SITECORE

BROCHURE

Sitecore Managed Cloud disaster recovery

JANUARY 2021

SITECORE

Table of contents

Introduction

This document describes the two approaches for disaster recovery provided by Sitecore Managed Cloud. The objective is to provide you with a general overview of each approach, including the workflow that occurs during a recovery event. For more details on Sitecore Managed Cloud's disaster recovery processes, please review the KB article here.

Definitions

Below are some key terms used throughout the document, along with their definitions.

- **Sitecore Experience Manager™ (XM)** – Sitecore's web content management product.
- **Sitecore Experience Database™ (xDB)** – Sitecore Experience Database (xDB) is a big data marketing repository that collects and connects customer interaction data to create a comprehensive, unified view of each customer to drive personalization.
- **Sitecore Experience Platform™ (XP)** – Builds on Sitecore XM with additional components, including xDB, Email Experience Manager, and Sitecore xConnect™, designed to deliver complete digital experiences.
- **Sitecore on Azure PaaS** – A term used to refer to Sitecore running on Azure PaaS services, e.g. App Service, Azure SQL, Azure Search. Sitecore is deployed to Azure using Azure Marketplace or Sitecore Azure Toolkit using ARM Templates found on GitHub.
- **Managed Cloud** – A managed hosting offering from Sitecore built on top of Microsoft's Azure PaaS technologies.
- **Business continuity planning (BCP)** – The process of creating systems of prevention and recovery in the event of a disaster. In addition to prevention, the goal is to enable ongoing operation, before and during the execution of disaster recovery.
- **Disaster recovery (DR)** – An area of BCP that aims to protect an organization from the effects of significant natural or human-induced disaster. DR allows an organization to maintain or quickly resume mission-critical functions following such an event.

- **High availability (HA)** – A characteristic of a system, which aims to ensure an agreed level of operational performance, usually uptime, for a long period of time.
- **Time to initiate** – The time between an error taking place and an alert being received, noticed, and acted upon.
- **Domain Name System (DNS)** – A hierarchical, decentralized naming system for computers, services, or other resources connected to the Internet or a private network. It associates various information with domain names assigned to each of the participating entities. Most prominently, it translates more readily memorized domain names to the numerical IP addresses needed for locating and identifying computer services and devices with the underlying network protocols. By providing a worldwide, distributed directory service, the Domain Name System has been an essential component of the functionality of the Internet since 1985.
- **Traffic Manager** – An Azure service that routes traffic based on its own configuration rules.
- **Service level agreement (SLA)** – A contract between a service provider (either internal or external) and the end customer that defines the level of service expected from the service provider. SLAs are output-based in that their purpose is to define what a customer receives.
- **Network Operations Center (NOC)** – The place where support is available 24 hours a day, every day to deal with emergency issues.
- **Paired region** – An Azure concept referring to two data centers that have the lowest latency to one another.
- **Primary data center** – The Azure data center in which the production Sitecore environment is initially set up.
- **Secondary data center** – The Azure data center appointed for the restoration of a Sitecore environment in the case of an outage at the primary data center.
- **Non-Sticky / No Affinity** – Refers to load balancers not associating requests with a particular server, but rather load balancing every request to any server.
- **SQL Availability groups & SQL Azure Failover groups** – An always-on approach that provides a public endpoint to connect to SQL; when the SQL server fails, another SQL server takes over the public endpoint, so there is no need to change connection strings when the public endpoint changed.
- **Recovery Time Objective (RTO)** – The maximum tolerable length of time that a computer, system, network, or application can be down after a failure or disaster occurs.
- **Recovery Point Objective (RPO)** – The amount of data at risk. It's determined by the amount of time between data protection events and reflects the amount of data that is lost during a disaster recovery.

Disaster recover options

Summary of options

The disaster recovery options offered by Sitecore Managed Cloud are summarized in the tables below. The full details of each option can be found in the following sections, but the main decision points between them are (1) proactive vs. reactive initiation of failover and (2) the recovery time in the event of an outage (RTO).

