Considerations for a High Availability Enterprise

Sean McGinnis – Solution Architect

SEE
WHAT
OTHERS
CAN'T

High Availability

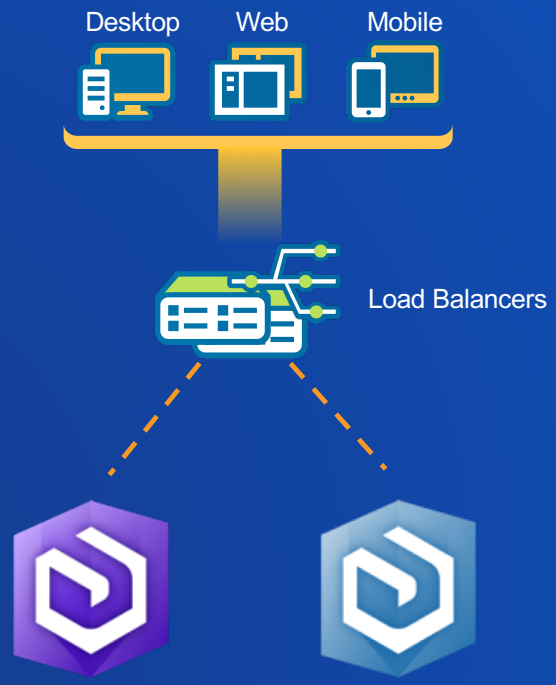
A design approach that targets a prearranged level of operational performance during a period of time

- Improve service delivery
- Reduce risks

Availability (%)	Downtime per year	Downtime per week
95.0	18.25 days	8.4 hours
99.0	3.65 days	1.68 hours
99.9	8.76 hours	10.1 minutes
99.99	52.56 minutes	1.01 minutes
99.999	5.26 minutes	6.05 seconds

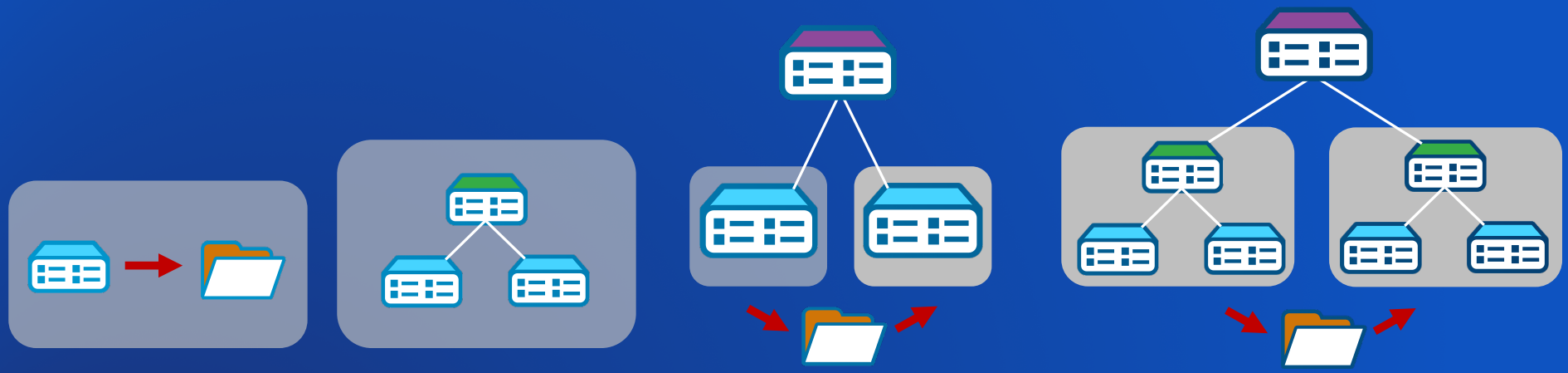
What is the acceptable downtime for your business workflows?

