Article
Anton Umnikov · Feb 11, 2020 19m read

InterSystems IRIS Deployment Guide for AWS using CloudFormation template

InterSystems IRIS Deployment Guide for AWS using CloudFormation template

Please note: following this guide, especially the prerequisites section requires Intermediate to Advanced level of knowledge of AWS. You'll need to create and manage S3 buckets, IAM roles for EC2 instances, VPCs and Subnets. You'll also need access to InterSystems binaries (usually downloaded via WRC site) as well as IRIS license key.

Aug 12, 2020
Anton Umnikov

Templates Source code is available here: https://github.com/antonum/AWSIRISDeployment

Table of Contents

Introduction

InterSystems provides the ***CloudFormation Template*** for users to set up their own InterSystems IRIS® data platform according to InterSystems and AWS best practices.

This guide will detail the steps to deploy the ***CloudFormation template.***

In this guide, we cover two types of deployments for the InterSystems IRIS ***CloudFormation template***. The first method is highly available using multiple availability zones (AZ) and targeted to production workloads, and the second method is a single availability zone deployment for development and testing workloads.

Prerequisites and Requirements

In this section, we detail the prerequisites and requirements to run and operate our solution.

Time

The deployment itself takes about **4 minutes**, but with prerequisites and testing it could take **up to 2 hours**.

Product License and Binaries

InterSystems IRIS binaries are available to InterSystems customers via https://wrc.intersystems.com/. Login with your WRC credentials and follow the links to Actions -> SW Distributions -> InterSystems IRIS. This Deployment Guide is written for the Red Hat platform of InterSystems IRIS 2020.1 build 197. IRIS binaries file names are of the format ISCAgent-2020.1.0.215.0-lnxrhx64.tar.gz and IRISHealth-2020.1.0.217.1-lnxrhx64.tar.gz

InterSystems IRIS license key – you should be able to use your existing InterSystems IRIS license key (iris.key). You can also request an evaluation key via the InterSystems IRIS Evaluation Service: https://download.intersystems.com/download/register.csp.

AWS Account

You must have an AWS account set up. If you do not, visit: https://aws.amazon.com/getting-started/

IAM Entity for user

Create an IAM user or role. Your IAM user should have a policy that allows AWS CloudFormation actions. Do not use your root account to deploy the CloudFormation template. In addition to AWS CloudFormation actions, IAM users who create or delete stacks will also require additional permissions that depend on the stack template. This deployment requires permissions to all the services listed in the following section.

Reference: https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/using-iam-template.html.

IAM Role for EC2

The CloudFormation template requires an IAM role that allows your EC2 instance to access S3 buckets and put logs into CloudWatch. See Appendix " IAM Policy for EC2 instance " for an example of the policy associated with such role.

Reference: https://docs.aws.amazon.com/IAM/latest/UserGuide/idroles.html .

S3 Bucket

Create an S3 bucket called " my bucket " , copy IRIS binaries files and iris.key:

```
BUCKET=<my bucket>

aws s3 mb s3://$BUCKET

aws s3 cp ISCAgent-2020.1.0.215.0-lnxrhx64.tar.gz s3://$BUCKET

aws s3 cp IRISHealth-2020.1.0.217.1-lnxrhx64.tar.gz s3://$BUCKET

aws s3 cp iris.key s3://$BUCKET
```

VPC and Subnets

The template is designed to deploy IRIS into an existing VPC and Subnets. In regions where three or more Availability Zones are available, we recommend creating three private subnets across three different AZ ' s. Bastion Host should be located in any of the public subnets within the VPC. You can follow the AWS example to create a VPC and Subnets with the

CloudFormation template: https://docs.aws.amazon.com/codebuild/latest/userguide/cloudformation-vpc-template.html.

EC2 Key Pair

To access the EC2 instances provisioned by this template, you will need at least one EC2 Key Pair. Refer to this guide for details: https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-key-pairs.html.

Knowledge Requirements

Knowledge of the following AWS services is required:

- Amazon Elastic Compute Cloud (Amazon EC2)
- Amazon Virtual Private Cloud (Amazon VPC)
- AWS CloudFormation
- AWS Elastic Load Balancing
- AWS S3

Account limit increases will not be required for this deployment.

