Annual Report of the

Shared Research Computing Policy Advisory Committee (SRCPAC)

Research Computing Executive Committee

Friday, May 15, 2015

SUMMARY
Yeti: A Catalyst for the Columbia HPC Community

Thus Far, Demonstrated Impacts On:

HPC Computing
Education
Inter-School and Cross-Campus Planning
Recruitment

Now, a Unique Opportunity to Build On and Beyond These First Steps

PRESENTATION OUTLINE

1. **The Basics**
   A. RCEC, SRCPAC, SRCF & RCS
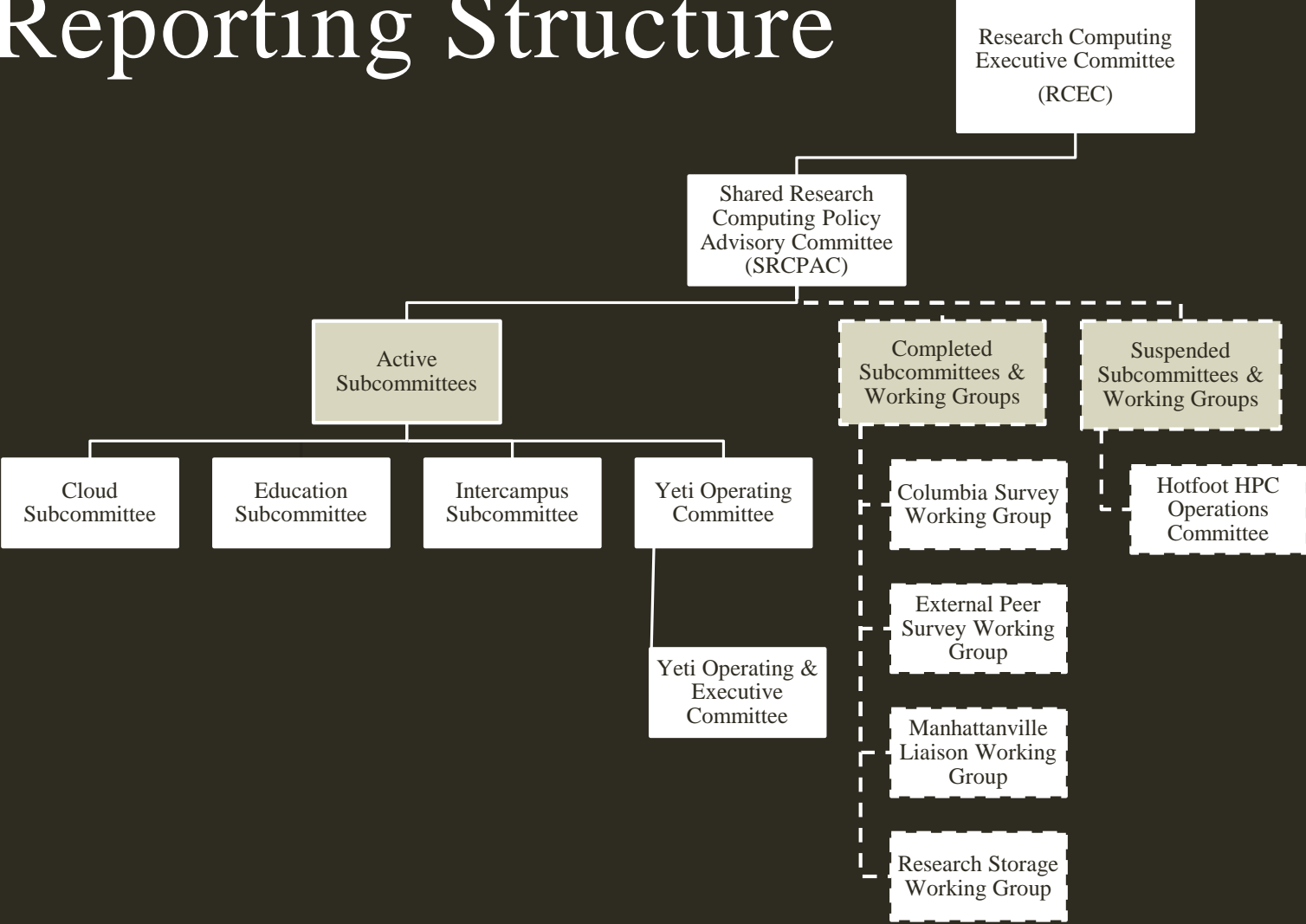   B. Hardware Report
   C. Infrastructure Report

