AWS SETUP

The AWS setup is very similar to PA0.

For launching multiple instances, there are three differences.

- Firstly, we will specify 5 instances instead of 1.
- One of these instances will run our Jupyter Notebook and Dask Scheduler, and the remaining 4 instances will run our Dask workers.
- Each of these instances will have 100GB SSD storage instead of 40GB.
- We will create a new security group for our 5 instances that allow each of the instances to communicate with each other. Follow the below steps one by one.
- Access your ETS account using single sign-on ID: https://ets-apps.ucsd.edu/individual/DSC102_SP23_A00/. To
 open the AWS console click "Click here to access AWS" at the bottom of the page. To
 get your AWS credentials for CLI / API usage click "Generate API Keys (for CLI/scripting)".

2. Open AWS Dashboard. We will first create a new security group so that we can apply it to all of our instances later. Click on "Security Groups" on the left menu.

| New EC2 Experience X Tell us what you think | Resources | | | | EC2 Global view 🖾 🛛 C | Account attributes |
|--|--|----------------------|----------------------------|----------------------|------------------------|---|
| 2 Dashboard | You are using the following Amazon EC2 n | esources i | n the US West (Oregon) Re | egion: | | Supported platforms |
| 2 Global View | ····· | | | | | |
| vents | Instances (running) | 1 | Auto Scaling Groups | 0 | Dedicated Hosts | 0 Default VPC P |
| igs | Elastic IPs | 0 | Instances | 1 | Key pairs | 1 vpc-0cfe9d74 |
| nits | Load balancers | | Placement groups | 0 | Security groups | Settings |
| tances | Lond Deterrors | | r tacement groups | v | Security groups | EBS encryption |
| tances | Snapshots | 0 | Volumes | 1 | | Zones |
| tance Types | | | | | | EC2 Serial Console |
| unch Templates | | | | | | Default credit specification |
| ot Requests | Launch instance | . de la de la se cul | stud serves in the sloud | Service health | C AWS Health Dashboard | Console experiments |
| vings Plans | To get started, launch an Amazon EL2 Instance, i | which is a vi | rtual server in the cloud. | | | |
| served Instances | | | 2 | Region | Status | Explore AWS |
| edicated Hosts | Launch Instance Migrate a | server L | <u> </u> | US West (Oregon) | ormally | |
| heduled Instances | Mater Vers instances will brough in the UC Meet / | Orman) De | | | | Get Up to 40% Better Price Performance |
| apacity Reservations | Note: Tool instances with faulter in the OS west (| oregon/ Re | gion | Zones | | T4g instances deliver the best price performance for burstable generation |
| nages | | | | | | purpose workloads in Amazon EC2. Learn more 🔀 |
| MIs | Scheduled events | | C | Zone name | Zone ID | Save up to 90% on EC2 with Spot Instancer |
| 1I Catalog | | | | us-west-2a | usw2-az1 | Optimize price-performance by combining EC2 purchase options in a |
| astic Block Store | US West (Oregon) | | | us-west-2h | usw7-az7 | single EC2 ASG. Learn more 🖸 |
| lumes | No scheduled events | | | | | Enable Best Price-Performance with AWS Graviton2 |
| apshots | | | | us-west-2c | usw2-az3 | AWS Graviton2 powered EC2 instances enable up to 40% better price |
| ecycle Manager | Migrate a server | | | us-west-2d | usw2-az4 | performance for a broad spectrum of cloud workloads. Learn more |
| twork & Security | | | | Enable additional Zo | nes | |
| ecurity Groups | Use AWS Application Migration Service to | simplify a | and expedite | | | Additional information [2] |
| stic IPs | migration from physical, virtual, and cloud | d infrastru | cture to AWS. | | | |
| cement Groups | Get started with AWS Application Migratio | on service | 6 | | | Getting started guide 🔀 |
| | | | | | | |

