

LIGHTNING STRIKES THE CLOUD

Lightning Round Rules



10 Minutes – Kind of Strictly Enforced



Countdown Notifications at 3 and 1 min remaining



Questions are not allowed during presentations, if we don't have time at the end, please follow up with the speaker during networking time

Click to add the title text



Amazon Lightsail for Research
Scott Friedman, Ph.D., Amazon



SAINT LOUIS
UNIVERSITY.

Lightning Strikes the Cloud at SLU
Shruthi Sreenivasa Murthy, St. Louis University



Unlocking Research Potential
on Google Cloud
Ezequiel Gioia, University of Central Florida



Automated Transcription for Social
Science Researchers
Alan Walsh, Indiana University



A dramatic background image of a dark, stormy sky with multiple bright, jagged lightning bolts striking down. The lightning is a mix of white and blue, illuminating the dark clouds. At the bottom, the dark silhouette of a forest or trees is visible against the lighter sky.

PART ONE

Amazon Lightsail for Research

Scott Friedman, Ph.D.

Amazon



Amazon Lightsail for Research

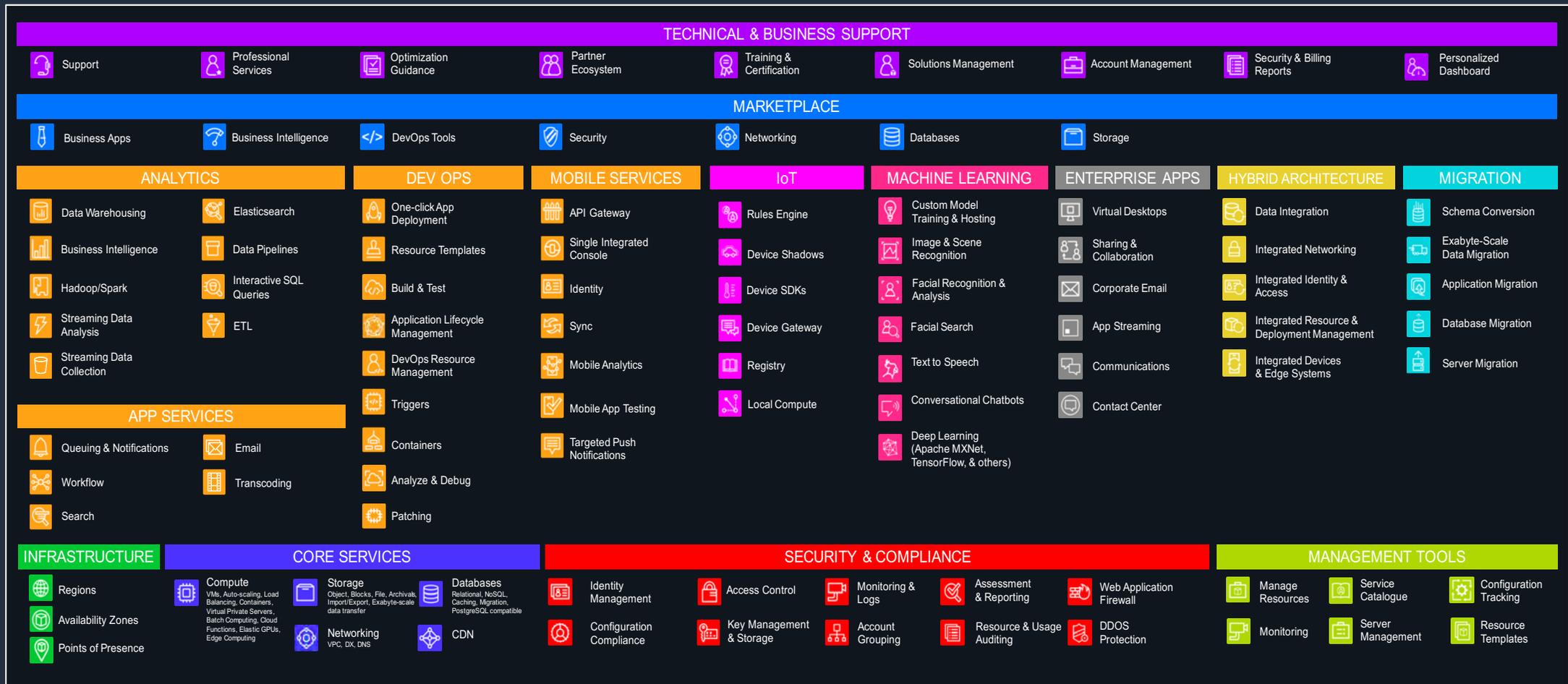
Scott Friedman, Ph.D. (he/him)

Higher Education Research
scofri@amazon.com

Research computing is not just HPC



AWS Functionality is broad and deep



Maybe too deep?

...and that's not even all of it







Lightsail for Research



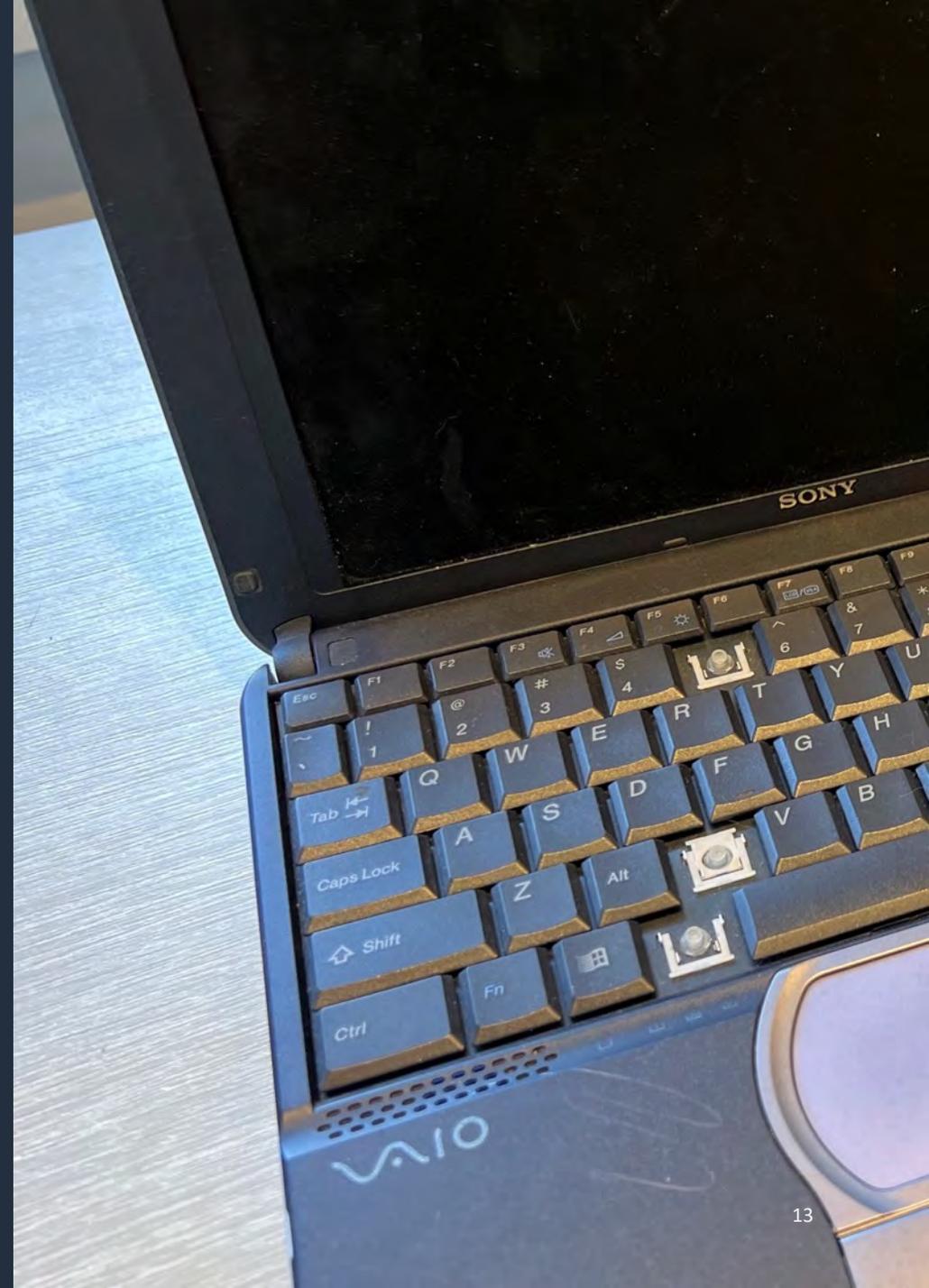
Key Features

- Lightsail for Research is an AWS service
- There is nothing to install
- Requires no cloud or IT skills to get started
- Simple to explain, understand, and use
- Offers bundled pricing, makes spending clear up front
- Has built-in cost controls, saving money

Researchers

Research lives on laptops...

- Papers & proposals written
- Data stored & visualized
- Analyses performed
- Collaboration platform

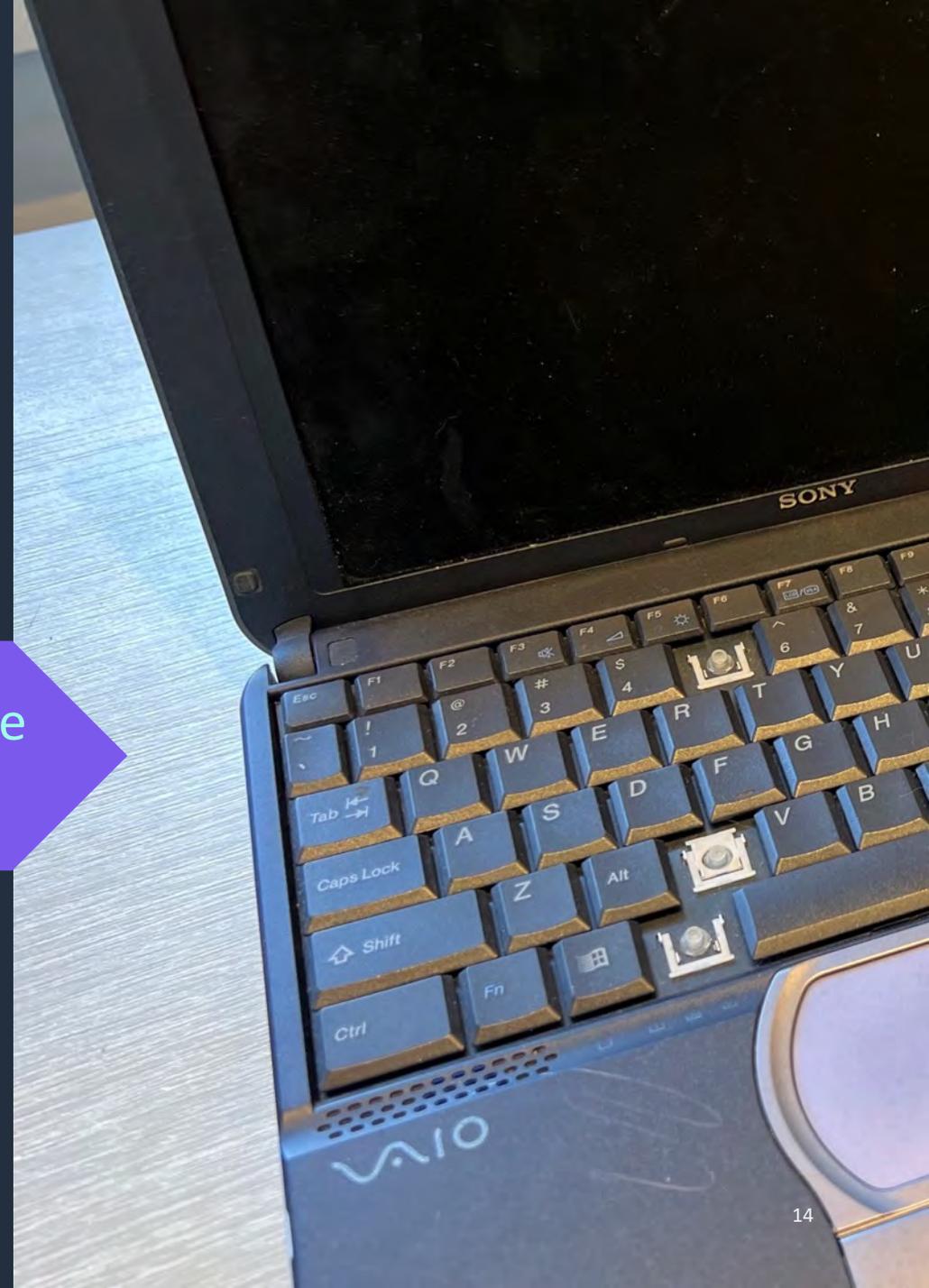


Researchers

Research lives on laptops...

- Papers & proposals written
- Data stored & visualized
- Analyses performed
- Collaboration platform

More and more
demanding!



Researchers

Modern research needs more...

- **Compute:** speed/cores/GPU
- **Memory:** limited problem size
- **Time:** long running/multiple analyses



Researchers

Most researchers exist in an IT abyss...

