GeoEvent Server: Scaling and Resiliency

Mark Bramer

Esri Professional Services

Vienna, VA

DevSummit DC
February 15, 2017 | Washington, DC

Agenda

- Real-time GIS at Esri
- Deployment Essentials
- Performance
- Event Hubs
- Kafka Connector for GeoEvent
- GeoEvent Resiliency and Scalability
- Stream Service Resiliency and Scalability
- Spatiotemporal Big Data Store Resiliency and Scalability

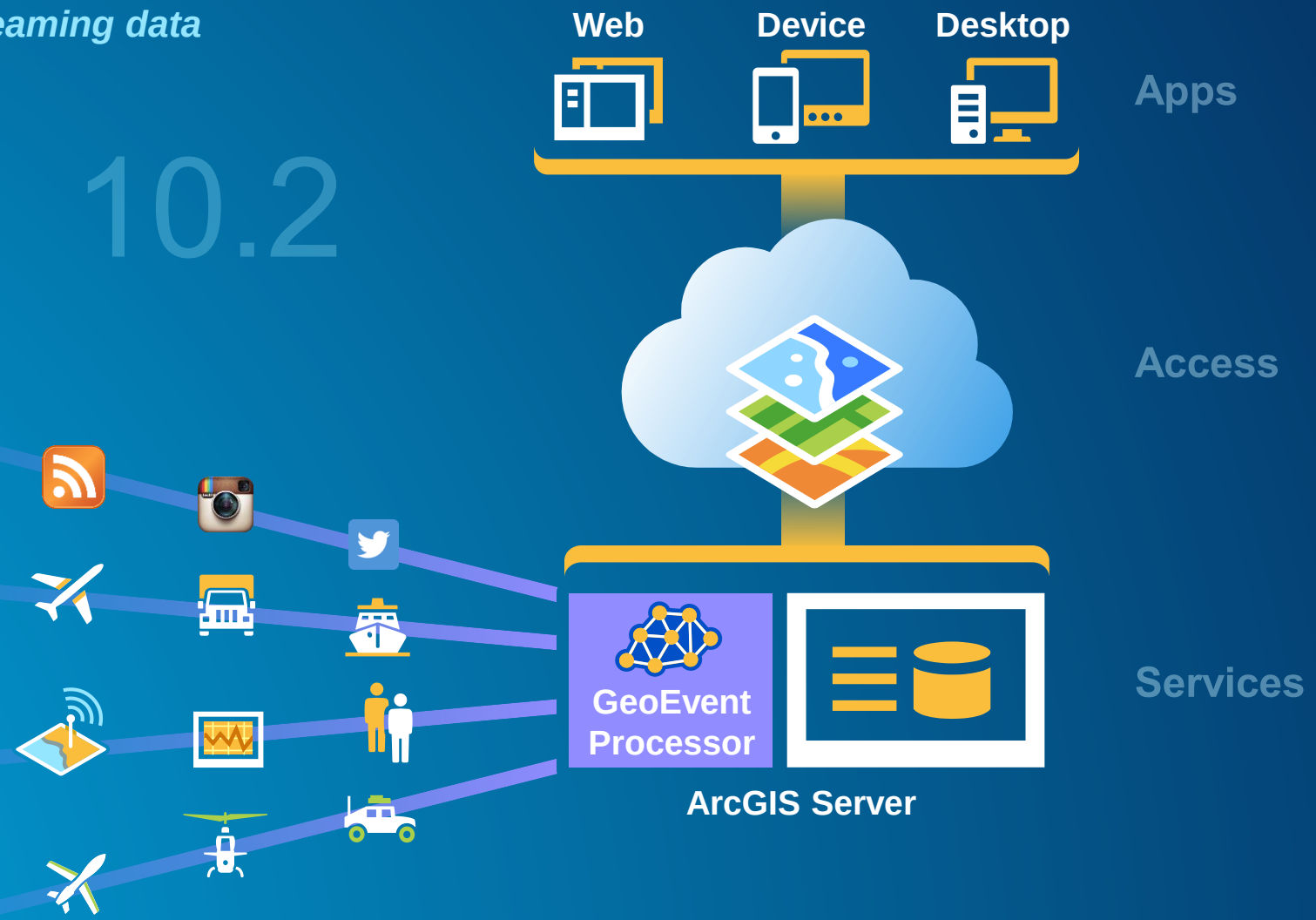
Real-Time GIS at Esri

Real-Time GIS

Integration and exploitation of streaming data

10.2

- 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

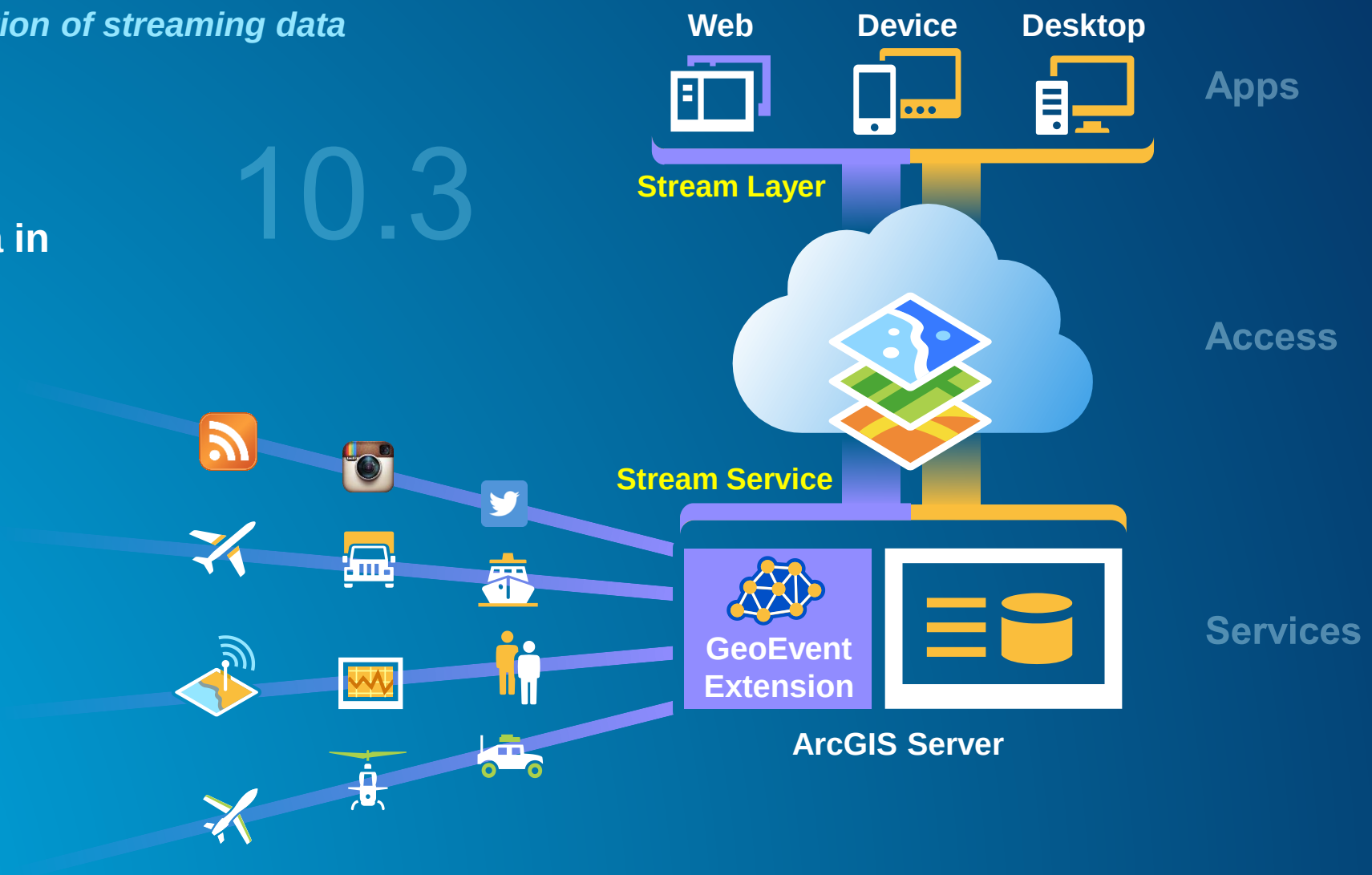


Real-Time GIS

Integration and exploitation of streaming data

- **Efficient real-time display of live data in web applications**

10.3

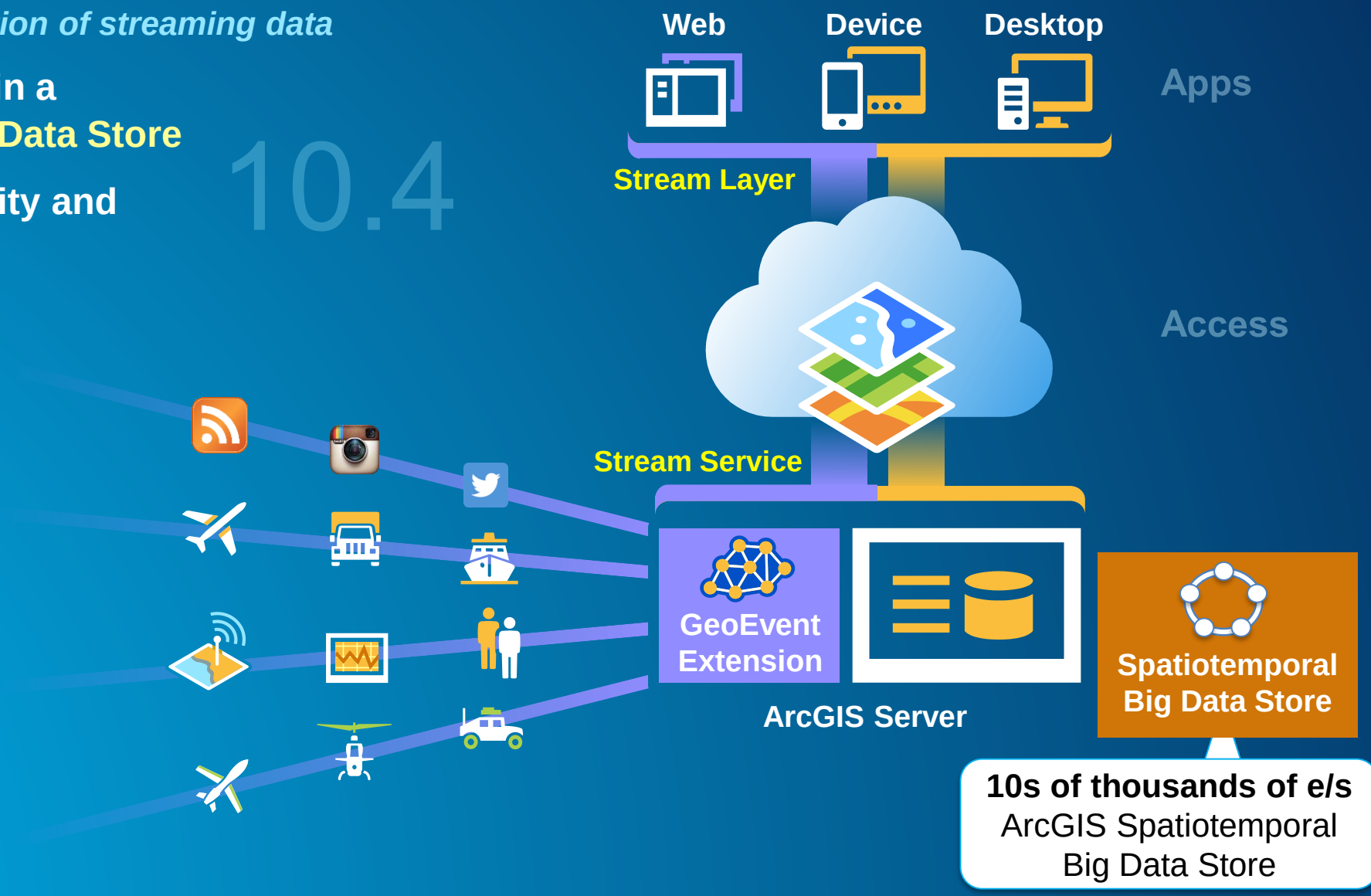


Real-Time GIS

Integration and exploitation of streaming data

- **Store observations in a Spatiotemporal Big Data Store**
- **Visualize high velocity and volume data:**
  - **as an aggregation**
  - **as discrete features**

10.4

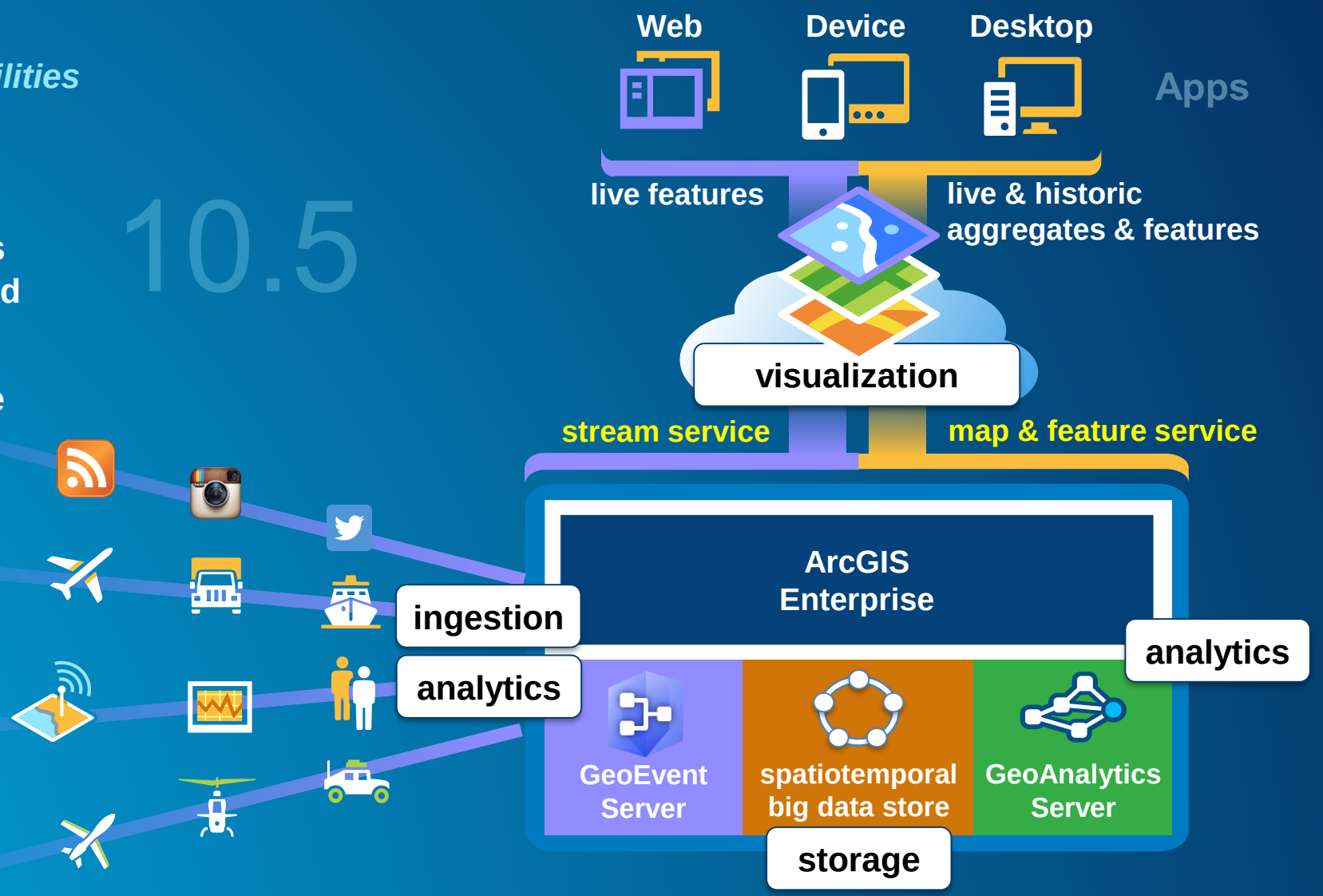


ArcGIS Enterprise

with real-time & big data 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
- Run analytics on stored observations
- Visualize high velocity & volume data:
  - as an aggregation
  - as discrete features
- Notify those who need to know about patterns of interest