3. Click on create security group.

DSC 102 : PA 1 : System Setup Tutorial

| 9 S3 | | | | | | | | | |
|----------------------|-------------------------|----------------------|-----------------------|----------------|---|--------------------------------|-----------------|-------------------------------|----------------|
| New EC2 Experience X | Security Groups (14 | 4) Info | | | | C Actions 🔻 | Export security | / groups to CSV 🔻 Create secu | rity group |
| EC2 Dashboard | Q Filter security group | 15 | | | | | | < | 1 > @ |
| EC2 Global View | Name | | Security group name 🛛 | VPC ID | ▽ | Description \bigtriangledown | Owner | | Outbound rul |
| vents | - | sq-0dad36f4f438c53ae | launch-wizard-4 | vpc-0cfe9d74 | | launch-wizard-4 create | 387848045596 | 1 Permission entry | 1 Permission e |
| gs | | sg-0f444add093aec6f0 | launch-wizard-3 | vpc-0cfe9d74 🛂 | | launch-wizard-3 create | 387848045596 | 1 Permission entry | 1 Permission e |
| nits | — - | sg-01f92ec3101646b64 | launch-wizard-7 | vpc-0cfe9d74 🛂 | | launch-wizard-7 create | 387848045596 | 1 Permission entry | 1 Permission e |
| stances | — - | sg-088cb3d1b6e844bc5 | launch-wizard-2 | vpc-0cfe9d74 🔀 | | launch-wizard-2 create | 387848045596 | 1 Permission entry | 1 Permission e |
| stances | — – | sg-0c80958d7032f617c | launch-wizard-8 | vpc-0cfe9d74 🖸 | | launch-wizard-8 create | 387848045596 | 1 Permission entry | 1 Permission e |
| tance Types | | sg-013e271bdd0d1f5d0 | launch-wizard-5 | vpc-0cfe9d74 🔀 | | launch-wizard-5 create | 387848045596 | 1 Permission entry | 1 Permission |
| unch Templates | | sg-0b8a858426ff31df7 | dask_security_group | vpc-0cfe9d74 🔀 | | dask_sg | 387848045596 | 2 Permission entries | 1 Permission e |
| ot Requests | | sg-0f640ecc746d3a1b4 | launch-wizard-10 | vpc-0cfe9d74 🛃 | | launch-wizard-10 crea | 387848045596 | 1 Permission entry | 1 Permission e |
| vings Plans | — - | sg-82bc33dd | default | vpc-0cfe9d74 🛃 | | default VPC security gr | 387848045596 | 0 Permission entries | 0 Permission e |
| served Instances | | sg-0f01f717cb3410d17 | launch-wizard-1 | vpc-0cfe9d74 🛂 | | launch-wizard-1 create | 387848045596 | 1 Permission entry | 1 Permission e |
| dicated Hosts | _ | | | | | | | | - |
| eduled Instances | | | | = | | | | | |
| pacity Reservations | | | | | | | | | |
| ages | | | | | | | | | |
| dis | | | | | | | | | |
| MI Catalog | | | | | | | | | |
| in controly | | | | | | | | | |

4. Choose any name and description for your security group. Change both the inbound and outbound rules to have "Type=All Traffic", "Source/Destination = Anywhere-IPv4". See image below.