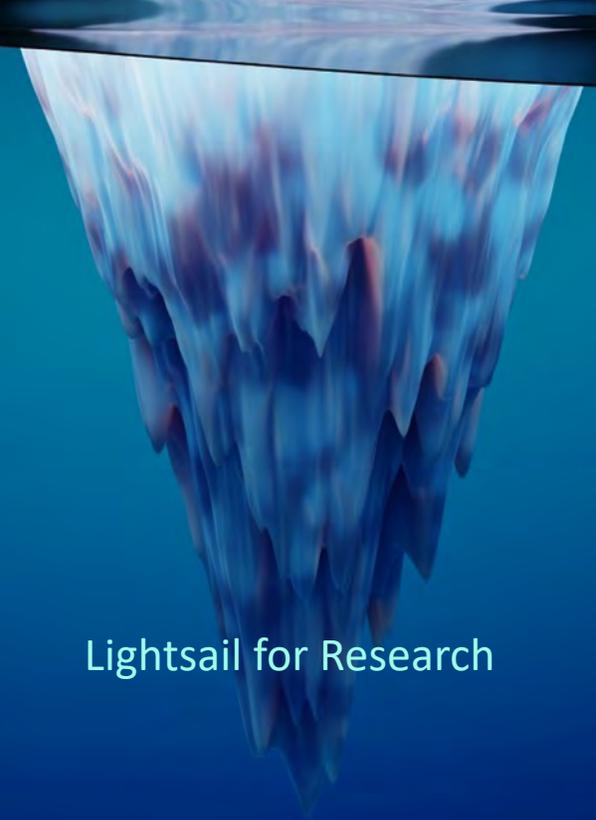
Little research IT support and enablement

Want...

to focus on their research
simple access to resources

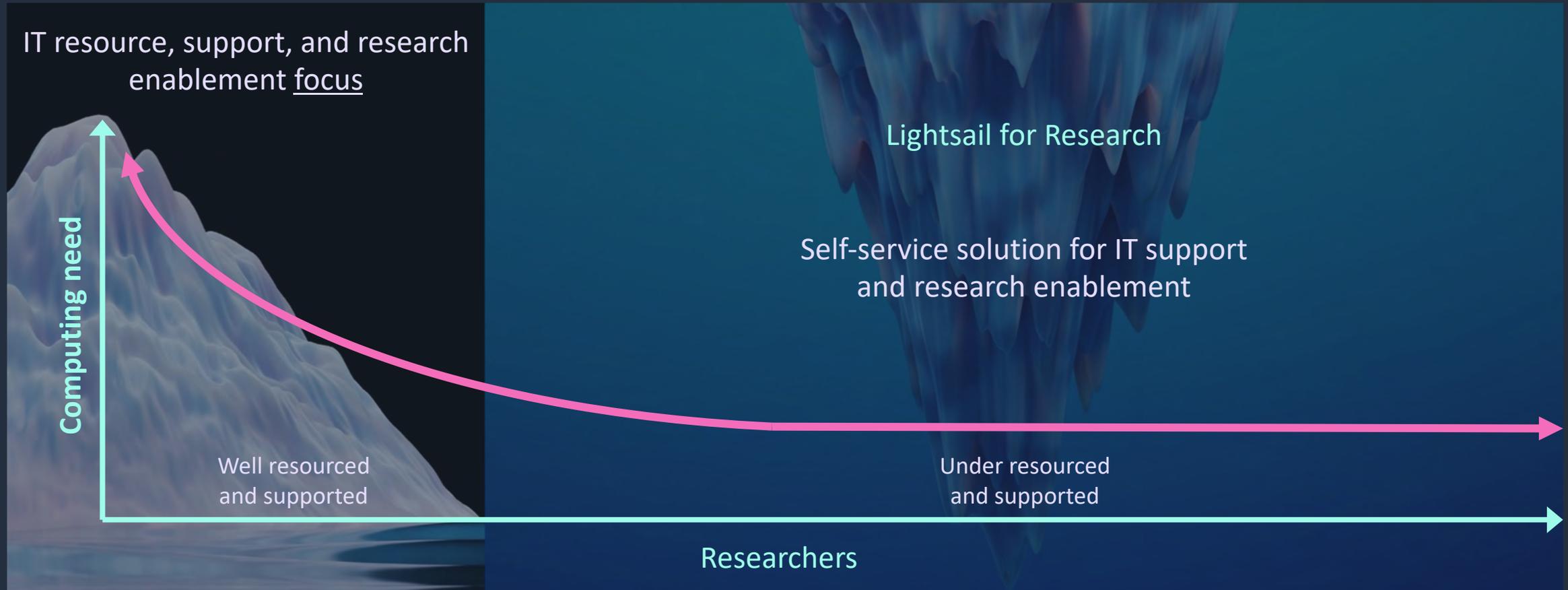
They are budget conscious

Power users



Lightsail for Research

Central and Research IT



NOTE: You will need an AWS account to follow along with the following steps.

1

Enter the URL above into the new browser tab.

The screenshot shows the AWS website for Amazon Lightsail for Research. The browser address bar contains the URL `aws.amazon.com/lightsail/research/`, which is highlighted in a pink box. A pink arrow points from this box to the 'Features' dropdown menu in the navigation bar. The main heading is 'Amazon Lightsail for Research' with the subtext 'Run complex research simulations in the cloud with ease'. Below this is a 'Get Started with Lightsail' button. Three key benefits are listed: 'Get started with a browser, no need for a powerful laptop.', 'Pay for what you use with bundled pricing.', and 'Access your preferred, pre-installed software environment at your convenience.' The 'How it works' section includes a paragraph and a four-step process flow diagram:

```
graph LR; A[Amazon Lightsail for Research] --> B[Create your virtual computer using your web browser]; B --> C[Launch your research application and upload your data]; C --> D[Run your simulations and analyze your data];
```

1
In your browser navigate to lfr.console.aws.amazon.com

2
Select a region close to you.

3
Choose an application bundle to continue

The screenshot shows the 'Choose application' page in the AWS Lightsail console. The browser address bar shows 'lfr.console.aws.amazon.com'. The page title is 'Choose application' and it includes a progress indicator for 'Step 1: Choose application' and 'Step 2: Configure virtual computer'. Under the 'AWS Region' section, the 'Region' dropdown is set to 'US West (Oregon) [us-west-2]'. The 'Application' section lists several options, each with a 'Choose' button. The 'RStudio' option is highlighted with a pink arrow pointing to the third instruction box.



Name your virtual computer Info

Virtual computer names must be unique within each AWS Region in your Lightsail account.

demo

Valid characters: A-Z, a-z, 0-9, hyphen (-), period (.) and underscore (_).

Select a plan Info

<input checked="" type="radio"/> Standard XL \$0.90 USD per hour	<input type="radio"/> Standard 2XL \$1.11 USD per hour	<input type="radio"/> Standard 4XL \$1.53 USD per hour
4 vCPUs	8 vCPUs	16 vCPUs
8 GB memory	16 GB memory	32 GB memory
50 GB storage	50 GB storage	50 GB storage
512 GB monthly data transfer	512 GB monthly data transfer	512 GB monthly data transfer

<input type="radio"/> GPU XL \$2.37 USD per hour	<input type="radio"/> GPU 2XL \$2.64 USD per hour	<input type="radio"/> GPU 4XL \$3.18 USD per hour
4 vCPUs	8 vCPUs	16 vCPUs
16 GB memory	32 GB memory	64 GB memory
50 GB storage	50 GB storage	50 GB storage
1 TB monthly data transfer	1 TB monthly data transfer	1 TB monthly data transfer

Summary

AWS Region

US West (Oregon) [us-west-2]

Application

RStudio
1.0 on Linux

Plan

Standard XL

4 vCPUs

8 GB memory

50 GB storage

512 GB monthly data transfer

Hourly total **\$0.90 USD**

Create virtual computer Cancel

By using this software, you agree to the [software's end user license agreement](#).

Your use of AWS services is subject to the [AWS Customer Agreement](#).

1
Enter a name for your virtual computer.

2
Choose a plan that matches your needs. You can change it later if you need to.

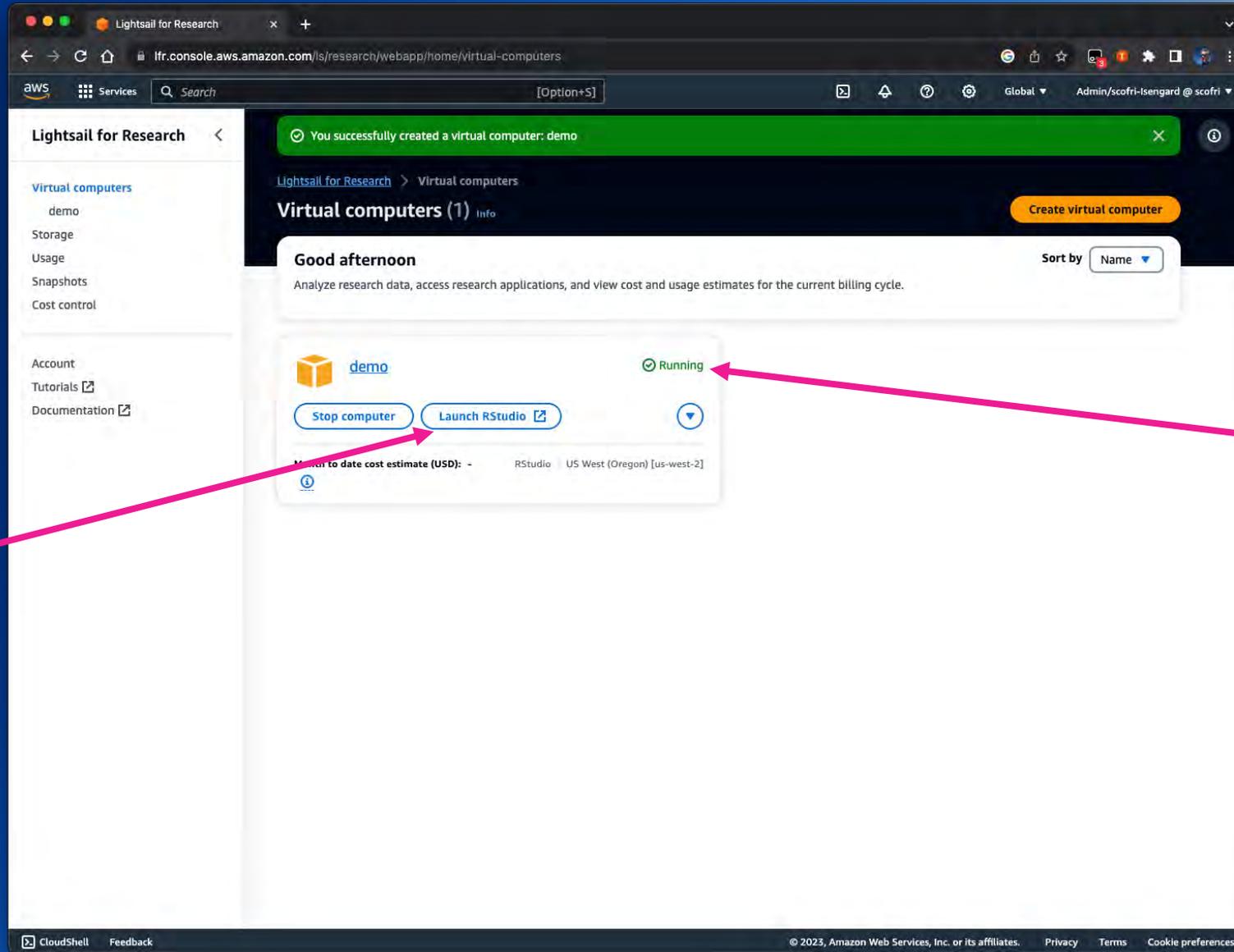
You can choose between CPU or CPU+GPU virtual computers.

3
When ready choose *Create virtual computer*.



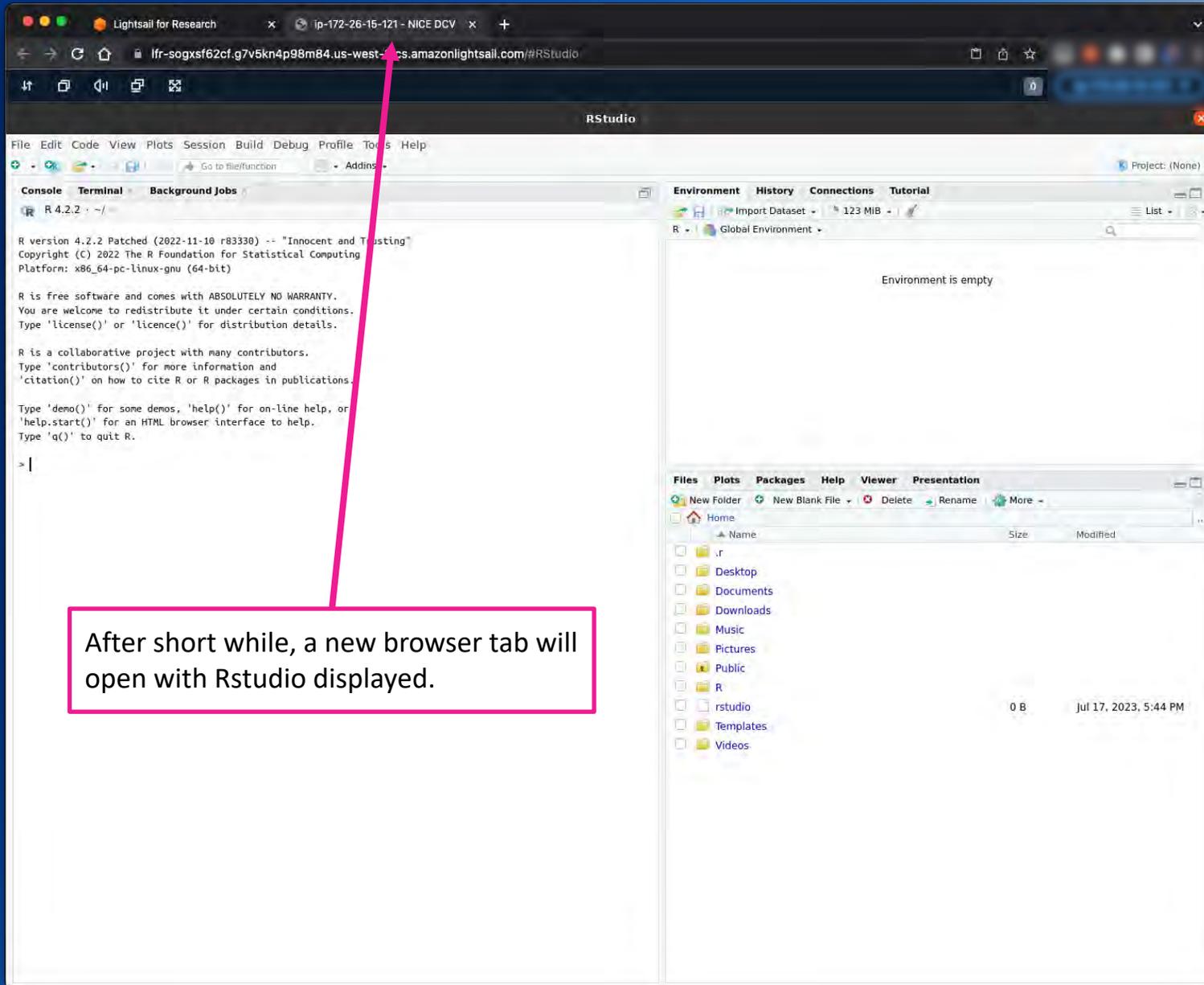
The screenshot shows the AWS Lightsail console for a 'Research' account. The main heading is 'Virtual computers (1)'. A 'Create virtual computer' button is visible in the top right. Below this, a 'Good afternoon' message is followed by a card for a virtual computer named 'demo'. The card shows the computer is in the 'Creating' state, with a progress bar at 52%. There are buttons for 'Stop computer' and 'Launch RStudio'. A 'Month to date cost estimate (USD): -' is shown for the 'RStudio' instance in 'US West (Oregon) [us-west-2]'. A pink arrow points from a text box to the 'Creating' label.

