

Azure Batch lab

Exercise 1. Running parallel image processing

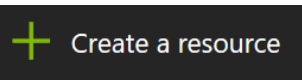
Duration: 45 minutes

In this exercise we will execute a simple ImageMagick resizing job on a set of input jpg files processed in parallel using Azure Batch service and Batch Explorer interface.

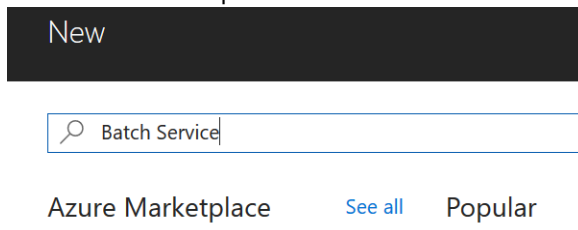
We will use Batch Explorer which is an open-source GUI application for managing Azure Batch accounts, designed to make the use of Azure Batch service easy and straight-forward.

Preparation step 1: Provision Azure Batch account








1. Navigate to Azure Portal in the browser.
2. Select **+ Create a resource**.



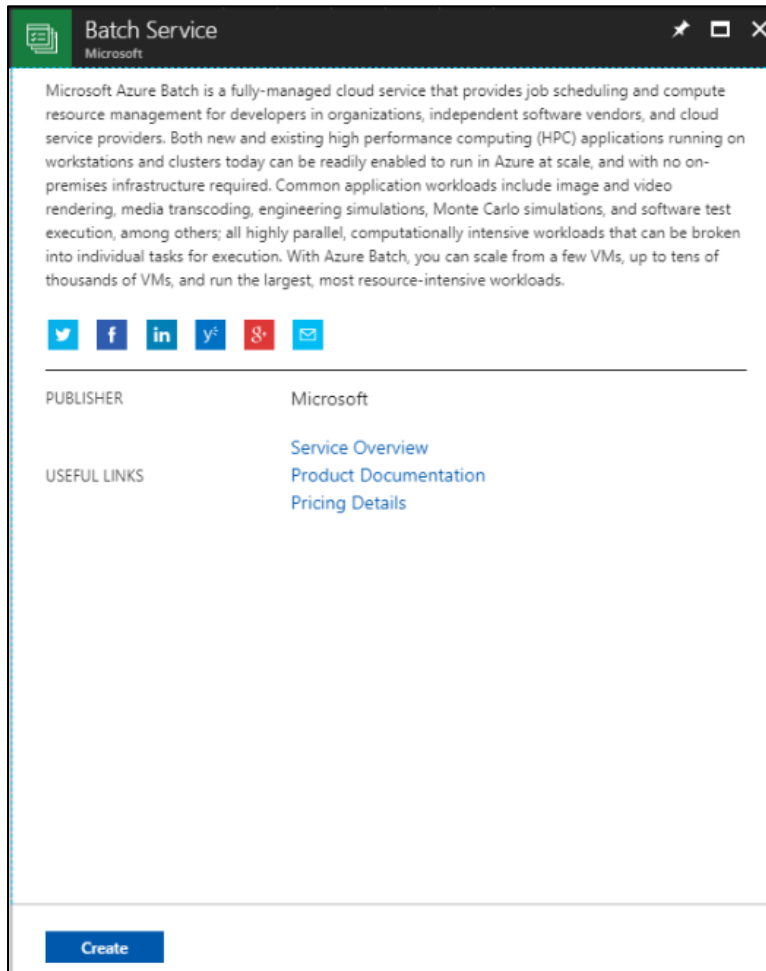
3. Search the Marketplace for **Batch Service**.



4. In the list, select **Batch Service**.

batch service			
Pricing	Operating System		Publisher
All	All	All	
Results			
NAME	PUBLISHER	CATEGORY	
 Batch Service	Microsoft		
 Azure Batch Rendering (CentOS 7.3)	Microsoft Azure Batch	Compute	
 Azure Batch Rendering For Windows 2016	Microsoft Azure Batch	Compute	
 Ubuntu (with RDMA) for Azure Batch container pools	Microsoft	Compute	
 Batch AI Service (Retiring)	Microsoft	Analytics	
 Azure ML Batch Execution Service Web App Template	Microsoft	Web	
 Ubuntu 16.04-LTS for Azure Batch container pools	Microsoft	Compute	

5. Select **Create** on the Batch Service blade.



6. On the New Batch Account blade, specify the following:
- Account name:** Provide a name for your new Batch Account. The name you choose must be unique within the Azure region where the account is created (see Location below).
 - Subscription:** Select the subscription in which to create the Batch account.
 - Resource group:** Select a new resource group and name it (e.g. partnersummit).

- d. **Location:** The Azure region in which to create the Batch account.

