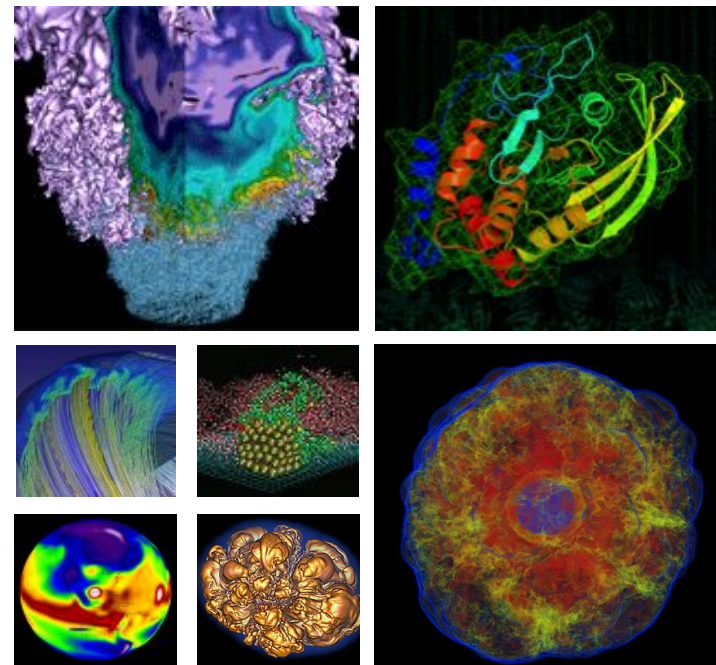# 2018 User Support & Outreach

NERSC

Rebecca Hartman-Baker

NERSC User Group Meeting
July 19, 2019

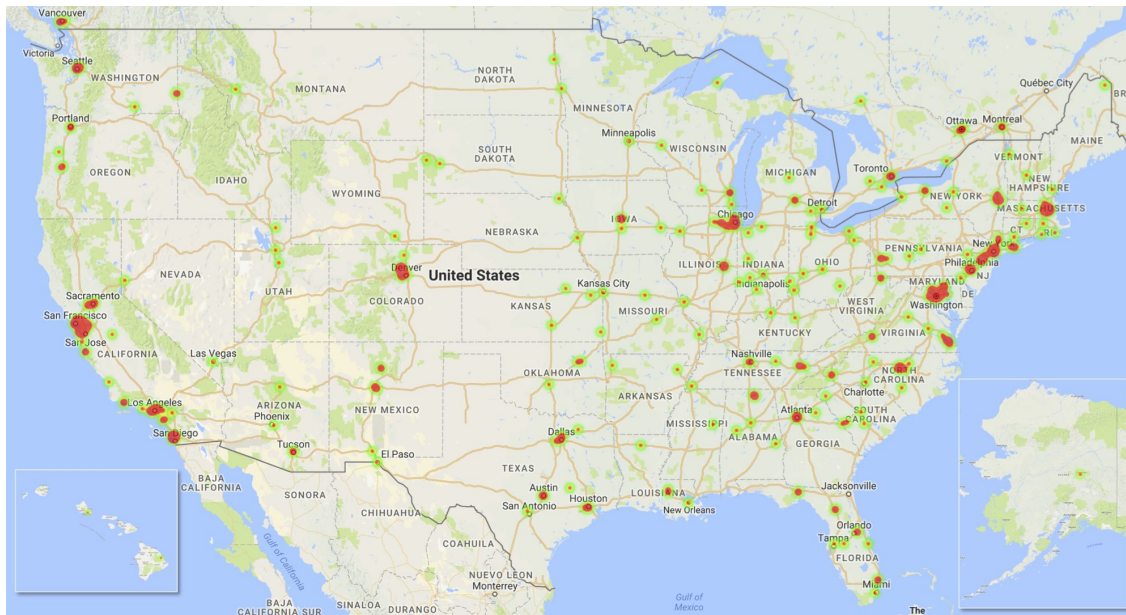# User Support & Outreach

I. **NERSC Users Overview**

- 7,000 users, 702 projects, all 50 states + international
- 4,285 users ran jobs on NERSC's Edison and Cori systems
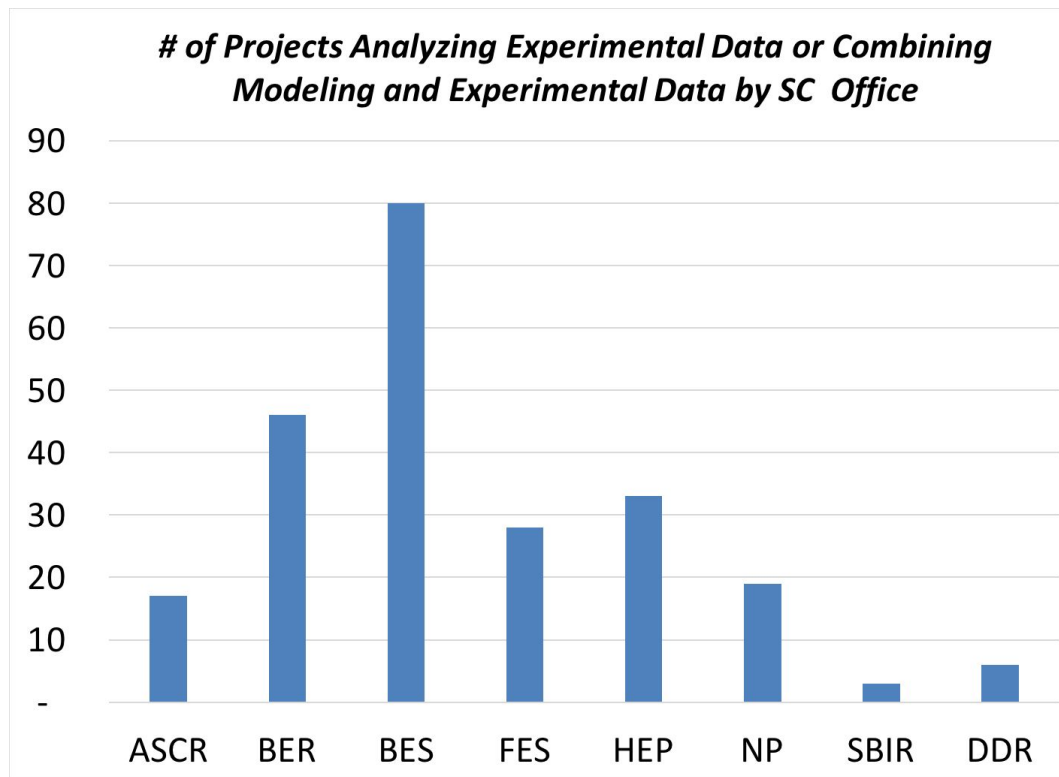
# NERSC Users Overview