The new virtual computer is being created.

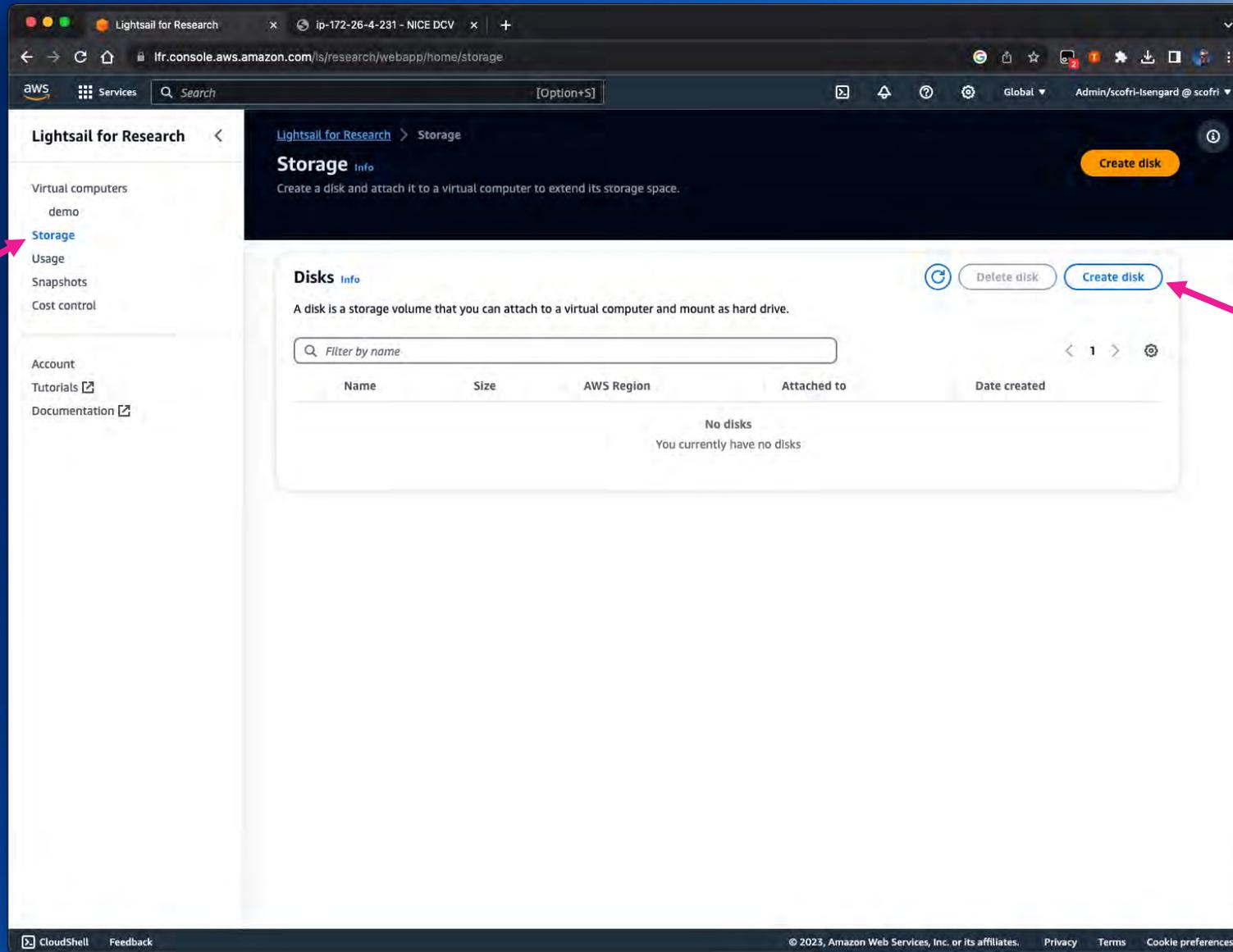


2
Choose *Launch Rstudio*.

1
After a short while your virtual computer will change its status to *running*.

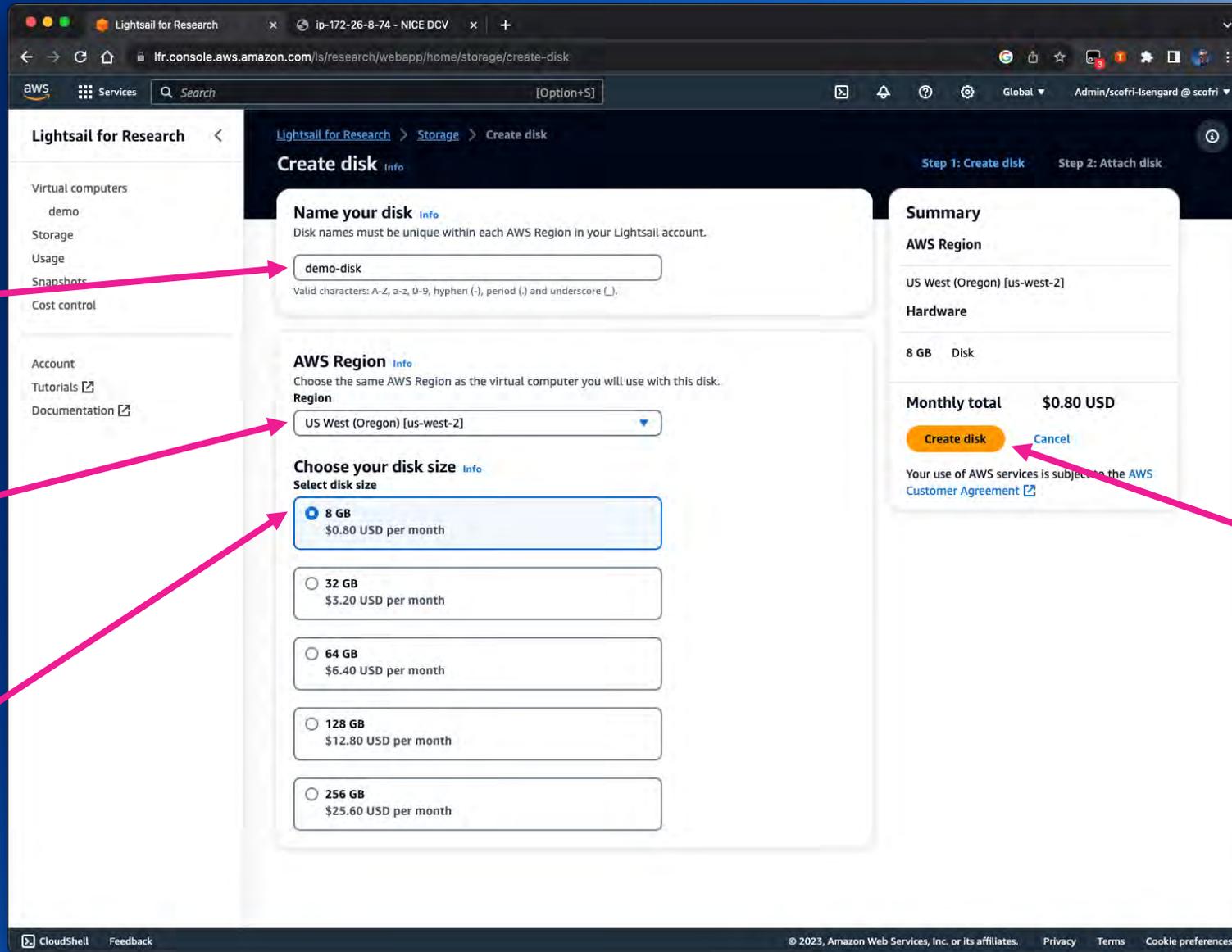


After short while, a new browser tab will open with Rstudio displayed.



1
Choose *Storage* to create a data disk.

2
Choose *Create disk*.



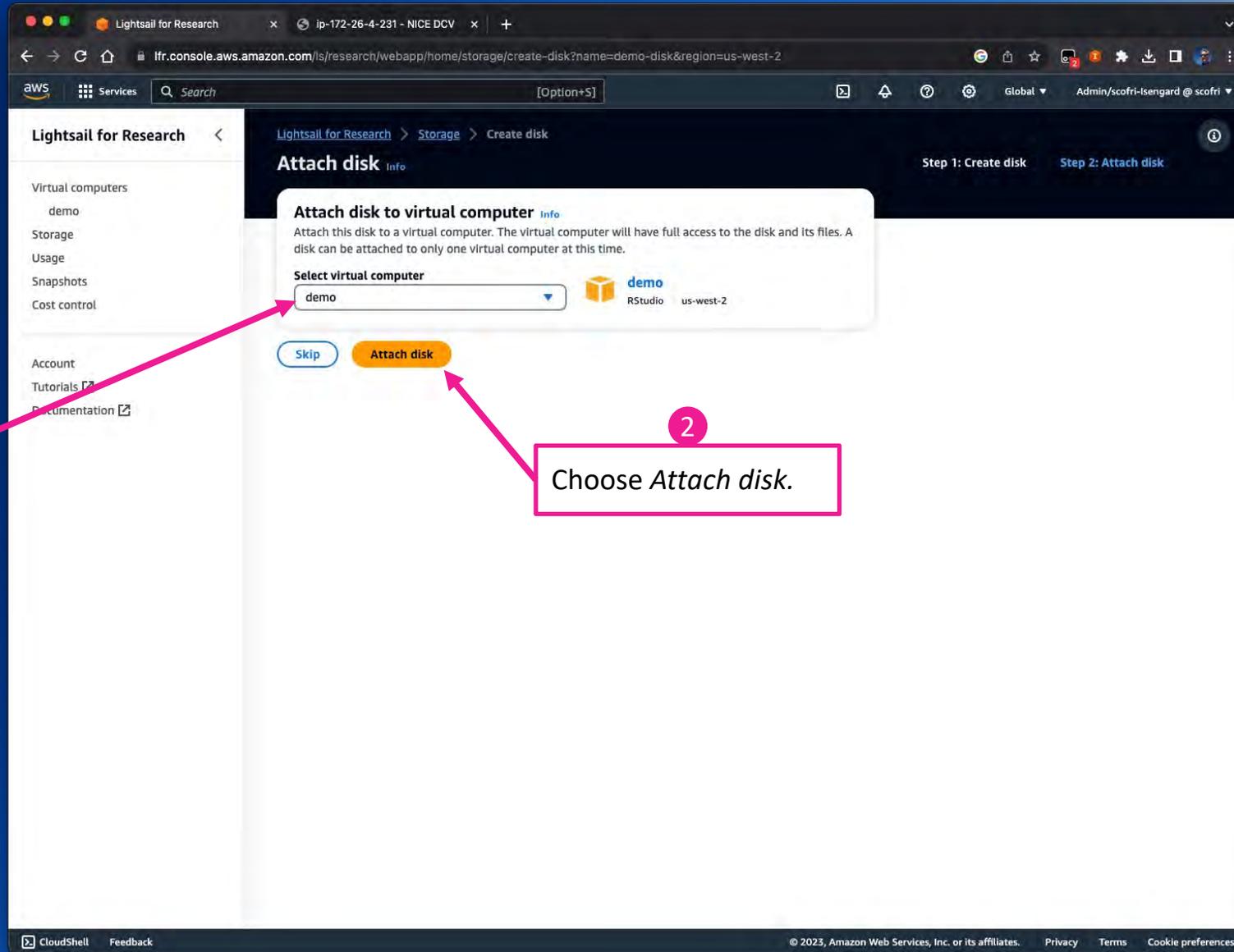
1
Enter a name for your new disk.