2. **2015 Achievements**
   A. New Purchase Round
   B. SRCPAC Subcommittees
   C. Recruitment

3. **2016 Prospects & Challenges**

1A: THE BASICS
Reporting Structure

Research Computing Executive Committee (RCEC)

Shared Research Computing Policy Advisory Committee (SRCPAC)

Active Subcommittees

Completed Subcommittees & Working Groups

Suspended Subcommittees & Working Groups

Cloud Subcommittee

Education Subcommittee

Intercampus Subcommittee

Yeti Operating Committee

Yeti Operating & Executive Committee

Columbia Survey Working Group

External Peer Survey Working Group

Manhattanville Liaison Working Group

Research Storage Working Group

Hotfoot HPC Operations Committee

RCEC Membership

G. Michael Purdy (Chair) Executive Vice President for Research	
David Madigan Executive Vice President for Arts & Sciences Dean of the Faculty of Arts & Sciences	*David Madigan's Alternate:* **Amber Miller** Dean of Science
Mary Boyce Dean The Fu Foundation School of Engineering & Applied Sciences	*Mary Boyce's Alternate:* **Shih-Fu Chang** Senior Vice Dean The Fu Foundation School of Engineering & Applied Sciences
Gaspare LoDuca Vice President of Information Technology Chief Information Officer	
Damon Jaggars Interim University Librarian	*Damon Jaggar's Alternate:* **Robert Cartolano** Associate Vice President for Digital Programs & Technology Services
Justin Pearlman Chief of Staff Office of the Provost	
Kathryn Johnston Chair of the Shared Research Computing Policy Advisory Committee (SRCPAC)	*Staff* **Victoria Hamilton** Director of Research Initiatives Office of the Executive Vice President for Research

1B: HARDWARE REPORT
Columbia's Shared Research Computing Systems

Hotfoot High-Performance Computing (HPC) System (2009 & 2011)

Yeti High-Performance Computing (HPC) System (2013 & 2014)

Why Are We Doing This?

1. **Meeting HPC Needs:**
   - Local Clusters
   - National Computing Centers
   - Cloud Computing

2. **Why _Share_ a Local Cluster?**
   - *Researchers Gain:*
     - Time
     - Local Expertise
     - Access to Larger Machine
     - Flexibility
   - *Columbia Gains:*
     - Energy & Space
     - Shared Staff & Hardware Costs
     - Recruiting Tool
     - Happy Faculty!

Yeti HPC History

Launched with the Shared Research Computing Facility (SRCF) in Fall 2013, then Expanded in Spring 2015.

Hardware Purchase – Total 2,672 Cores, 167 Nodes:

- 1,840 Cores From Columbia – 784 in Round I; 1,056 in Round II
- 832 Cores From NYSERDA Grant in Round I

Researchers Using Yeti:

- 24 Research Groups – 10 in Round I and 14 New in Round II
- 661 (107 faculty) Have Access: 277 Have Run Jobs in Past 12 Months

Support: Research Computing Services (RCS) Group

- 5 FTE's Supported by CUIT, A&S, SEAS, EVPR
- 1 Manager, 2 User Support, 1 Systems Administrator, 1 Software & Licensing
- Staff Hotfoot and Yeti Operating Committees
- Instigated Popular Workshops This Year for Novice Users

