ESNW

e-Science North West

Pulsar Astronomy on the Grid 1999-2006

E-Infrastructure Workshop

March 20 2006

John Brooke

(thanks to Jodrell Bank and the SKA projects for contributions of material)

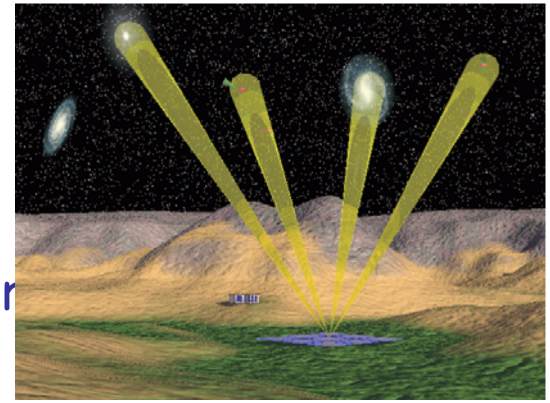
Radio-astronomy of 2015
Square Kilometre Array:
its science and technology

The SKA in a nutshell

- An extremely powerful survey telescope at metres to cm with the capability to follow up individual objects with high angular and time resolution
- ~ 1 km^2 collecting area
  - limited gains achievable by reducing receiver noise – need more microwave photons
- Frequency range 0.1 – 25 GHz (goal)
- 15-country international collaboration
    - executive, science, engineering, siting, simulation groups
- International funding
  - Cost goal ~ €/$ 1 billion
- Site selection in 2006; technology selection in 2008; initial operations 2015; full operations 2020

SKA: what's new?

- Sensitivity → ~ 50 x EVLA
- Large Field of View (FOV): 1 sq. deg. at 1.4 GHz
- Goal of multi-beam instrument, at least at lower frequencies
  - Re-use area: 2 < M < 8 looks feasible
    - operational and science advantage
- Innovative design to reduce cost
  - 10^6 sq metre → ~$1000 / sq metre
    - cf existing arrays ~$10000 / sq metr



Key Science Projects

1) the evolution of galaxies and large scale structure in the universe

2) probing the dark ages before the universe lit up

3) strong field tests of gravity using pulsars and black holes

4) the origin and evolution of cosmic magnetism

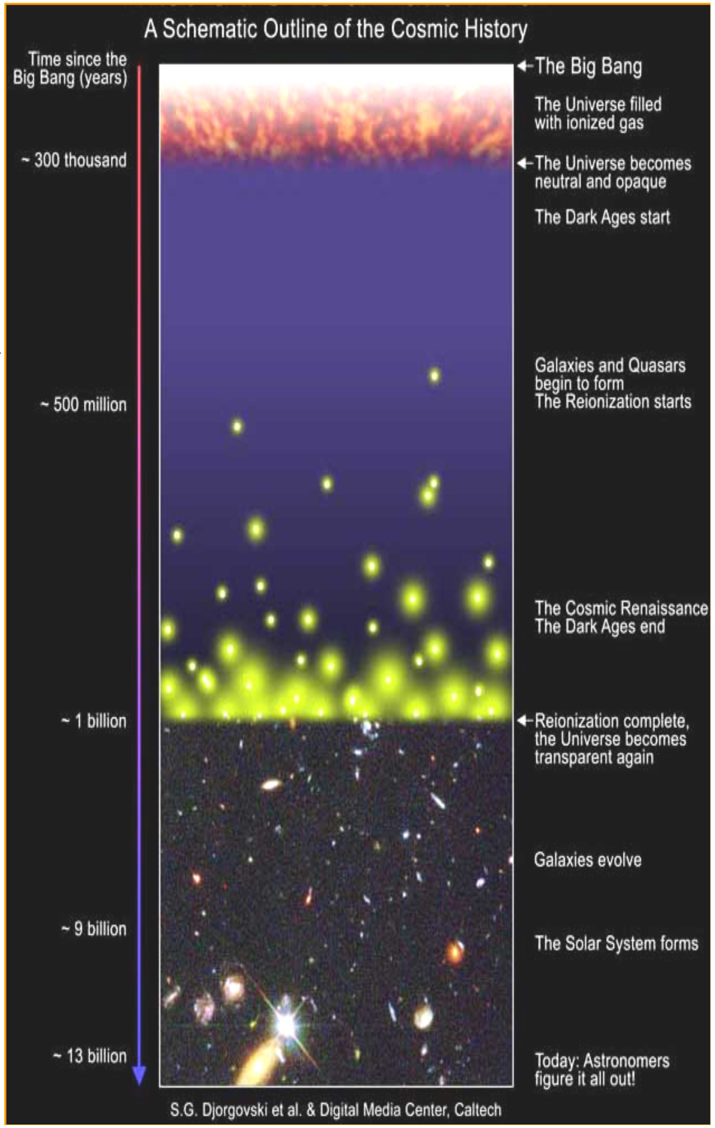
5) the cradle of life

+

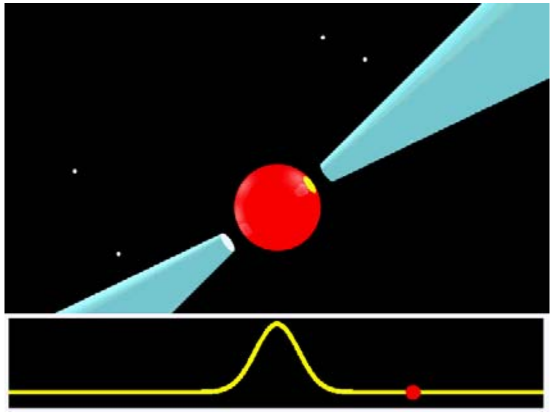
6) exploration of the unknown

•science case published in 2004

•science requirements on technology identified



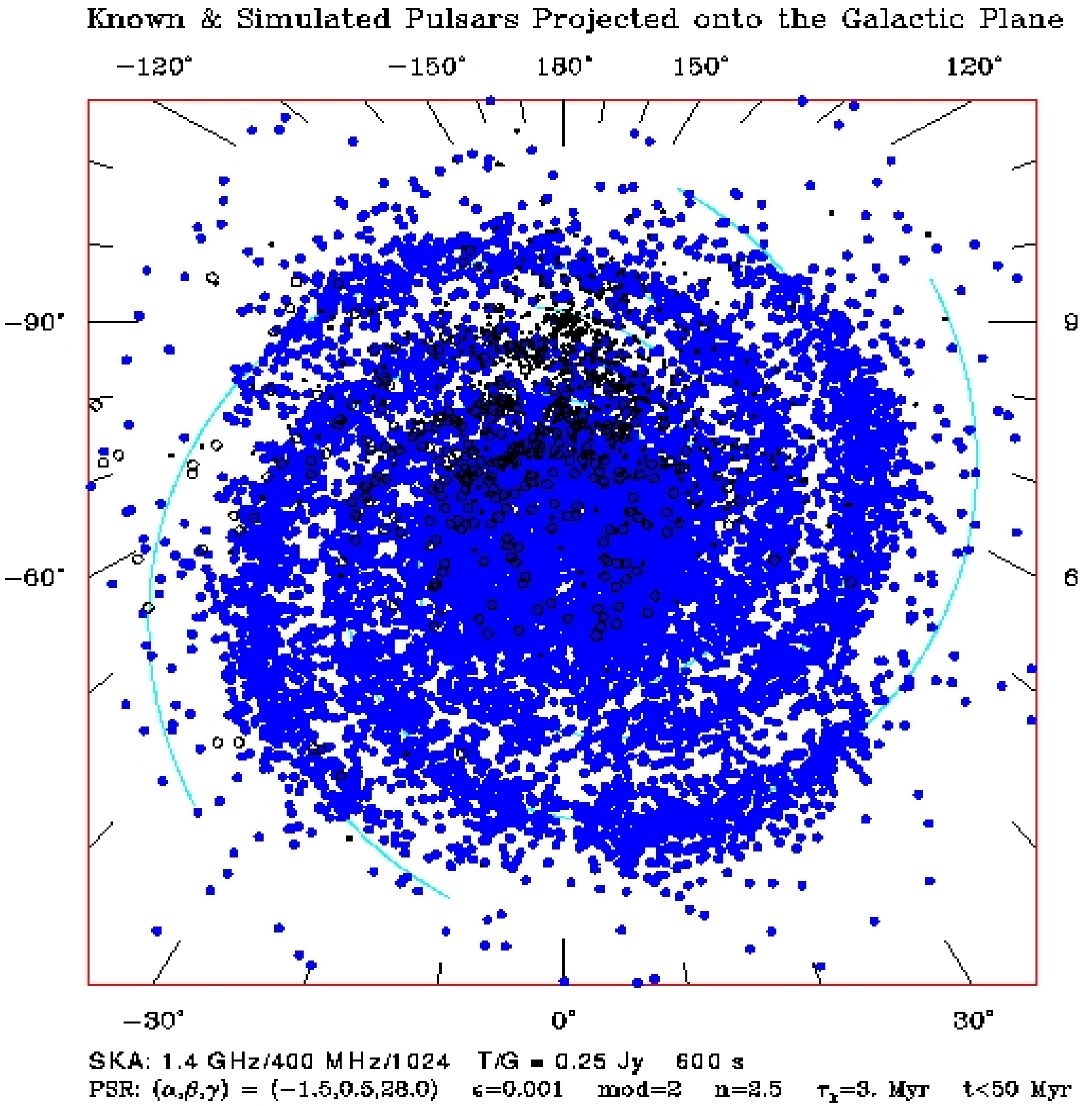
Pulsars…



- …almost Black Holes
- …objects of extreme matter
  - 10x nuclear density
  - $B \sim B_q = 4.4 \times 10^{13}$ Gauss
  - Voltage drops ~ 10^{12} volts
  - $F_{EM} = 10^9 F_g = 10^{11} F_{gEarth}$
  - High-temperature superfluid & superconductor
- …relativistic plasma physics in action
- …probes of turbulent and magnetized ISM
- …precision tools, e.g.
  - Period of B1937+21: **P = 0.0015578064924327±0.0000000000000004 s**
  - **Orbital eccentricity of J1012+5307: e<0.0000008**
- …testing ground for theories of gravity
- …cosmological gravitational wave detectors

