EUROEXA
FUNDED BY THE EUROPEAN UNION

LOFAR-IT HPC Activities
A distributed e-Infrastructure for data analysis

Ugo Becciani – Giuliano Taffoni

LOFAR

Project ID: 754337

EURO EXA

Outline

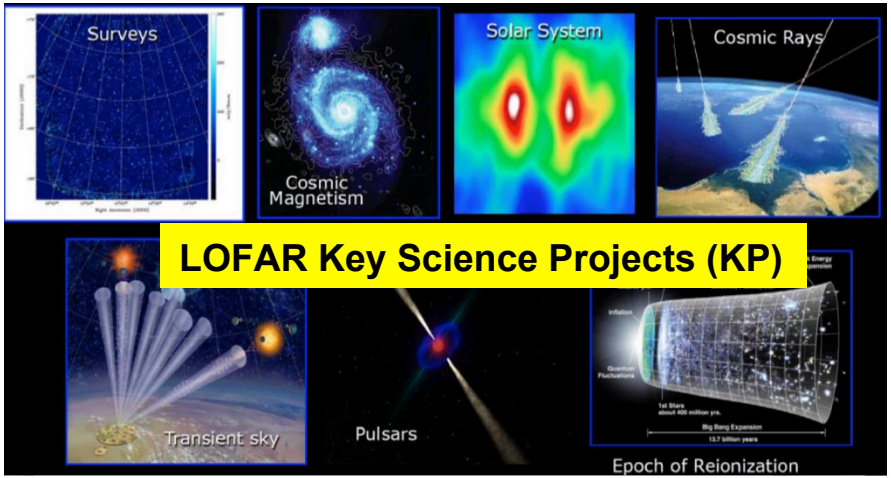
- The LOFAR.IT organization
- Italian participation to LOFAR
- The LOFAR Computing Model
- LOFAR-IT e-infrastructure:

  - Data analysis Requirements
  - Design and setup of the Data Analysis Infrastructure

Project ID: 754337

The International LOFAR Telescope

ILT consists of an interferometric array of dipole antenna stations, an it is distributed throughout 9 EU Countries: NL, Germany, France, Italy, Poland,UK, Sweden, Ireland, Latvia.



Project ID: 754337

LOFAR-IT Organization

Board: **Gianfranco Brunetti (Coordinator** INAF-IRA**) Ugo Becciani** (INAF-OA Catania) Segretario, **Federica Govoni** (INAF-OA Cagliari, UTGII), **Francesco Massaro** (UniTo), **Jader Monari** (INAF-IRA), **Roberto Scaramella** (INAF-OA Roma)

Science Advisory Committee: **Andrea Ferrara** (Chair), Matteo Murgia, Mauro Messerotti, Grazia Umana, Gianni Bernardi, Ettore Carretti, Isabella Prandoni, Laura Pentericci, Marta Burgay, Rossella Cassano, Andrea Chiavassa (UniTO)

Technological joint WG ASTRON-INAF: Established on March 2018 . Leaded by Astron. Primary objective: joint development of RCU for LOFAR 2.0 and eventually LBA2.0.

Data WG: **Giuliano Taffoni** (INAF-OA Trieste) - Chair, Alessandro Costa (INAF-OA Catania), Francesco Bedosti (INAF-IRA), Cristina Knapic (INAF-OA Trieste), Manuela Magliocchetti (INAF-IAPS Roma), Annalisa Bonafede (UniBo, Associata INAF IRA)

Project ID: 754337

Main activities of the Consortium

- Build a LOFAR 2.0 station in Medicina (2020-2022)
- Build a LOFAR data analysis infrastructure
- Implement a technological and scientific collaboration with ASTRON
- Develop a community that is able to work with LOFAR data (for science and technology)
- Participation of Italian community to Key Projects (surveys in particular)

Participation to the KPs, LOFAR guarantee time

Call For Interest on key projects

Close 30 Aug 2018

10 Proposal

Summary and people Involved

- KP SURVEYS => 5 + 17 + 1 TD new tot = 18
- KP EOR => 0 + 4 + 0 TD
- KP TRANSIENT => 0 + 5 + 2 TD
- KP MAG => 1 + 1 + 1 TD
- KP SUN => 0 + 6 + 0 TD

Project ID: 754337

Observation Time obtained : Italian Proposals

Cycle 11

Four Proposal approved by ILT TAC

Observation Time obtained : ***55 hours***

The following Single-Cycle proposals were reviewed

LC11_001	P. Zucca	72.0	46.0
LC11_002	A. Rowlinson	44.0	17.4
LC11_004	S. Purser	32.0	32.0
LC11_005	P. Leto	13.1	13.1
LC11_006	H. Vedantham	32.0	0.0
LC11_007	O. Wucknitz	9.0	9.0
LC11_008	A. Merloni	104.0	0.0
LC11_009	V. Cuciti	33.3	25.1
LC11_010	F. De Gasperin	67.0	67.0
LC11_011	A. Offringa	16.1	16.1
LC11_012	G. Bruni	25.0	0.0
LC11_013	V. Heesen	25.0	25.0
LC11_014	A. Ignesti	8.7	8.7
LC11_015	S. Mandal	50.0	17.0
LC11_016	F. Bempong-Manful	41.7	41.7
LC11_017	M. Murgia	8.0	8.0
LC11_018	M. Knapp	32.0	16.0
LC11_019	H. Intema	8.4	8.4
LC11_020	J. Callingham	174.0	174.0
LC11_021	J. Broderick	12.0	8.0

