



Mike Norman



Ed Lazowska



Kathy Yelick



The NSF CloudBank Pilot: Progress and Challenges Facilitating Cloud Adoption

Michael L Norman, Principal Investigator, CloudBank
Distinguished Professor of Physics, UC San Diego
mlnorman@ucsd.edu



Shava Smallen



Vince Kellen

A little about me

- HPC dinosaur
- Computational astrophysicist with 47 years HPC experience
- LLNL, LANL, MPA, NCSA heritage
- Director of the San Diego Supercomputer Center (2009-2021)
- PI of 3 NSF-funded national HPC systems (2011-present)
- Perennial reviewer of DOE ECP, LLNL Comp
- extreme-scale application code developer
- Teach UCSD's only graduate course in parallel scientific computing



Los Alamos Science 1985

How did I get mixed up in cloud computing?

NSF Cloud Access pilot

- Jim Kurose (NSF CISE AD) thought computer scientists should be using the public clouds to do their research
- Shopped the idea around the agency, but got no takers
- Launched Cloud Access program as a CISE initiative
- I led the team that won the proposal competition in 2019



Jim Kurose, NSF CISE AD (2015-2019)

What is CloudBank?

- The NSF funded CB as a **5-year pilot experiment** to *facilitate cloud adoption among their computer science researchers (CISE)*
- A **collaboration** of UC San Diego, San Diego Supercomputer Center, UC Berkeley, and U Washington
 - won the \$5M solicitation with proposal *“CloudBank: Managed Services to Simplify Cloud Access for Computer Science Research and Education”*
- I formulated the concept, name, and assembled the proposal team
 - *Idea: make using public clouds as easy as online banking*
- After 1 year of portal development, CB went into production in **Aug. 2020**
- ~2 ½ years later, we facilitate **200 active projects** led by NSF CISE PIs
- Cloud spend is approaching **\$100K/mo.**
- Predominantly **AWS & GCP**, but **Azure, IBM Cloud** also requested

Managed Services to Simplify Cloud Access for
Computer Science Research and Education

[REQUEST NSF CISE ACCESS \(DCL 22-087\)](#)

Eligible NSF Solicitations

[NSF 22-509: CISE Community Research Infrastructure...](#)

[NSF 21-583: Geoinformatics \(GI\)](#)

[NSF 22-632: Cyberinfrastructure for Sustained Scie...](#)

[NSF 22-631: Computer and Information Science and E...](#)

[NSF 21-551: Cyber-Physical Systems \(CPS\)](#)

[all solicitations >](#)

Latest News

[Cloud training opportunities from Internet-2](#)

[Advanced CLASS Cloud Training April 3-14](#)

[Internet2/Microsoft Azure Accelerator Presentation...](#)

[CloudBank Initiative Offers Researchers Access to...](#)

[Join CloudBank In-Person or Online @ SC'22](#)

[all news >](#)

Who are CB's users? NSF PIs in CISE

Computer and Information Science and Engineering (CISE)

Computing and Communication Foundations (CCF)

Computing and Communication Foundations advances computing and communication theory, algorithms for computer and computational sciences and architecture and the design of computers and software.

Computer and Network Systems (CNS)

Computer and Network Systems invents new computing and networking technologies, while ensuring their security and privacy, and finds new ways to make use of current technologies.

Information and Intelligent Systems (IIS)

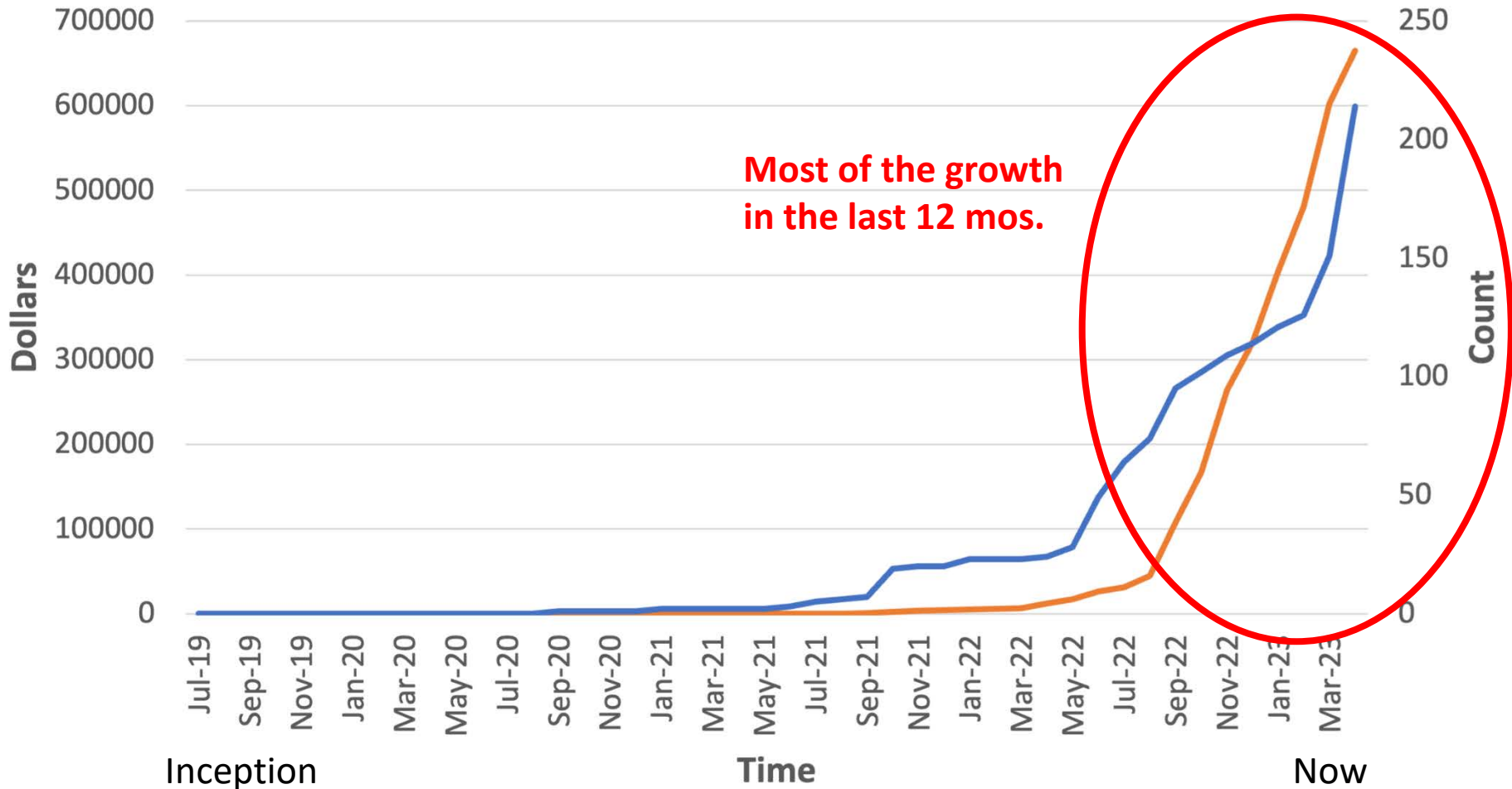
Information and Intelligent Systems studies the interrelated roles of people, computers and information to increase the ability to understand data, as well as mimic the hallmarks of intelligence in computational systems.

