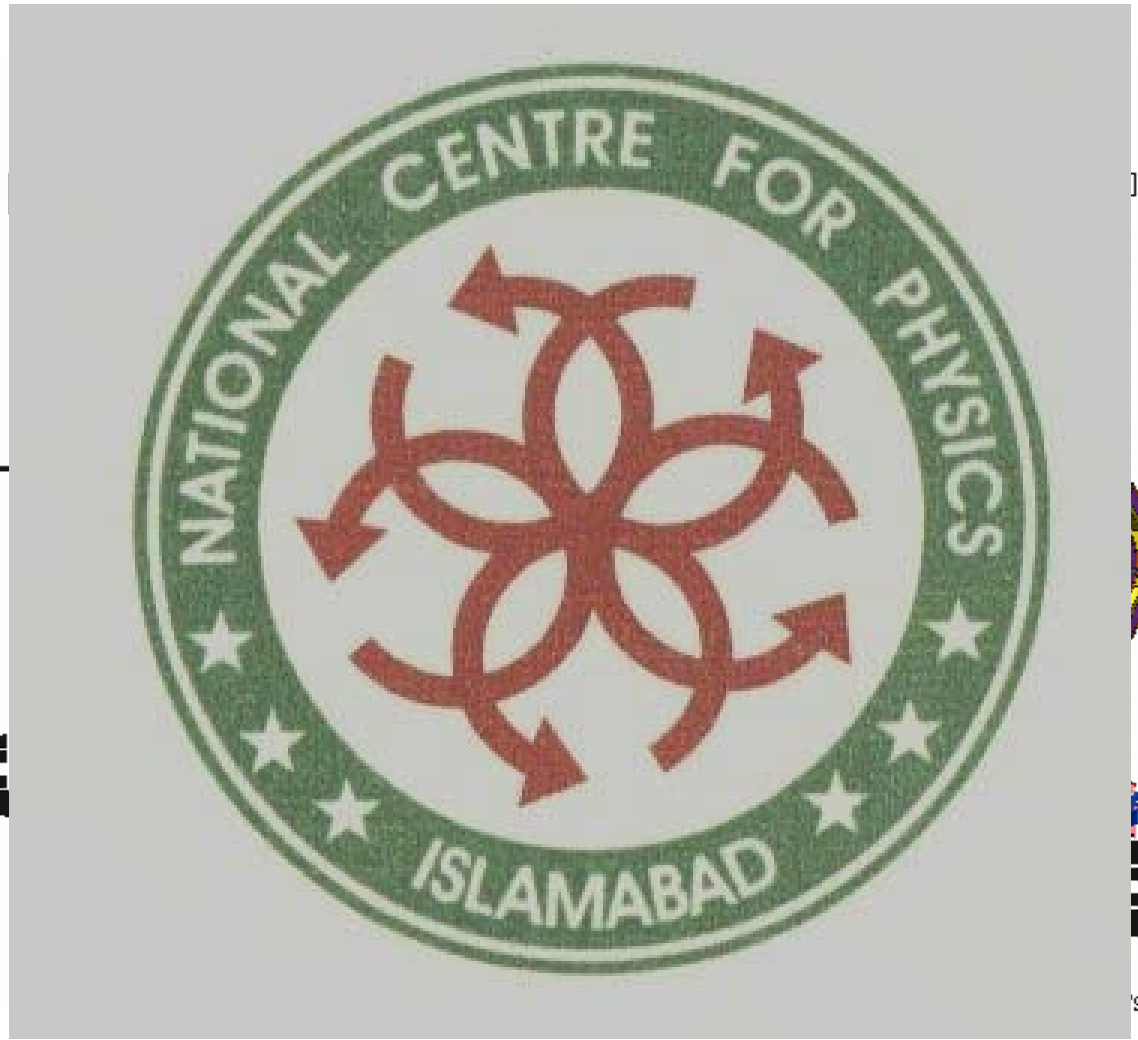
Advanced Scientific Computing

(ASC) at NCP

Presented By

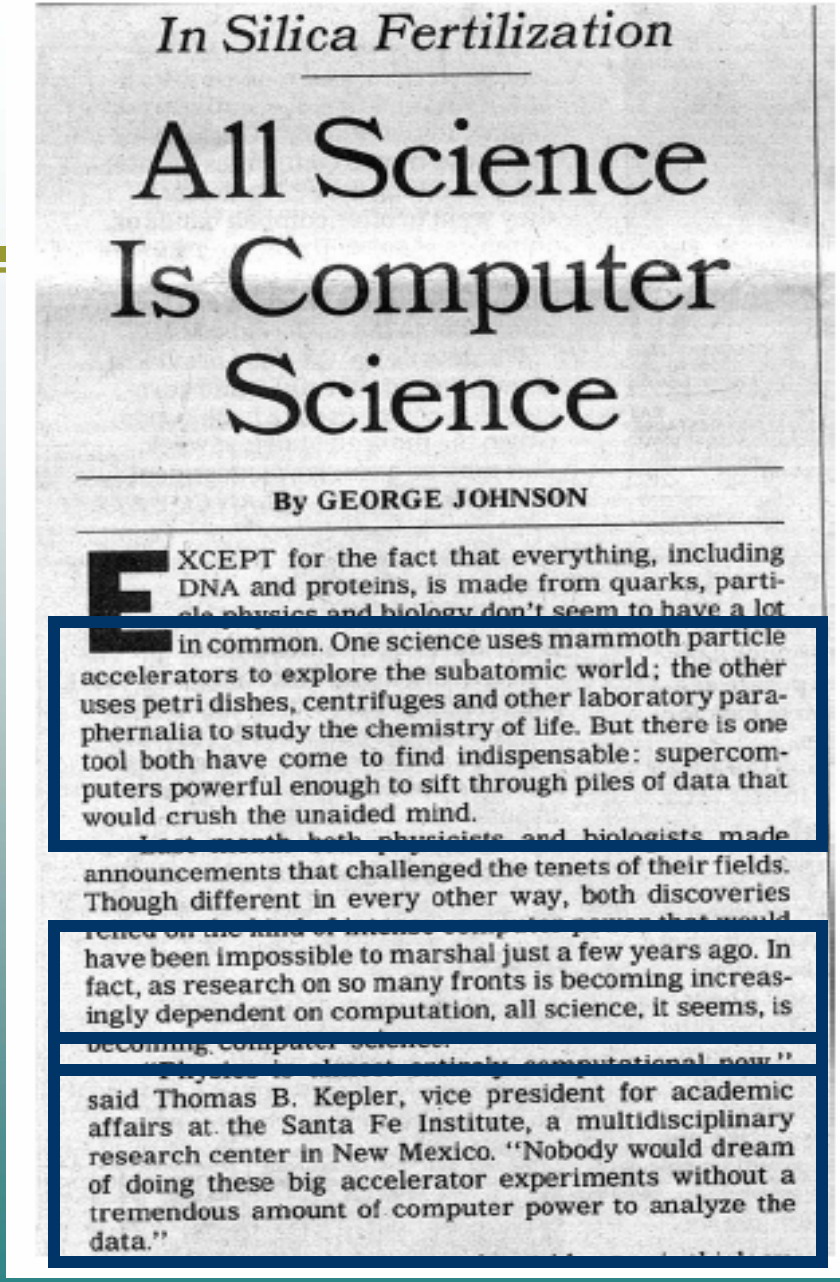
Mehnaz Hafeez



Total weight :
Overall diameter :
Overall length :
Magnetic field :

Y
Z
X

'97 JLB.PP



In Silica Fertilization

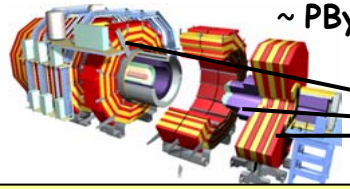
All Science Is Computer Science

By GEORGE JOHNSON

EXCEPT for the fact that everything, including DNA and proteins, is made from quarks, particle physics and biology don't seem to have a lot in common. One science uses mammoth particle accelerators to explore the subatomic world; the other uses petri dishes, centrifuges and other laboratory paraphernalia to study the chemistry of life. But there is one tool both have come to find indispensable: supercomputers powerful enough to sift through piles of data that would crush the unaided mind.

Last month both physicists and biologists made announcements that challenged the tenets of their fields. Though different in every other way, both discoveries relied on the kind of intense computer power that would have been impossible to marshal just a few years ago. In fact, as research on so many fronts is becoming increasingly dependent on computation, all science, it seems, is becoming computer science.

