MODELLING EO DATA PROCESSING WORKLOADS AND ENVIRONMENTS

Maximilian Schwinger, DLR

8.11.2022

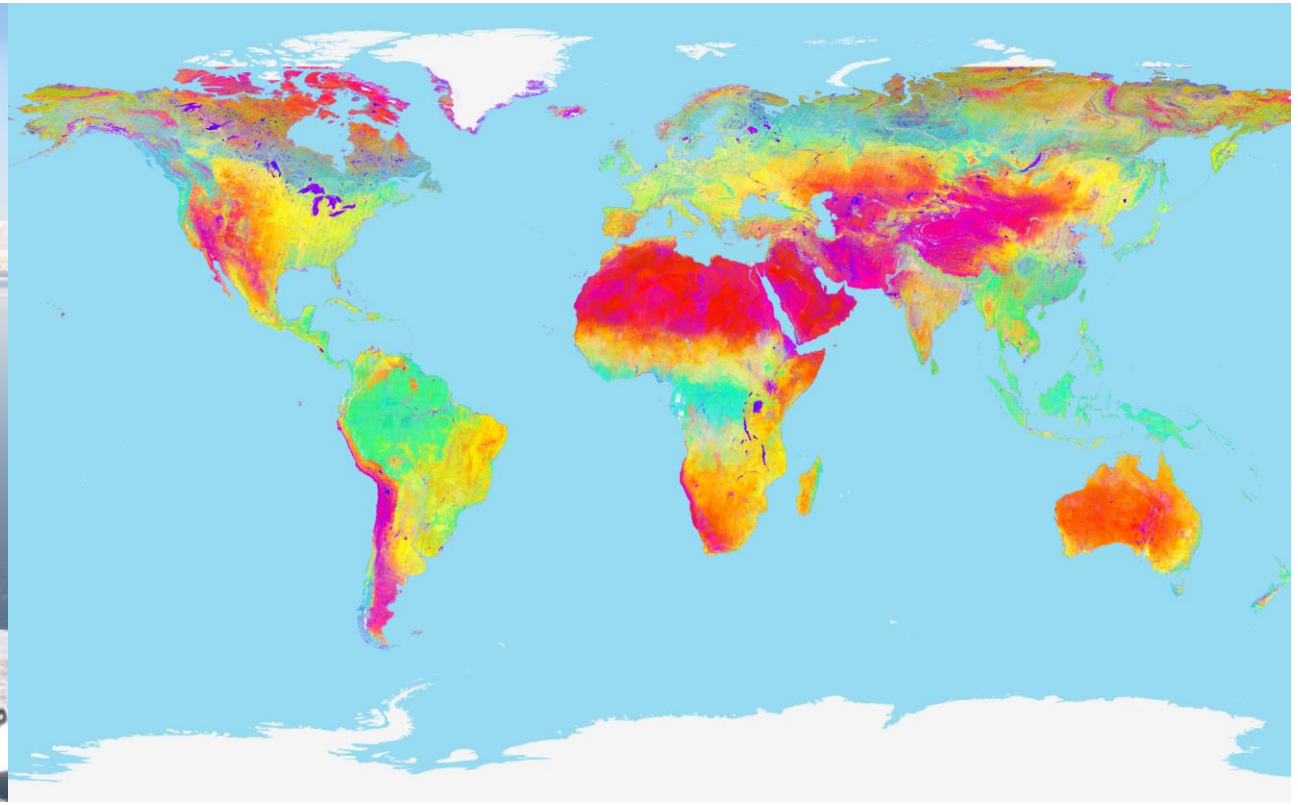
DLR

Earth Observation Center EOC