Advanced Cyberinfrastructure (OAC)

Advanced Cyberinfrastructure supports and coordinates the development, acquisition and provision of state-of-the-art cyberinfrastructure resources, tools and services essential to the advancement and transformation of S&E.

NSF Awards Supported by CloudBank

Cumulative Total Spend Cumulative # Awards Approved

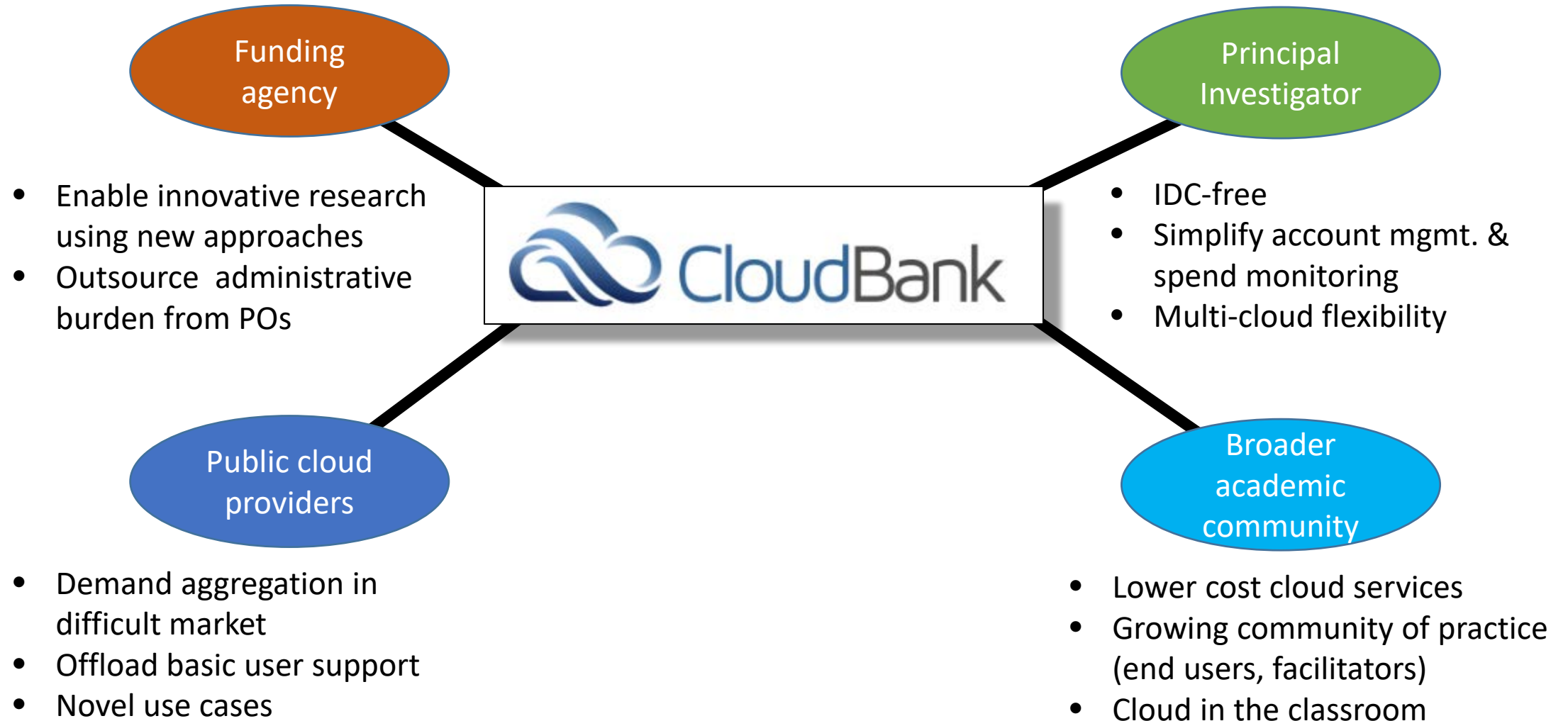


Topics

- What **value** does CloudBank offer?
- How does CloudBank **operate**?
- How are NSF PIs **using cloud**?
- What are our **learned experiences**?
- What is the **future** of CloudBank?