More information on proper policy and permissions can be found here: https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/using-iam-template.html.

Note: Individuals possessing the AWS Associate certifications should have a sufficient depth of knowledge.

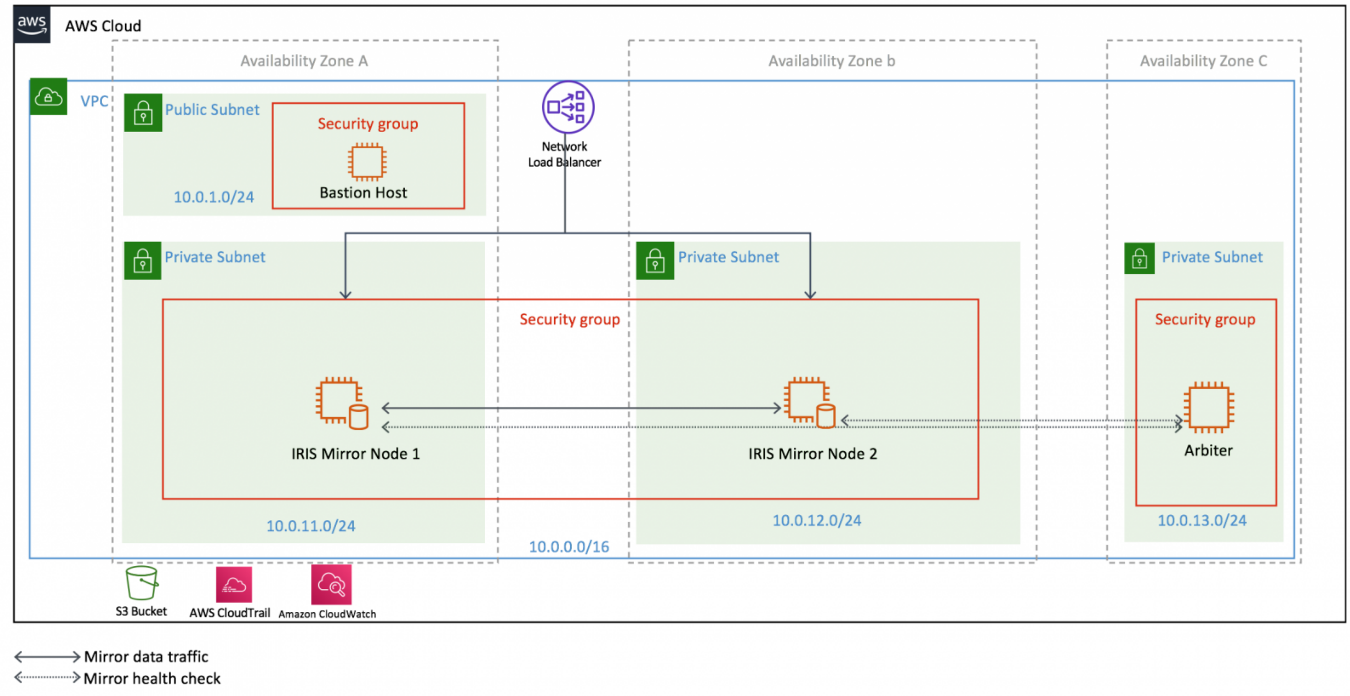
Architecture

In this section, we give architecture diagrams of two deployment possibilities, and talk about architecture design choices.

Multi-AZ Fault Tolerant Architecture Diagram (Preferred)

In this preferred option, mirrored IRIS instances are situated behind a load balancer in two availability zones to ensure high availability and fault tolerance. In regions with three or more availability zones, the Arbiter node is located in the third AZ.

