Big Data in Astronomy – Indian Perspective

Harinder P. Singh
(University of Delhi, india)
Ekaterinburg, September 5-7, 2016

Exciting time for Astronomy in India :

- **Participation in Mega Projects – TMT, LIGO, SKA, INO**
- **Space missions – ASTROSAT, ADITYA**
- **New Telescopes – 3.6m DOT, 4m ILMT, 2m NLST**
- **C-DAC HPC & National Knowledge Network backbone**
- **HRD Challenges**

Achnowledgements:

- **Eswar Reddy (IIA) – TMT**
- **Tarun Souradeep, Sanjit Mitra (IUCAA) - LIGO**
- **Yashwant Gupta (NCRA-TIFR) - SKA**
- **S. Seetha (ISRO) – ASTROSAT**
- **Dipankar Banerjee – Aditya, NLST**
- **Ajit Kembhavi (IUCAA) – VO, NKN**

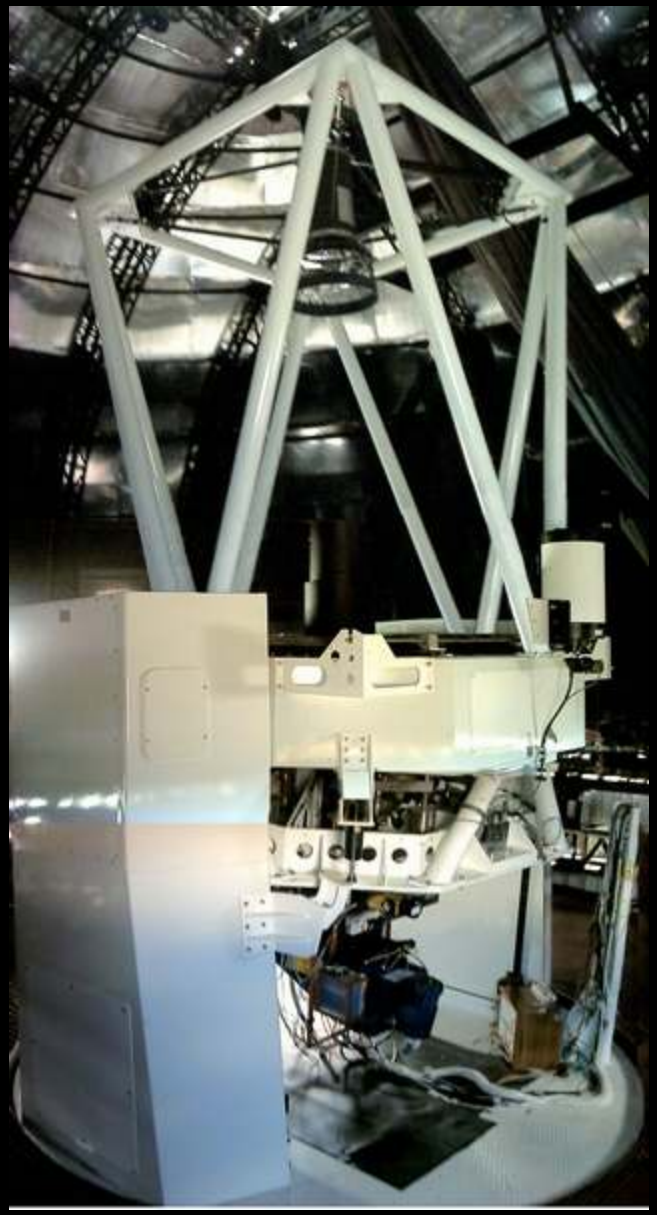
Existing optical facilities spread all across India

- **3.6 m DOT (ARIES)**
- **2.3m; 2m (IIA); 2m (IUCAA)**
- **1.3m, 1m (ARIES); 1.3m, 1m (IIA), 1.2m (PRL)**


Vainu Bappu Telescope,
2.3m, Kavalur, IIA

2m Himalayan Chandra Telescope

Indian institute if Astrophysics

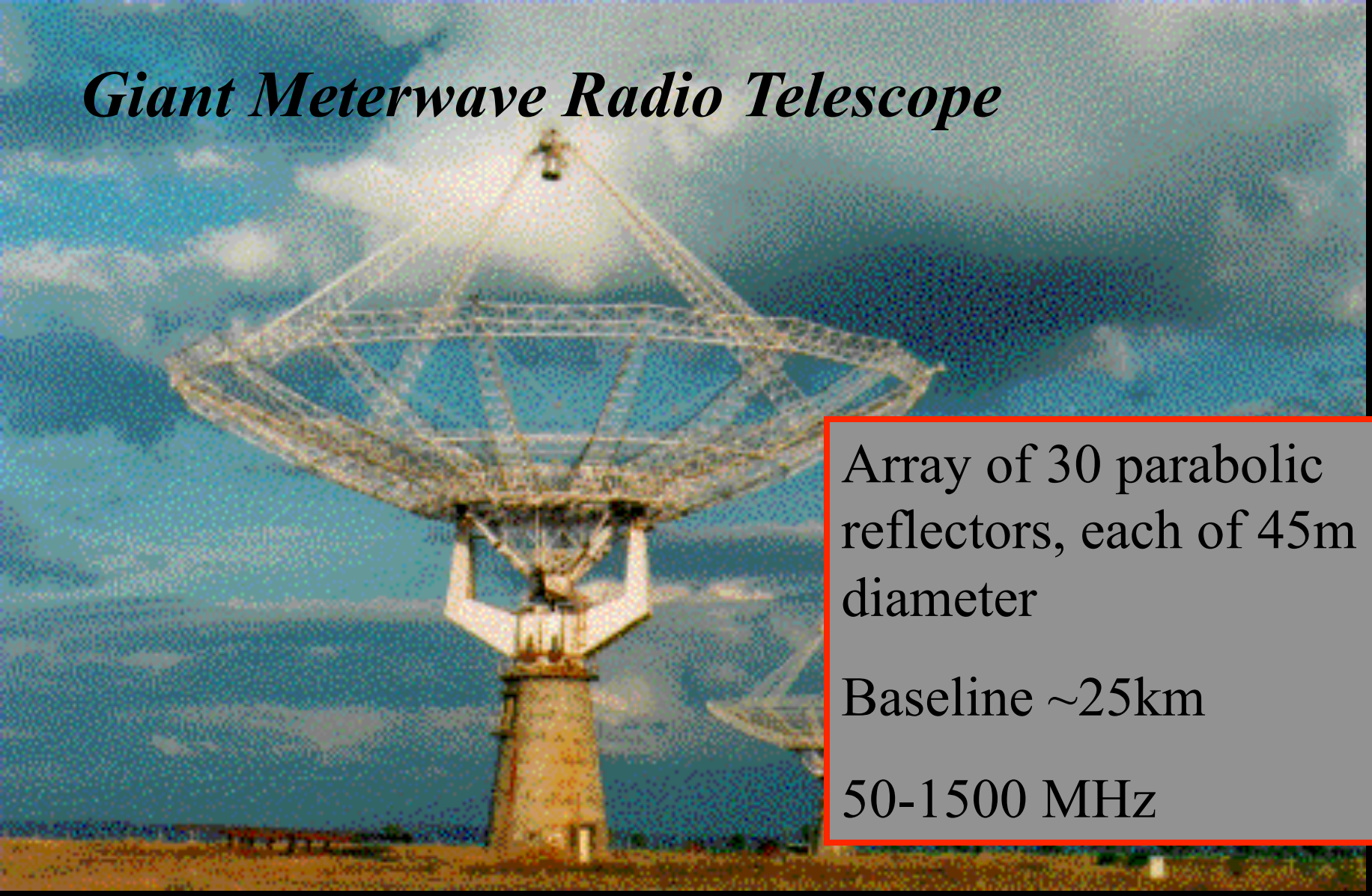


Hanle, Ladakh

IUCAA 2m, Giravali

Devasthal
3.6m Telescope
ARIES

Giant Meterwave Radio Telescope



Array of 30 parabolic reflectors, each of 45m diameter

Baseline ~25km

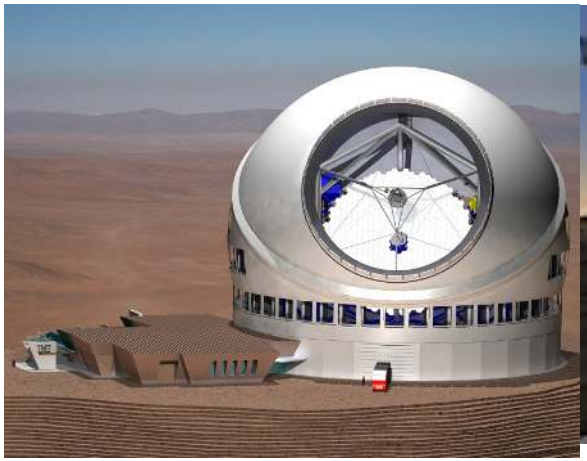
50-1500 MHz

GMRT, Narayangaon
NCRA-TIFR


Southern African Large
Telescope – IUCAA is a partner

Giant Segmented Mirror Telescopes (GSMTs)

24.5-m GMT



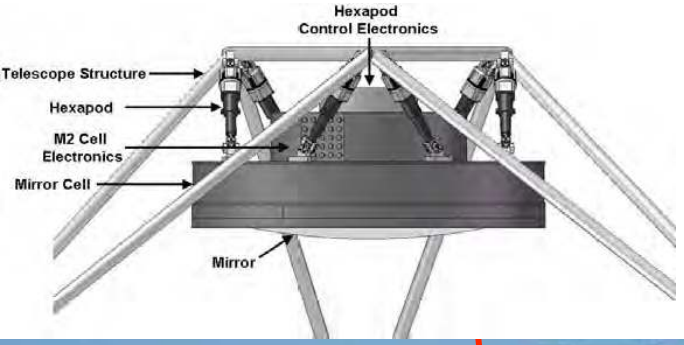
30-m TMT

39-m E-ELT

All the three projects invited India to participate.

