Parallel Job Processing

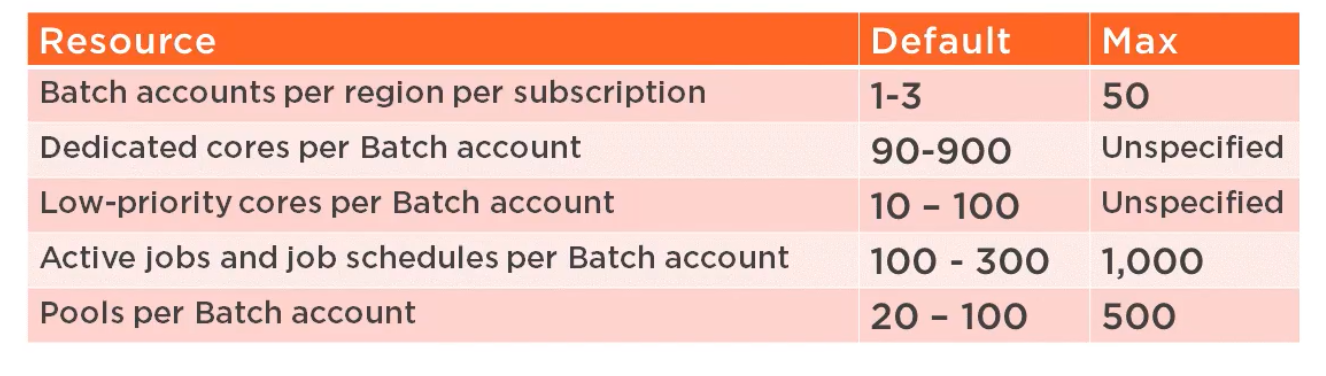
* Job
  + Requires processing/computing
  + Can take a long time to complete
  + Can be broken down into Tasks

Azure Batch core features:

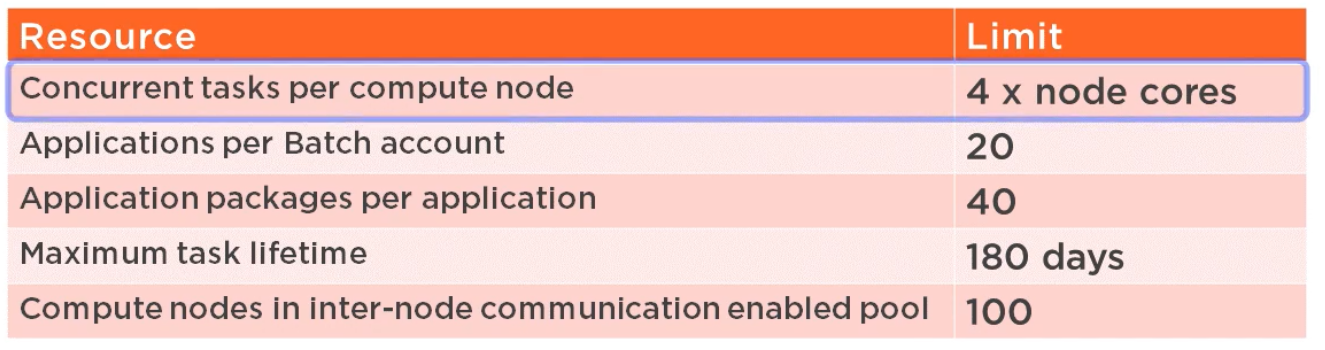
* **Resource Management**
  + Node management
    - Cloud Service Worker Roles
      * Stateless VMs
      * PaaS
      * Simple to work with
      * Limited OS
      * Temp local storage
      * Limited customization
    - VMs (recommended)
      * IaaS
      * Durable stateful storage
      * Many specific OSs, options
      * Full access to OS
    - Third Party Apps
      * Hourly billing
  + Autoscaling
    - Set number of nodes in a pool
    - Defined as a formula
      * Based on service variables, like CPU percentage, memory, number of tasks active/running/pending/failed
  + Low Priority Nodes
    - Cheaper
    - But tasks only processed when spare capacity is available
    - Cool for not time critical jobs, cost is important
  + Applications
    - Our own, third party, etc
    - Must be ZIP with an .exe (and additional DLLs)
    - Deployed to nodes when necessary
    - Can be owned by:
      * **Pools** (deployed to all nodes within)
      * **Tasks** (deployed to nodes running the task)
* Process Management
  + Job and Task management
    - Scheduling, prioritizing, retry, max durations on tasks, etc
  + Handle Task dependencies
  + Handle multi instance tasks
    - Subtasks which can be executed on separate node
    - MPI: Messaging Passing Interface
    - Inter-node communication is required within the Pool
* Resource and Process Monitoring