"Physics is almost entirely computational now," said Thomas B. Kepler, vice president for academic affairs at the Santa Fe Institute, a multidisciplinary research center in New Mexico. "Nobody would dream of doing these big accelerator experiments without a tremendous amount of computer power to analyze the data."



~ PByte/sec

Online System

~ 100 MByte/sec

Bunch crossing per 25 nsec.
100 triggers per second
Event is ~1 MByte in size

~622 Mb/sec
or Air Freight

Tier 0 +1

Offline Farm,
CERN Computer
Center > 20 TIPS

Tier 1

France Regional Center

UK Regional Center

Italy Regional Center

Fermilab
~ 4 TIPS

~ 2.4 Gbit/sec

Tier 2

Tier2 Center
~ ~ 1 TIPS

Tier2 Center
~1 TIPS

~622 Mbit/sec

Tier 3

Institute
~0.25TIPS

Physics data cache

100 - 1000
Mbit/sec

Tier

Workstations

Tier -2 Centre in Pakistan

Creation of Advance Scientific Computing group at NCP

is taking one step further , not only to strengthen

our scientific collaboration with CERN

but also heading towards getting an

OBSERVER STATUS *for Pakistan.*

ASC Mandate

- Solve computing issues for scientists and researchers
- Human resource development issues
- Initiate In-house Projects
- Develop National/International Collaborations in computing
  - NCP-CERN projects
    - Production
    - LCG
  - NCP-Caltech projects

ASC Mandate

- Investigation of new or alternative hardware solutions for scientific computing
- Give scientists and society the tools and the infrastructure they need to solve the most important problems of our time