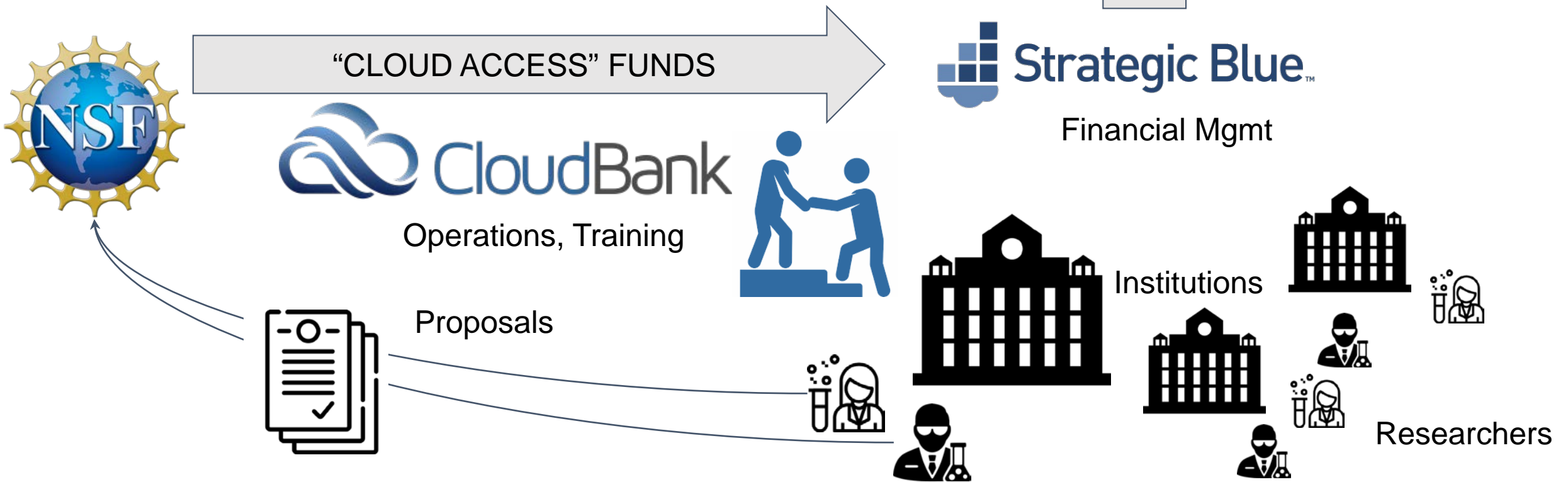


Value Proposition



Operations

Specific Solicitations



CloudBank project is organized into 3 main efforts

User portal, help desk



SDSC SAN DIEGO
SUPERCOMPUTER CENTER

Financial mgmt.



 Strategic Blue™

UC San Diego

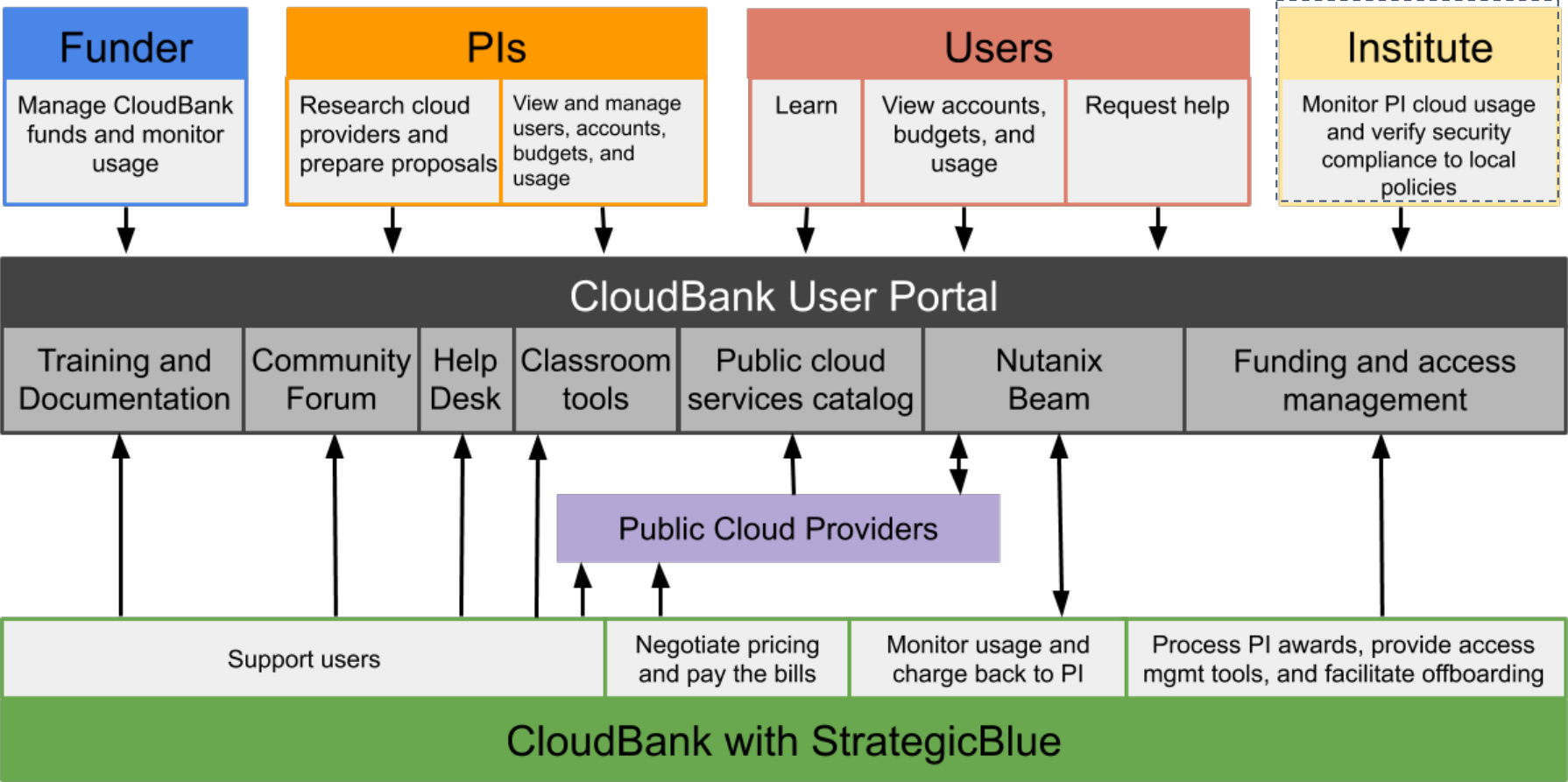
Education &
training



W
UNIVERSITY of
WASHINGTON

Berkeley
UNIVERSITY OF CALIFORNIA

Portal Architecture



Berkeley Data Stack

Dashboard

(Mimics online banking experience)