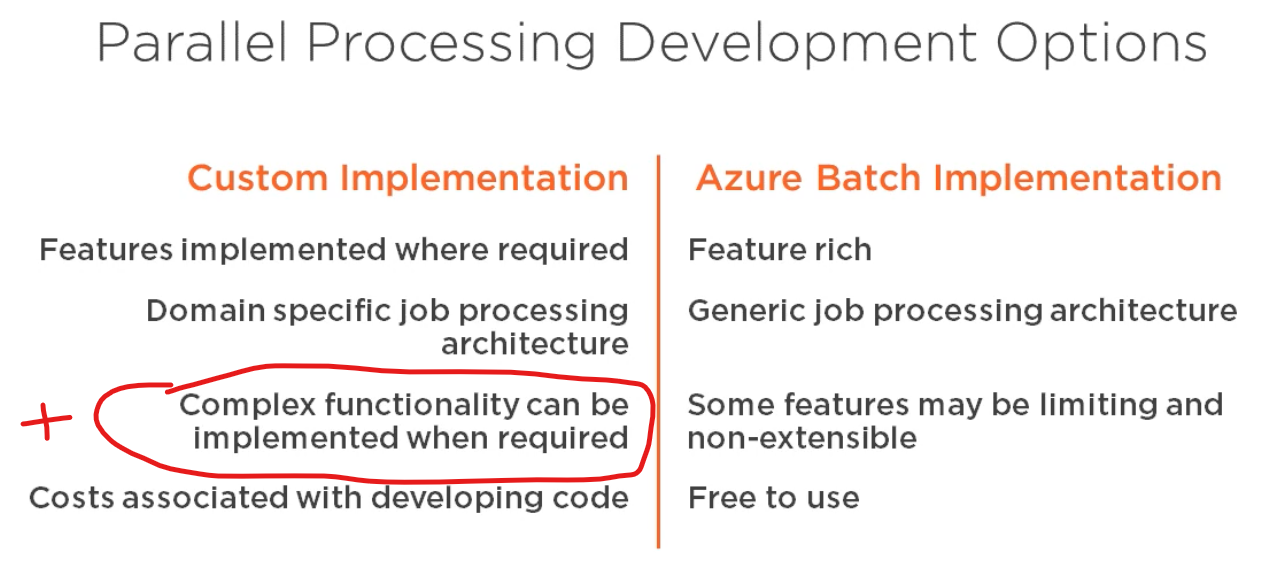
Billing is based on consumption:

