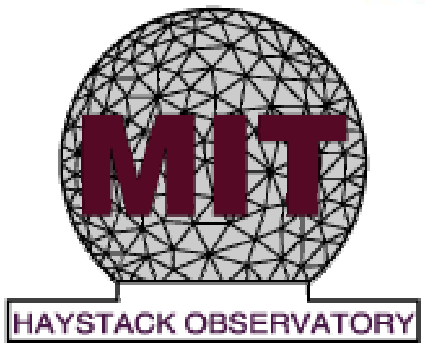
Networks and Processing for LOFAR

Marco de Vos (devos@astron.nl)

2nd e-VLBI workshop, Dwingeloo

The LOFAR Project

- International collaboration between
  - ASTRON
  - MIT Haystack Observatory
  - Naval Research Lab



- Membership applications:
  - LOIS consortium (Sweden)
  - LOPES project (Germany)

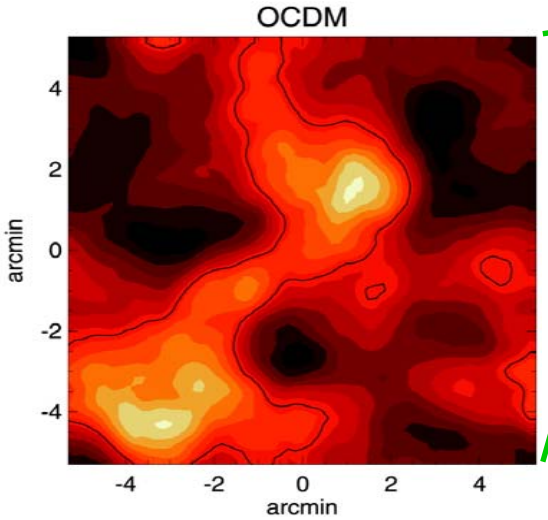
LOFAR

LOFAR Key features

- Major new array for 10-240 MHz range
- 400 km across, fixed dipole receptors
- Breakthrough in sensitivity
- Extreme agility in time/space/frequency
- Large instantaneous sky coverage
- Essentially digital

LOFAR

Scientific Versatility



Tozzi et al. (2000)

Redshifted HI from the Epoch of Reionization

High-z starbursts

Galaxy clusters and the IGM

Cosmic ray distribution, and airshower radio bursts

Steep spectrum and fossil radio galaxies

Supernova remnants and ISM energy budget

Interstellar recombination lines

Nearby pulsars, ghost nebulae

Extrasolar gas giant planetary radio emission

Stellar flares

Interstellar medium propagation effects

Transients, GRB and LIGO event counterparts, buffering

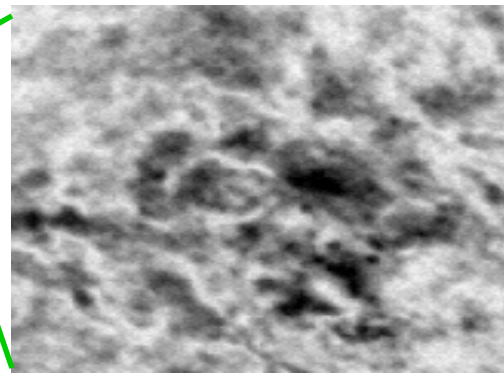
Solar radio studies

CME detection, mapping by IPS, scattering

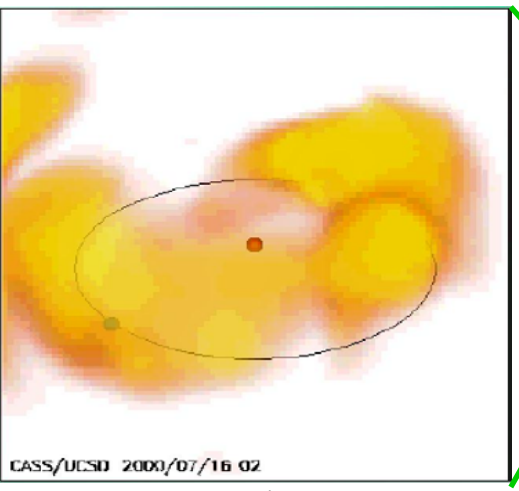
Extremely high resolution ionospheric tomography

Magnetospheric radar

Lane et al. (2001)