10.5

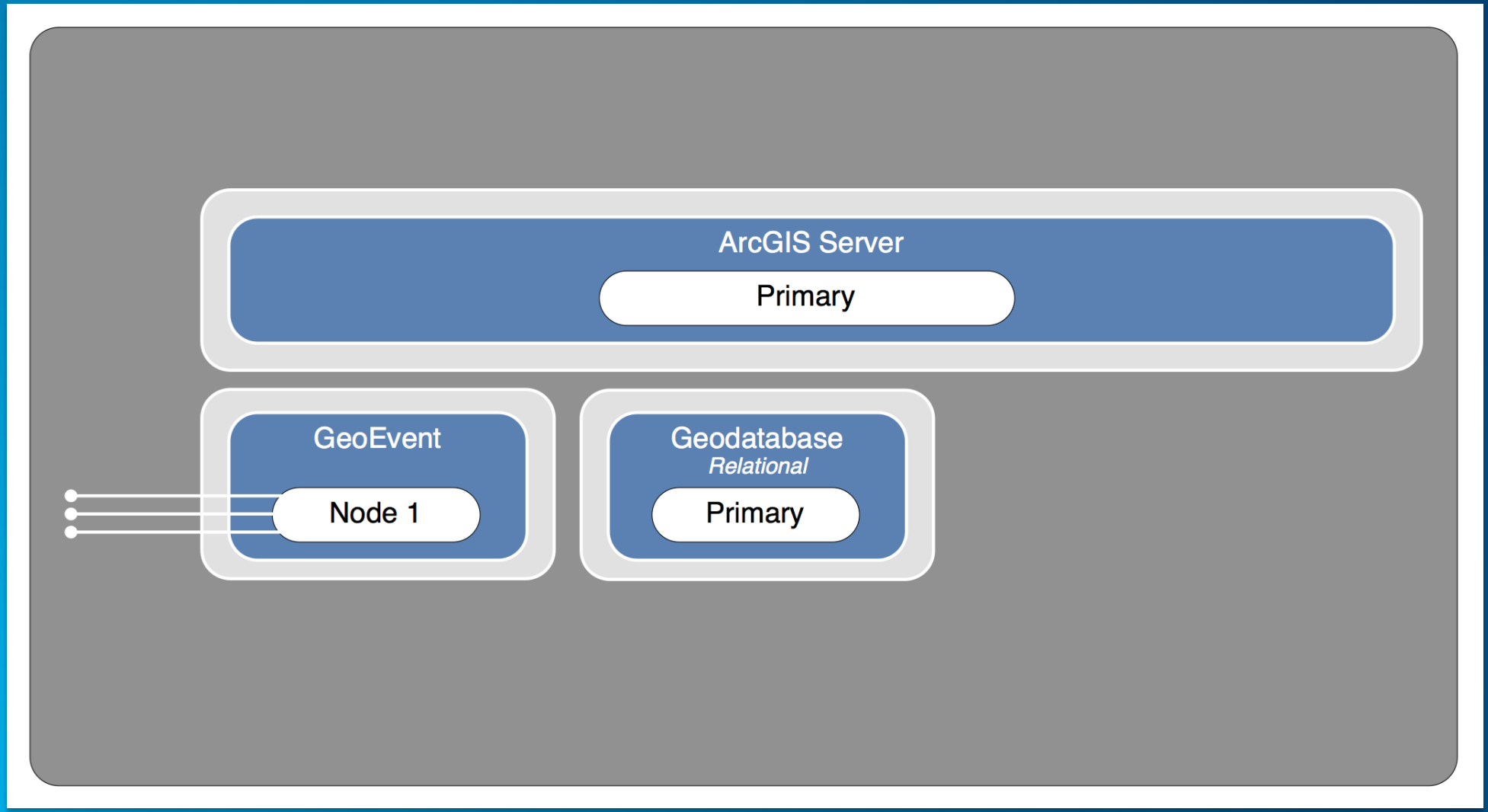


Deployment Essentials

Real-Time GIS Deployment

Basic

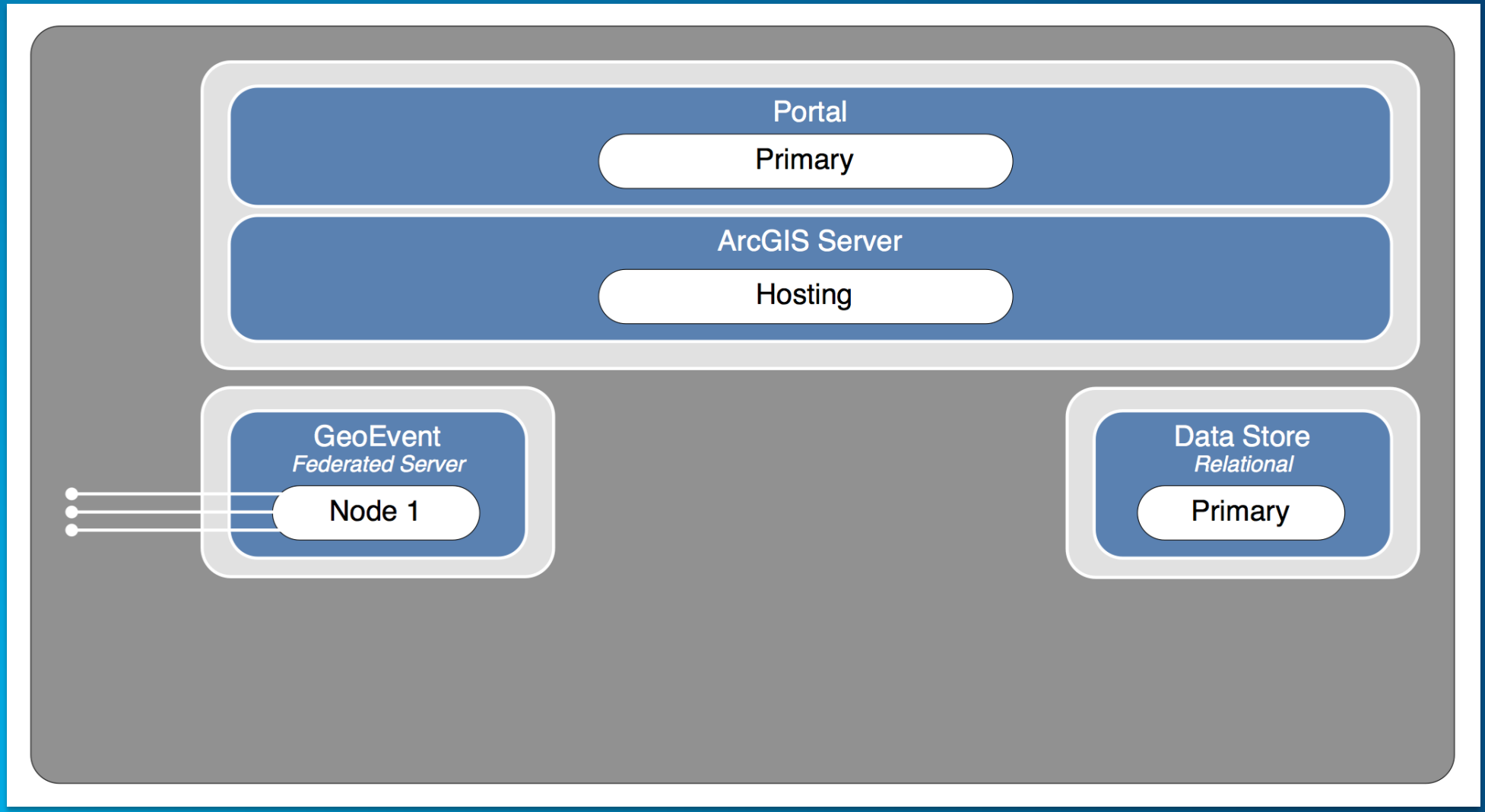
>= 10.2



Real-Time GIS Deployment

Standard

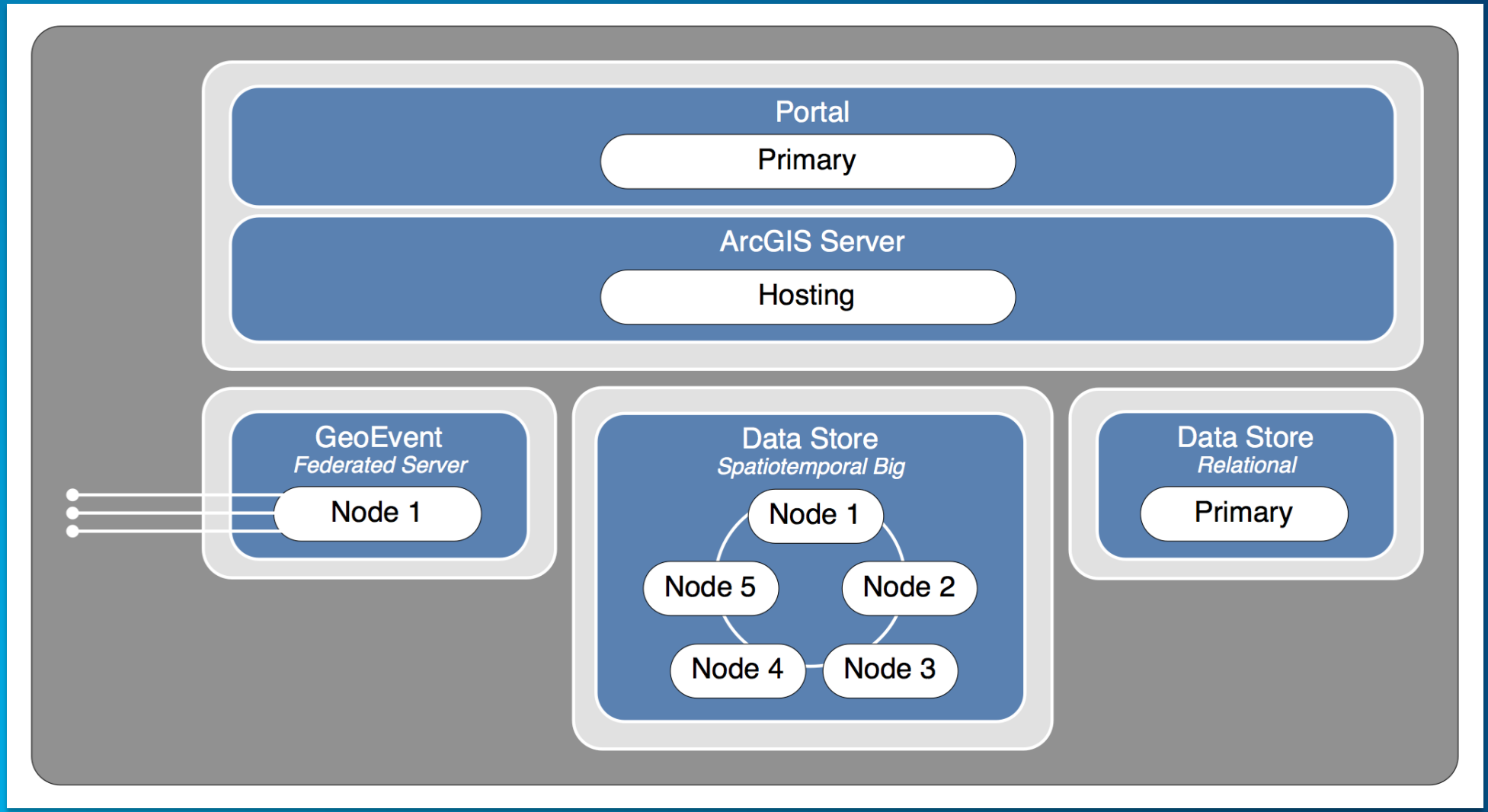
>= 10.3



Real-Time GIS Deployment

Standard

10.4+



Performance

Real-Time Throughput and Analytics Performance

What is performance?

- **“How many messages per second can GeoEvent support?”**
  - **There is no single or practical answer**
- **Performance is hard to measure**
- **Performance questions are harder to answer**
- **Many, many permutations of contributing factors**

Real-Time Throughput and Analytics Performance

Some basics to consider

- Message Size
  - Number of attributes
  - Size of geometry
- Inbound Transport Protocol
- Analytics
- Data Write/Persistence
- Outbound Transport Protocol
- Hardware
  - Network speed
  - RAM (size, generation, type)
  - CPU (type, speed, quantity)

Real-Time Throughput and Analytics Performance

Message Size

- **Much faster to process a small message**

- `1,38.8,142.8`
  - **12 byte message**

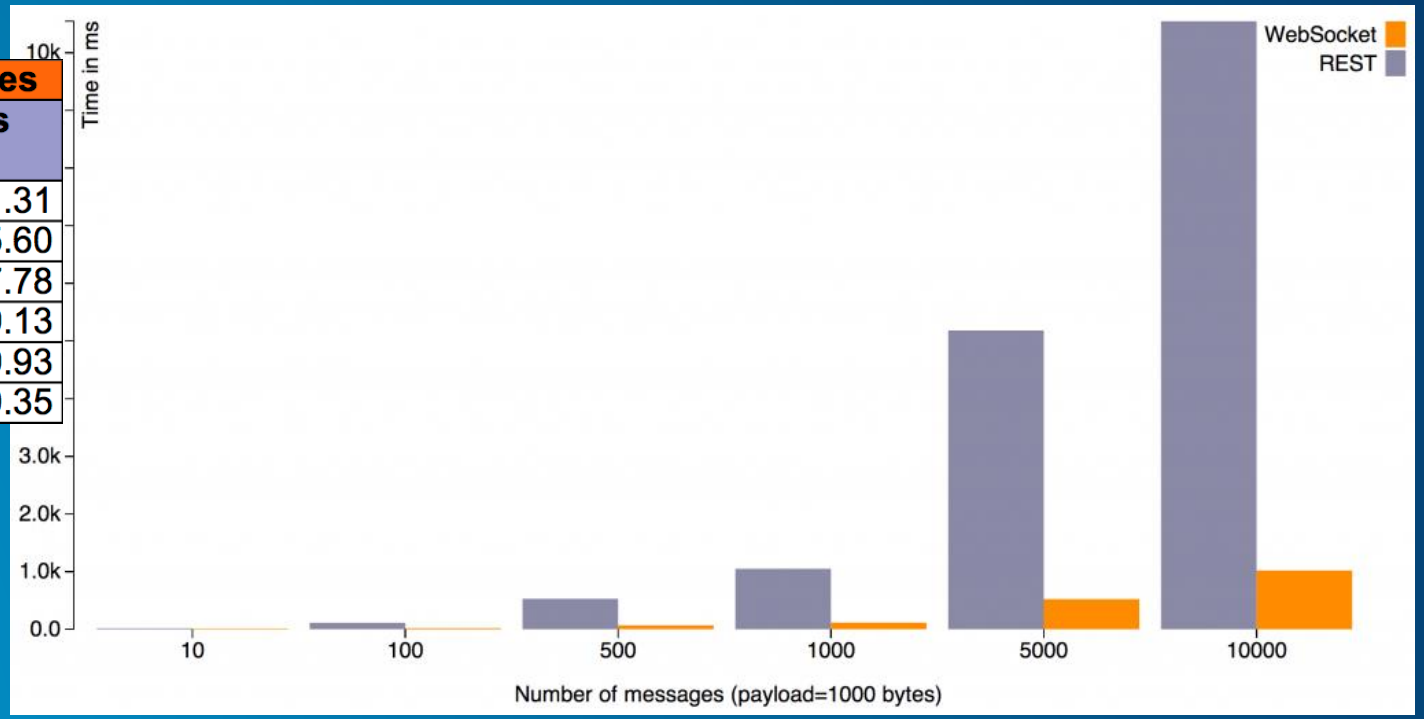
- `{"rings" : [ [ [-97.06138,32.837], [-97.06133,32.836], [-97.06124,32.834], [-97.06127,32.832], [-97.06138,32.837] ], [ [-97.06326,32.759], [-97.06298,32.755], [-97.06153,32.749], [-97.06326,32.759] ] ],"spatialReference" : {"wkid" : 4326}}`
  - **240 byte message (20x larger, and that's just geometry)**

Real-Time Throughput and Analytics Performance

Inbound Transport Protocol

- **Transport protocols vastly differ in performance**
- **TCP vs REST vs WebSocket vs etc**

Constant payload, increasing number of messages			
Messages	**REST (in ms)**	**WebSocket (in ms)**	**x times**
10	17	13	1.31
100	112	20	5.60
500	529	68	7.78
1000	1050	115	9.13
5000	5183	522	9.93
10000	10547	1019	10.35

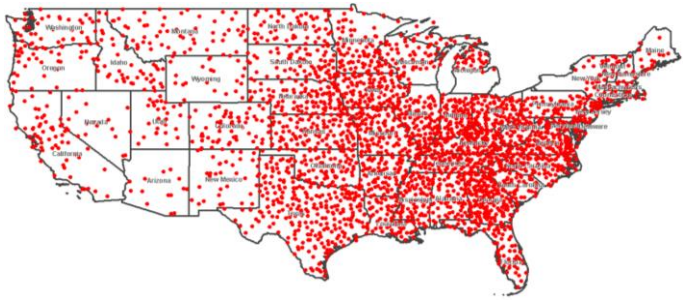


*** http://blog.arungupta.me/rest-vs-websocket-comparison-benchmarks/**

Real-Time Throughput and Analytics Performance

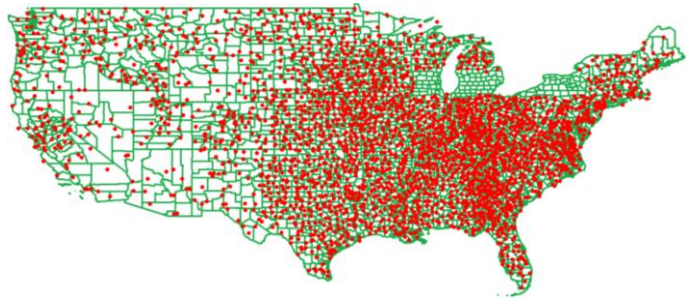
Analytics

Spatial Operator	US States Benchmarks	US Districts Benchmarks	US Counties Benchmarks
Disjoint Any	3,124 e/s	2,811 e/s	2,496 e/s
Inside Any	3,110 e/s	2,805 e/s	2,178 e/s
Intersect Any	3,108 e/s	2,810 e/s	2,495 e/s
Within Any	3,103 e/s	2,805 e/s	2,175 e/s
Touches Any	2,810 e/s	2,805 e/s	2,184 e/s
Outside Any	2,806 e/s	2,810 e/s	2,490 e/s



US States
51 geofences
1,617 vertices on average
78 min, 21,970 max

US Districts
436 geofences
512 vertices on average
24 min, 7,285 max



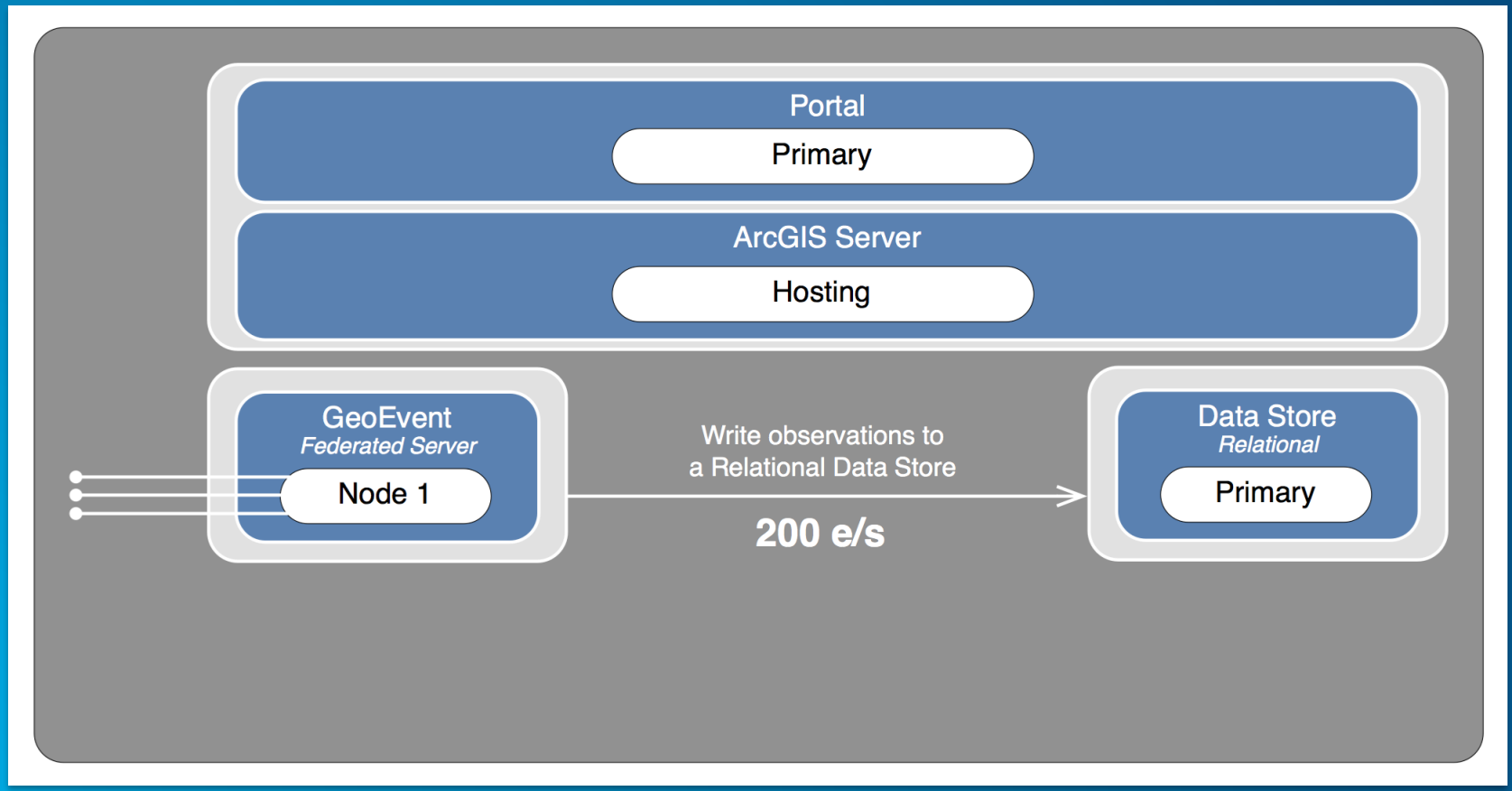
US Counties
3,143 geofences
166 vertices on average
9 min, 838 max

