

Building Scalable Services:

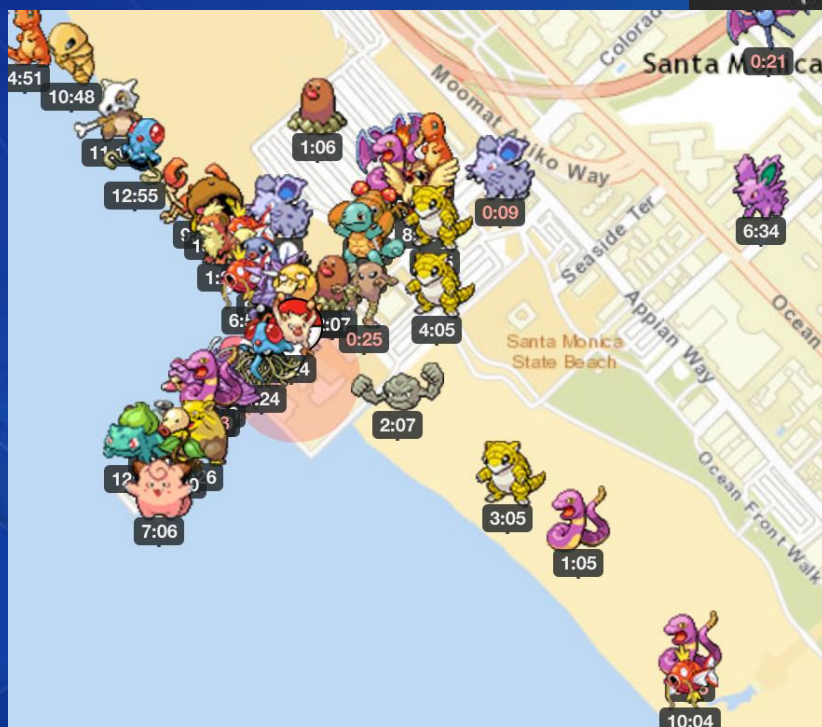
Surviving Pokemon and other Natural Disasters