This was in 2007

The Thirty Meter Telescope:



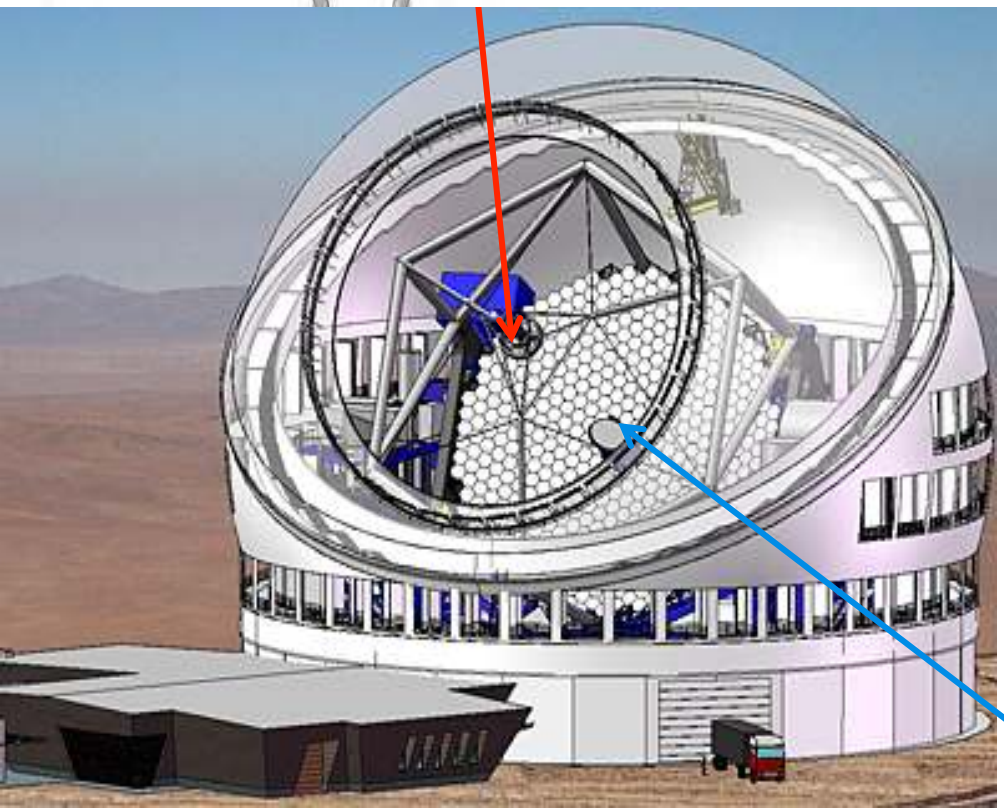
3.1m convex hyper boloidal secondary mirror

Ritchey-Chrétien optical design

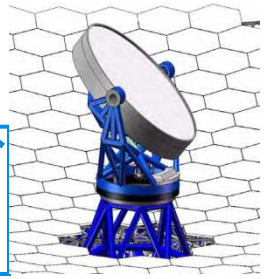
30m hyperboloidal f/1 primary mirror with 492 segments

20 arcmin field of view

f/15 final focal ratio



Flat 2.5m x 3.5m tertiary mirror



TMT Site: Mauna Kea

Clear/usable nights - 76%

Precipitable water - <2mm

Altitude - 4000m

Av. Temp. - 2.7C

Seeing statistics

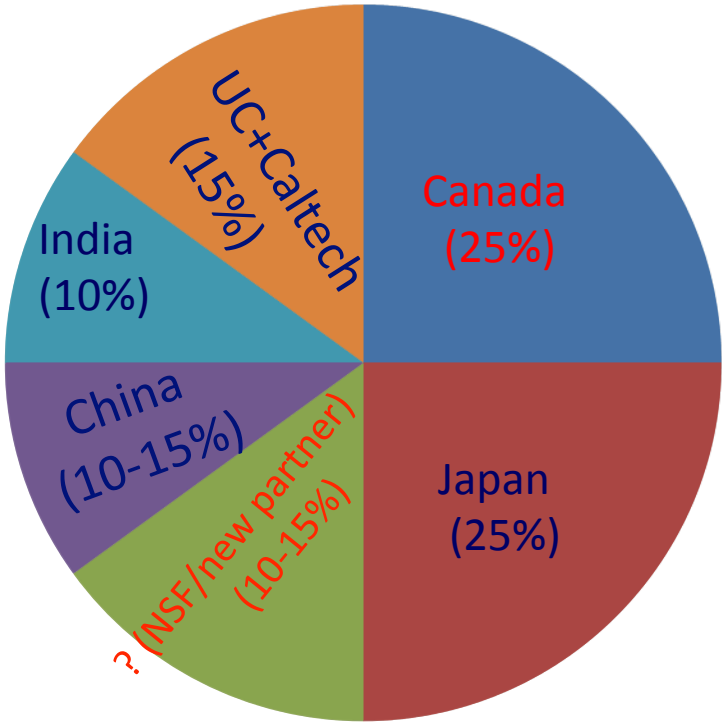
(50m above ground)

- **25% 0.40 arcsec**
- **50% 0.54 arcsec**
- **75% 0.75 arcsec**

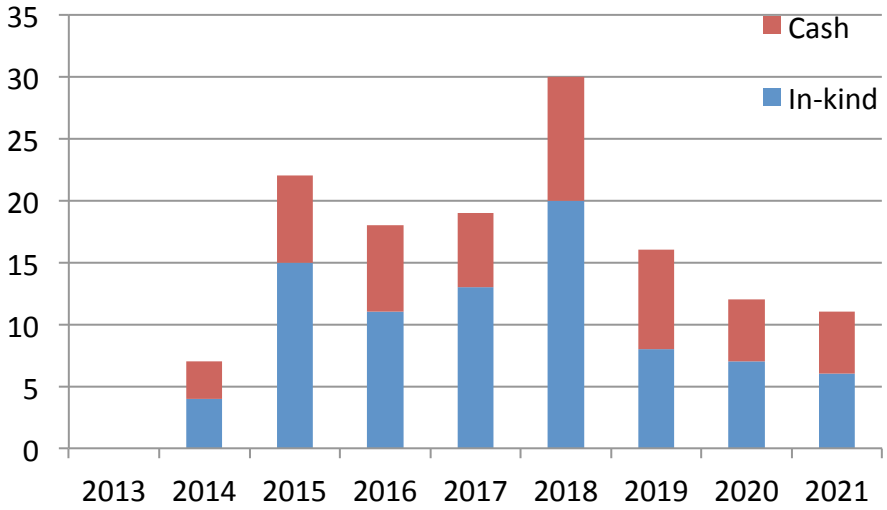
TMT Early Light Instruments

Instrument	Spectral Resolution	Science Case
Near-IR DL Spectrometer & Imager (IRIS)	≤4000	• Assembly of galaxies at large redshift • Black holes/AGN/Galactic Center • Resolved stellar populations in crowded fields • Astrometry
Wide-field Optical Spectrometer (WFOS)	300 - 5000	• IGM structure and composition 2<z<6 • High-quality spectra of z>1.5 galaxies suitable for measuring stellar pops, chemistry, energetics through peak epoch of gal form.
Multi-slit near-DL near-IR Spectrometer (IRMS)	2000 - 5000	• Near-IR spectroscopic diagnostics of the faintest objects • JWST followup
Mid-IR Echelle Spectrometer & Imager (MIRES)	5000 - 100000	• Physical structure and kinematics of protostellar envelopes • Physical diagnostics of circumstellar/protoplanetary disks: where and when planets form during the accretion phase
ExAO I (PFI)	50 - 300	• Direct detection and spectroscopic characterization of extra-solar planets
High Resolution Optical Spectrograph (HROS)	30000 - 50000	• Stellar abundance studies throughout the Local Group • ISM abundances/kinematics, IGM characterization to z~6 • Extra-solar planets!
MCAO imager (WIRC)	5 - 100	• Galactic center astrometry • Stellar populations to 10Mpc
Near-IR, DL Echelle	5000 - 30000	• Precision radial velocities of M-stars and detection of low-mass planets

Partnership Distribution and cost



India's contribution profile



Total Cost of the project
(construction)

USD 1195M (FY 2011)

USD 1522M (FY 2021)

10% share cost

USD 119M (FY 2011)

USD 152M (FY 2021)