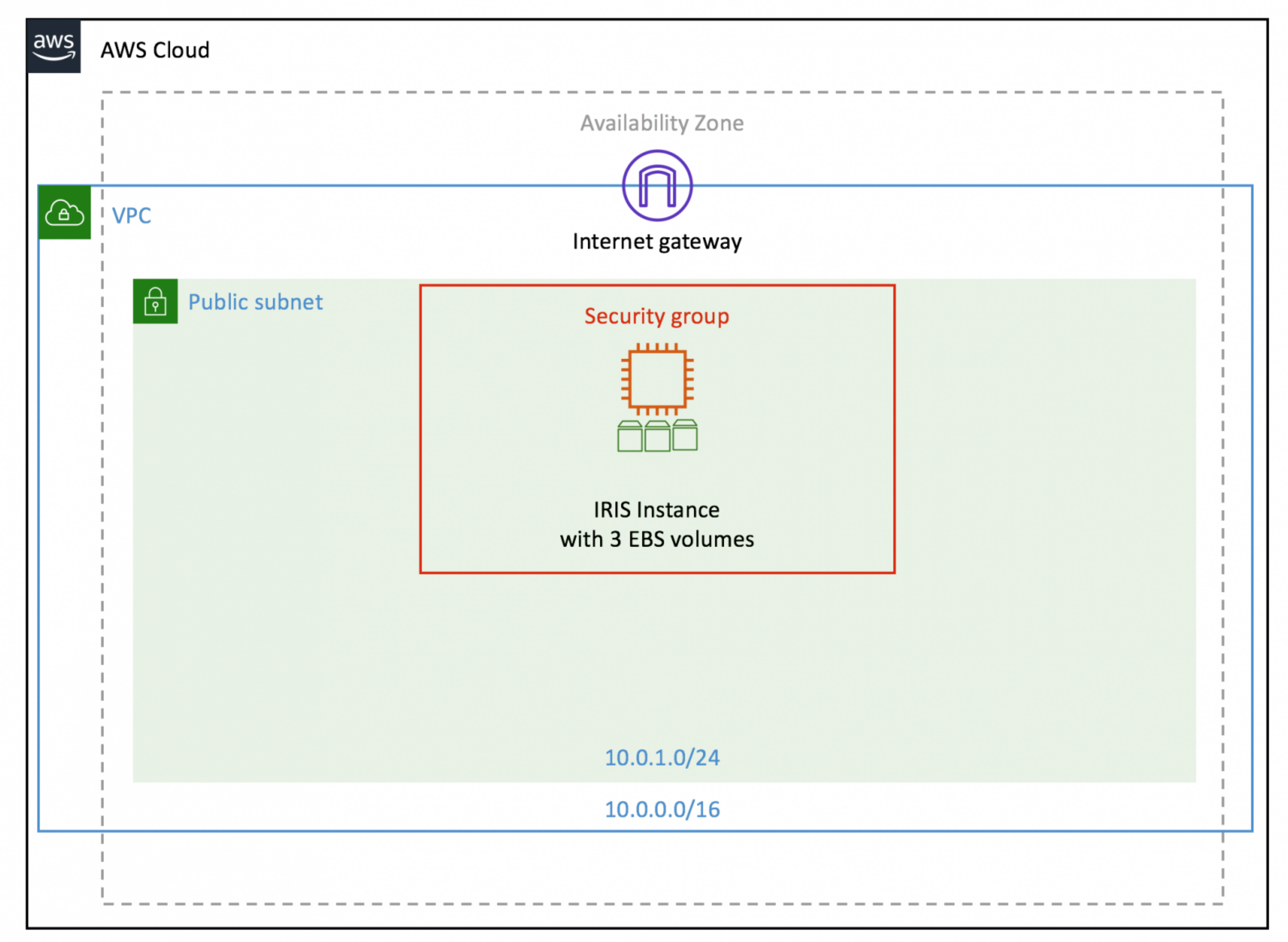
Database nodes are located in private subnets. Bastion Host is in a Public subnet within the same VPC.



1. Network Load Balancer directs database traffic to the current Primary IRIS node
2. Bastion Host allows secure access to the IRIS EC2 instances
3. IRIS stores all customer data in encrypted EBS volumes
    a. EBS is encrypted and uses the AWS Key Management Service (KMS) managed key
    b. For regulated workloads where encryption of data in transit is required, you can choose to use the r5n family of instances, since they provide automatic instance-to-instance traffic encryption. IRIS-level traffic encryption is also possible but not enabled by CloudFormation (see the Encrypting Data in Transit section of this guide)
4. Use of security groups restrict access to the greatest degree possible by only allowing necessary traffic

Single Instance, Single AZ Architecture Diagram (Development and Testing)

InterSystems IRIS can also be deployed in a single Availability Zone for development and evaluation purposes. The data flow and architecture components are the same as the ones highlighted in the previous section. This solution does not provide high availability or fault tolerance, and is not suitable for production use.



