Enabling Scientific Breakthroughs

Rob van Nieuwpoort

R.vanNieuwpoort@esciencecenter.nl

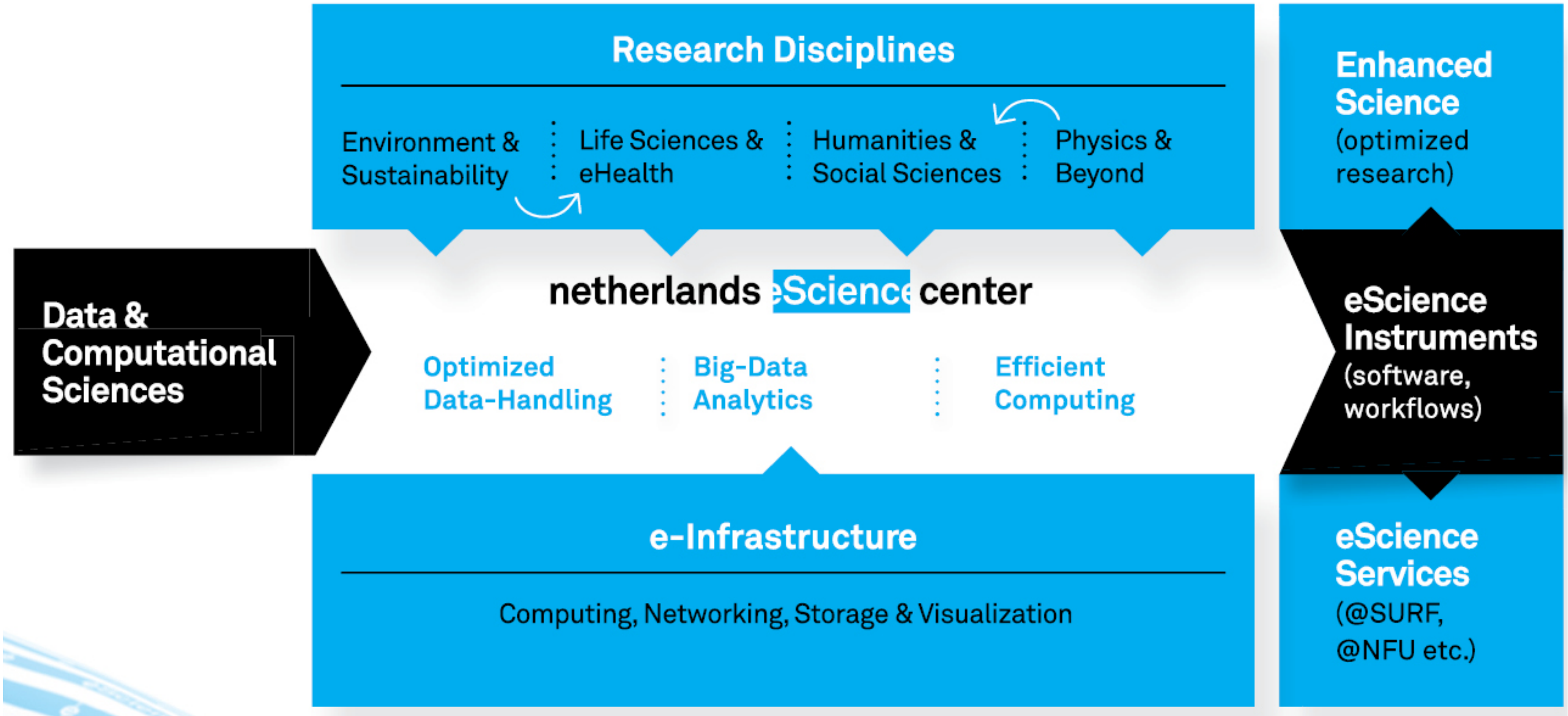
netherlands eScience center by SURF & NWO

ASTRON

VU UNIVERSITY AMSTERDAM

Faculty of Sciences

We work demand-driven



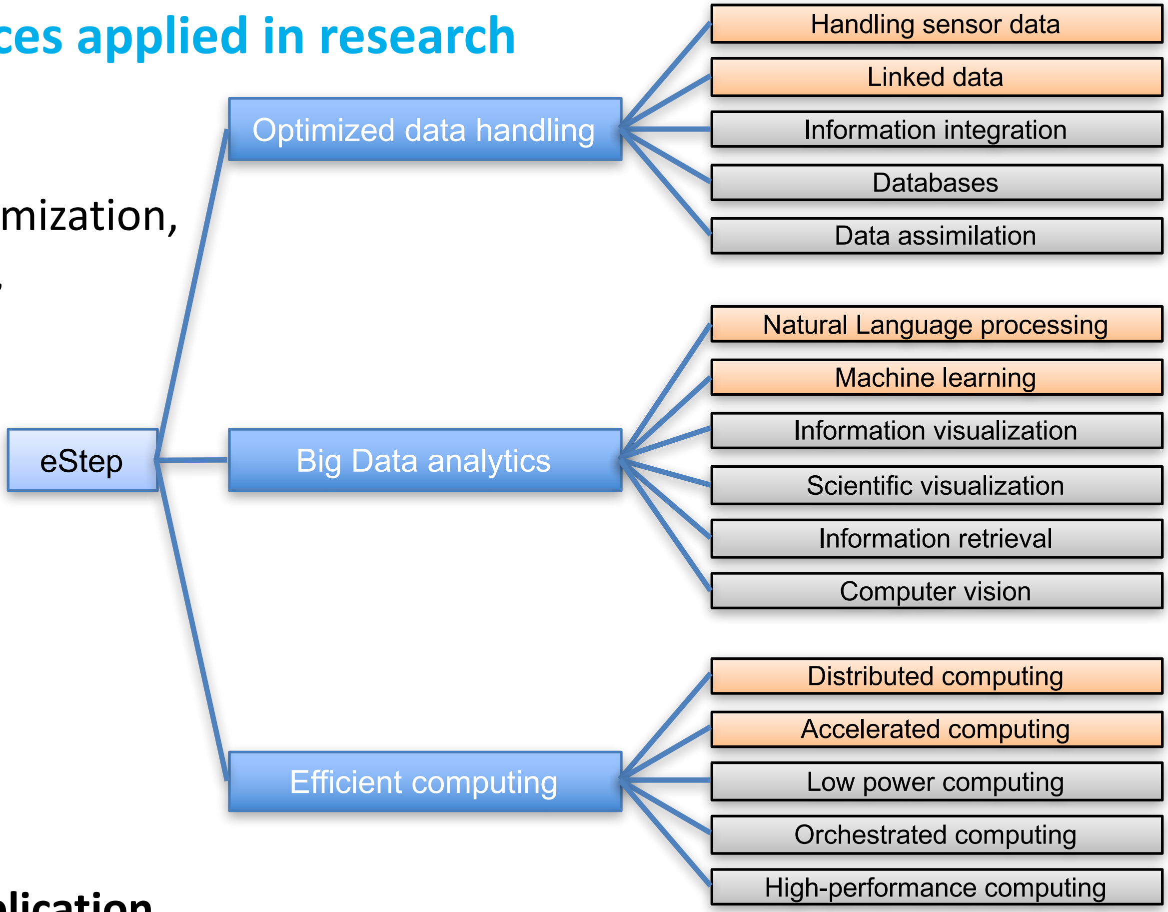
NLeSC eScience competences applied in research

1. Optimized data handling
Data integration, data base optimization, structured & unstructured data, real time data

2. Big data analytics
Statistics, machine learning, visualization, text mining

3. Efficient computing
Distributed & accelerated computing, efficient algorithms

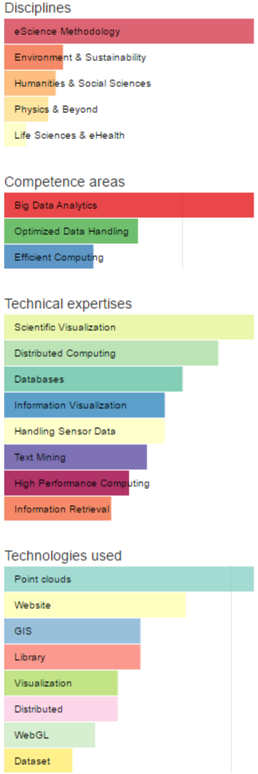
Prevent fragmentation and duplication



Software

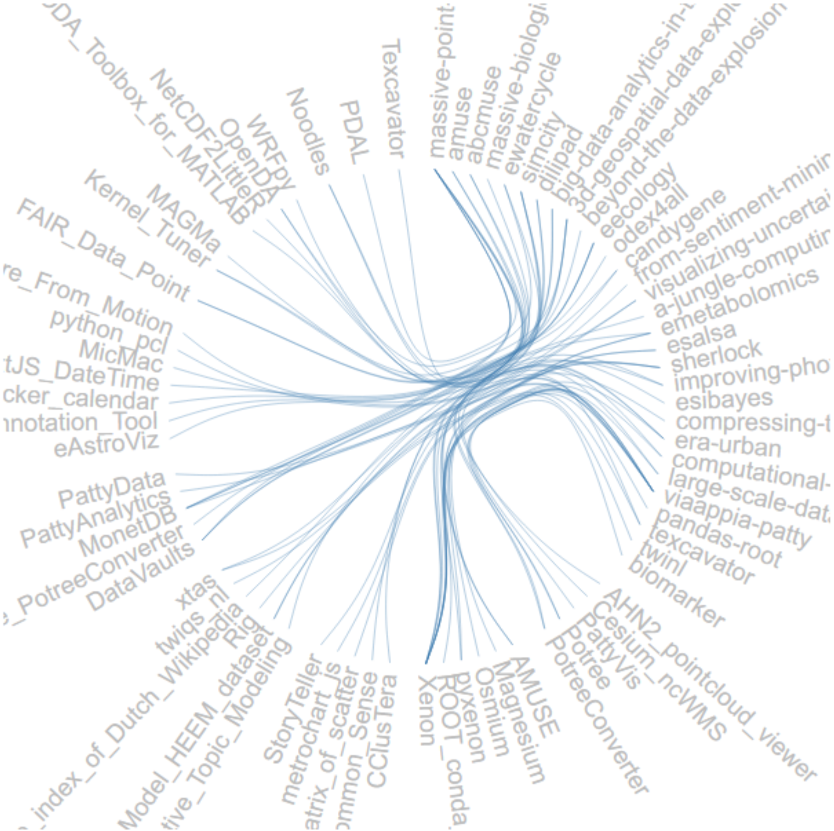
Click on the bars to find software projects.

49 selected out of 49 records
Reset All



Name	Description
AHN2 pointcloud viewer	WebGL point cloud visualization of AHN2
AMUSE	The Astrophysical Multipurpose Simulation Environment
CClusTera	A 3D web tool for interactive visualization of hierarchically clustered big data
Cesium-ncWMS	3D Globe Visualization of NetCDF data.
Common Sense	User-friendly web application for showing (GIS) data on a map.
Cross-perspective Topic Modeling	A Gibbs sampler that implements Cross-Perspective Topic Modeling
DataVaults	Technology of Attachment to a DBMS of large file repositories.
Differential Evolution	Differential Evolution global optimization algorithm, with Metropolis for uncertainty estimation
eAstroViz	This tool can convert and visualize radio astronomy measurement sets, as well as most LOFAR intermediate data producs. It also does RFI mitigation.
eEcology Annotation Tool	Visualize & annotate GPS measurements of bird movements
eEcology Tracker calendar	Calendar overview with daily statistics of GPS-tracker
eWaterLeaf	Web-based visualization for the eWaterCycle project
ExtJS-DateTime	DateTime form input field for ExtJS
FAIR Data Point	FAIR Data Point Metadata Service
GoogleEarth Toolbox for MATLAB	Export data from MATLAB to GoogleEarth's KML format.
Historic Embodied Emotions Model (HEEM) dataset	279 17th and 18th century Dutch theater texts with HEEM labels
Kernel Tuner	A simple CUDA/OpenCL kernel tuner in Python.

eStep Software used in Projects



Big Data & Big Compute in Radio Astronomy

Rob van Nieuwpoort

director of technology

netherlands eScience center by SURF & NWO

Universiteit van Amsterdam

ASTRON

VU University Amsterdam Faculty of Sciences

Two simultaneous disruptive technologies

- Radio Telescopes
  - New sensor types
  - Distributed sensor networks
  - Scale increase
  - Software telescopes

- Computer architecture
  - Hitting the memory wall
  - Accelerators

Two simultaneous disruptive technologies

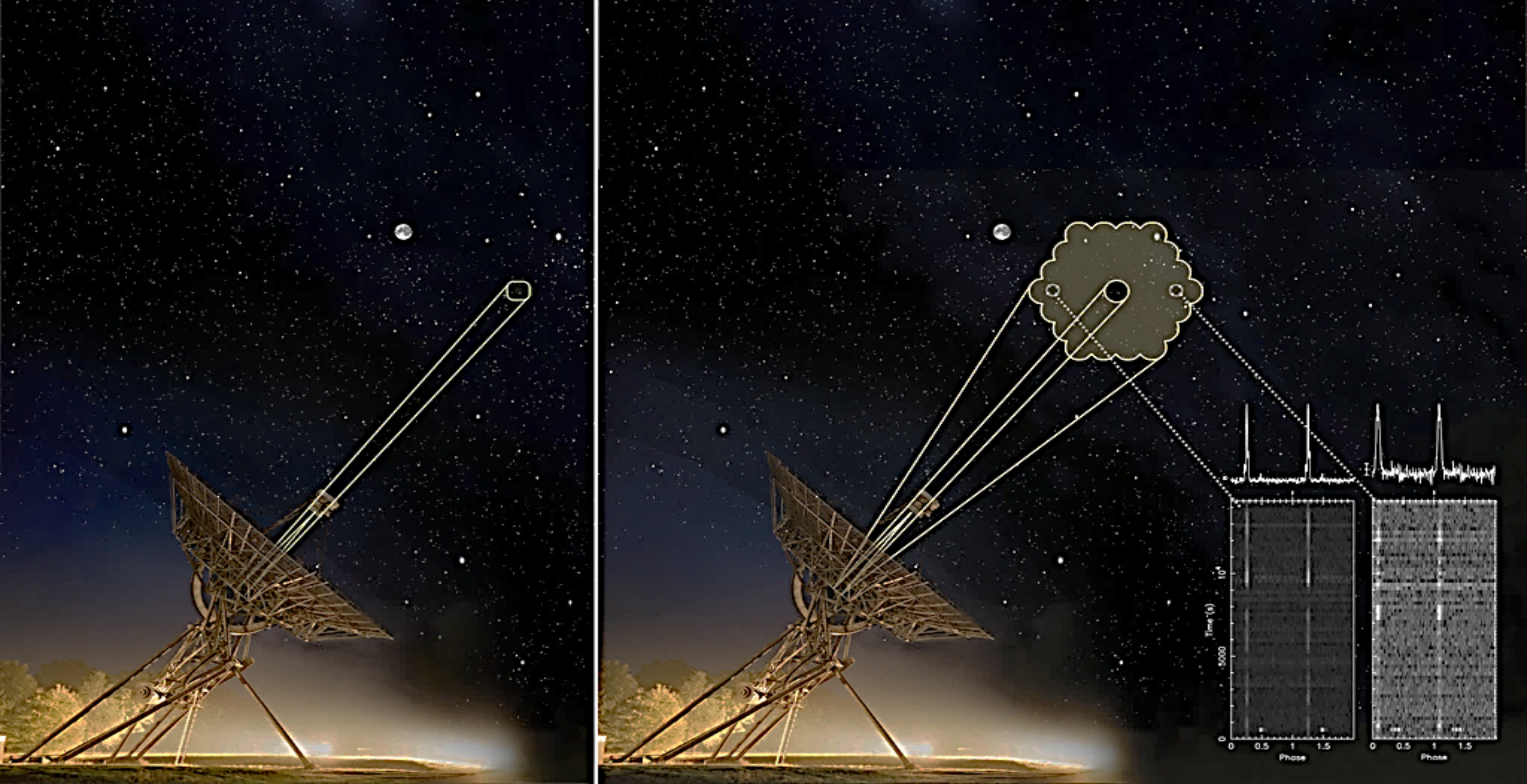
- Radio telescopes
  - New sensor types
  - Distributed sensor networks
  - Scale increase
  - Software telescopes
- Computer architecture
  - Hitting the memory wall
  - Accelerators

panic!

