CANVAS
BY INSTRUCTURE

ARCHITECTURE OVERVIEW

Engineering, Security, and Operations

March 2022

Table of Contents

CANVAS
BY INSTRUCTURE

Canvas Architecture

Over the past several years, it's no surprise that Software as a Service (SaaS) has become the 'in' thing. Companies the world over have come to understand and see the tremendous advantages that SaaS can bring their organizations, as opposed to traditional on-premise solutions. This is why Canvas - the world's leading Learning Management System - was born in the cloud. From its inception, our founders developed Canvas to be a cloud-native, multi-tenant solution architected to automatically scale and serve millions of simultaneous users around the world. Not only is Canvas delivered as a convenient SaaS model where our customers never need to worry about service packs, updates, versions, backups, and security, but we also developed it using open standards and technologies that all come together in a seamless, easy-to-use learning platform, enabling our users to focus their time where it matters most; on their ability to teach, learn, and engage across a wide variety of environments regardless of device, operating system, or location. The following document describes the Canvas architecture for those curious technical types who love getting into the detail of just how we make Canvas work its magic.

Hosting Regions

For US customers, Instructure uses two Amazon Web Services (AWS) regions, ensuring that client data is not stored outside of the United States:

- US East (Northern Virginia)
- US West (Oregon)

For international clients, Instructure uses the following AWS regions:

- Canada Central (Montreal)
- EU West (Ireland)
- EU Central (Germany)
- Asia Pacific (Sydney)
- Asia Pacific (Singapore)

In each region we operate, we utilize three (3) Availability Zones (AZ) for redundancy.

Product Security

Instructure holds the following certifications which are independently audited by a 3rd party:

- SOC 2 Type II
- ISO/IEC 27001:2013

The SOC 2 report can be made available under a Mutual Non-Disclosure Agreement (MNDA). The ISO 27001 compliance certificate can be made available to anyone upon request.

As one of the benefits of utilizing AWS cloud infrastructure, we also benefit from the following security certifications:

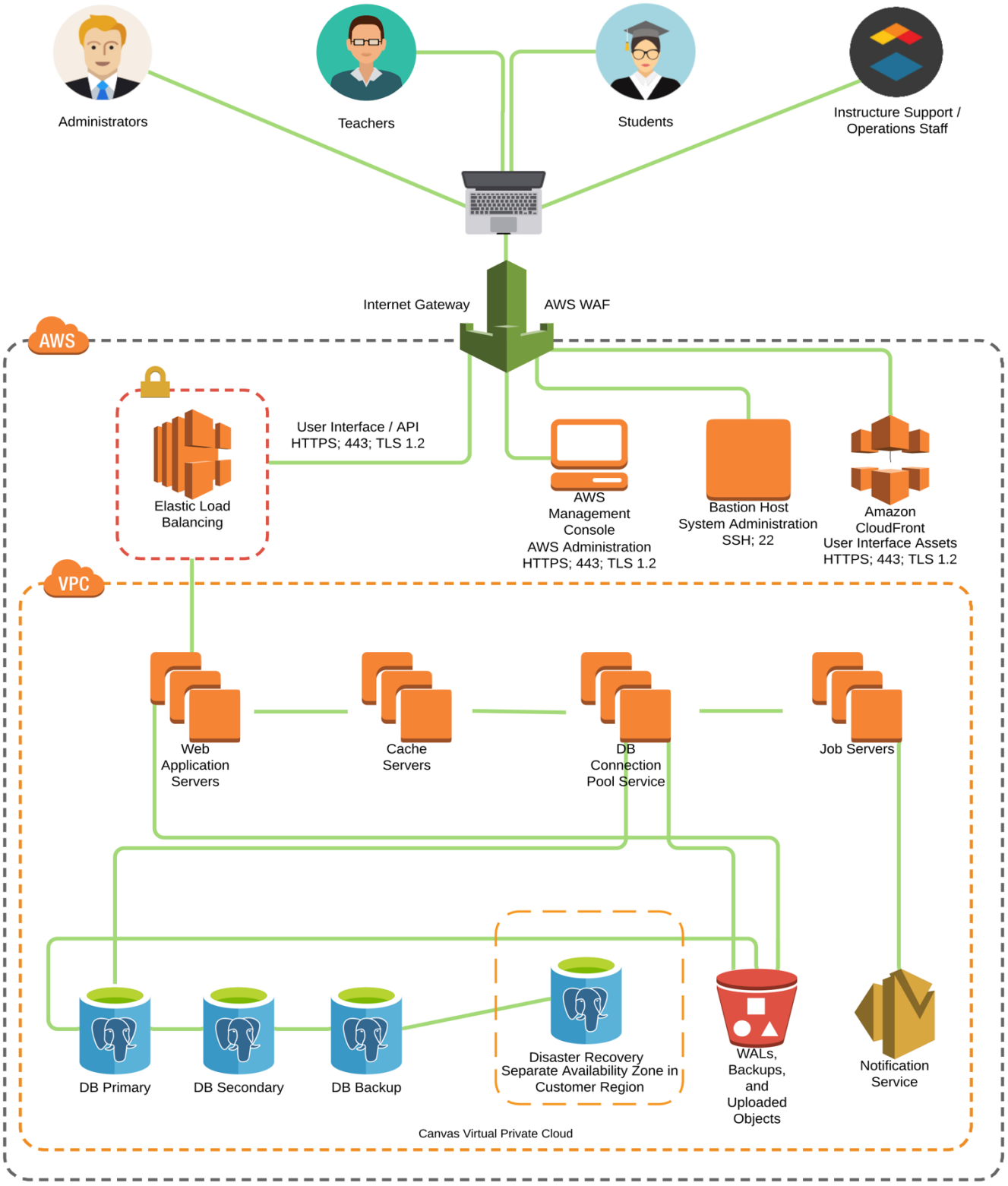
- SOC 1 Type II (ISAE 3402), SOC 2 Type II, and SOC 3 Type II reports
- ISO 9001, 27001 (CSA Star Level 2), 27017, and 27018 certified
- Level 1 PCI-DSS service provider
- FISMA-Moderate operation level
- GDPR ready, FERPA compliant (shared responsibility model)
- Cyber Essentials PLUS certification