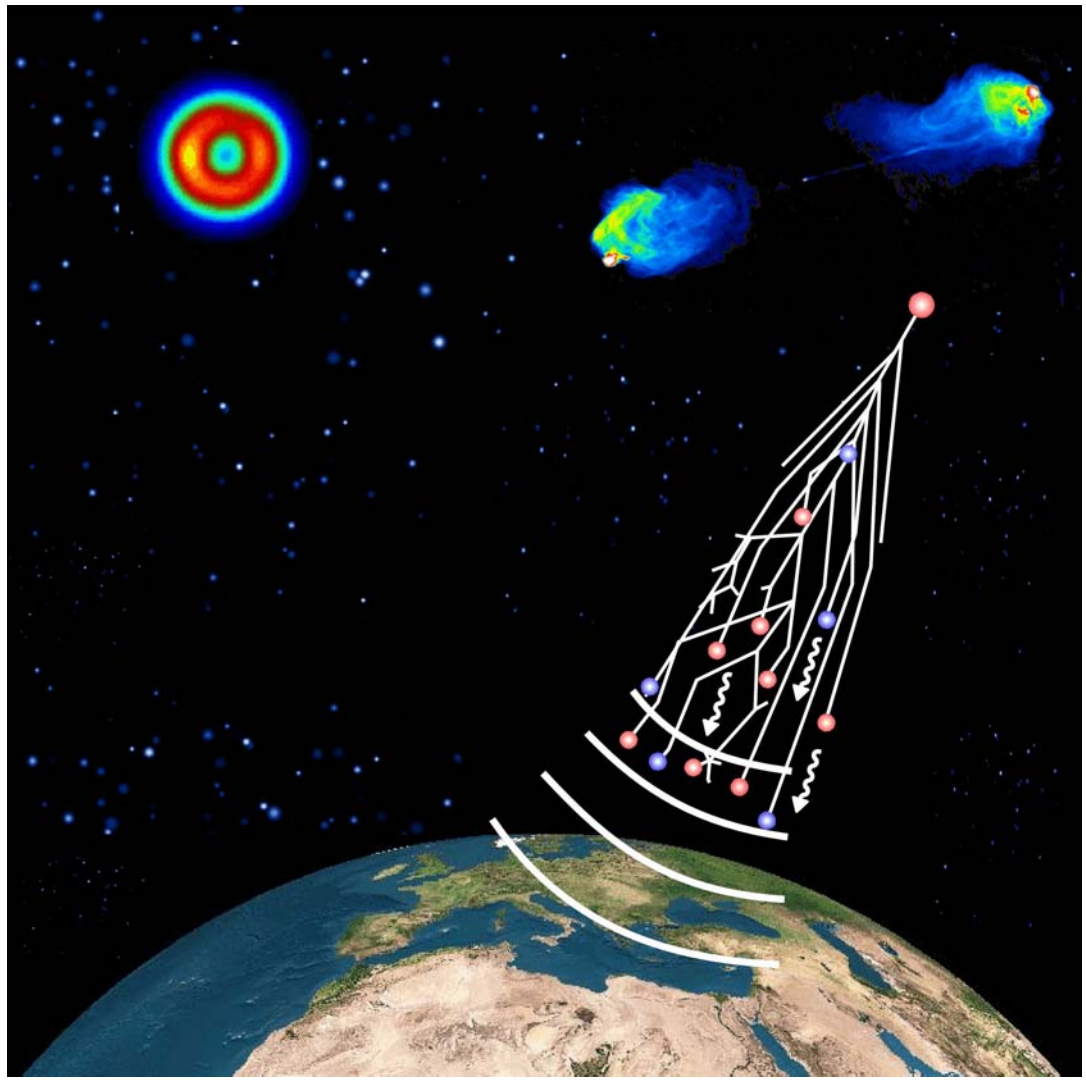


Courtesy: B. Gaensler



Courtesy: B. Jackson

Cosmic Air Showers



LOFAR Technical Concept

- LOFAR is a large distributed radiotelescope:
  - >100 phased array stations
  - Combined in aperture synthesis array
  - 13,000 small “LF” antennas
  - 13,000 small “HF” tiles
  - >20 Tbit/s datanetwork (if full scientific bandwidth)
  - >40 Tflop supercomputer
  - innovative software systems
  - new calibration approaches
  - full distributed control, VO and Grid integration
  - datamining and visualisation

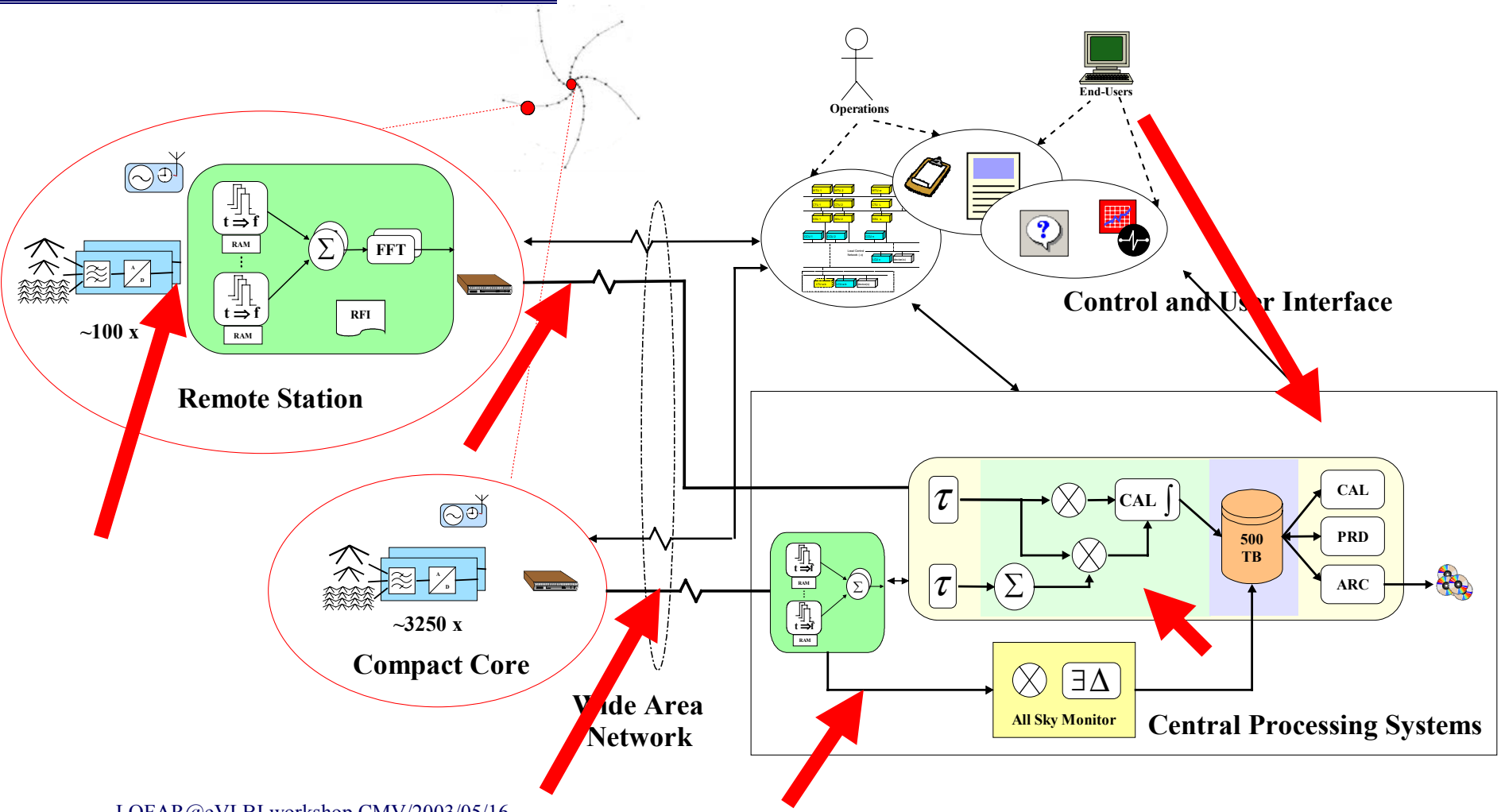
LOFAR

Key technologies

- LOFAR is realized as an aperture synthesis array using multi-beaming phased array techniques.
- LOFAR
  - uses **low-cost antennas** and
  - relies on **broad-band datalinks** and
  - on **high performance computers**
  - to implement the majority of its functionality in **(embedded) software**.