Growing number of users analyzing data

More than 35% of all NERSC projects in this category

# of Projects Analyzing Experimental Data or Combining Modeling and Experimental Data by SC Office

II. User Survey & Satisfaction

- Annual user survey since 1998 with major revisions in 2018
- Scale: 1 (very unsatisfied) to 7 (very satisfied); target: above 5.25
- 577 users responded to survey, representing 56% of hours used

Overall User Satisfaction with NERSC
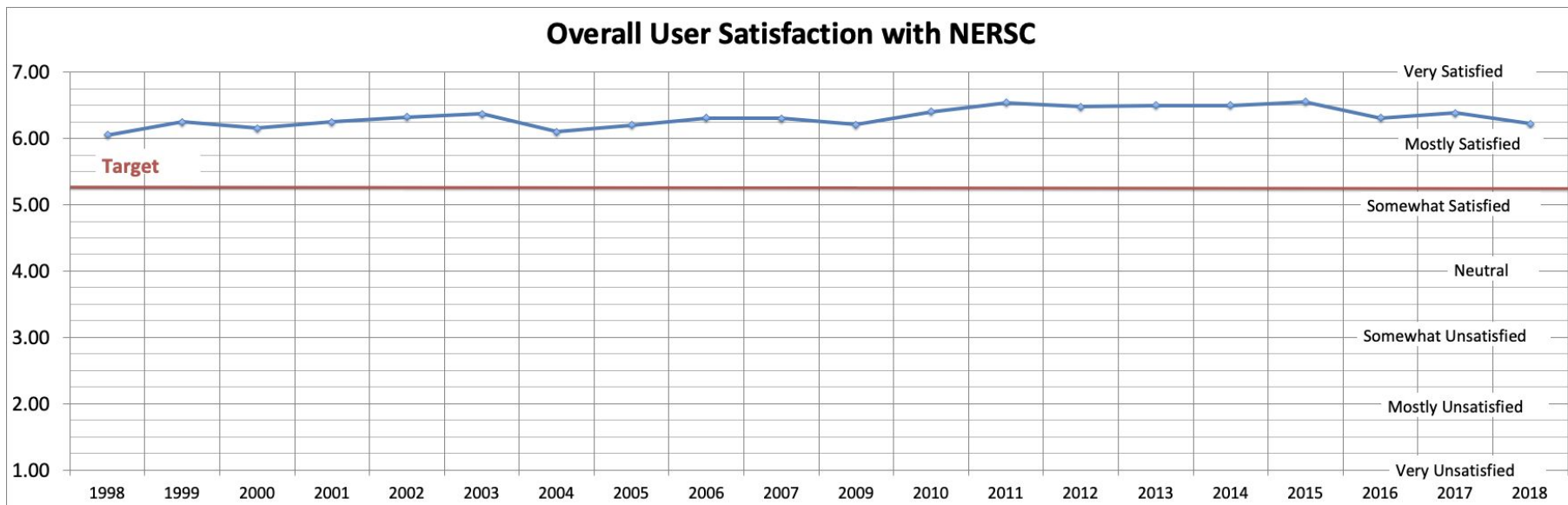
User Satisfaction Remained High in 2018

Survey Area	Target	2017 Actual	2018 Actual	Significant Change
Overall Satisfaction	5.25	6.38	**6.22**	-0.16
Avg. of User Support ratings (Services Overall)	5.25	6.42	**6.36**	-0.14

...Across all categories included in the survey

Overall Satisfaction

User Survey & Satisfaction
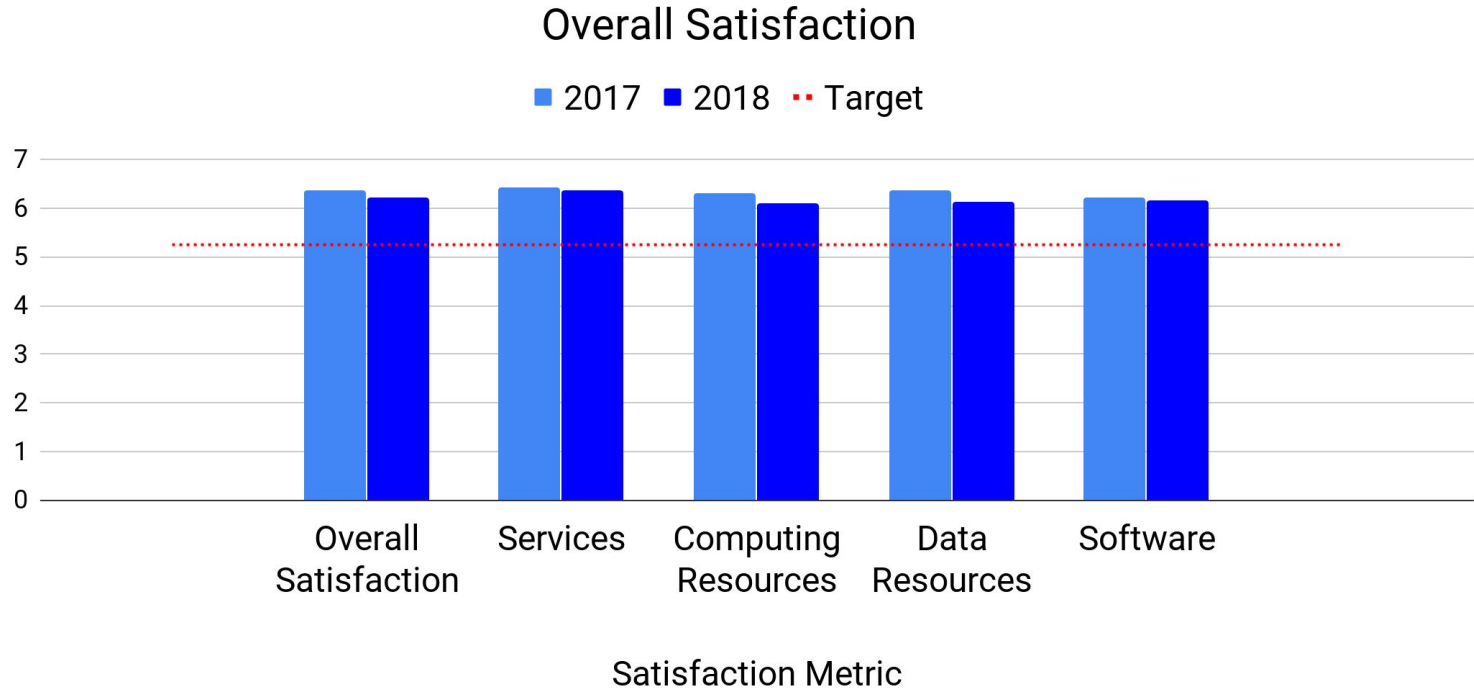
Factors Increasing Satisfaction