Table 1: Disaster Recovery Services

Service	Secondary environment	Failover type	Recovery process	Reason to choose
Basic DR	Created on-demand after event, this is commonly referred to as a Hot-Cold deployment	Manual – Customer request or permission received from customer	• Customer permission or request • Deploy environment • Restore data • Customer validate environment • Go live	Cost-effective option with a longer RTO
Managed DR	Created in advance with a full-size deployment in the secondary site, this is commonly referred to as a Hot-Warm or Hot-Hot deployment	Automated – Failover initiated from Infrastructure Monitoring	• Replication of data • Switch to secondary instance • Index re-build • Go live	Shorter RTO with automated failover

Table 2: Disaster Recovery RPO & RTO

Service	Backup technologies	Recovery point objective (RPO)	Recovery time objective (RTO) – Technology only	Failback time – Technology only
Basic DR	SQL Azure backup Azure APIs	SQL: 3 hours WebApp: 3 hours	4 hours	4 hours
Managed DR	SQL Azure geo-replication Azure APIs	SQL: 5 seconds WebApp: 3 hours	< 1 minute	10 minutes

NOTE: The Technology RTO values only account for the time it takes to restore the Sitecore platform. If manual steps involving the customer or partner are required, such as attempting to get in touch with authorized contacts for permission to failover or validation of customizations prior to going live, the effective RTO will be extended.

What is replicated between data centers?

A Sitecore environment is made up of five different Azure resource types:

- App Services
- Azure SQL
- Application Insights
- Azure Search / SOLR
- Redis Cache

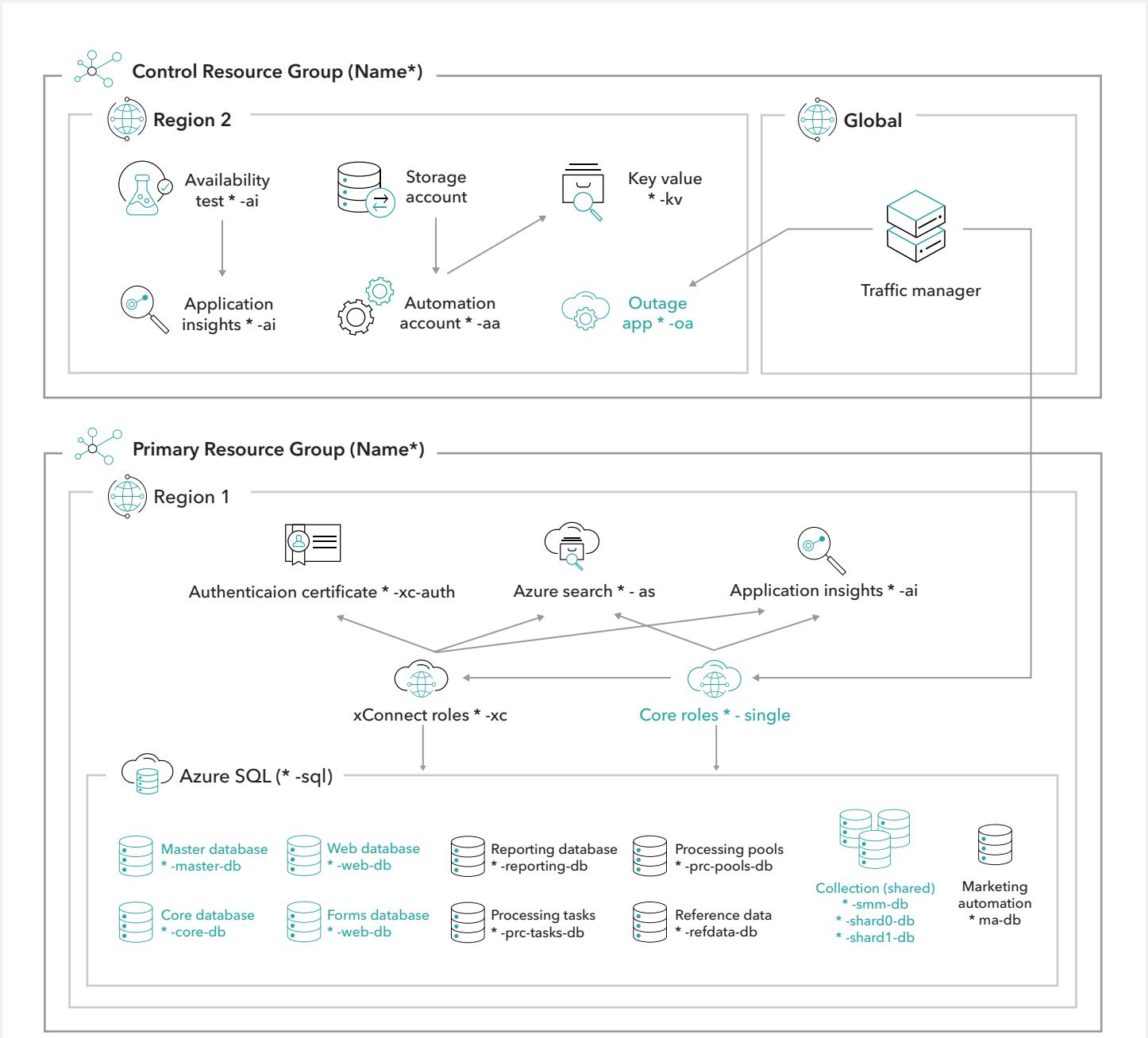
The sizes and instance counts of all of these resources are replicated to a secondary data center, but only App Services and Azure SQL have their files/data backed up and restored. The other services do not have their data replicated because it is transient or is not required for a successful restore:

