



CloudDat for AWS: User Deployment Guide

Introductory Material

Introduction

CloudDat uses the Emmy Award winning ExpeDat accelerated file transfer software to move files from anywhere in the world to and from your S3 bucket, EBS, or instance storage at up to **5 Gbps** per object. Upload terabytes of data in just a few hours for just a few dollars. Deployment takes only minutes, and includes easy to use clients for Windows, Mac, Linux, and more. Data Expedition, Inc. is an APN Technology Partner and a member of Marketplace since 2013.

Use Cases

Any transfer of data to or from S3, EBS, or Instance storage where speed and reliability are important is appropriate for CloudDat.

- Digital asset archiving
- Lift and shift cloud migration
- Video transcoding and distribution
- Backup and recovery
- Biomedical research
- Customer job submission
- Integration with third-party hybrid storage devices

Public case studies, including several involving CloudDat for AWS, can be found at: <https://www.dataexpedition.com/customers/casestudies.html>

Overview of Typical Customer Deployment

CloudDat for AWS is typically deployed as an AWS Marketplace AMI via a CloudFormation stack. Customers may choose among hourly, hourly-monthly, or annual payment models, with hourly the most common. All resources and services, including EC2 instances, S3 storage, and security profiles are owned and controlled by the customer at all times. DEI has no access to customer data.

Initial setup and configuration typically takes about 10 minutes, less for experienced AWS users. Customer specifies the deployment region, target S3 bucket, and initial client credentials. Once CloudFormation has launched the EC2 instance, users can download the client software from that EC2 instance and begin browsing the S3 bucket and transferring data. A step-by-step guide to deployment is available here: <https://www.dataexpedition.com/clouddat/aws/installation.html>

Prerequisites and Requirements

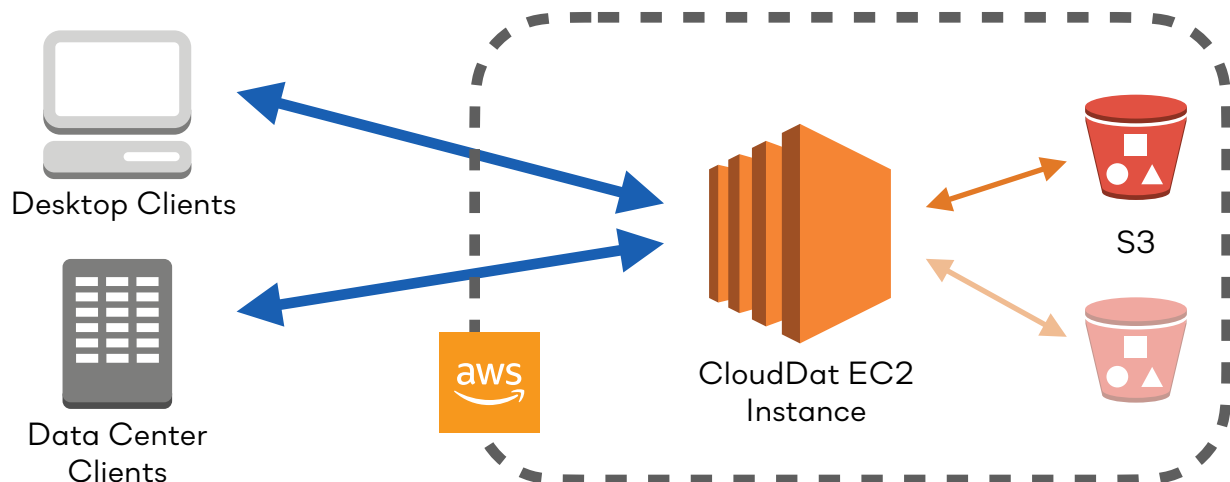
All of the CloudDat software, including server and clients, is contained in the marketplace AMI. Only novice AWS skills are required for basic deployment. Familiarity with the Amazon Linux command line environment is required for advanced configuration. To use CloudDat, customers will need:

- An Amazon Web Services account
- An S3 bucket
- EC2 SSH Keys for the bucket's region
- A high-speed Internet path to Amazon Web Services
- A Secure Shell (SSH) client (optional, for advanced configuration)

Client software is downloaded from the CloudDat EC2 instance and supports Windows, macOS, Linux, FreeBSD, and Solaris. Anyone familiar with FTP, Secure Copy, or similar file transfer software can use CloudDat.