- Consulting support
- Web-based documentation
- Large-scale resources
- Provision of software
- System stability and uptime

Factors Decreasing Satisfaction

- Queue wait times
- Storage space limits
- Requirement to use MFA
- Compute time/allocation limits
- System interrupts and maintenances

III. Problem Resolution

- Front-line consulting team is made up of staff with diverse backgrounds in computational science, almost all with advanced degrees
- Tickets that cannot be answered by consulting team are addressed by other experts within NERSC
- 6,736 consulting & account support tickets submitted by 2,177 different users in 2018
- Problem resolution target: >80% of user tickets addressed within 3 business days
- 2018 statistic: **90.5%** addressed within 3 business days
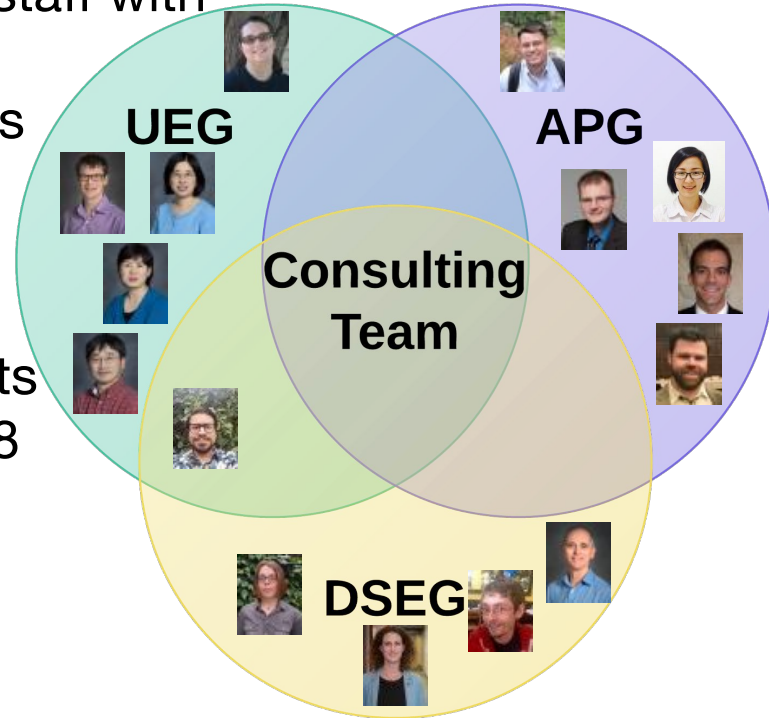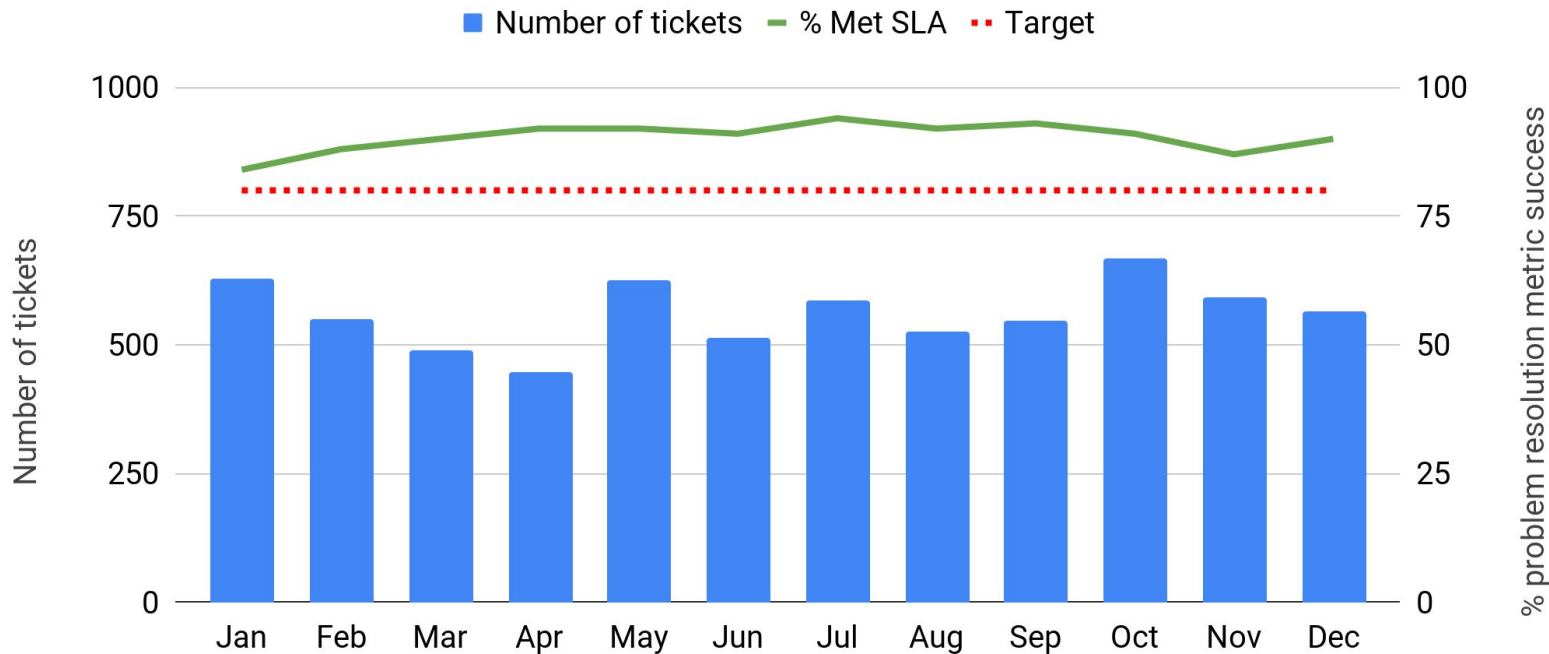
Problem Resolution Satisfaction Rates High