Jim Mason, Paul Dodd & Vivek Gupta

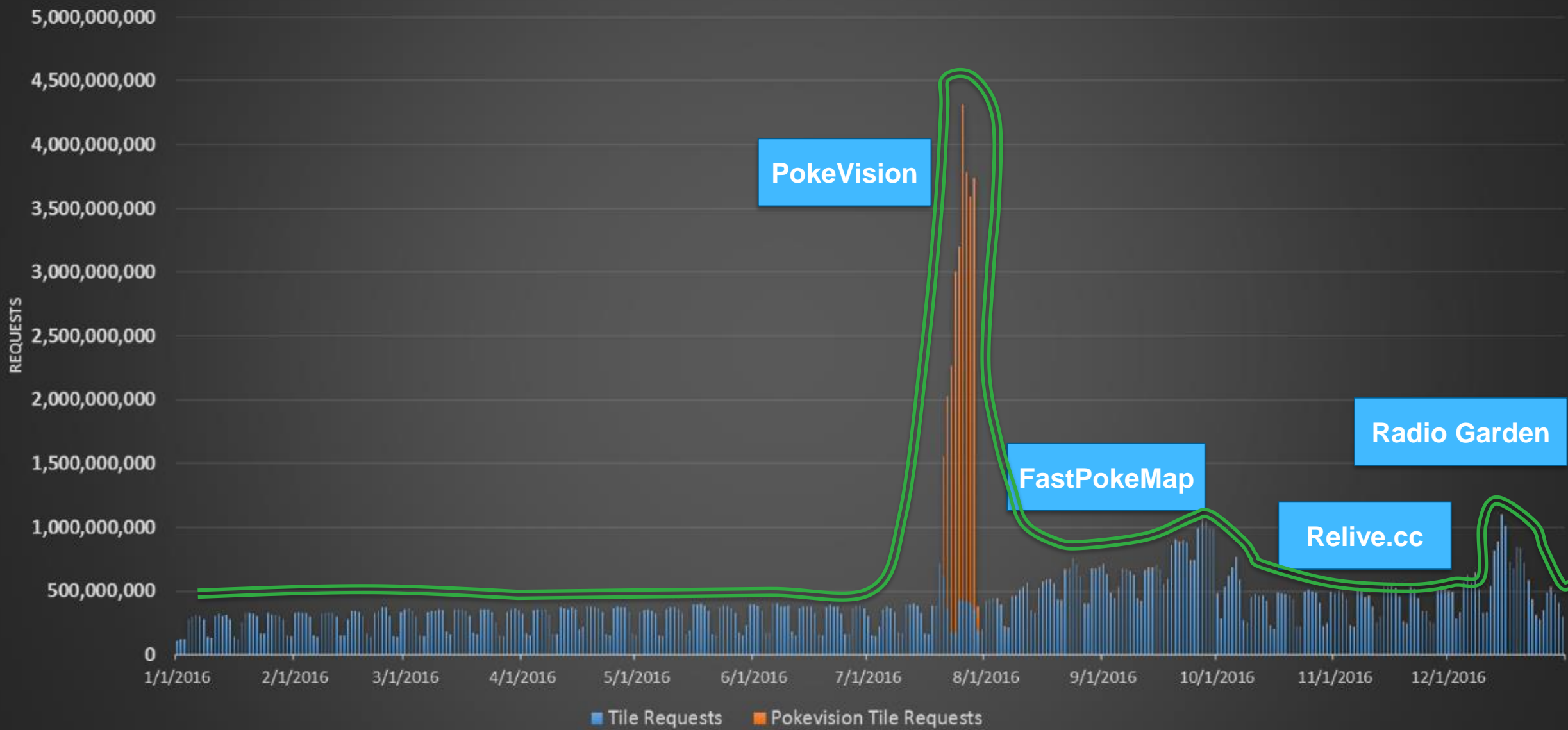
Pokemon GO



PokeVision



2016 Tile Requests by Day

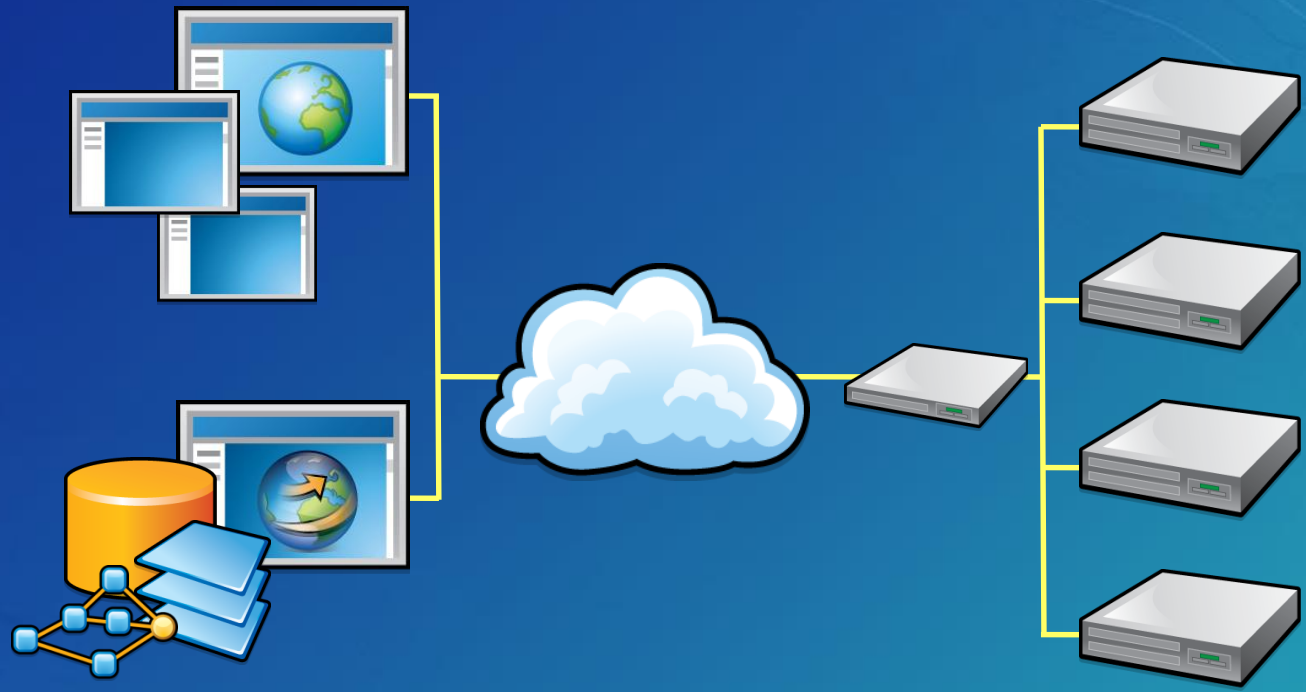


Responding to Natural Disasters



esri

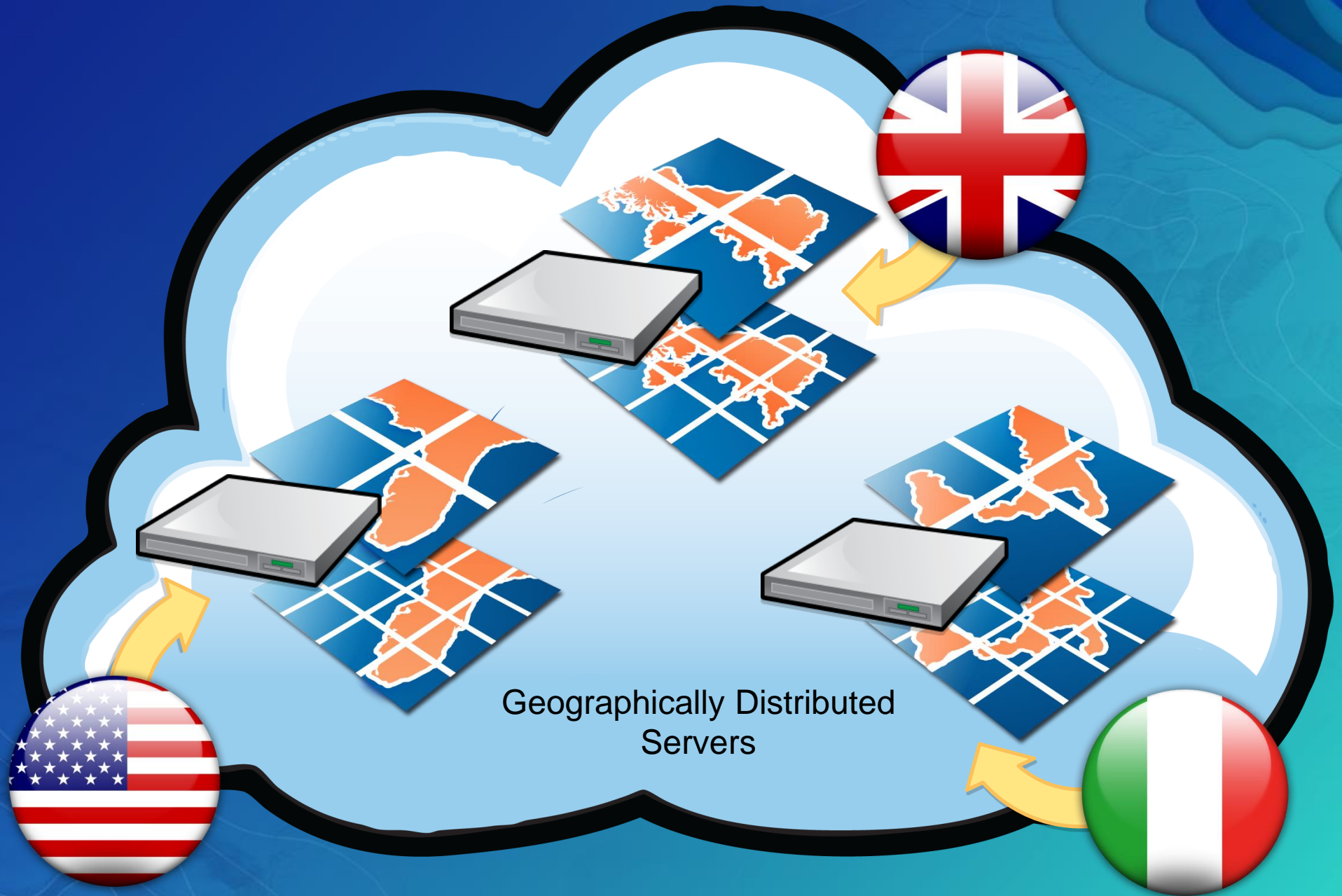
**THE
SCIENCE
OF
WHERE**



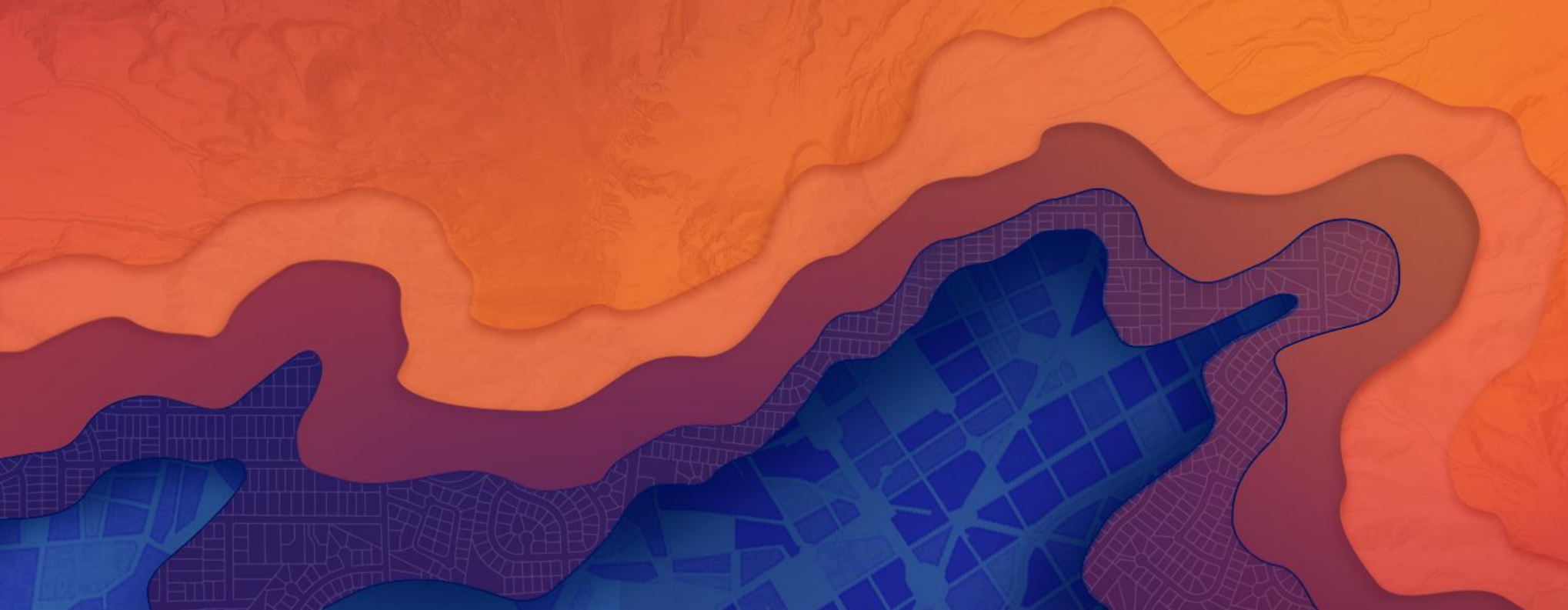
Users

Basemap
Servers





Key Considerations



Key Considerations

- **Data Optimization**
- **Infrastructure Setup**
- **ArcGIS Server/Service Configuration**
- **Geographic Distribution**
 - **Multiple Data Centers**
 - **Geo Load Balancing**
 - **Content Delivery Network**

Data Optimization - Tips

- **Reduce Chokepoints**

- **Replicate Databases** (LBs, AZURE SQL, AWS RDS)
- **Distribute I/O** (DAS, RAID, SAN)
- **Distribute Data/Cache Files** (multi-server, CDNs, AWS S3)