Yeti Participants: Round I

Large Purchase Groups

1. Center for Computational Learning Systems (CCLS)
2. Statistics
3. SSCC (Economics, Sociology, Social Work & SIPA)

Medium Purchase Groups

1. Ocean & Climate Physics
2. Astronomy & Astrophysics

"Toe -in-the-Water" Groups

1. Earth & Environmental Engineering
2. CIESIN
3. Psychology
4. Physics
5. Journalism

2 Renters (Mechanical Engineering & Neuroscience as Pilot)

Yeti Participants: Round II

Five Groups Invested Heavily

Applied Physics & Applied Mathematics

Chemical Engineering

Data Science Institute

Physical Oceanography

Quantum Mechanics Lab

Nine Groups or Labs Made Introductory/Small Investments

Brain Lab

Bussemaker Lab

Combustion Systems

Condensed Matter Theory
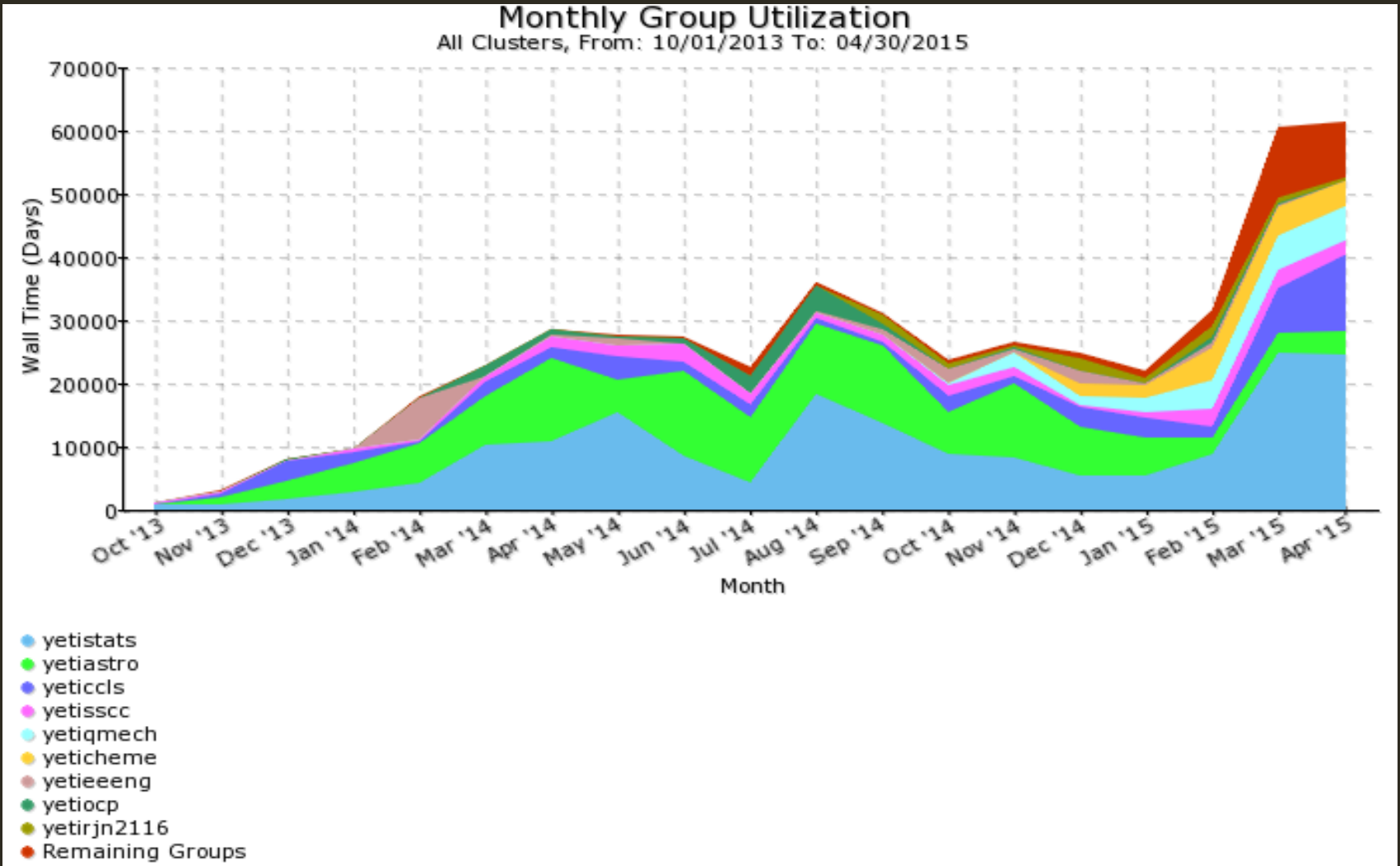
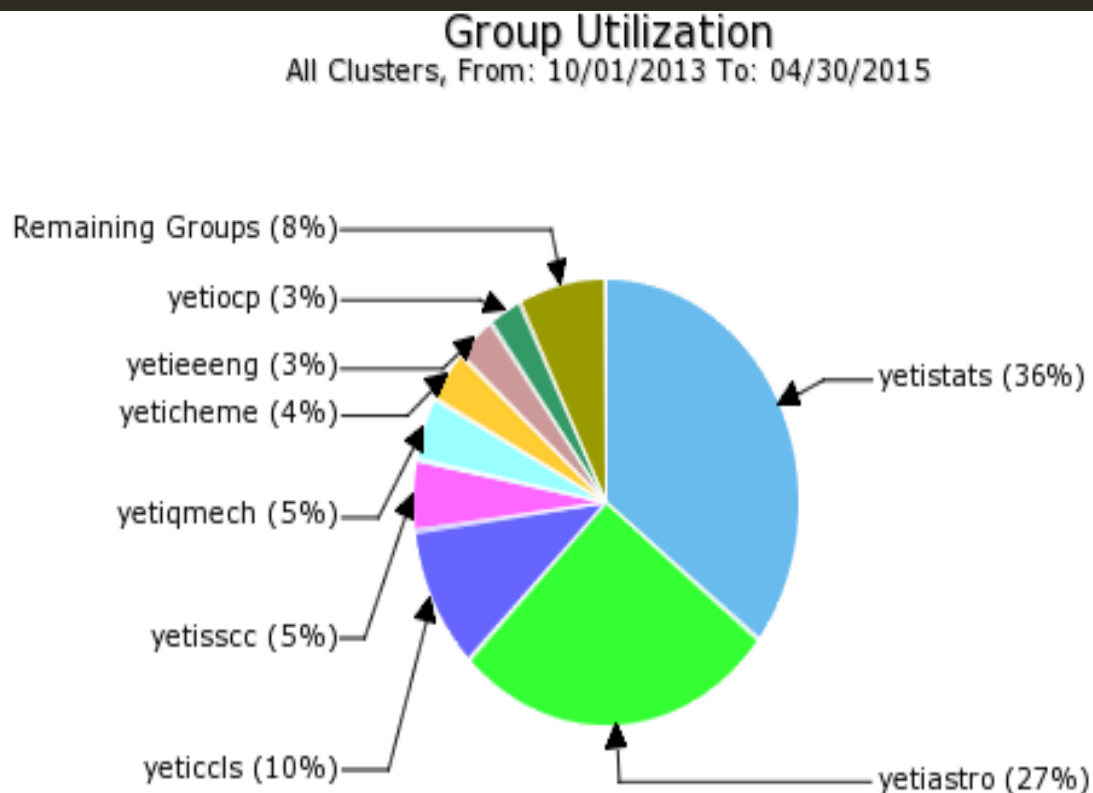Digital Video & Multimedia Lab