Real-Time Throughput and Analytics Performance

>= 10.3

Data Write/Persistence

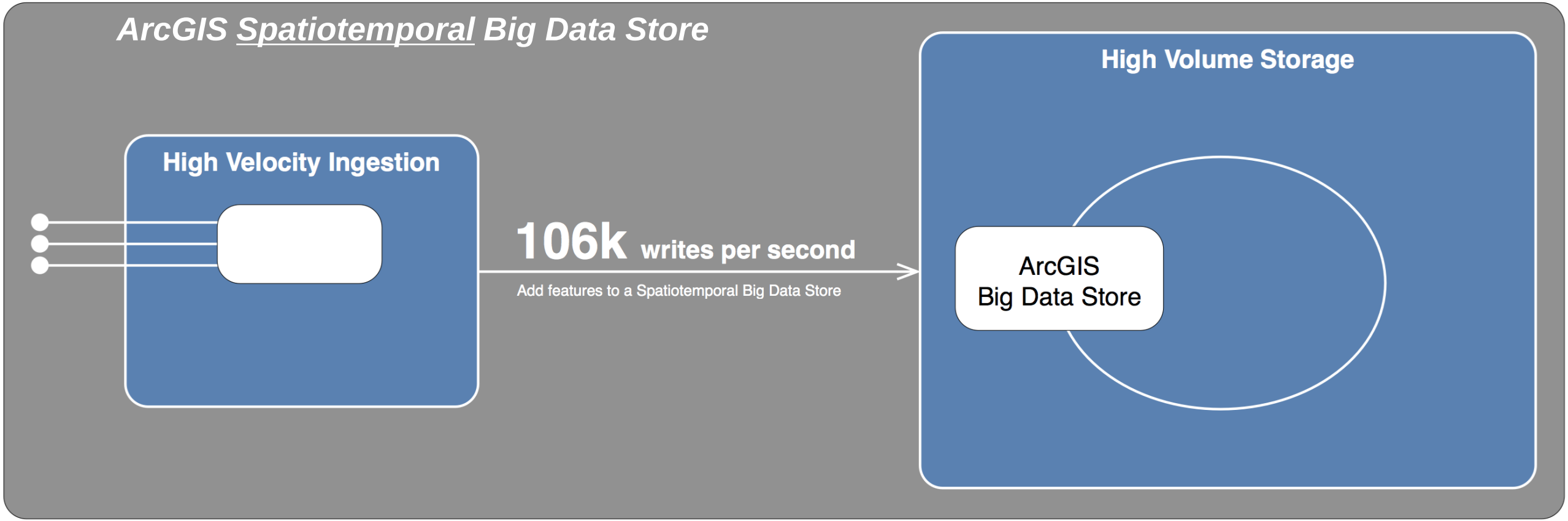
ArcGIS Relational Data Store



Real-Time Throughput and Analytics Performance

Data Write/Persistence

10.4



Storage Write Throughput	1 node
ArcGIS Big Data Store	106k
ArcGIS Relational Data Store	0.2k

Real-Time Throughput and Analytics Performance

Stream Services

- **Like inbound protocols, outbound protocols vastly differ in performance**
- **Example: sending outbound data via websockets will always be magnitudes faster then SMTP (emails)**
- **Stream services (websocket-based, no writing of data to disk) can sustain 3,000 msgs/sec**

Real-Time Throughput Performance

ArcGIS 10.4 Benchmarks

- **Real-Time GIS *throughput** performance:**
  - **ArcGIS 10.2: 600 e/s**
  - **ArcGIS 10.3: 3,000 e/s**
  - **ArcGIS 10.4: 4,000 e/s**

- **Benchmarks were captured by:**
  - **A physical machine running ArcGIS Server + GeoEvent Extension**
    - **Windows 7 (64-bit)**
    - **3.70 GHz, 4 physical cores (8 virtual cores)**
    - **16GB RAM**
    - **1 Gbps network**

*** Throughput being general raw geoevent processing, using fastest inputs**

Real-Time Throughput and Analytics Performance

What are the primary factors I should 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 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)*

Event Hubs

Event Hubs

What is an Event Hub?

- In context of this talk, an “event hub” is a messaging system, a.k.a.
  - Message Queue
  - Message Broker
- Examples:
  - RabbitMQ
  - ActiveMQ
  - Kafka
  - HornetQ
  - Amazon SQS
- For GeoEvent scaling and resiliency, Kafka is recommended

Message Queues

The Fundamentals

- Applications communicate via messages rather than direct communication
- Messages placed on queues (think mailboxes)
- Applications run independently of each other
  - Different speeds and times and locations
  - No logical connection between applications
- Scale out with more than one system
  - Applications compete to retrieve messages from queue
- Add resiliency with more than one system
  - If “A” becomes unavailable, “B” and “C” can continue processing messages

Kafka

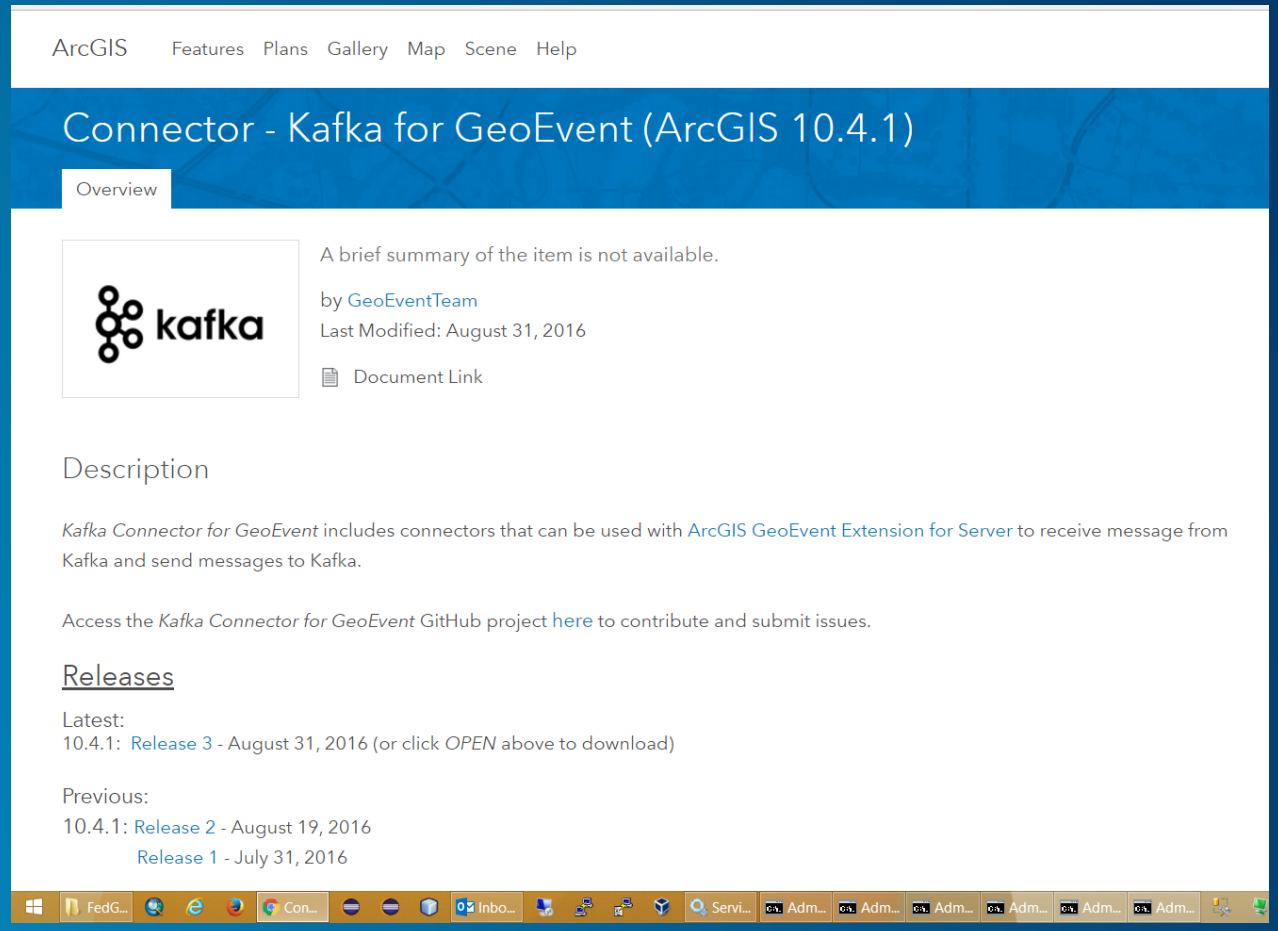
The Fundamentals

- Started at LinkedIn
- Distributed publish-subscribe messaging system
- Designed to be fast, scalable, and durable
- Maintains feeds of messages in topics
- Producers write data to topics
- Consumers read from topics
- Topics are partitioned and replicated across multiple nodes

Kafka Connector for GeoEvent

Kafka Connector for GeoEvent

- **Free download on GeoEvent Gallery**
- **Supports resiliency by subscribing to Kafka topics**
- **Able to be deployed as-is**
  - **No need to compile, build, etc**

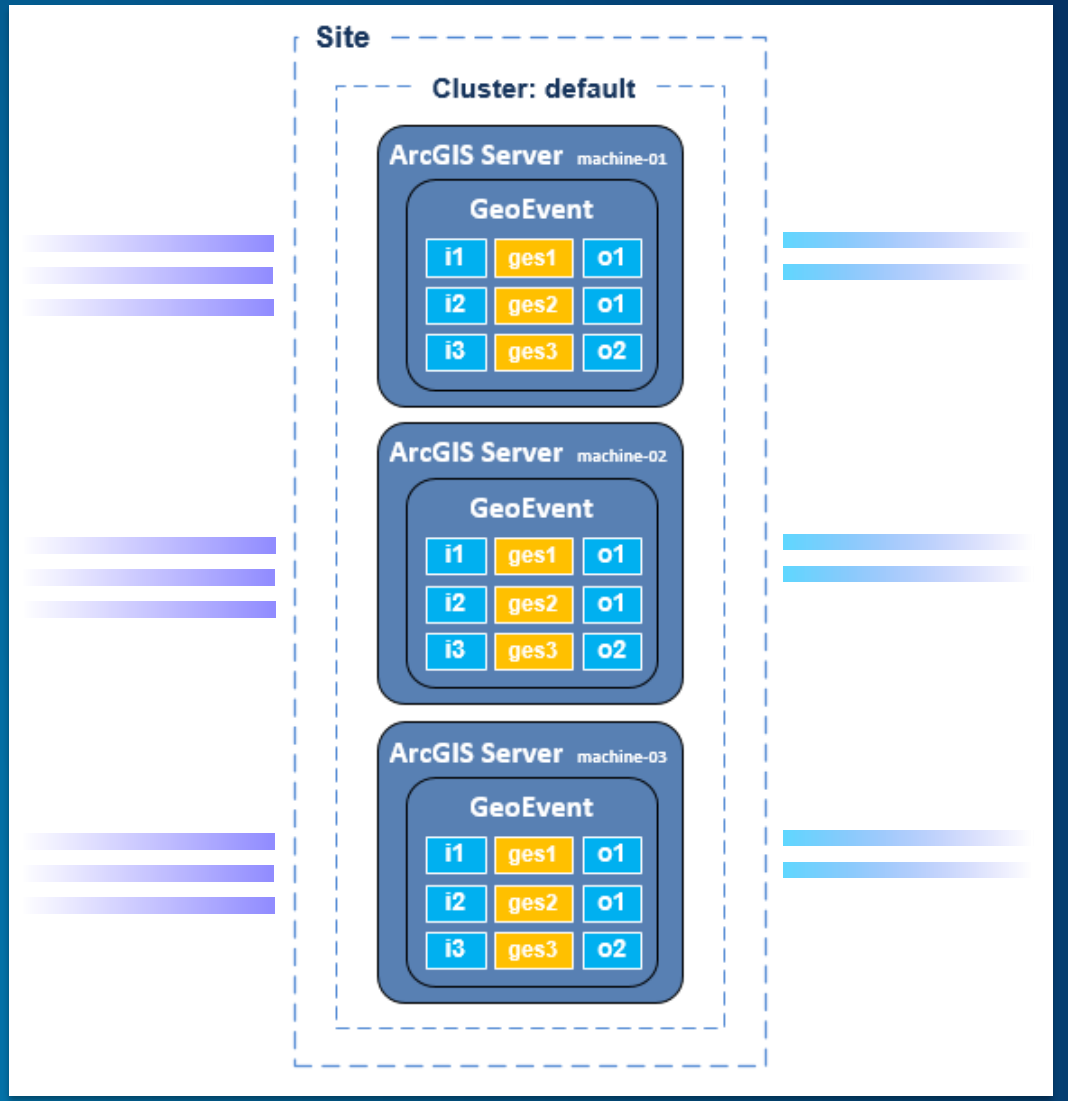


GeoEvent Resiliency and Scalability

GeoEvent Resiliency & Scalability

Clustered Deployment

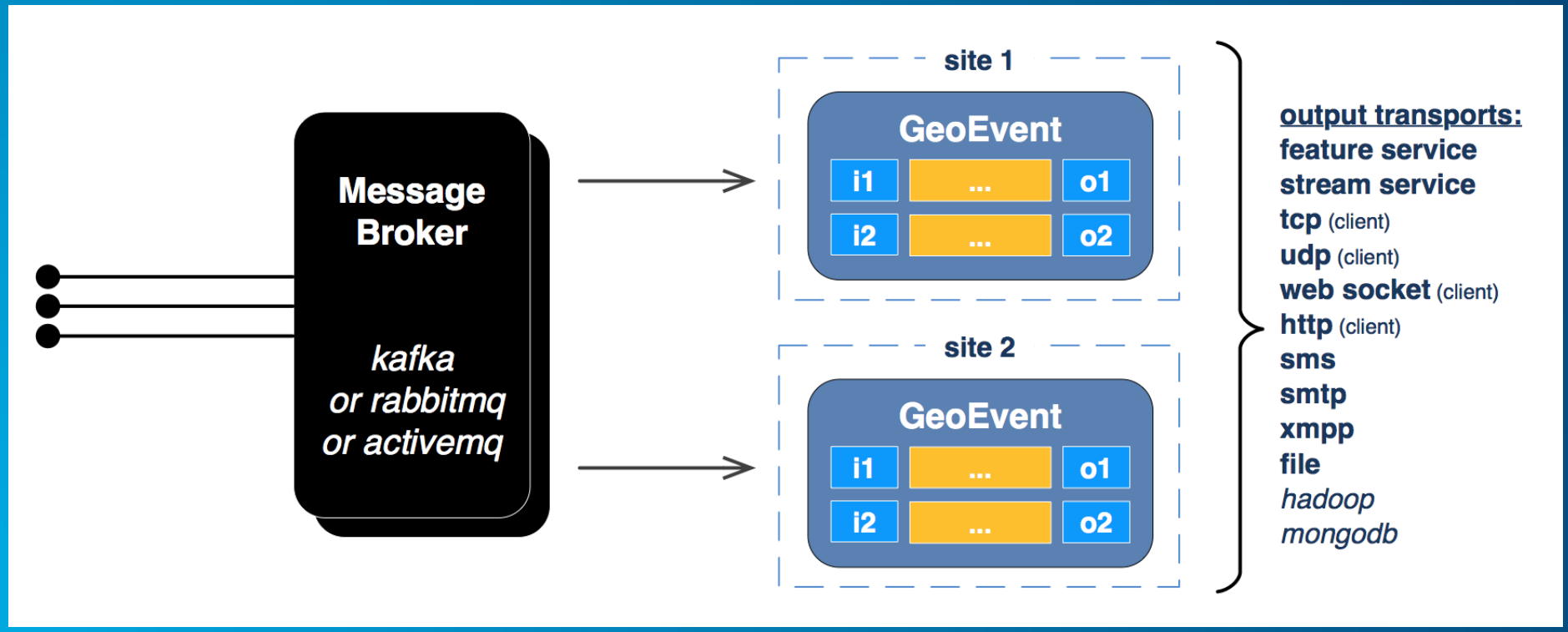
- GeoEvent clustering introduced at 10.3
- However, observed to be brittle and subject to failure
  - Must have at least 3 nodes in a cluster
    - If node goes down, must be brought back up ASAP
    - 2-node cluster can cause a split-brain condition in cluster
  - Must be physical hardware
    - Many customers only have virtualized environments
  - Network must be stable
    - Network instabilities can cause intermittent membership changes
- Given this, clustering GeoEvent should be avoided
- Instead, if resiliency or addt'l scalability is required:
  - An isolated deployment model should be employed
  - This is true at 10.3, 10.4 and 10.5



