Results of CC* Assessment

The purpose of the CC* Assessment was to ascertain current use of GPUs for research and/or teaching at NJIT and anticipated future use if public-access were to become available. The intention was to gather use-based support for an NSF grant application for funding of public-access GPUs.
NJIT faculty, researchers, and postdocs who have an account on any cluster were invited to participate in the assessment.
70 invitees completed the assessment and 1 completed all but the final written comment; thus the assessment contains data from 71 participants.

Content of the Assessment:

After specifying their NJIT status, school/college, department, how long they have been using IST HPC, and general and specific areas of research, the assessment focused on GPU use.
All participants were asked to list GPU software they used for their specific research areas(s), and whether they would use public access GPUs if available. Participants who responded “no” or don’t know” were invited to provide an optional comment, but no further input was collected.
Participants who responded “yes” were queried further about their anticipated use of public access GPUs in research and teaching, and, in addition to the optional comment, were asked to describe the importance of public access GPUs in their work.

Results:

Demographics

Chart 1 indicates participants’ position at NJIT.

Chart 1

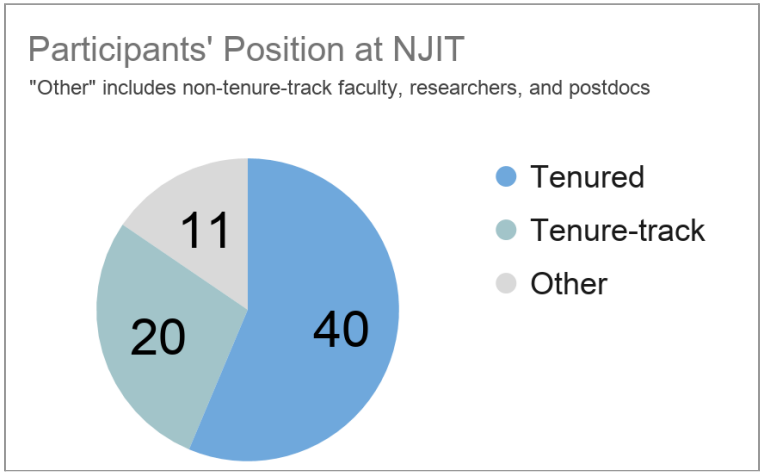


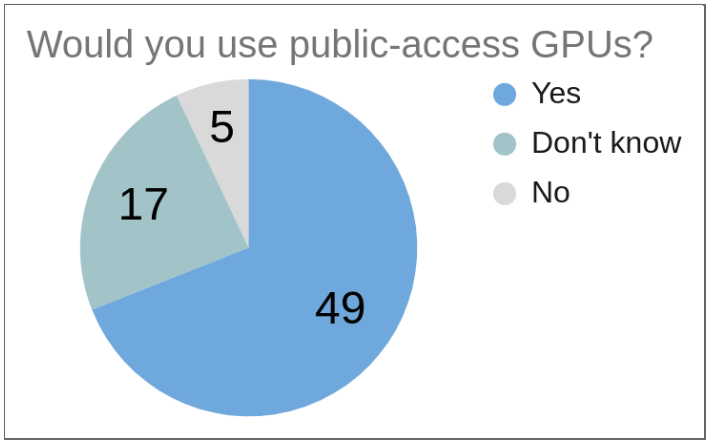
Table 1 shows participants’ approximate length of time for using IST-managed high performance computing (HPC) resources.

Table 1

Use of IST/HPC	Number of participants
< 6 months	5
6+ to 12 months	2
1+ to 2 years	14
2+ to 5 years	20
>5 years	27
Don’t know	3

All participants were asked whether they would use public- access GPUs if available. Chart 2 shows that 69% of participants (49 out of 71) responded YES.

Chart 2



NJIT positions were similar across the YES and DON'T KNOW response groups; tenured faculty accounted for 59% of the YES group and 53% of DON'T KNOW group.

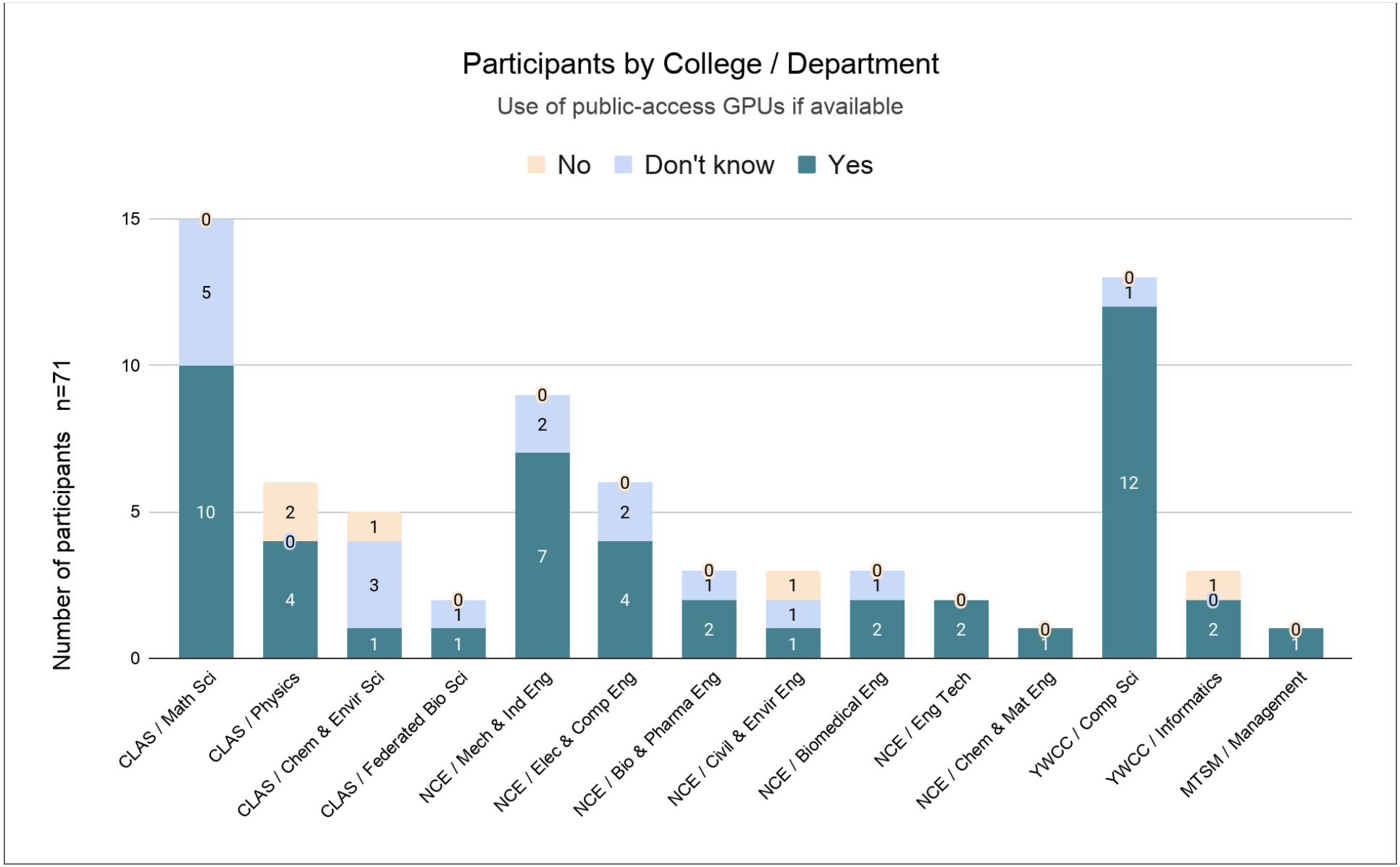
There were similarities and differences in departmental representation across the YES and DON'T KNOW groups:

- Similarities:
  - Participants from the Dept. of Mathematical Sciences in the College of Science and Liberal Arts accounted for 20% of the YES and 29% of the DON'T KNOW respondents
  - Participants from Mechanical and Industrial Engineering in the Newark College of Engineering accounted for 12% of both the YES and DON'T KNOW respondents.
- Differences:
  - Participants from the Dept. of Computer Science in the Ying Wu College of Computing accounted for nearly 25% of the YES group, but only 6% of the DON'T KNOW group.

Chart 3 shows the Yes/No/Don'tKnow responses (for whether they would use public-access GPUs) for all 71 participants by College (or School) and department.

Note that CLAS = College of Liberal Arts and Sciences; NCE = Newark College of Engineering; YWCC = Ying Wu College of Computing; MTSM = Martin Tuchman School of Management

Chart 3



The 49 YES respondents were asked whether public-access GPUs would have a *large*, *moderate*, *small*, or *no effect* on their research and teaching. Those whose responses were anything other than *no effect* were queried further about the nature of the effects.

Charts 4 and 5 show the reported effects of public-access GPUs. Table 2 presents responses to the further query, in which participants were asked whether the effects of public-access GPUs would allow *larger versions of current models*, allow *model types currently not feasible*, or some to-be-specified *other* effect. (Note that participants could choose more than one response for these queries, hence the number of tabulated responses exceeds the number of participants.)

Chart 4

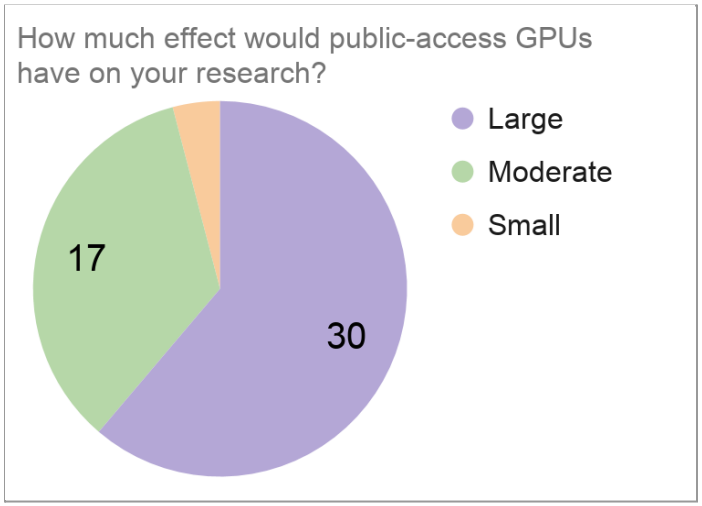


Chart 5

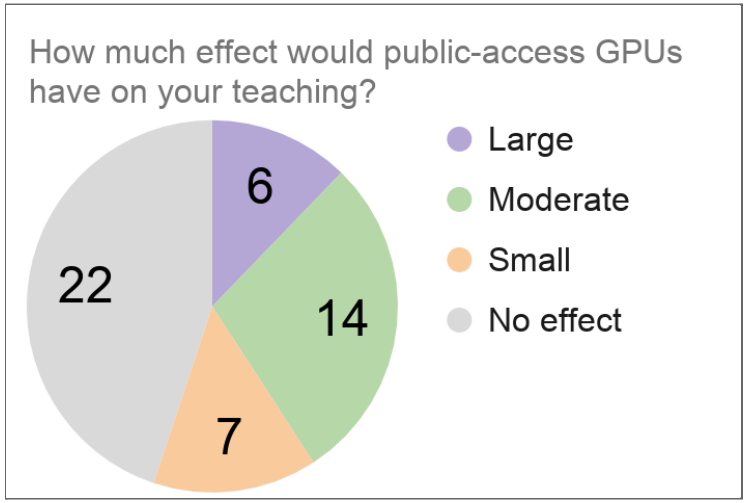


Table 2

Effects of public-access GPUs on RESEARCH			
	Large effect	Moderate effect	Small effect
Allow larger versions of current models	26	17	2
Allow different types of models currently not feasible	22	15	1
OTHER	5	0	0

OTHER responses:	(1) Develop more scalable algorithms; (2) permit proper ensemble averaging; (3) It would reduce run times; (4) increase computational parallelisms; (5) Permit training deep learning models and solving CO problems much faster		
Effects of public-access GPUs on TEACHING			
	Large effect	Moderate effect	Small effect
Allow larger versions of current models	3	9	1
Allow different types of models currently not feasible	0	10	2
OTHER	3	3	5
OTHER responses:	(1) Expose students to more assignments, practice, and a better learning experience; (2) I would incorporate more use of GPUs in the HPC class. In DMS; (3) it makes possible to teach dee learning methods in a class; current #GPU des aren't sufficient for all students in a class.	(1)Provide an HPC platform for teaching; (2)Allow students to have a try on implementing algorithms on GPU; (3) Introduce modelling to students in classes I teach.	(1) Allow for more practice oriented projects in my machine learning class; (2) Educate students on CFD; (3) Students would be more aware of the potential advantages of GPUs in computation, but are unlikely to use them outside research. (4) Unclear; (5) Provide student access to advanced analytics

The 49 YES respondents were asked to indicate the extent of their anticipated use of public-access GPUs, i.e. *heavy*, *moderate*, *light*, *none*, or *don't know*, in their research and teaching.

Charts 6 and 7 show the reported anticipated GPU use.

