

Distributed Data Processing (DDP) experiences @SURF

DDP projects: key elements

Data intensive projects processing instrument data from sensors, sequencers, telescopes, and satellites during the entire mission lifetime

Data volumes:
parallel processing
of large amounts of
data, from many
Terabytes to Petabytes

Processing pipelines:
steady production
workflows with semi-
continuous data flows

Project organisation:
international
collaborations working
on a shared set of data
and software

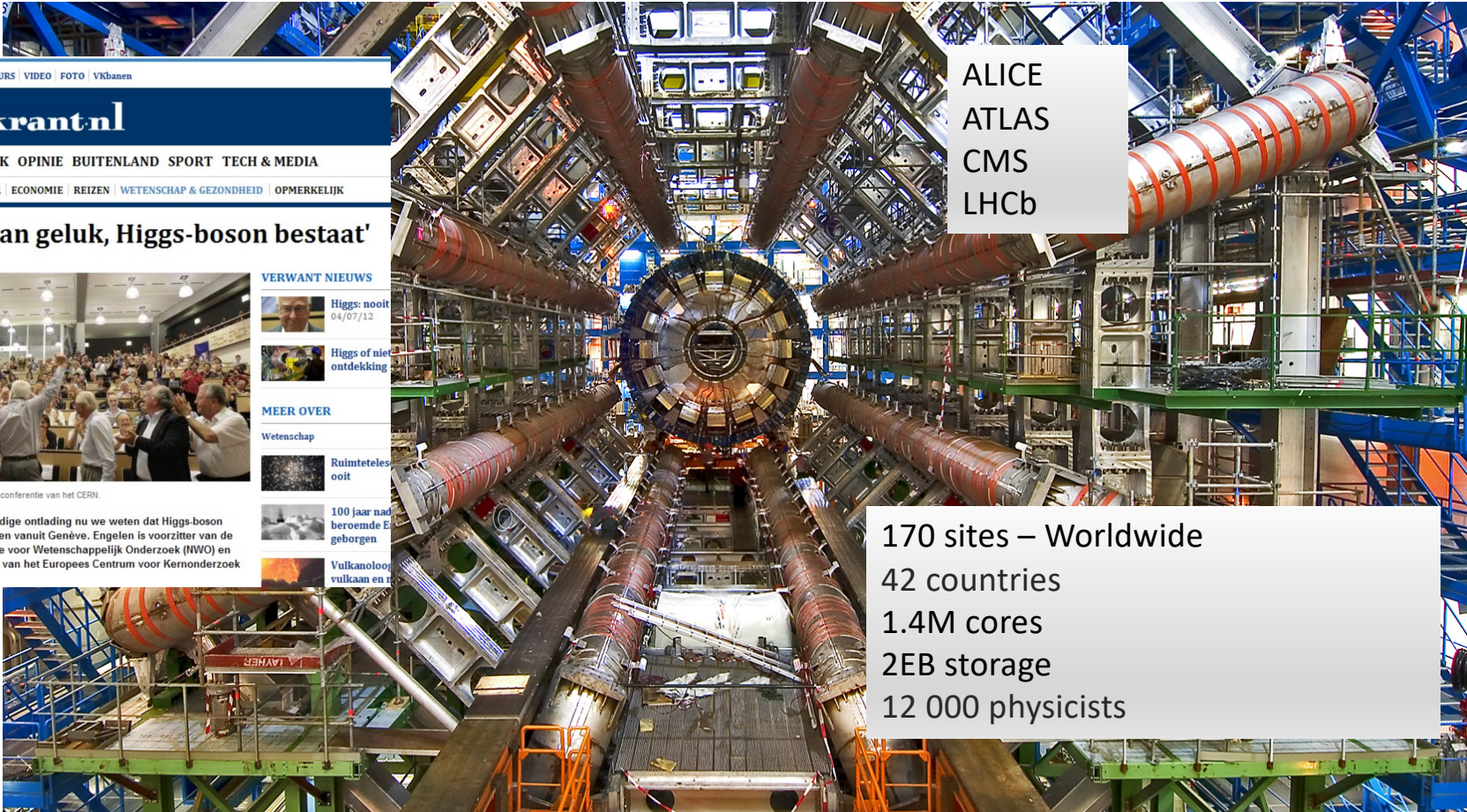
Ecosystem:
Portable and interoperable
solutions based on APIs
and industry standard
protocols

DDP services: Grid & Spider

Over 50PB of storage and over 100 million core hours are consumed together
by various scientific domains



WLCG/CERN



WEER | VERKEER | BEURS | VIDEO | FOTO | VKbanen

Volkskrant.nl

NIEUWS | POLITIEK | OPINIE | BUITENLAND | SPORT | TECH & MEDIA

BINNENLAND | CULTUUR | ECONOMIE | REIZEN | WETENSCHAP & GEZONDHEID | OPMERKELIJK

'Tranen van geluk, Higgs-boson bestaat'

04/07/12, 12:03 - bron: ANP

© EPA. Applaus tijdens de persconferentie van het CERN.

