Table of Contents

Introduction

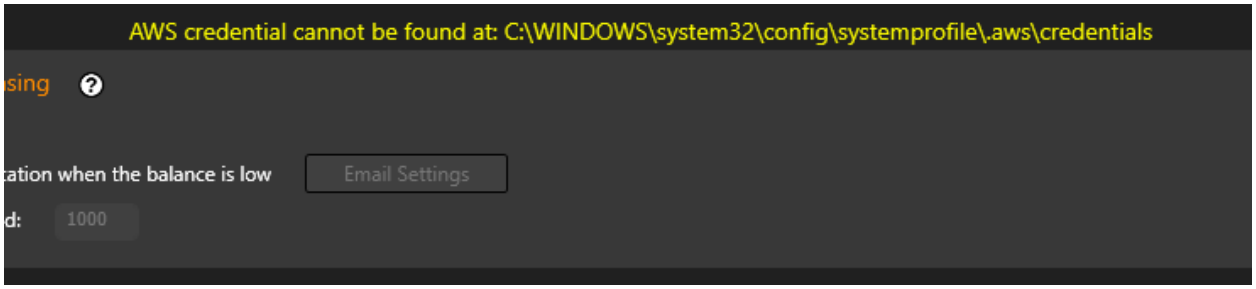
Cloud computing is a popular trend and is favored in many industries, including transcoding. With continuous improvement of cloud infrastructure, support, and user experience, such as AWS and Linode, many users have started to look into cloud transcoding solutions. It has low deployment cost, no long-term commitment, is flexible, etc. However, it is often very costly in the long term for workflows that have stable and continuous transcoding. Media files also take up significant storage space and require high internet bandwidth. Also, many users have security concerns, and their media assets cannot be stored in cloud environments. Hence, many workflows still favor transcoding locally with physical hardware.

For local transcoding solutions, the setup is often fixed and not very flexible to change in capacity. When there is a surge in transcoding needs, backlogs will be created temporarily. If not, the transcoding system has to be over-built to handle the surge as well, resulting in huge initial deployment costs.

Cambria Cluster with the Cloud Extend feature is here to solve the problem smartly and efficiently. When the local transcoding system is overloaded, it will launch new pre configured instances automatically. These instances will offload the additional transcoding jobs and shut down when the jobs are done. The integration is aimed at being as seamless as possible. With bidirectional S3 mapping, the instances will be able to access source files and write back transcoded output to local drives.

Amazon AWS Setup

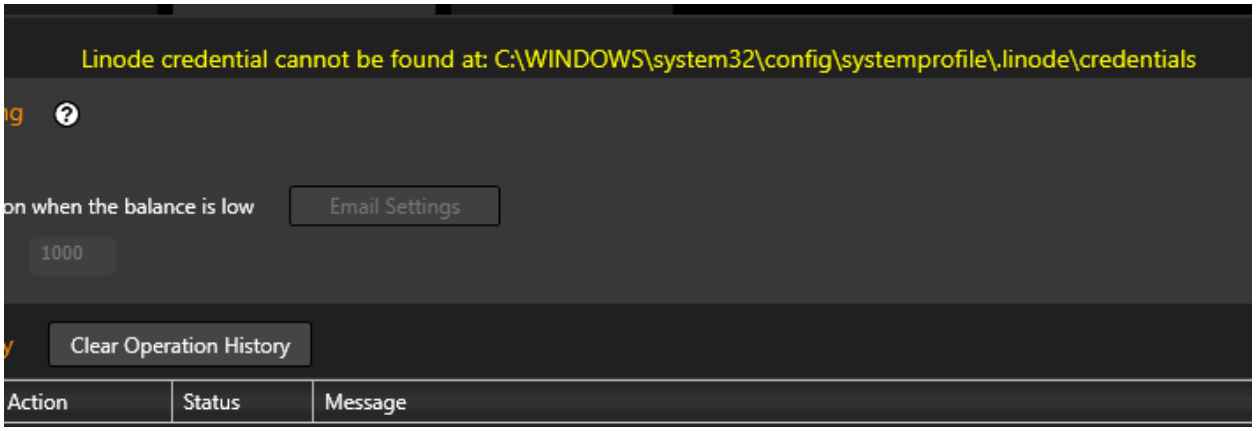
- **Credential and S3/EC2 Setup**
  - If you have not set up credentials for AWS yet, then you should see this warning. Follow the document below to learn how to set it up



