Real-Time and Big Data GIS
The Road Ahead

Sarah Ambrose
Suzanne Foss
Josh Joyner
Agenda:

1. GeoEvent Server: Current Status
2. GeoEvent Server: Roadmap
3. GeoAnalytics Server
4. Embracing the Internet of Things
5. ArcGIS Analytics for IoT
GeoEvent Server Current Status
ArcGIS Enterprise with real-time capabilities

- Apps
- Desktop
- APIs
- ArcGIS Enterprise
- ArcGIS GeoEvent Server
- spatiotemporal big data store
- visualization
- ingestion
- actuation
- live features
- stream services
- live & historic aggregates & features
- analytics
- storage
ArcGIS Enterprise

with real-time capabilities

**MINIMUM environment**
3 machines

---

<table>
<thead>
<tr>
<th>ArcGIS GeoEvent Server</th>
<th>10.2</th>
<th>10.3</th>
<th>10.4</th>
<th>10.5</th>
<th>10.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velocity throughput</td>
<td>up to 500 e/s</td>
<td>up to 2,000 e/s</td>
<td>up to 3,000 e/s</td>
<td>up to 4,000 e/s</td>
<td>up to 6,000 e/s</td>
</tr>
</tbody>
</table>

*measured in events per second (e/s)*
**ArcGIS Enterprise**

*with real-time capabilities*

**OPTIMIZED environment**

for a resilient & scaled out deployment
7 machines

---

**IoT**

---

**GeoEvent Server**

**spatiotemporal big data store**

---

<table>
<thead>
<tr>
<th>ArcGIS GeoEvent Server</th>
<th>10.2</th>
<th>10.3</th>
<th>10.4</th>
<th>10.5</th>
<th>10.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velocity throughput</td>
<td>up to 500 e/s</td>
<td>up to 2,000 e/s</td>
<td>up to 3,000 e/s</td>
<td>up to 4,000 e/s</td>
<td>up to 6,000 e/s</td>
</tr>
<tr>
<td>Resiliency &amp; Scalability</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
GeoEvent Server

resiliency, scalability & performance

- ArcGIS 10.5
  - Resiliency (high availability) & scalability is only possible if users “bring their own gateway”
    - Barrier to entry is HIGH & typically requires a professional services engagement for success
    - Loses flexibility of input types

OPTIMIZED environment
for a resilient & scaled out deployment

10.5
GeoEvent Server
resiliency, scalability & performance

- ArcGIS 10.6
  - Provides users with a resilient & scalable Real-Time GIS deployment OUT-OF-THE-BOX
    - Introduces a gateway process that is automatically configured as part of GeoEvent Server installation
  - Provides flexibility for all input types
  - Increased event throughput performance

OPTIMIZED environment
for a resilient & scaled out deployment

OPTIMIZED environment
for a resilient & scaled out deployment

ArcGIS Enterprise

IoT

GeoEvent Server

visualize

spatiotemporal big data store

store

ingest, analyze

18K e/s

6K e/s
ArcGIS Enterprise

with real-time GIS capabilities

OPTIMIZED environment
for a resilient & scaled out deployment
7 machines

<table>
<thead>
<tr>
<th>ArcGIS GeoEvent Server</th>
<th>10.2</th>
<th>10.3</th>
<th>10.4</th>
<th>10.5</th>
<th>10.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velocity throughput</td>
<td>up to 500 e/s</td>
<td>up to 2,000 e/s</td>
<td>up to 3,000 e/s</td>
<td>up to 4,000 e/s</td>
<td>up to 6,000 e/s</td>
</tr>
<tr>
<td>Resiliency &amp; Scalability</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes up to 5 practically, so can scale to up to 30,000 e/s</td>
</tr>
</tbody>
</table>

Resiliency & Scalability via multi-machine site

ArcGIS Enterprise with real-time GIS capabilities

Big data store

IoT

GeoEvent Server

visualize

Ingest, analyze

spatiotemporal big data store

Velocity throughput measured in events per second (e/s)

Resiliency & Scalability via multi-machine site

OPTIMIZED environment for a resilient & scaled out deployment 7 machines
GeoEvent Server

**best practices tutorial for multi-machine site deployment**

GeoEvent Server

**10.6.1 key improvements**

- Significant Performance improvements over 10.5.x and 10.6
  - Configurable multi-threaded writing to the spatiotemporal big data store
  - Single machine peak throughput velocity up to 10,000 events per seconds (e/s)

- Multi-machine deployments for improved resiliency with linear scalability
  - Up to 5 machines practically supporting up to 50,000 e/s

- Integrated Public Key Infrastructure (PKI) Authentication (Defense / Intel)

