# Real-Time GIS: Best Practices


RJ Sunderman
Product Engineer
GeoEvent Extension
rjsunderman@esri.com

C. Adam Mollenkopf
Real-Time GIS
Capability Lead
amollenkopf@esri.com
@amollenkopf

# What are we going to discuss?

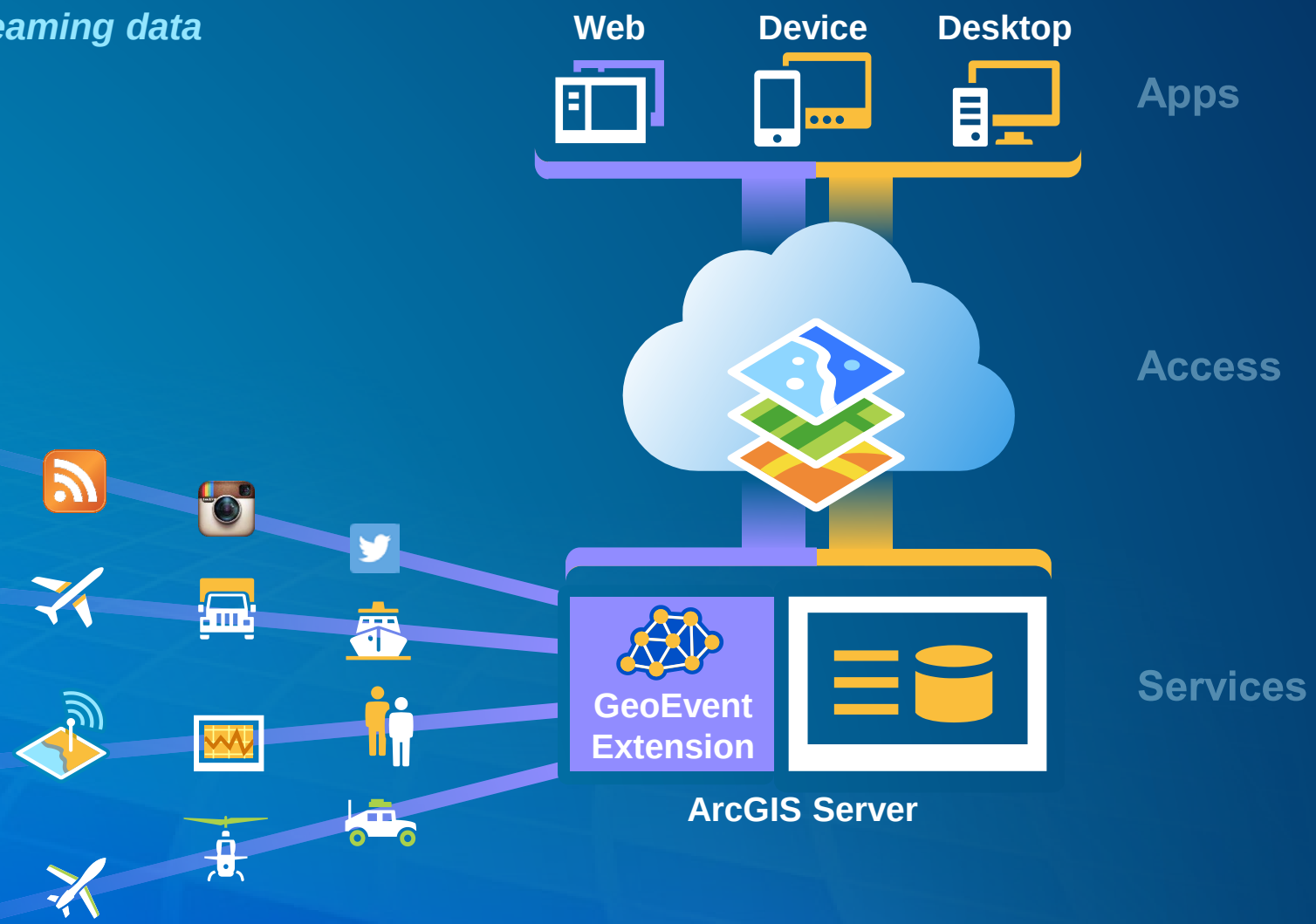
***Best Practices on***

- **Input and Output Connector usage**
- **Where, when, and why GeoEvent Definitions are created**
- **Feature Services and Stream Services workflows**
- **Performance**
- **High Availability & Scalability**

# Real-Time GIS

*Integration and exploitation of streaming data*

- Integrates real-time streaming data into ArcGIS
- Performs continuous processing and real-time analytics
- Sends updates and alerts to those who need it where they need it



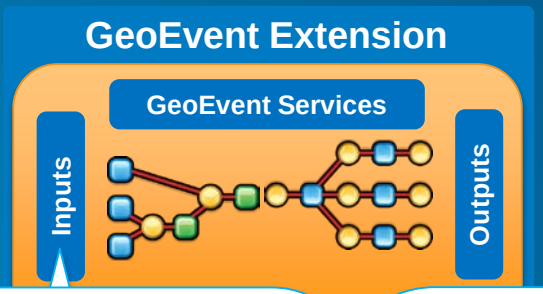
# Best Practice:

## How do I get my real-time data into ArcGIS?

# How do I get my real-time data into ArcGIS?

*Easily integrate real-time streaming data into ArcGIS using an Input Connector*

Use an existing connector




You can create your own connectors.

## Out of the Box

- Poll an ArcGIS Server for Features
- Poll an external website for GeoJSON, JSON, or XML
- Receive Features, GeoJSON, JSON, or XML on a REST endpoint
- Receive GeoJSON or JSON on a WebSocket
- Receive RSS
- Receive Text from a TCP or UDP Socket
- Subscribe to an external WebSocket for GeoJSON or JSON
- Watch a Folder for new CSV, GeoJSON, or JSON Files

## Esri Gallery

- ActiveMQ
- CAP
- Cursor-on-Target
- Exploitation Support Data
- Instagram
- KML
- Kafka *
- MQTT
- NMEA 0183
- RabbitMQ
- Sierra Wireless (RAP)
- Trimble (TAIP)
- Twitter

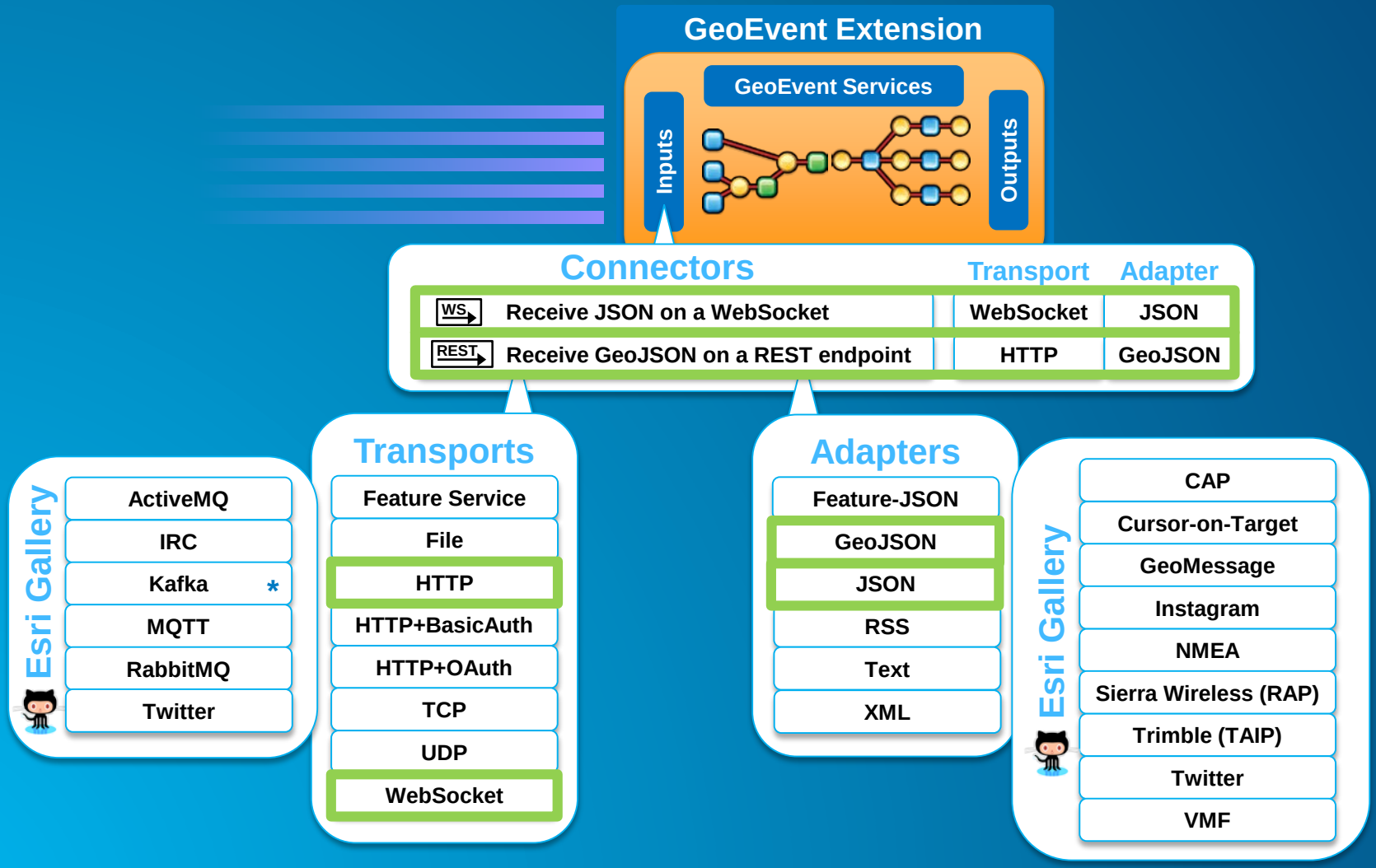
## Partner Gallery

- CompassLDE
- enviroCar
- exactEarth AIS
- FAA (ASDI) *
- GNIP *
- Networkfleet *
- OSIsoft *
- Valarm
- Zonar *

# How do I get my real-time data into ArcGIS?

*Easily integrate real-time streaming data into ArcGIS using an Input Connector*

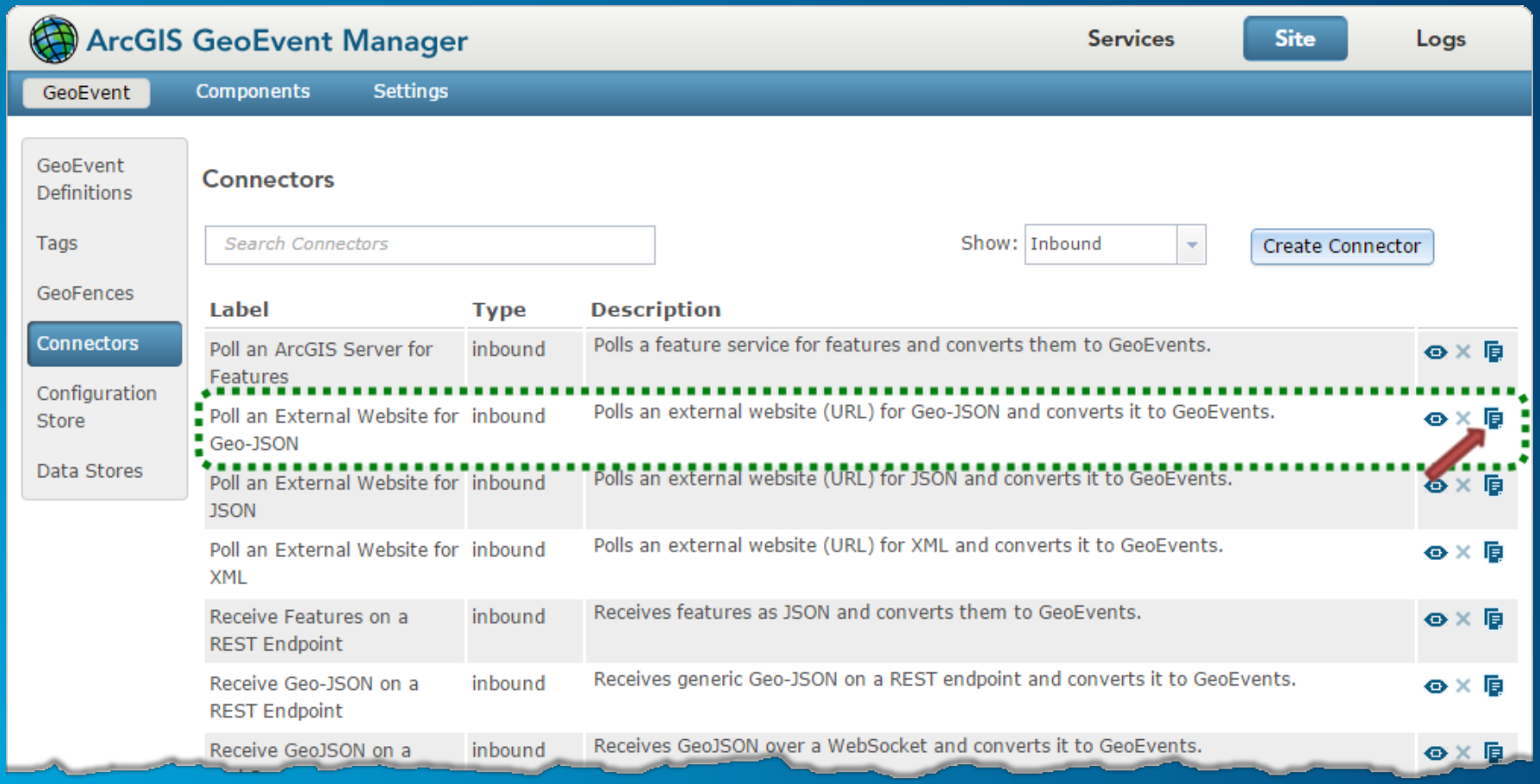
**Input Connector = Adapter + Transport**



# How do I get my real-time data into ArcGIS?

*Easily integrate real-time streaming data into ArcGIS using an Input Connector*

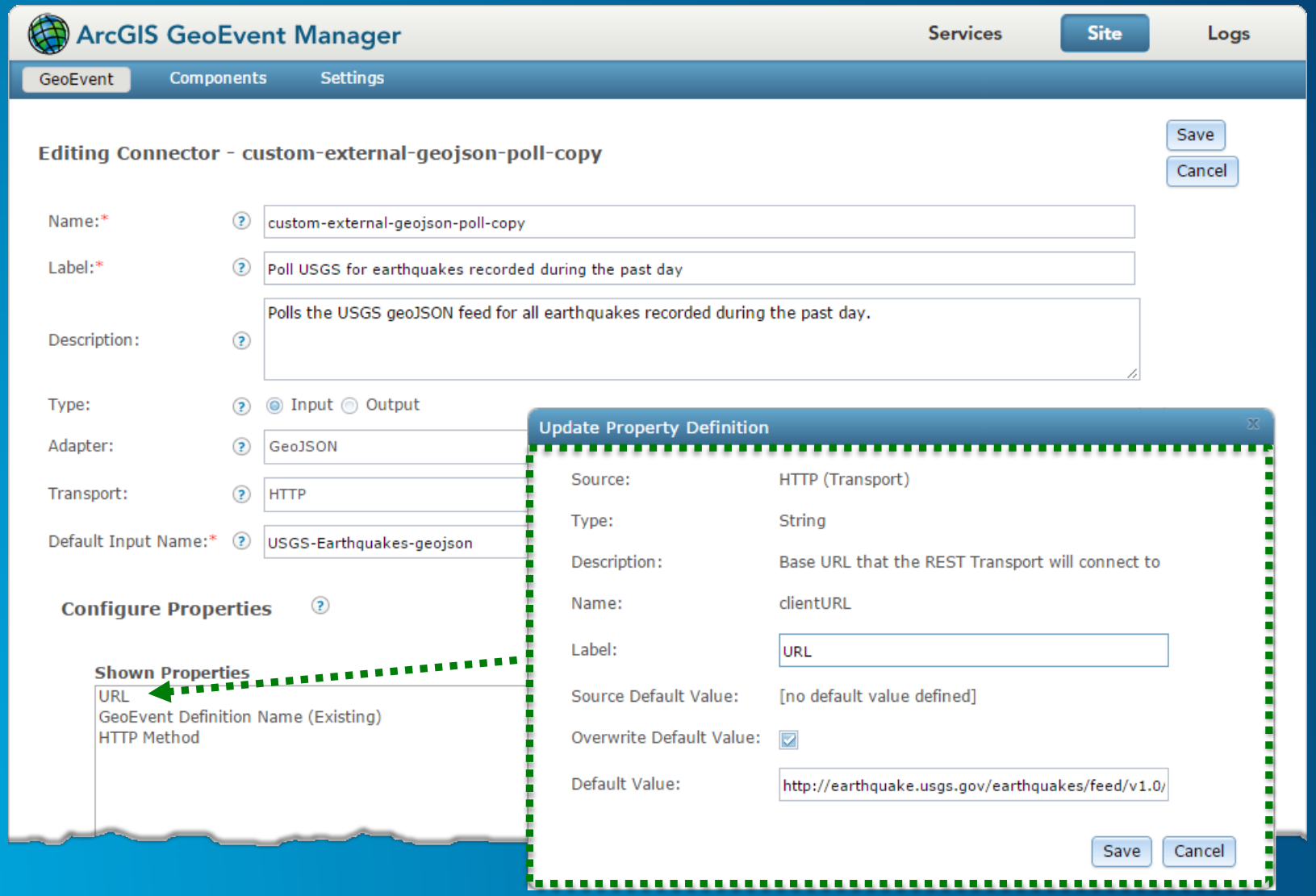
**Adjust an existing connector**




- **You can copy any of the out-of-the-box connectors and tailor their properties.**

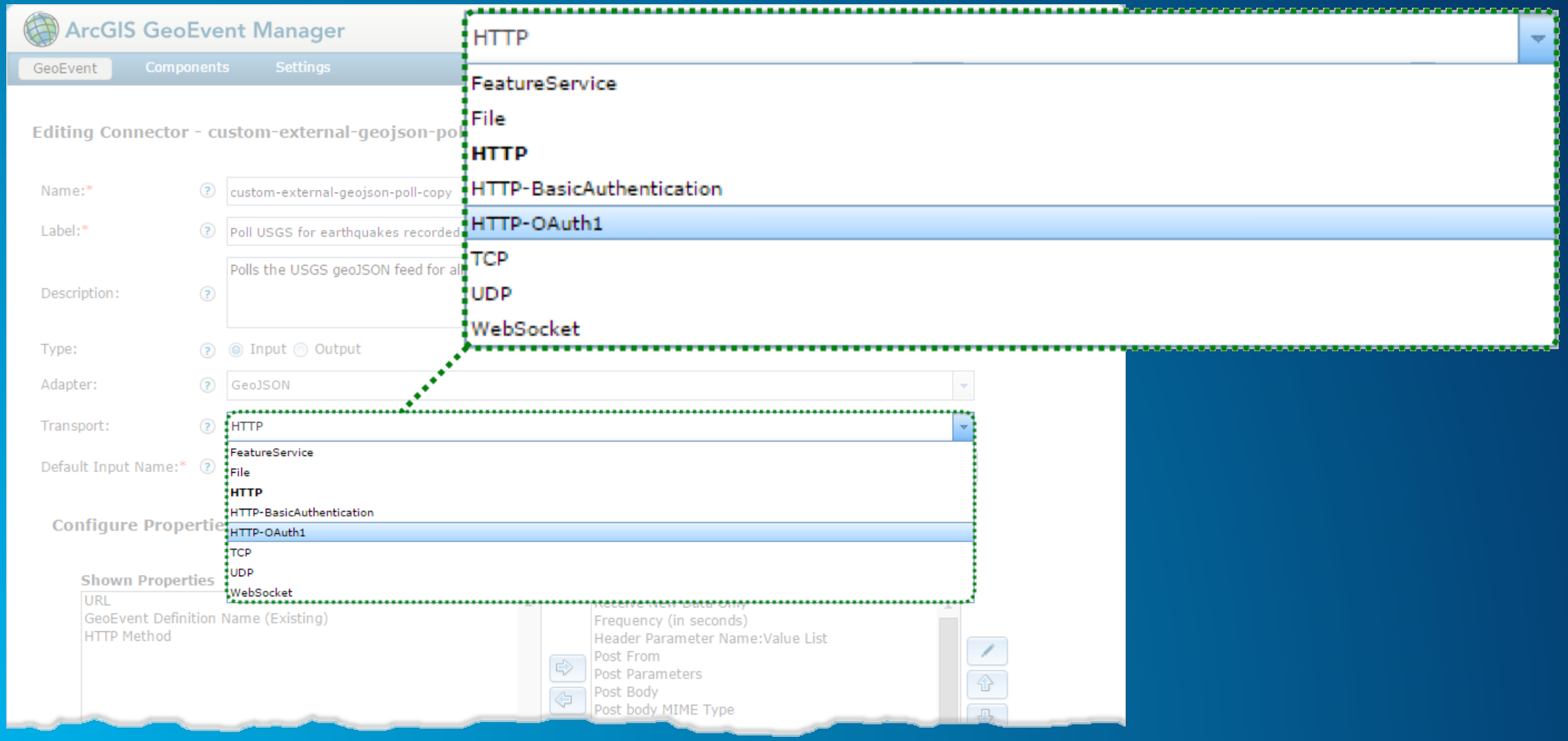
# How do I get my real-time data into ArcGIS?

*Easily integrate real-time streaming data into ArcGIS using an Input Connector*



# How do I get my real-time data into ArcGIS?

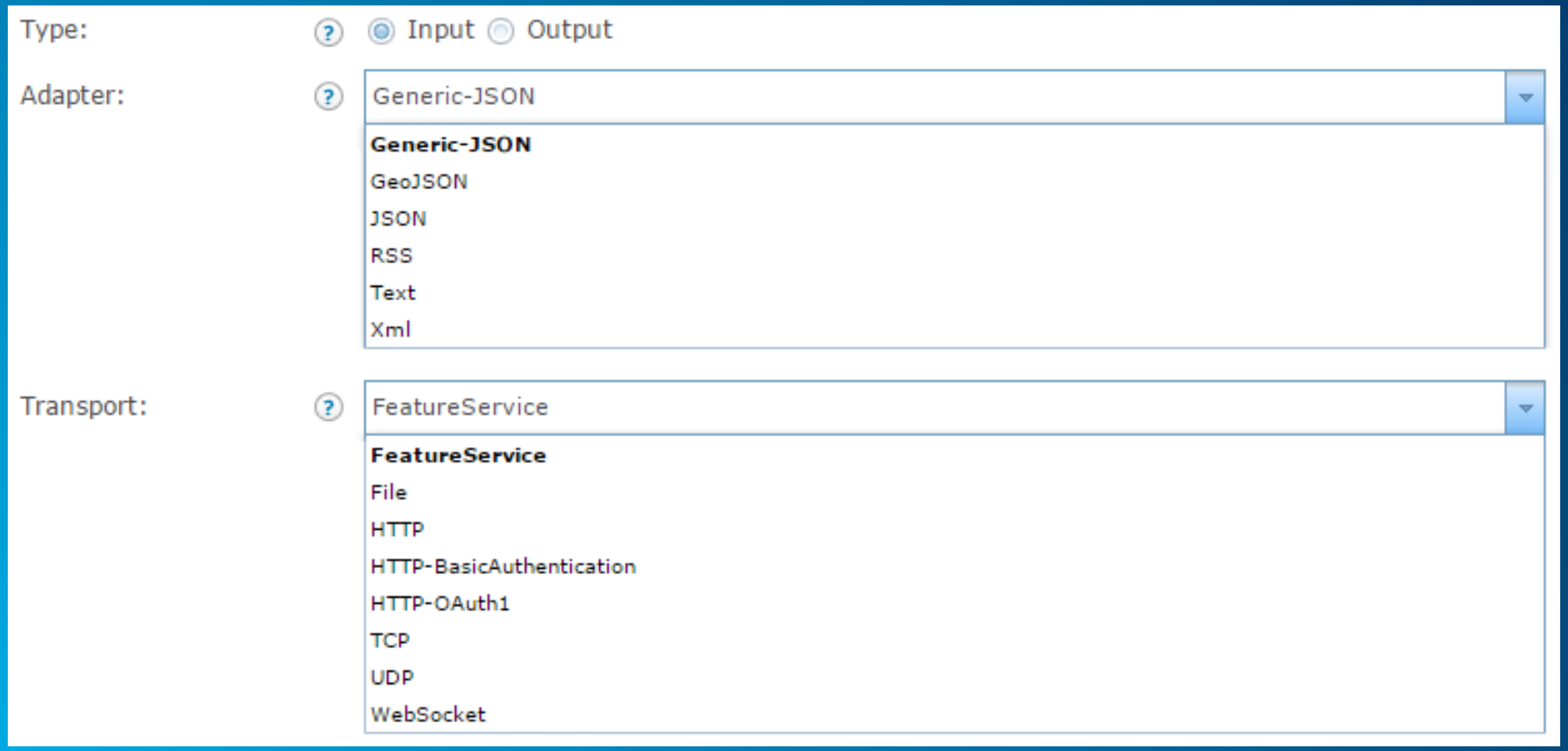
***Easily integrate real-time streaming data into ArcGIS using an Input Connector***



# How do I get my real-time data into ArcGIS?

*Easily integrate real-time streaming data into ArcGIS using an Input Connector*

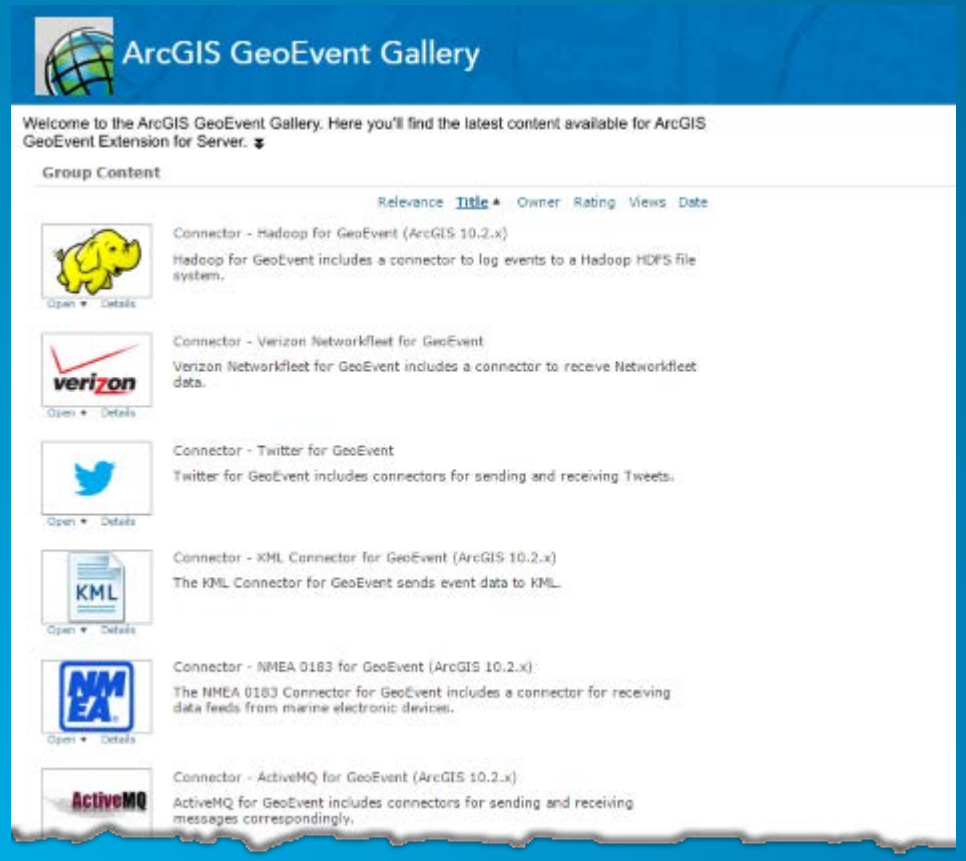
**Configure a new connector using an out-of-the-box adapter / transport**