LOFAR - COMPUTING MODEL

- 53 Stations (24 core (Exloo), 16 remote, 13 International)
- A LOFAR core station consists of 96 Low Band Antennas (LBAs), operating from 10 to 90MHz and 48 High Band Antenna (HBA) tiles that cover the frequency range from 110 to 250 MHz
- Remote stations in the Netherlands have the same number of HBA tiles, and LBAs
- International stations provide a single cluster of 96 HBA tiles and 96 LBAs (6 station in Germany, 3 in Poland, 1 in France, Ireland and UK)

Project ID: 754337

LOFAR - Data flow

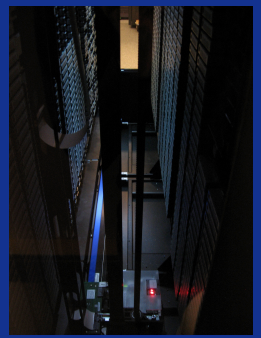
Central Processor system

Stations

Realtime system

Offline processing

Long Term Archive



240 Gbit/s

110 Gbit/s

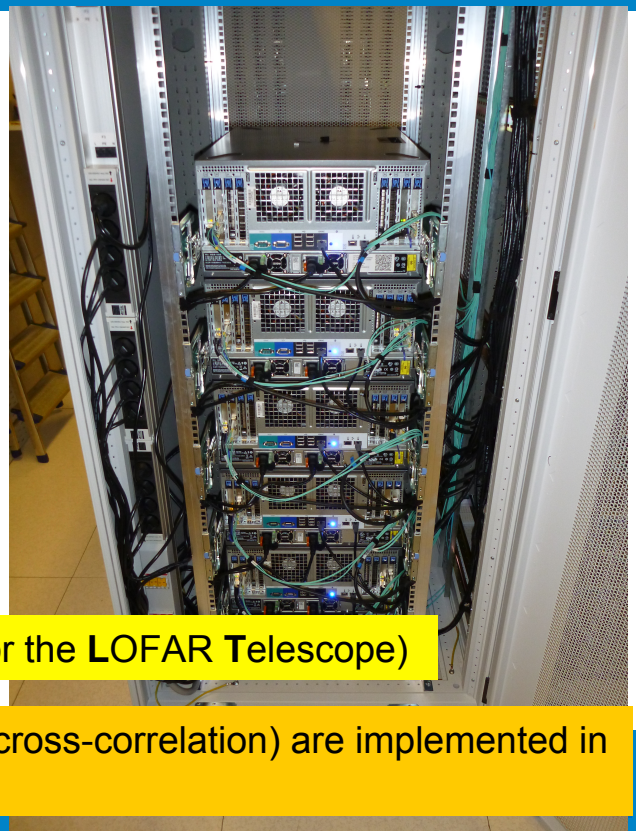
2 Gbit/s

Project ID: 754337

LOFAR COMPUTING MODEL

Central Processor System:The Correlator

- 9 Dell T620 nodes
  - Dual Intel Xeon E5-26xx
  - 2 Nvidia Tesla K10
  - 2 Dual port 10GbE
  - 2 FDR Infiniband HCA



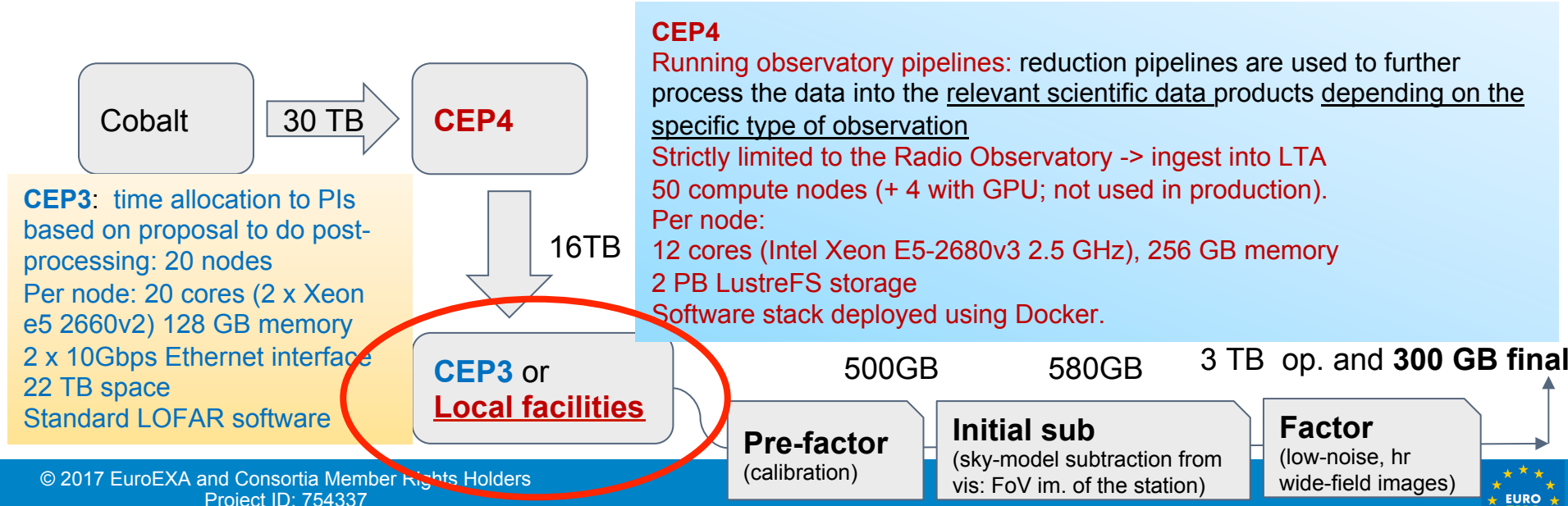
COBALT (**CO**rrelator and **B**eamformer **A**pplication for the **L**OFAR **T**elescope)

Both the F-stage (Fourier transfom) and the X-stage (cross-correlation) are implemented in GPUs

LOFAR COMPUTING MODEL
Central Processor System: Post Processing cluster

Two central processing (CEP) clusters in Groningen (i.e. near the correlator). Pipelines use locally-developed generic framework.

