

SDC@KIS

Rucio/dCache in Solar Physics



Sep 2021
4th Rucio WS21

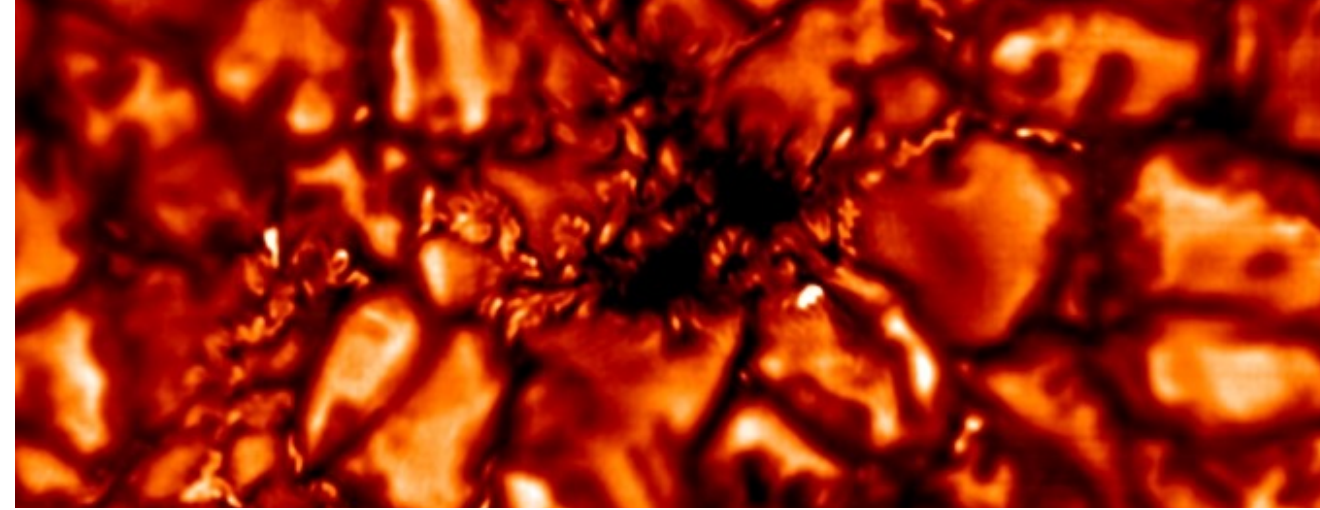
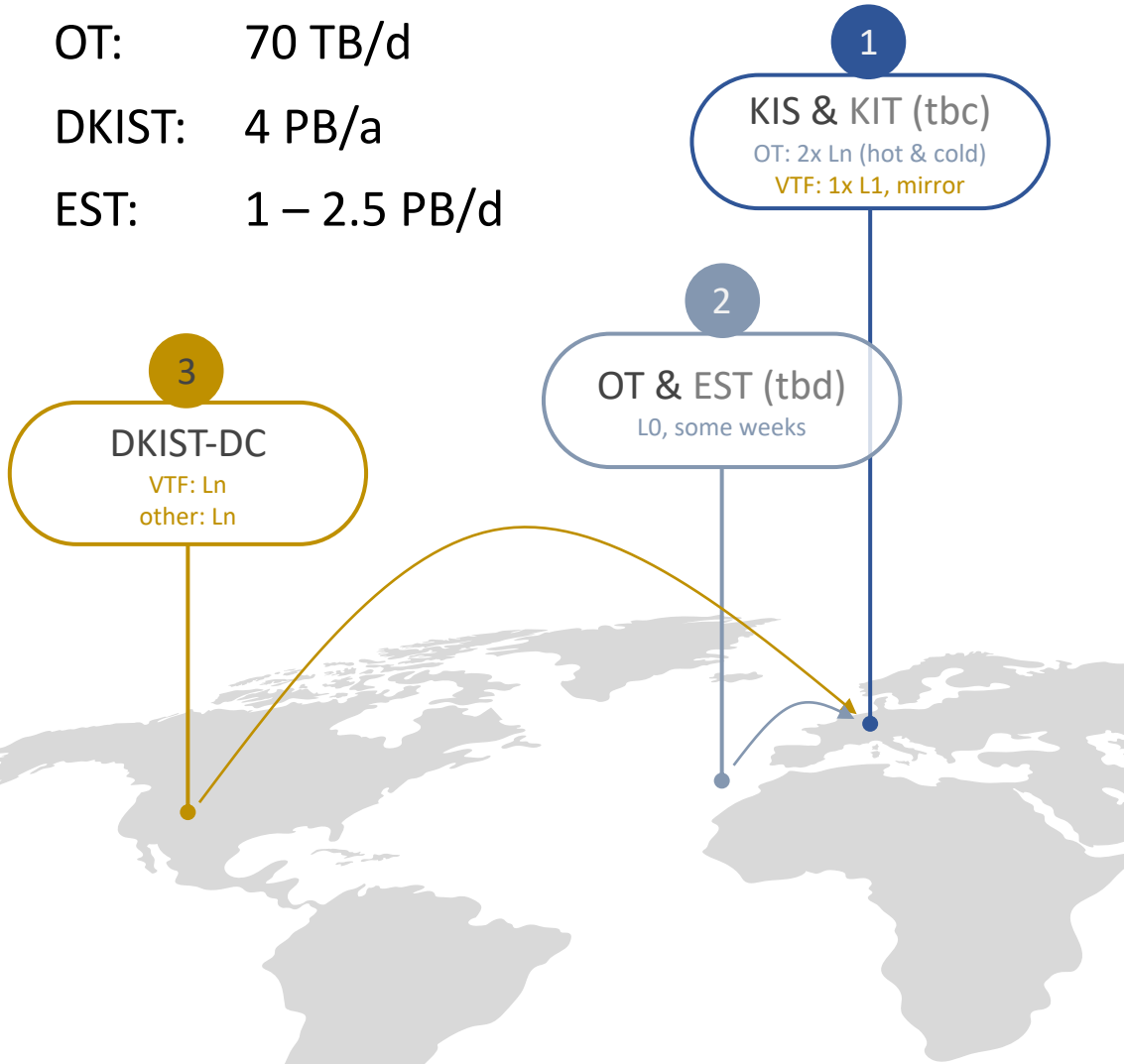
J. Beck, S. Berdyugina, P. Caligari, V. Gangadharan, N. Bello Gonzalez,
A. Gorobets, M. Günter, P. Kehusmaa, M. Knobloch, S. Müller, C. Schaffer,
T. Yakobchuk