German Remote Sensing Data Center and Institute for Remote Sensing Methodology

EO Payload Data Ground Segment

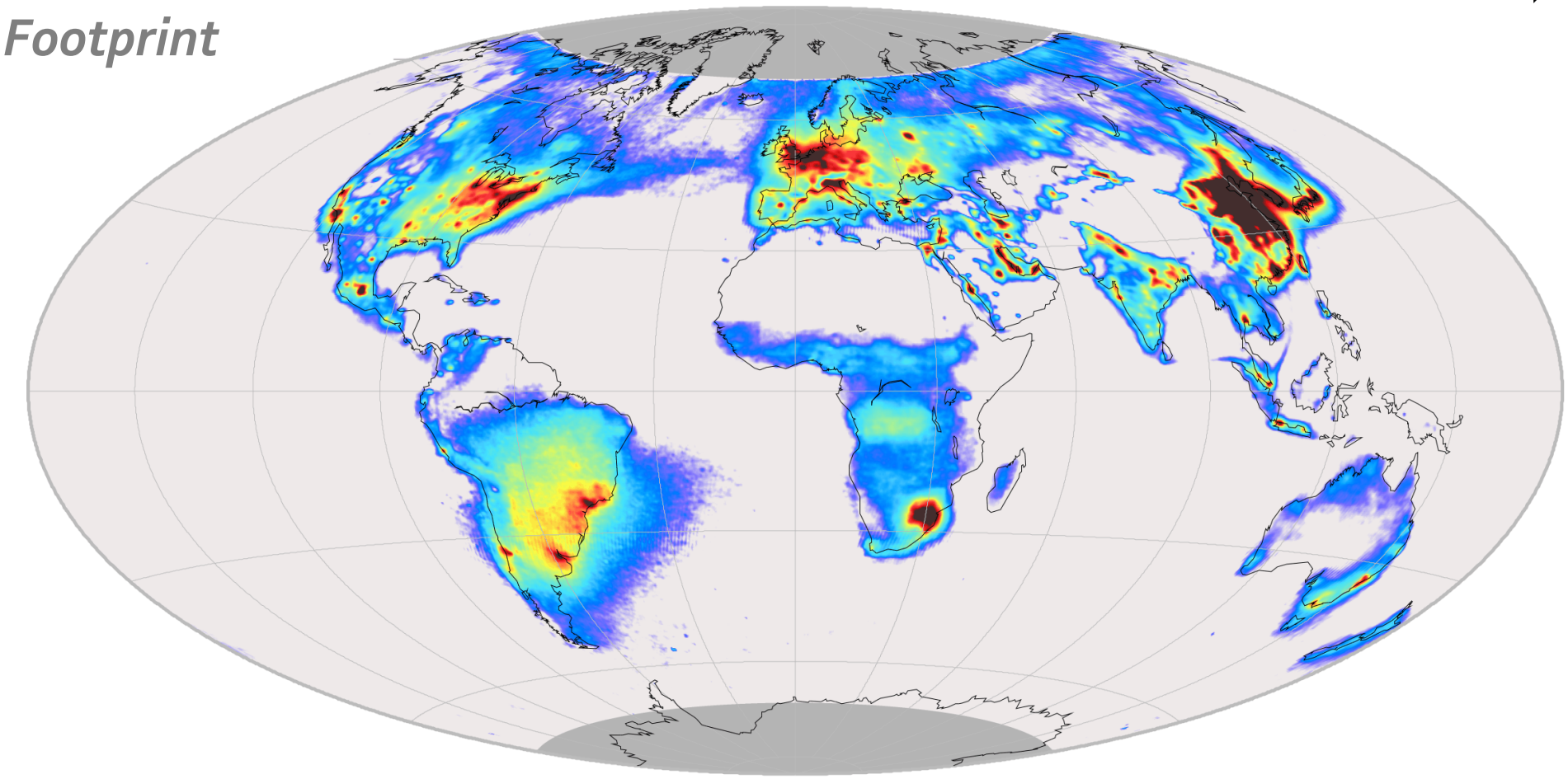
Global EO data



Why Satellite Based Earth Observation?

