GeoEvent Server:

Best Practices

Suzanne Foss
Real-Time & Big Data Product Engineer, Esri
sfoss@esri.com
@sfoss_esri

Sagar Ayare
Real-Time & Big Data Product Engineer, Esri
sayare@esri.com
@theTechieSagar

Agenda

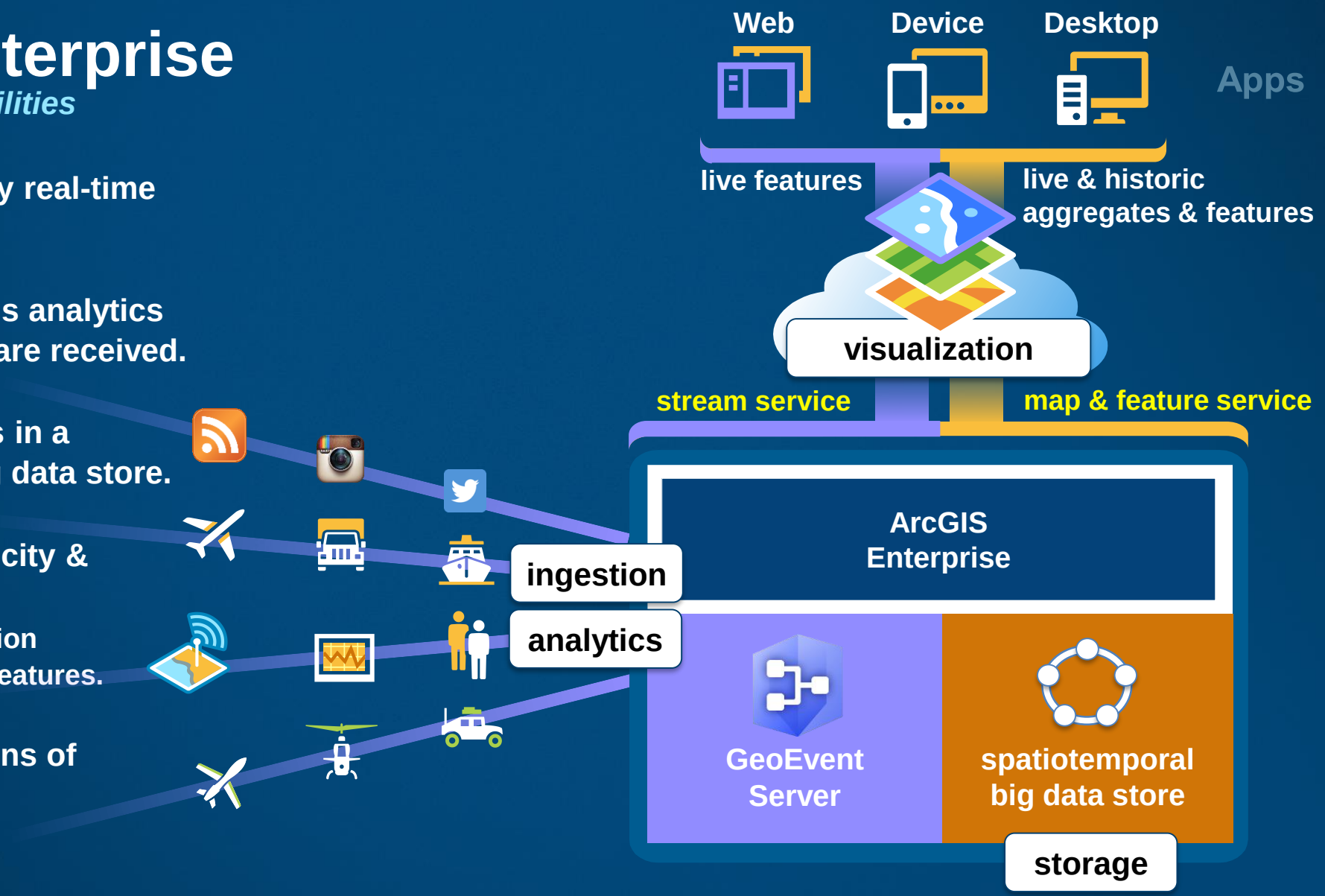
1. Ingestion & analytics
2. Dissemination & visualization
3. Performance
4. Resilience & scalability
5. Storage


1
Ingestion & analytics

ArcGIS Enterprise

with real-time capabilities

- Ingest high velocity real-time data into ArcGIS.
- Perform continuous analytics on events as they are received.
- Store observations in a spatiotemporal big data store.
- Visualize high velocity & volume data:
  - as an aggregation
  - or as discrete features.
- Notify about patterns of interest.



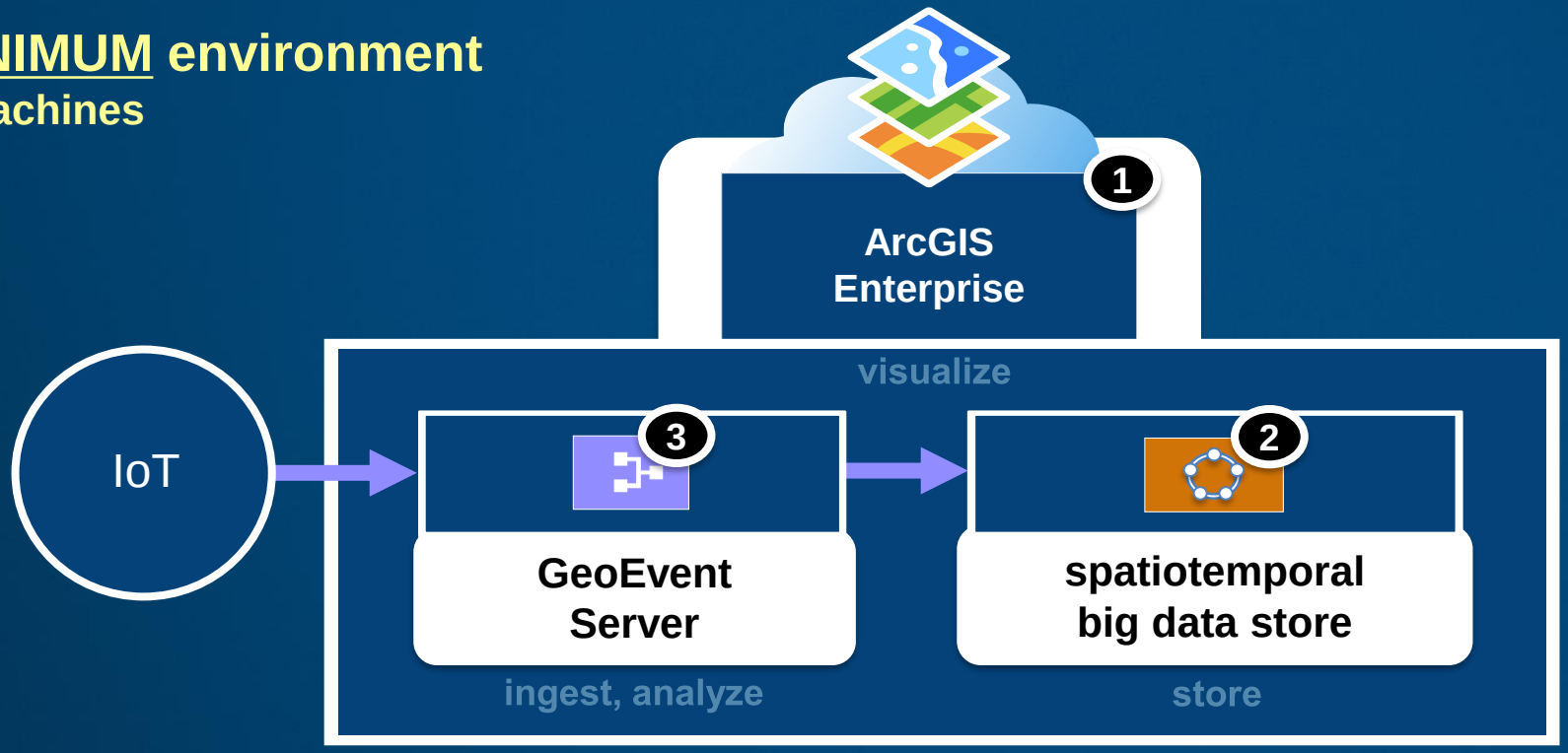
ArcGIS Enterprise

with real-time capabilities

10.5

MINIMUM environment

3 machines



functional servers & spatiotemporal big data store

SHOULD BE on ISOLATED machines!!!

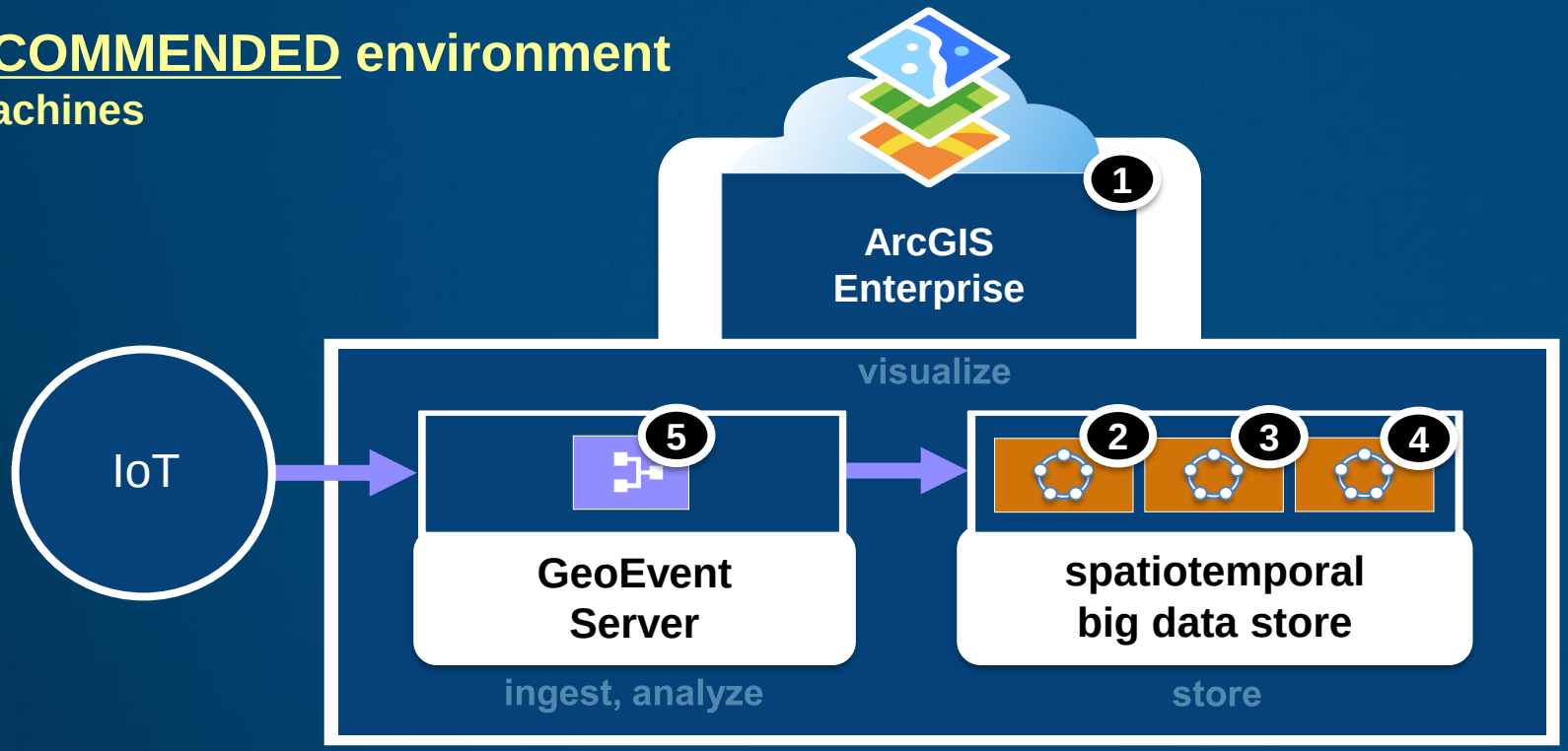
ArcGIS Enterprise

with real-time capabilities

10.5

RECOMMENDED environment

5 machines

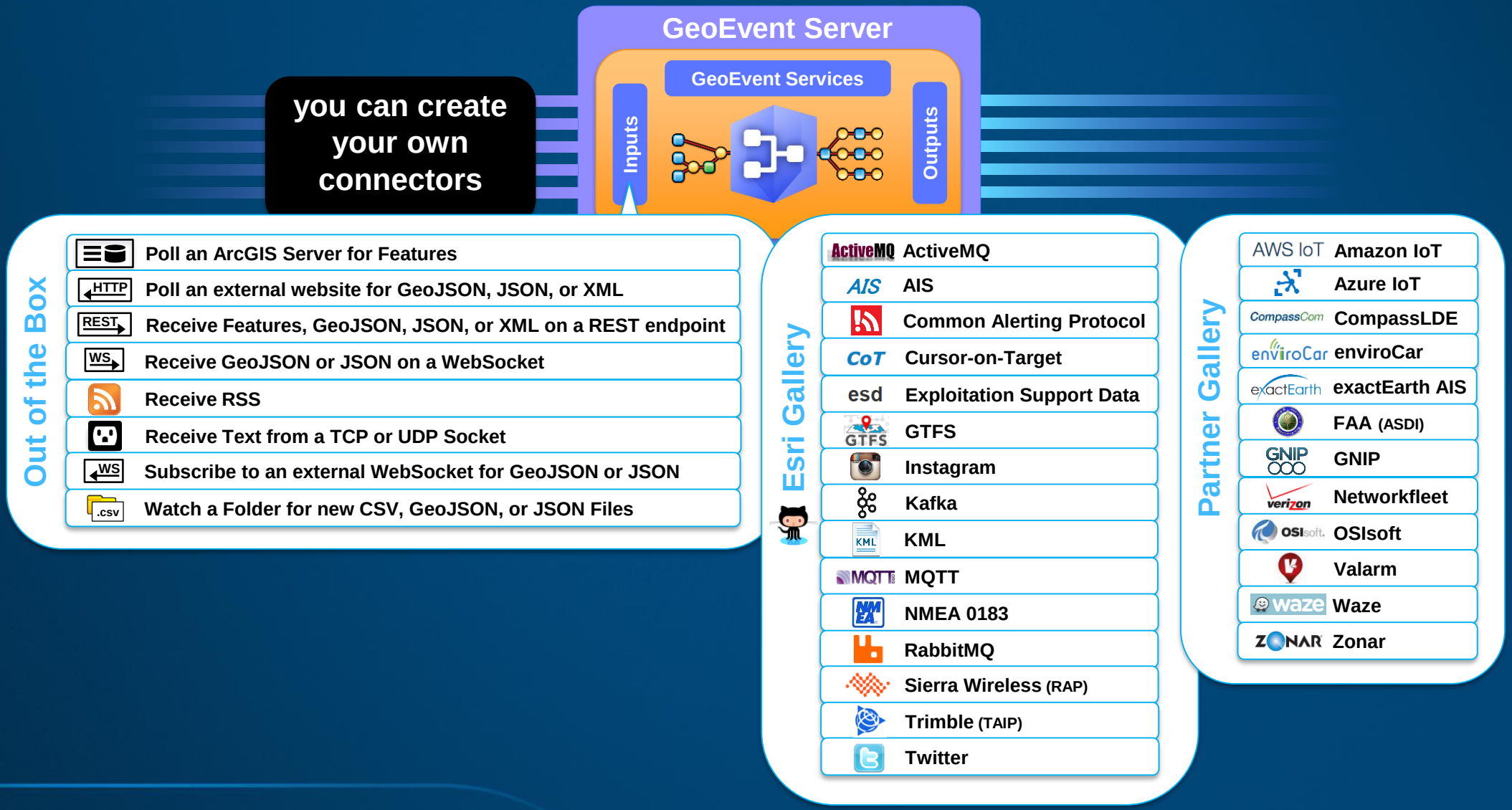


functional servers & spatiotemporal big data store

SHOULD BE on ISOLATED machines!!!

