Results of the Digital Research Alliance of Canada 2025 Resource Allocation Competition

List of awarded project for RAC 2025

Table of content

1. Summary

The Alliance Federation delivers Canada's national advanced research computing (ARC) platform in partnership with regional digital research infrastructure (DRI) organizations (Compute Ontario, Calcul Québec, ACENET, the BC DRI Group and the Prairie DRI Group) and institutions across Canada. Providing researchers with access to the infrastructure and expertise they need to accomplish globally competitive, data-driven and transformative research, this national ARC platform serves the needs of more than 20,000 researchers, including over 5,843 faculty based at Canadian institutions as of January 1, 2025.

For the 2025-2026 allocation period, the total available capacity of the national ARC platform for the Resource Allocation Competition is 700,192 CPUs, 106,304 vCPUs (virtual CPUs), 19,475 Reference GPU units (RGU-years) and 354.6 PB of storage on the following systems:

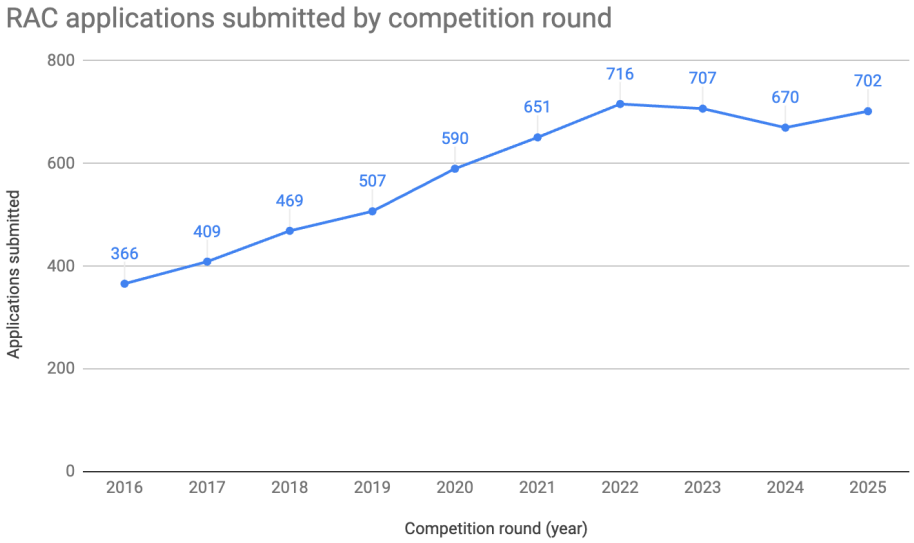
Host site	HPC system	Cloud system
Simon Fraser University	Fir	Cedar
University of Waterloo	Nibi	Graham Cloud
Calcul Québec	Rorqual Narval	Béluga Cloud
University of Toronto	Trillium HPSS	
University of Victoria		Arbutus

The 2025 RAC received 702 projects, 32 more applications than in the previous year, and it awarded, on average, 84% of the total compute requested, 85% of the total storage requested, 31% of the total RGU-years requested and 88% of the total vCPUs requested.

In 2024, the minimum amounts of resources required to be eligible to apply for RAC were increased to 200 core-years, 25 RGU-years, 41 TB of project storage or 101 TB of nearline storage, primarily to reduce the number of RAC applications received. As a result, some projects were ineligible to apply for RAC because they could get those resources, or a portion of them, through the Rapid Access Service (RAS). This resulted in a decrease of 37 applications submitted in 2024 compared to 2023. However, the number of applications increased again in 2025, which suggests that the demand for significant computational resources through RAC is continuing to grow.

Table 1: Applications Submitted to the Resource Allocation Competition

Year	Applications submitted	Year-over- year increase
2025	702	5%
2024	670	-5%
2023	707	-1%
2022	716	10%
2021	651	10%
2020	590	16%
2019	507	8%
2018	469	15%
2017	409	12%
2016	366	5%



If you have questions about the terminology used in this page, please consult the Technical Glossary. If you have any questions about the overall report, contact allocations@tech.alliancecan.ca

2. Computational Resources

2.1. Minimum size of RAC requests and opportunistic compute access

A minimum of compute resources (currently set at 200 core-years for CPU and 25 RGU-years for GPUs) is required to be eligible to submit a RAC application. These minimum values are set

in part to control the number of applications requiring peer-review. A minimum RAC award will also ensure higher job priorities than for non-RAC awardees.

