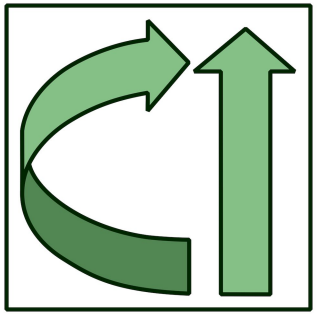


CILogon

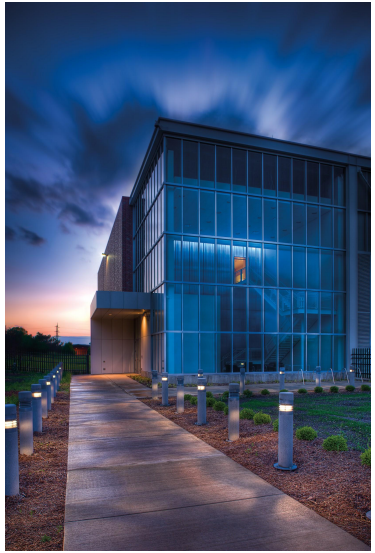
Enabling Federated Identity and Access Management for Scientific Collaborations

Jim Basney, Terry Fleury,
Jeff Gaynor, and Scott Koranda

June 28 2022 - AWS Lunch & Learn

NCSA

- National Center for Supercomputing Applications
- Established in 1986 as one of the original sites of the NSF's Supercomputer Centers Program
- A department of the University of Illinois at Urbana-Champaign
- Supported by the state of Illinois, the University of Illinois, the National Science Foundation, and other federal agencies
- Lead institution of the Extreme Science and Engineering Discovery Environment (XSEDE)



https://www.ncsa.illinois.edu/

XSEDE
Extreme Science and Engineering Discovery Environment

IAM for Research Collaborations

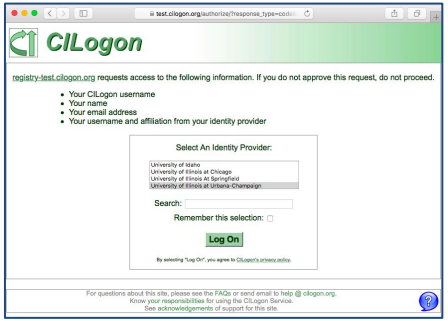
CILogon: 10+ year sustained effort to enable secure logon to scientific cyberinfrastructure (CI)

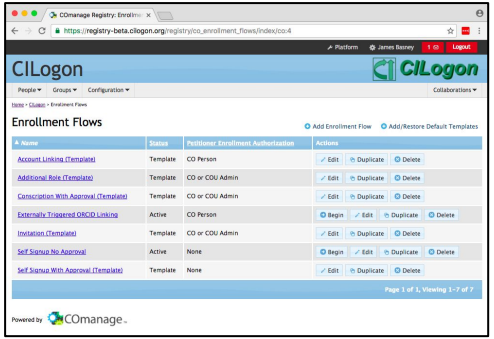
for seamless identity and access management (IAM)

using federated identities (SAML, OIDC, OAuth, JWT, X.509, LDAP, SSH, etc.) so researchers log on with their existing credentials from their home organization