Yearly Operations cost: ~5% of the total construction cost

TMT Software

Indian Institute of Astrophysics

Observatory Software

Inter-University Centre for Astronomy and Astrophysics

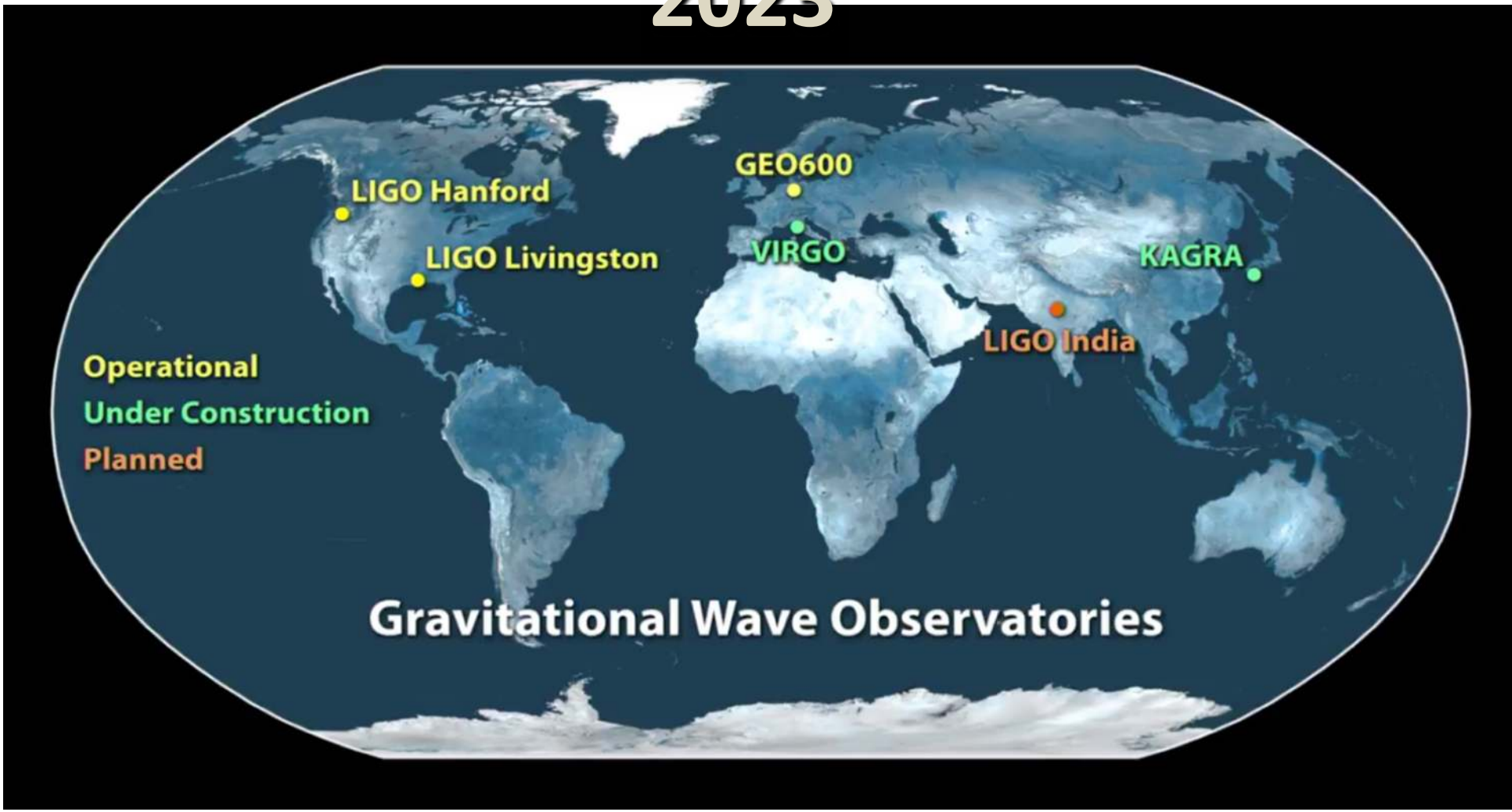
Telescope Control System

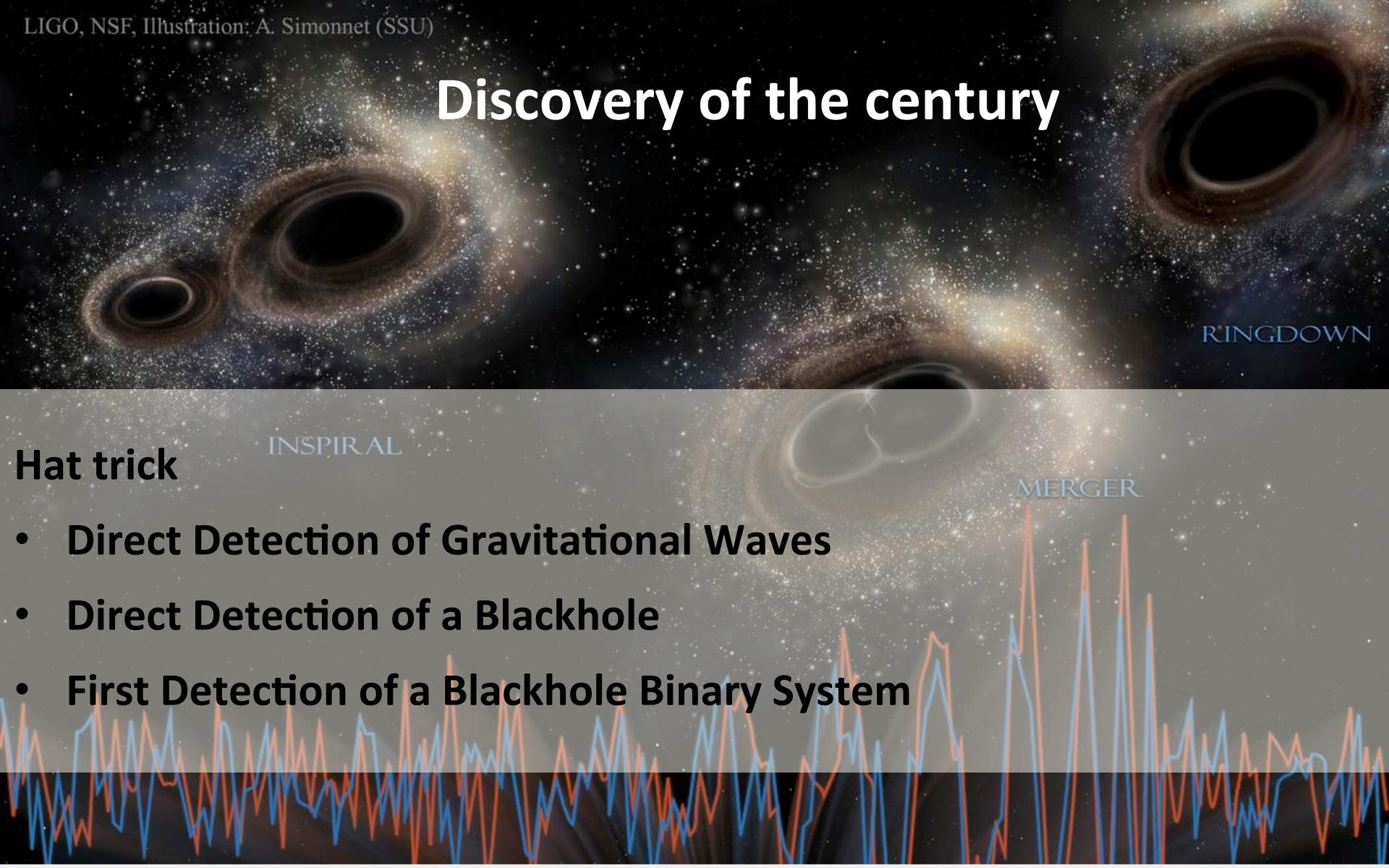
Partners: CDAC, PSL…

Global Network of GW Observatories

Largest baseline provided by LIGO-India

2023





Discovery of the century

Hat trick

- **Direct Detection of Gravitational Waves**
- **Direct Detection of a Blackhole**
- **First Detection of a Blackhole Binary System**

- **LIGO discovery paper has 37 authors from 9 Indian institutions,**
- **Over 60 Indian researchers in the Intl. LIGO Science Collab.**
- **The current FTE fraction in LSC is 5%, already 3rd largest outside USA.**

LIGO-India 'in principle' Approval
By Indian Union cabinet on Feb 17, 2016

Press Information Bureau
Government of India
Cabinet

17-February-2016 14:55 IST

Cabinet grants 'in-principle' approval to the LIGO-India mega science proposal

The Union Cabinet chaired by the Prime Minister Shri Narendra Modi has given its 'in principle' approval to the LIGO-India mega science proposal for research on gravitational waves. The proposal, known as LIGO-India project (Laser Interferometer Gravitational-wave Observatory in India) is piloted by Department of Atomic Energy and Department of Science and Technology (DST). The approval coincides with the historic detection of gravitational waves a few days ago that opened up of a new window on the universe to unravel some of its greatest mysteries.

The LIGO-India project will establish a state-of-the-art gravitational wave observatory in India in collaboration with the LIGO Laboratory in the U.S. run by Caltech and MIT.

The project will bring unprecedented opportunities for scientists and engineers to dig deeper into the realm of gravitational wave and take global leadership in this new astronomical frontier.