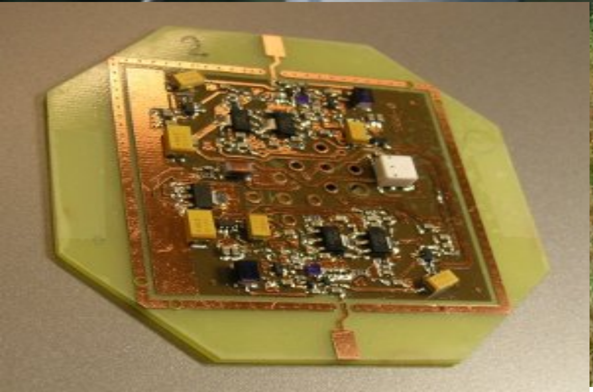
ASTRON

Image courtesy Joeri van Leeuwen, ASTRON

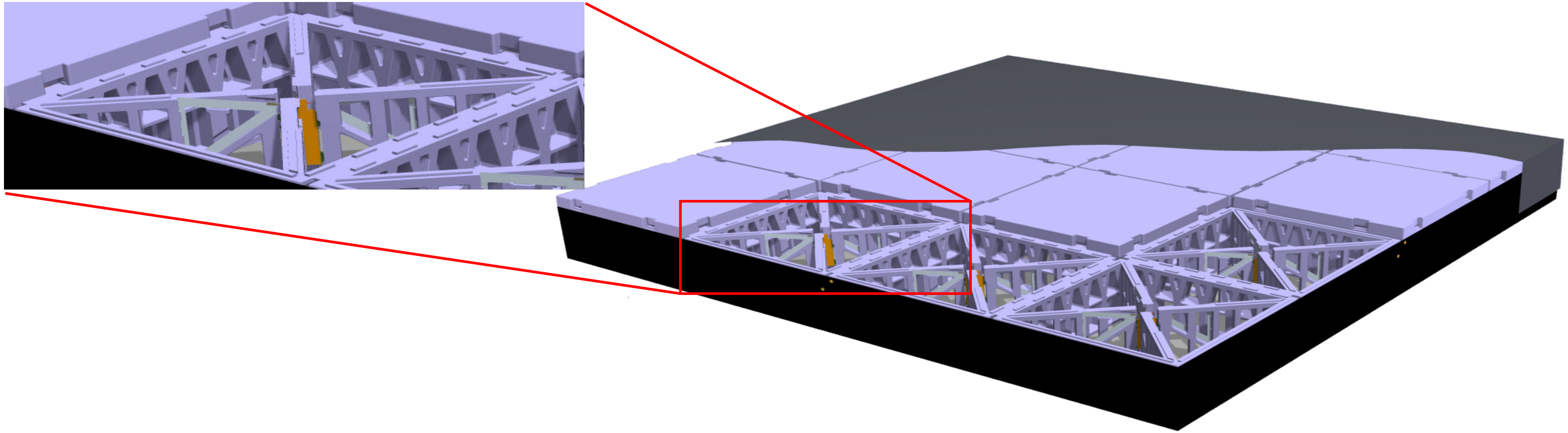
Next-Generation Telescopes: Apertif



LOFAR low-band antennas



LOFAR high-band antennas



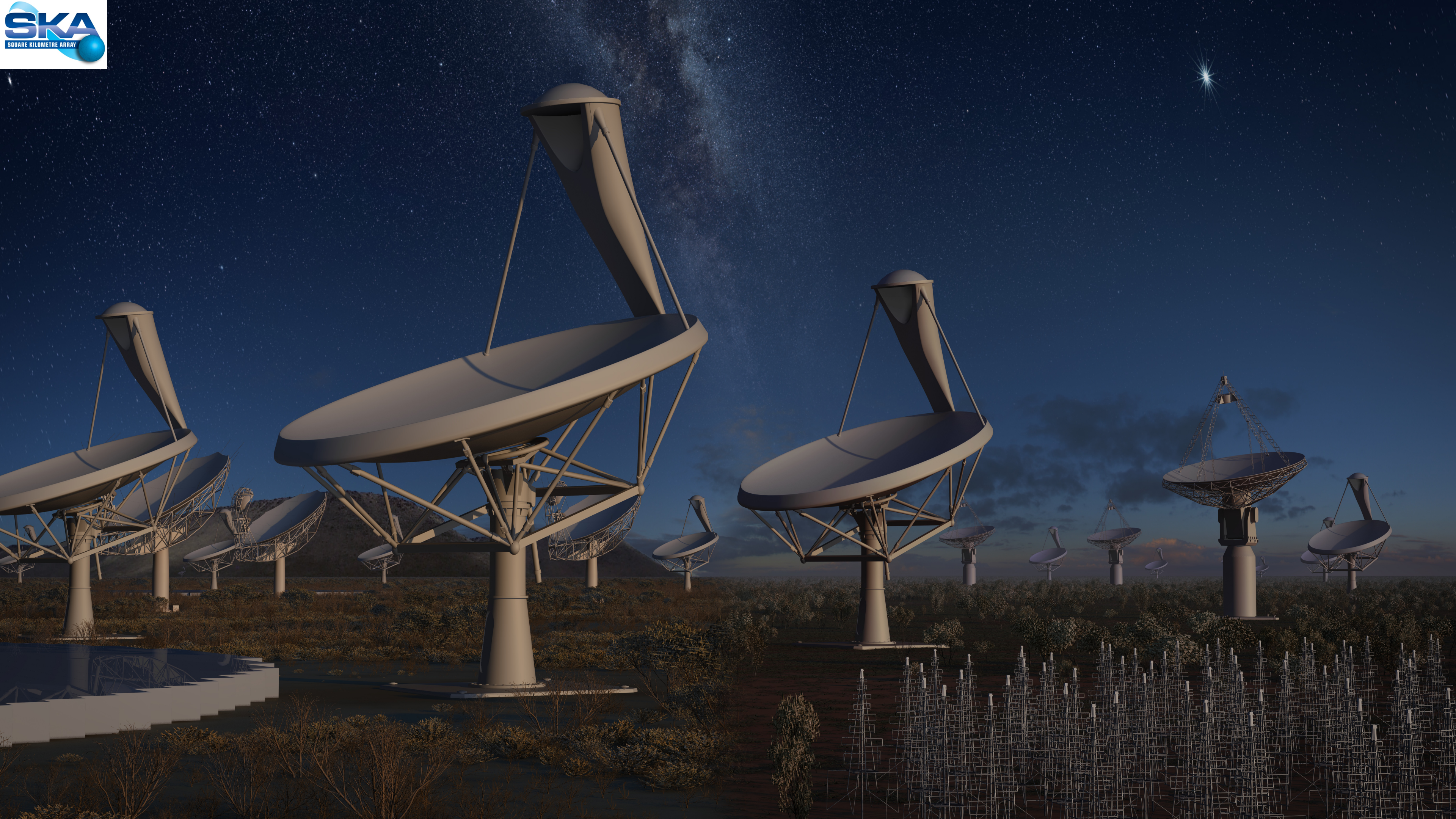
Station (150m)


2x3 km

LOFAR: The low-frequency array

- One of the largest telescopes in the world
- ~100.000 omni-directional antennas
- Ten terabit/s, 200 gigabit/s to supercomputer
- Hundreds of teraFLOPS
- 10–250 MHz
- 100x more sensitive

[John Romein et al, PPoPP, 2014]


SKA
SQUARE KILOMETRE ARRAY

~~Think Big~~ Think Huge: The Square Kilometre Array (SKA)

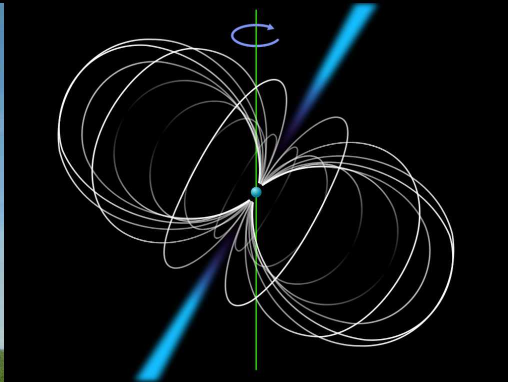
[Chris Broekema et al, Journal of Instrumentation, 2015]

Science Case

Pulsar Searching

Searching for Pulsars

- Rapidly rotating neutron stars
  - Discovered in 1967; ~2500 are known
  - Large mass, precise period, highly magnetized
  - Most neutron stars would be otherwise undetectable with current telescopes
- “Lab in the sky”
  - Conditions far beyond laboratories on Earth
  - Investigate interstellar medium, gravitational waves, general relativity
  - Low-frequency spectra, pulse morphologies, pulse energy distributions
  - Physics of the super-dense superfluid present in the neutron star core

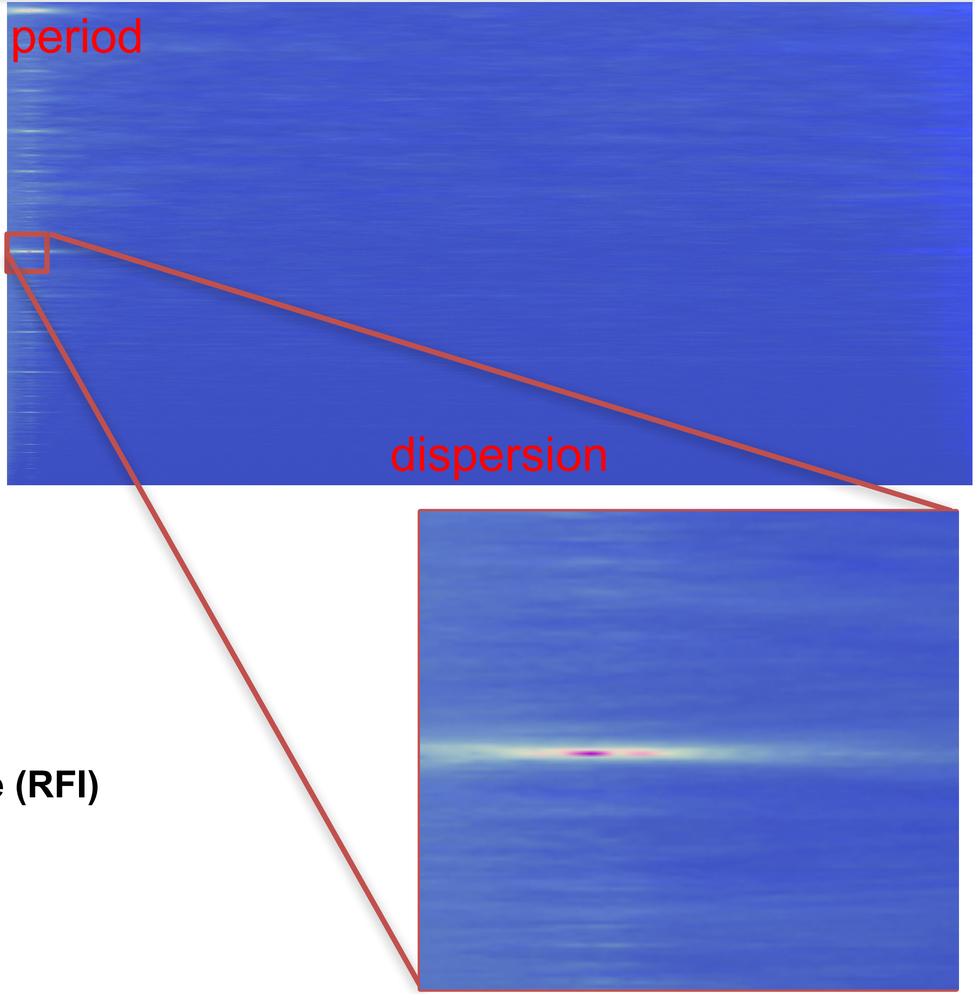


Alessio Sclocco, Rob van Nieuwpoort, Henri Bal,
Joeri van Leeuwen, Jason Hessels, Marco de Vos

[A. Sclocco et al, IEEE eScience, 2015]

Pulsar Searching Pipeline

- Three unknowns:
  - Location: create many beams on the sky
    [Alessio Sclocco et al, IPDPS, 2012]
  - Dispersion: focusing the camera
    [Alessio Sclocco et al, IPDPS, 2012]
  - Period
- Brute force search across all parameters
- Everything is trivially parallel (or is it?)
- Complication: Radio Frequency Interference (RFI)
  [Rob van Nieuwpoort et al: Exascale Astronomy, 2014]



Challenges

- Application becomes real-time because of the data rates
- Limited window of samples due to memory and compute constraints
  - Only fraction of a second, only limited statistics from the past
  - Only small number of frequency bands
- We can afford only few operations per byte
- Distributed system
  - Information distribution, synchronization, scheduling and load-balancing issues
- Limited power budget
- Investigate best platform, develop new algorithms

Potential of accelerators

- Example: NVIDIA K80 GPU (2014)

- Compared to modern CPU (Intel Haswell, 2014)
  - 28 times faster at 8 times less power per operation
  - 3.5 times less memory bandwidth per operation
  - 105 times less bandwidth per operation including PCI-e

- Compared to BG/p supercomputer
  - 642 times faster at 51 times less power per operation
  - 18 times less memory bandwidth per operation
  - 546 times less bandwidth per operation including PCI-e

- Legacy codes and algorithms are inefficient
- Need different programming methodology and programming models, algorithms, optimizations

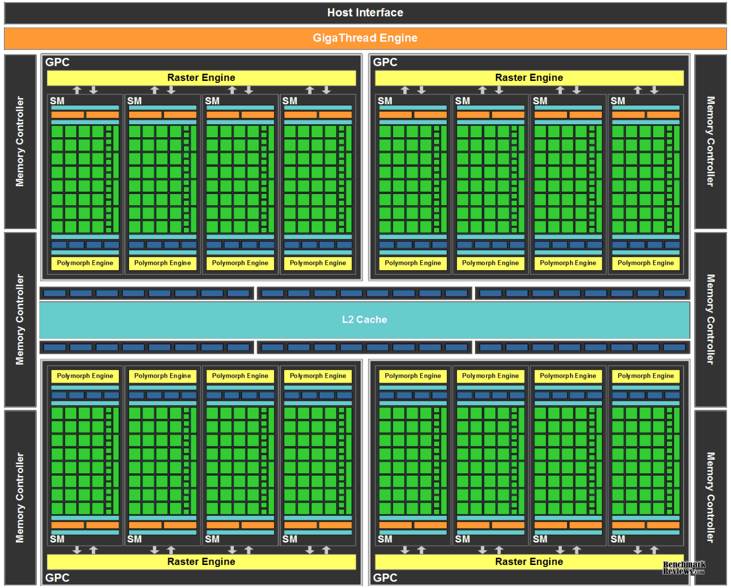
- Can we build large-scale scientific instruments with accelerators?

Systems become increasingly hierarchical

- Instruction-level parallelism, vectors, threads, warps, streaming multiprocessors, chips, multiple devices/node, islands, supercomputer, (hierarchical) distributed system

- Need to explicitly address parallelism on each level

- Communication
  - Explicit
  - Overlap communication and computation on all levels
  - Explicit caches, fast local memories, network on-chip, PCI-e, local interconnects, lightpaths

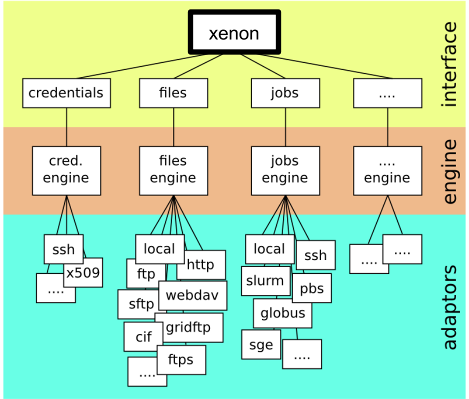


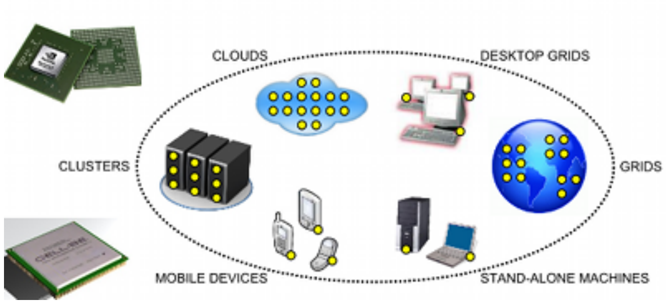
- Even supercomputers are becoming more and more heterogeneous: multiple generations of CPUs and accelerators in one system

Jungle computing with Ibis

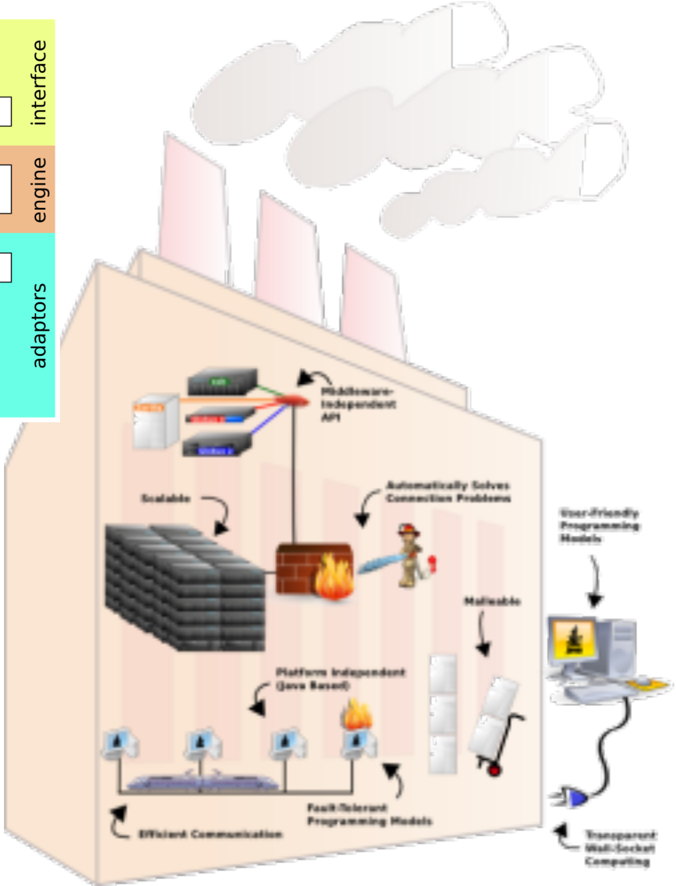
- **Xenon: middleware- independent deployment**
- **IPL: portable communication substrate**
  - **Steaming, malleable, asynchronous, upcalls**
  - **Solves connectivity issues**
  - **On top of TCP, UDP, MPI, infiniband, …**
- **High-level programming models**

[Henri Bal et al, IEEE computer 2010]



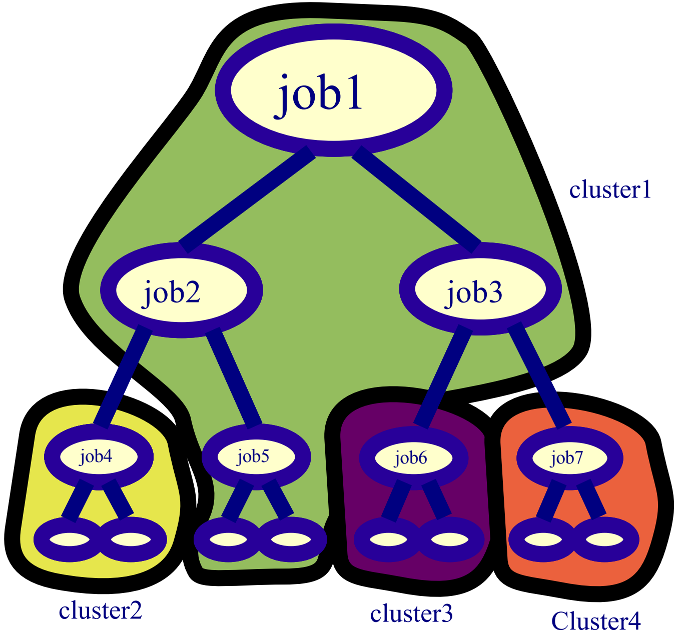


Multi-wide-area (jungle)
Multi-local-area (wide-area)
Multi-node (local-area)
Multi-processor (node)
Multi-core (processor)
Core



Programming hierarchical systems

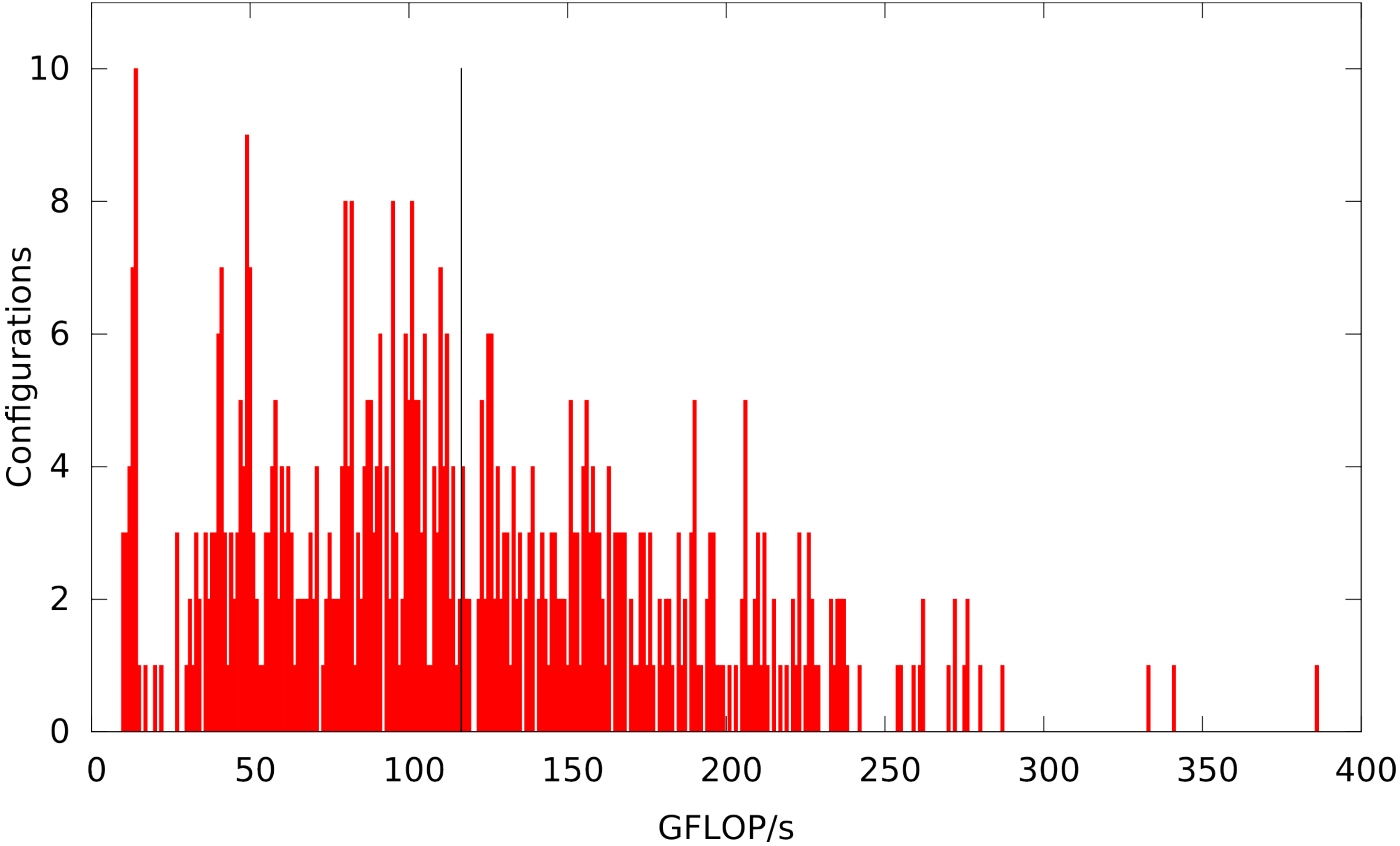
- Need truly hierarchical programming models
  - Hierarchy-aware MPI (point-to-point and collectives)
  - Example: divide-and-conquer
  - Generic model
  - Proven optimal for shared memory multiprocessors, uniform clusters (Cilk)
  - Shown to work extremely well in hierarchical distributed systems (Satin)
  - Fault-tolerance, malleability, adaptive, speculative parallelism, …
  - [Rob van Nieuwpoort et al, ACM TOPLAS, 2010]
- Cashmere integrates Satin & accelerators
  - Mixed programming models
  - Stepwise refinement for performance methodology
  - [Pieter Hijma et al, IPDPS 2015]
- My holy grail: one unified programming model to rule them all