UPDATE Wat een geweldige ontlading nu we weten dat Higgs-boson bestaat', zegt Jos Engelen vanuit Genève. Engelen is voorzitter van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) en was voorheen directeur van het Europees Centrum voor Kernonderzoek (CERN) in Genève.

VERWANT NIEUWS

- Higgs: nooit 04/07/12
- Higgs of niet ontdekking

MEER OVER

Wetenschap

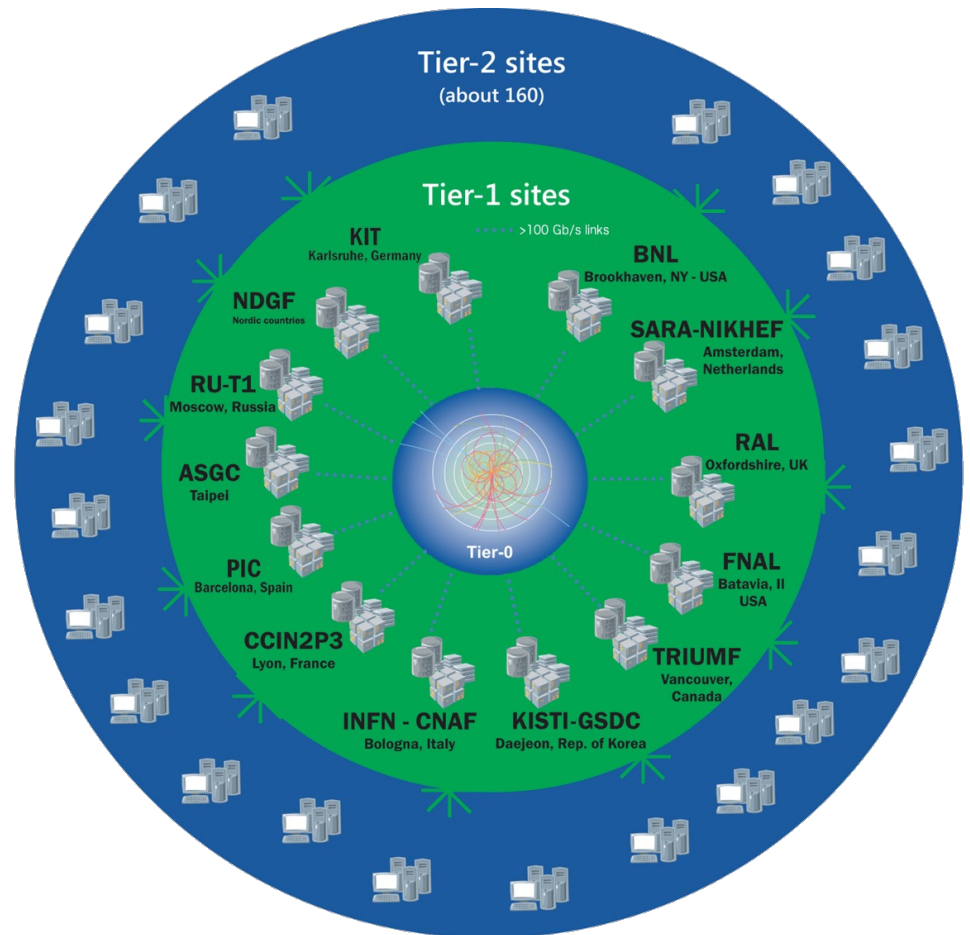
- Ruimteteloot
- 100 jaar nad beroemde E geborgen
- Vulkanoloog vulkaan en r