Survey Scores	Target	2017	2018
Resolution response time	5.25	6.41	**6.43**
Quality of technical advice	5.25	6.44	**6.43**
Time to solution	5.25	6.29	**6.27**
Overall satisfaction (Problem Resolution Average Score)	5.25	6.46	**6.47**

Problem Resolution Consistent in 2018

Number of Tickets & Problem Resolution Metric

User Queries Across Categories

Breakdown by Category

- Software 17.5%
- Running Jobs 19.2%
- Account Support 21.2%
- Data/IO 10.0%
- Other 7.3%
- Allocations 5.7%
- Programming 4.8%
- ERCAP 5.0%
- Performance 1.6%
- Login Node 4.6%
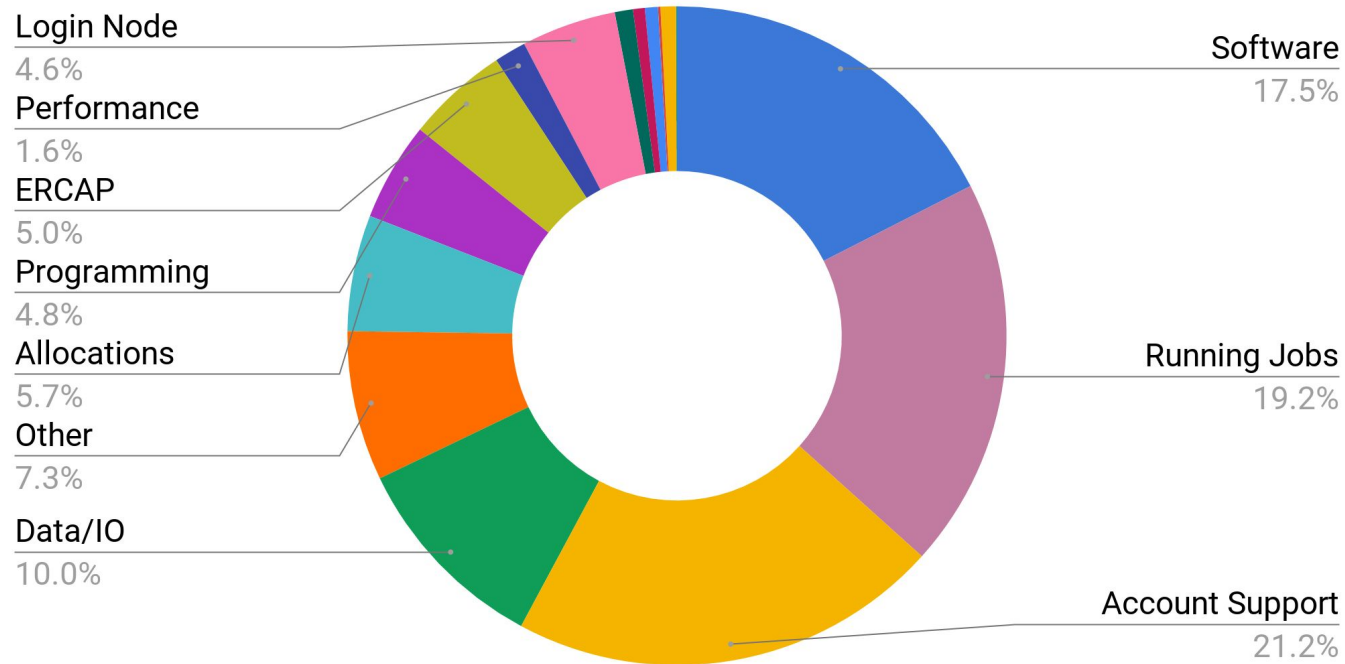
IV. User Support Success Stories: Training

- Catering to wide variety of users
- Beginning users how-to training
- Advanced trainings on tools, code optimization, hackathons
- Focus on machine learning and other techniques for data analysis
  - Inaugural Machine Learning for Science workshop
  - SC18 machine learning tutorial
- Focus on adoption of Cori KNL as advanced pre-exascale architecture
  - Performance optimization and tools trainings
  - Added emphasis because of Edison's imminent retirement

Transitioning Workload to Advanced Architectures

- Continued focus on modernization of user codes to exploit advanced architectures such as Cori
- NERSC assistance to users through NESAP program, training events, hackathons, user & code community engagement