ASC Mandate

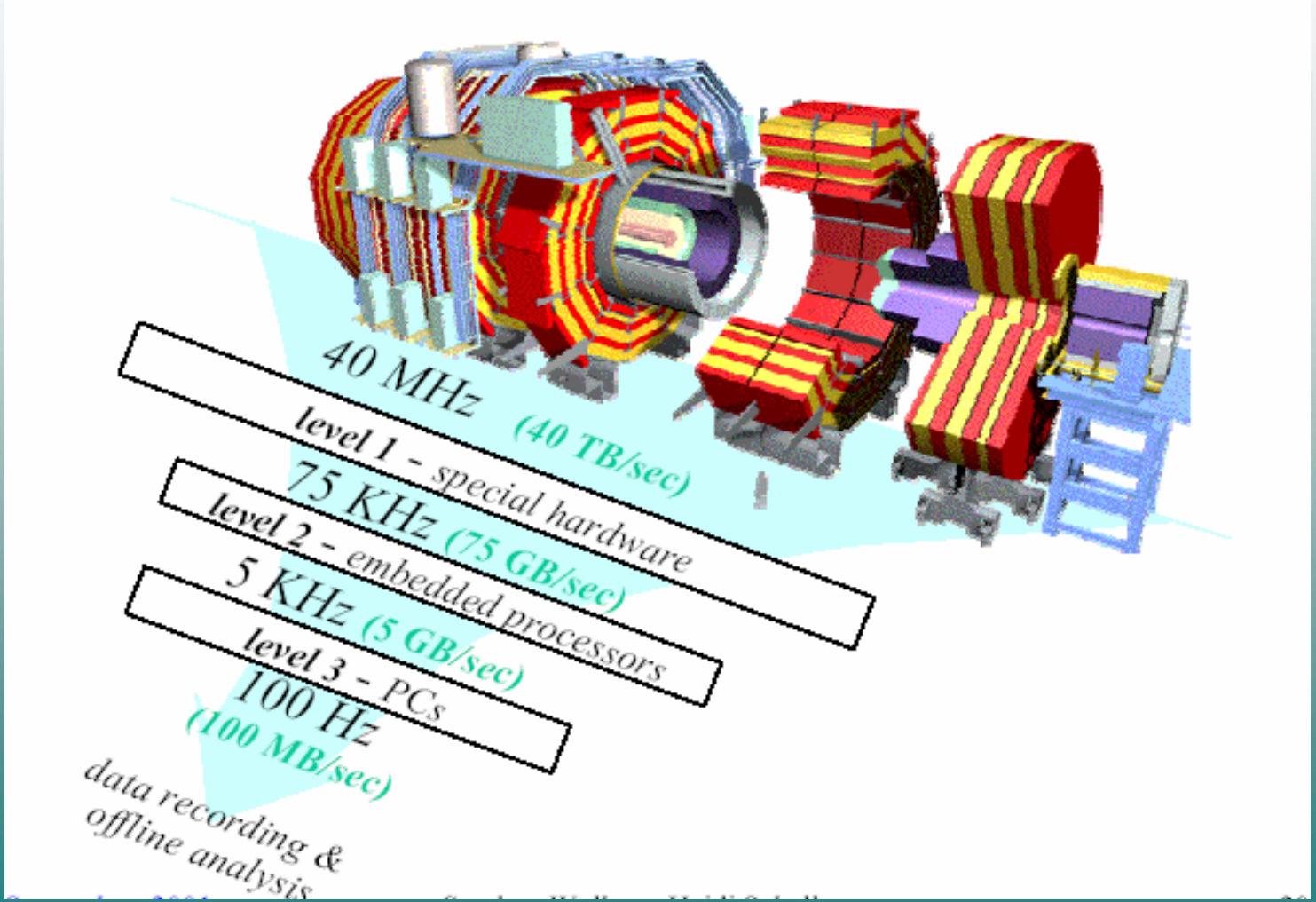
- Facilitate GRID computing environment
- Participate in the LCG to link communities together as grids
  - To build the network infrastructure
  - To develop Grid software
  - To work on the middleware tools

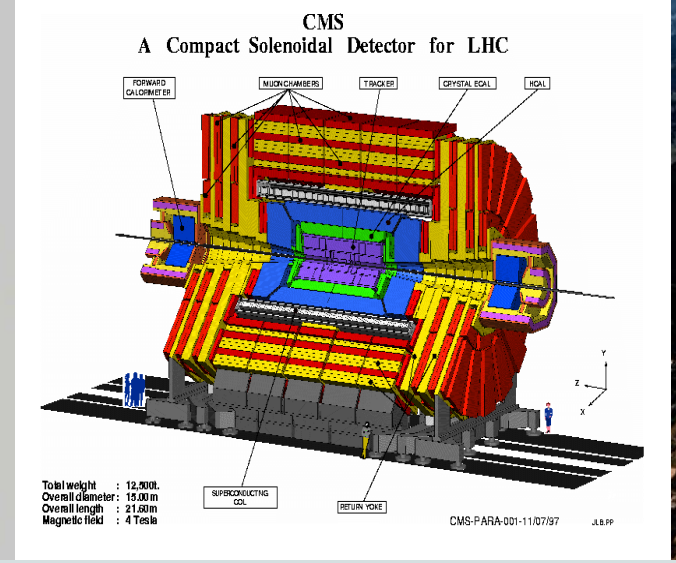
For our High Energy Physics (HEP) community in Pakistan in general and later bring in other research groups in Pakistan especially Medical Sciences groups.

Why we need Grid?

- Establishment of Grid Node is a part of computing required for the LHC (Large Hadron Collider) at CERN, Geneva
- CMS is one of Data Intense HEP (High Energy Physics) Detector of LHC
- Several hundred million channels are used for readout
- Each channel is one data bit

Why we need Grid?





NCP is committed to work on the establishment of a Grid node in Pakistan. It will allow Physicists in Pakistan to access the data from LHC in real time.

Who will use Grid?