ALICE
ATLAS
CMS
LHCb

170 sites – Worldwide
42 countries
1.4M cores
2EB storage
12 000 physicists

Dutch Grid contribution to CERN

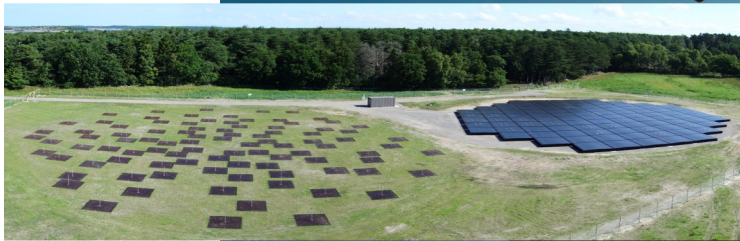
- Compute, storage, services
- SURF and NIKHEF are a tier 1 site
- Connected with dedicated, private, high-bandwidth network
- Allocations 2023:
Core hours: 37M @SURF, 40M @Nikhef
Storage: 37PB tape + 13PBdisk @SURF,
3.5disk @Nikhef



LOFAR

International LOFAR Telescope (ILT)

- Radio telescope
 - Data distributed to Long Term Archive (SURF, Jülich & Poznan)
 - Multi-purpose data: Surveys Key Science Projects, Transients, Pulsars, etc
 - Pathfinder for SKA
 - Largest astronomy archive in the world
- But looking forward to SKA 😊



Nançay

Dutch stations

LOFAR Core (NL)