Global NO2 Distribution

The Human Footprint

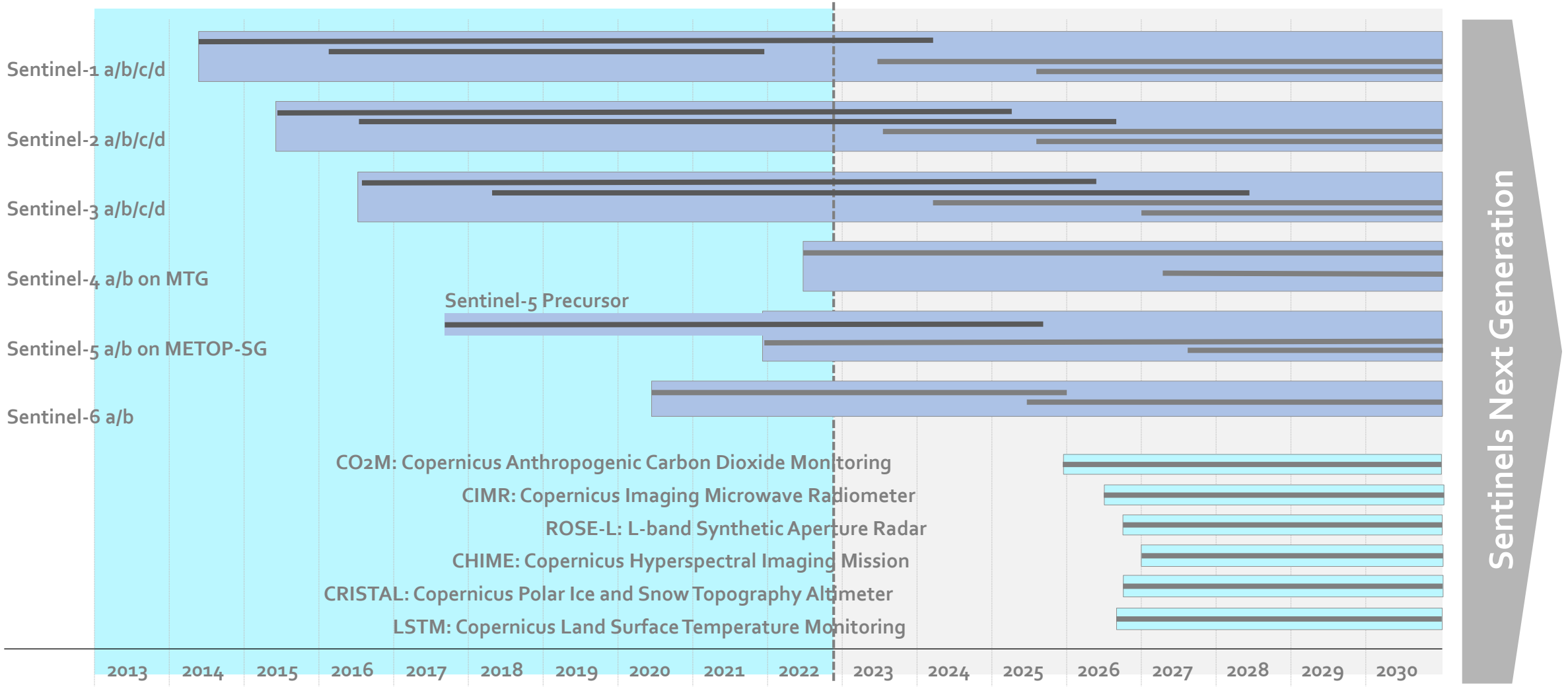


MetOp/GOME-2
Long-term mean
2007 -2016

NO_2 [10^{15} Moleküle/cm²]

