Imagery in the Cloud and Raster Analytics

Mike Muller
manage and process imagery into authoritative data sources that are appropriately and efficiently disseminated to those that need access

derive actionable information from imagery and rasters by performing analytics on massive volumes of data available from multiple sources

enable access to imagery and analysis through a wide range of integrated desktop, mobile, and web applications that are interactive, informative, and engaging

manage and process imagery into authoritative data sources that are appropriately and efficiently disseminated to those that need access
What is Raster Analytics?

- **ArcGIS has a new way to create and execute spatial analysis models and image processing chains which leverage distributed storage and analytics**

  - Raster Analytics works with your existing GIS data and imagery
    - register your data and go, importing existing data to distributed storage is not mandatory

  - Raster Analytics can optimize your data for distributed analytics
    - import your data into ArcGIS distributed storage which further improves the scalability of distributed analytics

  - Raster Analytics is designed to scale with your organization’s demands
    - scale up to get the job done, scale down when resources are no longer needed
Raster Analytics Foundational Concepts

- Raster Analytics adds to existing ArcGIS foundational concepts

Dynamic Raster Models
  - on-the-fly processing

Geoprocessing Models
  - powerful analytics

Server-based Distributed Raster Analytics with Distributed Raster Data Storage
  - (persistent) distributed analytics with optional distributed storage for even greater scalability

Web GIS Layers
  - rich geoinformation model

new

more

more
Solve New Problems with Raster Analytics

• run models against data that is too big for single desktop
  - small and medium scale global rasters (big geography)
  - large scale local or regional rasters (high resolution)

• run models against massive collections and scale it

• run models and meet time constraints
Raster Analytics is Powerful

- run a model based on a single function

<table>
<thead>
<tr>
<th>Math</th>
<th>Square</th>
<th>Not Equal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abs</td>
<td>Square Root</td>
<td>ArgStatistics</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>Times</td>
<td>Cell Statistics</td>
</tr>
<tr>
<td>Band</td>
<td>Bitwise And</td>
<td>Statistics</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>Bitwise Left</td>
<td>ACos</td>
</tr>
<tr>
<td>Calculator</td>
<td>Bitwise Not</td>
<td>ACosH</td>
</tr>
<tr>
<td>Divide</td>
<td>Bitwise Or</td>
<td>ASin</td>
</tr>
<tr>
<td>Exp</td>
<td>Bitwise Right</td>
<td>ATan</td>
</tr>
<tr>
<td>Exp10</td>
<td>Shift</td>
<td>ATan2</td>
</tr>
<tr>
<td>Exp2</td>
<td>Bitwise Xor</td>
<td>ATanH</td>
</tr>
<tr>
<td>Float</td>
<td>Boolean And</td>
<td>Cos</td>
</tr>
<tr>
<td>Int</td>
<td>Boolean Not</td>
<td>CosH</td>
</tr>
<tr>
<td>Ln</td>
<td>Boolean Or</td>
<td>Sin</td>
</tr>
<tr>
<td>Log10</td>
<td>Boolean Xor</td>
<td>SinH</td>
</tr>
<tr>
<td>Log2</td>
<td>Equal To</td>
<td>Tan</td>
</tr>
<tr>
<td>Minus</td>
<td>Greater Than</td>
<td>TanH</td>
</tr>
<tr>
<td>Mod</td>
<td>Greater Than</td>
<td>Equal</td>
</tr>
<tr>
<td>Negate</td>
<td>Equal</td>
<td></td>
</tr>
<tr>
<td>Plus</td>
<td>Is Null</td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td>Less Than</td>
<td></td>
</tr>
<tr>
<td>Round Down</td>
<td>Less Than</td>
<td></td>
</tr>
<tr>
<td>Round Up</td>
<td>Equal</td>
<td></td>
</tr>
</tbody>
</table>

- run a model by combining many functions

<table>
<thead>
<tr>
<th>Correction</th>
<th>Apparent Reflectance</th>
<th>Geometric Correction</th>
<th>Speckle Filtering (Lee,Frost,Kuan)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visualization &amp; Appearance</td>
<td>Contrast and Brightness</td>
<td>Convolution</td>
<td>Pan sharpening</td>
</tr>
<tr>
<td>Analysis: Density</td>
<td>Kernel Density</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis: Band Math &amp; Indices</td>
<td>NDVI / NDVI Colorized</td>
<td>SAVI / MSAVI / TSAVI</td>
<td>GEMI</td>
</tr>
<tr>
<td></td>
<td>GVI (Landsat TM)</td>
<td>PVI</td>
<td>Tasseled Cap (Kauth-Thomas)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Binary Thresholding</td>
</tr>
<tr>
<td>Analysis: Overlay</td>
<td>Weighted Sum</td>
<td>Weighted Overlay</td>
<td></td>
</tr>
<tr>
<td>Analysis: Zonal</td>
<td>Zonal Statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis: Interpolation</td>
<td>Natural Neighbor</td>
<td>Nearest Neighbor</td>
<td>Inverse Distance Weighted</td>
</tr>
<tr>
<td></td>
<td>Empirical Bayesian Kriging</td>
<td>Swath</td>
<td></td>
</tr>
</tbody>
</table>
Raster Analytics is Easy

• easy to get started, it is “out of the box analytics”
  - install on nodes -> start Raster Analytic services -> go

• ArcGIS Pro user experience
  - just works with layers
  - visual modeler to design simple and complex models

