INSTRUCTURE

Instructure Disaster Recovery Plan & Procedures

Instructure Security, Engineering, & Operations

February 2020

INSTRUCTURE

BRIDGE CANVAS MasteryConnect PORTFOLIUM

Table of Contents

1. Disaster Recovery Plan and Procedures

1.1 Overview

This document describes the plan and procedures that Instructure has established to recover from disasters affecting its production operations. We describe how the Software as a Service (SaaS) offering has been architected to recover from disaster scenarios, the steps to be taken when disasters are declared, the policies regarding notification of partners during disasters, and several example scenarios and how they affect the service. Our disaster recovery procedures address events which would affect an entire facility. Failures of individual components are recovered through architectural redundancies and fail-over mechanisms.

1.2 Policy and Practices

Definition of Disaster

A disaster is defined as any disruptive event that has potentially long-term adverse effects on the Instructure service. In general, potential disaster events will be addressed with the highest priority at all levels at Instructure. Such events can be intentional or unintentional, as follows:

Natural disasters: Tornado, earthquake, hurricane, fire, landslide, flood, electrical storm, and tsunami.

Supply systems: Utility failures such as severed gas or water lines, communication line failures, electrical power outages/surges, and energy shortage.

Human-made/political: Terrorism, theft, disgruntled worker, arson, labor strike, sabotage, riots, vandalism, virus, and hacker attacks.

Declaration of Disaster

All potential disasters will be escalated immediately to a designated officer who is authorized to declare a disaster. The incident officer will be responsible for assessing the event and confirming the disaster. Once the disaster is declared, the incident officer will be responsible for directing recovery efforts and notifications.

1.3 Key Organizational Resources

Disaster Recovery Team

The Disaster Recovery Team (DRT) is made up of key engineers and operations employees. The responsibilities of the DRT include:

- Establish communication between the individuals necessary to execute recovery
- Determine steps necessary to recover completely from the disaster
- Execute the recovery steps
- Verify that recovery is complete
- Inform the incident officer of completion

1.4 Notification

There are several parties that must be notified at various stages during disaster events.

Notifying Staff

The incident officer is responsible for making sure the DRT and any other necessary staff are notified of a disaster event and mobilized. Notification of staff will generally happen via cell phone.

Notifying Clients and Business Partners

Clients and business partners will be notified at various stages of disaster recovery using email and our official status page. If these methods are unavailable, notification will happen via alternative means (cell phone, etc.) as provided by each client institution. The stages of notification are:

Disaster Declaration: When a disaster is declared, the notification will include a description of the event, the effect to the service, and any potential impact to data.

Completion of Recovery: Once recovery is complete and the service is available, the notification will include general information about steps taken to recovery, and any data that may have been impacted. If the recovery is partial and the service is still in a degraded state, this notification will include an estimate of how long the degradation will continue.

Testing

A Disaster Recovery Plan is only useful insofar as it is tested regularly. The incident officer is responsible for ensuring that the plan is tested in its entirety at least annually and in part whenever major components are changed.

1.5 Disaster Recovery Solution

Current Operating Infrastructure

Instructure's software is based on a multi-tier cloud-based architecture. Each component is redundant with active monitoring for failure detection and failover. The different tiers are:

Load Balancers

All web traffic to instructure.com is served by two load balancers in an active/passive configuration. The load balancers are responsible for directing traffic to the next tier.

App Servers

App servers process incoming requests from clients from the load balancers. App servers implement all the business logic, but do not persist any important data. Asynchronous jobs also run on the app servers. The number of app servers varies based on demand, but will always be at least two in active/active configurations.

Caching

To improve website performance, Instructure's software aggressively caches data in a caching layer. The data stored here is strictly a performance cache. Any data loss resulting from the loss of any of these servers would be limited to a small number of page view statistics that may not have been flushed to persistent storage. The number of cache servers is variable and the cache data will be partitioned among all servers.

Databases

Most structured data—courses, user information, and assignments, for example—is stored in a database. This data is sharded between instances based on account and on demand. Each shard has a Master and a Slave database, located in geographically separate sites. The data from each Master is replicated asynchronously in near real-time to its corresponding Slave. Each Master is also backed up completely every 24 hours, and the backup is stored in a third geographically separate site. The infrastructure also includes an internal database proxy layer for the relational databases that enables the Operations team to perform maintenance on the relational database servers with minimal downtime.