| | | [Option+S] | | | | N 4 | Oregon • | OrganizationAccountAccessRole/rram |
|--|--|---|--|---|------------------------------|-----|----------|------------------------------------|
| | | | | | | | | |
| rity group acts as a virtual firewa | all for your instance to control inbou | und and outbound traffic. To create a new | v security group, complete the fields | below. | | | | |
| | | | | | | | | |
| asic details | | | | | | | | |
| scurity group name Info | | | | | | | | |
| dask_security_group | | | | | | | | |
| me cannot be edited after creation. | | | | | | | | |
| 2scription Info | | | | | | | | |
| dask_security_group | | | | | | | | |
| PC Info | | | | | | | | |
| ට, vpc-0cfe9d74 | | × | | | | | | |
| | | | | | | | | |
| bound rules lofe | | | | | | | | |
| | | | | | | | | |
| .ype Into | Protocol Info | Port range Info | Source Info | | Description - optional Info | | | |
| All traffic | ▼ All | All | Anywhere-IPv4 ▼ | Q | | | De | lete |
| | | | | | | | | |
| | | | | 0.0.0.0/0 × | | | | |
| | | | | 0.0.0.0/0 × | | | | |
| Add rule | | | | 0.0.0.0/0 × | | | | |
| Add rule | | | | 0.0.0.0/0 × | | | | |
| Add rule | | | | 0.0.0/0 X |] | | | |
| Add rule utbound rules Info Ype Info | Protocol Infe | Portrange Info | Destination Info | 0.0.0/0 X | Description - optional linfo | | | |
| Add rule utbound rules info Yype info All taffic | Protocol Info | Port range info | Destination Info | 0.000/0 X | Description - optional Info | | | |
| Add rule utbound rules info | Protocol Info | Portrange Iolo All | Destination Info Anywhere-IPv4 V | Q Q | Description - optional Info | | De | lete |
| Add rule utbound rules info ype info All traffic | Protocol Infe | Portrange Iolo All | Destination Info Anywhere-IPv4 ¥ | 0.00.0/0 X | Description - optional Info | | | lete |
| Add rule info | Protocol luño V All | Portrange lofe | Destination Info | Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q | Description - optional Info | | Da | lete |
| Add rule info | Protocol info | Portrange Info All | Destination Infe Anywhere-IPv4 ¥ | Q 0.000/0 X | Description - optional Infe | | De | tete |
| Add rule info | Protocol Info | Portrange Infe All | Destination Info Anywhere-IPu4 ¥ | Q 0.00.0/0 × | Description - optional Info | | De | iete |
| Add rule into Type into All traffic . | Protocol Info | Portrange Iolo All | Destination Info Anywhere-IPv4 v | Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q | Description - optional Info | | De | Aete |
| Add rule with the second of the second rule with the second rule with the second rule seco | Protocol Info All resource. Each tag consists of a key and a | Port range Info All | Destination Info Anywhere-IPv4 ¥ diffeer your resources or track your AWS co | Q Q 0000/0 X | Description - optional Infe | | | tete |
| Add rule utbound rules into Type into All traffic Add rule age a subst that you satight to an AWS tags associated with the resource | Protocol Info Protocol Info All resource. Each tag consists of a key and a 2. | Port range Info All In optional value. You can use tags to search and | Destination Info Anywhere-IPv4 V d filter your resources or track your AWS oc | Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q | Description - optional Info | | | iete |
| Add rule into into into into into into into into | Protocol Inflo All resource. Each tag consists of a key and a 8. | Port range Info All | Destination inte Anywhere-IPv4 V d filter your resources or track your AVIS co | 0.00.0/0 X | Description - optional Info | | De | tee |

5. Lastly, click on "Create Security Group" at the bottom right.

6. Now, we will create our 5 instances which will use this new security group. We have setup the Dask environment on an AMI with the name "dsc102-dask-environment-public" Go to "AMIs" (under "Images") in your EC2 dashboard, select private images, and then search by name to find it. Select this AMI.

DSC 102 : PA 1 : System Setup Tutorial

Golokesh Patra

| New EC2 Experience Tell us what you think | Amazon Machine Images (AMIs |) (1/1) Info 22 Image Builder Actions | Launch instance from AMI | | |
|--|--|--|--------------------------------|-----------------------------------|------------|
| EC2 Dashboard | Public images Q Search | | | < 1 | 1 > ⊚ |
| EC2 Global View | dsc X Clear filters | | | | |
| Events | | | | | |
| Tags | Name AMI ID | | AMI name 🗸 🗸 | Source | Owner |
| Limits | ✓ – ami-087ab | 6cf5ee2ab526 | dsc102-dask-environment-public | 035170873046/dsc102-dask-environm | 0351708730 |
| ▼ Instances | | | | | |
| Instances | | | | | |
| Instance Types | | | | | |
| Launch Templates | AMI IDi ami 097ab6af6aa2ab6 | 26 | = | | a v |
| Spot Requests | AMI ID: alli-067ab6ci5eezab5 | 20 | | | U A |
| Savings Plans | Details Storage Tags | | | | |
| Reserved Instances | | | | D (1) (1) | |
| Dedicated Hosts | AMI ID 고 ami-087ab6cf5ee2ab526 | image type machine | Platform details Linux/UNIX | Root device type EBS | |
| Scheduled Instances | | | | | |
| Capacity Reservations | AMI name dsc102-dask-environment-public | Owner account ID | x86_64 | Usage operation RunInstances | |
| ▼ Images | Root device name | Status | Source | Virtualization type | |
| AMIs | /dev/sda1 | Available | 🗇 035170873046/dsc102 | -dask- hvm | |
| AMI Catalog | | | environment-public | | |
| | Root mode | State reason | Creation date | Karnal ID | |
| | | | | | |

