Center for Advanced Research Computing
Internal Advisory Board
Wednesday, October 11, 2017, 3 p.m.

MINUTES

Present: **Patrick Bridges** - Interim Director of CARC; **Karl Benedict, Ph.D.** - Associate Professor, Director of Research Data Services. College of University Libraries and Learning Sciences; **Jed Crandall, Ph.D.** - Associate Professor, Computer Science; **Miguel Gandert, M.A.** - Director, Interdisciplinary Film and Digital Media; **Hua Guo, Ph.D.** - Distinguished Professor, Department of Chemistry and Chemical Biology, and Department of Physics and Astronomy; **Patricia Henning, Ph.D.** - Associate Vice President of Research; **Keith Lidke, Ph.D.** - Associate Professor, Physics & Astronomy; **Barbara McCrady**, Director of CASAA and Professor, Psychology; **Brian Pietrewicz, M.B.A.** - Interim Deputy CIO, Information Technologies; **Edl Schamiloglu, Ph.D.** - Distinguished Professor, Electrical and Computer Engineering; Associate Dean for Research, School of Engineering; **Gregory Taylor, Ph.D.** - Director, Long Wavelength Array; Director, Center for Astrophysical Research and Technology; Professor, Department of Physics and Astronomy; **Lee Taylor, Ph.D.** - Associate Professor, Biology; **Tracy Wenzl** - CARC Unit Administrator

1) CARC Strategic Plan & cost model
    a) Presentation by Patrick Bridges, Interim Director
        i) State of CARC
        ii) Strategic Plan
        iii) Presentation of cost model proposal

            IAB agreed that CARC should move forward with the cost model presented. It will be presented to additional stakeholders before adoption.
2) Additional IAB discussions, feedback
    a) Discussion of who “owns” research storage at UNM – parts currently managed by Information Technologies, University Libraries and CARC. UL and CARC have been meeting to discuss plans to manage short- and longer-term research data storage
    b) Staffing at CARC – suggested that graduate students continue to be used for first-touch technical support, with the development of annual boot camp to allow experienced students to share knowledge with new hires.
    c) Changes to the Computational Science and Engineering (CSE) certificate program were discussed, with an eye to streamlining the coursework and creating informal tracks or clusters of courses focused on a particular area of emphasis.

Center for Advanced Research Computing
Internal Advisory Board
October 11, 2017

AGENDA

1. CARC Strategic Plan & cost model
    - Discussion/Revision
    - Approval?
2. Additional IAB discussions, feedback

State of CARC

Strategic Planning, Service Baseline, Budget and Cost Models, and Challenges and Opportunities