- **Azure Search** – Does not provide a way to back up or geo-replicate indexes, so Sitecore indexes are rebuilt in the secondary data center instead
- **Redis Cache** – Contains user session data, which typically expires before a Sitecore site can be restored, so it is not included as part of the DR strategy
- **Application Insights** – Only contains health monitoring data, which is not required for the runtime of the Sitecore site

NOTE: Sitecore Managed Cloud only replicates the application default topology. Customer and Partner provisioned services are not included and would need to be maintained and replicated in the secondary site by the responsible party.

Option #1: Basic disaster recovery

In the event of an outage in the primary data center, a new Sitecore production environment is created in a secondary data center. During the creation of the secondary environment, a simple outage page is shown to customers to make them aware that the site is temporarily down. Because a new environment must be created in a secondary data center, this recovery option has the longer RTO but is the less expensive option.



Procedure

Sitecore Managed Cloud provides the following steps as part of the Basic DR offering.

Setup

The setup steps are as follows:

- Scheduled back up every 3 hours of the assets below into the secondary data center based on the designated RPO:
  - Back up the databases
  - Back up the WebApps
  - Back up the connection strings (for the credentials)
  - Back up the sizes/tiers of the resources
- Setup of outage page to be shown to the customers while there is a disaster
- Setup of Traffic Manager to switch between primary CD and outage page
- Setup of email alerts to notify Managed Cloud operations team when the availability tests fail

Initiating a failover

Sitecore Managed Cloud is continuously checking the health of the primary data center Content Delivery role by pinging it from five different data centers around the globe. If three out of the five data centers report an issue, then an alert is raised to the Managed Cloud Operations team.

The Sitecore Managed Cloud operations team begins to investigate the Sitecore environment in the primary data center to see if there is a legitimate issue and not a false positive. The operations team performs the following validation checks in the primary data center:

- Check for alerts raised by the Azure Resources used by the Sitecore site
- Check if Traffic Manager is reporting a degraded endpoint
- Check the Azure Status site for known data center issues https://azure.microsoft.com/en-gb/status/

Should the Cloud Operations team determine that there is an unrecoverable issue in part or all of the underlying infrastructure in the primary data center, then the failover confirmation process begins, and the customer is contacted.

Failover/recovery confirmation

Once a disaster happens, Sitecore notifies customer to confirm the decision on running the full recovery operation. Once confirmed, Sitecore triggers the recovery procedure via the following steps:

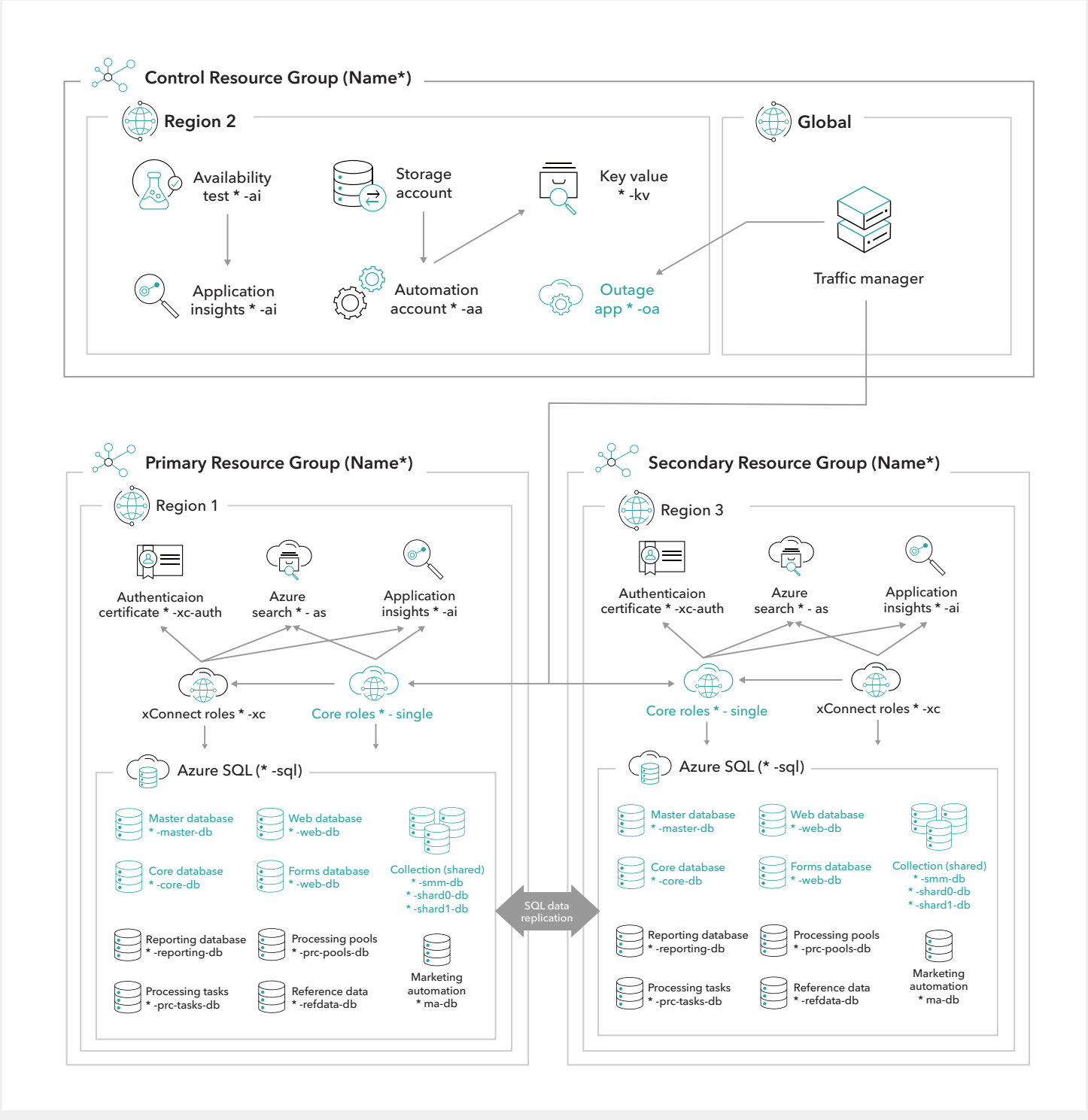
- Deploy a new Sitecore environment to the secondary data center
- Restore the WebApps from the last backup
- Restore the databases from the last backup
- Update the connection strings with the credentials from the primary Sitecore instance
- Re-index content and xDB indexes
- Switch the traffic manager to the secondary Sitecore instance

Failback

Return to the primary data center is performed at a time agreed to by the customer and the Sitecore Managed Cloud operations team. Since any data in the primary data center is now stale, the failover steps must be repeated from secondary to primary to bring the data up to date.

Option #2: Managed disaster recovery

For the Managed DR service, the secondary data center always has a complete Sitecore environment setup with all Azure Resources matching the sizes/instances count in the primary data center. This recovery option has a shorter RTO as the failover is automatic.



Procedure

Sitecore Managed Cloud provides the following steps as part of the Managed DR offering.

Setup

The setup steps are as follows:

- Deploy a new Sitecore environment in the secondary data center and shut down the Web Apps
- Sync the data between the primary and secondary Azure SQL using Active Geo-Replication
- Sync the sizes/tiers of all Azure Resources
- Sync the file contents of all WebApps
- Setup of outage page to be shown to the customers while there is a disaster
- Setup of Traffic Manager to switch between primary CD and outage page
- Setup of email alerts to notify Sitecore Managed Cloud operations team when the availability tests fail

Initiating a failover

The Sitecore Managed Cloud operations team is continuously checking the health of the primary data center Content Delivery role by pinging it from five different data centers around the globe. If three out of the five data centers report an issue, then an alert is raised to the Sitecore Managed Cloud Operations team and the automated failover is triggered.