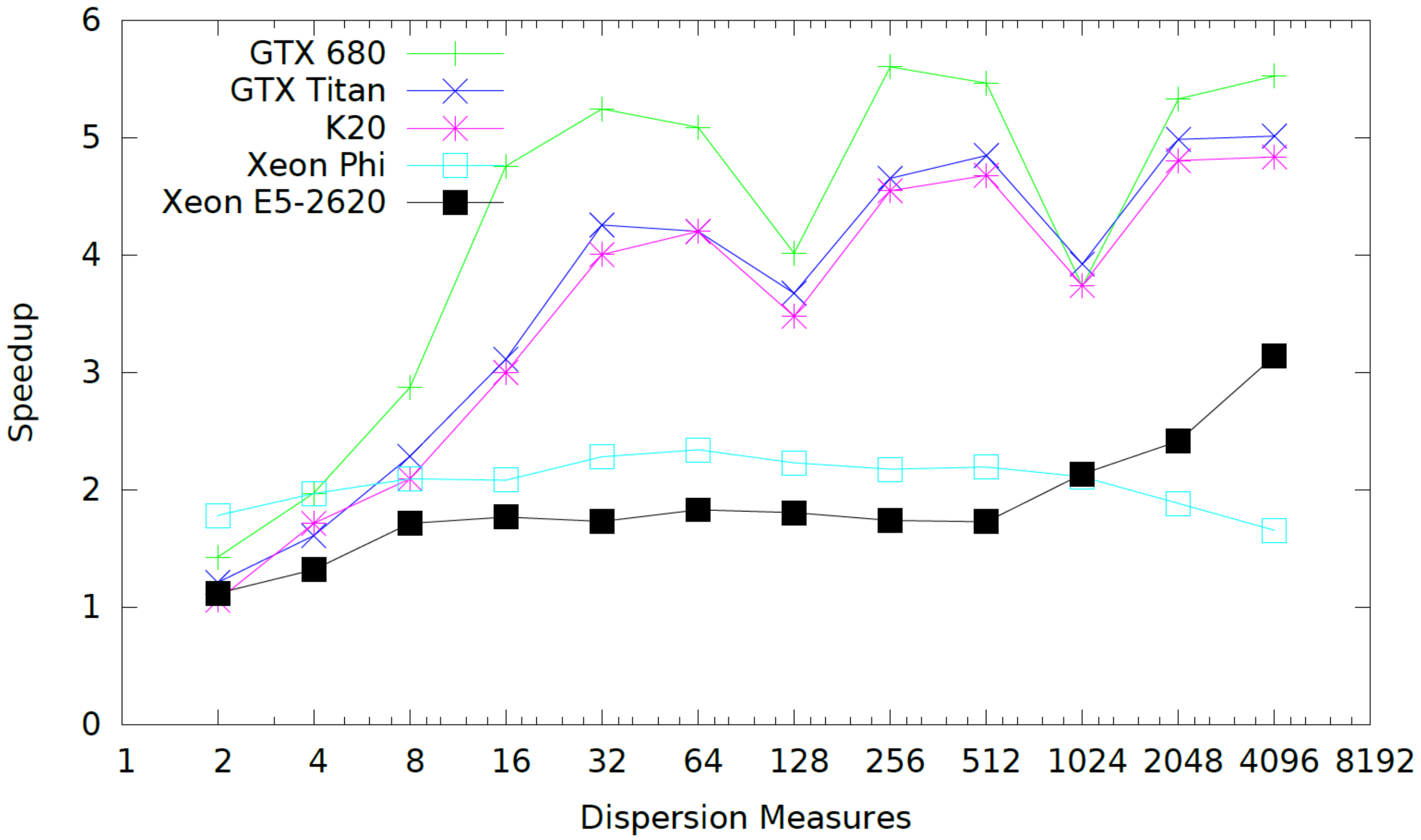
Our Strategy for flexibility, portability

- Investigate algorithms
- OpenCL: platform portability
- Observation type and parameters only known at run time
  - E.g. # frequency channels, # receivers, longest baseline, filter quality, observation type
- Use runtime compilation and auto-tuning
  - Map *specific problem instance* efficiently to hardware
  - Auto tune platform-specific parameters
- Portability across different instruments, observations, platforms, time!

Histogram: Auto-Tuning Dedispersion on AMD HD7970 for Apertif



Speedup over best possible fixed configuration

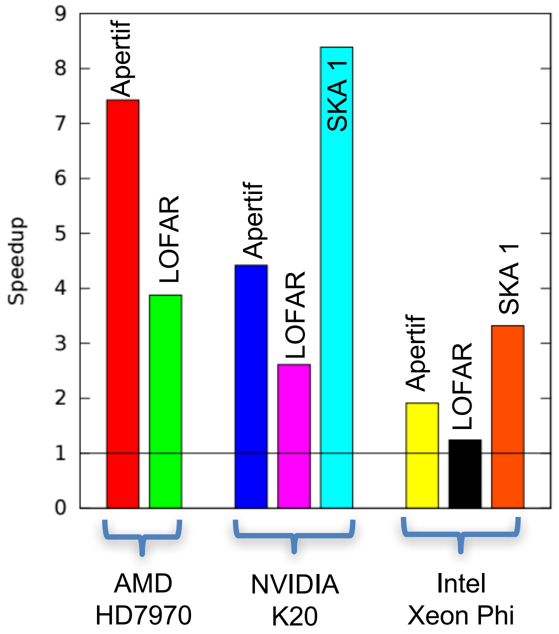


Apertif scenario

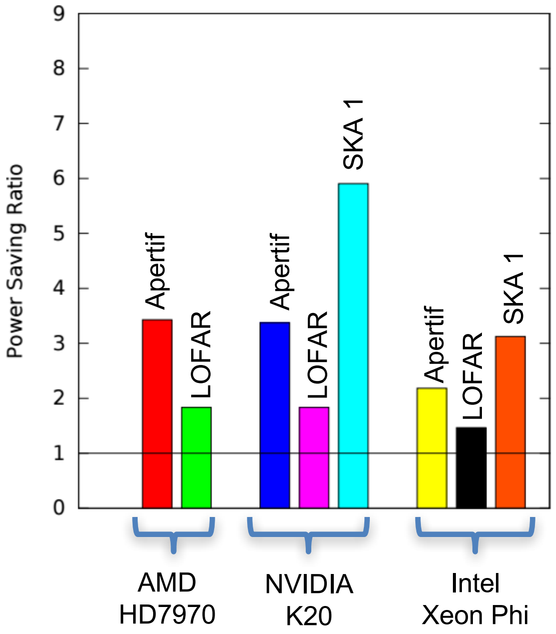
Pulsar pipeline

Apertif and LOFAR: real data
SKA1: simulated data

Speedup over Intel Xeon E5-2620 CPU, 2048x2048 case



Power saving over Intel Xeon E5-2620 CPU, 2048x2048 case



SKA1 baseline design, pulsar survey: 2,222 beams; 16,113 DMs; 2,048 periods.
Total number of GPUs needed: 140,000. This requires 30 MW. SKA2 should be 100x larger, in the 2023-2030 timeframe.

Conclusions

- Exascale changes everything
  - Offline versus streaming, best hardware architecture, algorithms, optimizations
  - All large compute platforms are becoming heterogeneous
  - Needed 8 years to make this work for one application
    We desperately need high-level programming models that incorporate the entire hierarchy!
  - Need new theory as well: from computational complexity to data access complexity

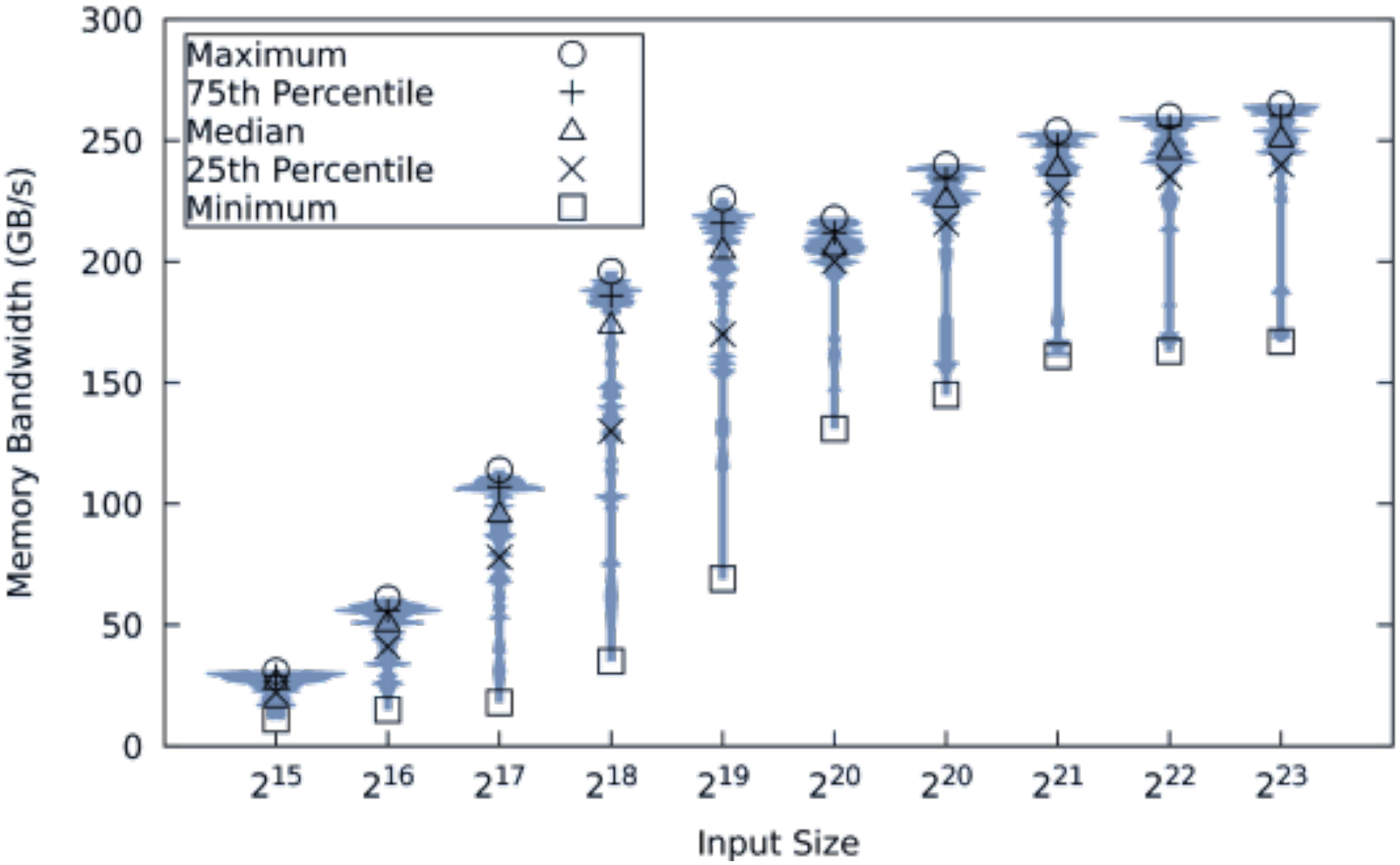
- eScience approach works!
  - Need domain expert for deep understanding & choice of algorithms
  - Need computer scientists for investigating efficient solutions
  - LOFAR has already discovered more than 25 new pulsars!

- Astronomy is a driving force for HPC, Big Data, eScience
  - Techniques are general, already applied in image processing, climate, digital forensics

An example of real time challenges

Investigate algorithms: RFI mitigation

Auto-tuning example, NVIDIA TitanX



Radio Frequency Interference

- RFI is a huge problem for many observations
- Caused by
  - Lightning, vehicles, airplanes, satellites, electrical equipment, GSM, FM Radio, fences, reflection of wind turbines, …
- Best removed offline
  - Complete dataset available
  - Good overview / statistics / model
  - Can spend compute cycles

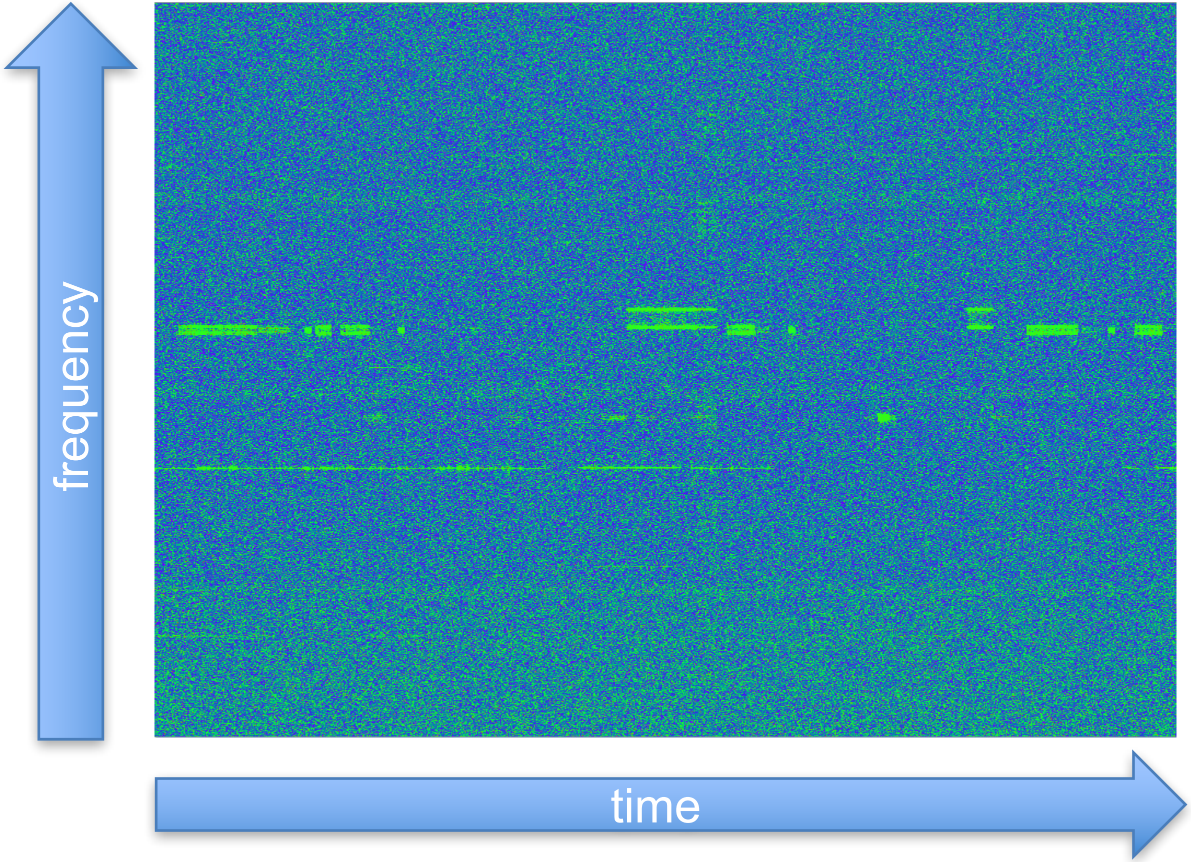


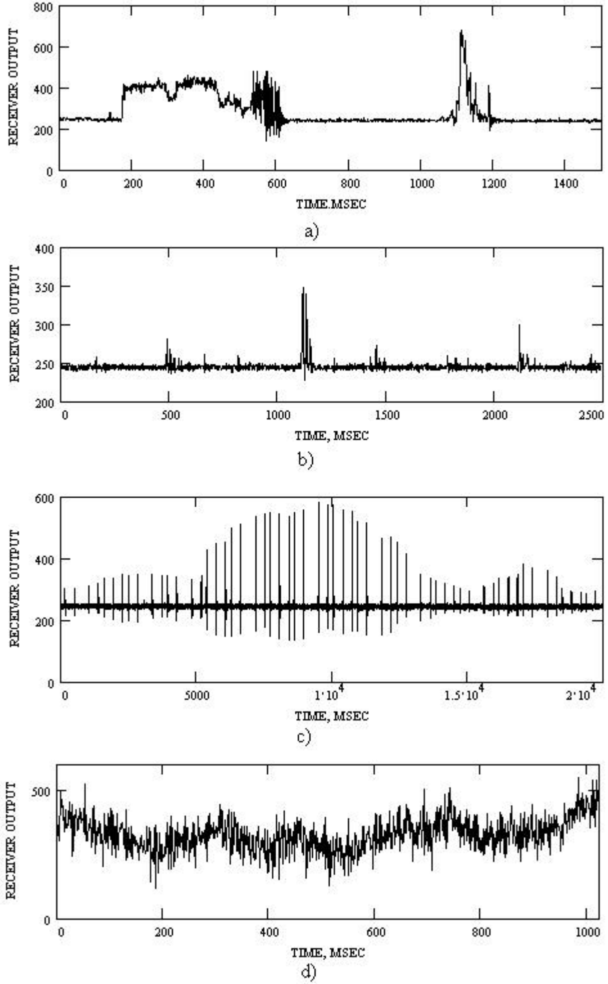

DAB+
Digital Audio Broadcasting



[R.V. van Nieuwpoort et al: Exascale Astronomy, 2014]