Prof. Patrick G. Bridges

Interim Director

September 18, 2017

CARC's Mission and Vision

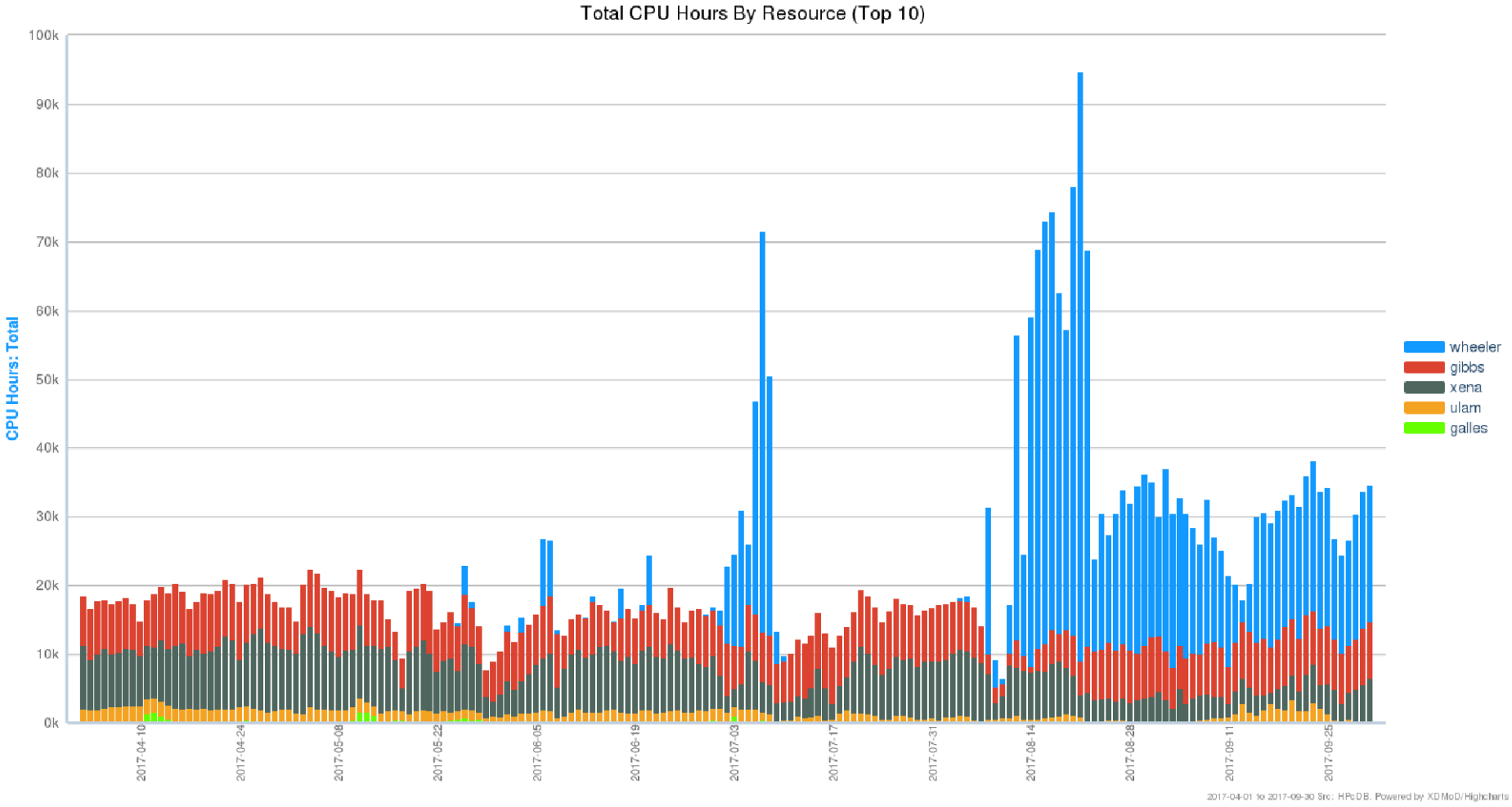
- **Vision**
  - The Center for Advanced Research Computing (CARC) is an interdisciplinary community at the University of New Mexico (UNM) that uses computational resources to create new research insights.
- **Mission**
  - To lead and grow the computational research community at UNM.
- **Note Two Aspects Implicit in CARC's Mission**
  - Service - providing cyberinfrastructure and associated support to the UNM community
  - Research – conducting research to improve the state of the art in cyberinfrastructure

CURRENT STATE AND RECENT CHANGES

Old systems retired, new systems coming online

- **Making way for new(er) systems**
  - Ulam, Pequena, Nano, and Metropolis retired
  - RSC phase 1 storage to be retired
- **Installing newer systems**
  - Wheeler phase 1 (140 8-core/48GB nodes) already available
  - Wheeler phase 2 (doubling system size) available next week
  - Taos condo cluster (initial investment by Biology/CETI) assembled, first availability to CETI users in 2-3 weeks.
  - 167TB of enterprise storage capacity being ordered (1/3rd for CARC home directories, 2/3^{rds} for Edgar research project)
- **Repurposing existing systems**
  - RSC phase 2 and 3 will be connected to the new CARC/Libraries storage system, with a portion held back as spares
  - Virtualizing contributed DataOne and CRF systems to increase capacity to support custom workloads