From Michael Kramer

SKA pulsar survey

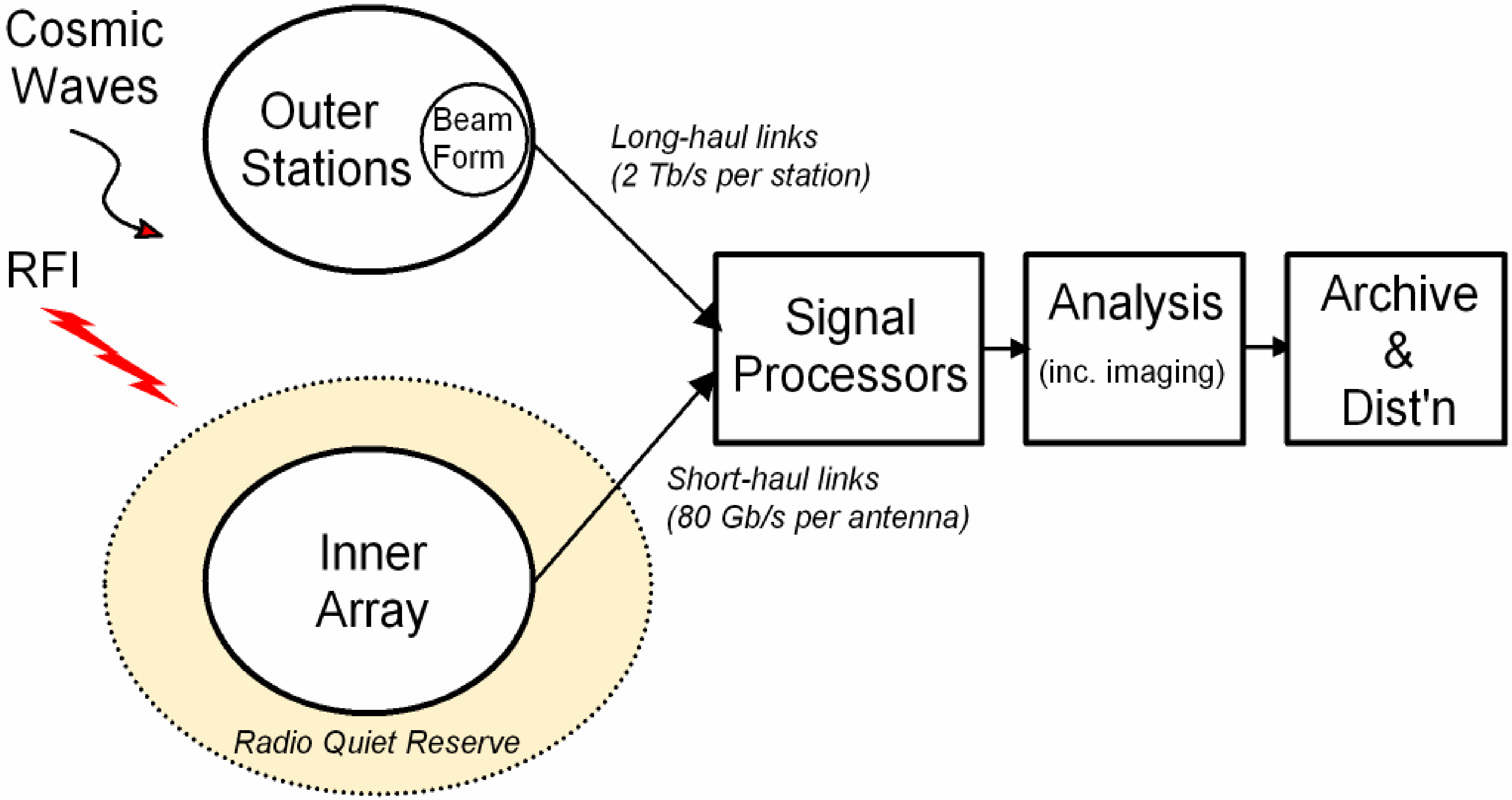


expect

~20000 psr's

incl. 1000 MSPs and BH-pulsar binaries

SKA Concept



from Peter Hall

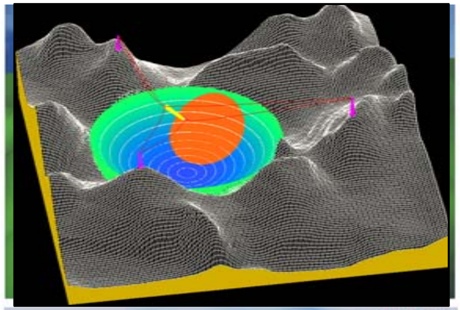
Cost Reduction Strategies

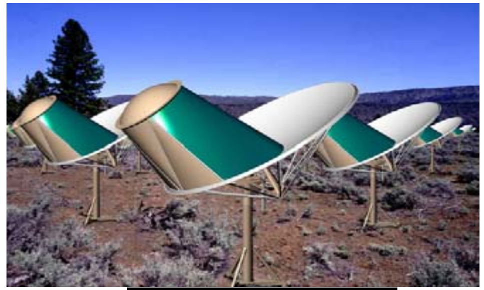
- Design a software telescope
  - Exploit convergence of radio and computing technologies, replacing hardware with firmware or software
  - But antennas still account for ~ 40% of array cost
- Exploit emerging technologies
- Learn from industry
  - Mass production is new to astronomy
- Plan evolution of SKA capability
  - eg ultimate signal processing capacity will not exist in 2015

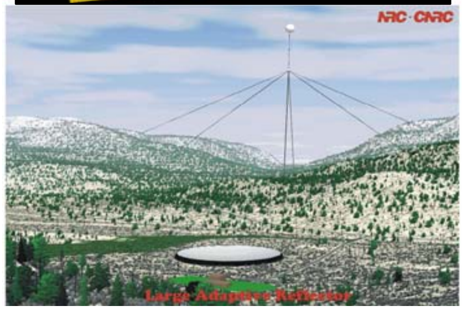
Antennas

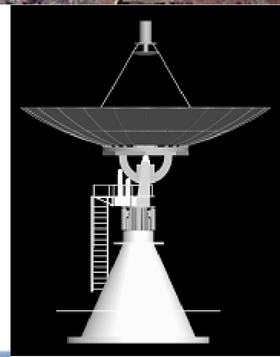
- Wide range of possible solutions
  - Aperture phased arrays
  - Flux concentrators (reflecting dishes, cylinders)

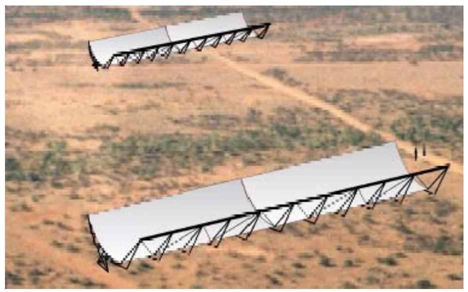
- SKA will likely use at least two antenna types
  - Cost effective high frequency solutions don't provide enough area at low frequencies
  - Want good efficiency at high frequency AND multi-fielding (or at least wide field of view) at low frequency

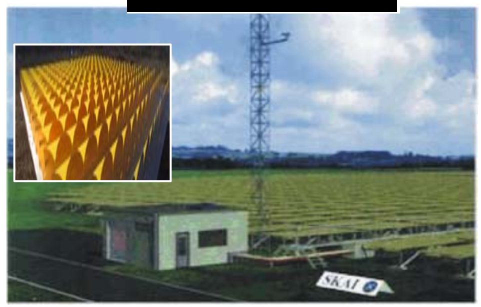






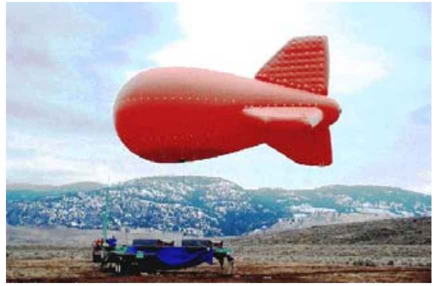






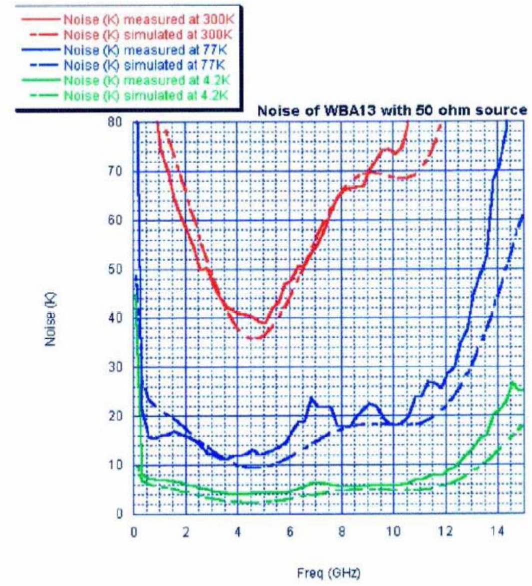
Antenna Innovations

- Low-cost dense arrays for aperture and focal planes
- Active surfaces for large reflectors
- Broadband feeds
- Suspended or airborne inertial feed platforms
- Cheap, accurate 12m dishes using hydroforming or preloading

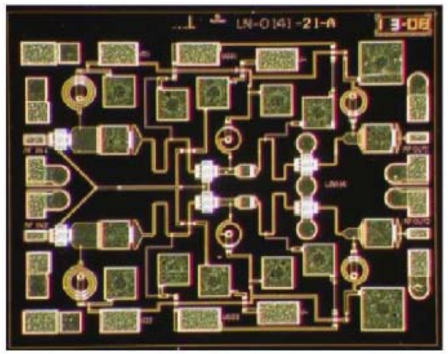


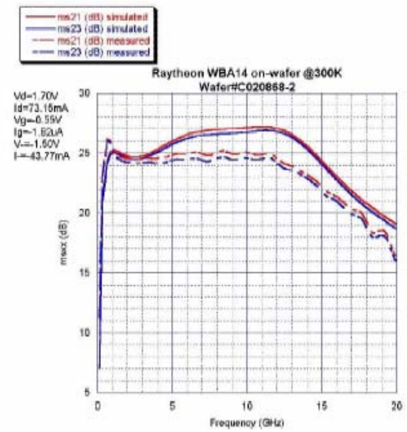
Receivers

- SKA receivers will be more like commercial systems than “conventional” astronomy receptors
- Wideband (e.g. 1 – 10 GHz)
- Low noise
- New, high-reliability cryo-coolers being tested for some SKA concepts



S. Weinreb & USSKAC





Data Transport

- High data rates
  - 1-2 Tb/s from stations desirable
  - 80 Gb/s from individual antennas in central array
  - Probably will need to compromise at a "commercially realistic" rate of ~ 100 Gb/s for longer links
- Digital fibre links throughout array
  - Dense wavelength division multiplexing on long links (100s of channels)
  - Aim for commercial compatibility on long links (e.g. ITU standards)
  - Use emerging astronomy standards on short links (better bandwidth efficiency)
- Information *transport* costs may dominate *processing* costs
- Local oscillator/timing is a challenge for a highly-distributed array