Deployment

1. Log into your AWS account with the IAM entity created in the Prerequisites section with the required permissions to deploy the solution
2. Make sure all the Prerequisites, such as VPC, S3 bucket, IRIS binaries and license key are in place
3. Click the following link to deploy CloudFormation template (deploys in us-east-1): https://console.aws.amazon.com/cloudformation/home?region=us-east-1#/stacks/new?stackName=InterSystemsIRIS&templateURL=https://isc-tech-validation.s3.amazonaws.com/MirrorCluster.yaml for multi-AZ, fault tolerant deployment
4. In ‘ Step 1 - Create Stack ’ , press the ‘ Next ’ button
5. In ‘ Step 2 - Specify stack details ’ , fill out and adjust CloudFormation parameters depending on your requirements
6. Press the ‘ Next ’ button
7. In ‘ Step 3 - Configure stack options ’ , enter and adjust optional tags, permissions, and advanced options
8. Press the ‘ Next ’ button
9. Review your CloudFormation configurations
10. Press the ‘ Create Stack ’ button
11. Wait approximately 4 minutes for your CloudFormation template to deploy
12. You can verify your deployment has succeeded by looking for a ‘ CREATE_COMPLETE ’ status
13. If the status is ‘ CREATE_FAILED ’ , see the troubleshooting section in this guide
14. Once deployment succeeds, please carry out Health Checks from this guide

Security

In this section, we discuss the InterSystems IRIS default configuration deployed by this guide, general best practices, and options for securing your solution on AWS.

Data in Private Subnets

InterSystems IRIS EC2 instances must be placed in Private subnets and accessed only via Bastion Host or by applications via the Load Balancer.

Encrypting IRIS Data at Rest

On database instances running InterSystems IRIS, data is stored at rest in underlying EBS volumes which are encrypted. This CloudFormation template creates EBS volumes encrypted with the account-default AWS managed Key, named aws/ebs.

Encrypting IRIS data in transit

This CloudFormation does not secure Client-Server and Instance-to-Instance connections. Should data in transit encryption be required, follow the steps outlined below after the deployment is completed.

Enabling SSL for SuperServer connections (JDBC/ODBC connections): https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GCASssltls#GCASssltlssuperserver .

Durable multi-AZ configuration traffic between IRIS EC2 instances may need to be encrypted too. This can be achieved either by enabling SSL Encryption for mirroring: https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GCASssltls#GCASssltlsmirroring or switching to the r5n family of instances which provides automatic encryption of instance-to-instance traffic.

You can use AWS Certificate Manager (ACM) to easily provision, manage, and deploy Secure Sockets Layer/Transport Layer Security (SSL/TLS) certificates.

Secure access to IRIS Management Portal

By default, the IRIS management portal is accessed only via Bastion Host.

Logging/Auditing/Monitoring

InterSystems IRIS stores logging information in the messages.log file. CloudFormation does not setup any additional logging/monitoring services. We recommend that you enable structured logging as outlined here: https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=ALOG.

The CloudFormation template does not install InterSystems IRIS-CloudWatch integration. InterSystems recommends using InterSystems IRIS-CloudWatch integration from https://github.com/antonum/CloudWatch-IRIS. This enables collection of IRIS metrics and logs from the messages.log file into AWS CloudWatch.

The CloudFormation template does not enable AWS CloudTrail logs. You can enable CloudTrail logging by navigating to the CloudTrail service console and enabling CloudTrail logs. With CloudTrail, activity related to actions across your AWS infrastructure are recorded as an event in CloudTrail. This helps you enable governance, compliance, and operational and risk auditing of your AWS account.

Reference: https://docs.aws.amazon.com/awscloudtrail/latest/userguide/cloudtrail-user-guide.html

InterSystems recommends monitoring of InterSystems IRIS logs and metrics, and alerting on at least the following indicators:

- severity 2 and 3 messages
- license consumption
- disk % full for journals and databases
- Write Daemon status
- Lock Table status

In addition to the above, customers are encouraged to identify their own monitoring and alert metrics and application-specific KPIs.