- **Optimize and Delivery**

- **Remove fields, check sizes and types, indexes**
- **Reduce Geometry storage**
 - **Precision** (coordinate resolution)
 - **Dissolve** (multipart and/or simplify)
- **Optimize Raster** (Tiled TIF / MRF, remote storage by Proxy)
- **Automate creation and delivery**
 - **Aggregated Live Feed (ALF) and Mosaic Dataset Configuration Script (MDCS)**

The Aggregated Live Feed methodology

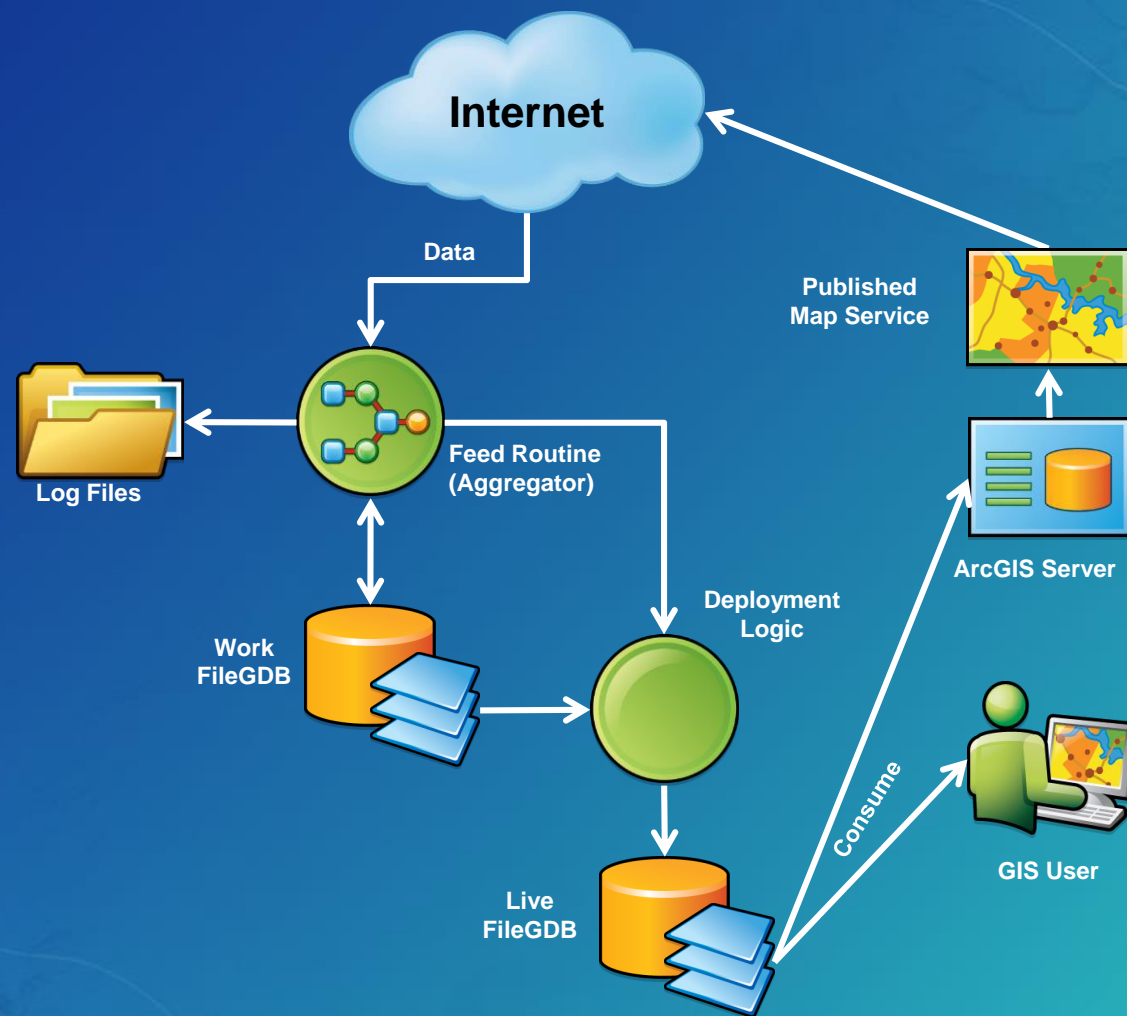
- What is it?
 - Processing workflow to handle near real-time data
 - Plus - Python production toolset
 - One more tool in your GIS toolbox
- Why would I use it?
 - Dealing with an unstable data source?
 - Need to reduce Internet load?
 - Internalize content!
 - Would you like to automate content refreshment?

