

# Distributed Computing

---

DIANE WOODBRIDGE, PH.D



# AWS Q&A

## How to copy my file to S3?

### 1. Check the permission.

Upload

Create Folder

Actions ▾

Search by prefix

NonePropertiesTransfers

All Buckets / usfca-msan694

	Name	Storage Class	Size	Last Modified
<input type="checkbox"/>	Italian_Stack_Exchange	--	--	--
<input type="checkbox"/>	README.md	Standard	3.7 KB	Thu Dec 22 06:21:56 GMT-800 2016
<input type="checkbox"/>	json_out	--	--	--
<input type="checkbox"/>	world_bank_project.json	Standard	2.8 MB	Fri Jan 06 03:01:31 GMT-800 2017

Bucket: usfca-msan694

Bucket: usfca-msan694

Region: Oregon

Creation Date: Thu Dec 22 06:20:58 GMT-800 2016

Owner: dwoodbridge

▼ Permissions

You can control access to the bucket and its contents using access policies. [Learn more.](#)

Grantee: dwoodbridge

☒ List ☒ Upload/Delete ☒ View Permissions ☒ Edit Permissions

x

Grantee: Everyone

☐ List ☒ Upload/Delete ☒ View Permissions ☐ Edit Permissions

x

+ Add more permissions

+ Add bucket policy

+ Add CORS Configuration

Save

Cancel

# AWS Q&A

---

How to copy my file to S3?

2. `aws s3 cp your_file destination`

```
[ML-ITS-603436:Answer dwoodbridge$ aws s3 cp supervisor_sf.csv s3://usfca-msan694
upload: ./supervisor_sf.csv to s3://usfca-msan694/supervisor_sf.csv
[ML-ITS-603436:Answer dwoodbridge$ █
```



# AWS Q&A

---

How to copy my file to S3?

2. `aws s3 cp your_file destination`

---




Upload

Create Folder

Actions ▾

Before

All Buckets / usfca-msan694

	Name	Storage Class	Size	Last Modified
<input type="checkbox"/>	 Italian_Stack_Exchange	--	--	--
<input type="checkbox"/>	 README.md	Standard	3.7 KB	Thu Dec 22 06:21:56 GMT-800 2016
<input type="checkbox"/>	 json_out	--	--	--
<input type="checkbox"/>	 world_bank_project.json	Standard	2.8 MB	Fri Jan 06 03:01:31 GMT-800 2017

# AWS Q&A

---

How to copy my file to S3?

2. `aws s3 cp your_file destination`

---






Upload

Create Folder

Actions ▾

After

All Buckets / usfca-msan694

	Name
<input type="checkbox"/>	 Italian_Stack_Exchange
<input type="checkbox"/>	 README.md
<input type="checkbox"/>	 json_out
<input type="checkbox"/>	 supervisor_sf.csv
<input type="checkbox"/>	 world_bank_project.json



# AWS Q&A

---

## How to check Spark status on EMR?

1. Add an inbound rule on your master to allow port 8088

Name of interface	URI
YARN ResourceManager	<code>http://<i>master-public-dns-name</i>:8088/</code>
YARN NodeManager	<code>http://<i>slave-public-dns-name</i>:8042/</code>
Hadoop HDFS NameNode	<code>http://<i>master-public-dns-name</i>:50070/</code>
Hadoop HDFS DataNode	<code>http://<i>slave-public-dns-name</i>:50075/</code>
Spark HistoryServer	<code>http://<i>master-public-dns-name</i>:18080/</code>
Zeppelin	<code>http://<i>master-public-dns-name</i>:8890/</code>
Hue	<code>http://<i>master-public-dns-name</i>:8888/</code>
Ganglia	<code>http://<i>master-public-dns-name</i>/ganglia/</code>
HBase UI	<code>http://<i>master-public-dns-name</i>:16010/</code>

# AWS Q&A

## How to check Spark status on EMR?

1. Add an inbound rule on your master to allow port 8088

Edit inbound rules

Custom TCP Rule	⌵	TCP	8443	Custom	⌵	54.240.230.176/29	✕
Custom TCP Rule	⌵	TCP	8443	Custom	⌵	205.251.233.160/28	✕
Custom TCP Rule	⌵	TCP	8443	Custom	⌵	54.240.230.240/29	✕
Custom TCP Rule	⌵	TCP	8443	Custom	⌵	54.240.230.184/29	✕
Custom TCP Rule	⌵	TCP	8443	Custom	⌵	205.251.233.32/28	✕
Custom TCP Rule	⌵	TCP	8443	Custom	⌵	205.251.233.48/29	✕
Custom TCP Rule	⌵	TCP	8443	Custom	⌵	205.251.233.176/29	✕
Custom TCP Rule	⌵	TCP	8443	Custom	⌵	205.251.234.32/28	✕
Custom TCP Rule	⌵	TCP	8082	Custom	⌵	0.0.0.0/0	✕
All UDP	⌵	UDP	0 - 65535	Custom	⌵	sg-a605c1de	✕
All UDP	⌵	UDP	0 - 65535	Custom	⌵	sg-a705c1df	✕
All ICMP	⌵	ICMP	0 - 65535	Custom	⌵	sg-a605c1de	✕

# AWS Q&A

## How to check Spark status on EMR?

2. On your browser call `http://master-public-dns-name:8088/`

ec2-54-202-189-5.us-west-2.compute.amazonaws.com:8088/cluster/nodes

hadoop

Nodes of the cluster

Log

Cluster

About

Nodes

Node Labels

Applications

NEW

NEW SAVING

SUBMITTED

ACCEPTED

RUNNING

FINISHED

FAILED

KILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealth Nodes
0	0	0	0	0	0 B	4 GB	0 B	0	8	0	2	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[MEMORY]	<memory:32, vCores:1>	<memory:2048, vCores:4>

Show 20 entries

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Mem Used	Mem Avail	VCores Used	VCores Avail
	/default-rack	RUNNING	ip-172-31-43-112.us-west-2.compute.internal:8041	ip-172-31-43-112.us-west-2.compute.internal:8042	Wed Jan 18 18:34:36 +0000 2017		0	0 B	2 GB	0	4
	/default-rack	RUNNING	ip-172-31-34-63.us-west-2.compute.internal:8041	ip-172-31-34-63.us-west-2.compute.internal:8042	Wed Jan 18 18:34:47 +0000 2017		0	0 B	2 GB	0	4

Showing 1 to 2 of 2 entries

First Previous 1

UNIVERSITY OF SAN FRANCISCO  
CHANGE THE WORLD FROM HERE

8



# AWS Q&A

---

## Pricing?

<http://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-web-interfaces.html>



# Reference

---

<http://docs.aws.amazon.com/cli/latest/reference>