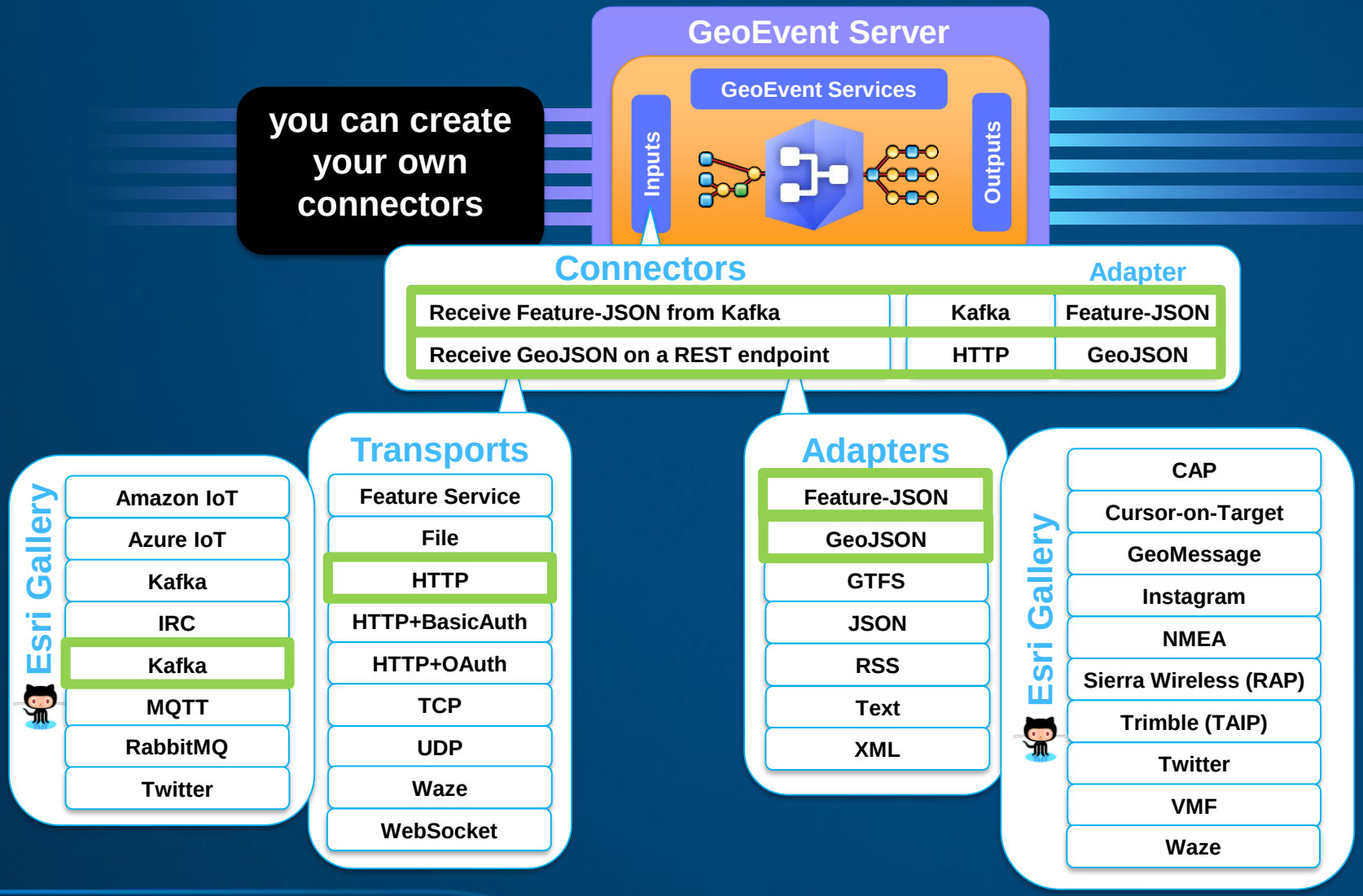
Ingestion of real-time data

input connectors



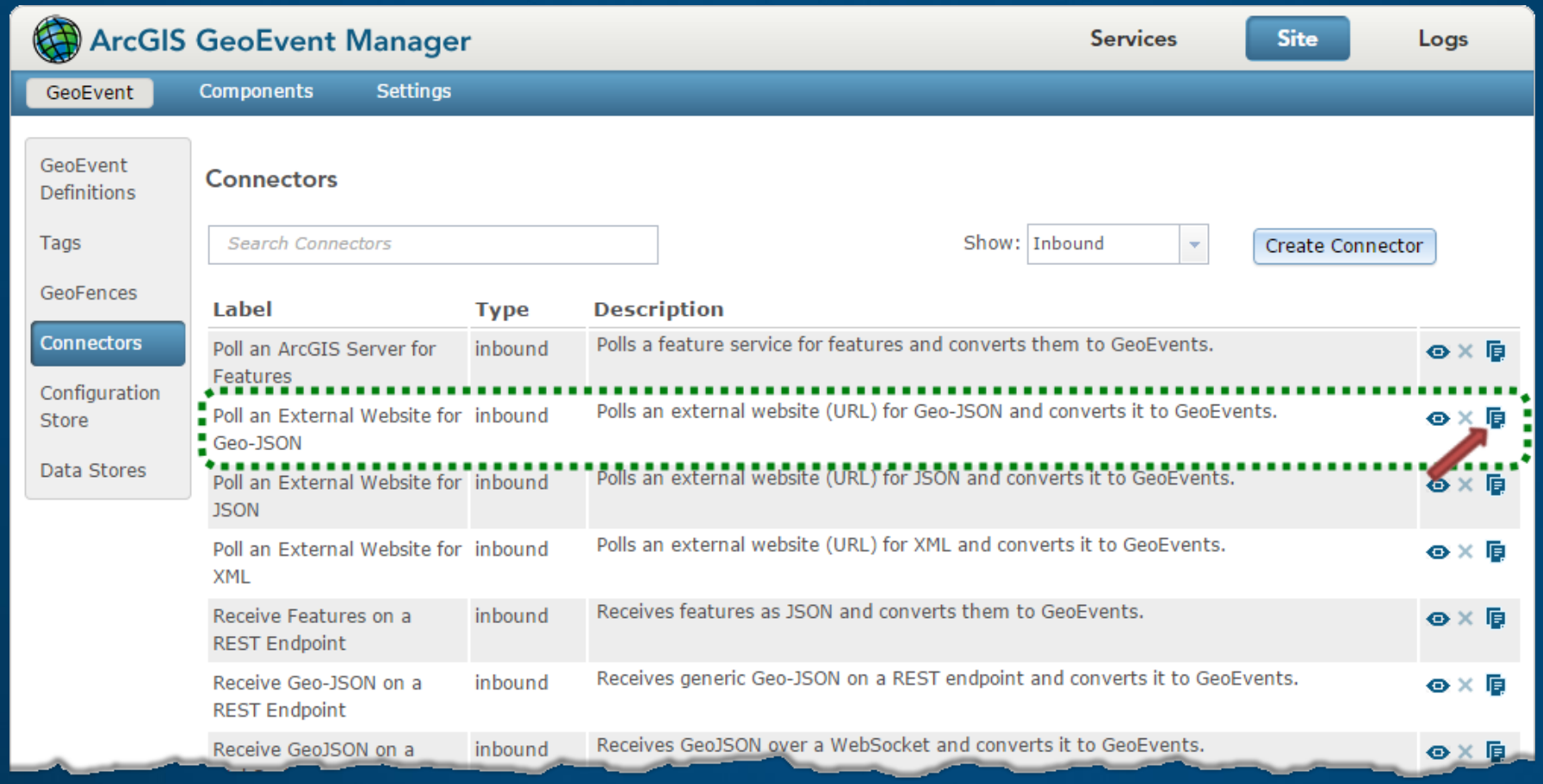
Ingestion of real-time data

configure a new input connector by pairing a transport & adapter together



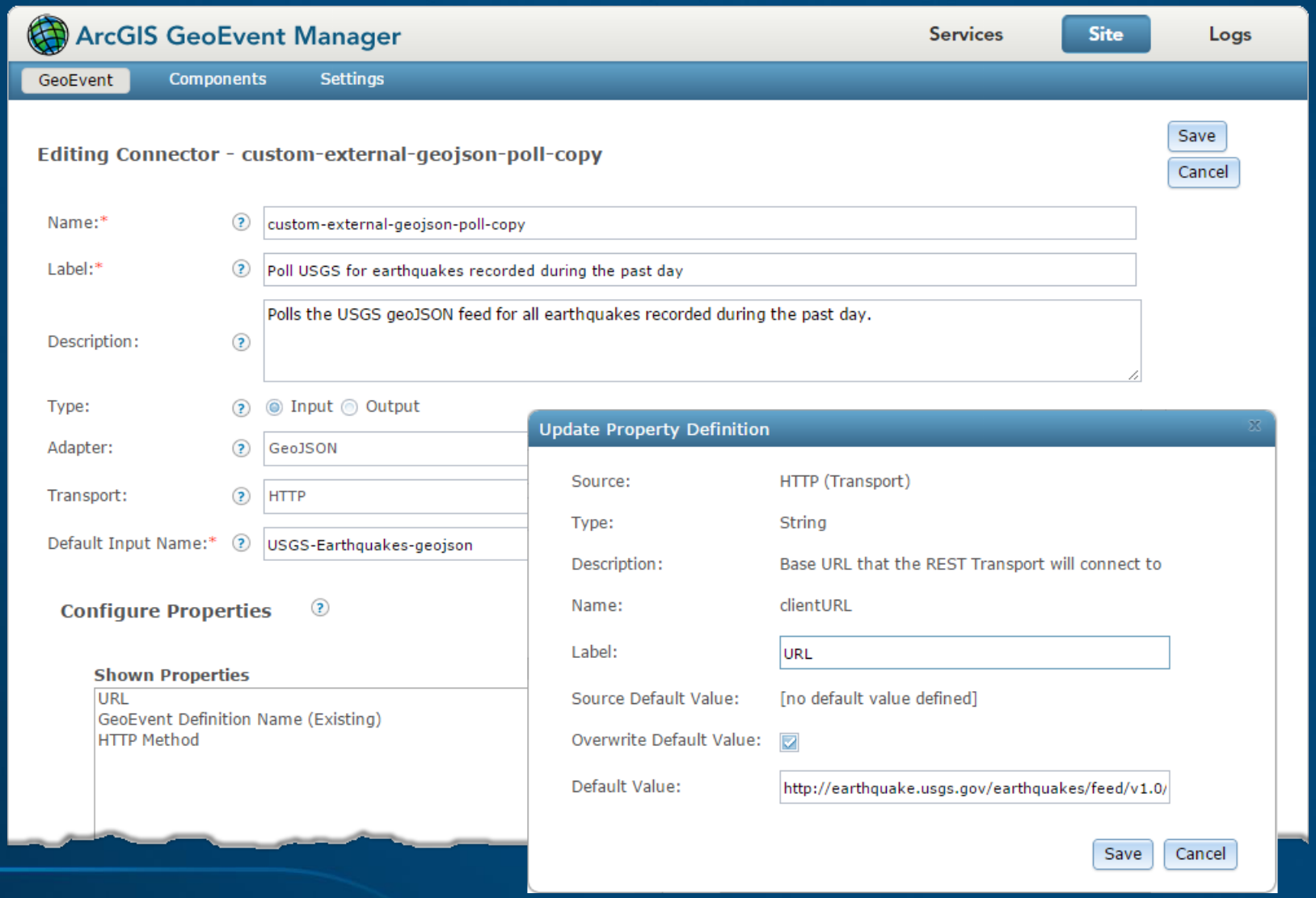
Ingestion of real-time data

clone an existing connector & tailor it's properties to make a more specific connector



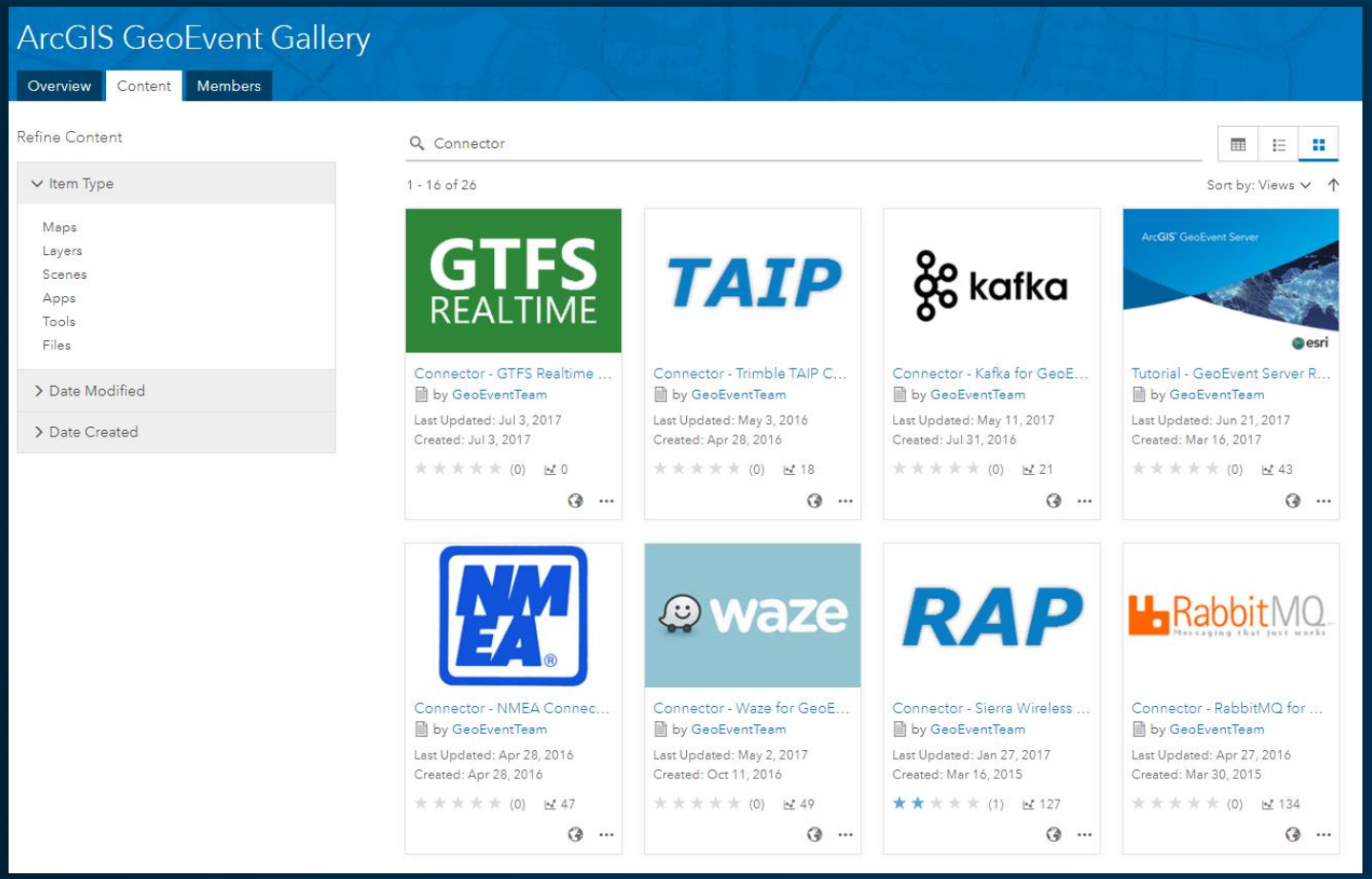
Ingestion of real-time data

clone an existing connector & tailor it's properties to make a more specific connector



Ingestion of real-time data

extend GeoEvent by downloading additional transports & adapters



solutions-geoevent-java

The solutions-geoevent-java repository includes custom connectors for use with ArcGIS GeoEvent Extension for Server.