Architecture Diagram

CloudDat client software on desktops and in data centers uses Data Expedition, Inc.'s proprietary data transport technology to send and retrieve object data across the internet to the CloudDat EC2 instance at maximum speed regardless of distance, latency, or congestion. The EC2 instance streams the data to one or more S3 buckets in the same region using multipart S3 APIs in realtime, without storing it on the instance.



Planning Guidance

Security

All components are owned and managed by the customer. Customer determines what, if any, VPCs, subnets, Elastic IPs, S3 buckets, etc. may be used. DEI does not have access to customer data.

The CloudFormation stack creates a default Security Group and IAM role to allow internet clients access to the EC2 instance and the EC2 instance access to S3. The security group allows ports with an option to restrict access to only certain client addresses:

- TCP/22 (SSH)
- TCP/80 (HTTP)
- UDP/8080 (ExpeDat)

The IAM role permits the following S3 actions on the specified bucket:

- AbortMultipartUpload
- DeleteObject
- GetBucketLocation
- GetObject
- ListBucket
- ListBucketMultipartUploads
- ListMultipartUploadParts
- PutObject

Customer may modify the Security Group and IAM role using standard AWS tools.

If deployed as a Marketplace Single AMI, the customer must assign their own Security Group and IAM role. The access specifications above are recommended but not required. The minimum requirement is one UDP port on any available port number.

Client access is controlled by a username and password, which is configured by the customer during initial deployment. Credentials are always encrypted between the client and CloudDat EC2 instance. Additional username/password pairs may be created via the advanced configuration steps shown here:

<https://www.dataexpedition.com/clouddat/aws/advanced.html#users>

Content encryption between the CloudDat Clients and the CloudDat EC2 instance may be enabled by checking "Encryption" in the graphical clients or via the "-K" option in the command line clients. User credentials are always encrypted regardless of this setting.

SSE encryption may be enabled via the advanced configuration steps shown here:

<https://www.dataexpedition.com/clouddat/aws/advanced.html#sse>

Costs

CloudDat Hourly from AWS Marketplace is billed at \$1.50 per instance hour. This allows up to 4 simultaneous operations per instance.

CloudDat Monthly from AWS Marketplace is billed at \$400 per month plus \$0.50 per instance hour. This allows up to 4 simultaneous operations per instance.

CloudDat Monthly High Capacity from AWS Marketplace is billed at \$400 per month plus \$1.00 per instance hour. This allows up to 32 simultaneous operations per instance.

AWS costs associated with a typical CloudDat deployment include:

- EC2 instance fees for the CloudDat EC2 instance based on the type
- EBS fees associated with the CloudDat EC2 instance
- EBS fees associated with any additional EBS storage accessed by CloudDat
- S3 fees associated with any S3 buckets accessed by CloudDat
- IOPS fees associated with any EBS or S3 storage accessed by CloudDat
- Egress fees associated with client data download

For current AWS fees, see: <https://aws.amazon.com/pricing/>

Sizing

Speed to and from each CloudDat instance is limited by the client bandwidth and the EC2 instance type. This varies considerably but as of this writing the maximum upload speeds and numbers of concurrent uploads to be expected for each instance type are:

Instance Type	Max Upload	Max Concurrency
c5n.large	1 Gbps	4
c5n.xlarge	2 Gbps	8
c5n.2xlarge	3 Gbps	16
c5n.4xlarge	5 Gbps	32
c5n.9xlarge	5 Gbps	64

Speeds are based on encrypted uploads over the internet to S3 storage. AWS throttles EC2 internet bandwidth to 5 gigabits per second per instance. Multiple instances may be used for faster total bandwidth.