- 7. After selecting the AMI, click "Launch Instance from AMI" as shown below.
- 8. Now, strictly follow the below instructions to launch the EC2 Spot instances.
 - a. Give any name for your instance.
 - b. Number of instances to launch is 5.
 - c. The instance type is "t2.xlarge".
 - d. Create a new key or use the key previously used for PAO.
 - e. Under "Network Settings", click "Select Existing Security Group" and choose the
 - name of the security group you just created.
 - f. Choose 100GB SSD gp2 storage.

g. Open advanced details. Select "Request Spot Instances". Then click on "customize" just on the right. Open the dropdown for "Request type" and select "One-time" and limit the "Set your maximum

price (per instance/hour) to as low as possible. (Start From 0.07)

9. Lastly, click "Launch Instance".

SETUP CLIENT AND SCHEDULER :

(For the 1st EC2 Instance)

Step 1 : Setup Client i.e Jupyter Notebook and Port Forwarding for Jupyter Notebook onto localhost.

| Name | ▼ Instance ID | Instance state abla Instance type abla Instance Instan | Status check | Alarm status Availability Zone 🔻 | Public IPv4 DNS 🛛 🗸 | Public IPv4 🔻 | 7 Elastic IP |
|------|---------------------|--|---------------------|----------------------------------|-----------------------|----------------|--------------|
| pa1 | i-08b49961db3e69fc2 | ⊘ Running ⊕⊖ t2.xlarge | ⊘ 2/2 checks passed | No alarms 🕂 us-west-2c | ec2-34-221-146-31.us | 34.221.146.31 | - |
| pa1 | i-03cb9d140d2fe869e | ⊘ Running ⊕⊖ t2.xlarge | ⊘ 2/2 checks passed | No alarms 🕂 us-west-2c | ec2-34-217-137-104.us | 34.217.137.104 | - |
| pa1 | i-0454ae152445fadc8 | ⊘ Running ⊕ ⊖ t2.xlarge | ⊘ 2/2 checks passed | No alarms + us-west-2c | ec2-35-92-24-86.us-we | 35.92.24.86 | - |
| pa1 | i-0c1b1002e1d2f8469 | ⊘ Running ⊕⊖ t2.xlarge | ⊘ 2/2 checks passed | No alarms + us-west-2c | ec2-54-203-131-59.us | 54.203.131.59 | - |
| pa1 | i-0517c357f8857adad | ⊘ Running ⊕⊖ t2.xlarge | ⊘ 2/2 checks passed | No alarms + us-west-2c | ec2-34-219-174-56.us | 34.219.174.56 | - |

- 1. Open a terminal and Do the following:
 - a. <u>SSH Into the Scheduler EC2 Instance :</u> ssh -i <pem_key_name>.pem <u>ubuntu@34.221.146.31</u>
 - b. <u>Activate the Dask Environment</u> : source dask_env/bin/activate
 - c. <u>Launch Jupyter Notebook on the EC2 :</u> jupyter notebook --port=8888



- 2. Open New Terminal And run the following commands -
 - Port Forwarding Jupyter Notebook running on port 8888 on the EC2 to port 8000 on local system (localhost:8000): ssh -i <pem_key_name>.pem ubuntu@34.221.4.2 -L 8000:localhost:8888