Features

Adapters

- Common Alert Protocol (CAP) Adapter
- CoT Adapter
- Exploitation Support Data (ESD) Adapter
- Geomessage Adapter
- regex Text Adapter

Processors

- Add XYZ Processor
- Bearing Processor
- Buffer Processor
- Ellipse Processor
- Event Volume Control Processor
- Event Joiner Processor
- Extract Point from Line
- Extract Point from Polygon

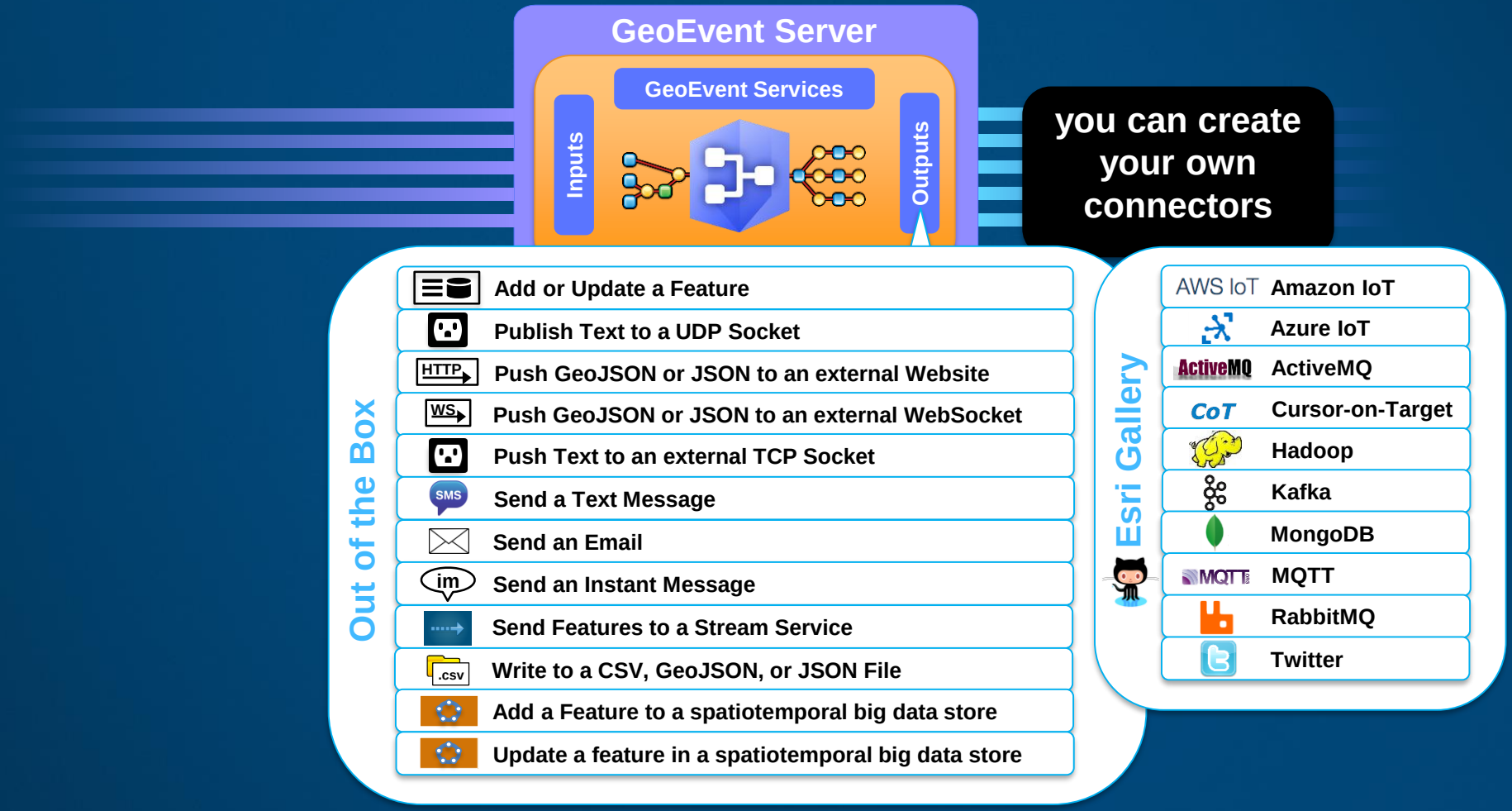
http://links.esri.com/geoevent-gallery

https://github.com/esri/solutions-geoevent-java

2 Dissemination & visualization

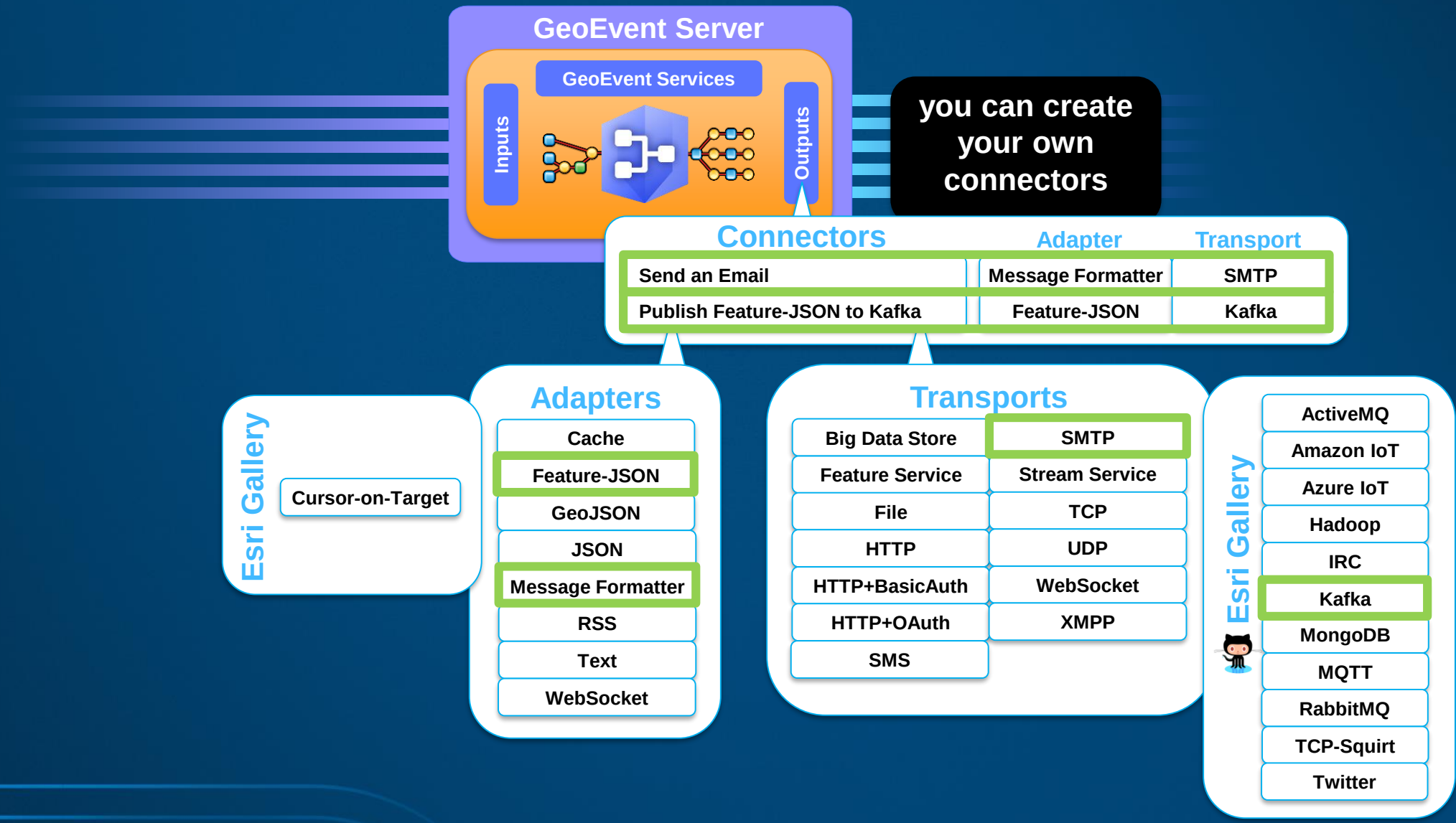
Dissemination of real-time data

output connectors



Dissemination of real-time data

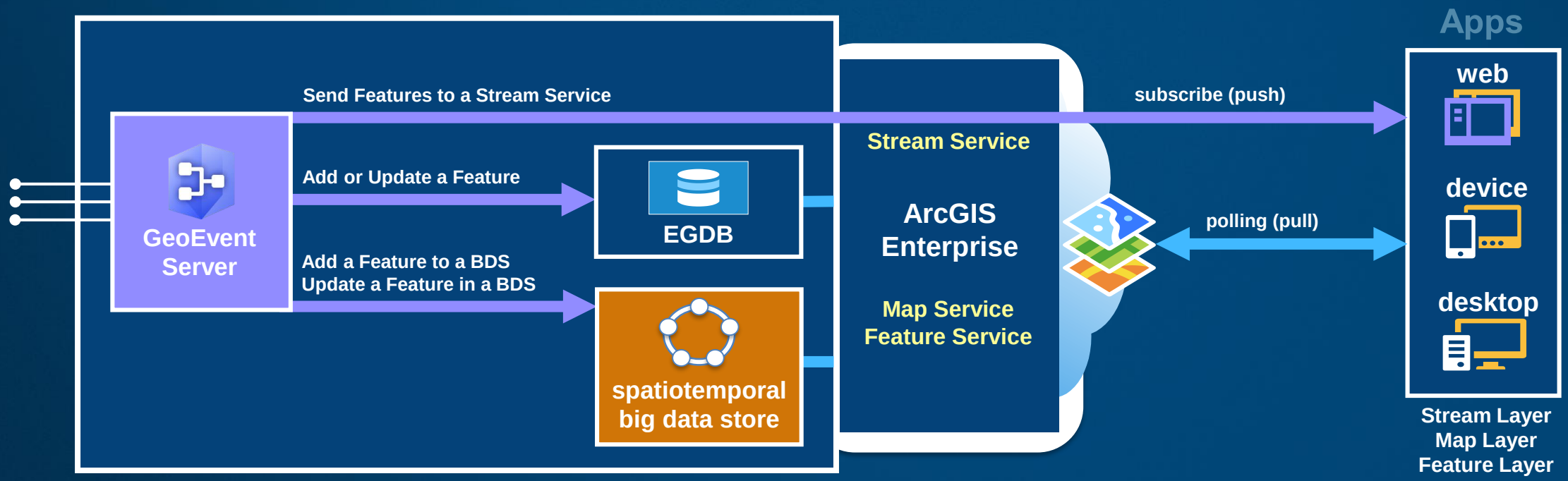
configure a new output connector by pairing an adapter & transport together



Visualization of real-time data

choosing a service type: stream service, feature service, map service

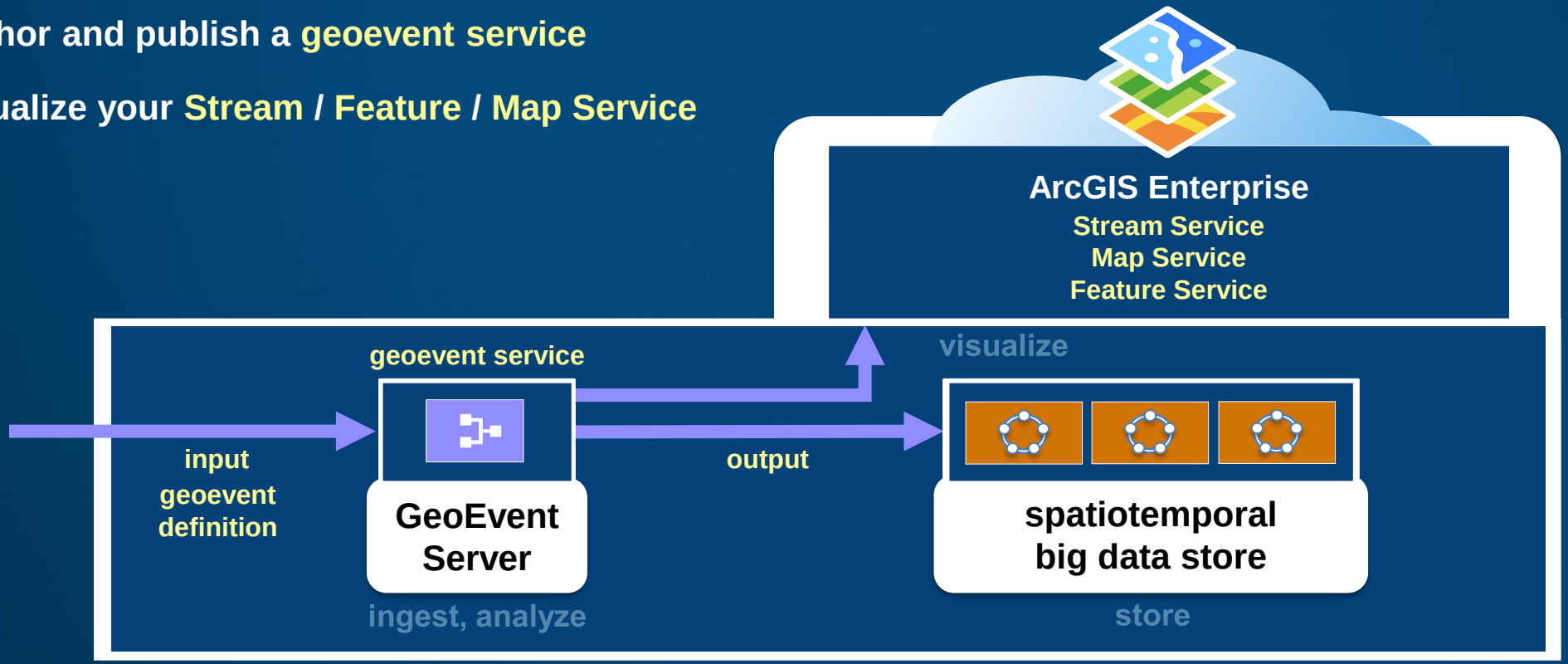
- Stream layers in apps **subscribe** to stream services to immediately visualize observations
  - does not require storage, low latency, no playback
- Map & Features layers in apps periodically **poll** to visualize most current observations
  - backed by an enterprise geodatabase (EGDB) or a spatiotemporal big data store (BDS)
  - history can be retrieved & queried for playback



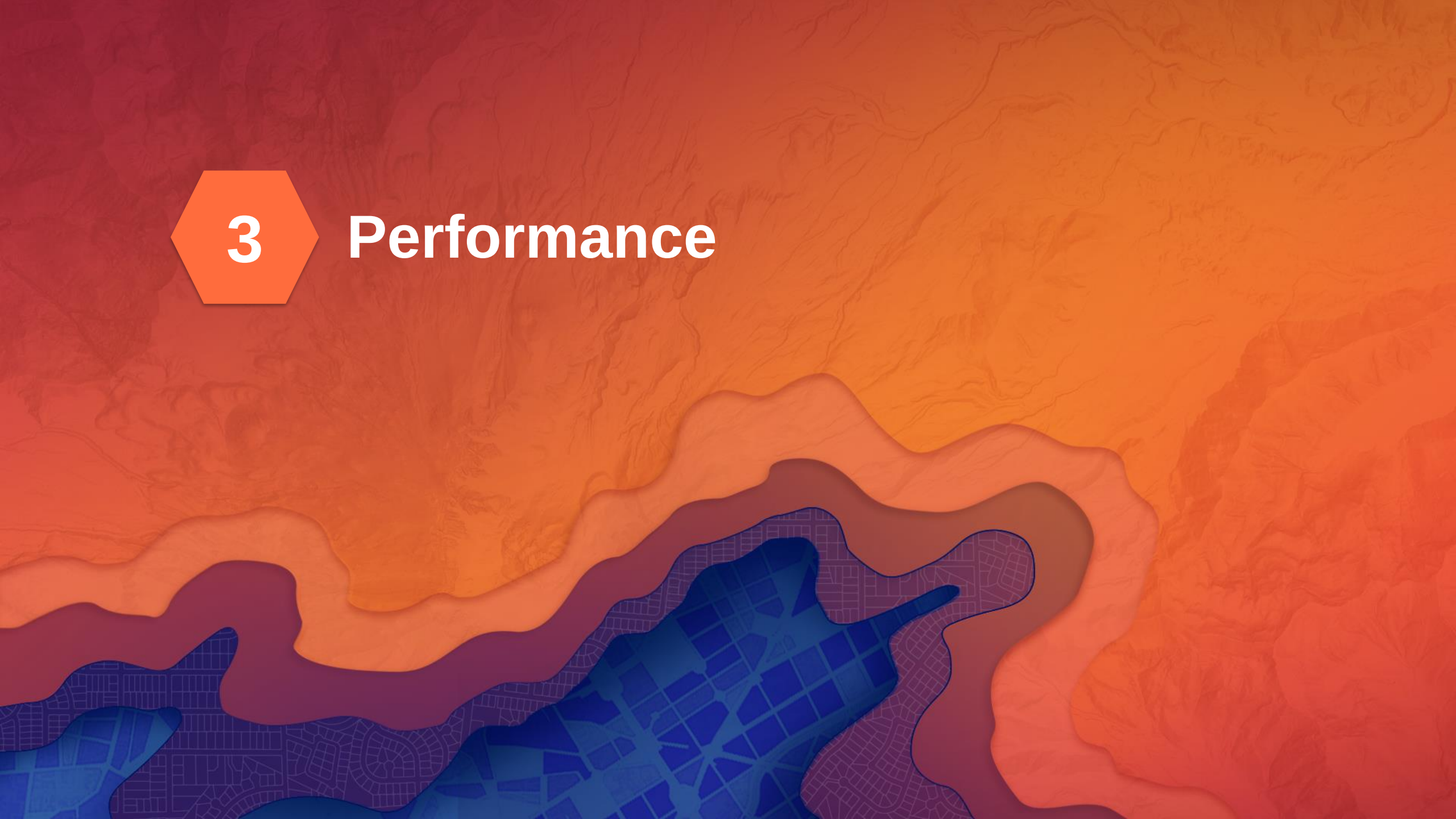
Visualization of real-time data

workflow for creating a real-time service using stream service, feature service, map service

- Configure an input to receive real-time data & define a geoevent definition accordingly
- Create an output and publish a Stream / Feature / Map Service using the geoevent definition
- Author and publish a geoevent service
- Visualize your Stream / Feature / Map Service



Demo


3
Performance

Performance

throughput: benchmarks

- GeoEvent Server throughput benchmarks by release:
  - ArcGIS 10.2: 600 e/s
  - ArcGIS 10.3: 3,000 e/s
  - ArcGIS 10.4: 4,000 e/s
  - ArcGIS 10.5: 4,000 e/s

- Benchmark environment consisted of:
  - A physical machine running ArcGIS Server + GeoEvent Extension
    - Windows 7 (64-bit)
    - 3.70 GHz, 4 physical cores (8 virtual cores)
    - 16 GB RAM
    - 1 Gbps network