- Improved User Interface for Processor configuration

- ArcGIS Monitor Support

- Stream Layer Support in Pro 2.2 (New render support for published Web Maps)
GeoEvent Server

**10.7 major epics**

- New Input and Output connectors coming with Support for Native Cloud Storage Types
  - Watch an Amazon S3 bucket for New Delimited Text / Parquet Files
  - Watch an Azure Blob for New Delimited Text / Parquet Files
  - Write to a Delimited Text / Parquet File in an Amazon S3 bucket
  - Write to a Delimited Text / Parquet File in an Azure Blob

- Automatic Configuration Backup to Azure Blob / Amazon S3 bucket

- Configurable Message level resilience for select inputs and outputs

- User Interface Improvements
  - Support for multi-machine control and statistics
  - Embedded help in Service Designer

- Additional Monitoring Support
Spatiotemporal big data store

10.7 major epics

• Enhance data retention policy with ability to automatically export prior to purge

• Enhance GeoEvent Manager and REST API with ability:
  - to export data on demand: including Delimited Text or Parquet to S3/Azure Blob
  - to import data on demand: including Delimited Text or Parquet from S3/Azure Blob

• Z Value Support for Polyline, Polygon, and Multi-Point Geometries

• Additional Symbology Support in ArcGIS Enterprise / Online JavaScript Web Maps

• Other incremental enhancements based on user feedback
3 GeoAnalytics Server
Tools for processing large vector and tabular data with both spatial (location) and temporal (time) components using distributed analytics and storage

- ArcGIS Enterprise server role released at 10.5
- Run analysis in Pro, Portal, or the ArcGIS Python API
- Distributes analysis over multiple machines for faster processing time
- Supports reading directly from common data sources
  - File shares, HDFS, Hive, Amazon S3, Azure Data Lake
- Most tools support spatiotemporal analysis
ArcGIS 10.7

Themes

- Enable users to run distributed machine learning analytics
- Support developers and data scientists that want to author python-based tools
- New GIS tools and enhancements to existing tools and user experiences
- Support for additional output data stores
ArcGIS 10.7

*Machine learning and Python*

- **Expand ML capabilities**
  - Find point clusters (HDBSCAN)
  - Geographically weighted regression (GWR)
  - Forest-based classification and regression, ordinary least squares (OLS)

- **Integration between GeoAnalytics and PySpark (python)**
  - Enable WebGIS datasets to work with Spark libraries (e.g., SparkML)
  - Expose spatiotemporal capabilities through python
ArcGIS 10.7

Developer API and Python integration

- Use the Python Script tool to run user defined Python on the GA Server within the PySpark framework
  - A user can submit a script that will run across the cluster

- Enable Hosted Notebooks to run on a GeoAnalytics Server
  - PySpark will be integrated to support execution across the cluster
ArcGIS 10.7

*New tools and enhancements*

- Implement new GIS functionality
  - Clip, Merge, Dissolve

- Continued work on interfaces and tool enhancements
• Users will have the ability to write the output of analysis to a big data file share
  - E.g., CSVs in HDFS, Parquet in S3, Azure Data Lake, Azure Blob Store
Embracing the Internet of Things (IoT)
ArcGIS & the IoT
enabling geospatial insights with your IoT

Weather
- warnings
- earthquakes
- precipitation
- icy conditions

Connected Cars
- autonomous driving
- traffic conditions
- holes
- parking meters
- road conditions
- slippery areas
- network improvements

Transit
- buses
- taxis
- rail
- trains
- crowds

People
- health monitoring
- social activity

Public Safety
- police
- fire
- surveillance

Public Health
- hospitals
- ambulances

Public Transit
- buses
- taxis
- rail
- trains
- crowds

Energy Usage
- electricity
- gas
- smart meters

City Workers
- sanitation
- snow plows

Telecommunications
- cell phone signals
- dropped calls

Environment
- noise
- gases
- atmospheric pressure
- pesticides
- rain gauges
- water quality

Internet of Things

ArcGIS Enterprise
with real-time & big data capabilities

Apps

Desktop

APIs

ArcGIS Enterprise
Blueprint for IoT solutions

- Ingestion
- Streaming Analytics & Policies
- Actions (including Actuation)
- Data Store
- Device Management

An IoT Platform & Enterprise consists of the following capabilities:
- Batch Analytics
- Management Console
- Visualization
- Dashboards
ArcGIS as an IoT Platform

enabling geospatial insights with your IoT solution

- An ArcGIS based IoT Platform & Enterprise consists of the following capabilities:
  - **Ingestion**: GeoEvent server input connectors
  - **Streaming Analytics & Policies**: GeoEvent Services
  - **Actions** (including Actuation): GeoEvent output connectors
  - **Data Store**: spatiotemporal big data store
  - **Device Management**: for those requiring this functionality another IoT platform can be complemented with ArcGIS.
  - **Batch Analytics**: GeoAnalytics Server
  - **Management Console**: Portal & GeoEvent Manager
  - **Visualization**: Map & Feature Services
  - **Dashboards**: Operations Dashboard, Insights, Story Maps