Sizing/Cost

This guide will create the AWS resources outlined in the **Deployment Assets** section of this document. You are responsible for the cost of AWS services used while running this deployment. The minimum viable configuration for an InterSystems IRIS deployment provides high availability and security.

The template in this guide is using the BYOL (Bring Your Own License) InterSystems IRIS licensing model.

You can access Pay Per Hour IRIS Pricing at the InterSystems IRIS Marketplace page: https://aws.amazon.com/marketplace/pp/B07XRX7G6B?qid=1580742435148&sr=0-3

For details on BYOL pricing, please contact InterSystems at: https://www.intersystems.com/who-we-are/contact-us/.

The following AWS assets are required to provide a functional platform:

- 3 EC2 Instances (including EBS volumes and provisioned IOPS)
- 1 Elastic Load Balancer

The following table outlines recommendations for EC2 and EBS capacity built into the deployment CloudFormation template, as well as AWS resources costs (Units $/Month).

Workload				
	Dev/Test	**Prod Small**	**Prod Medium**	**Prod Large**
EC2 DB*	m5.large	2 * r5.large	2 * r5.4xlarge	2 * r5.8xlarge
EC2 Arbiter*	t3.small	t3.small	t3.small	t3.small
EC2 Bastion*	t3.small	t3.small	t3.small	t3.small
EBS SYS	gp2 20GB	gp2 50GB	io1 512GB 1,000iops	io1 600GB 2,000iops
EBS DB	gp2 128GB	gp2 128GB	io1 1TB 10,000iops	io1 4TB 10,000iops
EBS JRN	gp2 64GB	gp2 64GB	io1 256GB 1,000iops	io1 512GB 2,000iops
Cost Compute	85.51	199.71	1506.18	2981.90
Cost EBS vol	27.20	27.20	450.00	1286.00
Cost EBS IOPS	-	-	1560.00	1820.00
Support (Basic)	-	-	351.62	608.79
Cost Total	127.94	271.34	3867.80	6696.69
Calculator link	Calculator	Calculator	Calculator	Calculator

AWS cost estimates are based on On-Demand pricing in the North Virginia Region. Cost of snapshots and data transfer are not included. Please consult AWS Pricing for the latest information.

Deployment Assets

Deployment Options

The InterSystems IRIS CloudFormation template provides two different deployment options. The multi-AZ deployment option provides a highly available redundant architecture that is suitable for production workloads. The single-AZ deployment option provides a lower cost alternative that is suitable for development or test workloads.

Deployment Assets (Recommended for Production)

The InterSystems IRIS deployment is executed via a CloudFormation template that receives input parameters and passes them to the appropriate nested template. These are executed in order based on conditions and dependencies.

AWS Resources Created:

- VPC Security Groups
- EC2 Instances for IRIS nodes and Arbiter
- Amazon Elastic Load Balancing (Amazon ELB) Network Load Balancer (NLB)

CloudFormation Template Input Parameters

General AWS

- EC2 Key Name Pair
- EC2 Instance Role

S3

- Name of S3 bucket where the IRIS distribution file and license key are located

Network

- The individual VPC and Subnets where resources will be launched

Database

- Database Master Password
- EC2 instance type for Database nodes

Stack Creation

There are four outputs for the master template: the JDBC endpoint that can be used to connect JDBC clients to InterSystems IRIS, the public IP of the Bastion Host and private IP addresses for both IRIS nodes.

Clean Up

- Follow the AWS CloudFormation Delete documentation to delete the resources deployed by this document
- Delete any other resources that you manually created to integrate or assist with the deployment, such as S3 bucket and VPC

Testing the Deployment

Health Checks

Follow the template output links to Node 01/02 Management Portal. Login with the username: SuperUser and the password you selected in the CloudFormation template.

Navigate to System Administration -> Configuration -> Mirror Settings -> Edit Mirror. Make sure the system is configured with two Failover members.

Verify that the mirrored database is created and active. System Administration -> Configuration -> Local Databases.