supporting 17,500+ active users from 450+ organizations around the world

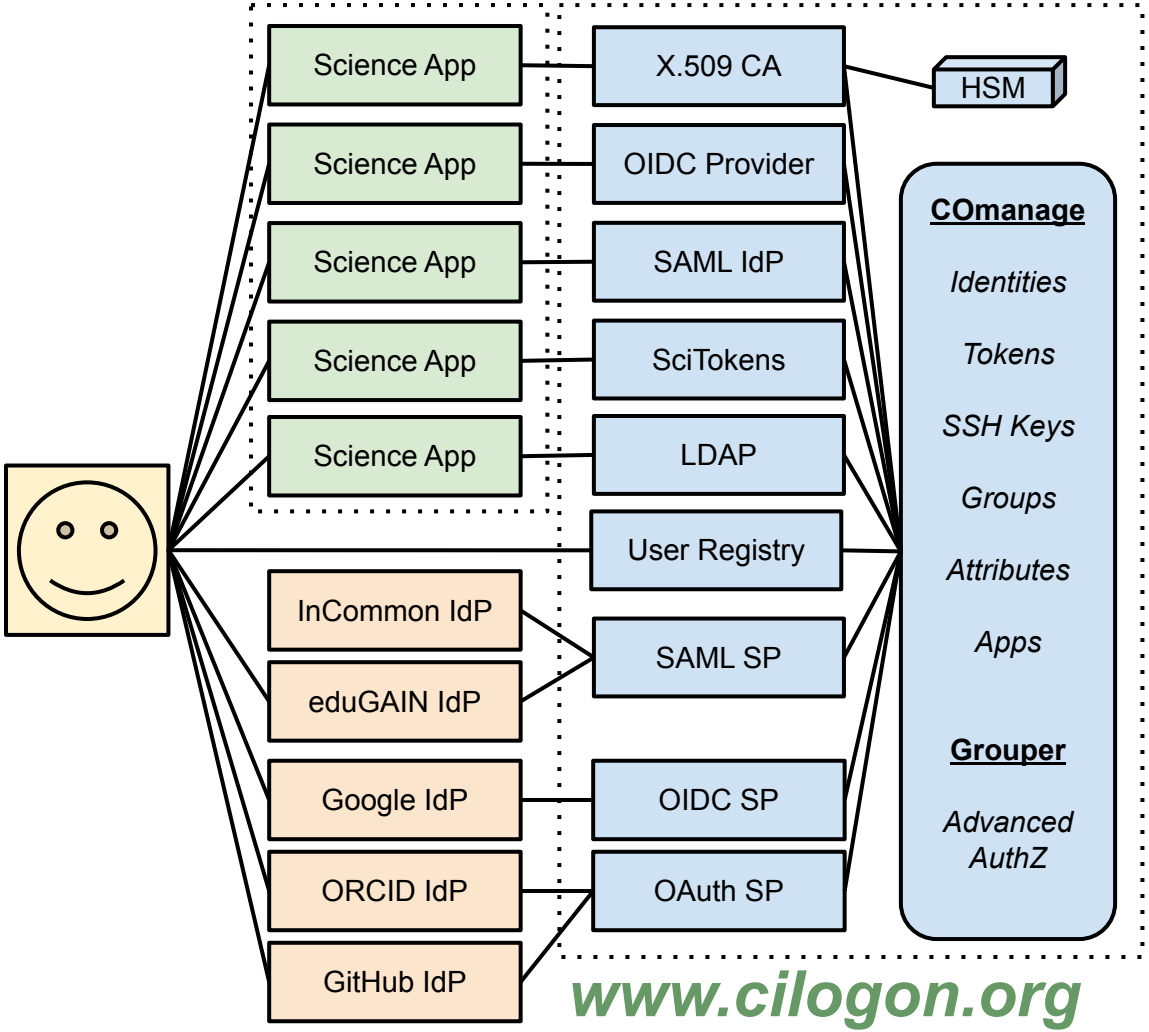
with onboarding/offboarding/attributes/groups/roles managed consistently across multiple applications





supporting access to science applications on HPC clusters, in Jupyter notebooks, using Globus, via REST APIs, and many other interfaces

using existing identity providers from the researcher's home organization (SAML/ADFS) or external sources (Google, GitHub, Microsoft, ORCID)



CILogon

www.cilogon.org

examples of CILogon-enabled sites

2i2c, Apache Airavata Test Drive, Ask.CI, ATLAS Connect,
Australian BioCommons, BNL Quantum Astrometry, Brainlife.io, CADRE, CERN
PanDA, Chem Compute, ClassTranscribe, CloudBank, Clowder, CMS Connect,
Connect.ci, Custos, CyberGISX, CyVerse, DataCite, Duke CI Connect,
Einstein Toolkit, FABRIC, Fermilab, Flywheel, GeoChemSim, Globus,
GW-Astronomy, HubICL, HTRC, ImPACT, JLab, LIGO, LROSE, LS-CAT, LSST,
Mass Open Cloud, MIT Engaging OnDemand, MSU HPCC OnDemand,
MyGeoHub, NEON, NIH ClinOmics, NIH KnowEnG,
Ocean Observatories Initiative, Open Science Chain, OSC OnDemand,
OSG Connect, Pacific Research Platform, QUBES, SciGaP, SCiMMA, SEAGrid,
SeedMeLab, SimVascular, Social Media Macroscope, UCLA JupyterHub,
Vanderbilt JupyterHub, and XSEDE