[View](#) [Revisions](#)



Manage CloudBank Funds

View and manage your CloudBank funds and grant access to associated billing accounts (IAM).



Access CloudBank Billing Accounts

View all of your cloud billing accounts and current spend. Access public cloud web portals to manage cloud services.



Monitor & Optimize Your Usage

Access Nutanix Beam to monitor usage for each of accounts and view recommendations to optimize your usage.



View & Administer Funders

View funders and grant administrator access and create new funds if enabled.



CloudBank Portal User Guide

Learn how to manage your funds and access your Cloud accounts.



CloudBank Catalog

Explore available public clouds and do an at a glance comparison of features.

CloudBank Catalog

HOME / LEARN

Big Data & Analytics

Big data is a combination of structured, semistructured and unstructured data collected by organizations that can be mined for information and used in machine learning projects, predictive modeling and other advanced analytics applications. [Source: TechTarget]. Big data analytics is the often complex process of examining large and varied data sets, or big data, to uncover information -- such as hidden patterns, unknown correlations, market trends and customer preferences -- that can help organizations make informed business decisions. [Source: TechTarget]

	Amazon Web Services	Google Cloud Platform	Microsoft Azure	IBM Cloud	Oracle Cloud Infrastructure
Data Sets	AWS OpenData , AWS Data Exchange	Google Cloud Public Datasets , Google Cloud Commercial Datasets	Azure Open Datasets , Azure Data Share	Watson Gallery	
Batch Data Processing	Amazon Elastic MapReduce	Google DataProc , Google DataFlow	Azure Batch , HDInsight	IBM App Connect	
Stream Data Processing	Amazon Kinesis , Amazon Managed Streaming for Apache Kafka (MSK)	Google DataFlow	Azure Stream Analytics	IBM Streaming Analytics	
Stream Data Ingest	Amazon Kinesis , Amazon Managed Streaming for Apache Kafka (MSK)	Google Cloud Pub/Sub	Azure Event Hubs , Azure Service Bus	IBM Event Streams	Streaming
Analytics	Amazon Redshift , AWS Lake Formation	Google BigQuery	Azure Data Lake Analytics , Azure Data Lake Store	IBM SQL Query	MySQL HeatWave Database Service , Autonomous Data Warehouse , Data Flow , Oracle Analytics Cloud
Workflow Orchestration	Amazon Data Pipeline , AWS Simple Workflow Service	Google Cloud Composer	Microsoft Flow , Azure Logic Apps	Apache Airflow on IBM Kubernetes Service	Oracle Cloud Infrastructure Process Automation , Workflow
Extract, Transform, Load	AWS Glue	Google DataProc , Google DataFlow , Cloud Data Fusion	Azure Synapse Analytics , Azure Data Factory	IBM DataStage	Oracle Cloud Infrastructure Data Integration

Compute

Cloud Computing Services provide information technology (IT) as a service over the Internet or dedicated network, with delivery on demand, and payment based on usage. Cloud computing services range from full applications and development platforms, to servers, storage, and virtual desktops. [Source: Dell]

	Amazon Web Services	Google Cloud Platform	Microsoft Azure	IBM Cloud	Oracle Cloud Infrastructure
	Amazon Elastic Compute Cloud	Google Compute Engine	Azure Virtual Machines	IBM Cloud Virtual Servers	Oracle Cloud Infrastructure (OCI) Compute
Virtual Machine	✓	✓	✓	✓	✓
Virtual Private Cloud	✓	✓	✓	✓	✓
Container Engine	✓	✓	✓	✓	✓
Managed Database	✓	✓ + <small>* select "SSD Persistent Disk" on any instance type</small>	✓	✓	✓
Serverless	✓	✓	✓	✓	✓
Spot Instance	AWS Spot Instances	Google Preemptible VM instances	Azure Spot Virtual Machines	IBM Transient (Spot Pricing)	OCI Preemptible Capacity
Marketplace	AWS Marketplace	GCP Marketplace	Azure Marketplace	IBM Third Party Software	Oracle Cloud Marketplace
Serverless Function	AWS Elastic Beanstalk	Google App Engine	Azure App Service , Azure Cloud Services	IBM PaaS	Oracle WebLogic Server for Oracle Cloud Infrastructure , Oracle WebLogic Server for OKE
Container Service	Amazon Elastic Container Service , Amazon Elastic Kubernetes Service (EKS) , AWS Fargate	Google Kubernetes Engine	Azure Container Service , Azure Service Fabric , Azure Kubernetes Service	Kubernetes Service , Red Hat OpenShift	Oracle Container Engine for Kubernetes (OKE)

Chargeback costs to PIs' allocations

Drill down to identify unused resources (volumes, IPs, etc.)

Monthly spend-to-date by resource type (VM, RDS, etc.)