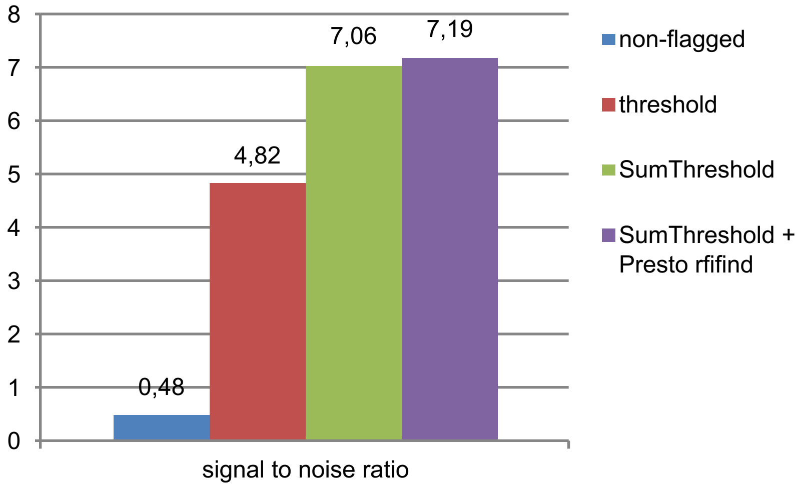
RFI mitigation

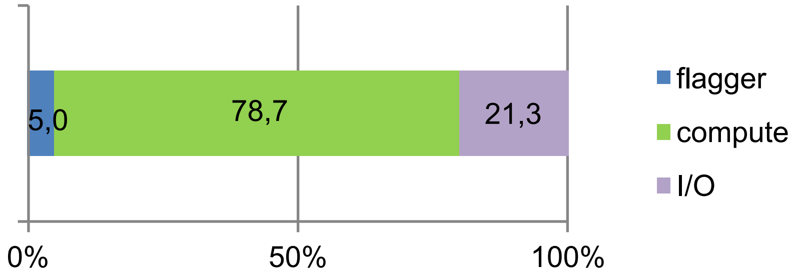




RFI mitigation results

- One robust algorithm for different scales (μs - hours)
  - Filters with exponentially increasing window sizes
- Scalable: linear computational complexity
- Quality almost as good as offline

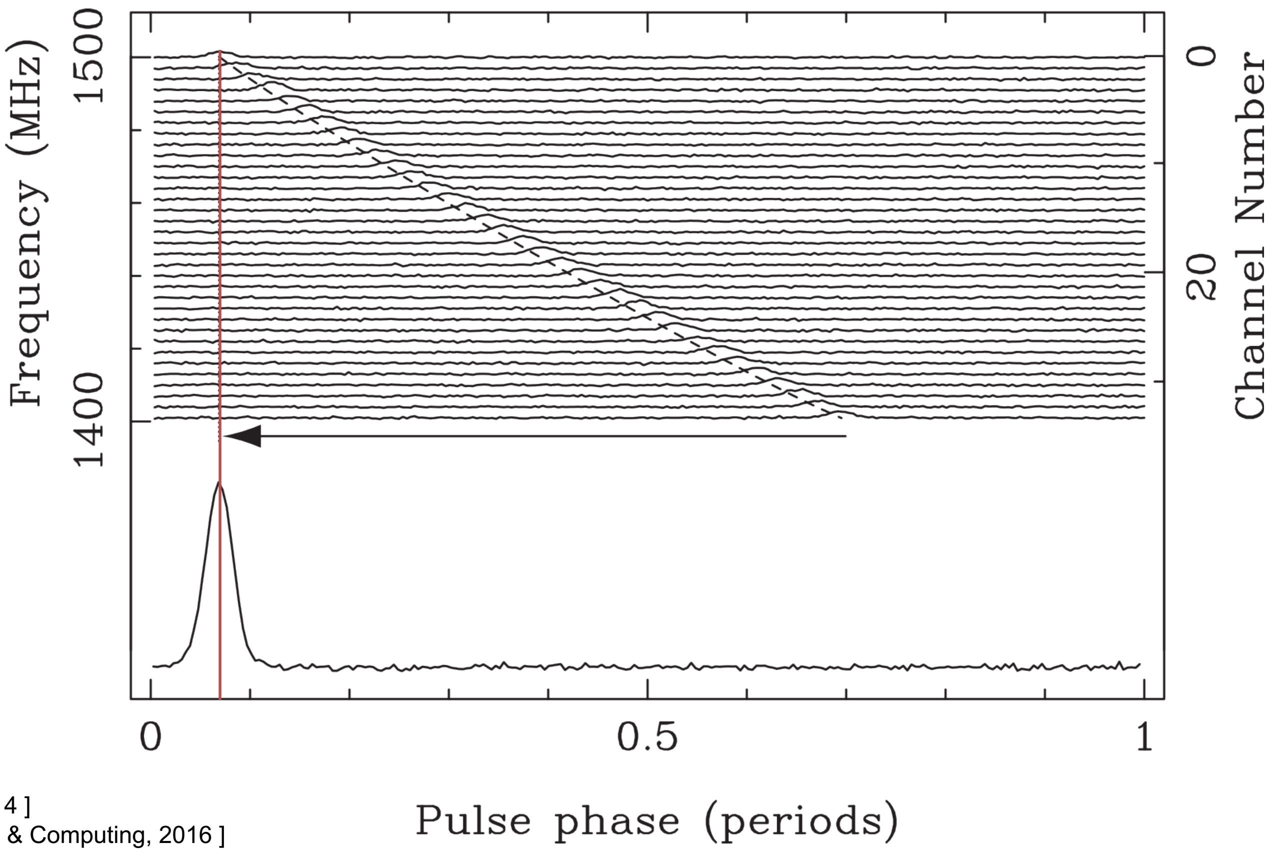




An example of real time challenges

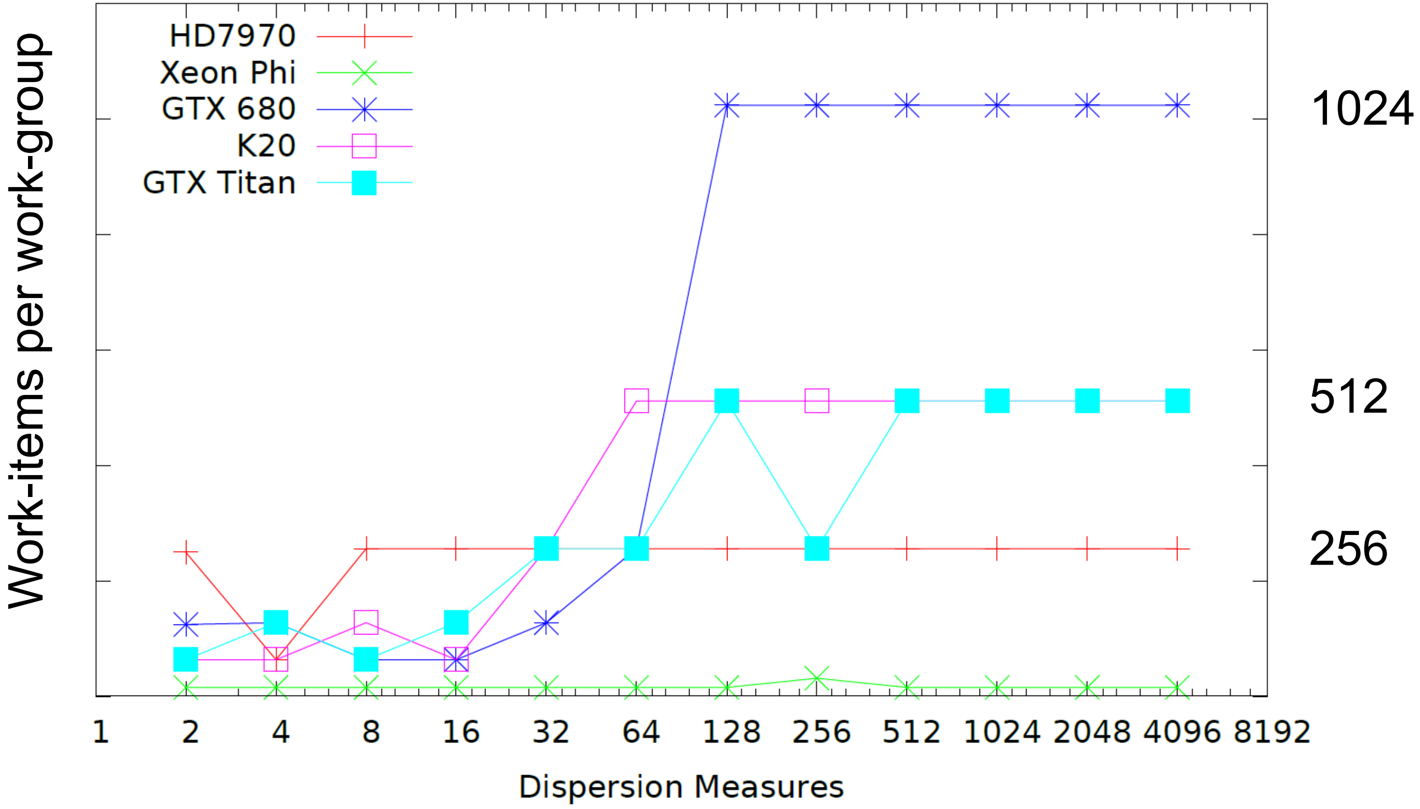
Auto-tuning: Dedispersion

Dedispersion



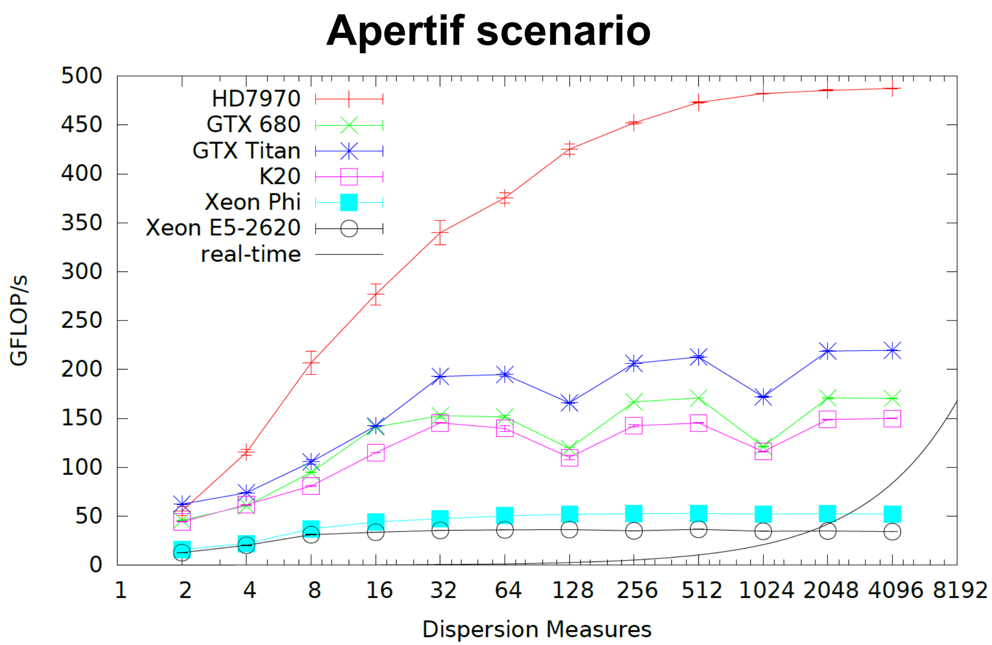
[A. Sclocco et al, IPDPS 2014]
[A. Sclocco et al, Astronomy & Computing, 2016]

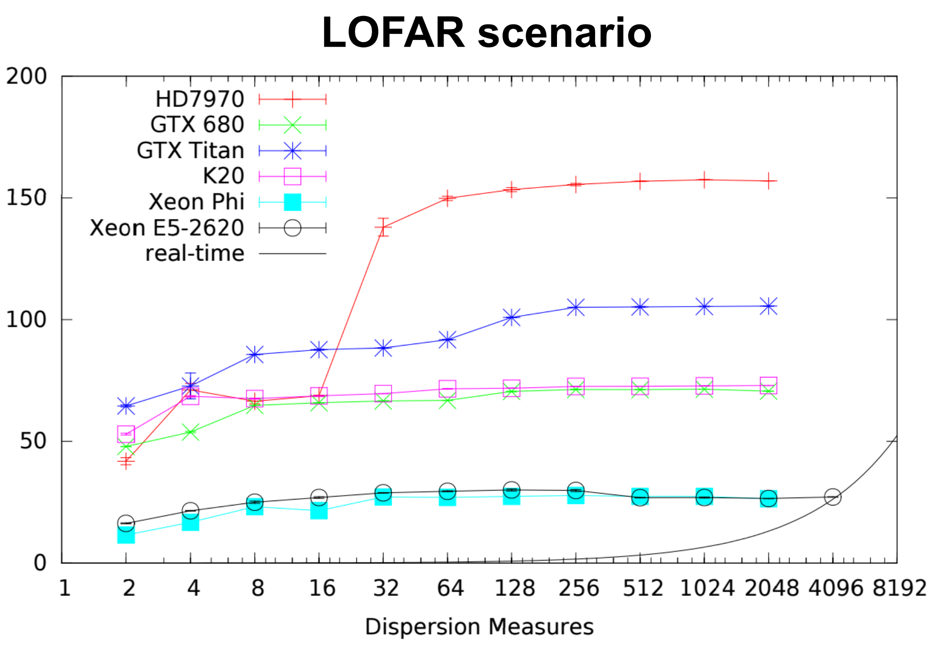
Auto-tuning platform parameters



Apertif scenario

Auto-tuned performance



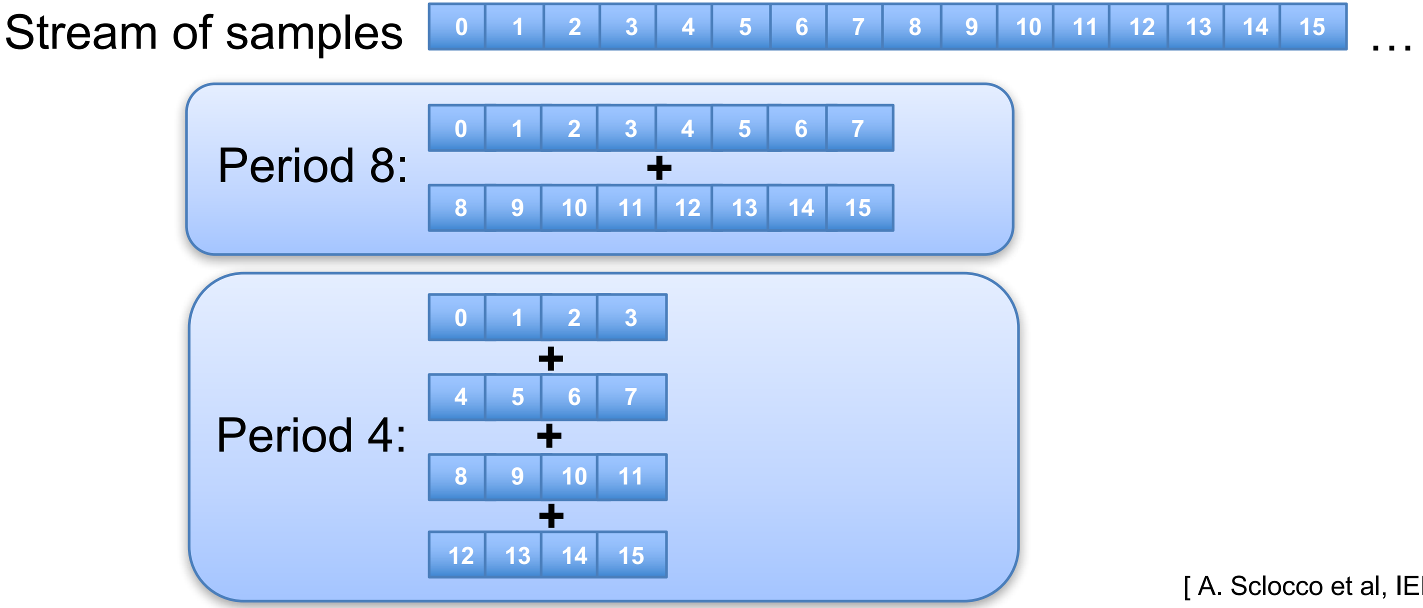


An example of real time challenges

Changing algorithms: Period search

Period Search: Folding

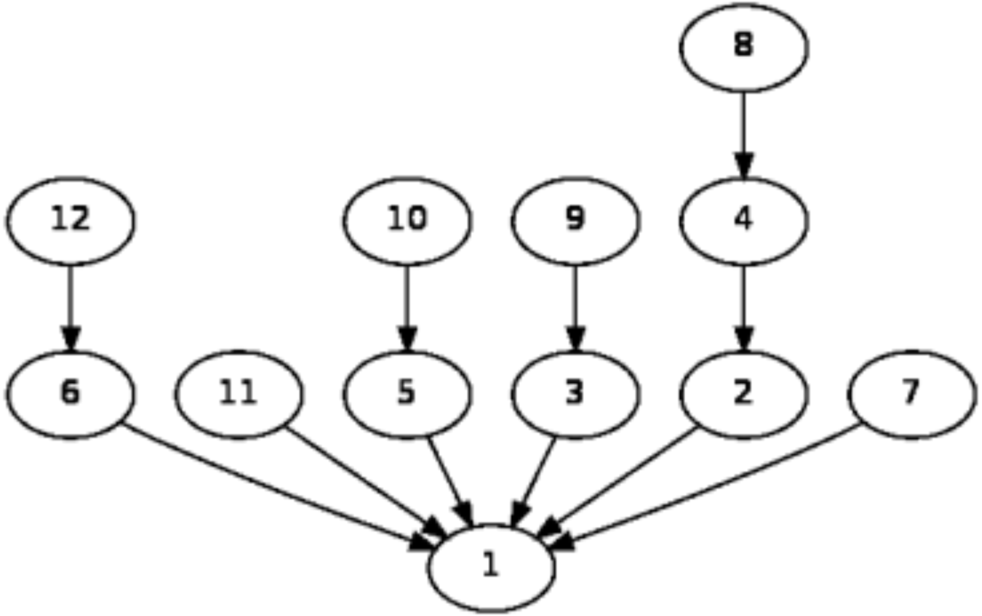
- Traditional offline approach: FFT
- Big Data requires change in algorithm: must be real time & streaming