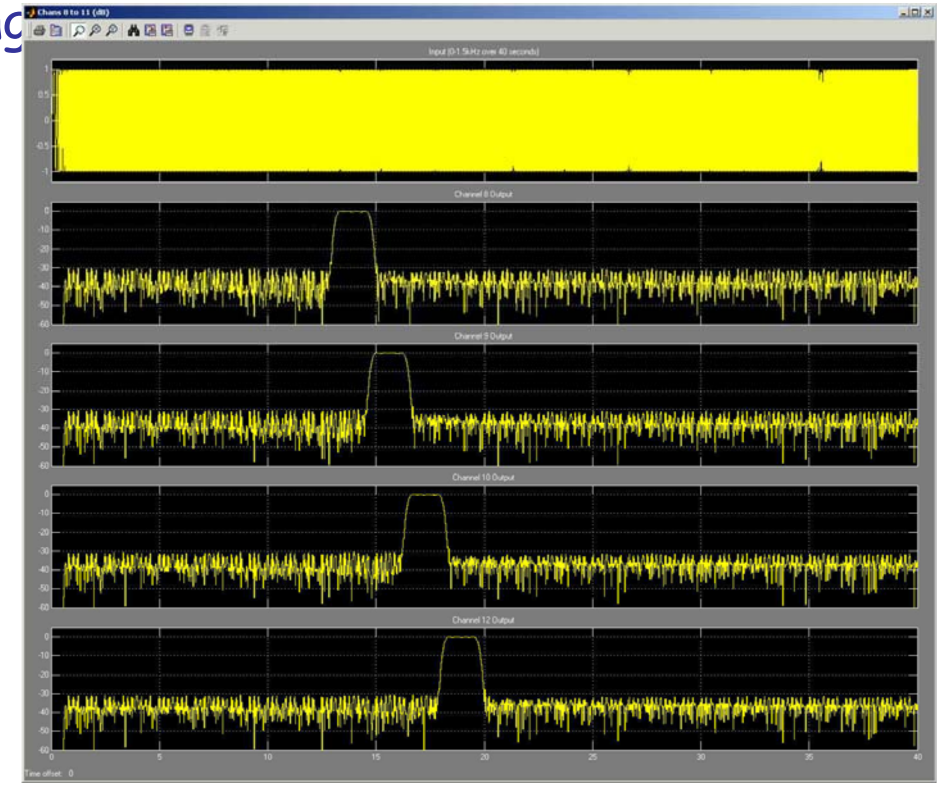
Signal Processing

- Reliant on Moore's Law extending beyond 2015

- Typical correlator
  - 2500 inputs, 4 GHz BW, 8-bit processing → n x 10^{15} ops per sec

- Extensive station-based processing (needed for calibration)

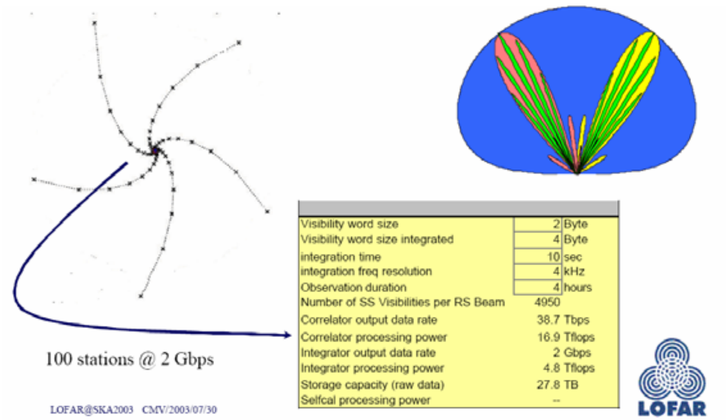
- Non-correlation processing for pulsars and transients required

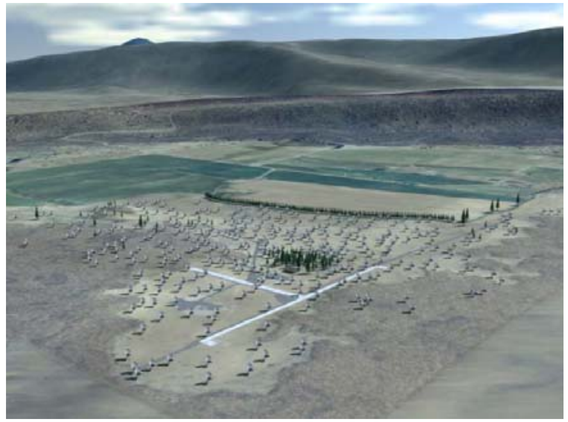


Post-processing

- Correlator output data rate can be very high
  - Typically 10^8 correlations per second
    - Most likely have post-processor at central site
- Computer capacity in 2015 may be inadequate but initial data rate can be reduced as necessary by reducing number of entities correlated
- 2015 capability should not unduly influence initial infrastructure design
  =>fibre should be adequate for 30-year SKA evolution
- Effective sharing and retrieval of SKA data will be a major challenge
  - Extensive internet connectivity is mandatory
  - Links to 'virtual observatory'

Large SKA Demonstrators

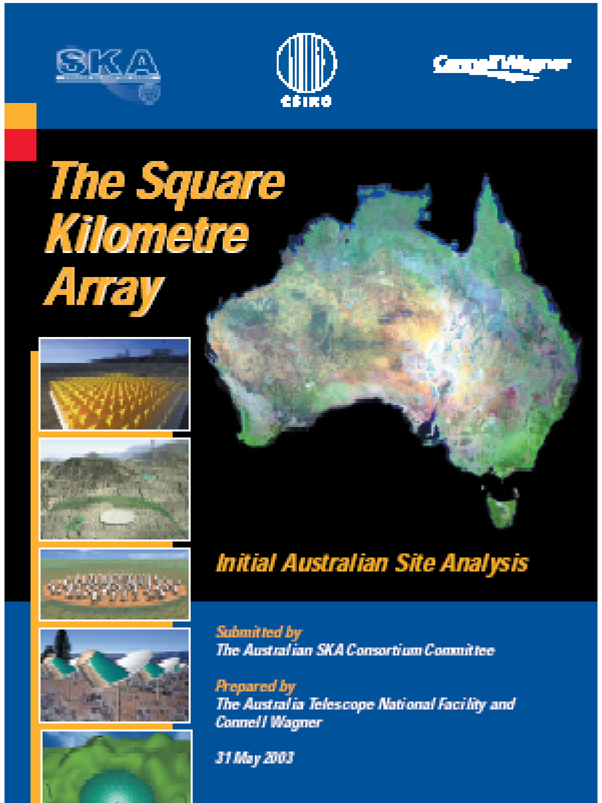


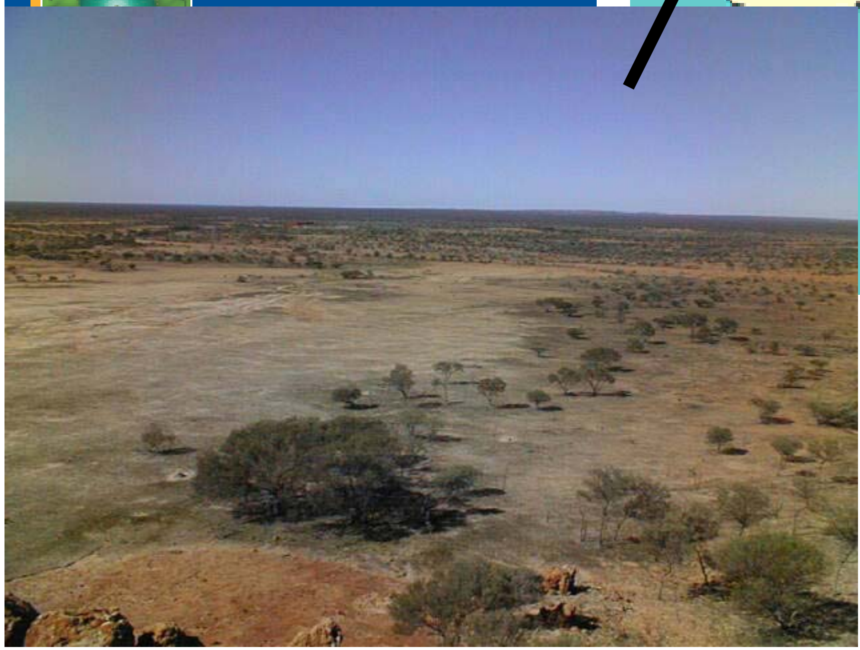


LOFAR

ATA

+ EVLA + DSN + US/TDP + EU/FP6 + …





SKA in Australia

Radio astronomy - research challenges

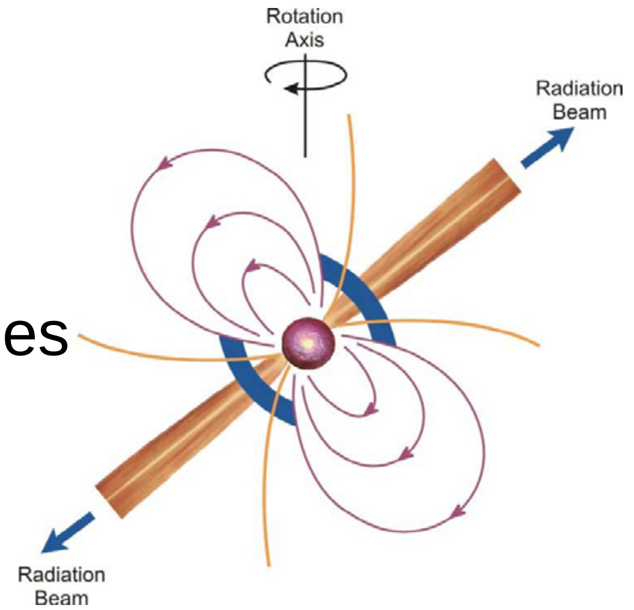
- System design - balance of software and hardware capabilities. What insights for Grid computing research?
- High performance networking linked to large processing requirements, challenges Grid research - ESLEA project.
- Image analysis and visualization, very high resolution possible by long baseline interferometry.
- Mathematics and algorithms for signal processing, more on this in next section.