Jülich

Effelsberg

Tautenburg

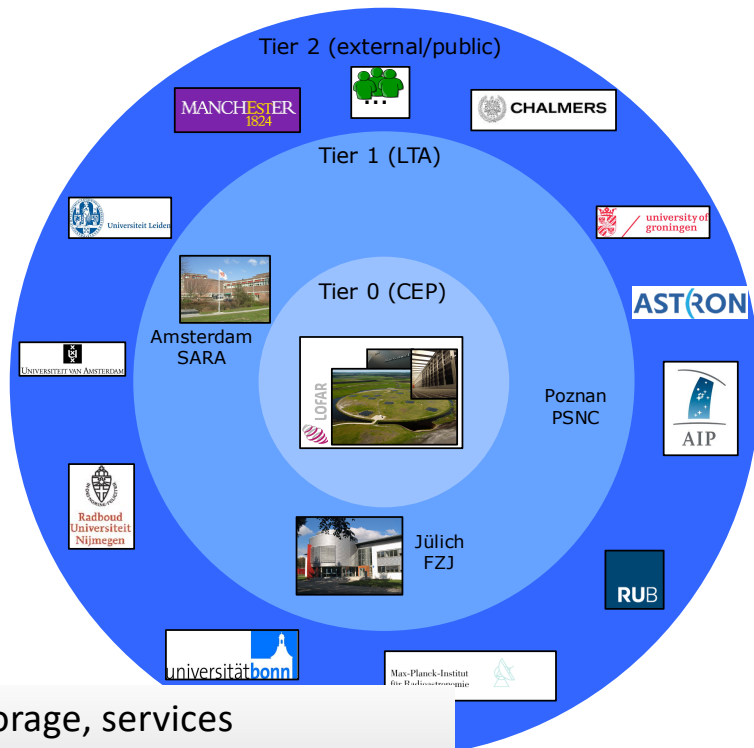
Unterweilenbach

Łazy

ASTRON

Netherlands Institute for Radio Astronomy

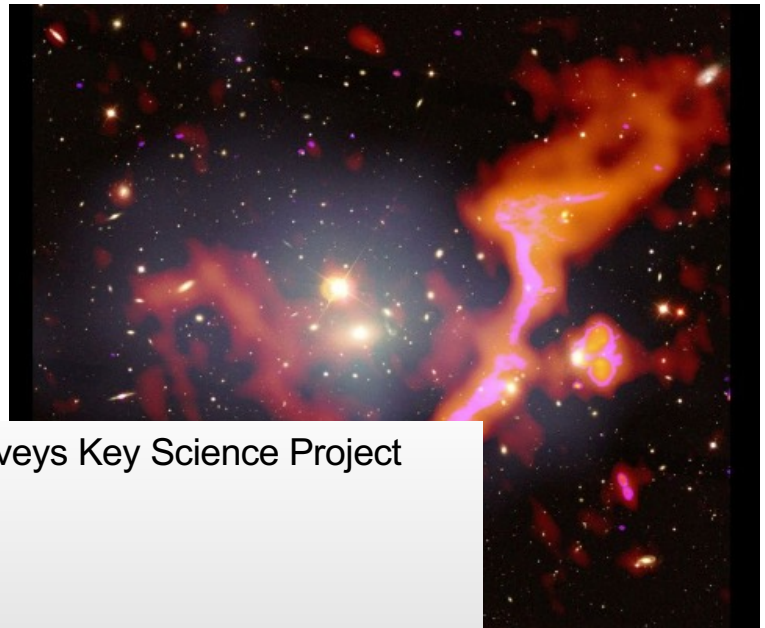
SURF contribution to LOFAR



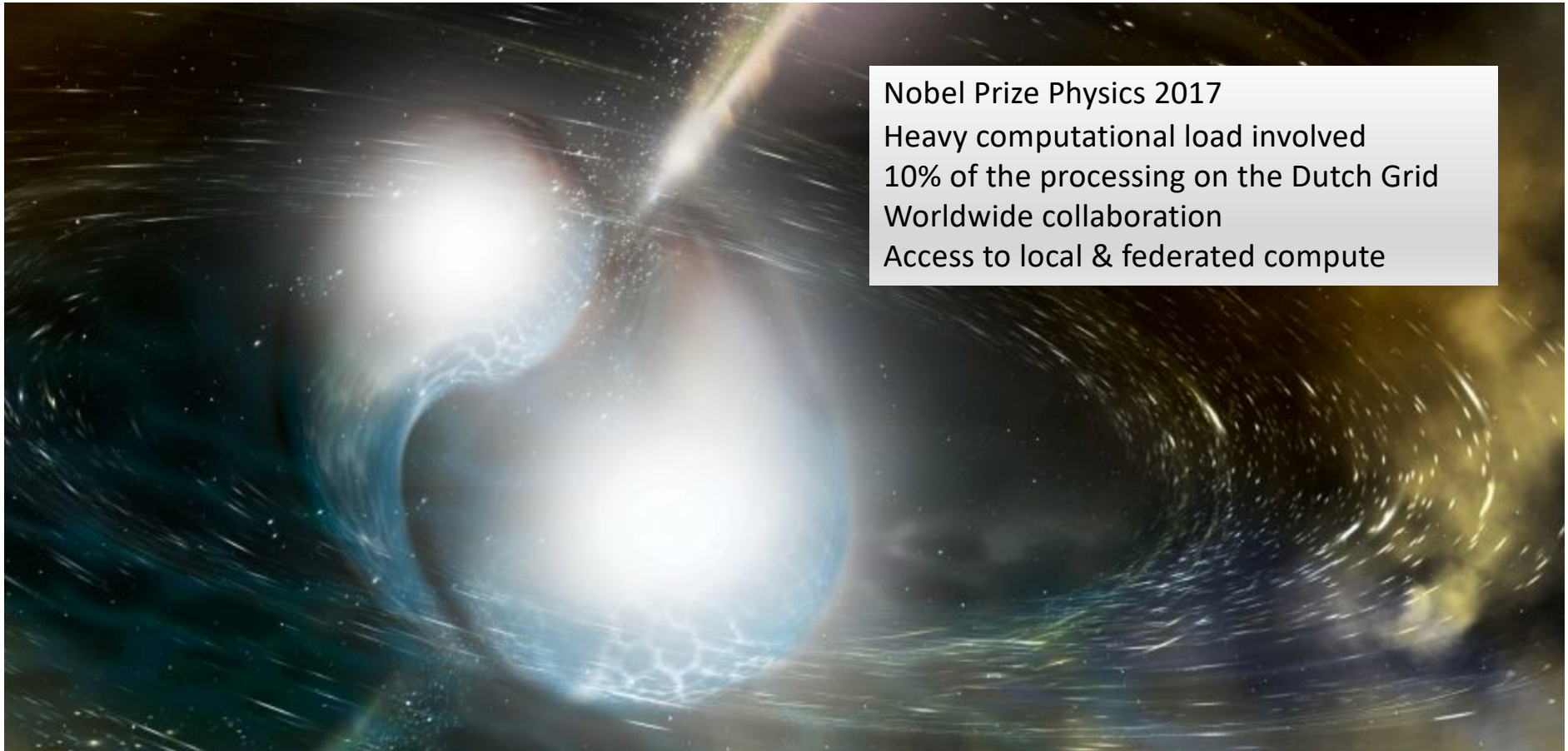
Compute, storage, services
30 PB stored at SURF
High-speed connection to the LTA
Processing power: Grid & Spider

LOFAR the Surveys Key Science Project

- 300 members
- 60 institutes
- 18 countries
- 175 papers
- Distribution of data products
- Processed over 1000 LOFAR observations (4M core hours, 300TB disk and 700TB tape)