GeoEvent Resiliency & Scalability

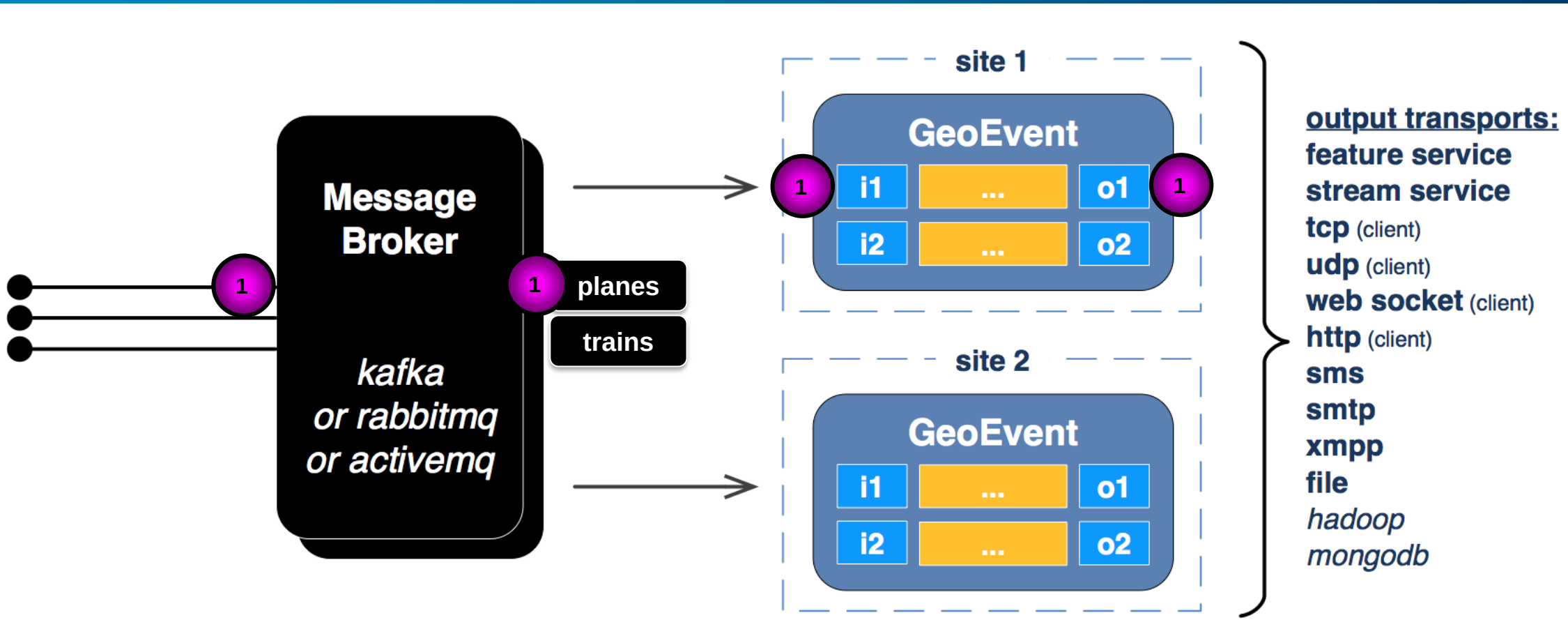
Isolated Deployment

- Each GeoEvent instance exists in its own site and shares a common configuration manually
- A message broker exists in front to provide a common event source
- GeoEvent instances share a common message broker and run active/active competing to consume events



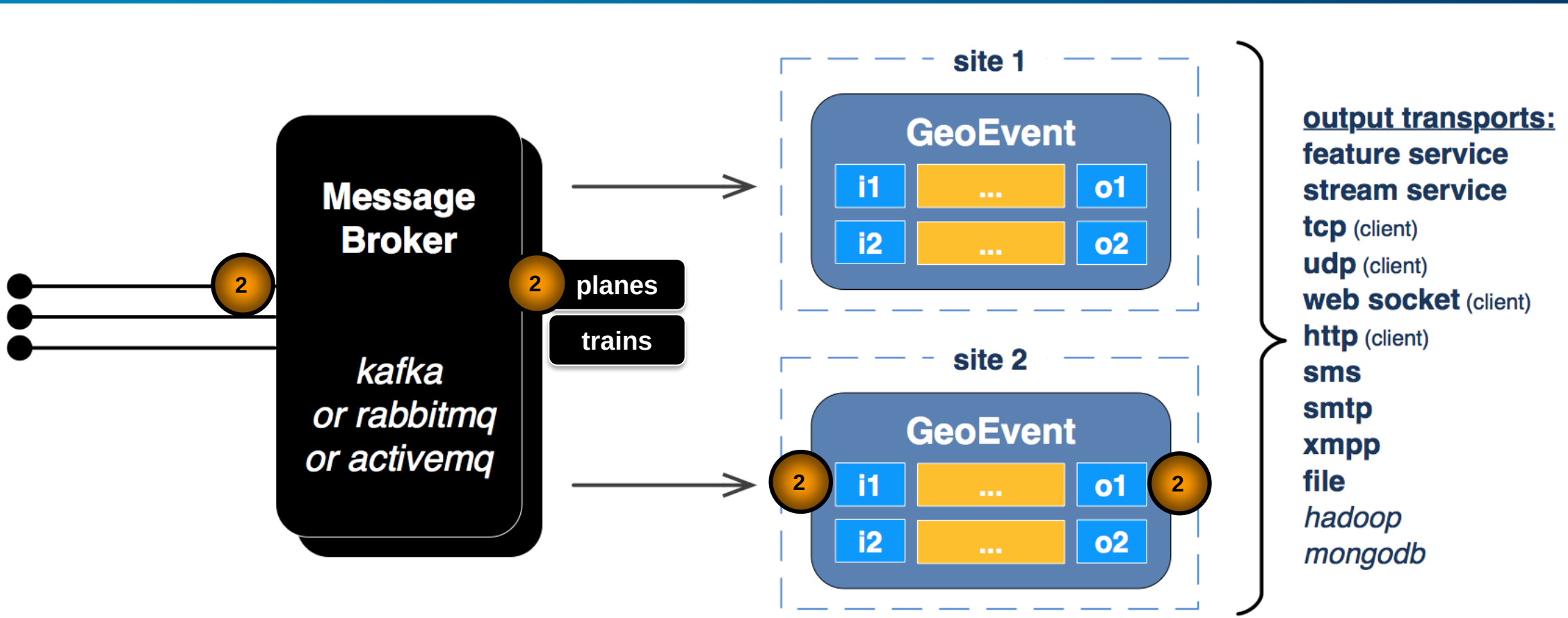
GeoEvent Resiliency & Scalability

Isolated Deployment



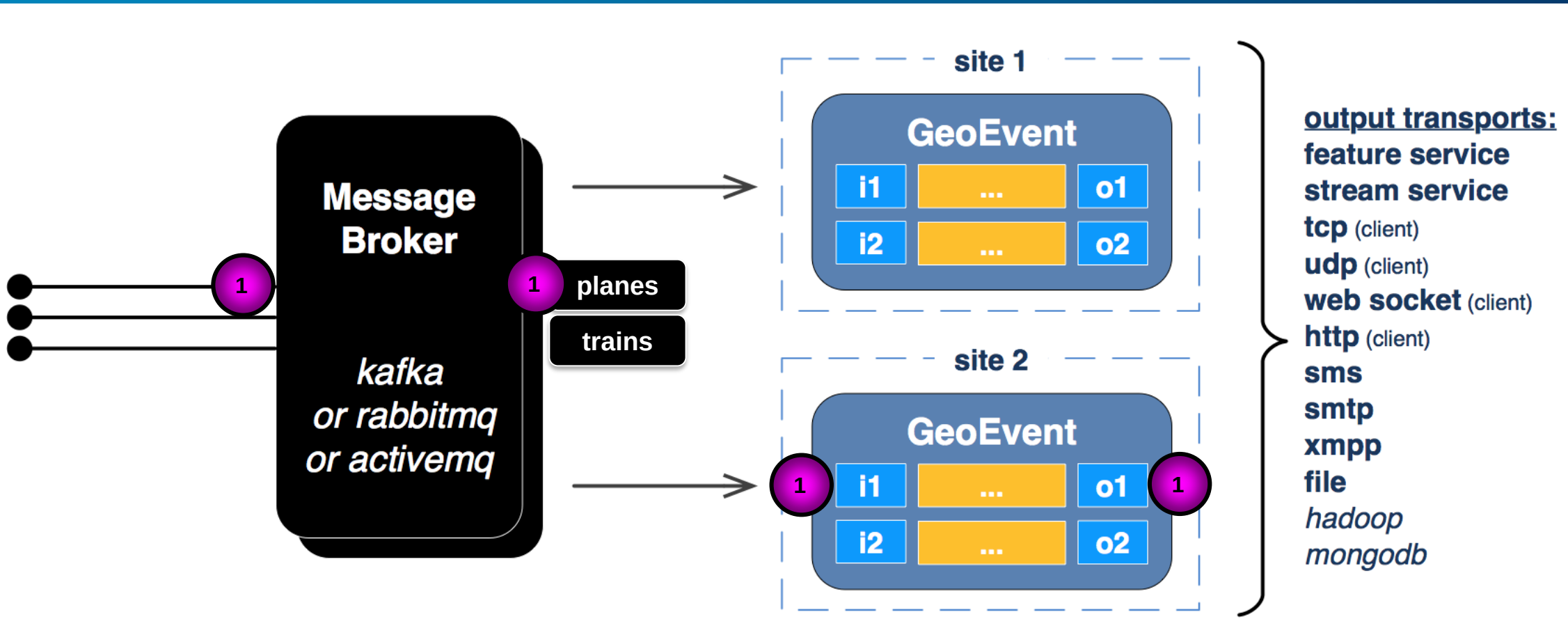
GeoEvent Resiliency & Scalability

Isolated Deployment



GeoEvent Resiliency & Scalability

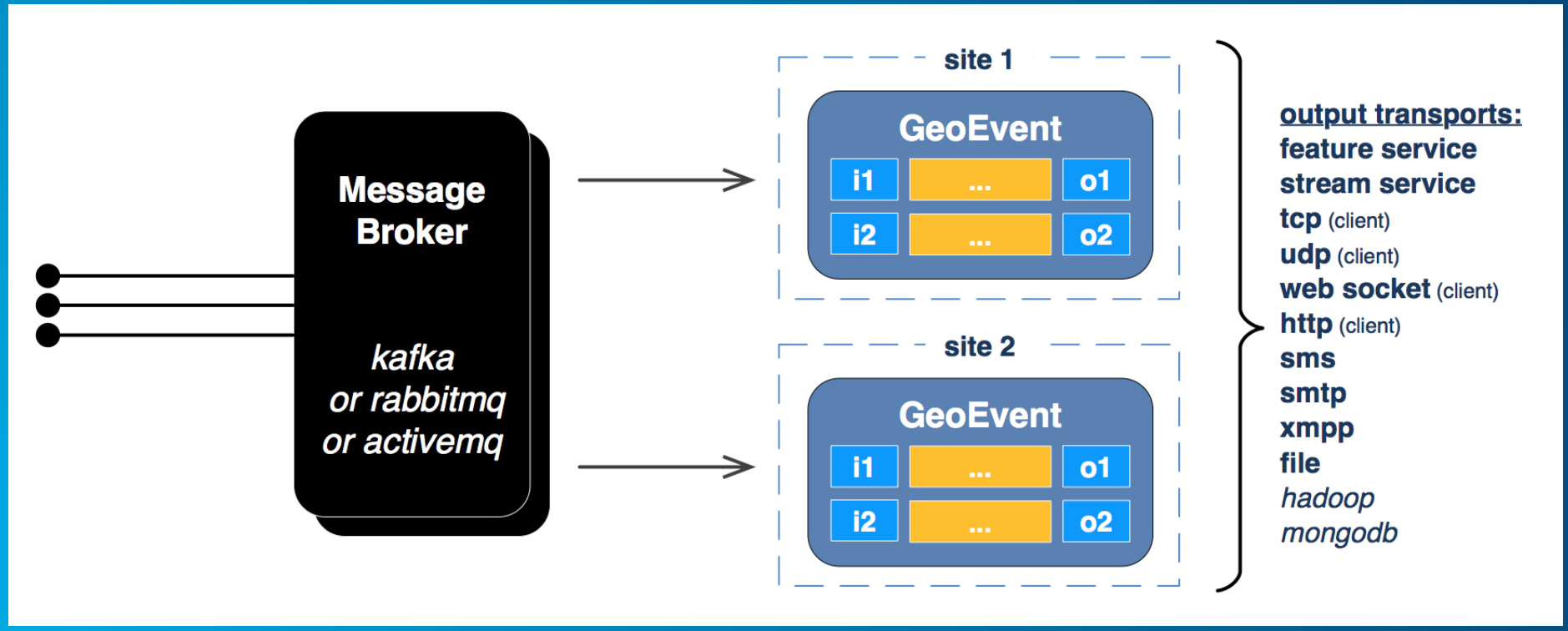
Isolated Deployment



GeoEvent Resiliency& Scalability

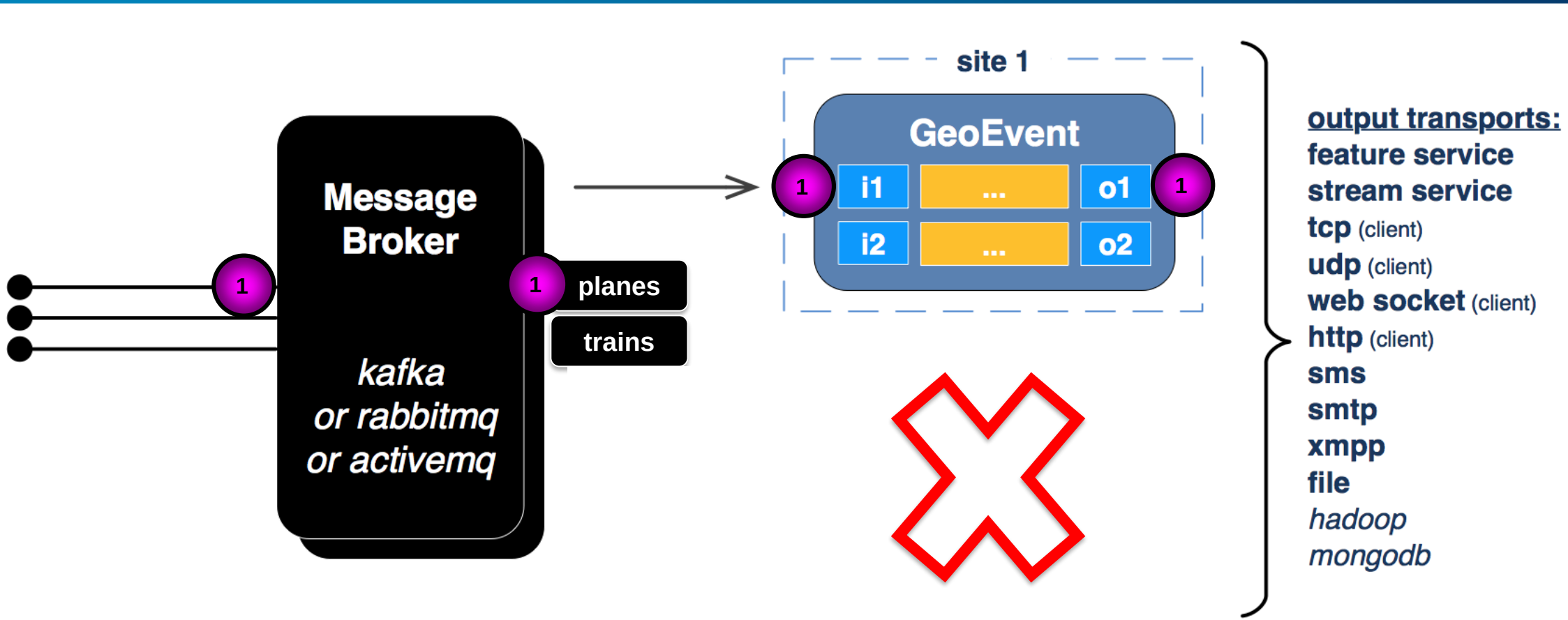
Isolated Deployment

- **Known constraints:**
  - **Does not support stateful processing: enter, exit, track gap detection, idle detector**
  - **Additional setup/config of message broker and sending data into it**
  - **Input to GeoEvent is the message broker only, lose flexibility of all the input connectors available**



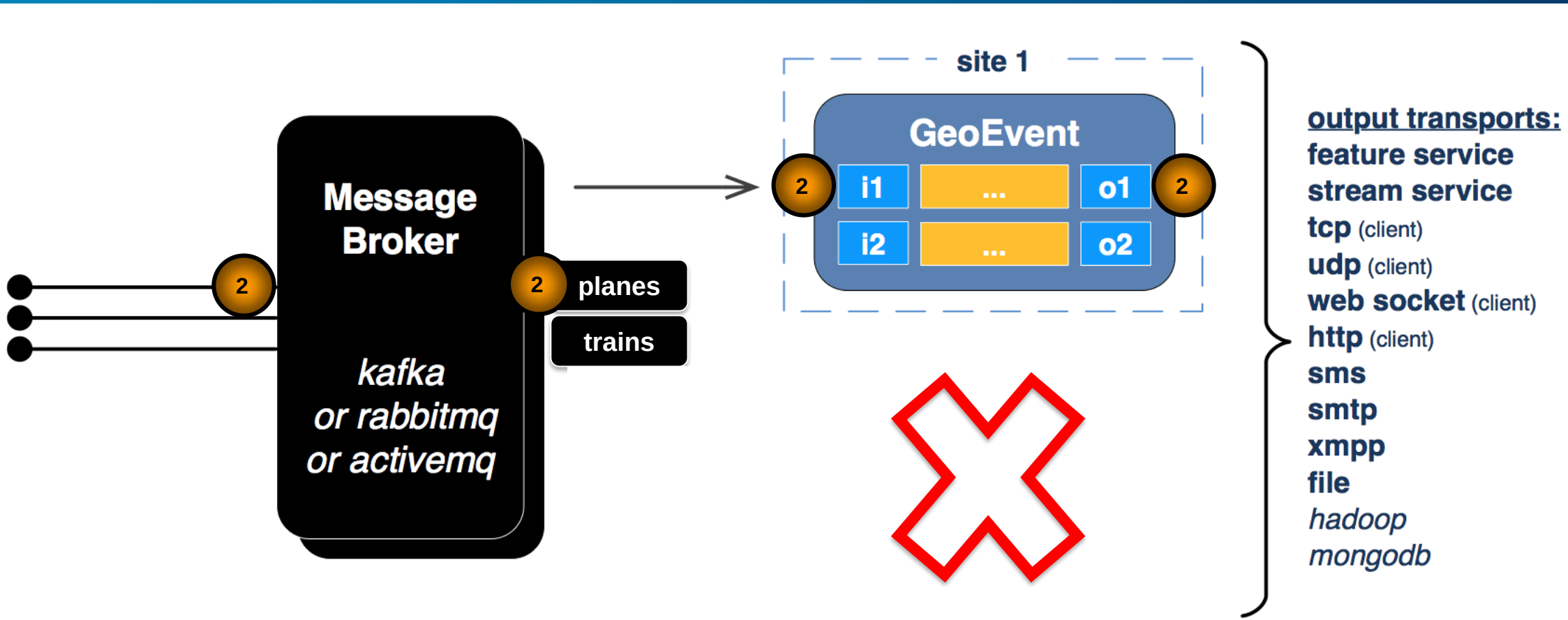
GeoEvent Resiliency & Scalability

Isolated Deployment



GeoEvent Resiliency & Scalability

Isolated Deployment

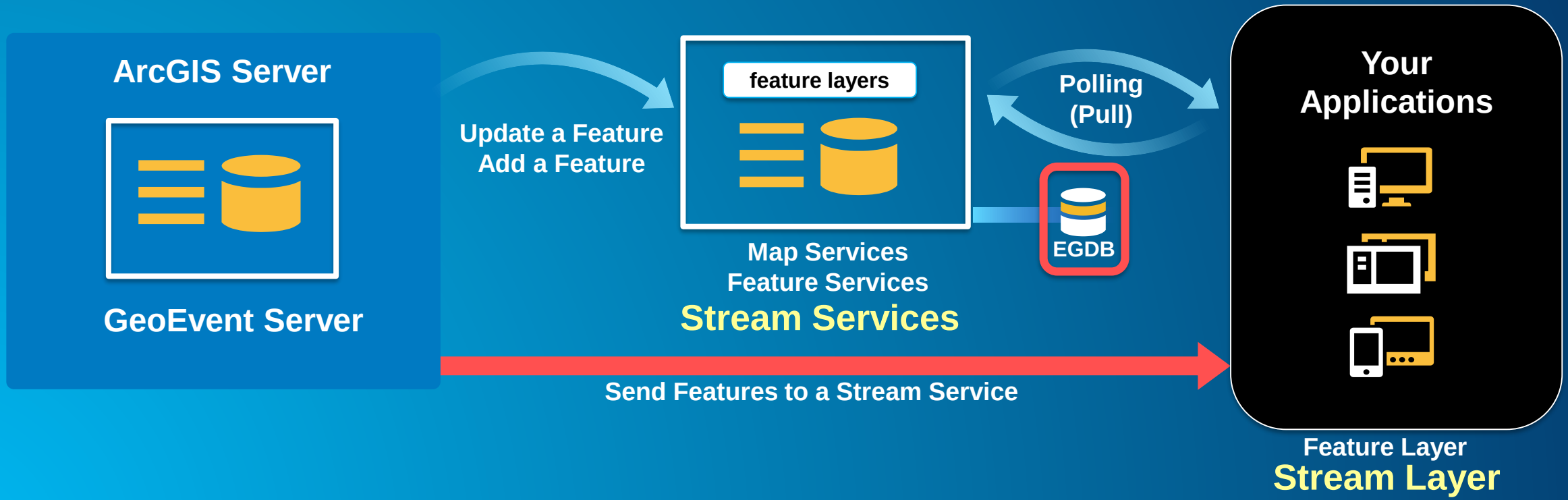


Stream Service Resiliency and Scalability

Stream Services

What are they?