Maximum concurrency is based on S3 buffering requirements and available instance memory.

Regions which lack support for c5n instances may use c5 instead.

Deployment Assets

The recommended deployment method is to select a CloudDat Marketplace offering and launch via CloudFormation using the steps below. Canonical instructions can be found here:

<https://www.dataexpedition.com/clouddat/aws/installation.html>

1. Verify that you have all of the requirements listed in Prerequisites and Requirements above.
2. Choose an S3 account and bucket to access.
Note which region hosts the bucket. Buckets in the "Standard" region are "US East (N. Virginia)".
3. Go to the [AWS Marketplace](#) and choose the Hourly, Monthly, or Monthly High Capacity product. Choose the Hourly product for the 7-day free trial. CloudFormation may not be available on the subscription management page: start from the AWS Marketplace page instead.
4. Click the Continue button at the top of the page to proceed to the Launch page.



Continue to Subscribe

5. If this is your first time subscribing to this product, or if you have previously cancelled your subscription, click the "Accept Software Terms" button.



Accept Software Terms

6. The next screen will verify that you are subscribed and may display other offers, terms, and details. Click "Continue to Configuration".



Continue to Configuration

7. Select a "Delivery Method" by clicking on the dropdown menu and choosing "CloudFormation Template". If drop-down menus interfere with your ability to make a selection, try making your browser window wider.

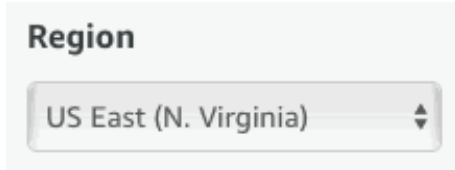
Delivery Method

CloudFormation Template

CloudDat

8. Verify that the latest "Software Version" is selected.

9. In the "Region" box, select the AWS region where your S3 bucket is located.



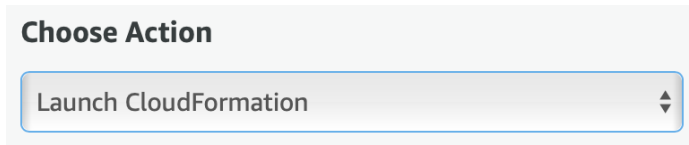
The image shows a dropdown menu titled "Region". The selected option is "US East (N. Virginia)".

10. Click "Continue to Launch" at the top of the page.



The image shows a blue button with the text "Continue to Launch".

11. In the "Choose Action" box, select "Launch CloudFormation".



The image shows a dropdown menu titled "Choose Action". The selected option is "Launch CloudFormation".

12. Click the "Launch" button.



The image shows an orange button with the text "Launch".

13. You will now be taken to the start of the launch workflow. Verify that the default Amazon S3 URL is selected as the Template source, then click "Next" to continue.

Template source

Selecting a template generates an Amazon S3 URL where it will be stored.



The image shows a selection box with a radio button next to the text "Amazon S3 URL". The radio button is selected.

[Continued on the following page]

14. Now you can configure your CloudDat instance. Most settings may be left with their default values or the first option in the drop-down menu.

Specify stack details

Stack name

Stack name

Stack name can include letters (A-Z and a-z), numbers (0-9), and dashes (-).

Parameters

Parameters are defined in your template and allow you to input custom values when you create or update a stack.

Instance Configuration

Instance Type
EC2 Instance type: larger instances provide more bandwidth.

Network Configuration

VpcId
VpcId of an existing Virtual Private Cloud (VPC).

Subnet
The Subnet in the Virtual Private Cloud (VPC).

Key Pair
Click to select an EC2 KeyPair for administrative SSH access to the CloudDat instance.

Restrict Client Access
Only allow client access to the CloudDat instance from systems matching this network pattern.

Restrict Admin Access
Only allow SSH (admin) access to the CloudDat instance from systems matching this network pattern.