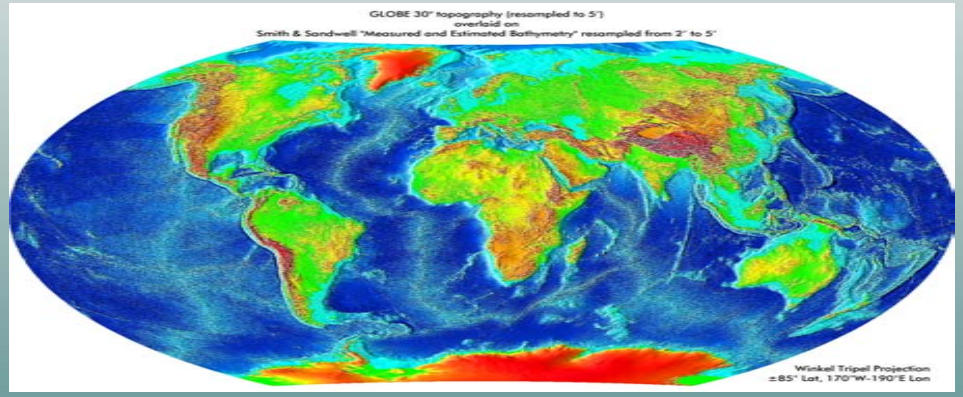
Computational Scientists

Experimental Scientists

Engineers

Associations

World



Consumers

Corporations

Training and education

Environment

Government

Status

Hardware Resources		
	CPU	19
	Memory (GB)	7.55
	Disk (TB)	1.55

Network Resources	
	NCP 128 kbps downlink / 64 kbps uplink (within premises) 512 kbps (offshore)

Status

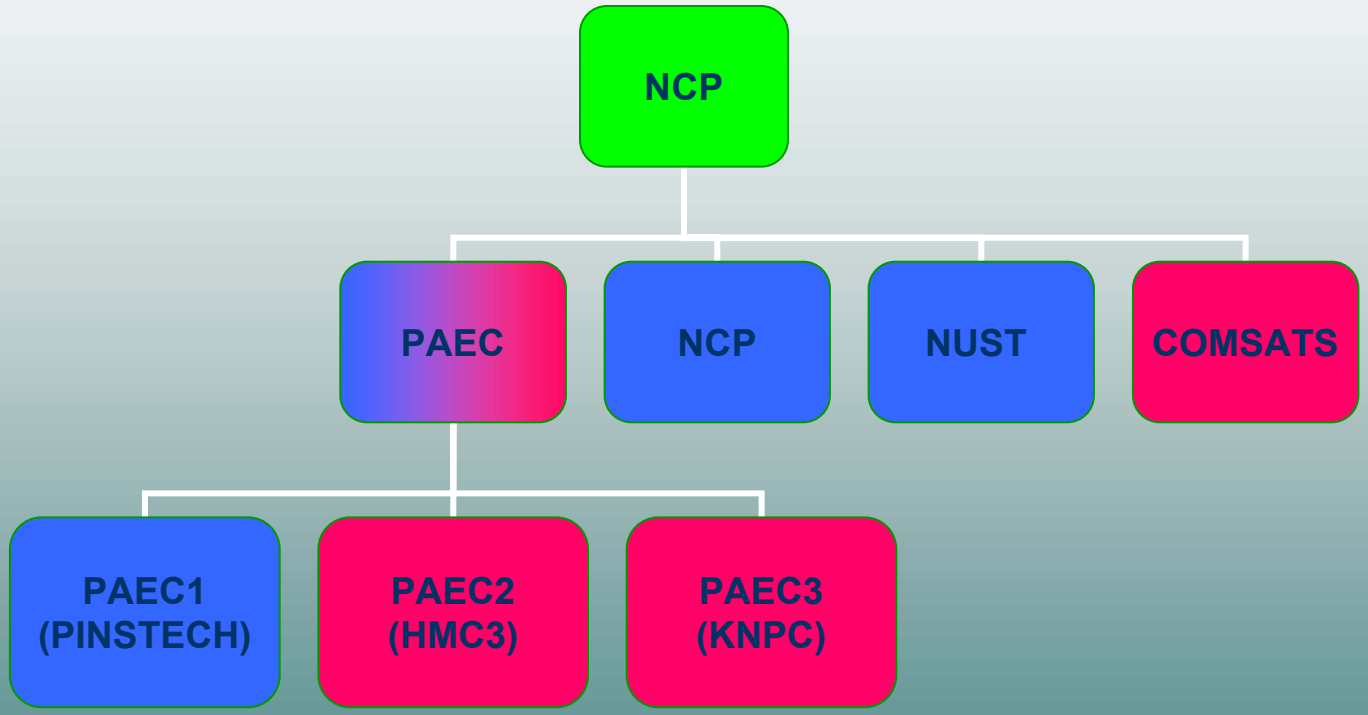
- ASC has taken a lead in CMS data production among all other of its working partners
- Produced 400k events using CMS valid production chain
- NCP is a Regional Center for CMS Production in Pakistan

Status

NCP Responsibilities

- Verify the resources committed by each PC
- Help PCs to get more resources for Production
- Request production assignments from CMS production coordinator
- Redistribute assignments among PCs
- NCP will be a Data Warehouse for LHC Data

Status


NCP
PAEC
NCP
NUST
COMSATS
PAEC1 (PINSTECH)
PAEC2 (HMC3)
PAEC3 (KNPC)

Status

- Regional Center Managar:
  Mehnaz Hafeez (hafeez@ncp.edu.pk)

- Production Center Managers:
  NCP:
  Mr. Ateeq Baig (ateeq@ncp.edu.pk)

  PAEC1: PINSTECH
  Ms. Naheed Batool
  (naheed@pinstech.org.pk)

Status

PAEC2: HMC3

Mr. Sajid Imtiaz
(hmc3@micro.net.pk)

PAEC3: KNPC

Mr. Rehan Siddiqui
(cern_rehan@hotmail.com)