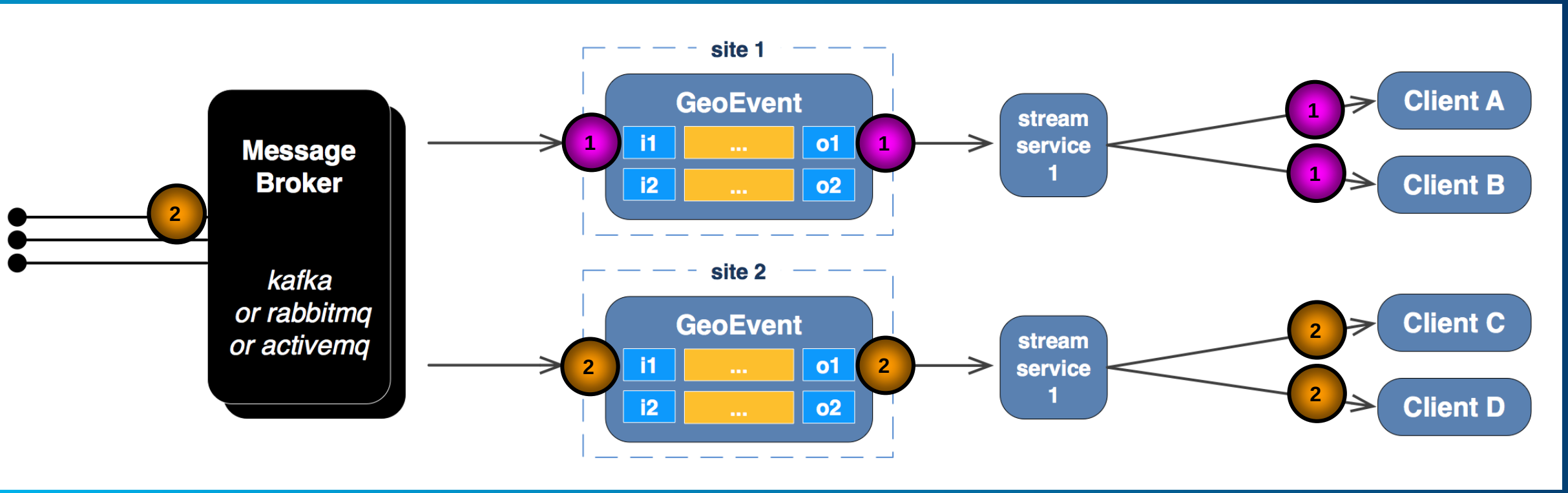
- Feature services persist their data in a Geodatabase
- Stream services broadcast their data without first persisting the data
  - Based on HTML5 websocket technology
  - Very fast



Stream Service Resiliency & Scalability

Isolated Deployment

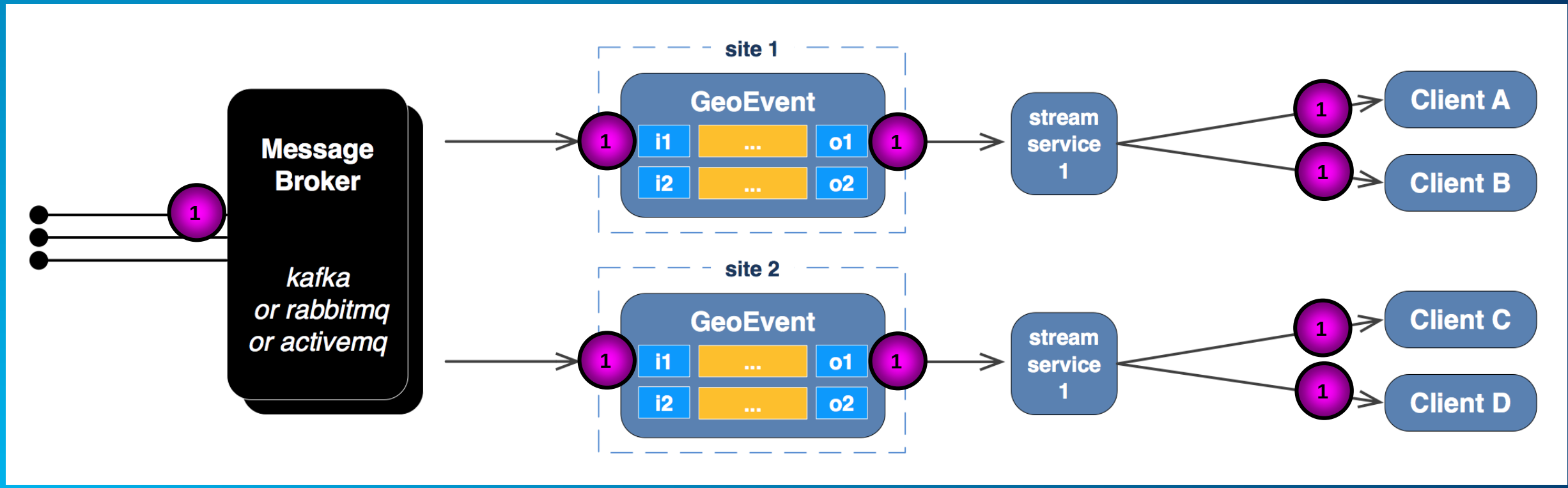
- **Isolated deployment of GeoEvent instances leads to challenges with Stream Services:**
  - **Client A & B see event 1, while client C & D do not**
  - **Client C & D see event 2, while client A & B do not**



Stream Service Resiliency & Scalability

Isolated Deployment

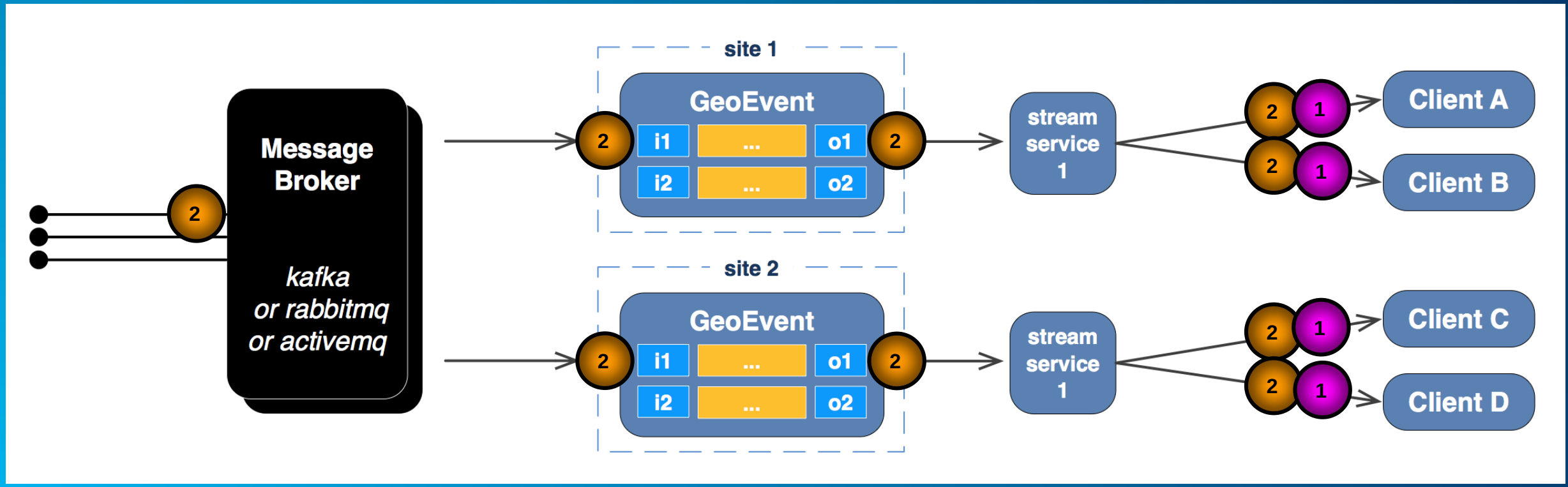
- **Using Kafka, GeoEvent instances can be configured to use separate consumer groups**
  - **With this configuration, all clients see all events**
  - ***Note: can also be accomplished with RabbitMQ fan-outs or ActiveMQ topics***



Stream Service Resiliency & Scalability

Isolated Deployment

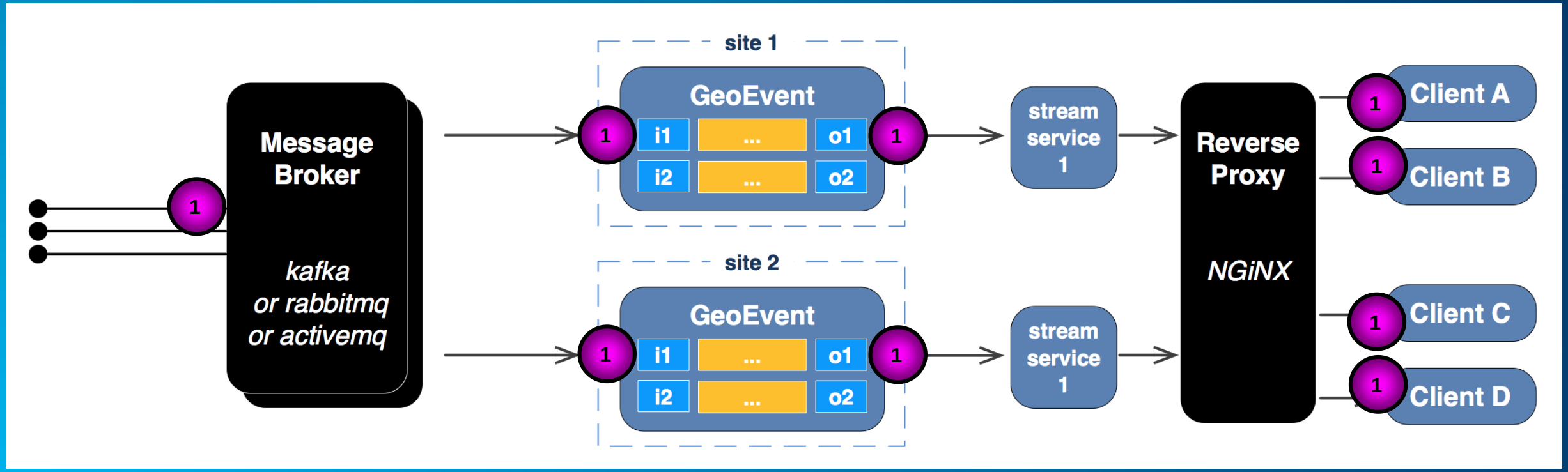
- **Using Kafka, GeoEvent instances can be configured to use separate consumer groups**
  - **With this configuration, all clients see all events**
  - ***Note: can also be accomplished with RabbitMQ fan-outs or ActiveMQ topics***



Stream Service Resiliency & Scalability

Isolated Deployment

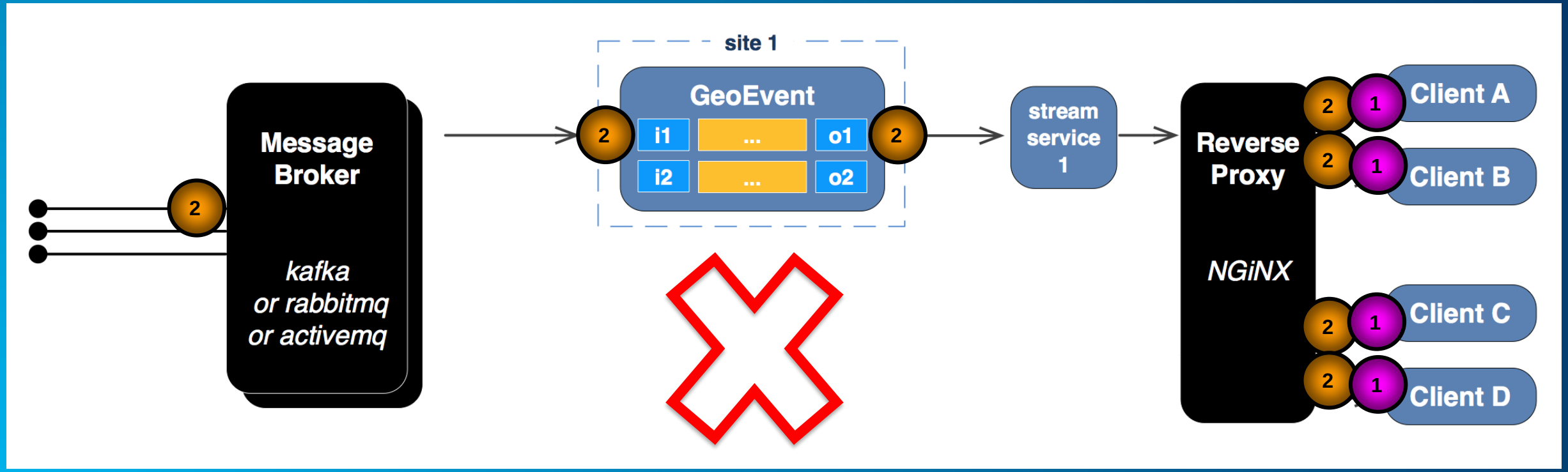
- A reverse proxy can be configured in between the clients and the stream services so that clients don't have direct knowledge of the servers they are connecting to.
  - NGiNX is one such reverse proxy



Stream Service Resiliency & Scalability

Isolated Deployment

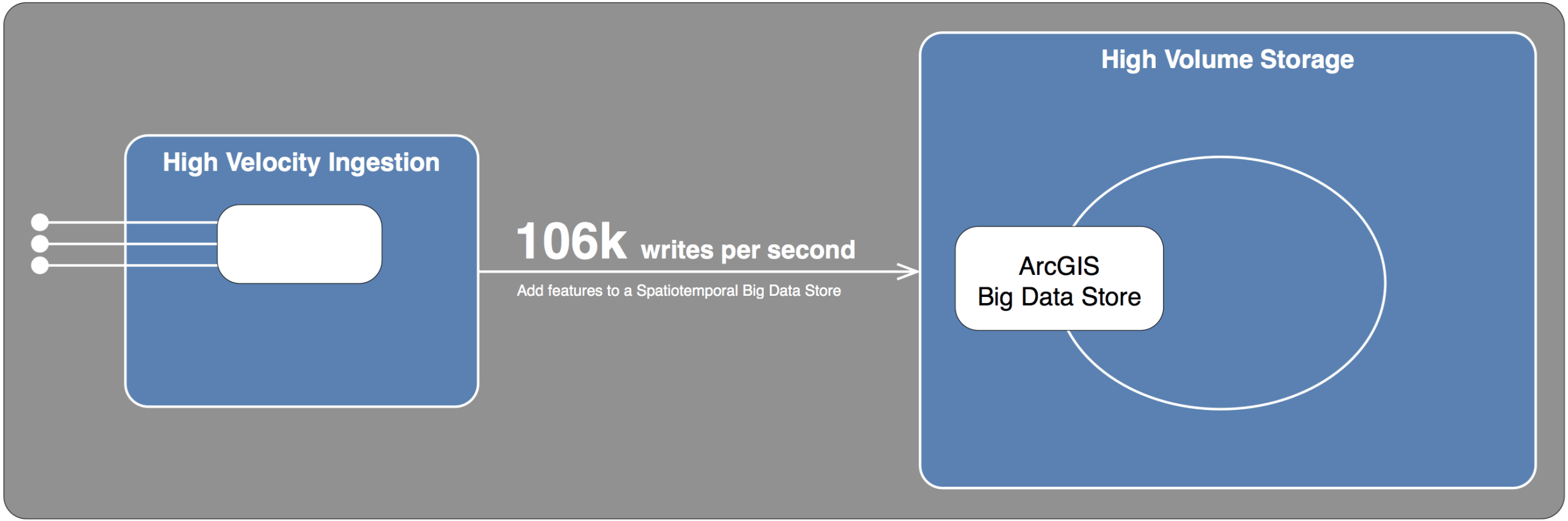
- **Using Kafka, GeoEvent instances can be configured to use separate consumer groups**
  - **With this configuration, all clients see all events**
  - ***Note: can also be accomplished with RabbitMQ fan-outs or ActiveMQ topics***