All researchers and their sponsored users with an active account can automatically make opportunistic use of CPU and GPU resources on any system. There is no guarantee on how much CPU or GPU can be consumed by non-RAC holders, as their use of the systems is purely opportunistic.

Historical utilization data shows that many groups are able to reach (or even exceed) the RAC minimums specified above. Non-RAC users who want to maximize their compute usage need to consider strategies that ensure:

- they regularly have jobs in the queue;
- are able to tolerate longer wait times for jobs to start; and,
- submit jobs with “optimal” characteristics.

For example, opportunistic jobs with short time limits that request a few cores on a general purpose (GP) system will generally run much sooner than those requesting hundreds dozens of cores. Please read this useful documentation about allocation scheduling priorities and job scheduling policies or contact support@tech.alliancecan.ca for advice on how to maximize usage for non-RAC awardees.

2.2. CPU allocations

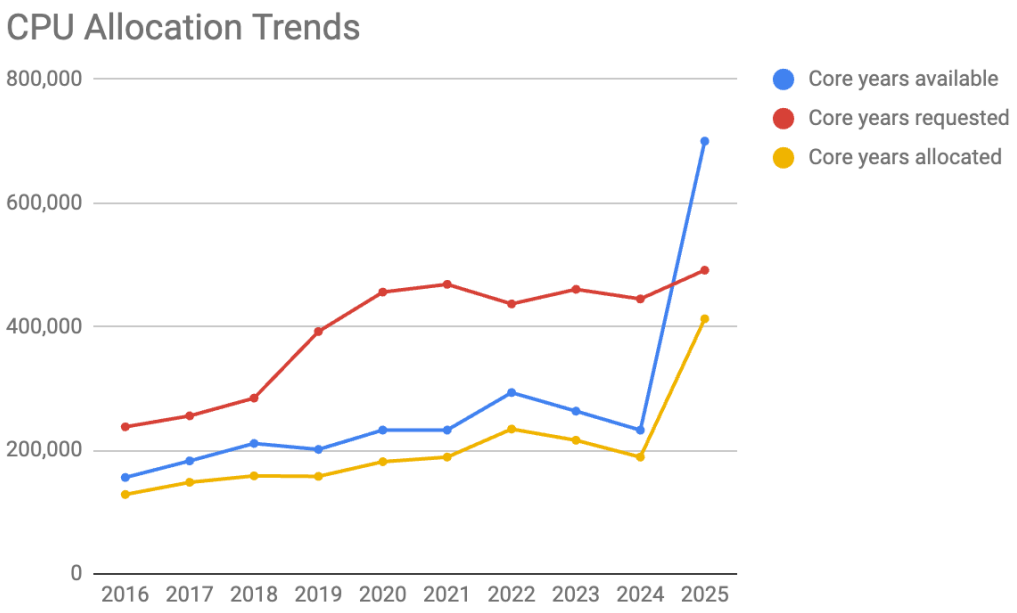
RAC 2025 was able to meet 84% of all of the CPU resources requested, 41% more than last year. Fir, Nibi, Narval, Rorqual and Trillium provide approximately 700,192 cores, 59% of which was allocated through the competition.

Table 2: 2025 CPU allocations per cluster

Cluster	CPU capacity (core-years)*	CPU requested (core-years)	CPU allocated (core-years)*	% of CPU capacity allocated
Rorqual	131,712	117,759	98,297	75%
Fir	166,272	112,224	91,229	55%
Nibi	136,320	56,312	46,404	34%
Narval	30,880	39,279	34,819	113%
Trillium	235,008	165,779	145,511	62%
Total	**700,192**	**491,353**	**416,260**	**59%**