LIGO-India will also bring considerable opportunities in cutting edge technology f
engaged in the construction of eight kilometre long beam tube at ultra-high vacuum on

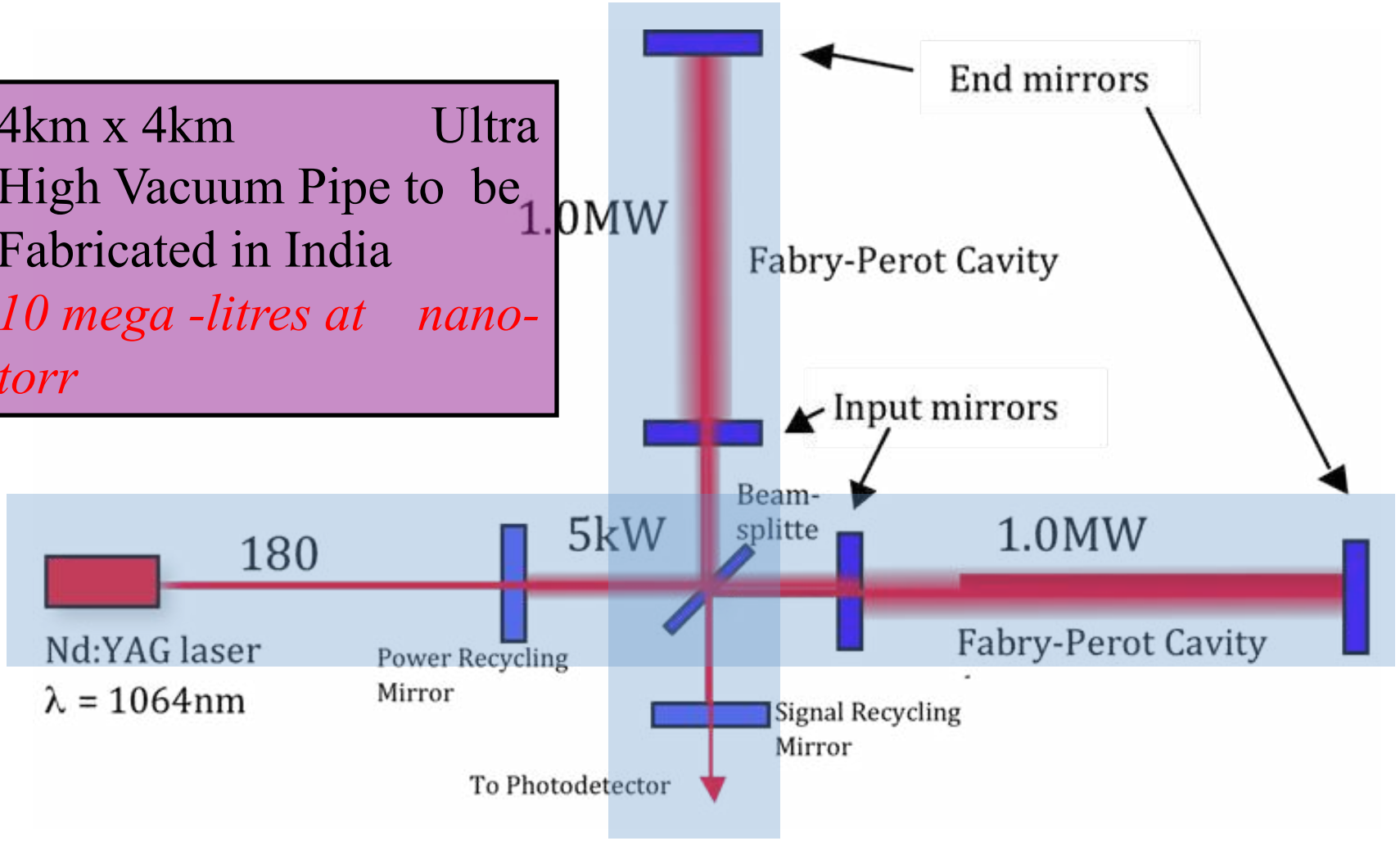
The project will motivate Indian students and young scientists to explore newer fronti
impetus to scientific research in the country.

***

Schematic of the Advanced LIGO Detector

4km x 4km Ultra High Vacuum Pipe to be Fabricated in India

10 mega -litres at nano-torr



The LIGO-India Advantage

Current:
Two US detectors
sky-localization
(620 square degree
2500 full moons)

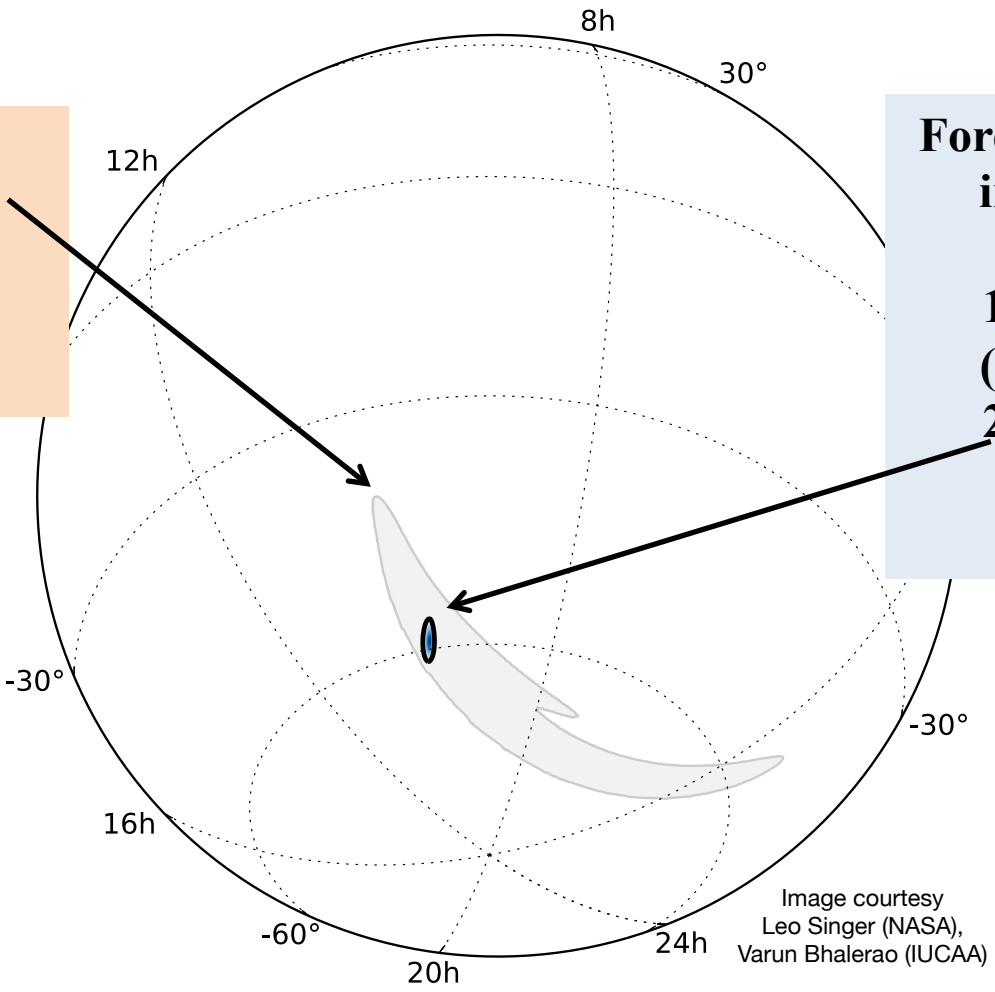


Image courtesy
Leo Singer (NASA),
Varun Bhalerao (IUCAA)

Forecast : LIGO-India
in joint operation
sky-localization
100 time smaller
(5 square degrees
20 full moons) !!!

LIGO-India Project

An Indo-US Collaboration

Funding agencies: NSF (USA) and jointly **DAE**(India) & **DST**(India)

Institutions: LIGO Laboratories, **Caltech** & **MIT** (USA), **IPR**, **IUCAA** & **RRCAT** (India)

Proposed Indian commitment

- **Construction and Operation of an Advanced LIGO Gravitiational-wave observatory on Indian soil in collaboration with the LIGO Laboratory**
- Infrastructure including 8 km of UHV system (10 million litres) with controls, installation of detector, as well as the team to build and operate the observatory.

Proposed US commitment

- **The entire hardware components of an advanced LIGO detector** along with facility designs and software provided by LIGO-USA and its UK, German and Australian partners..
- Close collaboration, providing required training in technology and sharing experience.