2
A region geographically will be selected for you by default. (it should be the same as earlier)

3
Choose a size for you disk.

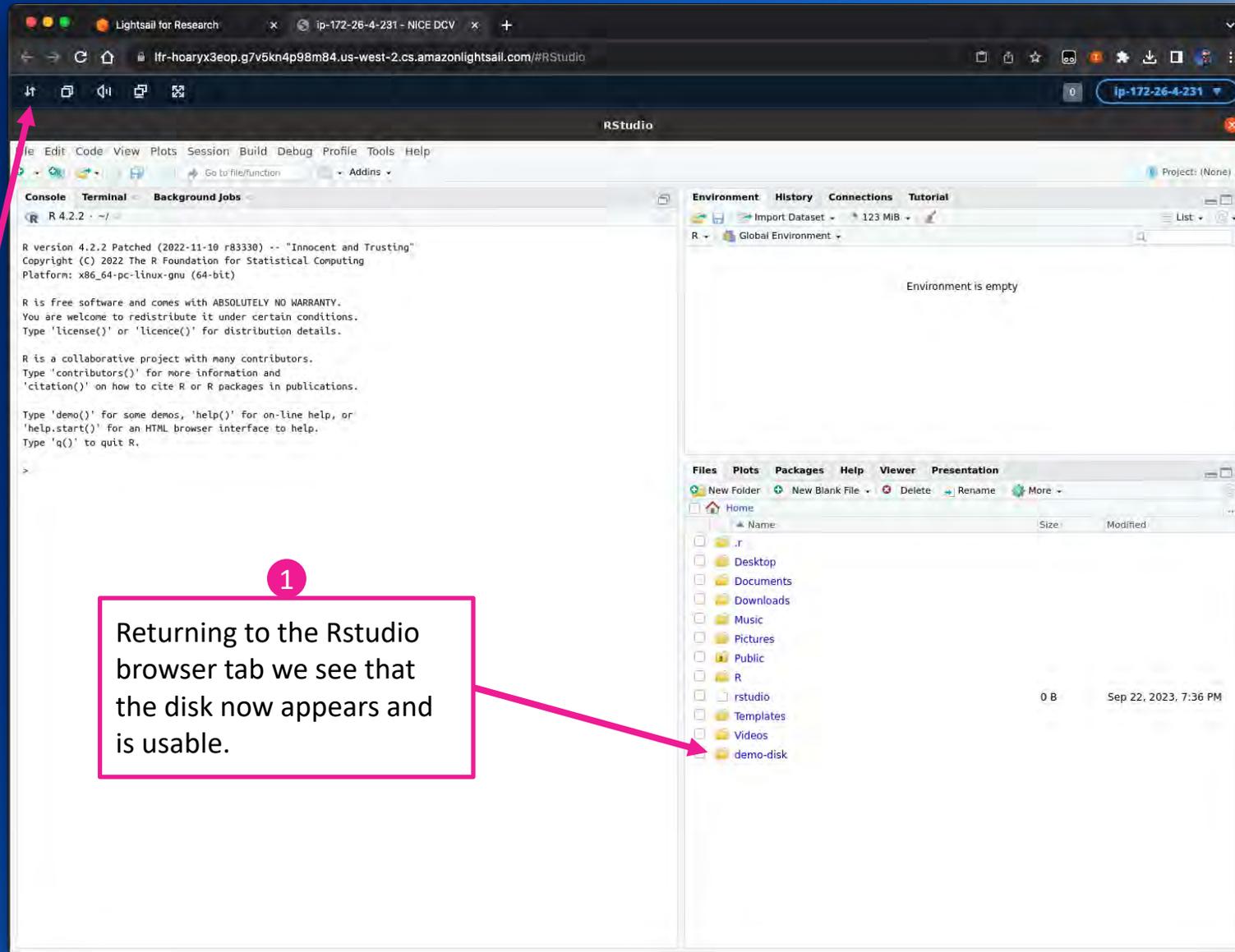
4
Choose *Create disk*.





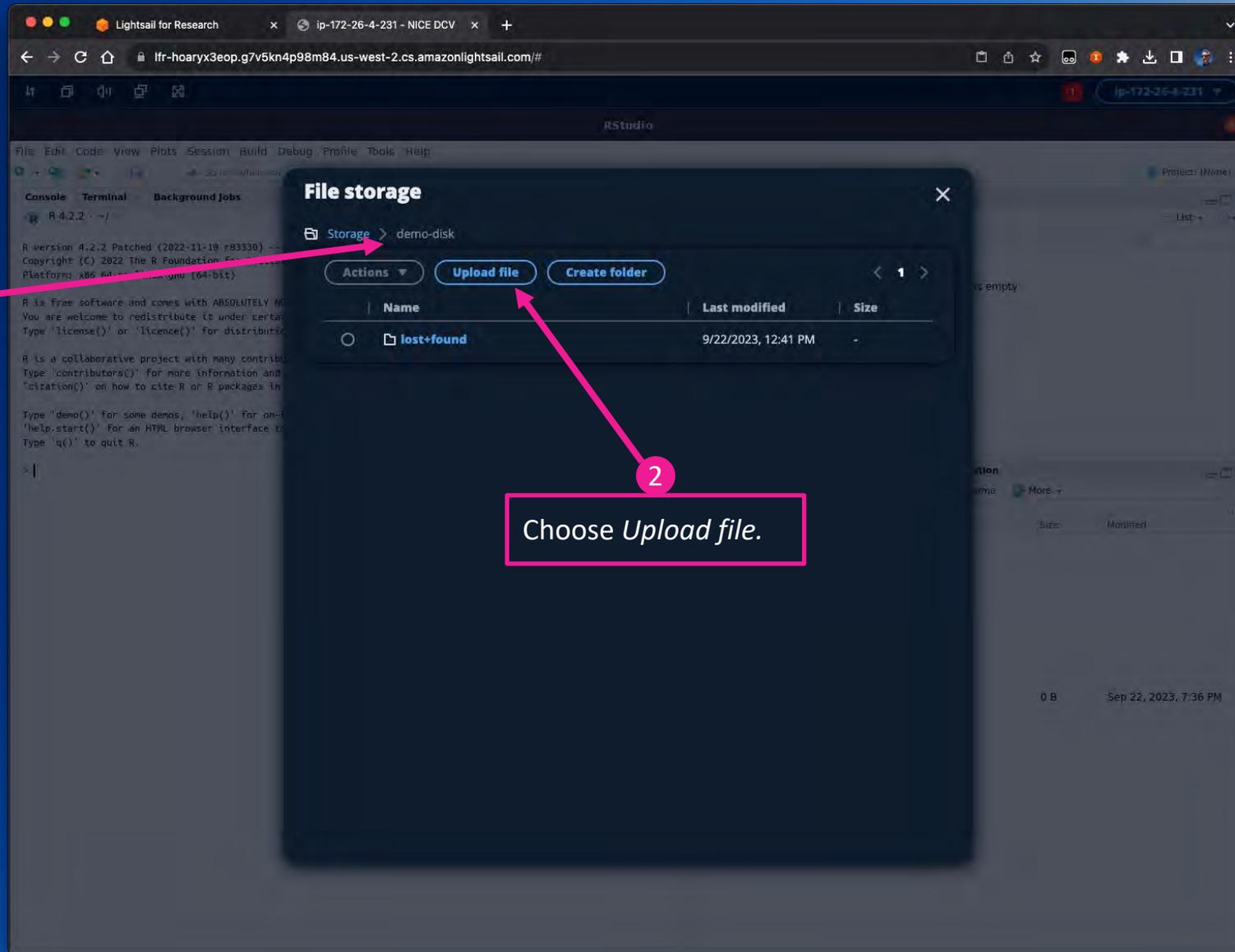
1
Select the virtual computer the new disk will be attached to.

2
Choose *Attach disk*.



2
To upload some data to our virtual computer choose the *double arrow* icon.

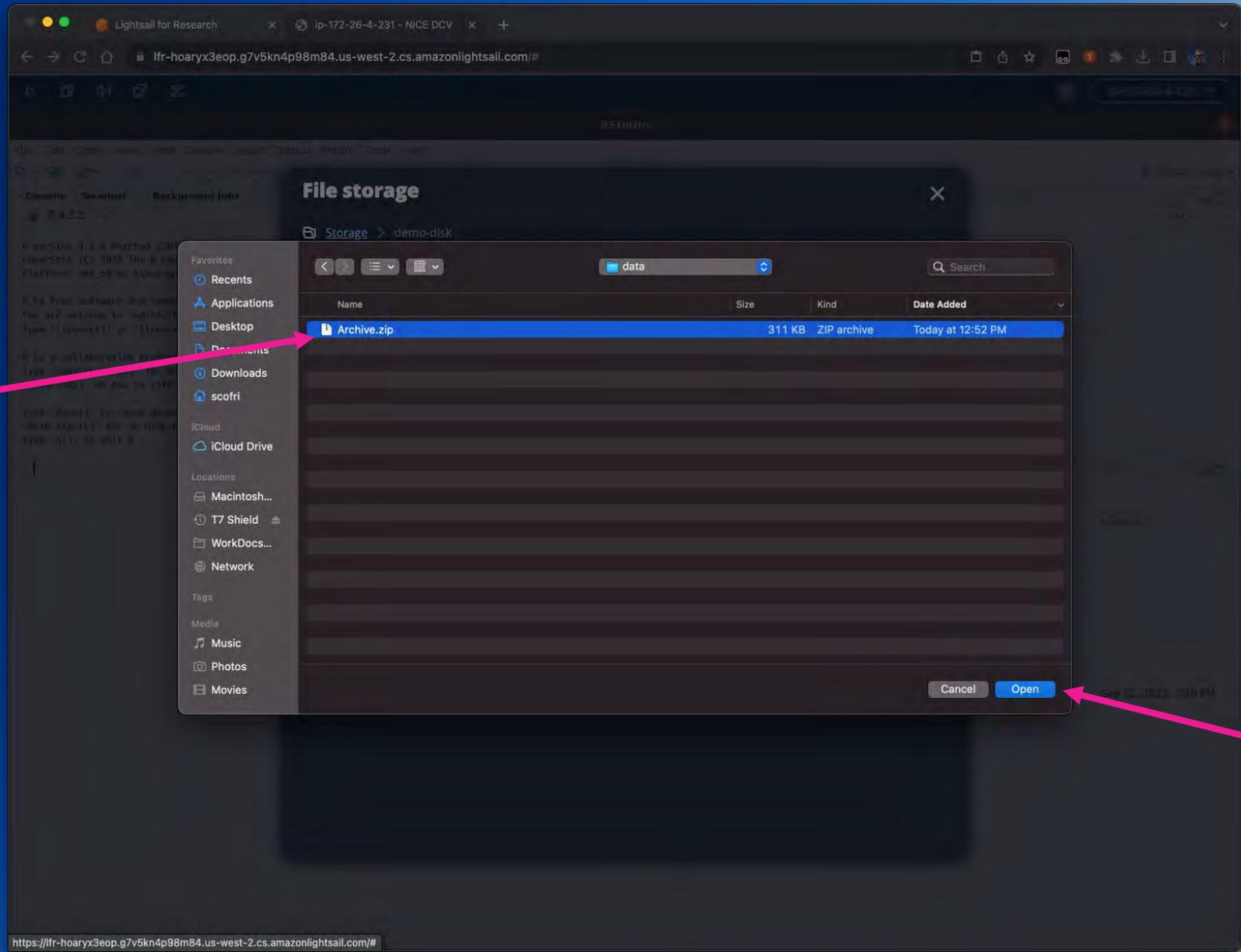
1
Returning to the Rstudio browser tab we see that the disk now appears and is usable.



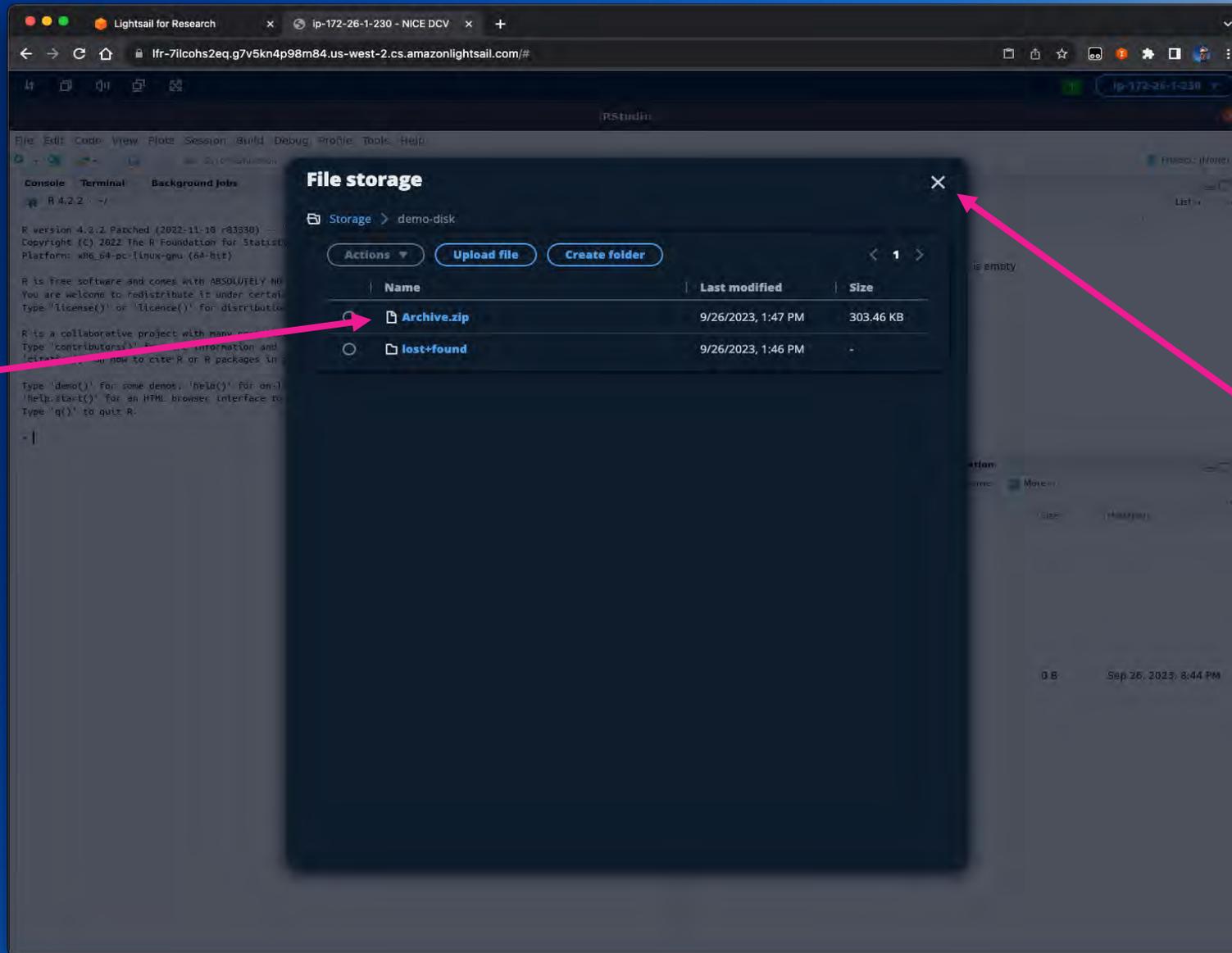
1
Navigate to the *demo-disk*.

2
Choose *Upload file*.

1
On your local computer navigate to and select the file you want to upload.



2
When ready, choose *Open*.



1

The file is now uploaded to the virtual computer.

2

When finished, click on the X to close the window.

The screenshot shows the RStudio interface running on an Amazon Lightsail instance. The browser address bar indicates the instance's IP address and the Lightsail console URL. The RStudio window title is "RStudio". The menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The main workspace is divided into several panes:

- Console:** Displays the R version 4.2.2 and its license information.
- Environment:** Shows "Environment is empty".
- Files:** Shows the file system structure. Under "demo-disk", there is a file named "Archive.zip" with a size of 303.5 KB and a modification date of Sep 22, 2023, 7:52 PM. A pink arrow points from a text box to this file.

