Managing and Serving Imagery in the Cloud Using ArcGIS Image Server

Peter Becker
ArcGIS Provides a Comprehensive Imagery Platform

Seamless integration and analysis of imagery with all spatial data

A scalable platform for working with all forms of imagery (drone, aerial, and satellite) to create valuable information products, integrated with GIS.
Why Manage and Serve Imagery in the Cloud

- You have large collection of imagery.
- Infrastructure is getting expensive
- Have heard that Cloud is: Cheap, Secure, Elastic, Resilient, Simple, … ??
- Organization is going “Cloud First”
Four Options

• Manage Locally → Publish Products to Cloud

• Manage Locally → Copy Data to Cloud → Publish in cloud

• Move Data to Cloud → Manage in Cloud → Publish in cloud

• Only Cloud (Data is in Cloud) → Manage in Cloud → Publish in cloud

Possibly Mix or transition over time

Don’t just copy your data there and expect it all to work
ArcGIS Support for All Imagery and Raster Formats

- Many different modes – Satellite, Aerial, Drones, Categorical, Scientific, ..
- Lots of different types
- Lots of different ‘Products’ from vendors (Imagery+Metadata)
- Many different formats, some better than others
Mosaic Datasets
Enable Management of Large Collections

- Optimum Data Model for Image Management
- Highly Scalable, from Small to Massive Volumes of Imagery
- Defined in GeoDatabase (File or Enterprise)
- References sources
- Maintains metadata
- Defines processing to be applied at Item and Service Level
- Defines mosaicking rules
- Defines overview to be used at small scale

- Access as Image or Catalog
Raster Types in ArcGIS

- Crawls for specific images
- Users Raster Types
- + Python Raster Types
Ways to Create Mosaic Datasets

- Context Menu
- Geoprocessing Tools
  - Data Management Tools / Raster / Mosaic Dataset
- Model Builder
- Ortho Mapping
- Imagery Workflows
- MDCS (Mosaic Dataset Configuration Scripts)
Imagery Workflows

- Preprocessed Orthos
- Elevation
- Landsat
- HighRes Satellite
- Scanned maps
- Scientific
- Historic
  - Landsat
  - Sentinel 2
  - Planet

Most Image Management Workflows use MDCS

http://www.esriurl.com/imagemanagement
Publishing Imagery to the Cloud – The simplest way

- May be all you need is to server as tile cache
- Create Tile Cache using Pro

- Publishing Tile Cache
  - To ArcGIS Online
  - To ArcGIS Enterprise

Using GeoProcessing Tools

From Content Pane
Which Cloud?

• ArcGIS Supported Clouds
  - AWS
  - Azure

• Image Services also work on: Google Cloud, Aliyun, Huawei
What is different in Cloud?

• Blob Storage (Https) vs Local (SAN/NAS)
• Autoscaling
• Security
• Implementation and Management of Infrastructure (Ec2, ELB, Storage)
• Cost : Pay for the use
• Test and Scale as required
Infrastructure options

- AWS:
  - Storage: EBS/S3 Standard/ S3 Infrequent Access /S3 Glacier

<table>
<thead>
<tr>
<th>AWS Storage</th>
<th>Cost</th>
<th>Capacity</th>
<th>Latency</th>
<th>Throughput</th>
<th>Reliability</th>
<th>Shared</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFS</td>
<td>$$$$</td>
<td>GB</td>
<td>Low</td>
<td>High</td>
<td>Med</td>
<td>Yes</td>
</tr>
<tr>
<td>S3 Standard</td>
<td>$</td>
<td>PB</td>
<td>High</td>
<td>Med</td>
<td>V.High</td>
<td>Yes</td>
</tr>
<tr>
<td>Glacier</td>
<td>$</td>
<td>PB</td>
<td>-</td>
<td>V.Low</td>
<td>V.High</td>
<td>Yes</td>
</tr>
<tr>
<td>EBS</td>
<td>$$$</td>
<td>TB</td>
<td>Med</td>
<td>High</td>
<td>Med</td>
<td>No</td>
</tr>
<tr>
<td>Ephemeral</td>
<td>Included</td>
<td>40/80GB</td>
<td>V.Low</td>
<td>V.High</td>
<td>Low</td>
<td>No</td>
</tr>
</tbody>
</table>