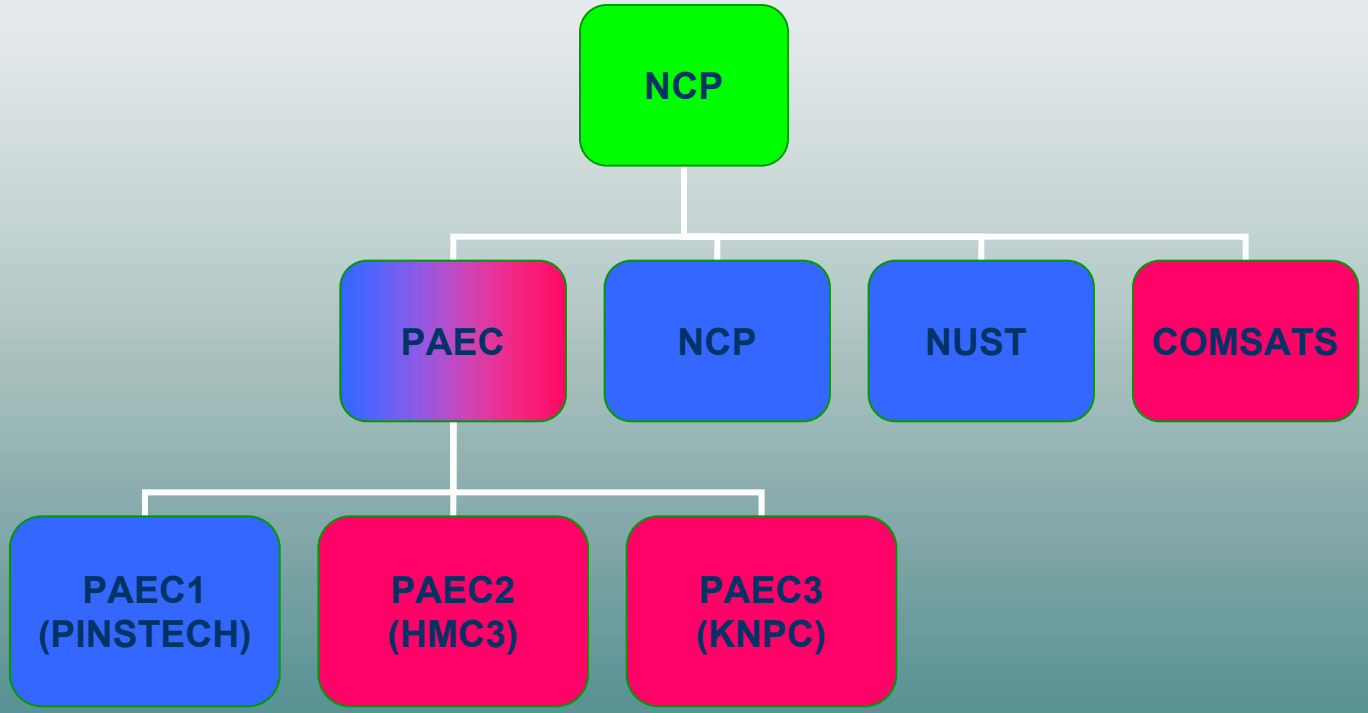
NUST:

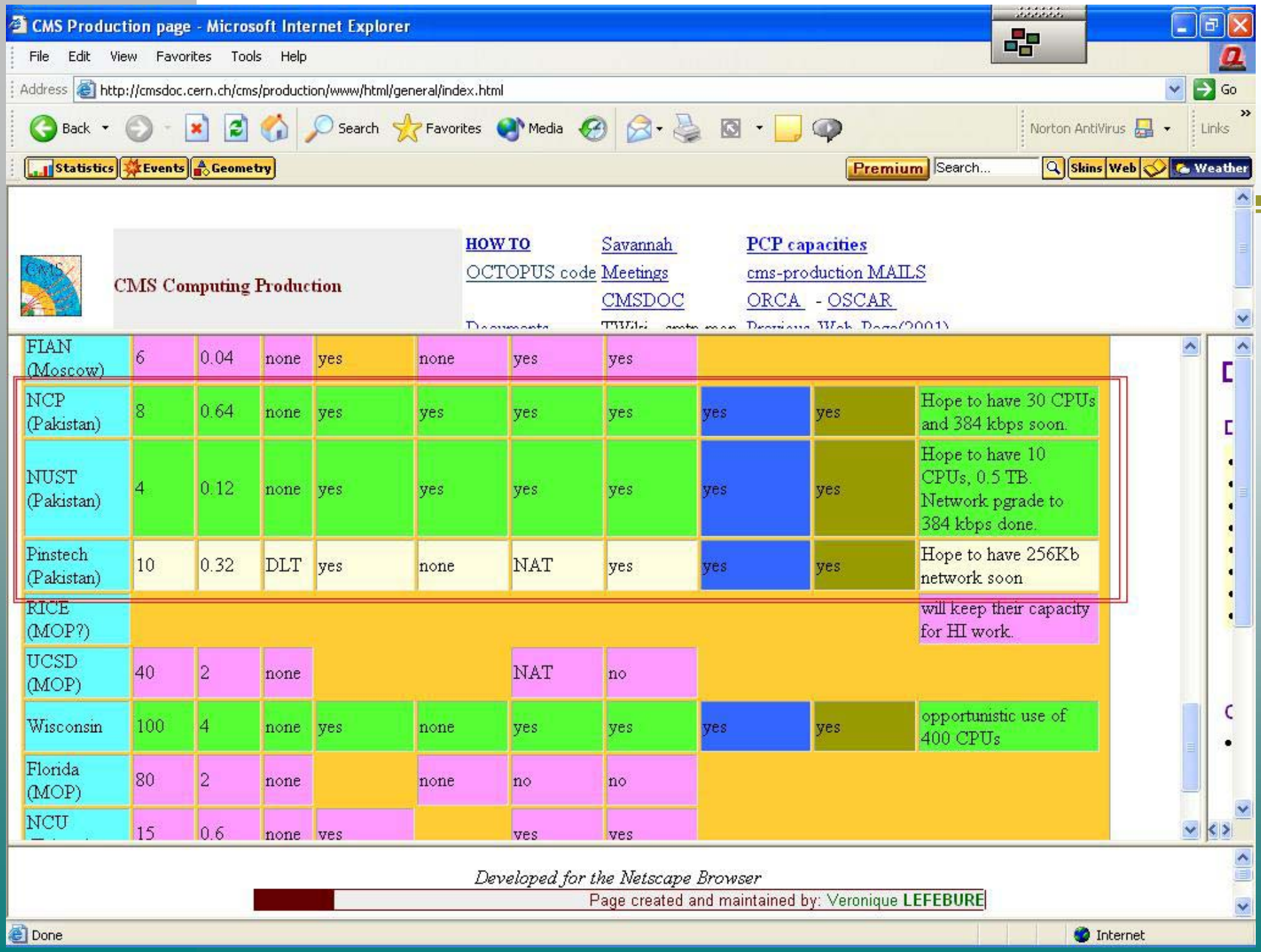
Mr. Kamran Munir
(kamran.munir@niit.edu.pk)