Distributed system built using a co-design approach (we know the algorithms and we design the HW)



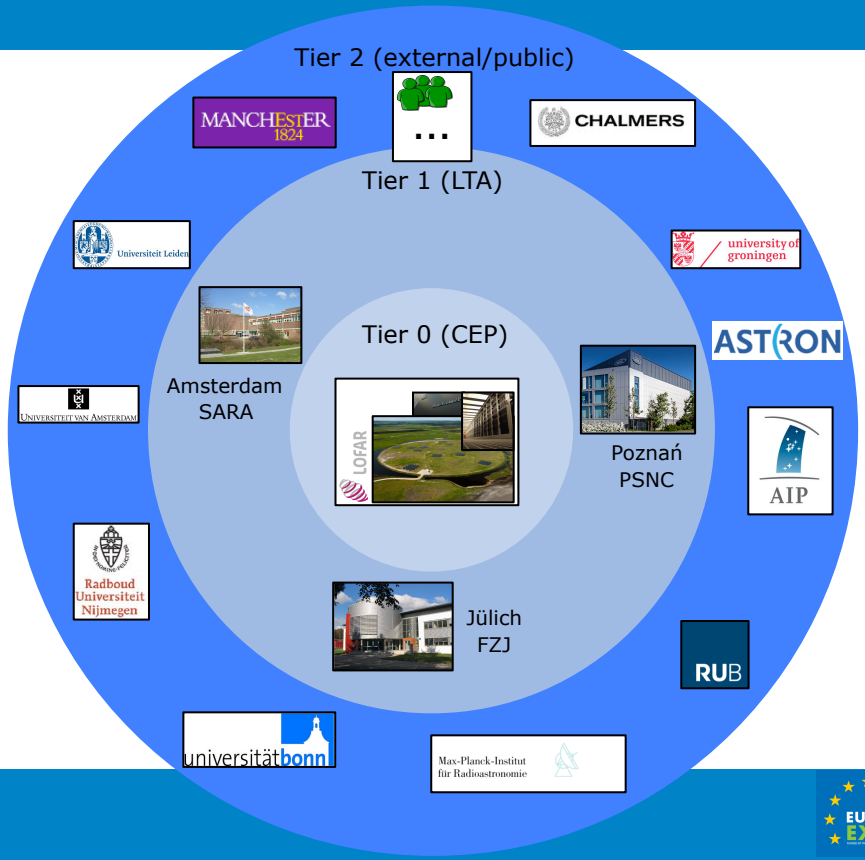
CEP3: time allocation to PIs based on proposal to do post-processing: 20 nodes
Per node: 20 cores (2 x Xeon e5 2660v2) 128 GB memory
2 x 10Gbps Ethernet interface
22 TB space
Standard LOFAR software

CEP4
Running observatory pipelines: reduction pipelines are used to further process the data into the relevant scientific data products depending on the specific type of observation
Strictly limited to the Radio Observatory -> ingest into LTA
50 compute nodes (+ 4 with GPU; not used in production).
Per node:
12 cores (Intel Xeon E5-2680v3 2.5 GHz), 256 GB memory
2 PB LustreFS storage
Software stack deployed using Docker.