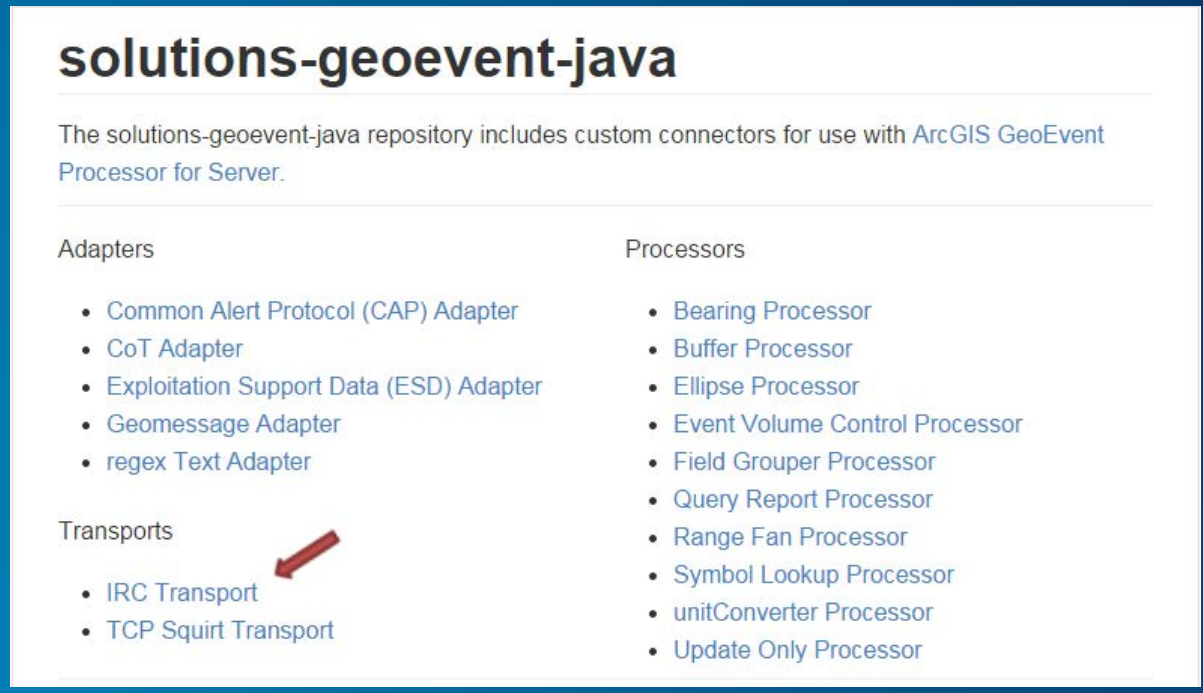


# How do I get my real-time data into ArcGIS?

*Easily integrate real-time streaming data into ArcGIS using an Input Connector*

**Download a connector, adapter, or transport from an Esri repository**



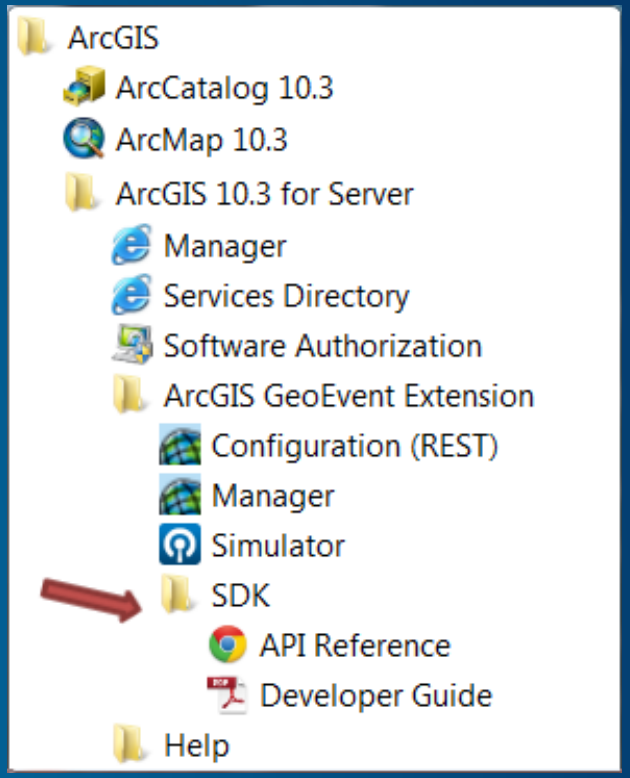




- **http://links.esri.com/geoevent-gallery**
- **https://github.com/Esri/solutions-geoevent-java**

# How do I get my real-time data into ArcGIS?

*Easily integrate real-time streaming data into ArcGIS using an Input Connector*

- Use an existing inbound connector available out-of-the-box
- Copy and customize an out-of-the-box connector to meet your needs
- Configure a new connector using available adapters and/or transports
- Download connectors, adapters, and/or transports from Esri Galleries and GitHub repositories
  - Source code is available from GitHub for customization or to use as a reference
- Develop your own adapter / transport using the GeoEvent SDK
  - The SDK and a Developer's Guide are included in the GeoEvent extension product's installation folder

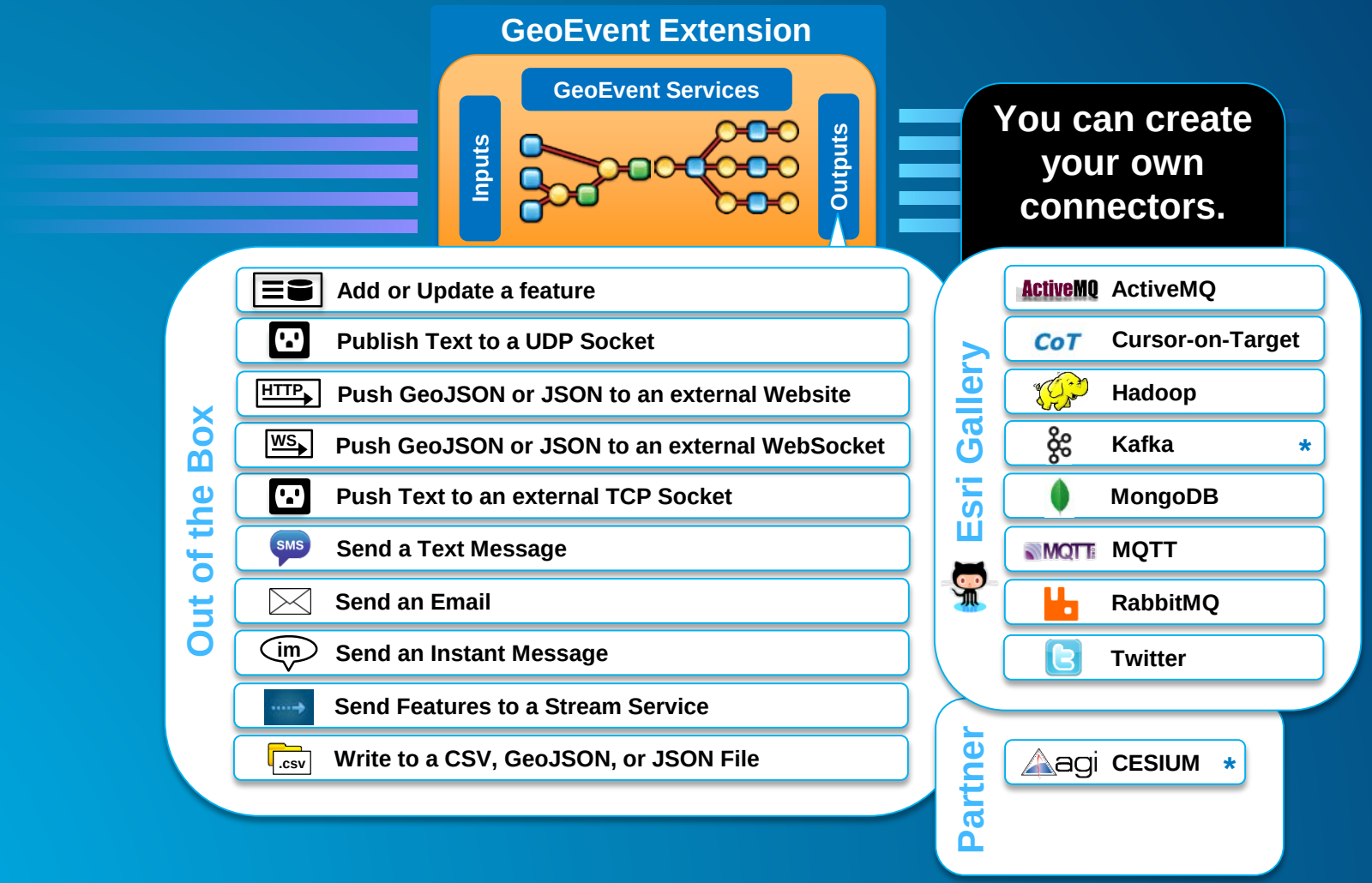


# Best Practice:

## How do I update & alert those who need it where they need it?

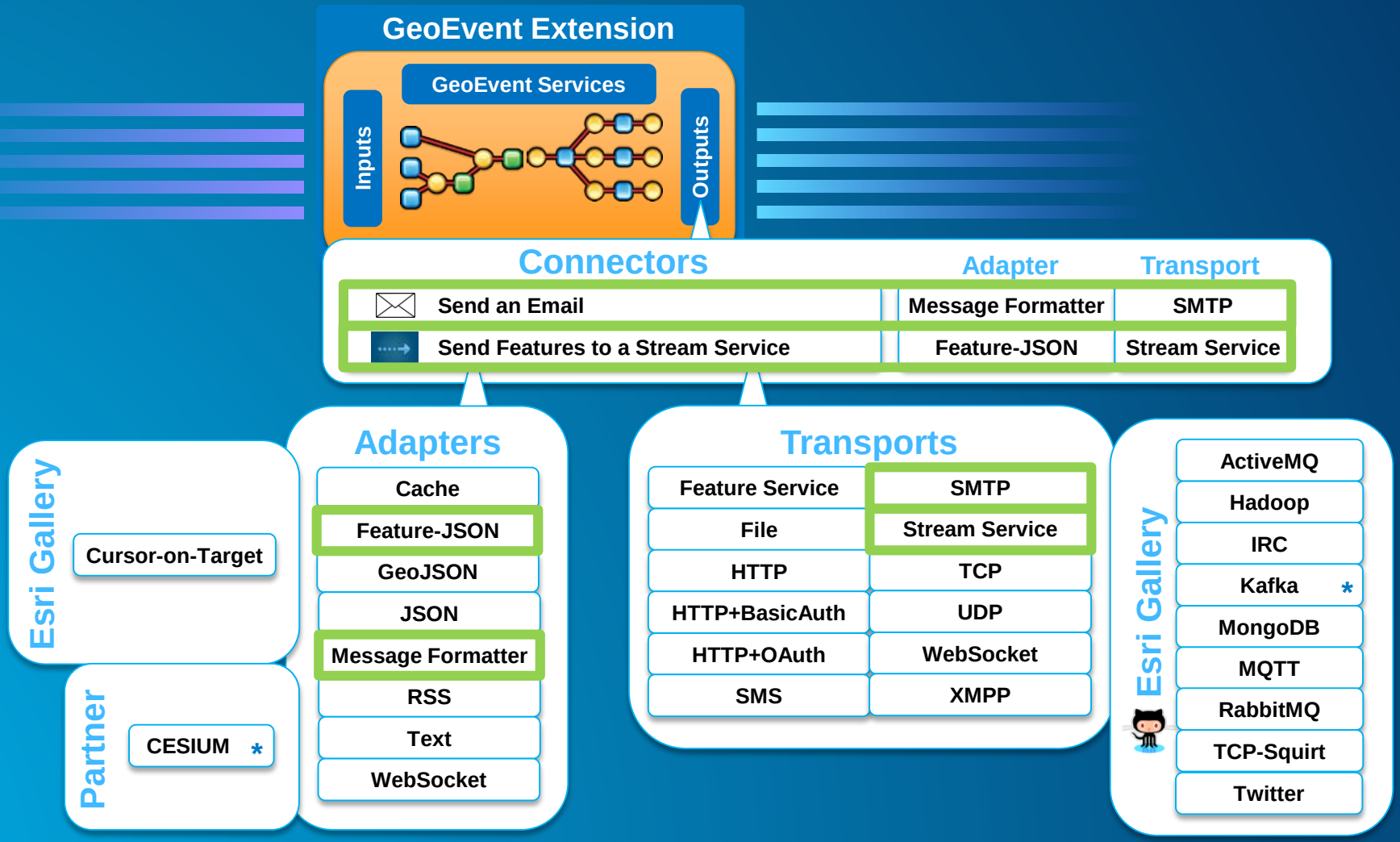
# How do I update and alert those who need it where they need it?

*Easily disseminate notifications, alerts, and updates using an Output Connector*



# How do I update and alert those who need it where they need it?

*Output Connector = Adapter + Transport*



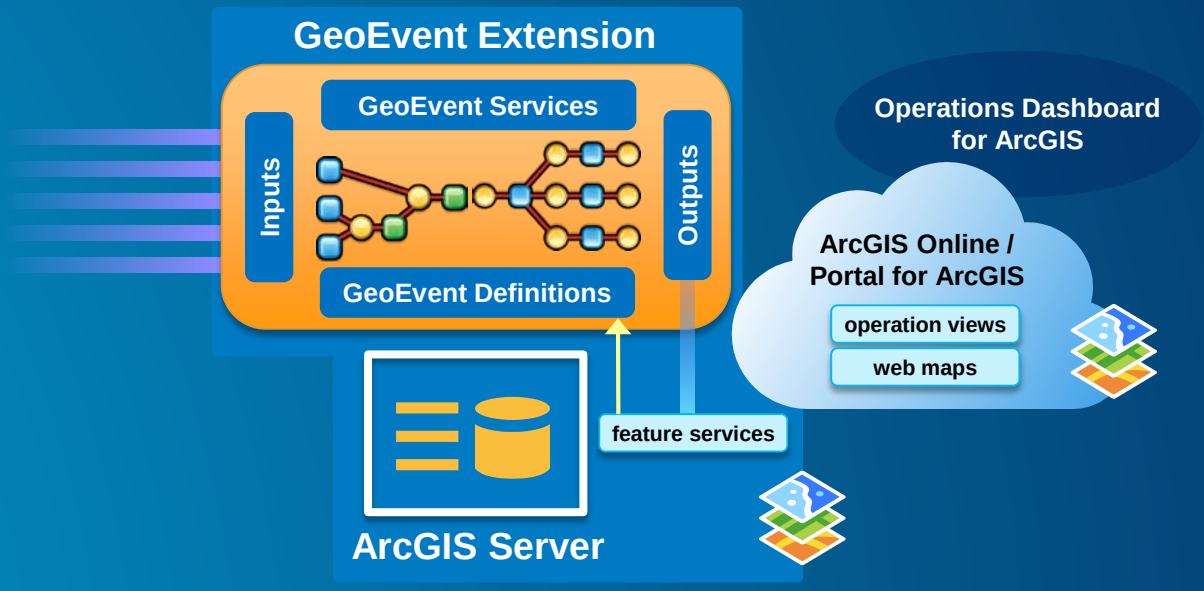
# Best Practice:

## When and why are GeoEvent Definitions created?

# When and why are GeoEvent Definitions created?

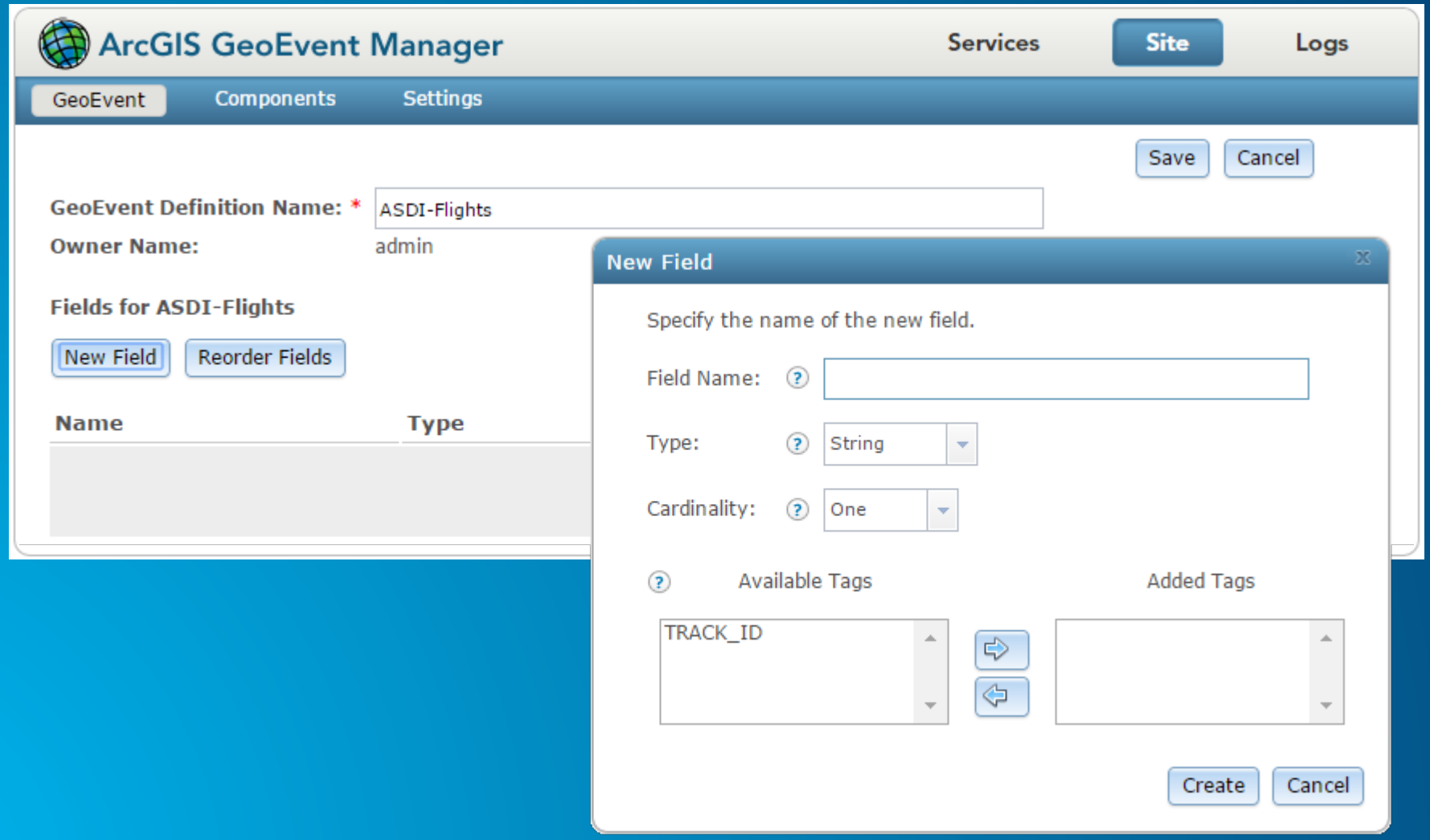
*Authoring GeoEvent Definitions – Importing from a feature service*

- Connect an output to your feature
- Import the schema of your feature as a GeoEvent Definition
- Configure an input to receive real-time data
- Author and publish a GeoEvent Service
- Visualize your real-time feature



# When and why are GeoEvent Definitions created?

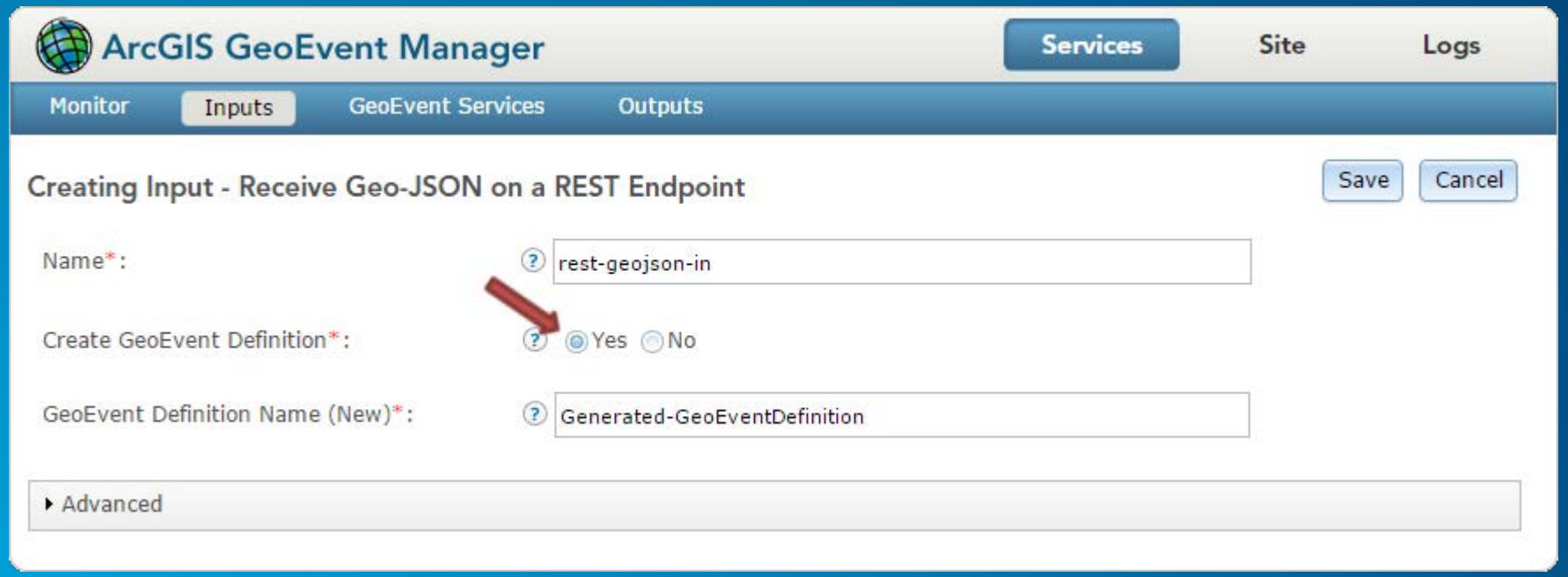
*Authoring GeoEvent Definitions – Creating using GeoEvent Manager*




You own GeoEvent Definitions you either import or create manually using the GeoEvent Manager

# When and why are GeoEvent Definitions created?

*Auto-generated GeoEvent Definitions – Created by an inbound connector*




- The adapter will make a best-guess based on data discovered in the first event received
- You can copy and tailor auto-generated GeoEvent Definitions to meet your needs
  - But you should not edit an event definition owned or created by an adapter

# When and why are GeoEvent Definitions created?

*Managed GeoEvent Definitions – Created by a processor from within a GeoEvent Service*

- **A *managed* GeoEvent Definition is created when a processor adds or removes a field from a GeoEvent**
  - **Field Calculator**
  - **Field Enricher**
  - **Geotagger**
  - **Buffer Creator**

  (etc…)
- **Publishing changes to a GeoEvent Service signals the owning processor to delete its *managed* GeoEvent Definition**
- **This can present you with issues you need to work around…**
  - **Field Mapping**
  - **Stream Services**
  - **Custom Processors**

**Be aware of the GeoEvent Definition associated with your GeoEvent(s) at every stage of your event processing**

*For more information see 'Understanding GeoEvent Definitions' blog:*
http://tinyurl.com/o83wjcv

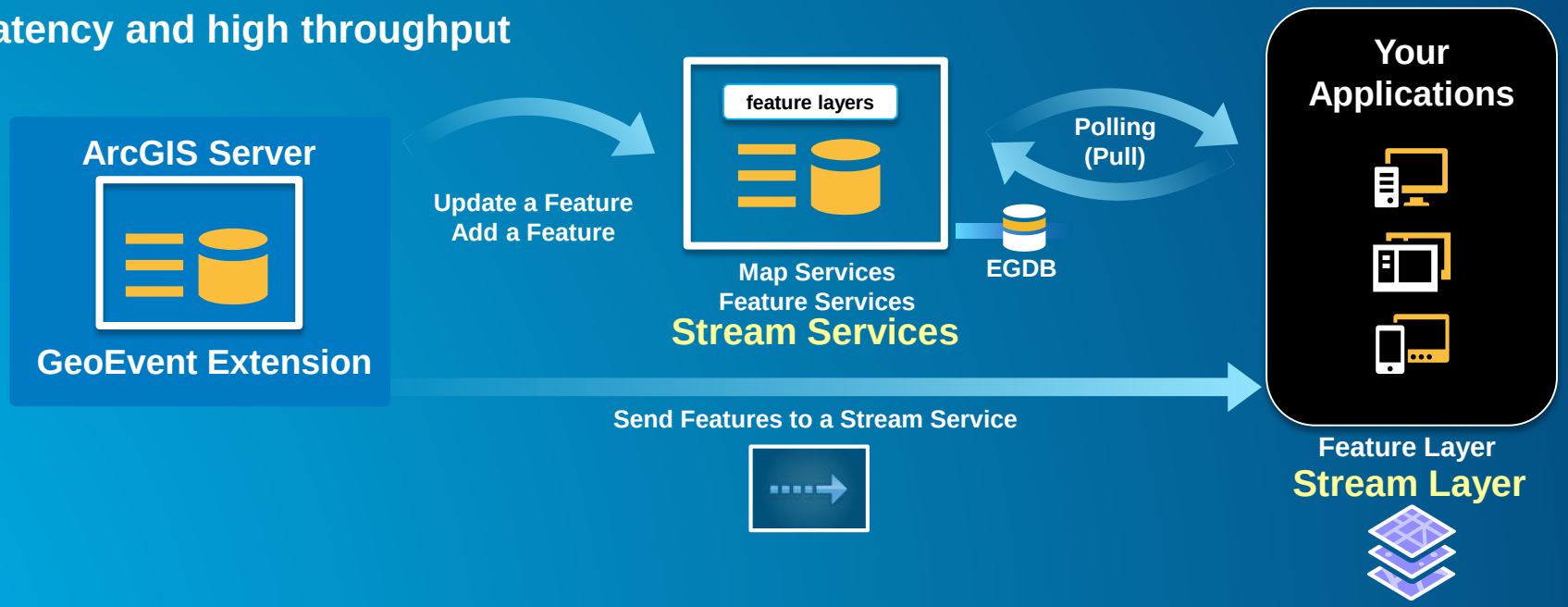
# Best Practice:

## When should I use a Feature Service vs. a Stream Service?

# When should I use Feature Services vs. Stream Services?

*Two patterns*

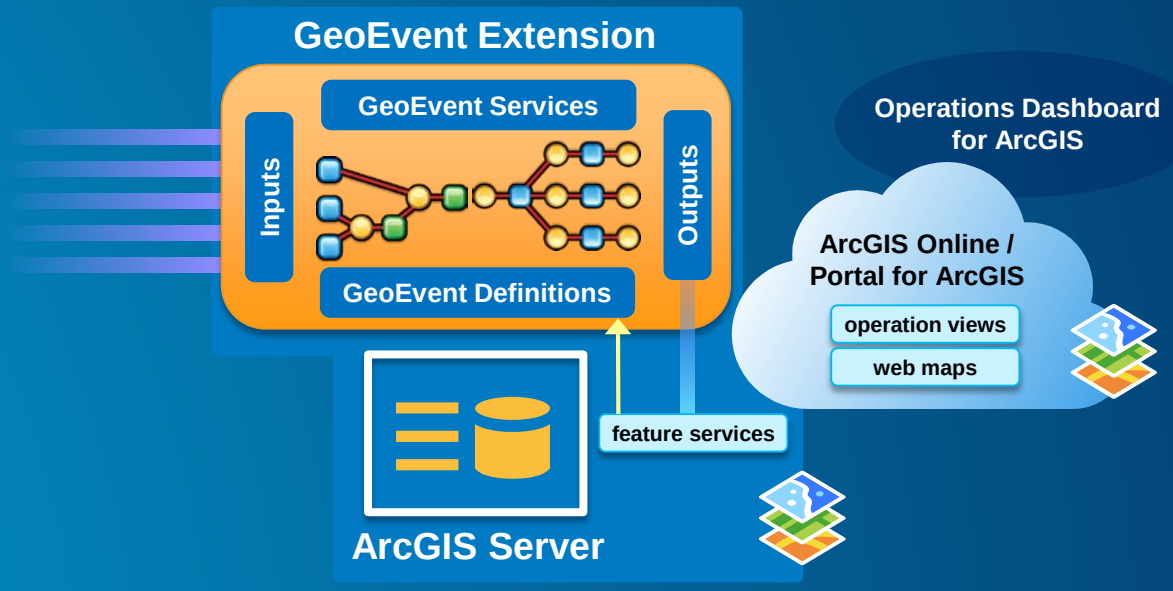
- Feature layers **pull** from feature services
  - Web apps poll to get periodic updates
  - Must be backed by an enterprise geodatabase (EGDB)
- Stream layers **subscribe** to stream services
  - Web apps subscribe to immediately receive data
  - Low latency and high throughput