Example of collaboration: pulsar signal processing via distributed computing 1999-2006

The University of Manchester
Jodrell Bank Observatory

MANCHESTER 1824

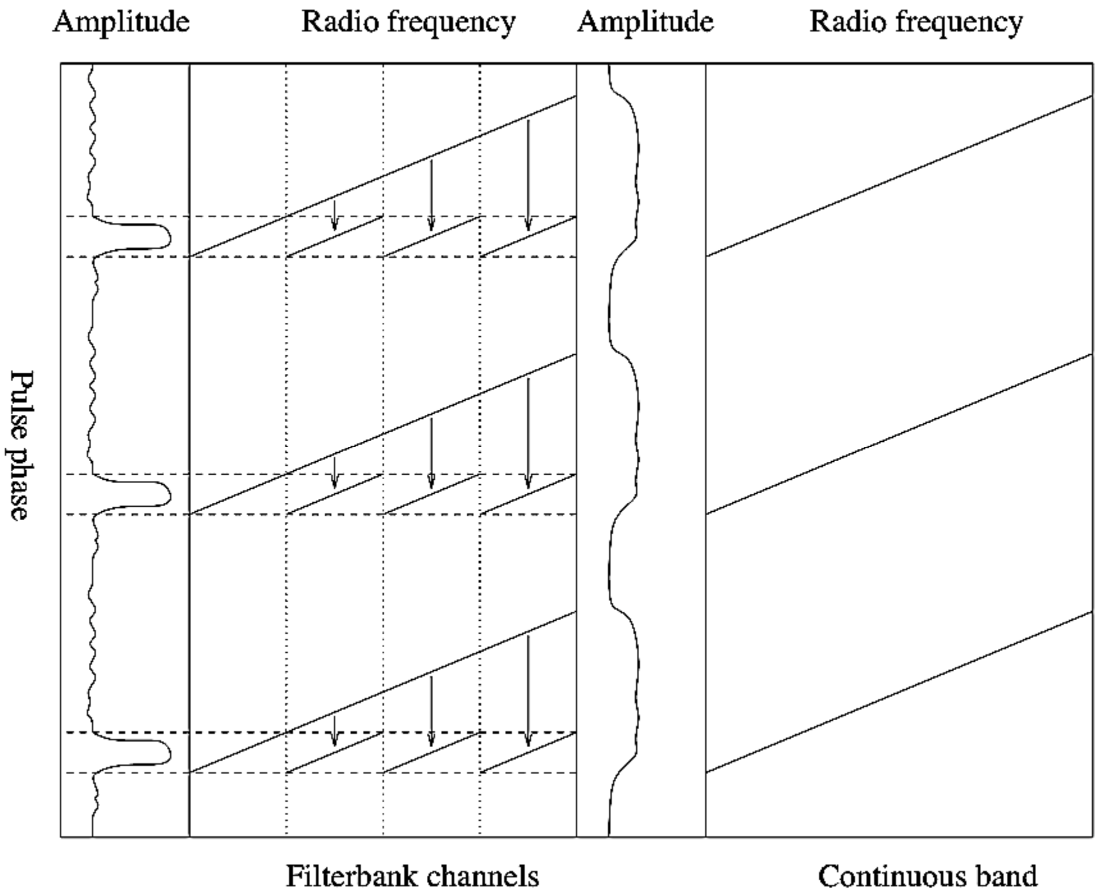
Pulsar Signals



- Pulsar signals are very weak
- Have a steep spectral index (~1.6)
- Emit radio signals across all frequencies simultaneously
- Pulses don't arrive at the same time for all frequencies
- Propagation speed is a function of frequency through ISM

$$v_g \approx c\left(1 - \frac{n_e e^2}{2\pi m \nu^2}\right)$$

Correcting the effects of interstellar dispersion



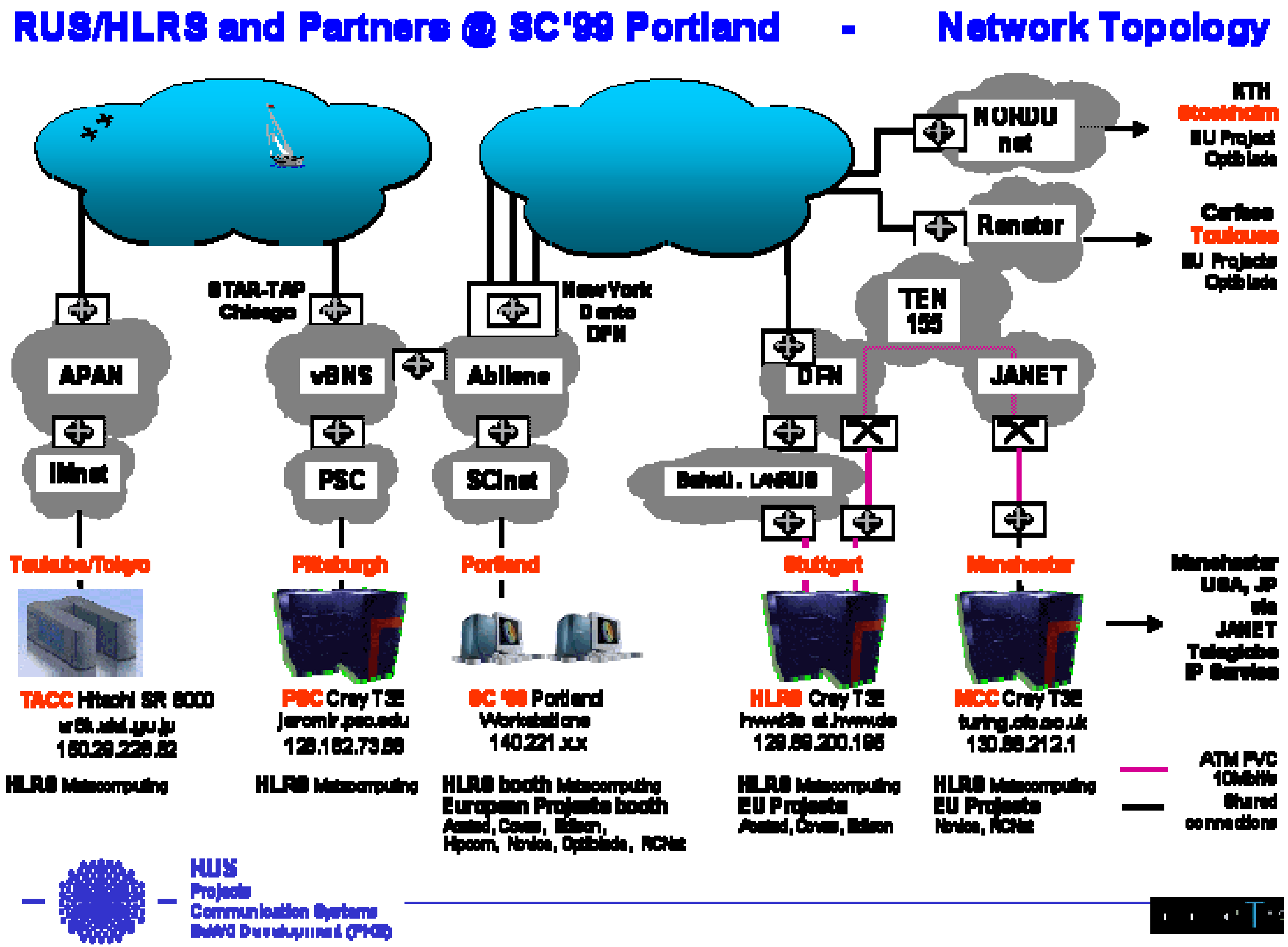
The Jodrell Bank De-Dispersion Code

- Jodrell Bank Radio Observatory searches for fast pulsars in binary systems.
- In order to obtain a high signal-to-noise ratio they observe across a wide band of radio frequencies and integrate.
- However the interstellar medium is dispersive (the speed of the radio waves depends on frequency)
- If the dispersion measure (DM) is known, coherent de-dispersion techniques can be used to correct for the dispersive effect.

Searching for pulsars

- A globular cluster comprises many stars in a compact volume. This means that the dispersion measures of all objects in the cluster should lie within a narrow range.
- If we know (or can estimate) the dispersion measure for one object in the cluster, we can search the cluster for pulsars by
- applying the same de-dispersion techniques at many trial dispersion measures, and
- examining the reconstructed signals.

Metacomputing

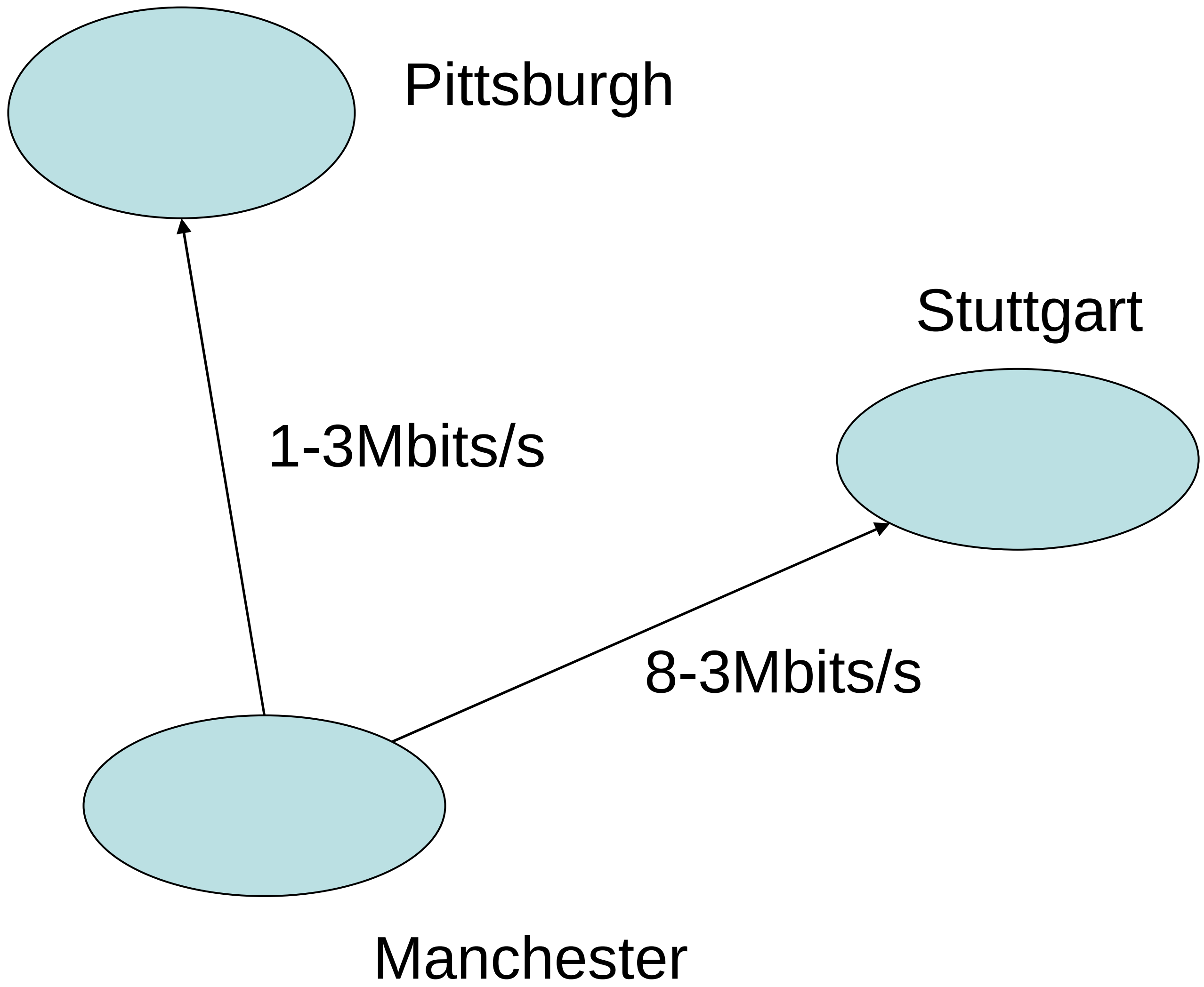


Machines used in the SC2000 global computer.