New Batch account
Provide basic Batch account info

[Basics](#) [Advanced](#) [Tags](#) [Review + create](#)

Microsoft Azure Batch is a fully-managed cloud service that provides job scheduling and compute resource management for developers in organizations, independent software vendors, and cloud service providers. Both new and existing high performance computing (HPC) applications running on workstations and clusters today can be readily enabled to run in Azure at scale, and with no on-premises infrastructure required. Common application workloads include image and video rendering, media transcoding, engineering simulations, Monte Carlo simulations, and software test execution, among others; all highly parallel, computationally intensive workloads that can be broken into individual tasks for execution. With Azure Batch, you can scale from a few VMs, up to tens of thousands of VMs, and run the largest, most resource-intensive workloads. [Learn more](#)

PROJECT DETAILS
Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

* Subscription: Microsoft Azure - internal usage
* Resource group: (New) partnersummit
[Create new](#)

INSTANCE DETAILS

* Account name: batchdemo00
westeurope.batch.azure.com

* Location: West Europe

7. Select **Storage account**.

STORAGE ACCOUNT

Specify an optional storage account. For best performance we recommend a storage account (general purpose v2) located in the same region as the associated Batch account. [Select a storage account](#)

8. In the Choose storage account, select **Create new**.

Choose storage account

These are the storage accounts in the selected subscription and location 'West US'.

+ Create new

None

9. On the Create storage account blade:

- Name:** Provide a unique name for the new Azure Storage account that will be attached to your Batch Service. E.g. use batchdemo<n> where n is the number on your desk.
- Account kind:** Select StorageV2
- Performance:** Leave this at Standard.
- Replication:** Leave this at Locally-redundant storage (LRS).

Create storage account

*

Name

batchdemo00

✓

.core.windows.net

Account kind

StorageV2 (general purpose v2)

Performance

Standard

Premium

Replication

Locally-redundant storage (LRS)

e. Select **OK** at the bottom of the blade.

10. Click "Next: Advanced" button at the bottom of the page:

Next: Advanced >

11. Make sure that Pool allocation mode is set to **Batch service**.

Basics

Advanced

Tags

Review + create

POOL ALLOCATION MODE

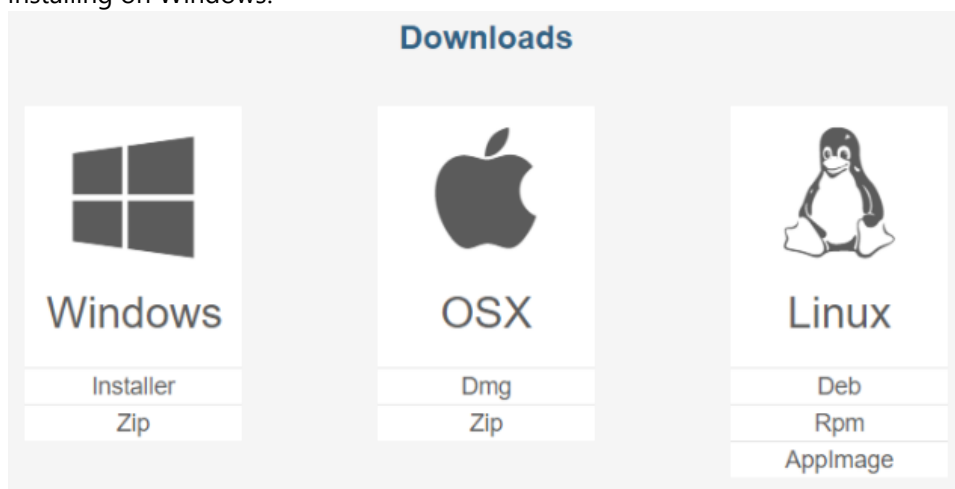
Pool allocation mode

☒ Batch service
 ☐ User subscription

12. Click **Review+Create** and then **Create** button to create the new Batch account. The provisioning should take around 1 minute.












Preparation step 2: Download Batch Explorer

- Using a web browser on your local machine, navigate to <https://azure.github.io/BatchExplorer/>.
- Scroll down until you see the **Downloads** section.
- Select the Zip download appropriate to your OS. The instructions that follow assume you are installing on Windows.

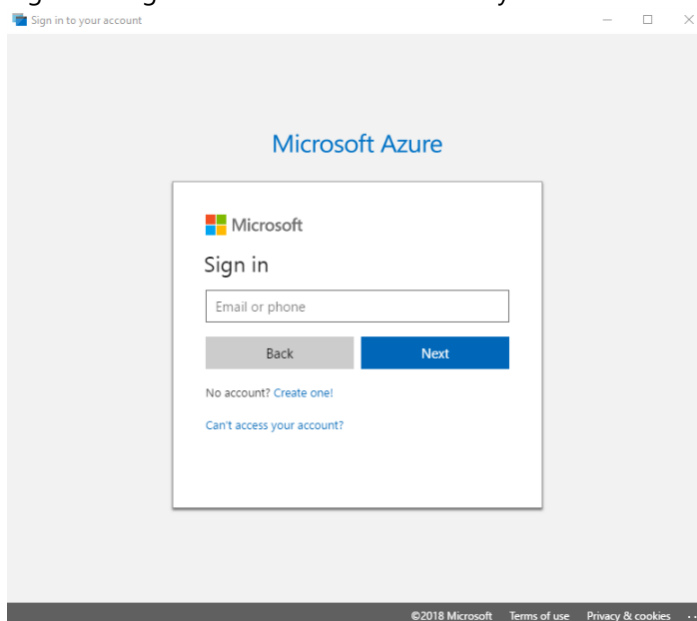


Further instructions assume that **Batch Explorer 0.19.1 (beta)** is installed.