  - Dropbox PDF: https://www.dropbox.com/s/3sciafbq5s0k588/Amazon%20AWS%20Credentials%20and%20S3_EC2%20Setup%20Guide.pdf?dl=0
  - If the setup of the "credentials" file is proper, you will observe no warning, otherwise, a warning will be shown:
- **AMI Image**
  - An AMI Image is a saved virtual machine that CloudExtend will use to create instances; allowing Cluster to connect to the machine and run queued jobs
  - An AMI Image with FTC will be needed to be created in order to use CloudExtend
  - Learn how to generate an AMI image follow this guide:
  - https://www.dropbox.com/s/e0nue54nn6hgqrk/Amazon%20AWS%20AMI%20Creation%20and%20Editing.pdf?dl=0
- **AWS Instances Settings**
  - Prepare following information to add AWS instances to Cluster:
  - AMI ID (starts with "ami-"). Ex. ami-010cd39cf6463e144
  - If testing, you can use this Capella created AMI: **ami-0d250373c58e8fadf** (This is updated after every release because we delete old versions. Currently this is for FTC 5.2)
  - Choose an Instance Type
    - Amazon EC2 Instance Types - Amazon Web Services
    - Use a fast instance type, recommend to start with c5.xlarge
    - ARM based CPUs is not supported
  - Enter a Security Group ID (eg, "sg-0586a52429f0b355a")
  - **To Create a Security Group using AWS:**

        - Open the Amazon EC2 at https://console.aws.amazon.com/ec2/. Make sure that you are logged in to your account.
        - In the navigation pane on the left, choose **Security Groups** under **Network & Security**.
        - Choose **Create Security Group**.
        - Enter a name for the security group (for example, my-security-group) and provide a description.
        - Specify your **Inbound rules** to Type: **All Traffic** and CIDR blocks: **0.0.0.0/0**
        - Click **Create security group**
        - Now look for the security group that you just created and enter it into the **Security Group ID** field.
    - Subnet ID
        - This section can be blank.
    - Key Pair Name
        - Allows remote access to instance
        - Use that as the Key Pair that you created when creating the AMI or make a new one
    - **To Create a Key Pair using AWS Console:**
        - In the navigation pane on the left, choose **Key Pairs** under **Network & Security**.
        - Choose **Create key pair**.
        - Enter a name for the key pair.
        - Set **key pair type** to RSA and **Private key file format** to **.pem.**
        - Then click **Create key pair.**
        - Make sure to keep this .pem file somewhere safe.
    - Instance Name
        - Eg “AWS-C5xLarge-*”, then new instances will be created and automatically named AWS-C5xLarge-0, AWS-C5xLarge-1 and so on
    - Region (eg, “us-west-2”)
        - Regions and Zones - Amazon Elastic Compute Cloud
        - If you are using the Capella created AMI please set region to “us-west-2”
    - Number of instances
    - When “Add” is clicked, instances will be automatically launched in AWS portal and added into Machines tab

Akamai Linode Setup



- If you have not set up credentials for Linode yet, then you should see this warning. Follow the document below to learn how to set it up
- **Credentials Setup**
    - Dropbox PDF: https://www.dropbox.com/s/gk9fr8wwkloj22i/Akamai%20Linode%20Credentials%20and%20Instance%20Setup%20Guide.pdf?dl=0
    - If the setup of the "credentials" file is proper, you will observe no warning, otherwise, a warning will be shown:
- **Linode Image**
    - Please contact Capella Support to receive information on how to setup a Linode Image
    - You will need to follow the Linux installation guide and save the Linode instance as an image
- **Linode Instance Settings**
    - Prepare following information to add Linode instances to Cluster:
    - For Instance Type and Region you can either use command prompt to find out the information, or look at the tables provided here on the third page: https://www.dropbox.com/s/gk9fr8wwkloj22i/Akamai%20Linode%20Credentials%20and%20Instance%20Setup%20Guide.pdf?dl=0
    - If you are using command prompt, follow this guide on how to install linode-cli to your computer https://www.linode.com/docs/products/tools/cli/guides/install/#install-python-3-and-pip3
    - Choose an Instance Type
        - To find the different types of instances in command prompt use **linode-cli linode types**

        - You can also use the table provided in the Setup Guide
    - Image ID
        - Image ID (starts with “private/”). Ex. private/14585314
        - To find your Image ID use command “linode-cli images list”
        - An alternative way is to go to **Images** on the Linode website and select the dropdown arrow on the image that you want to create an instance of
        - Next, you click **Deploy to New Linode** and the image ID will be in the URL
        - Ex. https://cloud.linode.com/linodes/create/?type=Images&imageID=**private/410965**
    - Root Password
        - Create a password for your instance
        - Password length must be 7-128 characters
    - Instance Name
        - Name the instance any unique name
        - You cannot have two instances with the same name, or else you will receive an error
    - Region
        - To find the different types of instances use command “linode-cli regions list”
        - You can also use the table provided in the Setup Guide
    - When “Add” is clicked, instances will be automatically launched in Linode portal and added into Linodes tab