Validate the JDBC connection by following the “ First Look JDBC ” document: https://docs.intersystems.com/irislatest/csp/docbook/DocBook.UI.Page.cls?KEY=AFLjdbc to validate JDBC connectivity to IRIS via the Load Balancer. Make sure to change the url variable to the value displayed in the template output, and password from “ SYS ” to the one you selected during setup.

Failover Test

On the Node02, navigate to the Management Portal (see “ Health Check ” section above) and open the Configuration->Edit Mirror page. At the bottom of the page you will see *This member is the backup. Changes must be made on the primary.*

Locate the Node01 instance in the AWS EC2 management dashboard. Its name will be of the format: MyStackName-Node01-1NGXXXXXX

Restart the Node01 instance. This will simulate an instance/AZ outage.

Reload Node02 “ Edit Mirror ” page. The status should change to: *This member is the primary. Changes will be sent to other members.*

Backup and Recovery

Backup

CloudFormation deployment does not enable backups for InterSystems IRIS. We recommend backing up IRIS EBS volumes using EBS Snapshot - https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSSnapshots.html - in combination with IRIS Write Daemon Freeze: https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GCDIbackup#GCDIbackupmethodsext .

Instance Failure

Unhealthy IRIS instances are detected by IRIS mirroring and Load Balancer, and traffic is redirected to another mirror node. Instances that are capable of recovery will rejoin the mirror and continue normal operations. If you encounter persistently unhealthy instances, please see our Knowledge Base and the “ Emergency Maintenance ” section of this guide.

Availability-Zone Failure

In the event of an availability-zone failure, temporary traffic disruptions may occur. Similar to instance failure, IRIS mirroring and Load Balancer would handle the event by switching traffic to the IRIS instance in the remaining available AZ.

Region Failure

The architecture outlined in this guide does not deploy a configuration that supports multi-region operation. IRIS asynchronous mirroring and AWS Route53 can be used to build configurations capable of handling region failure with minimal disruption. Please refer to https://community.intersystems.com/post/intersystems-iris-example-reference-architectures-amazon-web-services-aws for details.

RPO/RTO

Recovery Point Objective (RPO)

- **Single node Dev/Test** configuration is defined by the time of the last successful backup.
- **Multi Zone Fault Tolerant** setup provides Active-Active configuration that ensures full data consistency in the event of failover, with RPO of the last successful transaction.

Recovery Time Objective (RTO)

- Backup recovery for the **Single node Dev/Test** configuration is outside of the scope of this deployment guide. Please refer to https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-restoring-volume.html for details on restoring EBS volume snapshots.
- RTO for **Multi Zone Fault Tolerant** setup is typically defined by the time it takes for the Elastic Load Balancer to redirect traffic to the new Primary Mirror node of the IRIS cluster. You can further reduce RTO time by developing mirror-aware applications or adding an Application Server Connection to the mirror: https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GHAmirror#GHAmirrorsetconfigecp .

Storage Capacity

IRIS Journal and Database EBS volumes can reach storage capacity. InterSystems recommends monitoring Journal and Database volume state using the IRIS Dashboard, as well as Linux file-system tools such as df.

Both Journal and Database volumes can be expanded following the EBS guide https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-modify-volume.html. **Note:** both EBS volume expansion and Linux file system extension steps need to be performed. Optionally, after a database backup is performed, journal space can be reclaimed by running Purge Journals: https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GCDIjournal#GCDIjournaltasks .

You can also consider enabling CloudWatch Agent on your instances to monitor disk space (not enabled by this CloudFormation template): https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/Install-CloudWatch-Agent.html.

Security certificate expiration

You can use AWS Certificate Manager (ACM) to easily provision, deploy, manage, and monitor expiration of Secure Sockets Layer/Transport Layer Security (SSL/TLS) certificates.

Certificates must be monitored for expiration. InterSystems does not provide an integrated process for monitoring certificate expiration. AWS provides a CloudFormation template that can help setup an alarm. Please visit the following link for details: https://docs.aws.amazon.com/config/latest/developerguide/acm-certificate-expiration-check.html.

Routine Maintenance

For IRIS upgrade procedures in mirrored configurations, please refer to: https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GCIupgrade#GCIupgradetasksmirrors.