Project ID: 754337

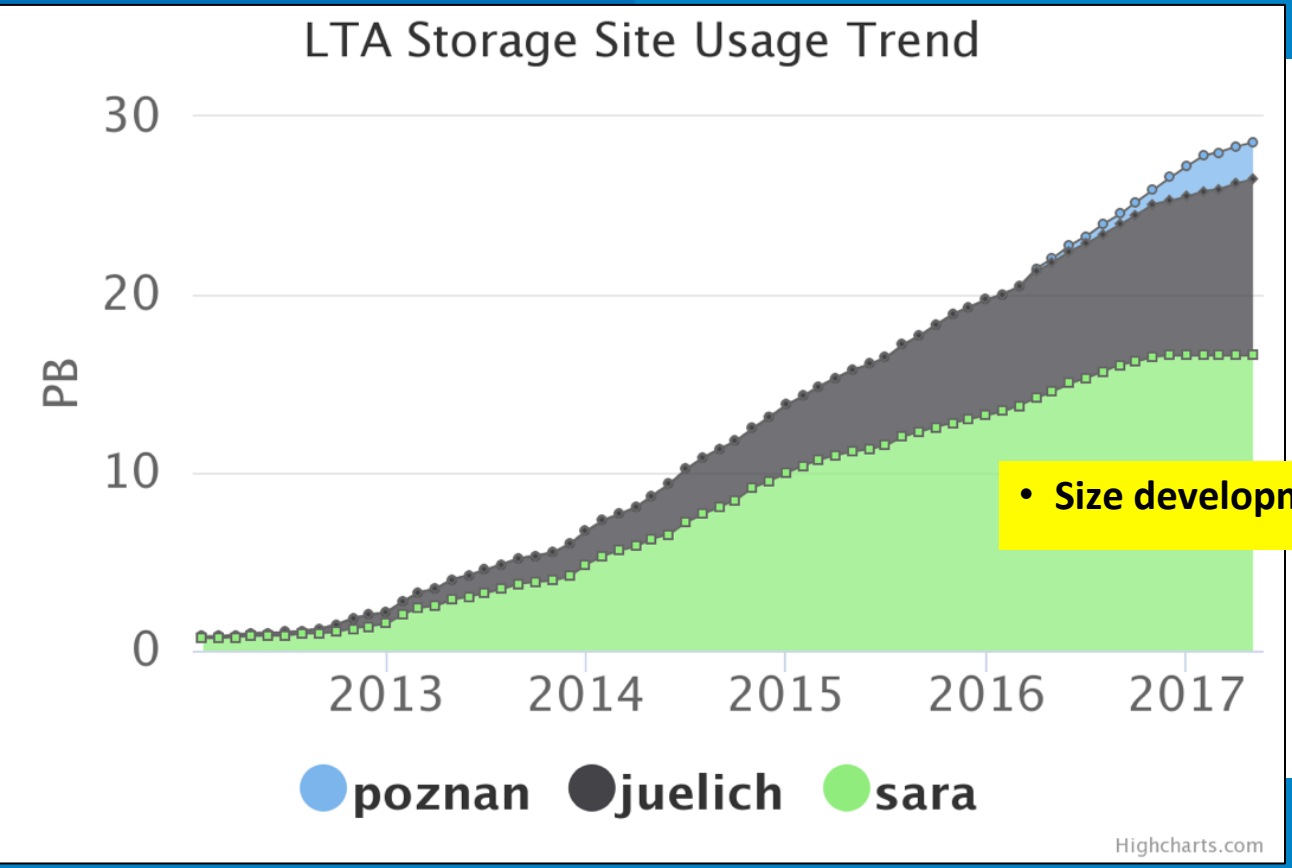
The LOFAR Long Term Archive

- **7 PB/yr**
- **10 Gbit/s connectivity between Tier 0 and each Tier 1 location**
- **27 PB**
- **~1.5 Bn files**



Project ID: 754337

LOFAR Long Term Archive

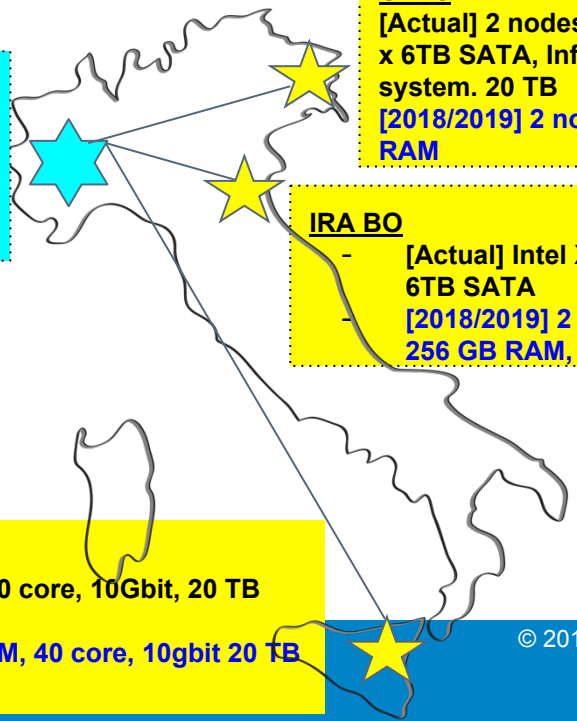


- **Size development in time**

LOFAR-IT e-infrastructure

Working Group WG-F03-01
UTGII “SKA Precursors and Pathfinders a bassa frequenza” (F03)

UNITO
3 FAT node on OCCAM
4 x Intel Xeon (12 core)
RAM 768 GB DDR4, 1 SSD 800GB, 1 HDD 2TB, 2x10Gb



OATs
[Actual] 2 nodes: Intel Xeon, 512 GB RAM DDR3 6 x 6TB SATA, Infiniband ConnectX®-3 - Hotcat system. 20 TB
[2018/2019] 2 nodes: 2 socket (40 core) 256 GB RAM

IRA BO
- **[Actual] Intel Xeon 512 GB RAM DDR3 6 x 6TB SATA**
- **[2018/2019] 2 nodes: 2 socket (40 core) 256 GB RAM, 10gbit ethernet**

OACT
- **[Actual] 2 nodes: 256GB RAM, 40 core, 10Gbit, 20 TB Storage**
- **[2018/2019] 2 nodes: 256 GB RAM, 40 core, 10gbit 20 TB storage Bee-GFS.**

Project ID: 754337

LOFAR WG and tests: people

Alessandro Costa

Gianmarco Maggio

Sara Bertocco

Luca Tornatore

Eva Sciacca

Fabio Vitello

Simone Riggi

Cristina Knapic

Francesco Bedosti

Annalisa Bonafede

Manuela Magliocchetti

Andrea Botteon

…and Giuliano Taffoni and Ugo Becciani

Project ID: 754337

LOFAR HW/SW Co-design

…learn from ILT experience, then….

Direction Independet Pipeline: does the first calibration of LOFAR data

> ***phase 1.a***: flag - calibration - transfer of solutions to the target and initial calibration of the target - averaging. Computing time: 3 - 4 days, RAM (core) at least 64GB.
>
> ***phase 1.b***: high and low resolution images and models for auto-calibration. Computing time: 4 -5 days, RAM (core) at least 64GB.