COMSATS:

Mr. Usman Ahmad (usman@ncp.edu.pk)

Status

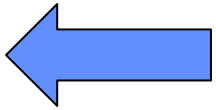

NCP
PAEC
NCP
NUST
COMSATS
PAEC1 (PINSTECH)
PAEC2 (HMC3)
PAEC3 (KNPC)


CMS Production page - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://cmsdoc.cern.ch/cms/production/www/html/general/index.html
Go
Back Search Favorites Media
Norton AntiVirus
Links
Statistics Events Geometry
Premium Search... Skins Web Weather
CMS Computing Production
HOW TO
Savannah
PCP capacities
OCTOPUS code
Meetings
cms-production MAILS
CMSDOC
ORCA - OSCAR
FIAN (Moscow) 6 0.04 none yes none yes yes
NCP (Pakistan) 8 0.64 none yes yes yes yes yes yes Hope to have 30 CPUs and 384 kbps soon.
NUST (Pakistan) 4 0.12 none yes yes yes yes yes yes Hope to have 10 CPUs, 0.5 TB. Network pgrade to 384 kbps done.
Pinstech (Pakistan) 10 0.32 DLT yes none NAT yes yes yes Hope to have 256Kb network soon
RICE (MOP?) will keep their capacity for HI work.
UCSD (MOP) 40 2 none NAT no
Wisconsin 100 4 none yes none yes yes yes yes opportunistic use of 400 CPUs
Florida (MOP) 80 2 none none no no
NCU 15 0.6 none yes yes yes
Developed for the Netscape Browser
Page created and maintained by: Veronique LEFEBURE
Done
Internet

Status

Production Status	egamma	muon	btau	jetmet	higgs	Total events
NCP	400k			13k	100k	513k
PINSTECH		1k	1k	120k		122k
NUST				200k		200k

CMS Data Challenges

- 2000/2001
  - Verify code, bring up production worldwide, prepare for DAQ TDR
- 2002
  - DAQ TDR massive production and analysis
- 2003/4 (DC04) 
  - First Year of Physics TDR, GEANT4 in Production
  - New Persistency, First truly GRID dependant challenge
  - Verify model and components for CMS Computing TDR
- 2004/5 (DC05)
  - Verify LCG2 Prototype in time for LCG TDR
- 2005/6 (DC06)
  - Final Readiness Check, all Software and Computing systems
- 2007
  - First Data. Ready for new Physics in first few fb-1

Data Challenge DC04

- As defined to the LHCC, the milestone consists of:
  - CS-1041 1 April 2004 5% Data challenge complete (Now called DC04)
  - *The purpose of this milestone is to demonstrate the validity of the software baseline to be used for the Physics TDR and in the preparation of the Computing TDR. The challenge comprises the completion of a "5% data challenge", which successfully copes with a sustained data-taking rate equivalent to 25Hz at a luminosity of 0.2×10^{34} cm^{-2} s^{-1} for a period of 1 month (approximately 5×10^{7} events). The emphasis of the challenge is on the validation of the deployed grid model on a sufficient number of Tier-0, Tier-1, and Tier-2 sites. We assume that 2-3 of the Tier-1 centers and 5-10 of the Tier-2 centers intending to supply computing to CMS in the 2007 first LHC run would participate to this challenge.*

Status

Immediate learning/benefits

- Our people are getting practical exposure of
  - Batch Scheduler
  - Manage data processing
  - Data placement activities
  - Monitoring and status reporting
  - File transfer services
- Network

  128kbps -> 512 kbps

Status

- Commodity Computing has a great deal to offer
  - Cheap CPU
  - Fast network I/O
  - Fast Disk I/O
  - Cheap Disk
- Open Source Software

Status

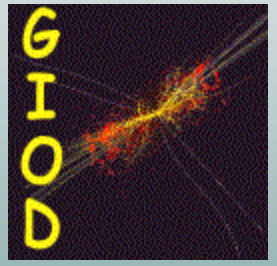
- Pakistan – an individual and a computer
- West – an individual in a multi-user environment
- Grid – an individual in a multi-institute environment

Status

- Major challenges associated with:
  - Communication and collaboration at a distance
  - Distributed computing resources
  - Remote software development and physics analysis
- R&D: New Forms of Distributed System

Status

- California Institute of Technology (Caltech)
  - Grid Enabled Analysis (GEA) Tools
  - Clarens



clarens

Status

- LCG official site partner
- Identified 6 – 7 different projects
  - Grid Deployment Group
  - Fabric Management
- Contribution in LCG middleware.

