High Availability and Disaster Recovery

Jonathan Quinn
Cherry Lin
Managing the Twin Risks to your Operations

- Data Loss
- Down Time
The Three Approaches

Backups
- Ability to go back in time

High Availability
- No single point of failure
- Machine redundancy

Disaster Recovery
- No single point of failure
- Environment redundancy
- Geographic Redundancy
Choosing Between Them

- Complementary
- Build On Each Other
- Cost and Capability
Backup and Restore
Backups are....

Simple
Highly Effective
Not Disruptive
Under appreciated
ArcGIS Enterprise Backups – WebGIS DR Tool

What the tool backs up

- **Settings**
  (Portal, Server, Data Store)

- **Portal Content**

- **Services**

- **ArcGIS Data Store Data**
  (relational, scene tiles)
ArcGIS Enterprise Backups – WebGIS DR Tool

What the tool doesn’t backup

EGDB or file based data

Traditional cache tiles
How to Backup an ArcGIS Enterprise deployment

Web GIS DR Tool

Property File
- Location
- Portal URL
- Credentials
- Scene Cache?

Automate
WebGIS DR Properties

• Questions:
  • Where is your content, (file system, bucket in S3, container in Azure)
  • Where do you want your backup stored, (file system, bucket in s3, container in Azure)

• Minimum parameters:
  • SHARED_LOCATION = where each backup will be staged
  • BACKUP_STORE_PROVIDER = where to store the backup (file system, or cloud)
  • PORTAL_ADMIN_URL = URL to connect to the portal
  • PORTAL_ADMIN_USERNAME = administrator’s username
  • PORTAL_ADMIN_PASSWORD = administrator’s password
  • BACKUP_RESTORE_MODE = defines if a full or incremental backup will be run
WebGIS DR Tool – Usage

• Backup
  - Component backups run concurrently
  - No downtime while exporting
  - Sample syntax

        C:\Program Files\ArcGIS\Portal\tools\webgisdr>webgisdr.bat -e -f webgisdr.properties

• Restore
  - Runs sequentially
    - Data Store → Server → Portal
  - Downtime while restoring
  - Sample syntax

        C:\Program Files\ArcGIS\Portal\tools\webgisdr>webgisdr.bat -i -f webgisdr.properties
Scheduling ArcGIS Enterprise backups - Windows
Scheduling ArcGIS Enterprise backups - Windows
Scheduling ArcGIS Enterprise backups - Windows

-e -f "C:\Program Files\ArcGIS\Portal\tools\webgisdr\webgisdr.properties"
Scheduling ArcGIS Enterprise backups - Windows

Create Basic Task Wizard

Summary

Create a Basic Task

Name: Backups

Trigger: Daily; At 3:43 PM every day
Action: Start a program: "C:\Program Files\ArcGIS\Portal\tools\webgisdr\webgisdr.bat"

When you click Finish, the new task will be created and added to your Windows schedule.
Scheduling ArcGIS Enterprise backups - Linux

- Creating a cronjob:
  ```bash
  [ags@wilson ~]$ crontab -e
  ```

- Cronjob syntax:
  ```plaintext
  * * * * * < command >
  ```

Examples:
- Run the WebGIS DR Tool at 12:00:00 AM every day:
  ```bash
  0 0 * * * /data/arcgis/portal/tools/webgisdr/webgisdr.sh -e -f /data/arcgis/portal/tools/webgisdr/webgisdr.properties
  ```

- Run the tool every 12 hours every day starting at 12:00:00 AM:
  ```bash
  0 */12 * * * /data/arcgis/portal/tools/webgisdr/webgisdr.sh -e -f /data/arcgis/portal/tools/webgisdr/webgisdr.properties
  ```
High Availability
Overview

- What is High Availability
- ArcGIS Enterprise High Availability
  - Components
  - Upgrade
- Other factors for High Availability
High Availability (HA)

• Definition:
  - A system or component that is continuously operational for a desirably long length of time. Availability can be measured relative to "100% operational" or "never failing." (SLAs)

• Shorter down time costs more

• Elimination of single points of failure.