Direction Dependet Pipeline produces low-noise, high-resolution wide-field images from LOFAR HBA data. Computing time: 1 month

Project ID: 754337

LOFAR tested pipelines

- **Latest versions** of pipelines for the analysis of the **LOFAR HBA observations** in the continuous produced by the **Survey Key Project** and **ASTRON**.
- Pipelines available within **Singularity containers**

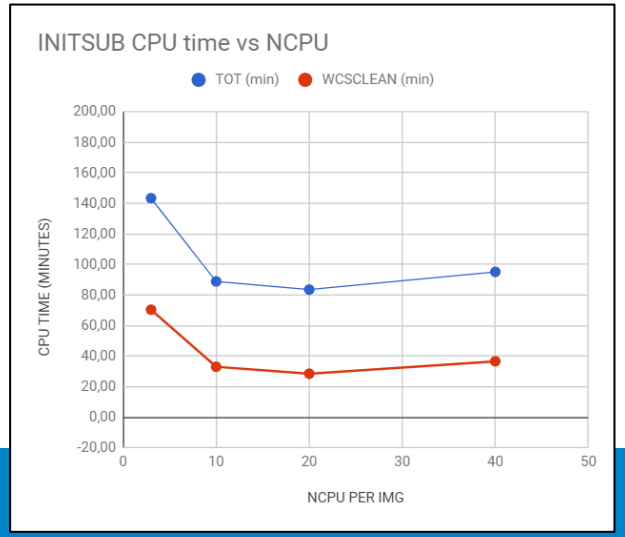
- Tests are a **collective work** of the data WG & friends performed at:
  - OACT: MUP node (24 core ht, 64 GB RAM) and LOFAR node (40 core, 256 GB RAM)
  - OATS: HOTCAT node (40 core, 256 GB RAM) with parallel File System

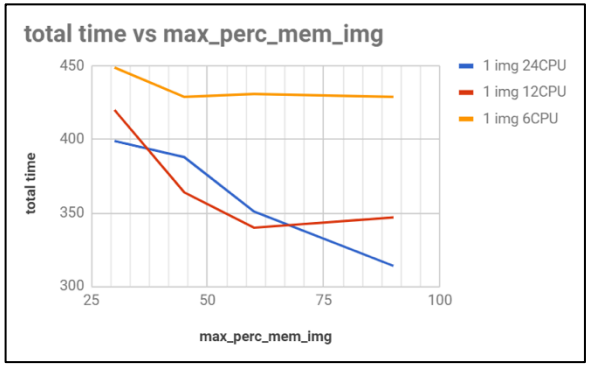
- LoTSS Data tested

Project ID: 754337

Pre-factor (phase 1.b) pipeline tests

- **INPUT Data**: 4 bands (2 MHz each)
- **Relevant input parameters**
  - **CPU Cores:** max_cpus_per_img, max_dppp_threads
  - **Memory:** max_percent_mem_per_img, max_imagers_per_node



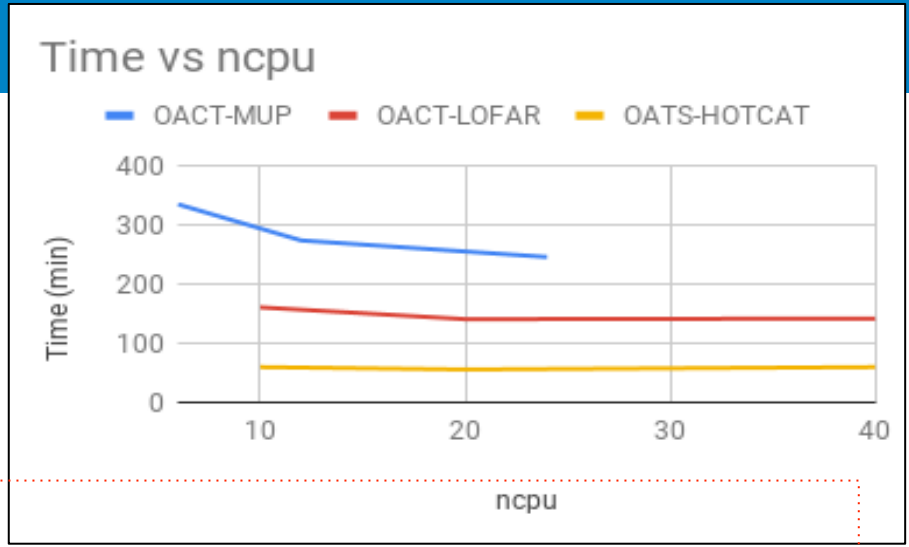


Conclusions:
the pipeline can run also on node with RAM < 256GB.
Recommend node with 3-4 GB per core and 6-10 cores per image.

Project ID: 754337

Factor pipeline tests

- **INPUT Data**: 6 bands (2 MHz)
- **Relevant input parameters:**
  - **CPU Cores:** ncpu
  - **Memory:** wsclean_fmem