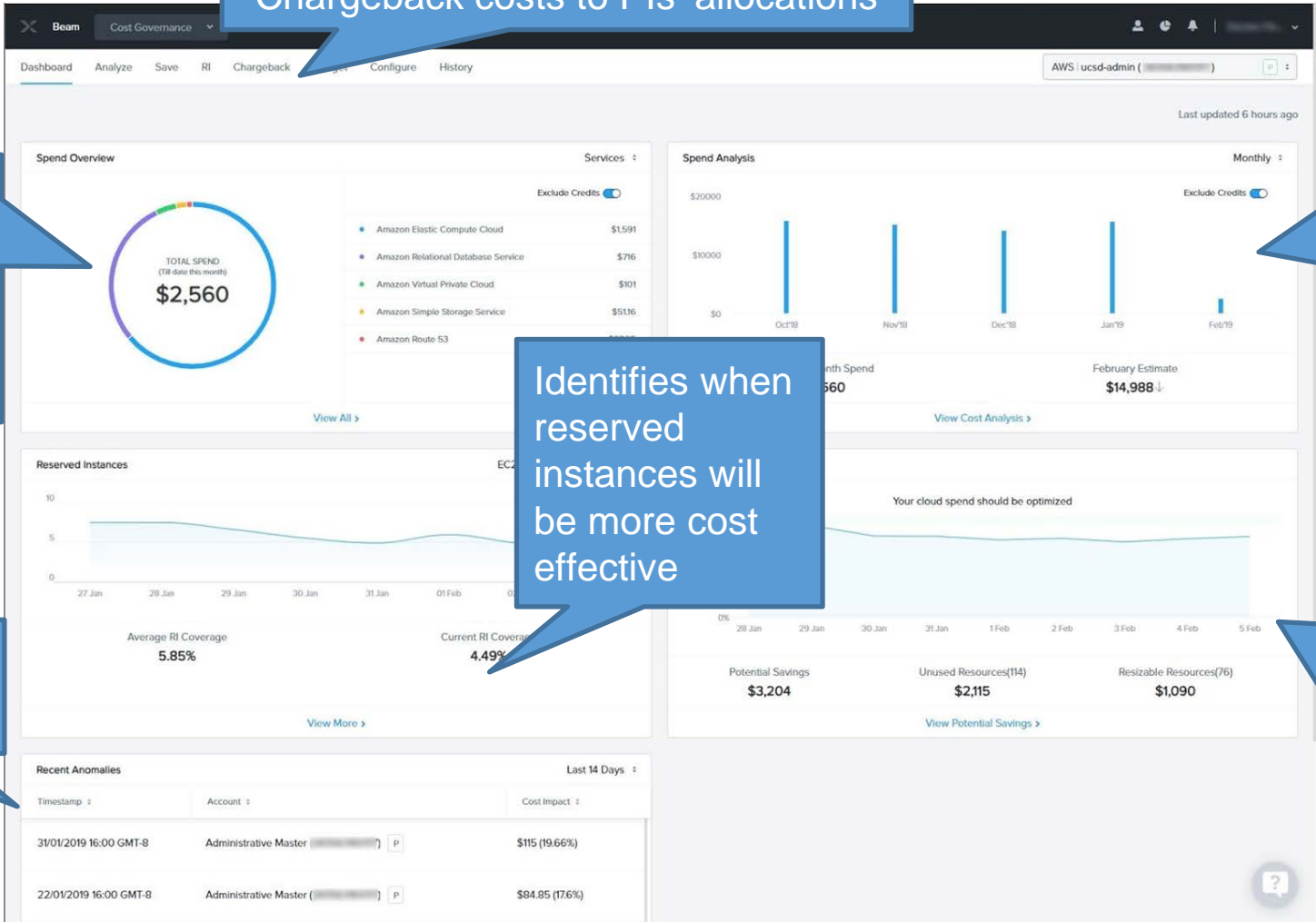
Historical monthly spending (also does forecasting)

Identifies when reserved instances will be more cost effective

Spend efficiency (summarizes allocated resources that are either unused or could be reduced in size based on historical usage)

Unusual spikes in cost

Daily cost analysis report emails



Beam

Nutanix Beam enables CloudBank to provide PI and user views across **multiple accounts** across **multiple cloud providers**



Beam Provides views of instances and compliance to security policies.

Suggests fixes to get an instance into compliance

Out of the box security policies (250+ security checks and can also define your own)

New Compliance view for all Regulatory policies and Standards. Can set Rule-based alerts to get notified on any specific audit non-compliance as and when they occur.



Beam Security Compliance

Dashboard | Compliance Summary | Compliance Remediation | Inventory | Configure | History

AWS - Research IT - PDE

Compliance Policies

Name	Created by	Last Modified At	Description
DR Policy	Product	Nov 30, 2018, 1:50:21 AM	Policy with Disaster Recovery specific audits
HIPAA (v5.0) Policy	Product	Nov 30, 2018, 1:50:21 AM	Automated Validation of HIPAA technical safeguards for AWS Services
Performance Policy	Product	Nov 30, 2018, 1:50:21 AM	Policy with Performance specific audits
Security Policy	Product	Nov 30, 2018, 1:50:21 AM	Policy with Security specific audits
Beam Compliance Policy	Product	Nov 30, 2018, 1:52:48 AM	Policy with all audits
PCI-DSS (v2.0) Policy	Product	Nov 30, 2018, 1:50:21 AM	PCI DSS Audits
CIS (v1.0.0) Policy	Product	Nov 30, 2018, 1:50:21 AM	A set of security configuration best practices for AWS

Security Compliance | AWS - cloudbank (941562322407)

Dashboard | Compliance | Remediation | Inventory | History | Reports

Regulatory Compliance View

Only the regulatory compliance policies like PCI-DSS, HIPAA are shown here

< | AWS CIS (v1.0.0) Policy | Status: All | Type: All | Automated Check: All

AWS CIS (v1.0.0) Policy | Total Checks: 93 | 62.50% Compliant

- 1. Data Protection** | Total Checks: 17 | 8 Pass | 3 Fail | 6 NA | View Details
- 2. Identity and Access Management** | Total Checks: 10 | 2 Pass | 5 Fail | 3 NA | View Details
- 3. Business Continuity** | Total Checks: 14 | 7 Pass | 1 Fail | 6 NA | View Details
- 4. Event Monitoring and Response** | Total Checks: 7 | 0 Pass | 0 Fail | 7 NA | View Details
- 5. Audit and Logging** | Total Checks: 11 | 1 Pass | 0 Fail | 10 NA | View Details

Financial Management

- Financial management is outsourced to **Strategic Blue**, a UK-based cloud services company.
- Strategic Blue specializes in **FinOps** and provides CB support on cloud billing, payment, and cost optimization
- UCSD CIO negotiates UC system-wide EDP with cloud vendors, which are leveraged for CloudBank project
- UCSD CIO partners with Strategic Blue for its **enterprise use of cloud**, and benefits from cost reductions through aggregation



User support: the Research Arc

● Necessary ● Good ● If Possible

Cloudbank

Anticipates...