* Free to create
* Only charged for resources being used

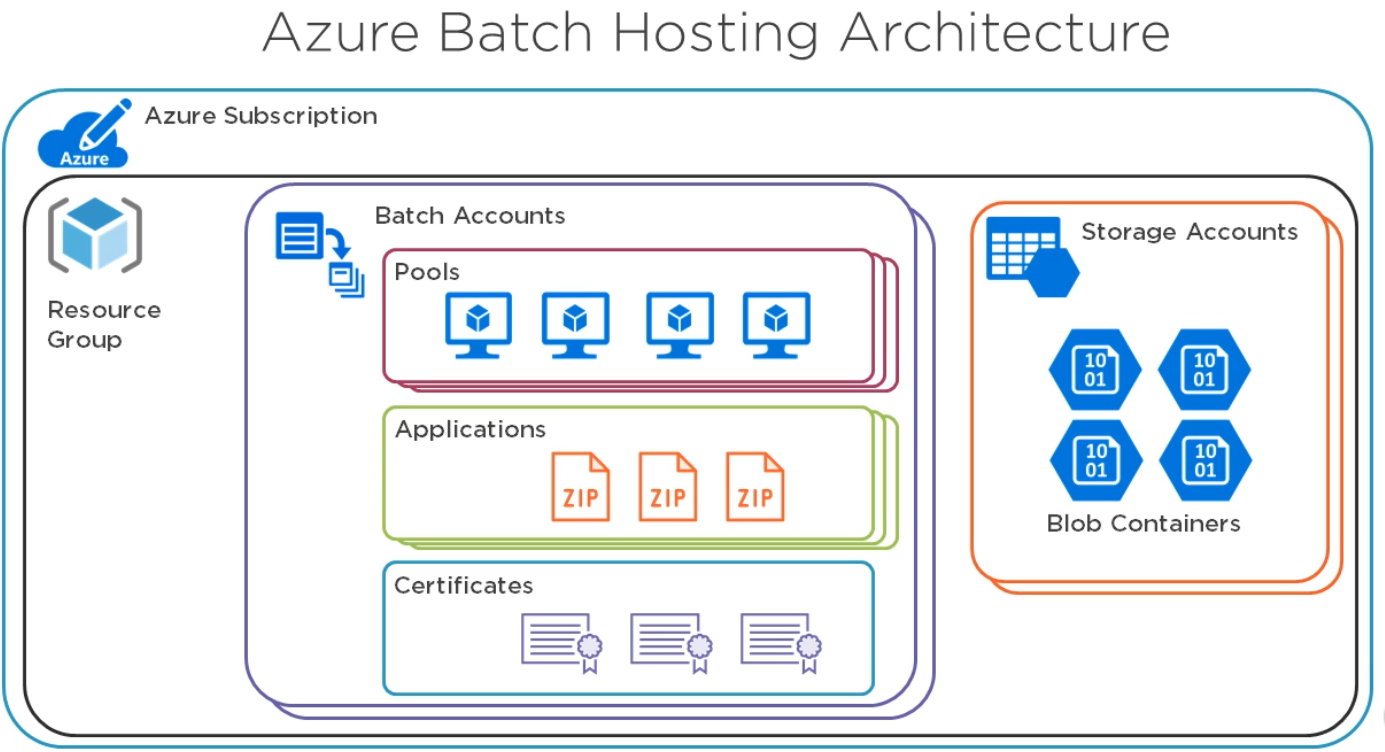