LOFAR

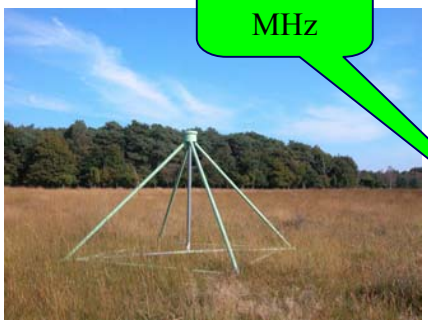
A distributed embedded system…

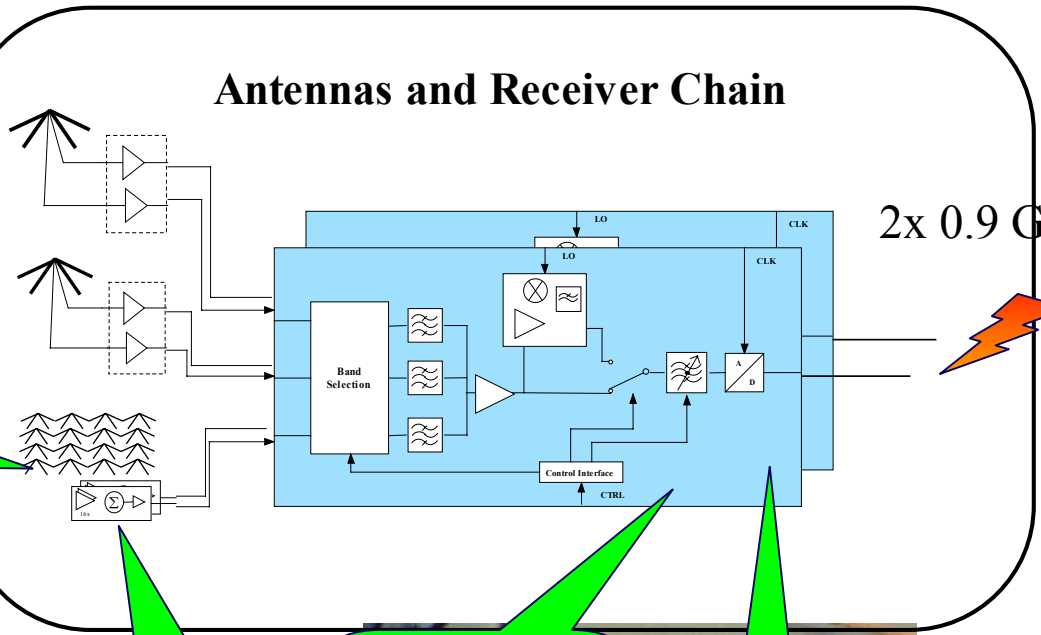


1.8 Gpbs per dual polarization antenna
220 Gbps per remote station

Antennas and receiver electronics

10..90 MHz





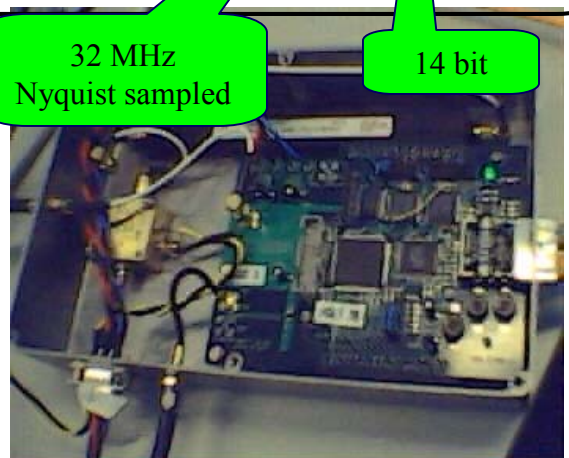
2x 0.9 Gbps

110..220 MHz (4x4 active dipoles)