sustainability

development supported by NSF/DOE

operational support from XSEDE

non-profit subscription model administered by NCSA/UIUC

supports long-term sustainability

provides contracted SLAs

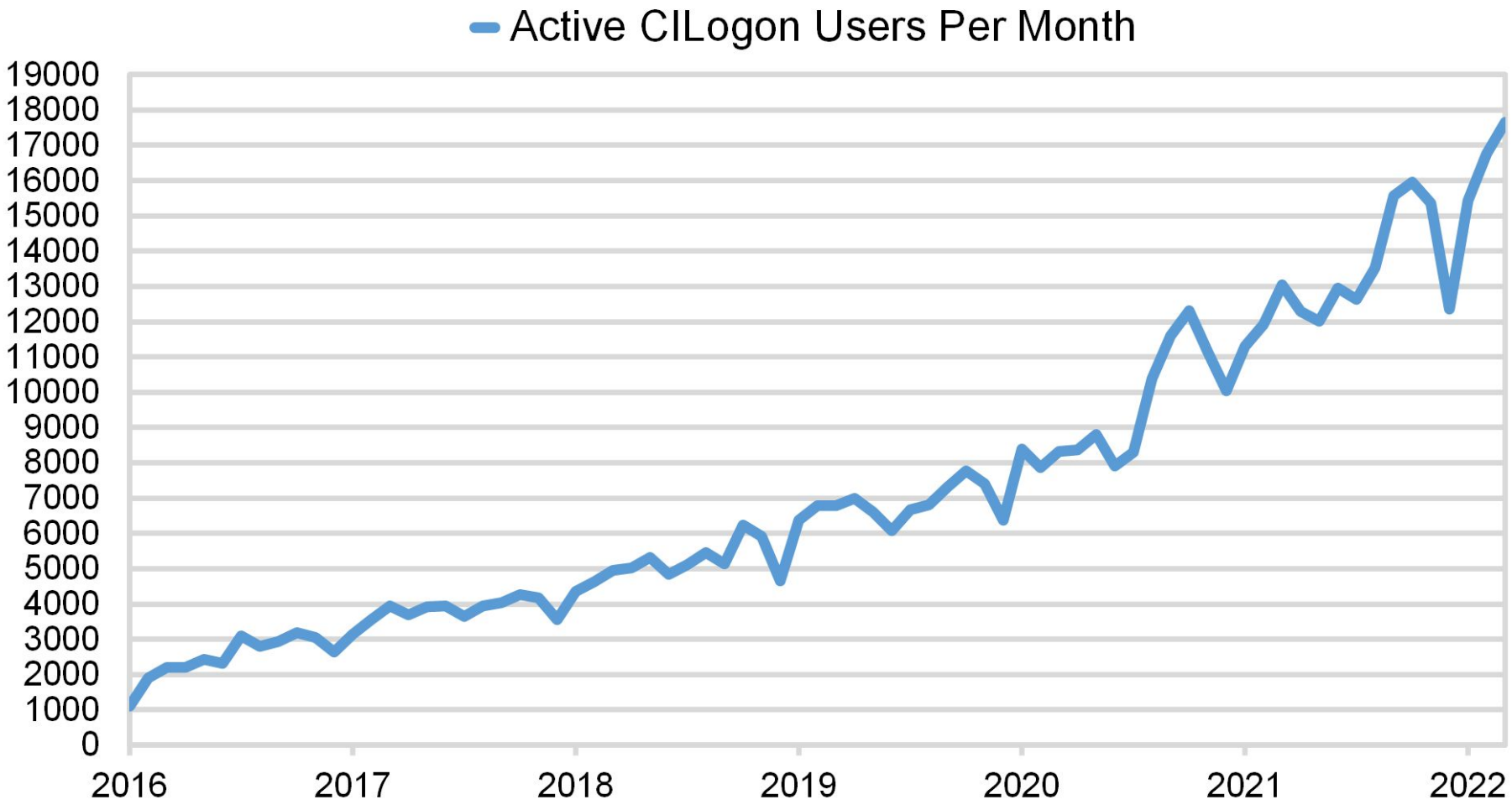
CILogon remains open source and focused on research & scholarship needs

https://www.cilogon.org/subscribe

Top 20 IdPs

(by # of unique active users in March 2022)

1315 Penn State
725 XSEDE
720 University of Illinois at Urbana-Champaign
665 Fermi National Accelerator Laboratory
619 National Institutes of Health
514 LIGO Scientific Collaboration
486 Northeastern University
483 University of Michigan
423 University of California-Los Angeles
420 Michigan State University
325 MIT
302 University of Chicago
297 Washington University in St. Louis
264 NCSA
255 Stanford University
251 Purdue University Main Campus
240 University of Wisconsin-Madison
216 Northwestern University
208 University of California-San Diego
189 Yale University