Table 3: Historical CPU ask vs. allocation

Year	CPU capacity (core-years)	CPU requested (core-years)	CPU allocated (core-years)	% of the demand allocated
2025	700,192	491,353	416,260	84%
2024	232,560	444,810	189,102	43%
2023	263,326	460,346	216,164	47%
2022	293,312	436,780	234,275	54%
2021	232,704	468,498	188,925	40%
2020	232,704	455,892	181,502	40%
2019	201,320	390,352	157,262	40%
2018	211,020	284,347	158,612	56%
2017	182,760	255,638	148,100	58%
2016	155,952	237,862	128,463	54%



2.2.1. Scaling CPU requests

While for RAC 2025 the capacity of CPU resources was greater than the demand, we still applied a function to scale CPU, although significantly less severely than in previous years. This function was established so that only applications with a score equal to or greater than 2.1 (out of 5) received an allocation. Applicants who did not receive a CPU allocation can still make opportunistic use of system resources via the Rapid Access Service. The average score of all of the applications submitted to the RAC 2025 was 3.9.

CPU requests are scaled based on the overall score of the application *and* the size of the request. Details and examples of the scaling function are available here. For further questions, contact allocations@tech.alliancecan.ca

2.3. GPU allocations

The Reference GPU Unit (RGU) was introduced in 2024 to request, allocate and measure the amount of GPU resources that are used. It represents the "cost" of utilizing a particular GPU model, whose RGU value varies based on performance. GPU allocations are from now on set in, and usage charged on, RGU-years and not in GPU-years. For more information about RGU, visit this page. Since this is a new unit, we are not able to provide historical numbers for GPU capacity, demand and supply, beyond 2024and that is why **Table 4** uses *RGU-years* and **Table 5** uses *GPU-years*.

The demand for GPU resources continues to be more competitive than for CPU resources. As **Table 5** shows, while the GPU capacity in RGU-years increased by 3x compared to the previous year, the request for GPU resources nearly doubled in the same period. In 2025, we allocated 31% of the total RGU-years requested, which represents an increase of 10% compared to the previous year.

Table 4: 2025 GPU allocation demand per cluster

GPU resource	GPU requested (RGU-years)	GPU allocated (RGU-years)	GPU Capacity (RGU-years)	% of GPU capacity allocated
Rorqual	9,582	3,200	3,953	81%
Fir	17,427	5,342	7,808	68%
Nibi	7,245	2,573	3,514	73%
Narval	5,842	1,281	1,272	101%
Trillium	3,766	1,277	2,928	44%
Total	**43,862**	**13,673**	**19,475**	**70%**

Table 5: Historical GPU allocation

Year	GPU capacity (RGU-years)	GPU requested (RGU-years)	GPU allocated (RGU-years)	% of the GPU demand allocated
2025	19,475	43,862	13,673	31%
2024	6,164	23,990	5,022	21%

2.3.1. Scaling GPU requests

GPU allocations are determined by the following factors:

- the overall score of the RAC application,
- the technical justification provided,
- evidence of previous GPU utilization,
- the research area of application for which GPUs are requested (e.g., Artificial Intelligence, machine learning, etc.),
- the size of the research group.

Keep in mind the following:

- GPU allocations are constrained, among other things, by the type of GPU requested and available in each system.
- The demand for GPUs for AI applications has increased considerably.
- In general, RAC applicants find it difficult to estimate their GPU needs, which in most cases are over-requested (and underutilized). We strongly encourage future RAC applicants to do two things *before* applying: start using the GPUs in order to get a better understanding of your needs *and* consult with our technical staff *before* submitting a RAC application. Our staff can provide advice on how to benchmark your codes and calculate your GPU needs as accurately as possible.

2.4. Storage allocations

Storage integrated with Arbutus, Béluga, Cedar, Graham, Narval and Niagara provided approximately 210.7 PB of storage capacity for 2024. This meant that, across all types of storage, 76% of the total storage capacity was allocated.