- research project elements
- cloud adoption/adaptation challenges

Builds the EOT agenda based on roles

- Principal Investigator
- Administrator
- Builder
- Researcher

Ideas, Questions!

Proposal ●

Costing ●

DMP ●

CB Onboard ●

Commit ●

Roles ●

Startup ●

Role-based training ●

Cost Mgmt ●

Load in ●

Work!

Tech Review ●

Networking ●

Provision ●

Exit Strategy ●

Curriculum ●

Open use (tools, data)

Ideas, Questions!

Training Resources

<https://cloudbank-project.github.io/cb-resources/>

Cloud 101

Our Cloud 101 video series is a conceptual introduction to the cloud from a researcher's perspective. The series is split into six parts, reflecting at a coarse grain the topics that we keep in mind when using the cloud for a project.

These videos cover concrete examples, but stay at a relatively high level to keep at an informative pace. For more detailed walkthroughs of technical content, see our [Cloudbank Solutions](#).



[Thinking Cloudily](#)



[Costing](#)



[Cloud Services](#)



[Collaboration](#)



[Security](#)



[Scale](#)

Best Practices

Once you have access to a cloud account, it can feel like there is a lot to keep track of! To help out your transition from cloud novice to master, Cloudbank has put together a downloadable onboarding document that organizes the salient points that cloud users and administrators should keep in mind throughout their work:



- [Cloudbank Onboarding Protocol \(pdf\)](#)
- [Appendix: Operations \(pdf\)](#)
- [Appendix: Security and Costing \(pdf\)](#)

Design Patterns

Here are some of the most common patterns we see appear in research cloud infrastructure. Follow the links for guides on how to map this work onto cloud services.



[Machine Learning](#)



[Science Gateways](#)



[Genomics](#)



[HPC and
computationally
intensive workloads](#)



[IoT](#)



[Meta](#)



[Security and Sensitive
data](#)

Solutions

Cloudbank Solutions are researcher-oriented tutorials written by the Cloudbank team. Follow the links below to access hands-on guides for how to get the most out of your cloud access.



[Containerization](#)



[Serve data using an
Application
Programming
Interface \(API\)](#)



[High Performance
Computing](#)



[Use a cloud virtual
machine as a Jupyter
notebook server](#)



[Optimize cloud spend](#)



[Quantum computing](#)



[Review and increase
quotas](#)



[Secure data](#)

Training Resources

<https://cloudbank-project.github.io/cb-resources/>

Platform-specific resources

Although all cloud platforms offer roughly equivalent functionality, the exact terminology and services available for a given use case differ across them. Below are references and guides tailored to the specific platforms supported by CloudBank.



[Amazon AWS](#)



Google Cloud

[Google GCP](#)



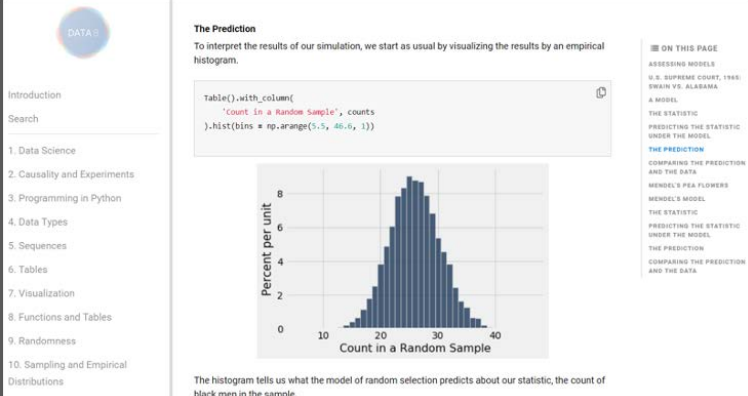
[Microsoft Azure](#)



[IBM Cloud](#)

Cloud in the classroom: the Berkeley Data Stack

- Scalable teaching both on campus and remote
- Popular Data Science curriculum -- fastest growing major at UC Berkeley
- Modern education: online textbooks, videos, interactive assignments, etc.



The screenshot shows a page from a data science textbook. On the left is a navigation menu with a 'DATA8' logo and a list of 10 topics: 1. Data Science, 2. Causality and Experiments, 3. Programming in Python, 4. Data Types, 5. Sequences, 6. Tables, 7. Visualization, 8. Functions and Tables, 9. Randomness, and 10. Sampling and Empirical Distributions. The main content area is titled 'The Prediction' and includes a code block for a histogram: `Table().with_column('Count in a Random Sample', counts).hist(bins = np.arange(5, 46, 6, 1))`. Below the code is a histogram with 'Count in a Random Sample' on the x-axis (ranging from 10 to 40) and 'Percent per unit' on the y-axis (ranging from 0 to 8). The histogram shows a bell-shaped distribution centered around 25. To the right of the histogram is a sidebar with a table of contents for the page, including sections like 'ASSESSING MODELS', 'U.S. SUPREME COURT, 1789-1803', 'ALABAMA', 'A MODEL', 'THE STATISTIC', 'PREDICTING THE STATISTIC UNDER THE MODEL', 'THE PREDICTION', 'COMPARING THE PREDICTION AND THE DATA', 'MENDEL'S PEA FLOWERS', 'MENDEL'S MODEL', 'THE STATISTIC', 'PREDICTING THE STATISTIC UNDER THE MODEL', 'THE PREDICTION', 'COMPARING THE PREDICTION AND THE DATA'.