CILogon www.cilogon.org

operational timeline

2010 Operations begin with servers in DCL.

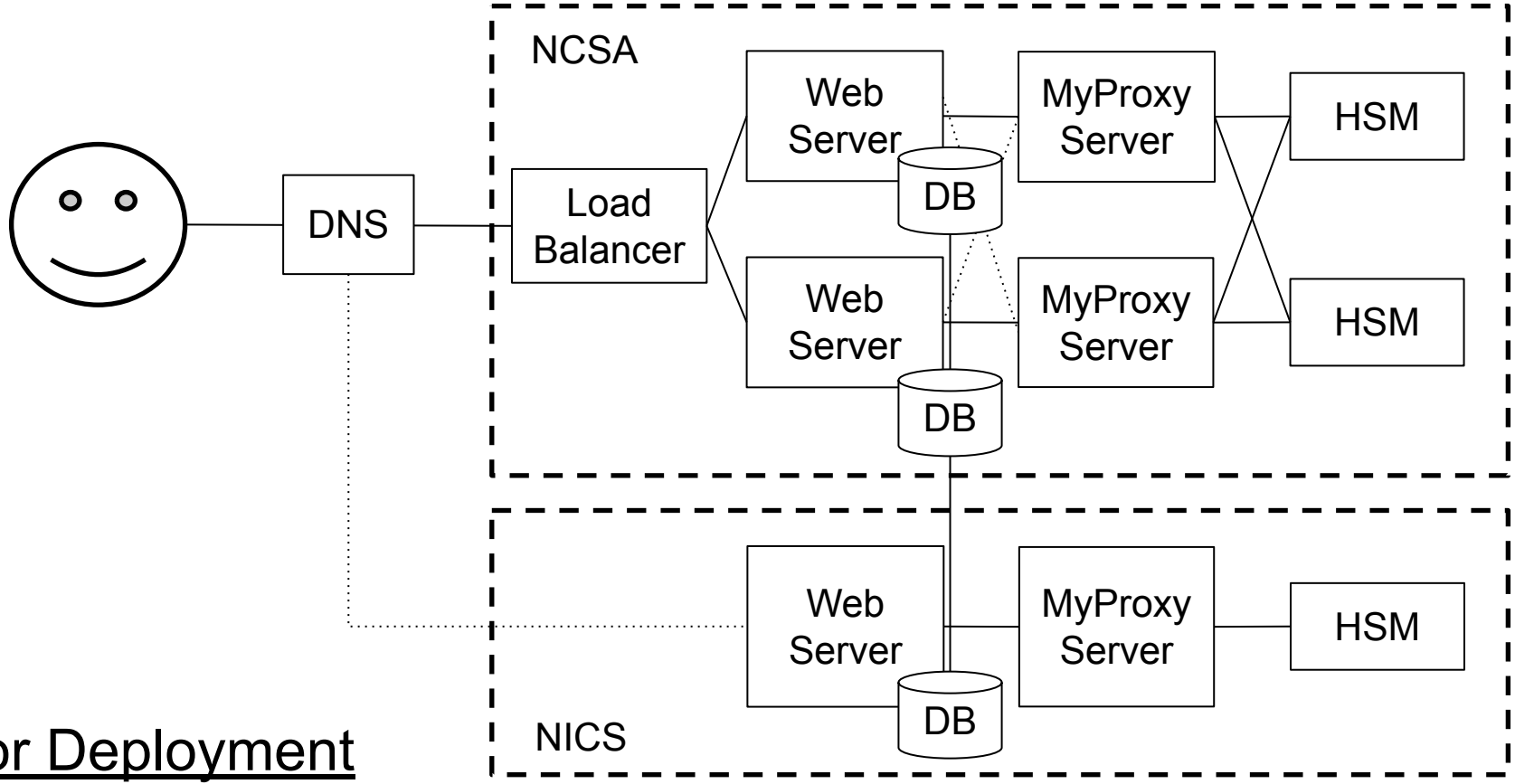
2013 Servers migrated from DCL to NPCF.

2014 Backup servers installed at ORNL/NICS.

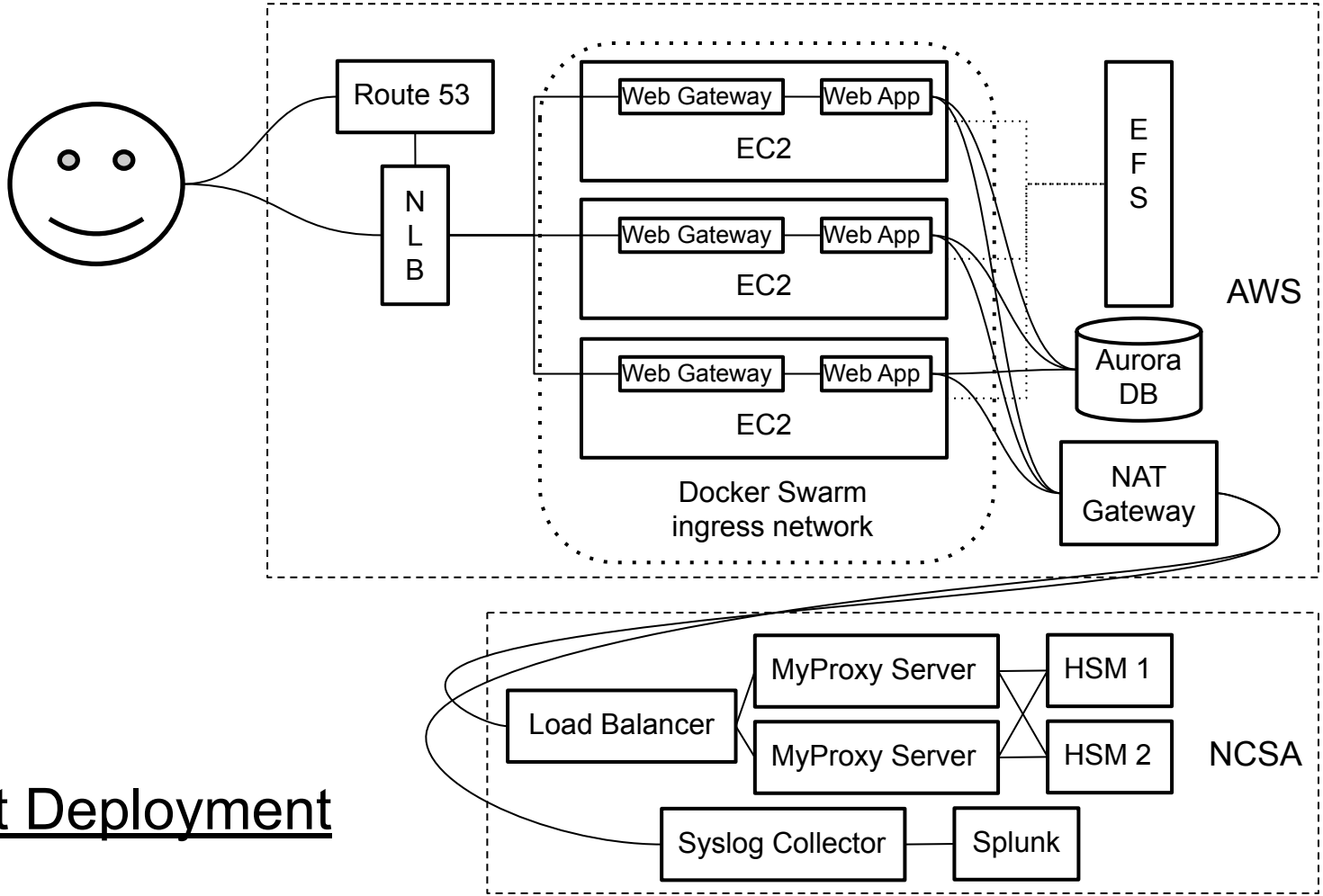
2017 New services deployed to AWS.