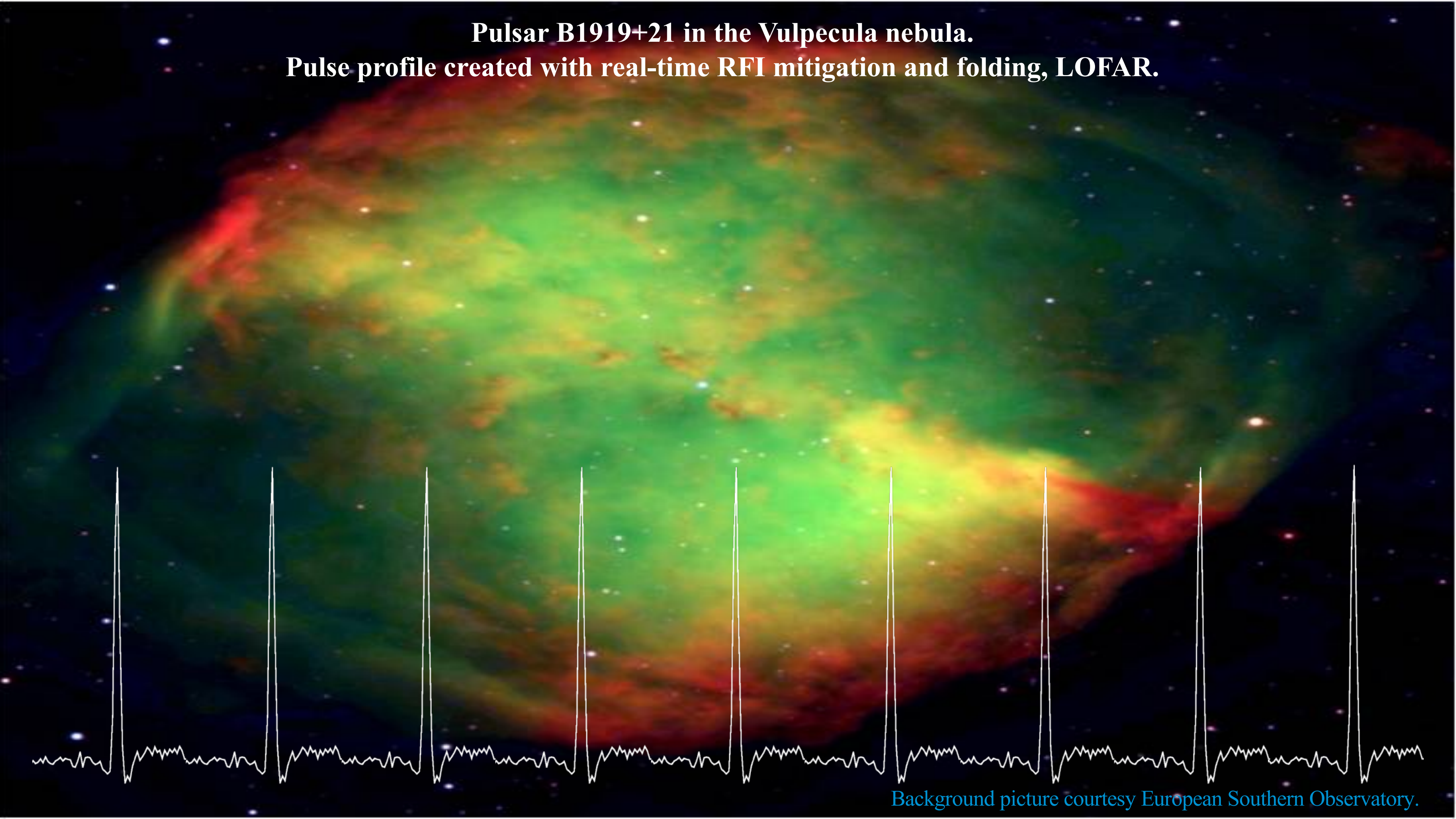


[A. Sclocco et al, IEEE eScience, 2015]

Optimizing Folding

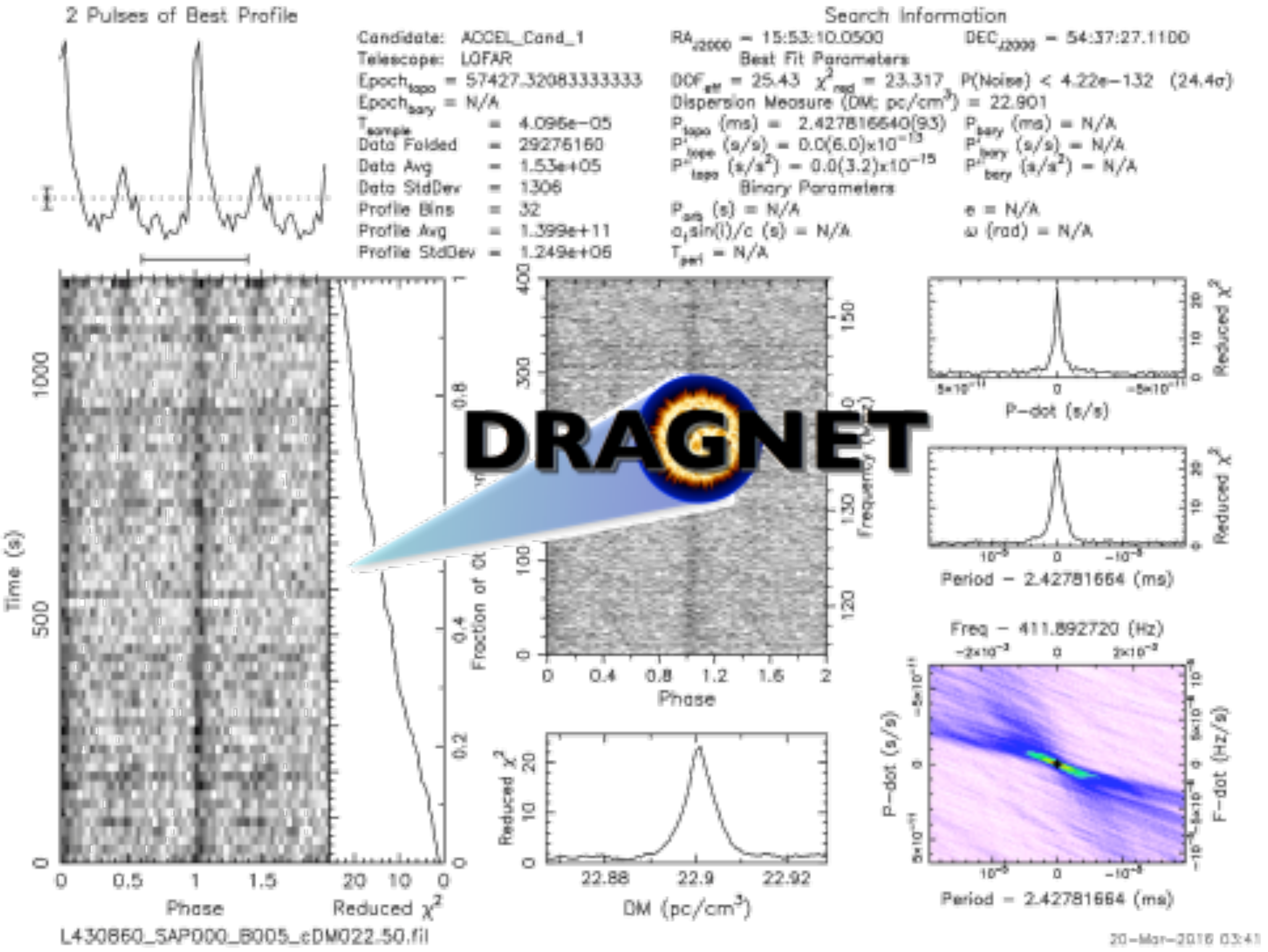
- Build a tree of periods to maximize reuse
- Data reuse: walk the paths from leafs to root




Pulsar B1919+21 in the Vulpecula nebula.
Pulse profile created with real-time RFI mitigation and folding, LOFAR.
Background picture courtesy European Southern Observatory.

Today's discovery

- Millisecond pulsar PSR J1552+54
- Discovered at 135 MHz
- Lowest observing frequency an MSP has been discovered.
- Non-detections at 1400MHz by Lovell and Nancay.
- Use of LOFAR indispensable.



Credits


Rob van Nieuwpoort


Alessio Sclocco


Pieter Hijma


Chris Broekema


Ana Lucia Varbanescu


John Romein


Ger van Diepen


Joeri van leeuwen

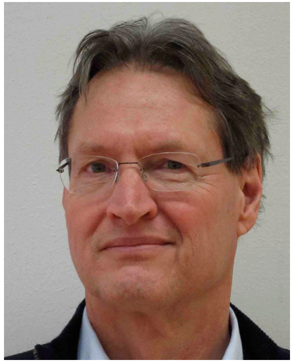

Jason Hessels


Souley Madougou


Tim Cornwell


Bruce Almegreen


Henri Bal

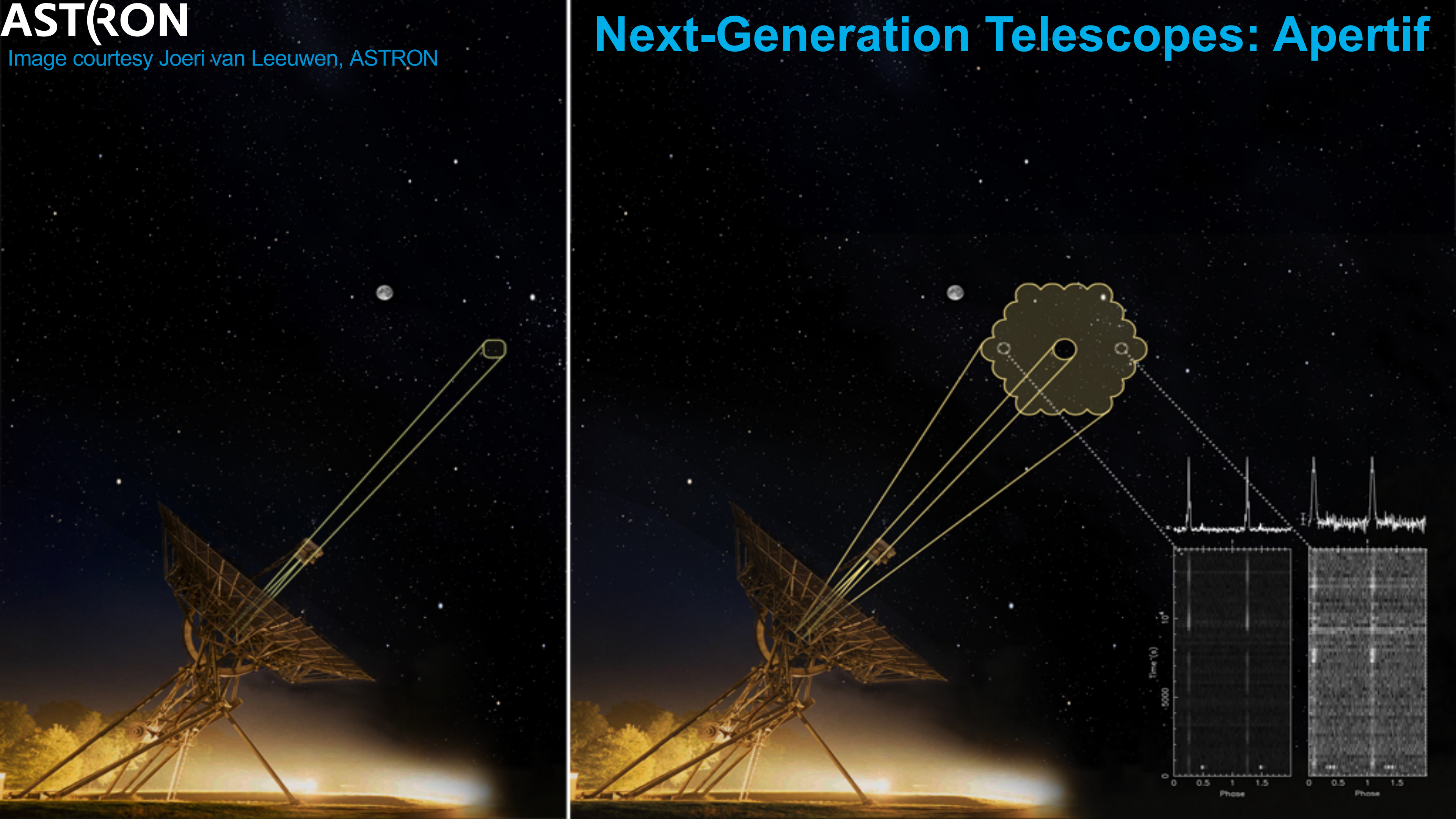

Henk Sips

and many others!

ASTRON

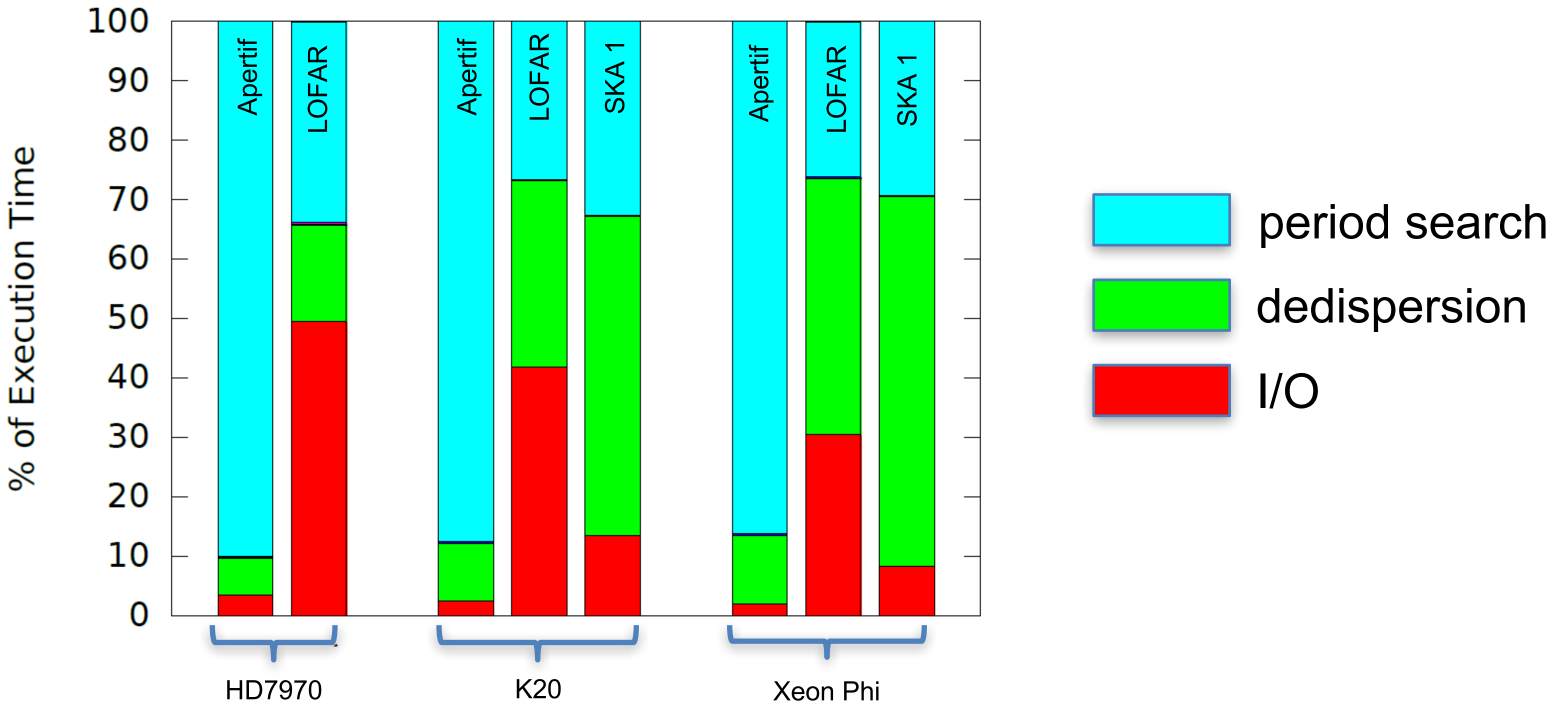
Image courtesy Joeri van Leeuwen, ASTRON

Next-Generation Telescopes: Apertif



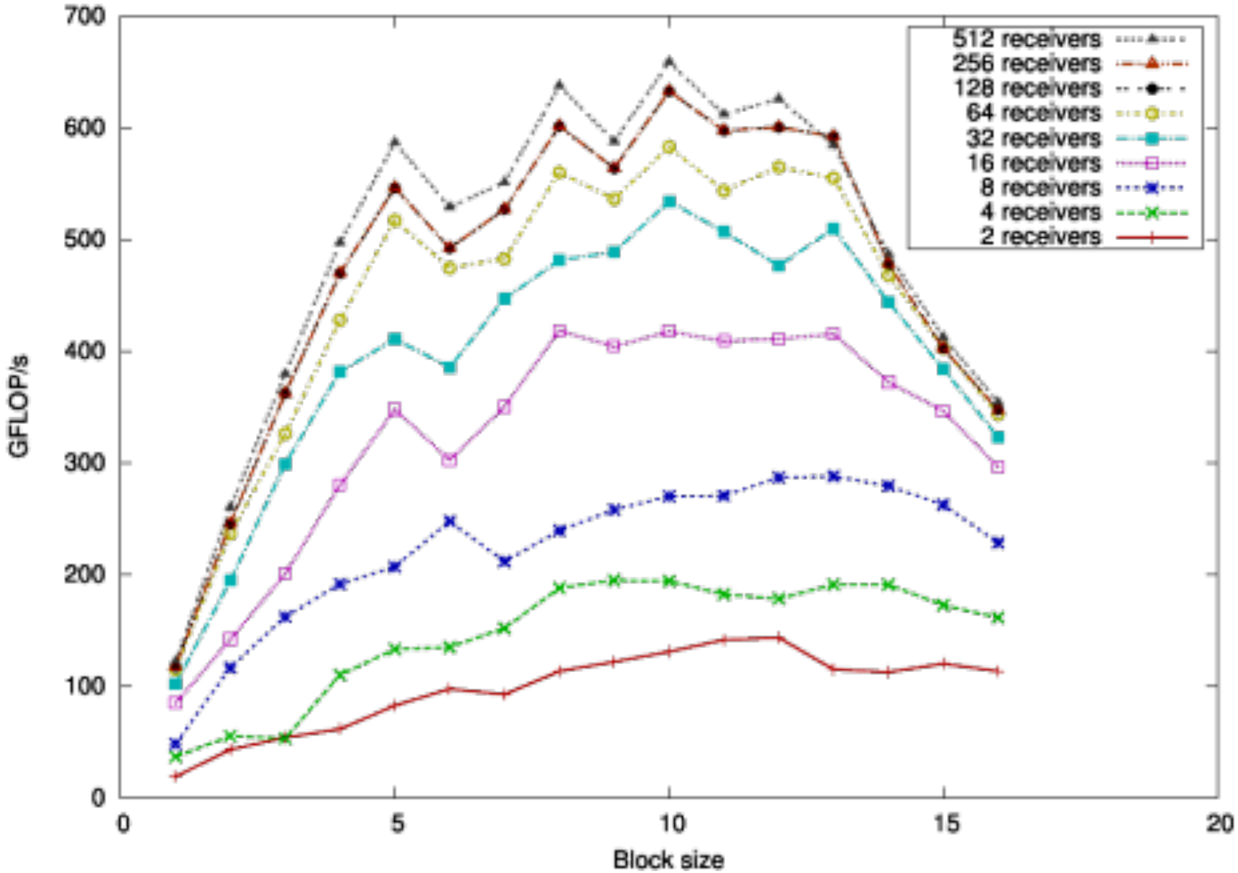
Backup slides

Pulsar pipeline Performance Breakdown

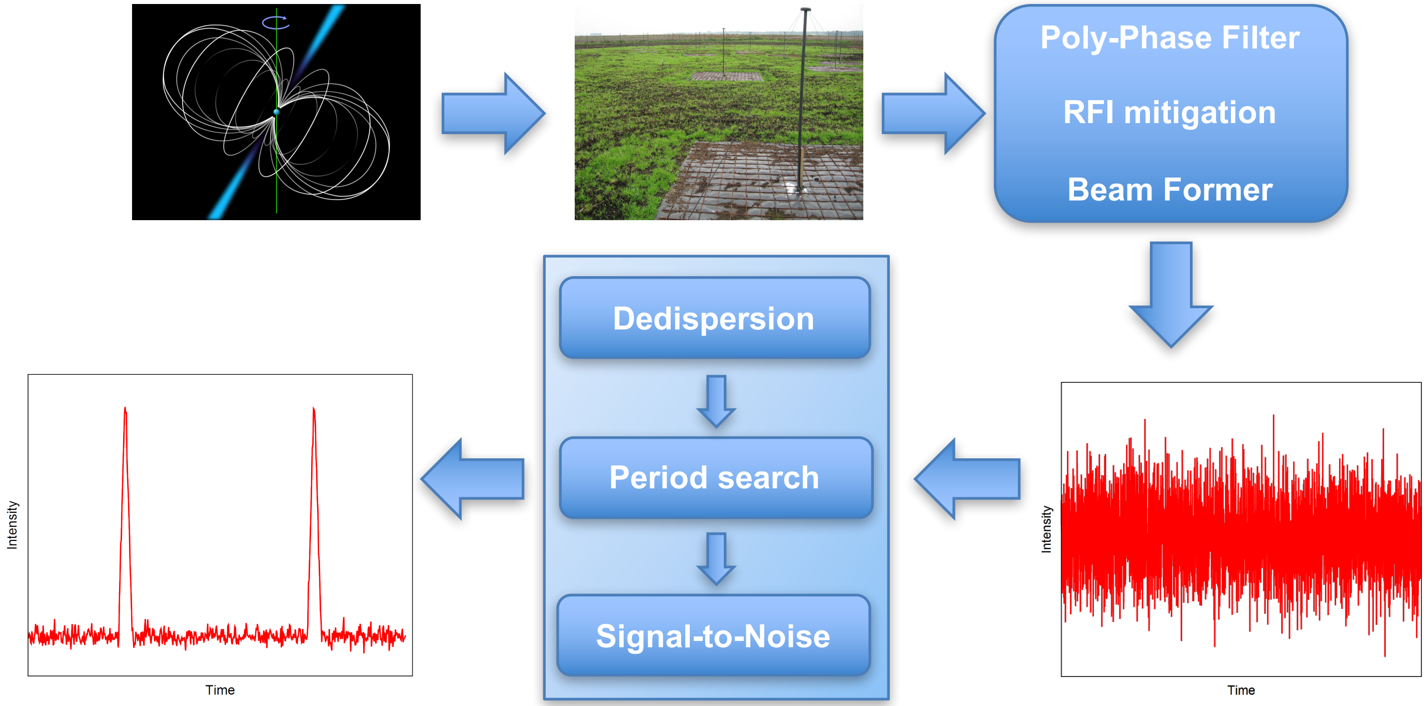


Auto-tuning example

Beam Forming, GTX580



Pulsar pipeline (real time)