---

**Edge**

- Sensors
- Devices (or Things)
- Actuators

**IoT Platform**

- GeoEvent Server
- GeoAnalytics Server
- ArcGIS Enterprise
- spatiotemporal big data store
- analytics
- data store
- policies & orchestration
- visualization

**Enterprise**

- Operations Dashboard for ArcGIS
- Insights for ArcGIS
- Esri Story Maps
- ArcGIS Earth
- ArcGIS Online
- ArcGIS Pro
- Collector for ArcGIS
- Web AppBuilder for ArcGIS
- AppStudio for ArcGIS
- dashboards
Complementing an IoT platform with ArcGIS

enabling geospatial insights with your IoT solution

- The Edge of an IoT broadcasts into an IoT platform such as: Azure IoT, Amazon IoT, Cisco IoT, IBM Bluemix, ...
- The IoT platform integrates with ArcGIS to expand its capabilities with spatiotemporal analytics, visualization & dashboards
Complementing an IoT platform with ArcGIS
enabling geospatial insights with your IoT solution

- The Edge of an IoT broadcasts into an IoT platform such as: Azure IoT, Amazon IoT, Cisco IoT, IBM Bluemix, ...
- The IoT platform integrates with ArcGIS to expand its capabilities with spatiotemporal analytics, visualization & dashboards.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Edge</th>
<th>Enterprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Devices (or Things)</td>
<td>Gateways</td>
<td></td>
</tr>
<tr>
<td>Actuators</td>
<td>IoT Hub</td>
<td>Power BI</td>
</tr>
</tbody>
</table>

Spatiotemporal Capabilities via ArcGIS
- Ingestion & actions
- Management console
- Policies & orchestration
- Visualization
- Batch analytics
- Data store
- Streaming analytics
ArcGIS Enterprise

*with real-time & big data GIS capabilities*

**MINIMUM environment**

- 4 machines
ArcGIS Enterprise

*with real-time & big data GIS capabilities*

OPTIMIZED environment
for a resilient & scaled out deployment
10 machines
ArcGIS Enterprise

with real-time & big data GIS capabilities on Microsoft Azure
ArcGIS Enterprise

with real-time & big data GIS capabilities on Amazon EC2
5 ArcGIS Analytics for IoT
ArcGIS Analytics for IoT

*Real-time and big data GIS as a service*

- A new class of customer is demanding real-time & big data analytic capabilities as a service
- A new class of customer is demanding MASSIVE real-time & big data analytic capabilities

**NEEDED environment**

10s to 1,000s of machines

Road Ahead content is provided for informational purposes only and is subject to change
ArcGIS Enterprise
real-time & big data GIS on-premise

ArcGIS Enterprise
visualization

ArcGIS Enterprise with
real-time & big data capabilities

up to millions

Ingestion +
real-time analytics

big data analytics

GeoEvent
Server

up to 10s of
thousands e/s

GeoAnalytics
Server

up to 100s
of millions

spatiotemporal
big data store

up to 100s of millions

Road Ahead content is provided for informational purposes only and is subject to change
ArcGIS Online

real-time & big data GIS as a service

ArcGIS Online

up to billions

Ingestion + real-time analytics

big data analytics

visualization

microservices as containers

storage

ArcGIS Analytics for IoT

up to billions

up to millions e/s
ArcGIS Online

*with real-time & big data GIS as a service*

- Launched as an application on ArcGIS Online
ArcGIS Online

with real-time & big data GIS as a service

• “ArcGIS for IoT” application enables authoring new ArcGIS Online items including:
  - Feed
  - Big Data Analytic
  - Real-Time Analytic
Feeds

*Real-time, streaming data*

- Configure & share
- Visualize directly in web maps
Feeds

*Real-time, streaming data*

- Configure & share
- Visualize directly in web maps
- Discover & use
  - Living Atlas
  - ArcGIS Marketplace
  - User community
Real-Time Analytics

*Continuous analysis*

- Consume a Feed
- Analyze & detect patterns of interest
- Alerts & actuation
- Store to map/feature services
- Share as item
Big Data Analytics

*Batch analysis*

- Access big data stores
- Analyze patterns
- Assess trends
- Store results
- Share as item

- Schedule analysis
**ArcGIS Analytics for IoT**

**Spatiotemporal operators**

### Summarize Data
- Aggregate Points
- Join Features
- Reconstruct Tracks
- Summarize Attributes
- Summarize Nearby
- Summarize Within