Example: Wheeler deployment dramatically increasing system capacity



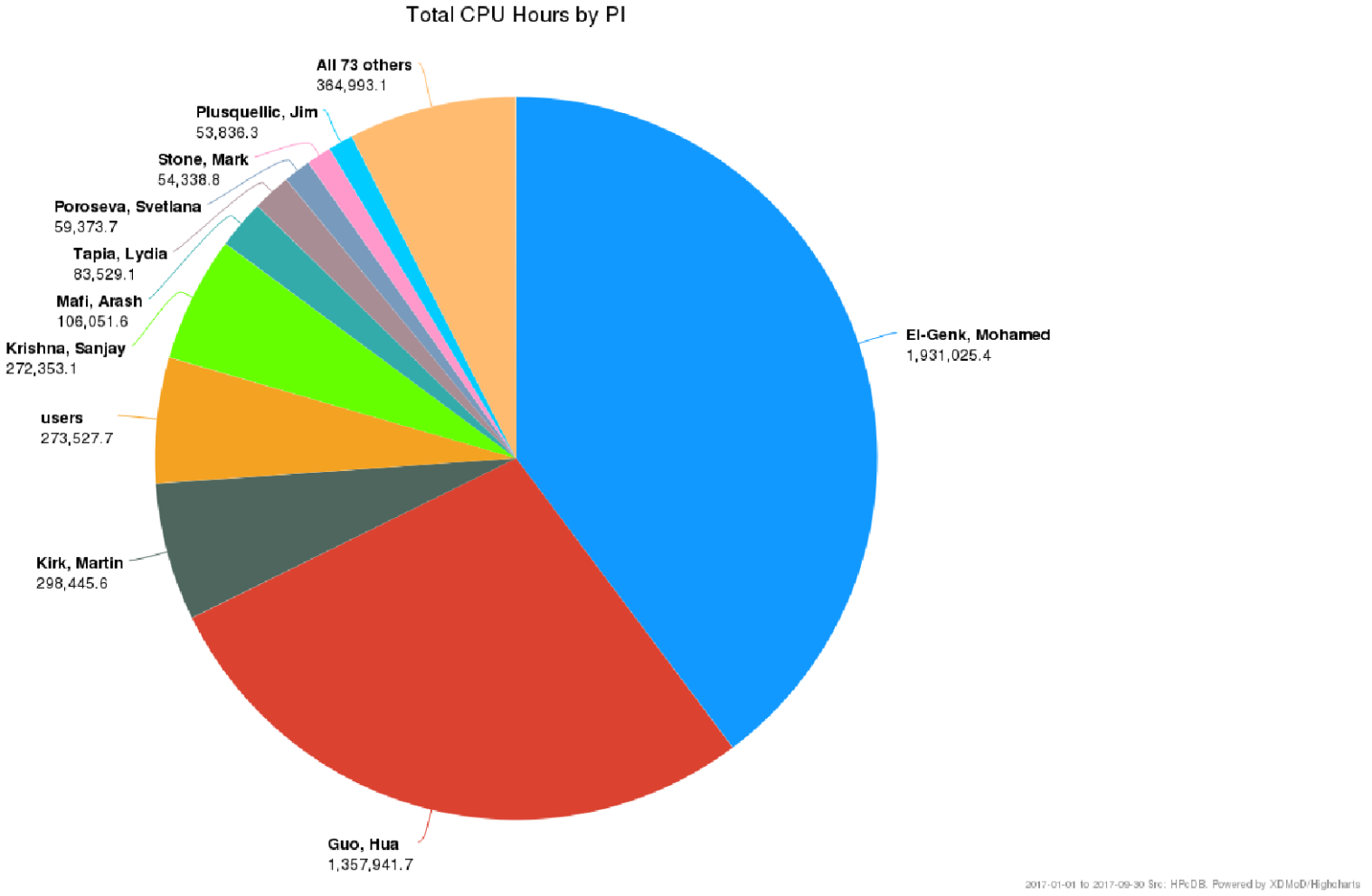
CARC Users: Present and Future

- **Majority of current CARC users are traditional HPC users**
  - "Advanced research computing" - E.G. fluid dynamics, material science, comp. biology, nuclear engineering, numerical analysis, etc.
  - Need standard systems and are generally computationally savvy;
    - Demand significant computational resources
    - Rarely need custom software environments
    - Need help starting, but then are relatively self-sufficient
- **The boundary of "advanced" research computing is increasingly blurry**
  - Big data research and cloud computing bringing "advanced" into social science, economics, literature, communications, medicine, etc.
  - Computation now essential to every domain of scientific research
  - Non-traditional research computing requires custom software
  - Non-traditional research computing requires frequent support

CARC Users: Present and Future (2)

- **CARC staffed and resourced to support approximately 80 research projects and 200 users**
  - Provided support is primarily for traditional HPC users
  - Currently have 275+ users, ~20% of which are courses or north campus (i.e. not budgeted for by F&A)
- **CARC needs to increase support for emerging research computing disciplines, tools, and techniques**
  - New programming models: Jupyter, Apache Spark
  - New tools: RedCAP, Web Portals (a la Galaxy)
  - HIPAA, FERPA, and Export Controlled-compliant research
- **UNM Research Strategic Plan Goals for CARC:**
  - Increase number of users
  - Develop funding model not purely dependent on F&A