- Unzip the installer file to a local folder and run the Batch Explorer executable:

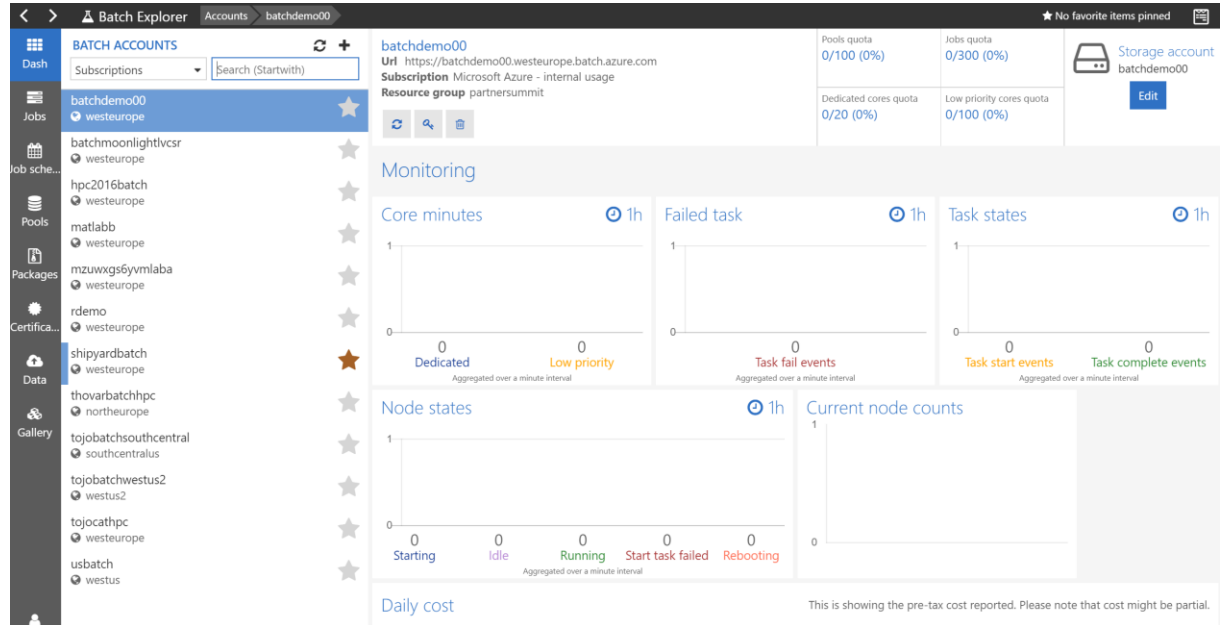
 api-ms-win-crt-multibyte-l1-1-0.dll	16.01.2019 01:43	Application extens...	19 KB
 api-ms-win-crt-private-l1-1-0.dll	16.01.2019 01:43	Application extens...	62 KB
 api-ms-win-crt-process-l1-1-0.dll	16.01.2019 01:43	Application extens...	12 KB
 api-ms-win-crt-runtime-l1-1-0.dll	16.01.2019 01:43	Application extens...	15 KB
 api-ms-win-crt-stdio-l1-1-0.dll	16.01.2019 01:43	Application extens...	17 KB
 api-ms-win-crt-string-l1-1-0.dll	16.01.2019 01:43	Application extens...	17 KB
 api-ms-win-crt-time-l1-1-0.dll	16.01.2019 01:43	Application extens...	13 KB
 api-ms-win-crt-utility-l1-1-0.dll	16.01.2019 01:43	Application extens...	11 KB
 BatchExplorer.exe	16.01.2019 01:43	Application	69 695 KB
 blink_image_resources_200_percent.pak	16.01.2019 01:43	PAK File	5 KB
 content_resources_200_percent.pak	16.01.2019 01:43	PAK File	1 KB

5. Sign in using the account associated with your Azure Subscription.



6. Wait for Batch Explorer to complete loading. When the main Batch Explorer window opens, in the left panel you should be able to select your Azure subscription and the just created Batch

account.