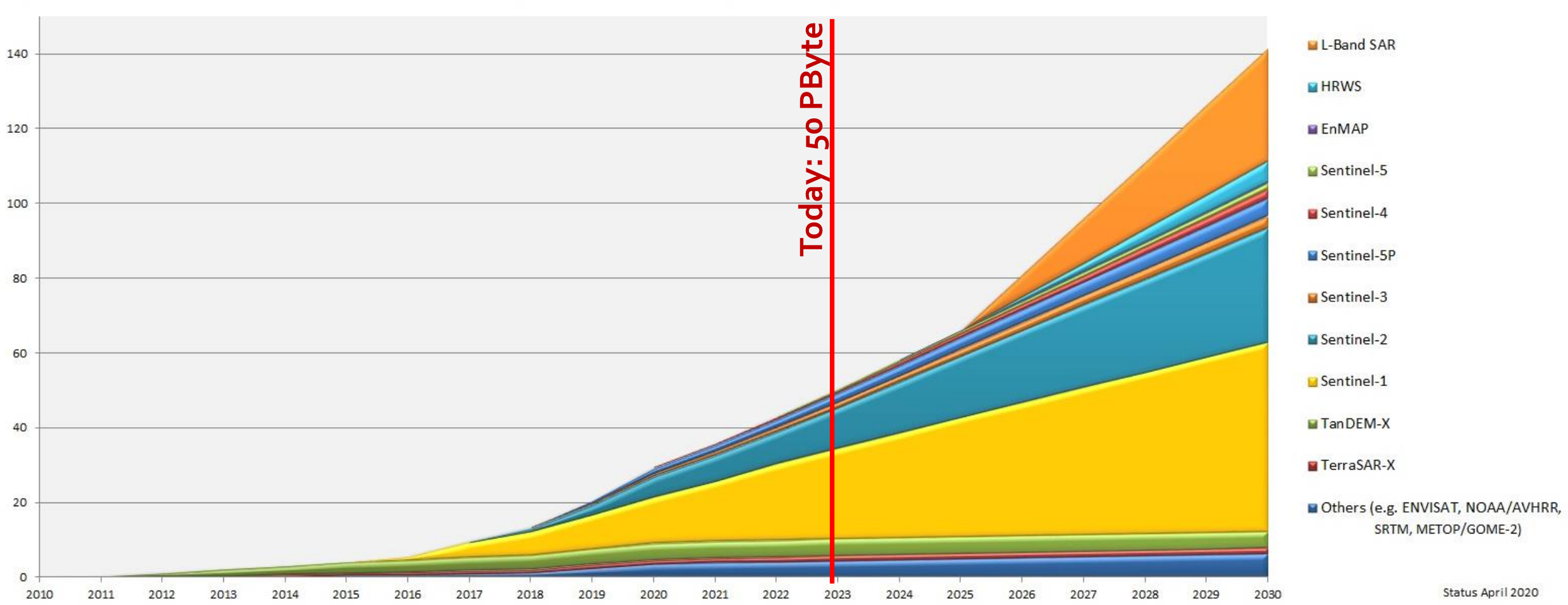
1 2 3 4 5

Copernicus Sentinel Missionen



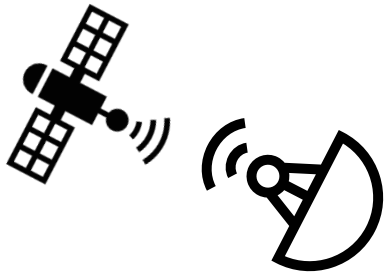
Data Volume in the German Satellite Data Archive

DFD Oberpfaffenhofen and Neustrelitz

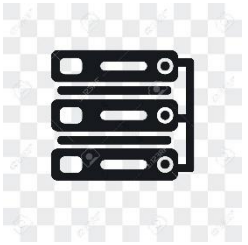


Introduction
EO data processing

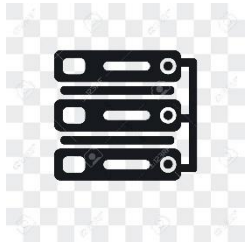
Mission Specific Processing



Data Reception



L1 Processing



L2 Processing

Dissemination

EO Science

Sensor Fusion

Information Extraction

Problem Statement/Research Direction

EO data processing is currently done mostly in dedicated infrastuctures due to

- I/O overhead
- Memory requirements
- Input data availability
- Legacy software character

Problem Statement/Research Direction

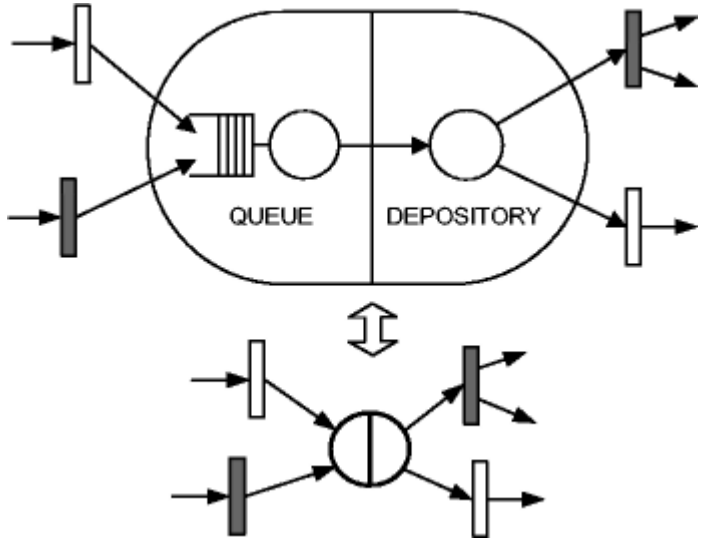
Useage of a shared infrastructure very desireable, but matching workloads, scenarios and infrastructure has to be done without empirical data of behaviour of combined workloads on shared infrastructure.