Significant *NESAP for Data* App Improvements

Jonathan Madsen
TomoPy (APS, ALS, etc)
- GPU acceleration of iterative reconstruction algorithms
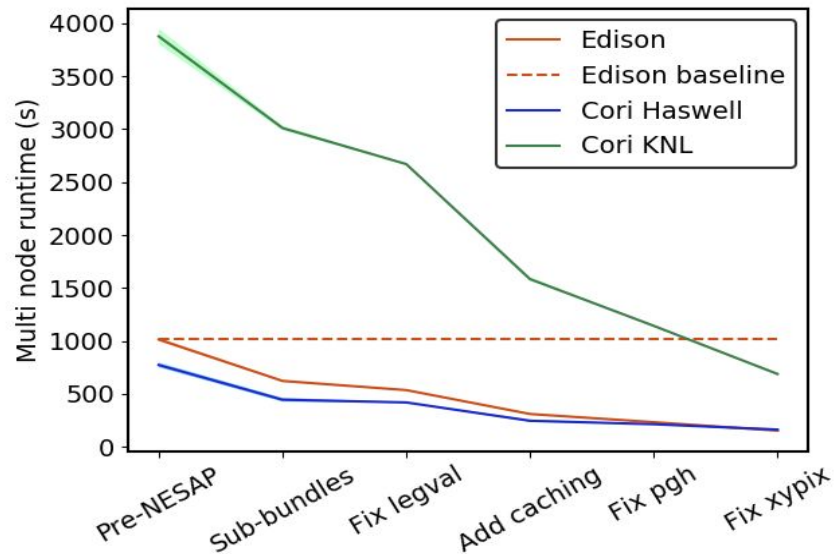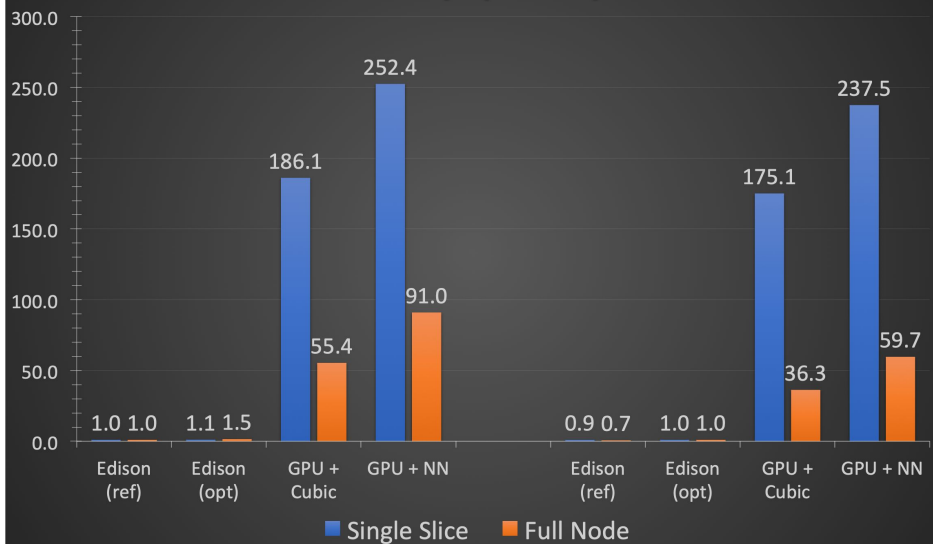- New results from first NERSC-9 hack-a-thon w/NVIDIA, >200x speedup!

Laurie Stephey
DESI Spectroscopic Extraction
- Optimization of Python code on Cori KNL architecture
- Code is 4-7x faster depending on architecture and benchmark
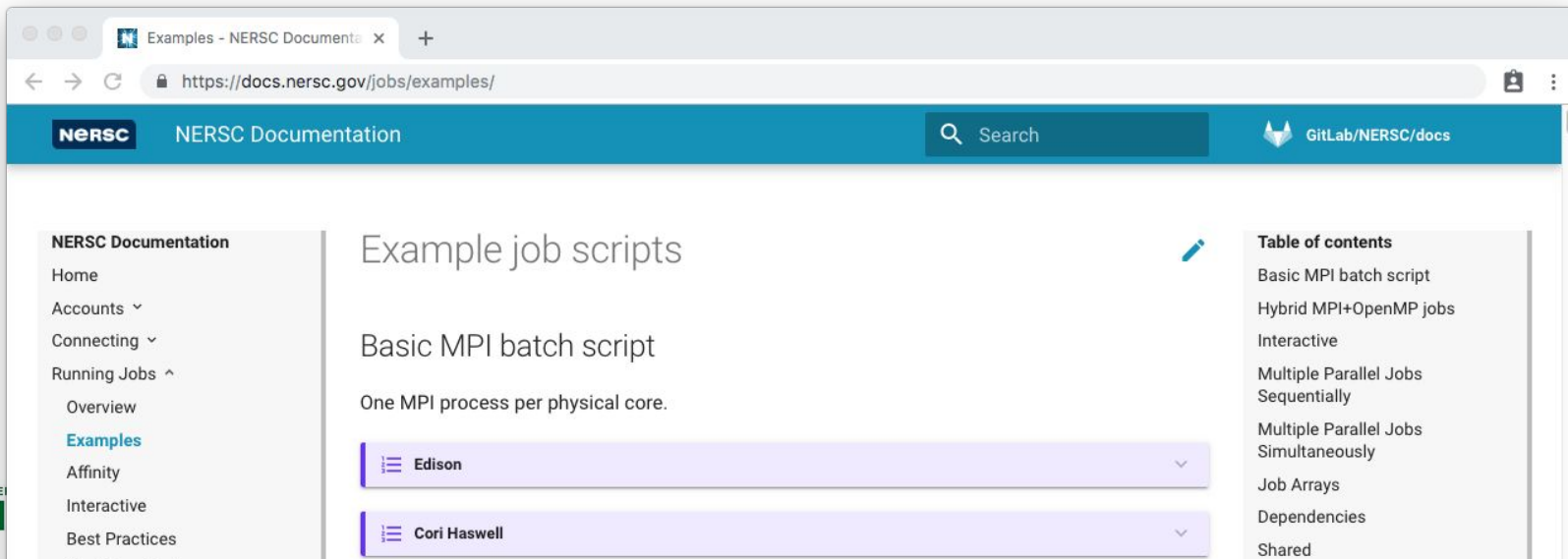
V. Communication & Outreach

- Communication to users: communicate updates, items of importance to users
- NERSC weekly email: weekly updates to users on Mondays
- Monthly NERSC User Group (NUG) webinars
- New in 2018: "NERSC User News" podcast
  - Featured content on topics of interest to NERSC users
  - First podcast about why maintenance takes so long
  - Other topics: HPSS, electrical maintenance, machine learning, Perlmutter, MFA (3 episodes)

Communication & Outreach

- Communication between NERSC and users
- Most user issues resolved via tickets
- New support model: virtual office hours
  - Some issues easier to fix with fast turnaround
  - Zoom teleconferencing allows screen sharing
  - Held office hours for ERCAP allocations program and MFA conversion
  - Helped more than 120 users during office hours