2019 Transition to subscription funding model.

2022 Core services migrated from NPCF to AWS.



Prior Deployment



Current Deployment

Why AWS?

XSEDE ends in 2022

- Need to migrate servers from ORNL/NICS

Multiple availability zones and regions

- Including international data centers

Hosted Services (Aurora DB, EFS, ELB, …)

CILogon AWS costs covered by subscribers

configuration management

Ansible Playbooks in private GitHub repos

Thanks to https://web.uillinois.edu/github !

Secrets managed by Ansible Vault

access management

Console access via Illinois and NCSA Shibboleth (MFA)

SSH access via EC2 Bastion Hosts

- Using SSH public keys + Duo MFA
- AWS security group rules limit access by IP

Ansible access via AWS REST API calls and SSH

Monthly Costs

EC2 is our primary cost

Savings Info

Total number of savings types	Total savings in USD
2	(USD 1,060.68)

Filter by savings type

1

Savings type	Amount in USD
Savings Plans Discounts	(USD 919.21)
Enterprise Discount Program Discounts	(USD 141.48)

Total active services	Total pre-tax service charges in USD
18	USD 1,156.70

Filter by service name

1

Service name | Region	Amount in USD
Savings Plans for AWS Compute usage	USD 460.52
Elastic Compute Cloud	USD 289.38
Relational Database Service	USD 144.46
Elastic File System	USD 103.03
Data Transfer	USD 72.81
Elastic Load Balancing	USD 45.59
CloudWatch	USD 12.25
Registrar	USD 12.00
Route 53	USD 7.30
Backup	USD 6.66
EC2 Container Registry (ECR)	USD 1.39
CodeBuild	USD 1.31
Simple Storage Service	USD 0.00
Key Management Service	USD 0.00
CloudTrail	USD 0.00
Elastic Container Registry Public	USD 0.00
Simple Email Service	USD 0.00
Simple Notification Service	USD 0.00

CILogon

budget alerts

Actual cost > 50%

Definition

When your actual cost is greater than **50% ($500.00)** of your **budgeted amount ($1,000.00)**, the alert threshold will be exceeded.

Threshold

Exceeded

Actions

-

Actual cost > 75%

Definition

When your actual cost is greater than **75% ($750.00)** of your **budgeted amount ($1,000.00)**, the alert threshold will be exceeded.

Threshold

Exceeded

Actions

-

Actual cost > 100%

Definition

When your actual cost is greater than **100% ($1,000.00)** of your **budgeted amount ($1,000.00)**, the alert threshold will be exceeded.

Threshold

Not exceeded

Actions

-

Actual cost > 150%

Definition

When your actual cost is greater than **150% ($1,500.00)** of your **budgeted amount ($1,000.00)**, the alert threshold will be exceeded.

Threshold

Not exceeded

Actions

-

Elastic Compute Cloud (EC2)

3 availability zones in us-east-2 (Ohio)

3 t3.small instances for SSH bastions

15 t2.large instances for Docker

- 6 prod, 5 test, 4 dev
- Memory is the scarce resource

Relational Database Service (RDS)

Aurora MySQL - 2 db.t3.medium instances

automatic replication & fail-over across availability zones

automatic backups

about 400 steady-state DB connections

good interactive performance

experimented with RDS Proxy but not using that now

Elastic File System (EFS)

NFS mounted into service containers, containing:

- container images & configs
- logs
- run-time service files (certificates, metadata, etc.)