→ Modelling workload and system behavior to gain insight :

- Requirements of different workloads and degrees of freedom
- Combination of workloads
- Systematic understanding of pros/cons in workload/infrastructure combination

Queueing Petri Nets (QPNs = QNs + PNs)

- Combine queueing networks and Petri nets
- Can depict scheduling strategies
- Allow for qualitative and quantitative analysis



S. Kounev, and A. Buchmann, "On the Use of Queueing Petri Nets for Modeling and Performance Analysis of Distributed Systems", in Petri Net, Theory and Applications. London, United Kingdom: IntechOpen, 2008 [Online]. Available: https://www.intechopen.com/chapters/715 doi: 10.5772/5317

Pros and Cons of QPNs

PRO:
Combine expressive power of QN and PNs.
Modelling of SW and HW-Behaviour

CON:
State space explosion problem: Exponential growth of complexity

Example: MultiSAR
Analysis Ready Data for multiple SAR-instruments

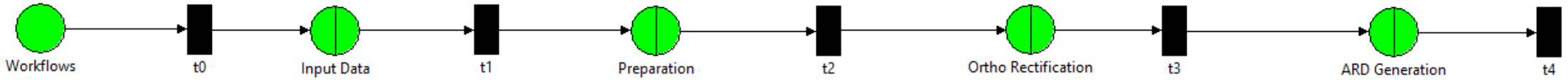
Generates Analysis Ready Data from SAR (Synthetic Aparture Radar) Data

Major steps of the workflow are:

- Inputpreparation
  - Metadataanalysis
  - Radiometric correction
  - Orbitdata preparation
- Orthorectification
  - Radiometric terrain flattening
  - Geocoded Incidence Angle Mask
- Enrichment of SAR Product with
  - Incodent Angle Mask
  - Shadow Layover Mask
  - Sigma-Gamma-ratio

Example: MultiSAR
representation of workflow and Environment

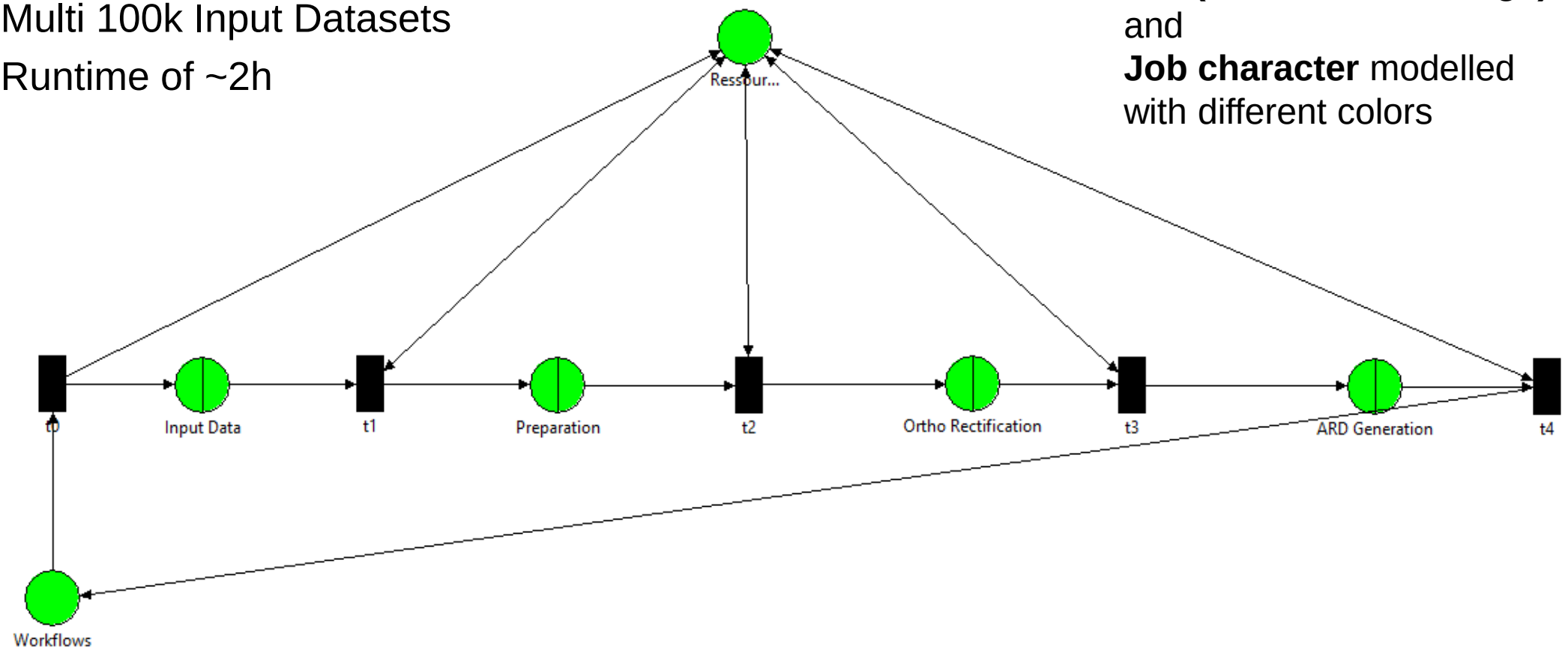
- Simple sequential workflow
- No intrinsic parallelization
- Multi 100k Input Datasets
- Runtime of ~2h