DSC 102 : PA 1 : System Setup Tutorial

ubuntu@ip-172-31-12-224: ~ (ssh) * Ubuntu Pro delivers the most comprehensive open source security and compliance features. https://ubuntu.com/aws/pro updates can be applied immediately. To see these additional updates run: apt list --upgradable Last login: Thu Jan 19 06:07:19 2023 from 24.43.123.81 ubuntu@ip-172-31-12-224:~\$ source dask_env/bin/activate (dask_env) ubuntu@ip-172-31-12-224:~\$ jupyter notebook --port=8888 19:04:18.101 NotebookApp] Writing notebook server cookie secret to /home/ubuntu/.local/share/jupyter/runtime/notebook_cookie_se I 19:04:20.632 NotebookApp] Serving notebooks from local directory: /home/ubuntu I 19:04:20.632 NotebookApp] Jupyter Notebook 6.4.10 is running at: 19:04:20.632 NotebookApp] http://localhost:8888/?token=fac098e07f8Z116c3c0b8c9539cb94eaf45d5e4a14ab773a 19:04:20.632 NotebookApp] or http://127.0.0.1:8888/?token=fac098e07f82116c3c0b8c9539cb94eaf45d5e4a14ab773a 19:04:20.632 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation). To access the notebook, open this file in a browser: file:///home/ubuntu/.local/share/jupyter/runtime/nbserver-12027-open.html Or copy and paste one of these URLs: http://localhost:8888/?token=fac098e07f82116c3c0b8c9539cb94eaf45d5e4a14ab773a or http://127.0.0.1:8888/?token=fac098e07f82116c3c0b8c9539cb94eaf45d5e4a14ab773a x ubuntu@ip-172-31-12-224: ~ (ssh) (base) golokeshpatra@Golokeshs-MacBook-Air ~ % cd ~/Downloads/Q2/DSC102/Assignments/PA1 (base) golokeshpatra@Golokeshs-MacBook-Air PA1 % ssh -i dsc102-pa1.pem ubuntu@34.221.146.31 -L 8000:localhost:8888 Welcome to Ubuntu 18.04.6 LTS (GNU/Linux 5.4.0-1092-aws x86_64) * Documentation: https://help.ubuntu.com * Management: https://landscape.canonical.com * Support: https://ubuntu.com/advantage System information as of Thu Feb 2 19:06:42 UTC 2023 System load: 0.01 Usage of /: 2.8% of 96.73GB Users logged in: Memory usage: 2% IP address for eth IP address for eth0: 172.31.12.224 Swap usage: 0% compliance features. https://ubuntu.com/aws/pro To see these additional updates run: apt list --upgradable New release '20.04.5 LTS' available. Run 'do-release-upgrade' to upgrade to it. *** System restart required *** Last login: Thu Feb 2 19:03:52 2023 from 24.43.123.92 ubuntu@ip-172-31-12-224:~\$ 🗌

Step 2 : Setup Dask Scheduler on the same EC2.

On the same node, we must start the dask-scheduler. To establish this, we need to follow these instructions:

- 1. Open a New Terminal :
 - a. SSH Into the EC2 Instance
 - b. Activate the Dask Environment : source dask env/bin/activate
 - c. Activate the dask-scheduler : dask scheduler --host 0.0.0.0

NOTE: The 'Scheduler at' shown above will be required while setting up workers

Step 3 : Dask UI/Dashboard Port forwarding .

- 1. Open a New Terminal:
 - a. Port Forward the dask dashboard UI from the EC2's 8787 port to local systems Port 8001 (localhost:8001):
 ssh -i dsc102-pa1.pem ubuntu@34.221.4.2 -L 8001:localhost:8787

Step 4 : Downlaod data from S3 :

- 1. In the same terminal opened or in a new terminal :
 - a. If New Terminal then activate the dask environment
 - b. Copy and paste the AWS_ACCESS_KEY and SECRET ACCESS KEYS
 - c. Download all the files from the S3 : aws s3 sync s3://dsc102-public /home/ubuntu/

Step 5 : Check Dask UI and Jupyter Notebook.