A few large users fill hardware capacity, many small users fill staff capacity



Example Recent Research Projects

- Mountain Lions on the Edge: Integrating Conservation into Urban Planning through Predictive Modeling – Prof. Bruce Milne, UNM Department of Biology
- A High-fidelity Model for Wind Farms – Prof. Sang Lee, UNM Department of Mechanical Engineering
- Performance Optimization of LANL Multi-Physics Applications – Prof. Patrick Bridges, UNM Department of Computer Science
- Economic Network Simulations – Prof. David Dixon, UNM Department of Economics
- Multiscale Mechanistic Model to Study Nanotherapy Delivery in Tumors – Dr. Elaine Bearer, UNM HSC Department of Pathology

BUDGET AND FUNDING

CARC Funding Model

- **Base funding from main campus sources to provide baseline level of free service to all CARC users.**
  - What services are in this baseline?
  - At what level is each baseline service provided?
- **Historical informal cost model for users needing above baseline-compute and storage capabilities**
- **External cyber-infrastructure funding to further increase center capabilities**
  - NSF CDES&E and CRI proposals in progress
  - SSI^2 and Cyber-training proposals delayed due to NSF withdrawal of opportunity
- **GOAL: Formalize informal models so that they can be budgeted and publicized**

Challenge: Growing Baseline and Funding

- **Free baseline support requires increased support with more users**
  - CARC (as with most of UNM) is already stretched thin on staffing
  - Economies of scale and grants help on hardware/hardware staffing
  - User support staffing is the critical problem – more users means more support needs, and there aren't easy scale efficiencies here
- **No easy or complete solutions, but some possible directions**
  - Top slice F&A – Limited pot that CARC already draws significantly from
  - Developing IT finance model – Details unknown; perhaps Banner tax replaced with IT tax?
  - Grant funding from NSF OAC, DOD, DOE, and other external agencies – non-NSF is mostly equipment funding
  - Resource sharing with other units (e.g. UNM IT, Libraries RDS, User support by department/college research and IT staff, etc.)

CARC Budget Overview

- Budget largely supported by top-slice F&A from OVPR
- Budget includes research and service center components
- Approximately 60% of budget is technical staffing costs
- Infrastructure costs (storage, networking, and other upgrades) are the large majority of misc. expenses

Category	Technical Staffing	Admin Staffing	Miscellaneous Expenses	Total
Service Center	**$329,417.61**	**$88,010.06**	**$119,297.30**	**$536,724.97**
Research Center	$49,408.00	$34,144.74	$45,551.70	$129,104.44
F&A Funding	$378,825.61	$122,154.80	$164,849.00	$665,829.41
Grant Funding	$23,086.14	$0.00	$0.00	$23,086.14
Total Budget	**$401,911.75**	**$122,154.80**	**$164,849.00**	**$688,915.55**

UNM Indirect Support for CARC

- CARC indirectly supported by PPD for power/cooling and by periodic refresh/upgrades of data center facilities
- Neither of these costs are part of the CARC budget
- Only have approximate estimates of these costs

Service	Facilities Depreciation	PPD Power Costs
Compute Nodes	$19,055.56	$231,789.60
Enterprise Storage	$64.81	$525.60
Working Storage	$648.15	$5,256.00
Server Colocation	$3,629.63	$29,433.60
Total	**$23,398.15**	**$267,004.80**

STRATEGIC PLANNING

CARC's Mission and Vision

- **Vision**
  - The Center for Advanced Research Computing (CARC) is an interdisciplinary community at the University of New Mexico (UNM) that uses computational resources to create new research insights.
- **Mission**
  - To lead and grow the computational research community at UNM.
- **Note Two Aspects Implicit in CARC's Mission**
  - Service - providing cyberinfrastructure and associated support to the UNM community
  - Research – conducting research to improve the state of the art in cyberinfrastructure

CARC Strategic Plan

- **To fulfill our mission, CARC:**
  - Researches and provides access to high-end computing resources and associated infrastructure;
  - Offers specialized expertise and technical support;
  - Coordinates and collaborates with other UNM programs that support the community; and
  - Grows the collaborative user community through research, education, workshops, and outreach events.
- **Strategic Objectives**
  - Increase the number of research projects, grants, awards, publications, and creative works supported at CARC
  - Increase the number of people actively engaged with CARC