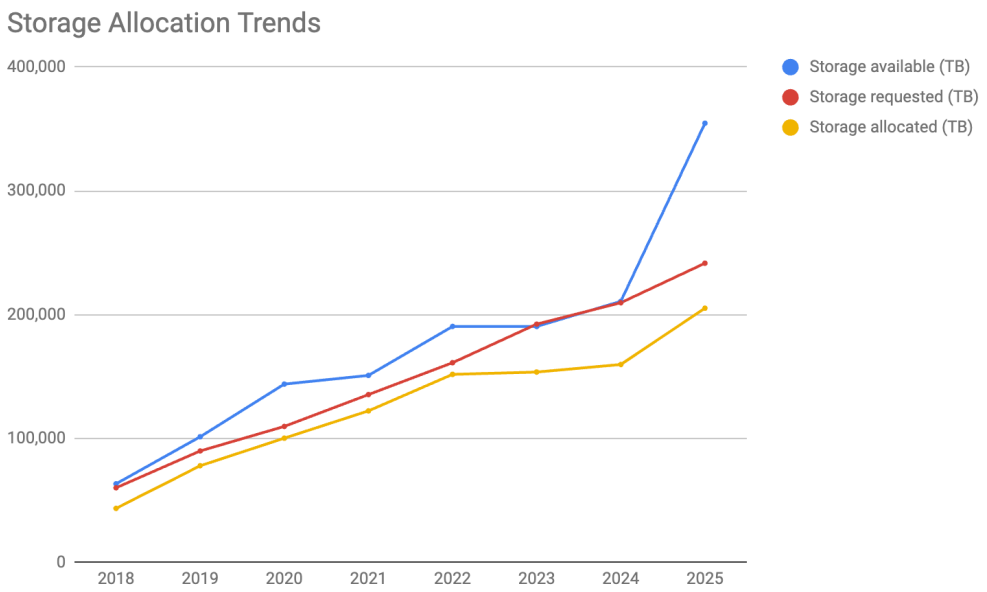


Table 6: Historical storage allocation

Year	Storage capacity (TB)	Storage requested (TB)	Storage allocated (TB)	% of storage demand allocated
2025	354,640	242,234	206,661	85%
2024	210,764	209,642	159,746	76%
2023	190,479	192,363	153,639	72%
2022	190,479	161,186	151,775	87%
2021	150,915	135,427	122,272	91%
2020	143,914	109,718	100,222	90%
2019	101,344	89,898	77,923	94%
2018	63,340	60,126	43,508	80%

Table 7: 2025 Storage demand by storage type

Category	Type	Storage capacity (TB)	Storage requested (TB)	Storage allocated (TB)	% of the storage capacity allocated
HPC	Project	168,684	103,467	82,903	49%
HPC	Nearline	136,300	102,333	92,640	68%
HPC	dCache	18,000	16,818	16,818	93%
Cloud	Volumes and snapshot	9,128	2,772	2,739	30%
Cloud	Object	8,192	9,870	5,067	62%
Cloud	Shared	14,336	6,974	6,494	45%
	Total	**354,640**	**241,559**	**205,261**	**58%**

2.5. Cloud allocations

The Arbutus cluster at the University of Victoria has 83,200 allocatable vCPUs (virtual CPUs). These are available via RAC and RAS and are also utilized for internal services such as software development and hosting. Relatively small cloud offerings are also implemented on Cedar, Graham and Béluga. For RAC 2025, the request for compute vCPUs decreased by an average of 21%. This apparent drop in requests is primarily the result of added contributed capacity on Arbutus for a number of major cloud users. In particular, the CANFAR/SKA astronomy project and the BELLE-II high-energy physics project have substantially more new contributed hardware, including a proportional amount of storage. Those contributed resources appear in the “available” column, but since they are no longer allocated by RAC, they do not appear in the “ask” or “allocated” columns.

Between Arbutus and the additional nodes on Cedar, Graham, and Béluga, this year's RAC was able to allocate 88% of the total virtual CPUs requested, 17% more than in the previous year.

Table 8: Historical vCPU demand

Year	vCPU capacity (vCPU-years)	vCPU requested (vCPU-years)	vCPU allocated (vCPU-years)	% of vCPU demand allocated
2025	106,304	36,550	32,103	88%
2024	56,405	46,894	33,213	71%
2023	56,405	35,618	27,313	77%
2022	62,549	34,536	27,444	79%
2021	62,549	30,323	24,443	81%
2020	50,501	18,330	18,229	99%
2019	29,147	19,479	18,511	95%
2018	24,854	12,480	11,829	95%