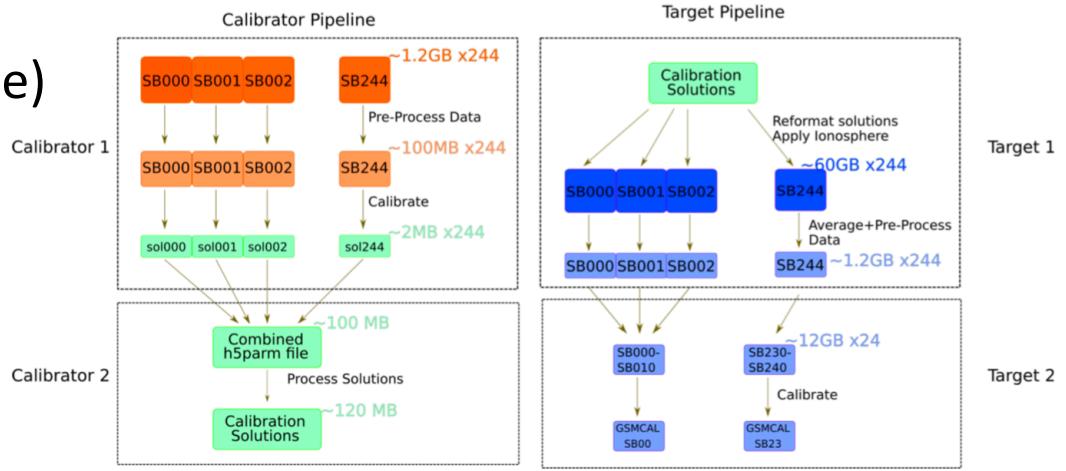


Conclusions: the pipeline does not scale with CPU cores/RAM memory but we have noticed it is very I/O time dependent therefore a parallel File Systems should be employed (e.g. BeeGFS as in OATS node). Recommend nodes with at least 10 cores and 256 GB RAM.

LOFAR HPC infrastructure

- Embarrassing parallel computation
- Multithreaded (on single node)
- Extremely IO demanding

Machev et al 2018



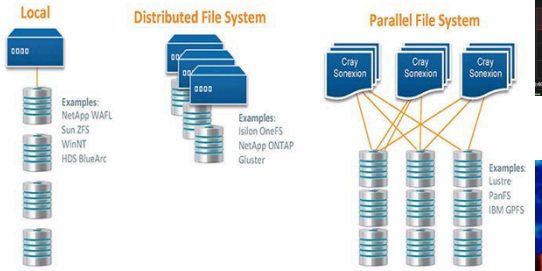
Project ID: 754337

How to build the infrastructure

- HTC cluster
- **Parallel filesystem** optimized for high throughput (~4GBs)
- Monitoring and Operations (telgraf+grafana)
- Containers (docker/singularity)

- **ObjectStorage** (new HPC approach to storage) testing

- Identify or train **“support” groups** for Italian Astronomers and for HW/SW contribution to ILT

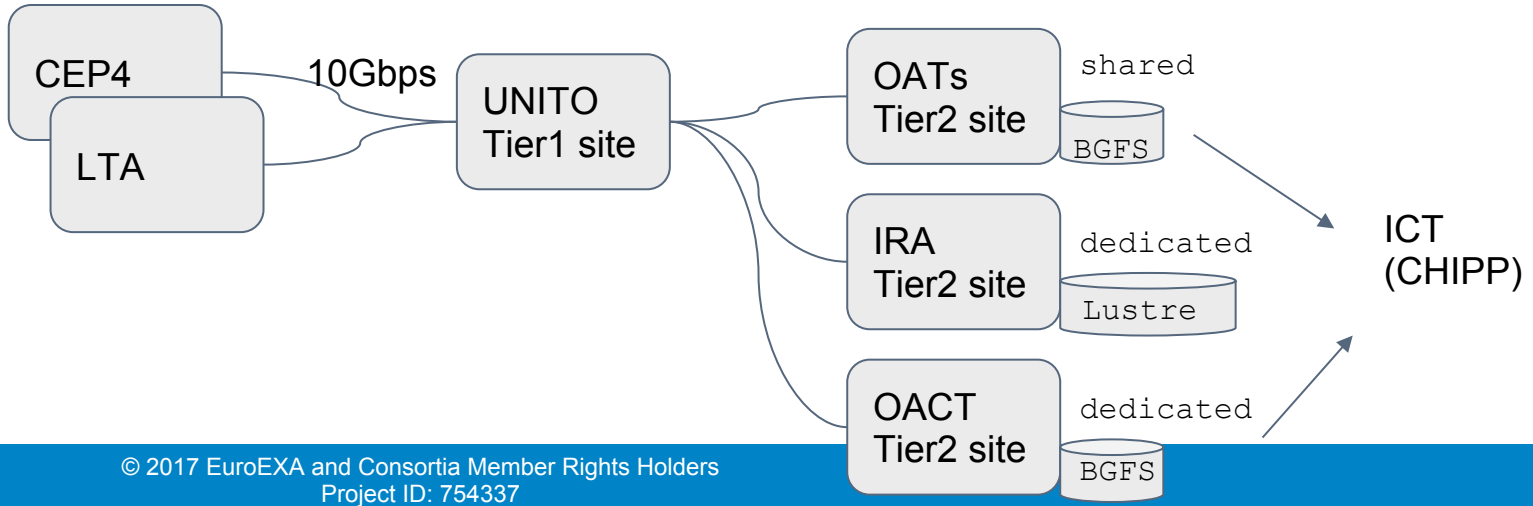


Project ID: 754337

LOFAR e-Infrastructure

It is not only HW but also **people**

We want to grow our knowledge to the LOFAR software infrastructure and to contribute to its **development and optimization.**