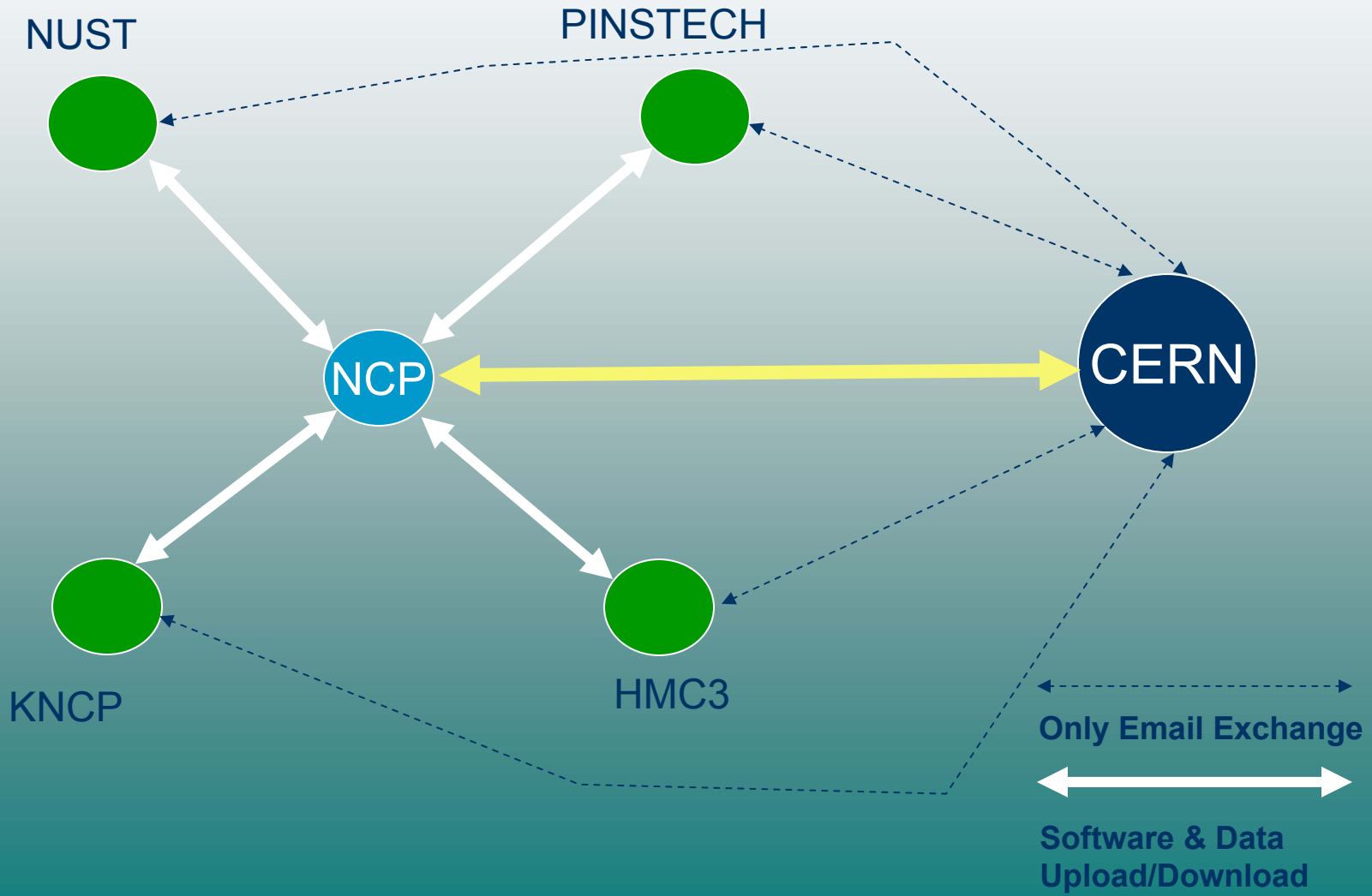
Status

- Video Conferencing, NCP VRVS reflector is up and running.
- Seminars
- Training Program
- Visitors program
  - Universities & Scientific Organizations
  - Egypt, person (stay for a month)

Status

Grid Technology Workshop

Future


NUST
PINSTECH
NCP
CERN
KNCP
HMC3
Only Email Exchange
Software & Data
Upload/Download