Example: MultiSAR
representation of workflow and Environment

- Simple sequential workflow
- No intrinsic parallelization
- Multi 100k Input Datasets
- Runtime of ~2h

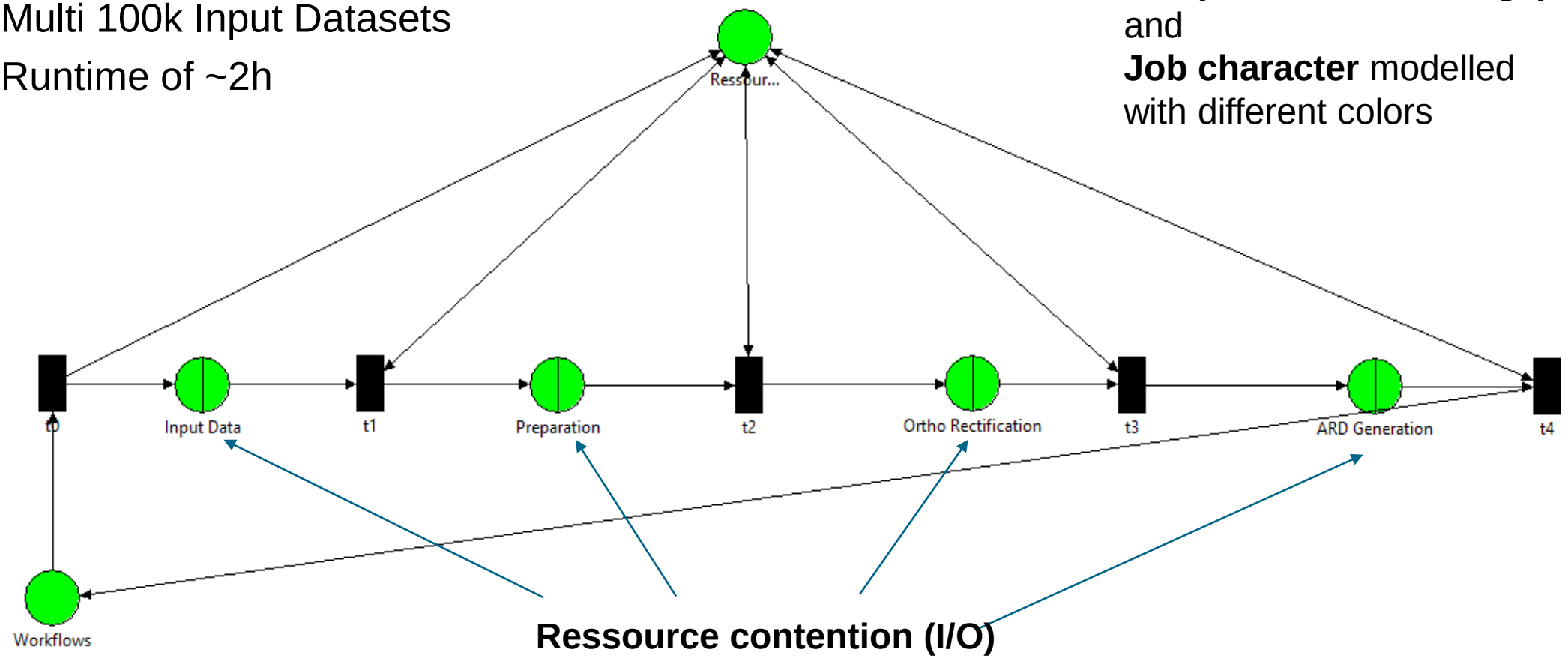
Guaranteed Ressources (RAM, CPU, Storage) and **Job character** modelled with different colors