- Compute – EC2 - Compute Options: Recommendation
  - M5.2xlarge (Portal) if using Feature service a lot use Higher Instance
  - M5d.xlarge (Image server) (150 GB SSD)
  - M5.large (File Server)
AWS Options to set up ArcGIS Enterprise

The below link gives an idea on what is ArcGIS Enterprise on AWS


- **Using AWS Cloud Formation Templates:**

- **Using ArcGIS Server Cloud Builder**
  - https://enterprise.arcgis.com/en/server/latest/cloud/amazon/arcgis-server-architectures-on-aws.htm#ESRI_SECTION1_04B5C34B31D54B8AB29AE14DD0F30F2F

- **Using AWS management Console (setup Manually).**
  - https://enterprise.arcgis.com/en/server/latest/cloud/amazon/arcgis-server-architectures-on-aws.htm#ESRI_SECTION1_E0C2D729F89048FF87A93D2114486A66

<table>
<thead>
<tr>
<th>Type</th>
<th>Advantages</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CloudFormation Templates</td>
<td>Easy for Automation, Standard from Amazon, have better control on working with various AWS services, Easy to setup</td>
<td>Need Better understanding on how AWS services works, It's for more advance users. Managing can become tricky in cases</td>
</tr>
<tr>
<td>ArcGIS Server Cloud Builder</td>
<td>Easy to setup, easy to start and stop the site, can templatize</td>
<td>Difficult to integrate with other AWS services, need Secrete Keys, advance customization might be difficult</td>
</tr>
<tr>
<td>Management Console (Manual)</td>
<td>Full control on how the system works,</td>
<td>Can be used only by advance users. Needs to setup all components and connect them. Need to be proficient in the ArcGIS Enterprise and AWS</td>
</tr>
</tbody>
</table>
Azure Options to set up ArcGIS Enterprise

The below link gives an idea on what is ArcGIS Enterprise on Azure


• Using ArcGIS Server Cloud Builder for Azure

• Automate your Azure Deployment
Workflow Overview

- Copy data to cloud / Optimize
- Create Mosaic Dataset (on Cloud or Locally)
- Publish
- Access
- Run Analytics
What Affects Performance

• Volume of data read
  - Only read what is required (tiling/pyramids)
  - ArcGIS will only read pixels required
  - Use of Footprints vs. NoData

• Process Efficiency
  - Process only pixels required
  - Concatenate functions to reduce I/O
  - Use raster functions to speed processing

• Latency
  - Reduce number of requests (especially duplicate)
  - Tile Level Caching

• Bandwidth
  - Reduce data volume through compression (but don’t increase CPU)

• Data structure
  - Ensure efficient data access
Formats - Importance of Data Structure

- Data format can have a huge effect on performance and costs
  - Tiling – Enable faster random access
  - Pyramids (Overviews or RRDs) – Enable faster access at lower resolutions
  - Compression – Reduces storage and transmission
  - Metadata access
  - Multi Write / Update?

![Diagram showing different data formats and their structures](image-url)
Formats - Importance of Data Structure

- Data format can have a huge effect on performance and costs
  - Tiling – Enable faster random access
  - Pyramids (Overviews or RRDs) – Enable faster access a lower resolutions
  - Compression – Reduces storage and transmission
  - Metadata access
  - Multi Write / Update?

CRF (Cloud Raster Format)

CRF Enables multi-write while minimizing number of files
Converting and Getting Imagery into Cloud

- Converting data
  - ArcGIS Export
  - Copy Raster (for CRF)
  - Use OptimizeRasters [https://github.com/Esri/OptimizeRasters](https://github.com/Esri/OptimizeRasters)
  - GDAL

- Uploading data
  - CloudBerry etc
  - Optimize Rasters (Convert and upload)
  - White Glove Services Snowball ($200/50TB) Snowmobile
  - Amazon cli Command or Console
Overview of Optimize Raster