ALF - Toolkit

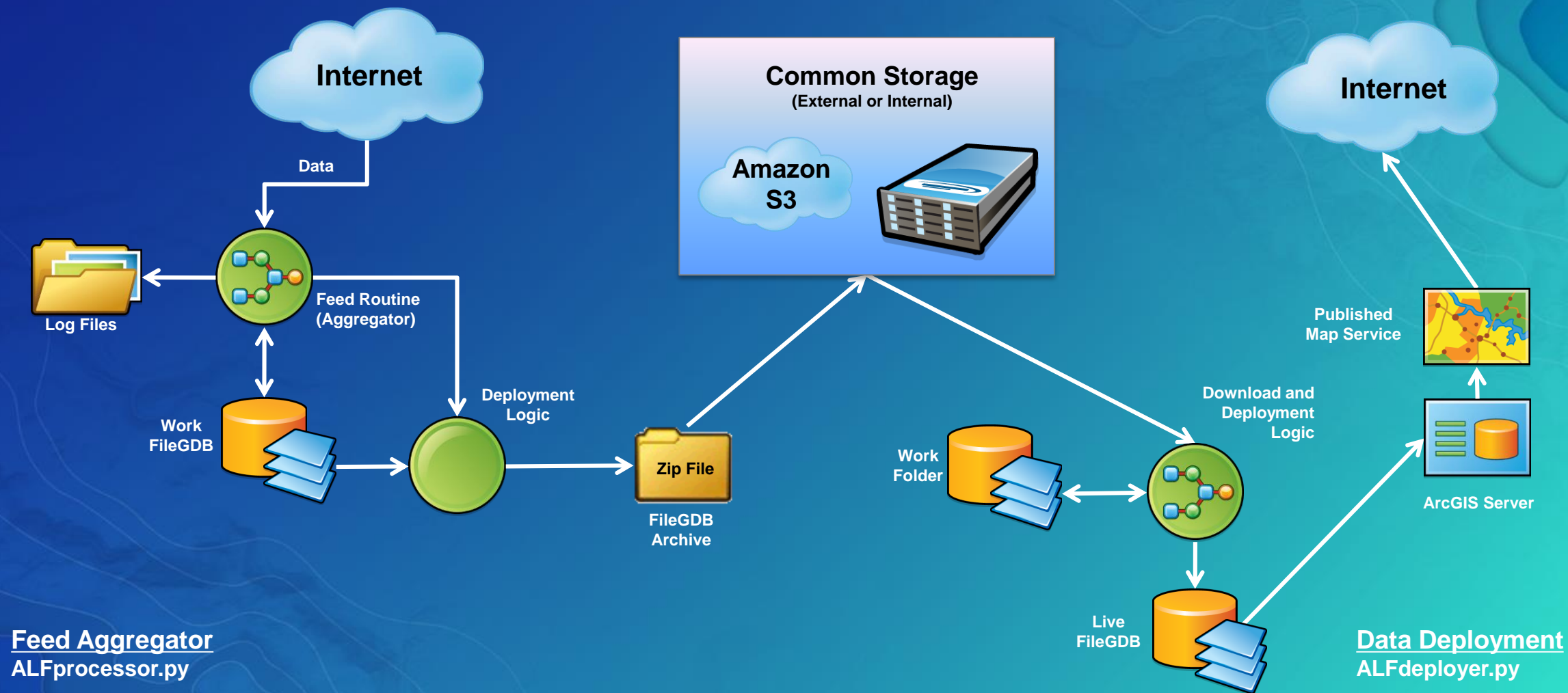
- **Benefits**

- **Multi-Processing support**
- **Process, Error, and Archive Logging (passive!)**
- **Integrated E-mail Altering / Status**
- **Production oriented**

ALF Methodology – Locally Consumed Feed



ALF Methodology – Cloud / Distributed Feed




```
SandyPrecip Routine Started: 11:42:05 Tue 03/19/2013
```

```
Importing ArcPy...
```

```
Processing: 'Sandy_Oct30_00.shp'
```

```
Preparing Workspace...
```

```
  Removing File Geodatabase...
```

```
  Creating File Geodatabase...
```

```
Importing data...
```

```
Creating Spatial Index...
```

```
Deploying...
```

```
  Gathering file list...
```

```
  Copying From:
```

```
    'C:\EPC_Sandy\Work\HurricaneSandy.gdb'
```

```
  to:
```

```
    'C:\EPC_Sandy\Live'...
```

```
Copy completed!
```

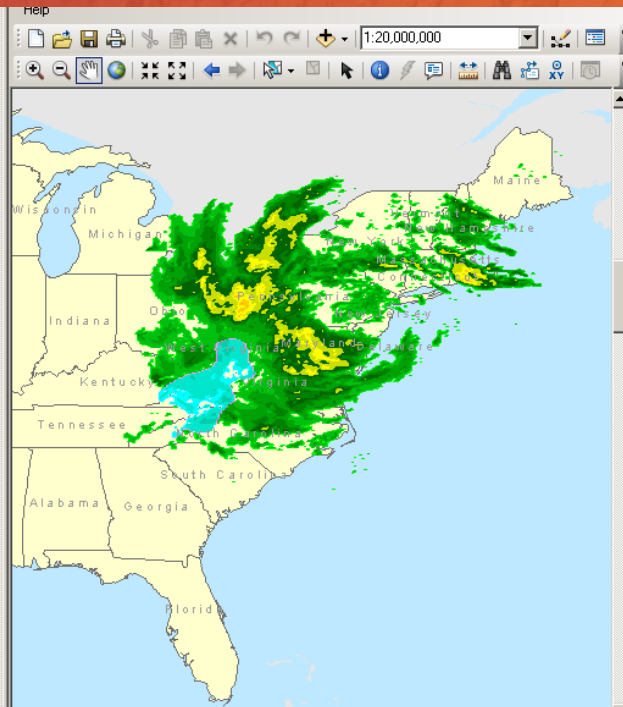
```
  Files Copied: 0
```

```
  Files Overwritten: 44
```

```
  Files Skipped: 0
```

```
Done!
```

```
C:\EPC_Sandy>
```