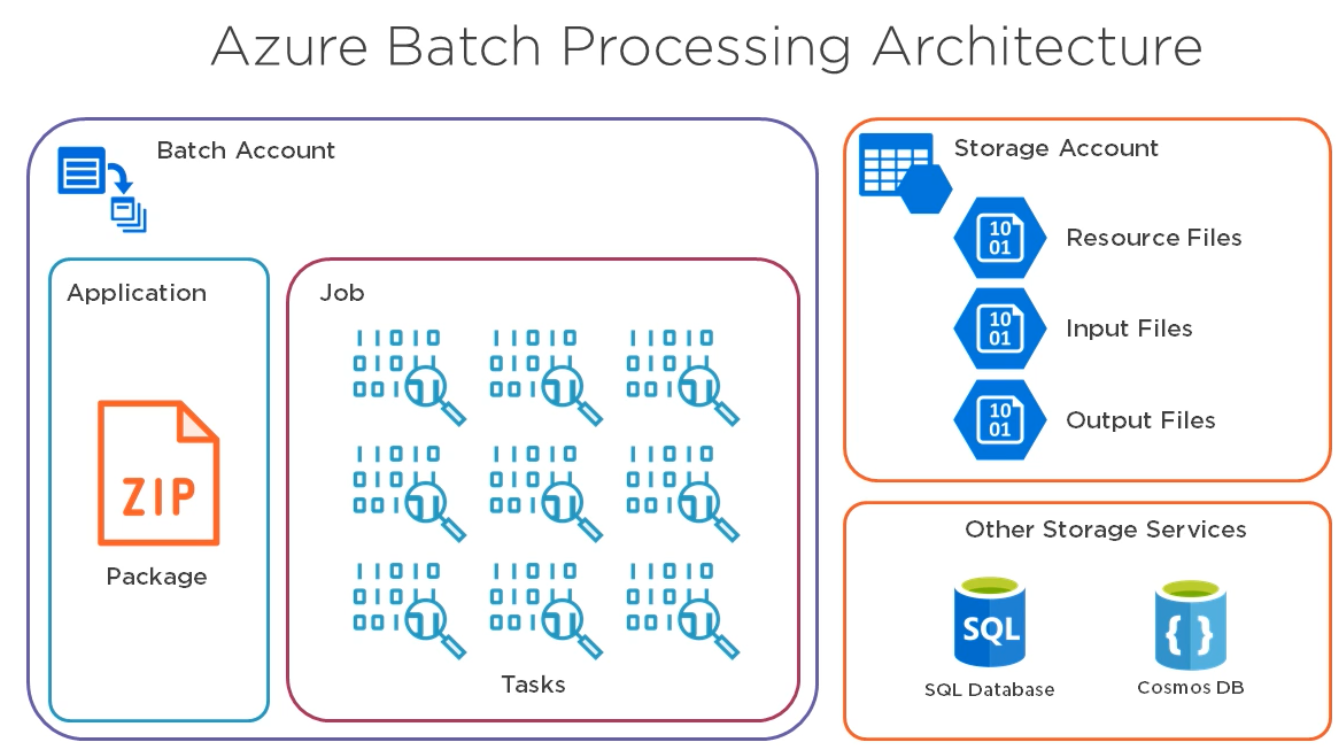
Hard limits:

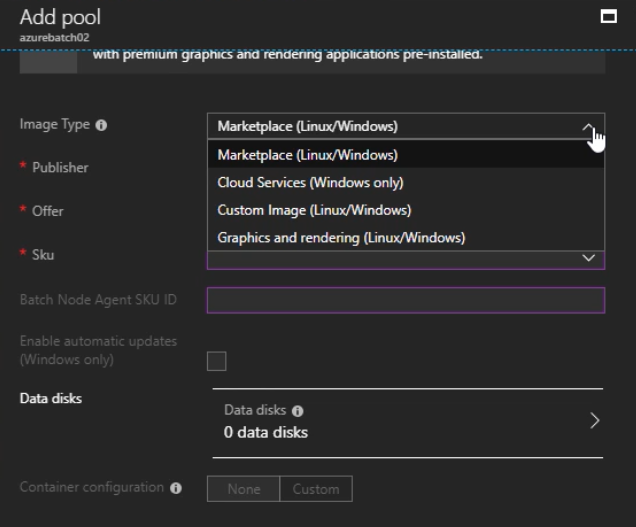
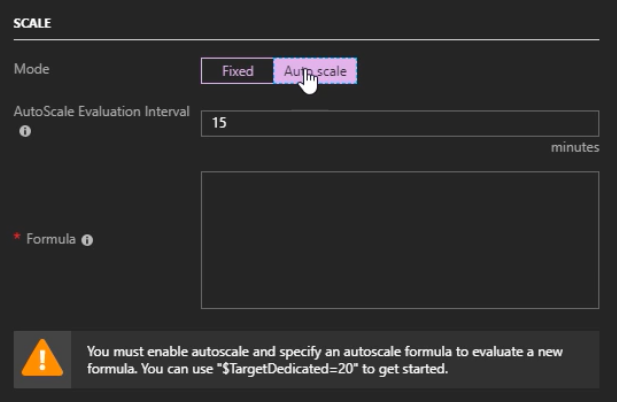