Strategies for Fulfilling Objectives

- Strategy 1: Conduct research to increase system accessibility, capability, ease of use, and supported disciplines.
- Strategy 2: Develop user support materials and courses offered by CARC and collaborating organizations.
- Strategy 3: Grow systems to better support users including adding condo and MRI-funded systems for specialized users, and inherited systems.
- Strategy 4: Create a collaborative user community that includes forums, seminars, colloquia, and other events that encourage collaboration.
- Strategy 5: Seek industry collaboration on research, education, and training.

Key Discussion Action Items (1)

Strategy 2: User Support (and others)

- Problem: Difficult to outreach to, engage, and support CARC users and to collaborate with main campus research IT personnel from current location
- Action Item: Relocate CARC staff to main campus to be closer to users; machine room and work space stays in Galles building
- Timeframe: Short (1-3 years)
- Status: In discussion with ECE chair about possibility of moving into ECE building (near Engineering, Psychology, Chemistry, Biology, Libraries RDS); other buildings are possible (ME, Farris, PAIS?)

Key Discussion Action Items [2]

Strategy 2: User Support

- Problem: Staffing levels at CARC are fragile to loss and insufficient to support current number of users
- Action Items: Actively collaborate and coordinate with campus/department/unit IT staff to support users
- Timeframe: Short (1-3 years)
- **Detail: Mainly focused on formalizing collaboration with front-line IT staff in A&S, Engineering, Libraries, etc.**
- **Related Question: Is the right model for user support to have multiple research staff members across campus supporting users in their disciplines, and how would that work?**
- Related action items: Deploying or utilizing existing collaboration tools (short), develop sustainable staffing plan (long)

Key Discussion Action Items [3]

Strategies 4 and 5: User Community and Industry collaboration

- Problem: CSE program, which could be a focal point of user engagement and industry outreach, is moribund
- Action Item: Overhaul the CSE certificate program by streamlining core requirements and creating pre-defined “tracks” in key areas
  - Define CSE Core as CS/MATH 442 (Parallel Programming), MATH 471 (Scientific Computing), and CS 467/567 (Big Data)
  - Require 1 core + 3 electives + thesis/project for MS/post-degree, 2 core + 4 electives + dissertation for Ph.D. students
  - Define tracks in CFD, Data Sciences, Comp. Biology, etc. in collaboration with departments/colleges
- Timeframe: Immediate (0-1 years)

Key Discussion Action Items [4]

Strategy 3: Grow and enhance systems

- Problem: Need to continually enhance capability of existing systems and grow system capacity in the face of flat budgets (A new 256-node cluster and associated storage will cost $1-2,000,000
- Action Items: Pursue research cyber-infrastructure grants (CARC's research mission) and instrumentation grants (CARC's service mission)
- Key Insight: Research leadership on the cyberinfrastructure itself) essential to competing for large cyber-infrastructure grants
- Timeline: Immediate
  - NSF Office of Advanced Cyberinfrastructure - CDES&E proposal in progress, others recently withdrawn and being re-worked at NSF
  - NSF Discipline-specific calls (e.g. CISE/CRI, Geosciences, etc.) – NSF CISE Research Instrumentation proposal in progress
  - DOE, DOD, and National Laboratory opportunities in the future

Key Action Items [5]

STRATEGY 3: Grow and Enhance Systems

BASELINE DEFINITION AND ABOVE-BASELINE COST MODELS

CARC Service Model: Free Baseline Research Computing Support

- **Goal: Generous baseline CARC services provided to the UNM community *free of charge***
- **Campus-provided baseline funding/staffing should support the vast majority of regular research on campus**
  - Basically all unsponsored and startup research computing
  - The vast majority of sponsored research computing
- **Define this baseline explicitly, enforce this baseline loosely**
- **Provide mechanisms for supporting extraordinary needs outside of this baseline**
  - Support for transitioning research to national resources (e.g. NSF XSEDE, NSFCloud, etc.)
  - Condo-like service center model for UNM researchers needing local resources well beyond the baseline