Future

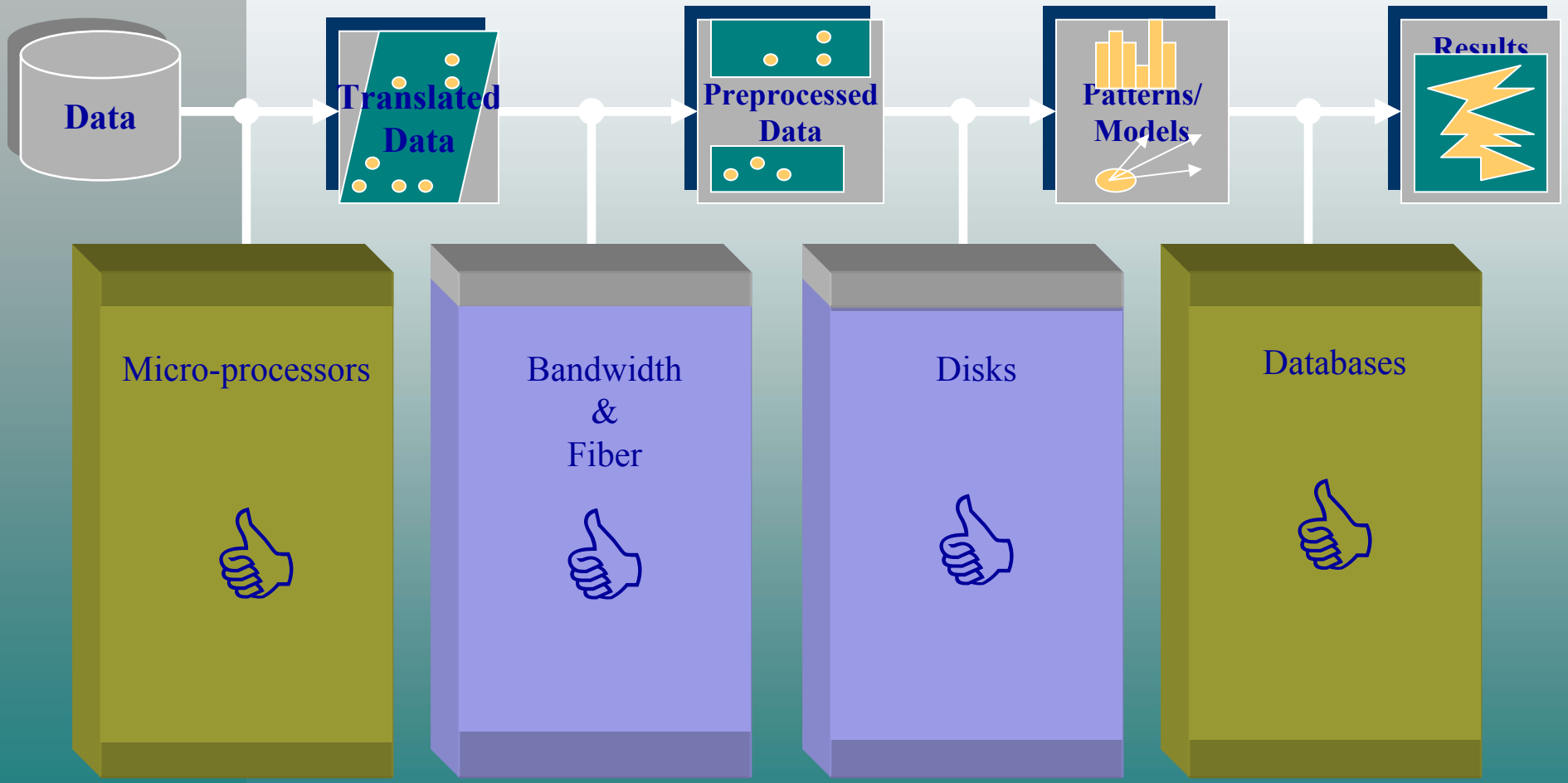
- Implementation of LHC Computing Grid (LCG)
- Management of LINUX based PC Cluster
- Development of Grid enabled software
- Contribution to LCG
- Expansion in hardware resources
- Connectivity

Advanced Scientific Computing

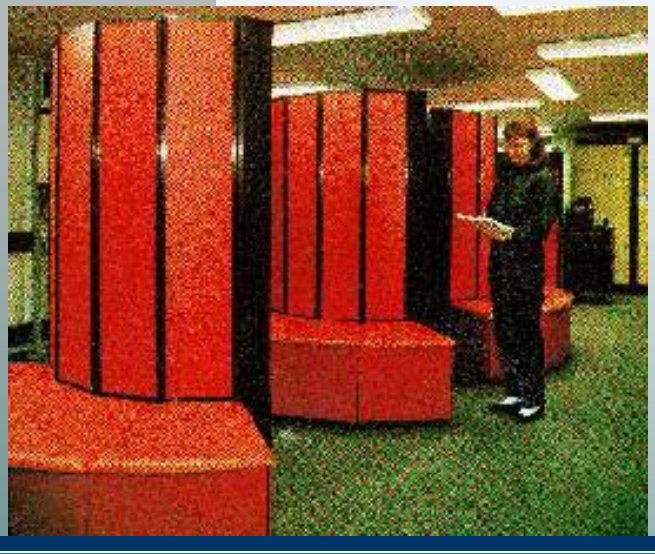
(ASC) at NCP

Thank You

What are Grids?



What are Grids?



with the 17 years difference

NCSA's first Cray X-MP

- Purchased: 1986
- Cost: $8 million
- Separate electrical line to power, 60,000 watts
- Special cooling
- No graphics
- Connected to other super-computer: 56 Kb/s

Today a child's Nindtendo-64 Video Game

- MIPS microprocessor
- Cost: $150
- 5 watts
- Interactive 3D graphics
- 64 Kb/s ISDN lines in homes

What are Grids?

These are persistent environments that enable software applications to integrate instruments, displays, computational and information resources that are managed by diverse organizations in widespread locations.

What are Grids?

"The vision behind GRID is a computing environment where anyone can plug in from anywhere and access any resource on the grid. And users would be able to rent processing power as well as software resources.."

Status

Human Resources		
	Administrators	2
	Software Engineers	3
	Technical Assistants	2