- Cray T3E-1200 at Manchester, UK
  - 576 application PEs (since upgraded to 788)
- Cray T3E-900 at Stuttgart, Germany 512
  - 512 application PEs
- Cray T3E-900 at Pittsburgh, USA
  - 512 application PEs.

- We tried to include a T3E at JAERI but the bandwidth was too low to make it feasible. Latency is not an issue in these experiments.

Configuration 1999/2000



Data handling and transmission

- All 10 Gigabytes of input data initially reside on the Manchester T3E. This represents roughly one hour of observations of the globular cluster M80.
- Processors on all hosts, both local and remote, are enlisted to reduce the turnaround time.
- During the run, huge volumes of data are transferred across the network to remote processors. We tried to use just enough remote processors to saturate the available bandwidth.
- At the end of the run, multiple transformed time series are sent back by FTP for post-processing in Manchester.

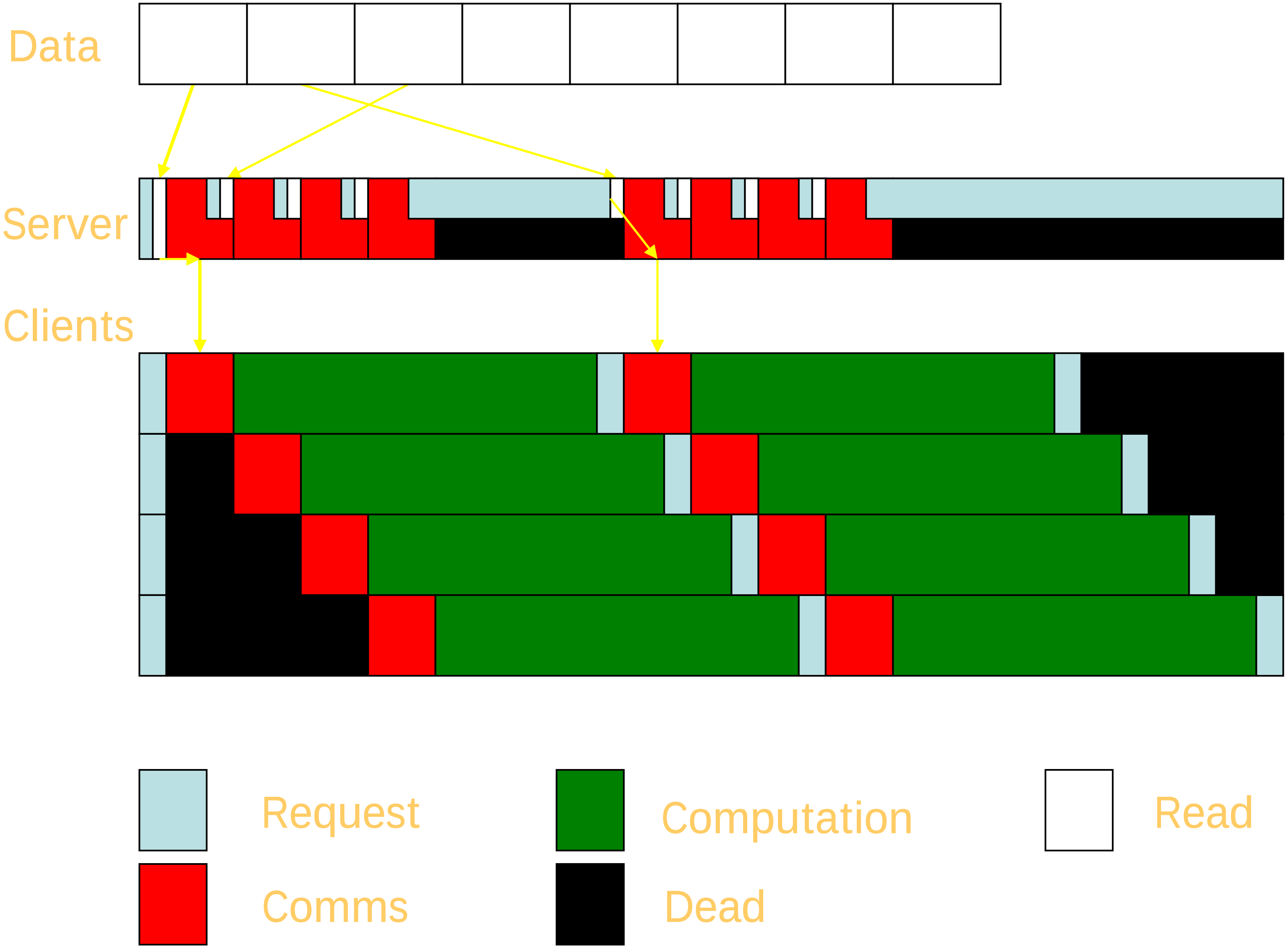
PACX-MPI - heterogeneous clusters

- In order to adapt our application to a metacomputer consisting of several coupled MPPs, we wanted a message passing harness that:
- makes efficient use of inter-system network bandwidth,
- is plug compatible with MPI on a single MPP,
- permits testing (including multiple host emulation) on a single MPP,
- does not require installation of additional software on each system, apart from the vendor's MPI and the application binary.
- We chose PACX-MPI, developed at HLRS (Stuttgart).

Data movement

- We adopt a client-server model to facilitate the movement of data.
- A handful of processors on the local host act as servers. These are responsible for accessing the input data on disk and sending it on to clients.
- The actual work is performed by client processors on possibly remote hosts. Whenever a client needs input data, it sends a request to its allocated server, and awaits the reply.
- Clients write their results to local filestore, for subsequent collection and analysis.

Client-server



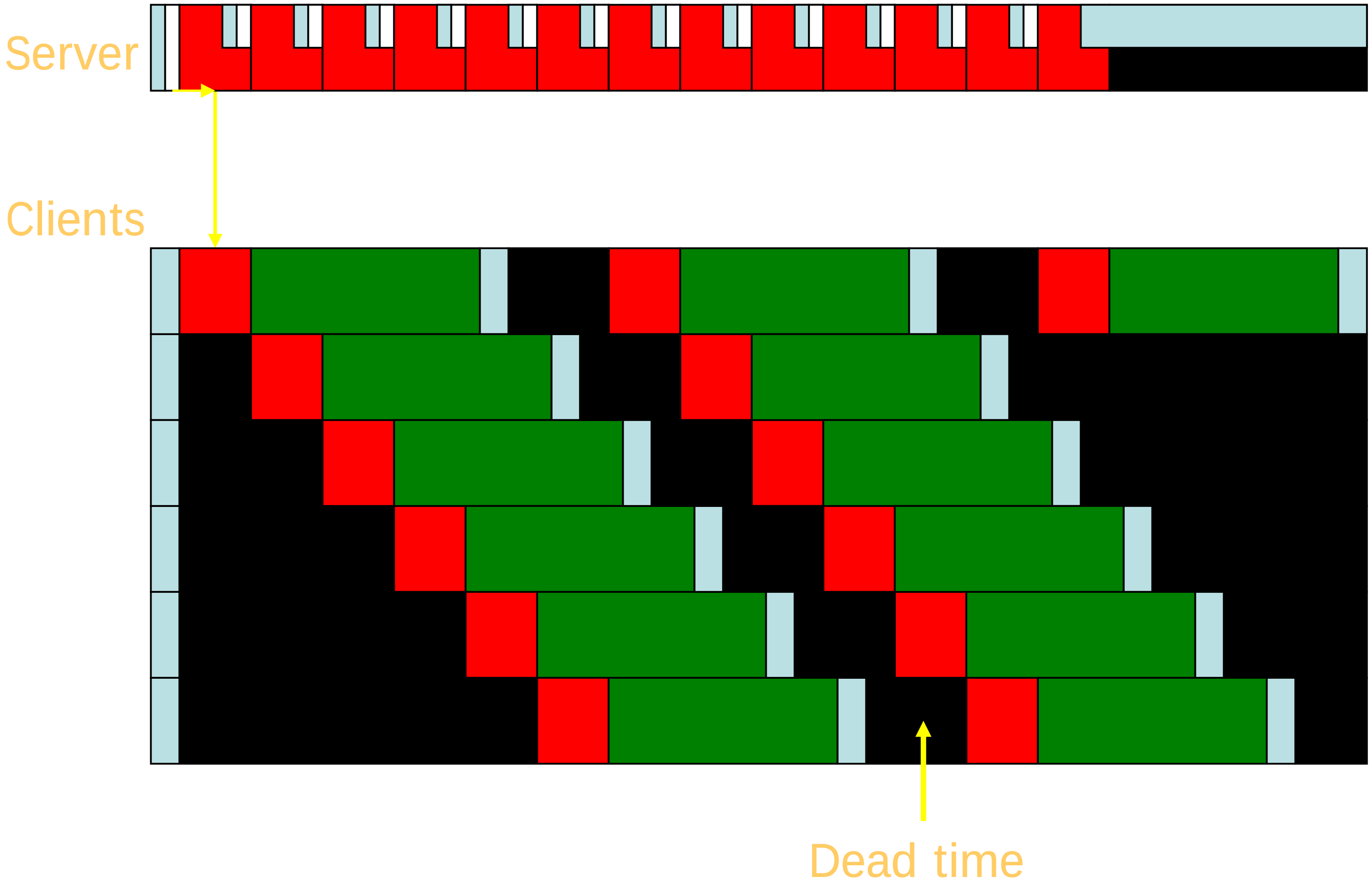
Static load balancing

- The problem is to determine what fraction of input-data should be allocated to each host, so that all hosts finish at the same time.
- We determine this at run-time, based on estimates of inter-host bandwidth, peak performance and number of PEs in each host, size of the data-set and measured time to process one work unit on the local host.
- Provided that the inter-host bandwidth is not saturated, we get a linear system which can be solved for the total elapsed time and the fractions to allocate to each host.