The text box containing the arrow has the text: "The file is visible here too."

Lightsail for Research

lfr.console.aws.amazon.com/lresearch/webapp/home/cost-control

Cost control (0) info

Create cost control rules to manage your spending. Cost control rules evaluate the average CPU utilization every 5 minutes

Resource	Rule name	Conditions
You currently have no cost control rules		
Create cost control rule		

Virtual computers

- demo
- Storage
- Usage
- Snapshots
- Cost control**
- Account
- Tutorials
- Documentation

1

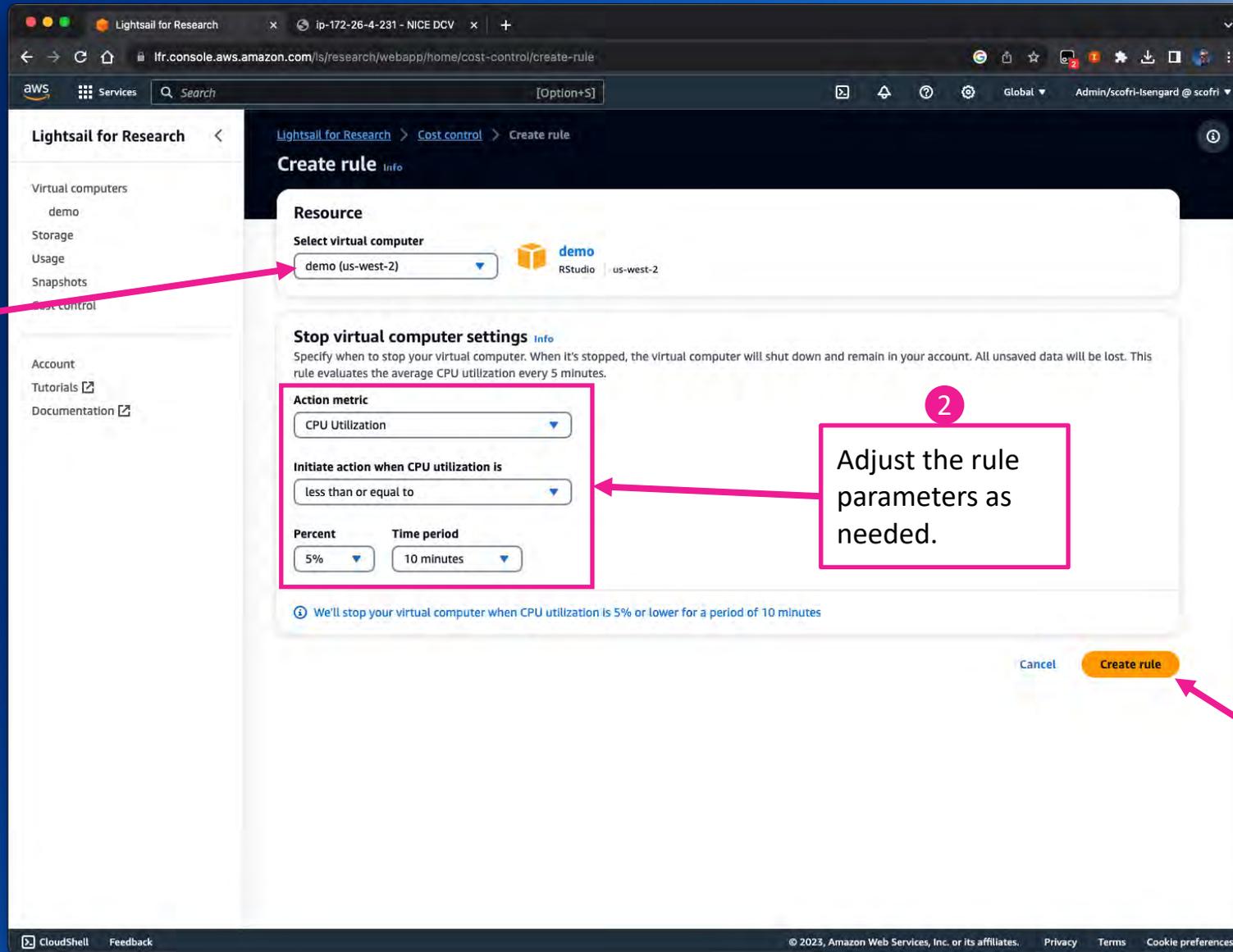
2

1

Choose *Cost control* to add a rule to your virtual computer.

2

Choose *Create cost control rule*.



1
Select the virtual computer you want to apply the rule to.

2
Adjust the rule parameters as needed.

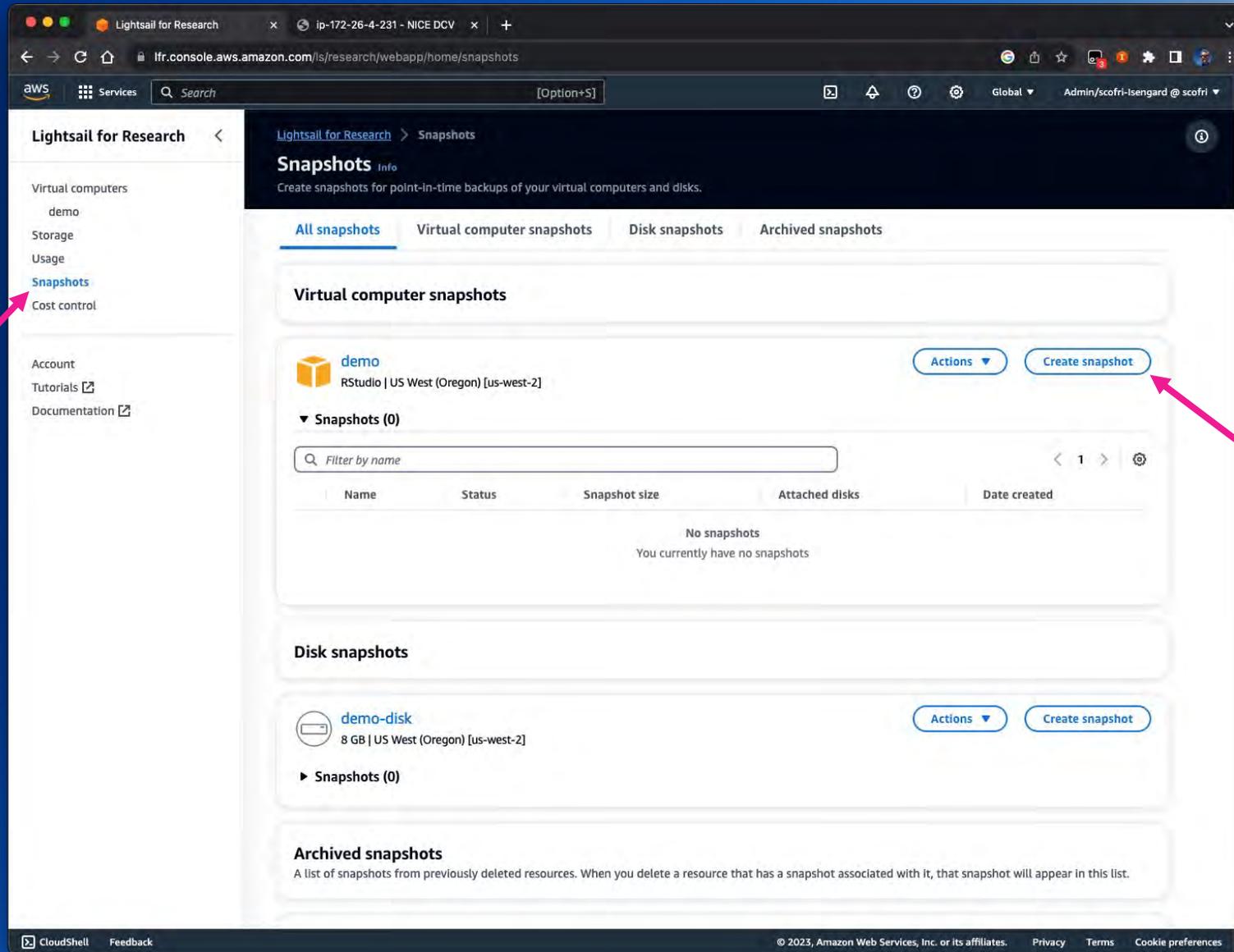
3
Choose *Create rule*.

The screenshot shows the AWS Lightsail console for a project named 'Lightsail for Research'. The 'Cost control' section is active, displaying a table with one rule. A pink arrow points from a text box to the resource 'demo (us-west-2)' in the table.

Resource	Rule name	Conditions
demo (us-west-2)	Stop virtual computer on idle	We'll stop your virtual computer when CPU utilization is 5% or lower for a period of 10 minutes

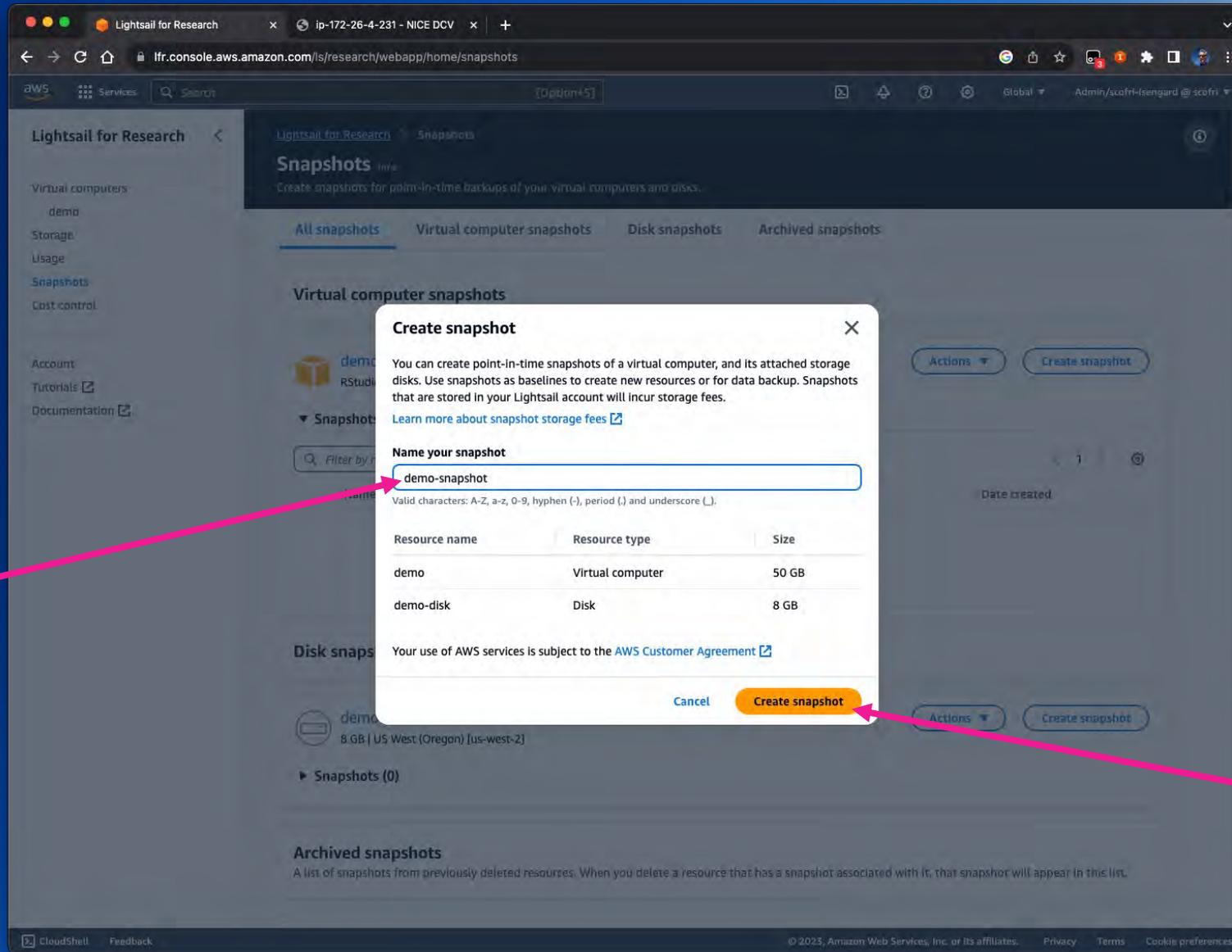
After a few moments the rule is created and attached to your virtual computer.