# What is the recommended workflow for working with Feature Services?

*Feature Service workflow*

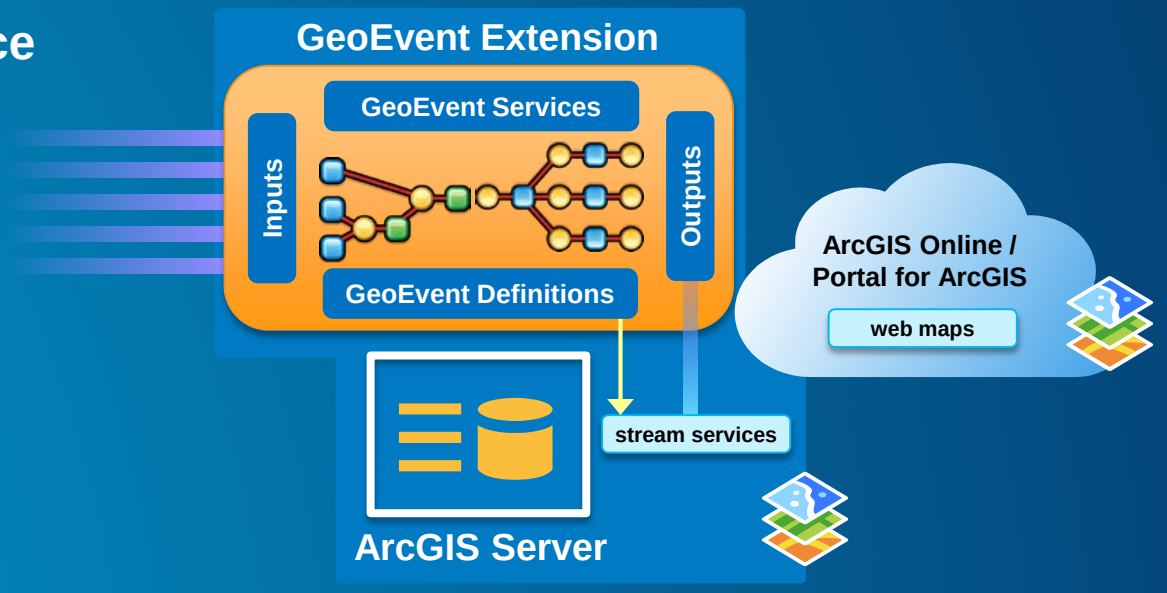
- Connect an output to your feature
- Import the schema of your feature as a GeoEvent Definition
- Configure an input to receive real-time data
- Author and publish a GeoEvent Service
- Visualize your real-time feature



# What is the recommended workflow for working with Stream Services?

*Stream Service workflow*

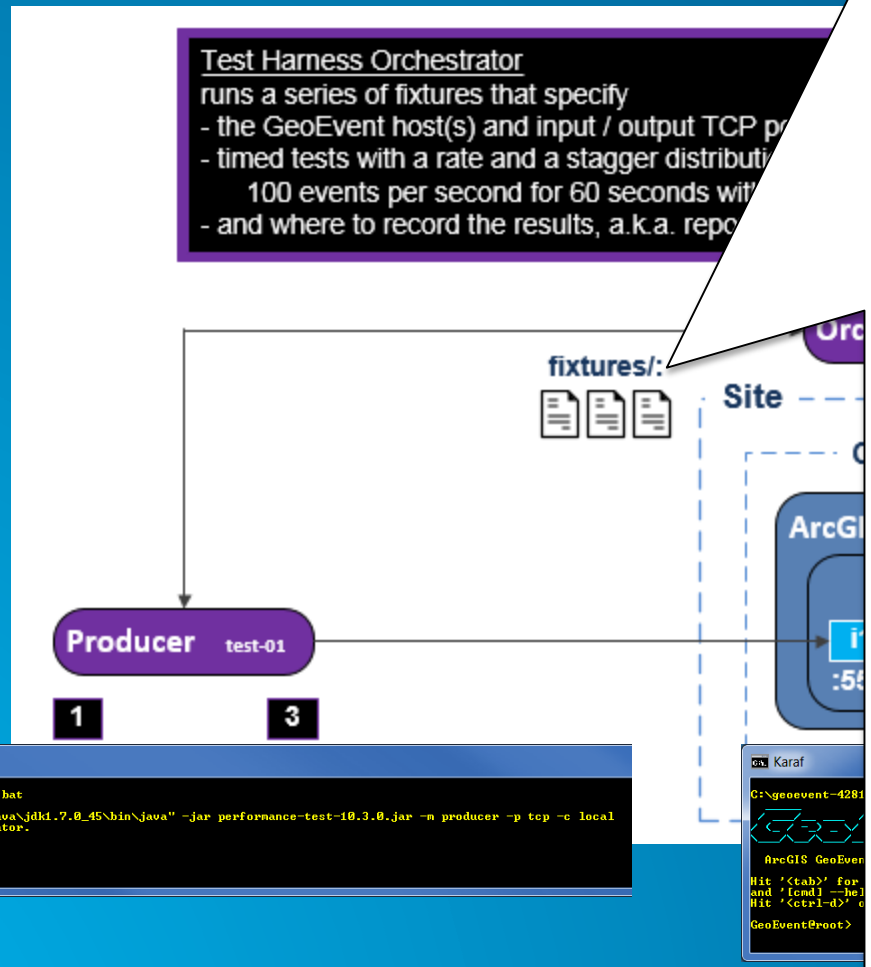
- Configure an input to receive real-time data
- Define a GeoEvent Definition
- When creating an output publish a Stream Service and connect the output to it
- Author and publish a GeoEvent Service
- Visualize your real-time stream service
  - Use an ArcGIS Online WebMap
  - Use a custom JavaScript web app



# Performance

# How do I measure Performance?

*performance-test-harness-for-geoevent*

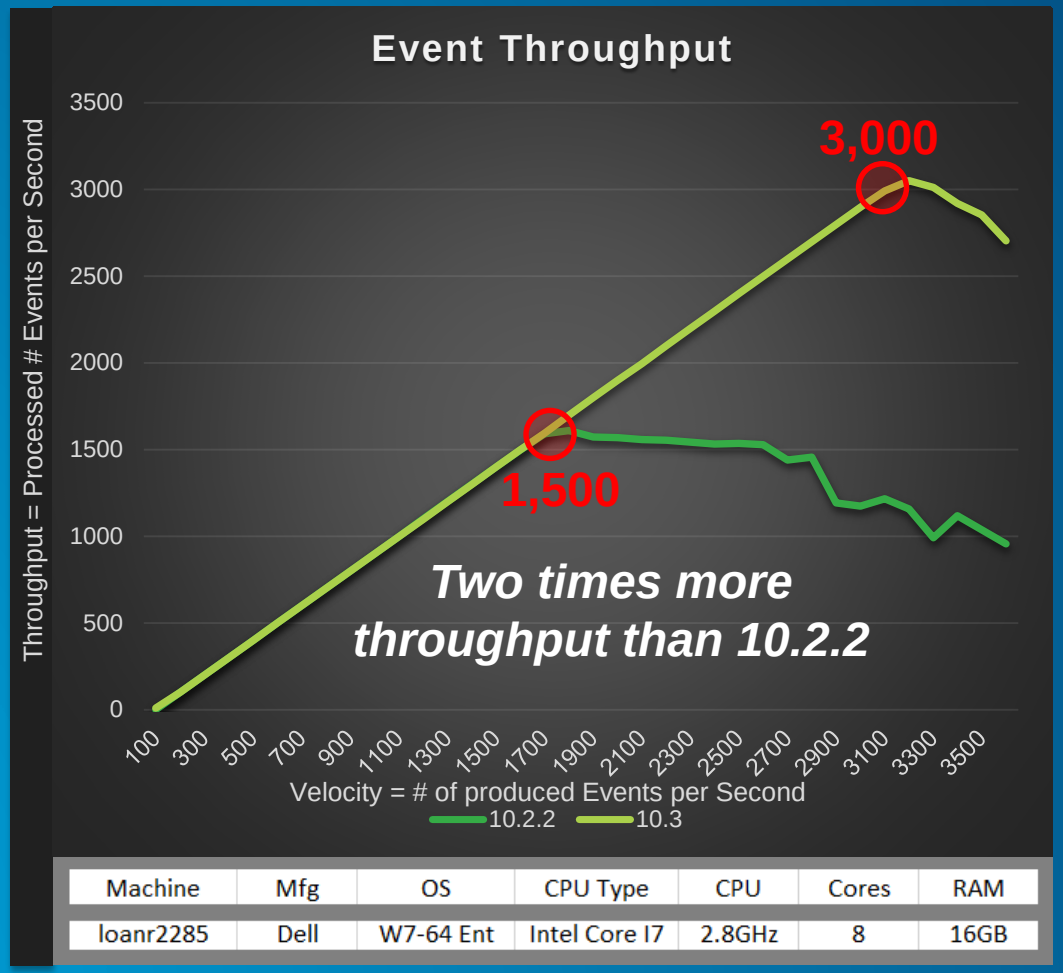



```
<Fixtures>
    <Report type="CSV">
        <ReportFile>reports/throughput-1000to3000step1000.csv</ReportFile>
    </Report>

    <DefaultFixture>
        <DefaultSharedConfig protocol="TCP">  <!-- also supports WS (Web Socket), RMQ (RabbitMQ)-->
            <Properties>
                <Property name="hosts">host1</Property>
            </Properties>
        </DefaultSharedConfig>

        <ProducerConfig commandPort="5010">
            <Properties>
                <Property name="port">5565</Property>
            </Properties>
        </ProducerConfig>

        <ConsumerConfig commandPort="5020" timeoutInSec="5">
            <Properties>
                <Property name="port">5575</Property>
            </Properties>
        </ConsumerConfig>

        <Simulation>
            <TimeTest eventsPerSec="1000" totalTimeInSec="30" staggeringInterval="10" />
        </Simulation>
    </DefaultFixture>

    <Fixture name="1000 e/s">
        <ProducerConfig>
            <Properties>
                <Property name="simulationFile">simulations/county_envelopes_points_1000.csv</Property>
            </Properties>
        </ProducerConfig>
        <Simulation>
            <TimeTest eventsPerSec="1000"/>
        </Simulation>
    </Fixture>

    ...

    <Fixture name="3000 e/s">
        <ProducerConfig>
            <Properties>
                <Property name="simulationFile">simulations/county_envelopes_points_3000.csv</Property>
            </Properties>
        </ProducerConfig>
        <Simulation>
            <TimeTest eventsPerSec="3000"/>
        </Simulation>
    </Fixture>
</Fixtures>
```

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

# Throughput Performance Benchmark @ 10.3

*Two times more throughput than 10.2.2*

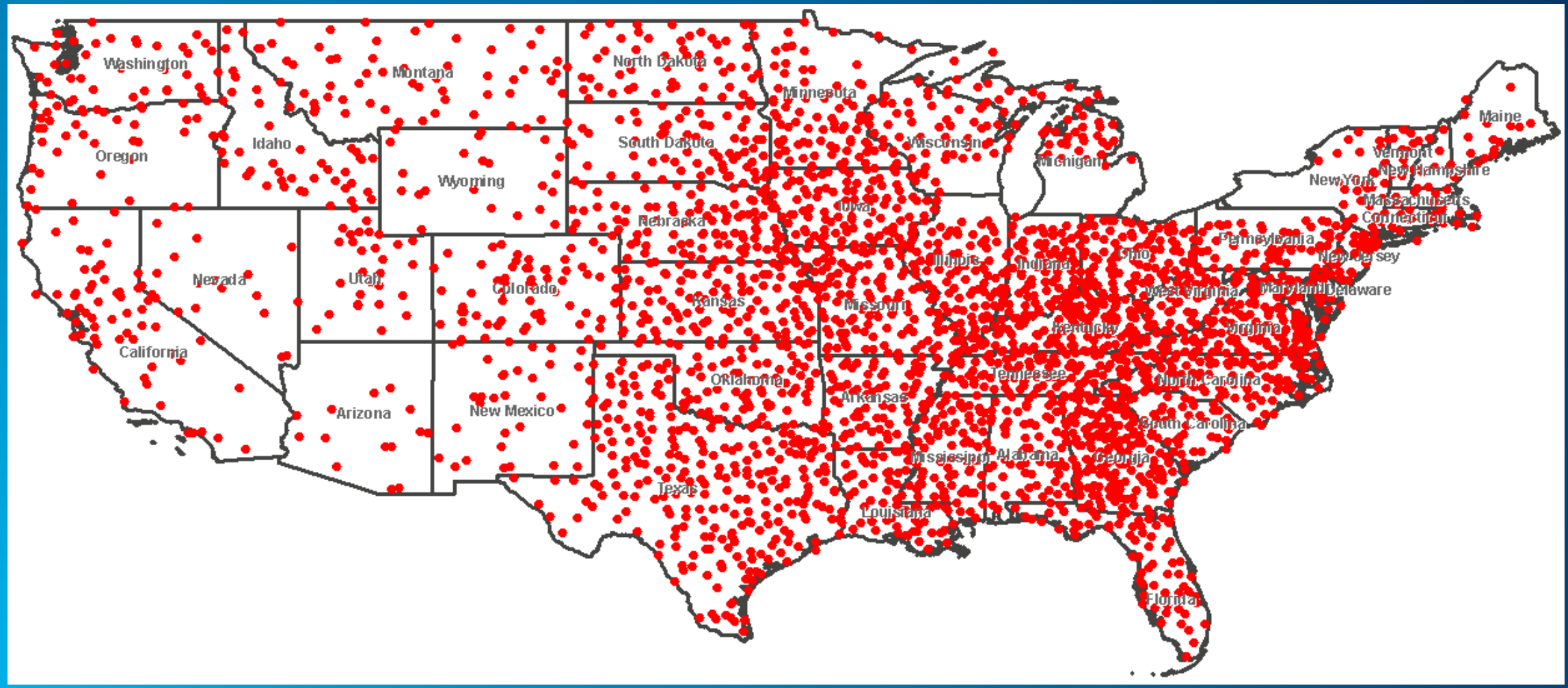



| Machine | Mfg | OS | CPU Type | CPU | Cores | RAM |
|---|---|---|---|---|---|---|
| loanr2285 | Dell | W7-64 Ent | Intel Core I7 | 2.8GHz | 8 | 16GB |

As captured on primary benchmarking machine using ArcGIS 10.3

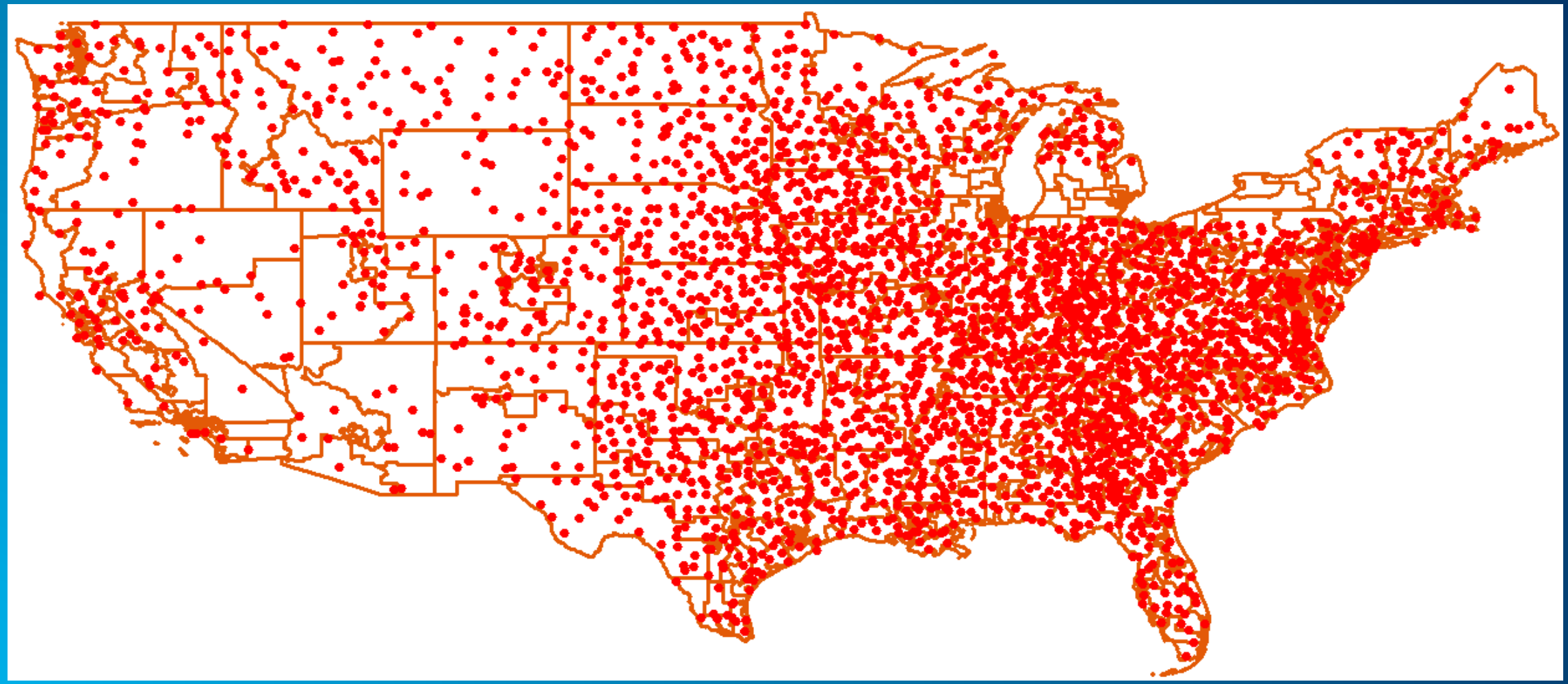
# Geofencing Performance Benchmark @ 10.3

*US States benchmark – 51 geofences with 1,617 vertices on average (78 min / 21,970 max)*



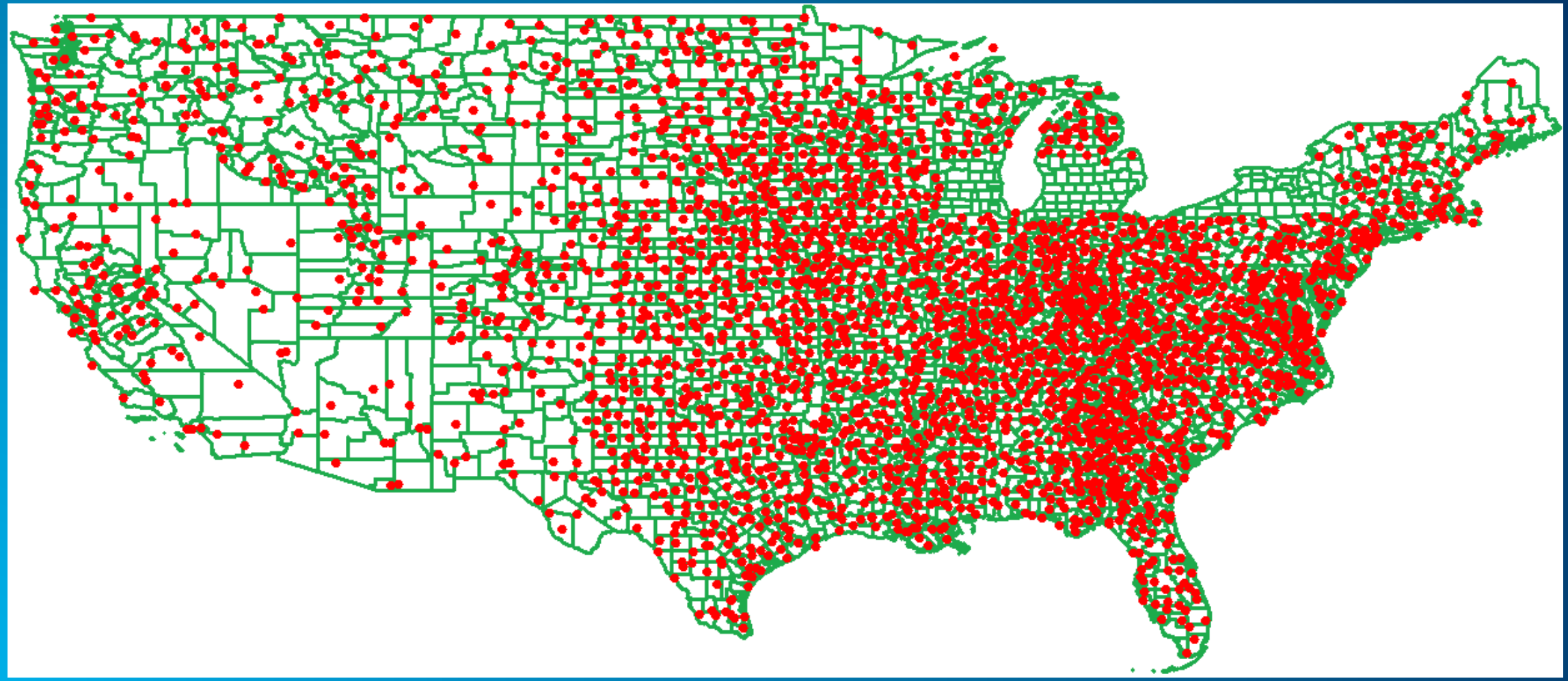
# Geofencing Performance Benchmark @ 10.3

*US Congressional Districts benchmark – 436 geofences with 512 vertices on average (24 min / 7,285 max)*



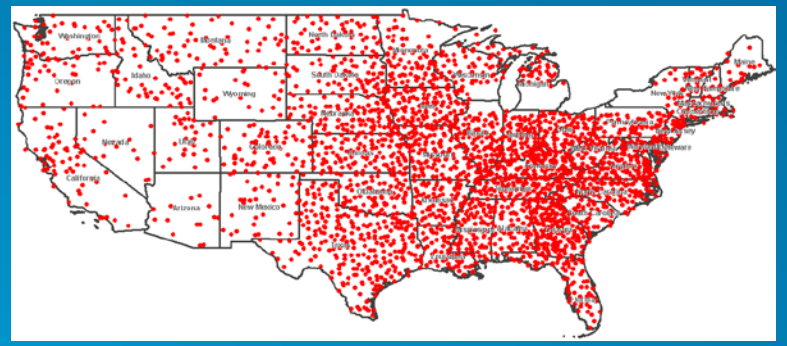
# Geofencing Performance Benchmark @ 10.3

*US Counties benchmark = 3,143 geofences with 166 vertices on average (9 min / 838 max)*

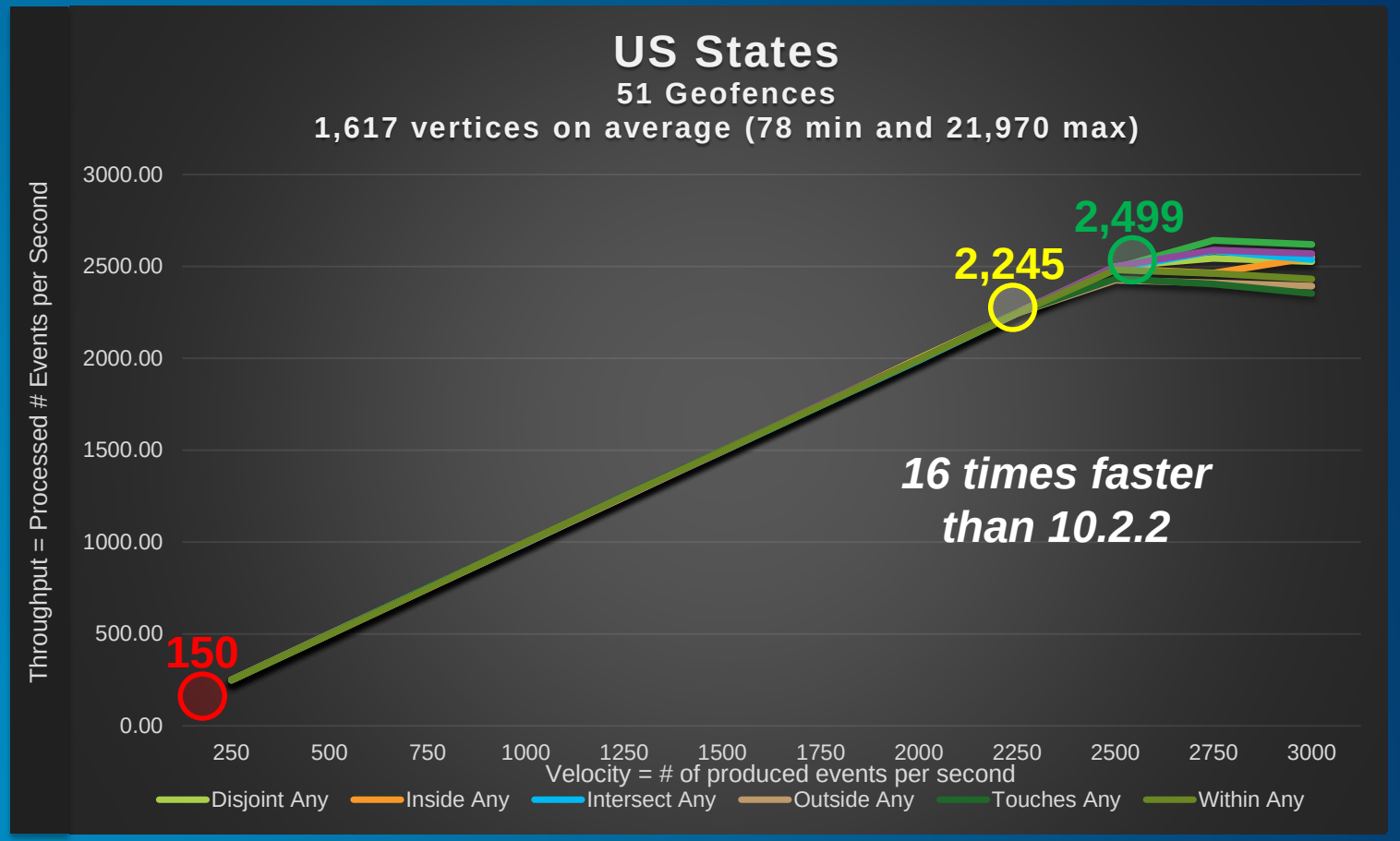


# Geofencing Performance Benchmark @ 10.3

*US States benchmark*



| Operator | 10.3 events per second | 10.2.2 events per second |
|---|---|---|
| Disjoint Any | 2,499 | |
| Inside Any | 2,488 | 150 |
| Intersect Any | 2,486 | |
| Within Any | 2,482 | |
| Touches Any | 2,248 | |
| Outside Any | 2,245 | |

