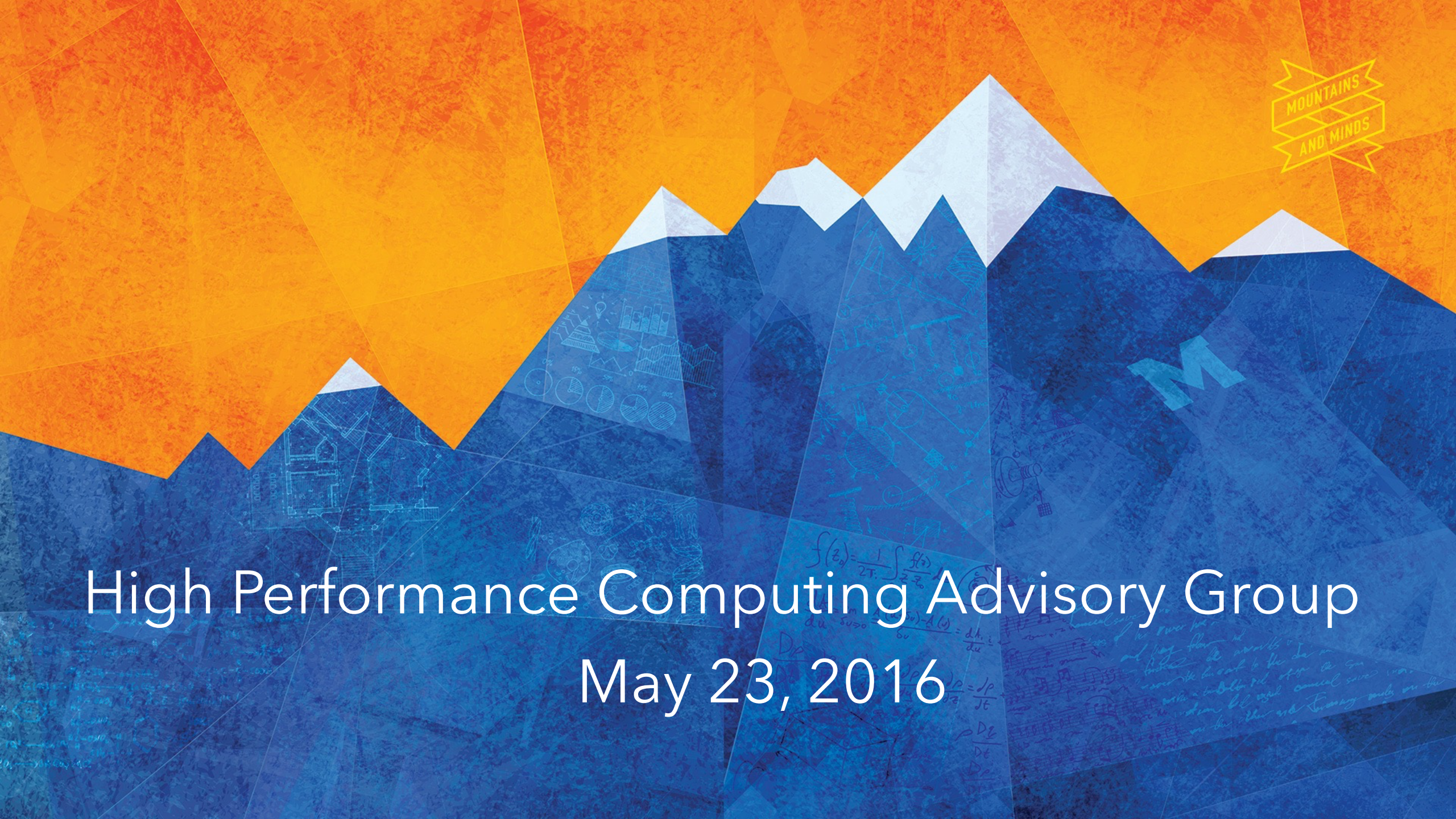
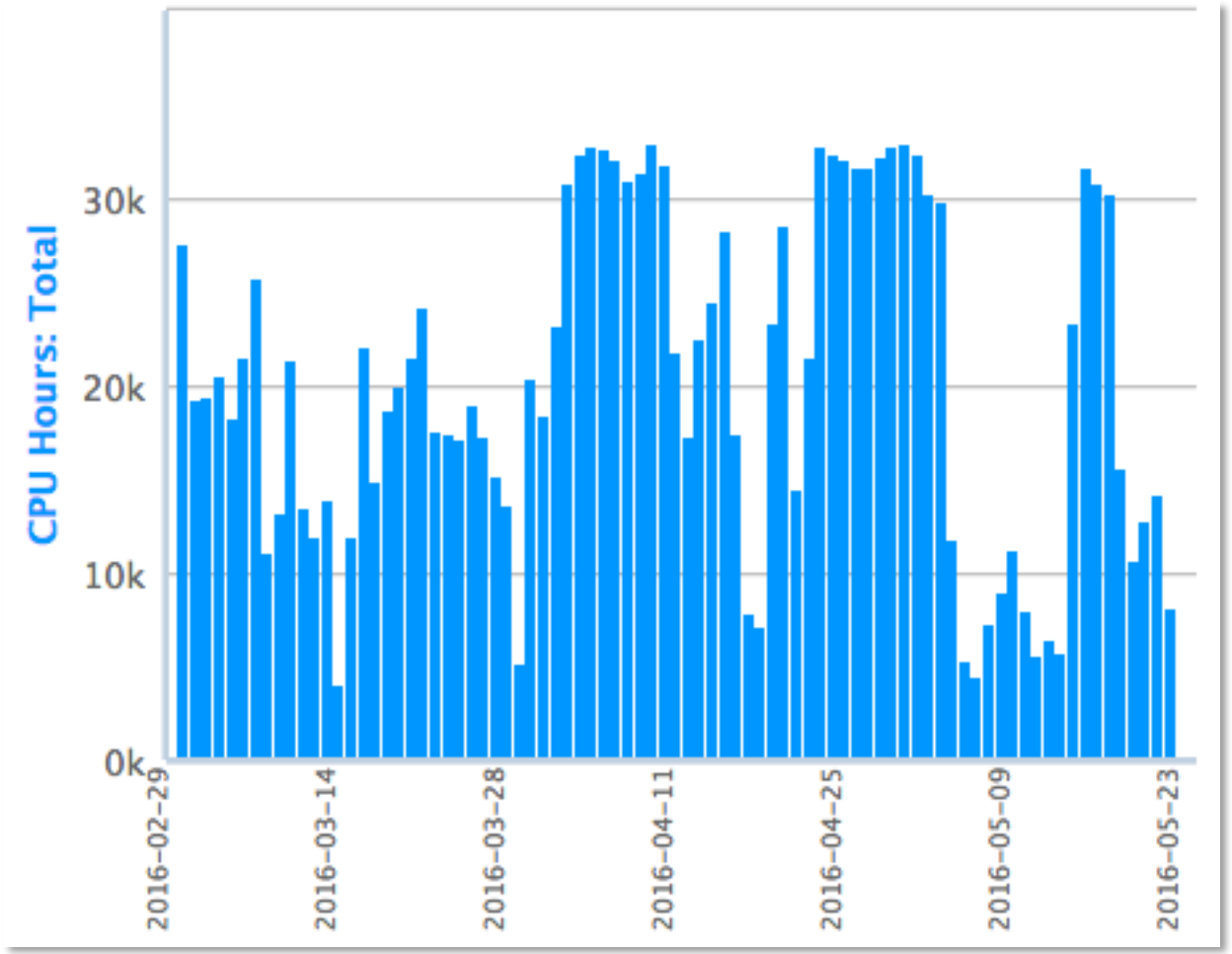