High Availability | Multi-Machine Redundancy



Redundancy and load
Balancing for high availability

High Availability | Multi-Machine Redundancy



Backup and Restore

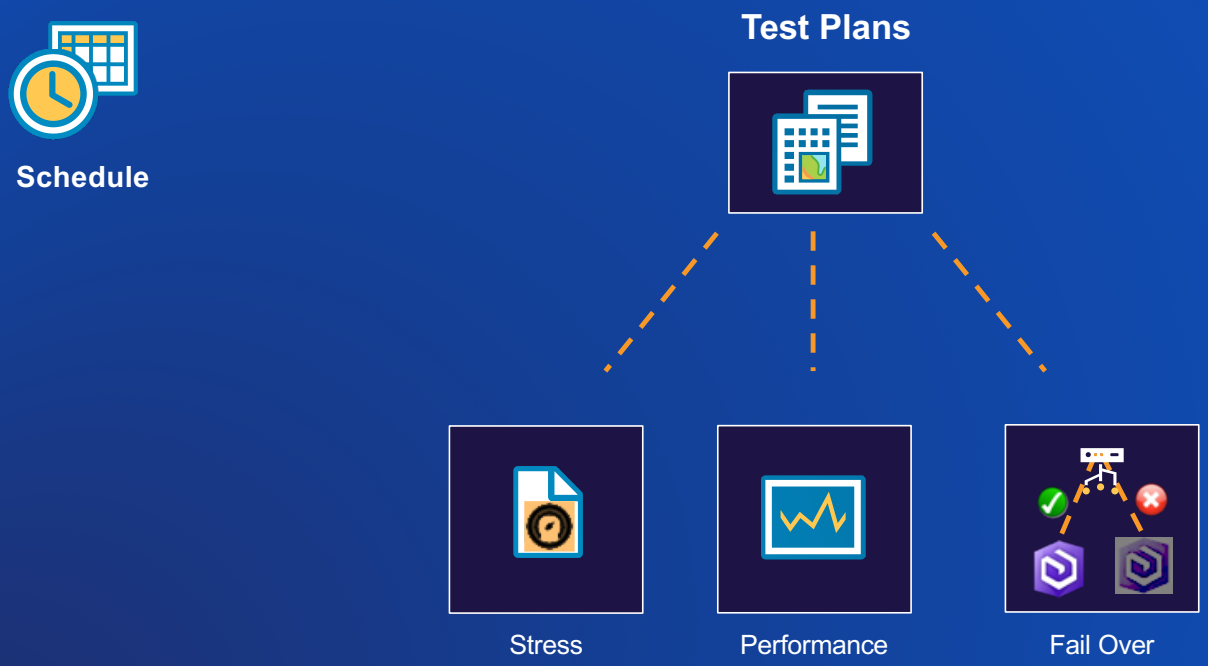
High Availability

Geographic Redundancy

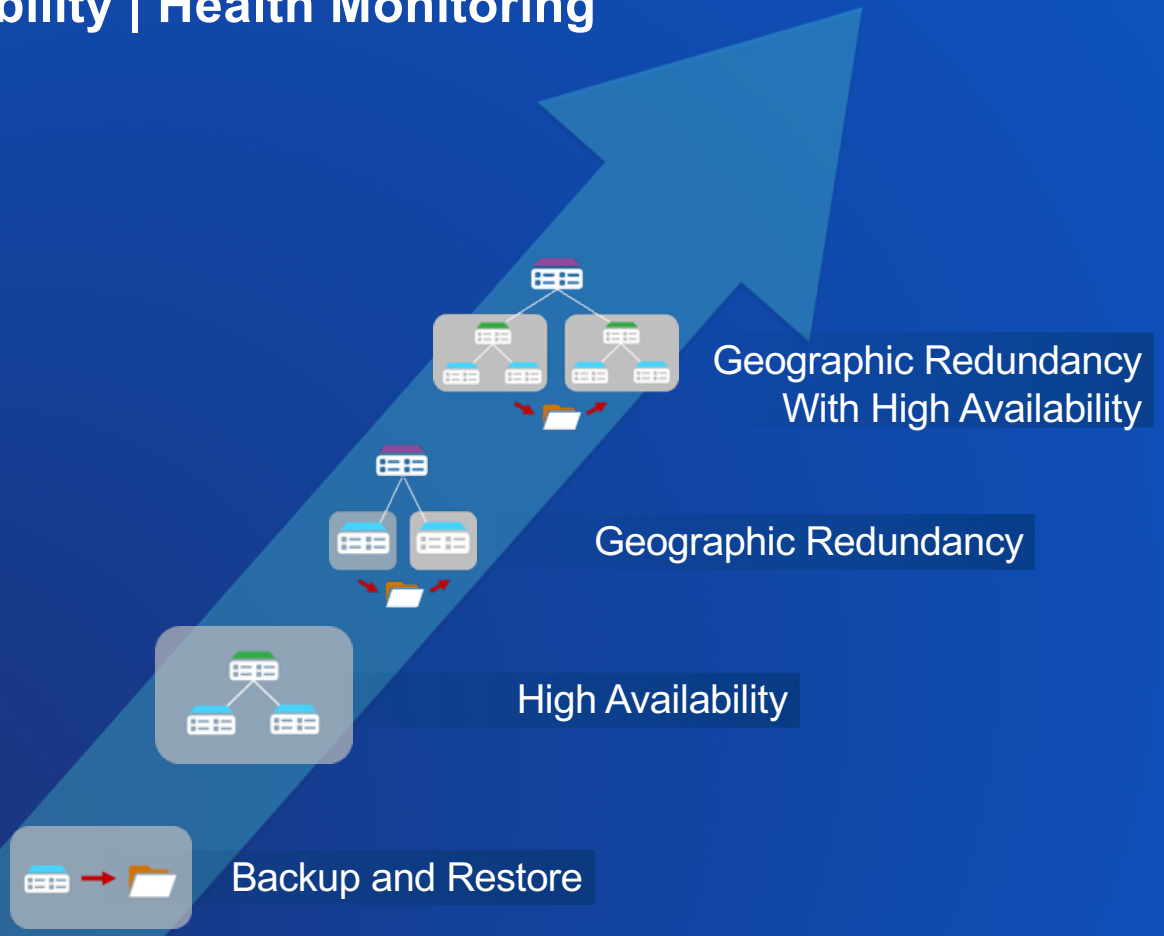
Geographic Redundancy
With High Availability