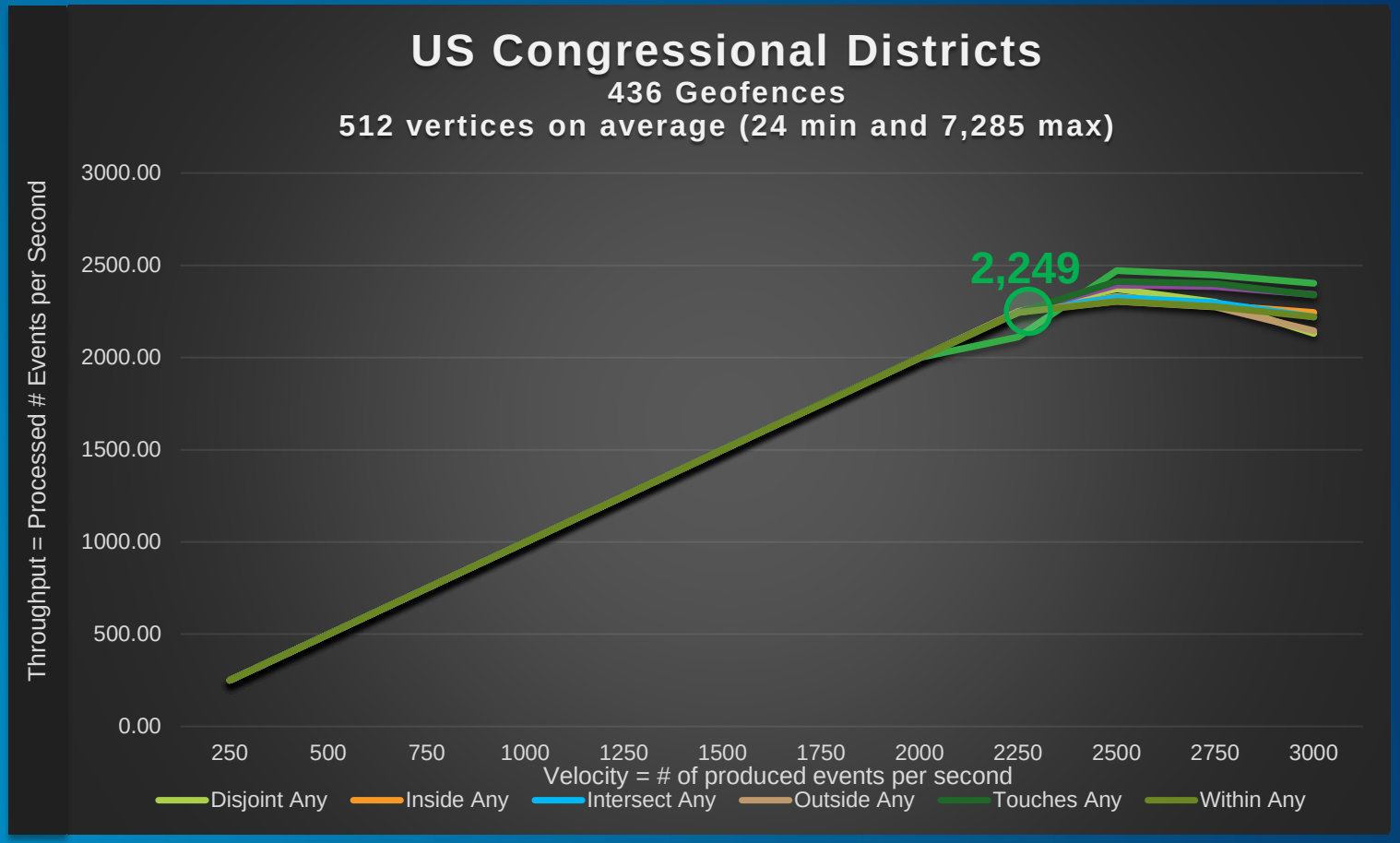



As captured on primary benchmarking machine using ArcGIS 10.3

# Geofencing Performance Benchmark @ 10.3

*US Congressional Districts benchmark*

| Operator | 10.3 events per second |
|---|---|
| Disjoint Any | 2,249 |
| Outside Any | 2,248 |
| Intersect Any | 2,248 |
| Touches Any | 2,244 |
| Within Any | 2,244 |
| Inside Any | 2,244 |

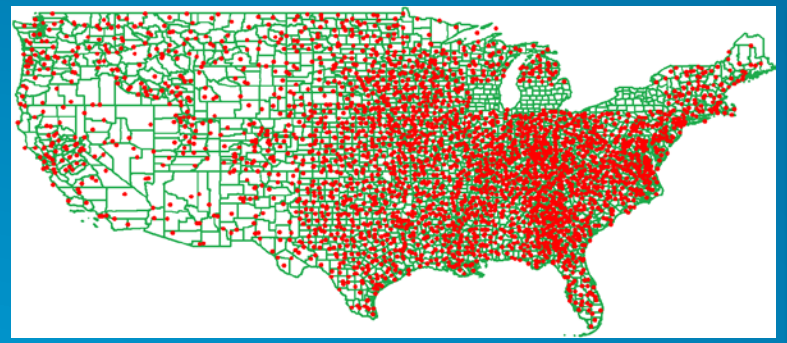



As captured on primary benchmarking machine using ArcGIS 10.3

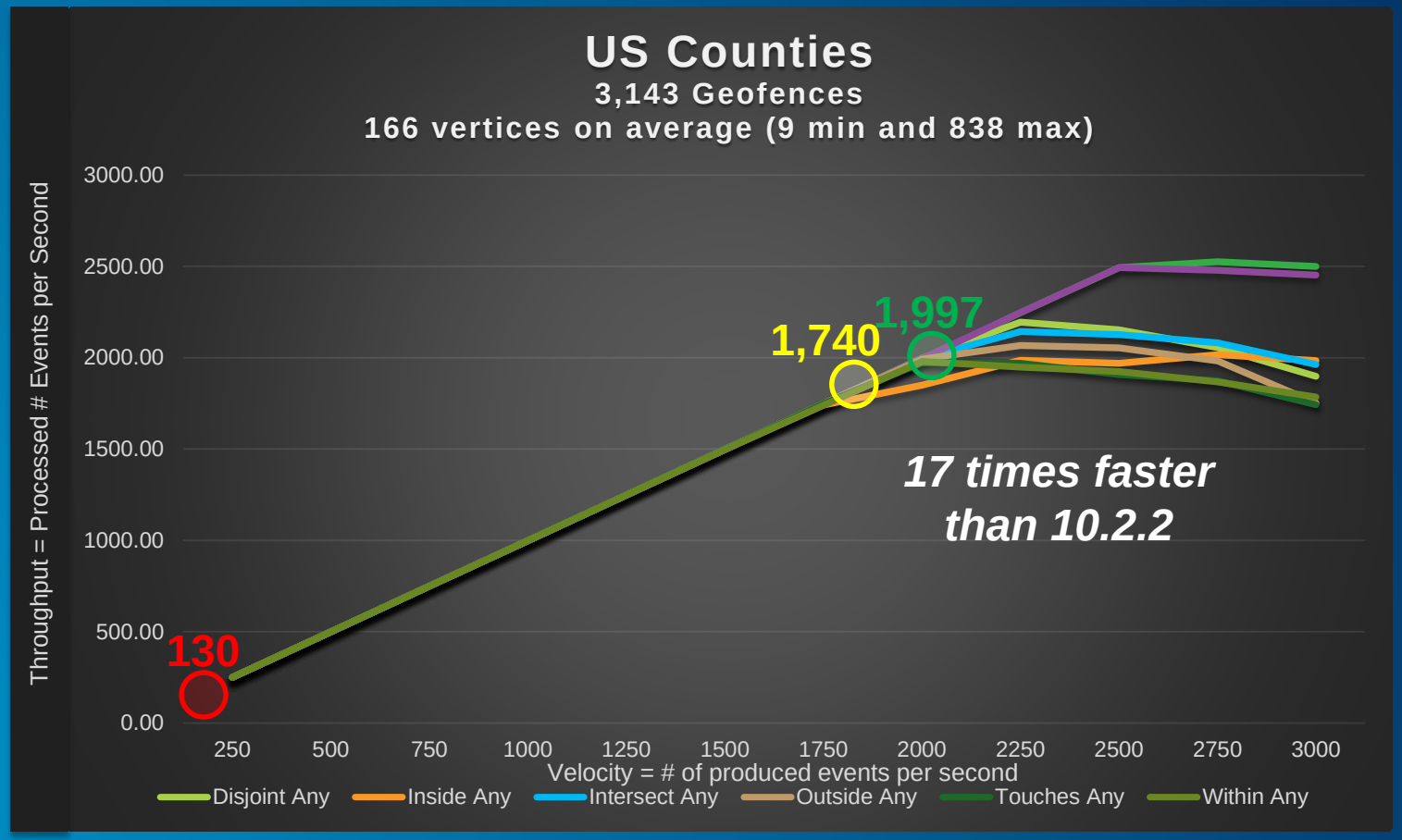
*note: this scenario was not benchmarked at 10.2.2*

# Geofencing Performance Benchmark @ 10.3

*US Counties benchmark*



| Operator | 10.3 events per second | 10.2.2 events per second |
|---|---|---|
| Disjoint Any | 1,997 | |
| Intersect Any | 1,996 | |
| Outside Any | 1,992 | |
| Touches Any | 1,747 | |
| Inside Any | 1,742 | 130 |
| Within Any | 1,740 | |




As captured on primary benchmarking machine using ArcGIS 10.3

# What are the primary factors I should consider for Performance?

*Operating environment, Network, RAM, and Processors*

- **Operating environment:** *Bare Metal = Windows 7 (64 bit) Enterprise*
  - Virtual Machines – beware!  resources need to be shared in an effective way, like EC2.
  - Bare-Metal machines – dedicated resources are much more deterministic.
- **Network**
  - Speed (Mbps) – the faster the better. *1 Gbps*
- **RAM**
  - size (GB) – 8GB is required at 10.3. *16GB, defaulted JVM max heap size of 4GB used*
  - 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*

# High Availability & Scalability

# High Availability & Scalability

*Clustering*

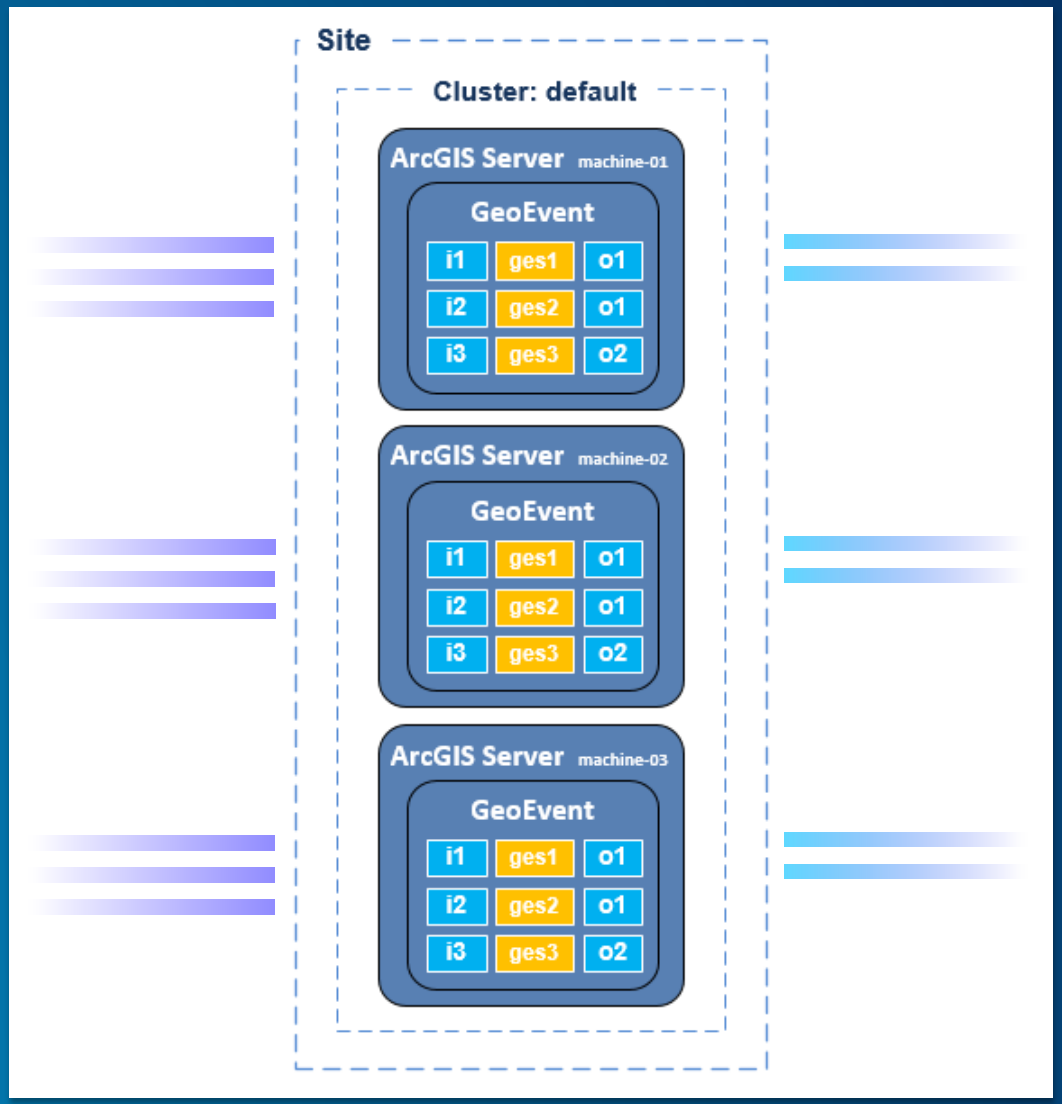
- Clusters administered via ArcGIS Server Manager
  - Site, Cluster(s), Machines
- Machines in a cluster share configuration
  - automatic provisioning upon joining a cluster
  - including custom components
- High Availability is achievable out-of-the-box
- Scale-out by adding machines to a cluster

**Machines**

| Name | Status |
|---|---|
| LOANR2285.ESRI.COM | Started |
| RTRUJILLO.ESRI.COM | Started |
| SOURCE.ESRI.COM | |
| XW8600-W7.ESRI.COM | Stopped |

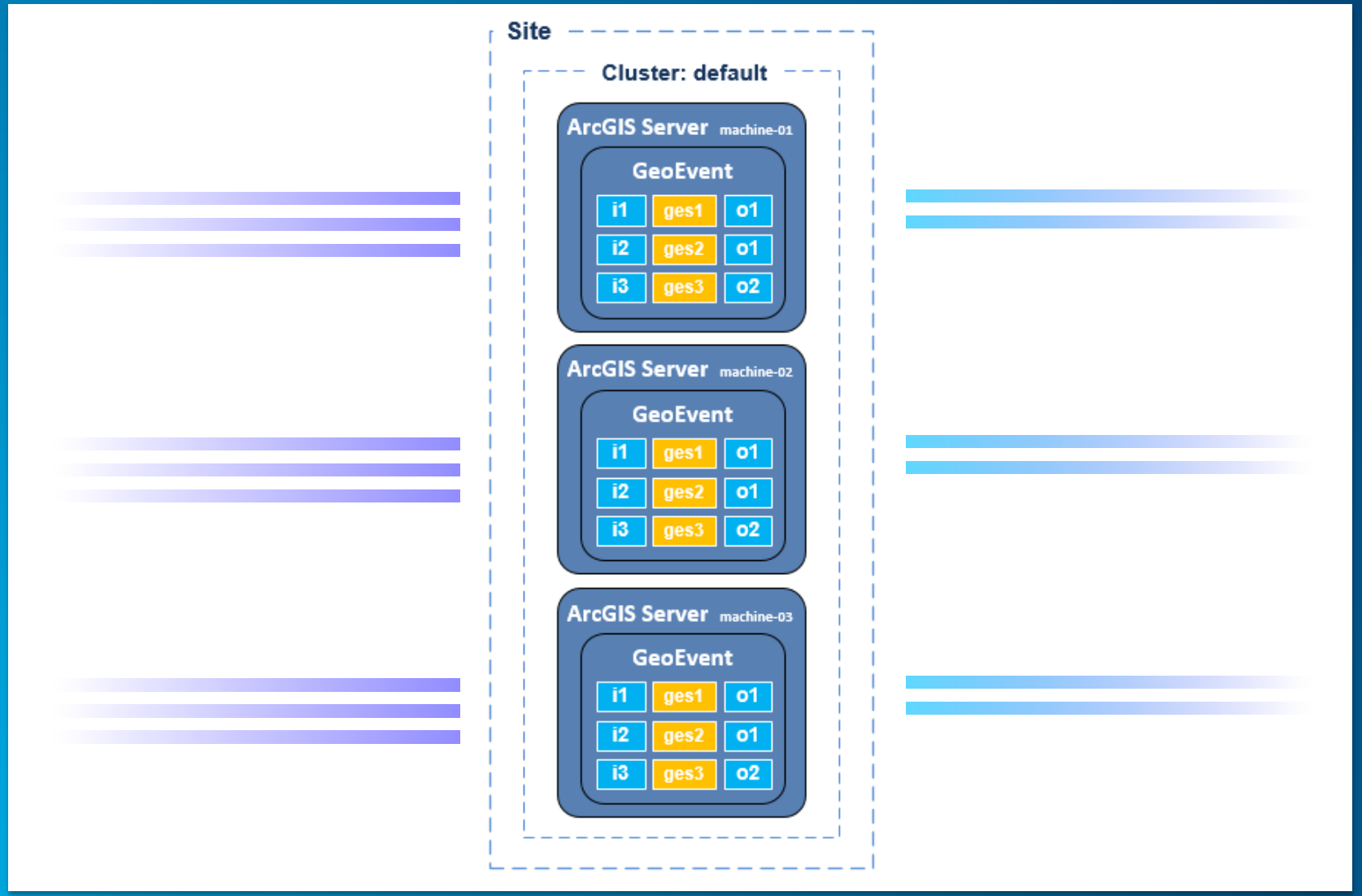
**Clusters**

| Name | Machines | Protocol |
|---|---|---|
| default | SOURCE.ESRI.COM<br>LOANR2285.ESRI.COM<br>RTRUJILLO.ESRI.COM | TCP port 4004 |