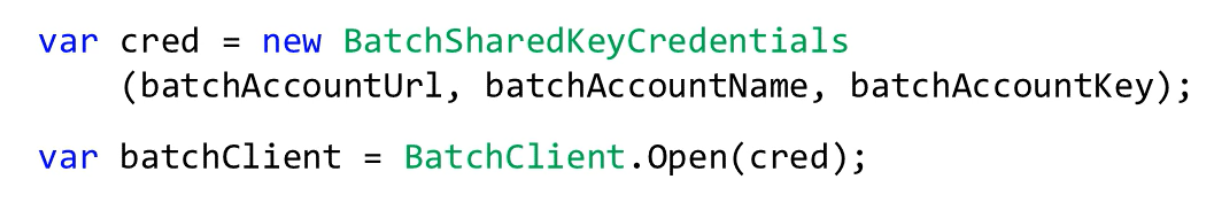
When creating a Batch Service, the creation of a Storage Account is also required

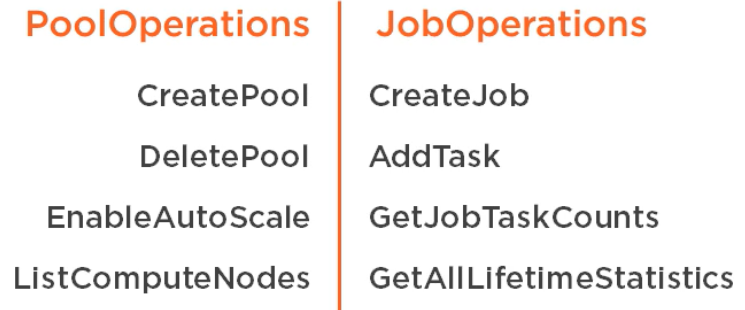


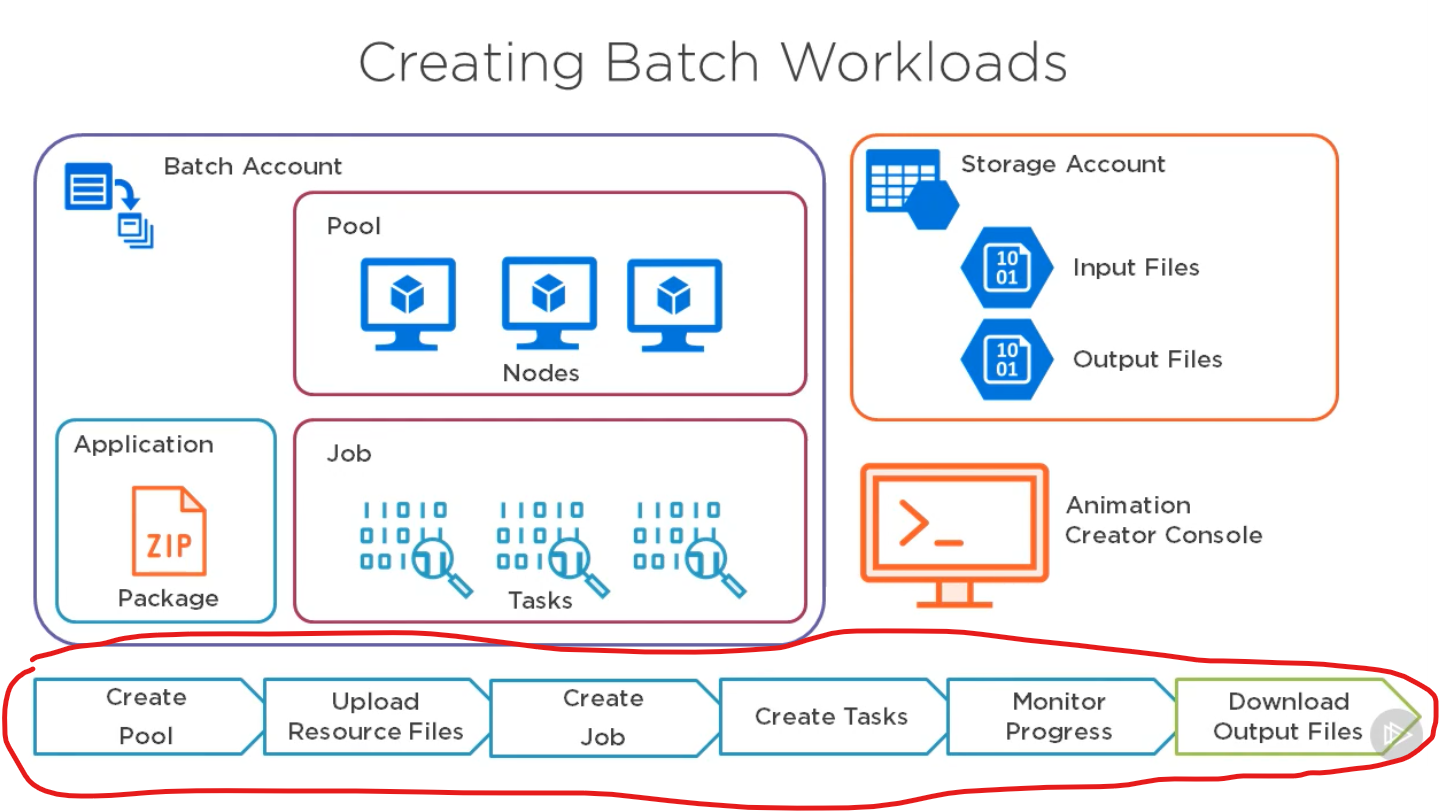


* Create Blob Container(s) in Storage Account (eg. Inputs, Output)
* Batch account -> Applications
  + 20 apps can be defined
  + Upload app package as ZIP
* Create Pool in Batch account
  + Used to manage resource as they are starting to process tasks
  + 
  + 
  + Select Pool’s app to the previously uploaded ZIP
  + Then create a new Job, assign it to use the previously created Pool
  + Then create tasks within the Job
    - SAS: Shared Access Signature: create a reference for the input Blob

Batch client (nuget package)