Increasing complexity and required resources

High Availability | System Operational Plans



High Availability | Health Monitoring



PREVENTION

ArcGIS Monitor

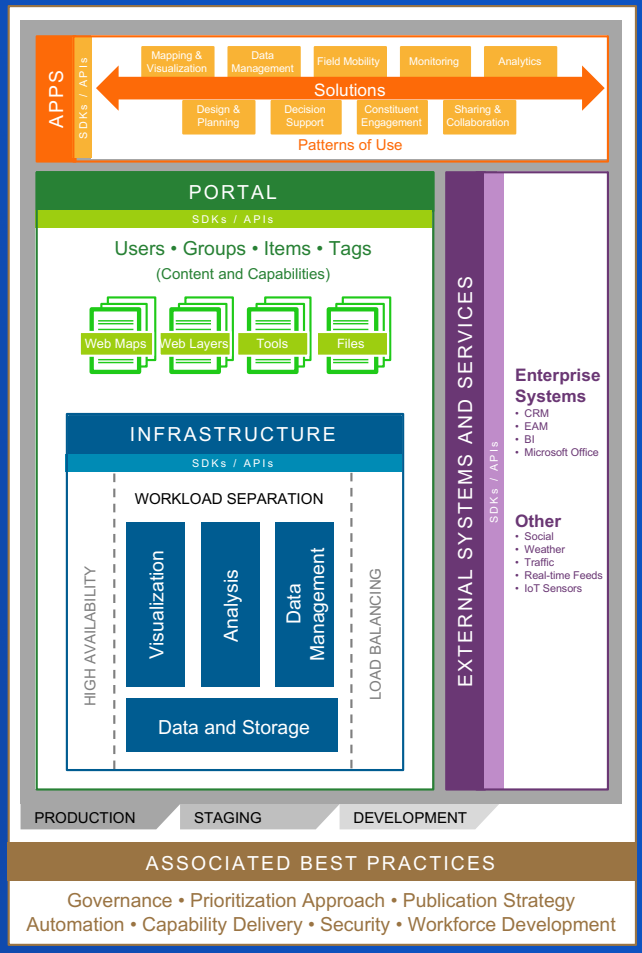
High Availability

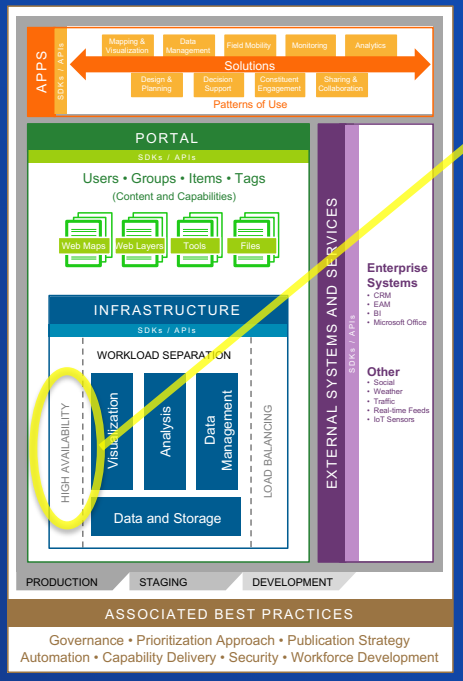
- Reduce single points of failure
- Develop and execute test plans
- Monitor the health of the system

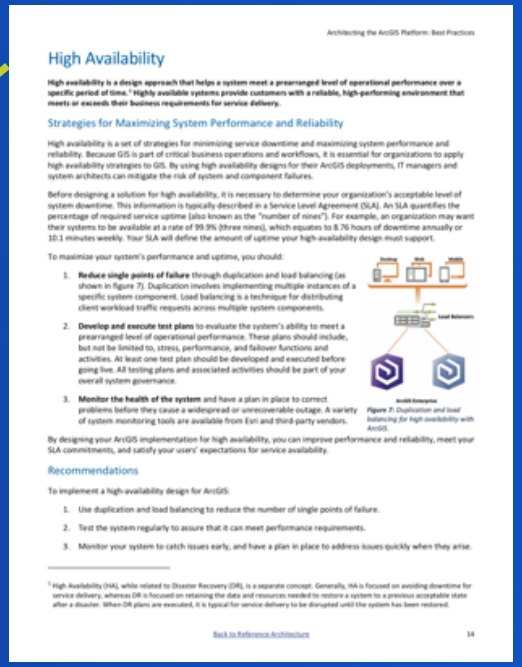
Availability (%)	Downtime per year	Downtime per week
95.0	18.25 days	8.4 hours
99.0	3.65 days	1.68 hours
99.9	8.76 hours	10.1 minutes
99.99	52.56 minutes	1.01 minutes
99.999	5.26 minutes	6.05 seconds

What is the acceptable downtime for your business workflows?

ArcGIS Conceptual Reference Architecture







Architecting the ArcGIS Platform: Best Practices

High Availability

High availability is a design approach that helps a system meet a prearranged level of operational performance over a specific period of time.[1] Highly available systems provide customers with a reliable, high-performing environment that meets or exceeds their business requirements for service delivery.