NOTE : Unlike PAO , when you go to the Info tab on Dask UI, you SHOULD NOT OBSERVE ANY WORKERS .

DASK UI :

← → c © localhost:8001/info/main/workers.html Scheduler tcp://172.31.4.65:8786

Logs Bokeh

Worker Name Cores Memory Memory use Occupancy Processing In-memory Services Logs Last seen

Jupyter Notebook :

| 💭 jupyter | Quit | Logout |
|--|---------------|-----------|
| Files Running Clusters | | |
| Select items to perform actions on them. | Upload | New - 2 |
| □ 0 💌 🖿 / Name 🐓 | Last Modified | File size |
| C dask_env | 10 days ago | |
| OutputSchema_PA1.json | 3 hours ago | 417 B |
| PA1.py | 3 hours ago | 1.22 kB |
| products_Release.csv | 3 hours ago | 9.79 GB |
| C results_PA1.json | 2 hours ago | 2.44 kB |
| user_reviews.csv | 3 hours ago | 28.5 GB |

SETUP DASK WORKER:

(For other 4 EC2 Instances)

Step 1 : SSH into the worker node

Step 2: Setup dask worker on the worker node.

- 1. To get the Dask Workers Up, we need to follow these instructions :
 - a. After SSH into the worker EC2, Activate the Dask Environment : source dask_env/bin/activate
 - b. <u>Now activate the dask worker with the following command:</u> dask worker tcp://172.31.152:8786 --nworkers 4

NOTE : The dask scheduler IP can be obtained from <u>Step 2 of Setting up</u> <u>Scheduler</u>