Separation of Tenant Data

Separation of tenants is accomplished in AWS via logical separation in natively multi-tenant software. Customer data is segregated via database sharding (horizontal partitioning). Horizontal partitioning is a design principle whereby rows of a database table are held separately, rather than splitting by columns (as for normalization). Each partition forms part of a shard. The advantage is the number of rows in each table is reduced, reducing index size, and improving performance.

Sharding is based on real-world aspect of the data (e.g., segmented by customer) and data cannot leak from one shard to another, nor can clients gain access to data in another shard as the method of inferring the client shard is accomplished after authentication. As client credentials are only valid for a single account, and therefore shard, user authentication is intrinsically tied to the shard identity. Validation of segregated client data occurs during weekly disaster recovery testing.

Architecture and Data Flow Diagram (AWS)



Scaling, Backup, Recovery, and Redundancy

AWS data center electrical and network systems are designed to be fully redundant and maintainable without impact to operations, 24 hours a day, seven days a week. Uninterruptible Power Supply (UPS) units are available in the event of an electrical failure for critical and essential loads in the facility. Data centers use generators to provide backup power for the entire facility.

The Canvas LMS architecture replicates data in near real-time and data is backed up on a daily basis. Instructure creates daily offsite database backups of Canvas data and content including course content, student submissions, student-created content, analytics, rubrics, learning outcomes, and metadata. Data is stored redundantly in multiple data centers and multiple geographic locations through Amazon S3. For further detail on backups, please see Instructure's Business Continuity & Disaster Recovery Paper.

The Canvas LMS architecture is horizontally scalable and uses a mix of in-house developed and AWS-provided technologies, enabling it to respond to usage spikes in real-time and accommodate expanded, long-term usage. Through automatic scaling and automated provisioning technology, Canvas adjusts cloud resources to handle large usage loads before they cause slowdowns. When concurrent user numbers grow, Canvas automatically adds resources so users don't experience outages or slowdown.

Assuring the recovery and redundancy of the Canvas LMS platform, we take advantage of multiple geographically separate sites and Availability Zones which provide resilience in the face of most failure modes including natural disasters or system failures. The Canvas application is designed to make full use of the real-time redundancy and capacity capabilities offered by AWS, running across multiple availability zones in regions throughout the world. Primary storage is provided by Amazon S3, which is designed for durability exceeding 99.99999999%.

The Canvas LMS architecture is also resilient to failure and capable of rapid recovery from component failure. The Canvas application, its media and file storage, and its databases are each independently redundant. If an application hosting node were to fail, all traffic would transfer to living nodes. If load increases, an automated provisioning system ensures that more hosting nodes are made available to handle the traffic—either in response to increased load or in predictive anticipation of future workloads. The database and file stores are also horizontally scalable, adding capacity for both additional storage and load as needed.

Predictive Scaling

Canvas is a Software-as-a-Service (SaaS), hosted by the most established and trusted cloud hosting provider in the world: Amazon Web Services. From the launch of Canvas in 2011 as the first cloud-native LMS, Instructure has exclusively provided and supported fully cloud-hosted, SaaS education technology platform. During this time, we have analyzed data and garnered usage trends, whereby allowing us to predict when a usage peak for any given customer is likely to occur.