High Performance Computing Advisory Group

May 23, 2016

XDMoD Stats - March 1 to May 23

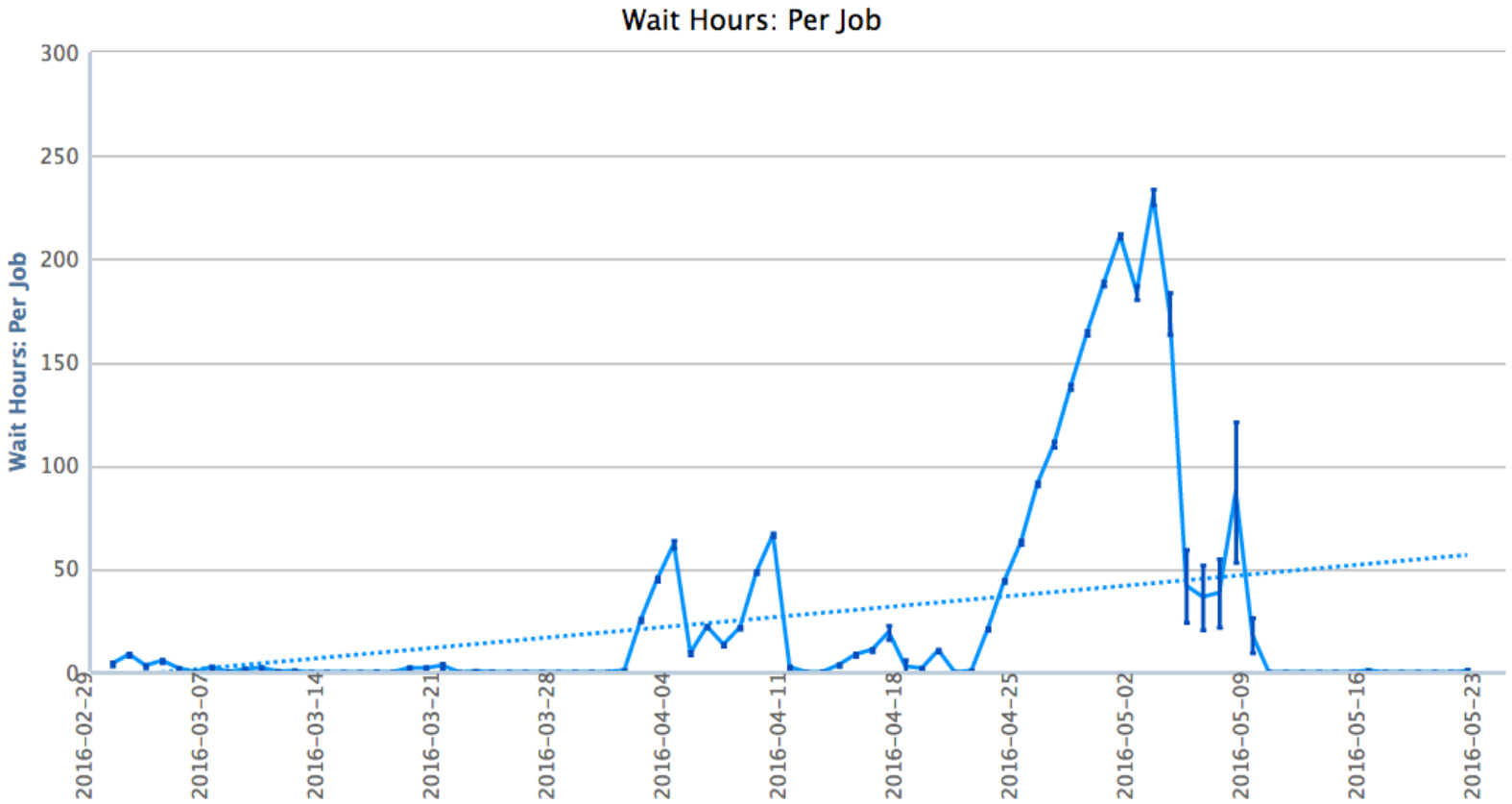
- XDMoD Stats
  - Active Users: 35
  - Total Jobs: 85,409
  - Total CPU Hours: 1,687,511
  - Average Job: 19.86 hours
  - Average Wait Time: 20 hours



