AWS FAQs

Answers to questions about AWS that we often see at Cornell.

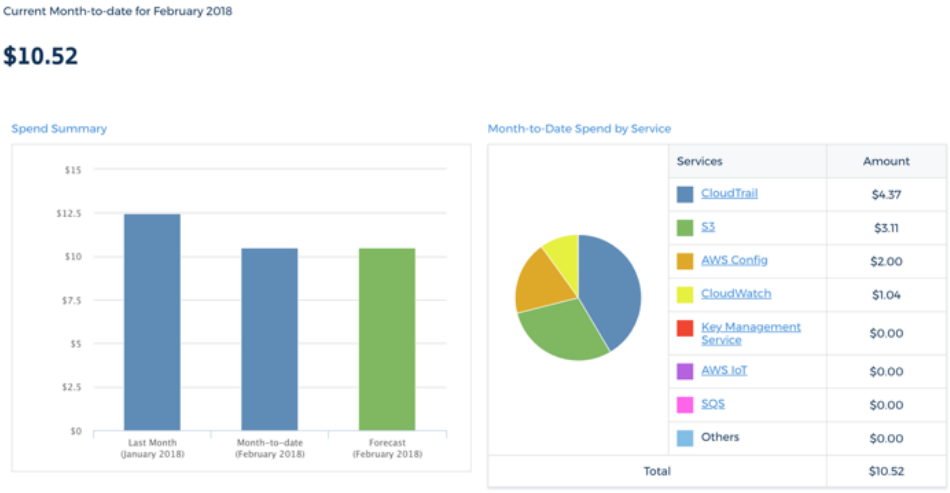
AWS Blended Pricing

Costs

What is the cost of the the Cornell standard AWS configuration?

The cost of Standard AWS Account Configurations can be broken into two parts:

- Basic auditing configurations (i.e., CloudTrail, AWS Config) for Cornell AWS accounts
  - This costs in the ballpark of about $11/month, even if you aren't actively using the AWS account.
    - Graph of Typical Costs for auditing configurations



- Cornell standard VPC and its associated NAT gateway
    - The total with additional costs for the VPC and NAT gateway are in the ballpark of about $40/month, even if you aren't actively using the AWS account. A big part of that is the NAT gateway which costs about $1/day.
    - You will see this larger baseline cost only if the Cloud Team has configured a Cornell Standard VPC.
    - AWS accounts used for research **generally don't** have a Cornell Standard VPC configured so their monthly costs are on the order of $11 / month.

Why do I see charges from AWS services in other parts of the world?

The auditing configuration created as part of our Standard AWS Account Configuration includes setting up auditing in all AWS regions. This is to log activity in an AWS account targeting other AWS regions. These costs should be minimal because most folks aren't using regions across the globe. However, it is important that the auditing configuration is in place globally to ensure that any malfeasance in other regions is logged.

What is the AWS Data Egress Waiver?

AWS offers a Data Egress Waiver that mostly eliminates Data Transfer charges for Cornell. See the AWS blog post about it for more details.

Is egress from all AWS offerings covered by the waiver?

As of February 2018, most services that advertise the $0.09/GB data egress fee on the pricing page are covered. Feel free to check with the Cloud Team if you aren't sure about a particular service.

One notable exception is CloudFront egress, which is **not** covered by the waiver. CloudFront egress shows as a separate line-item and is billed at a (slightly) cheaper rate than normal data egress. AWS is looking into adding CloudFront egress to the waiver. Any campus customers who have a current or future plan to generate significant CloudFront egress charges should forward their use-case to the Cloud Team for a consult with our AWS Representatives.

A second notable exception is Direct Connect egress, which is **not** covered by the waiver. Egress costs for our Direct Connect are $0.02/GB and those charges are billed to your AWS account.

According to AWS, their service offerings that use CloudFront under the covers, API Gateway for example, **are** covered by the egress waiver and a separate consult is not needed.

How much are Data Transfer Charges for Cornell? Are we close to the 15% cap for the AWS Data Egress Waiver?

In short, no, Cornell is not near the 15% cap of the Data Egress Waiver. As of March 2023, the last three months averaged 6.83% utilization and the last six averaged 6.16%. If more detailed information is needed, please contact cloud-support@cornell.edu.