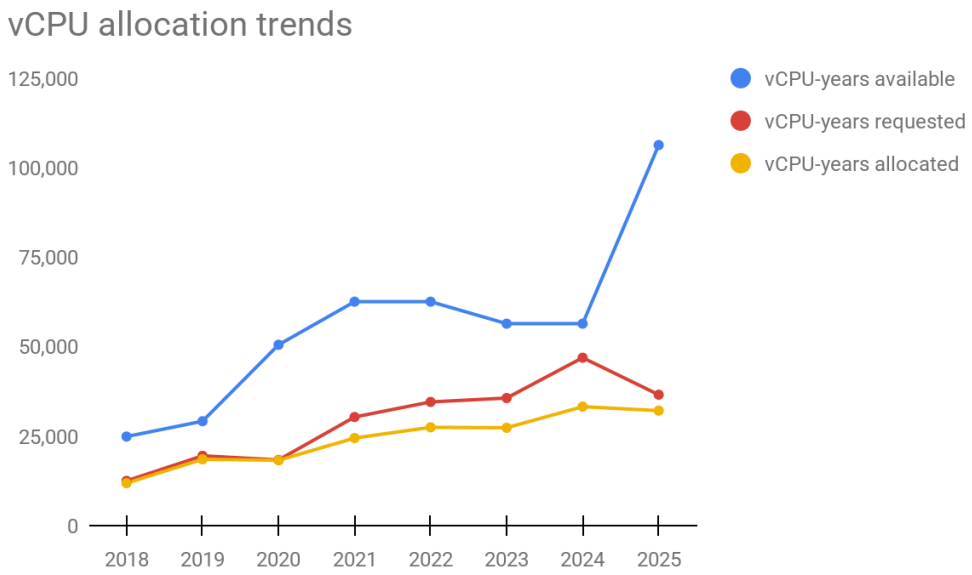
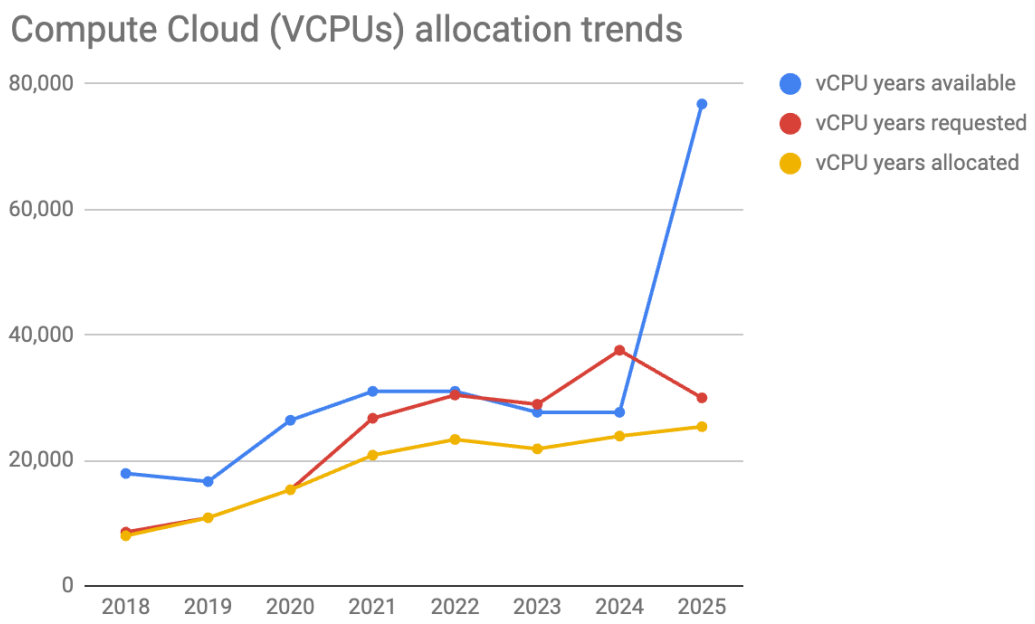


Table 9: Historical vCPU allocation broken down by compute and persistent instances