1
Choose *Snapshots* to save the state of your virtual computer.

2
Choose *Create snapshot*.



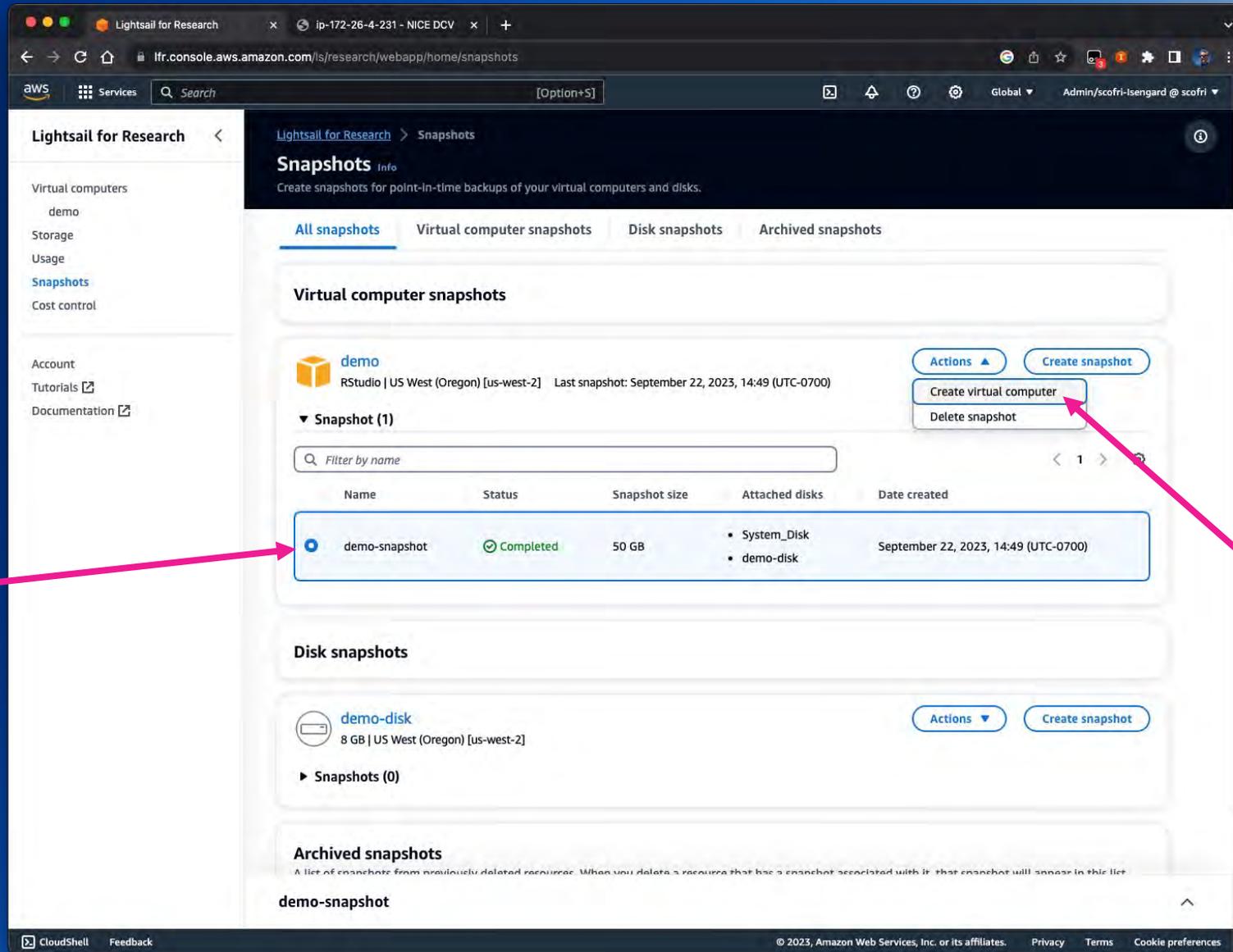
1

Enter a name for your snapshot.

2

Choose *Create snapshot*.





1

Once the snapshot is created, select it.

2

Choose the *Actions* menu and choose *Create virtual computer*.

1

Enter a name for the new virtual computer.

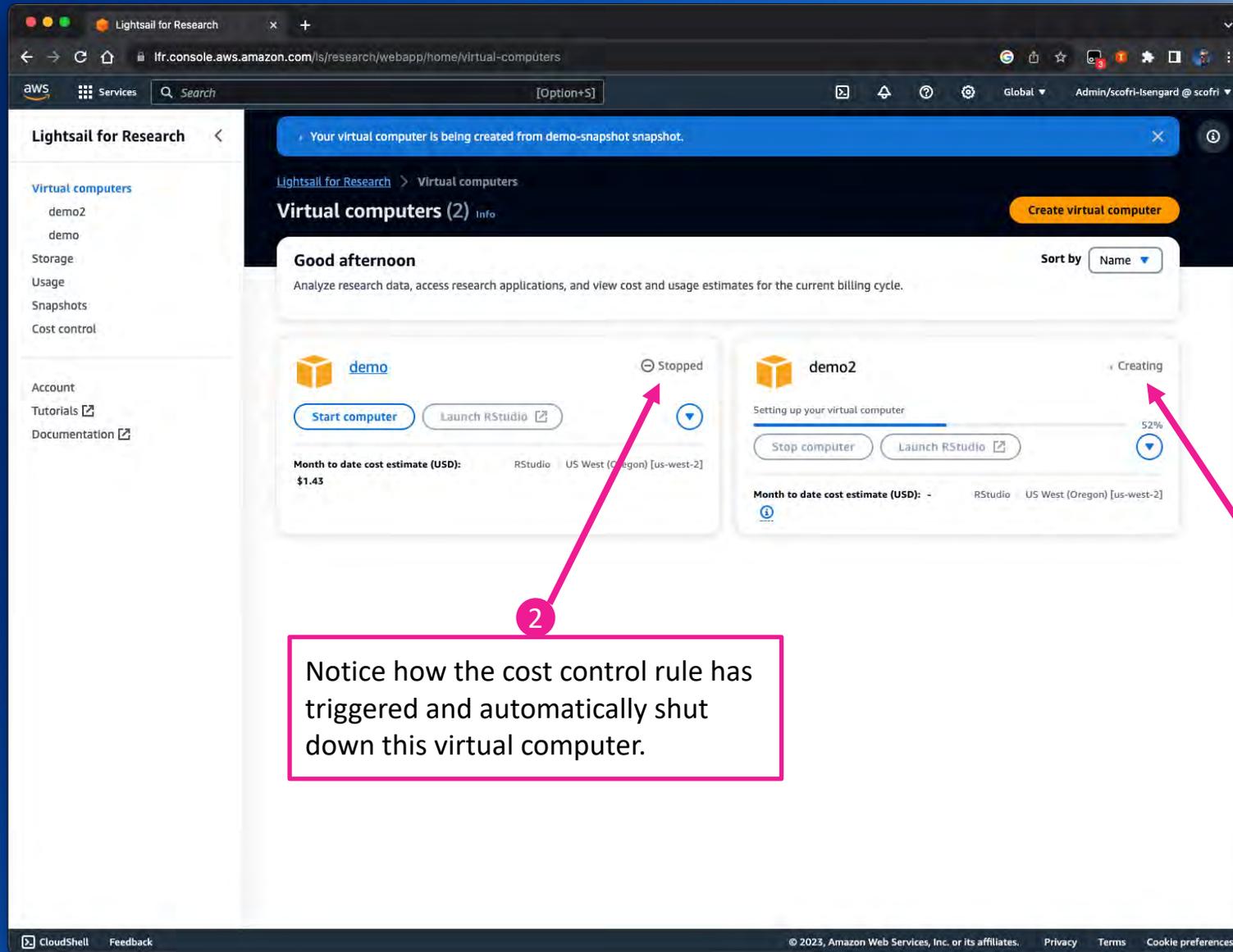
2

This time choose a different backing instance.

3

Choose *Create virtual computer*.





2

Notice how the cost control rule has triggered and automatically shut down this virtual computer.

1

The new virtual computer is being created.

The screenshot shows the AWS Lightsail console interface. On the left is a navigation sidebar with categories like 'Virtual computers', 'Storage', 'Usage', 'Snapshots', 'Cost control', 'Account', 'Tutorials', and 'Documentation'. The main content area is titled 'Virtual computers (2)' and includes a 'Create virtual computer' button. Below this, there's a 'Good afternoon' message and a table of virtual computers. The table has two entries: 'demo' (Stopped) and 'demo2' (Running). For 'demo2', a dropdown menu is open, showing options: 'Access operating system', 'Close session', and 'Delete virtual computer'. A pink arrow points from a text box to the 'Delete virtual computer' option.

Virtual computers (2) Info Create virtual computer

Sort by Name

demo Stopped demo2 Running

Start computer Launch RStudio Stop computer Launch RStudio

Month to date cost estimate (USD): \$1.43 RStudio US West (Oregon) [us-west-2] RStudio US West (Oregon) [us-west-2]

Access operating system
Close session
Delete virtual computer

To remove a virtual computer choose the action button and choose *Delete virtual computer*.

CloudShell Feedback © 2023, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

The screenshot shows the AWS Lightsail console interface. At the top, there are two green notification banners: "Your disk demo-disk has been detached from demo." and "Your virtual computer has been deleted: demo". Below these, the page title is "Virtual computers (1)" with a "Create virtual computer" button. A "Good afternoon" message is displayed, followed by a card for the virtual computer "demo2" which is in a "Running" state. The card includes buttons for "Stop computer" and "Launch RStudio". A dropdown menu is open over the computer, listing "Access operating system", "Close session", and "Delete virtual computer". A pink arrow points from a text box to the "Access operating system" option.

To access the virtual computer's Ubuntu desktop choose the action menu and select

The screenshot shows a web browser window with the URL `lfr-e6ksn8n93m.g7v5kn4p98m84.us-west-2.cs.amazonlightsail.com/#UbuntuConsole`. The browser's address bar shows the IP address `ip-172-26-4-5`. Below the browser, a terminal window is open, displaying the prompt `lightsail-user@ip-172-26-4-5:~$`. A pink arrow points from a text box below to the terminal. Another text box with a note is also present.

From the desktop you can do whatever you need to; install software, open a terminal, use the browser, etc.

NOTE: Do not forget to either add a cost control rule to this virtual computer or stop it before you finish. If you want to delete the virtual computer remember to also delete the demo-disk.



Thank you!

Scott Friedman

scofri@amazon.com



PART TWO

Lightning Strikes the Cloud at SLU

Shruthi Sreenivasa Murthy

St. Louis University



SAINT LOUIS
UNIVERSITY™
— EST. 1818 —

Lightning Strikes the Cloud at SLU

- Shruthi Sreenivasa Murthy

SCAER turns to RCG

The Sinquefeld Center for Applied Economic Research (SCAER) is one of the leading research centers at SLU. The center is a part of the SLU Research Institute and brings together all the researchers at SLU who have a common interest in applied economic research.

Research Computing Group (RCG) is a collaboration between the Office the VP of Research and the Information Technology services at SLU. RCG builds on SLU's ambition to develop and improve computing services, technology, services and support for research.



Project Discovery

Researchers at SCAER use large volumes of human mobility data to create useful research files for a wide range of studies in economics, public health, and finance related to patterns of social mobility. The underlying micro datafiles are very large, averaging approximately over 100 terabytes per year. Dataset includes human mobility data for the years 2019 to 2023.

Research Highlights



Time away from Home



Teacher turnover
expectation



Shopping Trends



More Work from Home

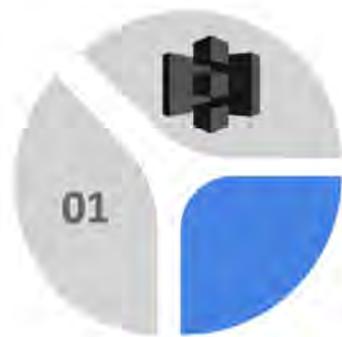


Immediate Imperative - Preprocessing

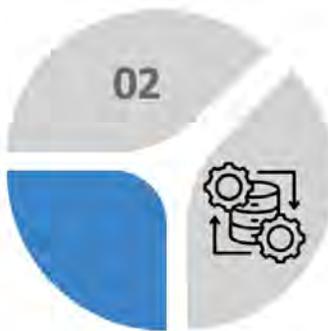
Pre-Processing Lifecycle

Veraset

Geospatial data is uploaded by Veraset to an S3 bucket in a Scaer AWS account. It is stored in a folder hierarchy organized by year, month, and day. Data is archived in S3 glacier and stored in parquet format.



02



Processing

Each record is given a time zone based on the geohash mapping contained in the time json file. Date and Time columns are then added based on the mapped time zone.

Geo Data Warehouse

Data that has been processed is then saved in a data warehouse (S3 bucket) that is divided up by a 3-char geohash, a year, a month, and a date. This would facilitate the speedy retrieval of particular files from the glacier and speed up the data input process during the post-processing stage.



03



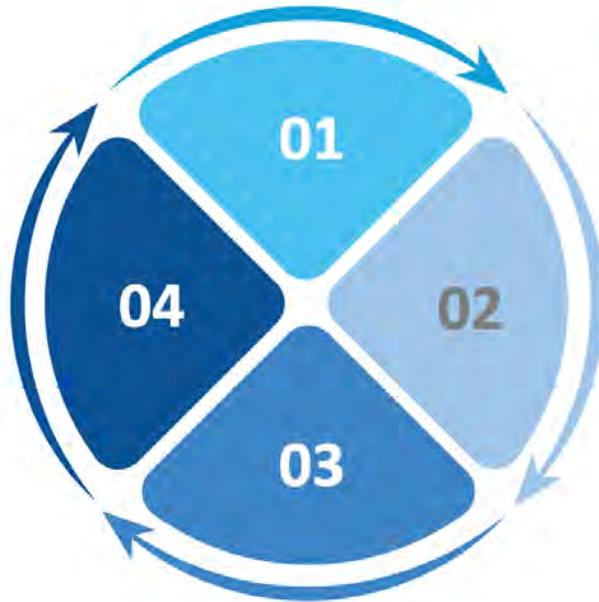
Deduplication

Geohash that is mapped to multiple timezones runs the risk of being duplicated. Before writing to the data warehouse, these duplicates are eliminated.