Project ID: 754337

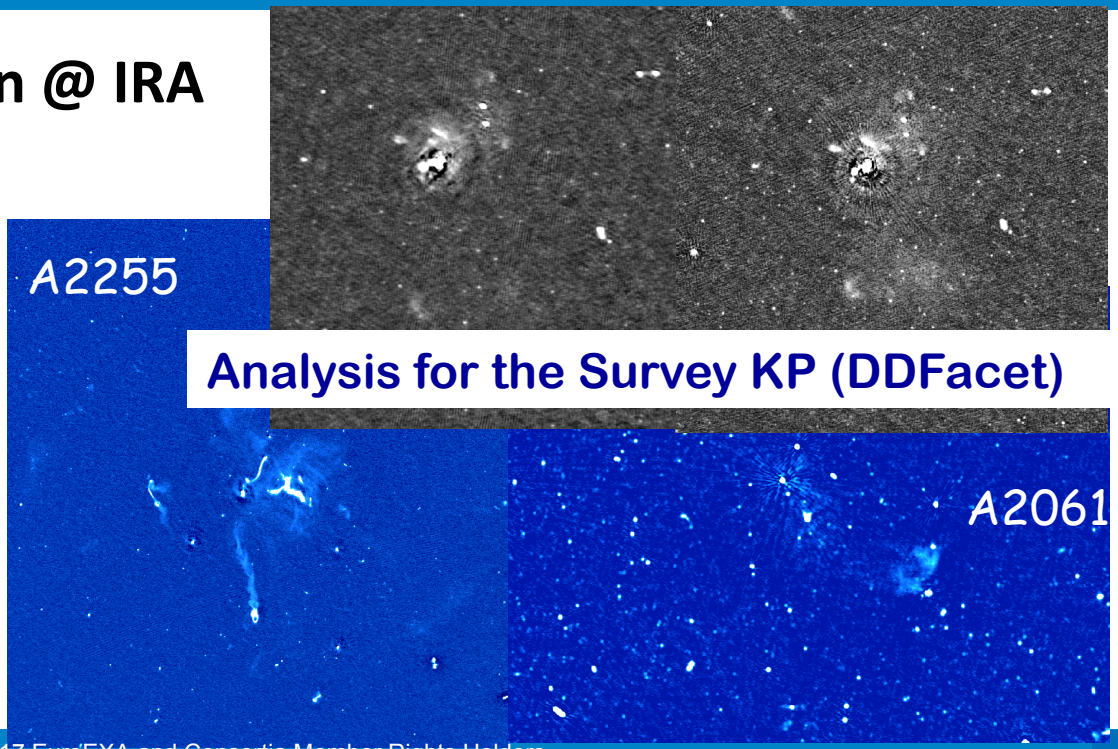
e-Infrastructure at work

A. Botteon @ IRA

FACTOR

killMS+DDFacet

Initially at IRA node now installed and operating also at the TS and CT nodes



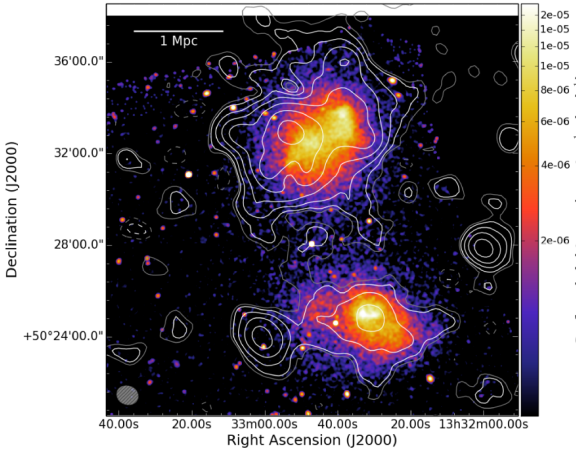
Project ID: 754337

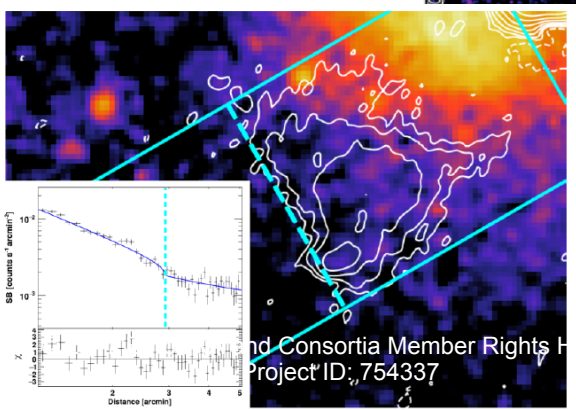
First Scientific Results

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS 478, 885–898 (2018)
Advance Access publication 2018 May 1
doi:10.1093/mnras/sty110

LOFAR discovery of a double radio halo system in Abell 1758 and radio/X-ray study of the cluster pair

A. Botteon,[1,2]★ T. W. Shimwell,[3,4] A. Bonafede,[1,2,5] D. Dallacasa,[1,2] G. Brunetti,[2] S. Mandal,[4] R. J. van Weeren,[4] M. Brüggen,[5] R. Cassano,[2] F. de Gasperin,[4] D. N. Hoang,[4] M. Hoeft,[6] H. J. A. Röttgering,[4] F. Savini,[5] G. J. White,[7,8] A. Wilber[5] and T. Venturi[2]