Failover/recovery

In the event of a failure disaster, Sitecore notifies a customer of the automated failover event. Sitecore triggers the recovery procedure via the following step:

- Traffic Manager automatically redirects traffic via a DNS change to the secondary Sitecore instance

Failback

Return to the primary data center is performed automatically by doing the failover steps again and directing the traffic to the original Sitecore instance once it has been detected as online.

Considerations

Disaster recovery introduces new considerations when building a Sitecore solution. This section highlights some of the most common ones.

Choosing your Azure region

Azure organizes its data centers into regions with a latency-defined perimeter and connected through a dedicated regional low-latency network. When choosing a secondary data center, it is recommended to choose one in the same region as the primary to ensure fast backups and consistent customer delivery speeds.

3rd party service API's

If the Sitecore implementation is using any 3rd party service APIs that limit access based on IP, then it is essential to register the IPs of the secondary data center with the service. Failure to register the IPs could result in a delay in bringing the secondary Sitecore environment online.

Outage page

Sitecore Managed Cloud uses Azure Storage to host an informational page in case of an outage. Using Azure Storage means the outage page needs to be static (i.e., pure HTML) which means no custom backend code can be executed for the page. It is recommended the outage page only contain the necessary information to ensure customers that the site will be back online soon, for example:

1. Mentioning the site is temporarily unavailable
2. Support staff is aware and is working on it
3. Approximate recovery time

The outage page returns a 503 (SERVICE UNAVAILABLE) HTTP Code with a custom page. If a search engine is crawling your site and sees a 503, then it understands that your site is down temporarily and comes back and re-indexes later.

Certificates

Any certificates used for the Sitecore site are stored in a Key Vault in the secondary data center. In the case of a failover, the certificates in the Key Vault are used as part of the new setup. If there are any changes to a certificate (e.g., expiration),

then the Sitecore Managed Cloud Operations team must be notified so the certificate in the Key Vault can be updated.

To help meet the Managed DR RTO, it is essential to keep the certificates up to date in the secondary data center. If you replace certificates at the primary site because an expiration date is upcoming, you should also update the certificates at the secondary site at the same time.

Supported Sitecore versions

The latest compatibility information for Sitecore versions is available at https://kb.sitecore.net/articles/768387.

No removal of Control Resource Group

The Control Resource Group contains all resources used to restore Sitecore successfully in a secondary data center. Deleting the Control Resource Group or its resources can lead to an inability to perform a successful recovery.

xDB excluded from RTO

For disaster recovery, the RTO does not cover the xDB index rebuild due to the significant amount of time it can take for a large content database. If the analytics indexes are not rebuilt, this should only affect functionality that depends on lists (e.g., EXM) and, generally speaking, should not affect the runtime site.

xConnect Search Indexer

Sitecore can only have one active xConnect Search Indexer WebJob across a solution. In the case of a failover and restoration of service, the indexer must be shut down.

Certificates in Azure

Only one website certificate is supported with Managed Cloud DR at this time, however wildcard certificates are supported.

Azure requirements & cost considerations

All disaster recovery options are dependent on Azure WebApp Backup and Traffic Manager, which require a minimum of the Standard Tier of Web Apps.

Unsupported scenarios

There are a small set of scenarios where it might not be possible to restore a production site into the secondary data center. Some examples of some of these are (not exhaustive):

