool (HEG)

The Land Data Operational Products Evaluation (LDOPE) Tools

PROJ.4 Cartographic Projections Library

GDAL - Geospatial Data Abstraction Library

NetCDF (Network Common Data Form) tools and libraries

....

Google Earth Engine

- a computing platform
- for petabyte-scale scientific analysis and visualisation of geospatial datasets

CODE EDITOR

web-based IDE
for writing and
running scripts

EXPLORER

a lightweight
web app for
exploring our
data catalog and
running simple
analyses

CLIENT LIBRARIES

Python and
JavaScript
wrappers around
our web API

Landsat

Sentinel

MODIS

1m resolution aerial image data US
National Agriculture Imagery Program
(NAIP)

Other sub-meter resolution sample images

SRTM 30-m

Landcover MODIS, MERIS

Surface temperature

Climate & Weather

World population

Google Earth Engine

DATA SET

Most of tools you need are here

Image

Math operation, convolutions, edge detection, spectral transformation, object-based methods, compositing & mosaicking, etc.

Reducer

Image reductions, statistics, raster-to-vector and vector-to-raster conversion, linear regression

Supervised classification

Landsat algorithms

Sentinel-1 algorithms

Google Earth Engine

Chart

Array

.....

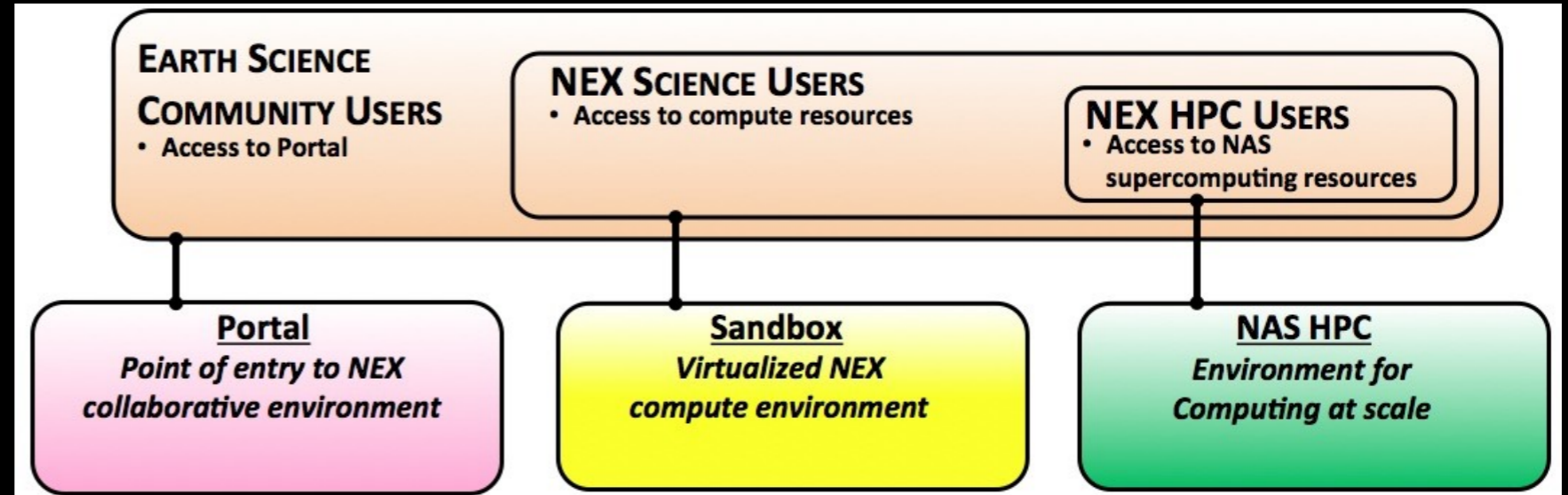
TOOLS

ACCESSIBILITY & LICENSE

NASA/NDC or OpenID account, need to signup and be approved



NEX sandbox
AWS, GEE, Azure, etc.