CloudDat Configuration

Existing S3 Bucket
Name of the existing S3 bucket which CloudDat will access.

ExpeDat User Name
Username for your initial user account.

ExpeDat Password
Password for your initial user account. See documentation for adding more accounts later.

Cancel Previous Next

Stack name: Optionally name this collection of CloudDat resources.

Instance Type: Select a c5n type based on your desired performance (see "Sizing" above). Some newer AWS regions may not support c5n, in which case c5 may be used instead.

VpcId: If you do not already have a VPC, choose any of the default values.

Subnet: If you do not already have a VPC, choose any of the default values.

Key Pair: Choose the SSH key which you will use to access the CloudDat instance for administration and advanced configuration.

Restrict Client Access: Set a subnet mask to limit file transfers, or enter "0.0.0.0/0" to allow file transfers from anywhere.

Restrict Admin Access: Set a subnet mask to limit SSH access, or enter "0.0.0.0/0" to allow admin access from anywhere.

Existing S3 Bucket: Type the name of your S3 bucket. Do not enter the endpoint URL, just the name.

ExpeDat Username: Choose a name for your initial file transfer user account.

ExpeDat Password: Select a password for your initial file transfer user account.

Click "Next" once you have completed the configuration form. If you receive errors, check that you have completed all form elements according to the instructions above and try again. Remember that you can always contact us for help.

Next

15. There are no changes required on the Options page, click "Next" again. Do not enter an IAM Role here: one will be created for you, which you can modify later.

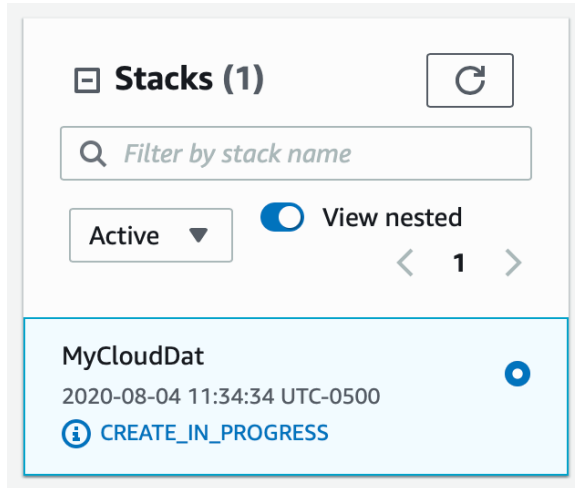
Next

16. On the Review page, click the checkbox at the bottom to allow the creation of an IAM role which will allow the CloudDat instance to access the bucket you named. Then click "Create".

I acknowledge that AWS CloudFormation might create IAM resources.