Example: MultiSAR
representation of workflow and Environment

- Simple sequential workflow
- No intrinsic parallelization
- Multi 100k Input Datasets
- Runtime of ~2h

Guaranteed Ressources (RAM, CPU, Storage)
and
Job character modelled
with different colors



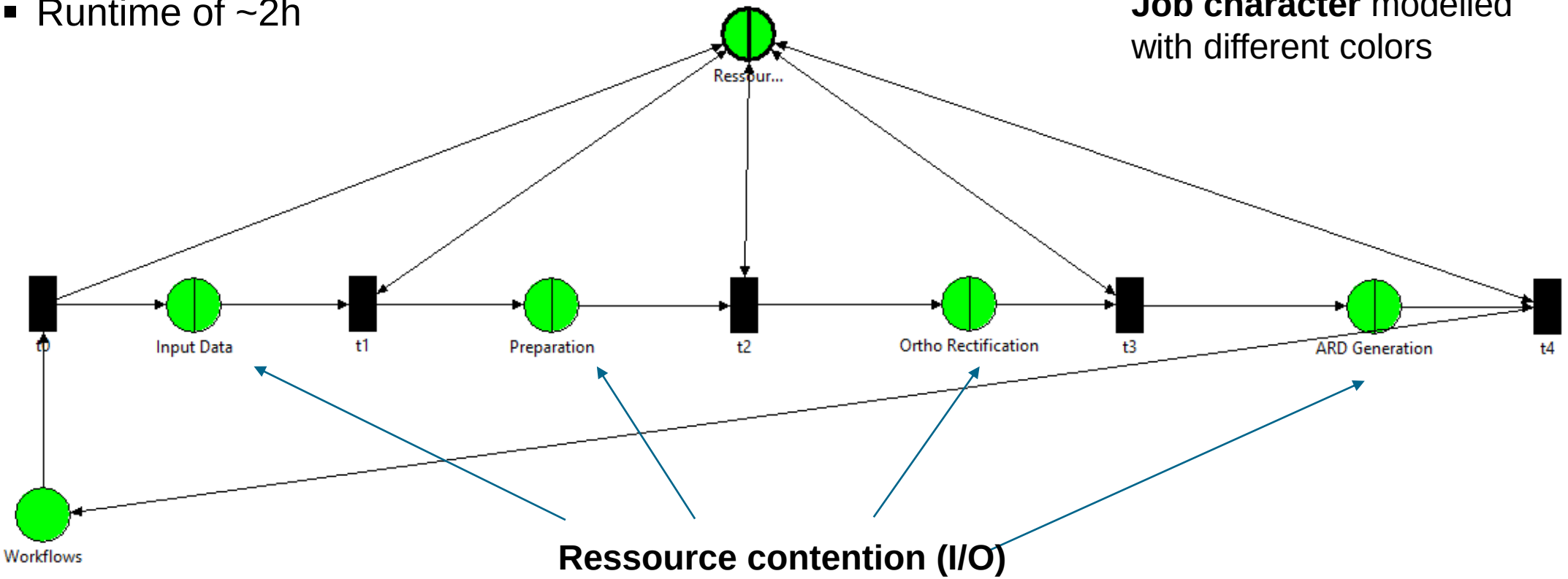
Ressource contention (I/O)
modelled
with shared queues

Example: MultiSAR
representation of workflow and Environment

- Simple sequential workflow
- No intrinsic parallelization
- Multi 100k Input Datasets
- Runtime of ~2h

Behavior of
environment

Guaranteed Ressources
(RAM, CPU, Storage)
and
Job character modelled
with different colors



Ressource contention (I/O)
modelled
with shared queues

terrabyte
a high performance data analytics (HPDA) platform

Project cooperation with the Leibniz Supercomputing Center (LRZ)

– DLR focus on EO research, services and user support

– LRZ focus on IT research and operation

Conclusion/Questions/Discussion

- Will QPNs modelling power be sufficient?
- Which risks are involved?

Maximilian Schwinger
DLR/DFD
maximilian.schwinger@dlr.de
+49 8153 28 1418