• Availability of a system depends on the availability of all components
ArcGIS Enterprise

Portal
GIS Services
Hosted Feature and Tile Data

Portal for ArcGIS
ArcGIS Server
ArcGIS Data Store
Portal for ArcGIS: High Available Deployment

Load Balancer

“Highly Available Portal”

Portal Machines

Portal Content (shared)
Highly Available Portal

- Two Portal machines
- Both Portal machines take requests
- Internally, there is a difference between the two machines’ role:
  - Primary
  - Standby
- Behavior after machine shuts down depends on role:
  - No interruption if standby machine becomes unavailable
  - Typically 30 seconds of unavailability at 10.6.1 and up
    - Improved from a few minutes
Portal for ArcGIS: Load Balancing Options

- Provided by Esri
  - Web-Tier Authentication
  - Availability dependent on web servers
- Not provided by Esri
  - Typically already fault tolerant
Portal for ArcGIS: High Availability Deployment Patterns

HA Portal with Load Balancer

- Load Balancer
- Portal Machines
- Portal Content *(shared)*

- Simpler
- Need certain settings on LB
- Doesn’t support Web Tier Authentication

HA Portal with Load Balancer & Web Adaptors

- Load Balancer
- Web Adaptors
- Portal Machines
- Portal Content *(shared)*

- More complex
- Web Tier Authentication
Portal for ArcGIS: Health Check

• Provided by Portal for ArcGIS
  - https://<webadaptor machine>.domain.com/<context>/portaladmin/healthCheck
  - https://<machine>.domain.com:7443/arcgis/portaladmin/healthCheck

• Check if Portal is ready to take request. Not individual component, e.g. service, item, etc.

• Or your own customized health check
Upgrade High Availability Portal for ArcGIS

- There is downtime
  - Plan
  - Practice
- Make a backup
- No need to take note of roles (new at 10.7)
- Similar other steps as standalone Portal
Portal for ArcGIS: Key Considerations for HA

• Two Portal machines
  - Primary
  - Standby

• Highly Available Load Balancer
  - Web Tier Authentication
  - No single Web Adaptor

• Health Check provided for Portal for ArcGIS

• Highly Available shared content store

• Upgrade: Downtime & steps in order
ArcGIS Server: Multiple-Machine Architecture

- Multiple machines
- Identical Roles
- No interruption when any machine is down
- The config-store and server directories need to be accessible to all machines.
ArcGIS Server: High Availability Deployment Patterns

Server Site with Load Balancer

- Load Balancer
- Server Machines
- Config-store Server Directories *(shared)*

Server Site with Load Balancer & Web Adaptors

- Load Balancer
- Web Adaptors
- Server Machines
- Config-store Server Directories *(shared)*
ArcGIS Server: Health Check

- Provided by ArcGIS Server
  - https://<.....domain.com>/<context>/rest/info/healthcheck
  - https://<machine>.domain.com:6443/arcgis/rest/info/healthcheck

- Server level health check. Not checking service.

- Or your own customized health check
Upgrade multi-machine ArcGIS Server

- Install and Upgrade
- Same on all machines
- Downtime for upgrade one machine
Portal for ArcGIS and ArcGIS Server: Federation

- Portal URL: 443
- Private Portal URL: 7443
- Administrative URL: 6443
- Services URL: 443
- Communication
- Administrative Communication
Portal for ArcGIS and ArcGIS Server: Federation

Portalurl: 443

Services URL: 443

privatePortalurl: 7443

Administrative URL: 6443
ArcGIS Server: Key Considerations for HA

- Highly Available shared config-store and server directories
- Health Check provided for ArcGIS Server
- Highly Available URLs when communicating with Portal
  - Portal URL
  - Private Portal URL
  - Services URL
  - Server Administrative URL
- Install and Upgrade on all machines
Spatiotemporal Big Data Store

Title: Data Store Management Best Practices

Date: 03/06/2019

Time: 10:30am – 11:30am

Location: Catalina/Madera

The session has already passed – watch the video later and/or please bring questions to the Expo.
ArcGIS Data Store: High Availability Architecture

Server Site: ArcGIS Data Store's Load Balancer

“Highly Available ArcGIS Data Store”

Primary

Standby

Backups (shared)
Primary ArcGIS Data Store stops working: Define Failure
- Computer crashes
- Gets unplugged
- Lose network connectivity
- etc

Not “gracefully” shutdown
- Data Store service stops

Upgrade High Availability ArcGIS DataStore

- Run setups on both machines
- On PRIMARY, run configure to upgrade
  - Relational
  - TileCache
ArcGIS Enterprise High Availability Deployment

Load Balancer

“Highly Available Portal”

Portal Content
(shared)

Site

Configuration Store
Server Directories
(shared)