ALF Demo

```
SandyPrecip_LastRun - Notepad
File Edit Format View Help

SandyPrecip Routine Started: 12:07:08 Tue 03/19/2013
-----

Importing ArcPy...

    arcpy: v10.1.1
Configuration: SandyPrecip.cfg, v1.0.0
    Processor: ALFprocessor.py, v0.8.0
    Library: ALFlib.py, v1.6.0

Processing: 'Sandy_Oct30_00.shp'

Preparing workspace...
    <directory>: 'C:\EPC_Sandy\work'
    <gdbName>: 'HurricaneSandy.gdb'
    <recreate>: True
Removing File Geodatabase...
Creating File Geodatabase...
    Returned: C:\EPC_Sandy\work\HurricaneSandy.gdb

Importing data...

-----
Executing: CopyFeatures Sandy_Oct30_00.shp C:\EPC_Sandy\work\HurricaneSandy.gdb\CurrentPrecip # 0 0 0
Start Time: Tue Mar 19 12:07:16 2013
Succeeded at Tue Mar 19 12:07:17 2013 (Elapsed Time: 1.00 seconds)
-----

Creating spatial index...

-----
Executing: AddSpatialIndex C:\EPC_Sandy\work\HurricaneSandy.gdb\CurrentPrecip 5 0 0
Start Time: Tue Mar 19 12:07:17 2013
Succeeded at Tue Mar 19 12:07:17 2013 (Elapsed Time: 0.00 seconds)
-----

Deploying...
    Gathering file list...

    Copying From:
        'C:\EPC_Sandy\work\HurricaneSandy.gdb'
    to:
        'C:\EPC_Sandy\Live'...

    Updating: 'HurricaneSandy.gdb\000000001.gdbindexes'
    Updating: 'HurricaneSandy.gdb\000000001.gdbtable'
```

ALF Feed Construction – What do I need?

- **Pre-requisites**

- Python
- ArcPy
- Others (.py, 7zip, degrib, ...)

- **Optional**

- IDE – PyScripter, VS, ...

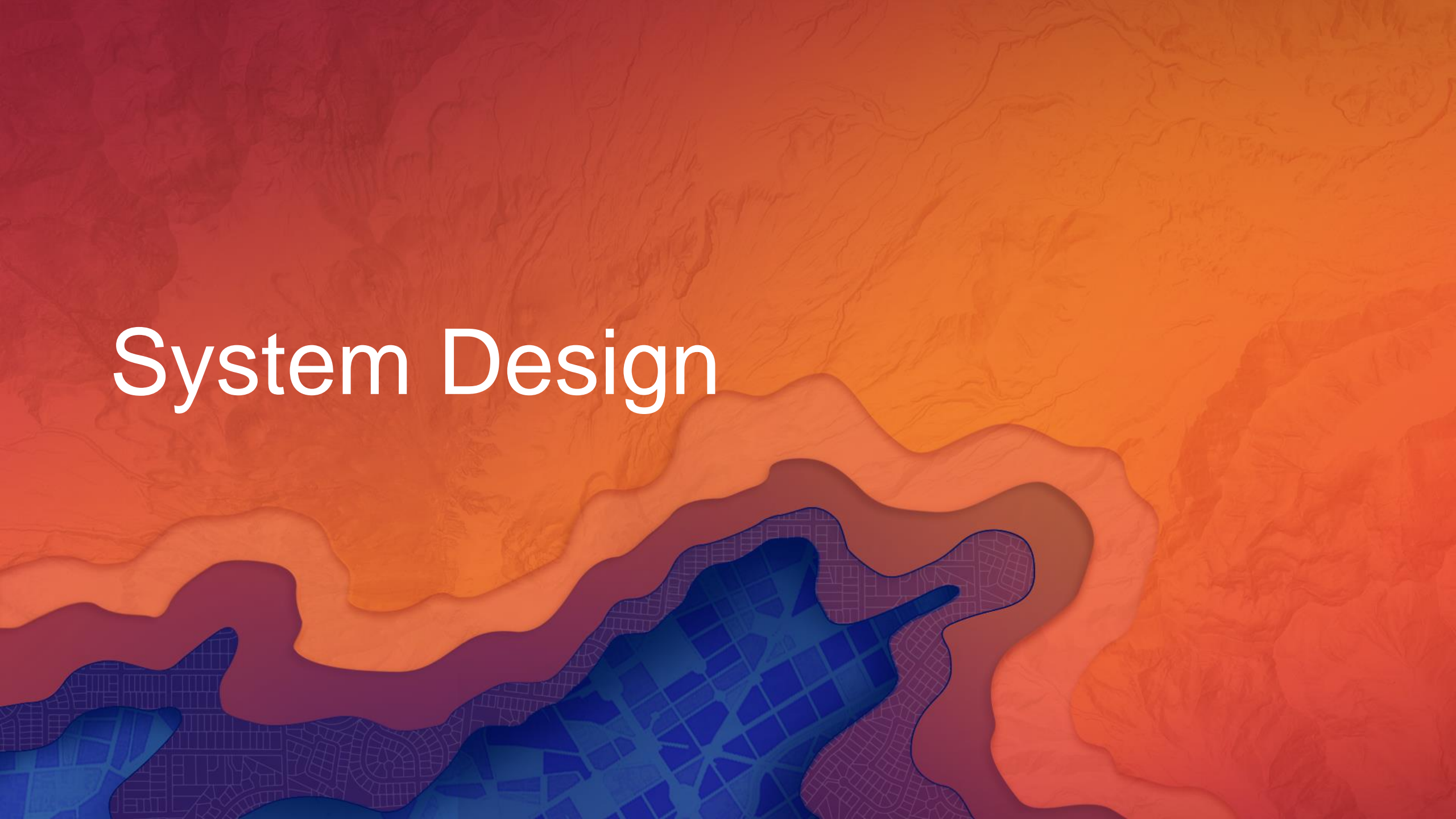
- **Components**

- ALFlib.py
 - Function / Class Library
- ALFprocessor.py
 - Feed Processor +
 - '<feed file>.cfg' (data processing logic)
- ALFdeployer.py
 - Archive Distribution