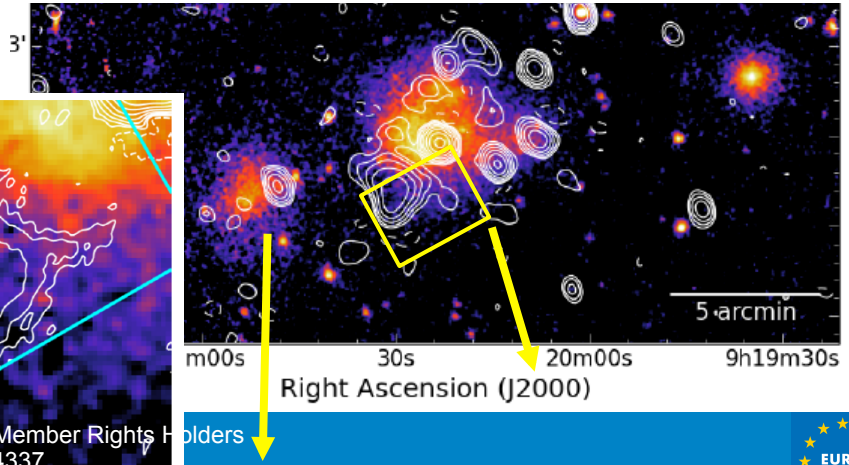




Astronomy & Astrophysics manuscript no. a781
November 21, 2018

The spectacular cluster chain Abell 781 as observed with LOFAR, GMRT, and *XMM-Newton*

A. Botteon[1,2], T. W. Shimwell[3,4], A. Bonafede[1,2,5], D. Dallacasa[1,2], F. Gastaldello[6], D. Eckert[7], G. Brunetti[2], T. Venturi[2], R. J. van Weeren[4], S. Mandal[4], M. Brüggen[5], R. Cassano[2], F. de Gasperin[4,5], A. Drabent[8], C. Dumba[8,9], H. T. Intema[4], D. N. Hoang[4], D. Rafferty[5], H. J. A. Röttgering[4], F. Savini[5], A. Shulevski[10], A. Stroe[11] and A. Wilber[5]



Towards ExaScale: 10^{18} Operation/s

- New platforms able to **ExaScale** (1000 times more the sustained performance of actual TIER-0 Facilities)
- 1 Billion Euro EU investment: EuroHPC and EPI
- **HW/SW Co-Design**

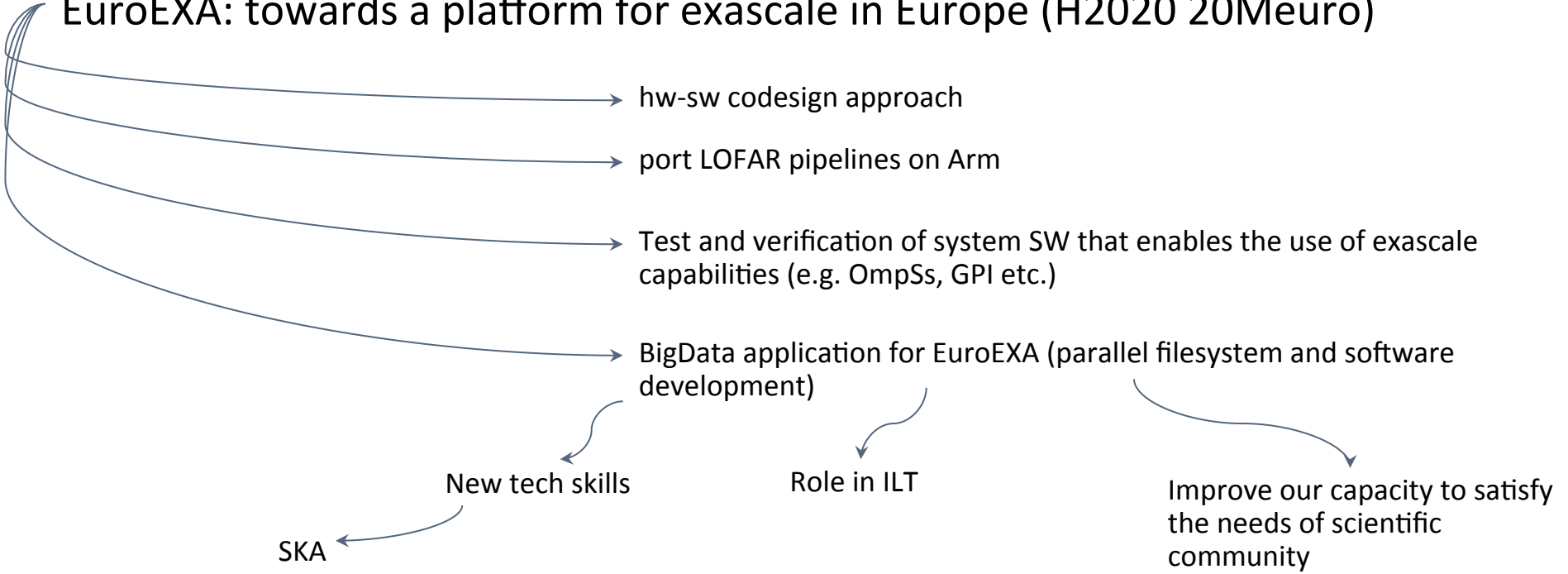
- Scientist must **change their algorithms** and approach to computing resources (new system software, massive use of accelerators, massive code optimization …)
- We must **change the way we program** our codes

- Already some tests and investments for SKA on UK
- INAF is the only A&A Institute in EU Involved in Exascale prototyping
    - LOFAR data reduction and analysis

Project ID: 754337

Towards the future

EuroEXA: towards a platform for exascale in Europe (H2020 20Meuro)



Project ID: 754337

Thanks for your attention

Alessandro Costa

Gianmarco Maggio

Sara Bertocco

Luca Tornatore

Eva Sciacca

Fabio Vitello

Simone Riggi

Cristina Knapic

Francesco Bedosti

Annalisa Bonafede

Manuela Magliocchetti

Andrea Botteon

…and Giuliano Taffoni and Ugo Becciani

Project ID: 754337