Spatiotemporal Big Data Store Resiliency & Scalability

Spatiotemporal Big Data Store Performance

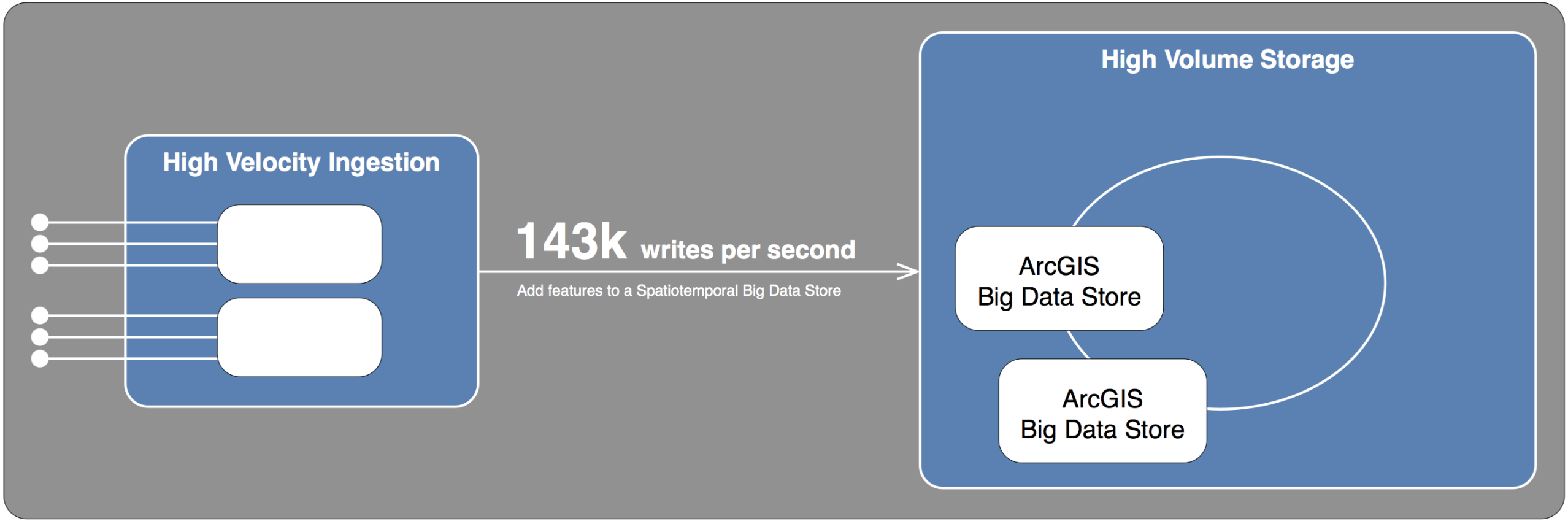
ArcGIS 10.4 Benchmarks *c4.2xlarge (Windows 2012 Server R2): 8 vCPU, 15 GiB, 100GB local SSD, 1,000 Mbps EBS*



Storage Write Throughput	1 node
ArcGIS Big Data Store	106k
ArcGIS Relational Data Store	0.2k

Spatiotemporal Big Data Store Performance

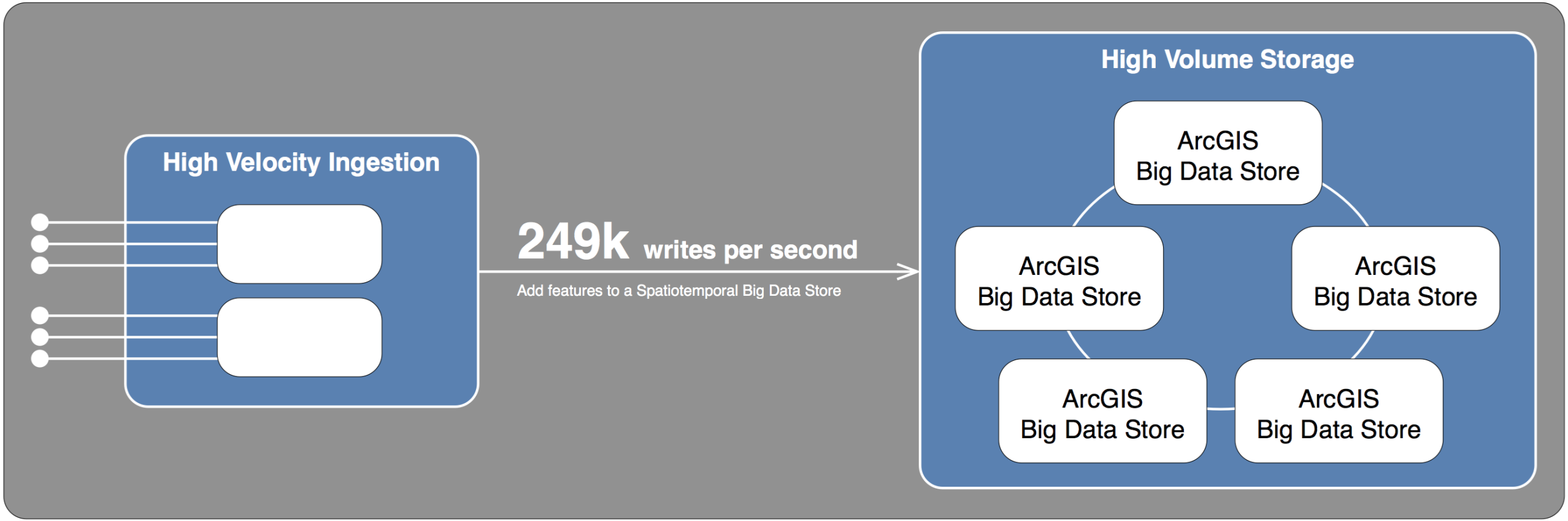
ArcGIS 10.4 Benchmarks *c4.2xlarge (Windows 2012 Server R2): 8 vCPU, 15 GiB, 100GB local SSD, 1,000 Mbps EBS*



Storage Write Throughput	1 node	2 node
ArcGIS Big Data Store	106k	143k
ArcGIS Relational Data Store	0.2k	n/a

Spatiotemporal Big Data Store Performance

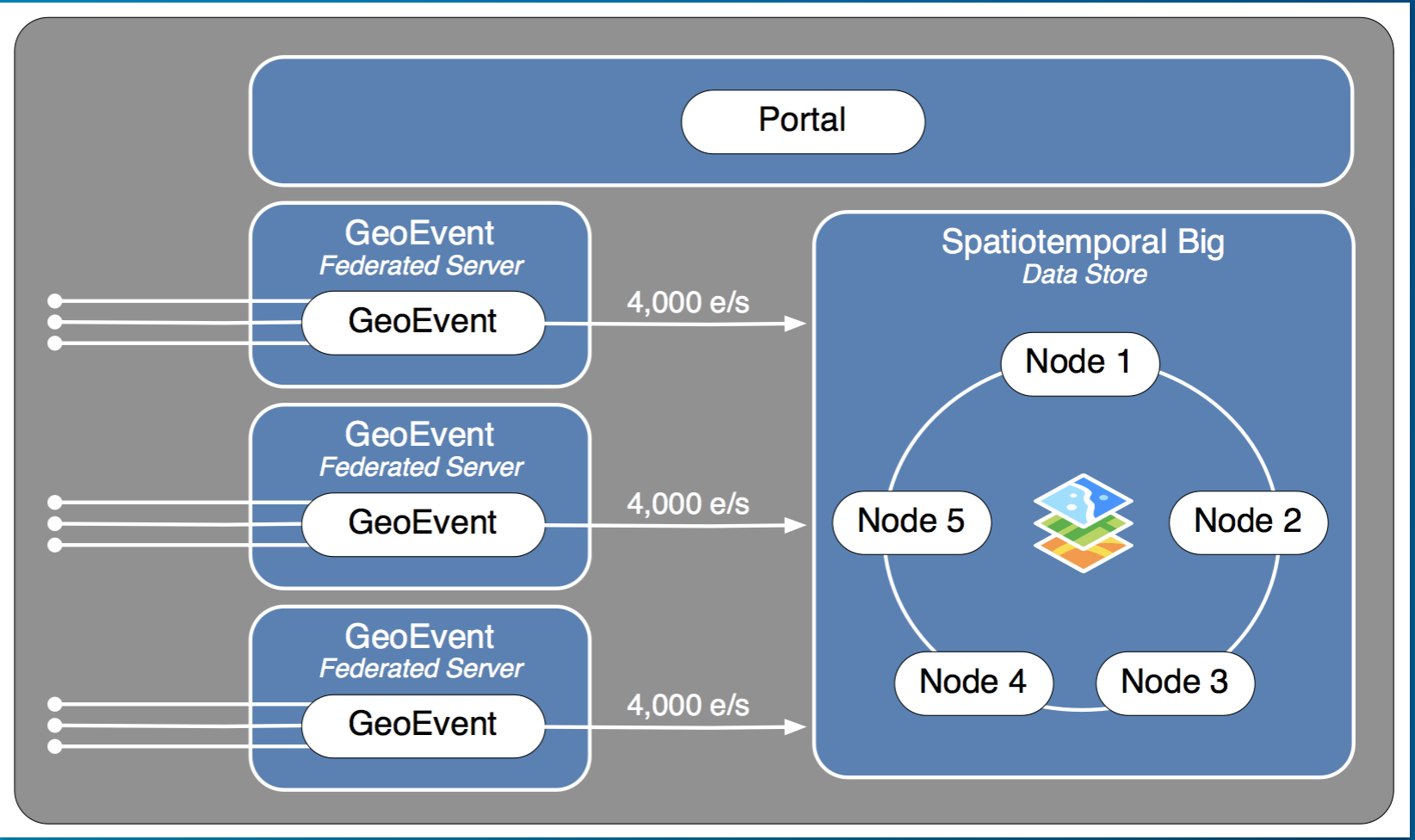
ArcGIS 10.4 Benchmarks *c4.2xlarge (Windows 2012 Server R2): 8 vCPU, 15 GiB, 100GB local SSD, 1,000 Mbps EBS*



Storage Write Throughput	1 node	2 node	3 node	4 node	5 node
ArcGIS Big Data Store	106k	143k	192k	224k	249k
ArcGIS Relational Data Store	0.2k	n/a	n/a	n/a	n/a

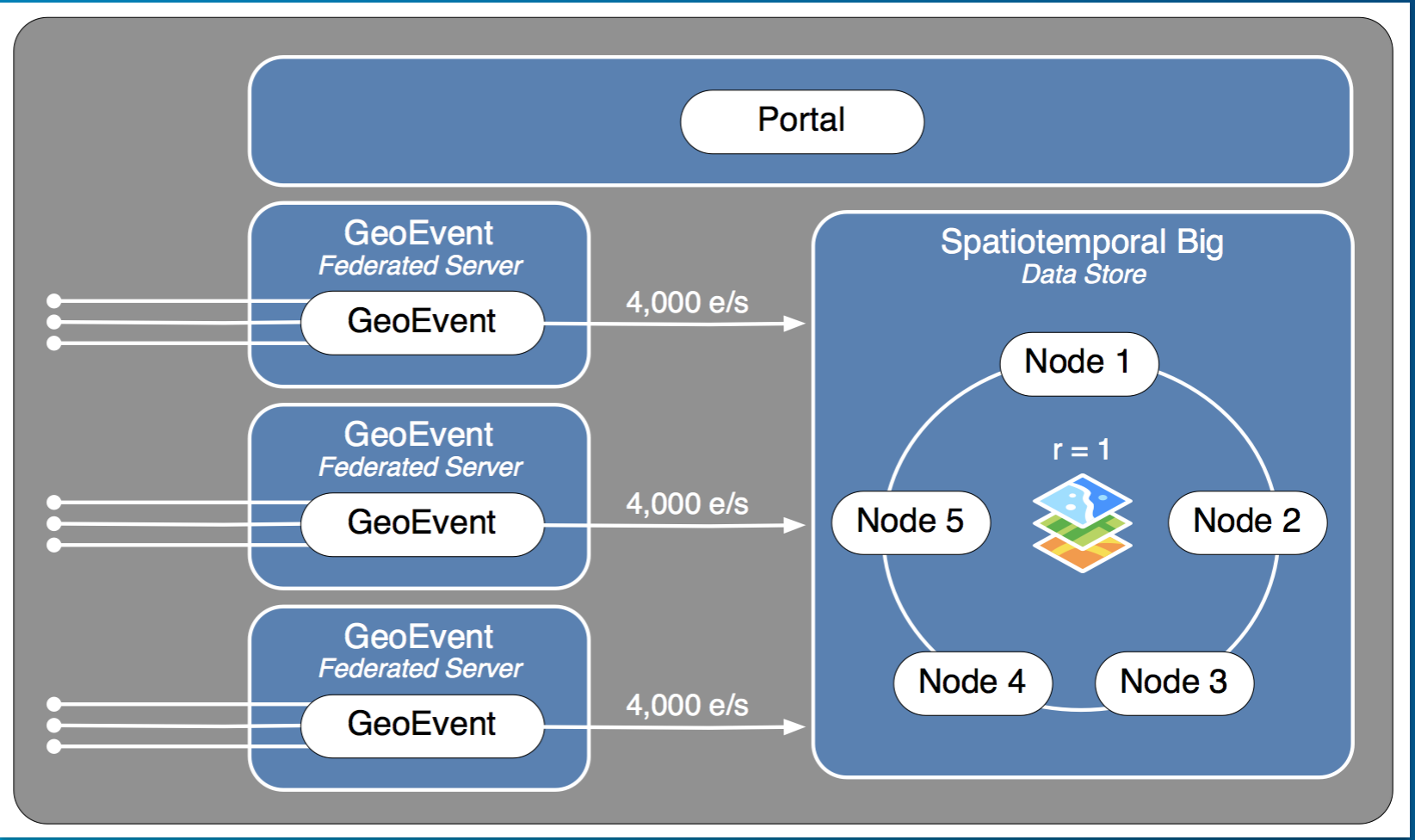
Spatiotemporal Big Data Store Resiliency

Keep data reliable by properly configuring the replication factor of a data source



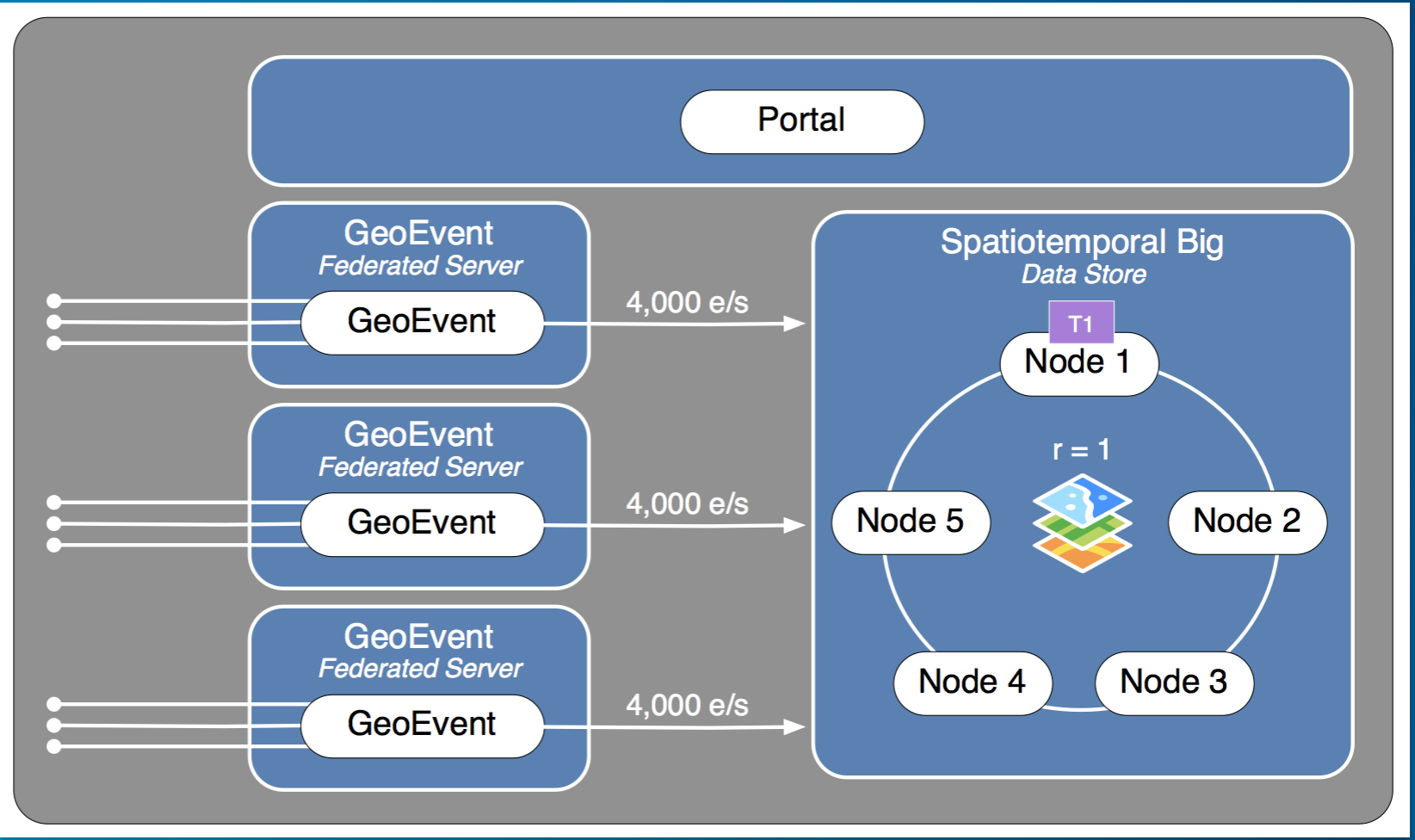
Spatiotemporal Big Data Store Resiliency

Keep data reliable by properly configuring the replication factor of a data source



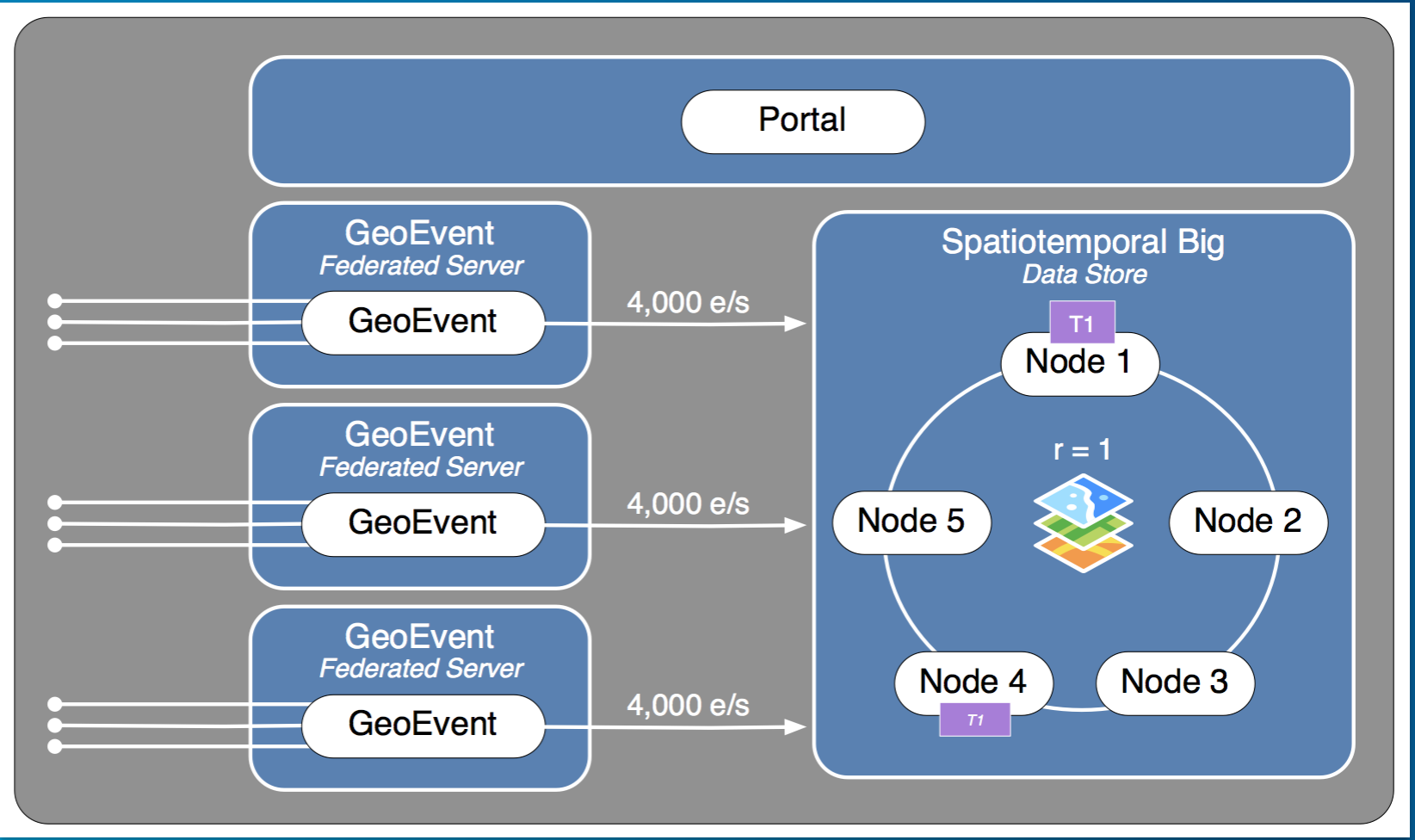
Spatiotemporal Big Data Store Resiliency