Performance

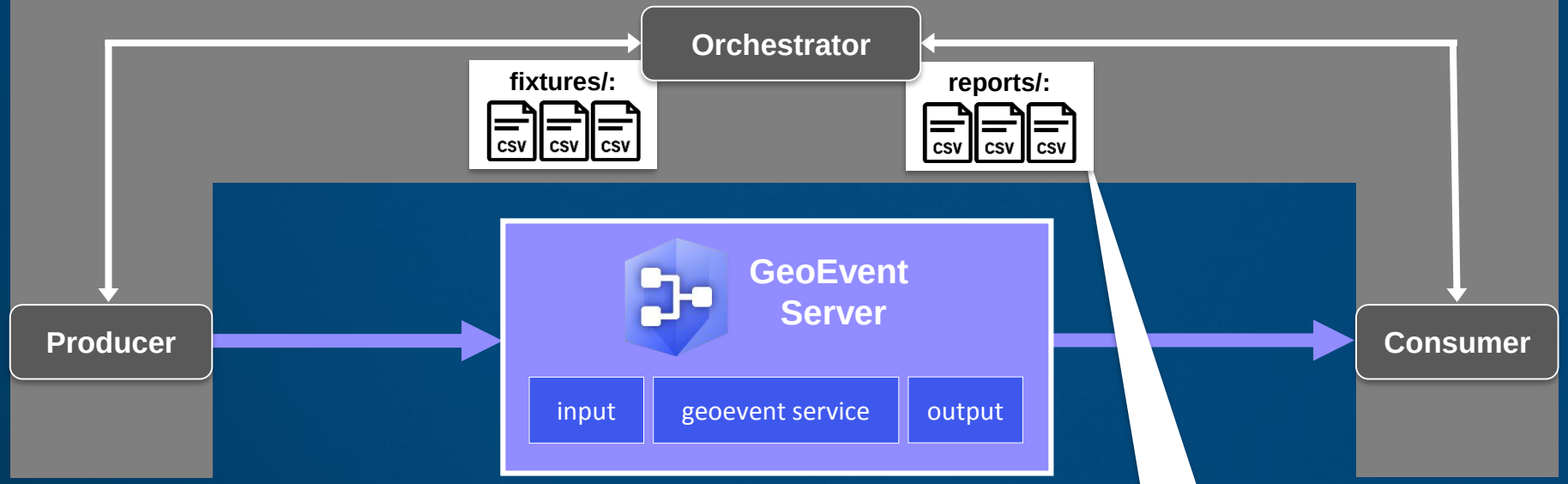
throughput: performance test harness for geoevent

Test Harness Orchestrator:
Runs a series of fixtures that specify:
- the GeoEvent host(s), input & output
- timed tests with a rate and stagger distribution
  - e.g. 100 e/s for 60 seconds with 10 staggers
- and where to record the resulting reports

https://github.com/Esri/performance-test-harness-for-geoevent

Esri / **performance-test-harness-for-geoevent**

Benchmark Metrics	
production time	= 3 - 1
consumption time	= 4 - 2
total time	= 4 - 1
latency of first receipt	= 2 - 1
latency of last receipt	= 4 - 3

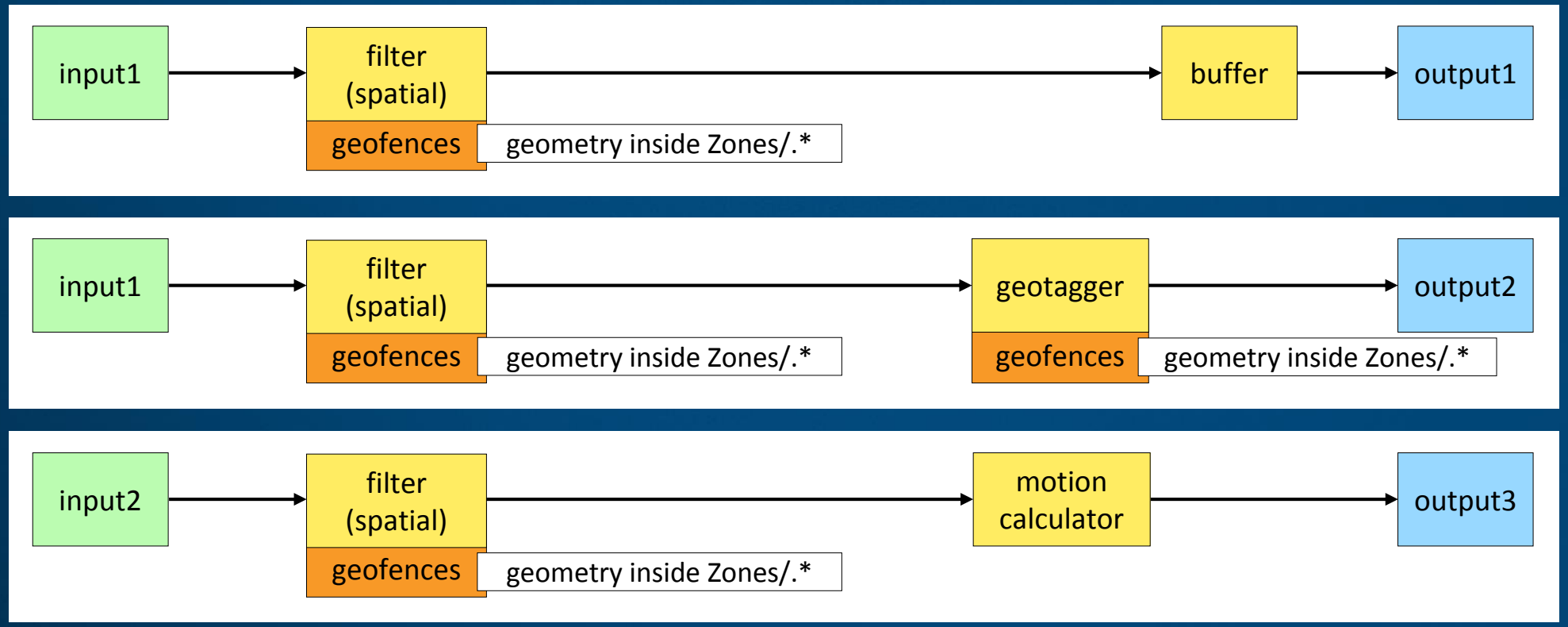


Test Name	Rate	totalEvents	successes	expectedResultCount	failures	avgTotalTime	avgEventsPerSec	%
1000 e/s	1000	30000	30000	30000	0	30278	990.818416	0.990818416
2000 e/s	2000	60000	59996	60000	0	30189	1987.346384	0.993673192
3000 e/s	3000	90000	89999	90000	0	30194	2980.691528	0.993563843
Total Testing Time:	01:55.666							

Performance

analytics: geoevent services

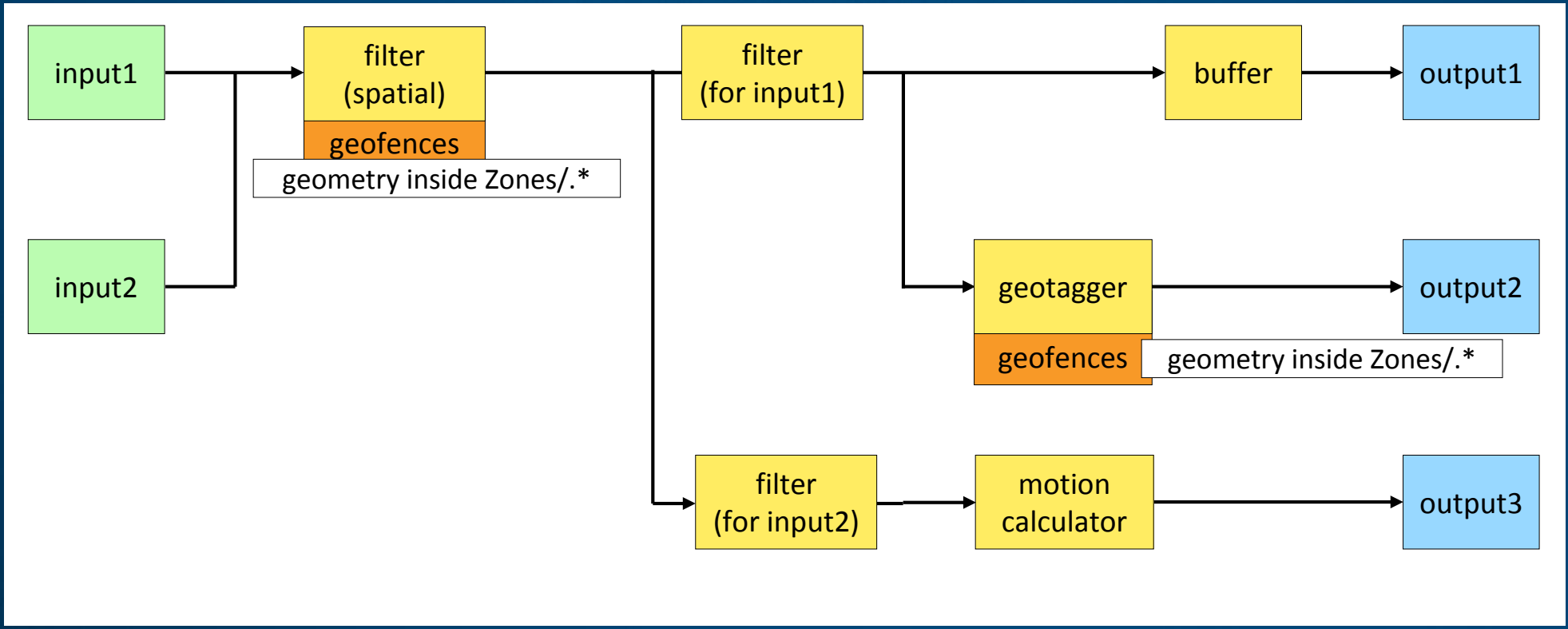
- **Which is better?**
  - **have fewer but more complex GeoEvent Services**
  - **or have more but simpler GeoEvent Services**



Performance

analytics: geoevent services

- **In general, it is better to have fewer and more complex GeoEvent Services**
  - **especially when working with Geofences**
  - **or Processors that change the schema of an event:** e.g. Field Enricher, Geotagger, Motion Calculator, …



Performance

throughput & analytics: primary factors to consider

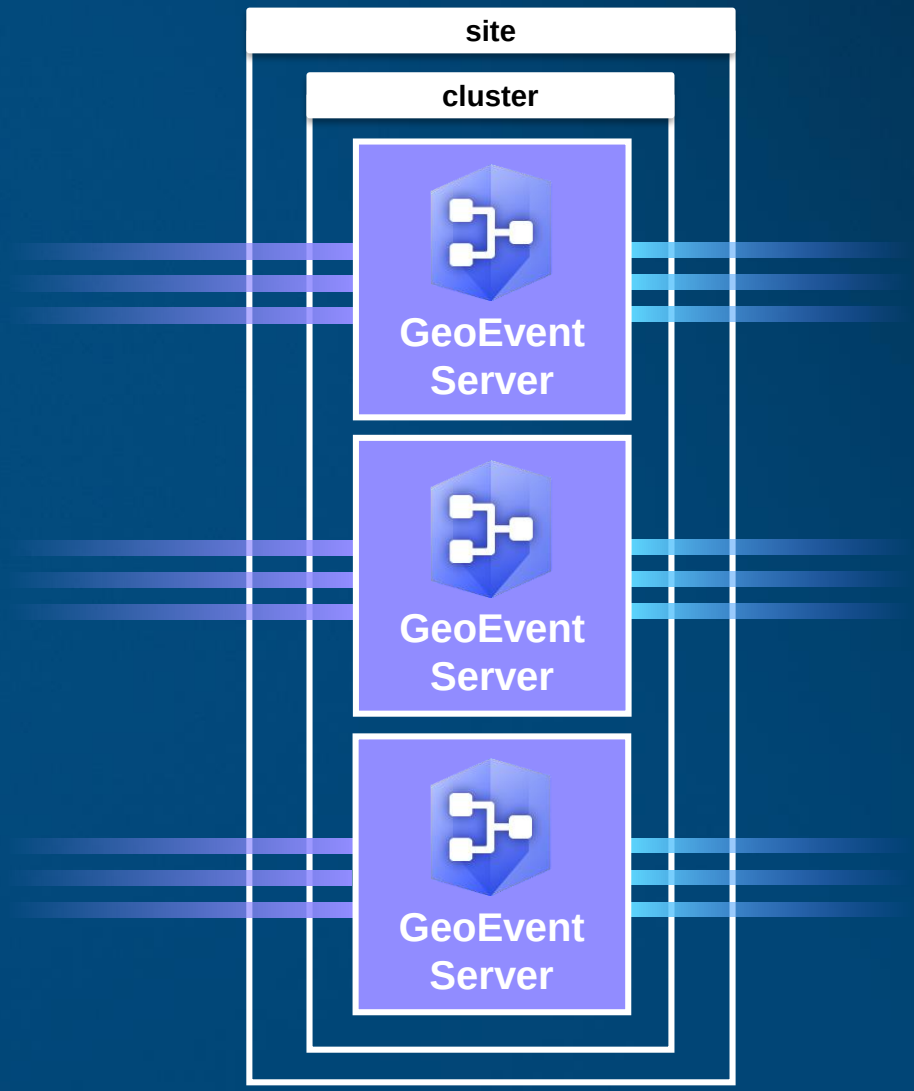
- **Operating environment:** *Bare Metal = Windows 7 (64 bit) Enterprise*
  - **Virtual Machines – beware! resources need to be shared in an effective way, like EC2 or Azure.**
  - **Bare-Metal machines – dedicated resources and public cloud instances are much more deterministic.**
- **Network**
  - **Speed (Gbps) – the faster the better.** *1 Gbps*
- **RAM**
  - **size (GB) – 8GB has been required since 10.3.** *16GB, default JVM max heap size is 4 GB*
  - **type – minimum of DDR3 is recommended.** *DDR3*
  - **clock speed (MHz) and transfer rate (Mbps) – the faster the better.**
- **Processors**
  - **speed (GHz) – the faster the better.** *3.70GHz, Intel Xeon E5-1620 v2*
  - **# of cores – the more the better.** *4 physical (8 Virtual)*

4 Resilience & scalability

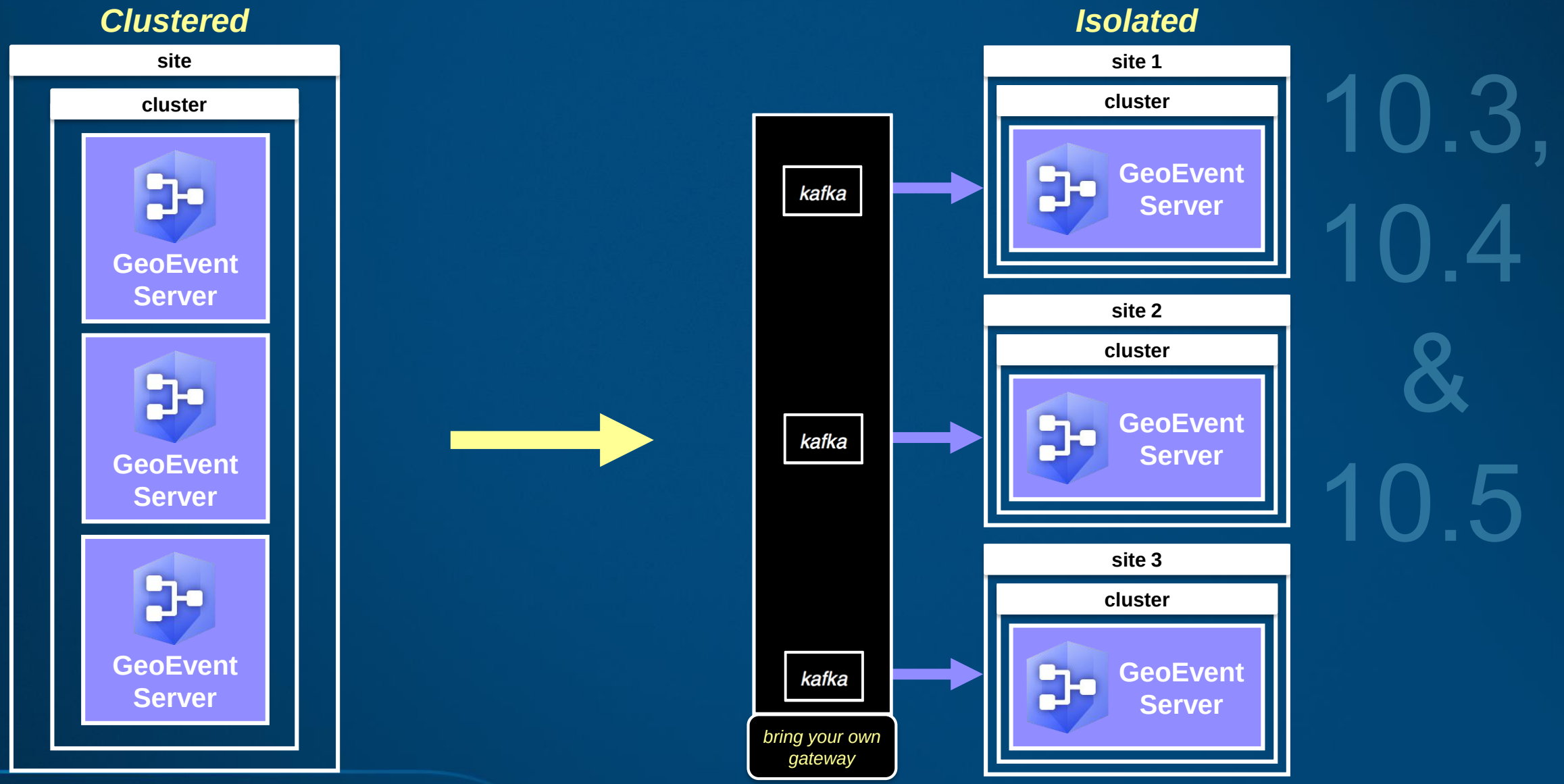
GeoEvent Server resilience & scalability

clustering is DISCOURAGED

- **Clustering was introduced at ArcGIS 10.3.**
- **However, observed to be brittle & subject to failure:**
  - **Required at least a 3-node cluster:**
    - When a node goes down, it had to be brought back up ASAP
    - as a 2-node cluster can cause a split-brain condition in cluster.
  - **Must be physical hardware:**
    - Many customers only have virtual environments.
  - **Network must remain stable:**
    - Network instabilities can result in intermittent cluster membership.
- **Given this, GeoEvent Server clustering:**
  - is strongly discouraged at ArcGIS 10.3, 10.4 & 10.5
  - and should be avoided.
- **Instead, if resiliency or additional scalability is required:**
  - isolated deployments should be employed.