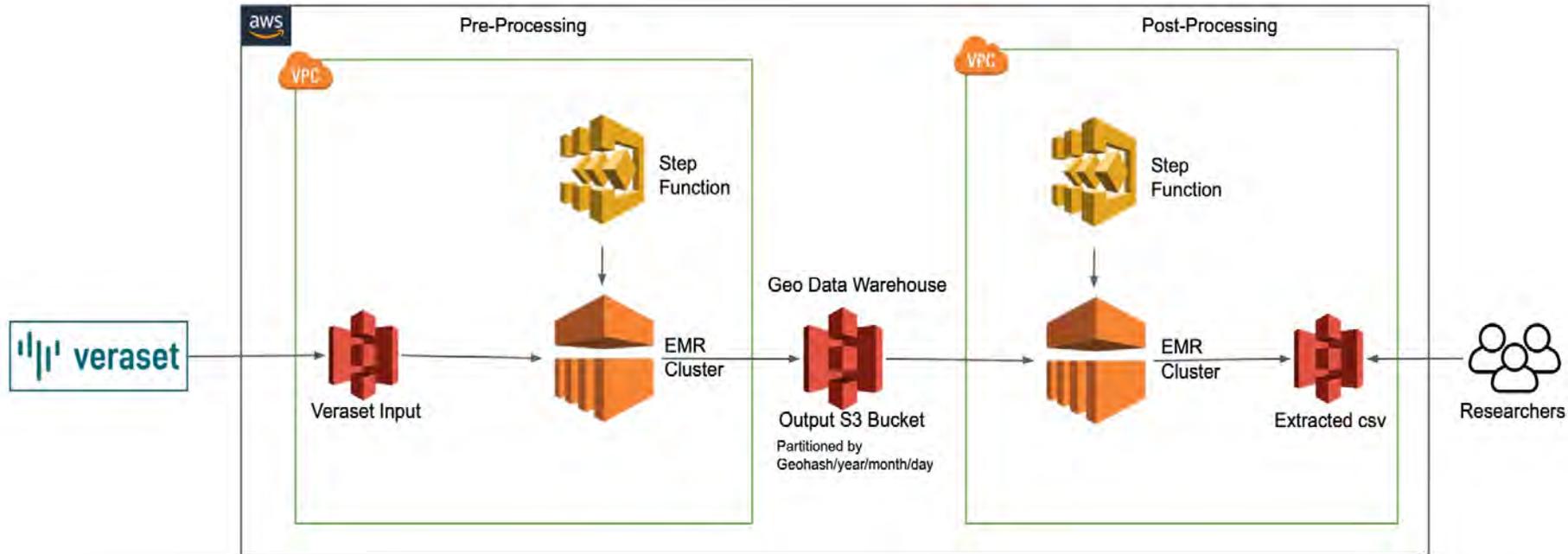
Immediate Imperative - Post Processing

Post Processing



-  **Input Shapefiles**
Fetch relevant information from the Geo Data Warehouse based on the researcher's requirements and the input provided in the shapefiles
-  **Single Pass**
Utilize the shape files' latitude and longitude to filter the data
-  **Double Pass**
Records with matching CAID data from single pass is filtered and extracted
-  **GEO Spatial Analysis**
One CSV file comprising all of the extracted data is created, offering a thorough output for additional analysis using programs like Geoda, STATA and ArcGis.

The solution we built



- The preprocessing phase is estimated to take approximately 30 minutes for a dataset of 110GB.
- To optimize cost and streamline the workflow, intermediate steps will be removed.
- EMR cluster will be terminated once the process is completed.
- Same EMR Cluster and Steps can be used for incremental quarterly data.
- Eliminated 95% of manual intervention.
- Improved efficiency of the post-process by organizing the output data in a hierarchical folder structure.



Putting the data into action

SCAER's research directly aligns with SLU's broader mission of serving humanity. Some use cases that we extensively use this data for:

- By investigating human mobility patterns during natural disasters, SCAER researchers aim to better understand how communities respond and adapt to such events.
- They are leveraging mobility data to provide a novel perspective on accessibility to dialysis centers for patients with end-stage renal disease (ESRD)
- They are utilizing this human mobility data to study food insecurity, by tracking movement patterns, we can identify regions where populations frequently travel longer distances to access food, indicating potential "food deserts" or areas lacking affordable, nutritious food options.
- Researchers at the TGI have used the mobility spend data to examine the effect of MLB games on spending patterns and visits to various businesses in the St. Louis metro area.

Q&A

Any questions?

THANK YOU!



A dramatic night sky with dark, heavy clouds and several bright, jagged lightning bolts striking downwards. The lightning is a mix of white and blue, illuminating the surrounding clouds. The bottom of the image shows a dark silhouette of a treeline or forest.

PART THREE

Unlocking Research Potential on Google Cloud

Ezequiel Gioia

University of Central Florida



Unlocking Research Potential on Google Cloud

Internet2 - Community Exchange 2024

Ezequiel Gioia

Systems Architect

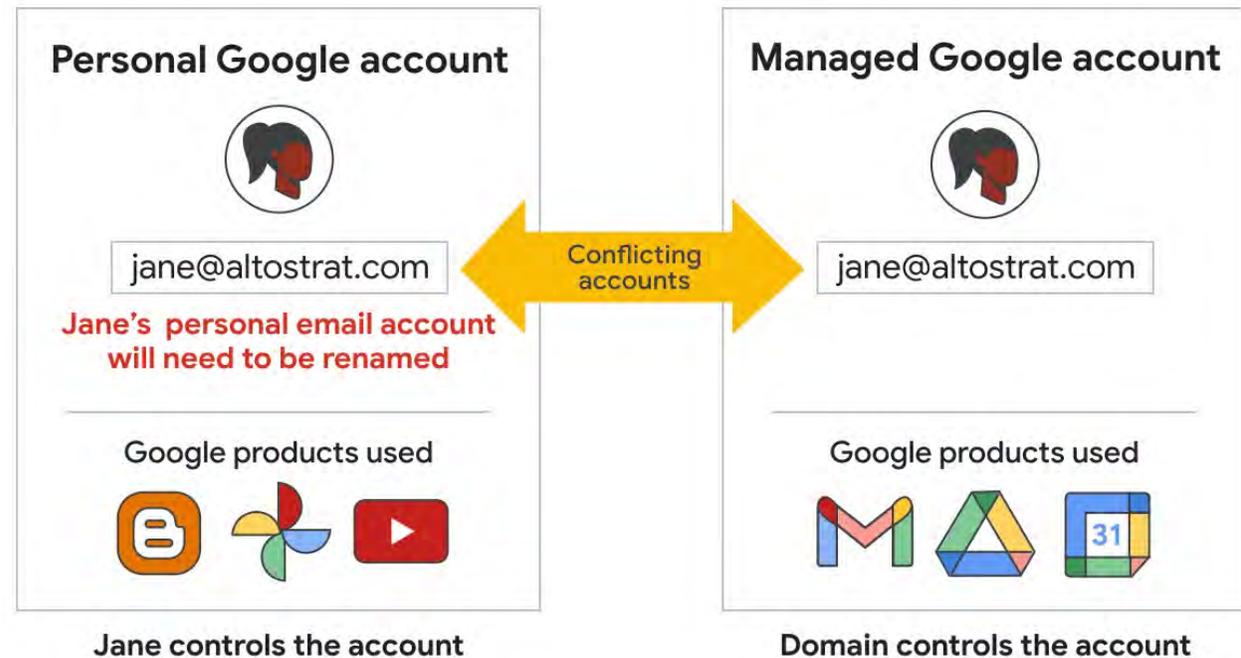
University of Central Florida

INTRODUCTION

- Why another cloud?
- Getting the contract in place
- Choosing a reseller

GCP IMPLEMENTATION AND CHALLENGES

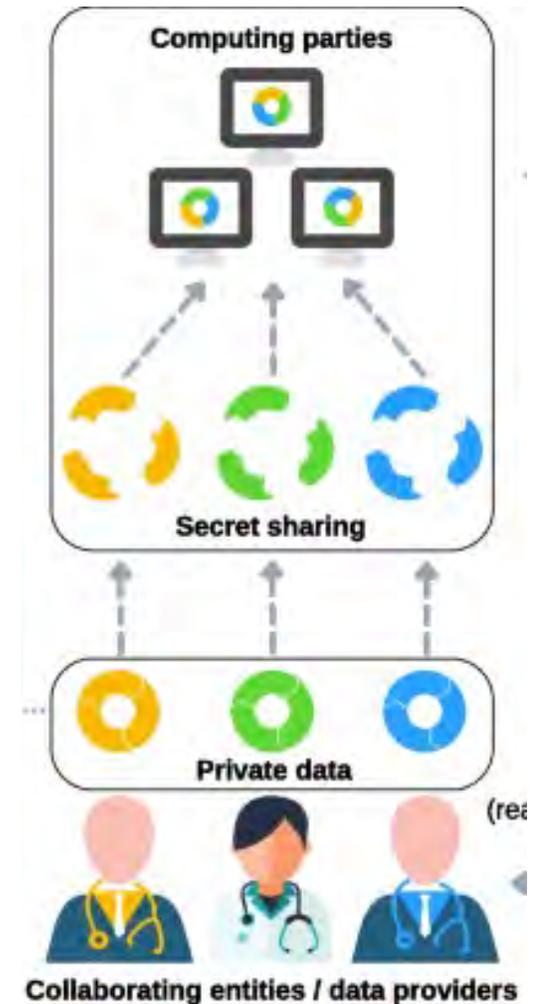
- Accessing Google Cloud without Workspace
- Abandoned projects?
- Conflicting unmanaged users accounts?



Source: google.com

PROJECT - MPC

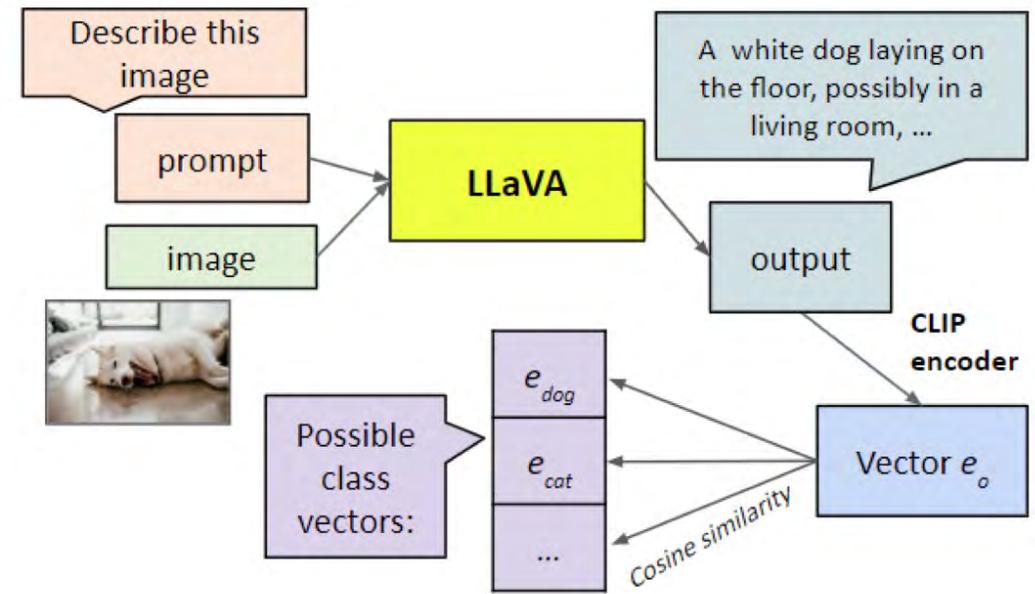
- Multi-party Computation (MPC) for Fairness
- Protocol for Collaborative Analysis
- Data Privacy
- Multi-university project to analyze data without revealing individual responses



Source: [biomedcentral.com](https://www.biomedcentral.com)

PROJECT - Visual-LLM Zero-Shot Classification

- Center for Research in Computer Vision
- Generate a rich set of descriptions and analyses for all aspects of a given image
- Combining a vision encoder and a transformer-based language model
- Using Tensor Processing Units (TPUs)



Source: Matsuura, M., Jung, Y. K., & Lim, S. N. (2018).
Visual-LLM zero-shot classification

PROJECT - Computational Radiology for Breas

- First HIPAA project on Google Cloud
- Mammogram false positives lead to:
 - Patient anxiety
 - Unnecessary biopsies and medical procedures
- Goal: Enhance breast cancer screening through accurate detection
- Deep learning algorithms will be trained on 5 million digital mammograms



Dexter Hadley

Assistant Professor of Pathology

<https://www.hadleylab.org>

FUTURE PLANS AND VISION

- Interconnect
- Logs centralization
- STRIDES, in partnership with the College of Medicine
- Training
- Achieve the same level of support we have on AWS



PART FOUR

**Automated Transcription
for Social Science Researchers**

Alan Walsh

Indiana University



INDIANA UNIVERSITY BLOOMINGTON

Automated Transcription for Social Science Researchers

Alan Walsh



A longstanding challenge in social science research



Gathering Qualitative Data

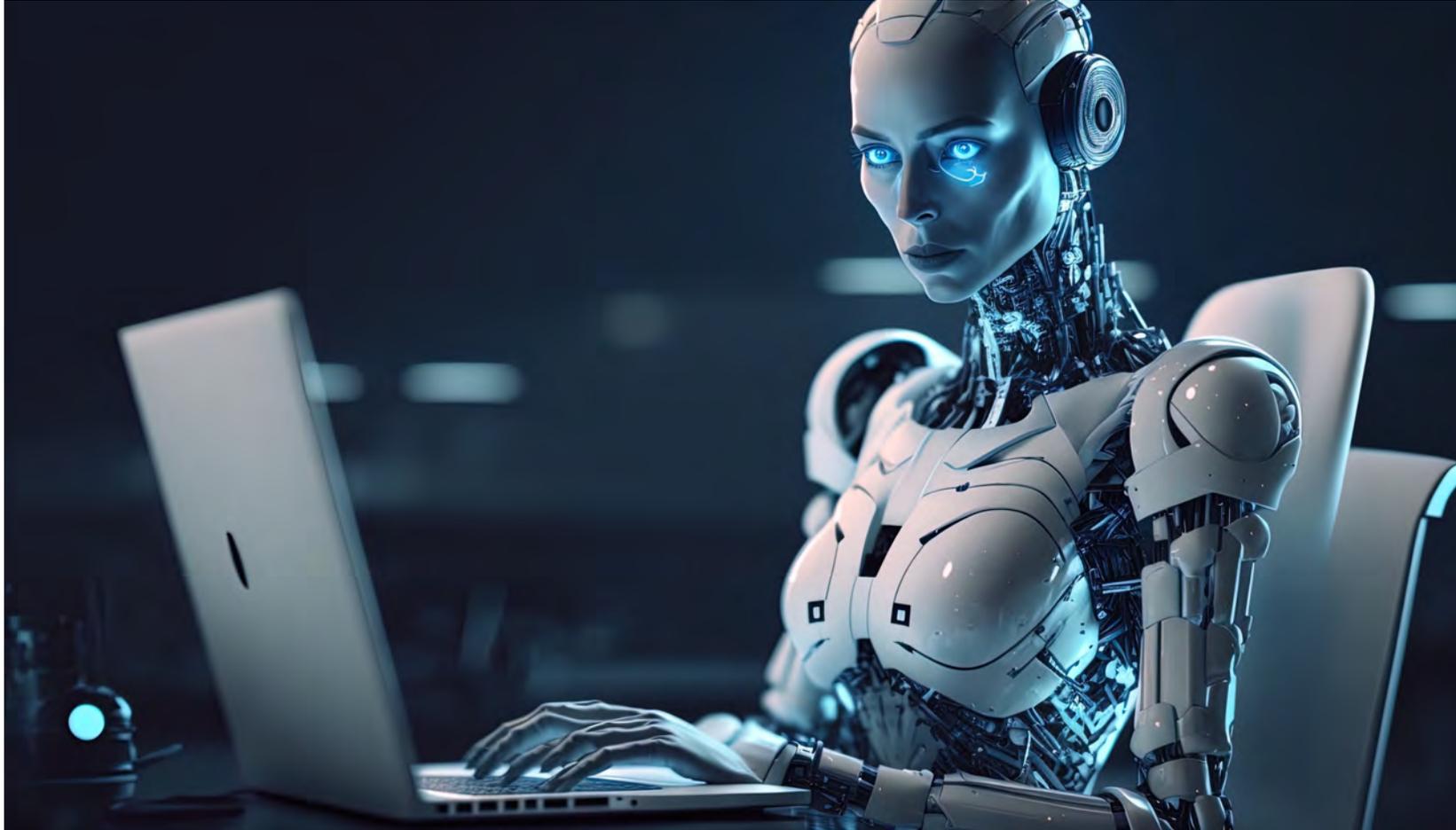
Interviews, focus groups,
and more

Audio Recording





Generating text transcripts for analysis



| New technology, new possibilities

Challenges



Quality



Cost



Languages



Security



Usability

Specify job details [Info](#)

Job settings

Name

MyTranscriptionJob

The name can be up to 200 characters long. Valid characters are a-z, A-Z, 0-9, . (period), _ (underscore), and - (hyphen).