CloudWatch

alerts

graphs

future: customer access to logs


PROD - CPUUtilization
Percent
64
37.2
10.5
23:00
05:00
11:00
17:00
i-039b896d13d6bde64 (cilogon-prod-b-docker-wkr-01)
i-0657102a8c1b98783 (cilogon-prod-b-docker-mgr-01)
i-0d7357443413b7ff8 (cilogon-prod-c-docker-wkr-01)
i-0b621e14be7eb59fc (cilogon-prod-c-docker-mgr-01)
i-0eda3ada2df3224a9 (cilogon-prod-a-docker-wkr-01)
DEV - CPUUtilization
Percent
6.72
4.29
1.86
23:00
05:00
11:00
17:00
i-018ed218456c88b65 (cilogon-dev-b-docker-mgr-01)
i-07c7560ac7ec98c95 (cilogon-dev-c-docker-mgr-01)
i-0e7307b9e579e1d14 (cilogon-dev-a-docker-wkr-01)
i-011a6ce9b2e3ba2df (cilogon-dev-a-docker-mgr-01)
TEST - CPUUtilization
Percent
81.2
49.9
18.6
23:00
05:00
11:00
17:00
i-0e48255b0877dafa0 (cilogon-test-b-docker-wkr-01)
i-0900e755fac785f24 (cilogon-test-b-docker-mgr-01)
i-0dce00bb3304408b4 (cilogon-test-a-docker-wkr-02)
i-0a9ce7d4b02bf2abf (cilogon-test-a-docker-mgr-02)
i-0c5ca7e68bec23466 (cilogon-test-a-docker-mgr-01)
PROD - NetworkIn
Bytes
723M
362M
1.73M
23:00
05:00
11:00
17:00
i-039b896d13d6bde64 (cilogon-prod-b-docker-wkr-01)
i-0657102a8c1b98783 (cilogon-prod-b-docker-mgr-01)
i-0d7357443413b7ff8 (cilogon-prod-c-docker-wkr-01)
i-0b621e14be7eb59fc (cilogon-prod-c-docker-mgr-01)
i-0eda3ada2df3224a9 (cilogon-prod-a-docker-wkr-01)
DEV - NetworkIn
Bytes
15.6M
8.24M
847k
23:00
05:00
11:00
17:00
i-018ed218456c88b65 (cilogon-dev-b-docker-mgr-01)
i-07c7560ac7ec98c95 (cilogon-dev-c-docker-mgr-01)
i-0e7307b9e579e1d14 (cilogon-dev-a-docker-wkr-01)
i-011a6ce9b2e3ba2df (cilogon-dev-a-docker-mgr-01)
TEST - NetworkIn
Bytes
230M
117M
4.08M
23:00
05:00
11:00
17:00
i-0e48255b0877dafa0 (cilogon-test-b-docker-wkr-01)
i-0900e755fac785f24 (cilogon-test-b-docker-mgr-01)
i-0dce00bb3304408b4 (cilogon-test-a-docker-wkr-02)
i-0a9ce7d4b02bf2abf (cilogon-test-a-docker-mgr-02)
i-0c5ca7e68bec23466 (cilogon-test-a-docker-mgr-01)
PROD - NetworkOut
Bytes
294M
149M
4.08M
23:00
05:00
11:00
17:00
i-039b896d13d6bde64 (cilogon-prod-b-docker-wkr-01)
i-0657102a8c1b98783 (cilogon-prod-b-docker-mgr-01)
i-0d7357443413b7ff8 (cilogon-prod-c-docker-wkr-01)
i-0b621e14be7eb59fc (cilogon-prod-c-docker-mgr-01)
i-0eda3ada2df3224a9 (cilogon-prod-a-docker-wkr-01)
DEV - NetworkOut
Bytes
5.31M
2.90M
501k
23:00
05:00
11:00
17:00
i-018ed218456c88b65 (cilogon-dev-b-docker-mgr-01)
i-07c7560ac7ec98c95 (cilogon-dev-c-docker-mgr-01)
i-0e7307b9e579e1d14 (cilogon-dev-a-docker-wkr-01)
i-011a6ce9b2e3ba2df (cilogon-dev-a-docker-mgr-01)
TEST - NetworkOut
Bytes
123M
64.6M
6.46M
23:00
05:00
11:00
17:00
i-0e48255b0877dafa0 (cilogon-test-b-docker-wkr-01)
i-0900e755fac785f24 (cilogon-test-b-docker-mgr-01)
i-0dce00bb3304408b4 (cilogon-test-a-docker-wkr-02)
i-0a9ce7d4b02bf2abf (cilogon-test-a-docker-mgr-02)
i-0c5ca7e68bec23466 (cilogon-test-a-docker-mgr-01)

Simple Email Service (SES)

user notifications

operational alerts

easy SPF, DKIM, DMARC config w/ Route 53

migrating core services to AWS

200k logins per month from 20k users
(from 500 organizations)

How will costs increase?

- Data Transfer, RDS I/O requests,
  NAT Gateway bandwidth,
  Route 53 queries, NLB capacity units