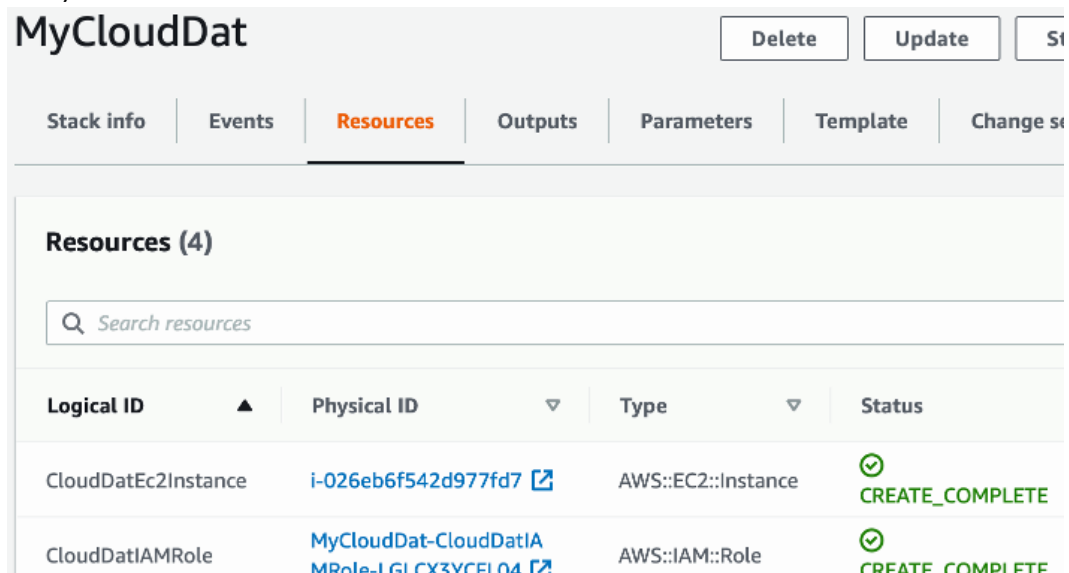
Create stack

17. You will now be taken to the Cloud Formation page. This page sometimes loads before the CloudDat stack has started and may say "Showing 0 stacks". If that happens, click the Reload button at the top right to find the new CloudDat stack.



The screenshot shows the AWS CloudFormation console. At the top, there is a 'Stacks (1)' section with a refresh button. Below it is a search bar labeled 'Filter by stack name'. There is a dropdown menu set to 'Active' and a toggle for 'View nested' which is turned on. A pagination indicator shows '< 1 >'. The main content area displays one stack: 'MyCloudDat', created on '2020-08-04 11:34:34 UTC-0500'. The stack's status is 'CREATE_IN_PROGRESS', indicated by a blue progress indicator and an information icon.

18. It may take several minutes for AWS to deploy the CloudDat resources. Once the Status shows "CREATE_COMPLETE", select the CloudDat stack and open the "Resources" section. Look for the the instance id in the "Physical ID" column and click on it.



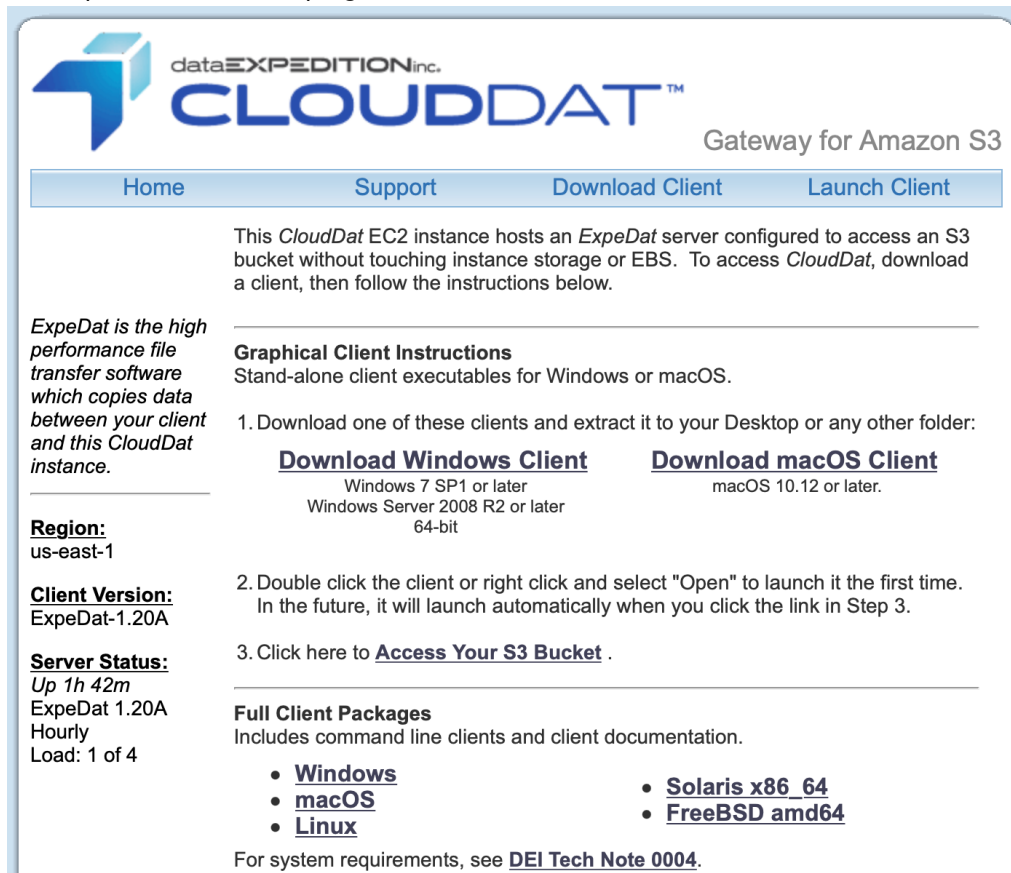
The screenshot shows the 'MyCloudDat' stack page in the AWS CloudFormation console. The 'Resources' tab is selected. The page title is 'MyCloudDat' with 'Delete', 'Update', and 'Stack' buttons. Below the title are tabs for 'Stack info', 'Events', 'Resources', 'Outputs', 'Parameters', 'Template', and 'Change set'. The 'Resources (4)' section has a search bar 'Search resources'. A table lists the resources:

Logical ID	Physical ID	Type	Status
CloudDatEc2Instance	i-026eb6f542d977fd7	AWS::EC2::Instance	✔ CREATE_COMPLETE
CloudDatIAMRole	MyCloudDat-CloudDatIAMRole-1GI6X3VCF104	AWS::IAM::Role	✔ CREATE_COMPLETE

19. You should now see the the CloudDat instance in the EC2 console. Once the Instance State reaches "running", copy the Public IP address to a web browser.

Instance ID	Instance Type	Availability Zone	Instance State	IPv4 Public IP
i-0eff77ec4af580541	c4.2xlarge	us-east-1d	● running	174.129.126.185

20. Your web browser should now show the CloudDat instance landing page. If the page does not load, check that the instance is running and that its security group allows TCP port 80. If the page displays yellow warnings, or the left-side "Server Status" is red, contact us for help. See below for an example of what the page should look like.



dataEXPEDITION inc.
CLOUDDAT™
Gateway for Amazon S3

Home Support Download Client Launch Client

This *CloudDat* EC2 instance hosts an *ExpeDat* server configured to access an S3 bucket without touching instance storage or EBS. To access *CloudDat*, download a client, then follow the instructions below.

ExpeDat is the high performance file transfer software which copies data between your client and this CloudDat instance.

Region:
us-east-1

Client Version:
ExpeDat-1.20A

Server Status:
Up 1h 42m
ExpeDat 1.20A
Hourly
Load: 1 of 4

Graphical Client Instructions
Stand-alone client executables for Windows or macOS.

- Download one of these clients and extract it to your Desktop or any other folder:
 - Download Windows Client**
Windows 7 SP1 or later
Windows Server 2008 R2 or later
64-bit
 - Download macOS Client**
macOS 10.12 or later.
- Double click the client or right click and select "Open" to launch it the first time. In the future, it will launch automatically when you click the link in Step 3.
- Click here to [Access Your S3 Bucket](#) .

Full Client Packages
Includes command line clients and client documentation.

- [Windows](#)
- [macOS](#)
- [Linux](#)
- [Solaris x86_64](#)
- [FreeBSD amd64](#)