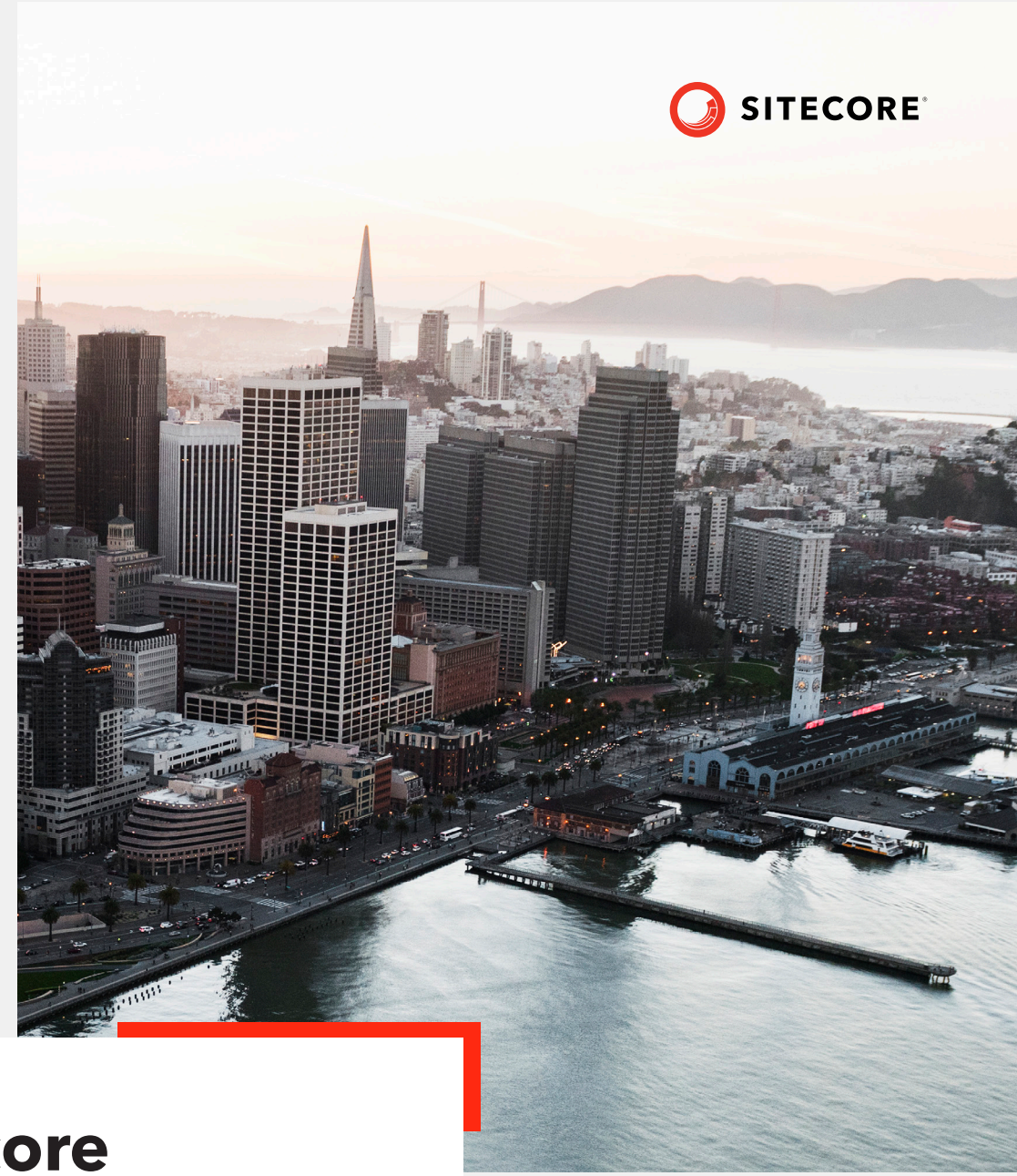
- A global Azure service such as authentication or Traffic Manager is down
- Both the primary and the "backup" (secondary) data centers are down at the same time
- Large scale global network failure/outage
- Network connection between primary and secondary data centers is down

Investigation and the RTO relationship

There exists a direct relationship between the amount of time investigating a failure or event prior to failover and the Return to Operations (RTO) time. The Base DR process includes time to investigate and confirmation from the Customer prior to provisioning the secondary site. In the Managed DR option, failover will be initiated automatically since the secondary site should be a full replica of the environment, which highlights the importance of maintaining any custom additions or modifications made to the default topology by the Customer or Partner.

Misconfiguration and custom code issues

Failures due to misconfiguration of a site or application code issues will not be resolved by a failover since the configuration is replicated exactly to the secondary site.



About Sitecore

Sitecore is the global leader in experience management software that enables context marketing. The Sitecore® Experience Platform™ manages content, supplies contextual intelligence, automates communications, and enables personalized commerce, at scale. It empowers marketers to deliver content in context of how customers have engaged with their brand, across every channel, in real time – before, during, and after a sale. More than 5,200 brands--including American Express, Carnival Cruise Lines, easyJet, and L'Oréal – have trusted Sitecore for context marketing to deliver the personalized interactions that delight audiences, build loyalty, and drive revenue.

Learn more at Sitecore.com.