MONTANA STATE UNIVERSITY

Deep Queue in April

- The queue was >9k jobs deep in April
- Mostly a single user
- Showed problems with fairshare algorithm
  - Users needed "accounts"
  - Should we set a "max queued jobs"?



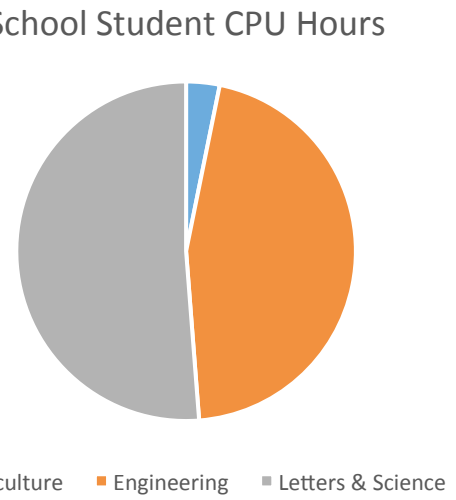
MONTANA STATE UNIVERSITY

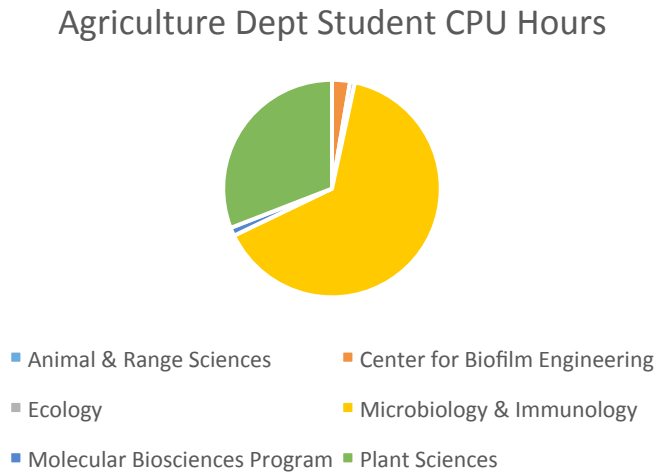
Student Jobs are 57% of cluster hours

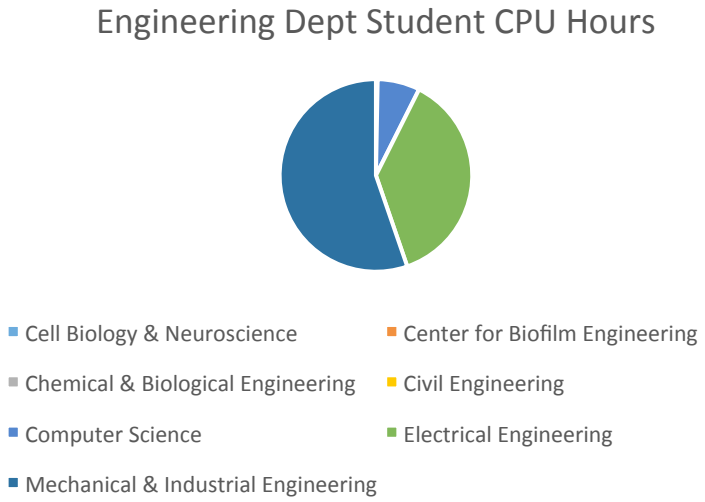
Total CPU Hours*	7,943,855.90	
Student CPU Hours	4,542,543.47	57.18%
Graduate Hours	4286823.095	53.96%
Undergrad Hours	255720.3727	3.22%

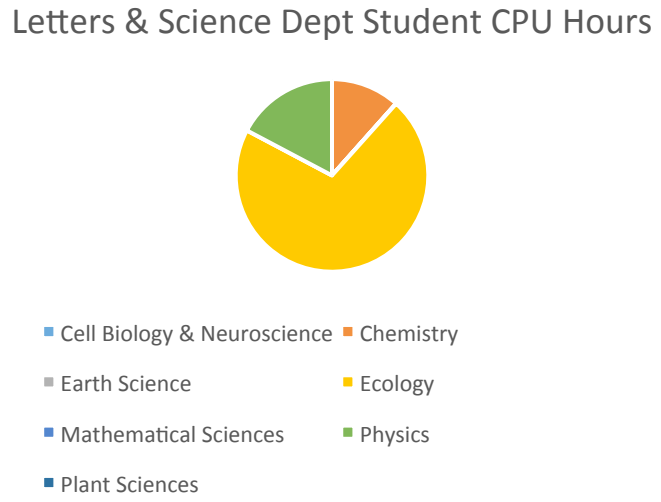
Agriculture	144795.6186
GR	143757.9059
UG	1037.7127
Engineering	2070747.756
GR	2070061.8
UG	685.9562
Letters & Science	2327000.094
GR	2073003.39
UG	253996.7038
Grand Total	4542543.468

