

# ITPB Cyberinfrastructure for computation working group report

## **Members:**

Warren B. Mori

Jim Davis

Bill Labate

Mitch Creem

Russ Caflisch

Jackson Beatty

Xua Hue



# What is our definition of CI for Computation?

- Intellectual
  - System Admin
  - Applications software support
- Data Center
  - Environmentals
- Compute Nodes
  - CPU's
  - Network
- Storage
- Data curation and archival is part of another working group

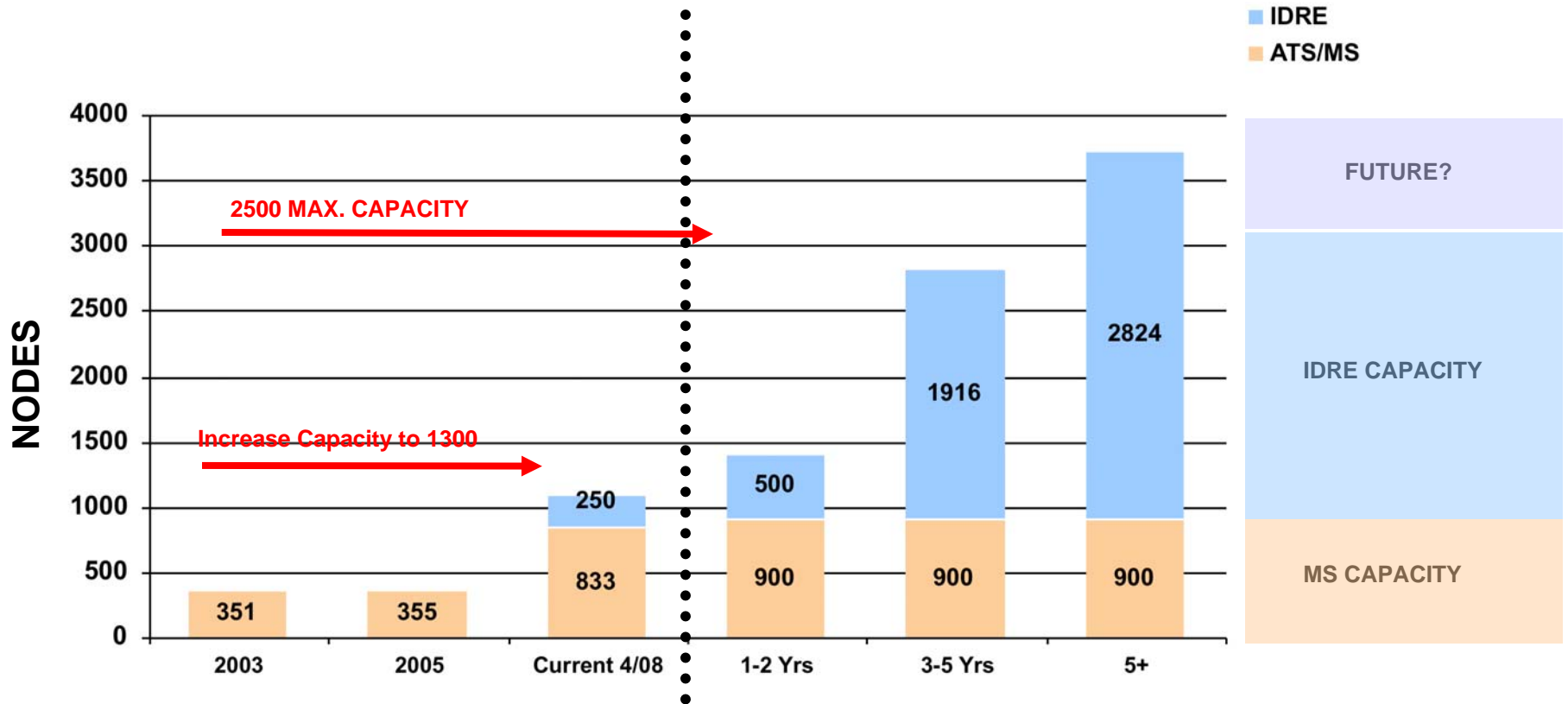


## Campus need for CI for computation

- **Researchers are acquiring nodes**
  - Where do they put them?
  - How are they managed and maintained?
- **Efficient use of scarce resources**
  - Power/cooling
  - unused cycles
  - space
  - people
- **Intellectual**
  - Optimized code uses less power/throughput
  - Need a pipeline to larger national resources
  - Grid is needed for linked UC data centers
  - Queuing systems for shared cluster
- **Recruitment/Retention**