LIGO Data

- Time series data sampled at 16Hz - 16kHz
  - Thousands of monitoring channels
  - "science channels": ~1% of total data
- LIGO data volume:
  - few MBps / detector x 1 year x 3 detectors
- Advanced LIGO data volume: ~1 petabyte / year

LSC Computing Facilities

- NEMO @ UWM (6914 cores)
- ATLAS @ AEI Hannover (6344 cores)
- Cluster at Tier-1 site Caltech (4236 cores)
- LHO (2300 cores), LLO (2032 cores)
- Clusters at Tier-2 sites MIT, UWM, PSU
- Clusters at GEO sites Birmingham, Cardiff
- SUGAR @ Syracuse (320 cores)

LDAS @ IUCAA

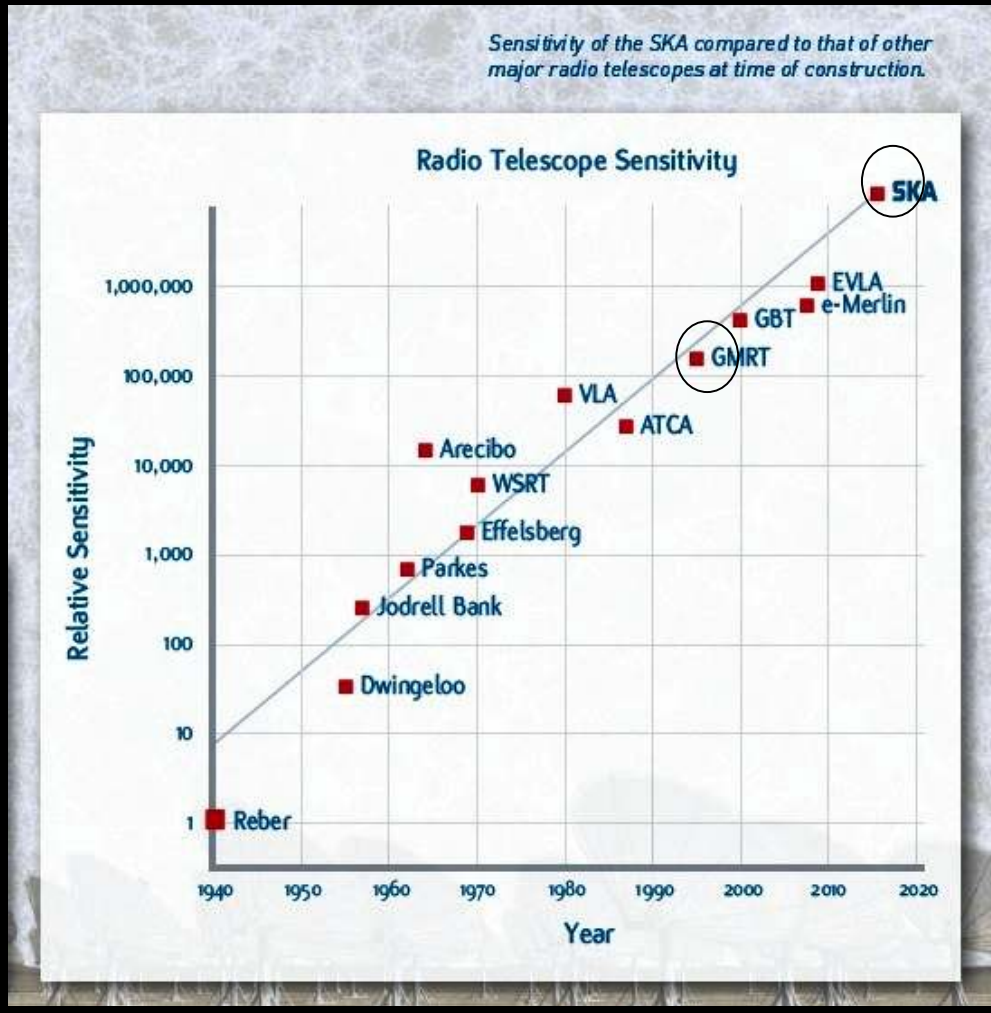
- LDAS = LIGO Data Analysis System
- 105 computing nodes x 24 cores per node
  - about 100TF peak performance
  - memory 128GB or 256GB/node and 250TB storage
  - 10G interconnect
- Scope for GPU based codes: 10 NVIDIA K40 cards
- Authentication method: GSI enabled OpenSSH

Future Plans

- Tier-1 Data Centre
- Part of NKN big data project
  - storage ~1 PetaByte
  - will serve to the national community
  - may lead to national astronomical database
- ≥ 100TF computing cluster

The SKA is the future of Radio Astronomy

- The SKA is the most ambitious radio astronomy project
- Science with the SKA will be truly revolutionary !
- SKA will drive the growth of many new & cutting edge technologies : from electronics to supercomputing to software
- All the major radio astronomy nations are members



Radio telescope sensitivities over the years

SKA will be 50x better than today's best !

Indian Participation in SKA : overview

- India has been part of SKA from the beginning & member of the formal organisation bodies (including the present "SKA Organisation")
- Indian astronomers have been active participants in various SKA committees
- With 3% of full SKA collecting area, the GMRT project is an important low frequency radio telescope -- provides useful test-bed for SKA work
- Presently ongoing upgrade activities at GMRT provide for natural overlap & synergy with SKA efforts ; e.g. monitor and control systems with industry collaboration, development of next-gen signal processing platforms
- RRI's involvement in the MWA pathfinder project provides valued experience
- Both main science goals of SKA Phase I are active areas of research in India
- NCRA, working with partners from software research groups & industry, has taken a lead role in some of the work packages for design of the SKA

ASTROSAT

Space astronomy observatory of India

- India's multiwavelength space astronomy mission is aimed at studying the celestial sources in X-ray (0.3 - 100 keV), optical (320 – 550 nm), far UV (130 – 180 nm) and near UV (200 – 300nm) spectral bands simultaneously.
- Spacecraft and payloads are healthy. Payloads developed and operated by different centers of ISRO, TIFR, IIA, IUCAA, CSA, UoL, NCRA, PRL, RRI
- All the payloads are operational. Observations under progress.
- The preliminary results are made available at the ISRO website (www.isro.gov.in).

Thirtieth consecutive success for PSLV

Lift-off from SHAR
at 10:00 hrs IST
on 28th Sep 2015

- The spacecraft was launched successfully on 28th September 2015 from SDSC Sriharikota. **Orbit: 650 km, near-equatorial; Inclination: ≤ 6 degree.**

Spacecraft – actual view

Large Area X-Ray Proportional Counter (LAXPC)

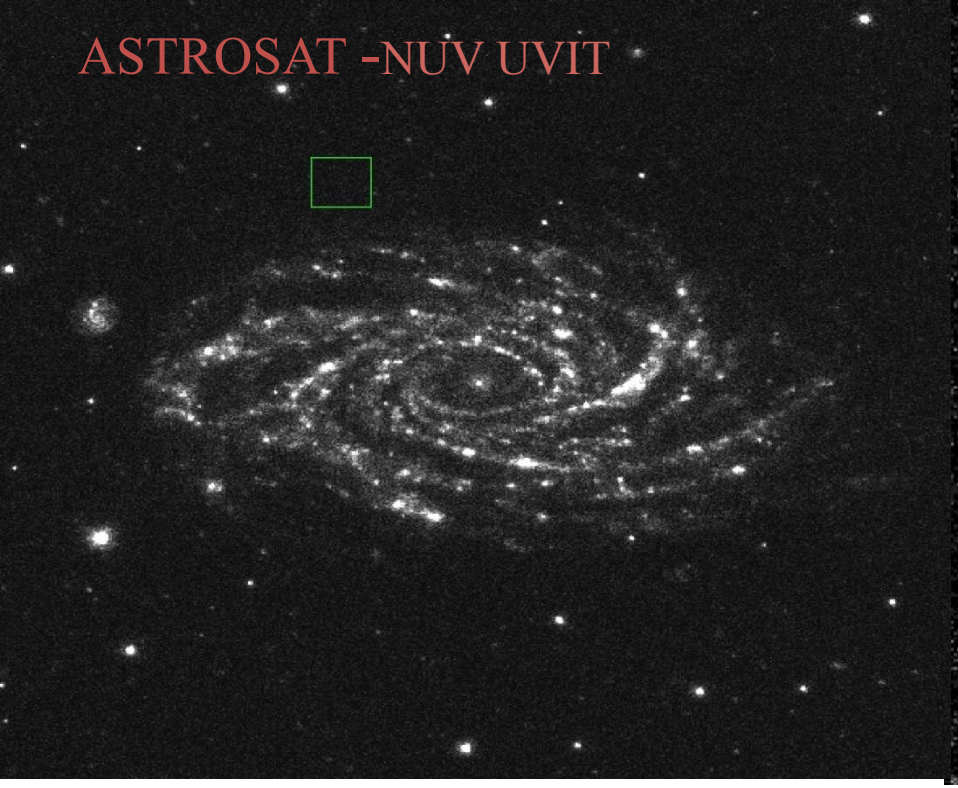
Ultra Violet Imaging Telescope (UVIT)