* For all time, since cluster installation









MONTANA STATE UNIVERSITY

New Hyalite Researcher

Professor Joe Atwood

- Agricultural Economics Professor
- Onboarded on Hyalite and Setup R environment and Rstudio modules
- The weather data contains monthly total prcp and monthly average daily temp data for the years 1900-2009
- Work is pleasantly parallel on millions of linear calculations

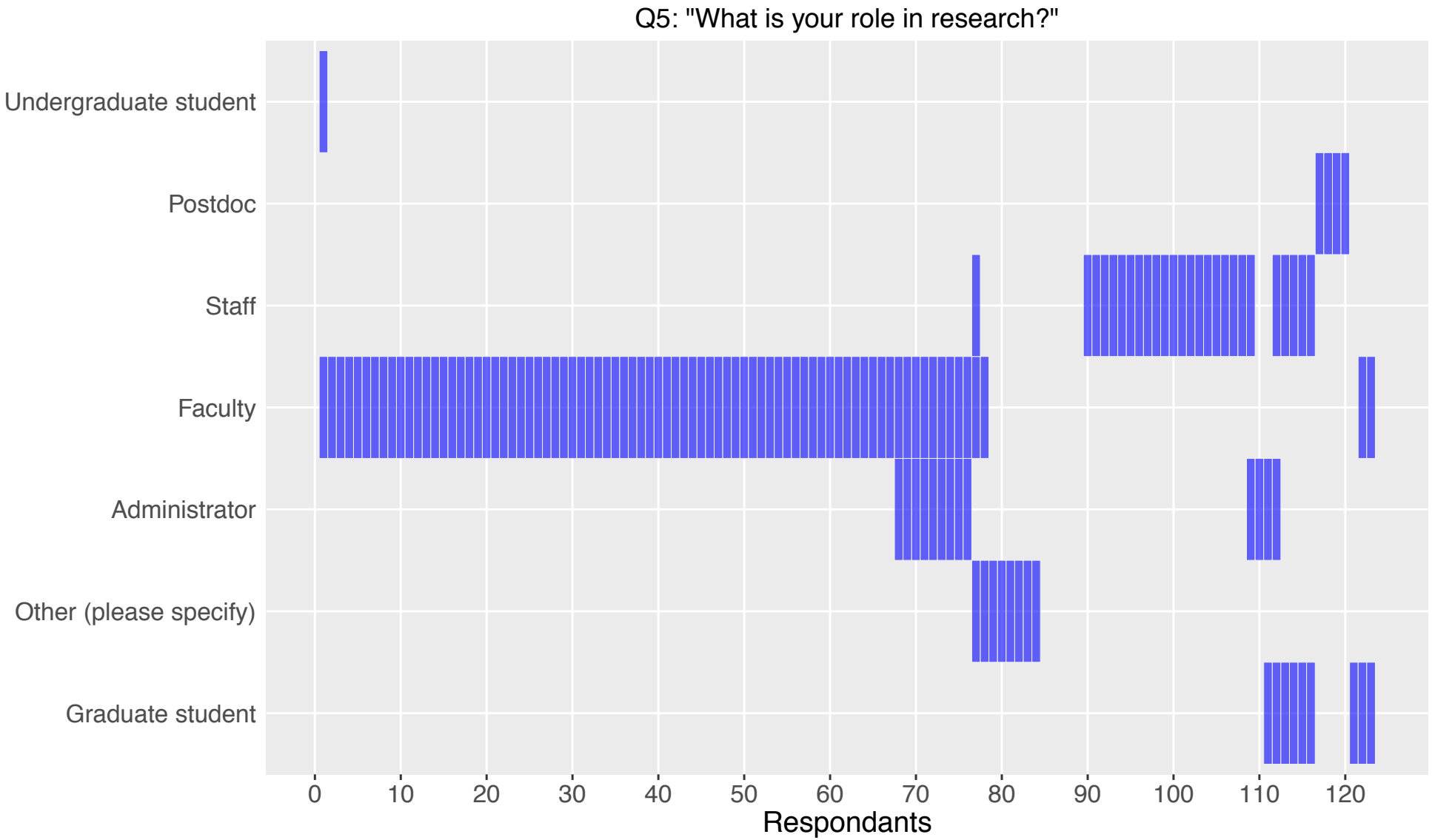
MONTANA STATE UNIVERSITY

Cluster Expansion Update

- 16 Nodes finalizing in New Jersey facility
  - Shipped late from Supermicro
  - Network Brackets were wrong
  - Should ship this week to us

