



CANVAS
BY INSTRUCTURE

ARCHITECTURE OVERVIEW

Engineering, Security,
Operations

Marzo 2022



Índice

Arquitectura del Canvas.....	3
Escalamiento, Backup, Recovery y Redundancia	6



CANVAS
BY INSTRUCTURE



Arquitectura del Canvas

En los últimos años, no sorprende que el software como servicio (SaaS) se haya convertido en la moda. Las empresas de todo el mundo han llegado a comprender y ver las enormes ventajas que SaaS puede brindar a sus organizaciones, en comparación con las soluciones locales tradicionales. Es por eso que Canvas, el sistema de gestión de aprendizaje líder en el mundo, nació en la nube. Desde sus inicios, nuestros fundadores desarrollaron Canvas para que fuera una solución multiusuario nativa de la nube diseñada para escalar automáticamente y servir a millones de usuarios simultáneos en todo el mundo. Canvas no solo se entrega como un modelo SaaS conveniente en el que nuestros clientes nunca tendrán que preocuparse por los paquetes de servicios, las actualizaciones, las versiones, las copias de seguridad y la seguridad, sino que también lo desarrollamos utilizando estándares abiertos y tecnologías que se unen en un entorno sencillo y sin inconveniente, plataforma de aprendizaje fácil de usar, que permite a nuestros usuarios concentrar su tiempo donde más importa; en su capacidad para enseñar, aprender y participar en una amplia variedad de entornos, independientemente del dispositivo, el sistema operativo o la ubicación. El siguiente documento describe la arquitectura de Canvas para aquellos tipos técnicos curiosos que aman entrar en los detalles de cómo hacemos que Canvas funcione su magia.

Regiones de Alojamiento

Para los clientes de EE. UU., Instructure utiliza dos regiones de Amazon Web Services (AWS), lo que garantiza que los datos del cliente no se almacenen fuera de los Estados Unidos:

- Este de EE. UU. (Virginia del Norte)
- Oeste de EE. UU. (Oregón)

Para los clientes internacionales, Instructure utiliza las siguientes regiones de AWS:

- Canadá Central (Montreal)
- Oeste de la UE (Irlanda)
- Centro de la UE (Alemania)
- Asia Pacífico (Sídney)
- Asia Pacífico (Singapur)

En cada región que operamos, utilizamos tres (3) zonas de disponibilidad (AZ) para redundancia.

Seguridad del Producto

Instructure posee las siguientes certificaciones que son auditadas de forma independiente por un tercero:

- SOC 2 Tipo II
- ISO/CEI 27001:2013



El informe SOC 2 puede estar disponible en virtud de un Acuerdo de confidencialidad mutua (MNDA). El certificado de conformidad con la norma ISO 27001 puede ponerse a disposición de cualquier persona que lo solicite.

Como uno de los beneficios de utilizar la infraestructura en la nube de AWS, heredamos de las siguientes certificaciones de seguridad:

- Informes SOC 1 Tipo II (ISAE 3402), SOC 2 Tipo II y SOC 3 Tipo II
- Certificación ISO 9001, 27001 (CSA Star Level 2), 27017 y 27018
- Proveedor de servicios PCI-DSS de nivel 1
- FISMA-Nivel de operación moderado
- Preparado para GDPR, compatible con FERPA (modelo de responsabilidad compartida)
- Certificación Cyber Essentials PLUS

Separación de datos de inquilinos

La separación de los inquilinos se logra en AWS a través de la separación lógica en el software de múltiples inquilinos de forma nativa. Los datos de los clientes se segregan mediante la fragmentación de la base de datos (partición horizontal). La partición horizontal es un principio de diseño mediante el cual las filas de una tabla de base de datos se mantienen por separado, en lugar de dividirse por columnas (como para la normalización). Cada partición forma parte de un fragmento. La ventaja es que se reduce el número de filas en cada tabla, lo que reduce el tamaño del índice y mejora el rendimiento.

La fragmentación se basa en el aspecto del mundo real de los datos (p. Ej., Segmentados por cliente) y los datos no pueden filtrarse de un fragmento a otro, ni los clientes pueden obtener acceso a los datos en otro fragmento, ya que el método de inferir el fragmento del cliente se logra después de la autenticación. Como las credenciales de cliente solo son válidas para una única cuenta y, por lo tanto, la autenticación de usuario está intrínsecamente vinculada a la identidad del fragmento. La validación de los datos de los clientes separados se produce durante las pruebas semanales de recuperación ante desastres.