Resources

- AGOL Community Page (ALF samples, scripts) : <http://esriurl.com/LiveFeed>
- Live Data Feeds : <http://blogs.esri.com/esri/arcgis/2014/07/09/new-live-feeds-added/>
- Optimizing Rasters : <https://github.com/Esri/OptimizeRasters>
- Image Management Workflows : <http://esriurl.com/6550>
- Author : pdodd@esri.com

System Design



System architectures

High Availability, High Capacity



ArcGIS Service Criteria

- Service Design
- Min/Max number of instances
- Max Wait time
- Max Startup time
- Max Usage time
- Recycle
- Low/High Isolation
- Stop/Delete Unwanted Services

ArcGIS Online Hosted Model

- **When you have unforeseen demand that cannot be met by running your own infrastructure**
- **When you need to broaden reach of your content outside of the enterprise firewall**
- **When you want to use the latest capabilities of the platform**
- **No need to manage infrastructure, maintenance, monitoring, upgrades etc.**

ArcGIS Online Hosted Service

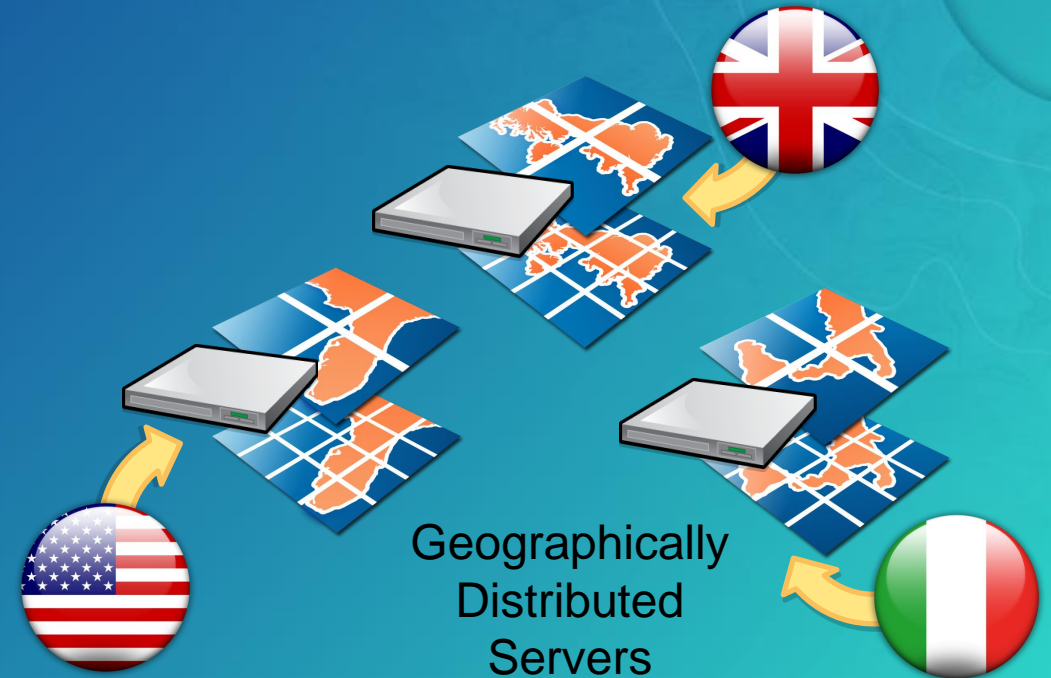
- Frequent update of data
- Can be updated using Desktop or Website
- Control Access for editing
- Export data
- Reference
 - <https://doc.arcgis.com/en/arcgis-online/reference/best-practices-layers.htm>
 - <https://blogs.esri.com/esri/arcgis/2016/11/01/69841/>

Geo Load Balancing

- Routes traffic based on location
- Used for:
 - Specialize content
 - Load balancing between data centers
 - Optimize network latency
- Amazon Route 53
- Akamai GTM

CDN [Content Delivery Network]

- Network of Geographically Dispersed Servers
- Edge Location
- Cache copies of content
- Close to end users
- Lowers latency



CDN / AWS CloudFront Applications

- **Static Content**
 - Media Content like audio, videos etc.
 - Software download
 - Web Site files like Images, CSS, JS
- **Dynamic Content**
 - News, Weather, Advertising etc.
- **Live Streaming**
 - Sporting events
 - Meetings, Gathering etc..

AWS CloudFront Components

- Distribution
 - Set of Rules
- Origin
 - Source of the content
- Behaviors
- Restrictions, Error pages
- WAF
- Edge Locations
- Price Classes

Origin Settings

Origin Domain Name

Origin Path

Origin ID

Origin Custom Headers

Header Name

Value

Default Cache Behavior Settings

Path Pattern

Default (*)

Viewer Protocol Policy

☒ HTTP and HTTPS

☐ Redirect HTTP to HTTPS

☐ HTTPS Only

Allowed HTTP Methods

☒ GET, HEAD

☐ GET, HEAD, OPTIONS

☐ GET, HEAD, OPTIONS, PUT, POST, PATCH, DELETE

Cached HTTP Methods

GET, HEAD (Cached by default)

Forward Headers

None (Improves Caching) ▼

Object Caching

☒ Use Origin Cache Headers

☐ Customize

[Learn More](#)

Minimum TTL

0

Maximum TTL

31536000

Default TTL

86400

Forward Cookies

None (Improves Caching) ▼

Query String Forwarding and Caching

None (Improves Caching) ▼

Smooth Streaming

☐ Yes

☒ No

Restrict Viewer Access (Use Signed URLs or Signed Cookies)