Strategies for Maximizing System Performance and Reliability

High availability is a set of strategies for minimizing service downtime and maximizing system performance and reliability. Because GIS is part of critical business operations and workflows, it is essential for organizations to apply high availability strategies to GIS. By using high availability designs for their ArcGIS deployments, IT managers and system architects can mitigate the risk of system and component failures.

Before designing a solution for high availability, it is necessary to determine your organization's acceptable level of system downtime. This information is typically described in a Service Level Agreement (SLA). An SLA quantifies the percentage of required service uptime (also known as the "number of nines"). For example, an organization may want their systems to be available at a rate of 99.9% (three nines), which equates to 8.76 hours of downtime annually or 10.1 minutes weekly. Your SLA will define the amount of uptime your high-availability design must support.

To maximize your system's performance and uptime, you should:

1. **Reduce single points of failure** through duplication and load balancing (as shown in figure 7). Duplication involves implementing multiple instances of a specific system component. Load balancing is a technique for distributing client workload traffic requests across multiple system components.
2. **Develop and execute test plans** to evaluate the system's ability to meet a prearranged level of operational performance. These plans should include, but not be limited to, stress, performance, and failover functions and activities. At least one test plan should be developed and executed before going live. All testing plans and associated activities should be part of your overall system governance.
3. **Monitor the health of the system** and have a plan in place to correct problems before they cause a widespread or unrecoverable outage. A variety of system monitoring tools are available from Esri and third-party vendors.

Figure 7: Duplication and load balancing for high availability with ArcGIS.

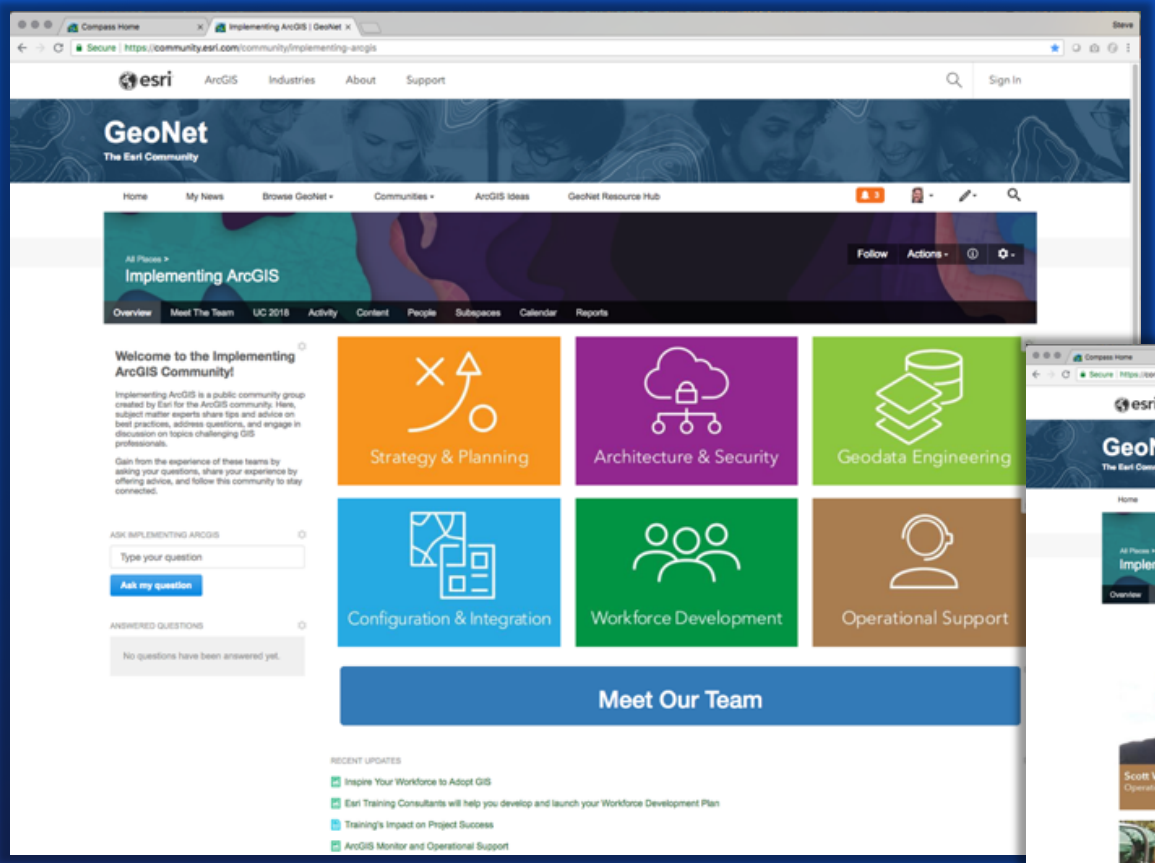
By designing your ArcGIS implementation for high availability, you can improve performance and reliability, meet your SLA commitments, and satisfy your users' expectations for service availability.