Instance Controller

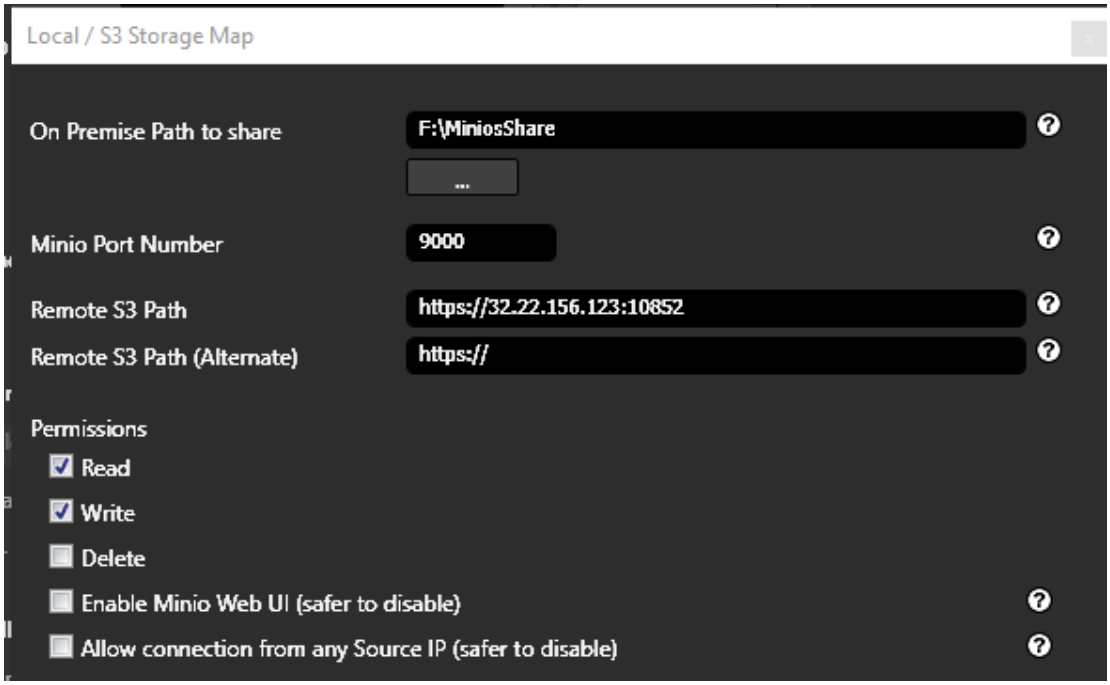
- A sample is already provided and installed with Cluster:
    - C:\Program Files (x86)\Capella\CambriaCluster\cpx64\AwsDynInstCtrl.exe
    - C:\Program Files (x86)\Capella\CambriaCluster\cpx64\LinodeDynInstCtrl.exe
- It facilitates automation of launching new instances when necessary, and stop/terminate the instances when unnecessary.
- Recommended to call every 30 minutes
    - Too short will have lots of overhead, as instances take time to launch
- Example with AwsDynInstCtrl (Linode works the same, just replace Aws with Linode)
- Usage: AwsDynInstCtrl [--create X --limit Y [--slot W]] [--stop Z] [--debug 1]
    - --create X: create instance if there are more than X queued jobs, to have sufficient total transcoding slots to handle all the queued jobs, until limit reached
    - --limit Y: limit to create Y instances only
    - --slot W: Default 2, control the slots of the new instance.

- --stop Z: Set all instances to 0 slot when there are Z or less queued jobs
- --debug 1: more debugging information

- eg: AwsDynInstCtrl.exe --create 20 --limit 1 --stop 5
  - Create 1 instance if there are more than 20 queued jobs
  - Set all instances to 0 slot when there are 5 or less queued jobs
  - Cluster service will automatically stop/terminate those instances if configured

CloudExtend Storage (S3 Mapping)

- This allows instances to access to local storage
  - This section can be skipped if your whole transcoding farm is purely in Amazon AWS (including Cluster), or, purely using direct read/write in S3 storage
- Configure via "CloudExtend Storage (On Premise to S3 mapping)", eg:



  - On Premise Path: The shared storage location (includes subdirectories)
    - **Note that source files/output files must be in subdirectories**
    - **Ex. F:\MiniosShare\source.mp4 is forbidden**
    - **What will work is F:\MiniosShare\Output\source.mp4**
  - Minio Port Number: the Minio server port. Ex. 9000.
  - The port number can be anything; it just needs to be opened up through your network settings. To do this you need to port forward your computer's IP Address with the port number.