Open infrastructure for introductory data science

Provide a shared computing environment

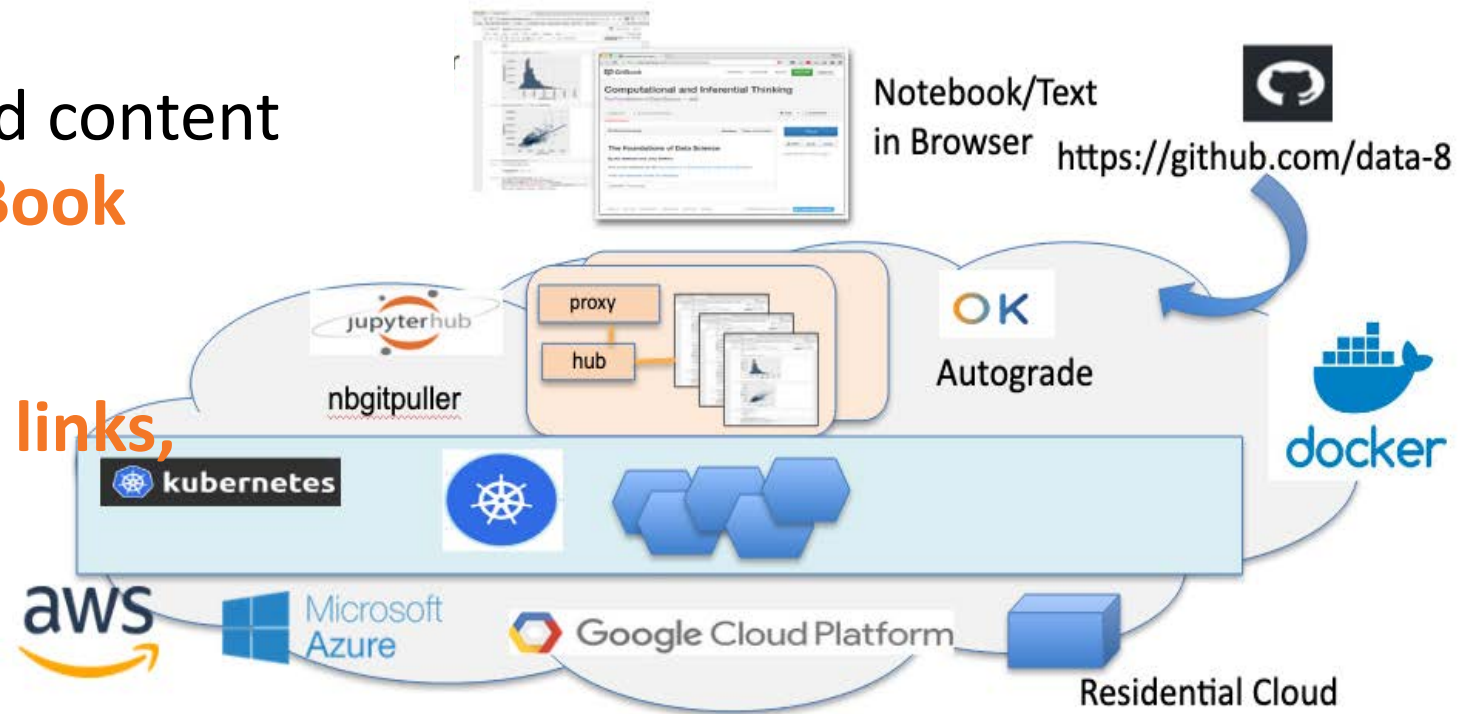
JupyterHub, BinderHub

Guide learners with structured content

Jupyter Notebooks, Jupyter Book

Tools for instruction

Online textbooks, interactive links, autograders, etc.