| (dask_env) ubuntu@ip-172-31-1-242:~\$ dask worker to | o://172.31.2.152:8786rw | workers 4 |
|--|--|---|
| 2023-04-26 20:58:54,824 - distributed.nanny - INFO - | Start Nanny at: | 'tcp://172.31.1.242:40655' |
| 2023-04-26 20:58:54,827 - distributed.nanny - INFO - | Start Nanny at: | 'tcp://172.31.1.242:36873' |
| 2023-04-26 20:58:54,830 - distributed.nanny - INFO | Start Nanny at: | 'tcp://172.31.1.242:45675' |
| 2023-04-26 20:58:54,838 - distributed.nanny - INFO | - Start Nanny at: | 'tcp://172.31.1.242:40845' |
| 2023-04-26 20:58:55,798 - distributed.worker - INFO | Start worker gt: | tcp://172.31.1.242:35981 |
| 2023-04-26 20:58:55.798 - distributed.worker - INFO | Listening to: | tcp://172.31.1.242:35981 |
| 2023-04-26 20:58:55.798 - distributed worker - INFO | dashboard at: | 172.31.1.242:40725 |
| 2023-04-26 20:58:55 798 - distributed worker - TNFO | - Waiting to connect to: | tcn://172_31_2_152:8786 |
| 2023-04-26 20:58:55 798 - distributed worker - TNFO | | |
| 2023-04-26 20:58:55,799 - distributed worker - TNFO | - Threads: | 1 |
| 2023-04-26 20:58:55,755 - distributed worker - INFO | - Memory: | 3 91 618 |
| 2023-04-26 20:58:55,759 - distributed worker - INFO | - Local Dinactory: | /two/dack_worker_rpace/worker_p69wc144 |
| 2023-04-26 20:50:55,755 - distributed worker - INFO | - Eocar Directory. | / uip/ dusk-worker -spuce/ worker -nobilist++ |
| 2023-04-20 20:56:55,799 - distributed worker - INFO | Ctart warken at | + //172 21 1 242-40407 |
| 2023-04-20 20:56:55,005 - distributed worker - INFO | - Start worker at: | tcp://1/2.31.1.242:4048/ |
| 2023-04-26 20:58:55,804 - distributed worker - INFO | Listening to: | tcp://1/2.31.1.242:4048/ |
| 2023-04-26 20:58:55,804 - distributed.worker - INFO | - aasnboara at: | 1/2.31.1.242:38999 |
| 2023-04-26 20:58:55,804 - distributed.worker - INFO | Maiting to connect to: | tcp://1/2.31.2.152:8786 |
| 2023-04-26 20:58:55,804 - distributed.worker - INFO | | |
| 2023-04-26 20:58:55,804 - distributed.worker - INFO | - Threads: | 1 |
| 2023-04-26 20:58:55,804 - distributed.worker - INFO | - Memory: | 3.91 GiB |
| 2023-04-26 20:58:55,804 - distributed.worker - INFO | Local Directory: | /tmp/dask-worker-space/worker-x4tdxup6 |
| 2023-04-26 20:58:55,804 - distributed.worker - INFO | | |
| 2023-04-26 20:58:55,808 - distributed.worker - INFO | Registered to: | tcp://172.31.2.152:8786 |
| 2023-04-26 20:58:55,808 - distributed.worker - INFO | | |
| 2023-04-26 20:58:55,810 - distributed.core - INFO - | Starting established conr | nection to tcp://172.31.2.152:8786 |
| 2023-04-26 20:58:55,810 - distributed.worker - INFO | Start worker at: | tcp://172.31.1.242:44943 |
| 2023-04-26 20:58:55,810 - distributed.worker - INFO | Listening to: | tcp://172.31.1.242:44943 |
| 2023-04-26 20:58:55,810 - distributed.worker - INFO | dashboard at: | 172.31.1.242:37215 |
| 2023-04-26 20:58:55,810 - distributed.worker - INFO | - Waiting to connect to: | tcp://172.31.2.152:8786 |
| 2023-04-26 20:58:55,811 - distributed.worker - INFO | | |
| 2023-04-26 20:58:55,811 - distributed.worker - INFO | - Threads: | |
| 2023-04-26 20:58:55,811 - distributed.worker - INFO | - Memory: | 3.91 GiB |
| 2023-04-26 20:58:55,811 - distributed.worker - INFO | Local Directory: | /tmp/dask-worker-space/worker-z758171z |
| 2023-04-26 20:58:55,811 - distributed.worker - INFO | | |
| 2023-04-26 20:58:55,814 - distributed.worker - INFO | Registered to: | tcp://172.31.2.152:8786 |
| 2023-04-26 20:58:55,814 - distributed.worker - INFO | | |
| 2023-04-26 20:58:55,815 - distributed.core - INFO - | Starting established conr | nection to tcp://172.31.2.152:8786 |
| 2023-04-26 20:58:55,817 - distributed.worker - INFO | Registered to: | tcp://172.31.2.152:8786 |
| 2023-04-26 20:58:55,817 - distributed.worker - INFO | | |
| 2023-04-26 20:58:55.818 - distributed.core - INFO - | Starting established conr | mection to tcp://172.31.2.152:8786 |
| 2023-04-26 20:58:55.826 - distributed.worker - INFO | Start worker at: | tcp://172.31.1.242:38621 |
| 2023-04-26 20:58:55.827 - distributed.worker - INFO | Listening to: | tcp://172.31.1.242:38621 |
| 2023-04-26 20:58:55.827 - distributed.worker - INFO | - dashboard at: | 172.31.1.242:43297 |
| 2023-04-26 20:58:55.827 - distributed worker - INFO | - Waiting to connect to: | tcn://172.31.2.152:8786 |
| 2023-04-26 20:58:55 827 - distributed worker - TNFO | | |
| 2023-04-26 20:58:55.827 - distributed worker - TNFO | - Threads | 1 |
| 2023-04-26 20:58:55,827 - distributed worker - TNFO | - Menory: | 3.91 GiB |
| 2023-04-26 20:58:55,827 - distributed worker - INFO | Local Directory: | /tmn/dask_worker_space/worker_9=725v2m |
| 2023-04-26 20:58:55 827 - distributed worker - INFO | cocar priectory: | and and the space worker system |
| 2022-04-26-20:50:55,027 - distributed worker - INFO | Pogistand to: | tco://172 21 2 152:0706 |
| 2023-04-20 20:35:55,655 - distributed worker - INFO | - Registered to: | CCp.//1/2.51.2.152.8/80 |
| 2023-04-20 20:35:55,633 - distributed worker - INFO | Stanting octoblished and | action to trait //172 21 2 152,0700 |
| 2025-04-20 20:58:55,854 - distributed.core - INFO - | stanting established con | eccton to tcp://1/2.31.2.152:8/86 |