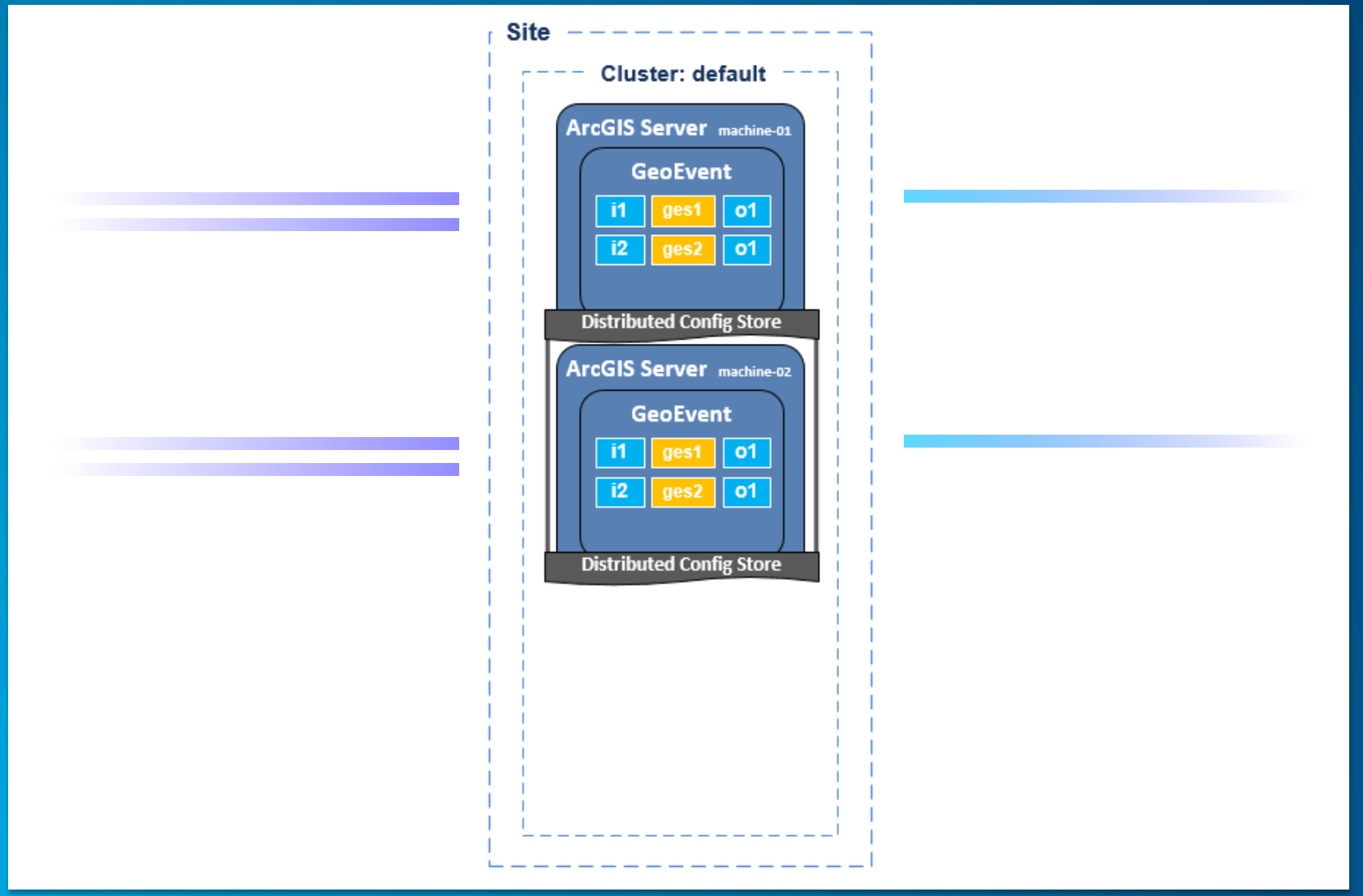
# High Availability

*Site, cluster(s), and machines*



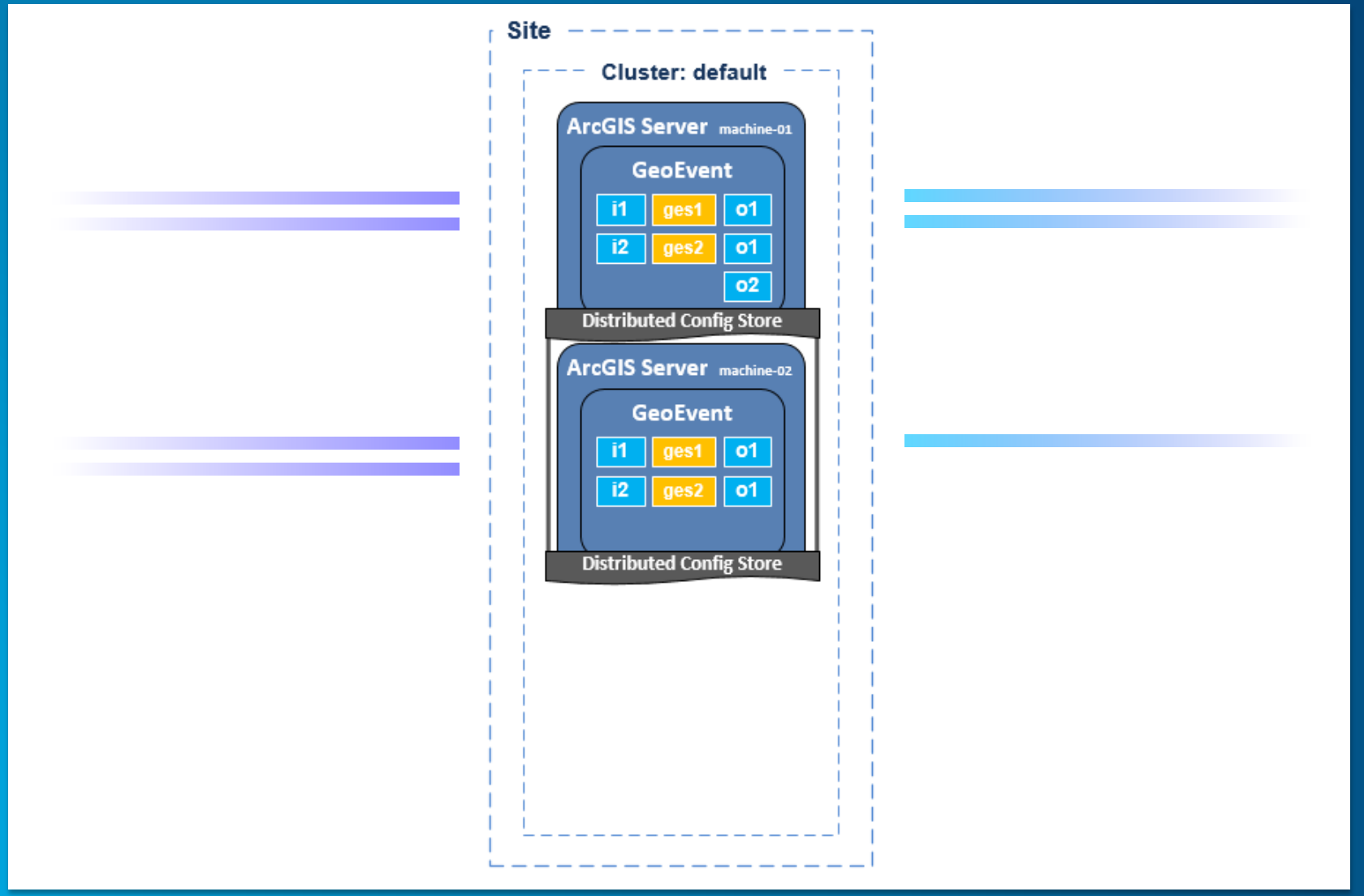
# High Availability

*Distributed configuration store*



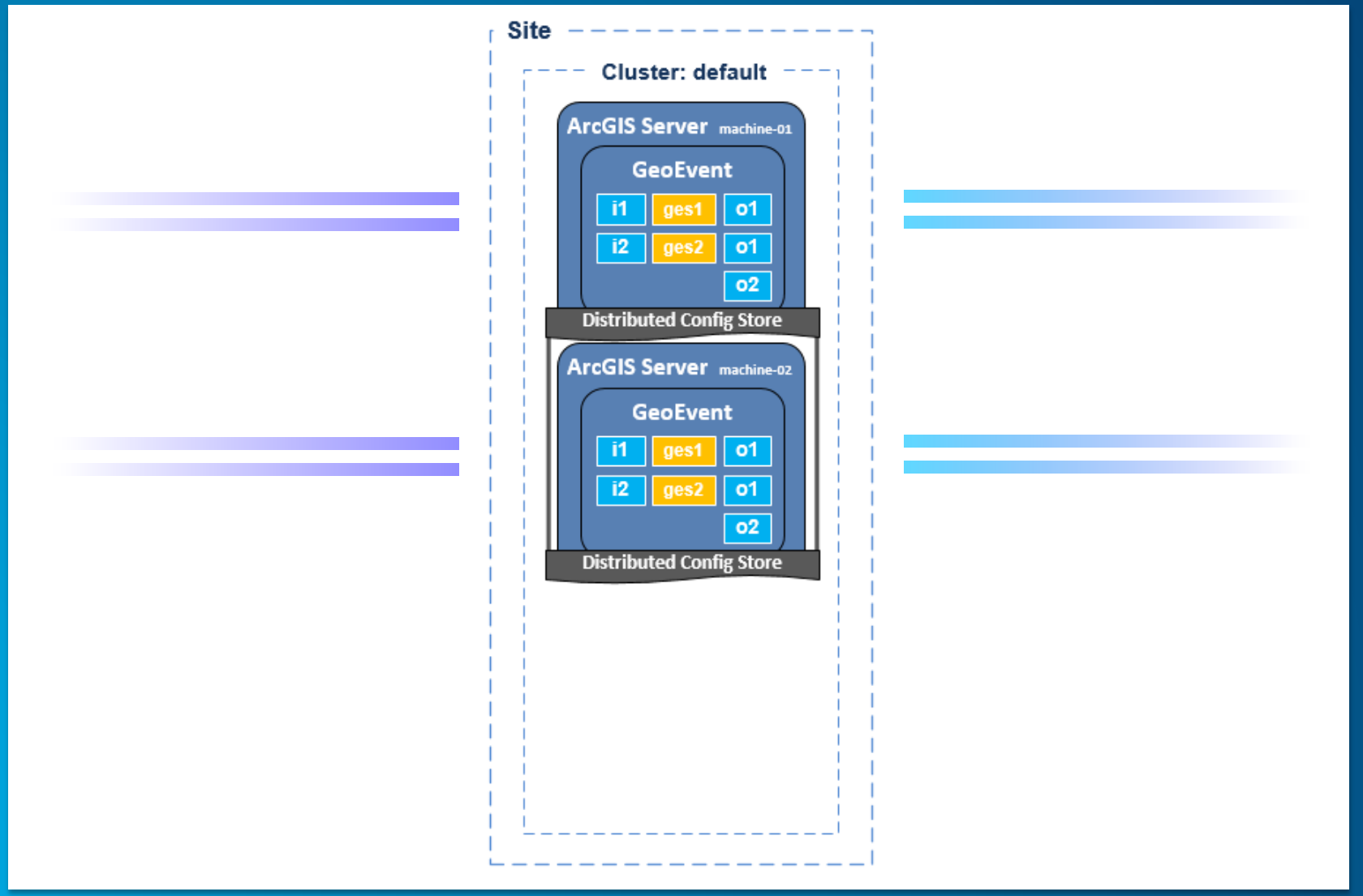
# High Availability

*Distributed configuration store*



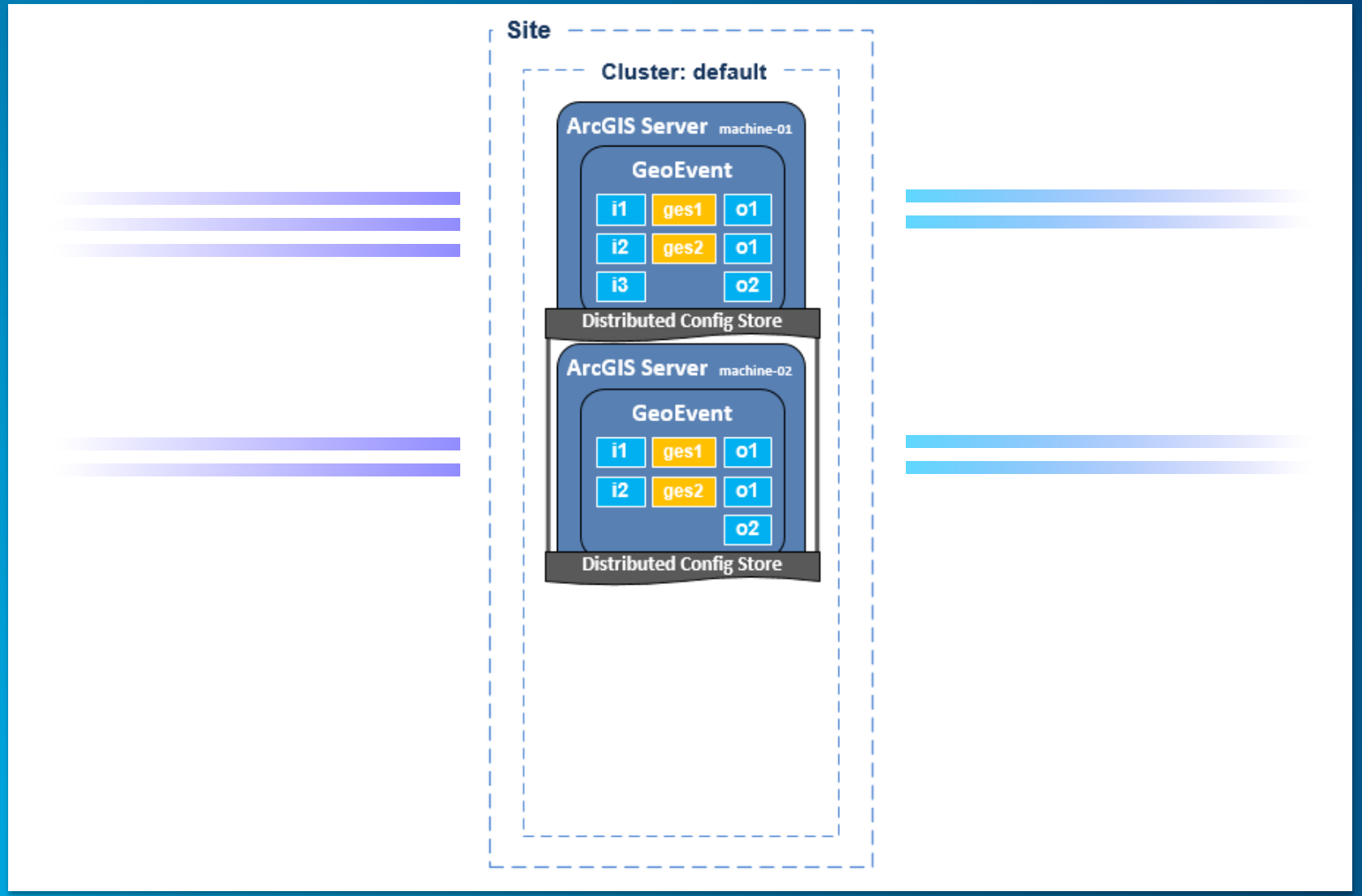
# High Availability

*Distributed configuration store*



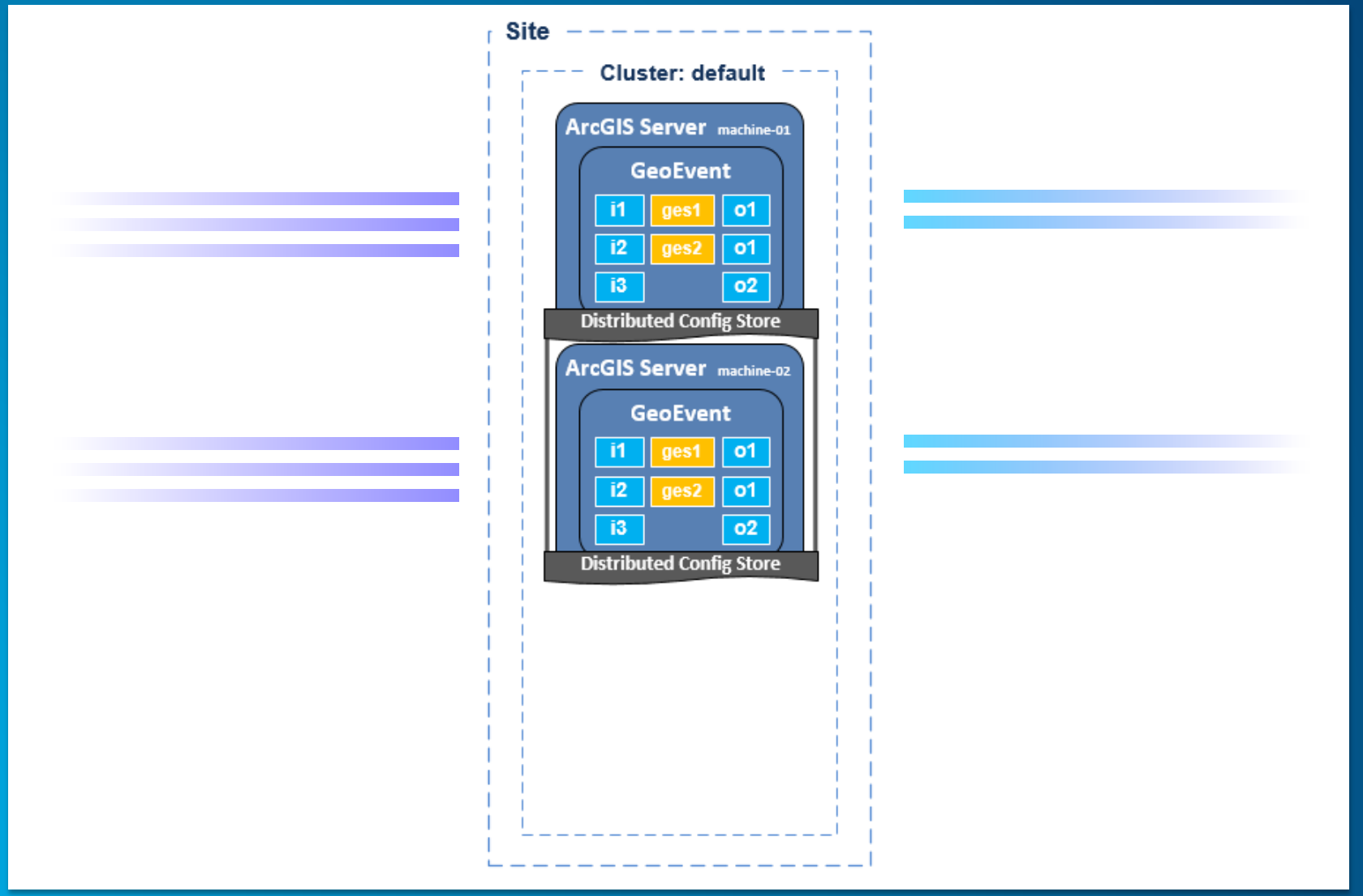
# High Availability

*Distributed configuration store*



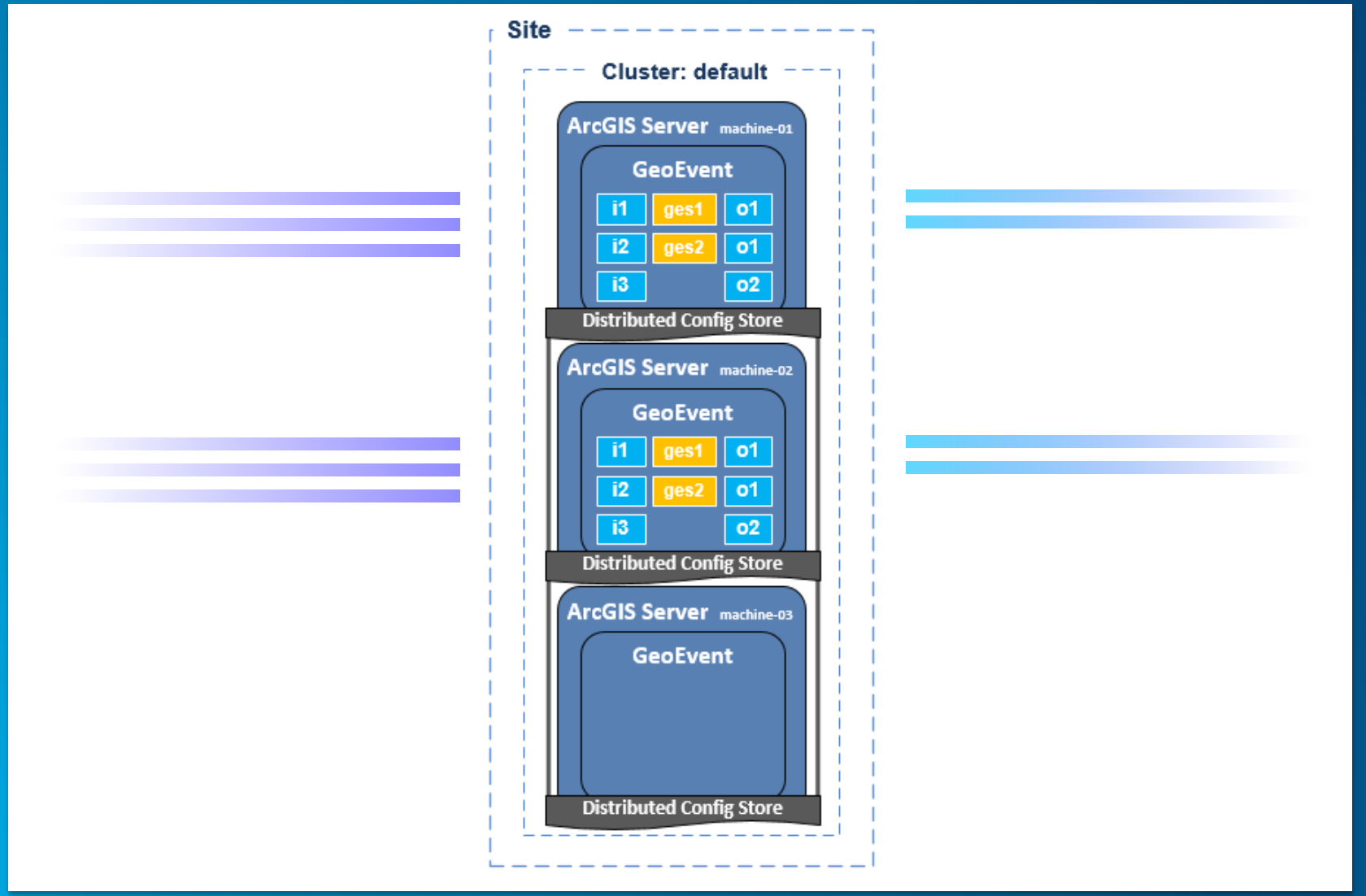
# High Availability

*Distributed configuration store*



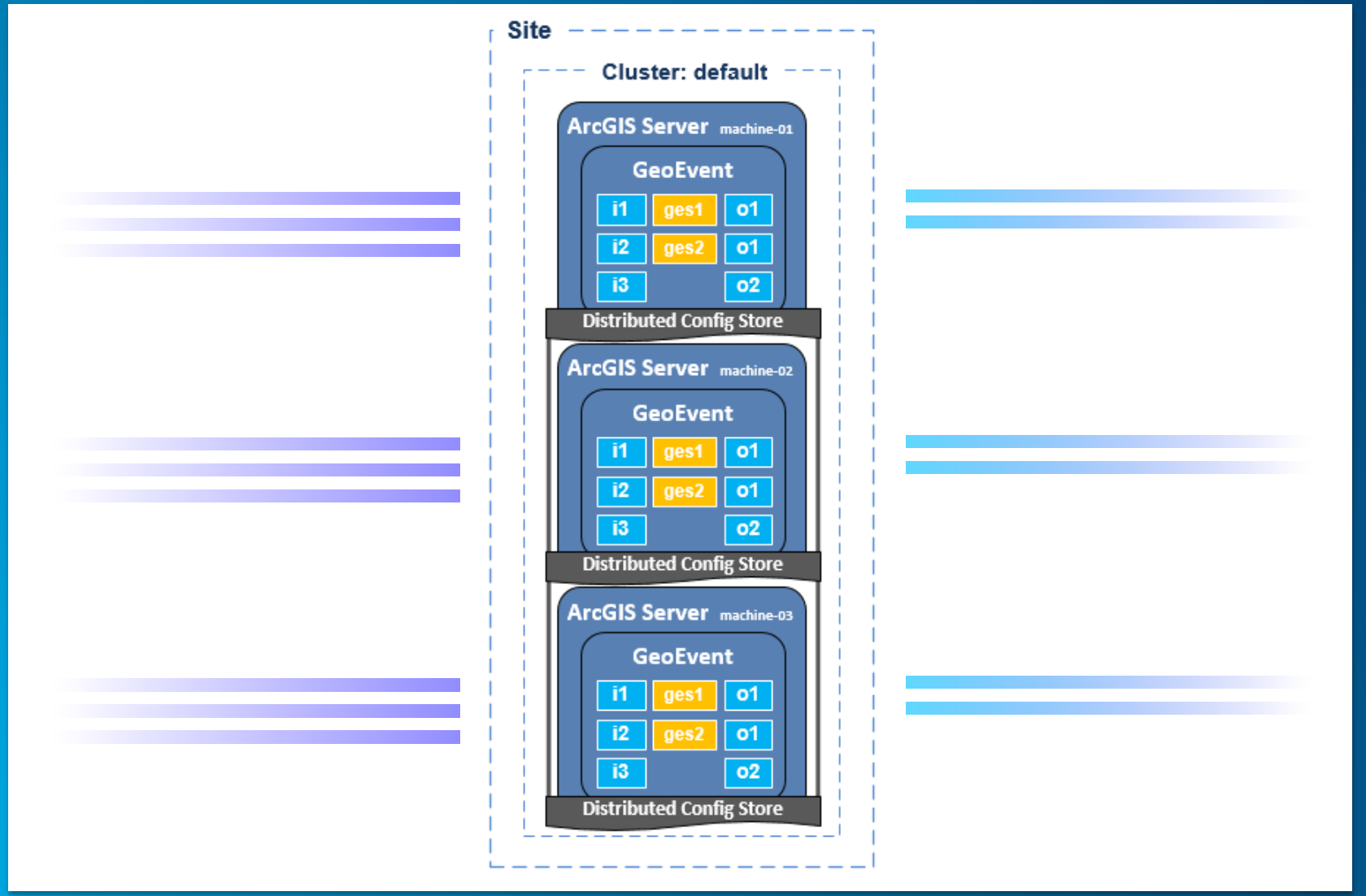
# High Availability

*Distributed configuration store*



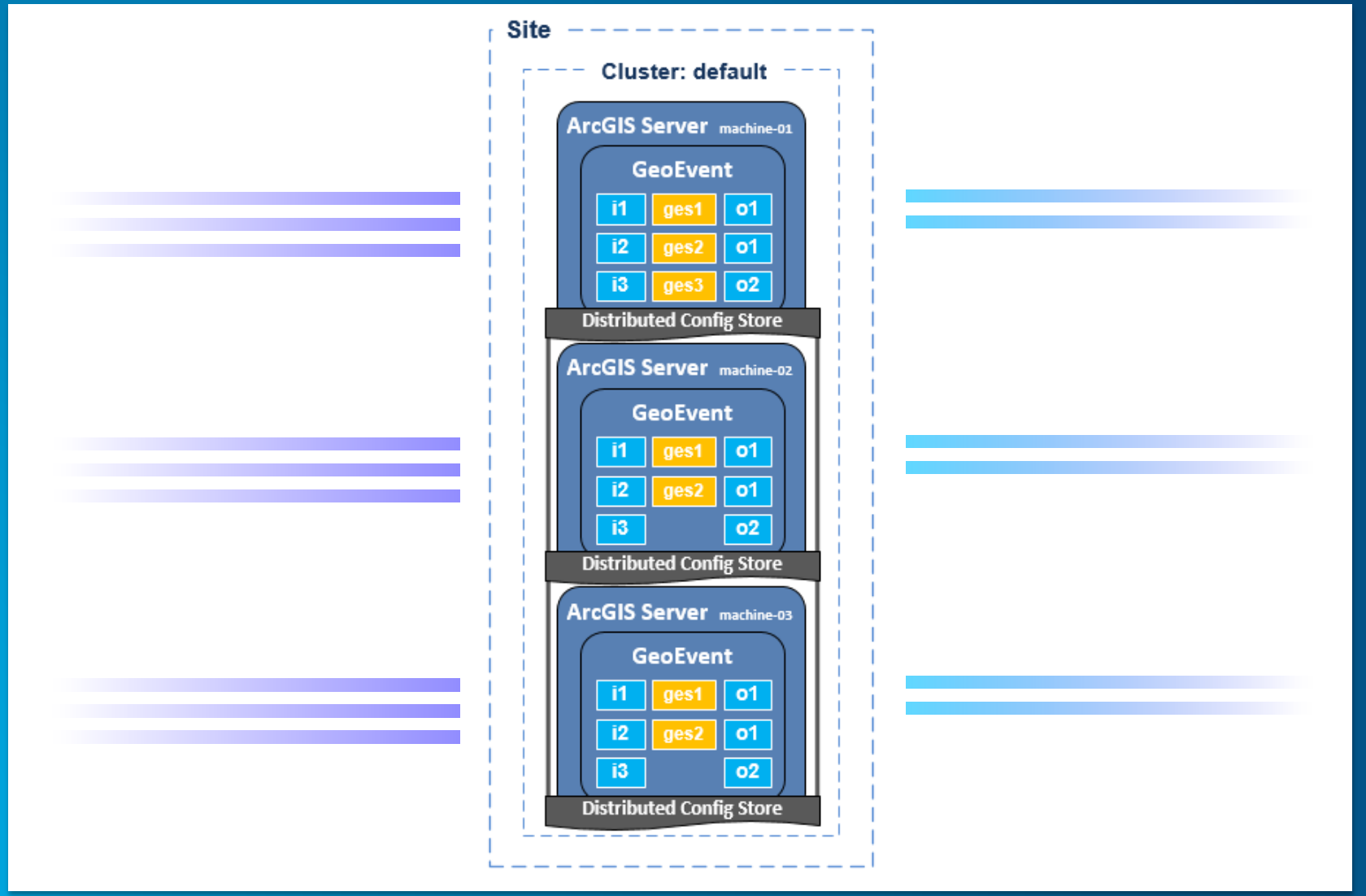