	Compute instances			Persistent instances		
Year	vCPU years capacity	vCPU years requested	vCPU years allocated	vCPU years capacity	vCPU years requested	vCPU years allocated
2025	76,800	29,968	25,377	29,504	6,582	6,726
2024	27,667	37,557	23,876	28,738	8,635	8,635
2023	27,667	28,938	21,834	28,738	6,040	5,479
2022	31,009	30,414	23,334	31,540	4,180	4,110
2021	31,009	26,714	20,834	31,540	3,609	3,609
2020	26,401	15,319	15,319	24,100	3,011	2,910
2019	16,629	10,858	10,858	12,518	8,621	7,653
2018	17,920	8,566	7,993	6,934	3,914	3,836



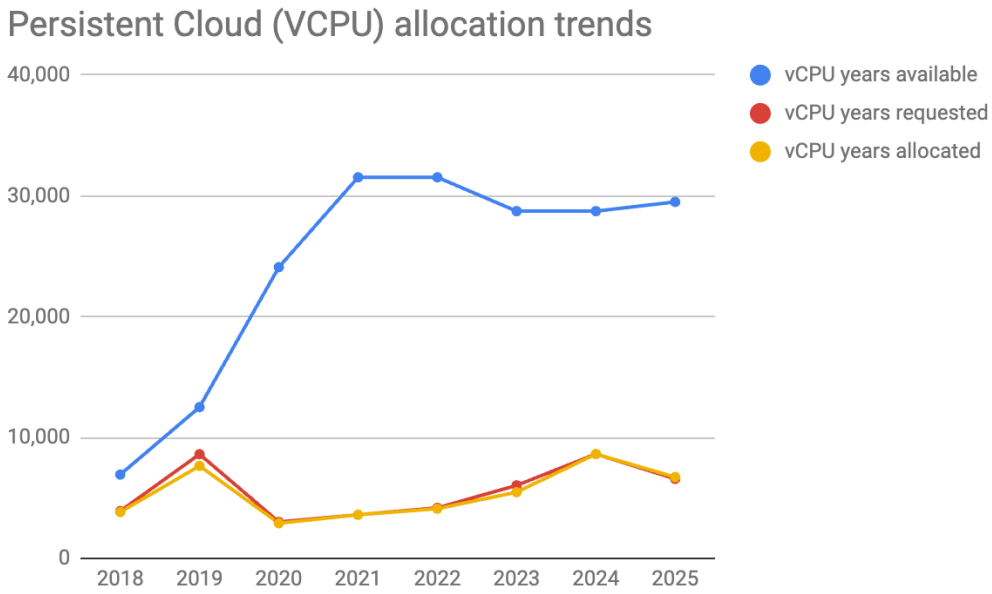
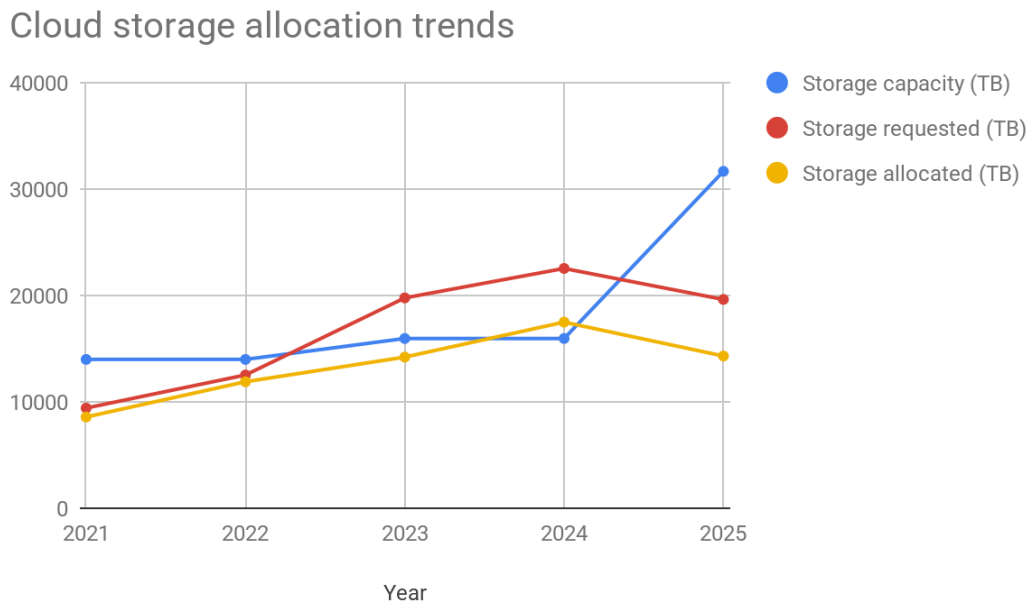