Key Documentation Improvements in 2018

- High-quality documentation is time-consuming to maintain
- Now developing documentation on, & serving from Gitlab platform:
  - Making it easier to add new content,
  - Encouraging a more open review process,
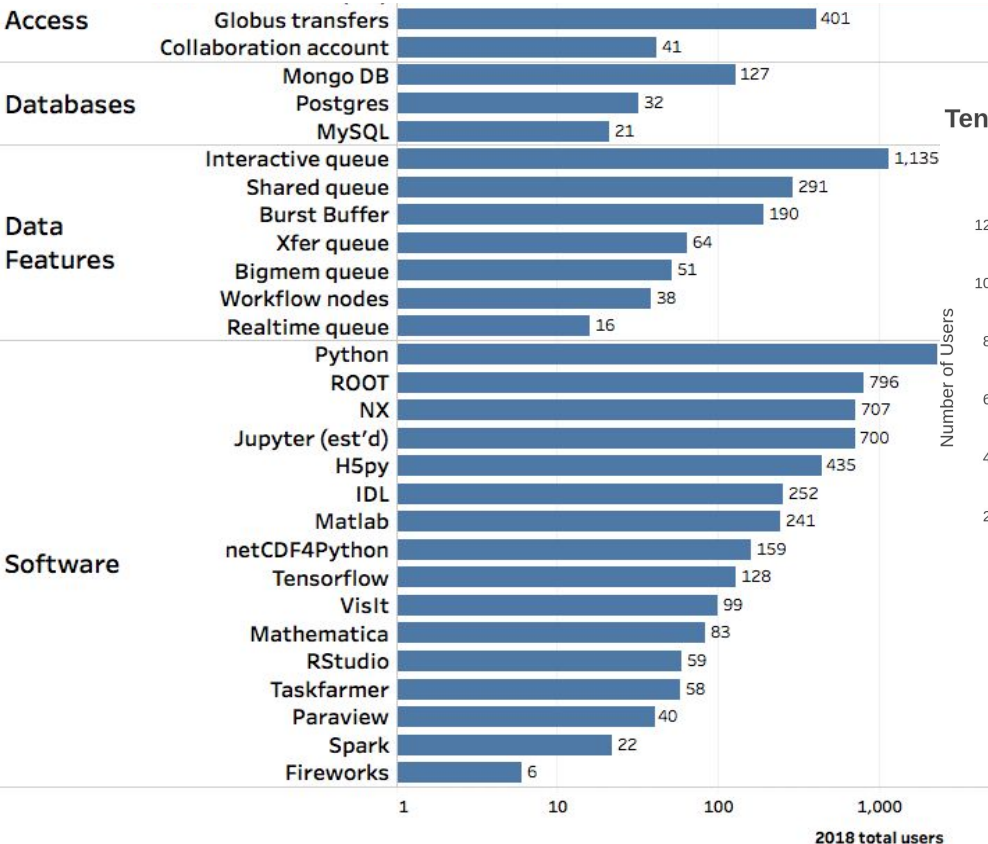 - Enabling others to easily contribute to the documentation

Strong Adoption of Data Software Stack

Access
- Globus transfers: 401
- Collaboration account: 41

Databases
- Mongo DB: 127
- Postgres: 32
- MySQL: 21

Data Features
- Interactive queue: 1,135
- Shared queue: 291
- Burst Buffer: 190
- Xfer queue: 64
- Bigmem queue: 51
- Workflow nodes: 38
- Realtime queue: 16

Software
- Python
- ROOT: 796
- NX: 707
- Jupyter (est'd): 700
- H5py: 435
- IDL: 252
- Matlab: 241
- netCDF4Python: 159
- Tensorflow: 128
- VisIt: 99
- Mathematica: 83
- RStudio: 59
- Taskfarmer: 58
- Paraview: 40
- Spark: 22
- Fireworks: 6

2018 total users

TensorFlow Usage on NERSC Systems

Number of Users — Year
- 2017: 35
- 2018: 128
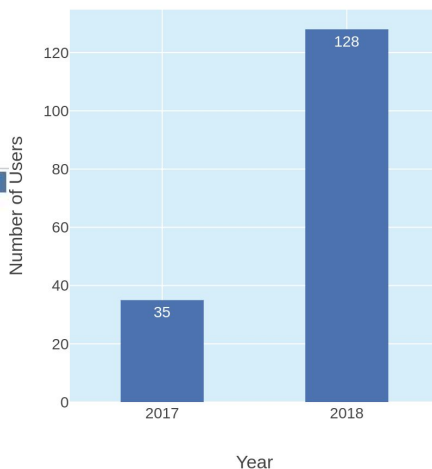
Jupyter Usage on NERSC Systems

Number of Users — Month
- 12/2017: 200
- 12/2018: 297
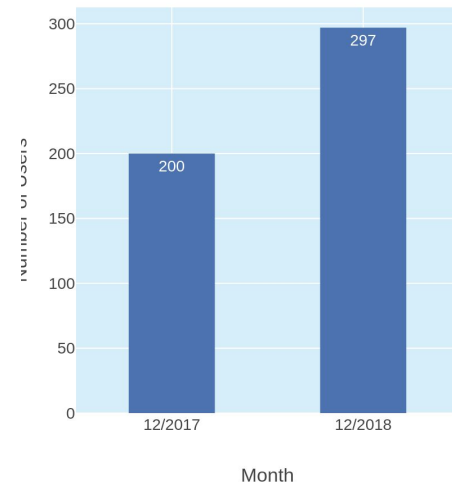
Outreach: Training Local Students on HPC

- NERSC helped Laney College (local community college) train team to compete in SC18 Student Cluster Competition
- Led to creation of Laney Supercomputing Club
- Most students still involved in HPC

Outreach to Next-Gen InfoSec Workforce

- NERSC local site for DOE Cyberforce competition
- Four teams participated locally, 66 teams across six DOE labs
- NERSC provided Red, Green, White teams
- Local winner UC Davis was second place overall