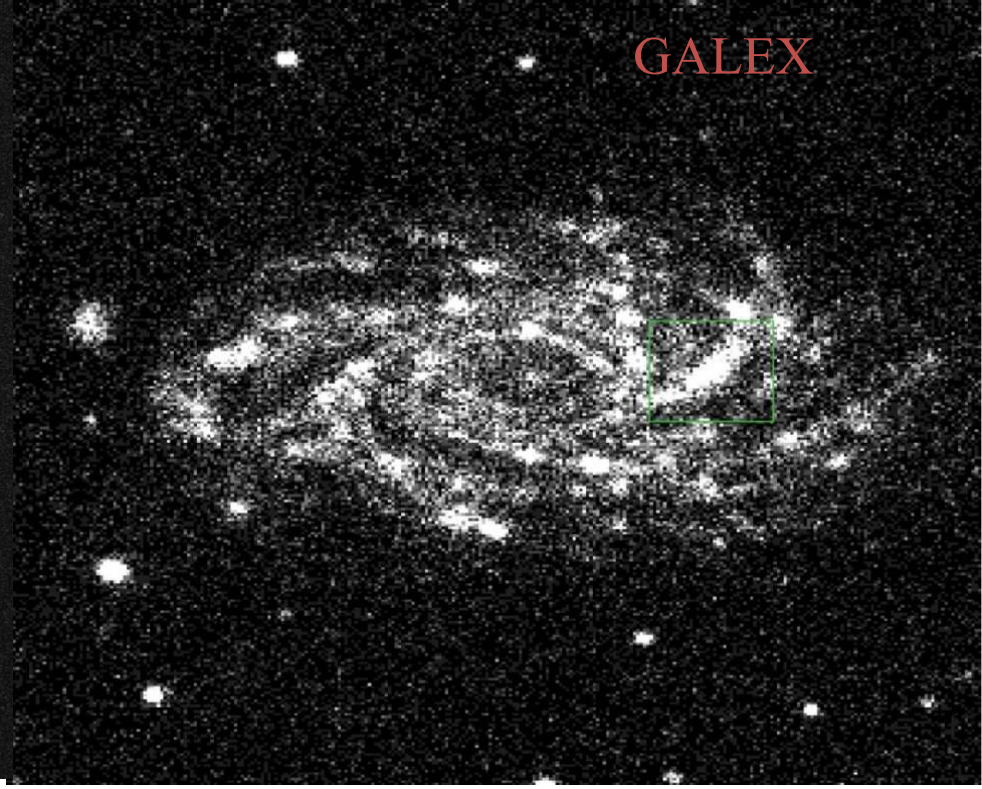
Cadmium Zinc Telluride Imager (CZTI)

Soft X-ray Telescope (SXT)

Scanning Sky Monitor (SSM)

S/C in tilted position during pre-launch test at SHAR





NGC 2336

First Indian satellite to be operated as an Observatory

- **Performance Verification (PV):** First 6 months after launch; payloads switched ON in a sequence and their performance verified and calibrated; all the data belongs to payload teams.
- **Guaranteed Time (GT):** 6 months period after PV phase completion; reserved for observations from the payload teams.
- **Regular observations**: After completion of PV and GT; based on Announcement of Opportunity (AO) cycles; open to national and international user community
- **Target of Opportunity (ToO)** aimed at new transient sources which cannot be predicted in advance
- 2% of the observing time is reserved for **calibration** and for maintaining the health of the instruments
- **Observation time is planned to be made open** to Indian Scientists after first year of performance verification and international users after second year, on proposal basis.

ASTROSAT Data Pipeline

- Level 0 Data: Raw Data, segregated by instrument, along with auxiliary data. Very instrument specific, not released for public use.
- Level 1 Data: Reorganized raw data in FITS format, along with auxiliary data needed for further processing. First released via the web to the instrument PI, then to other users.
- Level 2 Data: This contains standard science products obtained from Level 1 data. Released in FITS format in the same manner as Level 1 data.

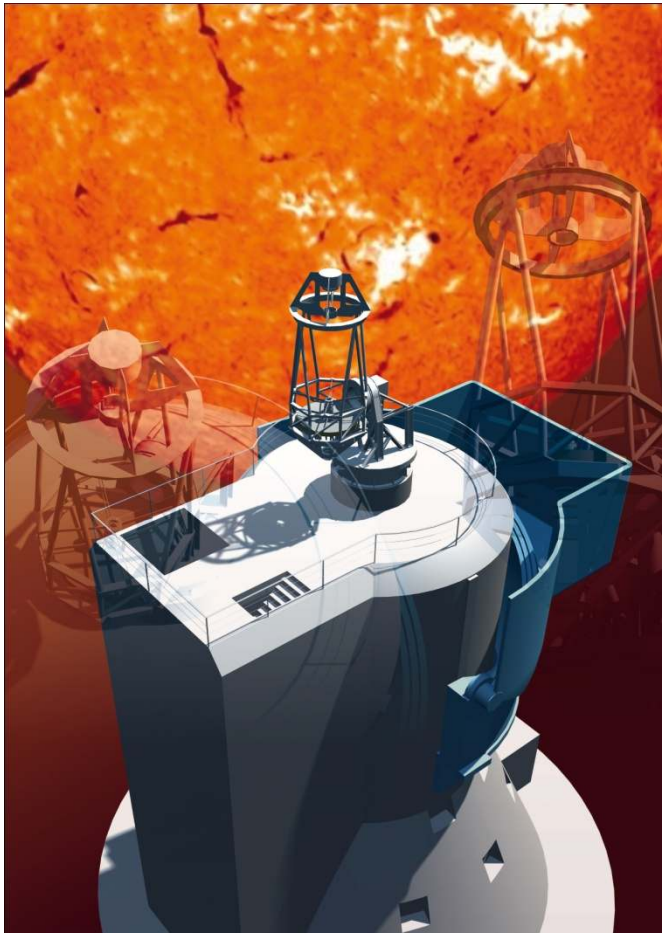
ASTROSAT Data Rate

Instrument	Level 0 Data	Level 1 Data	Level 2 Data
UVIT	50 GB	50 GB	10 GB
SSM	3GB	3GB	0.5 GB
AUX	1 GB	1 GB	0.5 GB
Total	**64 GB per day**	**64 GB per day**	**32.5 GB per day**

LAXPC: A typical event file for a 10 ks exposure will be about 0.5-2 GB, depending on the brightness of the source.

~12 TB yr^{-1}

THE NATIONAL LARGE SOLAR TELESCOPE (NLST)



NLST IN THE INTERNATIONAL CONTEXT
NEW GENERATION SOLAR TELESCOPES

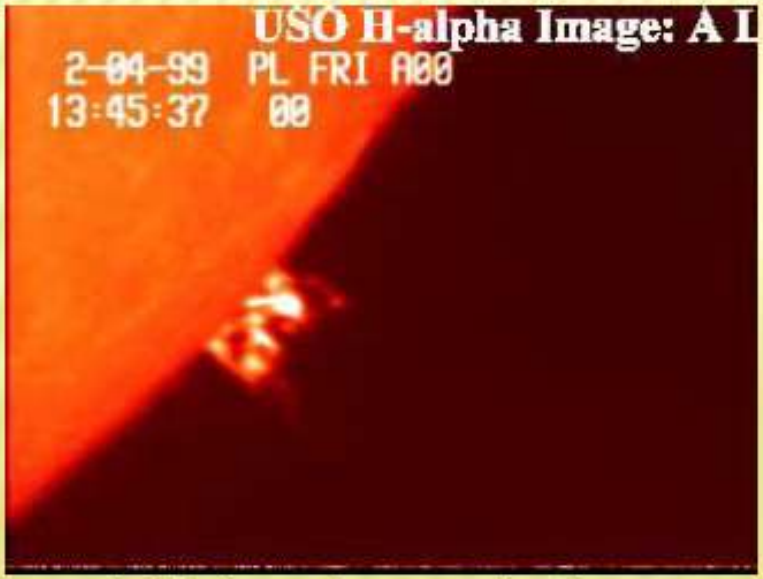
GREGOR	1.5 m	Germany	Tenerife (Spain)	2010
NST (New Solar Telescope)	1.6 m	U.S.A.	Big Bear (U.S.A.)	2009
NLST	2 m	India	Himalayas (India)	2020
ATST (Advanced Technology)	4m	USA	Haleakala (Hawaii, USA)	2018
EST (European Solar Technology)	4 m	European Consortium	Tenerife (Spain)	>2020
CGST (Chinese Giant Solar Telescope)	8 m	China	TBD	>2025 (?)

Aditya-L1 Mission

Proposed Indian solar observatory at L1

Mission objective: To design, realize and launch a satellite for solar studies to be placed in a halo orbit around the Lagrangian point 1 (L1) of the Sun-Earth system (1.5 million km from Earth; approximately 1/100th of AU).

Scientific objective: To study the solar dynamics in the chromosphere and corona with a suite of instruments including a coronagraph and a UV imager. The orbit around L1 provides continuous solar observations without any eclipse/ occultation and is an excellent outpost outside Earth's magnetic field to make in-situ measurements of incoming charge particles.



