

Screenshot of Azure Account showing username

The screenshot shows the Microsoft Azure portal dashboard. The user profile dropdown menu is open, displaying the user's name, email, and university affiliation. The user's name is Zhang, Siyu, and the email is SIZ24@pitt.edu. The university is the University of Pittsburgh. The user has a profile picture and a link to view their account. The dashboard also shows a list of recent resources, including a storage account and an HDInsight cluster.

Azure services

- Create a resource
- Storage accounts
- HDInsight clusters
- Virtual machines
- App Services
- SQL databases
- Azure Database for PostgreSQL
- Azure Cosmos DB
- Kubernetes service

Recent resources

Name	Type	Last Viewed
cs1660clusterhdstorage	Storage account	3 minutes ago
cs1660-cluster	HDInsight cluster	4 minutes ago

Navigate

- Subscriptions
- Resource groups
- All resources
- Dashboard

Tools

- Microsoft Learn: Learn Azure with free online training from Microsoft
- Azure Monitor: Monitor your apps and infrastructure
- Security Center: Secure your apps and infrastructure
- Cost Management: Analyze and optimize your cloud spend for free

Useful links

- Technical Documentation
- Azure Services
- Recent Azure Updates

Azure mobile app

Download on the App Store | GET IT ON Google Play

Screenshot of your implementation of Word Count on Azure

The screenshot shows the Azure portal interface for a storage account named 'cs1660-cluster-hw3'. The 'Overview' tab is selected, displaying a list of blobs. The 'Authentication method' is 'Access key (Switch to Azure AD User Account)' and the 'Location' is 'cs1660-cluster-hw3 / example / data'. The list of blobs includes folders like 'gutenberg', 'people.parquet', 'people.seq', and 'wordcountout', as well as files like 'fruits.txt', 'gutenberg', 'people.json', 'people.parquet', 'people.seq', 'sample.log', 'wordcountout', and 'yellowthings.txt'. The 'wordcountout' folder is highlighted with a red box.

cs1660-cluster-hw3

Authentication method: Access key (Switch to Azure AD User Account)
Location: cs1660-cluster-hw3 / example / data

Search blobs by prefix (case-sensitive)

Name	Modified	Access tier	Blob type	Size	Lease state
[...]					...
gutenberg					...
people.parquet					...
people.seq					...
wordcountout					...
fruits.txt	4/3/2020, 1:13:33 PM		Block blob	66 B	Available
gutenberg	4/3/2020, 1:13:31 PM		Block blob	0 B	Available
people.json	4/3/2020, 1:13:36 PM		Block blob	77 B	Available
people.parquet	4/3/2020, 1:13:39 PM		Block blob	0 B	Available
people.seq	4/3/2020, 1:13:42 PM		Block blob	0 B	Available
sample.log	4/3/2020, 1:13:44 PM		Block blob	95.59 KiB	Available
wordcountout	4/4/2020, 1:44:59 AM		Block blob	0 B	Available
yellowthings.txt	4/3/2020, 1:13:46 PM		Block blob	62 B	Available

My input data file (Uploaded to OneDrive)