Billing

I got an invoice for AWS from CloudCheckr. What should I do with it?

We use CloudCheckr to send informational invoice for each AWS account on the 15th of each month. Those invoices cover the prior calendar month. You do not need to do anything with those invoices.

On or about the 17th of each month, a KFS transaction is automatically created that charges the designated default KFS account for the charges in your AWS account. That KFS transaction also has a copy of the invoice for backup purposes.

Please contact cloud-support@cornell.edu if you need to change the KFS account used by default for your AWS account charges.

If you'd like to target the charges for specific AWS resources to specific KFS accounts, you can "Cost Center" tags to your AWS resources. See AWS Standard Tagging for details.

How come my AWS bill contains charges for EC2 when I haven't used EC2 at all?

In most cases, the EC2 charge you are seeing is a result of the standard configuration we use in your VPC. The private subnets in your Cornell standard VPC are connected to the world (for outgoing traffic) by a NAT gateway. That NAT gateway is really a small EC2 instance, though it won't appear in your EC2 instance list in the AWS Console. You can see the NAT gateway(s) configured for your account here: https://console.aws.amazon.com/vpc/home?region=us-east-1#NatGateways:sort=desc:createTime

The NAT gateway gives EC2 instances in your private subnets access to the world for things like Linux repos or Windows update servers. We do have some AWS account owners that do not find the $1/day cost of the NAT gateway to be worthwhile and turn it off. We advise caution around this because, with it off, your instances will not be able to do something as basic and critical as running "yum update" or "apt-get update" or get Windows updates.

Contact the Cloud Team if you'd rather not have the NAT gateway deployed for that VPC.

See NAT gateway pricing info here: https://aws.amazon.com/vpc/pricing/.

When will direct billing (though KFS) based on "Cost Center" tags be released?

Direct billing from AWS to KFS is now enabled. We have a default KFS account to bill for charges in each AWS account. If you'd like to target the charges for specific AWS resources to specific KFS accounts, you can "Cost Center" tags to your AWS resources. See AWS Standard Tagging for details.

How do I buy EC2 Reserved Instances?

As of June 2022, individual Cornell AWS accounts cannot buy Reserved Instances or Savings Plans. Cornell has a program that purchases those centrally. For more information, please contact the Cloud Team.

Licensing

Does the Cornell Microsoft Agreement cover Microsoft software in AWS?

In most cases, no. See Microsoft Licensing within AWS.

Users, Policies and Roles

How can I give Cornell users access and privileges to my AWS account?

You can create custom IAM roles that integrate with the Cornell Shibboleth so that access to those roles is granted according membership in an AD group. See Creating Custom Roles to use With Shibboleth.

Can I use a Holiding ID, Guest ID, or DOC (delegation of control) account to login to AWS?

No. Our Shibboleth implementation does not work with DOC accounts, Holding IDs, or Guest IDs. (More info about various Cornell account types: https://it.cornell.edu/cornellad/terms-and-conditions-cornellad and Cornell's Shibboleth Implementation.)

Please contact cloud-support@cornell.edu to discuss other options.

Networking

I deleted my "default" AWS VPC. How do I get it back?

You can create a new one. See https://aws.amazon.com/about-aws/whats-new/2017/07/create-a-new-default-vpc-using-aws-console-or-cli/.

Will AWS designate an existing VPC as the "default" VPC?

Direct from AWS tech support, here's what they have to say about this (as of 2017-02-08):

> *...existing VPC's can not be assigned as the default and we can only create a new Default VPC for you.*
>
> *Please note that when we create a default VPC, we do the following to set it up for you:*
>
> - *Create a default subnet in each Availability Zone.*
> - *Create an Internet gateway and connect it to your default VPC.*
> - *Create a main route table for your default VPC with a rule that sends all traffic destined for the Internet to the Internet gateway.*
> - *Create a default security group and associate it with your default VPC.*
> - *Create a default network access control list (ACL) and associate it with your default VPC.*
> - *Associate the default DHCP options set for your AWS account with your default VPC.*

What is AWS Direct Connect and how does Cornell use it?