Chart 6

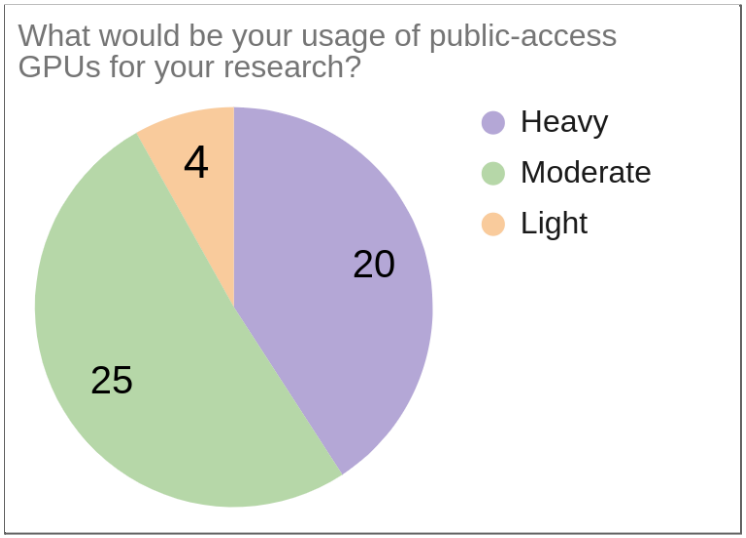


Chart 7

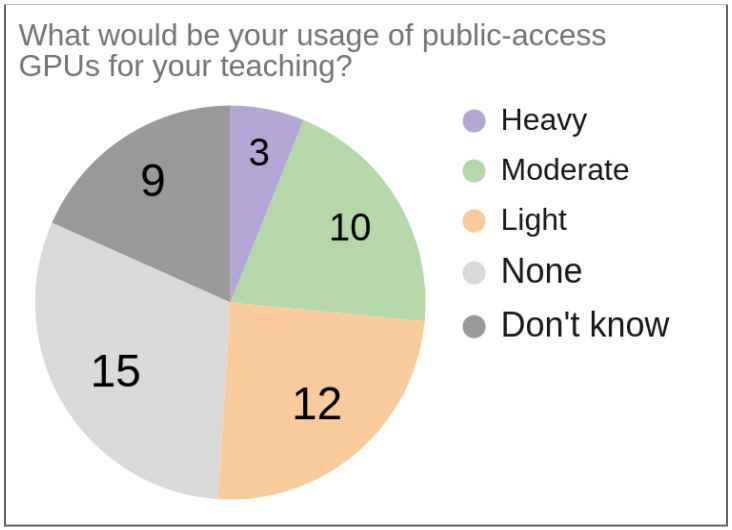


Table 3 lists the general research area(s), specific research area(s), current use of GPU software in those specific research areas, and the anticipated effect on research as reported by all 71 participants. Each row represents one participant's report of a single research area-- but as participants were able to report more than one research area, any participant could be represented by several rows, contiguous or not. Hence the table contains many more rows than 71 (the total number of participants).

Table 3

General area Number of participants			
	Specific area	GPU software	Public access effect
Neural networks, genetic algorithms 22 participants			
1	Learning encoding, decoding, channel modeling, random access strategies for	Python, Matlab, Tensorflow	Large

	next generation communication systems (6G and beyond, Internet-of-Things). Also, I have been working on privacy issues in learning and specifically, in federated learning, where distributed NN training is studied under adversarial influence. Another application is data-driven private file retrieval and computation, where NNs are used to model query generation, and encoding and decoding blocks.		
2	We design large neural nets to detect software bugs/faults and auto fix them.	tensorflow, keras, pytorch, MXnet, Caffe, PaddlePaddle	Large
3	Pytorch, TensorFlow	We develop and write our own algorithms for the methods we study.	Large
4	I am working on the integration of optimization and deep learning (e.g., neural networks, LSTM) approaches.	CPLEX, Phyton, C++	Large
5	genetic algorithms and other optimization methods, in particular data assimilation, to perform parameter estimation for biological models	MATLAB	Large
6	I did it on my own machines in the lab	Python and tensor flow	Large
7	Deep neural networks for image, medical data, and text classification	Python, scikit-learn, Keras (Tensorflow), and Pytorch deep learning libraries	Large
8	Use deep neural networks to mine solar images.	Python, TensorFlow, Keras, PyTorch, Jupyter Notebooks, SunPy, HelioPy, Astropy	Large
9	Various types of DNN and Genetic Search algorithms	Tensorflow, Pytorch, CUDA, python, C/C++	Large
10	We use neural networks to learn relationships between concepts in medical terminologies.	So this is done by my student and I don't know the details.	Large
11	We use neural networks and genetic algorithms to make classifications and parameter estimation of questions in Neuroscience.	*Do not currently use GPU software*	Large
12	We use neural networks and genetic algorithms to solve nonlinear optimization problems of human movement.	Pytorch, TensorFlow	Large
13	Deep neural nets for image/signal classification	python and matlab	Moderate
14	Image identification for biodiversity applications	*Do not use GPU software*	Moderate

15	Image and graph characterizations for computational social science.	PyTorch	Moderate
16	Reinforcement learning	Tensorflow, Keras	Moderate
17	Deep learning models.	*Do not use GPU software*	Small
18	using them for solar physics and gene prediction and classification	tensorflow, python	N/A -- DON'T KNOW for public-access GPU use
19	neural nets for computational fluid dynamics simulations	tensorflow	N/A -- DON'T KNOW for public-access GPU use
20	Simulations of radar systems implemented by neural networks.	Matlab, C	N/A -- DON'T KNOW for public-access GPU use
21	Potential energy surface (PES) of metallic organic frameworks	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
22	We use genetic algorithms to parametrize force fields that allow long-time molecular dynamics simulations of material systems.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
Computational fluid dynamics 14 participants			
23	we co-develop ClimateMachine, a large GPU based software for climate forecast using machine learning.	in house development of ClimateMachine (with Caltech, MIT, NPS)	Large
24	Particle simulations for Fluid-Structure Interaction, Turbulence and Waves	ANSYS FLUENT, Converge CFD, In house code	Large
25	Simulation of fluid flow in a variety of mixing systems, such as dissolution testing equipment, mixing vessels, pharmaceutical reactors, etc.	Ansys, M-Star	Large
26	Simulation of time-dependent physical systems. This often includes one or more effects of chemistry, electromagnetism, and elasticity (e.g., soft matter) coupled to	*Do not currently use GPU software*	Large

	fluid transport.		
27	We study the transport of particles and drops fluids. The simulations are performed using a finite element code developed by my group. We also do simulations of particulate suspension subjected to uniform and non-uniform electric fields.	*Do not currently use GPU software*	Large
28	numerical solutions of PDE's using finite volume and finite difference methods	C++ and fortran compilers	Large
29	We use open source or commercial CFD code (such as OpenFOAM or ANSYS Fluent) for respiratory and blood flow analysis.	*Do not currently use GPU software*	Large
30	We need to run Fluent and OpenFOAM cases on Kong.njit.edu.	*Do not currently use GPU software*	Moderate
31	Moving boundary problems for the evolution of drops, cells and vesicles	*Do not currently use GPU software*	Moderate
32	direct numerical simulations of complex flows	in house	N/A -- DON'T KNOW for public-access GPU use
33	We used Ansys Fluent software to conduct large eddy simulation to investigate initial droplet size for an oil jets emanating from a 10 mm pipe. We also collaborated with Johns Hopkins University and compared their experimental results to our findings. The findings are under review of the Journal of Heat and Fluid Flow	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
34	My research group uses computational approaches, ranging from molecular dynamics, Monte Carlo simulations and computational fluid dynamics in conjunction with the theoretical formalisms of equilibrium and non-equilibrium statistical mechanics, and high-performance computing in massively parallel architectures, to solve problems of transport in in heterogeneous materials with applications in enhanced oil recovery and targeted drug delivery.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
35	Interfacial behaviors in two-phase viscoelastic fluids	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
36	The PDEs described above arise in fluid dynamical contexts (free surface film flow leads to the 4th order nonlinear parabolic PDE, and fluid filtration leads to other PDE systems).	C++, Fortran, Matlab	N/A -- DON'T KNOW for public-access GPU