Key Support Highlights: ACTIV pipeline I/O

- ACTIV: neurological disorder pipeline
- Large-scale I/O taking large fraction of time
- NERSC patch to h5py reduced I/O time by factor of 20
- 1.5x speedup for entire application
- Patch accepted into h5py
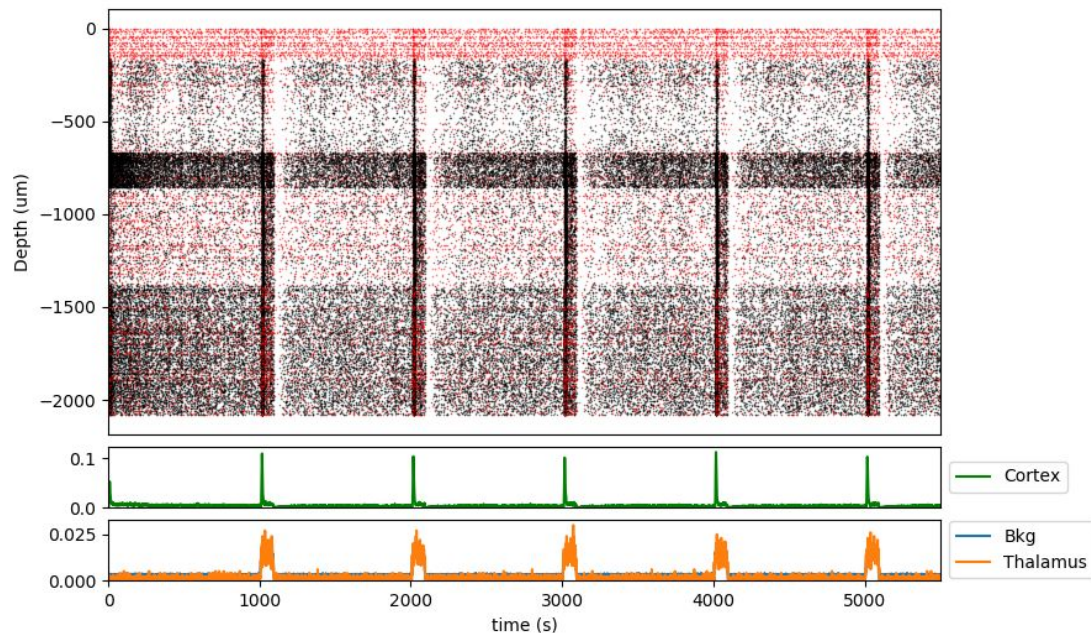
Microscale neural activity modeled by ACTIV. In the top panel dots represent spiking neurons, the middle panel shows population-averaged spike rate, and the bottom panel shows input to the simulated network.

Key Support Highlights: Particle Collision Data at Scale

- BNL STAR nuclear datasets: PB scale
- Reconstruction processing takes months at BNL computing facility
- With help from NERSC consultants & storage experts, & ESNet networking experts, built highly scalable, fault-tolerant, multi-step data-processing pipeline
- Reconstruction process reduced from months to weeks or days
- Scaled up to 25,600 cores with 98% end-to-end efficiency
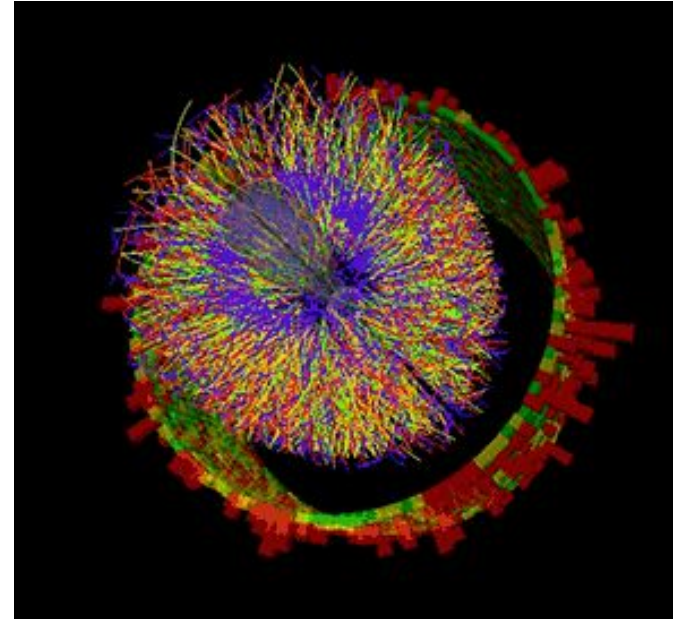
A series of collision events at STAR, each with thousands of particle tracks and the signals registered as some of those particles strike various detector components.

User Support & Outreach: Summary

- NERSC supported 7000 users in 703 projects overall
- NERSC exceeded all user satisfaction metrics
- NERSC resolved 90% of user tickets within 3 business days, exceeding problem resolution metric
- NERSC provided extensive training and significant support for user code performance improvements
- NERSC keeps its users informed via weekly email, podcast
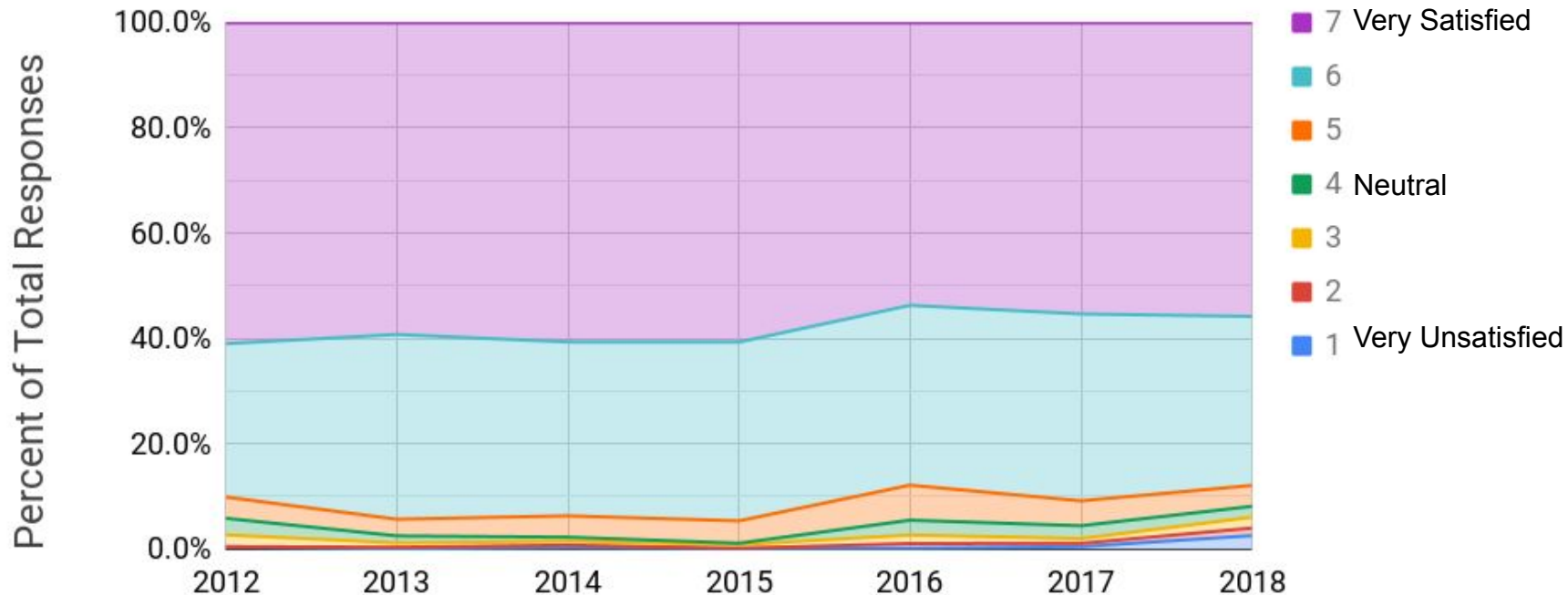- NERSC interacts with users via support and virtual office hours