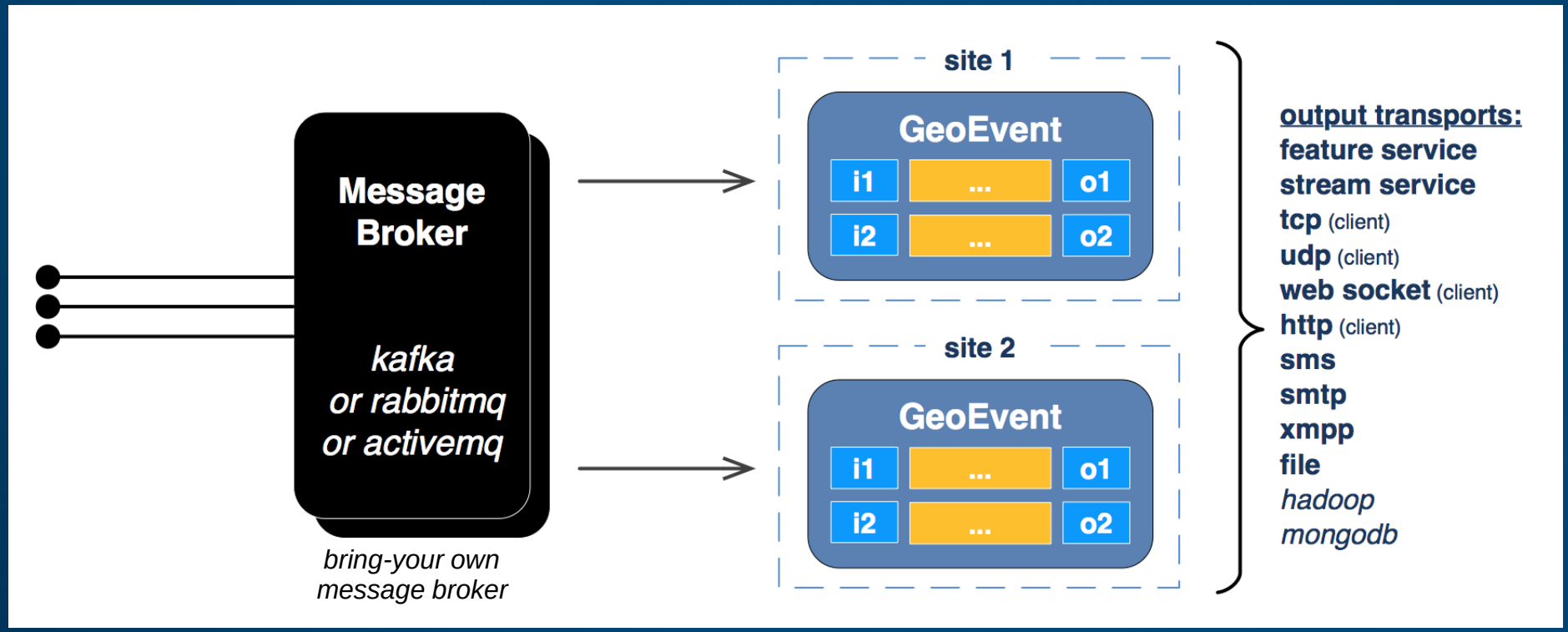
GeoEvent Server resilience & scalability



GeoEvent Server resilience & scalability

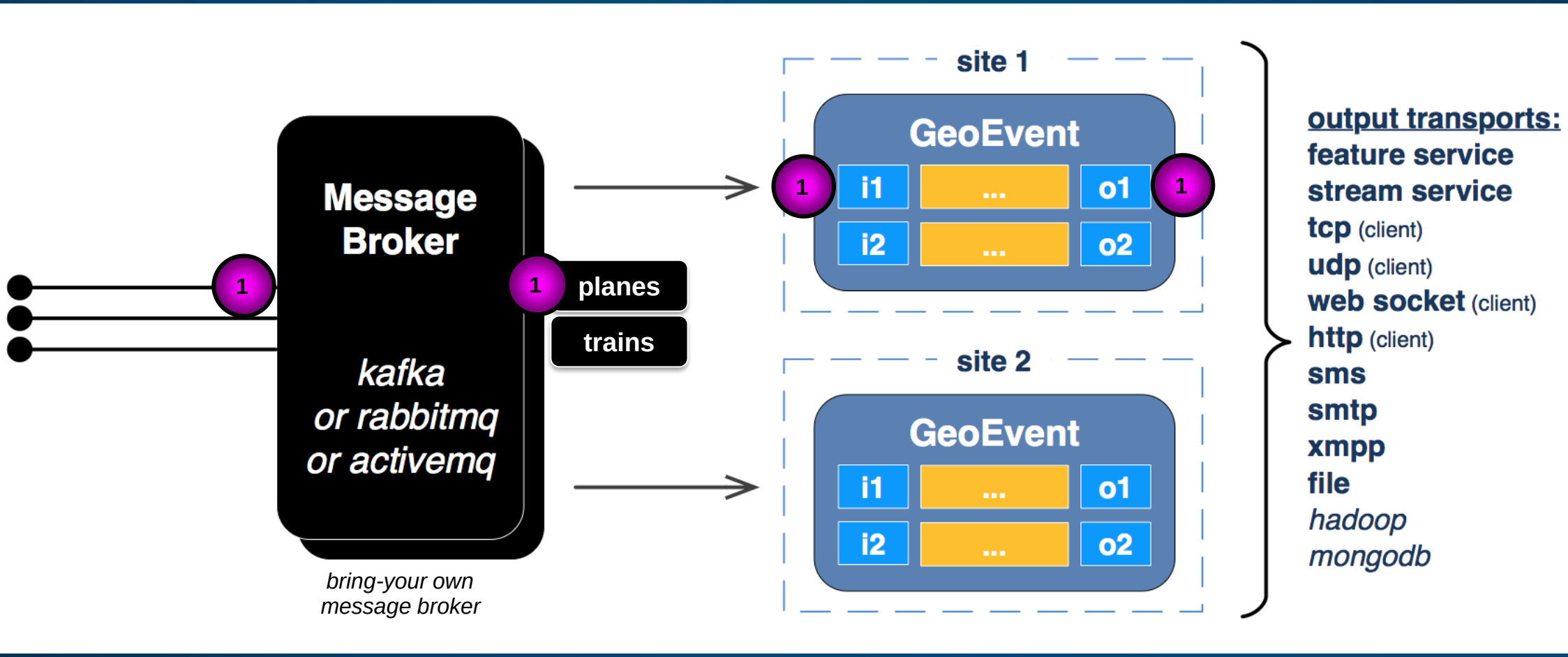
isolated deployment

- Each GeoEvent instance exists in it's own site and shares a common configuration manually.
- A message broker must be stood up in front to provide a common event source.
- GeoEvent instances share a common message broker and run active/active competing to consume events.



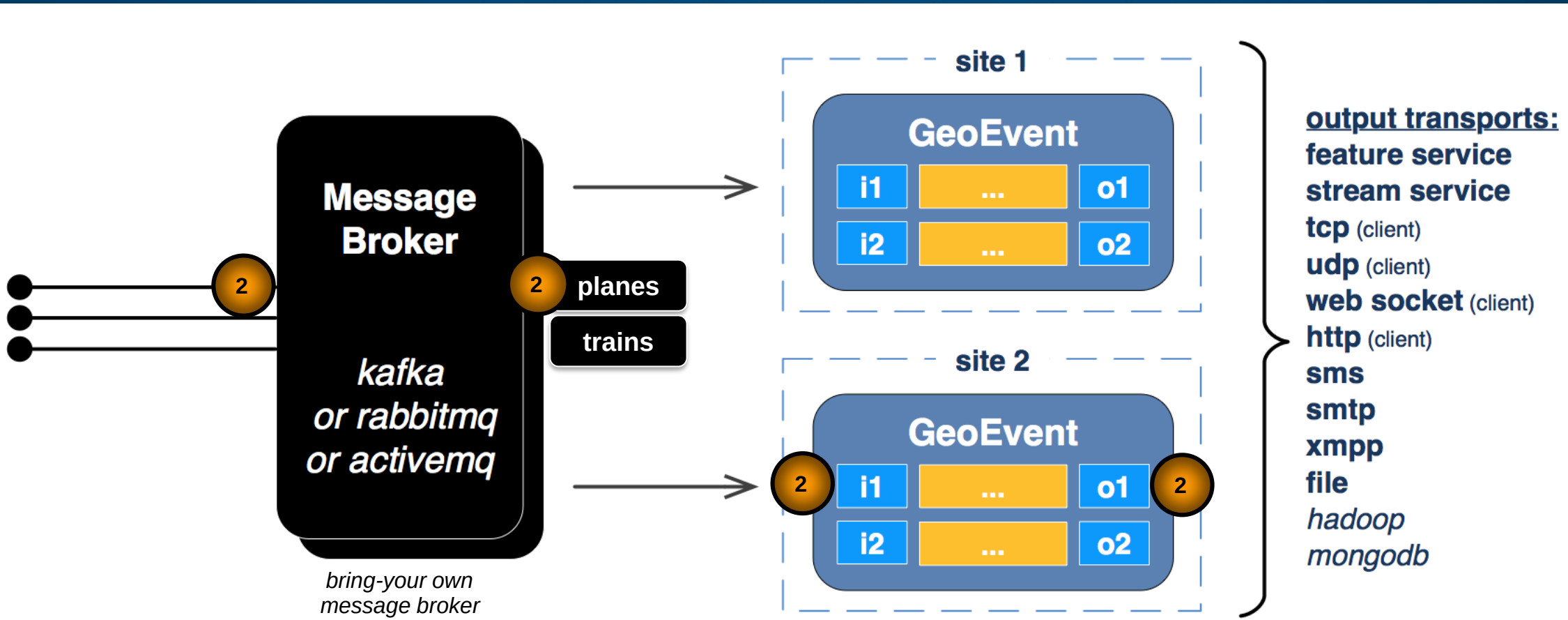
GeoEvent Server resilience & scalability

isolated deployment



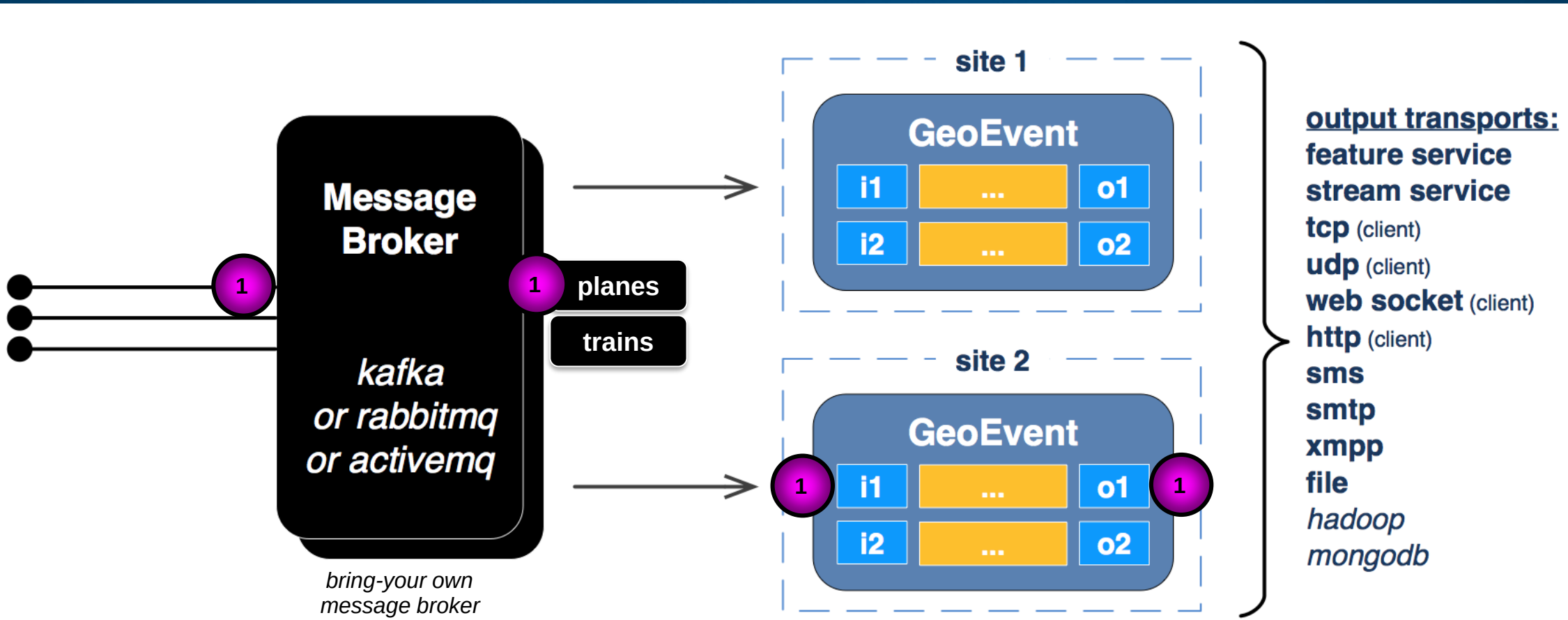
GeoEvent Server resilience & scalability

isolated deployment



GeoEvent Server resilience & scalability

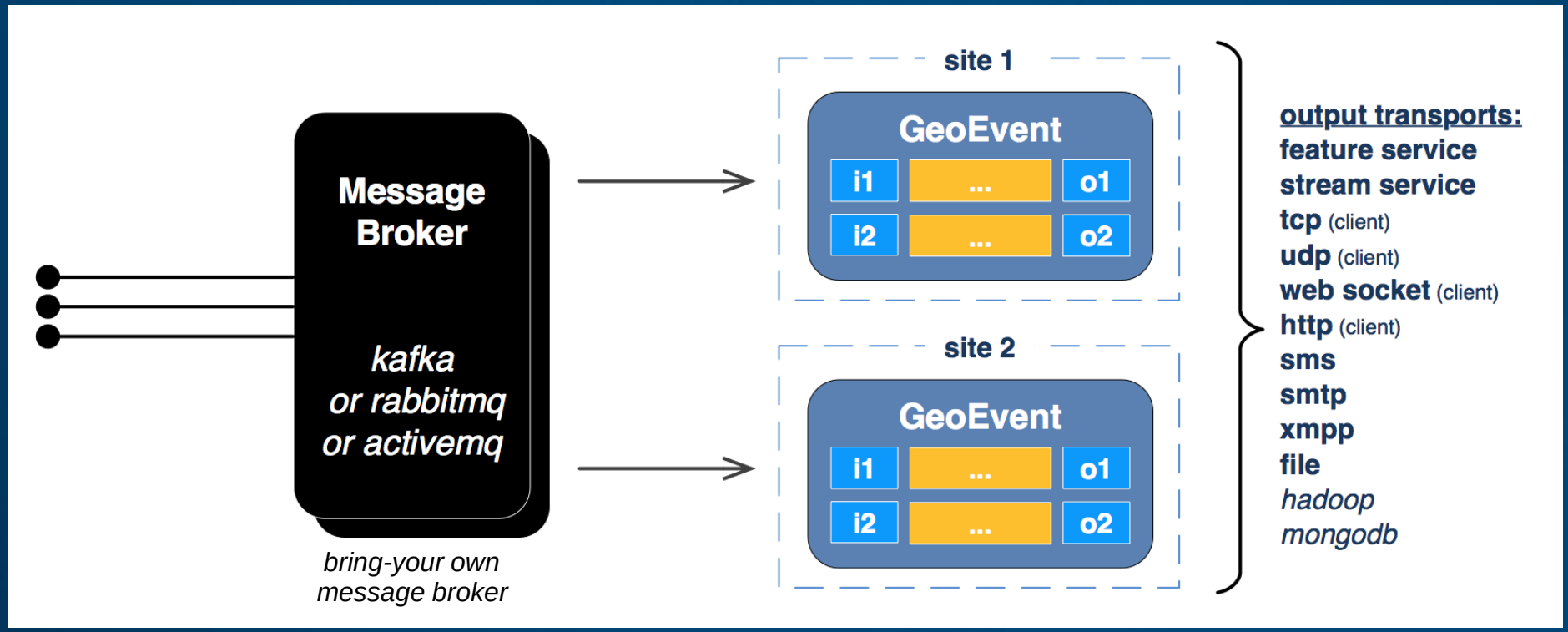
isolated deployment



GeoEvent Server resilience & scalability

isolated deployment

- **Known constraints:**
  - **project team/IT staff needed to setup/configure message broker and get all sensor events into it**
  - **input to GeoEvent is the message broker only, lose flexibility of all the input connectors available**
  - **does not support failover of state: enter, exit, track gap detection, idle detector**