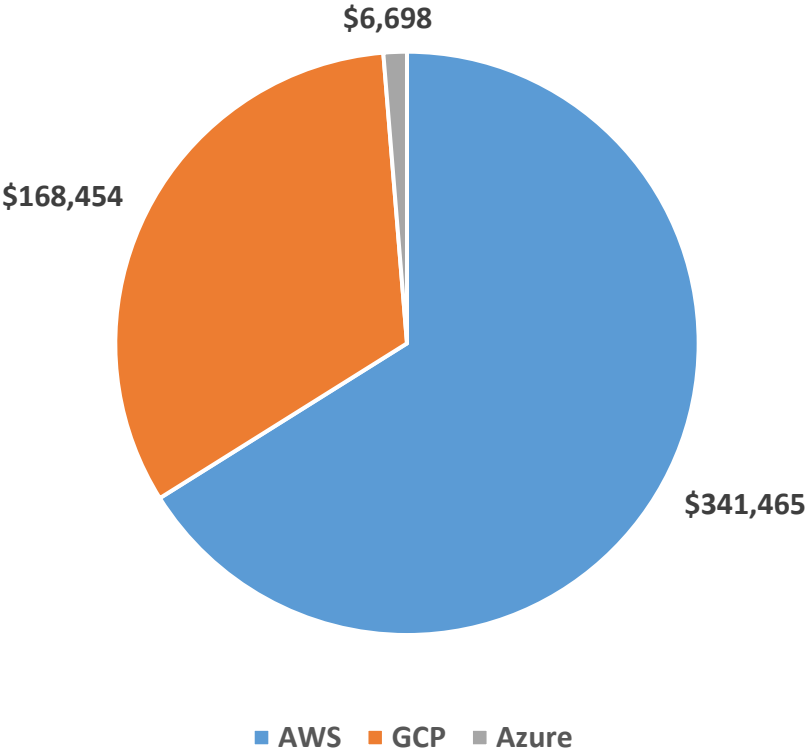


PI	University	NSF Award Title	Category
Vanessa Frias-Martinez	University of Maryland College Park	SCC-IRG Track 1: Inclusive Public Transit Toolkit to Assess Quality of Service Across Socioeconomic Status in Baltimore City	IoT and edge computing
MD Rahman	Morgan State University	CISE-MSI: DP: IIS:III: Deep Learning Based Automated Concept and Caption Generation of Medical Images Towards Developing an Effective Decision Support System (DSS)	Machine learning
Samee Khan	Mississippi State University	Collaborative Research: CNS Core: Small: HARMONIA: New Methods for Colocating Multiple QoS-Sensitive Jobs	Quantum computing
Valerio Pascucci	University of Utah	OAC: Piloting the National Science Data Fabric: A Platform Agnostic Testbed for Democratizing Data Delivery	Data processing
David Brooks	Harvard University	SHF: Medium: A Cloudless Universal Translator	Chip design
Mohammad Husain	California State Polytechnic University, Pomona	CISE-MSI:RCBP-ED:CNS:Big Data and Cloud Computing Project Hatchery	Classroom training
Kerstin Lehnert	Columbia University	Collaborative Research: Facility: Next Generation Interoperable Data Infrastructure for Geoscience Sample Data (EarthChem, LEPR/traceDs, SESAR): IEDA Re-invented	Geosciences

Award #	Award Title	PI Name	PI Inst.	Spend thru Jan 1, 2023	Summary
2153502	High-Efficiency Serverless Computing Systems for Deep Learning: A Hybrid CPU/GPU Architecture	Hao Wang	LSU	\$4444	PI is using AWS EC2 services to validate the performance of tools that facilitate machine learning on serverless platforms (Freyr+ and Libra), as well as development of a federated learning framework called Fedcod and an application to analyzing backdoor attacks.
2107150	III: Medium: Collaborative Research: Collaborative Machine-Learning-Centric Data Analytics at Scale	Chen Li	UC Irvine	\$9,079	PI described how they are using Google Compute, Storage, and Database services to host their data analysis and workflow tools. They utilize 30 VMs and 250 GB of data in the Google cloud. PI says previously "It was difficult to store this data locally, which hindered the scale of our experiments."
2006886	RI: Small: Exploiting Symmetries of Decision-Theoretic Planning for Autonomous Vehicles	Lantao Liu	Indiana University	\$7,825	PI is using AWS EC2 instances with multiple GPUs to develop a new unsupervised domain adaptive semantic segmentation method for images. With CloudBank, they are able to access AWS "computation resources with sufficient computation cores and memory to support our training with high-resolution image data."
2106446	OAC Core: SMALL: DeepJIMU: Model-Parallelism Infrastructure for Large-scale Deep Learning by Gradient-Free Optimization	Liang Zhao	Emory University	\$11,691	PI has 3 students using AWS EC2, ElastiCache/EFS, and Lambda services to do large-scale Graph Neural Network model training research. They submitted paper to ICLR 2023 and ICML 2023 and plan to submit to NeurIPS 2023. This PI also requested and received a supplement of \$8K (original request was \$10K) to complete their experiments given their excellent results.
2105329	CRII: III: Knowledge Graph Completion with Transferable Representation Learning	Muhao Chen	University of Southern California	\$16,933	PI developed and tested 1) a new Sharpness-Aware Minimization method, 2) a new abstractive summarization technique, and 3) large vision-language model workflow leveraging AWS EC2 services with GPUs. The results were submitted in 3 papers to EMNLP 2022 and EACL 2023.

AWS spend by service

Last 12 months of CloudBank spend by cloud



	A	O
1	Services	Total(USD)
2	Amazon Elastic Compute Cloud	289442.95
3	Leap,Ñ Quantum Cloud Service	20400
4	Amazon Elastic Container Service	12948.54
5	Amazon Simple Storage Service	6427.2
6	Amazon SageMaker	2624.06
7	Amazon Elastic File System	2268.32
8	Amazon ElastiCache	1992.43
9	AWS Support (Business)	1291.78
10	Amazon Relational Database Service	1187.87
11	Amazon Elastic Container Service for Kubernetes	1026.18
12	AWS Lambda	635.11
13	Amazon Virtual Private Cloud	539.84
14	Elastic Load Balancing	338.07
15	AWS Backup	329.84
16	AWS Global Accelerator	326
17	AWS Support (Developer)	284.76
18	Amazon Registrar	253
19	AmazonCloudWatch	212.68
20	Amazon WorkSpaces	195.39
21	Amazon Neptune	160.97
22	AWS DeepRacer	159.76
23	Amazon GuardDuty	104.86
24	Amazon Lightsail	103.9
25	Amazon FSx	80.68
26	Amazon EC2 Container Registry (ECR)	71.54

Learned experiences

Lesson	Response
PIs don't know CloudBank exists	Advertise in CISE newsletter Direct email to funded PIs
PIs don't know how to use cloud	Provide use cases, design patterns, community forum, security training
PIs don't know how to estimate cost	Simple online tool
PIs don't want cloud costs to impinge on their budget for people, travel, etc.	Formulate as a supplement to current awards
Some PIs have startup funds, cloud credits	PIs can manage multiple accounts on multiple clouds regardless of funding source

Future of CloudBank

- At the end of the pilot (8/24), we expect to be renewed for another 5 years
- *Goal:* grow user base, expand to other NSF directorates
- Leverage growing cloud spend to offset operations costs
 - *Sustainability strategy:* tap into universities **administrative spend**
- Toughest scaling parameter is user support
 - Research is notoriously **one-off and high touch**
 - Cloud offerings are **bewilderingly complex and constantly changing**
 - strategy is to leverage cloud vendor training materials as much as possible and grow the community of practice via **user forum**