



Guidelines for requesting cloud computing resources through CloudBank

The Principal Investigator (PI) of a proposal to the National Science Foundation (NSF) should not request the budget amount for cloud computing resources through CloudBank in the PI's main budget request to the NSF. Instead, this amount should be specified in a supplementary document. Most cloud providers including Amazon Web Services (AWS) provide an online resource [pricing calculator](#) that can be used to develop yearly budget requests for the supplementary document. In addition, the PI may explain how the data for a project with cloud computing resources would be managed in the Data Management Plan. Below we illustrate the Data Management Plan and the Supplementary Document through examples.

Example of a Data Management Plan using AWS:

Note that the PI may include additional details in the Data Management Plan for handling other aspects of the project that are within the scope of the overall project besides the following related to the cloud computing component:

The bulk of the data in this project will be collected by individuals and organizations. This data will be made available to the public by working with relevant stakeholders and data owners once the data products become accessible and validated. One option is to implement the AWS data lifetime management policy to the data product, which comprises a set of rules with predefined actions that the block storage (Amazon S3) will perform on objects during its 10-year lifetime. For example, if the data are unused for 3-4 months, these policies will automatically move the data to a storage archive (Amazon Glacier). Similarly, if the data are unused for 10 years, these policies can take expiration actions. AWS has allocated 1 PB of Amazon Glacier (costs \$122k at \$12.2k/year to retain data in archive) for a period of 10 years. These configurations will also follow institutions' data retention policies. In most cases, the data are stored in standard data formats such as csv and HDF5.

Example of a Supplement document:

The following is a sample supplementary document to be submitted with the application for CloudBank. It is only for illustrative purposes.

Technical Plan: As part of the three-year grant with CloudAccess via CloudBank, we present here an estimate of AWS-based cloud resources usage using the [AWS pricing calculator](#), [Amazon EC2 Spot Pricing](#) and [Amazon Braket pricing for quantum resources](#) estimated for the course of the project and associated budget. Our research team (2 faculty, 3 postdocs, 4 graduate students, 3 undergraduates) will work with organizations to collect datasets for machine learning, use remote desktops, run supervised machine-learning models using containers orchestrated with Amazon Kubernetes Services, and compare results with quantum support vector machines (QVM) on Rigetti and IonQ using Amazon Braket. We will use dedicated development instances, maintain a PostgreSQL database for meta-data and keep results in object store. The following table represents estimates for episodic clusters, quantum compute resources, object storage, and a database including costs outlined in the Data Management Plan. We include a miscellaneous category to cover hosted notebooks, data provisioning, egress (after including [data egress discount for researchers](#) by AWS) and other cloud-related expenses.



Cloud cost estimated budget, three years, staff of 12: \$266,418 (from AWS rates as of 10/2020)*

*This estimated budget is for illustrative purposes.

Resource Type	Year 1	Year 2	Year 3
12 Remote Linux Desktops: Amazon Workspaces	\$4,176	\$4,176	\$4,176
Aurora PostgreSQL-Compatible (db.r5.xlarge). Standard Reserved Instances 1 year upfront	\$2,794	\$2,794	\$2,794
912 Experiments for QVM with 1000 shots on IONQ	$912 * (0.3 + 1000 * 0.001) = \1186	\$1,186	\$1,186
912 Experiments for QVM with 1000 shots on Rigetti Aspen 8	$912 * (0.3 + 1000 * 0.00035) = \593	\$593	\$593
Compute VM + EBS: Operating system: Linux, EC2 Instance Savings Plans each year full upfront. Storage for each EC2 instance: General Purpose SSD, Storage amount: 1 TB, Instance type: c5n.18xlarge, 72 vCPUs, 192GiB Memory, and 100 Gigabit of network performance.	\$21,255	\$21,255	\$21,255
50 TB Object Storage Amazon S3	\$14,132	\$14,132	\$14,132
1 PB Data Archive for 10 years	\$40,670	\$40,670	\$40,670
Misc (Hosted Notebooks, Post Project Data, Egress beyond 15% discount)	\$4,000	\$4,000	\$4,000
Total	\$88,806	\$88,806	\$88,806
Data egress discount (15%) included from AWS for researchers	\$13,320	\$13,320	\$13,320

Learn More

To learn more about CloudBank, watch this [on-demand presentation](#). Please [contact us](#) or email help@cloudbank.org for help.