InterSystems recommends following the best practices of AWS and InterSystems for ongoing tasks, including:

- Access key rotation
- Service limit evaluation
- Certificate renewals
- IRIS License limits and expiration https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GCMdashboard
- Storage capacity monitoring https://docs.intersystems.com/irislatest/csp/docbook/Doc.View.cls?KEY=GCMdashboard .

Additionally, you might consider adding CloudWatch Agent to your EC2 instances: https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/Install-CloudWatch-Agent.html.

Emergency Maintenance

If EC2 instances are available, connect to the instance via bastion host.

Note: The public IP of the bastion host may change after an instance stop/start. That does not affect availability of the IRIS cluster and JDBC connection.

For command line access, connect to the IRIS nodes via bastion host:

```
$ chmod 400 <my-ec2-key>.pem

$ ssh-add <my-ec2-key>.pem


$ ssh -J ec2-user@<bastion-public-ip> ec2-user@<node-private-ip> -L 52773:1<node-private-ip>:52773
```

After that, the Management Portal for the instance would be available at: http://localhost:52773/csp/sys/%25CSP.Portal.Home.zen User: SuperUser, and the password you entered at stack creation.

To connect to the IRIS command prompt use:

```
$ iris session iris
```

Consult InterSystems IRIS Management and Monitoring guide: https://docs.intersystems.com/irislatest/csp/docbook/DocBook.UI.Page.cls?KEY=GCM.

Contact InterSystems Support.

If EC2 instances are not available/reachable, contact AWS Support.

NOTE: AZ or instance failures will automatically be handled in our Multi-AZ deployment.

Support

Troubleshooting

I cannot " Create stack " in CloudFormation

Please check that you have the appropriate permissions to " Create Stack " . Contact your AWS account admin for permissions, or AWS Support if you continue to encounter this issue.

Stack is being created, but I can ' t access IRIS

It takes approximately 2 minutes from the moment EC2 instance status turns into " CREATE COMPLETED " to the moment IRIS is fully available. SSH to the EC2 Node instances and check if IRIS is running:

```
$iris list
```

If you don ' t see any active IRIS instances, or the message " iris: command not found" appears, then IRIS installation has failed. Check $cat /var/log/cloud-init-output.log on the instance to identify any problems with the IRIS installation during instance first start.

IRIS is up, but I can ' t access either the Management Portal or connect from my [Java] application

Make sure that the Security Group created by CloudFormation lists your source IP address as allowed.

Contact InterSystems Support

InterSystems Worldwide Response Center (WRC) provides expert technical assistance.

InterSystems IRIS support is always included with your IRIS subscription.

Phone, email and online support are always available to clients 24 hours a day, 7 days a week. We maintain support advisers in 15 countries around the world and have specialists fluent in English, Spanish, Portuguese, Italian, Welsh, Arabic, Hindi, Chinese, Thai, Swedish, Korean, Japanese, Finnish, Russian, French, German, Hebrew, and Hungarian. Every one of our clients immediately gets help from a highly qualified support specialist who really cares about client success.

For Immediate Help

Support phone:

+1-617-621-0700 (US)

+44 (0) 844 854 2917 (UK)

0800615658 (NZ Toll Free)

1800 628 181 (Aus Toll Free)

Support email:

support@intersystems.com

Support online:

WRC Direct

Contact support@intersystems.com for a login.

Appendix

IAM Policy for EC2 instance

The following IAM policy allows the EC2 instance to read objects from the S3 bucket ‘ my-bucket ’ , and write logs to CloudWatch:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "S3BucketReadOnly",
      "Effect": "Allow",
      "Action": ["s3:GetObject"],
      "Resource": "arn:aws:s3:::my-bucket/*"
    },
    {
      "Sid": "CloudWatchWriteLogs",
      "Effect": "Allow",
      "Action": [
        "logs:CreateLogGroup",
        "logs:CreateLogStream",
     "logs:PutLogEvents",
        "logs:DescribeLogStreams"
    ],
      "Resource": "arn:aws:logs:*:*:*"
    }
  ]
}
```

#AWS #Cloud #Databases #Deployment #InterSystems Business Solutions and Architectures #Mirroring #InterSystems IRIS

Source URL:https://community.intersystems.com/post/intersystems-iris-deployment%C2%A0guide-aws%C2%A0using-cloudformation-template