Thank You

Overall Satisfaction Score

Overall Satisfaction — Avg 6.23
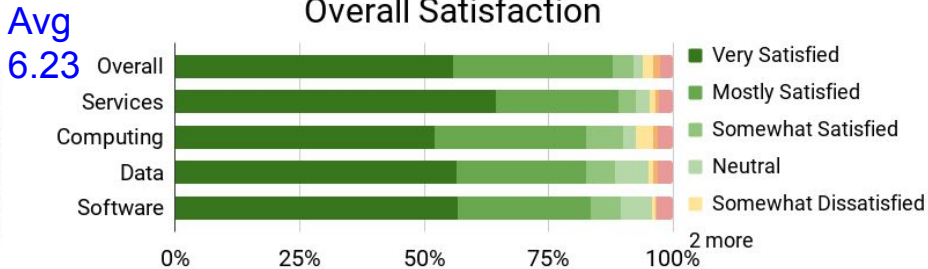
1M+ users (150 responses) — Avg 6.27
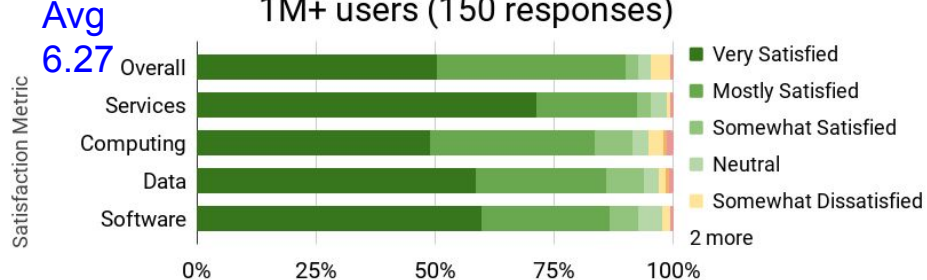
100M+ users (10 responses) — Avg 6.9
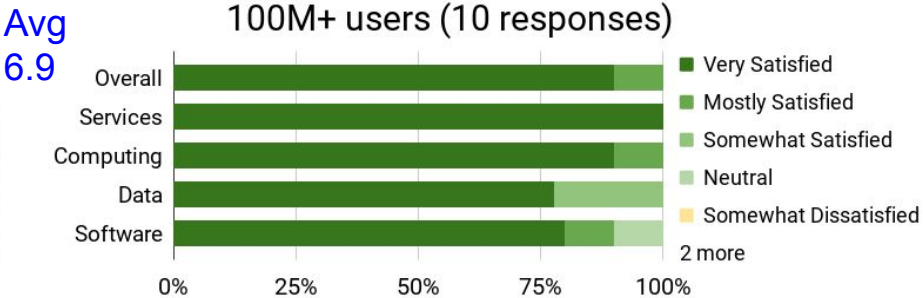
< 1M+ users (273 responses) — Avg 6.10
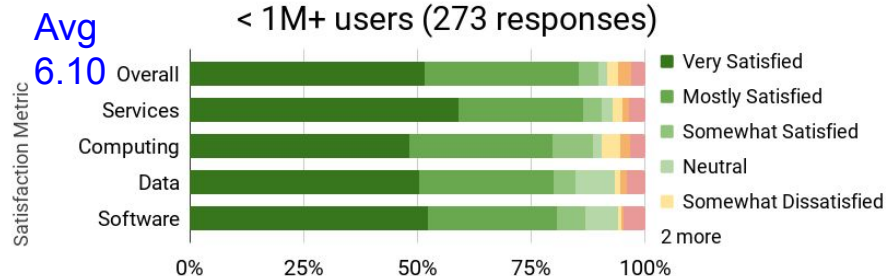
10M+ users (60 responses) — Avg 6.55
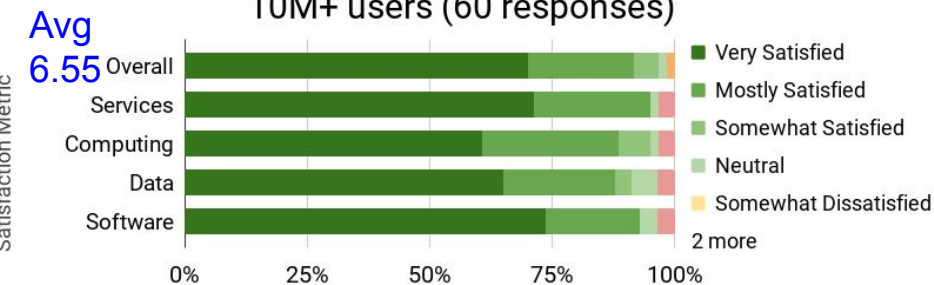
indirect users (used no time, PIs or others) (149 responses) — Avg 6.33

Legend: Very Satisfied, Mostly Satisfied, Somewhat Satisfied, Neutral, Somewhat Dissatisfied, 2 more

Satisfaction Metric (y-axis): Overall, Services, Computing, Data, Software