Recommendations

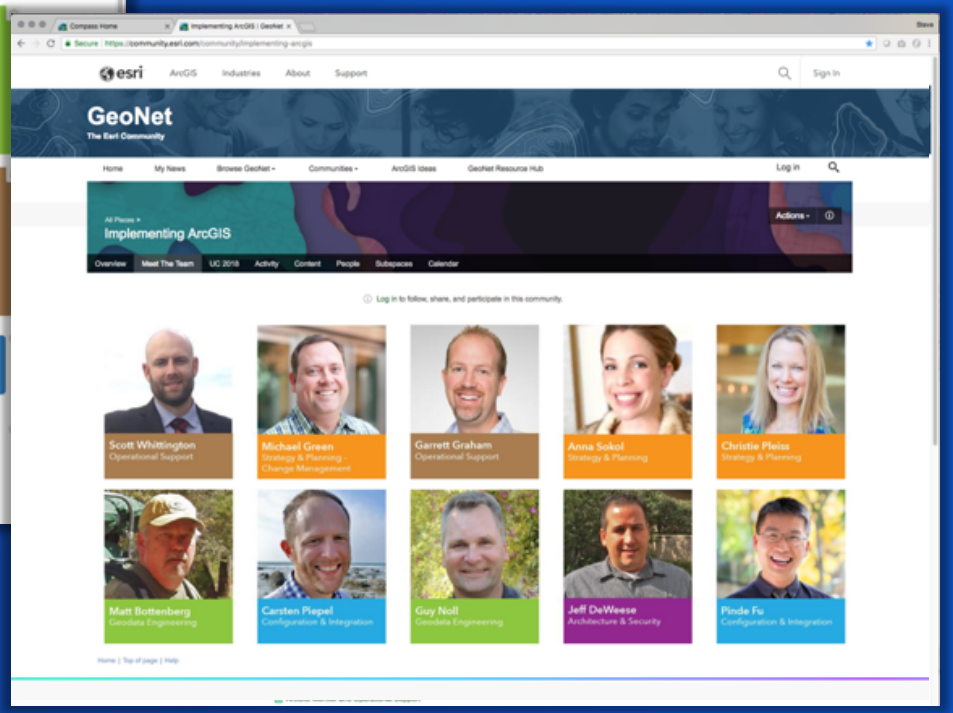
To implement a high-availability design for ArcGIS:

1. Use duplication and load balancing to reduce the number of single points of failure.
2. Test the system regularly to assure that it can meet performance requirements.
3. Monitor your system to catch issues early, and have a plan in place to address issues quickly when they arise.

[1] High Availability (HA), while related to Disaster Recovery (DR), is a separate concept. Generally, HA is focused on avoiding downtime for service delivery, whereas DR is focused on retaining the data and resources needed to restore a system to a previous acceptable state after a disaster. When DR plans are executed, it is typical for service delivery to be disrupted until the system has been restored.



Thought leaders sharing best practices on Implementing ArcGIS throughout the year



go.esri.com/ImplementingArcGIS