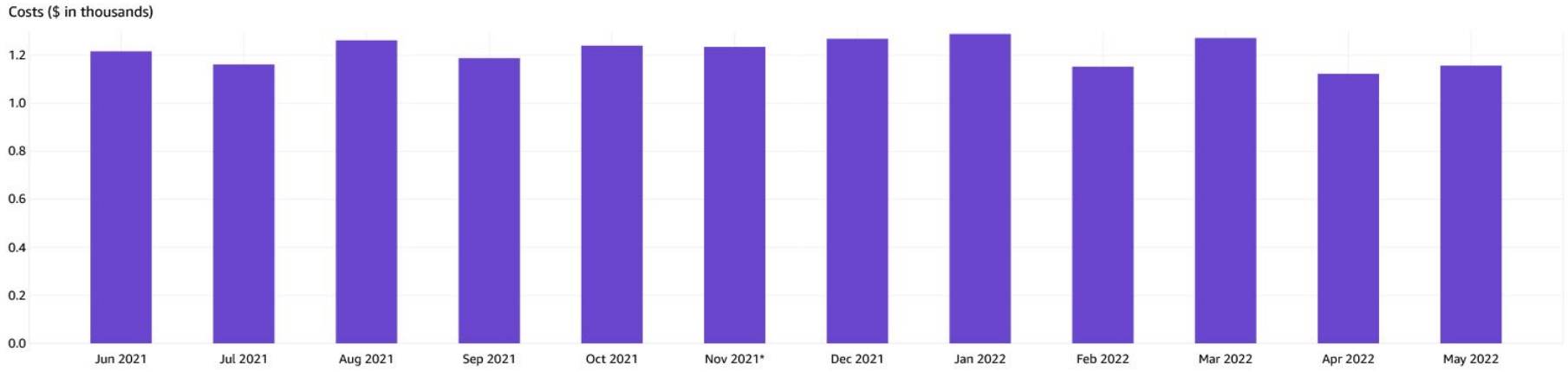
Will our services be overloaded? (EC2, EFS, RDS)

Database Migration Service

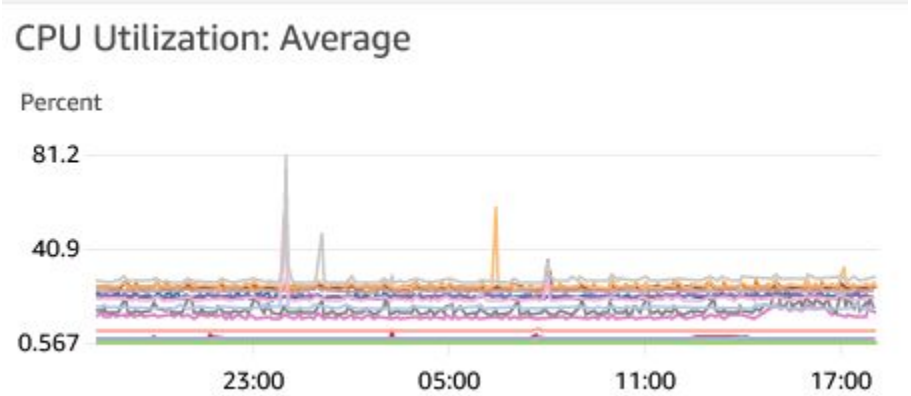
seamless migration from NCSA/NICS on-prem MariaDB instances to AWS RDS Aurora MySQL

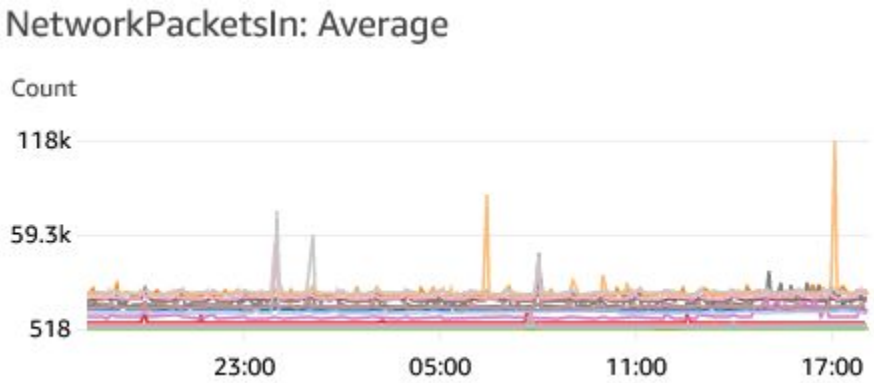
live database replication enabled swap-over without downtime