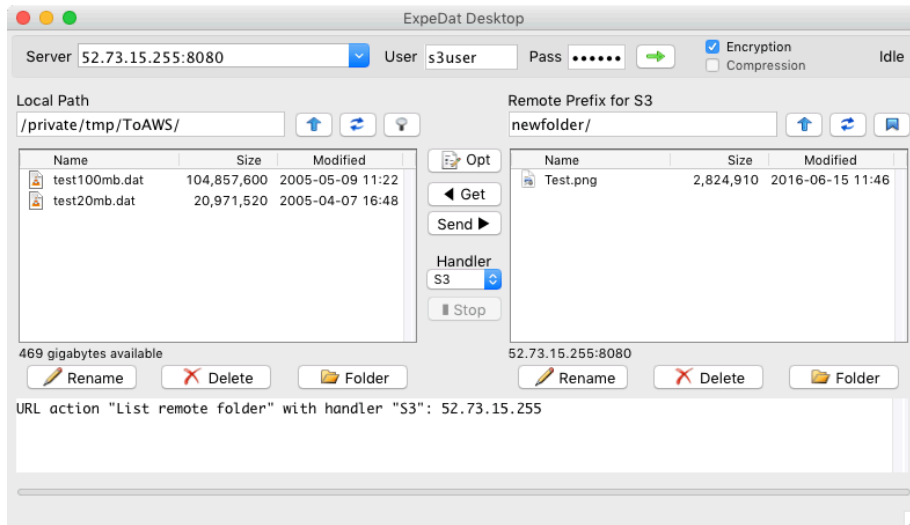
For system requirements, see [DEI Tech Note 0004](#).

21. Download a client.

Choose Windows or Mac for the graphical client and extract the contents of the ZIP package after download. A command line client is included in the "Full Client Package" links near the bottom of the page.

22. Run the client.

No installation is required. You may move the client executable to any location on your system before running it.



23. Back in the web browser, click "Access your S3 Bucket". The ExpeDat Desktop client will prompt you to enter the password you created earlier. For command line clients, see the Client Use page.

After clicking the Access link, the client will list the contents of your S3 bucket (if any) and you can begin transferring files. You may wish to bookmark the link for easy access later. The next page describes more about using clients. See the Support page if the client has trouble reaching the CloudDat instance.

Additional guidance for usage of the client software can be found here:
<https://www.dataexpedition.com/clouddat/aws/clients.html>

Operational Guidance

Health Check

Use the CloudDat EC2 instance landing page, as shown in step 20 above, to verify that the CloudDat instance is functional and monitor its load level.

Logging and diagnostic data CloudDat server activity can be found on the CloudDat EC2 instance at the path `/var/log/servedat.log`. Logging and diagnostic data for S3 access can be found in `/tmp/s3logs`.

Additional guidance for troubleshooting CloudDat operation can be found here:
<https://www.dataexpedition.com/clouddat/aws/support.html>

Backup and Recovery

Basic configuration entails only the information entered in step 14 shown above. A screenshot or other recording of this information is sufficient to recreate a new CloudDat instance of the same basic configuration.

Basic and advanced configuration of CloudDat is stored in the following files on the CloudDat EC2 instance:

```
/etc/servedat.cf  
/etc/svpasswd  
/usr/local/expedat/s3handler.sh
```

Make backups of these three files, plus any additional custom s3handler files you may have created. Restore these files to a new CloudDat instance to recover functionality.

Basic configuration stores no user data on the instance. If the configuration is modified to allow uploading and downloading to instance or EBS storage, additional steps must be taken to backup that data.

Routine Maintenance

No routine maintenance is required.

Is recommended that CloudDat EC2 instances be migrated to new versions as they become available using the methods discussed in Backup and Recovery above. This ensures that both the Amazon Linux operating system and the CloudDat software receive the latest updates.

Emergency Maintenance

In the event a CloudDat EC2 instance becomes inoperable and unrecoverable, create a new instance and recover the settings using the methods discussed in Backup and Recovery above.

Troubleshooting & Support

CloudDat is fully supported by Data Expedition, Inc. for the duration of the subscription. A troubleshooting guide and additional documentation can be found on DEI's website: <https://www.dataexpedition.com/clouddat/aws/support.html>

Customers needing assistance can contact DEI here: <https://www.dataexpedition.com/support/>

Customers will typically receive a response within one business day.

Support Costs

Support is included with the CloudDat Marketplace subscription.

Accessibility

Reference Materials

The following page contains complete deployment and operating instructions for CloudDat for AWS:

<https://www.dataexpedition.com/clouddat/aws/>

Localization

This user guide is available in English.