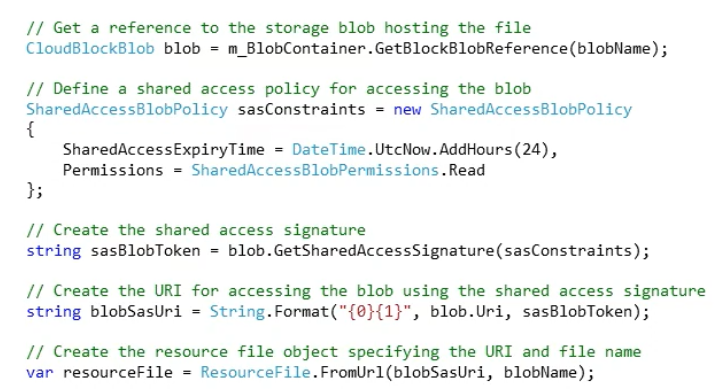




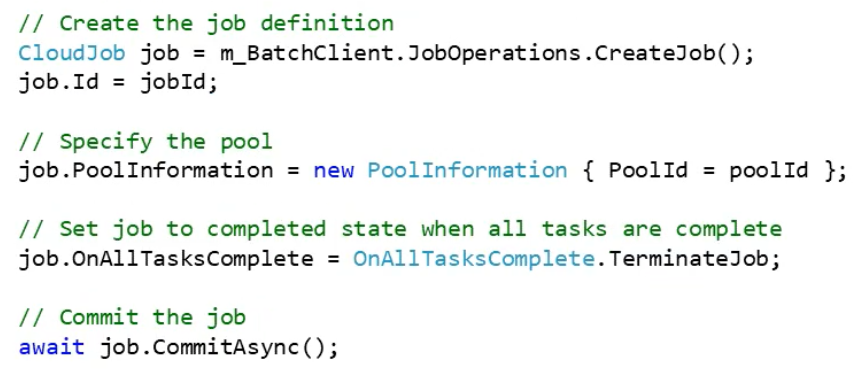


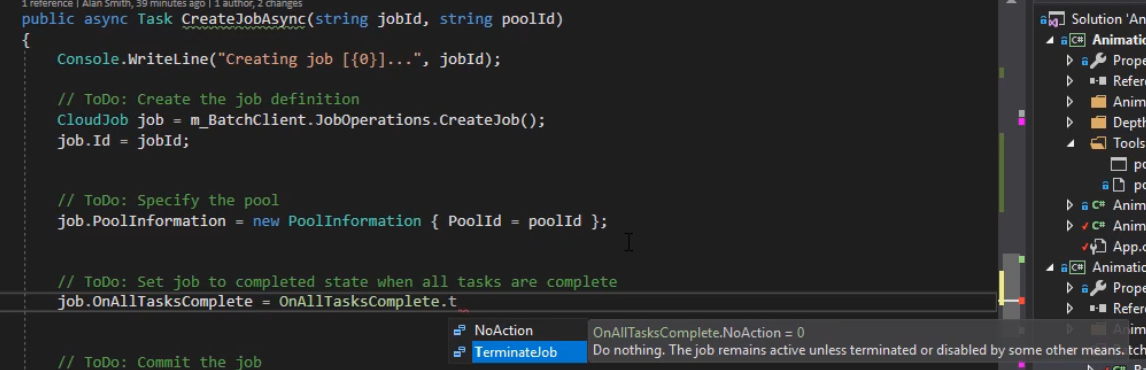
Create pool:

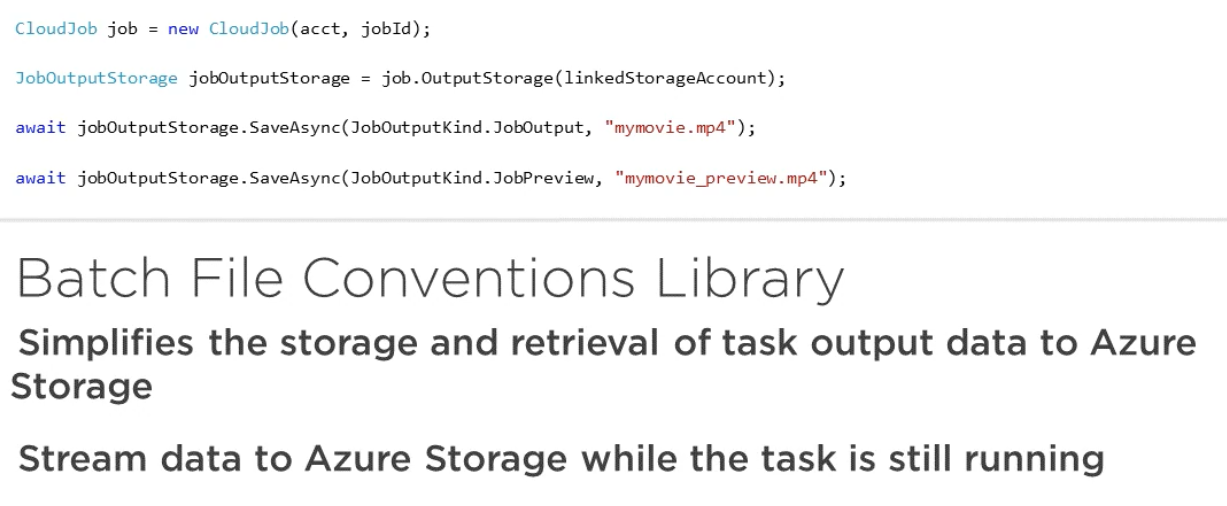
Create resource file:



Create job:







Create task:



