## ***NSF CloudBank Supplement Template***

This document contains a template for CloudBank supplements to be submitted alongside supported NSF project proposals, and is followed by an example proposal using vendor-neutral language. Your NSF CloudBank Supplement is a document no longer than two pages which outlines:  
  
**(a)** anticipated annual and total costs for accessing the desired cloud computing resources, based on pricing currently available from the public cloud computing providers; and  
  
**(b)** a technical description of, and justification for, the requested cloud computing resources.

Keep in mind that what follows is a template, and you should add or remove sections of it as applicable for your specific use case. It is useful in your actual proposal to include as much meaningful detail about your platforms of choice as possible. This does not constitute a commitment, but rather helps the reviewer understand the intention and feasibility of your computing plan.

To emphasize this point: if your proposal receives funding, you are under **no obligation** to use the specific cloud platform described in your supplement. You can switch or use multiple cloud platforms as you deem appropriate.

Proposers should include "CloudAccess" (one word without space) at the end of the Overview section as a keyword (before the section on Intellectual Merit) of the Project Summary page if incorporating this request into the proposal.

CloudBank can be contacted for consulting in preparation of such a document. Please email [help@cloudbank.org](mailto:help@cloudbank.org) with any questions you may have.

*DELETE THIS PAGE* ✂

## 

## Proposal Supplement for Cloud Computing with CloudBank

**Proposal: [PROPOSAL NAME]**

**Investigator(s): [PI NAME AND INSTITUTION(S)]**

**Anticipated cloud platform(s): [Amazon AWS / Google Cloud Platform / Microsoft Azure]**

**Anticipated total cloud cost: [TOTAL COST]**

We propose using **[CLOUD PLATFORM]** for research computing needs over **[AWARD TERM]** as described in our proposal “**[PROPOSAL TITLE]**”. Per the Project Description, we begin with **[1-2 SENTENCE DESCRIPTION OF INITIAL DATA SOURCE]**.

**[1-2 SENTENCE DESCRIPTION OF OVERALL PROJECT GOAL]**

**[1-2 SENTENCE DESCRIPTION OF HOW OUTPUTS OF CLOUD COMPUTATION WILL BE USED]**

**[BULLET LIST OR 2-4 SENTENCE DESCRIPTION OF CLOUD RESOURCES TO BE USED AND THEIR APPROXIMATE TIMELINE *(include specific service names and tiers where possible, eg. VM sizes and data storage type)* ]** We expect approximately **[# OF PEOPLE]**  people on our team to be responsible for managing these resources, and **[# OF PEOPLE]** to be accessing them regularly.

Data accumulation rate estimate (GB/month): **[RATE *(if applicable)*]**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Item | Cost (Year 1) | Cost (Year 2) | Cost (Year 3) | Cost (Year 4) | Cost (Year 5) | Cost (Total) |
| **[STORAGE FOR DATASET]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** |
| **[WORKSTATIONS FOR PURPOSE ‘\_\_\_’]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** |
| **[WEB SERVER FOR PURPOSE ‘\_\_\_’]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** |
| **[MOVING DATA *OUT* OF THE CLOUD (*if applicable)*]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** |
| **[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** |
| **[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** |
| **[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** | $**[\_\_\_\_]** |
| **Total** | **$[\_\_\_\_]** | **$[\_\_\_\_]** | **$[\_\_\_\_]** | **$[\_\_\_\_]** | **$[\_\_\_\_]** | **$[\_\_\_\_]** |

**[LINK TO, OR COPY OF, OFFICIAL CLOUD CALCULATOR ESTIMATE *(if applicable)*]**

*DELETE THIS PAGE* ✂  
***Example:***

## Proposal Supplement for Cloud Computing with CloudBank

**Proposal:** *Random Forests For Thin Layer Detection in Ocean Profiles*

**Investigator(s):** Steve Zissou (University of North Alaska)

**Anticipated cloud platform(s):** Macrozon Vapor Cloud

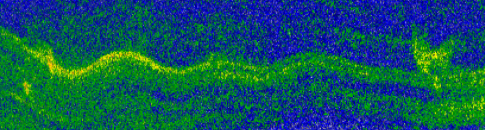
**Anticipated total cloud cost:** $100700

We propose using Vapor Cloud (imaginary for illustration purposes) for research computing needs over three years as described in our proposal “Random Forests For Thin Layer Detection in Ocean Profiles”. Per the Project Description, we begin with a synoptic workflow on data from the [OOI manifest](https://ooi-data.github.io/catalog.yaml); plus ARGO, cytometry and NASA remote sensing data.

This open source project will import, organize, clean,and analyze many (~20) independent streams of ocean sensor time series data. The initial outputs are correlation products and sharp signal annotation products. We will use Vapor Cloud to develop the processing workflows, perform the analysis and disseminate the results as a science gateway. We will use cloud VMs both in Development (“light”) mode for software development and documentation; and in Analysis (“heavy”) mode for processing, switching between modes to optimize spend.

Cost factors:

* Source data: Initialization block, accumulation rates
* Workflow development
* Data reduction to streamlined derived data products
* Data gateway: Lifetime 5 years + 5 years post-project per Data Management Plan
* Data transfer to cold storage
* Serverless Function development for real-time cruise planning



Data accumulation rate estimates, GB/month: 1.6 (ARGO), 2.1 (NASA), 2.9 (OOI), 12.2GB (OOI Broadband), 6.0 (FC). Sum 25 GB/month. Source data to be transformed into data products on a self-scaling Science Gateway. Gateway and associated storage will persist for five years after the Period of Performance. The project will also host output from the Regional Ocean Model . Project team members (5) will operate a number of workstations in either Development mode or Analysis mode.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Item | Cost (Year 1) | Cost (Year 2) | Cost (Year 3) | Cost (Year 4) | Cost (Year 5) | Cost (Post-PoP) | Cost (Total) |
| object + cold storage: source data | $200 | $200 | $200 | $200 | $200 | --- | $1000 |
| workstations + block storage | $6420 | $6420 | $6420 | $6420 | $6420 | --- | $32100 |
| NFS 4TB + backup cold | $1040 | $1040 | $1040 | $1040 | $1040 | --- | $5200 |
| data science gateway | --- | $1275 | $1275 | $2550 | $3650 | --- | $7000 |
| derived product storage | $200 | $400 | $800 | $1600 | $2400 | $12000 | $17400 |
| ROM estuary model data hosting | $3200 | $3200 | $3200 | $3200 | $3200 | --- | $16000 |
| post-Period of Performance data gateway hosting (5 years) | --- | --- | --- | --- | --- | $22000 | $22000 |
| **Total** |  |  |  |  |  |  | **$100700** |

*DELETE THIS PAGE* ✂