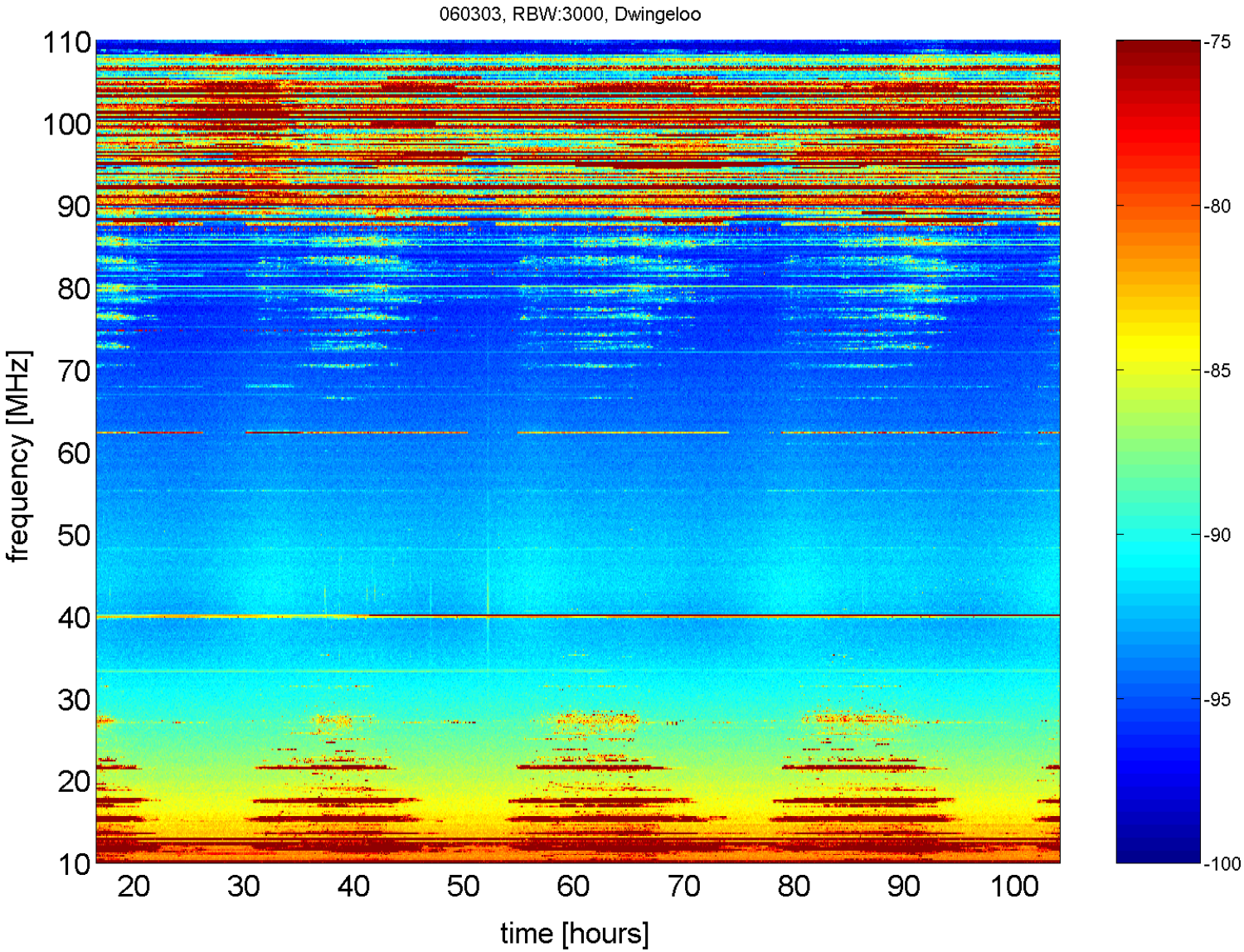
RF Beamforming

32 MHz Nyquist sampled

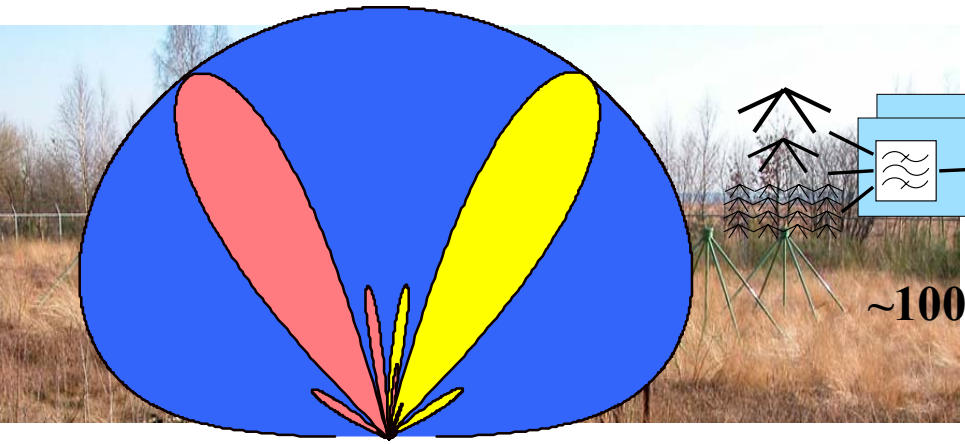
14 bit

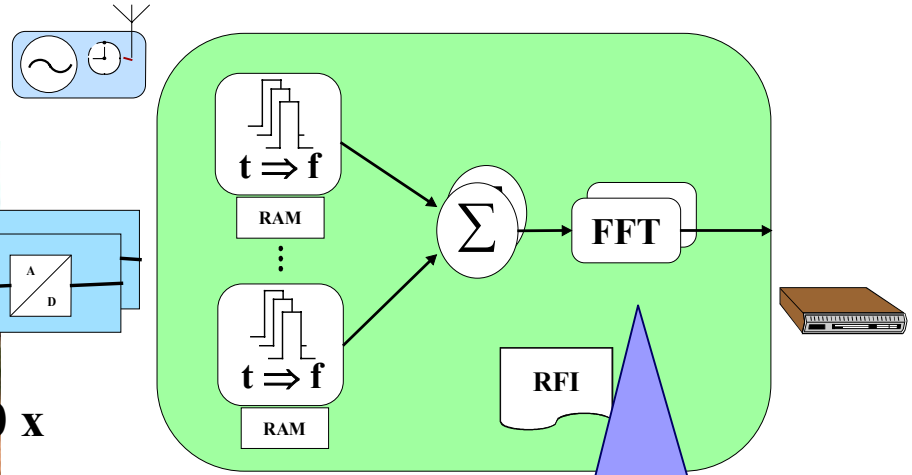


LOFAR


060303, RBW:3000, Dwingeloo
frequency [MHz]
110
100
90
80
70
60
50
40
30
20
10
20
30
40
50
60
70
80
90
100
time [hours]
-75
-80
-85
-90
-95
-100

LOFAR Remote Stations

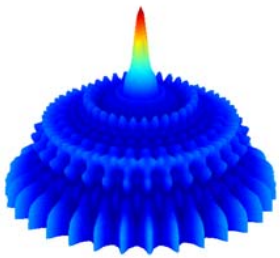


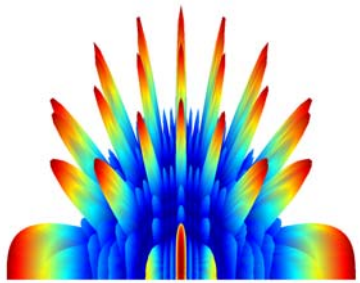


~100 x

Remote Station

Decrease datarate by
(spectral & spatial) selection and
compression (after RFI mitigation)