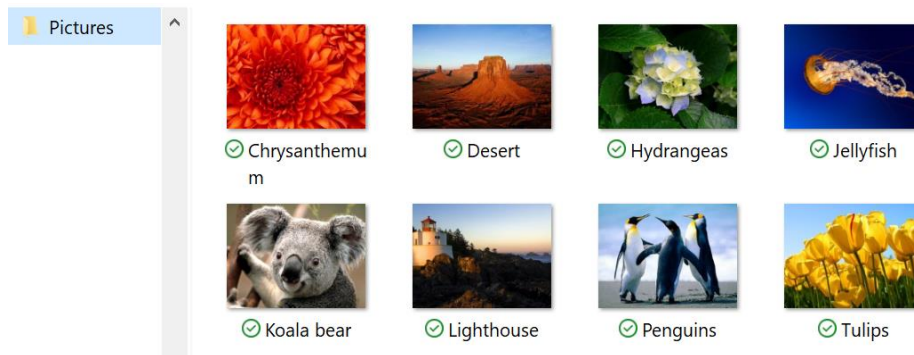


7. Your instance Batch Explorer is ready for use.

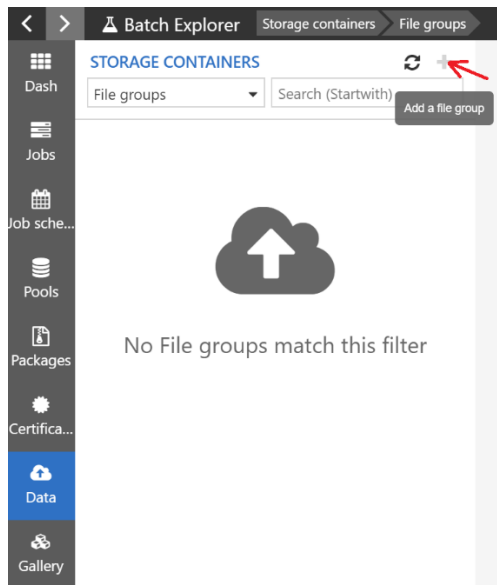
Now we will use the just installed Batch Explorer to process in parallel a set of input jpeg files with ImageMagick application.

Step 1: Upload input files to Batch Labs input file group

1. Download sample set of input jpeg files from <https://github.com/tojozefi/scfelab/blob/master/pictures.zip>
2. Uncompress the content to any local folder, preserving the internal directory structure – folder 'Pictures' with 8 jpeg files.



3. Navigate to **Data** panel in Batch Explorer and click **Add file group (+)** to create a new storage container:



4. Select **From local folder (file group)** from the drop-down list:

Empty container

Empty file group

From local folder(File group)

5. Fill out the **Create file group** form that opens:

Create file group

Upload files into a managed storage container that is used to supply resource files for your jobs and tasks. File groups require a 'fgrp-' prefix which is added automatically on creation.

1 General info

Basic information about the file group

File group name

fgrp- imagemagick-input

☐ Create an empty file group

2 Files

Drag and drop directories or files that you want to include in the file group

File or directory path

C:\Users\tojozefi\OneDrive - Microsoft\Events\FY19\PartnerSummit\pictures

Drag and drop files or folders here or select items with the buttons below

Select a folder

Select a file

☒ Include sub directories

3 File options

How data should be laid out in storage

Prefix

☐ Flatten

Create

Create and close

Close

In (1) **General info** provide the name for the input file group, e.g. **imagemagick-input**

In (2) **Files** select the *pictures* directory with sample jpeg files from local disk and check **Include sub directories** box.

Leave all options in (3) **File options** intact.

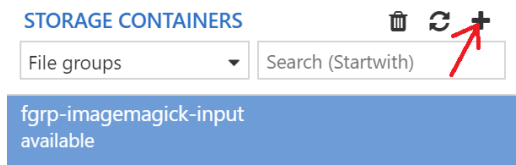
6. Click **Create and close** button and wait for the file upload to complete.

On successful upload you should see all jpeg files in the **fgrp-imagemagick-input** file group in Batch Explorer Data panel:

Name	Size	Last Modified
Chrysanthemum.jpg	879 kB	Jan 16th, 2019, 07:35:56
Desert.jpg	846 kB	Jan 16th, 2019, 07:35:56
Hydrangeas.jpg	595 kB	Jan 16th, 2019, 07:35:56
Jellyfish.jpg	776 kB	Jan 16th, 2019, 07:35:57
Koala bear.jpg	781 kB	Jan 16th, 2019, 07:35:57
Lighthouse.jpg	561 kB	Jan 16th, 2019, 07:35:57
Penguins.jpg	778 kB	Jan 16th, 2019, 07:35:57
Tulips.jpg	621 kB	Jan 16th, 2019, 07:35:57

Step 2: Create output file group

- Click again **Add file group (+)** to create an output storage container:



- Select **Empty file group** from the drop-down list:

Empty container

Empty file group

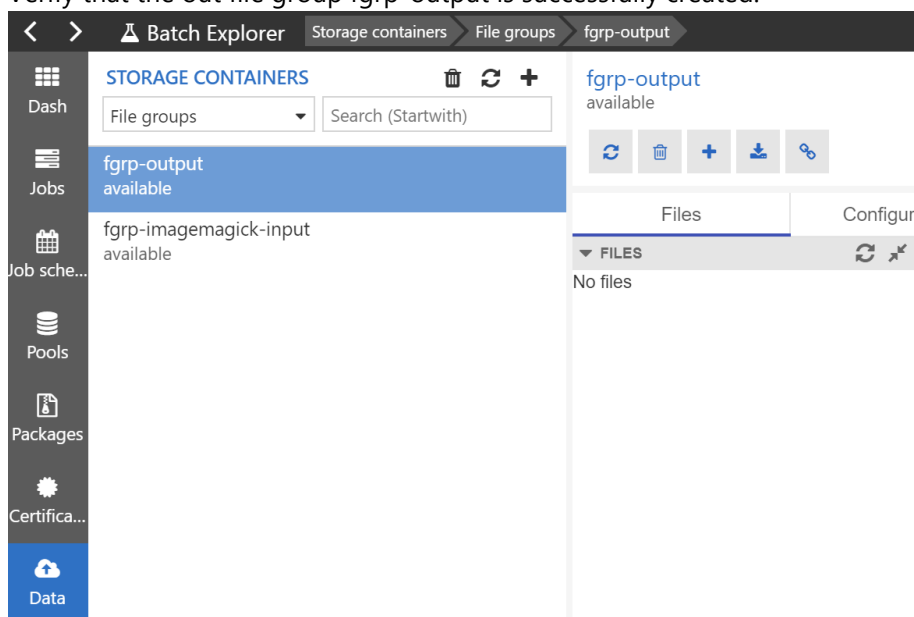
From local folder(File group)

- Provide a name for the output file group, e.g. **image-magick-output** and confirm.