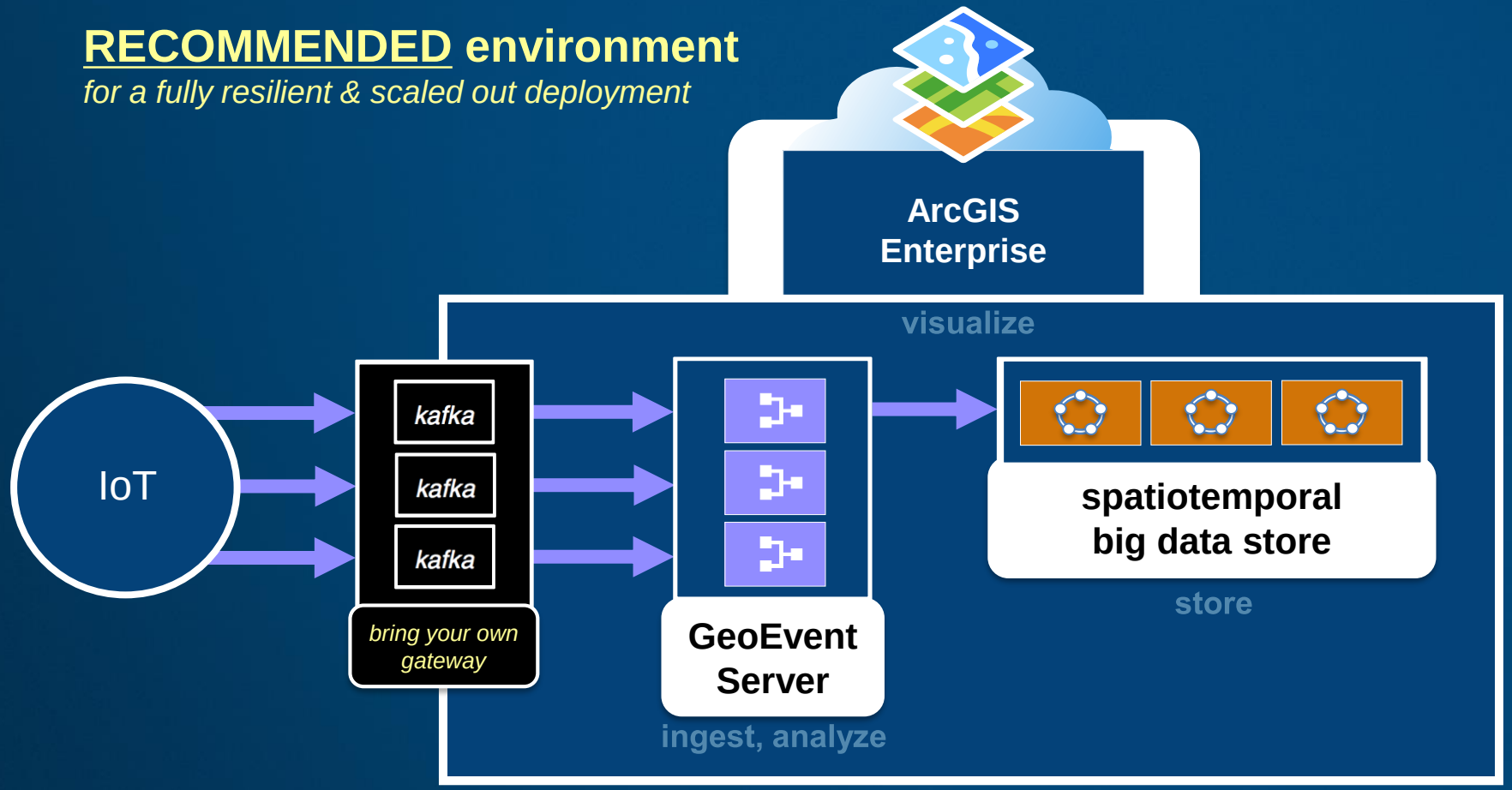
ArcGIS Enterprise

with real-time capabilities

10.5

RECOMMENDED environment

for a fully resilient & scaled out deployment



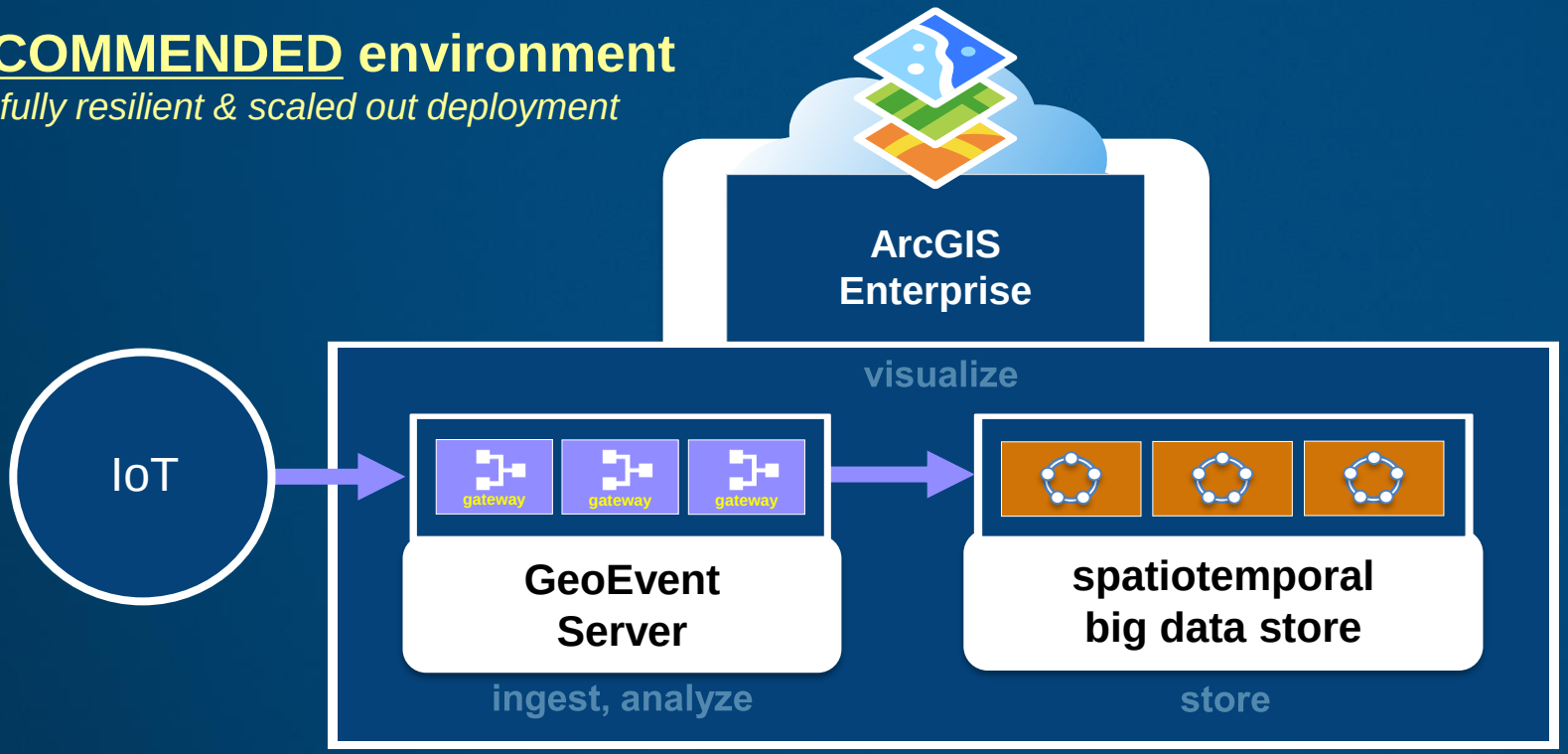
ArcGIS Enterprise

with real-time capabilities

10.6

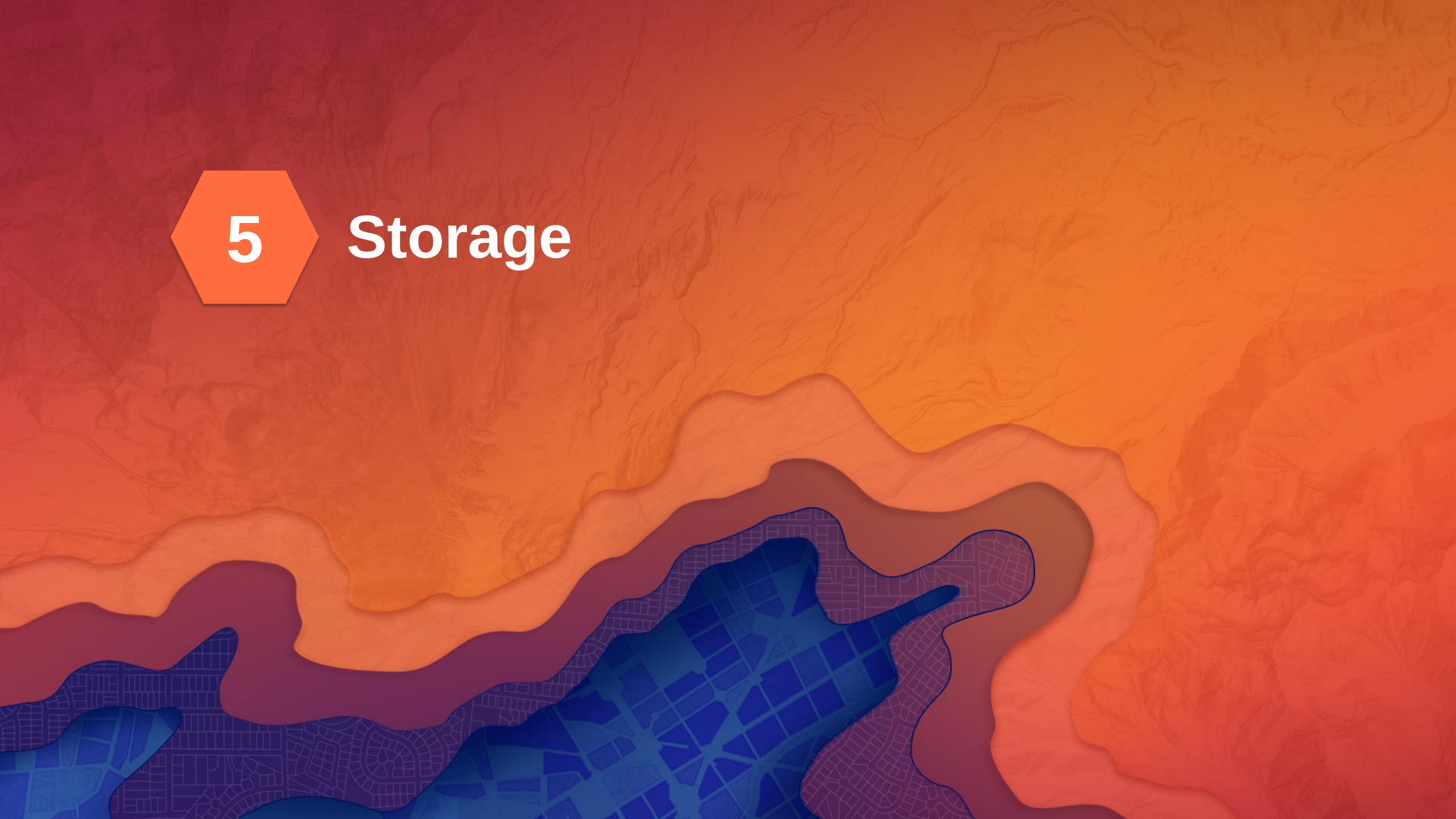
RECOMMENDED environment

for a fully resilient & scaled out deployment



Clustering is supported again!

5 Storage



Spatiotemporal Big Data Store Performance

What are the primary factors I should consider?

- **Operating environment:** *Amazon EC2 c4.2xlarge*
  - Virtual Machines – beware! resources need to be shared in an effective way, like EC2 or Azure.
  - Bare-Metal machines – dedicated resources and public cloud instances are much more deterministic.
- **Disk**
  - Speed (Mbps) – the faster the better *1,000 Mbps EBS, note: local SSD is much better*
- **Network**
  - Speed (Gbps) – the faster the better. *10 Gbps*
- **RAM**
  - size (GB) – minimum of 16GB is recommended. *15GiB, the Big Data Store was allocated 8GiB*
  - type – minimum of DDR3 is recommended.
  - clock speed (MHz) and transfer rate (Mbps) – the faster the better.
- **Processors**
  - speed (GHz) – the faster the better.
  - # of cores – the more the better. *8 vCPU*

spatiotemporal big data store

operating system requirements

resource	minimum	RECOMMENDED	notes
64-bit operating system (Windows)	Windows Server: 2012 R2 Standard & Datacenter (Sep 2016) 2012 Standard & Datacenter (Sep 2016) 2008 R2 Standard, Enterprise & Datacenter (sp1) 2008 Standard, Enterprise & Datacenter (sp2)	Windows Server: 2012 R2 Standard & Datacenter (Sep 2016)	'_' characters are not allowed in machine names
64-bit operating system (Linux)	Red Hat Enterprise Linux Server 7 (update 8) Red Hat Enterprise Linux Server 6 (update 2) CentOS Linux 7 (7.2) CentOS Linux 6 (6.8) Scientific Linux 7 (7.2) Scientific Linux 6 (6.8) Ubuntu Server LTS (16.04.1) SUSE Linux Enterprise Server 12 (sp1) SUSE Linux Enterprise Server 11 (sp4) Oracle Linux 7 (update 2) Oracle Linux 6 (update 8)	Use your Linux flavor of choice with the latest update version Esri has certified.	

http://server.arcgis.com/en/data-store/latest/install/windows/arcgis-data-store-system-requirements.htm
http://server.arcgis.com/en/data-store/latest/install/linux/arcgis-data-store-system-requirements.htm

spatiotemporal big data store

resource requirements: processors

resource	minimum	RECOMMENDED	notes
processors	8 cores	16 cores	More processors results in faster writes (indexing time), higher write throughput & faster query response time.

http://server.arcgis.com/en/data-store/latest/install/windows/arcgis-data-store-system-requirements.htm
http://server.arcgis.com/en/data-store/latest/install/linux/arcgis-data-store-system-requirements.htm

spatiotemporal big data store

resource requirements: processors, memory

resource	minimum	RECOMMENDED	notes
processors	8 cores	16 cores	More processors results in faster writes (indexing time), higher write throughput & faster query response time.
storage	400 GB	1 TB or more	Storage need is dictated by data velocity & size. SSD results in much faster writes & queries.

http://server.arcgis.com/en/data-store/latest/install/windows/arcgis-data-store-system-requirements.htm
http://server.arcgis.com/en/data-store/latest/install/linux/arcgis-data-store-system-requirements.htm

spatiotemporal big data store

resource requirements: processors, memory, storage

resource	minimum	RECOMMENDED	notes
processors	8 cores	16 cores	More processors results in faster writes (indexing time), higher write throughput & faster query response time.
storage	400 GB	1 TB or more	Storage need is dictated by data velocity & size. SSD results in much faster writes & queries.
memory	16 GB	32 GB	More memory results in higher likelihood of query cache hits. 32 GB is the maximum due to JVM limitations.