Table 10: Historical allocation of storage resources available in the cloud*

Year	Capacity (TB)	Requested (TB)	Allocated (TB)
2025	31,656	19,615.7	14,300
2024	15,947	22,529	17,484
2023	15,947	19,760.5	14,198
2022	13,980	12,517	11,883
2021	13,980	9,409.10	8,568

**The numbers in this table represent the sum of the following resources: Object, Shared and Volumes and Snapshot storage.*



3. Assessment process

The RAC involves two review processes each year:

- a scientific review, which is a peer-review process involving more than 100 discipline-specific experts from Canadian academic institutions. These volunteers assess and rate the merits of the computational research projects submitted. The scientific review results in a single score that provides a critical and objective measure to guide allocation decisions; *and*
- a technical review that is undertaken by staff who are responsible for verifying the accuracy of the computational resources needed for each project based on the technical requirements outlined in the application and for making recommendations about the national system to which the resources should be allocated to meet the project's needs.

The overall process is overseen by the Resource Access Program Administrative Committee, which includes representatives from each region and national system host sites.

Note that while new applications receive both scientific and technical reviews, applications submitted via the Fast Track process and Research Platforms and Portals (RPP) with a multi-year award receive only a technical review.

3.1. Guiding principles

RAC is guided by the following principles:

- all applications are given fair consideration through both a scientific and technical review process;
- resources are awarded based on the merits of the computational research *project* presented, rather than the merits of the overall research *program*;
- there is no direct correlation between the amount of computational resources needed and the quality (excellence) of the research outcomes of a project - important research can be done with a small amount of computational resources; and
- the challenges arising from the shortage of resources and other constraints within the system are shared among all applicants.

3.2. Technical review

The technical review is conducted by technical experts who:

- ensure the appropriate system is requested by the PI;
- ensure that the required software is available;
- evaluate application efficiency and scalability;
- identify groups that may need help with application and workflow optimization;
- identify discrepancies between the online request and the complete description of the project;
- identify special software requirements; and,
- provide a technical opinion on the reasonableness of the request.

Technical reviewers are required to sign a Non-Disclosure Agreement prior to accessing any RAC application.

3.3. Science review

New applications submitted to the RAC are peer-reviewed and scored. Scientific reviewers are required to sign a Non-Disclosure Agreement and accept the Conflict of Interest Policy prior to accessing any RAC application.

The final RAC score is based on the following:

- the scientific excellence of the specific research *project* for which computational resources are being requested;
- the scientific and technical feasibility of the proposed research project;
- the appropriateness of the resources requested to achieve the project's objectives; *and,*
- the likelihood that the resources requested will be efficiently used.

Applications are reviewed in one of the committees below:

- Astronomy, Astrophysics and Cosmology
- Bioinformatics

- Chemistry, Biochemistry and Biophysics
- Computer Sciences and Mathematics
- Engineering
- Environmental and Earth Sciences
- Humanities and Social Sciences
- Nano, Materials and Condensed Matter
- Neurosciences, Medical Imaging and Medical Physics
- Subatomic Physics, Nuclear Physics and Space Physics

4. Monetary value of the 2025 allocations

These values represent an average across the national ARC platform's facilities and include total capital and operational costs incurred to deliver the resources and associated services. These are not commercial or market values. For the 2025 competition, the value of the resources allocated was calculated using the following rates:

Table 9: Financial value of RAC awards

Resource	Financial value
1 core-year	$109.67
1 RGU-year (for GPU and VGPU)	$1,167.07
1 TB of project storage	$60.36
1 TB of nearline storage	$27.04
1 vCPU-year	$39.89
1 TB of cloud storage (Ceph)	$51.82
1 TB of object storage	$51.82
1 TB of shared file system storage	$51.82