Create a new empty file group

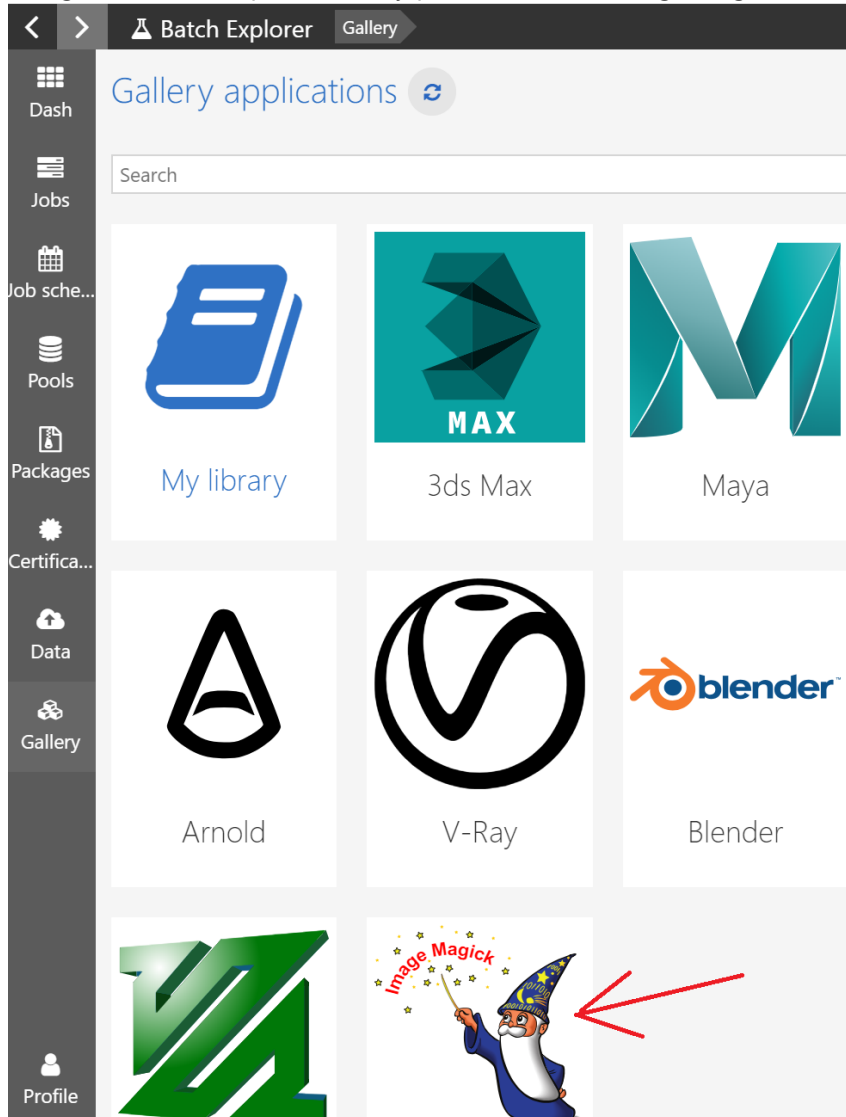
Confirm Cancel

- Verify that the out file group fgrp-output is successfully created:

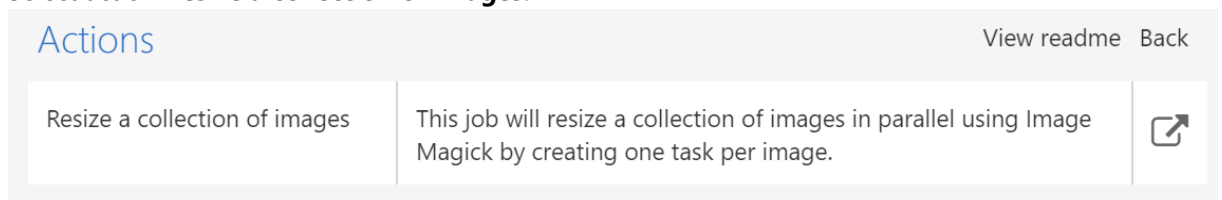


Step 3: Start Image Magick resize job for provided input files

11. Navigate to Batch Explorer Gallery panel and select Image Magick tab:



12. Select action **Resize a collection of images**:



Note: you can click on **view readme** to visit a website with ImageMagick Batch template repository.

Actions

View readme Back

Resize a collection of images

This job will resize a collection of images in parallel using Image Magick by creating one task per image.