☐ Yes

☒ No

Compress Objects Automatically

☐ Yes

☒ No

[Learn More](#)

Lambda Function Associations

Event Type

Lambda Function ARN

Behaviors

- **Path Pattern Matching**
- **Origin Selection**
- **Headers**
- **Query Strings / Cookies**
- **Time to Live**
- **GZIP Compression**
- **Protocol Enforcement**
- **SSL certificates**

Distribution Settings

Price Class	Use All Edge Locations (Best Performance) ▼	?
AWS WAF Web ACL	None ▼	?
Alternate Domain Names (CNAMEs)	<input type="text"/>	?
SSL Certificate	<input checked="" type="radio"/> Default CloudFront Certificate (*.cloudfront.net) Choose this option if you want your users to use HTTPS or HTTP to access your content with the CloudFront domain name (such as https://d1111111abode9s.cloudfront.net/logo.jpg). Important: If you choose this option, CloudFront requires that browsers or devices support TLSv1 or later to access your content.	
	<input type="radio"/> Custom SSL Certificate (example.com): Choose this option if you want your users to access your content by using an alternate domain name, such as https://www.example.com/logo.jpg . You can use a certificate stored in AWS Certificate Manager (ACM) in the US East (N. Virginia) Region, or you can use a certificate stored in IAM.	
	No certificates available ▼	
	Request or Import a Certificate with ACM	
	Learn more about using custom SSL/TLS certificates with CloudFront. Learn more about using ACM.	
Supported HTTP Versions	<input checked="" type="radio"/> HTTP/2, HTTP/1.1, HTTP/1.0 <input type="radio"/> HTTP/1.1, HTTP/1.0	?
Default Root Object	<input type="text"/>	?
Logging	<input type="radio"/> On <input checked="" type="radio"/> Off	?
Bucket for Logs	<input type="text"/>	?
Log Prefix	<input type="text"/>	?
Cookie Logging	<input type="radio"/> On <input checked="" type="radio"/> Off	?
Enable IPv6	<input checked="" type="checkbox"/>	?
	Learn more	
Comment	<input type="text"/>	?
Distribution State	<input checked="" type="radio"/> Enabled <input type="radio"/> Disabled	?

Path Pattern Matching

- Route requests to specific origins
- Set Protocol
- Set Header, Caching options
- Set cookie and Query String Forwarding
- Restrict Access
- Set Compression

AWS CloudFront Benefits

- Scalable
- Secured
 - SSL/TLS delivery
 - DDos Protection , WAF (Web Application Firewall)
- Robust Real Time Reporting
- Cost effective
 - Pay as you go model

Akamai – Designating Origin Server

If

Hostname

is one of

server.arcgisonline.com

Behaviors

Add Behavior

Origin Server

Due to the security risk, it is **strongly** recommended that you consult with the security expert at your organization before configuring the Origin Server SSL Certificate Verification settings. If you *are* the security expert at your organization and you have any questions, contact your account team.

Origin Type

Your Origin

Origin Server Hostname
(variable support)

origin-services-geo.failover.arcgisonline.com.akadns.net

Forward Host Header

Incoming Host Header

Cache Key Hostname

Origin Hostname

Supports Gzip

Yes

Akamai – Pattern Matching for Edge Caching

Caching

Caching Option **No Store**

If

Path

matches one of

Behaviors Add Behavior

Caching

Caching Option

Force Revalidation of Stale Objects

Max-age

Cache Key Query Parameters

Behavior

Parameters

Exact Match

CloudFront Reporting: Access Logs

- **Log delivered to S3**
- **Permissions Controlled**
- **Scheduled delivery**
- **Metrics**
 - **Cache Statistics**
 - **Usage Charts**
 - **Popular Objects**
 - **Browser, OS, Devices, Locations, Referrers etc..**

Key Considerations (in conclusion)

- **Data Optimization**
- **Infrastructure Setup**
- **ArcGIS Server/Service Configuration**
- **Geographic Distribution**
 - **Multiple Data Centers**
 - **Geo Load Balancing**
 - **Content Delivery Network**

Please Take Our Survey on the Esri Events App!

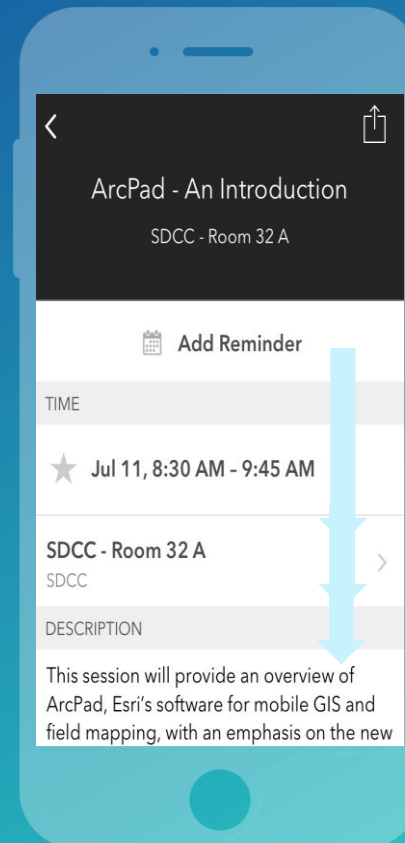
Download the Esri Events app and find your event



Select the session you attended



Scroll down to find the survey



Complete Answers and Select "Submit"





esri

THE
SCIENCE
OF
WHERE