See Cornell AWS Direct Connect.

What is the Cornell Standard VPC?

See The Cornell “Standard” AWS VPC.

Why can't I connect to my EC2 instance?

You might want to look at the diagrams on Cornell AWS Direct Connect Routing Diagrams

Can I coordinate VPC Availability Zones between AWS accounts?

In short, yes. To ensure distribution of load across their infrastructure, AWS creates an independent mapping of Availability Zone designations (ie: "us-east-1a", "us-east-1d") for each account. Within the same Region, if you need to guarantee the Availability Zone that *you* see as "zone A" lives in the same back-end environment as "zone A" seen from a different AWS account you will need to utilize the Availability Zone ID. For more information about zones and regions, see the AWS documentation on Regions and Availability Zones.

How can I request a cucloud.net subdomain for use in Route 53?

The process for creating a **cucloud.net** Hosted Zone in your AWS account and requesting DNS delegation can be found in Route 53 Subdomain Delegation.

How can confirm that a peering connection is being used for 10-space traffic instead of the Direct Connect.

Suppose I have two AWS VPCs that are both connected to campus networks with Direct Connect. How can I tell if traffic between those two VPCs are using an AWS peering connection or traveling the Direct Connect and making a u-turn on-campus? The answer is to look at the results of `traceroute` between two IPs, one in each VPC.

Here is the traffic pattern `traceroute` would return when Direct Connect is being used

traceroute results with Direct Connect show on-campus nodes involved in the Direct Connect

```
> traceroute 10.92.131.194
traceroute to 10.92.131.194 (10.92.131.194), 30 hops max, 60 byte packets
 1  * * *
 2  aws1-mx-vl3302.net.cornell.edu (10.22.223.4)  10.856 ms  10.799 ms  10.741 ms
 3  aws-bgp-vl3334.net.cornell.edu (10.22.223.85)  10.722 ms  10.676 ms  10.629 ms
 4  * * *
 5  * * *
 6  * * *
 7  * * *
 8  * * *
```

Here is the traffic pattern `traceroute` would return when AWS traffic is traversing peered VPCs:

traceroute shows just a single hop when traffic uses a peering connection

```
> traceroute 10.92.168.117
traceroute to 10.92.168.117 (10.92.168.117), 30 hops max, 60 byte packets
 1  10.92.168.117 (10.92.168.117)  5.174 ms  5.201 ms  3.095 ms
```

If there is any chance that Network ACLs or Security Groups are blocking ICMP traffic, you can use the TCP (-T) and port (-p) switch with {{traceroute}. The example below proves that the instance where the `traceroute` is run in in a VPC that is peered directly to the VPC containing the AWS Active Directory server `ad10.cornell.edu`. Note that you will need to pick a port that you know to be open for the target system. This example uses port 389 because the Active Directory server has port 389 (LDAP) open.

```
$ traceroute -T -p 389 ad10.cornell.edu
traceroute to ad10.cornell.edu (10.92.36.80), 30 hops max, 60 byte packets
 1  ip-10-92-36-80.ec2.internal (10.92.36.80)  7.740 ms  7.711 ms  9.136 ms
```

Do I need multiple NAT Gateways?

VPCs created by the Cloud Team for Cornell AWS customers generally contain only a single NAT Gateway. This NAT Gateway provides access to the public internet for private subnets in the VPC. All private subnets in the VPC are configured to use the same NAT Gateway, regardless of the Availability Zone of the private subnet. This means that the NAT Gateway is a single point of failure because the resources in your private subnets may not be able to reach the internet if the AZ where the NAT Gateway resides experiences network issues.

If you require high availability and resiliency for the deployments in your private subnets, you may want to consider adding additional NAT Gateways to your VPC. You would want one NAT Gateway in each Availability Zone where your private subnets reside.

The downside of multiple NAT Gateways is that each one costs about $1/day to run, and some Cornell AWS customers do not consider the high availability worth that cost.

Email cloud-support@cornell.edu if you'd like help setting up additional NAT Gateways in your Cornell AWS account.

Working with Data

When should I use Direct Connect and when should I use the public internet to transfer data?