View template on github

13. Select **Run job with auto pool** in (1) Mode selection:

Run resize-images from imagemagick

1 Mode

Run job with auto pool

Run job with existing pool

Create pool for later use

The job will be executed in automatically managed Azure Batch pool – i.e. the pool will be automatically deployed before the job start and deleted after the job is finished.

14. Fill out the remaining configuration parameters:

Run resize-images from imagemagick

1 Mode

Run job with auto pool

Run job with existing pool

Create pool for later use

2 Pool

Pool name

image-magick-pool

The ID of Azure Batch pool

Vm count

8

The number of virtual machines in the Azure Batch pool where the job will run

Node size

Standard_F1

The size of the virtual machines that run the application

3 Job

Job name

image-magick-resize

The job name

Input filegroup

fgrp-imagemagick-input

The file group where the input images are stored. Any non-image files will fail to convert

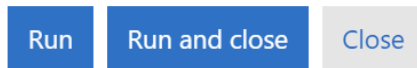
Run

Run and close

Close

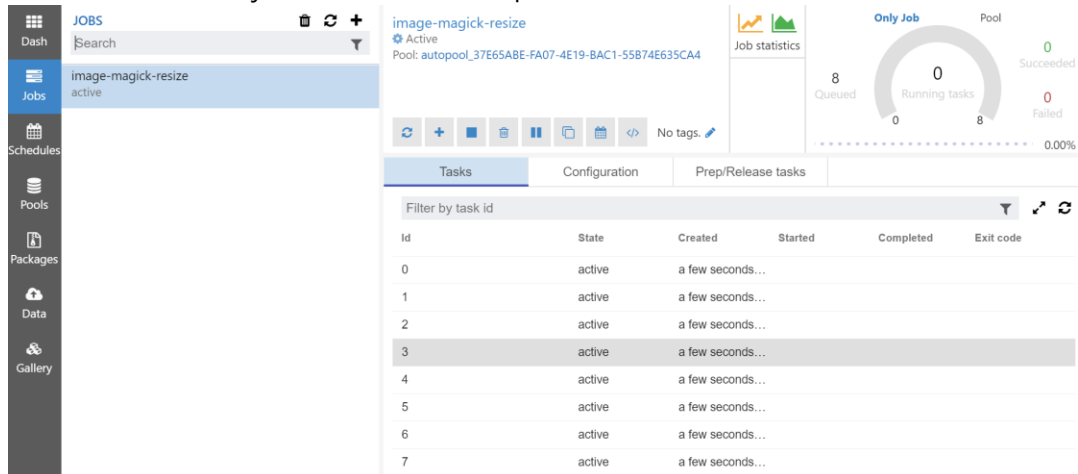
In (2) **Pool** definition provide a Batch pool name, e.g. **image-magick-pool**, **VM count** equal to the number of input files (in our case **8**) and select **VM size STANDARD_F1** from the list. In (3) **Job** definition provide the job name, e.g. **image-magick-resize**, and select the existing input and output file groups from the drop-down list – **fgrp-imagemagick-input** and **fgrp-image-magick-output** respectively. You may leave the **Resize** factor of **50%** or adjust.

15. When configuration is ready, click **Run and close** button to start the job:

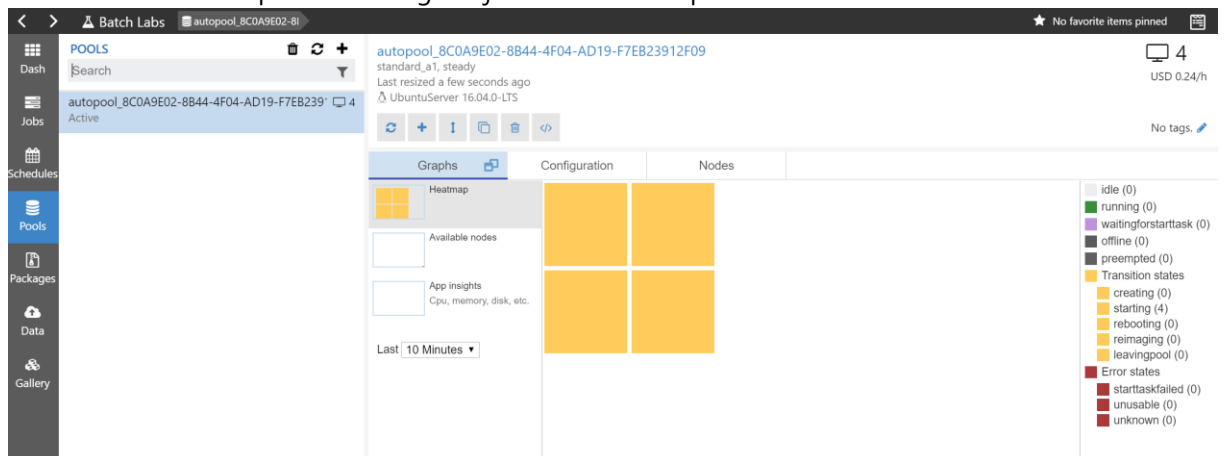


16. Monitor the job until it's finished.

You can observe the job status in the **Jobs** panel:



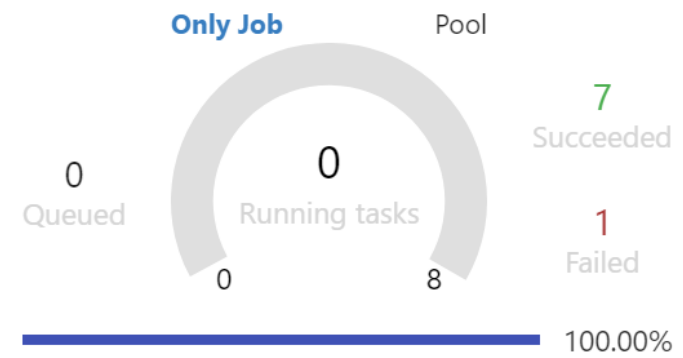
and the status of the pool executing the job on the **Pools** panel:



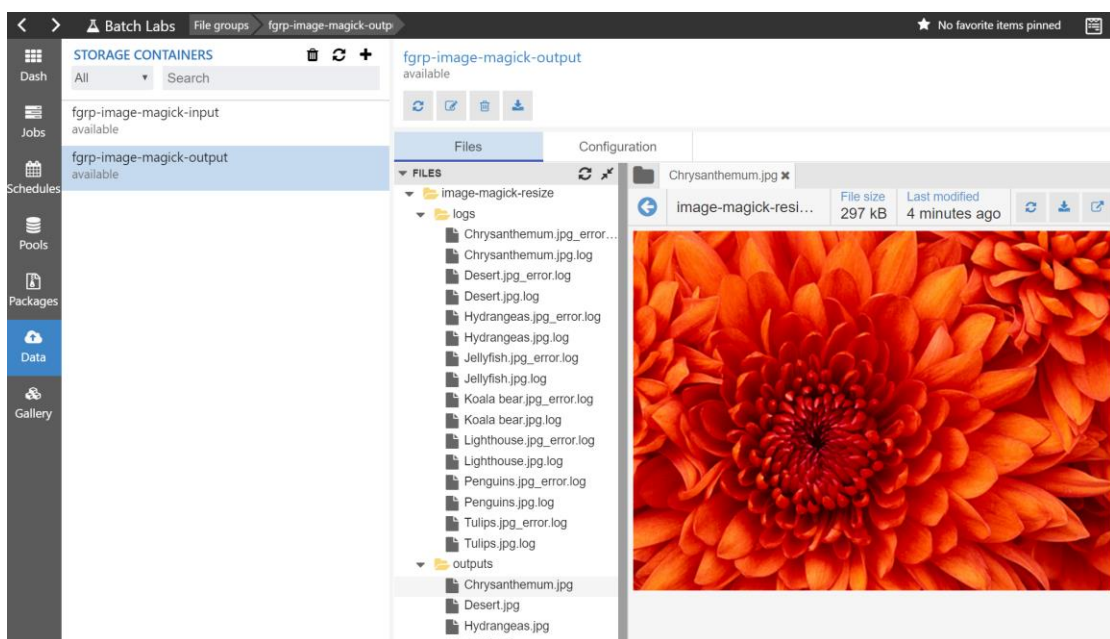
The job should take about 5 minutes to complete.

Step 4: Check result of Image Magick resize job

17. When the job is finished you should see the following status on the Jobs panel (top-right corner):

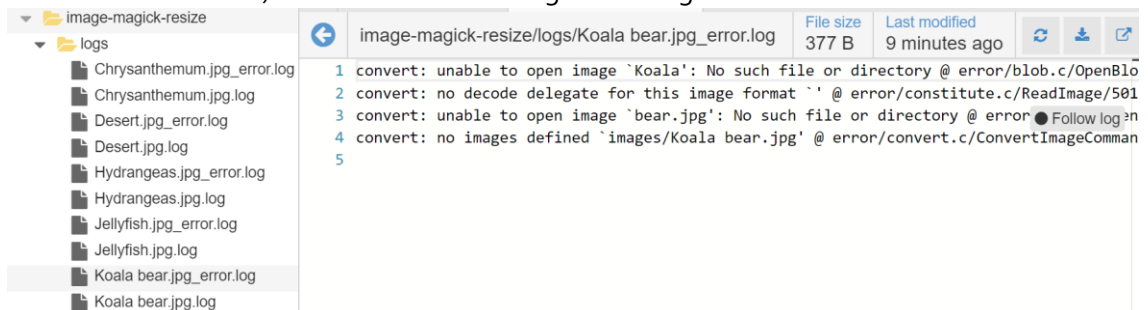


18. Navigate to Data panel in Batch Explorer main window and open fgrp-output file group. You should find image-magick-resize job's folder with log and output subfolders:

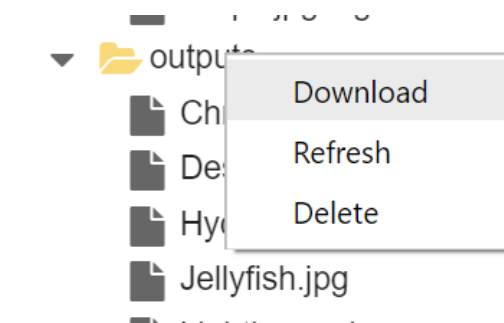


Check that the output jpeg files have reduced sizes as expected.