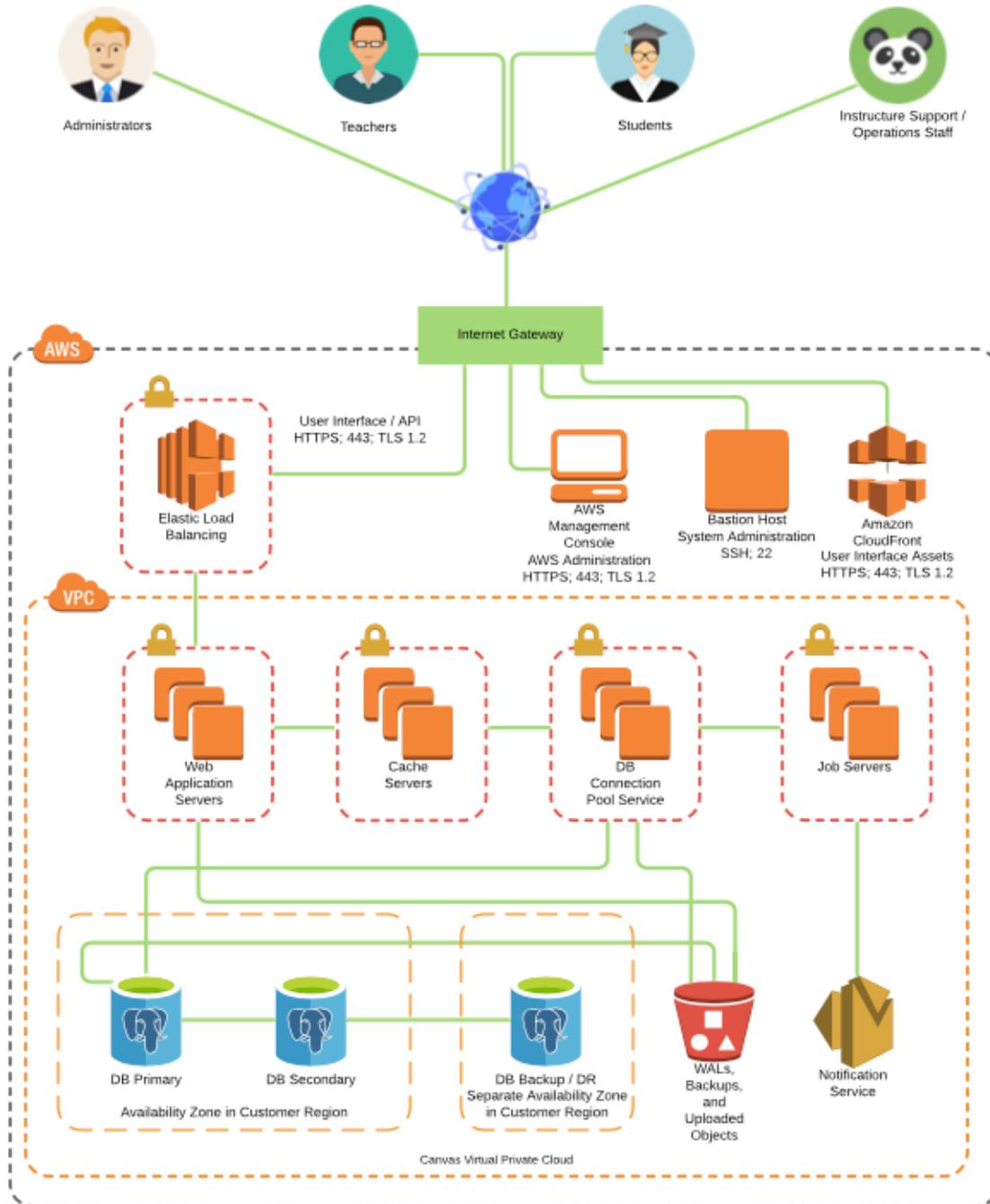


Diagrama de arquitectura y flujo de datos



Canvas Amazon Web Services (AWS) Architecture

INSTRUCTURE
Commercial in Confidence
Last Revision December 2020



Escalamiento, Backup, Recovery y Redundancia

Los sistemas eléctricos y de red del centro de datos de AWS están diseñados para ser completamente redundantes y fáciles de mantener sin afectar las operaciones, las 24 horas del día, los siete días de la semana. Las unidades de suministro de energía ininterrumpida (UPS) están disponibles en caso de una falla eléctrica para las cargas críticas y esenciales en la instalación. Los centros de datos utilizan generadores para proporcionar energía de respaldo a toda la instalación.

La arquitectura de Canvas LMS replica los datos casi en tiempo real y se realiza una copia de seguridad de los datos a diario. Instructure crea copias de seguridad diarias de bases de datos externas de los datos y el contenido de Canvas, incluido el contenido del curso, las presentaciones de los estudiantes, el contenido creado por los estudiantes, análisis, rúbricas, resultados de aprendizaje y metadatos. Los datos se almacenan de forma redundante en varios centros de datos y varias ubicaciones geográficas a través de Amazon S3.