Third-Party Object Store

Content—such as documents, PDFs, audio, and video—is stored in a third-party scalable object store.

Objectives

In the context of a disaster recovery scenario, there are two terms which are commonly used to describe how the data may be affected: Recovery Time Objective (RTO) and Recovery Point Objective (RPO). The RTO is how long it will take to make access to the data available again, and the RPO is how much of the most recent data will be preserved. For example, if it takes 12 hours for a service to recover, but on a failure up to 24 hours of data may be lost, the RTO is 12 hours and the RPO is 24 hours.

The platform has been architected to achieve an exceptionally low RPO and RTO in the common case due to the distributed and resilient nature of its infrastructure. For the vast majority of failure scenarios, the need to “failover” to another cloud region is obviated. In the event of a catastrophe, which would necessitate the need to move hosting regions, it would in all likelihood require multiple days for Instructure to restore service to an acceptable level.

1.6 Backup and Recovery Practices

Instructor, student, course, assignment data from database

Backup	**Data is replicated asynchronously in near real-time to remote site (monitored, etc.).** **Nightly backups of every database are stored at a remote site.**
Recovery	When slave is caught up: • Promote slave database to master, following replication docs • Provision new database using provisioning tools • Establish new database as new slave, following replication docs

Backup	**Data is replicated asynchronously in near real-time to remote site (monitored, etc.).** **Nightly backups of every database are stored at a remote site.**
	• When slave is > 24 hours behind • Copy last nightly backup to slave database • Load slave with nightly backup • Provision new database using provisioning tools • Establish new database as new slave, following replication docs

Static assets from courses and assignments such as documents and other content files

Backup	**Files are stored on a scalable, protected, geographically redundant storage system (Amazon S3)**
Recovery	Recovery in case of failures is built into the scalable storage system

Web applications

Backup	**Web application source code is stored in versioned source control and backed up to multiple locations** **There is no state stored on the application servers that would need to be backed up**
Recovery	Not applicable

1.7 Sample Disaster Scenarios

Following are several different possible disaster scenarios and their RPO/RTO, services affected, and recovery overview. Note that these are intended only to convey magnitude of impact and recovery efforts required under different situations.

Complete Loss of a Master Database

Services Affected	**Most accounts hosted on the affected database**
Recovery Overview	When the slave database is up-to-date (common case): The slave is promoted to be the new master according to the steps described above When the slave database is inconsistent: The slave is populated with the latest nightly snapshot and brought online as the new master

Services Affected	Most accounts hosted on the affected database
RPO	5 minutes (consistent slave, common case), 24 hours (inconsistent slave)
RTO	1 hour (consistent slave, common case), 6 hours (inconsistent slave)
Likelihood	Once a year

Simultaneous Complete Loss of Master and Slave Databases

Services Affected	Most accounts hosted on the affected database.
Recovery Overview	New master and slave databases are brought online in separate locations The master database is populated with data from the offsite backup App servers pointed to new master database Replication re-established with the new slave database
RPO	24 hours
RTO	6 hours

Services Affected	Most accounts hosted on the affected database.
Likelihood	Once every 20 years (the master and slave databases are hosted in geographically separate locations, which makes simultaneous failure very unlikely)

Database Destruction by Hacker

Services Affected	Most accounts hosted on the affected database.
Recovery Overview	The master database is restored from the most recent complete backup Replication is re-established with the slave database
RPO	24 hours
RTO	6 hours
Likelihood	Once every 10 years

Complete Loss of Primary Hosting Facility

Services Affected	Platform for most accounts

Services Affected	Platform for most accounts
Recovery Overview	New load balancers and app servers are brought up in the secondary site with the slave database The old slave database is promoted to master database. A new database slave is brought up at a third site and replication re-established DNS is pointed to the new load balancers at the recovery site and services are restored
RPO	4 hours
RTO	Commercially Reasonable
Likelihood	Extremely Unlikely