• results are immediately available as services
  - no publishing workflow required
• **using your own registered data**
  - registered data can be used as input but not output
  - models running against single rasters can be parallelized by block (*as long as the model allows it*)
  - models running against a collection of rasters will be parallelized per raster in the collection
  - performance can be susceptible to underlying image format (TIFF vs. JP2)

• **using ArcGIS distributed storage**
  - easy to use import tool gets your data into Raster Analytics optimized storage
  - CRF (Cloud Raster Format)
    - multi-band, block based, multiple readers, multiple writers, fast
  - CRF is a format optimized for Raster Analytics computations

• **all outputs of Raster Analytics are written in parallel to ArcGIS distributed storage**
  - running models on new Web GIS layers is inherently optimized
Raster Analytics in Your Infrastructure

- **deployed as Enterprise / Web GIS on-premise**

- **your infrastructure can be…**
  - your hardware
  - your Amazon
  - your Azure

- **deployment tools**
  - Amazon CloudFormation Templates
  - ArcGIS Enterprise Cloud Builder for Microsoft Azure
Raster Analytics can power systems that need to execute spatial analysis and image processing models in a distributed and scalable environment. It is designed for users, developers, and system integrators.

Results are stored in distributed storage and are immediately available as new Web GIS Layers which are already optimized for further analytics.
Raster Analytics Test Case: Terrain Suitability

Global SRTM 90m

Raster Analytics Processors
- 16GB RAM, 8 cores, NAS storage

Terrain Suitability Model
- Compute slope
- Compute aspect
- Remap
- Overlay

Global Terrain Suitability Raster

13.12 hours
80 minutes
Raster Analytics Test Case: Solar Power Plant Suitability

WebGIS (Image Server cluster) on Amazon
- 8 c3.2xlarge instances (8 vCPUs, 16GB RAM)

Mean Rainfall
Mean Temperature
Elevation
Landcover

30m National Solar Plant Suitability Raster

ArcGIS Pro
5 hours 45 minutes

Raster Analytics
9 minutes
**Raster Analytics Test Case: Landsat Processing**

**Infrastructure**
- ArcGIS Enterprise GIS on AWS
- Distributed Raster Analytics (Image Server) Cluster
  - single node
  - AWS c3.8xlarge
  - 60GB RAM, 32 cores, 500GB SSD
  - 200 Raster Analytics Processors

**Input Collection**
- Landsat GLS 1990
  - 7422 Multispectral Scenes
  - S3 storage

**Processing**
1. (foreach) input scene
2. mask no data
3. top of atmosphere correction
4. modified soil adjusted vegetation index
5. remap to classes
6. output thematic raster

**Output**
- Thematic Rasters
  - 7422 Thematic Rasters
  - Distributed Raster Datastore

**Timeline**
- 2 hours 48 minutes
- 44 scenes per minute
- ¾ scene per second
ArcGIS 10.4
Comparison
ArcGIS 10.4 Raster Analytics and Image Processing

- ArcGIS 10.4 has scalable high performance analysis of big rasters and imagery for visual analytics
- on-the-fly processing of massive images and massive image collections
- desktop and server
- visual results can be exported

Dynamic Raster Models

Mosaic Datasets published

Image Services

Desktop can be run on the desktop

Server can be run on the server

Raster Functions
ArcGIS 10.4 Raster Analytics and Image Processing

- ArcGIS 10.4 has scalable high performance analysis of **standard rasters and imagery for persistent analytics**
- processing of single images or spatial subsets of massive images or mosaicks
- desktop and server
- persistent results

**Geoprocessing Models**

*GP / SA / Data* published *GP Services*

*can be run on the desktop*  
*can be run on the server*
What differentiates Raster Analytics from ArcGIS 10.4?

• Raster Analytics are “out of the box” and “ready to use” within your ArcGIS system
  - today you have to explicitly author and publish the specific analytics you need
  - faster prototyping and R&D

• Raster Analytics gives you tools and operations that work against existing layers and future layers within your ArcGIS system – built for Web GIS

• Raster Analytics helps you get “big jobs” done faster
  - you don’t have to partition the job yourself – built for big jobs, big data
  - elasticity and scalability that doesn’t come with desktop workflows

• Raster Analytics is for massive collection processing with persistent results
  - product generation, automated production systems (TCP\text{PED})
  - for systems that can’t rely on visual analytics and on-the-fly product generation
Please Take Our Survey on the Esri Events App!

Download the Esri Events app and find your event

Select the session you attended

Scroll down to find the survey

Complete Answers and Select “Submit”

* Download the Esri Events app and find your event
* Select the session you attended
* Scroll down to find the survey
* Complete Answers and Select “Submit”
Print Your Certificate of Attendance
Print stations located in the 140 Concourse

<table>
<thead>
<tr>
<th>Monday</th>
<th>Tuesday</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>12:30 PM – 6:30 PM</strong></td>
<td><strong>10:45 AM – 5:15 PM</strong></td>
</tr>
<tr>
<td><strong>GIS Solutions Expo,</strong></td>
<td><strong>GIS Solutions Expo,</strong></td>
</tr>
<tr>
<td><strong>Hall B</strong></td>
<td><strong>Hall B</strong></td>
</tr>
<tr>
<td><strong>5:15 PM – 6:30 PM</strong></td>
<td><strong>6:30 PM – 9:30 PM</strong></td>
</tr>
<tr>
<td><strong>Expo Social,</strong></td>
<td><strong>Networking Reception,</strong></td>
</tr>
<tr>
<td><strong>Hall B</strong></td>
<td><strong>Smithsonian National Air</strong></td>
</tr>
<tr>
<td></td>
<td><strong>and Space Museum</strong></td>
</tr>
</tbody>
</table>