La arquitectura de Canvas LMS es escalable horizontalmente y utiliza una combinación de tecnologías desarrolladas internamente y proporcionadas por AWS, lo que le permite responder a picos de uso en tiempo real y adaptarse al uso ampliado a largo plazo. Mediante el escalado automático y la tecnología de aprovisionamiento automatizado, Canvas ajusta los recursos de la nube para manejar grandes cargas de uso antes de que causen ralentizaciones. Cuando aumenta el número de usuarios simultáneos, Canvas agrega recursos automáticamente para que los usuarios no experimenten interrupciones o desaceleración.

Para garantizar la recuperación y la redundancia de la plataforma Canvas LMS, aprovechamos múltiples sitios y zonas de disponibilidad geográficamente separados que brindan resistencia frente a la mayoría de los modos de falla, incluidos desastres naturales o fallas del sistema. La aplicación Canvas está diseñada para hacer un uso completo de las capacidades de capacidad y redundancia en tiempo real que ofrece AWS, que se ejecuta en múltiples zonas de disponibilidad en regiones de todo el mundo. El almacenamiento primario lo proporciona Amazon S3, que está diseñado para una durabilidad superior al 99,99999999%.

La arquitectura de Canvas LMS también es resistente a fallas y capaz de recuperarse rápidamente de fallas de componentes. La aplicación Canvas, sus medios y almacenamiento de archivos, y sus bases de datos son cada uno independientemente redundantes. Si un nodo de alojamiento de aplicaciones fallara, todo el tráfico se transferiría a los nodos activos. Si la carga aumenta, un sistema de aprovisionamiento automatizado garantiza que haya más nodos de hospedaje disponibles para manejar el tráfico, ya sea en respuesta al aumento de carga o en anticipación predictiva de cargas de trabajo futuras. La base de datos y los almacenes de archivos también son escalables horizontalmente, lo que agrega capacidad para almacenamiento y carga adicionales según sea necesario.

Escalado Predictivo

Canvas LMS es un software como servicio (SaaS), alojado por el proveedor de alojamiento en la nube más establecido y confiable del mundo: Amazon Web Services. Desde el lanzamiento de Canvas en 2011 como el primer LMS nativo de la nube, Instructure ha proporcionado y respaldado exclusivamente la plataforma de tecnología educativa SaaS totalmente alojada en la nube. Durante este tiempo, hemos analizado datos y recopilado tendencias de uso, lo que nos permite predecir cuándo es probable que ocurra un pico de uso para un cliente determinado.

Es con estos datos que aprovechamos las tecnologías EC2 Auto Scaling de AWS para llevar el escalado a un nivel completamente nuevo para manejar ciclos a veces imprevistos de mayor volumen. El uso del escalado predictivo nos permite predecir cuándo es probable que se produzca un pico de uso para un cliente determinado. Aprende de patrones anteriores y lanza instancias antes de la



demanda prevista, dando tiempo a las instancias para calentarse y estar listas de manera preventiva antes de que exista una situación de alta demanda en lugar de en respuesta a una. Además, proporciona una reducción de escala flexible que garantiza que los recursos del sistema no se eliminen demasiado rápido cuando la carga comienza a disminuir.

Yendo un paso más allá, también utilizamos la propia tecnología de escalado de Instructure llamada HotTub. HotTub es un mecanismo de escalado automático reactivo específicamente para Canvas LMS que puede escalar nuestros clústeres de aplicaciones en respuesta a saltos inesperados en la actividad del usuario hasta 20 veces más rápido que el propio servicio de escalado automático de Amazon. Dado que podemos mirar hacia atrás a días o semanas anteriores y predecir qué recursos se necesitarán con anticipación, nuestro escalador HotTub puede tener un grupo de servidores de aplicaciones precalentados que están listos para ponerse en servicio en cualquier momento. Entre estos dos servicios, Canvas proporciona una estabilidad y escalabilidad inigualables, independientemente de la carga de usuarios.

Load Balancers

AWS Elastic Load Balancers se implementan en una configuración active/active muy disponible, que manejan las solicitudes de comunicación y las conexiones de las conexiones de conexión incluso a aplicaciones de aplicación. La carga de equilibrio mantiene una lista dinámica de aplicaciones de aplicaciones disponibles para dispatch. La carga de equilibrio de los mensajes de correo ordinario de la mente de la red de correo simple para comprobar el servidor de aplicaciones es saludable, disponible y capaz de recibir trabajos adicionales. El load balancer no funcionará a unresponso aplicaciones de aplicación. Adicionalmente se agrega automáticamente la carga de equilibrio del tráfico y la demanda de tráfico.