Modern Computer Architectures

	IBM BG/p Super	CPU Haswell (2015)	GPU NVIDIA K80 (2014)	
Peak (gflops)	13.6	307	8736	(28x CPU)
Memory bandwidth (GByte/s)	13.6	59	480	(8x CPU)
Operations / byte	1.0	5.2	18.2	(3.5x CPU)

Modern Computer Architectures

	IBM BG/p Super	CPU Haswell (2015)	GPU NVIDIA K80 (2014)
Peak (gflops)	13.6	307	8736 (28x CPU)
Memory bandwidth (GByte/s)	13.6	59	480 (8x CPU)
Operations / byte	1.0	5.2	18.2 (3.5x CPU)

Including data movement (PCI-e): **546 (105x CPU)**

Modern Computer Architectures

	IBM BG/p Super	CPU Haswell (2015)	GPU NVIDIA K80 (2014)	
Peak (gflops)	13.6	307	8736	(28x CPU)
Memory bandwidth (GByte/s)	13.6	59	480	(8x CPU)
Operations / byte	1.0	5.2	18.2	(3.5x CPU)

Including data movement (PCI-e): **546 (105x CPU)**

gflops / Watt	0.57	3.61	29.1	**(8.1x) CPU**

51x BG/p

Modern Computer Architectures

	IBM BG/p Super	CPU Haswell (2015)	GPU NVIDIA K80 (2014)	
Peak (gflops)	13.6	307	8736	(28x CPU)
Memory bandwidth (GByte/s)	13.6	59	480	(8x CPU)
Operations / byte	1.0	5.2	18.2	(3.5x CPU)
Including data movement (PCI-e):			**546 (105x CPU)**	
gflops / Watt	0.57	3.61	29.1	**(8.1x) CPU**

51x BG/p

- Huge performance potential and increase in power efficiency
- Legacy codes are inefficient on modern architectures
- Need completely different optimizations, algorithms, programming models
- Can we build large-scale scientific instruments with accelerators?

Big Data == Big Compute

- We need "Big Compute" for processing Big Data
  - Currently petaflops
  - SKA will be exascale

    [Chris Broekema et al, Journal of Instrumentation, 2015]
- Large-scale parallelism
- Accelerators
  - GPUs NVIDIA, AMD; Intel Xeon Phi, FPGAs, ASICs, DSPs, …

Modern Computer Architectures

	IBM BG/p Super	CPU Nehalem (2009)	CPU Haswell (2015)	GPU NVIDIA GTX Titan (2014)	GPU NVIDIA K80 (2014)	
Peak (gflops)	13.6	85	307	4500	8736	(28x CPU)
Memory bandwidth (GByte/s)	13.6	25.6	59	288	480	(8x CPU)
Operations / byte	1.0	3.3	5.2	15.6	18.2	(3.5x CPU)

Modern Computer Architectures

	IBM BG/p Super	CPU Nehalem (2009)	CPU Haswell (2015)	GPU NVIDIA GTX Titan (2014)	GPU NVIDIA K80 (2014)
Peak (gflops)	13.6	85	307	4500	8736 (28x CPU)
Memory bandwidth (GByte/s)	13.6	25.6	59	288	480 (8x CPU)
Operations / byte	1.0	3.3	5.2	15.6	18.2 (3.5x CPU)

Including data movement (PCI-e): **546 (105x CPU)**

Modern Computer Architectures

	IBM BG/p Super	CPU Nehalem (2009)	CPU Haswell (2015)	GPU NVIDIA GTX Titan (2014)	GPU NVIDIA K80 (2014)
Peak (gflops)	13.6	85	307	4500	8736 (28x CPU)
Memory bandwidth (GByte/s)	13.6	25.6	59	288	480 (8x CPU)
Operations / byte	1.0	3.3	5.2	15.6	18.2 (3.5x CPU)
Including data movement (PCI-e):					**546 (105x CPU)**
gflops / Watt	0.57	0.65	3.61	18.0	29.1 **(8.1x) CPU**

Modern Computer Architectures

	IBM BG/p Super	CPU Nehalem (2009)	CPU Haswell (2015)	GPU NVIDIA GTX Titan (2014)	GPU NVIDIA K80 (2014)
Peak (gflops)	13.6	85	307	4500	8736 (28x CPU)
Memory bandwidth (GByte/s)	13.6	25.6	59	288	480 (8x CPU)
Operations / byte	1.0	3.3	5.2	15.6	18.2 (3.5x CPU)
Including data movement (PCI-e):					**546 (105x CPU)**
gflops / Watt	0.57	0.65	3.61	18.0	29.1 **(8.1x) CPU**

- Legacy codes are inefficient on modern architectures
- Need completely different optimizations, algorithms, programming models
- Can we build large-scale scientific instruments with accelerators?

My Interests: Efficient Computing

- Generic hierarchical programming models
- Efficiently mapping challenging scientific applications to these complex platforms
  - Performance
  - Power
  - Programmability
- This talk: example from astronomy

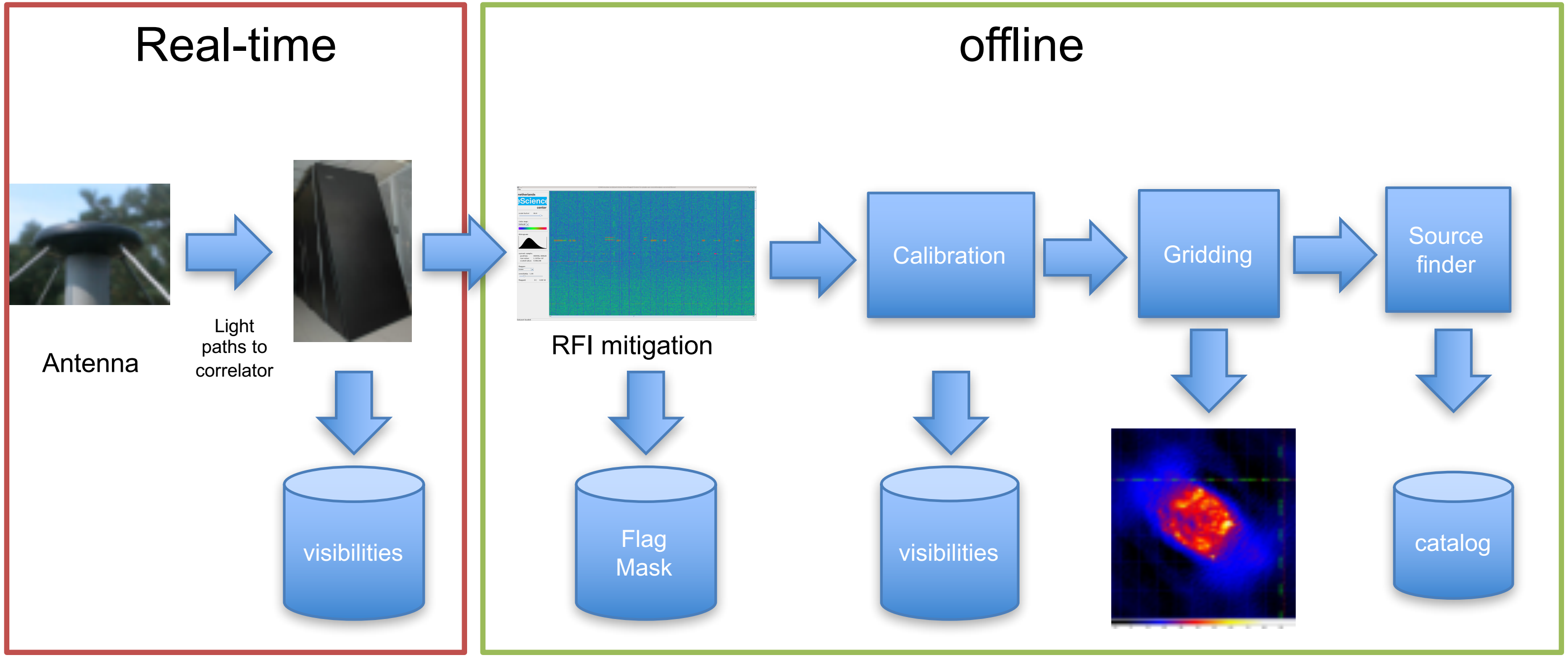
Big Data in Astronomy

- Start of the pipeline: huge volume, structured, 99.999% noise
- Intermediate: huge^2
- Final product
  - Can be 1 bit (pulsar)
  - Can be image cubes: 2D sky, frequency, time
  - Can be source catalogs
- Complexity of data and algorithms increases

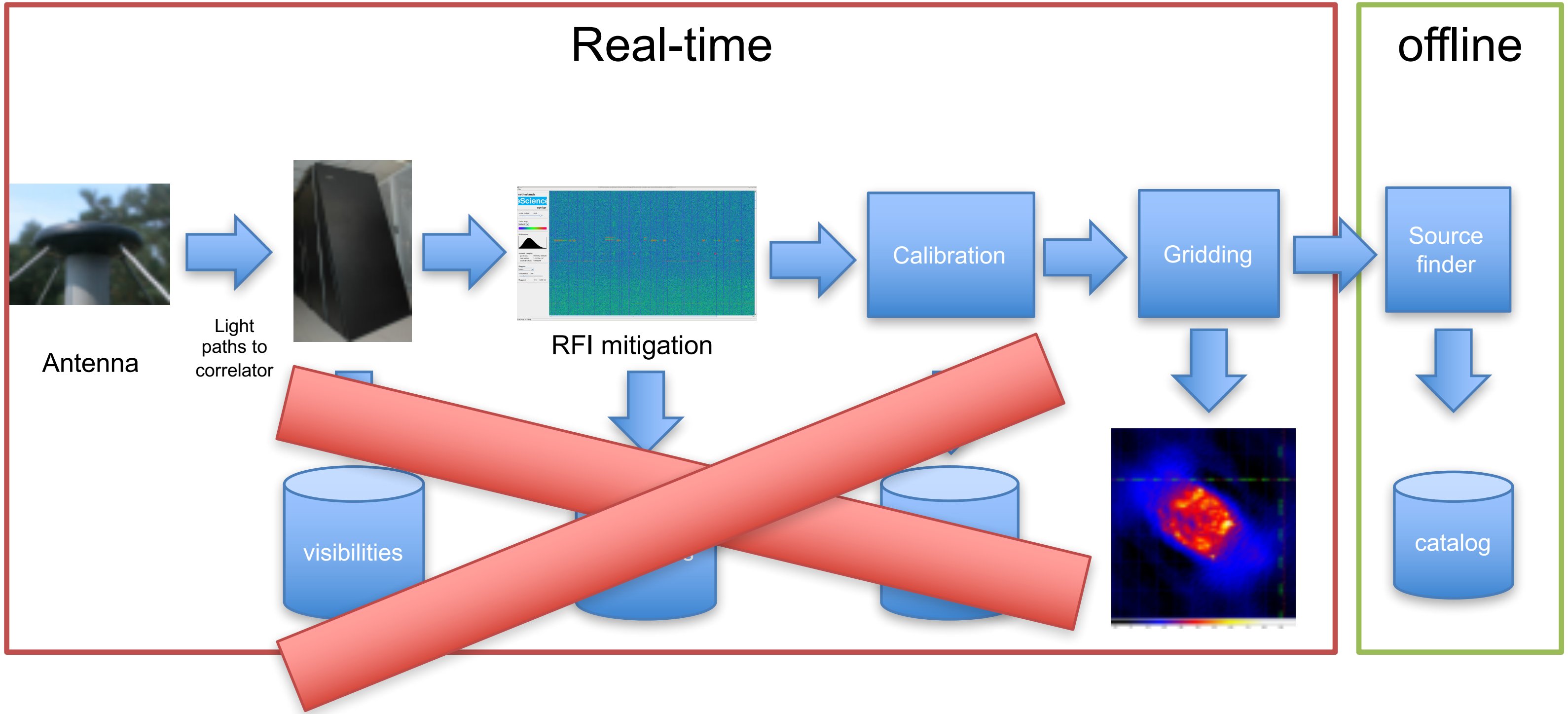
SKA1 details

	SKA1 mid	SKA1 low	SKA1 survey
Number of receivers	254 (190 + 64)	262,144 (1024 x 256)	96 (64 + 36)
Receiver diameter	15 m (13.5 m)	35 m	15 m (12 m)
Maximum baseline	100 km	70 km	50 km
Input bandwidth	34 Tbps	73 Tbps	47 Tbps
Red'q Compute capacity	52 PFLOPS	25 PFLOPS	72 PFLOPS

Imaging pipeline

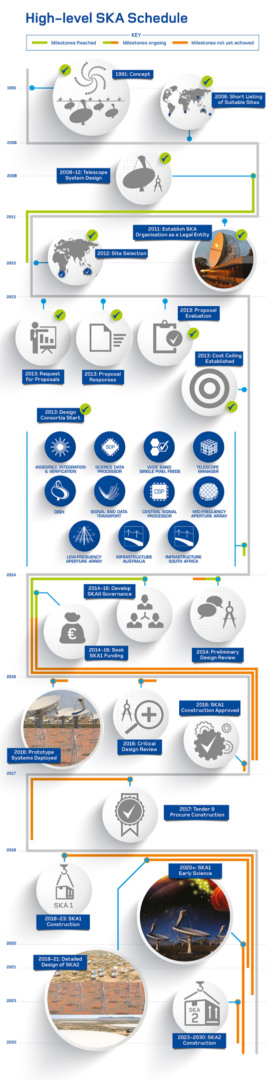


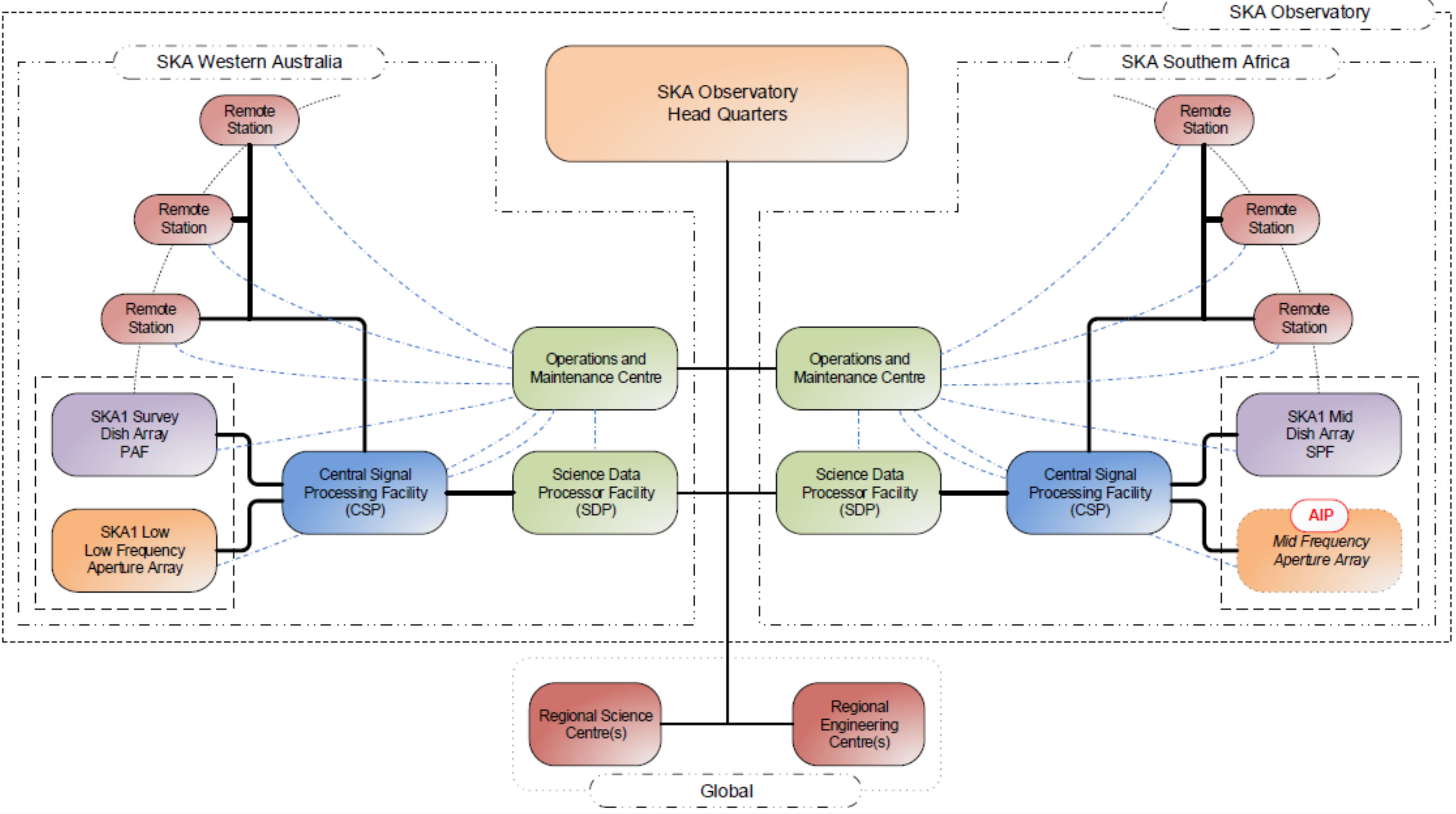
Imaging pipeline: scaling up



Data distribution

- SKA1: construction 2018-2023; early science 2020+
- SKA2: construction 2023 - 2030
- SKA is distributed instrument by design
  - Western Australia and South Africa
  - Central archive?
  - Replicate?
- Distribute image cubes to SKA data science centers
  - Image cubes can be large: ~ 20K x 20K x 1K x double
  - Rough estimate: 100 Gbit/s
- Infrastructure?
- Bring processing to the data?