			use
Monte Carlo 14 participants			
37	Numerical computation of error probabilities for next generation communication systems (6G and beyond, Internet-of-Things), training of reinforcement learning systems with randomized training data	C, C++, Python, OpenGL, Matlab	Large
38	Monte Carlo Tree Search	*Do not currently use GPU software*	Large
39	Research on Monte Carlo simulations was conducted several years ago using the computing network. Physical phenomenon under study was 'density relaxation'	*Do not currently use GPU software*	Large
40	DFT calculations for force field development, classical (force-field based) Molecular Dynamics and Monte Carlo simulations	*Do not currently use GPU software*	Large
41	We do experiments with randomized algorithms for global optimization.	*Do not currently use GPU software*	Moderate
42	Reinforcement Learning	Tensorflow	Moderate
43	Because we develop detection and estimation methods, before applying them to real data, we need to test them on multiple realizations. Thus, we repeat processes N times (N large) for every algorithm and for numerous sets of parameters.	*Do not currently use GPU software*	Moderate
44	My research involves designing new efficient Monte Carlo methods. As part of this work, my students will run many (thousands of) Monte Carlo experiments to study the efficiency of the methods that we devise.	*Do not currently use GPU software*	Moderate
45	Monte Carlo simulations of wireless and radar systems.	Matlab	N/A -- DON'T KNOW for public-access GPU use
46	I run uncertainty analyses using Monte Carlo and Advanced Mean Value methods for my physics-based simulations.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
47	My research group uses computational approaches, ranging from molecular dynamics, Monte Carlo simulations and computational fluid dynamics in	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU

	conjunction with the theoretical formalisms of equilibrium and non-equilibrium statistical mechanics, and high-performance computing in massively parallel architectures, to solve problems of transport in in heterogeneous materials with applications in enhanced oil recovery and targeted drug delivery.		use
48	Some of the filtration problems we consider deal with transport of large particles through pore networks. Monte Carlo methods can be used to simulate these stochastic processes. We also use Monte Carlo methods to address problems in non-Newtonian Hele-Shaw flow, which can be treated as a random walk in 2 space dimensions.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
49	Design of Nuclear Reactor Fuel Rods	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
50	Study of the diffusion in porous materials	*Do not currently use GPU software*	N/A -- NO for public-access GPU use
Statistical analysis 12 participants			
51	We have been studying eigensubspace method for investment portfolio design. This study requires empirical correlation matrix generation for stock returns in real time and multiple sampling frequencies. It is one of the major applications of the bigdata finance.	We develop and write our own algorithms for the methods we study.	Large
52	We mainly use this to analyze the relations among different variables	*Do not currently use GPU software*	Large
53	Social Network Analysis; Gravity Model estimation	R, Python	Moderate
54	Simulations to study the operating characteristics of estimators in statistical models for analysing censored data	*Do not currently use GPU software*	Moderate
55	Data and behavioral characterization in online social platforms	*Do not currently use GPU software*	Moderate
56	Senior undergraduate student is using high performance comping for the research and analysis of data.	R programming language.	N/A -- DON'T KNOW for public-access GPU use

57	Processing large data sets, analyzing them; running large scale simulations	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
58	Magnetic field distributions	*Do not currently use GPU software*	N/A -- NO for public-access GPU use
59	Phylogenetic reconstruction	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
60	Simulations of wireless and radar systems.	Matlab	N/A -- DON'T KNOW for public-access GPU use
61	My research group uses computational approaches, ranging from molecular dynamics, Monte Carlo simulations and computational fluid dynamics in conjunction with the theoretical formalisms of equilibrium and non-equilibrium statistical mechanics, and high-performance computing in massively parallel architectures, to solve problems of transport in in heterogeneous materials with applications in enhanced oil recovery and targeted drug delivery.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
62	Modelling of Parking Lots in NYC and Nature of Dynamics within the Parking Lot - Number of Parking Spots over time in various parts of NYC.	*Do not currently use GPU software*	N/A -- NO for public-access GPU use
Computational PDE 8 participants			
63	Numerical solution of partial-integro-differential equation model for granular dynamics	Developing and writing custom code	Large
64	Various techniques for the solution of transport and related problems: boundary integral methods, spectral methods for solution of nonlinear PDEs, numerical construction of conformal maps, simulation of time-dependent phenomena.	*Do not currently use GPU software*	Large
65	We use the HPC for computational PDE, including computational fluid dynamics for air and blood flow, biomechanical finite element analysis, and bioheat analysis.	An in house developed finite volume CUDA code to solve the bioheat equation on Cartesian	Large

		grids.	
66	numerical solutions of PDE's using finite volume and finite difference methods	C++ and fortran compilers	Large
67	I use and develop fast algorithms to solve integral equation formulations of boundary value problems, including moving boundaries	*Do not currently use GPU software*	Moderate
68	Running simulations of the Cahn-Hilliard equation, a PDE that describes phase separation processes.	*Do not currently use GPU software*	Moderate
69	We are implementing a semi-discrete scheme for numerical solving an integro-PDE. Part of this process is to determine conditions to insure numerical stability.	*Do not currently use GPU software*	Moderate
70	We solve numerous PDEs using computational models. These simulate sound propagation in ocean environments.	*Do not currently use GPU software*	Moderate
Computational physics and chemistry 8 participants			
71	Modeling and data analysis	*Do not currently use GPU software*	Large
72	Molecular dynamics simulation of small compounds in solution	GROMACS	Large
73	My research in the area of 'granular science' falls under the general category of Computational Physics	Discrete element code	Large
74	DFT calculations for force field development, classical (force-field based) Molecular Dynamics and Monte Carlo simulations	*Do not currently use GPU software*	Large
75	Molecular structure, bond energies, thermochemistry and physics, kinetics, rate constants, equilibrium, energy transfer, elementary reaction mechanisms	Gaussian, home built software,	N/A -- DON'T KNOW for public-access GPU use
76	I develop physics-based simulations of human movement.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
77	Simulations of Band Structure, electronic, optical and transport properties of semiconductors, semiconductor alloys	Density Functional Theory, VASP	N/A -- NO for public-access GPU use
78	Electronic structure, molecular dynamics	*Do not currently use GPU*	N/A -- DON'T KNOW