Servidores de Aplicación

Los servidores de aplicaciones procesan las solicitudes entrantes de los equilibradores de carga. Son responsables de ejecutar la lógica empresarial, renderizar HTML y devolver algunos activos estáticos al navegador web del usuario de Canvas LMS. Además, estos servidores están equilibrados en varias zonas de disponibilidad para garantizar la máxima tolerancia a fallos.

Los servidores de aplicaciones se supervisan constantemente de forma individual para obtener información sobre carga y capacidad. Cuando todos los servidores de aplicaciones alcanzan un determinado umbral de carga, se aprovisiona e implementa automáticamente un nuevo servidor de aplicaciones. La automatización interna de Instructure puede programar de forma dinámica e inteligente nuevos servidores de aplicaciones en previsión de tiempos de alta carga, como durante el comienzo y el final de los semestres.

Servidores de Caché

La capa de almacenamiento en caché proporciona una optimización del rendimiento. Una caché en buen estado significa que los servidores de aplicaciones deben realizar menos viajes a la base de datos, lo que acelera los tiempos de respuesta. La capa de almacenamiento en caché está formada por numerosas máquinas que ejecutan Redis. Los datos se distribuyen uniformemente en todas las



máquinas. Además, Amazon Cloudfront (una CDN de almacenamiento en caché) se utiliza para entregar rápidamente activos estáticos a los usuarios de Canvas. Estos extremos de CDN se distribuyen globalmente, lo que hace que la ruta de red para estas solicitudes sea lo más eficiente posible.

Los servidores de caché se supervisan constantemente. Cuando falla un servidor de caché, se aprovisiona y se implementa uno nuevo para que ocupe su lugar. Cuando falla un servidor de caché, los datos que se habrían almacenado en él simplemente se recuperan de la base de datos.

Los servidores de caché están completamente basados en memoria. El uso de la memoria se monitorea continuamente. Cuando la tasa de aciertos de caché cae por debajo de un umbral aceptable, se aprovisionan e implementan nuevos servidores de caché.

Servidores de base de datos

Los datos del curso y del usuario se almacenan en bases de datos relacionales. Las bases de datos están particionadas por la institución del cliente para fines de rendimiento y aislamiento de datos. Cada institución utiliza un par de bases de datos: una base de datos primaria y una base de datos secundaria en una zona de disponibilidad separada.

También hay un tercer servidor de respaldo en cada región y (si está disponible) en una zona de disponibilidad separada. Todos los cambios de la base de datos se transmiten en tiempo real entre sí y a una capa de datos duradera (S3). Esto significa que la información de la base de datos de Canvas LMS para los clientes de EE. UU. y Latinoamérica. Se almacena en tres ubicaciones separadas geográficamente. Los clientes canadienses se benefician de dos ubicaciones separadas geográficamente. Además, las copias de seguridad de la base de datos (una forma diferente de redundancia de datos para diferentes propósitos) se prueban semanalmente. Si la base de datos primaria falla, la secundaria se promoverá a primaria y se aprovisionará e implementará una nueva base de datos secundaria.

En caso de falla de la base de datos secundaria, se aprovisiona e implementa una nueva base de datos secundaria. En el improbable caso de que se produzcan fallos simultáneos de componentes o daños en los datos, se puede utilizar el servidor de respaldo en espera para crear un nuevo par de bases de datos.

Las bases de datos se controlan constantemente para determinar el uso de recursos y el tiempo de respuesta. Si alguna de las bases de datos se acerca a la carga máxima, los clientes individuales serán reubicados en clústeres con capacidad disponible.

Almacenamiento de archivos distribuidos

Los medios del curso, incluidos videos, archivos de imagen, grabaciones de audio, etc., y los archivos cargados por los estudiantes, como tareas, documentos y artefactos de aprendizaje, se almacenan fuera de la base de datos en un depósito de Amazon Simple Storage Service (S3) independiente y escalable diseñado para durabilidad superior al 99,99999999%. Todos los objetos dentro de los depósitos de S3 están cifrados y replicados entre sitios geográficamente separados y tienen



habilitado el control de versiones para que las versiones anteriores de un objeto se puedan restaurar con un esfuerzo mínimo.

Conclusión

Siguiendo las mejores prácticas, hemos creado una aplicación web nativa de la nube dinámica y altamente escalable que se ha convertido en el sistema de gestión de aprendizaje más confiable del mundo y es utilizado por instituciones de aprendizaje respetadas como las ocho escuelas de la Ivy League, la Universidad de Oxford, Universidad de Birmingham, KTH Royal Institute of Technology, London Business School, Universidad de Amsterdam y muchos más. Al tener mucho cuidado y diligencia en la creación de la mejor aplicación SaaS de su clase, nuestra arquitectura se destaca en la industria de la tecnología educativa por su robustez, endurecimiento, escalabilidad y confiabilidad.





INSTRUCTURE

© 2022 Instructure Inc. All rights reserved.