Dead time

- The amount of data that can be processed on a remote host is limited by the bandwidth to that host.
- This in turn limits the number of processors on a given host that can be fully utilised.
- We therefore try to avoid using more remote processors than necessary.
- If we get this wrong, the symptom will be “dead time” between work units: one client waiting for other clients to receive their data.

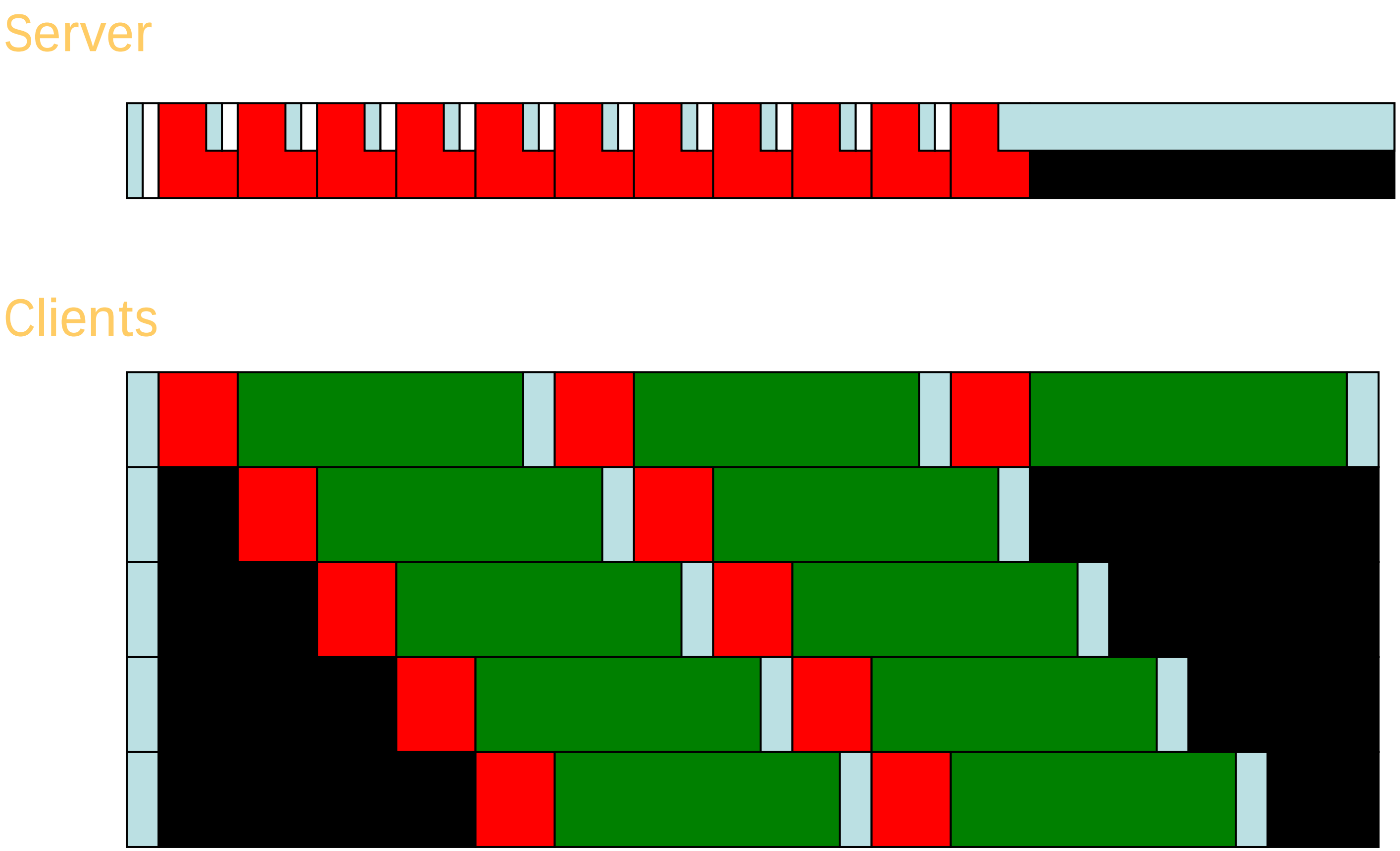
Client-Server: dead time



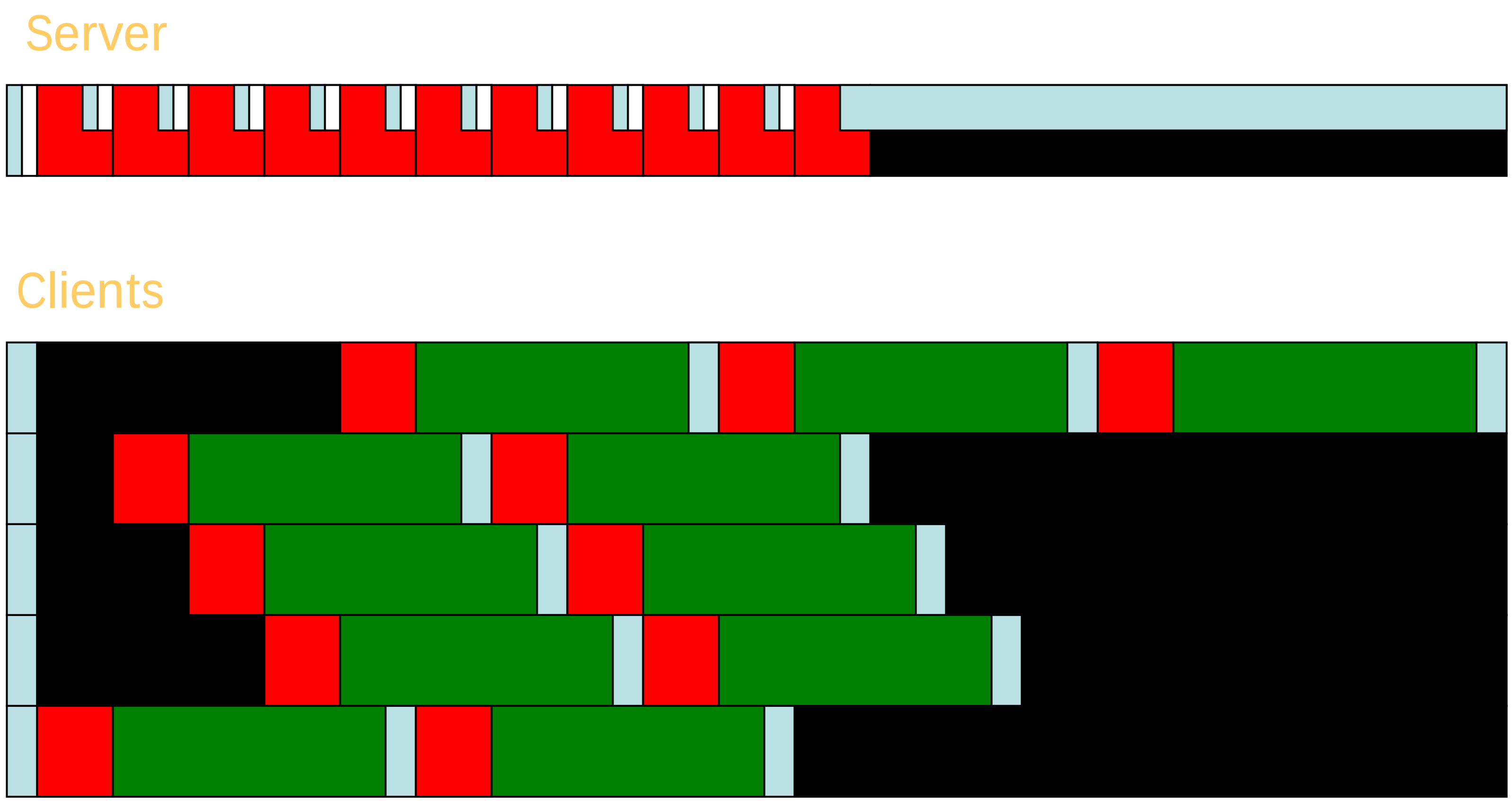
Run down time

- As implemented at SC99, our client-server method exhibited excessive run-down times.
- This is because: some clients get one more work unit than others on the same host, servers process requests in FIFO order.
- We cured this by imposing a schedule on the servers, incidentally eliminating all requests!
- We call the improved method “scheduled service”.