		software	for public-access GPU use
Materials research 8 participants			
79	Materials Modeling	VASP	Large
80	Using molecular dynamics simulations we study the mechanical behavior of two dimensional nano structures.	Do not currently use GPU software	Large
81	HPC is used to investigate the synthesis mechanism of PN materials and their reaction pathways.	DFT	Large
82	We do materials discovery and drug discovery research	Rosetta	Large
83	DFT calculations for force field development, classical (force-field based) Molecular Dynamics and Monte Carlo simulations	*Do not currently use GPU software*	Large
84	For modeling of plastic based nanocomposites	*Do not currently use GPU software*	
85	First principles study are being performed to analyze the structural, elastic, electronic and optical properties of perovskites. The effects of alkaline earth metal cation substitution on their overall properties are further investigated. The calculations are performed using the projector augmented wave (PAW) within GGA-PBE and HSE06 formalism.	DFT, VASP	N/A -- NO for public-access GPU use
86	We run molecular dynamics simulations of layered structures in dry and hydrated conditions to understand water dynamics in nonporous materials.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
Computational biophysics 6 participants			
87	Molecular dynamics simulation of protein aggregation	GROMACS	Large
88	We study the transport of biological cells in fluids. The simulations are performed using a finite element code developed by my group.	*Do not currently use GPU software*	Large
89	We use the HPC for computational biophysics, including simulations of respiratory and blood flow and tissue biomechanics.	*Do not currently use GPU software*	Large
90	We work with Computational Neuroscience where models are biophysically	*Do not currently use GPU*	Large

	grounded. The work involve simulations of differential equations (neurons) interconnected and demands a lot of CPU and RAM memory.	*software*	
91	My research group uses computational approaches, ranging from molecular dynamics, Monte Carlo simulations and computational fluid dynamics in conjunction with the theoretical formalisms of equilibrium and non-equilibrium statistical mechanics, and high-performance computing in massively parallel architectures, to solve problems of transport in in heterogeneous materials with applications in enhanced oil recovery and targeted drug delivery.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
92	Collective behaviors in confined domains	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
Bioinformatics 4 participants			
93	Deep learning models on genome sequence data	Python, scikit-learn, Keras (Tensorflow), and Pytorch deep learning libraries	Large effect
94	We are doing Medical Informatics, which is not the same as Bioinformatics, but close. We are working with large medical terminologies.	Google BERT.	Large effect
95	deep learning computation; reads alignment, variant calling for tens, hundreds, or even thousands of samples (one sample/one job).	PyTorch, Keras, bioinformatics tools such as GATK, BWA, fastqc, ANNOVAR, htseq, plink,	Moderate
96	Phylogenetics, gene annotation, genome assembly simulations,	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
Software verification, static analysis 4 participants			
97	We build static approaches to analyze programs and conduct testing to detect bugs/faults	tensorflow, keras, pytorch, MXnet, Caffe, PaddlePaddle and some existing academic tools for parsing software	Large effect
98	Perform static analysis for algorithm design and performance, etc.	*Do not currently use GPU software*	Large effect

99	Automated program repair	*Do not currently use GPU software*	Moderate
100	I run finite element models.	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
Electromagnetism, wave propagation 2 participants			
101	Solve the wave equation in complex geometries	*Do not currently use GPU software*	Large effect
102	Semiconductor device simulation using TCAD	*Do not currently use GPU software*	Moderate
Granular science 2 participants			
103	Monte Carlo simulations, discrete element modeling of multi-body dynamics of macroscopic particles, computation of transport quantities, application of discrete dynamical systems methods	Element code	Large effect
104	Discrete element method for submicron aggregates	*Do not currently use GPU software*	Large effect
Biophysics 2 participants			
105	We work with Computational Neuroscience where models are biophysically grounded. The work involve simulations of differential equations (neurons) interconnected and demands a lot of CPU and RAM memory.	*Do not currently use GPU software*	Large effect
106	blood clotting and engineered tissue culturing simulations	*Do not currently use GPU software*	N/A -- DON'T KNOW for public-access GPU use
Transportation analysis 2 participants			
107	I did it on my own machines in the lab	*Do not currently use GPU software*	Large effect

108	Traffic Management and Modeling in NYC.	*Do not currently use GPU software*	N/A -- NO for public-access GPU use
Computational chemistry 2 participants			
109	We are using Molecular dynamics (MD) simulations to study the stability of nano bubbles	*Do not currently use GPU software*	N/A -- NO for public-access GPU use
110	Density Functional Theory-based optimizations of organic structures and energy computations	*Do not currently use GPU software*	N/A -- NO for public-access GPU use
Condensed matter physics 1 participant			
111	Band structure calculations of semiconductors and semiconductor alloys.	DFT, VASP	N/A -- NO for public-access GPU use
OTHER 10 participants **General** Specific			
112	**Data Science** Graph Analytics, data science, national security, cybersecurity	rapids.ai	Large
113	**Machine learning and deep learning research** Training algorithms and experimental performance studies of sign activation deep neural networks	Python, scikit-learn, Keras (Tensorflow), and Pytorch	Large
114	**mathematical neuroscience and computational biology** simulating ODE models of neurons and other biological oscillators	MATLAB	Large

115	**Profile the running time (speedup) of various algorithms on CPU vs. GPU** Same as *General*	Tensorflow, Pytorch, CUDA, python, C/C++	Large
116	**We do system research and are interested in improving the performance of data science software on GPGPUs. We use the platform to verify the effectiveness.** We don't have specific type. All possible types of computation may be used as workloads.	*Do not currently use GPU software*	Large
117	**I and my group develop efficient solution algorithms to tackle hard combinatorial optimization (CO) problems, which require the use of HPC. Solving such CO problems (e.g., with applications to airline/rail operations scheduling) could take weeks, months even years without the use of powerful computers** The CO problems I try to solve involve exponential number of decision variables (huge action and state spaces), which require the use of HPC.	*Do not currently use GPU software*	Large
118	**Big data: cleaning, processing and analyzing data that are too big to handle efficiently on my local computer** Same as *General*	R, Python	Moderate effect
119	**Functional Image Processing** Functional Near Infrared Spectroscopy data analysis to examine neuronal changes after a TBI/Concussions	*Do not currently use GPU software*	Moderate
120	**mathematical modeling in fluid dynamics** Running simulations of dynamical systems (ordinary differential equations, integro-differential equation and iterated maps) that describe interacting between flapping wings, and the pilot-wave dynamics of bouncing droplets. The models contain a number of parameters, and the goal is to perform parametric sweeps to assess the impact of changing individual parameters on the observed phenomenology.	*Do not currently use GPU software*	Moderate
121	**We apply global optimization algorithms to problems including image processing and data clustering**	*Do not currently use GPU software*	Moderate

	We derive analytical convergence rate bounds for global optimization algorithms. The HPC allows us to test conjectures about bounds, and also allows us to explore the practicality of variations including parallel versions of the algorithms.		

At the end of the Assessment, all participants were invited to provide optional comments. Participants who responded YES (as opposed to NO or DON'T KNOW) to whether they would use of public-access GPUs were, in addition, asked to "describe the role of GPUs in your research and/or teaching, and how public-access GPUs would materially support that research and/or teaching".