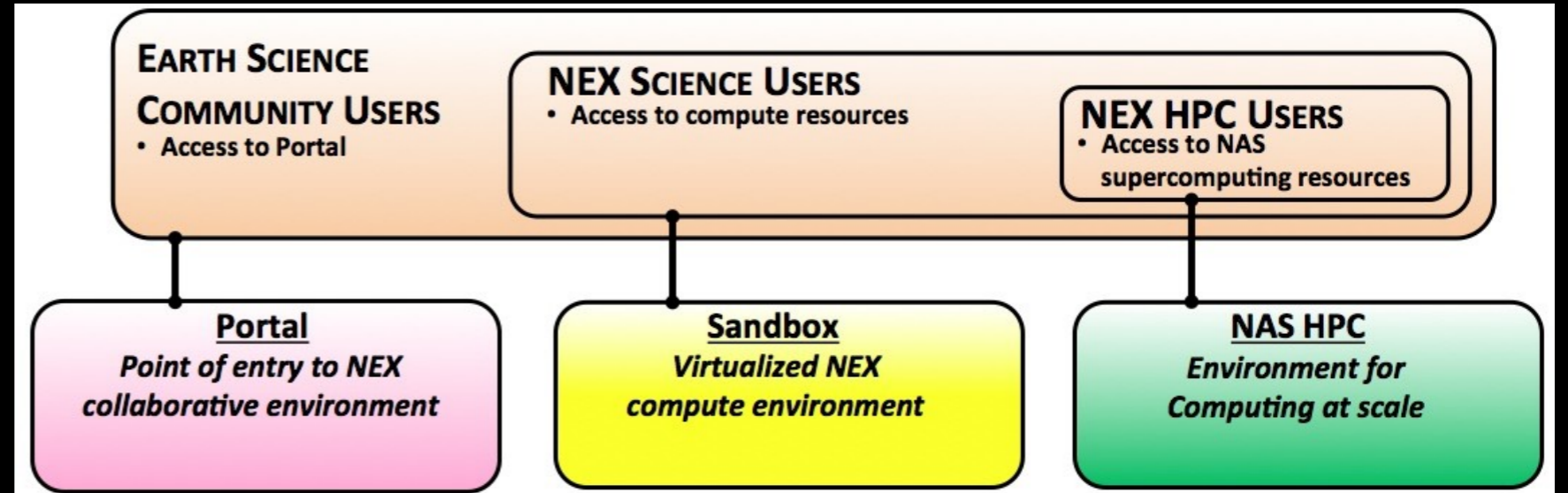


ACCESSIBILITY & LICENSE

NASA/NDC or OpenID account, need to signup and be approved



NEX sandbox
AWS, GEE, Azure, etc.



<https://earthengine.google.com/signup>. Google evaluates the request
Free for research, education, and nonprofit use.

Google Earth Engine

- Can import your own data, download data & images to work with your own tools
- You own your written algorithms and results on Google EE

- Cloud-computing platforms/resources are currently available
- Google Earth Engine is growing very fast, both basic and advanced remote sensing processing tools are available, free access
- NEX is more than just a computing platform, is a community and knowledge based system



- Cloud-computing platforms/ resources are currently available
- Google Earth Engine is growing very fast, both basic and advanced remote sensing processing tools are available, free access
- NEX is more than just a computing platform, is a community and knowledge based system



Ground truth?

Regional, local context?

Users (ready?)

Crowd-2-Cloud



REMOTE SENSING

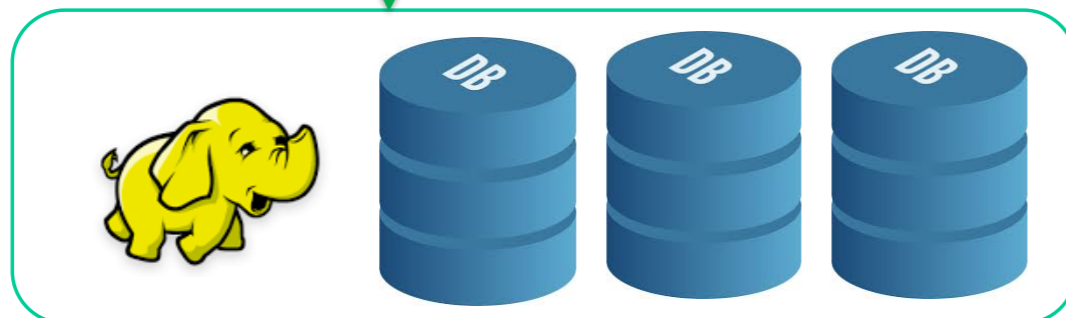
CROWD-SOURCING



**PROCESSING SERVICES
ON CLOUD COMPUTING**

Training

Validation



Applications

- Disaster recovery monitoring
- Damage mapping

BIG DATA ANALYTICS

DISASTER

URBAN

ECOLOGY

AGRICULTURE

PUBLIC HEALTH

- Urban green space
- Urban construction monitoring

- Megafauna index

- Oil palm plantation management

- Health care delivery
- Disaster Health

- Remote sensing processing services on the cloud
- Crowd-sourcing data quality assessment

on-demand, scalable processing capabilities
efficiently co-located with large image databases
rapid deployment of new algorithms
sharing, collaborative platform

**CLOUD
COMPUTING
PLATFORM**

on-demand, scalable processing capabilities
efficiently co-located with large image databases
rapid deployment of new algorithms
sharing, collaborative platform

**CLOUD
COMPUTING
PLATFORM**

Promote open resources (data, tools, etc.)

NEX and GEE

on-demand, scalable processing capabilities
efficiently co-located with large image databases
rapid deployment of new algorithms
sharing, collaborative platform

**CLOUD
COMPUTING
PLATFORM**

Promote open resources (data, tools, etc.)

NEX and GEE

Potential or limitation depends on user's

**OBJECTIVES
TECHNICAL LEVELS**

on-demand, scalable processing capabilities
efficiently co-located with large image databases
rapid deployment of new algorithms
sharing, collaborative platform

CLOUD COMPUTING PLATFORM

Promote open resources (data, tools, etc.)

NEX and GEE

Potential or limitation depends on user's

OBJECTIVES TECHNICAL LEVELS

Google Earth Engine

PaaS and SaaS



potential to deploy to various
SaaS on demand

on-demand, scalable processing capabilities
efficiently co-located with large image databases
rapid deployment of new algorithms
sharing, collaborative platform

CLOUD COMPUTING PLATFORM

Promote open resources (data, tools, etc.)

NEX and GEE

Potential or limitation depends on user's

OBJECTIVES TECHNICAL LEVELS

Google Earth Engine

PaaS and SaaS



potential to deploy to various
SaaS on demand

Many other infrastructures, platforms were not discussed here including ESA,
AIST (Japan), other national programmes

THANK YOU