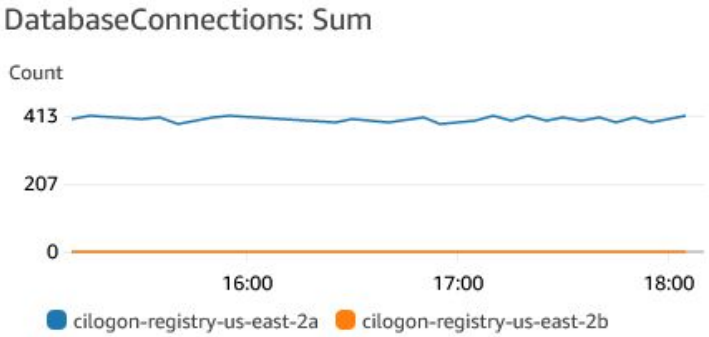
costs held steady

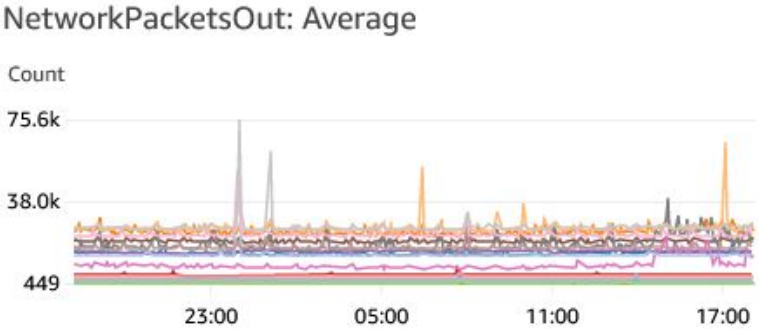


generally steady service load

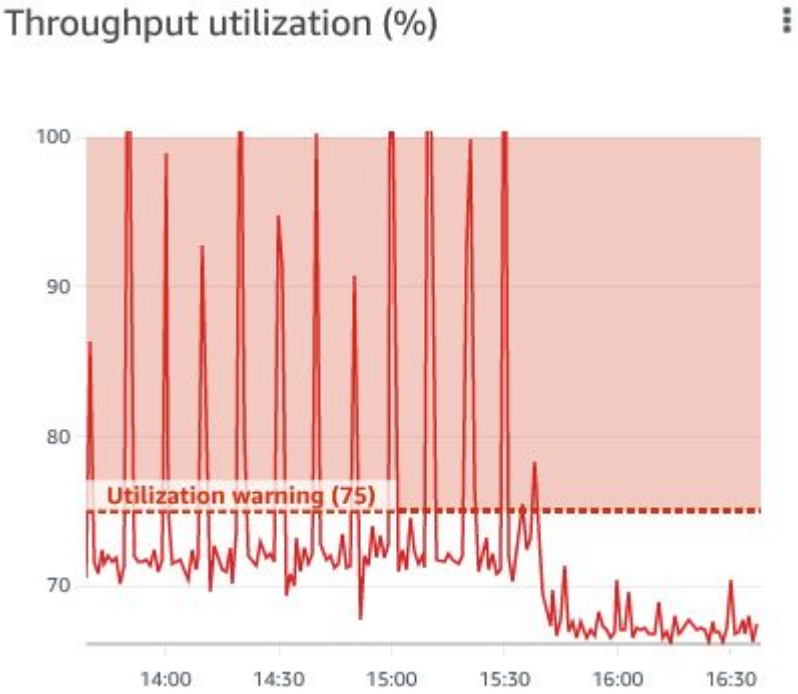


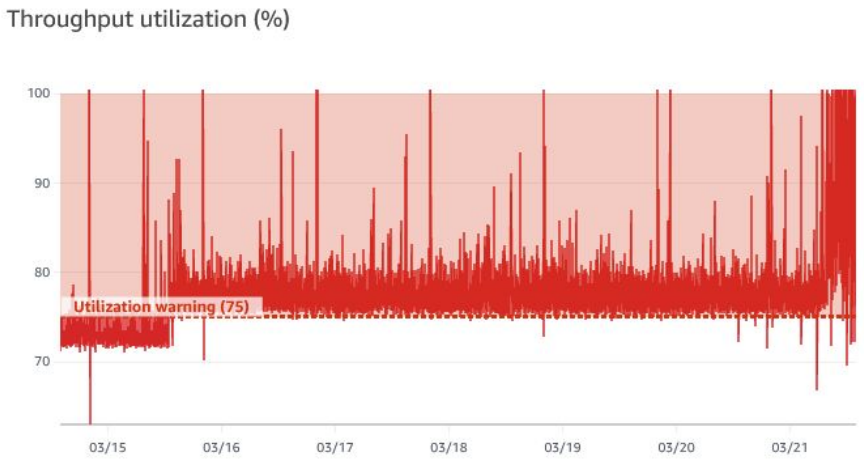


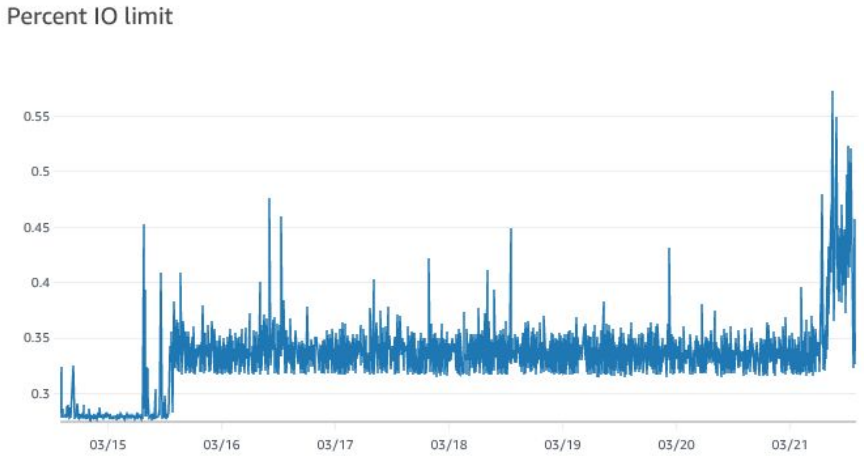




EFS surprise!







CILogon

What's next?

IPv6

Elastic Kubernetes Service (EKS)

New Region: Sydney

Thanks!

contact:

help@cilogon.org

subscribe for updates:

https://groups.google.com/a/cilogon.org/g/announce