Table 4 lists all written responses.
(Note that because one YES responder did not submit a comment, there are 48 rather than 49 comments describing the role of GPUs.)

Table 4

Participants who WOULD USE public-access GPUs				
	School/College *Department*	If public-access GPUs were available... Effect on research	 Use for research	*Optional comments* Description of how public-access GPUs support that research and/or teaching.
1	**Newark College of Engineering** *Mechanical and Industrial Engineering*	Large	Heavy	We are part of a multi-institutional development team building a next generation model for climate forecast. This new model is designed to perform effectively and efficiently on massively parallel GPU based architectures. This model will harness more data than ever and will provide a new level of accuracy to predictions of droughts, heat waves, and rainfall extremes.

2	**Newark College of Engineering** *Mechanical and Industrial Engineering*	Large	Heavy	*Optional comment: Without access to public GPU's, my research program will effectively be ended.* Public-access GPU's are essential (necessary) in order for me to pursue and continue my research agenda. Without this resource, I will no longer be able to continue the research program that I have been building over 34+ years at NJIT. Over the course of my tenure at NJIT, I have worked with (and mentored) more than 90 undergraduate students on various research projects, most of which have involved computations using the research computing system. The majority of these students have been at the very top of their classes, a percentage completed MS /Theses under my supervision, and a subset have gone on to obtain PhD's at very prestigious US universities. Public research-computing has played a critical role here in preparing students for their careers.
3	**Newark College of Engineering** *Mechanical and Industrial Engineering*	Large	Heavy	The DFT computation is very expensive. Without GPU, it is impossible to finish many essential tasks. GPU nodes are expensive. Public-access GPUs will be very helpful.
4	**Newark College of Engineering** *Mechanical and Industrial Engineering*	Large	Heavy	*Optional comment: NJIT ARCS group is the best department at NJIT. The management and the team has been wonderful. Always quite prompt to respond to my questions and address my computing needs for research. I do cite NJIT ARCS in my publications in the acknowledgments section.* GPUs are the main computing tools for training and testing deep learning algorithms. I need to use those powerful GPUs to train my neural network algorithms and solve combinatorial optimization problems, which cannot be solved with the state-of-the-art computing power. NJIT ARCS access of those GPU resources would certainly positively impact of my research and help with the timely completion of student doctoral degree requirements.
5	**Newark College of Engineering** *Mechanical and Industrial Engineering*	Large	Heavy	We study the transport of particles, drops, and biological cells in fluids. Our code at present does not use GPUs, but we are planning to modify our code to take their advantage. This will allow us to simulate larger problems, including the problems we cannot solve using our current code.
6	**Newark College of Engineering** *Electrical and Computer Engineering*	Large	Heavy	Emerging communication systems, as 6G cellular networks and the Internet-of-Things (IoT) strive for extremely high data rates, while ensuring low complexity processing at the same time. In order to achieve these high data rates, standard approaches used in 4th and even 5th generation cellular systems do not work anymore. For example, at a projected data rate of 10Gbit/s, a current transmission cost of 6 nJ/bit and a typical battery capacity of 7 Wh for a cell phone battery, the resulting battery life is less than 6 minutes. To combat this complexity bottleneck, cooperative and opportunistic communication techniques and improved multiple access techniques need to be leveraged. One promising approach is the introduction of deep learning techniques for channel estimation, non-orthogonal massive multiple access and scheduling, in particular for IoT, and for

				encoding/decoding and modulation/demodulation. Public-access GPUs would be beneficial for speeding up training of these potentially quite high dimensional models, and thus directly have an impact on the performance of these systems.
7	**Newark College of Engineering** *Electrical and Computer Engineering*	Large	Heavy	*Optional comment: Just to mention that one of my PhD students has been stuck to deal with the memory limitation problem in the current compute infrastructure and I am reaching out to NVIDIA for their temporary help through a former PhD student.* I have developed the first graduate level course in 2011 entitled ECE641 Laboratory for High Performance Digital Signal Processing that uses FPGA and GPU devices to implement DSP algorithms. We have educated at least 250 MS and PhD students with this skill set over a decade at NJIT.
8	**Newark College of Engineering** *Biological and Pharmaceutical Engineering*	Large	Heavy	I think that GPUs will have a significant impact on my computational research. So far we use my own GPU. I sincerely hope that you receive this grant.
9	**Newark College of Engineering** *Biomedical Engineering*	Large	Heavy	Public-access GPUs will have a profound impact in my research as I am currently using GPU clusters to train neural networks to solve nonlinear optimal control problems of human movement. These GPUs will allow me to run large models with more muscles and more complicated reflex control algorithms and many models in parallel. It will speed up the simulations and increase my research productivity considerably.
10	**Newark College of Engineering** *Engineering Technology*	Large	Heavy	GPU cards have a central role in my research in that they allow parallel processing within the codes that I use, therefore significantly less computation time. Availability of GPUs will support research by allowing multiple complex model to run at once in a simulation, and to also run several simulations at once.
11	**Newark College of Engineering** *Chemical and materials engineering*	Large	Heavy	GPUs are used every day by my students for their thesis work to investigate the material synthesis pathways and chemical reaction mechanisms. In teaching, GPUs will be gradually incorporated into graduate courses, especially in ChE750, one chapter from this course focuses on the theoretical calculations.
12	**Ying Wu College of**	Large	Heavy	GPUs play an important role in teaching machine and deep learning at NJIT. The deep learning course

	Computing *Computer Science*			CS 677 has assignments and projects that require GPUs to run deep learning libraries such as Keras and Pytorch. These assignments and projects are necessary for students to learn the material since both machine and deep learning are experimental sciences (see Langley in Machine Learning 1988 and Drummond in AAAI 2006). Aside from the deep learning course GPUs are also used in CS 698 Medical AI and CS 732 Advanced Machine Learning. Both courses expose students to machine and deep learning models in foundations and applications. Undergraduate and graduate students working on deep learning thesis and projects also use GPUs for their work. These students have obtained industrial positions in these areas thanks to their hands-on experience gained by project work on the GPU machines.
13	**Ying Wu College of Computing** *Computer Science*	Large	Heavy	Please see my research papers on my website.
14	**Ying Wu College of Computing** *Computer Science*	Large	Heavy	My research heavily uses high-performance GPUs. These public-access GPUs would allow my group to take advantage of recent, new deep learning models and use the models to develop new deep learning tools for data analysis.
15	**Ying Wu College of Computing** *Computer Science*	Large	Heavy	DNNs have been extensively used in today's computer science research. Running DNNs on CPU is significantly slower, if not feasible, compared to running on GPUs. Having a public-access cluster with many GPUs would enable the use of more powerful DNN-based models, that is currently not possible using our own small machines. Of course, it would increase the breadth and depth of our CS research.
16	**Ying Wu College of Computing** *Computer Science*	Large	Heavy	*Optional comment: My students H and S graduated, who really did the bulk of the GPU work.* *But my student U is learning it and she will become a heavy user.* *H always told me he could not run programs at all, or he had to break them into small batches and running them took many hours or even days. If one experiment takes two days it's not possible to do many experiments. And we are competing with research groups that have much better computational resources.* *Especially with tuning hyperparameters such runtimes are really difficult for getting results.* Please refer to papers of Hao Liu. You will need a ton of reference for a good NSF proposal. Here's a starting point. https://pubmed.ncbi.nlm.nih.gov/33098987/ https://pubmed.ncbi.nlm.nih.gov/32308910/ https://pubmed.ncbi.nlm.nih.gov/32308894/ https://pubmed.ncbi.nlm.nih.gov/30815117/ Essentially we want to continue this kind of research an ORDER of magnitude faster and with much larger data sets. Ideally we would like to run experiments with training data on the order of