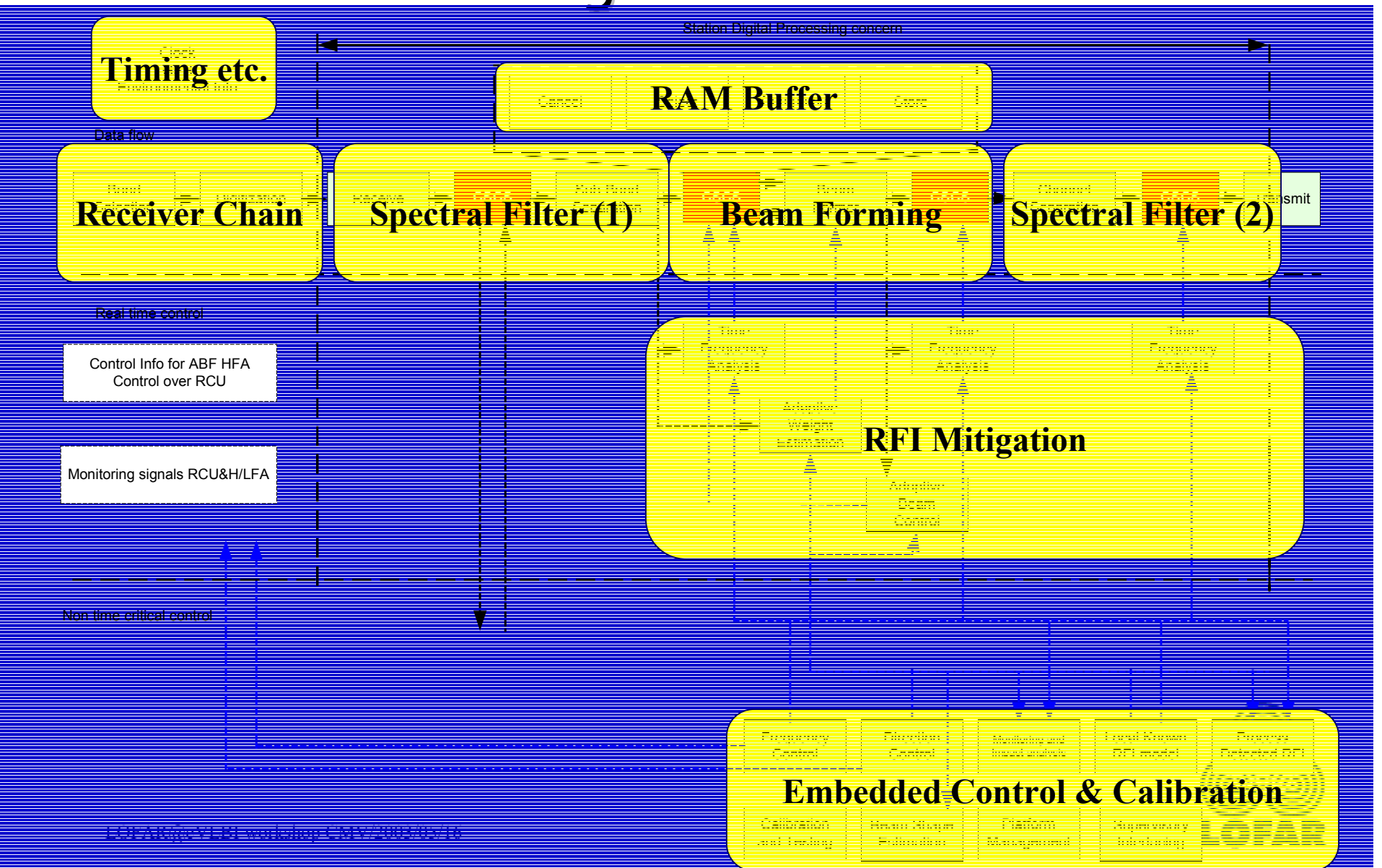




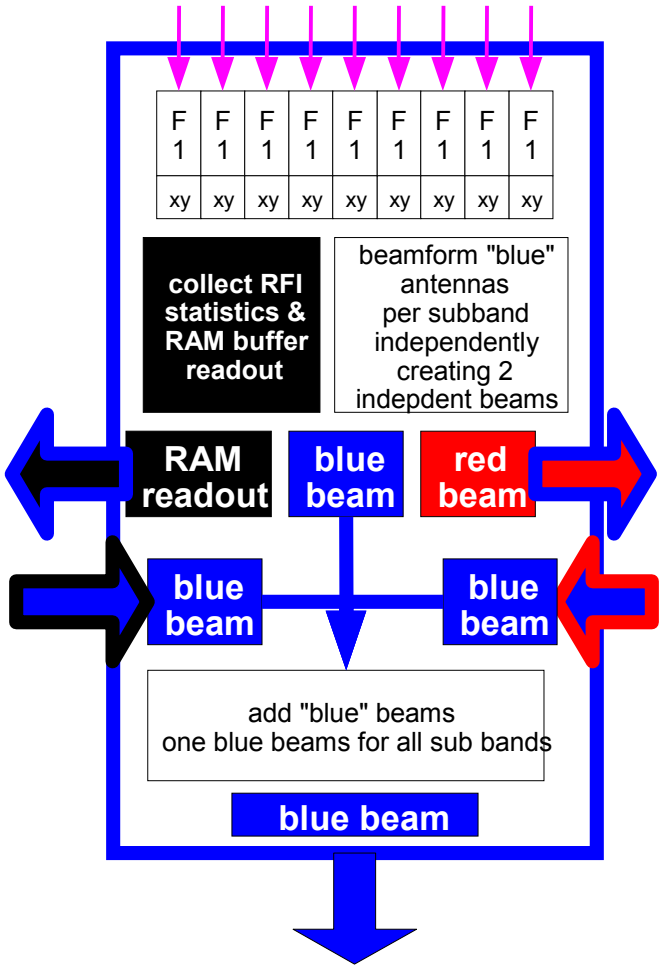
20 … 2 Gbps per Remote Station
(in order not to shock you, could be much more…)