It is with this data that we leverage AWS' EC2 Auto Scaling technologies to take scaling to a whole new level to handle sometimes unforeseen cycles of higher volume. Using Predictive Scaling allows us to predict when a usage peak for any given customer is likely to occur. It learns from past patterns and launches instances in advance of predicted demand, giving instances time to warm up and be ready preemptively before a high-demand situation exists rather than in response to one. Additionally, it provides flexible downscaling ensuring that system resources are not removed too quickly when load begins to fall.

Taking this a step further, we also utilize Instructure's own scaling technology called HotTub. HotTub is a reactive autoscaling mechanism specifically for Canvas that can scale up our application clusters in response to unexpected jumps in user activity up to 20 times faster than Amazon's own autoscaling service. Since we are able to look back to previous days or weeks and predict what resources will be needed ahead of time, our HotTub scaler can have a pool of pre-warmed application servers that are ready to be put into service at a moment's notice. Between both these services, Canvas provides unmatched stability and scalability regardless of user load.

Load Balancers

AWS Elastic Load Balancers are deployed in a highly available active/active configuration, which handles incoming requests and dispatches the underlying connections evenly to available application servers. The load balancer maintains a dynamic list of available application servers for dispatch. The load balancer sends regular heartbeats—a simple network message—to verify the application server is healthy, available, and capable of receiving additional work. The load balancer will not dispatch work to unresponsive application servers. Additional capacity is automatically added to the load balancing pool as traffic and demand increases.

Application Servers

Application servers process incoming requests from the load balancers. They are responsible for executing the business logic, rendering HTML, and returning some static assets to the Canvas LMS user's web browser. Additionally, these servers are balanced across multiple availability zones to ensure maximum fault tolerance.

Application servers are constantly monitored individually for load and capacity information. When all application servers reach a certain load threshold, a new application server is automatically provisioned and deployed. Instructure's in-house automation can dynamically and intelligently schedule new application servers in anticipation of high load times, such as during the beginning and end of semesters.

Cache Servers

The caching layer provides performance optimization. A healthy cache means the application servers need to make fewer trips to the database which speeds up response times. The caching layer is made up of numerous machines running Redis. Data is spread out evenly across all machines. Additionally, Amazon CloudFront (a caching CDN) is used to quickly deliver static assets to Canvas LMS users. These CDN endpoints are globally distributed, thereby making the network path for these requests as efficient as possible.

Cache servers are constantly monitored. When a cache server fails, a new one is provisioned and deployed to take its place. When a cache server fails, the data that would have been stored on it, is simply retrieved from the database instead.

Cache servers are completely memory based. Memory usage is monitored continuously. When the cache hit rates falls below an acceptable threshold, new cache servers are provisioned and deployed.

Database Servers

Course and user data are stored in relational databases. The databases are partitioned by client institution for performance and data isolation purposes. Each institution utilizes a pair of databases: A Primary database and a Secondary database in a separate availability zone.

There is also a third Backup server in each region and in a separate availability zone. All database changes are streamed in real-time to each other, and to a durable data layer (S3). This means Canvas LMS database information for all customers is stored in three separate geographically separated locations. Additionally, database backups (a different form of data redundancy for different purposes) are tested weekly.

If the Primary database fails, the Secondary will be promoted to Primary and a new Secondary database provisioned and deployed. Upon failure of the Secondary database, a new Secondary database is provisioned and deployed. In the unlikely event of simultaneous component failure or data corruption, the standby backup server can be used to create a new database pair.

Databases are constantly monitored for resource usage and response time. If either database approaches peak load, individual customers will be relocated to clusters with available capacity.

Distributed File Storage

Course media, including videos, image files, audio recordings, etc. and student-uploaded files, like assignments, documents, and learning artifacts are stored outside the database in a separate and scalable Amazon Simple Storage Service (S3) bucket that is designed for durability exceeding 99.99999999%. All objects within the S3 buckets are encrypted and replicated between geographically separate sites and have version control enabled so previous versions of an object can be restored with minimal effort.

Conclusion

Following industry best practices, we have created a dynamic and highly scalable, cloud-native web application which has become the most reliable Learning Management System in the world, used and respected by prestigious global learning institutions such as all eight of the Ivy League schools, University of Oxford, KTH Royal Institute of Technology, University of Amsterdam, National University of Singapore, Royal Melbourne Institute of Technology, and countless more. In taking great care and diligence in creating a best-of-breed SaaS application, our architecture stands on its own in the edu-tech industry for its robustness, hardening, scalability and reliability.