“Highly Available ArcGIS Data Store”

Primary

Backups
(shared)

Standby
Upgrade ArcGIS Enterprise High Availability Deployment

- Upgrade Order
Cloud Storage Support

- S3
  - Portal content store
  - Cloud Storage in Server Manager
    - Caching
    - GeoAnalytics Data Input
    - Raster Analytics
  - ArcGIS Spatial Temporal DataStore backups
  - Webgisdr backups
- DynamoDB & S3
  - ArcGIS Server config-store
- User-defined compatible storage
  - Caching
ArcGIS Enterprise HA: Part of Your HA Architecture

• Your Data
  - Enterprise GeoDatabase
  - File based Data

• Software
  - Web Server
  - Software Load Balancer

• Hardware
  - File Server
  - Network

• People
  - HA?
  - IT strong?
ArcGIS Enterprise HA: IT Governance

• Ensure the effective and efficient use of IT

• Policies and procedures highly disciplined
  - Planned and updated in a timely manner
  - Documented clearly
  - Tested properly
  - Exercised with staff
ArcGIS Enterprise HA: Spectrum, Not a Switch

Downtime (decreasing)

Cost

Business Interruption

Workday Interruption

Momentary Interruption

Days

Hours

Minutes

Seconds
Disaster Recovery
Geographic Redundancy
Agenda

- What is geographic redundancy
- Using the Web GIS DR tool
- Roadmap to being geographically redundant
Overview

- Geographically separate data centers
- Duplicated configurations and data between the two data centers
- Components within data centers are typically highly available
- WebGIS DR Tool is used to move snapshots of data from primary to standby
- Complex disaster recovery option
Geographic Redundancy

Traffic Manager

Public Portal URL - https://mysite.esri.com/portal
Services URL – https://mysite.esri.com/server

East coast data center (primary)

West coast data center (standby)

Public portal URL and services URL need to be the same
Geographic Redundancy

Traffic Manager

East coast data center (primary)

West coast data center (standby)
Geographic Redundancy

Traffic Manager

East coast data center (primary)

West coast data center (standby)
Geographic Redundancy

East coast data center (primary)

Traffic Manager

West coast data center (standby)
Geographic Redundancy – Cloud deployments

Traffic Manager

Central Region*

East Coast Region

West Coast Region
Roadmap for geographic redundancy

1. Duplicate the deployment between primary and standby data centers

2. Create snapshots of the primary data center

3. Apply snapshots to the standby data center

4. Monitor your standby data center
Step 1 - Duplication

- Number of machines should be the same
- Identical URLs between data centers
  - Public Portal URL
  - Services URL
- Identical paths to file based data
  - Enterprise data stores can be different, relying on the data store name for the mapping
Duplication – What needs to be the same?

10.4-10.4.1
- Public facing and internal URLs
- By reference data stores
- Server site directory paths
- Machine names
- Security information

10.5-10.6
- Public facing URLs
- By reference data stores
- Server site directory paths
- Security information

10.6.1-10.7
- Public facing URLs
- Registered data stores
Step 2 - Creating snapshots

• Full snapshot
  - Create an initial snapshot of all of the data within the ArcGIS Enterprise
  - Internally defines a base time that will be used for an incremental snapshot

• Incremental snapshot
  - Creates a snapshot of all of the data that has been created or modified since the last full backup
  - Decreases the time it takes to synchronize content, services, and data between primary and standby
Creating incremental snapshots

- Creates a snapshot of all data added or modified since the last full snapshot
Step 3 - Applying snapshots

- Two approaches
  - As snapshots are ready
  - At the time of a failure
- Define an automated schedule that works for your organization
Step 4 - Monitoring

- QC process on standby ArcGIS Enterprise
  - Checking the index within Portal
  - Validating federated Servers
  - Validating data stores through the Server Admin
  - Checking important services or applications

- Detecting when components fail within a data center
  - Monitoring the healthCheck URLs of Portal and Server

- Failing over data centers can be expensive
Takeaway points

- Important to understand the requirements of geographic redundancy as a disaster recovery option

- Take advantage of the Web GIS DR tool to move snapshots of the deployment from primary to standby

- Geographic redundancy is a complex disaster recovery option
Success Stories with HA or DR

- Let us know if you have a success story to share
Please Take Our Survey on the App

Download the Esri Events app and find your event

Select the session you attended

Scroll down to find the feedback section

Complete answers and select “Submit”