Keep data reliable by properly configuring the replication factor of a data source



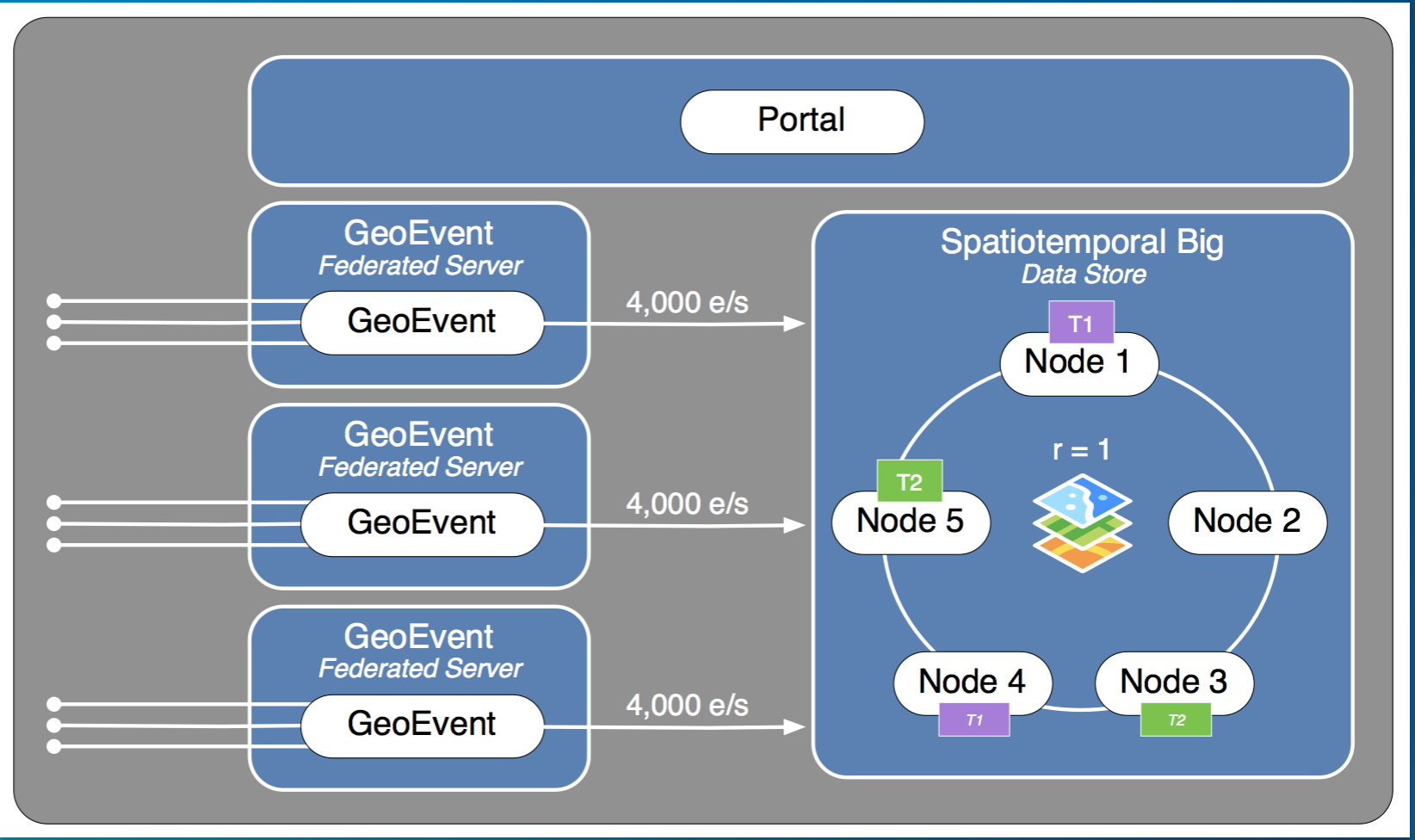
Spatiotemporal Big Data Store Resiliency

Keep data reliable by properly configuring the replication factor of a data source



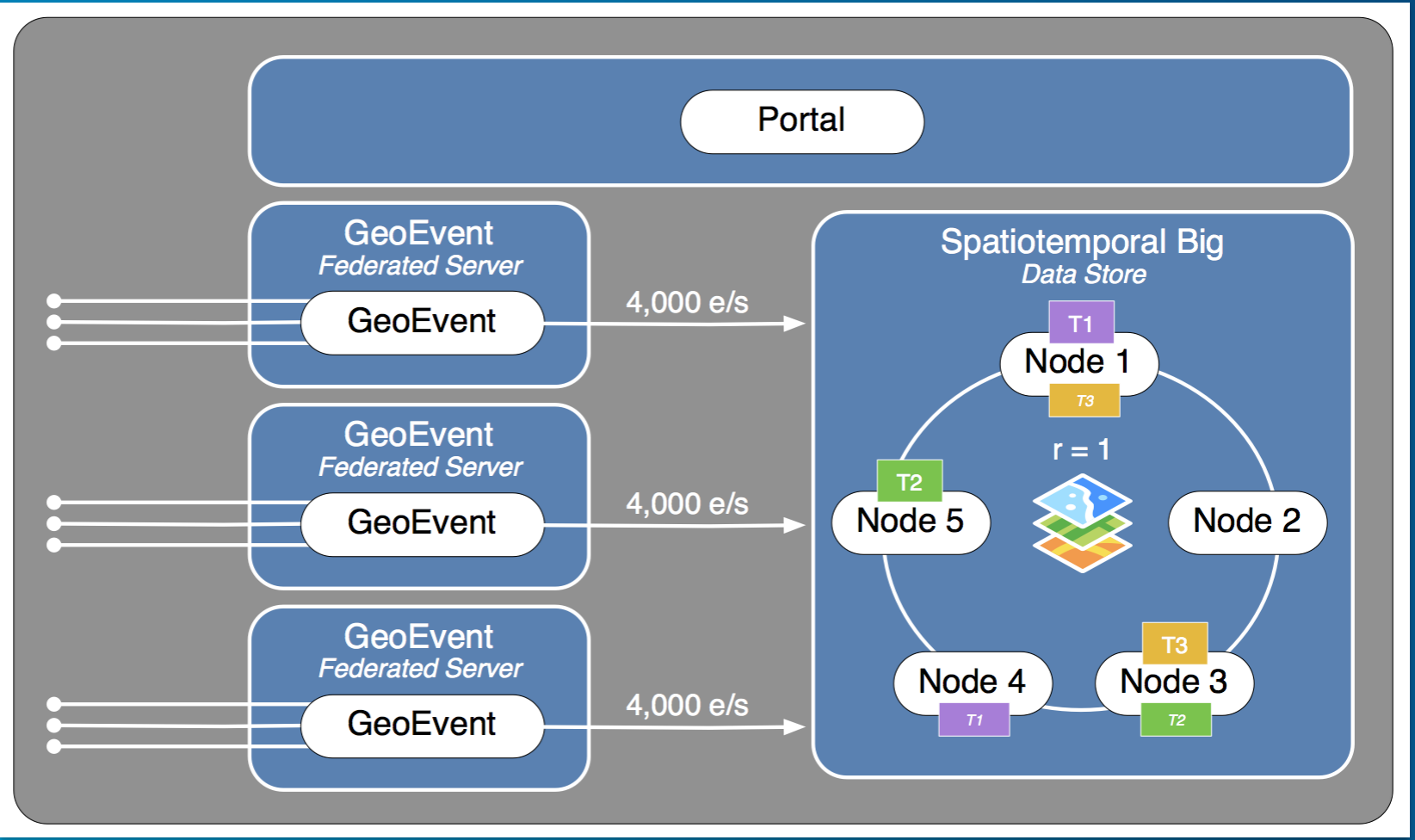
Spatiotemporal Big Data Store Resiliency

Keep data reliable by properly configuring the replication factor of a data source



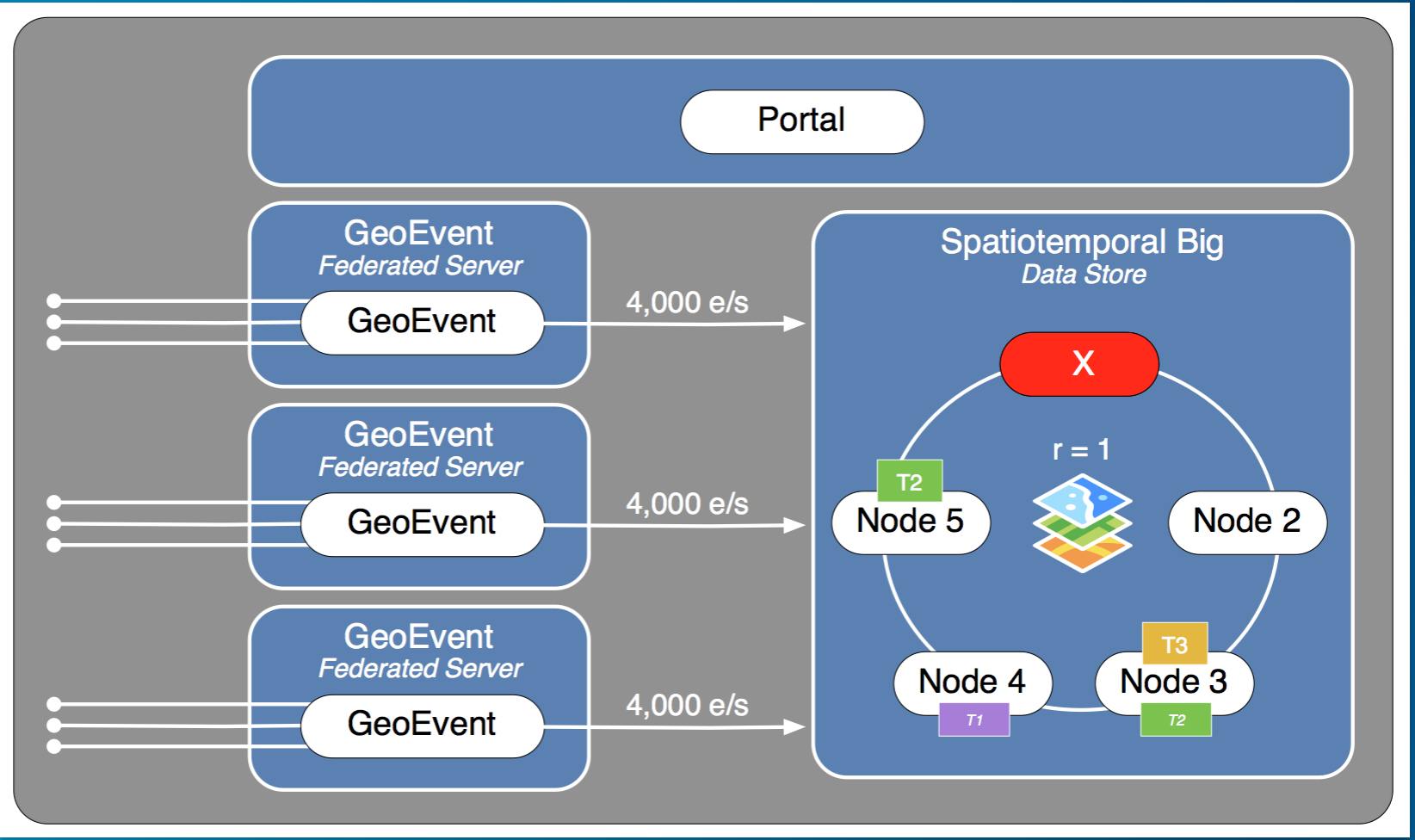
Spatiotemporal Big Data Store Resiliency

Keep data reliable by properly configuring the replication factor of a data source



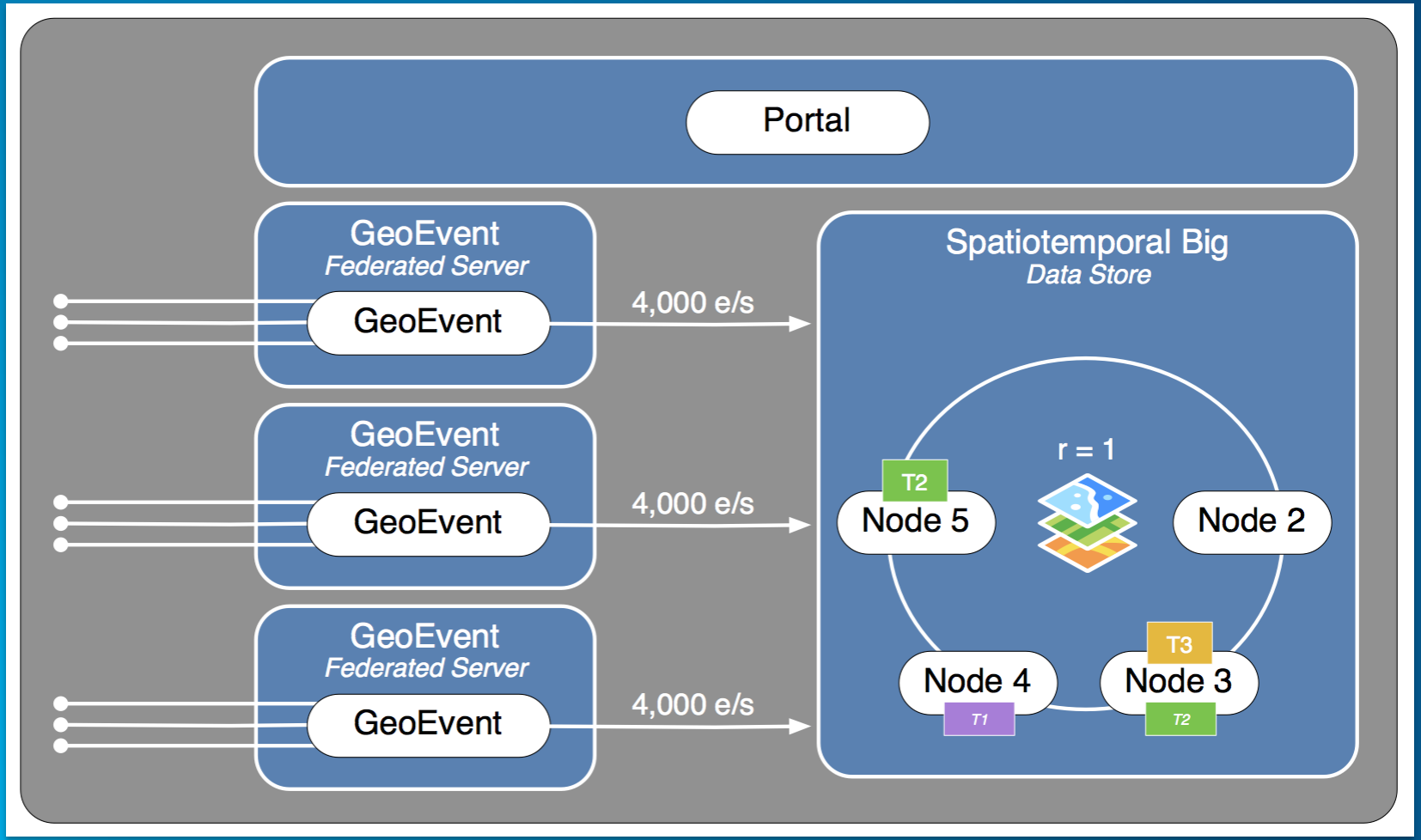
Spatiotemporal Big Data Store Resiliency

Auto-rebalancing of data upon node membership changes, + or -, in the Big Data Store



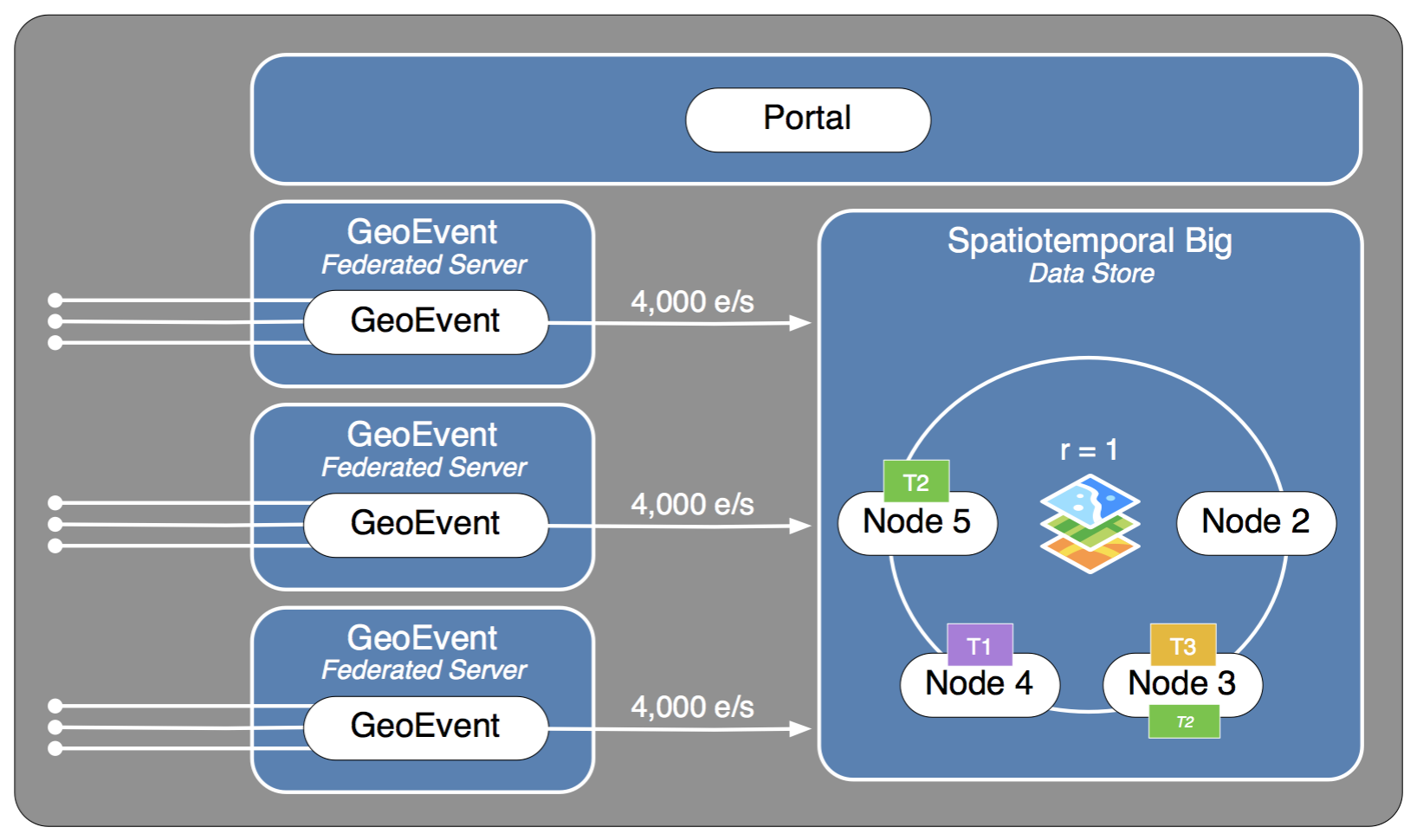
Spatiotemporal Big Data Store Resiliency