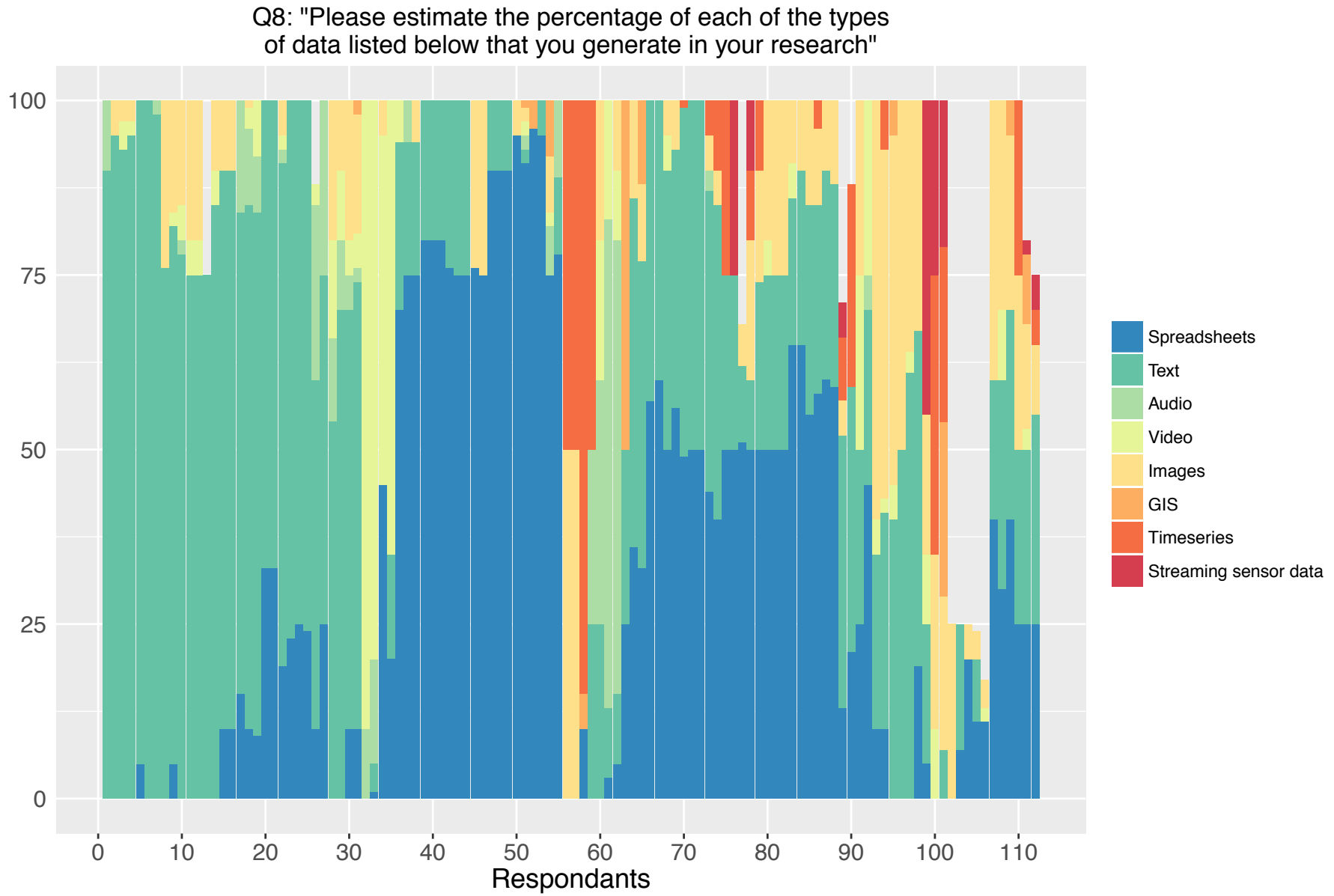
- High CPU, RAM, I/O Node Quote
  - Quad E5-4648v3 CPU (96 HT)
  - 512GB RAM (1TB max)
  - 1TB Local SSD
  - $18,000
  - ~4x power, ~3.5x cost of one node