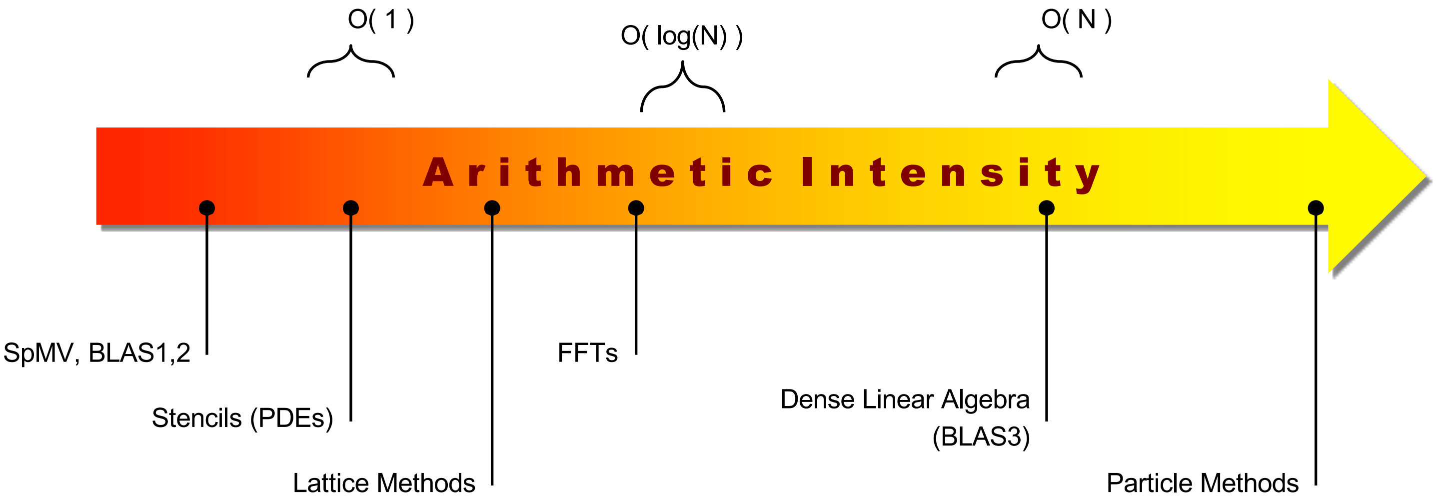


SKA Observatory
SKA Western Australia
SKA Observatory Head Quarters
SKA Southern Africa
Remote Station
Remote Station
Remote Station
Remote Station
Remote Station
Remote Station
Operations and Maintenance Centre
Operations and Maintenance Centre
SKA1 Survey Dish Array PAF
Central Signal Processing Facility (CSP)
Science Data Processor Facility (SDP)
Science Data Processor Facility (SDP)
Central Signal Processing Facility (CSP)
SKA1 Mid Dish Array SPF
SKA1 Low Low Frequency Aperture Array
AIP
Mid Frequency Aperture Array
Regional Science Centre(s)
Regional Engineering Centre(s)
Global

Flexibility and Portability

- **Many different instruments**
- **Many different science cases, observation types, and parameters**
- **Life time of an instrument is much longer than life time of compute hardware**

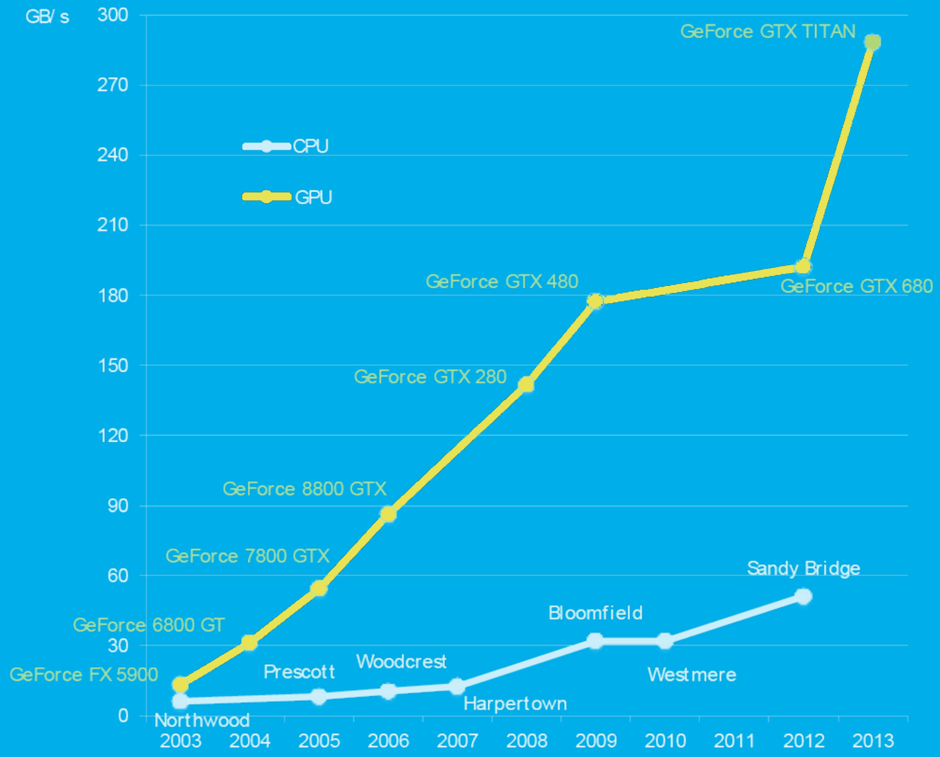
The balance has shifted



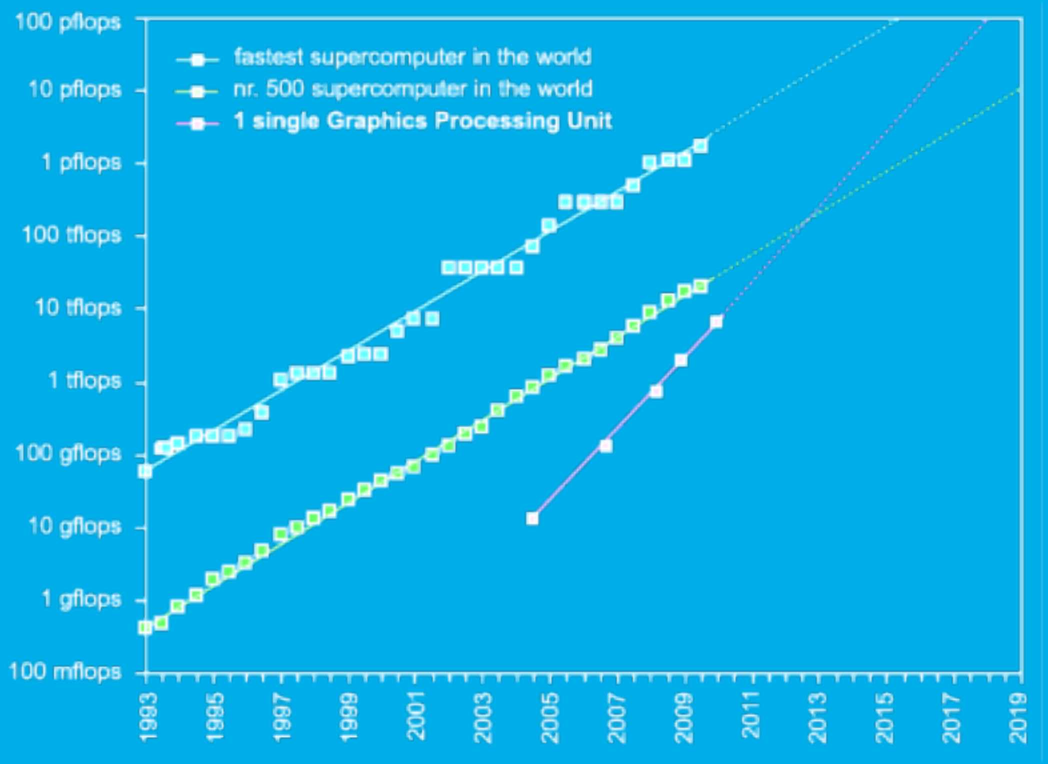
- **Legacy codes are inefficient on modern architectures**
  - **Need completely different optimizations, algorithms**

CPU versus GPU Compute Performance

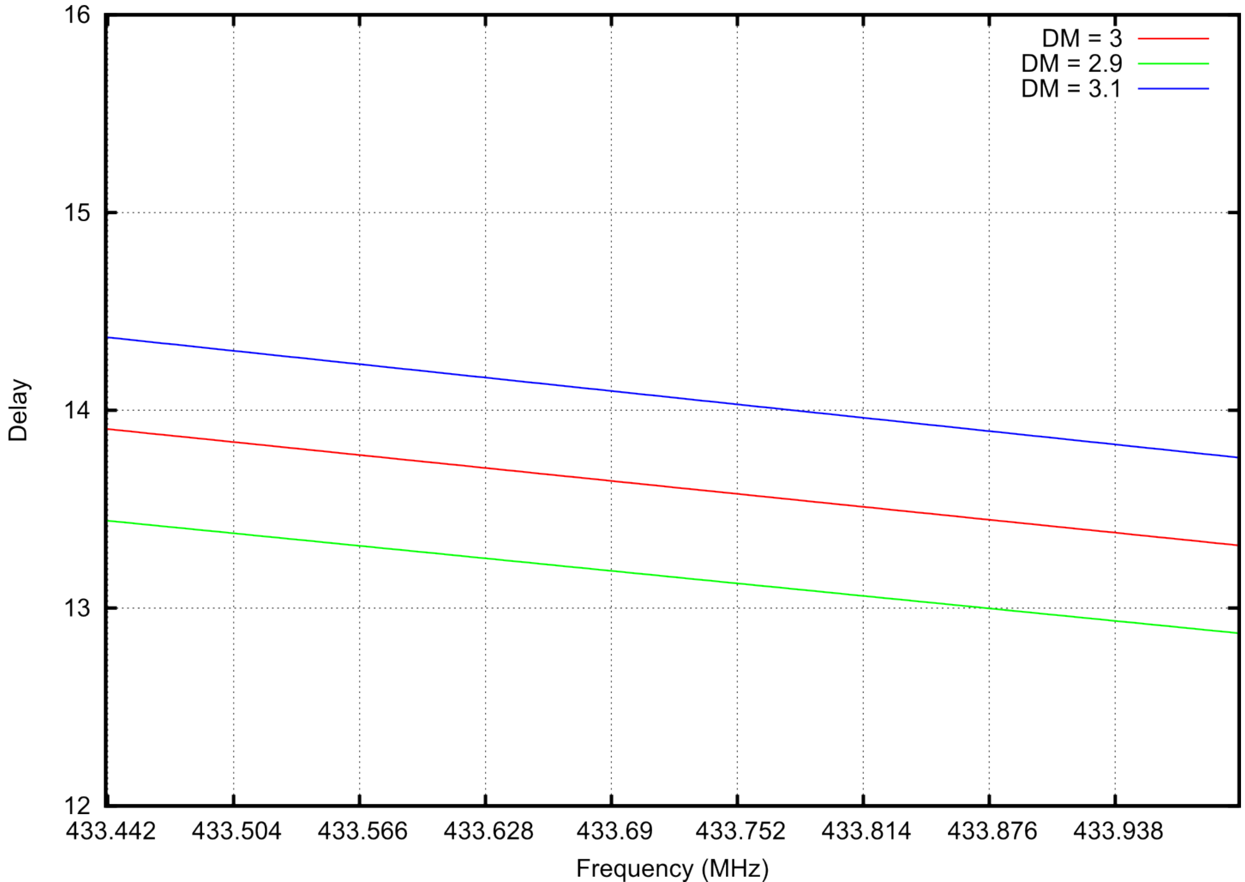
CPU versus GPU memory performance



Supercomputers & Accelerators

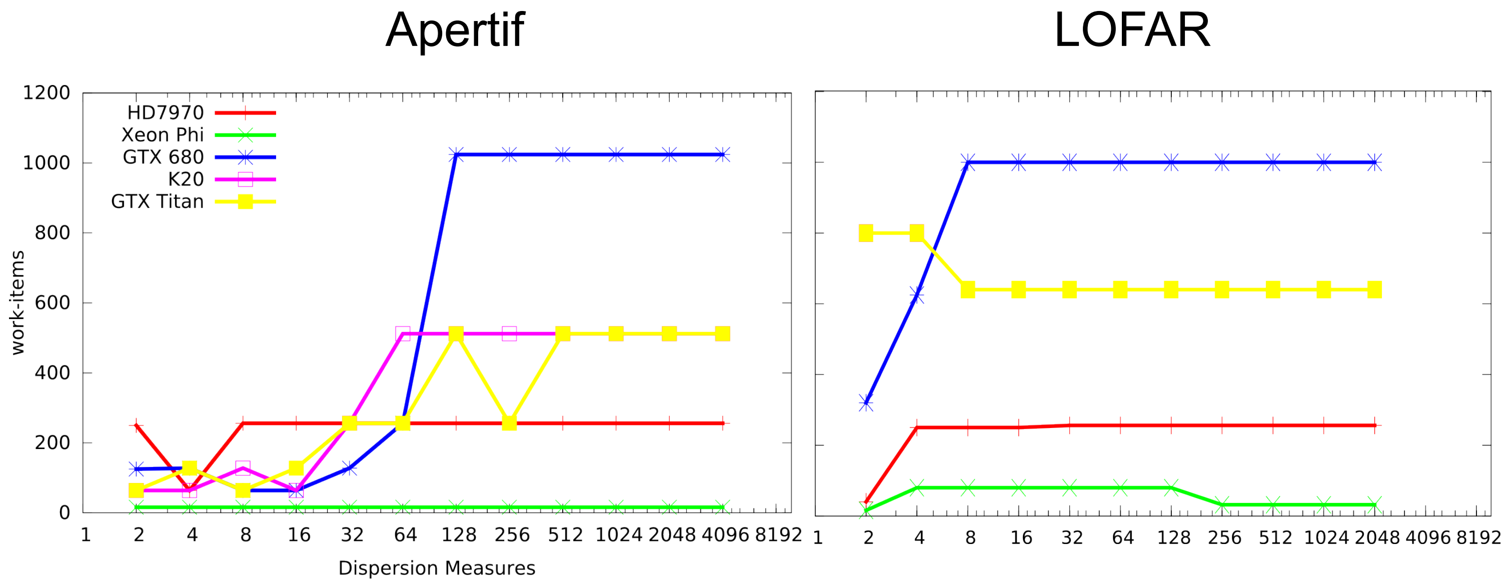


Data reuse

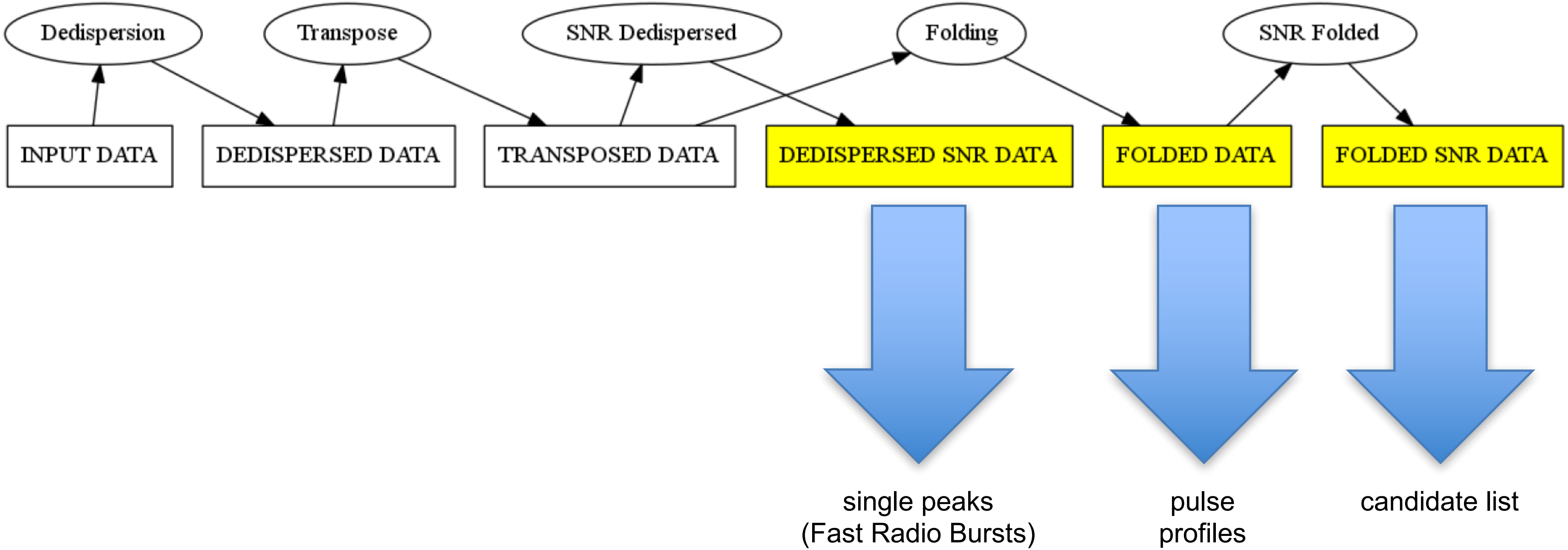


DM = 3
DM = 2.9
DM = 3.1
16
15
14
13
12
Delay
433.442
433.504
433.566
433.628
433.69
433.752
433.814
433.876
433.938
Frequency (MHz)

Auto-Tuning Dedispersion



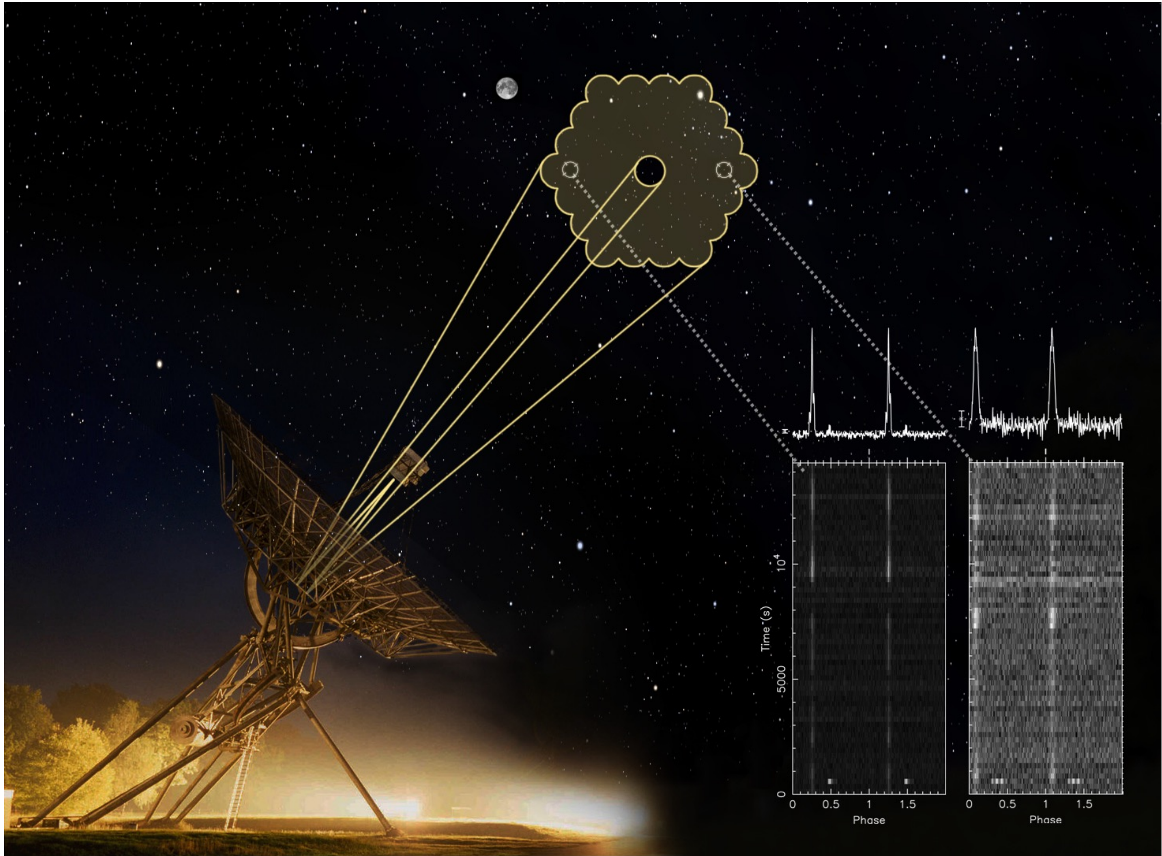
GPU Pulsar pipeline schematic



A Real-Time Radio Transient Pipeline for ARTS

ARTS, the Apertif Radio Transient System, is the system Astron is buiding to find FRBs in real-time