http://server.arcgis.com/en/data-store/latest/install/windows/arcgis-data-store-system-requirements.htm
http://server.arcgis.com/en/data-store/latest/install/linux/arcgis-data-store-system-requirements.htm

spatiotemporal big data store

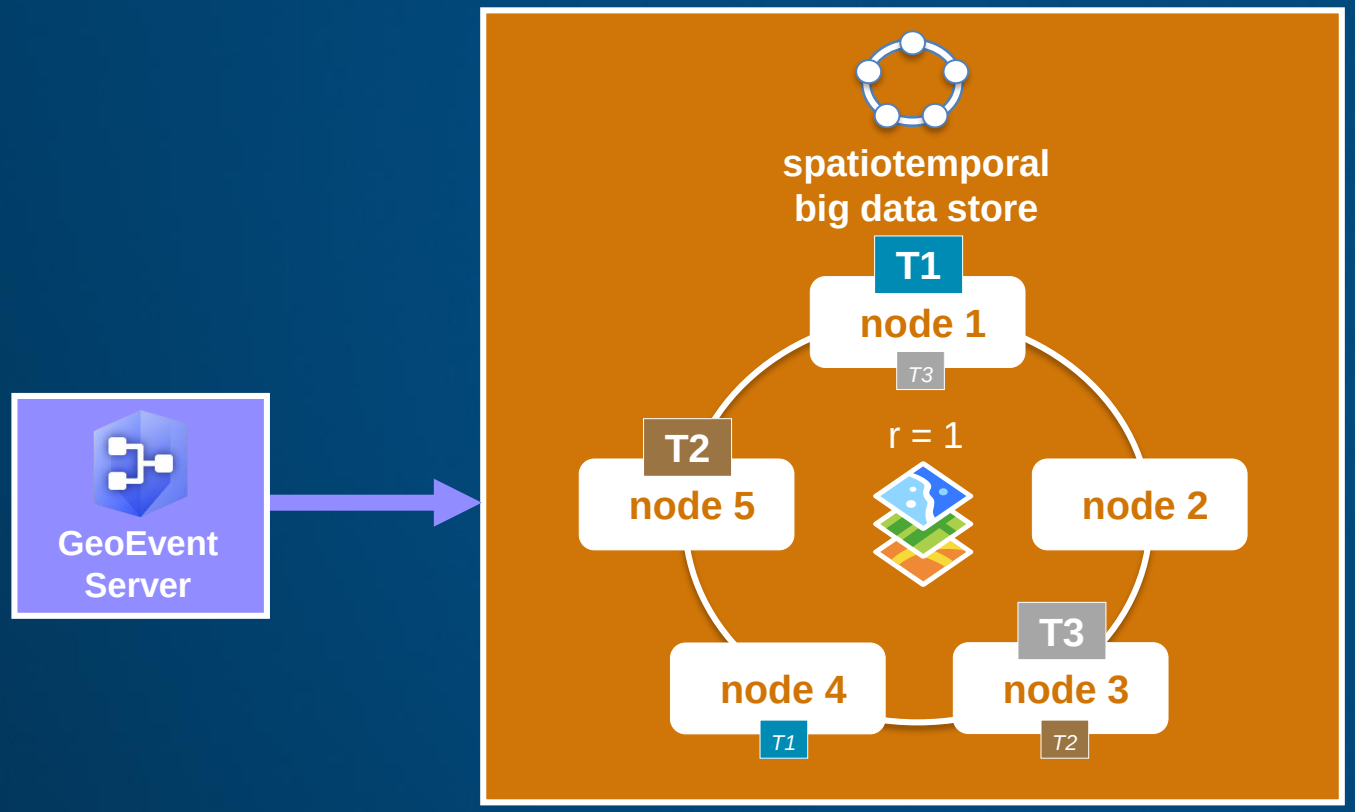
resource requirements: processors, memory, storage & instances

resource	minimum	RECOMMENDED	notes
processors	8 cores	16 cores	More processors results in faster writes (indexing time), higher write throughput & faster query response time.
storage	400 GB	1 TB or more	Storage need is dictated by data velocity & size. SSD results in much faster writes & queries.
memory	16 GB	32 GB	More memory results in higher likelihood of query cache hits. 32 GB is the maximum due to JVM limitations
# of **ISOLATED** instances	1	3 or more	With one instance there is risk of data loss. With three instances replicas are written increasing data reliability. The # of instances will vary based on the velocity of incoming data as well as the data retention policy configured on each data source.

http://server.arcgis.com/en/data-store/latest/install/windows/arcgis-data-store-system-requirements.htm
http://server.arcgis.com/en/data-store/latest/install/linux/arcgis-data-store-system-requirements.htm

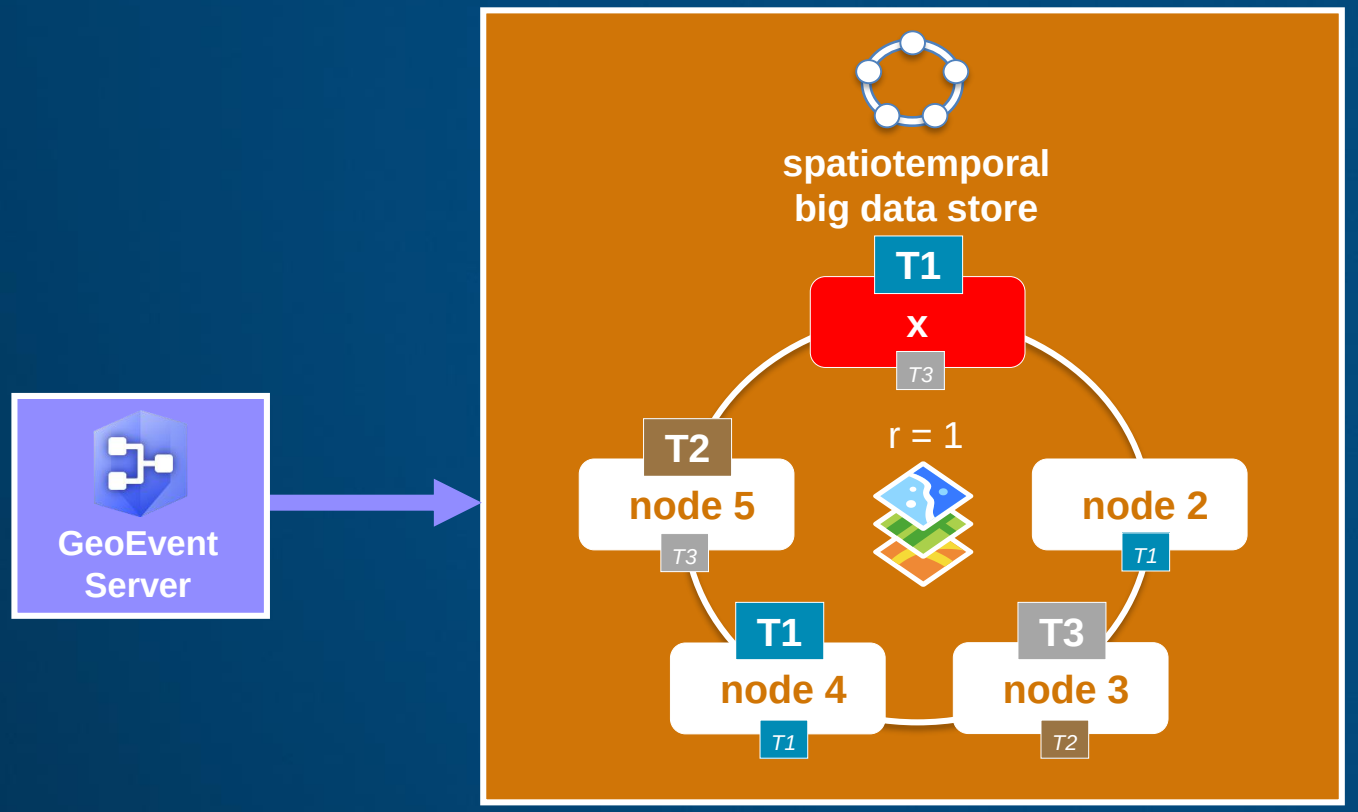
spatiotemporal big data store

shards & replication factor



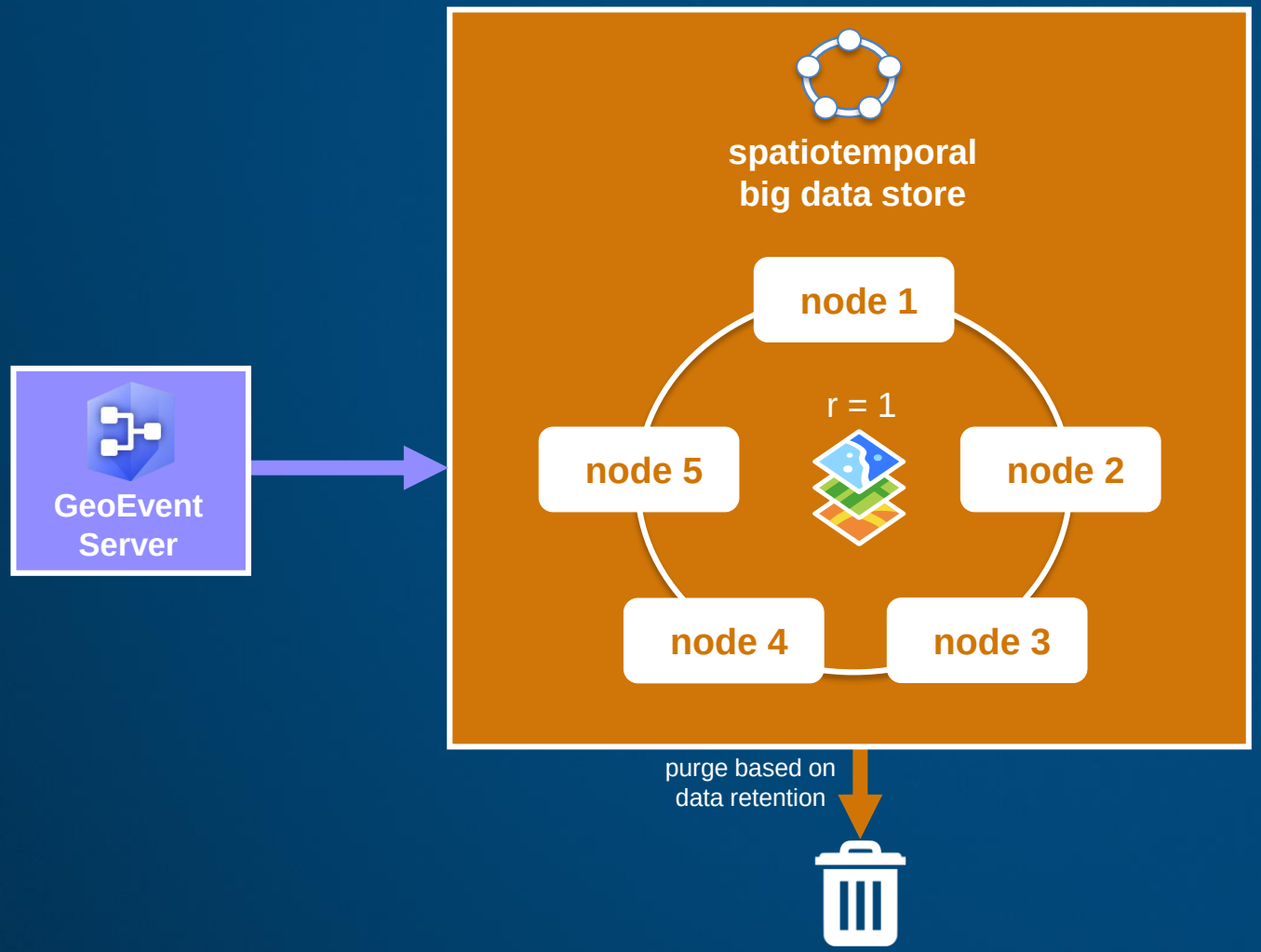
spatiotemporal big data store

auto-rebalancing of data upon node membership changes, + or -, in the big data store