- Converts Raster from one format
- Transfer data to and from cloud storage (AWS, Azure, Google, Alibaba) or enterprise storage
- Create Raster Proxies
- Runs with in ArcGIS (ArcGIS Pro or ArcMap) or as a standalone utility
- Supports for Windows and Linux
- Parallel Processing is supported by default
- Logging support
- Converts the data using the best practices (templates provided)

https://github.com/Esri/OptimizeRasters
Storage Security Options

• Obfuscate
• IAM Role (Recommended)
• Bucket Polices
• Access Control List
• Cross Origin Resource Sharing (CROS): used for allowing web application to access the data
• …
Storage Security Recommendations

Most Common Options

- **Public Bucket – Anyone Access – Eg Public Data**
  - Adv: Simple to set up. Can use vsicurl (no credentials)
  - Dis: No control of access (use Obfuscate), Egress can be expensive

- **Public Bucket – Requestor Pays (outside region)**
  - Adv: No Egress costs
  - Dis: User must have account and provide minimal credential (Environment Variable). Need VSIS3

- **Secured Bucket – User must have credentials**
  - Adv: Change security when required
  - Dis: Need to set policies and roles, Need VSIS3

- **Secured Access Signature (SAS)**
  - Adv: Allow access for a specific period of time, or till revoked
  - Dis: Need to manage permissions/credentials, Need VSIS3

• List all data, List Buckets/Containers
• Read data, Write data
• Requestor Pay to read the data
• Read Permissions, Modify the Permissions

Set up soon to be documented on Imagery Workflows
Accessing Imagery from Cloud Storage

• Direct Access using \vsicurl \vsis3 (Using Table Raster)
• Raster Proxies
  - XML file that references file on cloud storage
  - Enables local caching
  - Create using Optimize Rasters

• Set Environment Variables
  - AWS_ACCESS_KEY_ID=<accesskey>
  - AWS_REQUEST_PAYER=requester (if it is user payer only then needed)
  - AWS_SECRET_ACCESS_KEY=<secret_key>
  - GDAL_HTTP_UNSAFESSL=TRUE

• Cloud Storage Connection file (.ACS)
Scaling Mosaic Dataset

- By default Mosaic Datasets reference files on disk
- Embed URL into Mosaic dataset
- Use ACS or Raster Proxy
- If using Raster Proxy Embed
  - Use Table Raster type with Raster field that defines content of Raster Proxy
  - Use MDTool Embed (On Image Management Workflows)
  - Create Python Raster Type

- File Geodatabase or Enterprise (e.g. Postgres or RDS)
- Use RDS when:
  - Update of the source data is frequent
  - Dynamic Image service are needed to be created at runtime as soon as the data comes in.
  - Use Raster Analytics (irrespective of Mosaic dataset size)
Recommended Cloud Implementation

Enterprise + ArcGIS Server + Image server + RDS

- Elastic Load Balancer
- Imagery S3 Storage
- Elastic IP for Portal
- VPC
- Portal: m5.2xlarge 1 machine
- Image server: m5.2xlarge 3 machine
- File Store: m5.xlarge 1 machine
- Postgres: RDS
- Load Balance for Image server stack
- Elastic IP for the portal machine
- All machines in same VPC

Dynamic Image Services
- EC2
- Image Services
- Auto Scaling group
- VPC

Hosted Image Services
- EC2
- Image Services
- Raster Analytics

Configure Store
- EC2
- RDS
- Postgres RDS

S3 Storage User
- Raster Store, zip, fgdb

Federated with
- Portal
- AGS

More info on how to do ArcGIS enterprise Deployment
More info on how to do this Imageserver Deployment

ArcGIS Enterprise Cloud Formation Template
Image server Cloud Formation Template

- Used when we need to do raster analytics
- Mosaic sit in Postgres (RDS)
- All server machines are in cluster
- Config store is on a different EC2 machine
Setting up AWS Infrastructure

- Amazon account with full access to EC2 and other resources.
- If you are using ArcGIS Amazon Image (AMI) for the first time accept terms in the AWS Marketplace.
- A valid domain name for your site.
- A TLS (SSL) certificate for your domain, obtained from a certifying authority.
- An Elastic IP address that you will associate with the EC2 instance; you must map your domain name to the Elastic IP address.
- License file for Portal, ArcGIS server, Image server
- If the amazon account is new we recommend creating the below things
  - VPC, ElasticIP, KeyPair. Elastic Load Balancer
- Also note down the VPC id, SiteEIPAllocationID for the elasticIP, keypair