# High Availability

*Distributed configuration store*



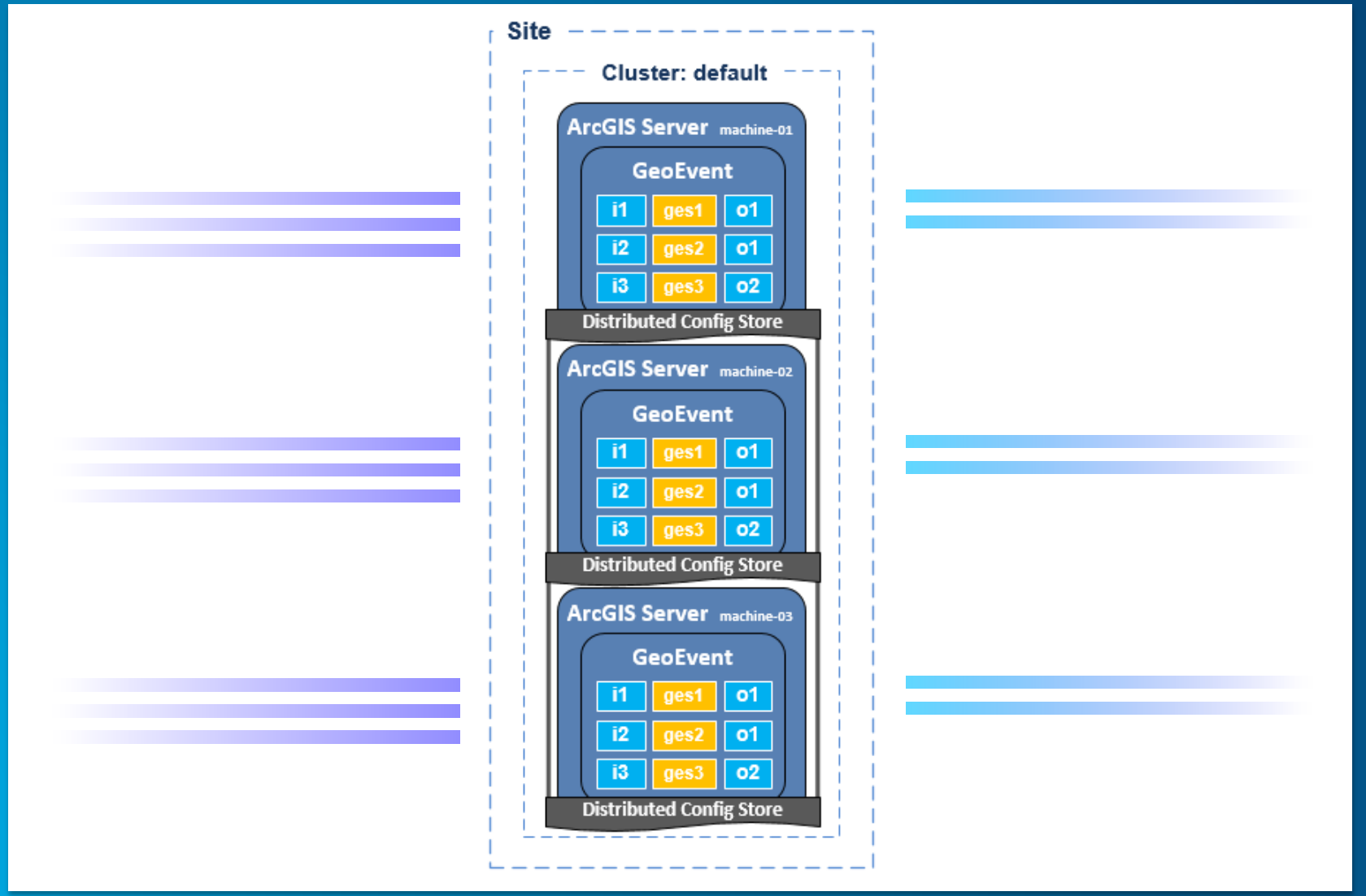
# High Availability

*Distributed configuration store*



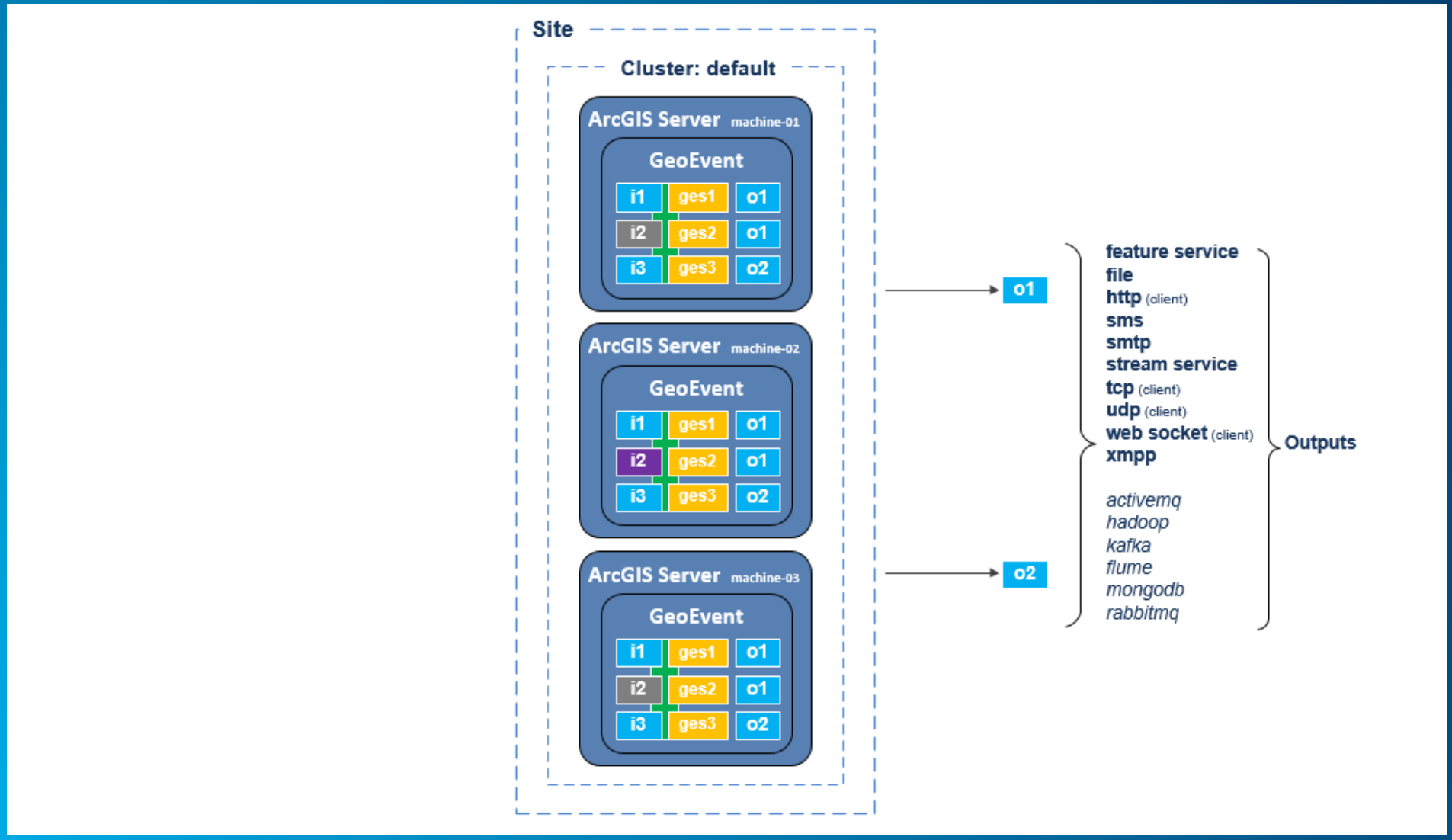
# High Availability

*Distributed configuration store*



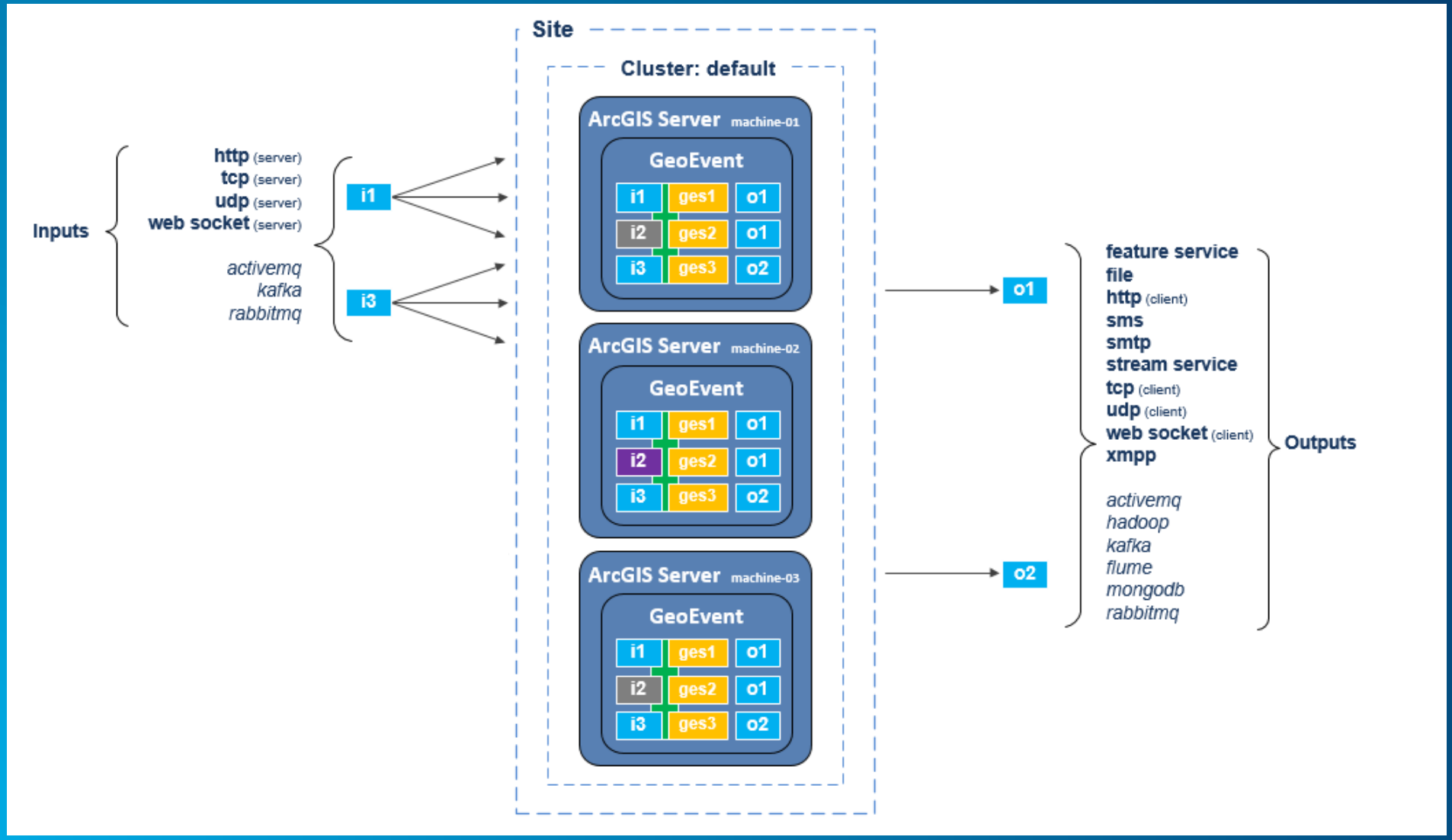
# Scalability

*Output transports*



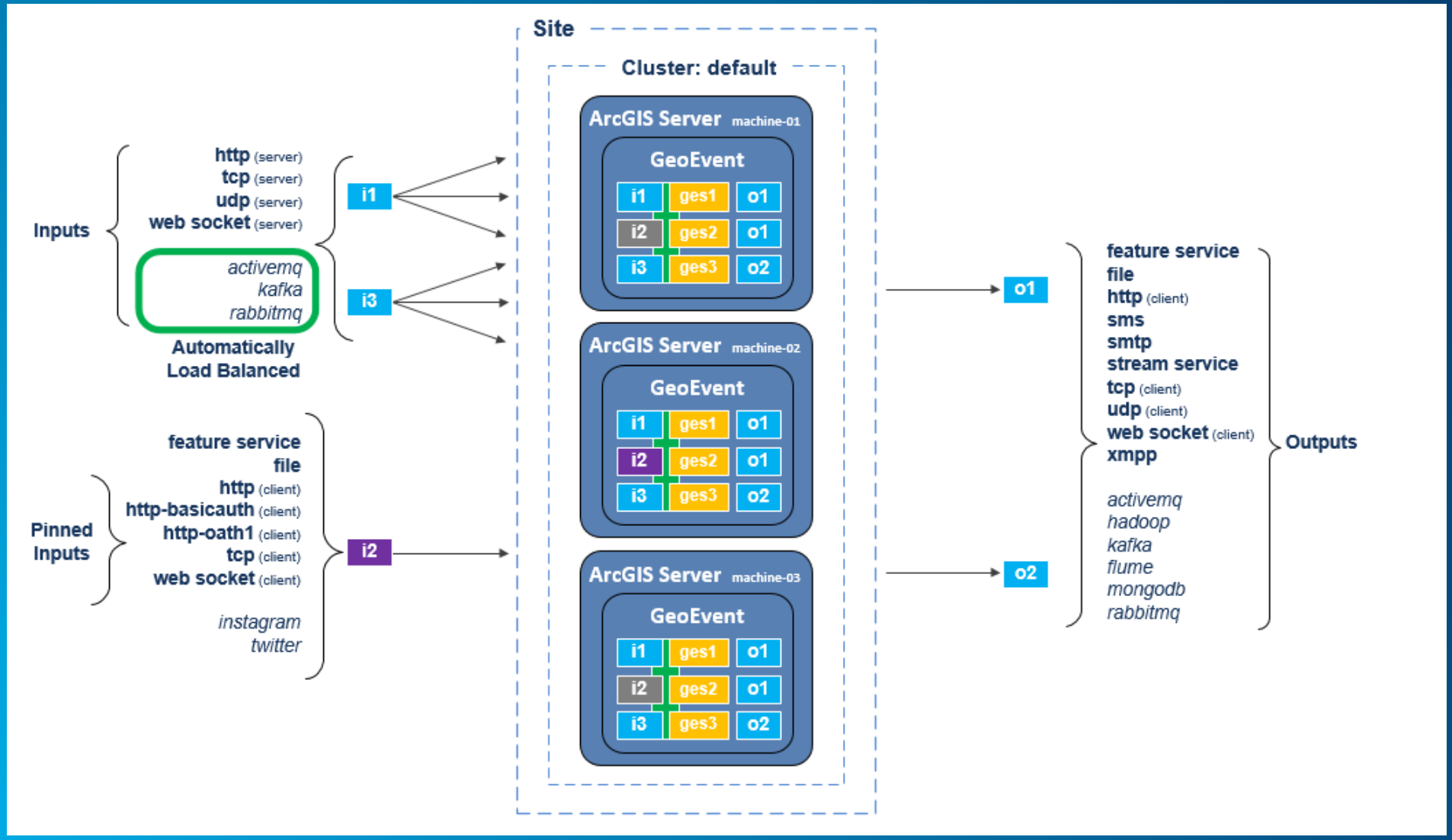
# Scalability

*Input transports*



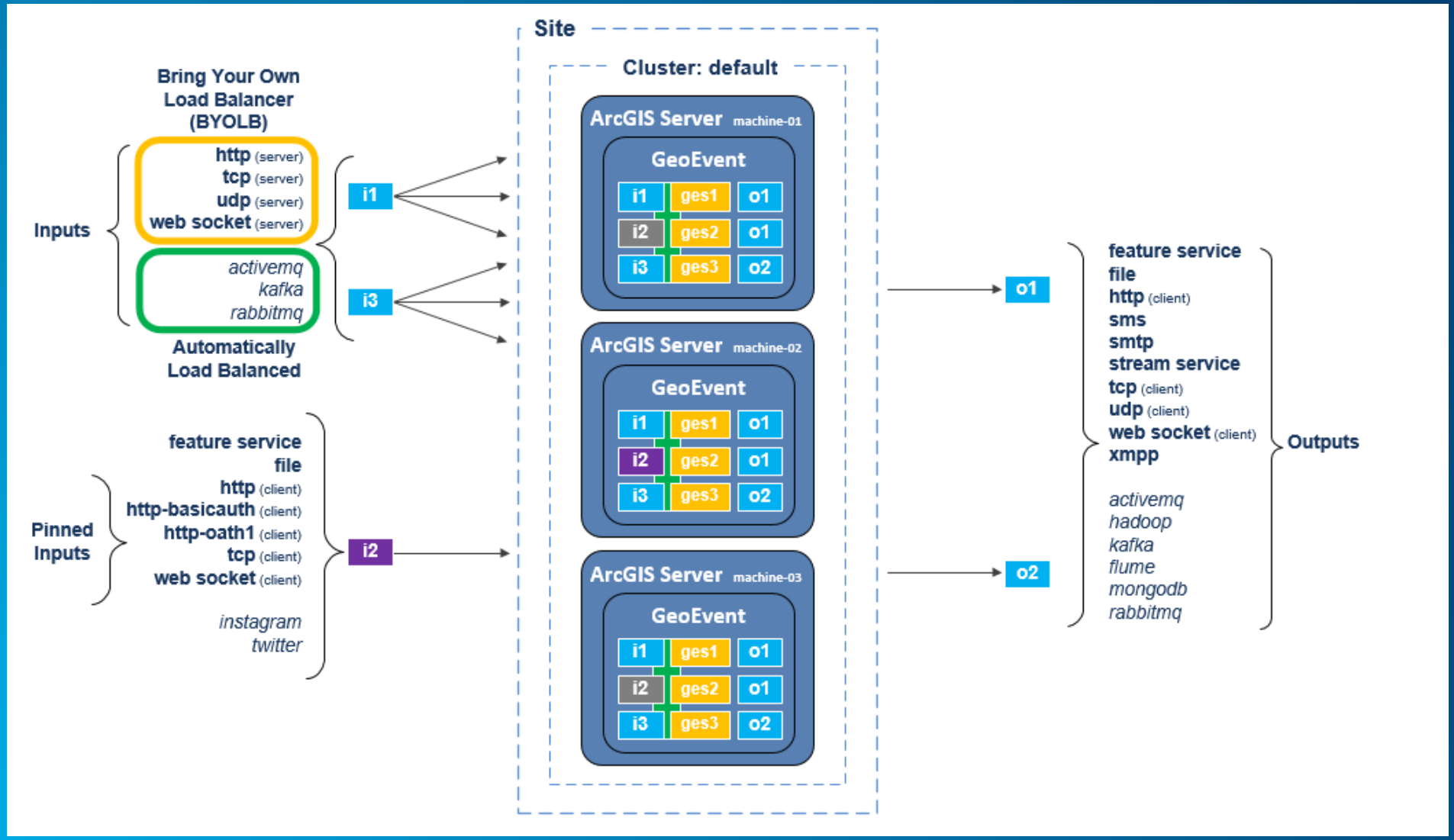
# Scalability

*Input transports that are automatically load balanced*



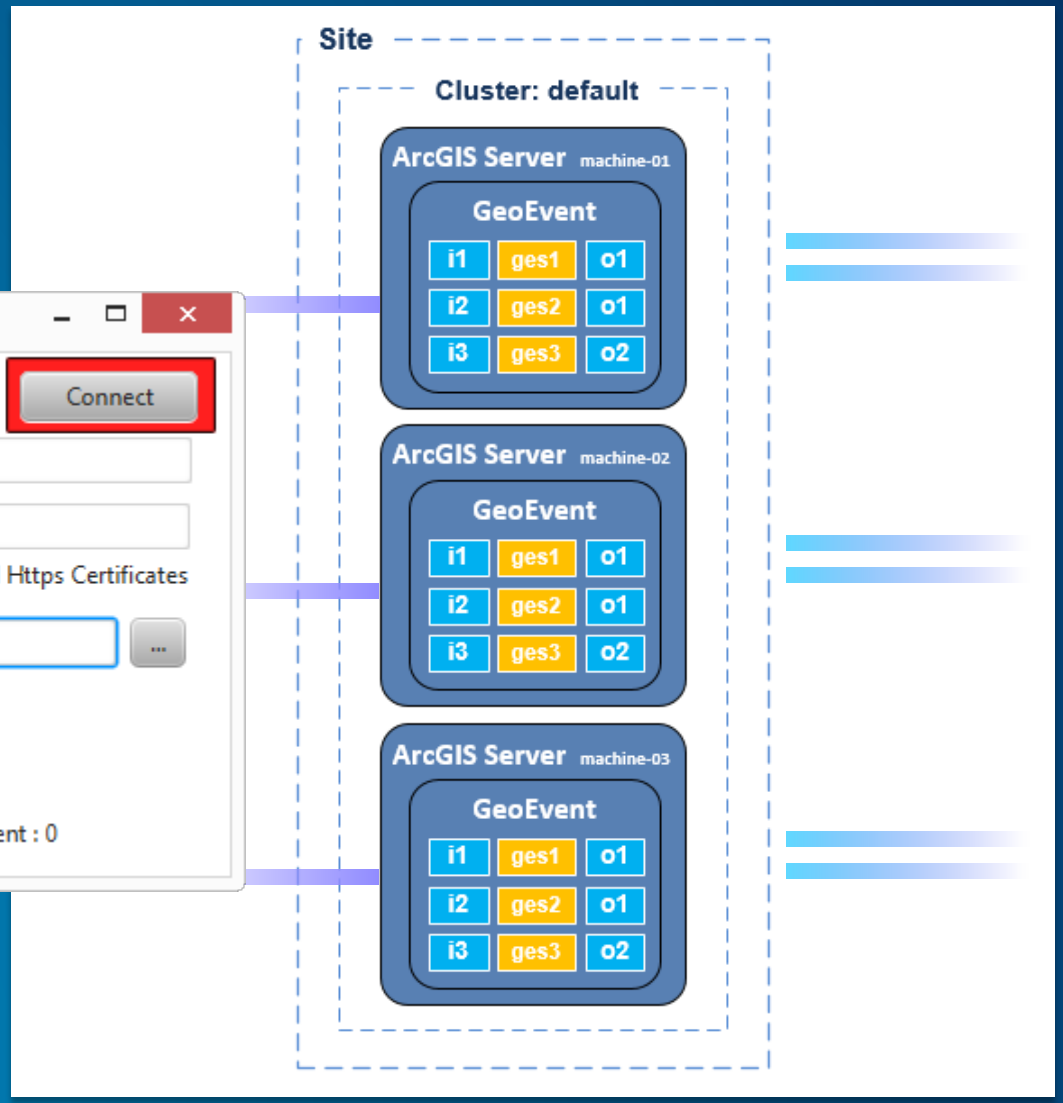
# Scalability

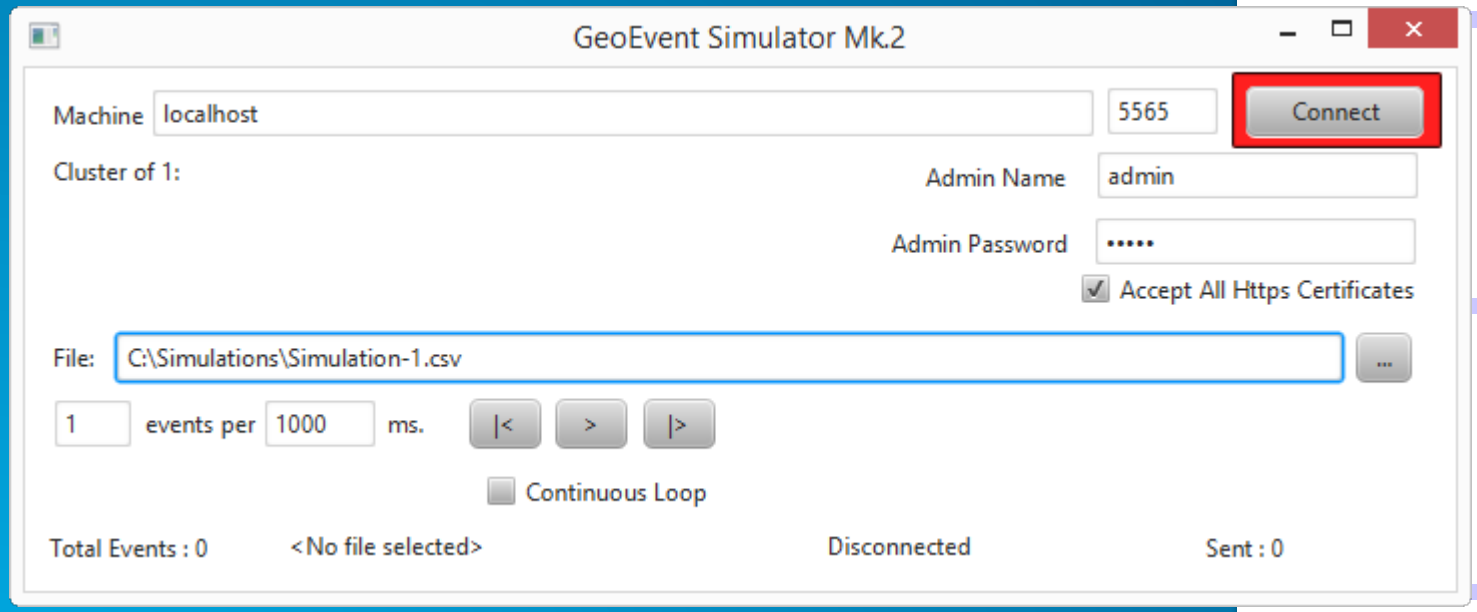
*Inputs transports that require you to bring your own load balancer*



# Scalability

*cluster-simulator-for-geoevent*

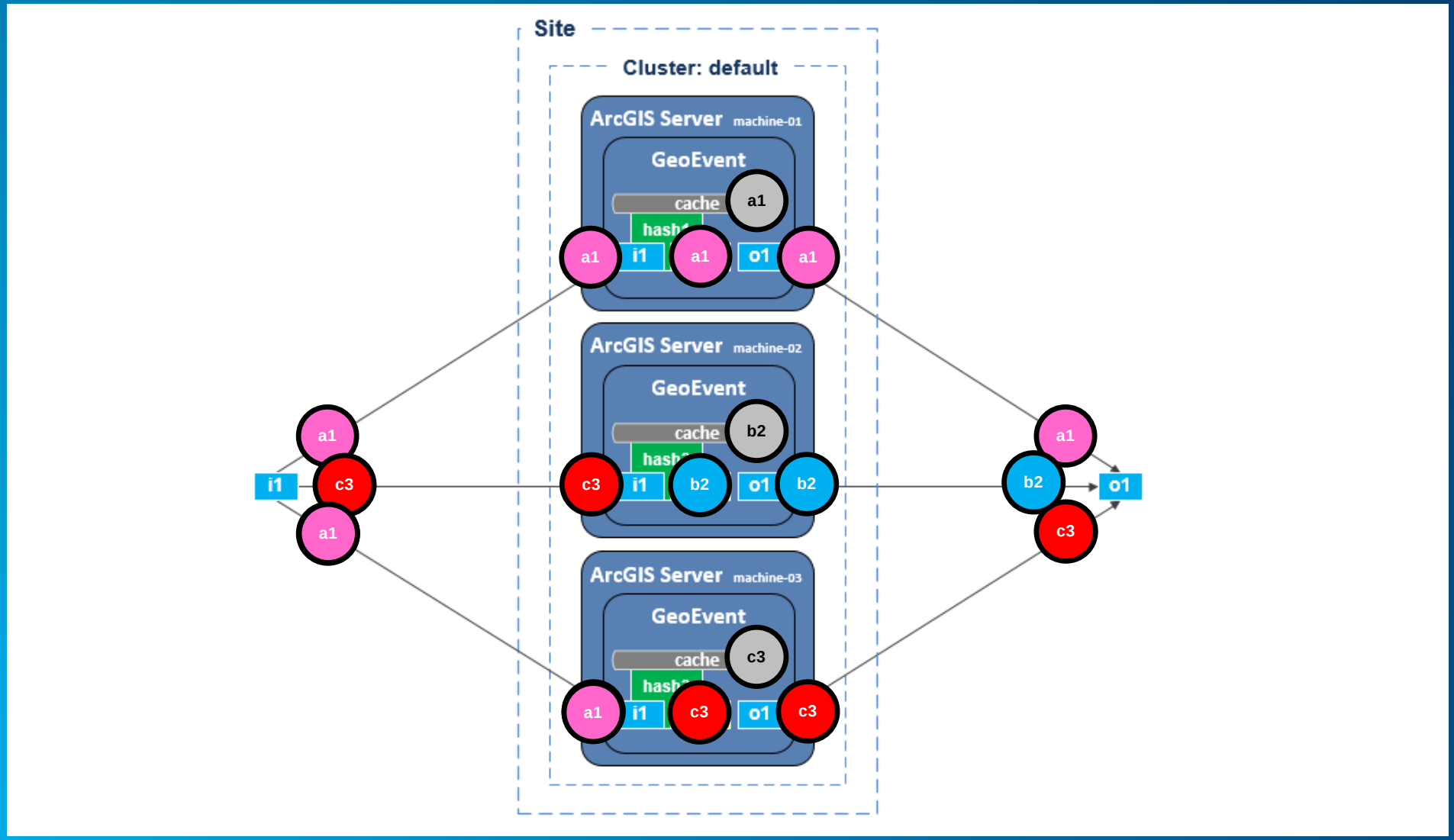



https://github.com/Esri/cluster-simulator-for-geoevent

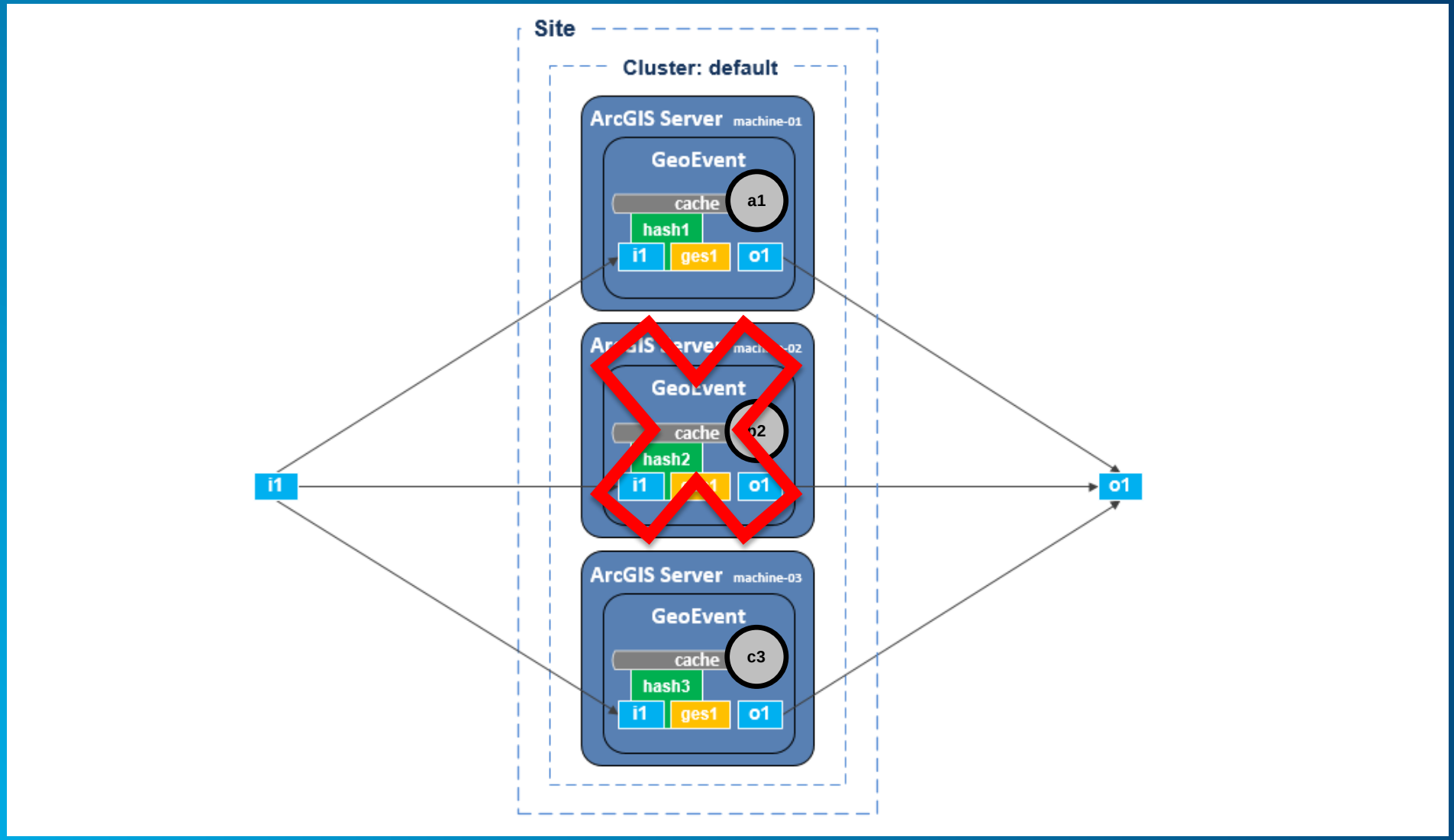
# Scalability

*Inputs and distributed stream processing*



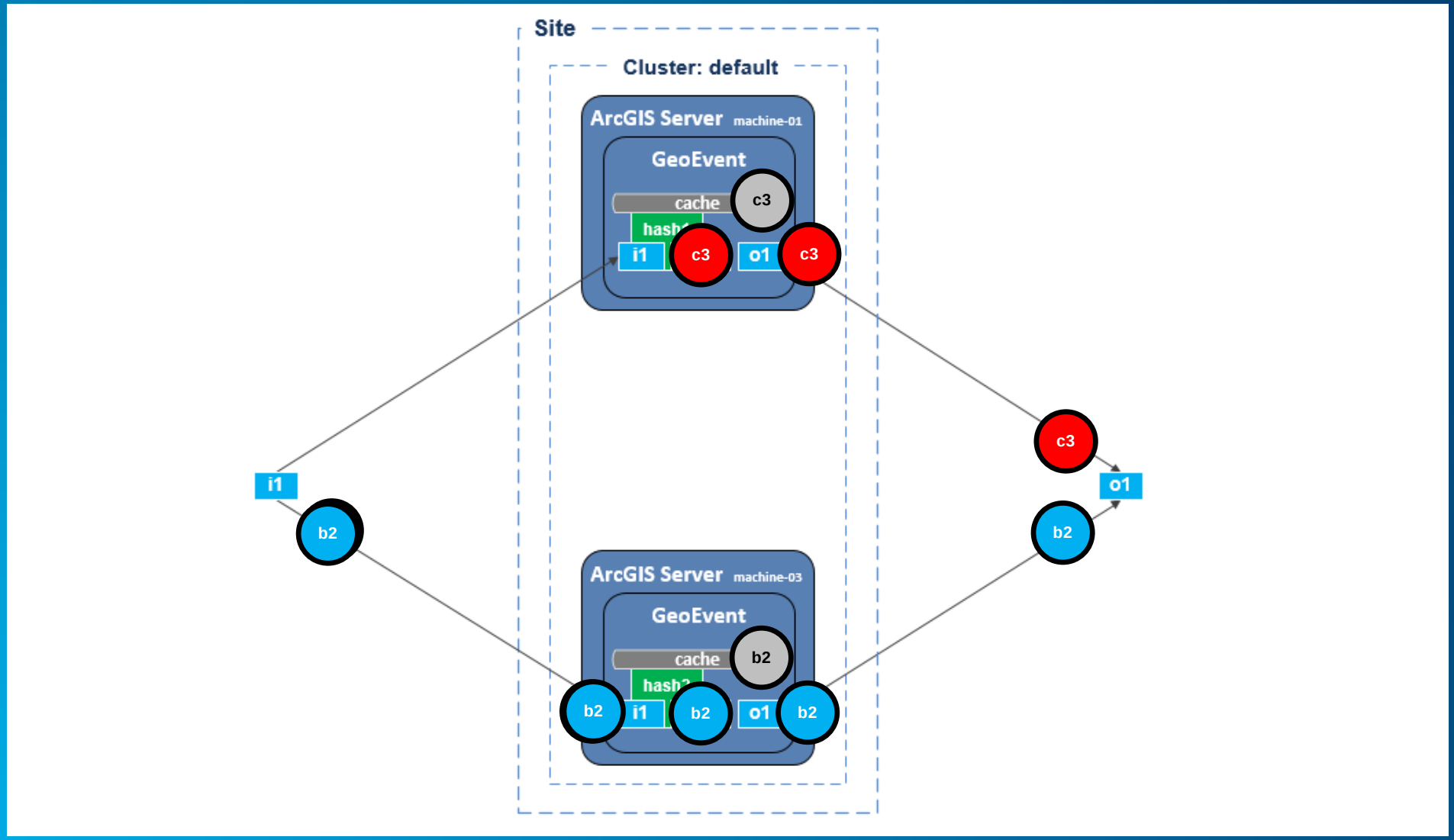
# High Availability

*Inputs and distributed stream processing*



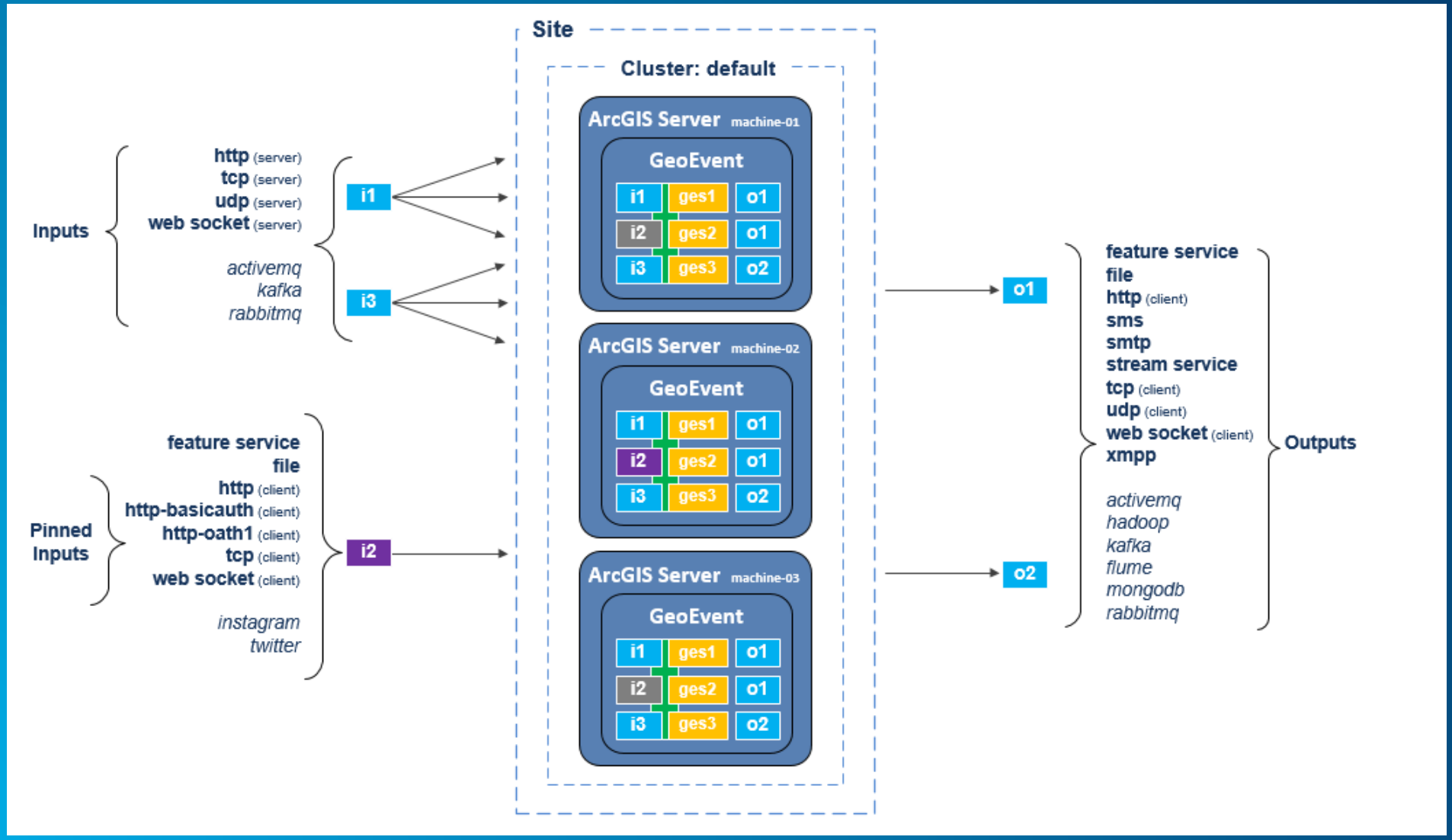
# High Availability

*Inputs and distributed stream processing*



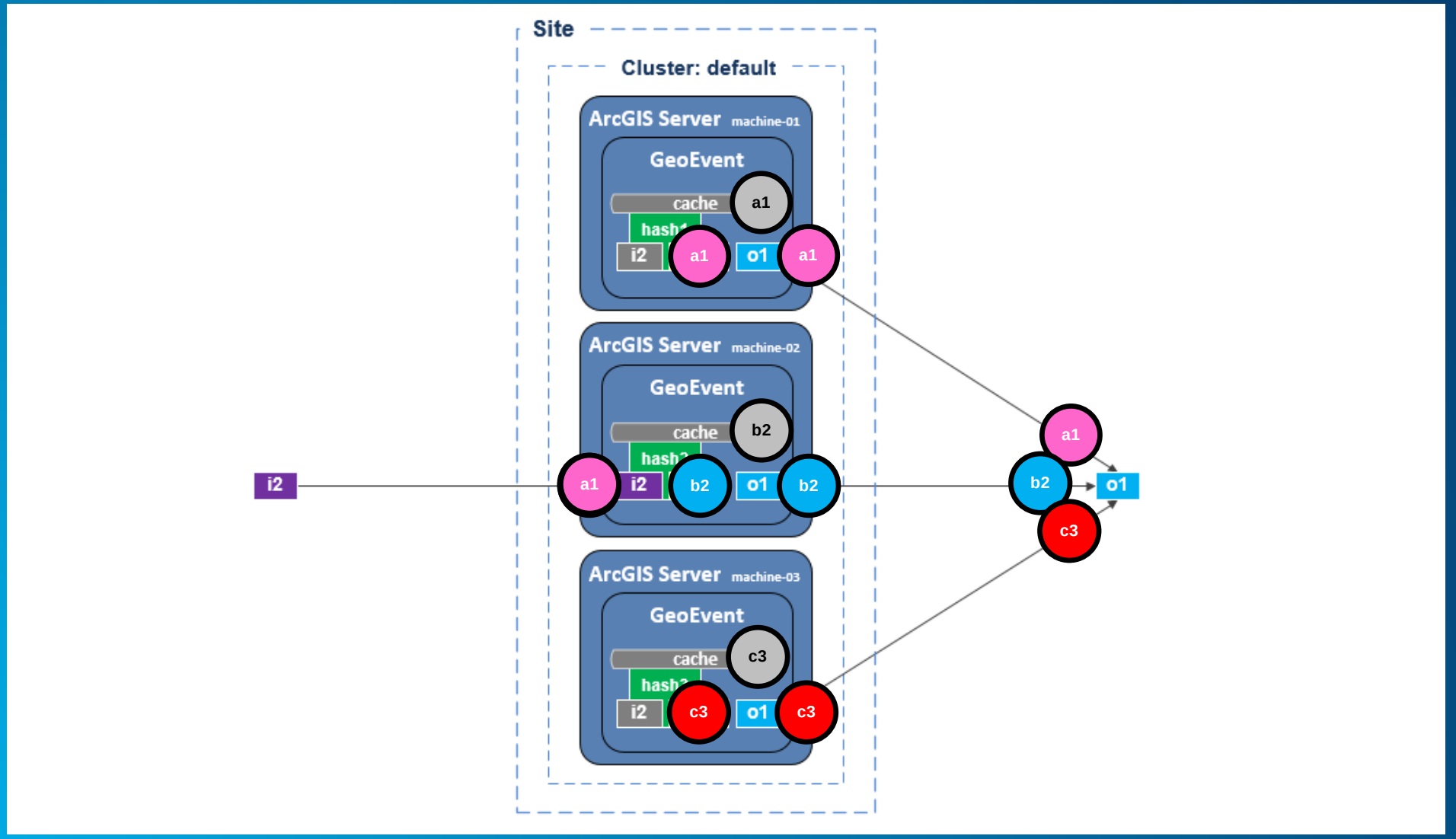
# Scalability

*Pinned inputs*



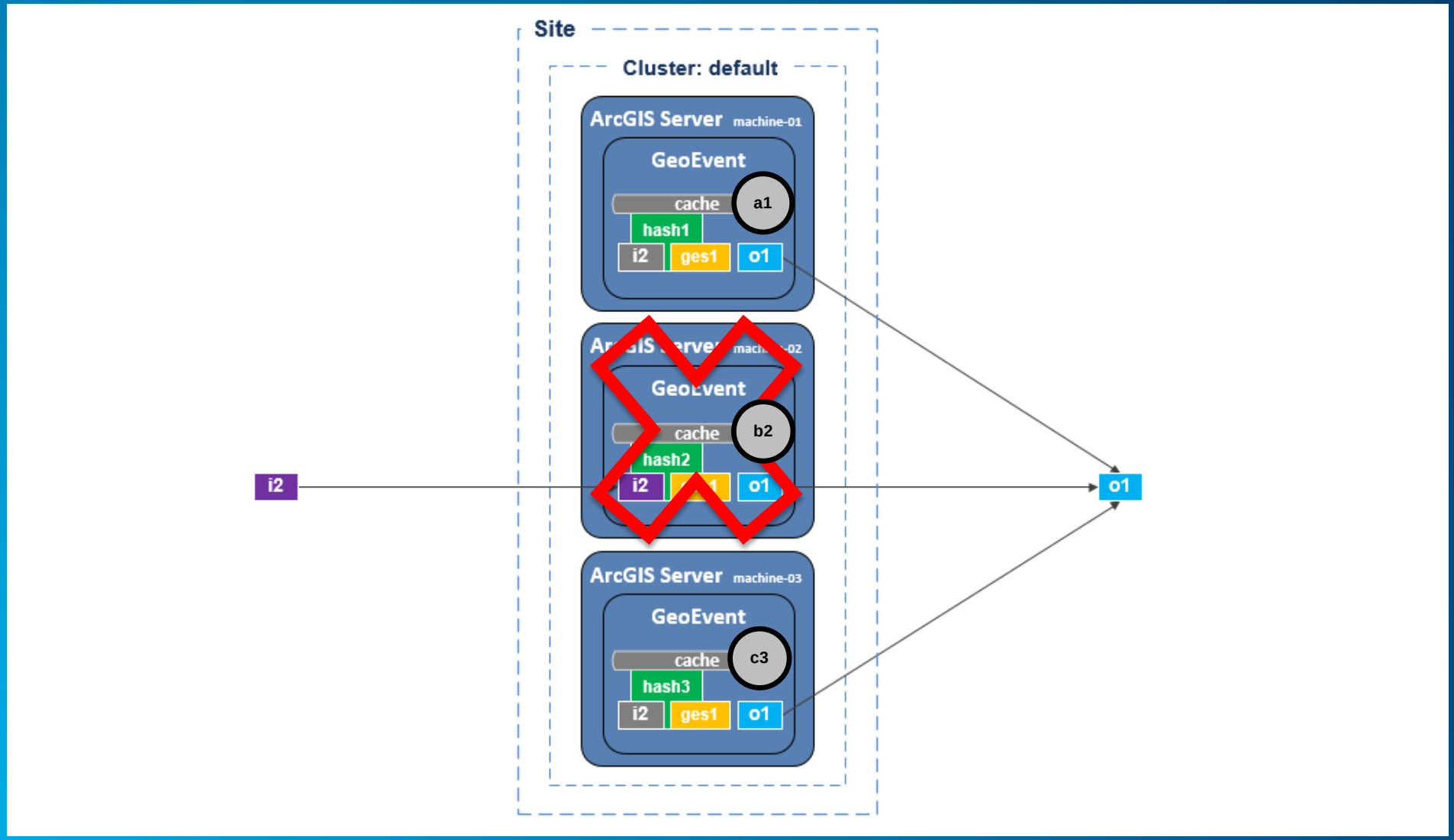
# Scalability

*Pinned inputs and distributed stream processing*



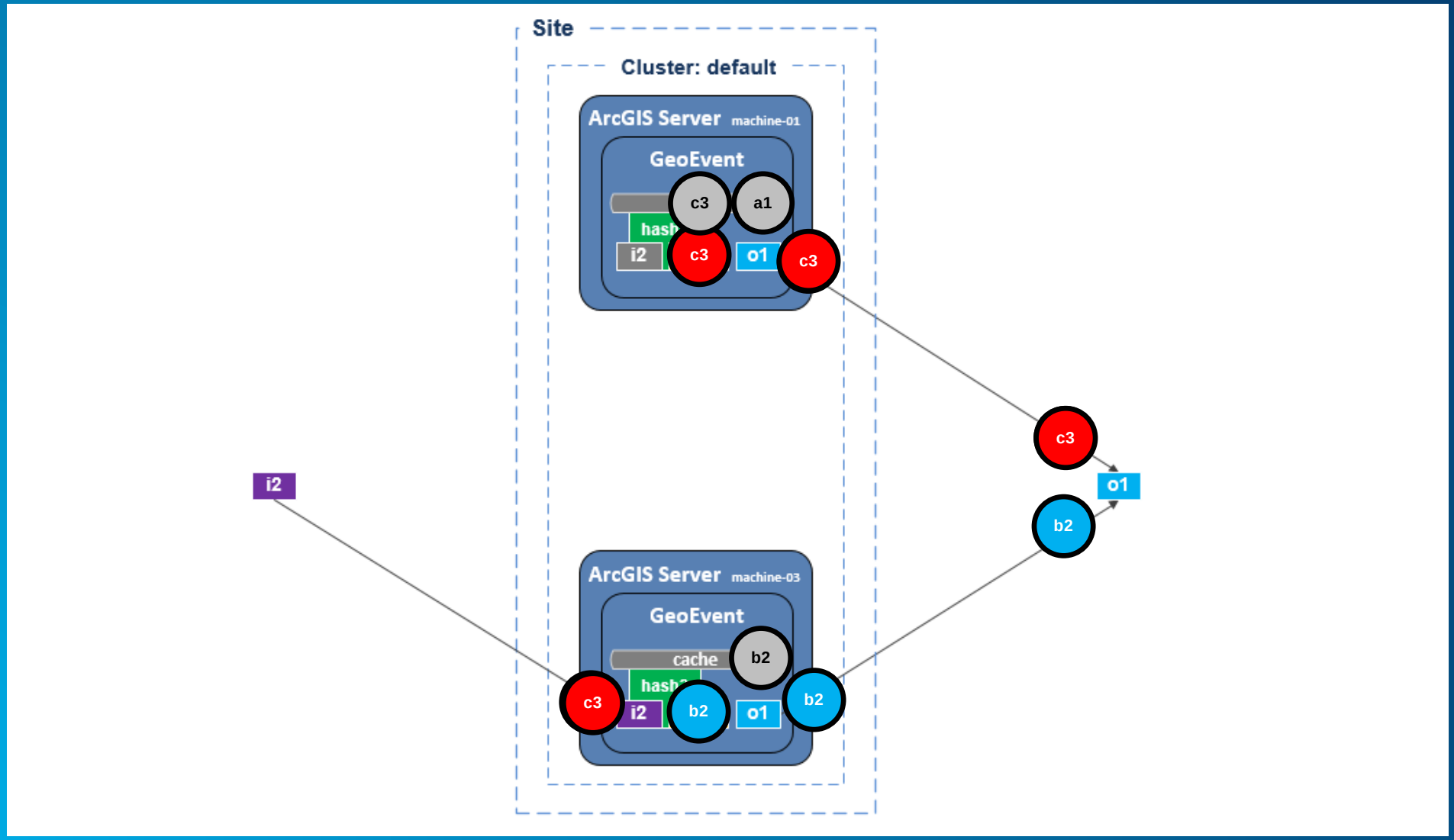
# High Availability

*Pinned inputs and distributed stream processing*



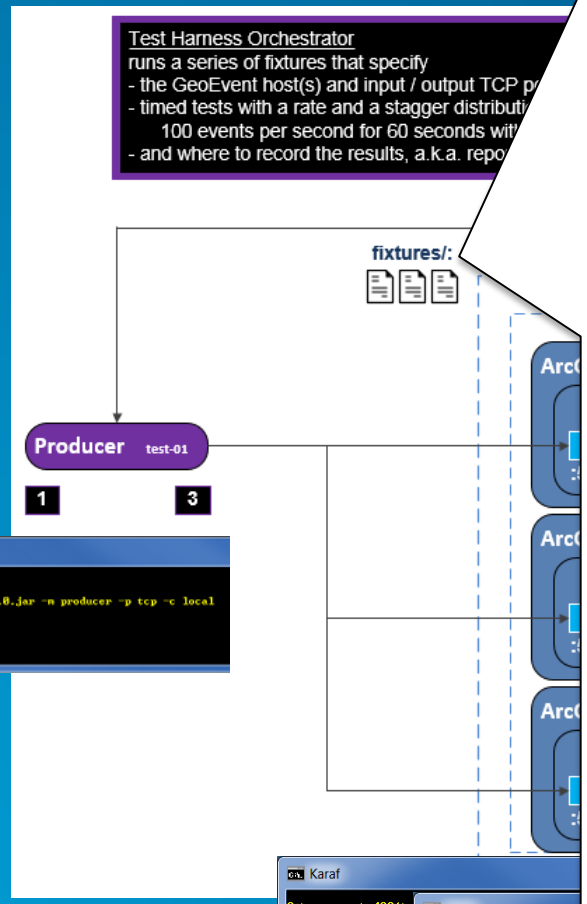
# High Availability

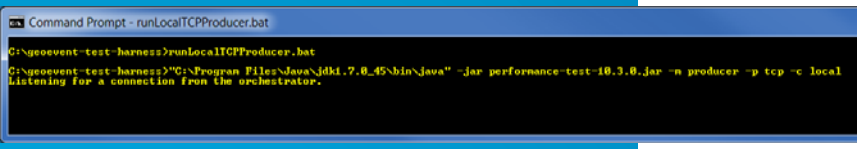
*Pinned inputs and distributed stream processing*

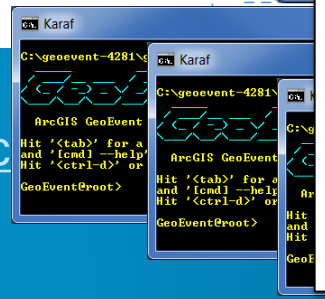


# How do I measure Cluster Perfor

*performance-test-harness-for-geoevent*




https://github.c

```
<Fixtures>
    <Report type="CSV">
        <ReportFile>reports/throughput-1000to4000step1000-cluster.csv</ReportFile>
    </Report>

    <DefaultFixture>
        <DefaultSharedConfig protocol="TCP">  <!-- also supports WS (Web Socket), RMQ (RabbitMQ)-->
            <Properties>
                <Property name="hosts">host1,host2,host3</Property>
            </Properties>
        </DefaultSharedConfig>

        <ProducerConfig commandPort="5010">
            <Properties>
                <Property name="port">5565</Property>
            </Properties>
        </ProducerConfig>

        <ConsumerConfig commandPort="5020" timeoutInSec="5">
            <Properties>
                <Property name="port">5575</Property>
            </Properties>
        </ConsumerConfig>

        <Simulation>
            <TimeTest eventsPerSec="1000" totalTimeInSec="30" staggeringInterval="10" />
        </Simulation>
    </DefaultFixture>

    <Fixture name="1000 e/s">
        <ProducerConfig>
            <Properties>
                <Property name="simulationFile">simulations/county_envelopes_points_1000.csv</Property>
            </Properties>
        </ProducerConfig>
        <Simulation>
            <TimeTest eventsPerSec="1000"/>
        </Simulation>
    </Fixture>

    ...

    <Fixture name="4000 e/s">
        <ProducerConfig>
            <Properties>
                <Property name="simulationFile">simulations/county_envelopes_points_4000.csv</Property>
            </Properties>
        </ProducerConfig>
        <Simulation>
            <TimeTest eventsPerSec="4000"/>
        </Simulation>
    </Fixture>
</Fixtures>
```

# Scalability

*Clustering for increased throughput*

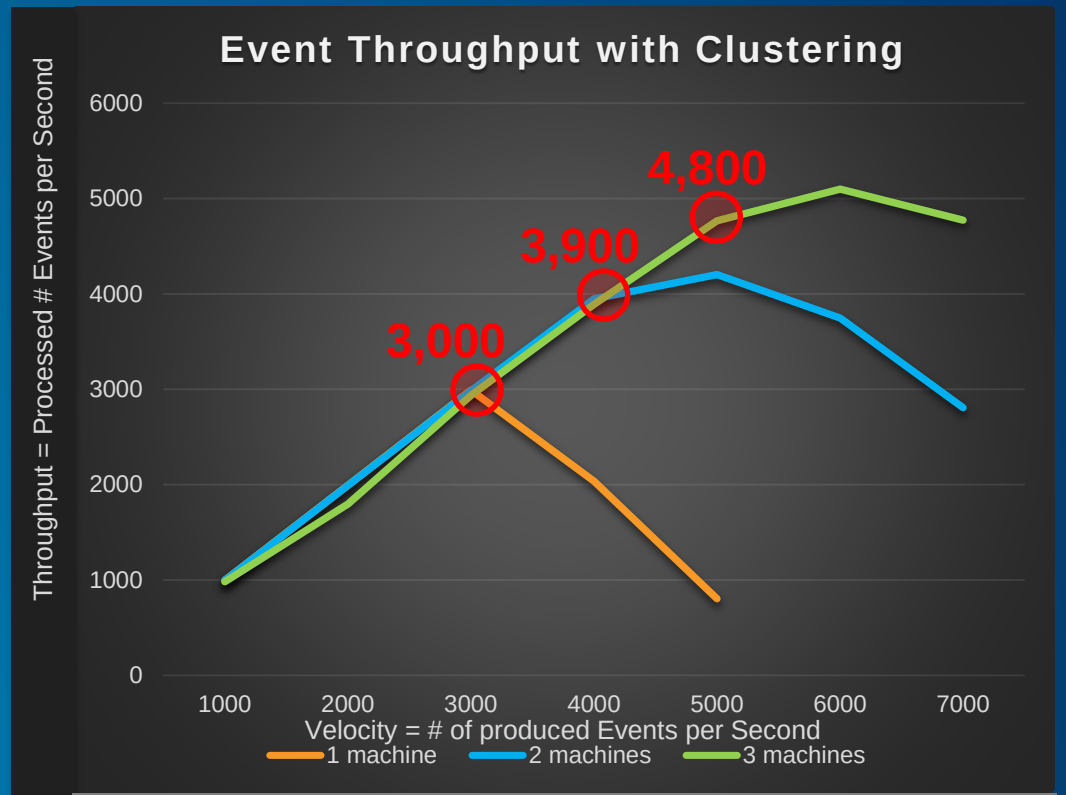
- **Clusters administered via ArcGIS Server Manager**
  - **Site, Cluster(s), Machines**
- **Scale-out by adding machines to a cluster**

Machines

| Name | Status |
|---|---|
| LOANR2285.ESRI.COM | Started |
| RTRUJILLO.ESRI.COM | Started |
| SOURCE.ESRI.COM | |
| XW8600-W7.ESRI.COM | Stopped |

Clusters

| Name | Machines | Protocol |
|---|---|---|
| default | SOURCE.ESRI.COM<br>LOANR2285.ESRI.COM<br>RTRUJILLO.ESRI.COM | TCP port 4004 |




| Machine | Mfg | OS | CPU Type | CPU | Cores | RAM | Role |
|---|---|---|---|---|---|---|---|
| loanr2285 | Dell | W7-64 Ent | Intel Core I7 | 2.80GHz | 8 | 16GB | geoevent-01 |
| rtrujillo | Dell | W7-64 Ent | Intel Xeon | 3.70GHz | 8 | 16GB | geoevent-02 |
| source | Dell | W8.1 Ent | Intel Core il7 | 2.80GHz | 8 | 16GB | geoevent-03 |

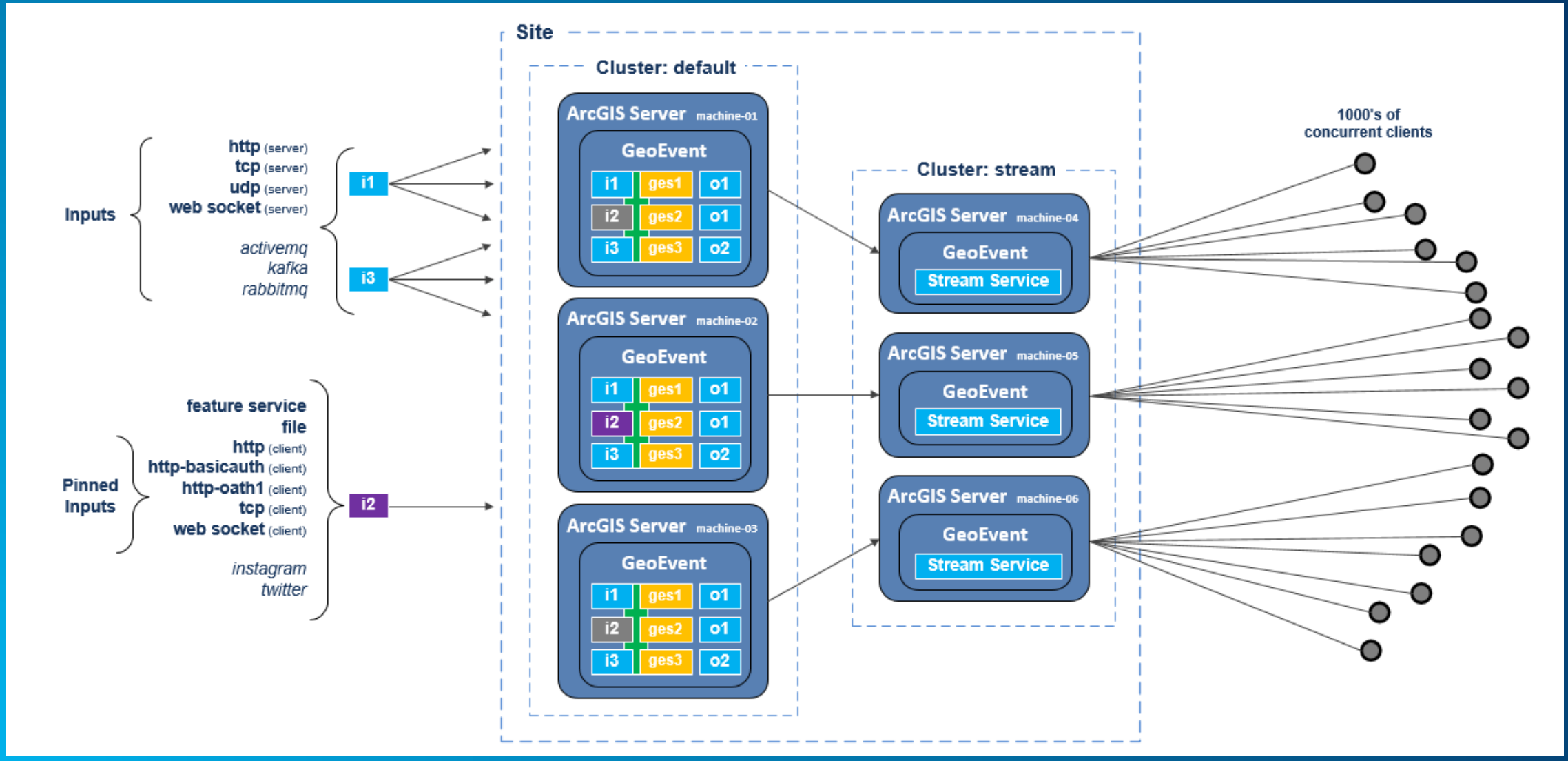
As captured on 10.3 benchmarking cluster using ArcGIS 10.3