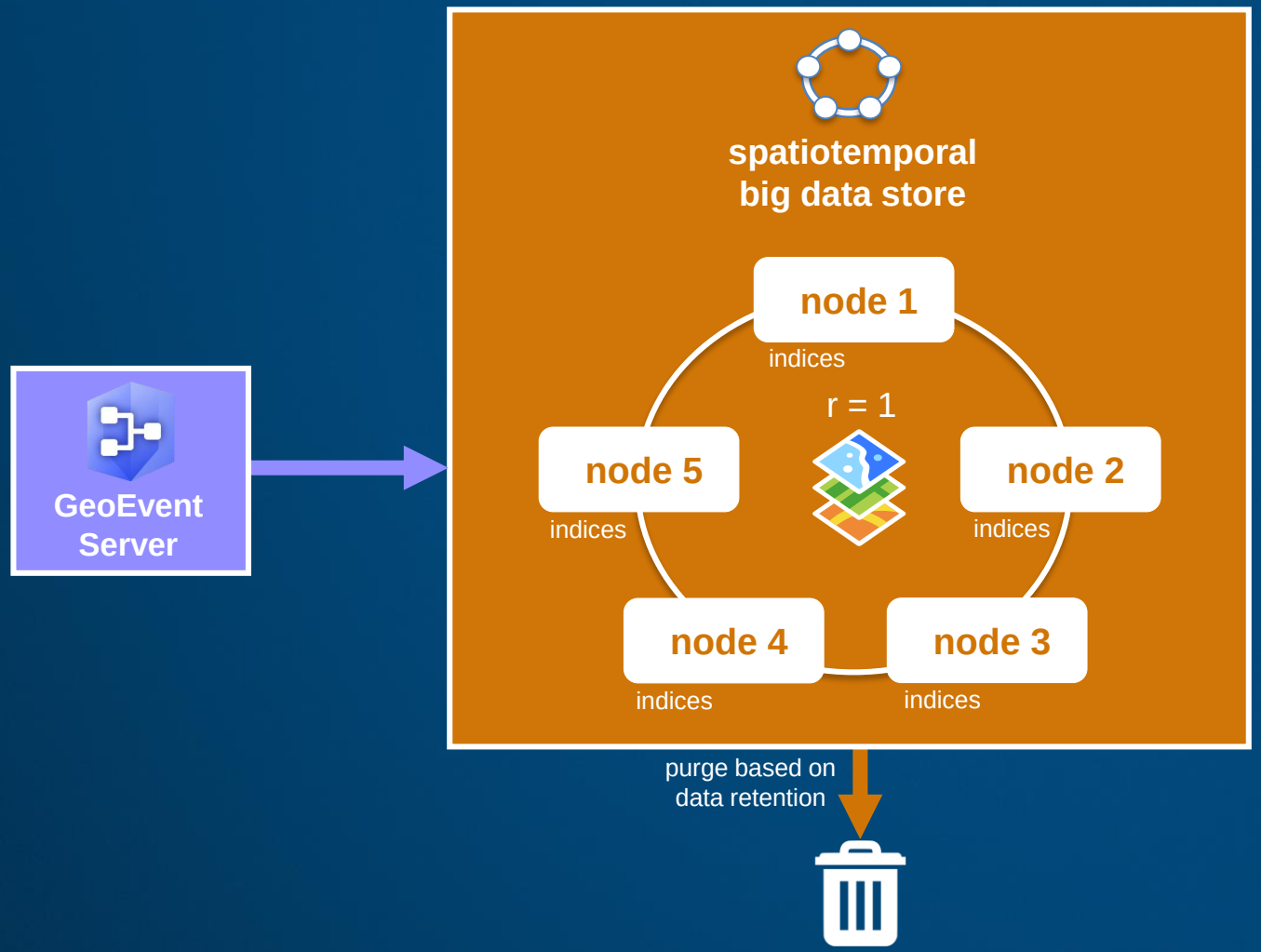
spatiotemporal big data store

data retention policies, configured per data source



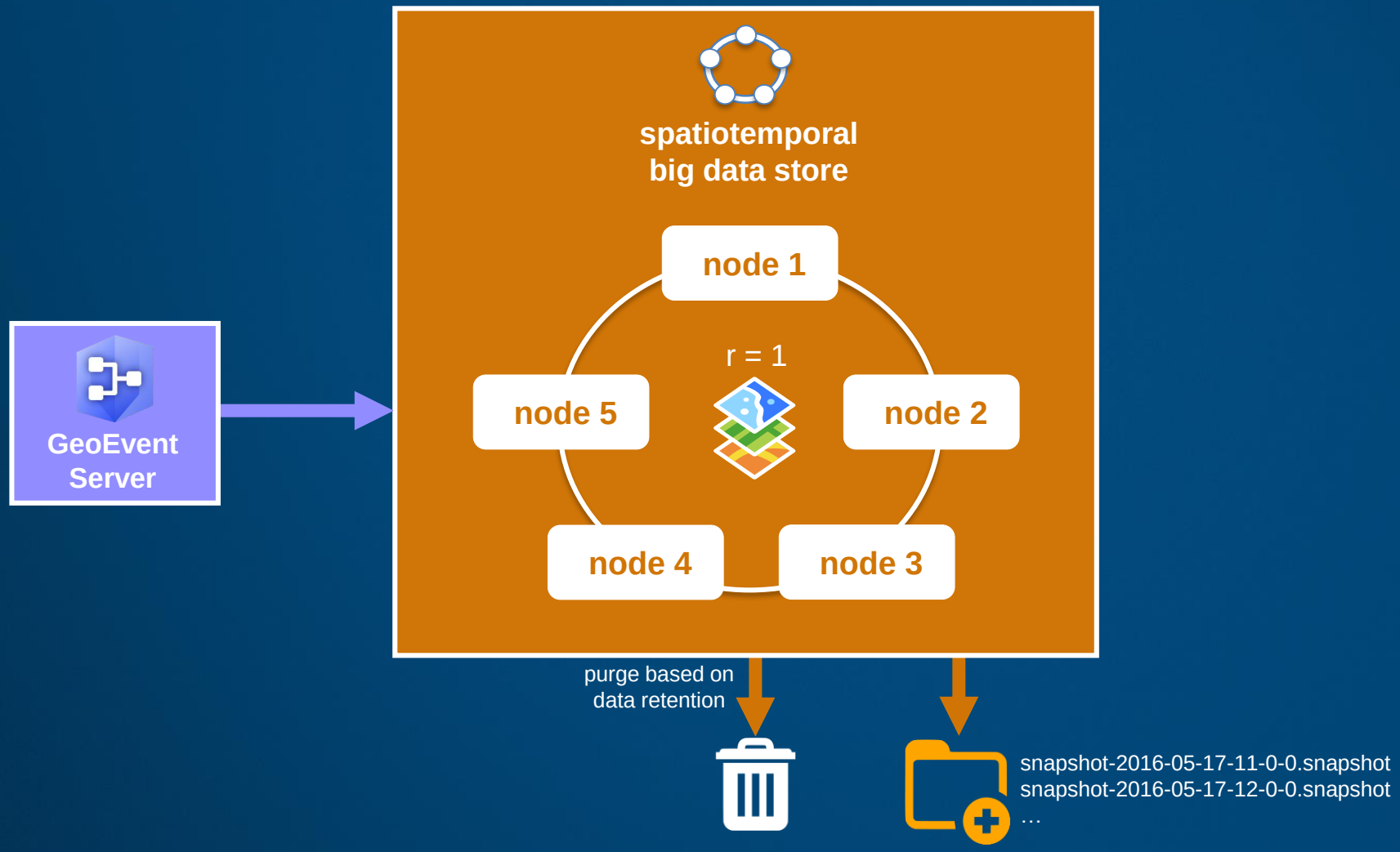
spatiotemporal big data store

***rolling index option**, set appropriately to the velocity of your observation data*



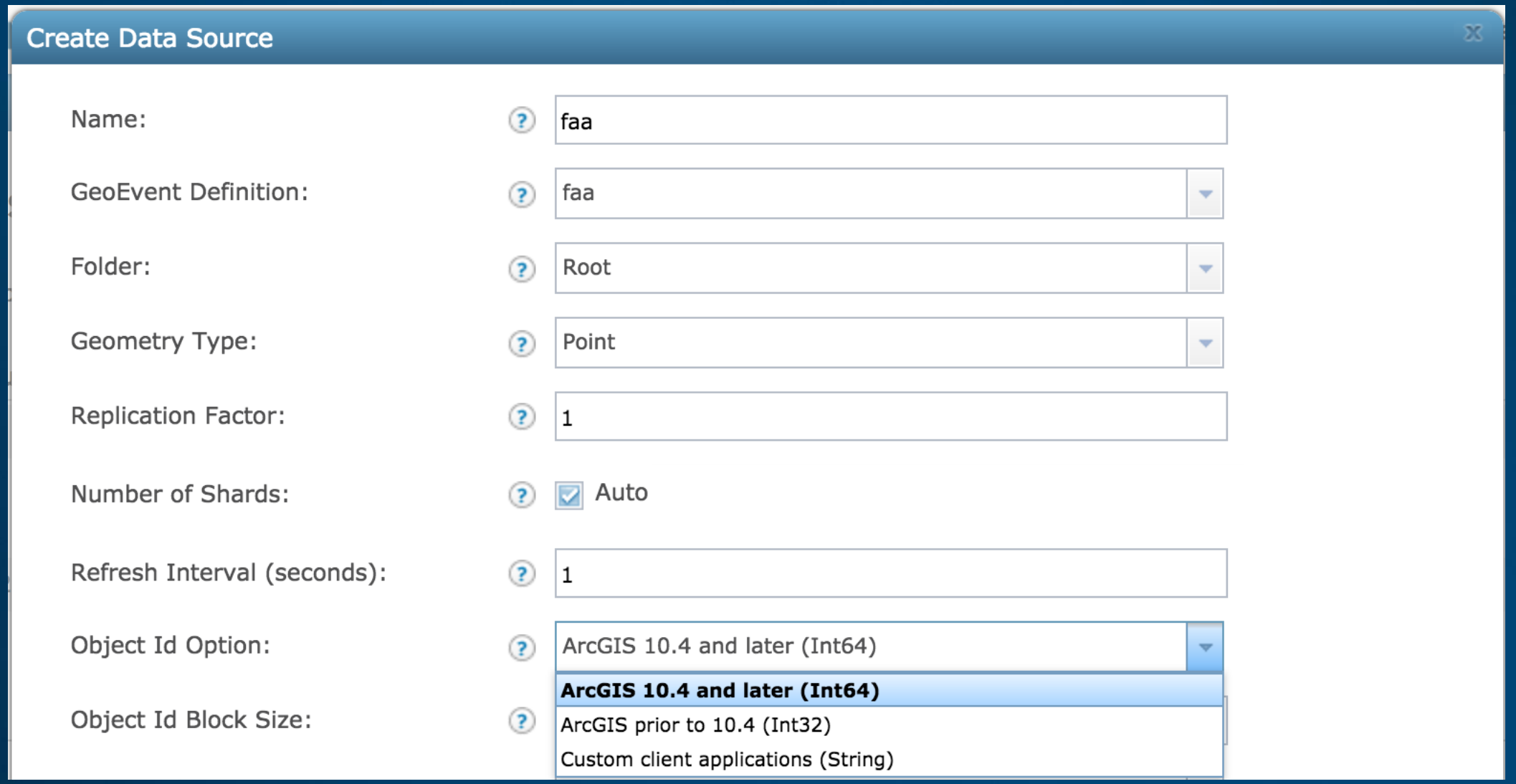
spatiotemporal big data store

automatic data backups using periodic snapshots, including ability to restore from a snapshot



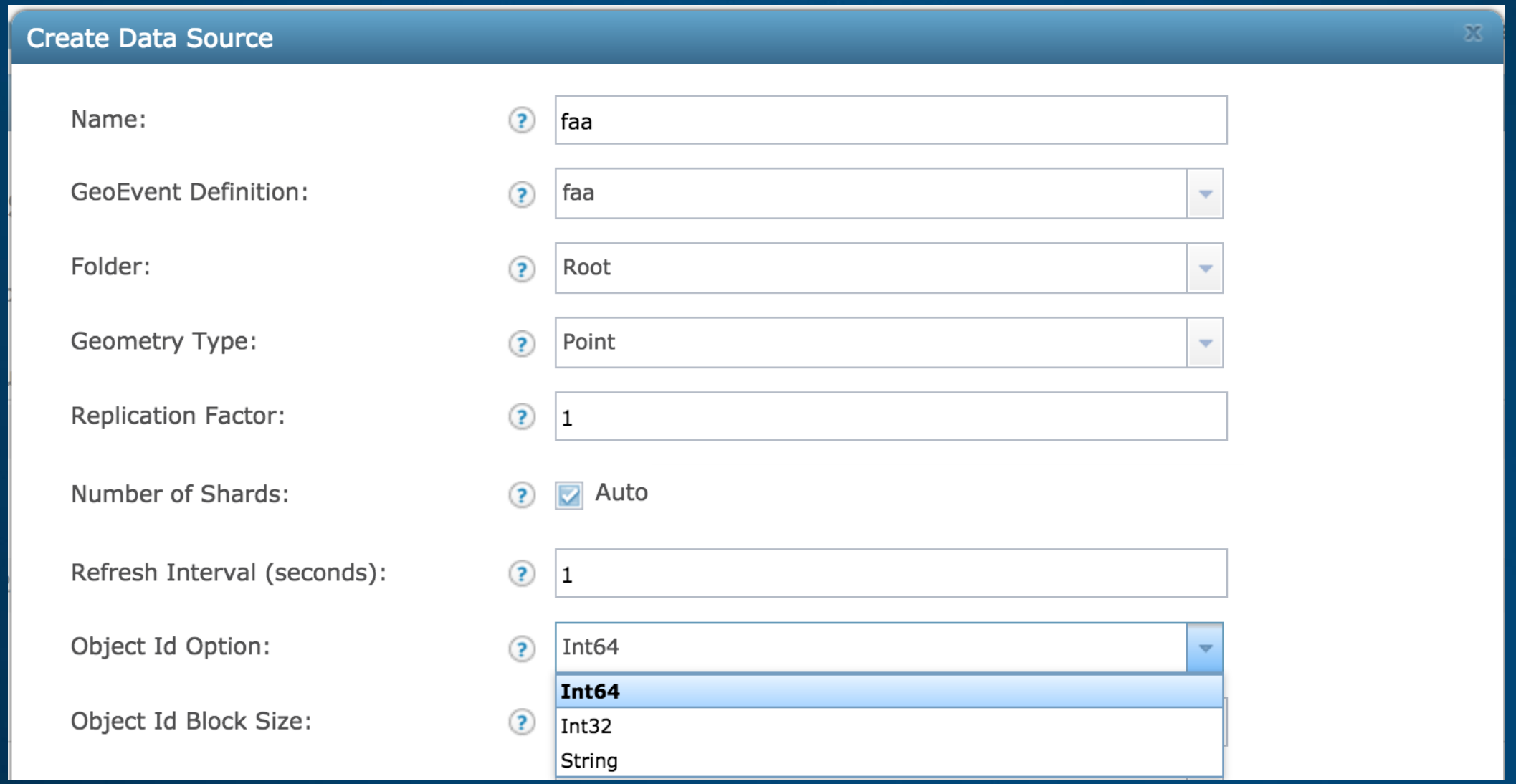
spatiotemporal big data store

choosing an Object Id option



spatiotemporal big data store

choosing an Object Id option



Demo

spatiotemporal big data store

choosing an Object Id option

	Max Value	# of IDs	ArcGIS Clients
Int32	2,147,483,647	2.1 billion	Pro, Desktop, Ops Dashboard, …

	events per day	Int32
1,000 e/s	86,400,000	25 days
10,000 e/s	864,000,000	2.5 days
100,000 e/s	8,640,000,000	6 hours
1,000,000 e/s	86,400,000,000	36 minutes
10,000,000 e/s	864,000,000,000	4 minutes

spatiotemporal big data store

choosing an Object Id option

	Max Value	# of IDs	ArcGIS Clients
Int32	2,147,483,647	2.1 billion	Pro, Desktop, Ops Dashboard, …
Int64 (signed)	9,223,372,036,854,775,807	9.2 quintillion	JavaScript, custom apps

	events per day	Int32
1,000 e/s	86,400,000	25 days
10,000 e/s	864,000,000	2.5 days
100,000 e/s	8,640,000,000	6 hours
1,000,000 e/s	86,400,000,000	36 minutes
10,000,000 e/s	864,000,000,000	4 minutes

spatiotemporal big data store

choosing an Object Id option

	Max Value	# of IDs	ArcGIS Clients
Int32	2,147,483,647	2.1 billion	Pro, Desktop, Ops Dashboard, …
Int64 (signed)	9,223,372,036,854,775,807	9.2 quintillion	JavaScript, custom apps

	events per day	Int32	Int64 (signed)
1,000 e/s	86,400,000	25 days	292,472,000 years
10,000 e/s	864,000,000	2.5 days	29,247,200 years
100,000 e/s	8,640,000,000	6 hours	2,924,720 years
1,000,000 e/s	86,400,000,000	36 minutes	292,472 years
10,000,000 e/s	864,000,000,000	4 minutes	29,248 years

spatiotemporal big data store

choosing an Object Id option

	Max Value	# of IDs	ArcGIS Clients
Int32	2,147,483,647	2.1 billion	Pro, Desktop, Ops Dashboard, …
Int64 (signed)	9,223,372,036,854,775,807	9.2 quintillion	JavaScript, custom apps
UniqueStringID	n/a	unlimited	JavaScript, custom apps

	events per day	Int32	Int64 (signed)
1,000 e/s	86,400,000	25 days	292,472,000 years
10,000 e/s	864,000,000	2.5 days	29,247,200 years
100,000 e/s	8,640,000,000	6 hours	2,924,720 years
1,000,000 e/s	86,400,000,000	36 minutes	292,472 years
10,000,000 e/s	864,000,000,000	4 minutes	29,248 years

Summary

GeoEvent Server: best practices

- **When applying Real-Time GIS there are many best practices for**
  - **ingestion & analytics**
  - **dissemination & visualization**
  - **storage**
  - **resilience & scalability**

- **To learn more:**
  - **See the tutorial:**
    ‘ArcGIS GeoEvent Server: Resiliency’

Real-Time & Big Data GIS

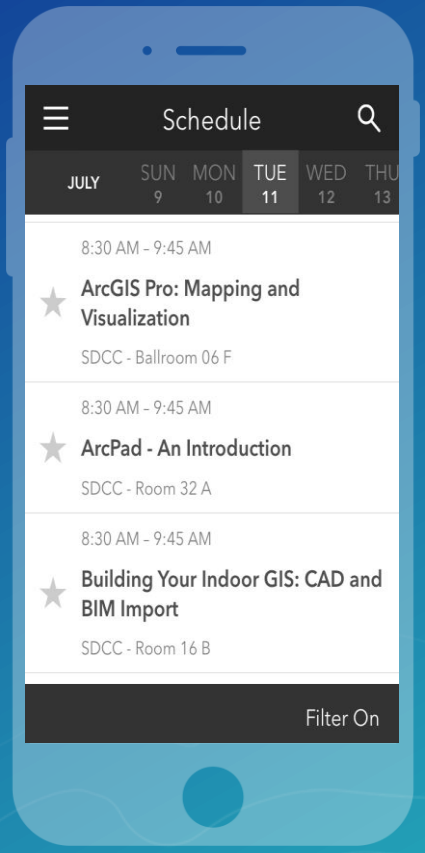
other sessions

• **GeoEvent Server: An Introduction**	Tue, 10:15-11:30am, Room 10 Thu, 1:30-2:45pm, Hilton – Sapphire Ballroom I
• **Real-Time & Big Data: Leveraging the Spatiotemporal Store**	Tue, 10:15-11:30am, Room 15 A Thu, 1:30-2:45pm, Room 15 A
• **GeoEvent Server: Applying Real-Time Analytics**	Tue, 1:30-2:45pm, Room 17 B Thu, 3:15-4:30pm, Room 14 A
• **Real-Time & Big Data GIS at a Massive Scale**	Wed, 3:15-4:30pm, Room 3 Fri, 9:00-10:15am, Room 8
• **GeoEvent Server: Leveraging Stream Services**	Wed, 3:15-4:30pm, Room 14 B
• **GeoEvent Server: Best Practices**	Thu, 10:15-11:30am, Room 9
• **GeoEvent Server: Internet of Things (IoT)**	Thu, 10:15-11:30am, Room 14 B
• **GeoEvent Server: Making 3D Scenes Come Alive**	Wed, 1:30-2:15pm, Demo Theater 05 - Real-Time
• **GeoAnalytics Server: An Introduction**	Wed, 10:15-11:30am, Room 4 Thu, 10:15-11:30am, Hilton – Sapphire Ballroom E

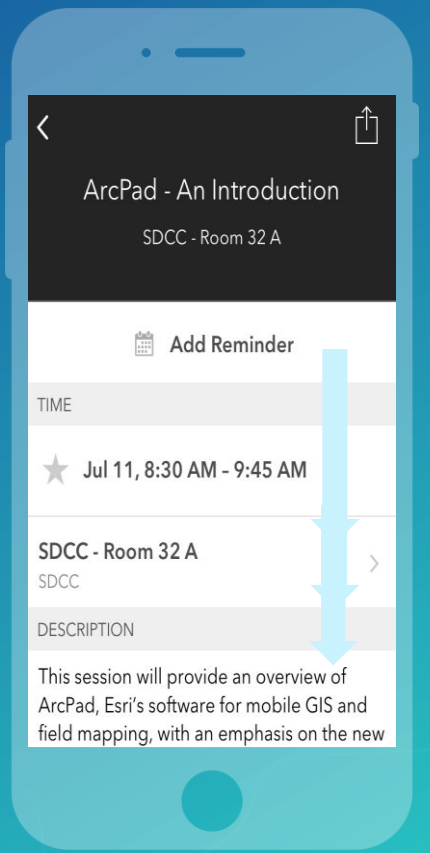
Please Take Our Survey on the **Esri Events App!**

Download the Esri Events app and find your event

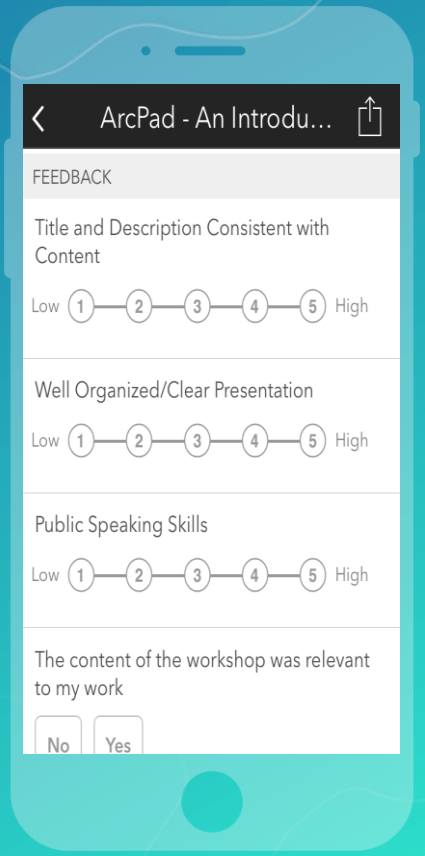
Select the session you attended



Scroll down to find the survey



Complete Answers and Select "Submit"



Questions / Feedback?

Suzanne Foss
Real-Time & Big Data Product Engineer, Esri
sfoss@esri.com
@sfoss_esri

Sagar Ayare
Real-Time & Big Data Product Engineer, Esri
sayare@esri.com
@theTechieSagar

esri | THE SCIENCE OF WHERE