SDC@KIS: Context

Data Volumes

OT: 70 TB/d
DKIST: 4 PB/a
EST: 1 – 2.5 PB/d



Solar photosphere by GREGOR (Kleint et al. 2020)

• KIS Mission

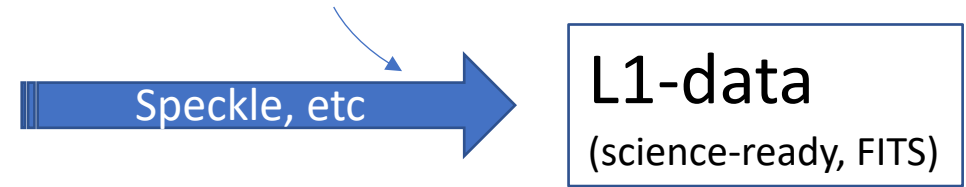
- fundamental research in solar & stellar astrophysics, ground-based (mainly)
 - operating the German Solar Telescopes on Tenerife (OT) GREGOR (largest in Europe) & VTT for: KIS, AIP, MPS, IAC, international time
 - building (parts of) optical instruments, e.g. VTF@DKIST (NSO), PHI@Solar Orbiter & Sunrise (MPS)
- ## • SDC is a distributed & tiered archive for
- *solar observational data*: mainly OT, planned partial mirroring from DKIST, a future central hub for EST DC
 - solar and stellar multi-D *simulations*
 - *higher-level data-products* thereof

Solar Observational Data

(instrument types, data heterogeneity)

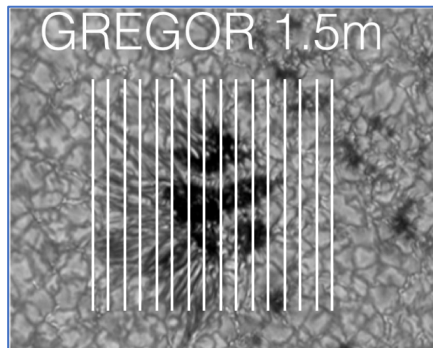
- **L0-data (raw): 5 dimensional data “cubes”**
 - 2D spatial
 - $\lambda \Rightarrow T, |B|, \text{etc. with height}$
 - polarimetric Information (I, Q, U, V) \Rightarrow **B vector**
 - time

terrestrial atm. data/model



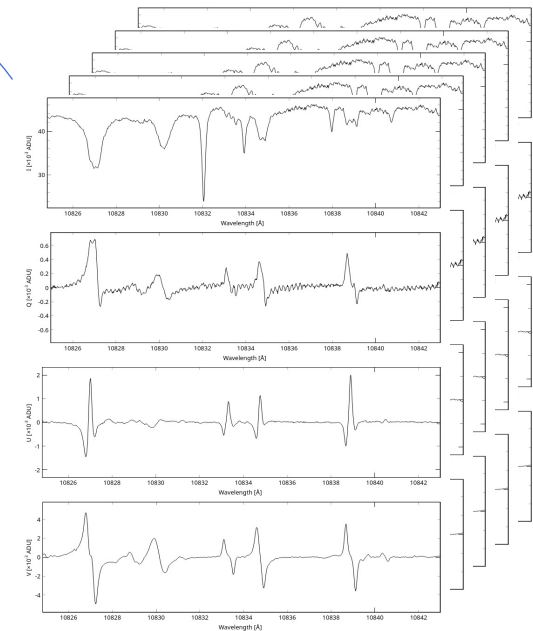
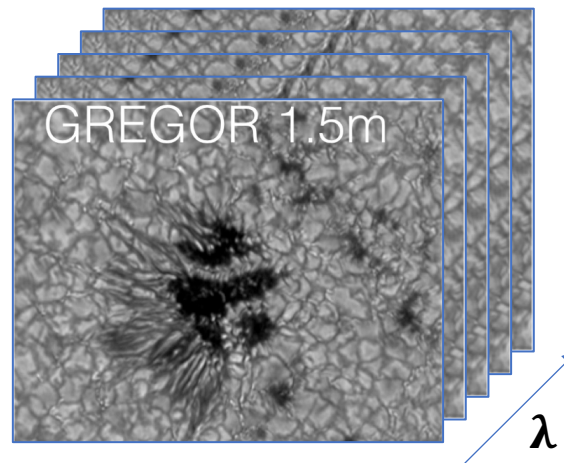
+ calib.