Research Data Census 2 (RDC2): Demographics



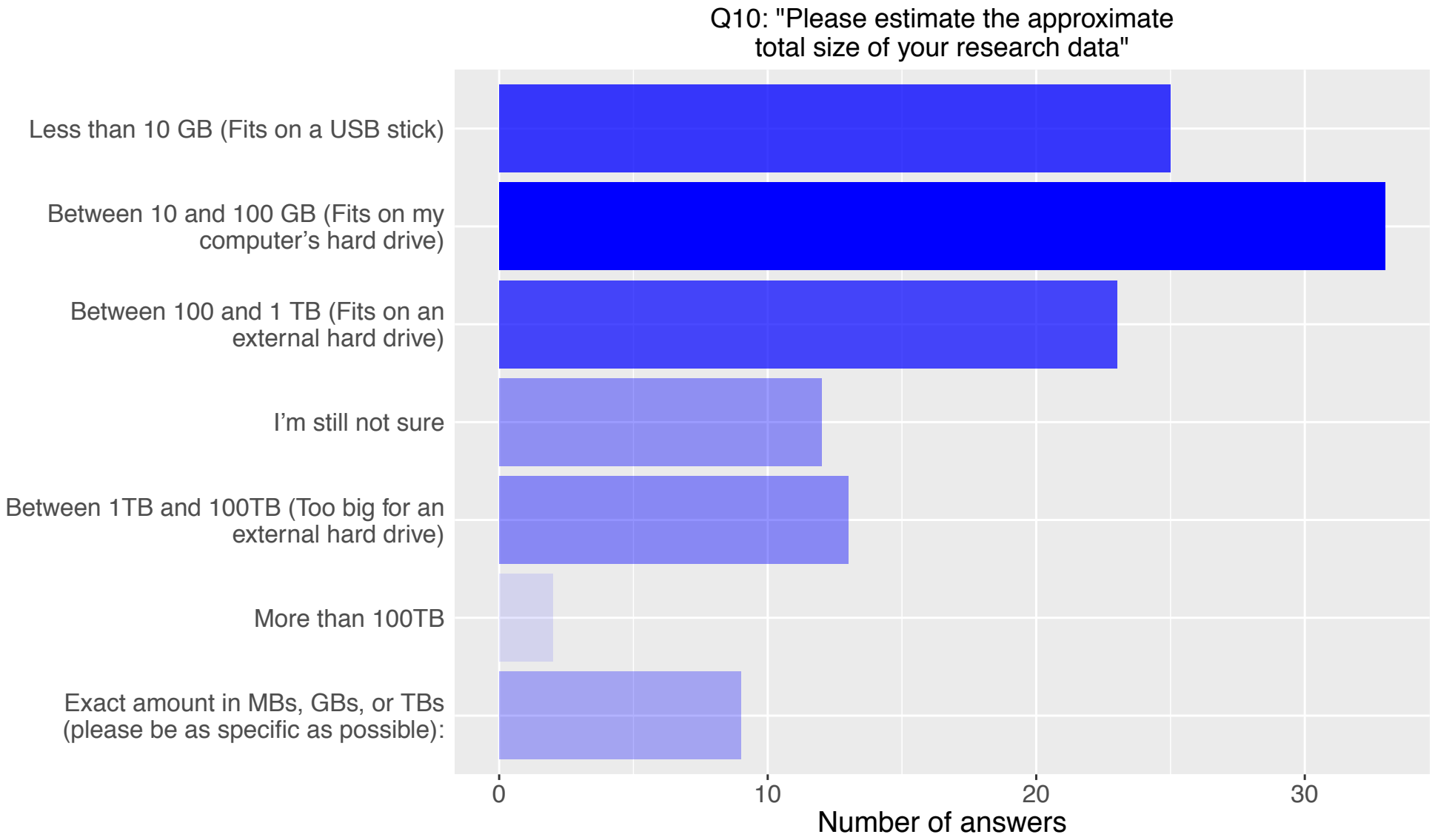
MONTANA STATE UNIVERSITY

RDC2: Types of Data



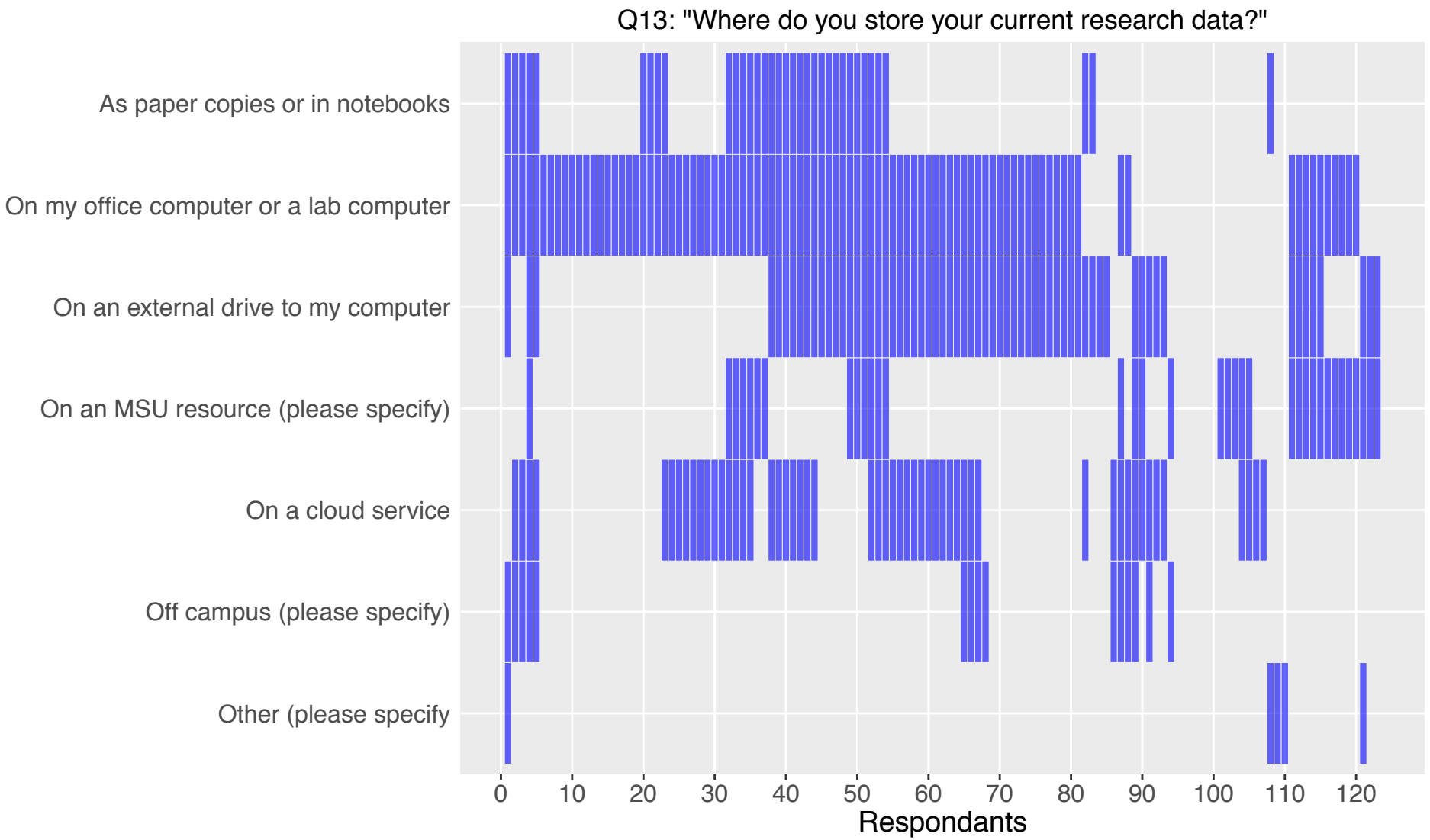
MONTANA STATE UNIVERSITY

RDC2: Digital Research Data Size



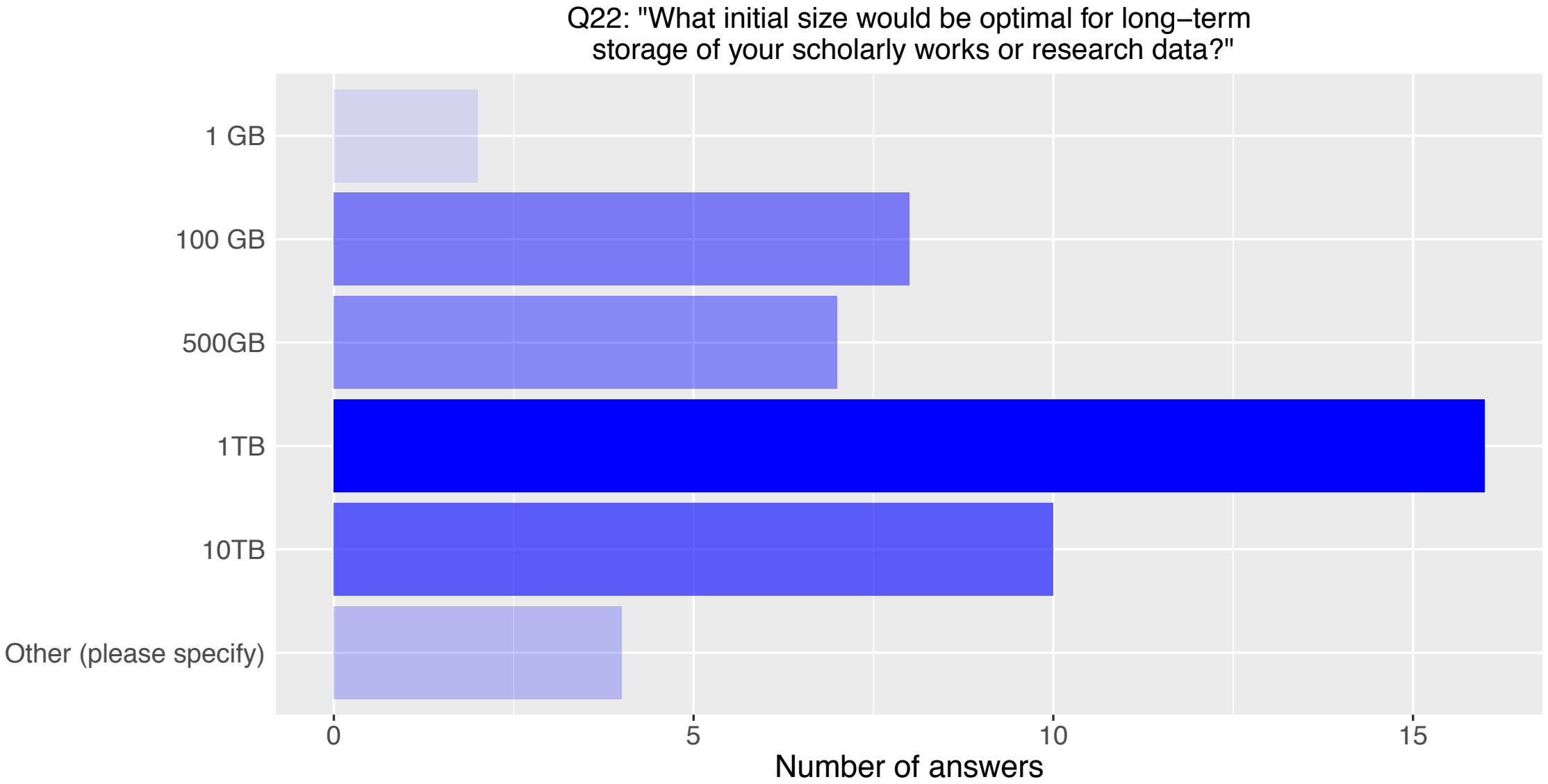
MONTANA STATE UNIVERSITY

RDC2: Where Do You Store Your Data



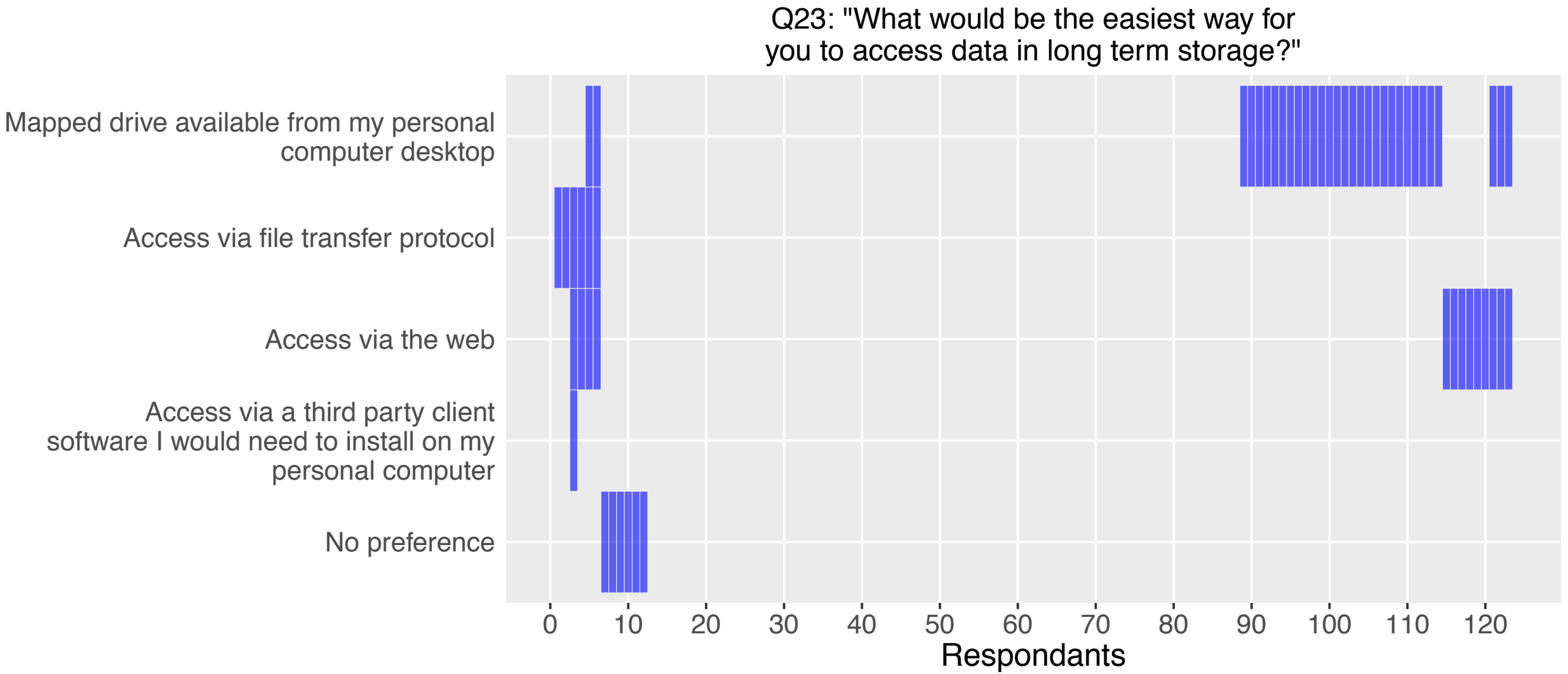
MONTANA STATE UNIVERSITY

RDC2: Long Term Storage Archive Size



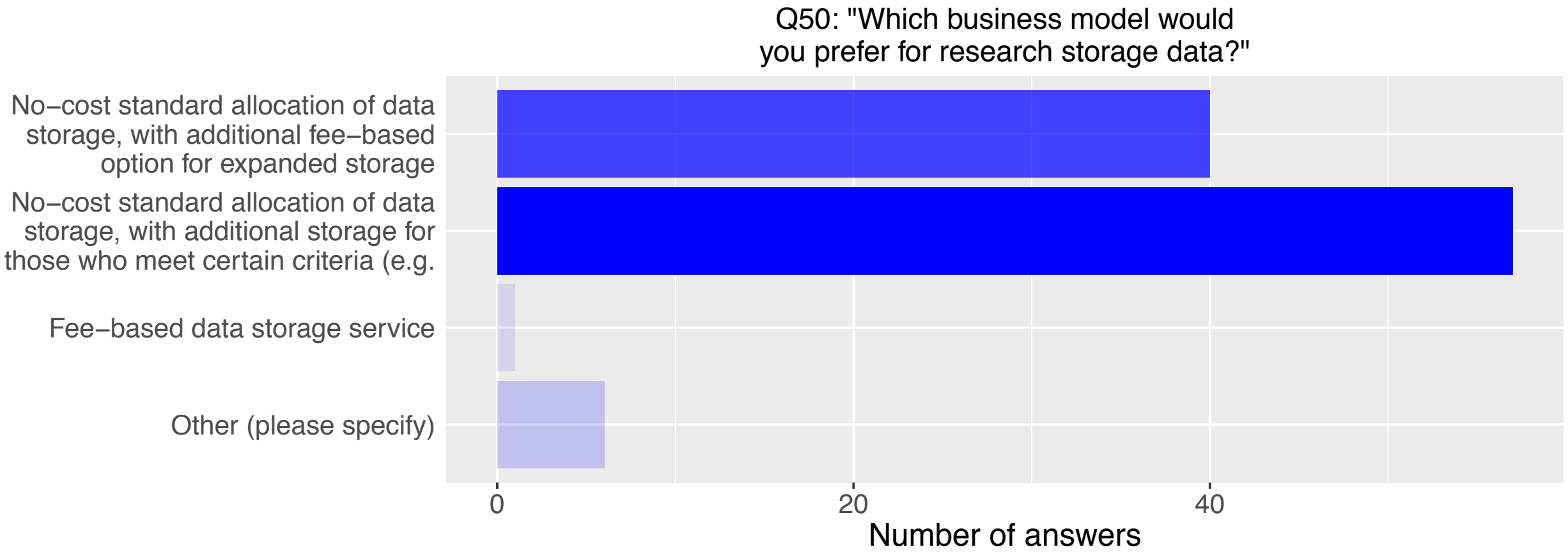
MONTANA STATE UNIVERSITY

RDC2: How Access Long Term?



MONTANA STATE UNIVERSITY

RDC2: Business Model



MONTANA STATE UNIVERSITY

Research Focus Group

- Two focus groups were held in May, predominantly faculty participants.

- Insights
  - Concern expressed about “finding” institutional data assets.
  - Consistent view with RDC2 that long term archive should be centrally supported.
  - How data was uploaded seemed to be of little concern.

- Next Steps
  - Synthesis of RDC and Data Focus Groups.
  - RFP Team named in June, development of technical requirements.
  - July target data for RFP to be published.

Data Science Competition

- Open to students from all disciplines
- Students are invited to solve a data-rich problem through visualization and/or number crunching

- Currently working with the Office of Sustainability
- Competition around optimum deployment of new and existing bike racks across campus, taking into account pedestrian traffic, bike traffic, sidewalks and building locations, building occupancy

- ETA: Early Fall 2016

MONTANA STATE UNIVERSITY