				61,250,000,000 data points. (This is 350,00 to the square divided by 2. A network with 350,000 nodes and we want all pairs of them) I know it won't happen in my lifetime, but the closer we can get, the better off we are. Feel free to check with me -- Jim Geller
17	**Ying Wu College of Computing** *Informatics*	Large	Heavy	*Optional comment:* The GPUs are not enough. It would be better to TPU for even larger models, like BERT-large and even some distilled Bert models. My main research is to build AI (especially deep learning) solutions to improve software reliability, quality, and intelligence. We heavily rely on GPUs/TPUs to run our large models. Also, deep learning is a popular and important topic in teaching. Currently, it is limited to run any large models on our current infrastructure. The public access GPUs would enable more people to get involved and run larger models.
18	**College of Science and Liberal Arts** *Mathematical Sciences*	Large	Heavy	GPU simulations using our limited local resources turned out to be so much more efficient than similar CPU based simulations that we were able to bring the corresponding research project to a completely new level. Availability of additional GPUs would increase our research productivity significantly.
19	**College of Science and Liberal Arts** *Physics*	Large	Heavy	Microsecond-long simulations performed in our lab can take up to two months to be executed using CPU-nodes with a high density of cores and up to two weeks using GPU-based nodes. Thus, access to the GPU nodes of the planned acquisition will provide a significant improvement of our group's workflow wherein students will need to wait less time to obtain results from simulations. This is critical during the development of new projects when several test simulations need to be performed. Notice that the open source suite GROMACS, which is used in our lab, is optimized to perform simulations on GPU wherein the whole time-step of the molecular dynamics integrator can run on the GPU avoiding significant CPU and PCIe bottlenecks
20	**College of Science and Liberal Arts** *Mathematical Sciences*	Large	Moderate	The GPUs will speed up my research codes.
21	**College of Science and Liberal Arts** *Mathematical*	Large	Moderate	*Optional comment: I very much hope this proposal is successful. NJIT needs up-to-date HPC facilities to carry out its core mission in research and teaching.* It would enable faster turn around of results (i.e., reduced run times) and (as indicated above) enable solution of larger versions of current problems as well as problems that are not now feasible.

	Sciences			
22	**College of Science and Liberal Arts** *Mathematical Sciences*	Large	Moderate	My research group would use public-access GPUs to efficiently perform simulations of biological models with many different parameter (hundreds of thousands). Then we would use machine learning algorithms (also running on GPUs) to analyze the parameter combinations that produce various types of model output, providing insight into the biological mechanisms underlying the behaviors corresponding to those outputs.
23	**College of Science and Liberal Arts** *Physics*	Large	Moderate	*Optional comment: As a recipient of a faculty seed grant I have investigated the feasability of using GPU for my research. Although this approach has been validated, lack of access to dedicated GPU resources prevented this prototype to be be transformed in a functional research resource.* Some components of the modeling and data analysis frameworks employed in my research rely on embarrassing parallel algorithms whose computational speed mainly depend on the availability of parallel threads. Thus, access to GPU resources would greatly decrease computational the speed and increase research productivity.
24	**College of Science and Liberal Arts** *Physics*	Large	Moderate	The brain contains billions of neurons connected with trillions of synapses. Such a number is way beyond the current set of computational resources we have in hand so researchers usually downsize simulations to a few millimeters of the brain. The more neurons and synapses the more computational power in terms of RAM memory and CPUs are necessary. However, the use of CPUs creates constraints due to their limited capabilities. The inclusion of GPUs in the research of computational neuroscience opens new venues that could be explored given its enormous computational capabilities when compared to CPUs. More specific and detailed brain areas could be explored giving rise to a new set of questions. I can only encourage the inclusion of public-access GPUs.
25	**College of Science and Liberal Arts** *Chemistry and Environmental Science*	Large	Moderate	During the last two years, my group has been heavily using the HPC for discrete dipole approximation (DDA) calculations. The DDA simulates scattering of light by arbitrarily shaped particles, such as soot aggregates. The DDA models the particle as a collection of dipoles on a lattice, and is computationally much more expensive than the methods applied to regularly shaped particles. On a single computer the calculation for a soot particle of a realistic size at one specific wavelength can take several days. It has been shown that porting DDA for GPUs can result in a speed-up by a factor of 5 in double precision and a factor of 15 in single precision for cloud ice particles (https://doi.org/10.1016/j.jocs.2011.05.011). In my projects, porting the calculations to GPU and spreading the work load over multiple GPUs will allow calculating the scattering properties of larger, more realistic soot aggregates and also expand the range of effects and morphologies that can be investigated.
26	**Newark College of Engineering** *Biological and*	Large	Moderate	I have never been using GPUs, but I understand that for some of my work they would work well.

	Pharmaceutical Engineering			
27	**Newark College of Engineering** *Biomedical Engineering*	Large	Moderate	We will use computational drug discovery to elucidate targets, design inhibitors/ binders and self-assembling drugs.
28	**Newark College of Engineering** *Electrical and Computer Engineering*	Large	Moderate	In our recent publications, in order to verify the new approaches we derive, people always ask for validation on large data sets. This is where we need GPUs.
29	**Ying Wu College of Computing** *Computer Science*	Large	Moderate	We do system research and are interested in improving the performance of data science software on GPGPUs. We use the platform to verify the effectiveness. I teach programming courses. With public-access GPUs, I can include some GPGPU programming components in my courses, which will be very helpful since a large number of students may work in data science areas after graduation.
30	**College of Science and Liberal Arts** *Mathematical Sciences*	Moderate	Moderate	If GPUs are as easy to program as shared memory machines and OpenMP, I could see using it for computational PDE, perhaps tailoring fast algorithms to their architecture. My research is all about algorithm development, so if there are interesting issues involved in optimizing fast algorithms for GPUs I could see getting more involved in that. Good luck!
31	**College of Science and Liberal Arts** *Mathematical Sciences*	Moderate	Moderate	More computing access helps the completion of extensive simulations that are required for strong papers.
32	**College of Science and Liberal Arts** *Mathematical*	Moderate	Moderate	Access to GPUs would allow me to run computational codes far more efficiently, and would allow me to simulate processes that I currently cannot. For example, I could write my code to take advantage of parallel architecture, and thus simulate the dynamics of large numbers of interacting particles. Understanding the emergent properties of large collections of particles is an active area of research, and one that receives substantial funding from grant agencies.