Initiation of mass ejection

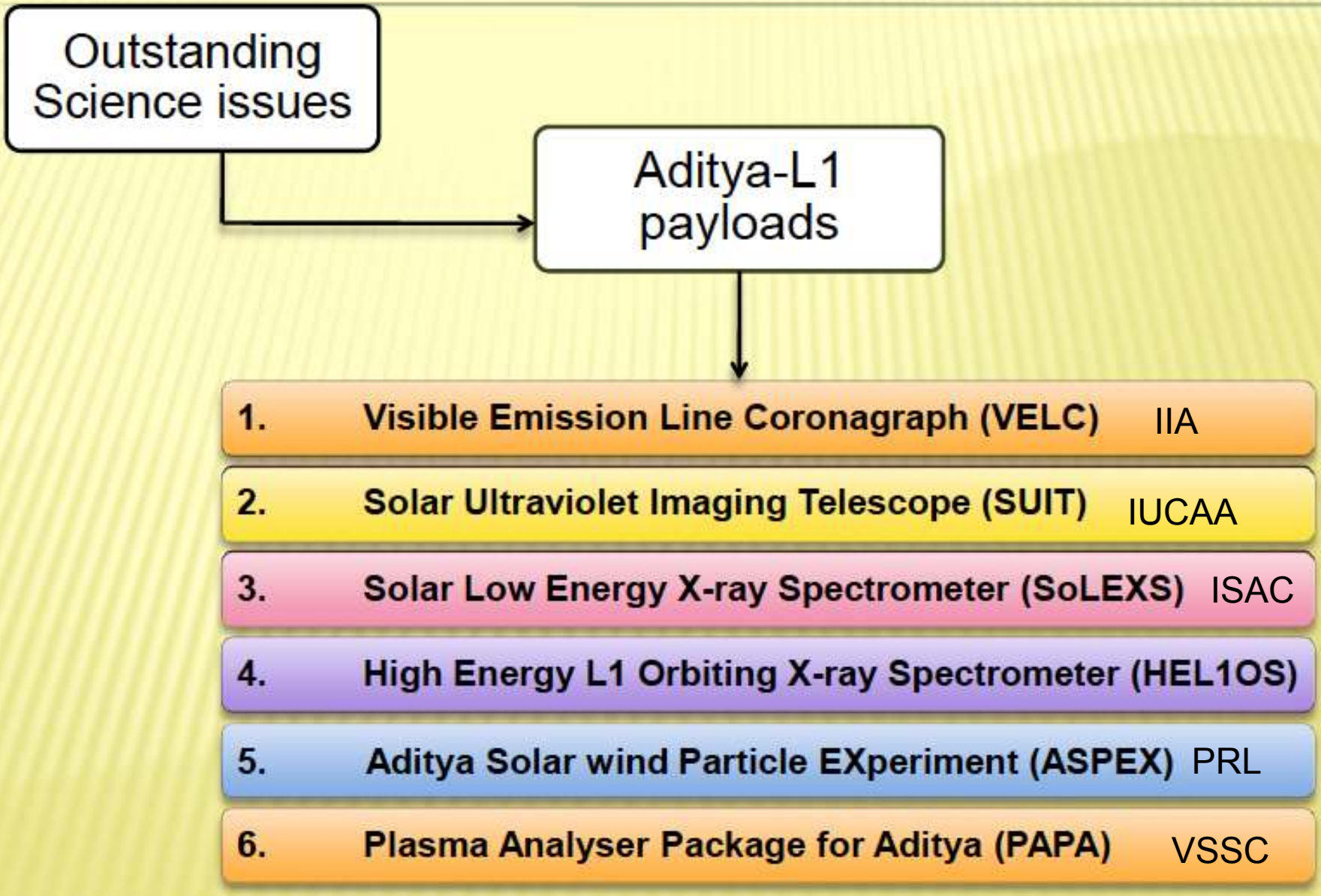
Launch : PSLV-XL

Target : 2018 – 2019

L1 Orbit insertion : ~ 100 days

Mission life : 5 years (nominal)

Payloads selected for Aditya-L1



C-DAC - High Performance, Grid and Cloud Computing (HGCC)

High Performance Computing (HPC) plays an important role in both scientific advancement and economic competitiveness of a nation - making production of scientific and industrial solutions faster, less expensive, and of higher quality. HPC is a key component in many applications: designing vehicles and airplanes; designing high-rise buildings and bridges; discovery of drugs; discovery and extraction of new energy sources like oil and natural gas; weather forecasting; and many more.

Supercomputing Systems

C-DAC embarked on its first HPC mission in 1988...

National Supercomputing Facilities

C-DAC has commissioned and operates three national...

Supercomputing Applications

The supercomputing systems and facilities of C-DAC...

Supercomputing Solutions

Drawing from a large pool of resources and skill sets...

Grid Computing

Garuda is India's national grid infrastructure of HPC...

Cloud Computing

C-DAC's research focus in cloud computing includes design...

Products & Services

- High Performance Computing, Grid & Cloud Computing
- Multilingual Computing & Heritage Computing
- Professional Electronics, VLSI & Embedded Systems
- Software Technologies including FOSS
- Cyber Security & Cyber Forensics
- Health Informatics
- Education & Training

Related Links

- **Download brochure** (PDF 4.69MB)

C-DAC HPC Supercomputing Systems

C-DAC embarked on its first HPC mission in 1988. Since then, it has delivered a series of supercomputing systems called PARAM series of supercomputers. These include:

- PARAM 8000, India's 1st Giga-scale supercomputer in 1990
- PARAM 10000, 100 Gigaflop supercomputer in 1998
- PARAM Padma, 1Teraflop supercomputer in 2002. This was India's first supercomputer to enter the Top500 list of supercomputers of the world (ranked 171 in June 2003)
- PARAM Yuva, a 54 Teraflop supercomputer in 2008 (ranked 69 in November 2008)
- PARAM Yuva II, a 529 Teraflop supercomputer in 2013 (ranked 69 in June 2013)
- PARAM Biochrome is an HPC cluster for Bioinformatics applications. The cluster has a compute capacity of 5 TeraFlops.
- PARAM Bio Blaze: A supercomputing facility with peak compute power of 10.65 TF, was launched on February 18, 2014 to address the challenges in bioinformatics

PARAM Yuva II (File Type: .pdf, File Size: 2.48MB, Updated on: Jan 2013), the latest in the series, is a eight-core, dual-socket node based hybrid compute cluster with multiple interconnects, compute co-processor, hardware accelerator, high performance storage and supporting softwares for parallel computing. It incorporates C-DAC's in-house technologies including PARAMNet-3 - a High Speed System Area Network, FPGA based hardware accelerators called Reconfigurable Computing System (RCS), a range of Systems and Applications Software components, and Industrial Design and Engineering of the whole system. PARAM Yuva II has sustained performance of 386.71 TF for LINPACK Benchmark and peak performance of 529.4TF. It has high bandwidth storage of 200 Terabytes.

Products & Services

- High Performance Computing, Grid & Cloud Computing
- Multilingual Computing & Heritage Computing
- Professional Electronics, VLSI & Embedded Systems
- Software Technologies including FOSS
- Cyber Security & Cyber Forensics
- Health Informatics
- Education & Training

Related Links

National Knowledge Network

- In March 2010, Govt. of India allocated ~1 billion USD for implementation of NKN
- To connect educational institutions, R & D institutions, health service facilities, agricultural institutions and libraries of the country

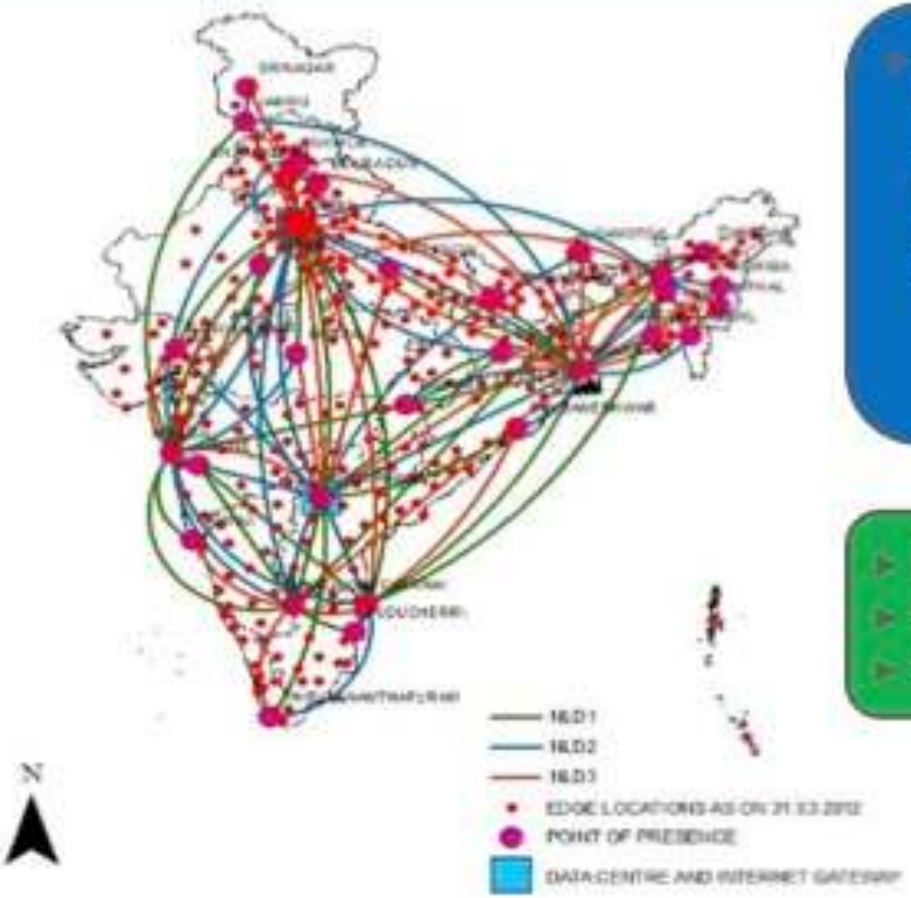
Technical Overview

- The backbone of the network starts from 2.5 Gbps and progressively moves onto 10 Gbps connectivity between 7 Supercore (fully meshed) locations across India.
- The network is further spread out through Core locations with multiple of 2.5/10 Gbps partially meshed connectivity with Supercore locations.
- The network architecture and governance structure allows users with options to connect to the distribution layer as well.
- For special interest groups, NKN enables Virtual Private Networks
- NKN provides international connectivity to its users for global collaborative research. Presently, NKN is connected to:
  - Trans Eurasia Information Network (TEIN3)
  - Similar connectivity to few other research networks is in the pipeline