+ calib.



Schlichenmaier et al. (2016)

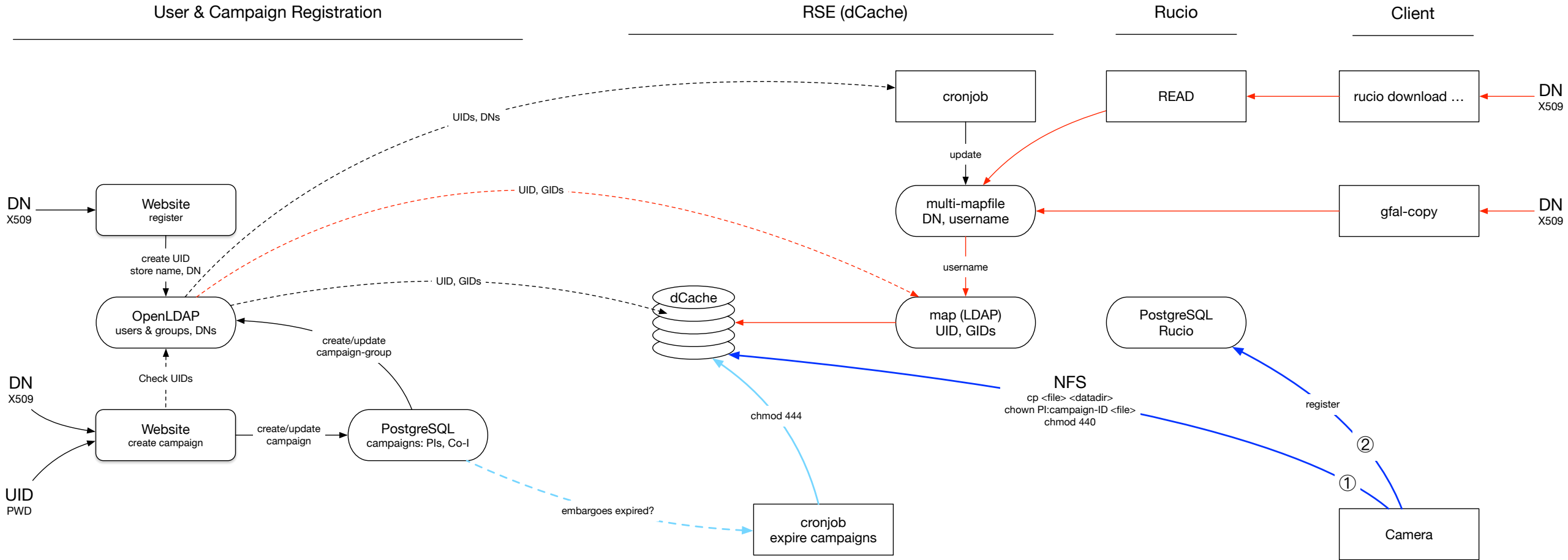
+ calib.



Embargoes

- Data and **data-products on all sites** are subject to embargoes
 - **OT: default 1 yr (2 yrs with Ph.D.)** from day of observation
- Varies for data from
 - technical campaigns
 - opportunistic campaigns (subject to target availability and atm. conditions)
 - 3rd party financed campaigns (EU SOLARNET)
 - data from other sources than OT
- During embargo data are
 - searchable by everyone (incl. metadata & previews) , but
 - only readable for: PI & Co-I (group must be subsequently customisable)
- Must be protocol agnostic
 - i.e.: respected by: rucio download ..., xrdcp, gfal-copy, WebDAV, nfs, ...

Embargoes Implementation (x509 only!)



Summery

- Now start building (@KIS); switching existing system (based on Mongo-DB) to Rucio
- Building SDC-nodes @OT early 2022
- Using on k8s and dCache on VMs based on Debian, ZFS
- We will start with disks only, no tapes (yet)
- Some questions are not finally decided, e.g.:
 - metadata implementation (within Rucio, separate DB)?
 - file naming (carrying *observations* to Rucio DSDs)?

Thanks to the Rucio team for this great product and their continuous support and help they provide to SDC