### Use Proximity
- Create Buffers
- Create Service Areas
- Calculate Distance
- Snap to Network

### Analyze Patterns
- Calculate Density
- Calculate Journeys
- Create Space Time Cube
- Emerging Hot Spots
- Find Hot Spots
- Find Point Clusters

### Find Locations
- Find Similar Locations
- Detect Gaps
- Detect Idle Locations
- Detect Incidents

### Manage Data
- Calculate Fields
- Dissolve Boundaries
- Filter
- Map Fields
- Merge
- Overlay Layers
- Project
- Select Fields

### Enrich Data
- Lookup Geoname
- Calculate Motion Statistics
- Reverse Geocode

### Identify Relationships
- Geographically Weighted Regression
- Logistic Regression
- Ordinary Least Squares
- Principal Component Analysis
- Forest-based Classification
- Self Organizing Maps

---

*capabilities in italics slated for later release*
Street Lighting
Visualizing & analyzing street light status
Optimizing maintenance priority based on context
Real-Time Analytic

*receiving and updating streetlight status*

<table>
<thead>
<tr>
<th>Feed</th>
<th>Receive streetlight status from Azure IoT Hub</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select Fields</td>
<td>Select fields of interest (STLID, LIGHTING_STATUS, …)</td>
</tr>
<tr>
<td>Calculate Field</td>
<td>Calculate if status represents outage (…)</td>
</tr>
<tr>
<td>Output</td>
<td>Keep latest feature in streetlight_status feature service</td>
</tr>
</tbody>
</table>
### Big Data Analytic

*finding bus stops near streetlight outages*

<table>
<thead>
<tr>
<th><strong>Input</strong></th>
<th>Load streetlights where LIGHTING_STATUS = ‘Off’</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Join Features</strong></td>
<td>Join streetlights to bus stops (Near Geodesic, 50 m)</td>
</tr>
<tr>
<td><strong>Calculate Field</strong></td>
<td>Calculate WEIGHT field (…)</td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>Keep latest feature in streetlight_priority feature service</td>
</tr>
</tbody>
</table>
Big Data Analytic

**analyzing businesses near streetlight outages**

**Input**
Load streetlights where LIGHTING_STATUS = ‘Off’

**Join Features**
Join streetlights to *businesses* (Near Geodesic, 50 m)

**Calculate Field**
Calculate WEIGHT field (…)

**Output**
Keep latest feature in streetlight_priority feature service
Big Data Analytic

analyzing human movement near streetlight outages

**Input**
Load streetlights where LIGHTING_STATUS = ‘Off’

**Join Features**
Join streetlights to summarized human movement data (Safegraph)

**Calculate Field**
Calculate WEIGHT field (…)

**Output**
Keep latest feature in streetlight_priority feature service
Street Lighting

Visualizing & analyzing street light status

Optimizing maintenance priority based on context
ArcGIS Analytics for IoT

*summary of planned real-time & big data GIS as a service capabilities*

- **Feeds**
  - Configure a real-time data stream & share it
  - Add a Feed to a web map and immediately visualize as live layer
  - Find & use publicly shared Feeds, *e.g.* Waze, Weather, Transit, …

- **Analytics**
  - Real-Time Analytics
    - Work on a Feed, perform continuous analysis
    - Alert on patterns of interest, actuate device behavior, store observations
  - Big Data Analytics
    - Work on data in ArcGIS, a big data cloud share, etc
    - Understand patterns and trends over time
    - Schedule to run on a recurring basis, *e.g.* once an hour, once a day, every 5 minutes
  - Authoring an Analytic is a simple guided user experience
  - Analytic results are shared *as Stream Services, Feature Services and/or Map Services*
Wrap Up
Real-Time GIS: Road Ahead

- Real-Time & Big Data GIS:
  - Enables real-time data to be ingested, analyzed, stored & visualized within ArcGIS
  - At 10.6, supports reliability & scalability out-of-the-box via deployment of a multi-machine site
  - At 10.7, support for native cloud storage for input/output and backup

- You can embrace the Internet of Things:
  - Today, by deploying a reliable & scalable on-premise real-time & big data GIS
  - Today, by complementing your ArcGIS deployment with popular IoT cloud platforms
  - In the future, as a service on ArcGIS Online
Help us improve the Real-Time & Big Data GIS Capabilities

http://esriurl.com/RealTimeSurvey
Questions / Feedback?

Suzanne Foss
ArcGIS Analytics for IoT
Product Manager
sfoss@esri.com

Josh Joyner
ArcGIS GeoEvent Server
Product Manager
jjoyner@esri.com

Sarah Ambrose
GeoAnalytics Server
Product Engineer
sambrose@esri.com