	Sciences			
33	**College of Science and Liberal Arts** *Mathematical Sciences*	Moderate	Moderate	*Optional comment: Although public-access GPUs would have a moderate to light impact on my research projects and teaching, but I know from numerous colleagues that they would have significantly more positive impact on their research and teaching.* GPUs would certainly improve the simulation part of our research on analyzing a certain class of integro-PDEs. On the other hand, GPUs could be used to enhance some of the applied aspects of my teaching.
34	**College of Science and Liberal Arts** *Mathematical Sciences*	Moderate	Moderate	I am currently entering the area of machine learning for marine sediment classification. Convolutional neural networks are excellent tools for this task but are very computational onerous. GPUs would facilitate the design and implementation of such networks.
35	**College of Science and Liberal Arts** *Physics*	Moderate	Moderate	I would be able to use more efficient test particle codes to investigate particle dynamics in the given fields, and thus, to improve our understanding of the process in space plasmas, like solar wind and planetary magnetospheres.
36	**Ying Wu College of Computing** *Computer Science*	Moderate	Moderate	The current state of the cluster has resulted in research computations for program correctness enhancement taking months. Adding additional resources would have significantly sped up the research activities.
37	**Ying Wu College of Computing** *Computer Science*	Moderate	Moderate	My research has been focused on deep learning which heavily depends on the usage of GPU. The number of GPUs is not sufficient and it is not possible to provide all students with GPU for teaching purposes.
38	**Ying Wu College of Computing** *Computer Science*	Moderate	Moderate	We develop global optimization algorithms that are natural to implement as parallel algorithms. Our applications in image processing and data clustering are large-scale problems. These two factors make GPUs ideal for numerical experiments. In our Monte Carlo simulation work, parallel algorithms and long run times are also typical. Both global optimization and Monte Carlo simulation are part of my courses.
39	**Ying Wu College of Computing**	Moderate	Moderate	Public-access GPUs would enhance research possibilities in Deep Reinforcement Learning and greatly contribute to helping students master the more advanced skills in various artificial intelligence specialties that are in demand today.

	Computer Science			
40	**Ying Wu College of Computing** *Informatics*	Moderate	Moderate	I work in crisis informatics, political science, and computational social science as they pertain to online behavior. This work requires characterizing complex online behaviors and data in tractable representations. Of particular value are abilities to characterize images using computer vision techniques and large networks of interactions using graphical neural models. These approaches enable characterization of information type and priority during crises and disasters (e.g., whether an image contains descriptions of damaged infrastructure) and characterization of online coordination during misinformation campaigns.
41	**Newark College of Engineering** *Electrical and Computer Engineering*	Moderate	Moderate	With the GPU upgrade, semiconductor device simulation can be extended from 2D to 3D and benefit both research and education.
42	**Newark College of Engineering** *Engineering Technology*	Moderate	Moderate	Functional Near Infrared Spectroscopy analysis is currently limited by the availability of computing resources to test various methods of analysis. Unfortunately, analysis of a single participant data file is not sufficient to yield meaningful data, thus data from at least 40+ participants must be utilized to test various methods. It may take up to 24 hours for each file to be analyzed, data from which can later be used for group level analysis. Access to additional GPUs will enable this work to occur at a faster rate, especially when involving new students. Access to public-access GPUs will additionally enable me to formulate new classes that can utilize this computation power and teach different data analysis methods.
43	**Newark College of Engineering** *Civil and Environmental Engineering*	Moderate	Moderate	The public-access GPUs will significantly enhance the simulation speed of our simulations. In addition, some large-scale simulations will be feasible with GPUs.
44	**Martin Tuchman School of Management**	Moderate	Moderate	Accounting and finance research are rapidly adopting new methods in recent years that require high performance computing. Those methods include but are not limited to machine learning, big data, textual analysis and etc. High performance computer have enabled a few of my projects that analyze social networks of directors and its implications in several areas in accounting and finance. It is also important for my on-going and future projects in machine learning applications in accounting and finance. As far as teaching, the Martin Tuchman School of Management is quickly adapting to make sure students have enough exposure to new technology and are ahead in the future work force. We launched the B.S. in Fintech, added two accounting data analytics courses (one under and one graduate) and many more. I will be teaching the two accounting data analytics courses, which surely

				will benefit from this initiative.
45	**College of Science and Liberal Arts** *Mathematical Sciences*	Moderate	Light	GPUs are critical in the fast computation of solutions to classification problems that I am interested in. Public-access GPUs would simplify the process and make the tools produced more accessible to students.
46	**College of Science and Liberal Arts** *Federated Biological Sciences*	Moderate	Light	One of my research directions is the use of automated image identification for ecology and biodiversity research. In particular we are developing 'specialized' ID systems to solve constrained but challenging problems, such as separating closely related bee species based on images of their wings. My collaborators (in CS) and I take trained 'deep' CNNs and use transfer learning to retarget them to the constrained problem. But despite not having to train a network from scratch, the best results occur when we are able to adjust feature extraction parameters at all levels (rather than just training new classification layers). This requires substantial GPU resources to be feasible, and these are in limiting supply, meaning that progress has been slow. A publicly-accessible GPU cluster resource would dramatically increase the scope for this aspect of my research. Going forward, I hope to promote more broadly the use of technology to address environmental challenges, and this proposal would boost that effort.
47	**Ying Wu College of Computing** *Computer Science*	Light	Light	*Optional comment: A couple years ago, one of my students was running many Monte Carlo experiments in Matlab on kong, but kong is very outdated and not very reliable. He designed the experiments to run on multiple cores, but not infrequently, one of the cores would fail and he didn't get any results from it. Also, the version of Matlab at NJIT that he had to use to run on multiple cores was several years out of date. These issues held up our progress on the research.* NJIT needs to have modern, up-to-date computing systems to support research. Unfortunately, kong is really outdated. I don't always need to use high-powered computing to do my research, but there are times when it would be very helpful.
48	**Ying Wu College of Computing** *Computer Science*	Light	Light	I teach deep learning, so the availability and ease of use of these systems is significant. The availability of these systems is expected to accelerate the experimentation cycle for some of my PhD students who need to train neural network models.
Participants who DO NOT KNOW if they would use public-access GPUs				
	School/College			

	Department	*Optional comments*
49	**College of Science and Liberal Arts** *Mathematical Sciences*	*I am fully support of this proposal effort and am happy to participate in any capacity - shahriar@njit.edu*
50	**College of Science and Liberal Arts** *Mathematical Sciences*	*The methods described are being implemented by PhD students in my research group.*
51	**College of Science and Liberal Arts** *Chemistry and Environmental Science*	*The computer services team at NJIT has been just excellent in keeping the NJIT computer systems operational, and incorporating complex computer programs and maintaining their performance.*
52	**Newark College of Engineering** *Civil and Environmental Engineering*	*I think Fluent can be used with GPU nodes, but I did not try it*