Connected Sites Under Initial Phase of NKN

R S Mani

NKN: On Conclusion of Final Phase



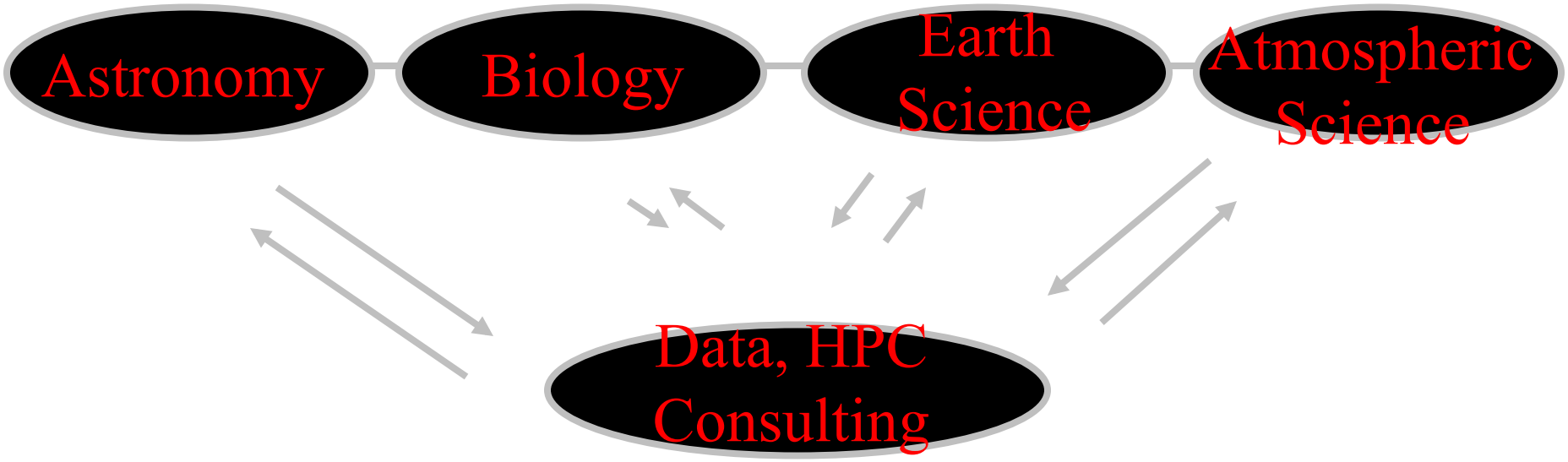
- At the conclusion of the final phase, NKN shall have presence in more than 500 districts of India with connectivity to major research and education institutions

- Points of Presence (PoP) : 31
- Backbone Links : 89
- Edge links : 1500

R S Mani

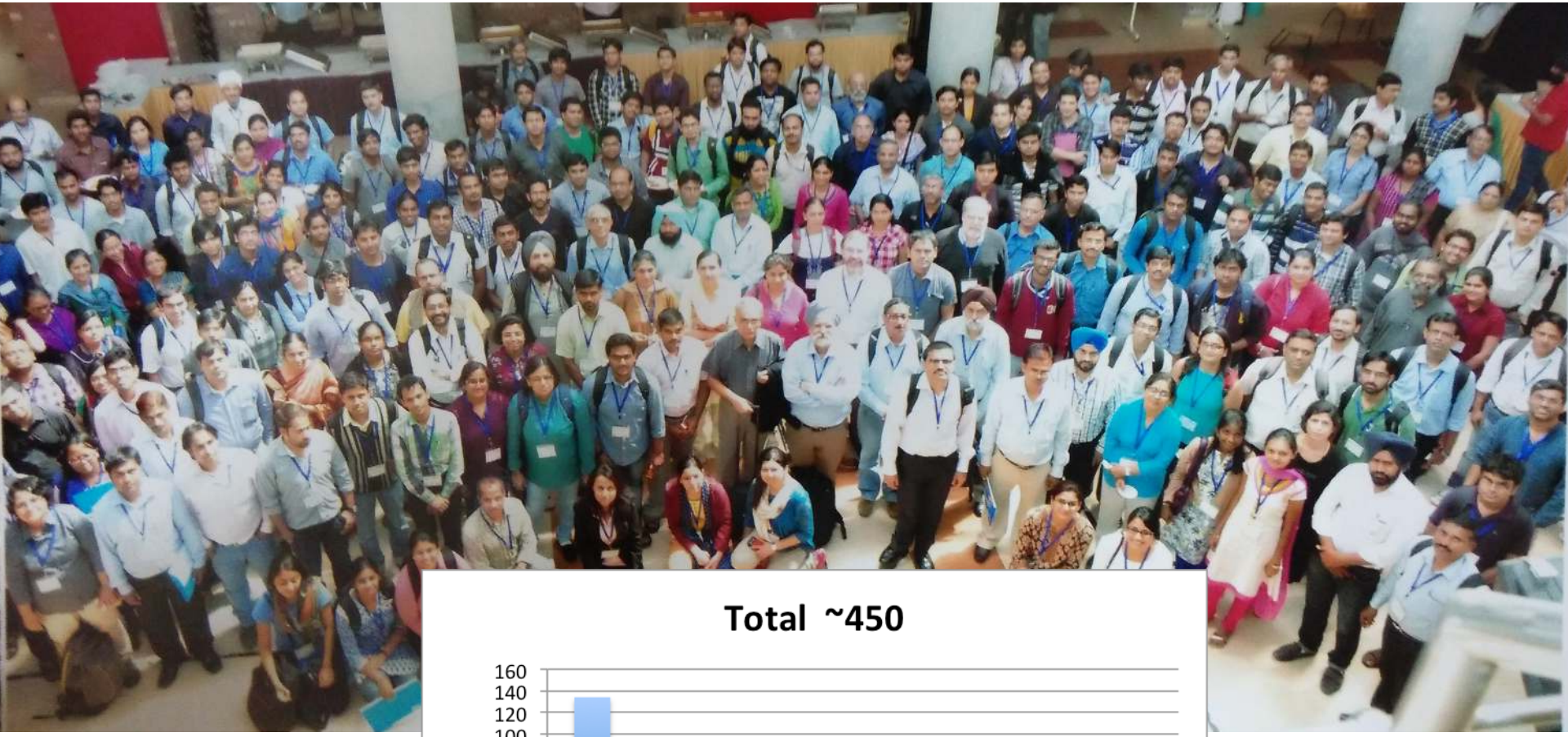
Big Data Infrastructure

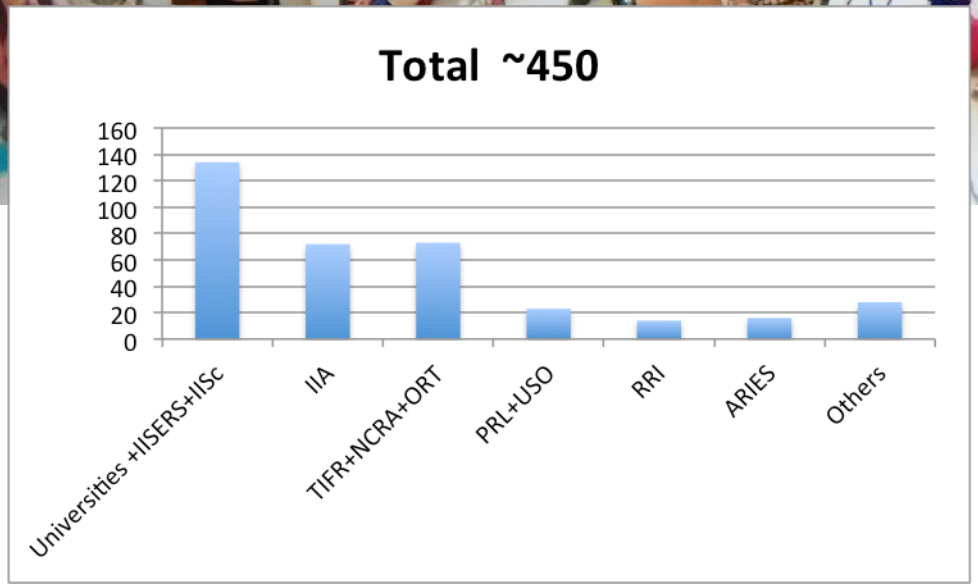
Domain Specific Data Centres
Few PByte Level



Main Data Centre
Many PByte Level

Pilot study by Ajit Kembhavi & Group





Human Resource

IAU Membership profile of BRICS

Country	IAU membership
Brazil	206
Russia	440
India	281
China	659
South Africa	118

HRD requirements in Astronomy in the next 5-10 years

- LIGO – TMT – SKA – Aditya etc
- 800 Additional Active Astronomers

Ways:

- Popularization at different levels
- More introductory workshops, advanced workshops
- Internships like Indigo summer internships
- Joint graduate schools for large groups – 20-50 students/yr
- Instrument and software development centers at universities
- Introduction of observatory fellowships
- Student fellowships at MS level like SKA
- Post-Doc Fellowships (SALT/TMT Post Doc. Fellowships)
- TMT faculty for universities
- Intake expansion of PhDs
- Job Opportunities

Thank You !