Heat Lab

Stockwell Lab

Structure Function Imaging Lab

Zuckerman Institute

Yeti Usage (Slide 1 of 2)

Yeti Usage (Slide 2 of 2)

Group Utilization
All Clusters, From: 10/01/2013 To: 04/30/2015

Remaining Groups (8%)
yetiocp (3%)
yetieeeng (3%)
yeticheme (4%)
yetiqmech (5%)
yetisscc (5%)
yeticcls (10%)
yetistats (36%)
yetiastro (27%)

57 Yeti Publications To Date

Physics & Astronomy: 19

Social Sciences: 13 (Plus 2 In Review & 3 Forthcoming)

Biomedical Sciences: 13

Statistics, Computer Science & Engineering: 12

1C: INFRASTRUCTURE REPORT

As of Today, Nearly 60% is Already Installed or Pledged.

Anticipated Future Demand:
- Existing Users Wishing to Expand (2 Inquiries Within the Past Month).
- New Groups Joining.
- Recruits (1,024 Cores Promised in Start-Up Packages Over the Next 2 Years).
- Proposals (Another ~1,000-Core Proposal Pending from 1 PI).

Semi-Annual Estimate Based on (Short) History
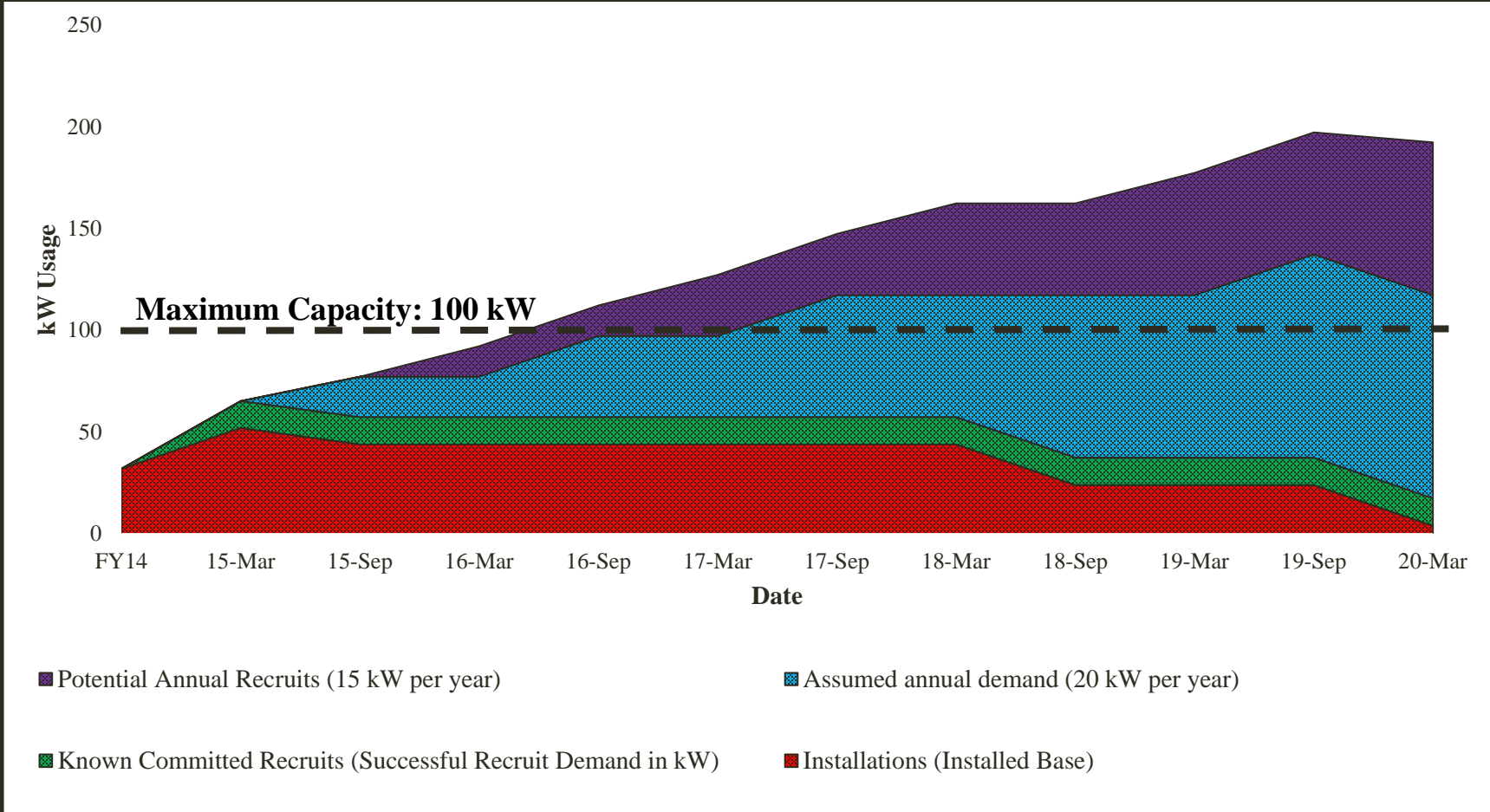
Options to Deal With Impending SCRF Capacity Constraint

Explore Alternative Paths to Expand Capacity – Budget & Timeline. For Example:

- Increase Cooling in SRCF;
- Partner With Other Data Centers On Campus That Might Have Temporary or Permanent Capacity;
- Expand Off-Site (In Syracuse, For Example);
- Partner With Other Private Consortiums Such as the Massachusetts Green Data Center;
- Use Professional Resources Like Amazon to Expand the SRCF;
- Develop Expertise to Advise Faculty About Alternatives Such as Amazon.

Establish Ad Hoc RCEC Subcommittee to Allocate Resources When Demand Exceeds Supply.

1C: INFRASTRUCTURE REPORT

Installed and Promised HPC is Over 50% Today and Could Exceed Cooling Capacity Constraint in FY16.

Anticipated Future Demand:

- Existing Users Wishing to Expand (2 Enquiries in the Past Month);
- New Groups Joining;
- Recruits (1,024 Cores Promised in Start-Up Packages Over the Next 2 Years);
- Proposals (Another ~1,000-Core Proposal Pending from 1 PI).

Possible Actions?

- $3M to Increase Cooling – Perhaps Triple to Quintuple Capacity;
- Use of Other Data Centers Across Campus;
- Cloud or Other External Consortia (With Associated Local Staff to Advise).

2. 2015 Achievements

A. New Purchase Round
  - Intense Interest! Researchers ARE Happy.

B. SRCPAC Subcommittee Work
  - Yeti as Aa Focal Point for an "HPC Community" of Researchers, Educators, and Administration.