- Remote S3 Path: how the instance can communicate to this Cluster Minio mapped S3 storage
  - You can get the ip address by looking at **What is my IP Address** on Google.
  - The port number that is attached to your IP address will be the port that you opened up.
  - Ex. instance will communicate with Cluster Minio Server via 32.22.156.123:10852
  - The router at 32.22.156.123 will route 10852 to 9000 of the Cluster machine
  - Remote S3 Path (Alternate): how an instance can communicate to the Cluster Backup (if configured). This follows same spec as Remote S3 Path, eg 32.22.156.123:10853
- The router at 32.22.156.123 will route 10853 to 9000 of the Cluster Backup machine

Pay As You Go

- Once you have a PAYG license, you will be able to see your current balance in the following locations:
  - Cluster CloudExtend Tab under **Pay As You Go licensing**
  - In the **Selected Machine AWS Info** summary of any of the AWS machines in your machine list in Cluster
- How are credits deducted from your Pay As You Go balance?
  - Amount of credits deducted depends on source duration (in seconds)
    - Amount of credits deducted changes if using stitching or applying source filters that change the source's duration
    - Amount of credits deducted is doubled if encoding to multiple targets or multiple layers
  - The credits are deducted only after a job has finished
  - The Cluster UI takes 2-3 minutes to update the new balance
  - Queue jobs with a balance of 0
    - You should get an error: No balance: License Error: Balance is empty for pay-as-you-go license

Remarks and Troubleshooting

- Instances take a while to start up. You can refer to the Machines tab to know the estimated ready time. Cambria Cluster enforces a minimum 10 minutes waiting time for each newly launched instance to avoid jobs to be assigned prematurely and causing errors.
- When creating an instance make sure that it shows up in the Machine tab
- Make sure the instance is shown in the Machine tab. It should show up as the Instance Name that you put in CloudExtend.

- If the machine is showing offline, make sure you have waited long enough ("!" icon will show estimated time)
- Make sure available slot is more than 0
    - Be default, new instances added has 2 default slots
- The machine is required to be set as Online in order to be assigned with jobs
    - By default, new instances added via Cluster UI is added to Cluster (so it is either Online or Offline, but will not be Standby/Cold Standby)
- Instances, if left idle, or even stopped, will still incur charges
    - "Terminate"or "Delete" instances when it is not used
    - Setup usage warning/monitoring notification in AWS portal
- The temp output folder for S3 Mapping redirection is C:\Users\Public\Documents\CapellaOutput, so C: has to be large
- HTTP GET from Cluster machine to instances via "Machine IP" should return proper xml, eg: https://34.123.186.99:8648/CambriaFC/v1/SystemInfo
    - Make sure you have configured Security Group properly, such that the instance is reachable from Cluster
    - Make sure that the internet connection is "Private" or "Domain" and not "Public". Otherwise, during FTC/Cluster installation, you will need to enable firewall rules for "Public" as well. You can verify that this is the culprit by disabling all Windows Firewall for Private, Domain and Public.
- Source files/output files are not mapped properly
    - Or, encountered these errors:
        - Unable to connect to endpoint
    - Make sure CloudExtend S3 Mapping is correct
    - **As an example**, we have such local directories:
        - F:\MiniosShare\Source → source files
        - F:\MiniosShare\Output → transcoded outputs
        - 32.22.156.123:10852 → Cluster public IP, firewall configured properly
        - Hosting Minio web server at port 9000
        - Routing external port 10852 to Cluster machine IP and port 9000
- Slow transcoding performance
    - Make sure Instance Type is set to a strong enough machine with sufficient CPU cores and system memory
    - Lower concurrent transcoding slots

- Have fast internet connection to the instances
- Make sure the the Region is set close to you

Q/A

- How to revoke the configured “Cloud Extend Storage”?
  - The snapshot of them are stored in encrypted form in client’s job
  - Remove the configuration from the Cluster. Existing credentials that are stored in the client's job will be invalidated.
  - Hence, if you remove the configuration, and add the “same” configuration again, the old one will still be invalidated.
- The “Cloud Extend Storage” apparently does not apply the settings in Credential Manager for “On Premise Path”
  - This is the limitation as of now. To workaround, logon “Capella CpClusterServiceManager” with an administrator account that has access to the “On Premise Path”