Client-Server: run-down time

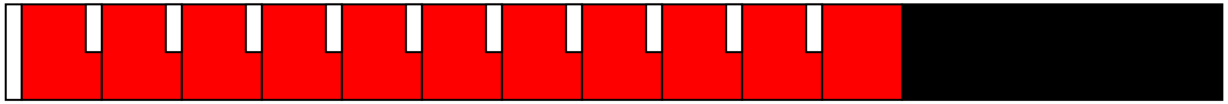


Client-Server: run-down time II

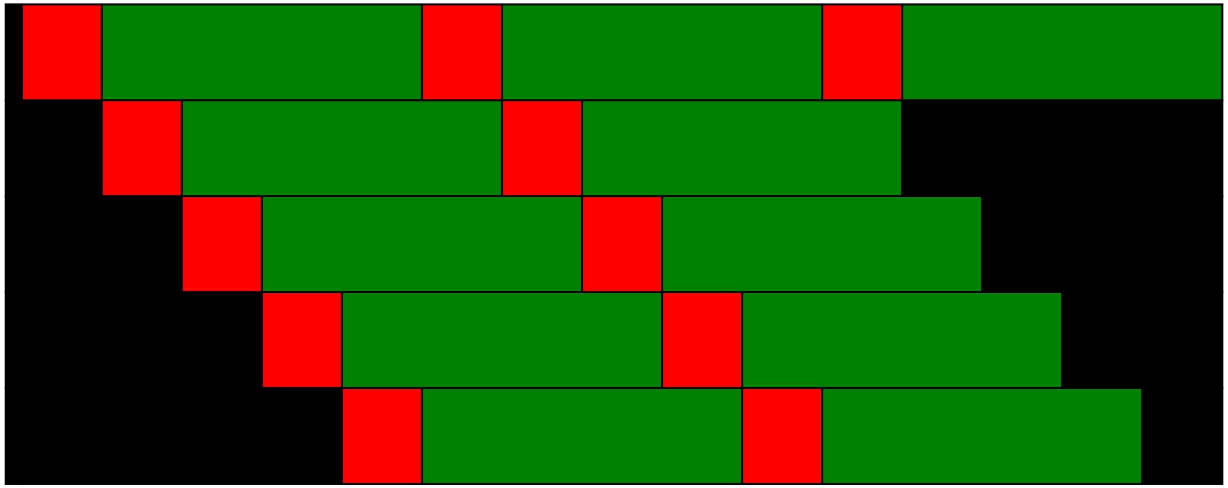


Scheduled Service

Server



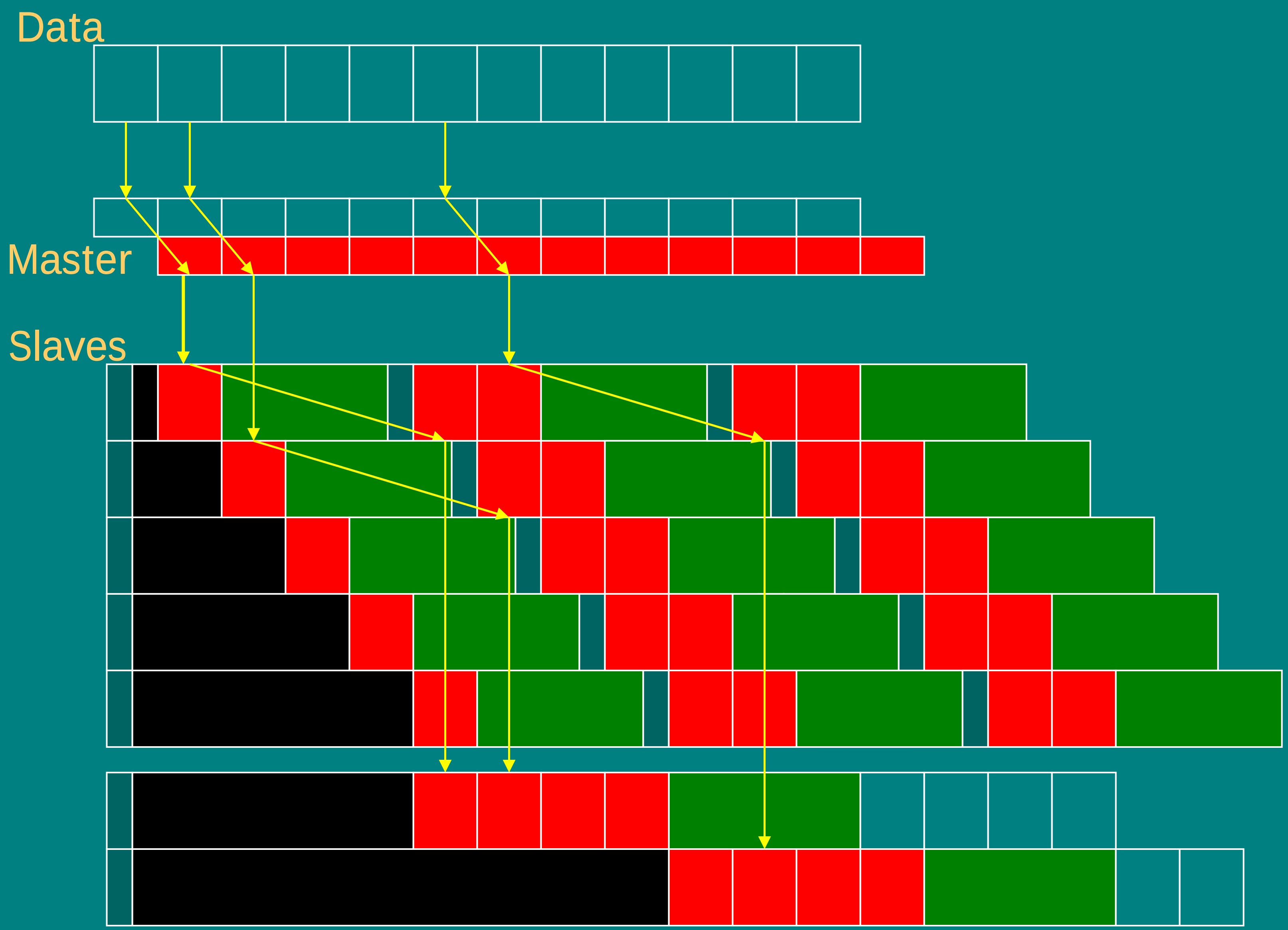
Clients



Parallelisation for real time processing

- Processing data as it is acquired raises new problems for parallelisation on distributed memory machines.
  - Can't seek on the input stream!
- To keep up with the input data, we must:
- get processors working as quickly as possible.
  - this means allocating work in the smallest possible units.
  - the problem of "wings" can no longer be avoided.
- make sure we have sufficient processing power
  - too much is now better than not enough!.

Real-time: master-slave pipeline hybrid



Developments since 1999

- Jodrell installed a 180 processor cluster to directly process the signal captured by 10 special cards each backed by ~10 processors
- They can now de-disperse the signal in real time provided the de-dispersion measure is known.
- However other interesting effects due to gravitational acceleration also call for signal processing searches.
- Discovery of first binary double pulsar makes tests of general relativity possible, theories of gravity are distinguished in strong field limits.

COBRA

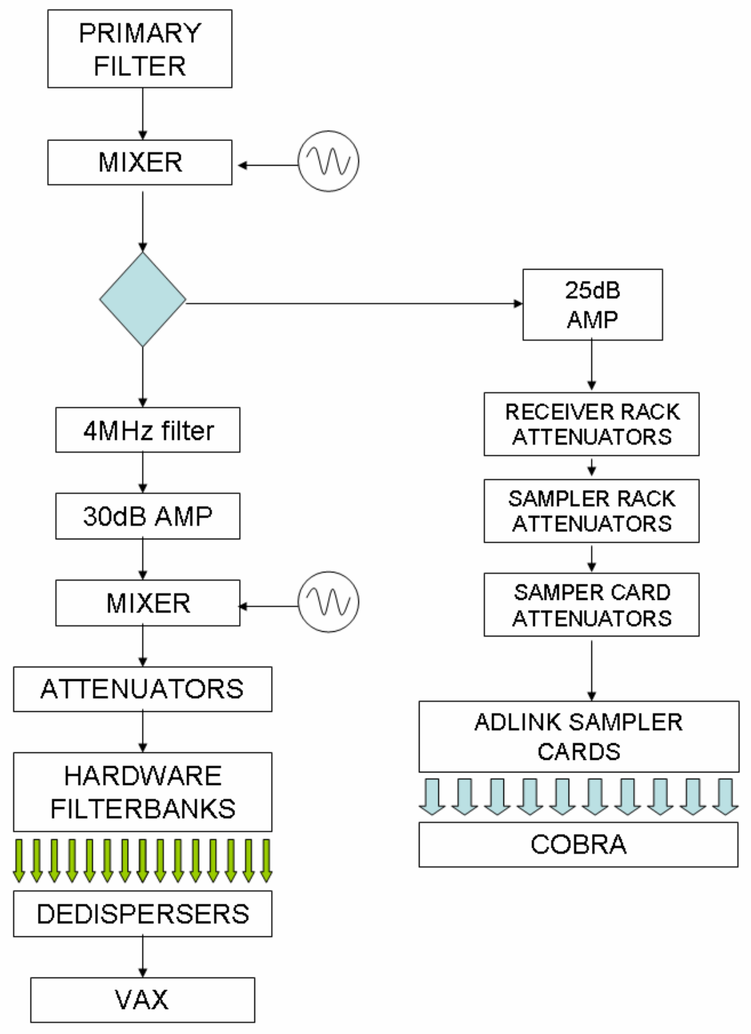
- **C**oherent **O**nline **B**aseband **R**eceiver for **A**stronomy

- A Beowulf supercomputing cluster
- 186 Intel Pentium processors
- 10 A to D sampling cards (8-bit)
- 5 MHz bandwidth across each sampler card