C. Recruitment
  - At Least 2 Top Hires This Year That Would Not Have Happened Without Yeti.

2A: NEW PURCHASE ROUND

Expanded From 1,600 to 2,700 Cores – Definite Interest

Lessons Learned

- Fall Purchase Timing
- Trial-with-Intent-to-Purchase (TIP)
- SEAS Incentive Program Spurred Participation

Looking Forward

- Already Interest for Next Round From Existing Users & Recruits
- "Technical Advice" Faculty Subcommittee Forming
- Motivates Investment in SRCF & Exploring Cloud to Allow Growth

2B: SUBCOMMITTEES
Cloud

7 Members:
- *Faculty*: APAM, Astronomy, Physics, Computer Science, Social Work
- *Staff*: CUIT

Charge: To Discuss if the Cloud Presents a Realistic Alternative To or Extension Of SRCF.

FY15 Meetings: 2

Outcomes
- RCS-Supported "Cloud Trial" (In Process)
- Cloud Committee Currently Suspended

Looking Forward
- Review Status of Subcommittee Every Fall
- Amazon Vendor Agreement (In Process)
- Cost Negotiation – Opportunity for Institutional Discount?

2B: SUBCOMMITTEES
Education

7 Members:
- *Faculty*: Physics, APAM, Lamont
- *Staff*: Libraries, CUIT, Psychology

Charge:
- Formulate Policy for Class Access to Yeti;
- Discuss Other Courses/Workshops for Yeti Users

FY15 Meetings: 2 (Plus Outreach to Other Columbia Educators)

Associated Activities & Outcomes
- Education Access Policy for Use of Yeti Adopted by SRCPAC;
- Workshops on Linux and Introduction to HPC by RCS;
- Recommendation for Course Coordination Across Departments.

Moving Forward
- Continued Workshops
- Work Study for Coordination of Classes?
- Expansion to Include Basic Introductions for Complete Novices?

2B: SUBCOMMITTEES
Intercampus

9 Members:

- *Faculty*: Physics, APAM, Biological Sciences/Systems Biology, Psychiatry
- *Staff*: Zuckerman, Business School, Lamont, CUIT
- *Invited Guest*: Data Sciences Institute

Charge: Examine Whether the Different Campuses Could Coordinate so as to Leverage Investments in Infrastructure, Equipment, and Staff.

FY15 Meetings: 4

Outcomes: *MANY RECOMMENDAITONS*

- Dedicated Staff Person to Identify and Track University Options for External Resources;
- Identifying Solutions for Affordable Remotely Backed-Up Research Storage Options and Space for Hosting Research Computing Equipment;
- Creating a Columbia "Research Computing" Portal.

Moving Forward: Unclear Whether to Continue Meeting or How to Implement Recommendations.

2B: SUBCOMMITTEES
Yeti Operating Committee

9 Executive Committee Members:
- 4 Representatives of Large Purchasers: Statistics, APAM, Lamont, CCLS
- 4 Representatives of Small Purchasers (1 Vacant Seat): Journalism, Biological Sciences, Chemistry
- Faculty Chair Representing Renters & Free Tier

Charge: Allow Users to Guide Operating Decisions.

FY15 Meetings: 2

Outcomes:
- Endorsed Recommendation of Education Committee to Provide 5% of Yeti for Education;
- Allowed Yeti II Purchasers to Join Yeti Prior to Installation of New Cores;
- Decided on Queueing Policy.

Moving Forward:
- Continue the Same

2C: RECRUITMENT

2 Success Stories This Year

Reminder of Benefits Beyond Recruitment:

- Support New Faculty in Immediately Concentrating on Research;
- Save Money in Renovations and Ongoing Power/Cooling Costs of Decentralized Clusters;
- Start-Up Has Impact Beyond Single Researcher.

Moving Forward – Only Finite Space Left in SRCF

- Need for Allocation Policy (See Handout/Attachment)
- Motivates Investment in Growth

3. 2016 Prospects & Challenges

SRCF is *Finite* - How To Satisfy Demand for Capacity?

Recruitment - Given Finite Space, How to Make Future Promises?

Funding Model – Success Motivates Longer-Term Plan?

The Cloud – What Can We Move Off-Campus?

Data - How (And Who) to Tackle These Issues?

HPC and Novice-Users – Can We (*Should* We) Include All?

Building a Sustainable "HPC Community" – How to Implement Subcommittee Recommendations?

SUMMARY
Yeti: A Catalyst for the Columbia HPC Community

So Far, Demonstrated Impacts On:

HPC Computing
Education
Inter-School and Cross-Campus Planning
Recruitment

Now, a Unique Opportunity to Build On and Beyond These First Steps