# Need for IDRE data center: Anticipated growth requires increased capacity

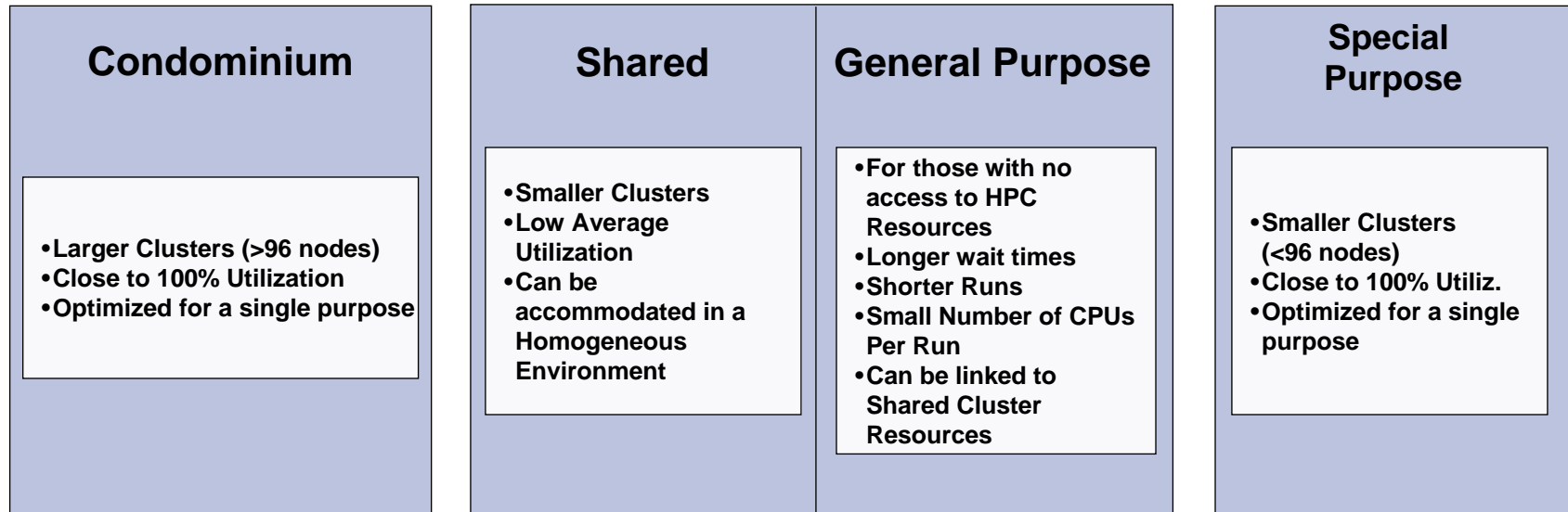
Committed/Waiting ← → Projected



# Data center concept

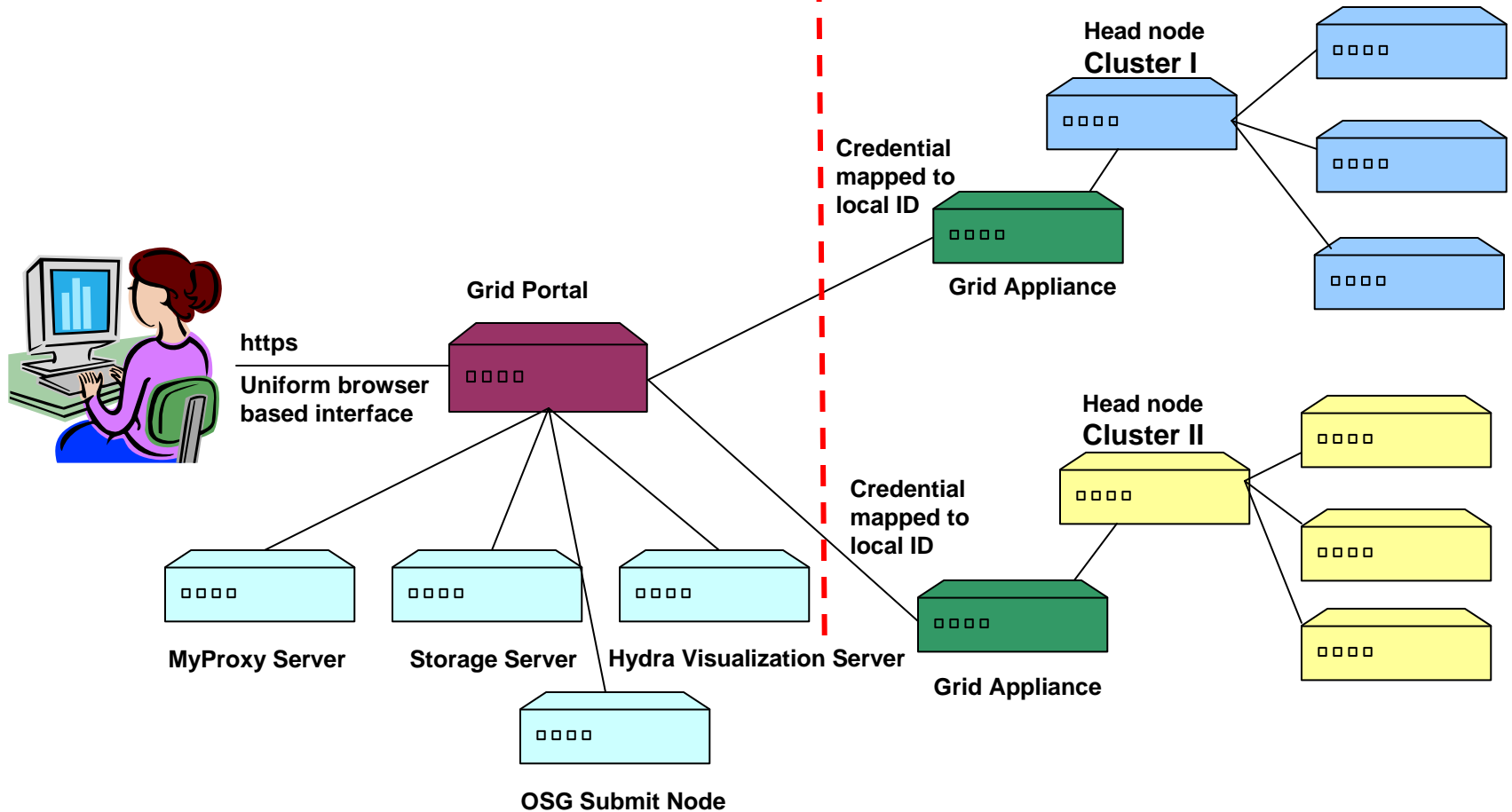
improved efficiency and increased visibility

- Hierarchy of node types:



# Single campus concept

IDRE shared cluster plus individual clusters linked via Grid



## Potential of shared cluster: HoffmanII partnership between campus and researchers



Configuration for Hoffman2 when completed:

- Peak performance of over 73 TFlops
- Total memory of 12.4 terabytes
- 512TB of storage
- 13TB of scratch space
- Total power of 600 kilowatts
- Over 11 miles of cabling
- Over \$600K a year for power
- 200 tons of air conditioning
- 2,500 square feet of floor space
- 1,600 nodes
- 10,400 processor cores
- Total weight of over 32 tons in 40 racks
- DDR Infiniband interconnect for message passing traffic
- Gigabit Ethernet for storage administration usage
- Total campus/researcher investment of approximately \$9.3M



# Objectives that drive policy for CI

- UCLA must be a leader
- Limit stand alone systems
- Limit unused cycles
- Make efficient use of “resources”
- Adhere to standard configurations
- Increase visibility to the campus
- Emphasize value added
  