Fast Radio Burst (FRB): “high energy astrophysical phenomenon manifested as a transient radio pulse lasting only a few milliseconds” (Wikipedia)



Alessio Sclocco, Joeri van Leeuwen, Henri E. Bal, Rob V. van Nieuwpoort
Global Conference on Signal & Information Processing 2015

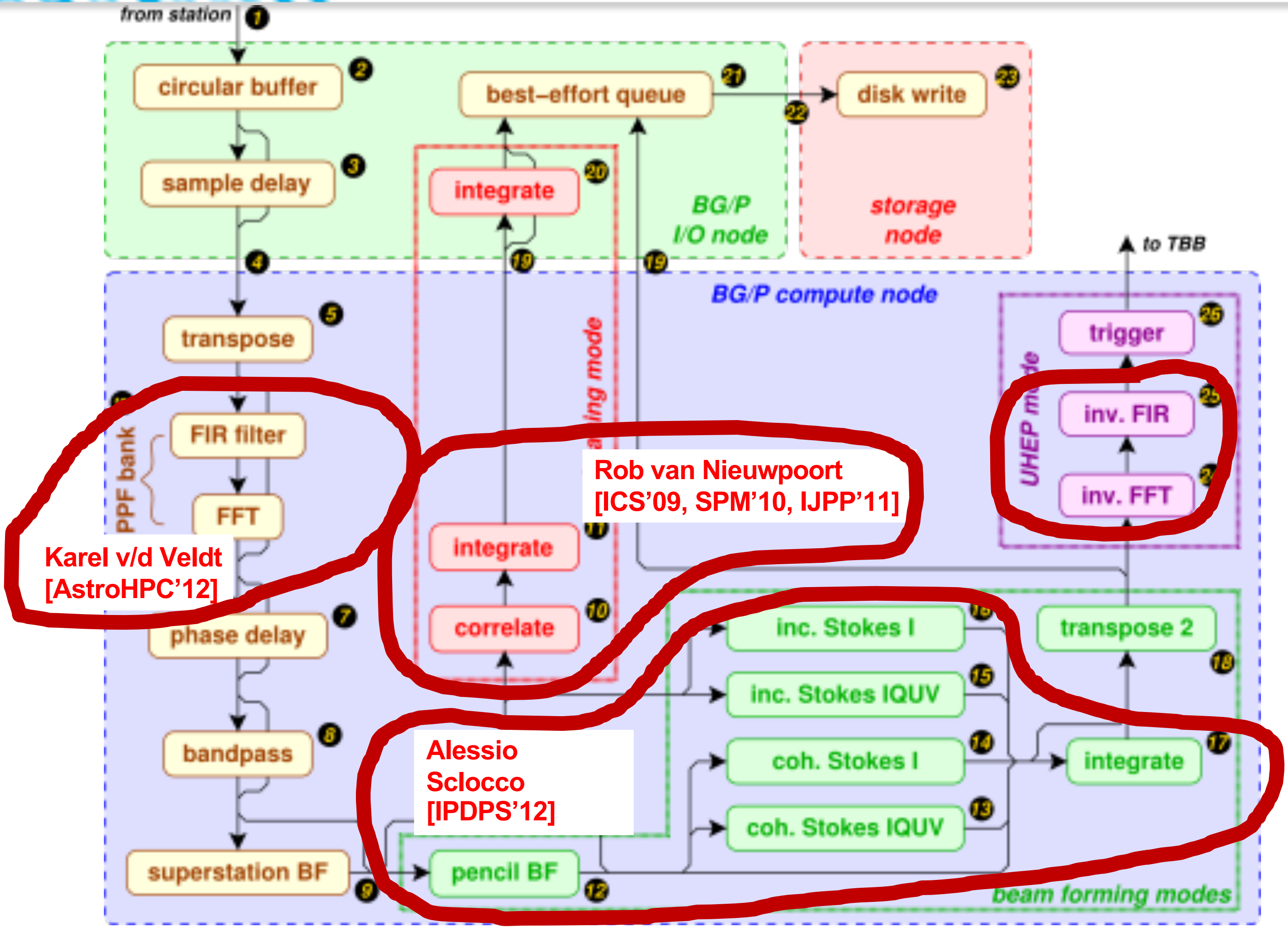
Data reuse

- Data reuse
- Automatically optimize for occupancy
  - (keep compute cores busy)
- Automatically optimize for memory bandwidth

OpenCL: The Khronos group

OpenCL: Open Compute Language

- Architecture independent
- Explicit support for many-cores
- Low-level host API
  - Uses C library, no language extensions
- Separate high-level kernel language
  - Explicit support for vectorization


from station
circular buffer
sample delay
best–effort queue
disk write
integrate
BG/P I/O node
storage node
to TBB
BG/P compute node
transpose
trigger
inv. FIR
inv. FFT
UHEP mode
FIR filter
FFT
PPF bank
Rob van Nieuwpoort [ICS'09, SPM'10, IJPP'11]
Karel v/d Veldt [AstroHPC'12]
integrate
correlate
phase delay
inc. Stokes I
inc. Stokes IQUV
coh. Stokes I
coh. Stokes IQUV
transpose 2
integrate
bandpass
Alessio Sclocco [IPDPS'12]
superstation BF
pencil BF
beam forming modes

eAstronomy peer reviewed Publications

1. Swinbank, J. - Staley, T. - Molenaar, G. - Rol, E. - Rowlinson, A. - Scheers, L.H.A. - et al
   The LOFAR Transients Pipeline
   2015 - Astronomy and Computing, 11, p.25–48 [Journal, NLeSC in acks]
2. Alessio Sclocco, Henri E. Bal, Rob V. van Nieuwpoort
   Finding Pulsars in Real-Time.
   11th IEEE International Conference on eScience, 31 August - 4 September, 2015, Munich, Germany.
3. Alessio Sclocco, Henri E. Bal, Jason Hessels, Joeri van Leeuwen, Rob V. van Nieuwpoort.
   Auto-Tuning Dedispersion for Many-Core Accelerators.
   28th IEEE International Parallel & Distributed Processing Symposium (IPDPS), May 19-23, 2014, Phoenix (Arizona), USA.
4. Alessio Sclocco, Henri E. Bal, Rob V. van Nieuwpoort.
   Real-Time Pulsars Pipeline Using Many-Cores.
   AAS Exascale Radio Astronomy Meeting, 30 March - 4 April, 2014, Monterey (California), USA.
5. Rob V. van Nieuwpoort and the LOFAR team:
   Exascale Real-Time Radio Frequency Interference Mitigation.
   Exascale Radio Astronomy, AAS Topical Conference Series Vol. 2. Proceedings of the conference held 30 March - 4 April, 2014 in Monterey, California. Bulletin of the American Astronomical Society, Vol. 46, #3, #403.01
6. Alessio Sclocco, Rob V. van Nieuwpoort.
   Pulsar Searching with Many-Cores.
   Facing the Multicore-Challenge III, September 19-21, 2012, Stuttgart, Germany.
7. Alessio Sclocco, Ana Lucia Varbanescu, Jan David Mol, Rob V. van Nieuwpoort. Radio Astronomy
   Beam Forming on Many-Core Architectures.
   26th IEEE International Parallel & Distributed Processing Symposium (IPDPS), May 21-25, 2012, Shanghai, China.
8. Alessio Sclocco, Joeri van Leeuwen, Henri E. Bal, Rob V. van Nieuwpoort:
   A Real-Time Radio Transient Pipeline for ARTS.
   IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE Signal Processing Society. Held in Orlando, Florida, USA, December 14-16, 2015.
9. P. Chris Broekema, Rob V. van Nieuwpoort and Henri E. Bal:
   The Square Kilometre Array Science Data Processor Preliminary Compute Platform Design.
   Journal of Instrumentation, Volume 10, July 2015.
10. Alessio Sclocco, Joeri van Leeuwen, Henri E. Bal, Rob V. van Nieuwpoort:
    Real-Time Dedispersion for Fast Radio Transient Surveys, using Auto Tuning on Many-Core Accelerators.
    accepted for publication in Astronomy and Computing, 2016

Masters theses

- **Rene van Klink:**
  In progress
  Working title: Auto-tuning memory layouts
  Vrije Universiteit Amsterdam,
  Netherlands eScience Center, 2016.

- **Linus Schoemaker:**
  Removing Radio Frequency Interference in the LOFAR using GPUs.
  Vrije Universiteit Amsterdam,
  Netherlands eScience Center, 2015.

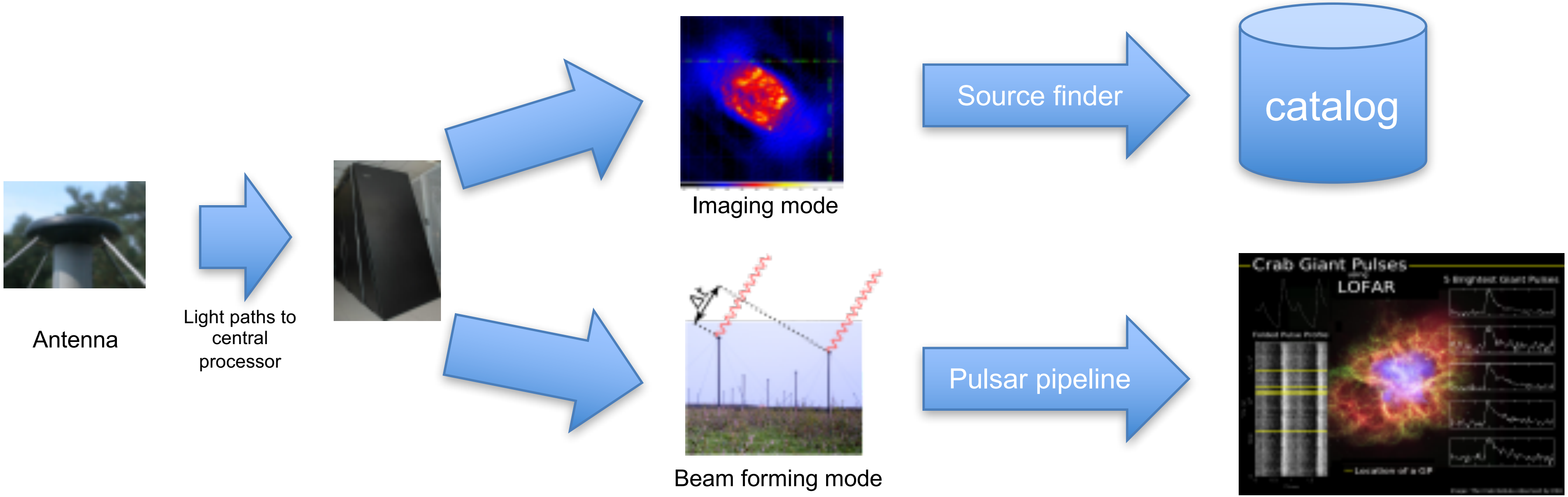
- **Jan Kis:**
  Auto-tuning a LOFAR radio astronomy pipeline in JavaCL.
  Vrije Universiteit Amsterdam,
  Netherlands eScience Center, 2013.

Grants / “spinoffs”

PI	Title	Call	Amount (K euro)
Martin Kersten	Big data for the big bang	Dome	200
Martin Kersten	Compressing the sky	NLeSC	50
Joeri van Leeuwen	ARTS — the Apertif Radio Transient System	NOVA	730
Joeri van Leeuwen	ARTS — the Apertif Radio Transient System	NWO-M	590 + 540 matching
John Romein	Radio-Telescope Algorithms for Many-Core Processor Architectures	NWO Open Competitie	140

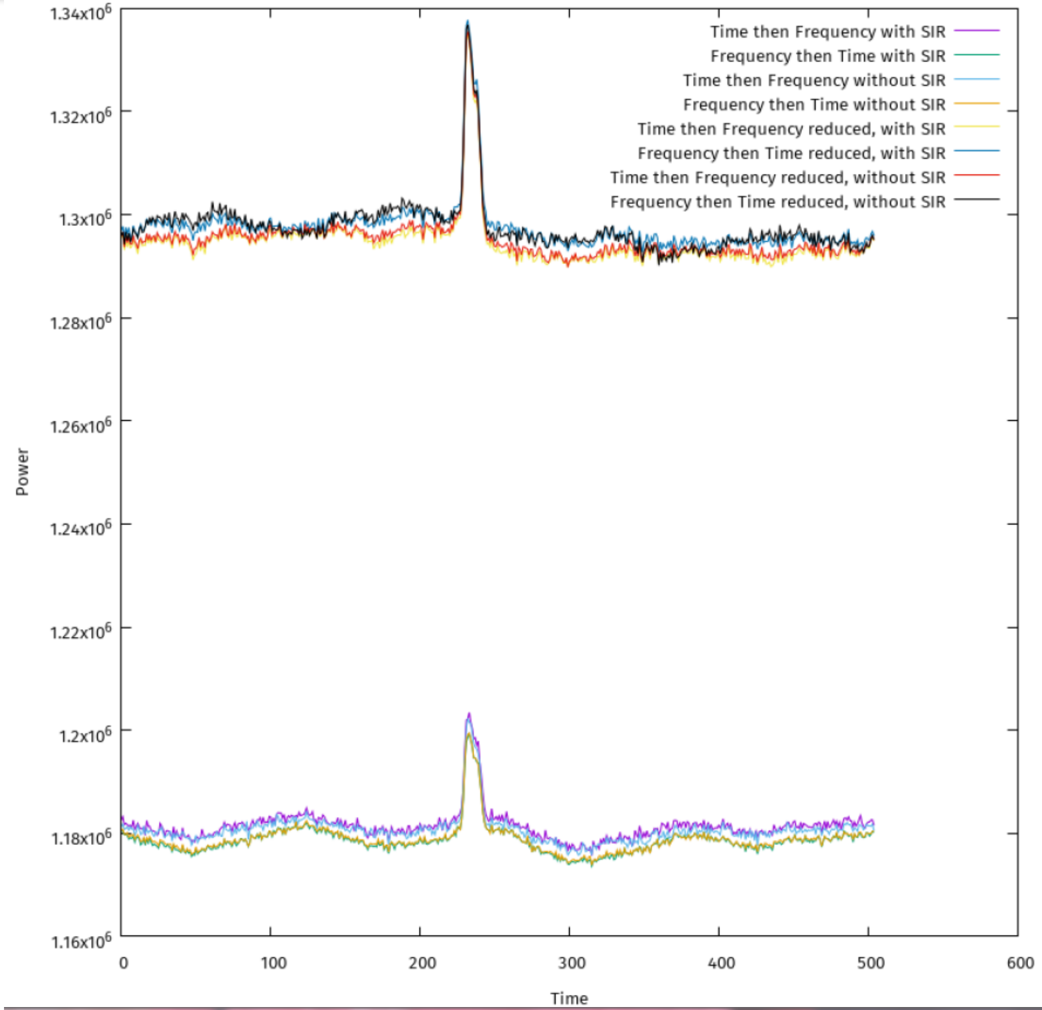
Three LOFAR pipelines

Transient pipeline:
imaging in real-time

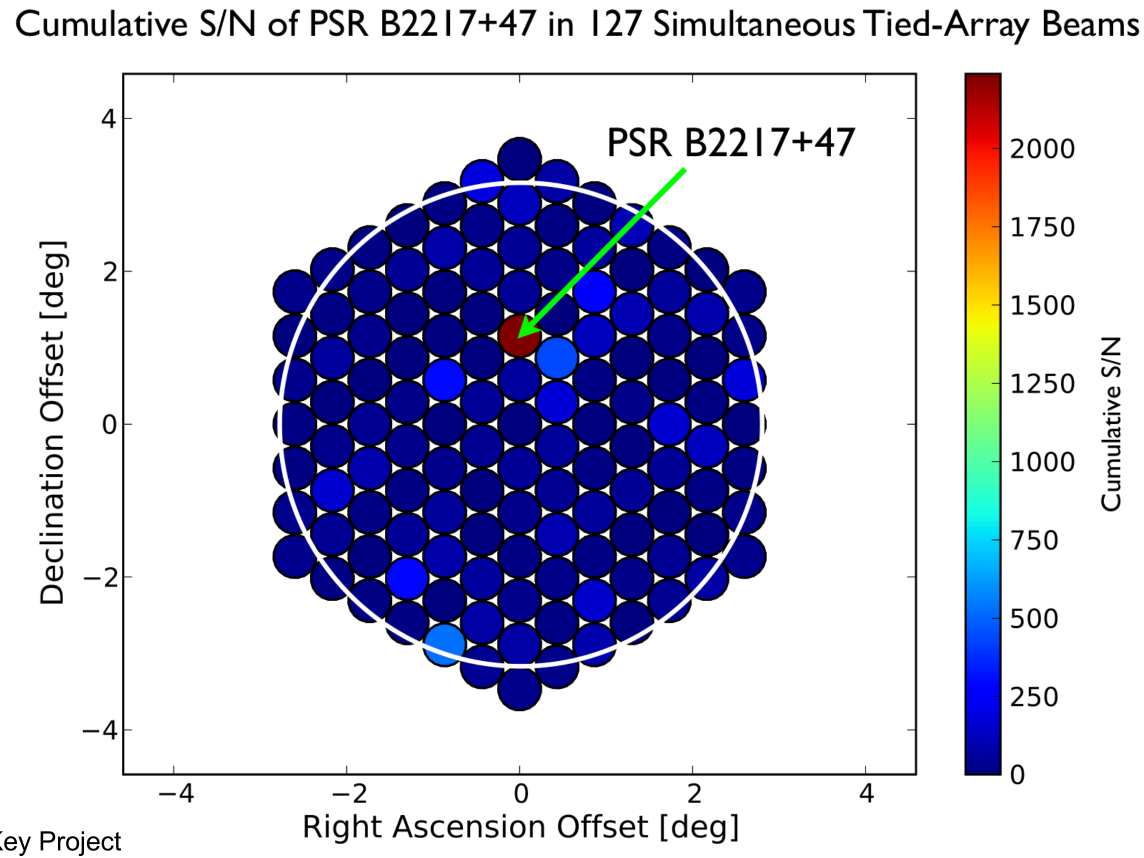


RFI GPU Results

- Ported to GPUs
- Up to 200 LOFAR stations in real time on a single GPU



127-beam tied-array observation using the LOFAR Superterp



Courtesy LOFAR Transients Key Project