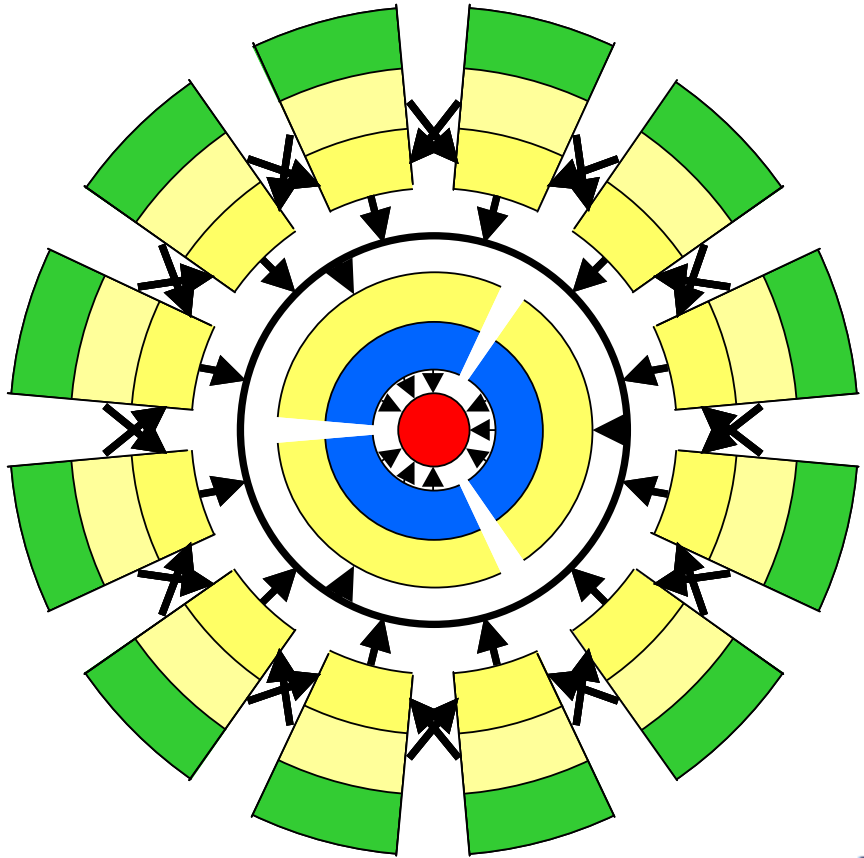
LOFAR

Station Processing Functions



Station Processor





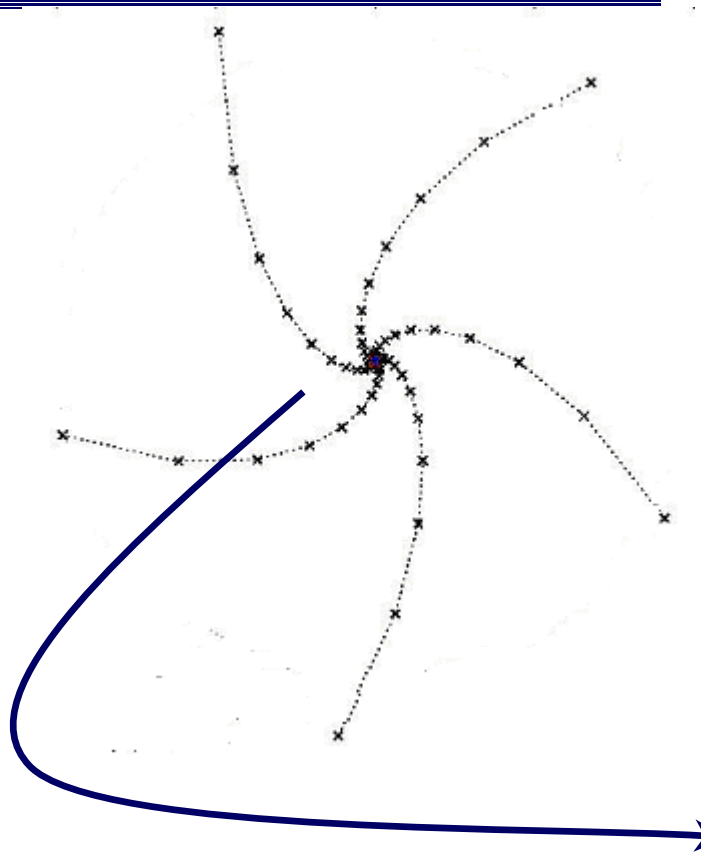
LOFAR

Transient buffering

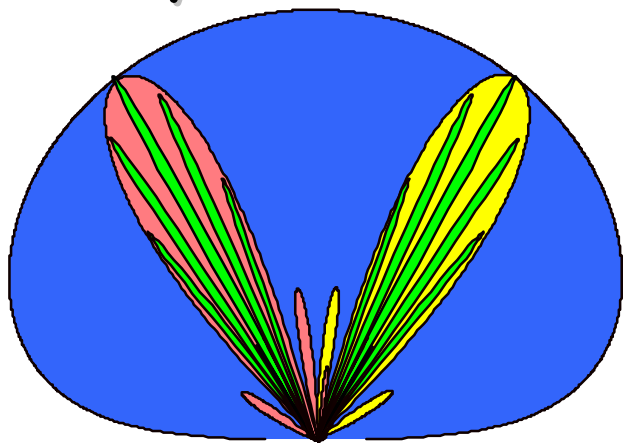
- Large cyclical RAM buffer per antenna
  - Buffer contains full FOV information
  - Upon trigger, can point LOFAR post-facto
  - Tradeoff between duration, sensitivity
- Implement subsystem, add more RAM later
  - Current cost estimate for full array:
    - 3×1012 bits/sec for 4 MHz, 2 pols, 14 bits/sample
    - 100 seconds of buffer $\Rightarrow$ \$600k in RAM chips, 2005 prices

LOFAR

Combining stations into an array



100 stations @ 2 Gbps



Visibility word size	2	Byte
Visibility word size integrated	4	Byte
integration time	10	sec
integration freq resolution	4	kHz
Observation duration	4	hours
Number of SS Visibilities per RS Beam	4950	
Correlator output data rate	38.7	Tbps
Correlator processing power	16.9	Tflops
Integrator output data rate	2	Gbps
Integrator processing power	4.8	Tflops
Storage capacity (raw data)	27.8	TB
Selfcal processing power	--	

LOFAR

Data Transport Network

- Long-range, modest bandwidths (2 .. 20 Gbit/s/station)
- Short-range, high bandwidths (multi Tbit/s)
  - Both: dedicated network (at least in virtual sense)
- Proposed architecture based on 10Gbps
  - Follows industrial/commercial trends
  - Low-cost, short lead times, high market penetration

- A network is more than technology!
  - Even dedicated p2p networks like the LOFAR WAN
  - Optical performance, reliability, maintenance, …
  - Integration for dual use, …

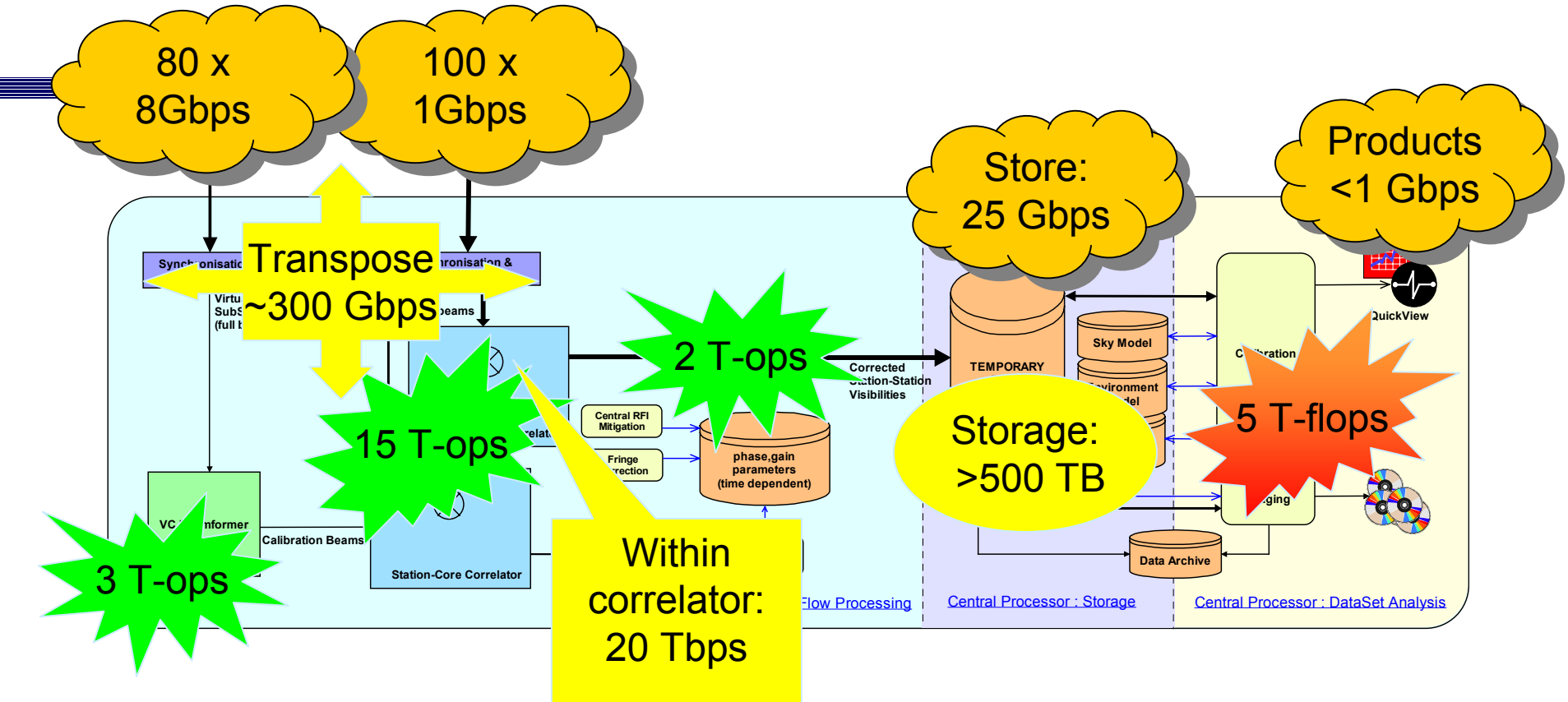
LOFAR

Central Processing Facility

- Handles:
  - Correlation
  - Beamforming (virtual core and tied array)
  - Calibration and associated tasks
  - All Sky Monitor imaging
  - Other heavy-duty processing (e.g. pulsars)
- Massive processing power
  - Tens to hundreds of Tera-ops (details still under study)
  - Fortunately most problems are “embarrassingly parallel”
  - >500 Tbytes of storage
- Implemented via PC clusters
  - Coprocessor boards provide the muscle
  - Special backbone provides the veins

LOFAR

CEP Performance



- ~1 Tbps I/O
- 25 T-ops processing power

LOFAR

Correlation

- This is “X” part of FX architecture
  - “F” part was done at the stations
- ~5000 baselines for each spectral channel
- Crude calibration applied
  - Data averaging to 1 sec, 10 kHz
  - Output data rate of order 1Gbyte/sec
  - Data storage for further calibration
- Successive stages of calibration, integration
  - Final product - smaller data sets, or images