Proposed CARC Service Baseline

Service	Unit	Per User	Per Project	Average Utilization/ Uptake	Center Total Need
Compute Nodes	Node	2	16	25.00%	420
Enterprise Storage	TB	0.2	2	25.00%	50
Working Storage	TB	2	20	25.00%	500
Server Colocation	Rack Units	0	16	5.00%	64
System Administration	Images	0	1	12.50%	10
Storage Administration	Partitions	0	1	25%	20

- The baseline assumes continued use of older compute nodes (from gifts, old grants, or surplus) to provide compute capacity. Newer nodes would result in a lower compute node baseline
- Should custom system administration be part of the baseline? It's very labor intensive but very important for emerging research areas

Cost of Providing this Baseline

- Assuming staffing for 200 users, this baseline breaks down the CARC service center budget as follows
- 5 year term result from 5 year hardware warranties

Resource	Unit	Term (Years)	Base F&A Amount	Term Service Cost	Yearly F&A Center Cost
Compute Nodes	Node-Years	5	420	$519.14	$43,607.93
Enterprise Storage	TB-Years	5	50	$361.14	$3,611.38
Working Storage	TB-Years	5	500	$36.11	$3,611.38
Server Colocation	Rack Unit-Years	1	80	$46.71	$3,002.16
System Administration	Image-Years	1	10	$9,362.84	$93,628.41
Storage Administration	Partition-Years	1	20	$1,460.12	$29,202.39
User Support	Users	1	200	**$1,800.31**	$**360,061.32**
					$536,724.97

Service Above Baseline

- **Goal: Bridge the gap between baseline CARC-provided resources and national resources**
  - Supports specialized research needs without the large waiting times on national resources
  - Provides systems which support UNM baseline research computing at grant termination
- **Split costs between capital equipment (not charged overhead) and service center costs (overheaded) for staff and infrastructure**
- **Keep infrastructure and staffing costs low for items purchased in bulk (nodes, storage capacity)**
- **Note: most existing centers appear to only charge actual hardware costs for condo compute and storage**

Service Center Costs

Resource	Node Type	Term (Years)	CARC Services per Unit	Capital Hardware	Infrastructure per Unit	CARC Term Cost
Compute Nodes	Compute Node	5	$519.14	$4,900.00	$611.47	**$6,030.61**
	Large Memory Node	5	$519.14	$17,100.00	$611.47	**$18,230.61**
	GPU Node	5	$519.14	$21,500.00	$1,070.07	**$23,089.21**
	Storage Node	5	$519.14	$14,700.00	$1,070.07	**$16,289.21**
Enterprise Storage	-	5	$361.14	$800.00	$44.38	**$1,205.52**
Working Storage	-	5	$36.11	$100.00	$11.10	**$147.21**
Server Colocation	-	1	$37.53	$0.00	$191.46	**$228.99**
VM Hosting	-	5	$64.89	$612.50	$191.79	**$869.18**
System Administration	-	1	$9,362.84	$0.00	$0.00	**$9,362.84**
Storage Administration	-	1	$1,460.12	$0.00	$0.00	**$1,460.12**

SUMMARY

The CARC Value Proposition

- CARC's capabilities provide significant value to UNM researchers
- Use of previously grant-funded and gifted computational resources a significant portion of this value
- Implication: UNM needs to continue to pursue cyberinfrastructure grants aggressively.

	Yearly Cost	Yearly Value		
Service	Center Budget	Per User	Per Project	Center
Compute Nodes	$43,607.93	$625.75	$5,005.97	**$525,626.60**
Enterprise Storage	$3,611.38	$12.56	$125.64	$12,563.80
Working Storage	$3,611.38	$16.48	$164.79	$16,478.53
Server Colocation	$3,002.16	$0.00	$465.82	$37,265.85
System Administration	$93,628.41	$0.00	$1,170.36	$93,628.41
Storage Administration	$29,202.39	$0.00	$365.03	$29,202.39
User Support	$360,061.32	$1,800.31	$0.00	$360,061.32
Total	**$536,724.97**	**$2,455.09**	**$7,297.60**	**$1,074,826.90**

Summary

- CARC's dual service and research missions provide significant value to the UNM community
- Generous free-of-charge research computing baseline provides a high value to campus users
- Staffing challenges due to continued growth of computing-based research and the emergence new computational research areas
- Service center model provides model for above-baseline support for users and research with extraordinary needs
- Leadership on research in cyberinfrastructure is essential to sustaining CARC's long-term value proposition