Direct Connect is mostly useful when a reliable latency is needed to be maintained between systems on campus and in AWS. Another use case could be that you are required to use a private network due to some policy, or you must access a system on campus that will not allow access via the public internet due to firewall rules that cannot be changed or because the system is only in campus 10-Space.

In the majority of other scenarios, the Cloud Team recommends using the public internet to transfer all data and updating firewall configurations to allow access to/from the internet with trusted systems that you run in AWS. The available bandwidth to the internet is much greater than the 1Gbps Direct Connect that is shared among many units at Cornell.

We also recommend using end-to-end encryption whenever transferring data over the internet. If you are using AWS provided CLI or SDKs (or 3rd party tools that utilize these) to transfer data to AWS, your connections will be encrypted by default.

How do I transfer a large file (>1GB) to Amazon S3?

Amazon S3 supports individual objects up to 5TB in size. However, when uploading large files, you run the risk of that transfer being interrupted and having to start over. Each individual connection to S3 also only gets 100Mbps from AWS.

We recommend using the AWS CLI or a 3rd party tool to utilize "multipart uploads" when transferring large files. Most tools also multithread when uploading the parts of your file, so you will be able to utilize the full bandwidth of your machine (usually 1Gbps on campus).

The following tools support multipart uploads:

- S3 Browser (Pro Only)- Windows
- CloudBerry Explorer (PRO only) - Windows
- Cyberduck 4 - macOS / Windows
- s3cmd (version 1.1.0-beta2 +) - Linux / macOS (plus others)
- aws s3 - AWS CLI - Linux/macOS/Windows
- rclone - Linux/macOS/Windows

STS Token use for manual data transfers with existing shibboleth IAM roles

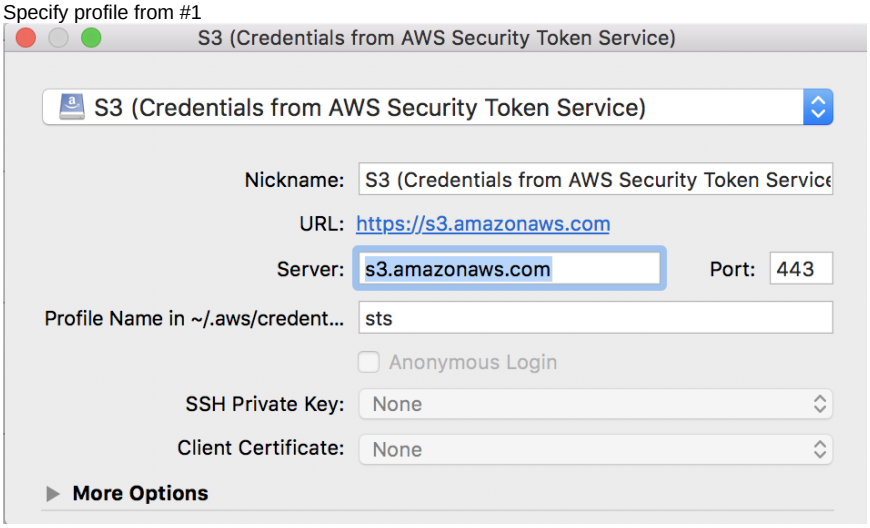
There are some options here:

1. Install the aws login tool (Access Keys for AWS CLI Using Cornell Two-Step Login - Shibboleth
2. Docker with the aws login tool with other helpful cloud utilities (https://github.com/CU-CommunityApps/ct-cloud-utils-dockerized)
3. Install the aws cli (https://docs.aws.amazon.com/cli/latest/userguide/cli-chap-install.html) using 'aws sts get-session-token' with a new or existing IAM user (https://docs.aws.amazon.com/cli/latest/reference/sts/get-session-token.html)

    a. Create a new or use default profile
    b. "aws configure --profile {name}"

- AWS CLI
    1. Access Keys for AWS CLI Using Cornell Two-Step Login - Shibboleth
- rclone
    1. rclone config
        a. set id, secret and session token (under advanced config)
- Cyberduck
    1. Copy ID, Secret and Token from ~/.aws/credentials {name}
        a. aws_access_key_id = [paste ID]
        b. aws_secret_access_key = [paste key]
        c. aws_session_token = [paste token]
    2. Download Cyberduck STS token profile
    3. Open Connection - S3 (Credentials from AWS Security Token Service)

a. Specify profile from #1



- Mountain Duck now available with similar process as outlined above with CyberDuck.