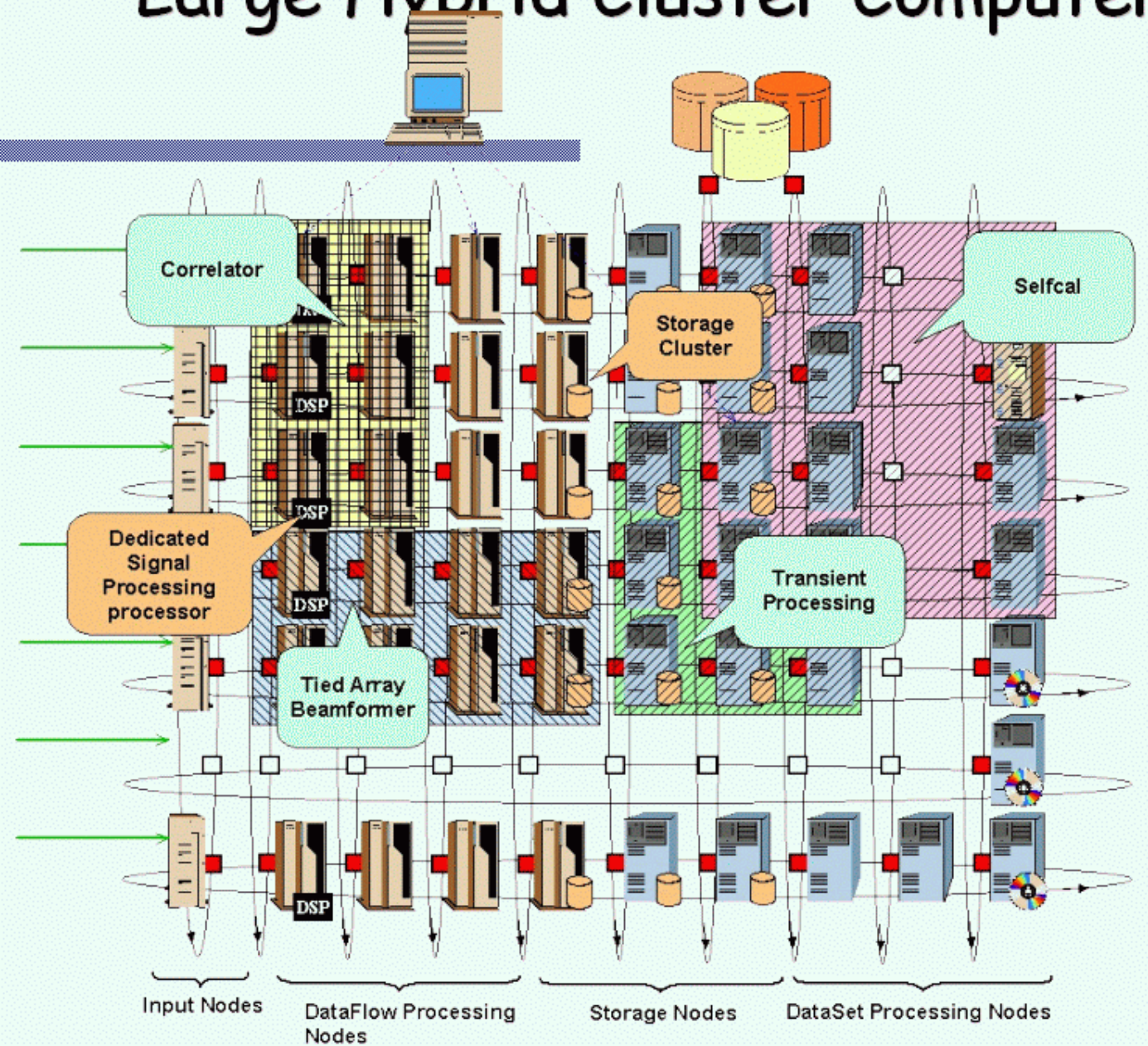
LOFAR

All Sky Monitor

- Image up to 4 steradians every second
  - 2km “virtual core” only, few arcmin resolution
  - Correlate, calibrate and image on the fly
  - Examine images for transient sources
  - Time average images to cover many timescales
- Large computational burden – options:
  - Reduce bandwidth to 256 kHz initially?
  - Cluster and beamform VC antennas?
  - Special-purpose number-cruncher?
- Automatic event detection/notification
  - Requires sophisticated algorithms (e.g. reject RFI)