https://1drv.ms/t/s!Aov0m369Nhu-n3RBf97tgsg_AvWM?e=hCWbmo

Screenshot of the Word Count Output

```
sshuser@hn0-cs1660:~$ yarn jar wordcountjava-1.0-SNAPSHOT.jar org.apache.hadoop.examples.WordCount /example/data/gutenberg/hw3-input.txt /example/data/hw3-wordcountout
20/04/03 19:31:29 INFO client.AHSProxy: Connecting to Application History server at headnodehost/10.0.0.17:10200
20/04/03 19:31:29 INFO client.RequestHedgingRMFailoverProxyProvider: Looking for the active RM in [rm1, rm2]...
20/04/03 19:31:29 INFO client.RequestHedgingRMFailoverProxyProvider: Found active RM [rm2]
20/04/03 19:31:30 INFO input.FileInputFormat: Total input paths to process : 1
20/04/03 19:31:30 INFO lzo.GPLNativeCodeLoader: Loaded native gpl library
20/04/03 19:31:30 INFO lzo.LzoCodec: Successfully loaded & initialized native-lzo library [hadoop-lzo rev b5efb3e531bc1558201462b8ab15bb412ffa6b89]
20/04/03 19:31:30 INFO mapreduce.JobSubmitter: number of splits:1
20/04/03 19:31:30 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1585890432593_0003
20/04/03 19:31:30 INFO impl.YarnClientImpl: Submitted application application_1585890432593_0003
20/04/03 19:31:30 INFO mapreduce.Job: The url to track the job: http://hn1-cs1660.qh34qfohggtulc4xi2diz0ams.yx.internal.cloudapp.net:8088/proxy/application_1585890432593_0003/
20/04/03 19:31:30 INFO mapreduce.Job: Running job: job_1585890432593_0003
20/04/03 19:31:47 INFO mapreduce.Job: Job job_1585890432593_0003 running in uber mode : false
20/04/03 19:31:47 INFO mapreduce.Job: map 0% reduce 0%
20/04/03 19:32:09 INFO mapreduce.Job: map 67% reduce 0%
20/04/03 19:32:20 INFO mapreduce.Job: map 100% reduce 0%
20/04/03 19:32:25 INFO mapreduce.Job: map 100% reduce 100%
20/04/03 19:32:26 INFO mapreduce.Job: Job job_1585890432593_0003 completed successfully
20/04/03 19:32:26 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=466761
    FILE: Number of bytes written=1261915
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    WASB: Number of bytes read=55826853
    WASB: Number of bytes written=376540
    WASB: Number of read operations=0
    WASB: Number of large read operations=0
    WASB: Number of write operations=0
  Job Counters
    Launched map tasks=1
    Launched reduce tasks=1
    Rack-local map tasks=1
    Total time spent by all maps in occupied slots (ms)=30048
    Total time spent by all reduces in occupied slots (ms)=2954
    Total time spent by all map tasks (ms)=30048
    Total time spent by all reduce tasks (ms)=2954
    Total vcore-milliseconds taken by all map tasks=30048
    Total vcore-milliseconds taken by all reduce tasks=2954
    Total megabyte-milliseconds taken by all map tasks=46153728
    Total megabyte-milliseconds taken by all reduce tasks=4537344
  Map-Reduce Framework
    Map input records=1284720
    Map output records=10054280
    Map output bytes=95511920
    Map output materialized bytes=466761
    Input split bytes=173
    Combine input records=10054280
    Combine output records=32956
    Reduce input groups=32956
    Reduce shuffle bytes=466761
    Reduce input records=32956
```

Microsoft Azure | Search resources, services, and docs (G+)

Home > cs1660clusterhdistorage | Containers > cs1660-cluster-hw3 > example/data/hw3-wordcountout/part-r-00000

cs1660-cluster-hw3 Container

Authentication method: Access key (Switch to Azure AD User Account)

Location: cs1660-cluster-hw3 / example / data / hw3-wordcountout

Search blobs by prefix (case-se...)

Show deleted blobs

Name

- [] [-]
- [] _SUCCESS
- [] part-r-00000

Overview Snapshots Edit Generate SAS

The file 'example/data/hw3-wordcountout/part-r-00000' may not render correctly as it contains an unrecognized extension.

```
1 "(Lo)cra" 40
2 "1490 40
3 "1498," 40
4 "35" 40
5 "40," 40
6 "AS-IS" 40
7 "A_ 40
8 "Absoluti 40
9 "Alack! 40
10 "Alack!" 40
11 "Alla 40
12 "Allegorical 40
13 "Alpine-glow" 40
14 "And 80
15 "Antoni 40
16 "At 40
17 "B_ 40
18 "Bathers 40
19 "Bononiae 40
20 "By 40
21 "Come 40
```