Mechanical Turk (MTurk)

Can I use Mechanical Turk with my Cornell AWS account?

- Mechanical Turk requester accounts can use the same email address and password as AWS root accounts.
  However, in order to keep these concerns separate, we recommend using different email accounts for each of AWS, Amazon.com retail store, and Mechanical Turk.

- As of December 2020, MTurk accounts can be linked to AWS accounts for billing purposes.
  MTurk accounts linked like that have their charges included in the charges for the AWS account. Please contact cloud-support@cornell.edu to link your MTurk account.
  - With this linkage, research awards/credits issued by AWS to an AWS account can be used for paying MTurk charges.
  - **Only one MTurk account can be linked to each AWS account.**
  - In order to establish this linkage, the root credentials for the AWS account must be used.
    If the Cloud Team manages the root credentials for an AWS account, we will be happy to help establish this linkage. Please contact cloud-support@cornell.edu.

- **As of August 2023 ,** New MTurk Requester accounts that are created and linked to AWS accounts , the Requester UI is ONLY available using the ROOT login for the account , not another email account.
  - One can specify on the MTurk Account page a Requster name and alternate email for Contact by workers , and a Display Name .
  - One can still use an SDK / CLI script management of the Requester using User Access Keys in the AWS Account that have Mechankical Turk permissions thru attached policies.
    - https://docs.aws.amazon.com/AWSMechTurk/latest/AWSMechanicalTurkGettingStartedGuide/SetUp.html#toolsinstallation
  - Alternatively one can use a 3rd party solution like CloudResearch that can provide a UI and use AccessKeys to manage the MTurk Requester functions; Creting HITs, setup Sandbox, and interact with Workers, vet workers etc.

Can I use tagging in Mechanical Turk?

As of 21 Jul 2022 Mechanical Turk resources (projects, etc.) **cannot be tagged.**
Therefore it is not possible to use "Cost Center" tagging to direct MTurk charges to KFS accounts other than the default KFS account configured for the AWS account.

RDS

How is the OS hosting my RDS patched?

RDS is a fully-managed service at Amazon, meaning you do not have access to the underlying operating system. Patching to the database engine or underlying operating system is handled as a scheduled maintenance event:

http://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/USER_UpgradeDBInstance.Maintenance.html

Maintenance to RDS can happen "immediately" or in a regularly-scheduled 30-minute window. An RDS instance can also be configured to automatically apply "minor" updates during the standing maintenance window without prior approval. For instances supported by the CIT DBA team, they can likely give you the details on your current configuration(s); the Cloud Team can also help fill in gaps if needed.

It is certainly likely that this activity will result in a brief outage. Use of a Multi-AZ deployment for RDS helps mitigate that for most activities since AWS will first perform maintenance on the standby instance, promote that to primary, then perform maintenance on the "old primary/new standby". Emphasis is on

"most" because some major changes, like modifying the database engine, may require updating both primary and standby at the same time. Multi-AZ also gives you increased availability in the case of an Availability Zone going offline. Keep in mind that Multi-AZ deployments come with additional cost. You will need to find the balance between desired availability and total monthly spend.

Web Hosting

What are my options for hosting a web site in AWS?

While we can help you setup almost any type of web site in AWS, your best bet may be to start with the CIT Custom Web Development team, which brokers a varierty of web site hosting options. See Hosting Comparison Chart for options. In short the options are:

Basic
Static Web Hosting

Content Management Systems
CampusPress (WordPress)
Pantheon (WordPress and Drupal)
Acquia (Drupal)

Highly Customizable
Media3 (ColdFusion, LAMP)
Amazon Web Services (AWS) (CIT Cloud Services)
Managed Server

Miscellaneous

What kind of email does AWS send to the root account email address or the security contact address configured in an AWS account?

See Examples of Email Sent to AWS Root Account Addresses and AWS Security Contacts