Language settings

You can transcribe your audio file in a language that you specify or have Amazon Transcribe identify and transcribe it in the predominant language.

- Specific language** [Info](#)
If you know the language spoken in your source audio, choose this option to get the most accurate results. The options available for additional processing vary between languages.
- Automatic language identification** [Info](#)
If you don't know the language spoken in your audio files, choose this option. You have access to fewer options for additional processing than if you choose **Specific language**.
- Automatic multiple languages identification** [Info](#)
If there are multiple languages spoken in your audio files and you're not sure what these languages are, choose this option. This selection provides limited additional processing options compared to **Specific language**.

Language options for automatic language identification - *optional*

To improve accuracy, choose at least two languages spoken the most often in your audio library. Amazon Transcribe chooses from one of the languages you've specified to transcribe each audio file. Leave this field empty if you're unsure about which languages to select.

Select languages

AWS Web Interface

Create transcription jobs for one file at a time

```
{"jobName":"BOHP_Broide_Mace_2021-09-17
_excerpt1a.mp3-55e6dc7e-3eae-4612-9793-9177b10d5fbc","accountId":"445593358808","results":{"transcripts"
:[{"transcript":"This is Danielle Emmerling. I am, it is September 17, 2011. I am conducting an oral
history, uh, for the I U School of Journalism. Could you please tell me your full name, the year you
graduated? And what discipline your degree was in? My name is Mace. Hi, Brody. I graduated in 1947. A
degree in journalism and government. Where was your hometown? And what brought you to I U, my home cat
was brewing in Vermont and I came to I U because I had an opportunity to go to a good journalism
education and at a reasonable rate tuition in those days was very reasonable even for an out stayer, $75
a semester. And I got a scholarship for half of that."},"speaker_labels":{"channel_label":"ch_
0","speakers":2,"segments":[{"start_time":"0.4","speaker_label":"spk_
0","end_time":"13.479","items":[{"start_time":"0.409","speaker_label":"spk_
0","end_time":"0.699"},{"start_time":"0.709","speaker_label":"spk_
0","end_time":"0.879"},{"start_time":"0.889","speaker_label":"spk_
0","end_time":"1.539"},{"start_time":"1.61","speaker_label":"spk_
0","end_time":"2.059"},{"start_time":"2.809","speaker_label":"spk_
0","end_time":"3.0"},{"start_time":"3.009","speaker_label":"spk_
0","end_time":"3.14"},{"start_time":"3.15","speaker_label":"spk_
0","end_time":"3.569"},{"start_time":"3.579","speaker_label":"spk_
0","end_time":"3.94"},{"start_time":"3.95","speaker_label":"spk_
0","end_time":"4.86"},{"start_time":"4.869","speaker_label":"spk_
0","end_time":"6.94"},{"start_time":"7.25","speaker_label":"spk_
0","end_time":"7.44"},{"start_time":"7.449","speaker_label":"spk_
0","end_time":"7.599"},{"start_time":"7.61","speaker_label":"spk_
0","end_time":"8.22"},{"start_time":"8.229","speaker_label":"spk_
0","end_time":"8.319"},{"start_time":"8.329","speaker_label":"spk_
0","end_time":"8.71"},{"start_time":"8.72","speaker_label":"spk_
0","end_time":"9.609"},{"start_time":"10.0","speaker_label":"spk_
0","end_time":"10.01"},{"start_time":"10.02","speaker_label":"spk_
0","end_time":"10.479"},{"start_time":"10.489","speaker_label":"spk_
0","end_time":"10.939"},{"start_time":"10.949","speaker_label":"spk_
0","end_time":"10.96"},{"start_time":"10.97","speaker_label":"spk_
0","end_time":"11.1"}, {"start_time":"11.109","speaker_label":"spk_
0","end_time":"11.68"}, {"start_time":"11.689","speaker_label":"spk_
0","end_time":"11.71"}, {"start_time":"11.72","speaker_label":"spk_
0","end_time":"12.52"}]}],{"start_time":"14.85","speaker_label":"spk_
0","end_time":"21.709","items":[{"start_time":"14.979","speaker_label":"spk_
0","end_time":"15.409"}, {"start_time":"15.42","speaker_label":"spk_
```

AWS JSON Output

Text, transcripts,
speaker
identification, and
word confidence
information



| How it works



Share files effortlessly

At Indiana University, Secure Share provides everyone a way to securely share data with others via a web interface.

Your Secure Share files will automatically disappear 30 days after you upload them (but you may delete them sooner). We maintain backups of files uploaded to Secure Share; when you delete a file, there is no way to recover it. Do not use Secure Share for files you cannot afford to lose.

Upload a file

Create encrypted text file

Create

After your upload is complete, you will be able to access documents and send invitations to download.

Select file

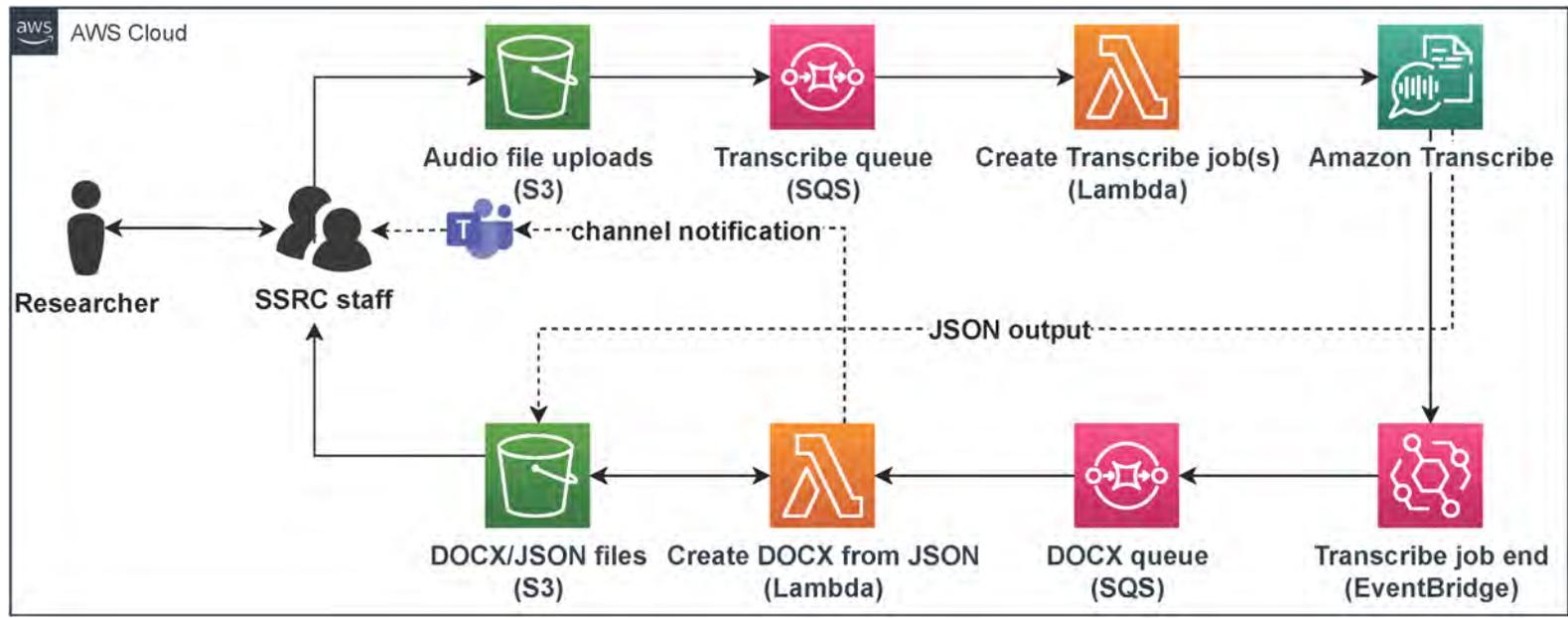
No file selected

Description (optional)

Encryption password

Secure file transfer

IU Secure Share
Secure Teams



JSON to DOCX (Word)

AWS Sample Python3 Script for CLI

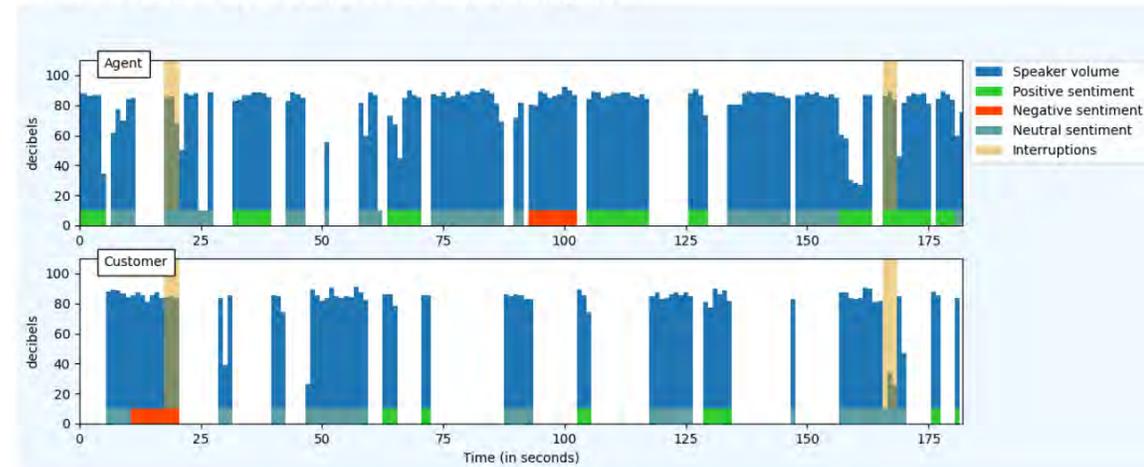
<https://github.com/aws-samples/amazon-transcribe-output-word-document>

Removed statistics

Changed formatting

Added MS Teams notification

Conversation Volume Levels with Sentiment and Interruptions



0:02:14.84	11.6s	Agent	😊	Okay I've got that and [ACTION] I will email you a confirmation is [PII]. Still the email address for you.
0:02:27.24	0.4s	Customer	😐	correct?
0:02:28.14	8.2s	Agent	😊	Okay great so to [OUTCOME] review we've got that card canceled, we're sending it to the [PII] and I've emailed you a confirmation. Is that everything that you needed today?

Amazon Transcribe Audio Source

Job Name	GMT20221117-211655_Recording.mp4-b9af90db-9072-48ec-98b4-daa4cd486f0b
Audio Duration	0m 55.58s
Audio Identification	Speaker-separated
Language(s)	en-US
File Format	mp4
Sample Rate	32000 Hz
Job Created	Thu 16 Nov '23 at 22:12:15
Average Confidence	97.92%

Audio Transcription

WORD CONFIDENCE: >= 90% in black, < 90% in yellow highlight

[00:00:02] Speaker 1: OK, we're recording. Um How about you? Read the first paragraph and I'll read the second one.

[00:00:13] Speaker 2: OK. Become the best version of yourself at I U. Develop the skills you need for rewarding career. And follow in the footsteps of more than 760,000 I U alumni who are leading organizations innovating the future and making an impact on their communities and the world.

[00:00:33] Speaker 1: Pursue your passion or explore the possibilities from nursing to criminal justice, epidemiology to environmental science, social work to cybersecurity and more than 900 academic programs in between I U gives you the power to study whatever sparks your interest with the support of world class professors and faculty who care about your success.

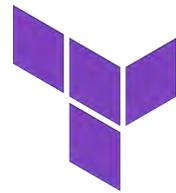
Word Confidence Scores

Confidence	Count	Percentage
98% - 100%	113	95.76%
90% - 97%	2	1.69%
80% - 89%	0	0.0%
70% - 79%	1	0.85%
60% - 69%	0	0.0%
50% - 59%	1	0.85%
40% - 49%	1	0.85%
30% - 39%	0	0.0%
20% - 29%	0	0.0%
10% - 19%	0	0.0%
0% - 9%	0	0.0%

ATS output



100% serverless



HashiCorp

Terraform

Automated Transcription Service



Quality: Accuracy, speaker recognition (diarization), time stamps, word confidence scores



Cost: \$0.024 per audio minute (plus minimal storage/processing costs and staff time); **free to IU researchers for the next two years**



Languages: 39 languages/variations (Amazon); additional languages supported by Google Cloud Speech-to-Text



Security: With review and approval of IU Data Stewards (existing contracts and Business Associate Agreements make this possible)



Usability: Automated workflows implemented by UITS RT; researchers submit audio to SSRC; SSRC returns Word transcripts



Thank you!



INTERNET²

2024
COMMUNITY
exchange

THANKS!