LIGO-Virgo-KAGRA (LVK)



Nobel Prize Physics 2017
Heavy computational load involved
10% of the processing on the Dutch Grid
Worldwide collaboration
Access to local & federated compute

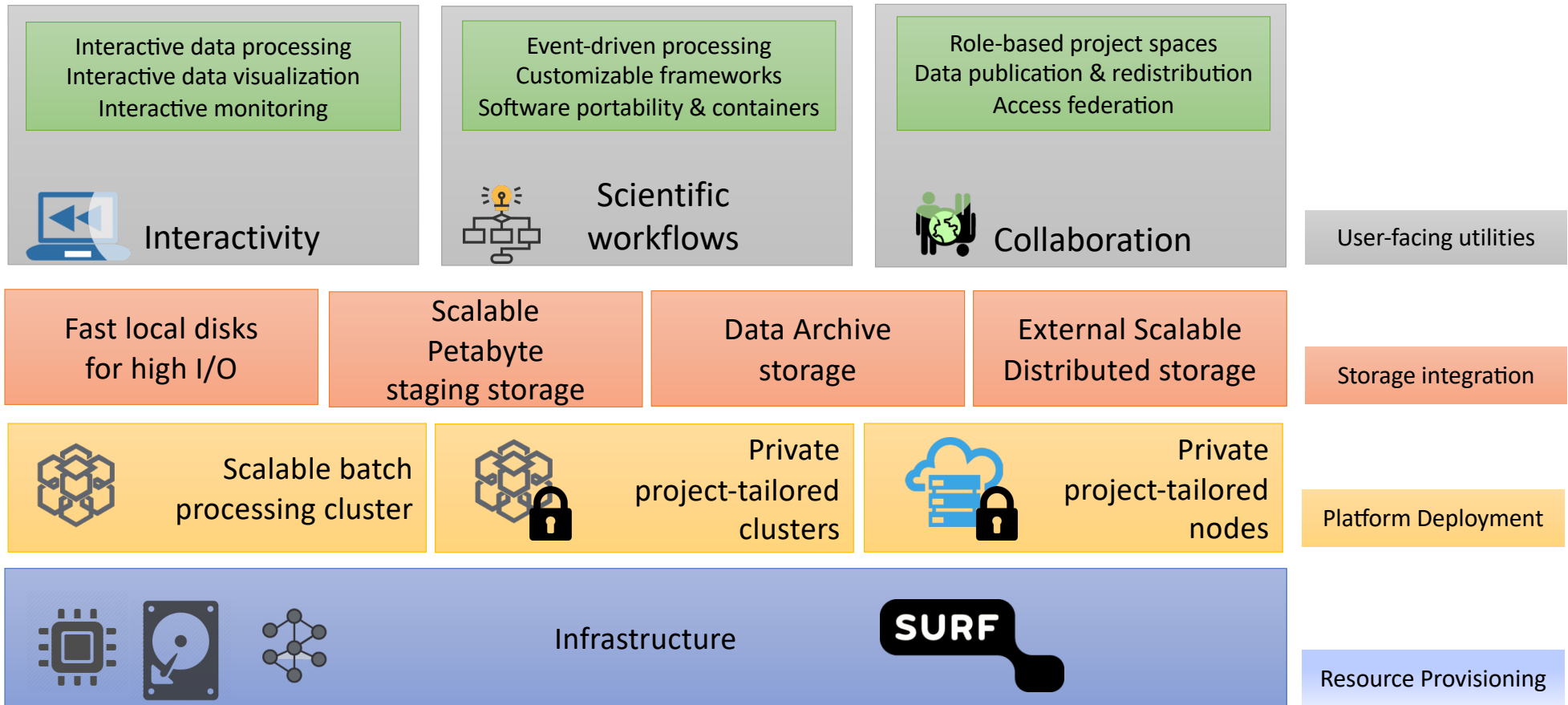
Dutch Grid contribution to LVK

- Compute, storage, services
 - Data retrieved from international cache servers
 - Hosting data distribution server for derived data
-
- Allocation 2023:
Core hours: 9M @SURF, 9M@Nikhef
Storage: 200TB disk @SURF



The four gravitational wave detectors in the LVK scientific collaboration: LIGO at Hanford in the USA (top left), LIGO at Livingston in the USA (bottom right), Virgo in Italy (bottom left) and KAGRA in Japan (top right).

DDP Architecture



More about DDP services at SURF [here](#)