The University of Manchester
Jodrell Bank Observatory

MANCHESTER 1824

Signal Flow



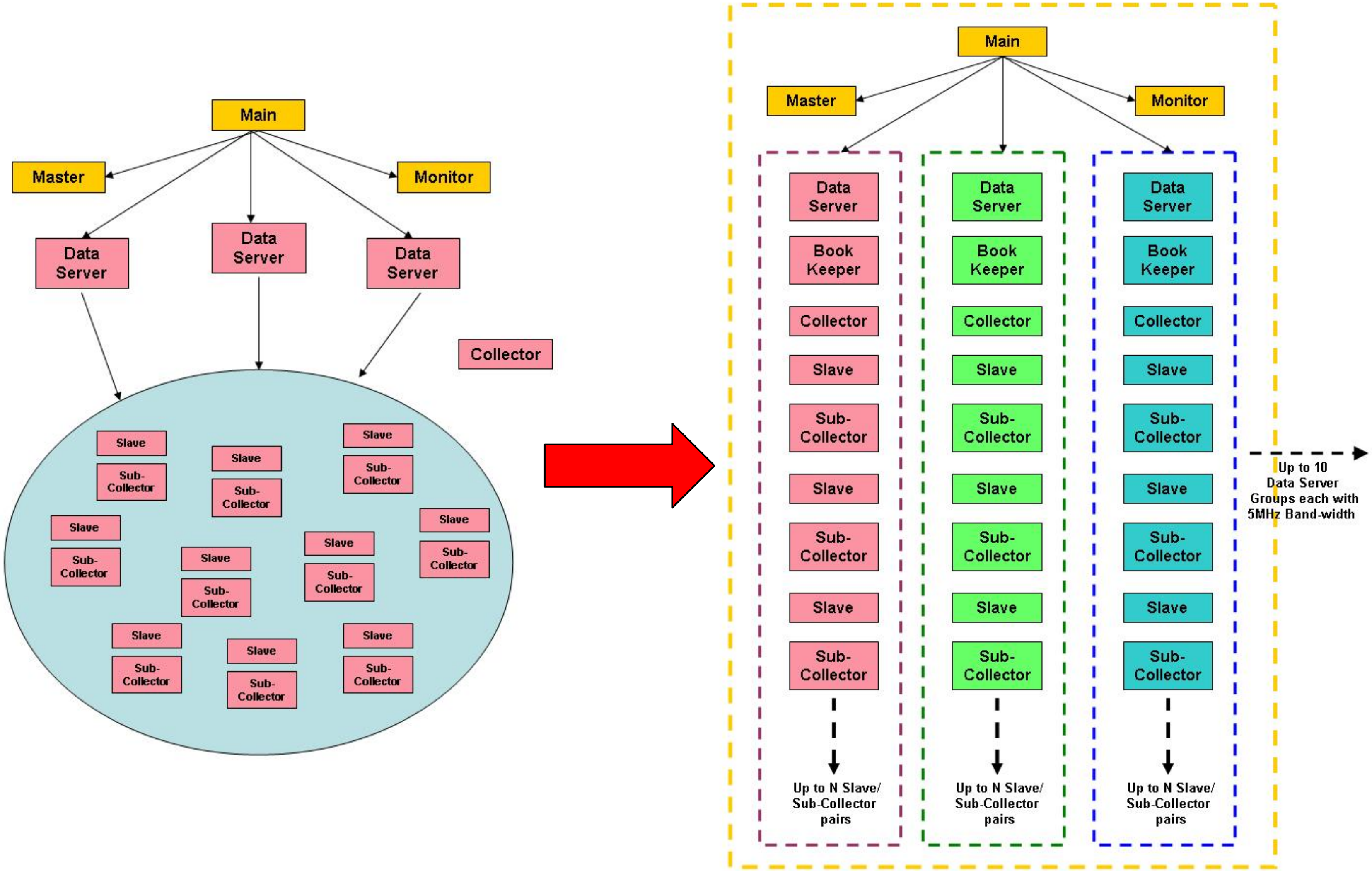
COBRA Software

- Development of legacy code
- Written in C, MPI, TCL and shell scripts

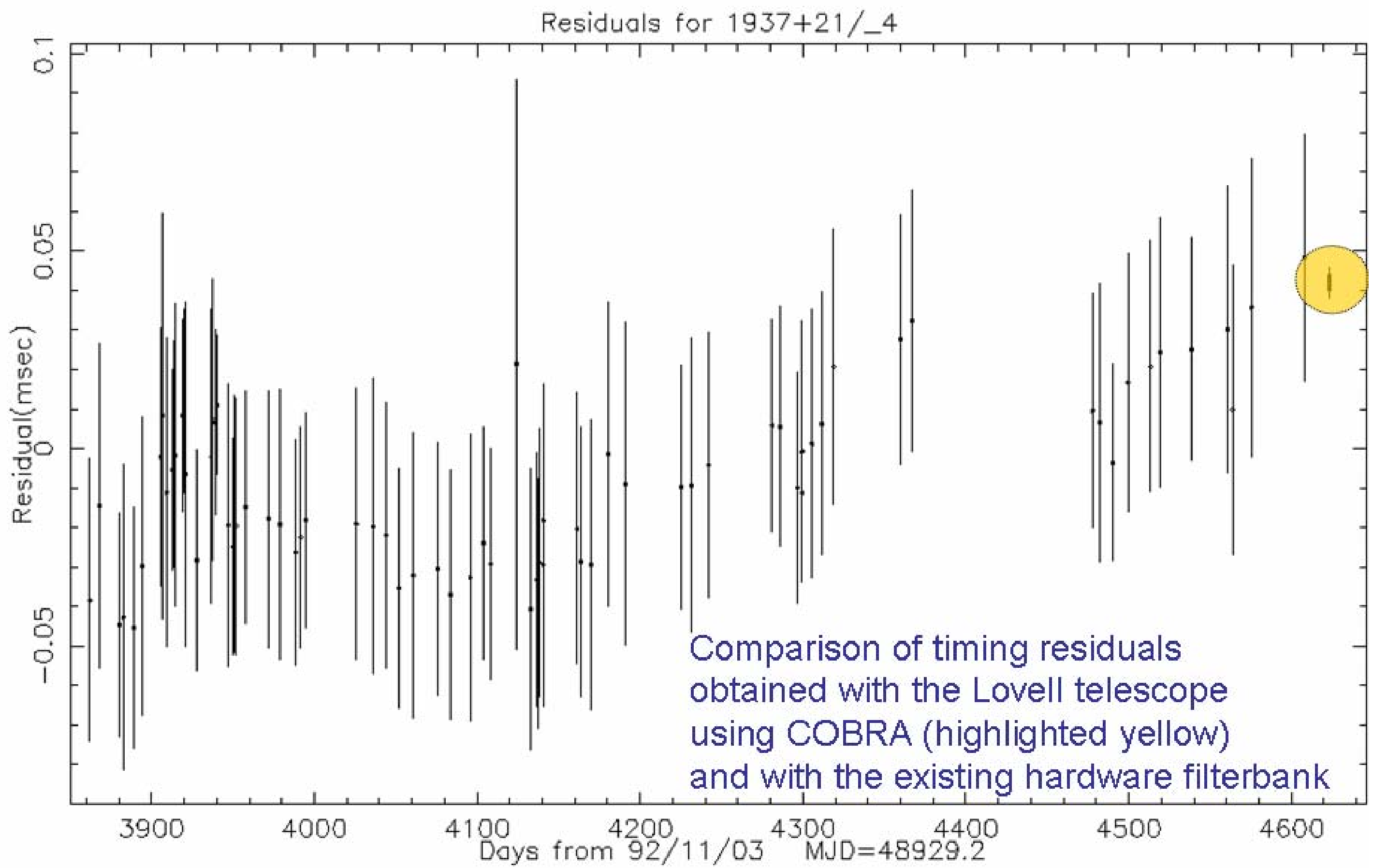
Message Passing Interface (MPI)

- Provides commands for writing parallelised programs
- Treated as a library within C
- Runs multiple instances of same code across multiple processors
- Cobra code written to make extensive use of MPI features

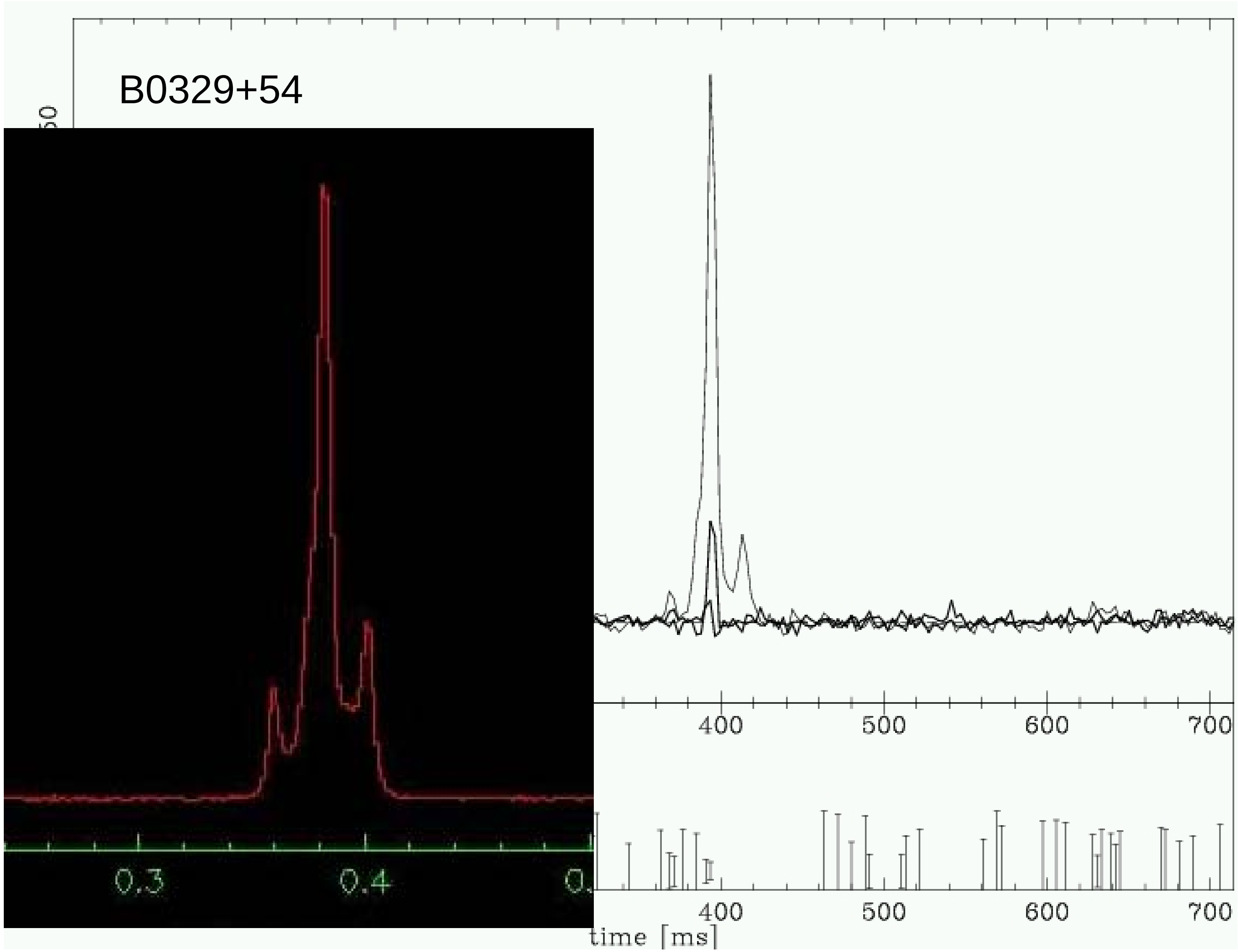
Code Structure



Timing Results



Dedispersed Pulse Profile



Pulsar Virtual Observatory

- Grid1D project funded in 2004 PPARC eScience round
- Encapsulate data so that it can only be accessed via methods that capture knowledge of pulsar astronomer.
- Use Triana as a workflow tool but need to predict in advance processing required for compute intensive modules, then use resource broker to reserve resources.
- Methods generalise to gravitational wave analysis.

Pulsar Virtual Observatory

- Need advanced prediction of resource requirements as user queries data.
- Need to provide non-expert users tools to construct correct workflows or reuse workflows for given form of search
- Need to capture knowledge of pulsar observation experts, control axis to data so that incorrect conclusions based on naïve methods of signal processing are avoided.

Summary

- SKA work in systems design, interaction of hardware and software, grid computing, networking and signal processing.
- Pulsar and gravitational work, algorithms for distributed signal processing, grid computing.