Now if we check the Worker Info on the DASK UI then -

Scheduler tcp://172.31.8.102:8786

| Logs Bokeh | | | | | | | | | | |
|--------------------------|--------------------------|-------|---------|-------------|-----------|------------|-----------|-----------|------|-----------|
| Worker | Name | Cores | Memory | Memory use | Occupancy | Processing | In-memory | Services | Logs | Last seen |
| tcp://172.31.5.152:35911 | tcp://172.31.5.152:35911 | 1 | 4.20 GB | $(\ \)$ | 0.00 us | 0 | 0 | dashboard | logs | 510.11 ms |
| tcp://172.31.5.152:38641 | tcp://172.31.5.152:38641 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 510.62 ms |
| tcp://172.31.5.152:41199 | tcp://172.31.5.152:41199 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 508.22 ms |
| tcp://172.31.5.152:41637 | tcp://172.31.5.152:41637 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 514.77 ms |

Step 3: Download the data from s3 to the worker nodes too.

1. Open a new Terminal:

← → C ③ localhost:8001/info/main/workers.html

Repeat the <u>Step 4 in the scheduler setup section</u>

Step 4 : Check the DASK Dashboard/UI for Information of the workers.

We should observe 16 Worker nodes i.e 4 Workers * 4 nodes = 16 Nodes For Reference -

| Scheduler tcp: | //172.31.12.224 | :878 | 6 | | | | | | | |
|---------------------------|---------------------------|-------|---------|------------|-----------|------------|-----------|-----------|------|-----------|
| Logs Bokeh | | | | | | | | | | |
| Worker | Name | Cores | Memory | Memory use | Occupancy | Processing | In-memory | Services | Logs | Last seen |
| tcp://172.31.12.248:33483 | tcp://172.31.12.248:33483 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 766.98 ms |
| tcp://172.31.12.248:36339 | tcp://172.31.12.248:36339 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 771.40 ms |
| tcp://172.31.12.248:42139 | tcp://172.31.12.248:42139 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 770.26 ms |
| | tcp://172.31.12.248:46435 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 769.32 ms |
| tcp://172.31.12.96:33965 | tcp://172.31.12.96:33965 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 719.89 ms |
| tcp://172.31.12.96:34225 | tcp://172.31.12.96:34225 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 720.44 ms |
| tcp://172.31.12.96:39611 | tcp://172.31.12.96:39611 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 721.26 ms |
| cp://172.31.12.96:41441 | tcp://172.31.12.96:41441 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 722.73 ms |
| tcp://172.31.13.222:39279 | tcp://172.31.13.222:39279 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 316.18 ms |
| tcp://172.31.13.222:40447 | tcp://172.31.13.222:40447 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 316.58 ms |
| tcp://172.31.13.222:40543 | tcp://172.31.13.222:40543 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 318.89 ms |
| tcp://172.31.13.222:46063 | tcp://172.31.13.222:46063 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 318.49 ms |
| tcp://172.31.2.228:33005 | tcp://172.31.2.228:33005 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 714.36 ms |
| tcp://172.31.2.228:33185 | tcp://172.31.2.228:33185 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | logs | 717.33 ms |
| tcp://172.31.2.228:35339 | tcp://172.31.2.228:35339 | 1 | 4.20 GB | | 0.00 us | 0 | 0 | dashboard | logs | 717.03 ms |
| tcp://172.31.2.228:39095 | tcp://172.31.2.228:39095 | 1 | 4.20 GB | (| 0.00 us | 0 | 0 | dashboard | loas | 721.38 ms |

Step 5 : Repeat All the above steps on each and every Worker EC2 Instance.