- These can be accomplished with incentives that have minimum cost



## Value Added of the shared model

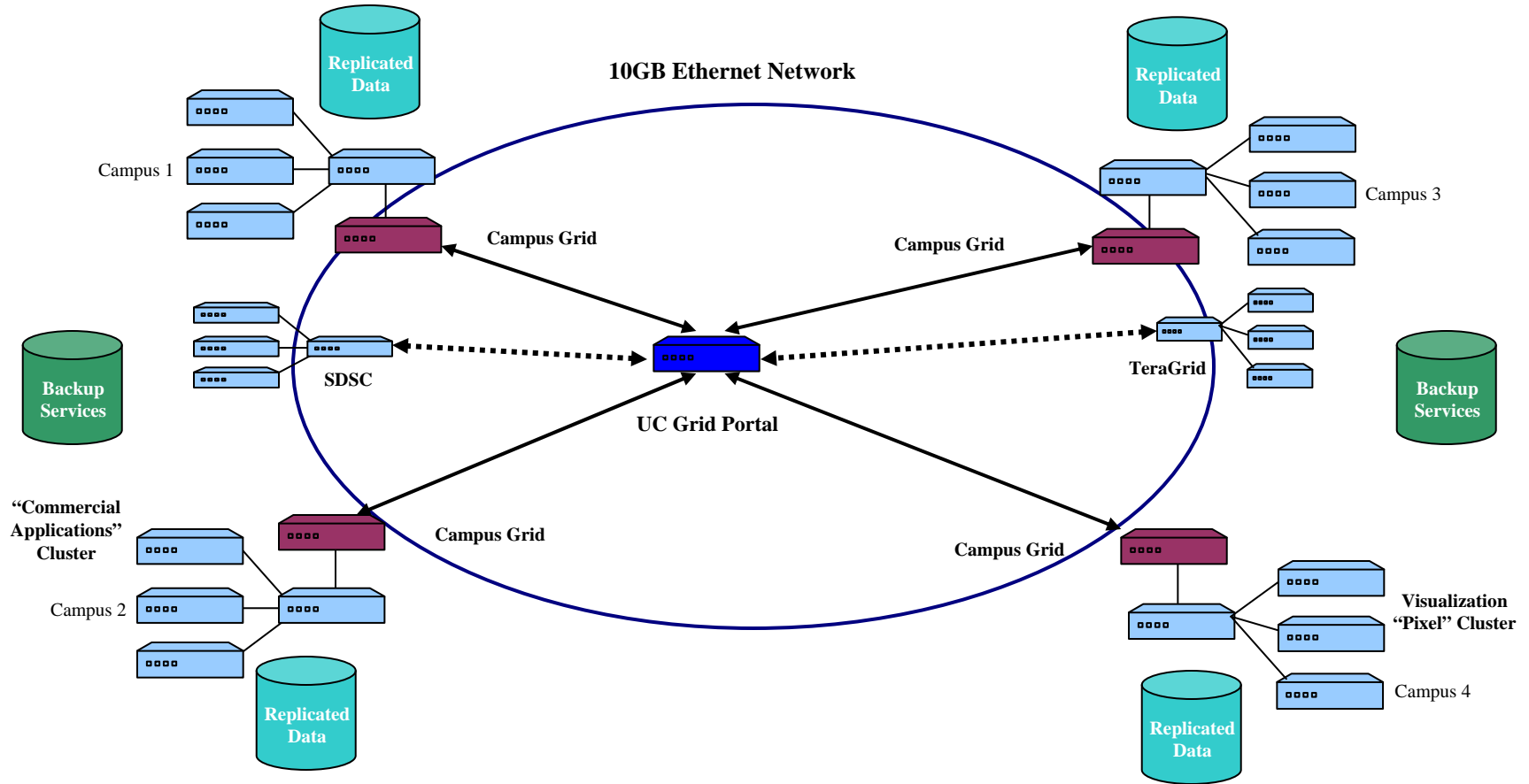
- A large cluster is more visible: TOP 500 list
- Unused cycles are available for others
- Maximize impact of research dollars
- Access to larger system for larger simulations
- Access to a wider spectrum of applications



# UC Concept

- Regional data centers connected via a Grid
- Goal is to take advantage of economy of scales and make efficient use of resources
- Overflow or surge capacity--computational or storage capacity
- UCLA is positioned for leadership. This should be exploited for the benefit of UCLA.

# UC-Wide Cyberinfrastructure Vision





# Business Model: Who pays for what?

1. Researchers

Nodes  
Storage

2. Recurring recharges

Minimized as much as possible but needed  
for stand alone systems

3. Campus

environmentals  
System Admin  
Networking  
Occasional seed contributions for nodes

4. State

Regional data centers  
General purpose nodes

5. Donors/Industry

Strategic fund raising campaigns



# Recommendations:

1. Continue to pursue the shared cluster model.
2. Continue to be active and leading participants in the UC regional data center model.
3. Continue to be active and leading participants in the UC grid portal development.
4. Continue to explore business models.
5. Begin to investigate cloud computing.