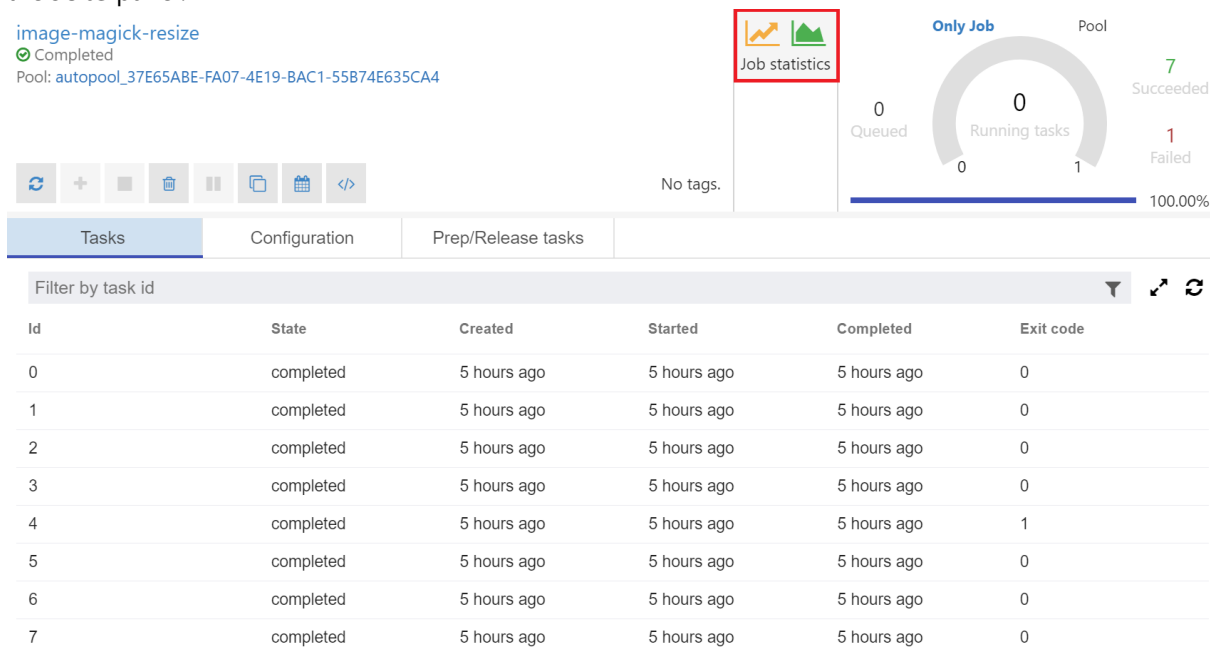
19. You may have observed that one of the tasks failed (for file Koala bear.jpg – due to space character in the name). Check the error message in the log file:



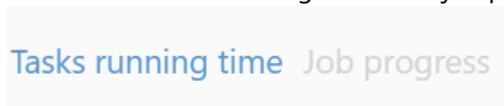
20. You can download the output files to local disk by right-clicking the **output** folder and selecting download action:



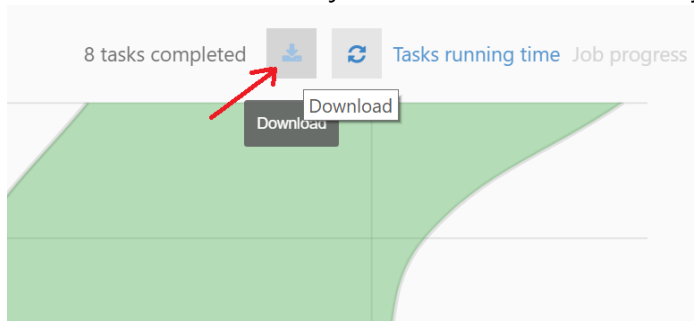
21. You can check job execution statistics by clicking **Job statistics** for *image-magick-resize* job on the **Jobs** panel:



You can view task running times and job progress diagrams:



You can also download the job execution data to csv file by clicking download icon:



Step 5: Clean up resources

22. As the last step of the exercise clean up the created resources by deleting the resource group. Click **"Delete resource group"** button on your resource group tab:

The screenshot shows the Azure portal interface. On the left, the 'partnersummit' resource group is selected. The main area displays the 'Overview' tab with a list of resources: 'batchdemo00' (Batch account) and 'batchdemo00' (Storage account). The 'Delete resource group' button is visible. A confirmation dialog is open on the right, asking 'Are you sure you want to delete "partnersummit"?' and listing the affected resources.

NAME	TYPE	LOCATION
batchdemo00	Batch account	West Europe
batchdemo00	Storage account	West Europe

23. Enter the resource group name to confirm resource deletion and click **Delete**:

The screenshot shows the 'Delete resource group' confirmation dialog. The resource group name 'partnersummit' has been entered in the text field. The dialog lists the affected resources: 'batchdemo00' (Batch account) and 'batchdemo00' (Storage account).

NAME	TYPE	LOCATION
batchdemo00	Batch account	West Europe
batchdemo00	Storage account	West Europe