# Stream Services:

## High Availability & Scalability

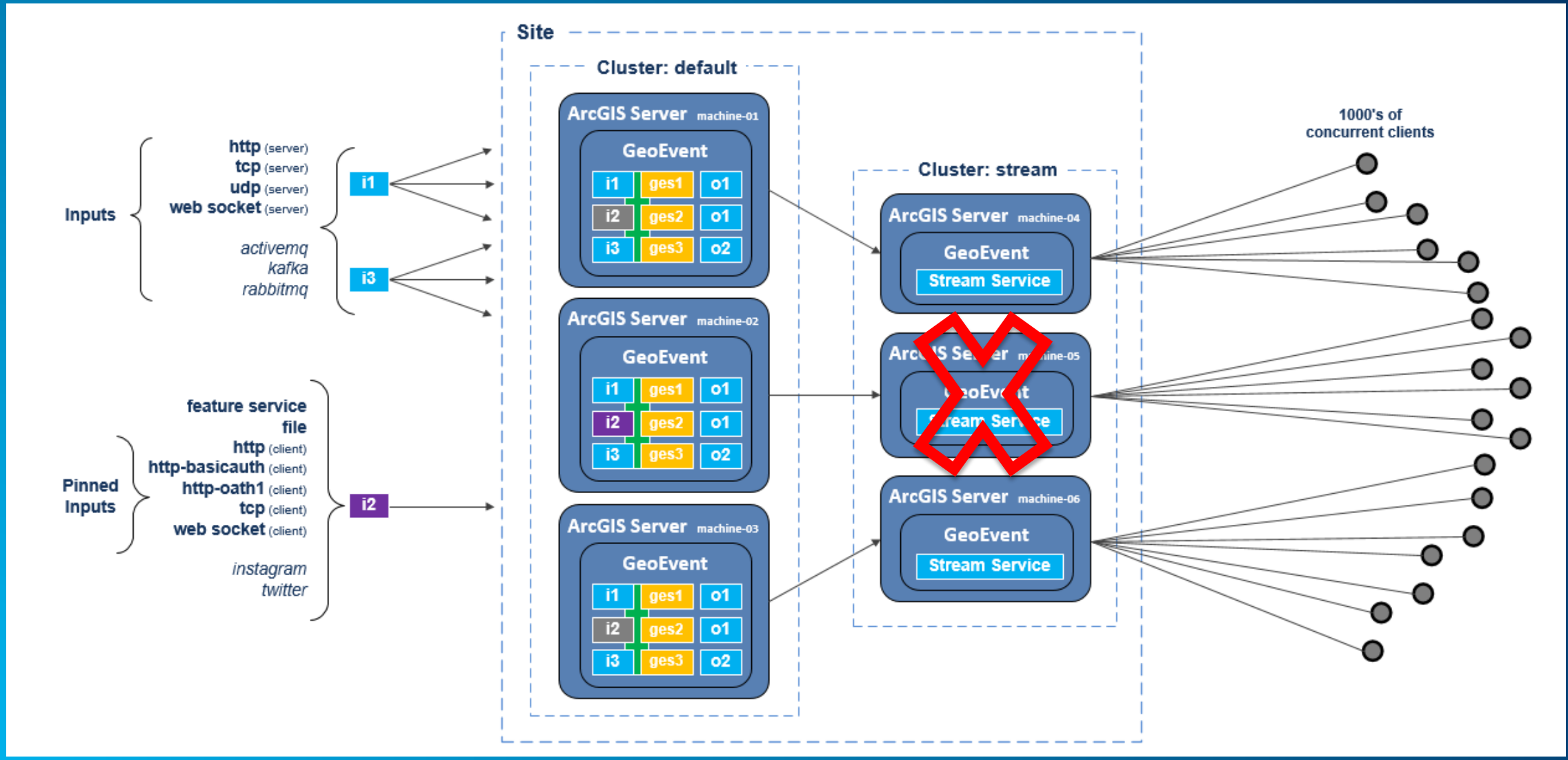
# Scalability

*Scaling out Stream Services to support an increased # of concurrent users*



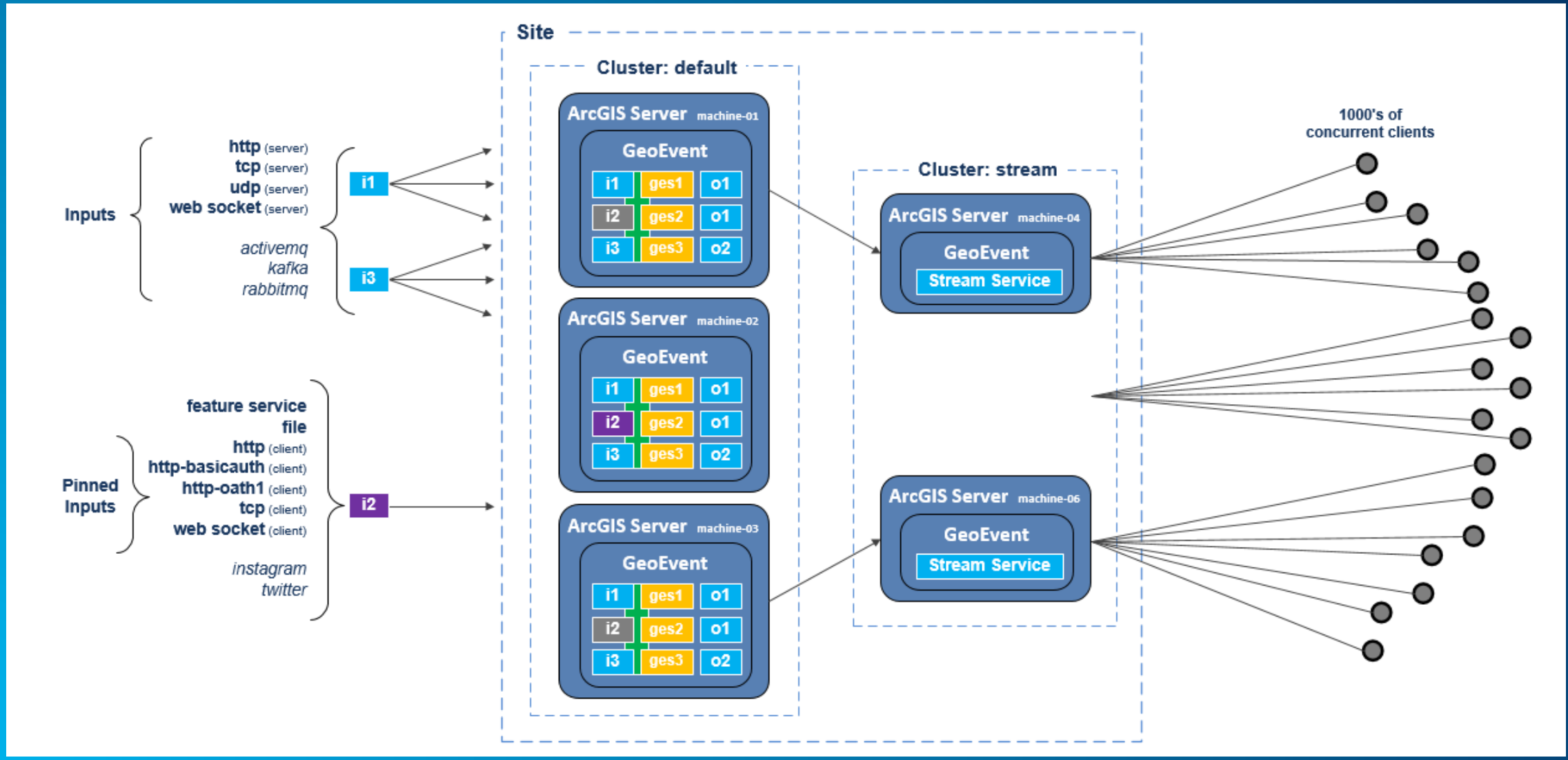
# High Availability

*Stream Services concurrent user failover*



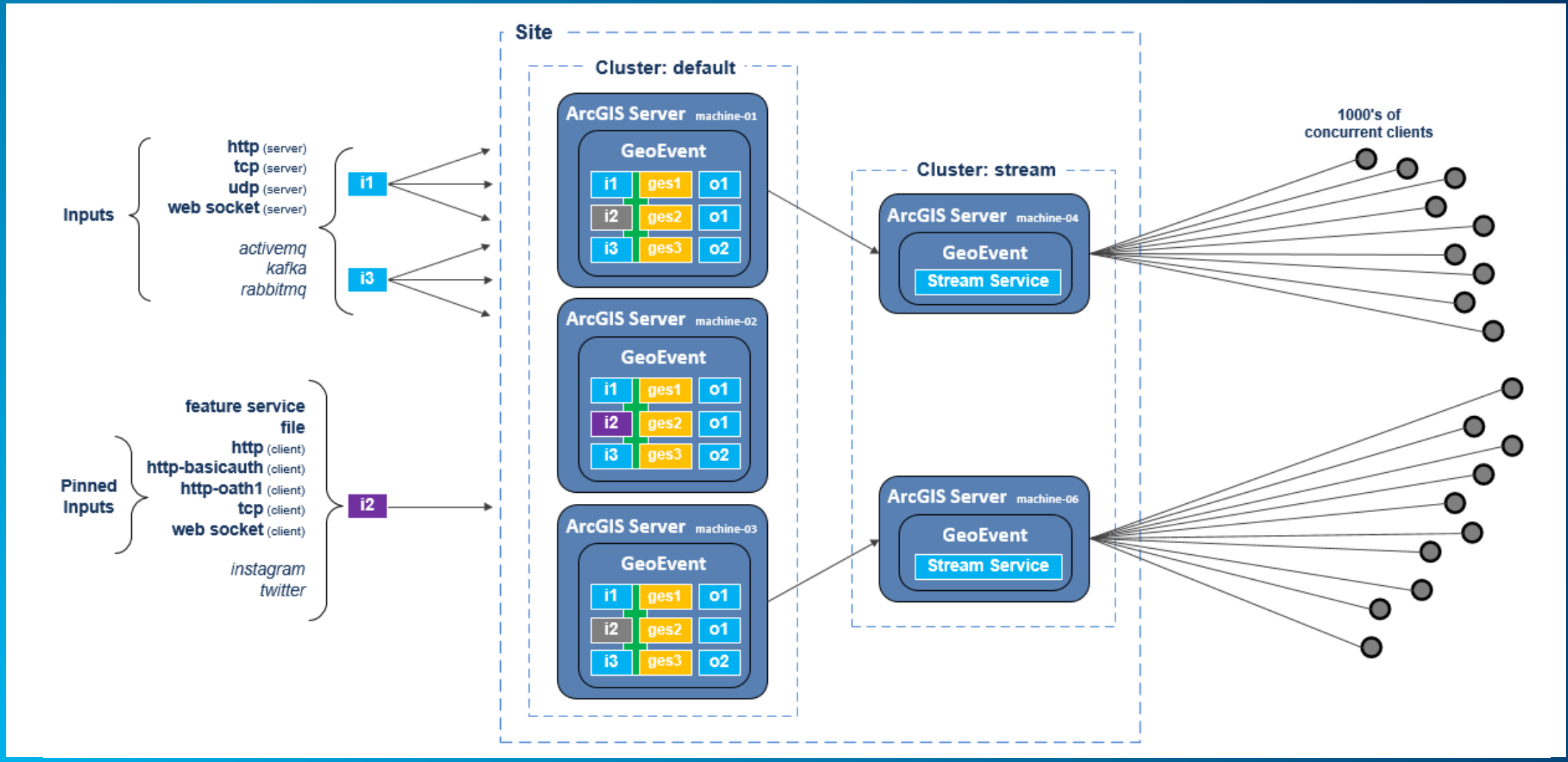
# High Availability

*Stream Services concurrent user failover*



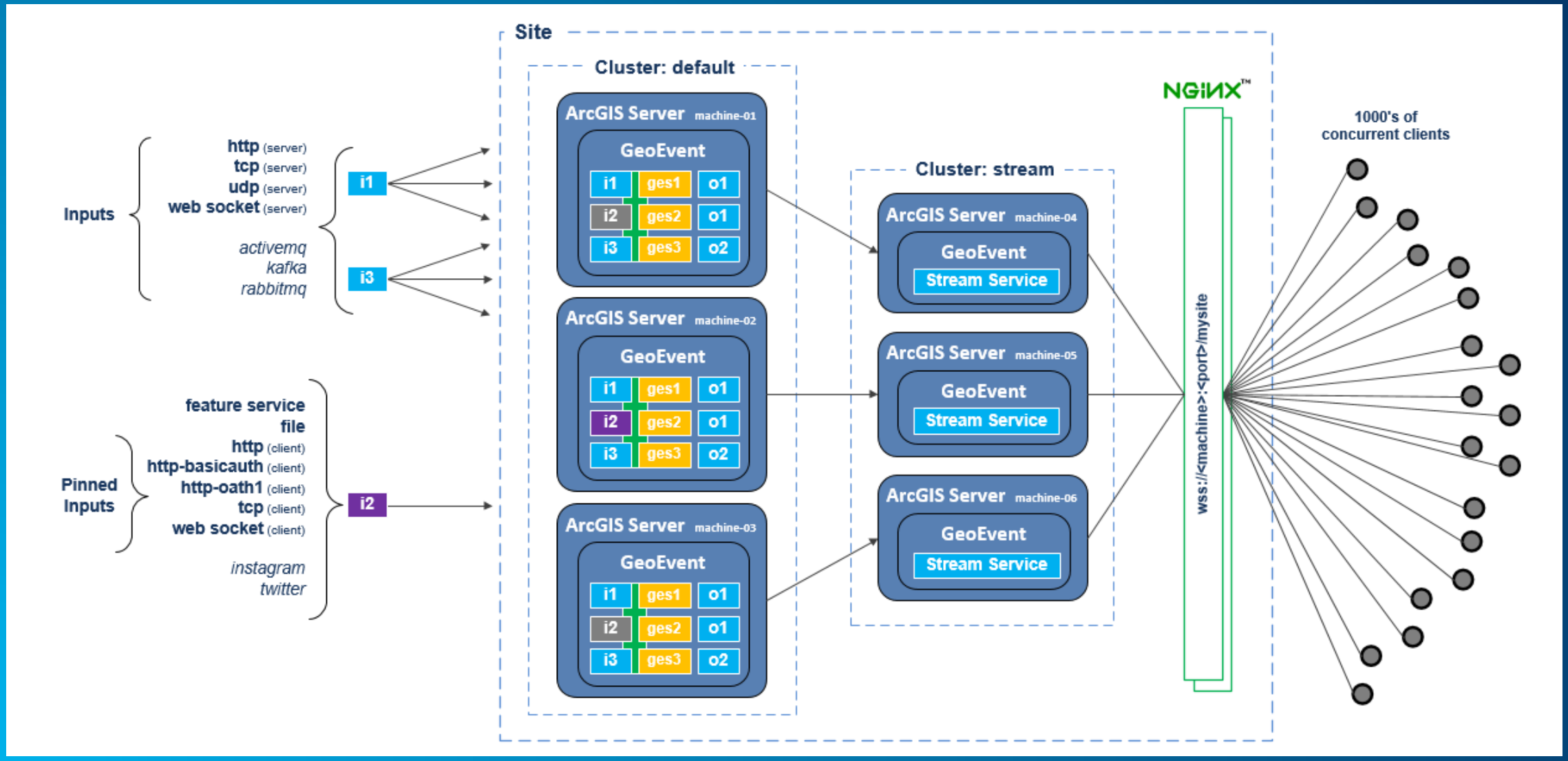
# High Availability

*Stream Services concurrent user failover*



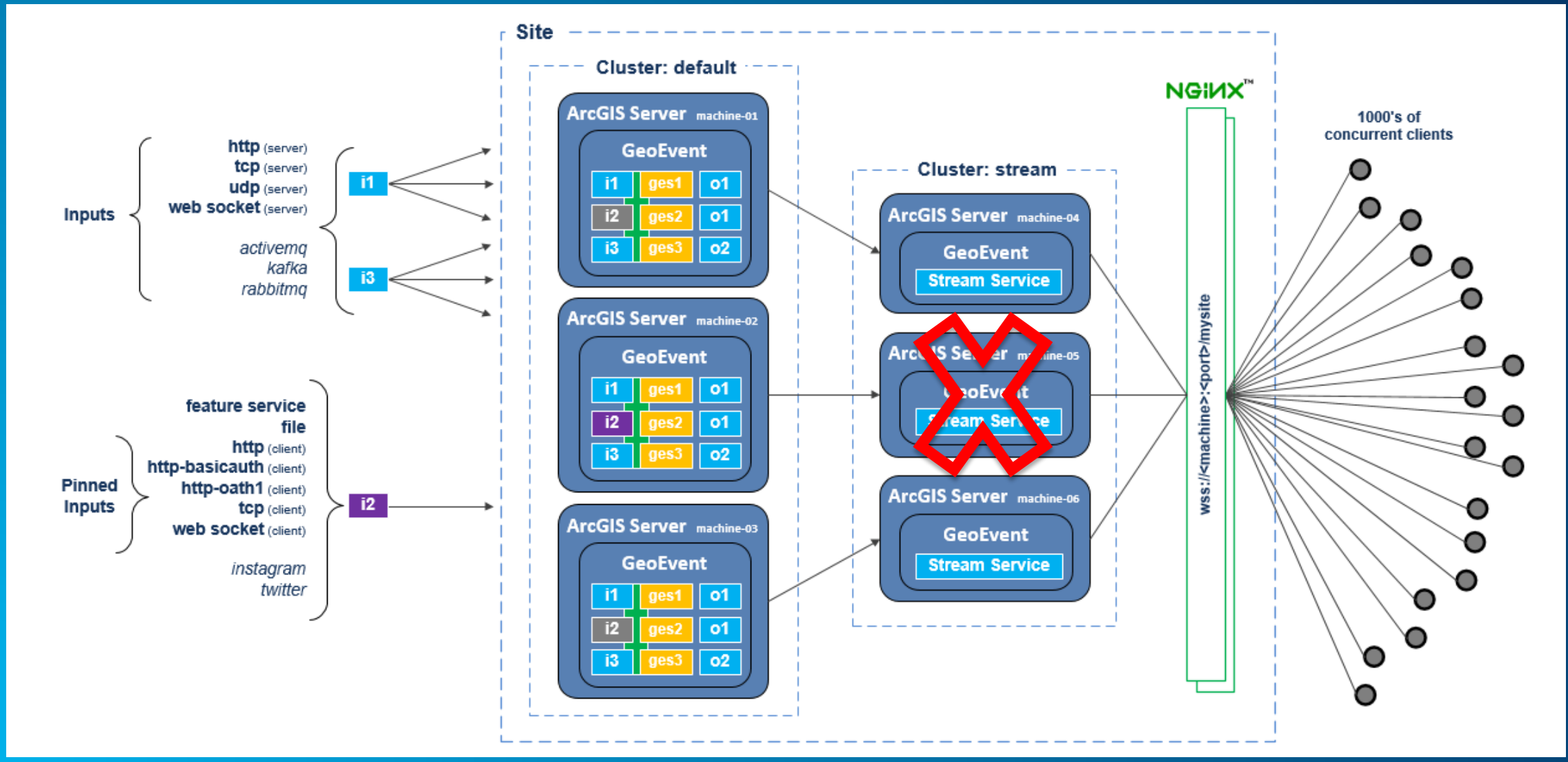
# Scalability

*Scaling out Stream Services to support an increased # of concurrent users with a reverse proxy*



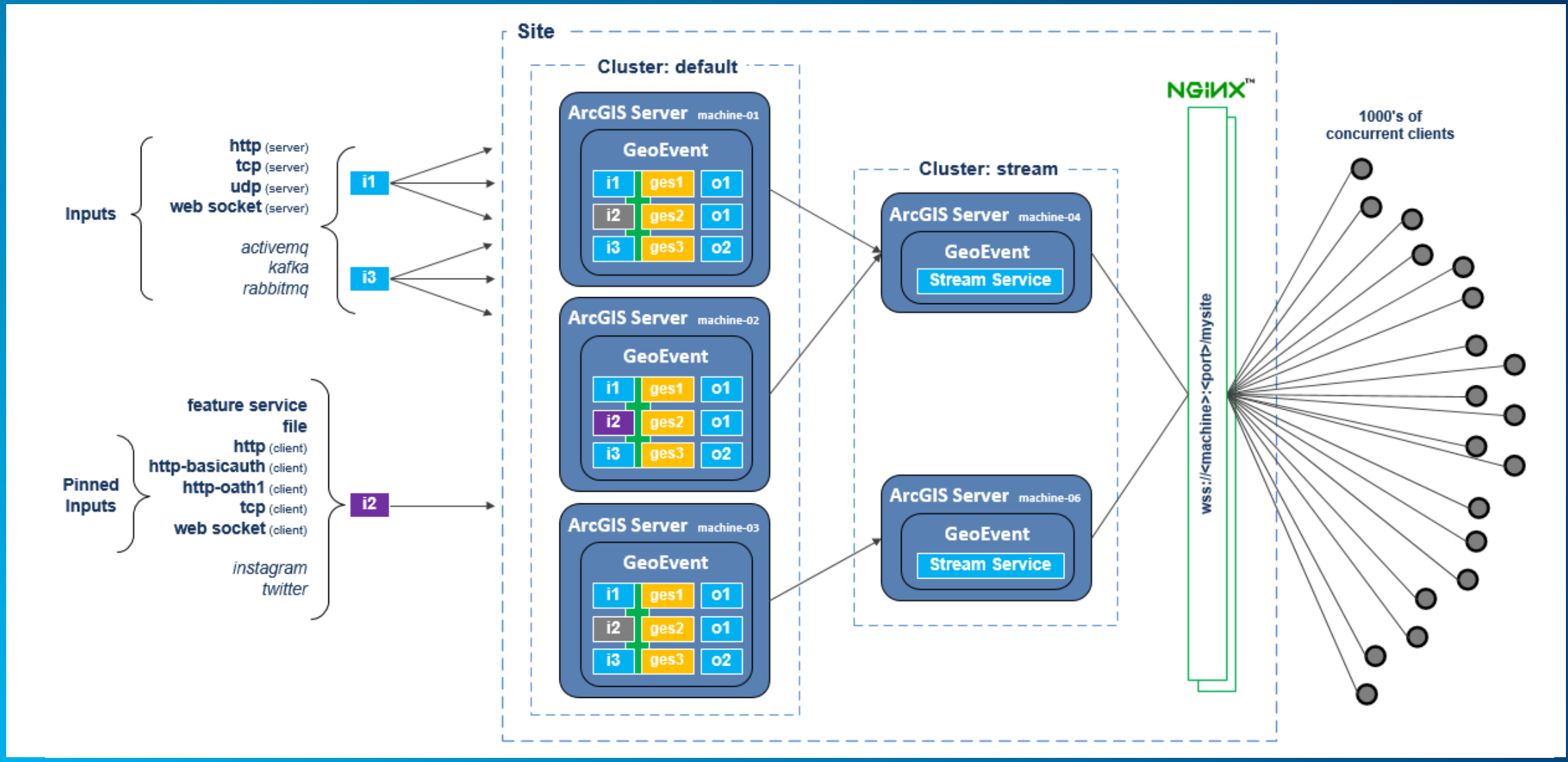
# High Availability

*Stream Services concurrent user failover with a reverse proxy*



# High Availability

*Stream Services concurrent user failover with a reverse proxy*



# GeoEvent Extension: Best Practices

*Summary*

- **ArcGIS is a dynamic platform that enables continuous analytics and real-time visualization for better understanding of our world.**

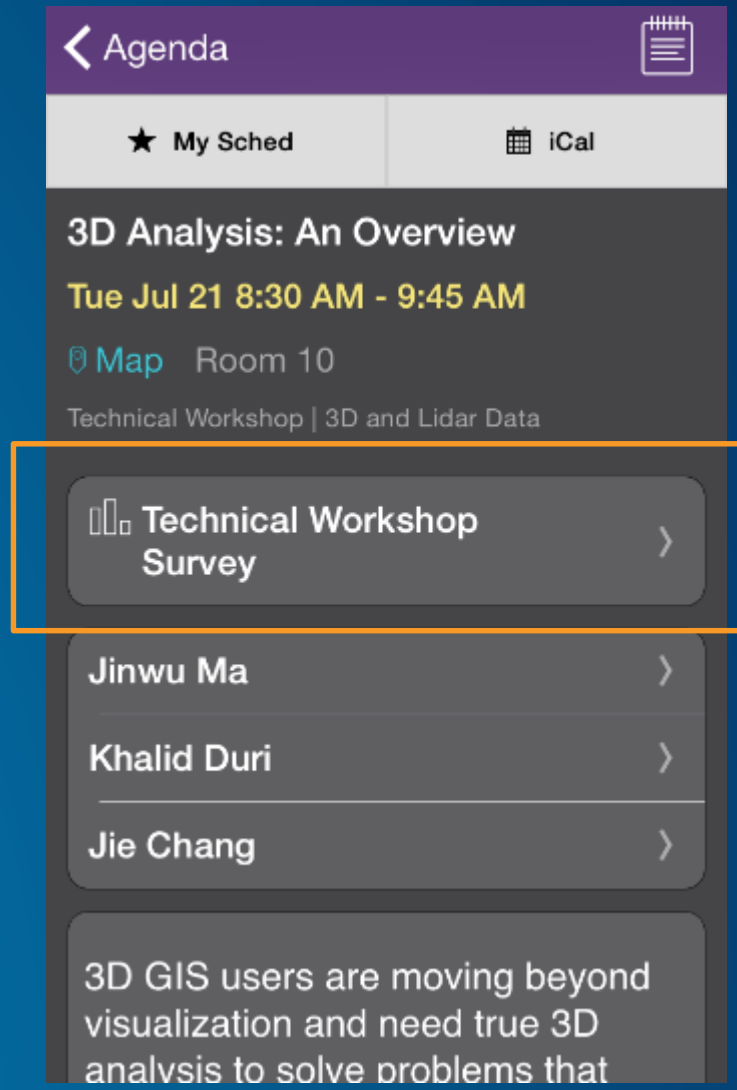
- **The ArcGIS GeoEvent Extension for Server allows you to:**
  - **know what is happening, as it happens**
  - **react and make smarter decisions faster**
  - **be notified when interesting events occur**

# Other Real-Time GIS Sessions

- **Real-Time GIS: GeoEvent Extension**
  - Tue 10:15-11:30am, Room 17 A
  - Wed 10:15-11:30am, Room 14 B
- **Real-Time GIS: Leveraging Stream Services**
  - Tue 8:30-9:45am, Room 01 A/B
  - Wed 8:30-9:45am, Ballroom 06 D
- **Real-Time GIS: Applying Real-Time Analytics**
  - Tue 10:15-11:30am, Room 15 B
  - Wed 8:30-9:45am, Room 14 B
- **Real-Time GIS Use Cases and Implementation Patterns**
  - Tue 2:30-3:15PM, Demo Theater 6 - Geodata
- **ArcGIS Intelligence: Discern Activities of Interest through Advanced Analytics**
  - Wed 10:15-11:30am, Omni Ballroom A/B
- **Real-Time GIS: The Road Ahead**
  - Wed 1:30-2:45pm, Room 14 B
- **Real-Time GIS: Best Practices**
  - Thu 8:30-9:45am, Room 14 B
- **Real-Time GIS for Asset Readiness, Event Preparation, and Intervention**
  - Thu 8:30-9:45am, Room 29 C

# Thank you…

- Please fill out the session survey in your mobile app
- Select session in the Mobile App
  - Use the Search Feature to quickly find this title
- Click “Technical Workshop Survey”
- Answer a few short questions and enter any comments



# Questions / Feedback?

**To learn more:**
**http://links.esri.com/geoevent**
**https://links.esri.com/geoevent-forum**


RJ Sunderman
Product Engineer
GeoEvent Extension
rjsunderman@esri.com



C. Adam Mollenkopf
Real-Time GIS
Capability Lead
amollenkopf@esri.com
@amollenkopf


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