ArcGIS Enterprise Deployment + Image Server

• **Enterprise**
  - M5.2xlarge (1 instance)
    - Readme: [https://s3.amazonaws.com/arcgisstore106/8321/docs/ReadmeAllInOneVPC.html](https://s3.amazonaws.com/arcgisstore106/8321/docs/ReadmeAllInOneVPC.html)

• **Image Server**
  - M5.2xlarge (1 instance) (in autoscaling mode) (for image server stack)
    - db.r4.xlarge (RDS) for high 20gb+ Mosaic datasize, or db.m4.xlarge
    - You will need an ELB, you can use the below Cloudformation template to create an ELB
      - Then map the ELB name to the required domain name
    - Obtain following: Domain Name, SSL certificate, Image server License File (note you edit the file to mention your name and details)
    - If using RDS select DBengine and select postgres for it.
    - ReadMe: [https://s3.amazonaws.com/arcgisstore106/8321/docs/ReadmeHAServerVPC.html](https://s3.amazonaws.com/arcgisstore106/8321/docs/ReadmeHAServerVPC.html)

• **Raster Analytic Server**
  - Similar to Image server
  - No need to setup RDS, use the above one.
Setting up Image Server and RDS

- User CloudFormation Template or ArcGIS Enterprise Cloud Builder CLI for AWS
- This template will deploy 4 machines,
  - 2 Image services (need to define site admin password in the stack)
  - One RDS (same user name password as site admin)
  - One File server
- Generate the login password for all the machines, enable the RPP
- Copy the RDS instance name
Elastic Deployment using Mosaic Dataset on RDS

Client consuming Imagery Items

ArcGIS Online

Image Services shared on AGOL

Elastic Load Balancer

EC2

Image Services

Image Server

Auto Scaling group

RDS

Postgres RDS

Public S3

S3 storage for other files (zip, etc)

Log S3 - ESRI

EC2

ArcGIS Desktop

AGS

Mosaic Dataset Tools

Local Dev
Elastic Deployment using Mosaic Dataset on S3

Client consuming Imagery Items

ArcGIS Online

Image Services shared on AGOL

Elastic Load Balancer

Client consuming Image Services

Log S3 - ESRI

S3 storage for other files (zip, etc)

Image Services

EC2

Image Services

Image Server

Auto Scaling group

VPC

Image Services

EC2

Image Services

Image Server

Imagery S3

EC2

ArcGIS Desktop

AGS

Mosaic Dataset Tools

Local Dev
Managing the mosaic and Updating the services

- Uses MDCS to build the mosaic
- The mosaic can be build in Postgres or FGDB
- Build footprint, Set Properties, Calculate Statistics, updateGeometry
- Stretch, Color correction
- Use AFL Deplorer to update the service
  - [http://www.arcgis.com/home/item.html?id=fa5c4dfa07094bcfbe4a16c28f43be85](http://www.arcgis.com/home/item.html?id=fa5c4dfa07094bcfbe4a16c28f43be85)
  - [http://www.arcgis.com/home/item.html?id=b01b18bbf0e34338b6a2c71609ea1373](http://www.arcgis.com/home/item.html?id=b01b18bbf0e34338b6a2c71609ea1373)
Post Install Script for Image server

• This script will run once the machine is setup
• It is used is you need to run some of your scripts on the startup of the server
• You must upload the zip file to the Amazon S3 bucket you specified for the "DeploymentBucket".
• There should be a deploy.bat file
Scaling to Raster Analytics

Enterprise GIS
Web GIS Layers
GIS Data & Imagery
New Web GIS Layers
Gdb Files
WCS Services
ArcGIS Services

Desktop
Web
Device

ArcGIS Pro
Design & Run Models
Portal UX
Run Portal Analysis
Developers & System Integrators
Device & Run Model (Parameterize)

Model Execution Distribution

analysis results as a new Web GIS Layers

Integration to Additional Deep Learning etc
Please Take Our Survey on the App

Download the Esri Events app and find your event.

Select the session you attended.

Select the Feedback tab.

Complete answers and select “Submit.”