Auto-rebalancing of data upon node membership changes, + or -, in the Big Data Store



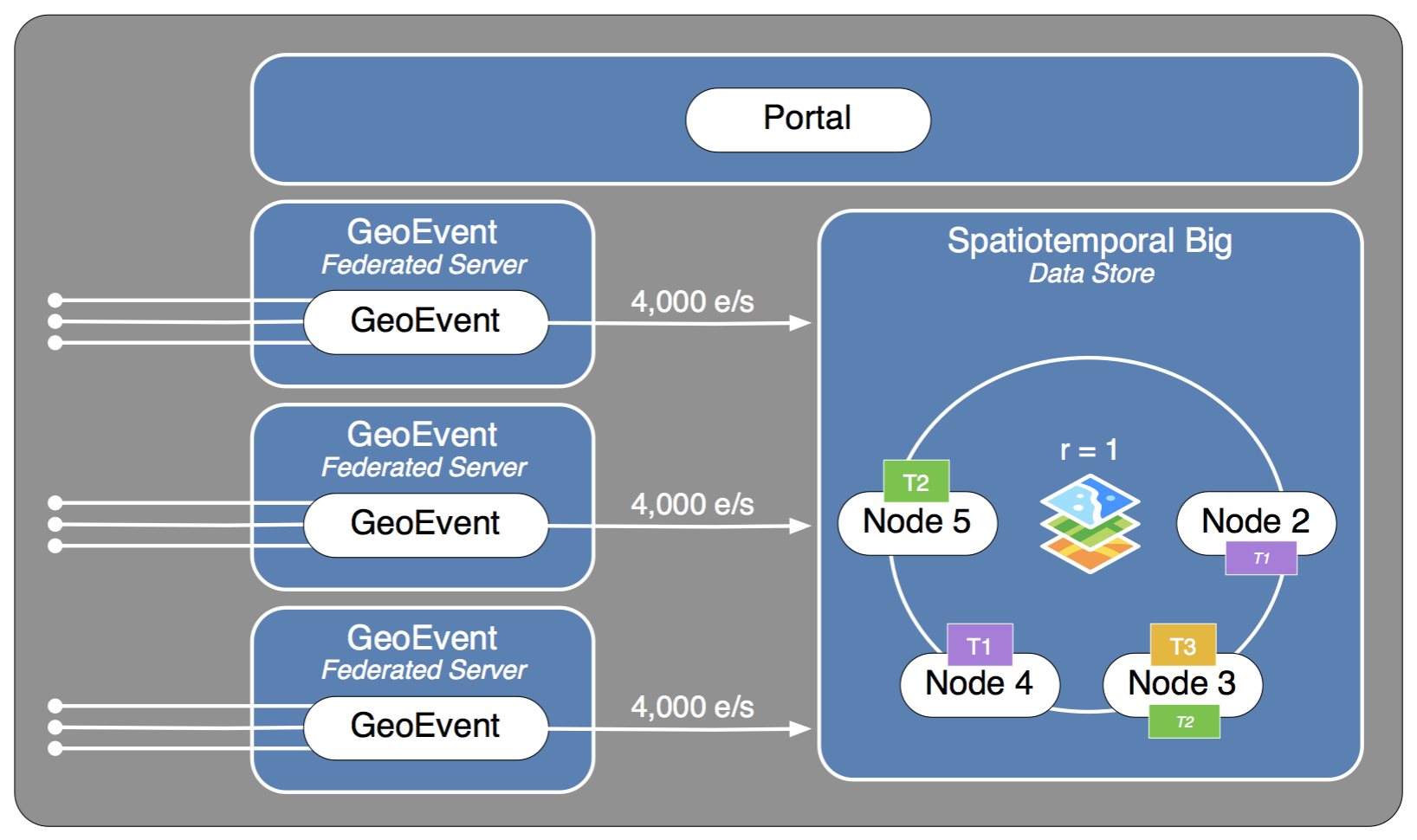
Spatiotemporal Big Data Store Resiliency

Auto-rebalancing of data upon node membership changes, + or -, in the Big Data Store



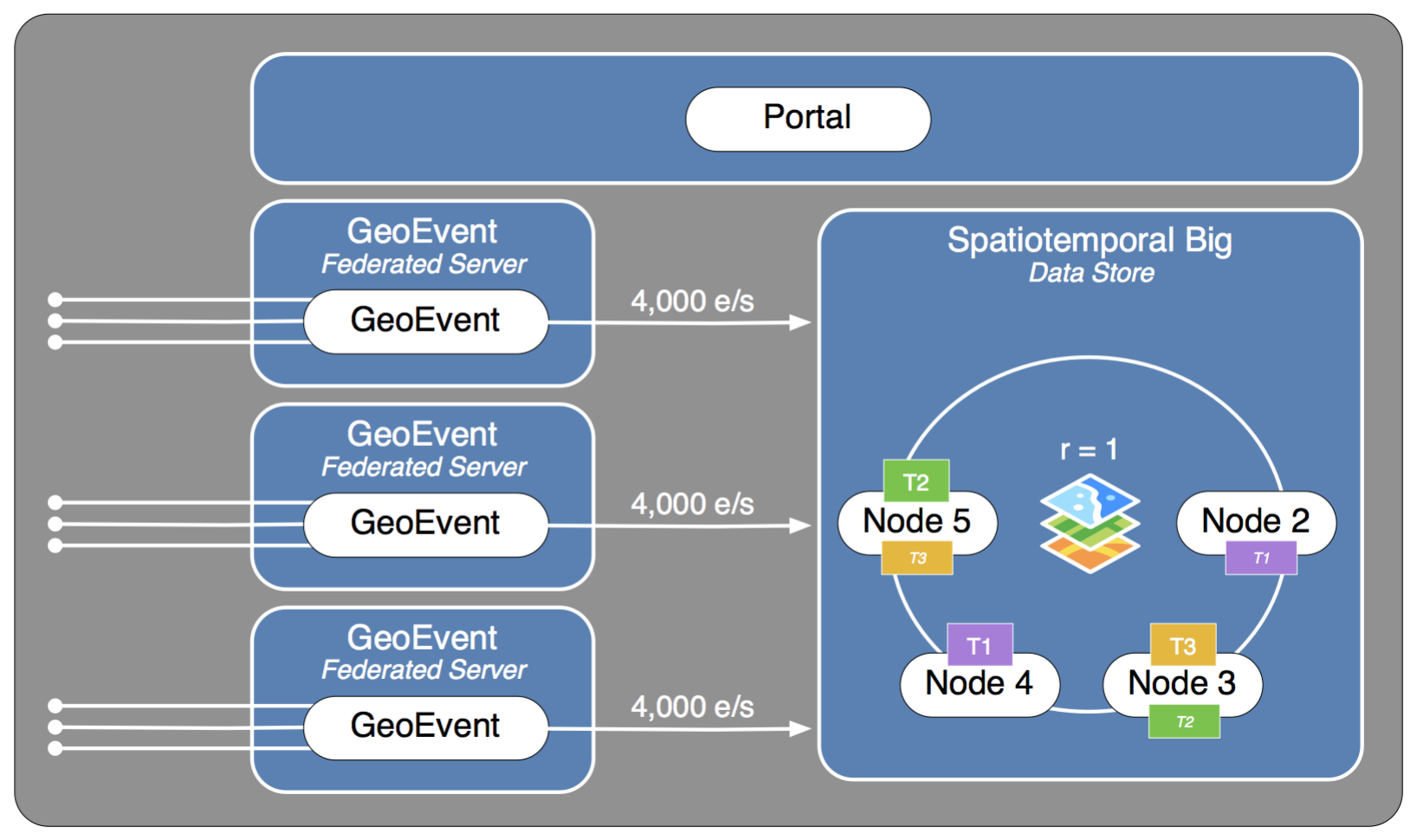
Spatiotemporal Big Data Store Resiliency

Auto-rebalancing of data upon node membership changes, + or -, in the Big Data Store



Spatiotemporal Big Data Store Resiliency

Auto-rebalancing of data upon node membership changes, + or -, in the Big Data Store



Spatiotemporal Big Data Store Resiliency

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 Resiliency

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 Resiliency

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 Resiliency

Resource requirements: processors, memory, storage & instances

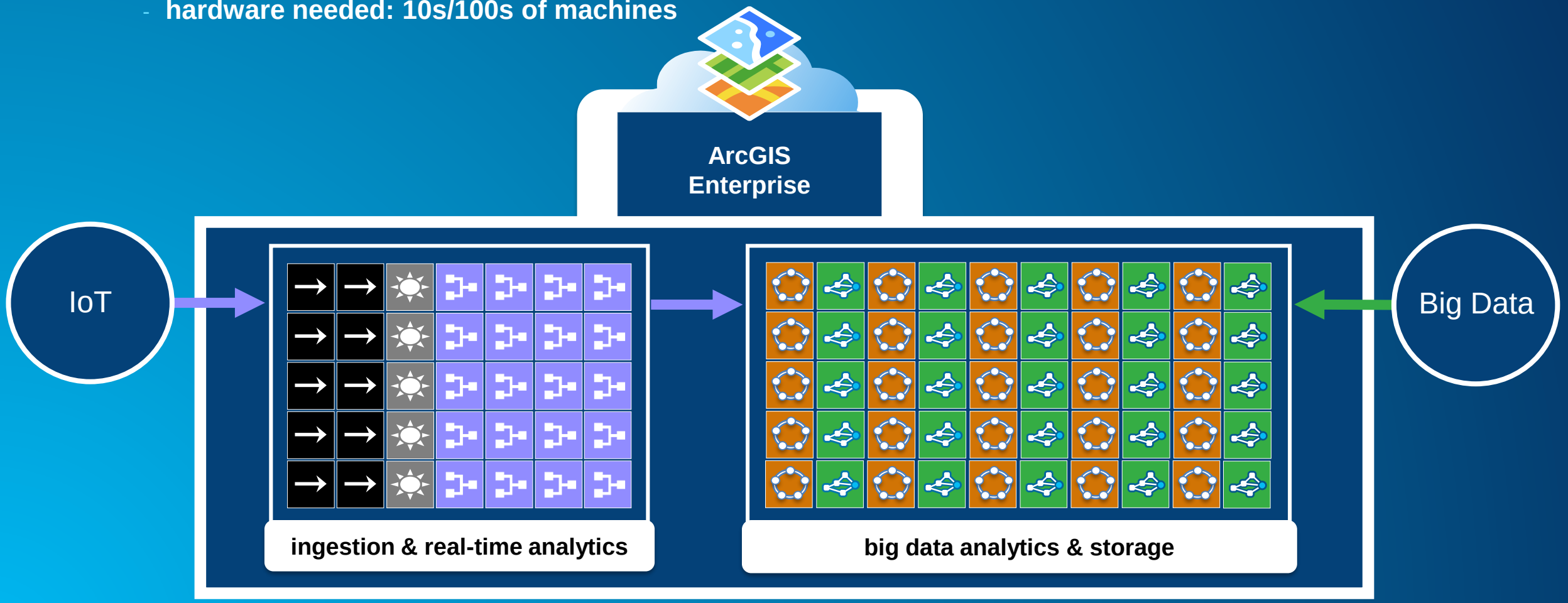
resource	minimum	RECOMMENDED	notes
processors	8 cores	16 cores	More processors results in faster writes (indexing time), higher write throughput & faster query response time.
storage	400 GB	1 TB or more	Storage need is dictated by data velocity & size. SSD results in much faster writes & queries.
memory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

ArcGIS Enterprise

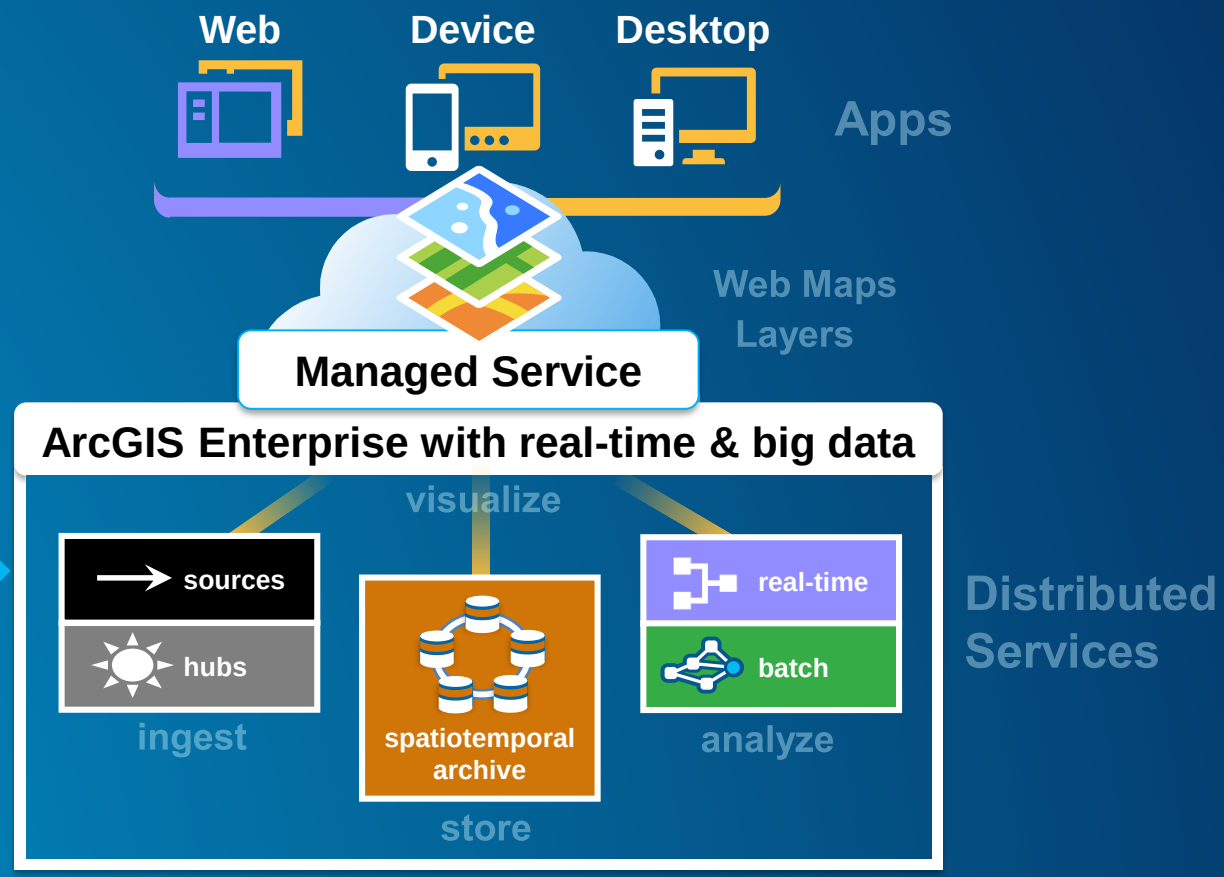
With real-time & big data capabilities

- **A new class of customer is demanding MASSIVE real-time & big data analytic capabilities.**
  - **hardware needed: 10s/100s of machines**



Trinity
as a Managed Service
Bringing geospatial insights to your IoT

- **Trinity as a managed service is a capability that enables existing customers and prospects with the ability to:**
  - **Reliably ingest, analyze & store up to millions of sensor events per second.**
  - **Quickly visualize, replay & explore observations that are billions of features.**
  - **Perform fast analytics on billions of observations.**



Trinity
as a Managed Service
bringing geospatial insights to your IoT

Web Device Desktop Apps

Task Type:	**ArcGIS Enterprise**
Task Name:	**ae1**
Instances:	**1**
Cores:	4
Memory:	8GB
Disk Space:	200GB

on Amazon C2S

on Amazon

on Azure

Managed Service

ArcGIS Enterprise with real-time & big data
880 cores, 3TB ram, 50TB ssd, 80 nodes

DC/OS

ArcGIS Enterprise	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd
8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd
8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd
8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd
8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd
16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd
16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd
16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd	16 cores 56 GB ram 1 TB ssd

sources Scala | hubs kafka | real-time analytics Spark | spatiotemporal archive elasticsearch | batch analytics Spark

Trinity
as a Managed Service
bringing geospatial insights to your IoT

Web Device Desktop Apps

on Amazon C2S

on Amazon

on Azure

Managed Service

ArcGIS Enterprise with real-time & big data
880 cores, 3TB ram, 50TB ssd, 80 nodes

DC/OS

ArcGIS Enterprise							8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd	8 cores 28 GB ram 400 GB ssd
3	3	3	4	4	5	6 7	6 7	6 7	6 7
2	2	2	2	2	2	2	2	2	2
1 1	1 1	1 1	1 1	1 1	2	2	2	2	2
1	1	1	1	1	1	1	1	1	1

sources Scala | hubs kafka | real-time analytics Spark | spatiotemporal archive elasticsearch | batch analytics Spark

Summary

- Performance is dependent on many variables.
- Scaling can be complex.
- Resiliency can be complex.
- Message broker technology is key to GeoEvent scaling.
  - Also for stream service resiliency
- Spatiotemporal Big Data Store is key to persistent data scaling and resiliency.
- Real-time and big data as a service is next generation of scaling and resiliency.
- You can still scale and implement resiliency on your own with careful, thoughtful design.
  - We are soon releasing a tutorial on this configuration.

Questions?