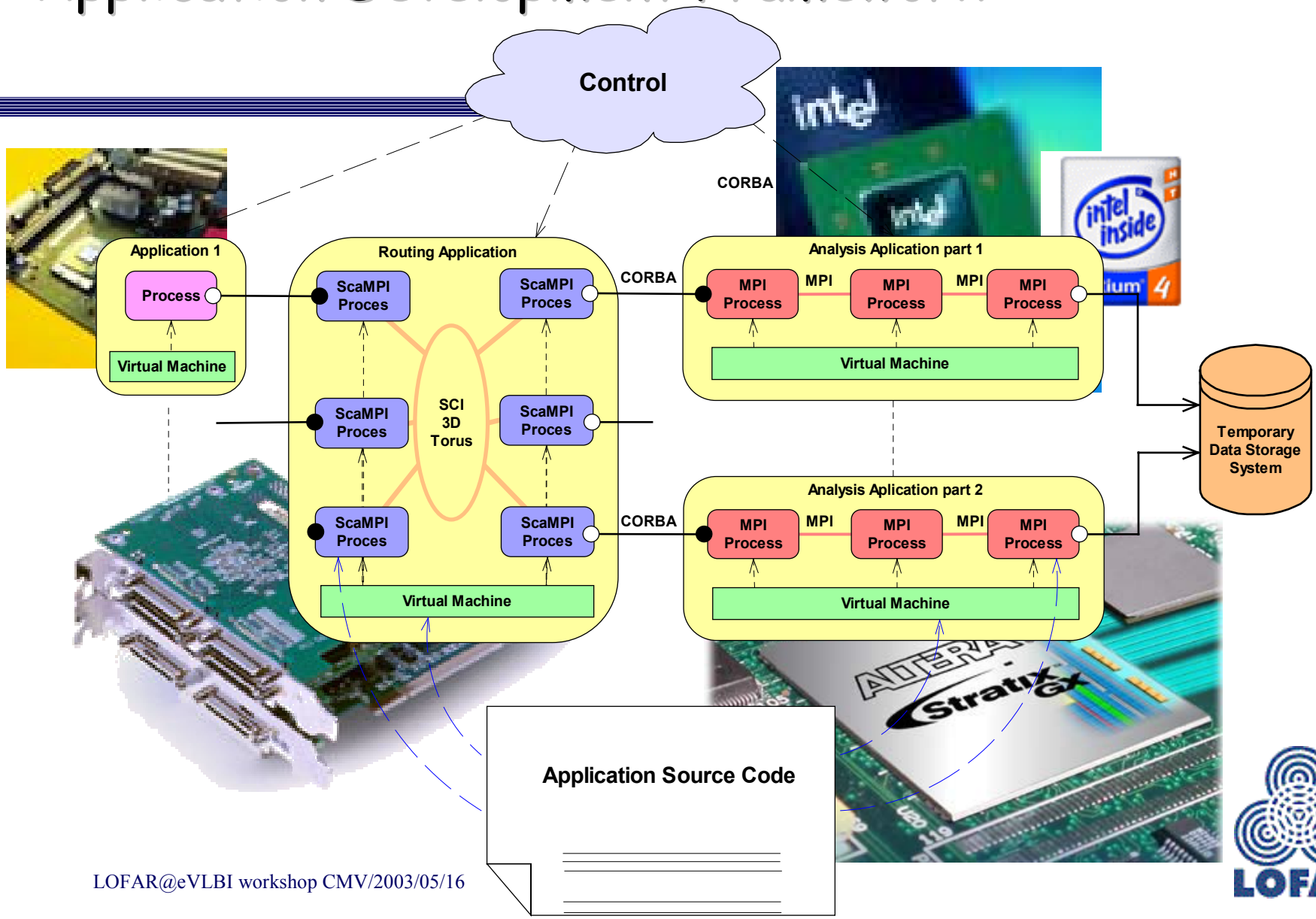
LOFAR

Large Hybrid Cluster Computer



LOFAR

Application Development Framework



Where does LOFAR end?

- Observation Modes
  - Implemented through “Virtual Instruments”
    - Synthesis Imaging including SelfCal processing
    - Phase Array Operation (no detailed processing)
    - Transient Processing (insufficiently defined yet!)
    - Hooks and stubs for many more
- Dataproducts
  - Regular and special purpose
    - Global Sky Model (develops in time)
    - Event Lists
    - Sky Models + Residual Maps for synthesis mode
    - Hooks and stubs for anything else

LOFAR

Remote Operations

- LOFAR is an agile, multibeaming instrument
  - Can do many things at once
  - Efficiency demands multiple simultaneous users
  - Resource sharing via internet is integral to LOFAR concept
- General principle for LOFAR operations
  - Bandwidth > internet capacity, CEP task
  - Bandwidth < internet capacity, Science center task
- Limit: functional independence of LOFAR beams
  - Design and cost issue

LOFAR

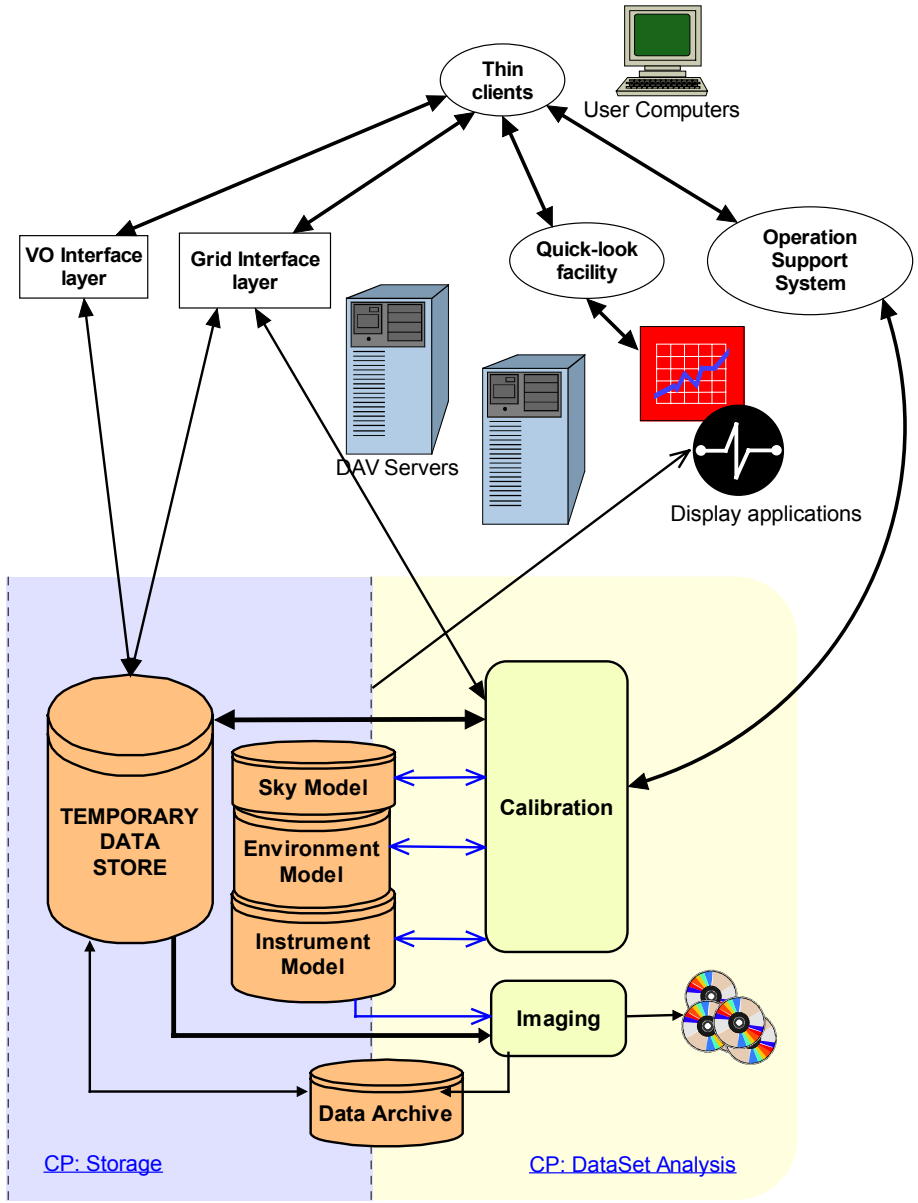
User Data Products

- Data volumes too large for ~2010 internet
  - Lots of online/nearline processing/analysis
    - VO/Grid model: *Transport information, not data*
  - Sufficient processing to reduce data to manageable volume, suitable for transport for routine operations
- Imaging applications
  - Widefield (e.g. survey) imaging – process on-site
  - Narrow field (targeted) imaging – export of visibilities possible
- Non-imaging applications
  - Needs vary by application
    - Voltage time-series might be exported in some cases
    - Some applications (e.g. pulsars) need CEP power

LOFAR

End-users

- Transparent access to data and processing resources



LOFAR Science Centres

- Responsible for the scientific operation of LOFAR and for the support of the broader scientific community.
- Teams consisting of expert LOFAR-users
  - qualified to assessing the instruments behavior and data products
  - committed to obtaining the ultimate quality in calibration and scientific integrity.

LOFAR Development Plan

- Major milestones
  - 2003Q2: 10 antenna array
  - 2003Q3: Initial Test Station
    - 100 antenna array + recorder
  - 2004Q1: Full Test Station
    - 100 antenna array + digital processing + MAC
  - 2004Q2: HFA integration
  - 2004Q4: Start procurement Virtual Core
    - 25% of antenna's, 25% of Central Processor
  - 2006: Initial Operating Conditions
  - 2008: Ultimate Operating Conditions

LOFAR

Software Development Approach

- Time-boxing in 6 months periods
- Large number of well-defined functional modules
  - high-cohesion, low coupling