



Los aspectos principales en el

diseño de un entorno de AI/ML listo para la producción

Contenido

1 Obtenga valor empresarial de los datos

2 Diseñe un entorno de IA/ML listo para la producción

- 2.1 Contenedores
- 2.2 Organización de los contenedores
- 2.3 Gestión del ciclo de vida de las aplicaciones
- 2.4 Prácticas de MLOps
- 2.5 Plataforma de nube híbrida
- 2.6 Implementaciones en el extremo de la red

3 Comience con una base abierta y flexible

4 Casos de éxito

5 Empiece a utilizar IA/ML



Obtenga valor empresarial de los datos

Se espera que, para 2026, la cantidad de datos creados supere los 221 000 exabytes¹. En el mundo digital actual, los datos pueden brindarle una ventaja competitiva fundamental, pero recolectarlos solo es el primer paso; lo que realmente permite marcar la diferencia es el uso que hace de ellos.

La inteligencia artificial (IA), el aprendizaje automático (ML) y el aprendizaje profundo (DL) emplean los datos para ofrecer información empresarial relevante, automatizar las tareas y optimizar las funciones de los sistemas. Se trata de tecnologías que pueden transformar todos los aspectos de la empresa, desde los relacionados con los clientes y los empleados hasta el desarrollo y las operaciones. Cuando incorpora las tecnologías

de IA/ML a sus aplicaciones, puede alcanzar varios resultados empresariales que pueden medirse:

- ▶ Mayor grado de satisfacción de los clientes
- ▶ La posibilidad de ofrecer servicios digitales diferenciados
- ▶ Optimización de los servicios empresariales actuales
- ▶ Automatización de las operaciones empresariales
- ▶ Aumento en los ingresos
- ▶ Mejoras en el proceso de toma de decisiones
- ▶ Aumento en la eficiencia y reducción de los costos

Tecnologías fundamentales

En este ebook, se describen varias tecnologías que posibilitan el análisis procesable de los datos:

- ▶ La **inteligencia artificial (IA)** implica que las máquinas imiten el comportamiento humano para llevar a cabo tareas que generalmente requieren su intervención.
- ▶ El **aprendizaje automático (ML)** es un subconjunto de la IA que permite realizar tareas sin instrucciones explícitas por medio de algoritmos y modelos de estadística.
- ▶ El **aprendizaje profundo (DL)** es un subconjunto del ML que utiliza capas para extraer progresivamente funciones de nivel superior a partir del ingreso de datos sin procesar, por lo que resulta un mecanismo similar al del cerebro humano. La IA generativa, por ejemplo, puede utilizar modelos de DL entrenados para crear texto, imágenes y otros tipos de contenido de alta calidad.
- ▶ Las **operaciones de aprendizaje automático (MLOps)** incluyen todas las herramientas, las plataformas y los procesos que se necesitan para crear, entrenar, implementar, supervisar y mejorar permanentemente los modelos de IA/ML para utilizarlos en aplicaciones desarrolladas en la nube.

Casos prácticos de las tecnologías de IA/ML en diferentes sectores

Las tecnologías de IA/ML permiten generar resultados empresariales más rápido en diferentes sectores.



Diseñe un entorno de IA/ML listo para la producción

La implementación de la inteligencia artificial y el aprendizaje automático en la producción es un proceso constante que implica mucho más que simplemente crear modelos de estas tecnologías. Estos son los pasos principales en el ciclo de vida de IA/ML:

1. Establecer objetivos empresariales para su iniciativa de IA/ML y compartirlos con todas las partes interesadas
2. Reunir y preparar los datos necesarios para la iniciativa
3. Desarrollar modelos que se adapten a sus objetivos
4. Implementarlos en el proceso de desarrollo de las aplicaciones
5. Implementar las aplicaciones inteligentes e impulsadas por ML y comenzar a hacer inferencias
6. Supervisar y gestionar los modelos para garantizar la precisión con el tiempo

Ciclo de vida de IA/ML y MLOps



Desafíos de la implementación de las tecnologías de IA/ML

Las empresas enfrentan varios desafíos cuando diseñan un entorno de IA/ML:

- ▶ **Escasez de personal especializado.** Debido a que existe una cantidad limitada de especialistas en IA/ML, es difícil encontrar ingenieros y analistas de datos, ingenieros de ML, desarrolladores de software y otros miembros del personal capacitados en estos temas, y luego, poder conservarlos.
- ▶ **Falta de datos listos para utilizar.** No es suficiente con que las empresas recopilen grandes cantidades de datos; también deben encontrar los adecuados para cada iniciativa de IA/ML, prepararlos y protegerlos.
- ▶ **Tecnología y equipos dispares.** Es difícil lograr la colaboración entre los equipos y la implementación de IA/ML en la empresa cuando la infraestructura y las operaciones son lentas y manuales y están desconectadas.
- ▶ **Demoras en la disponibilidad de los recursos.** Las demoras a la hora de disponer de la infraestructura y las herramientas pueden dificultar el desarrollo del modelo, su integración y su implementación en las aplicaciones.

Aun así, si aplica un enfoque de desarrollo de aplicaciones en la nube en su ciclo de vida de IA/ML, podrá superar estos desafíos.

Con una arquitectura abierta y adaptable, puede adoptar IA/ML y MLOps con mayor eficiencia y cumplir sus objetivos empresariales.

Existen varias tecnologías y funciones que son esenciales para tener una arquitectura de IA/ML lista para la producción:

- ▶ **Las herramientas de MLOps e IA/ML** permiten que los analistas de datos, los ingenieros de ML y los desarrolladores de aplicaciones creen, implementen y gestionen las aplicaciones y los modelos de ML.
- ▶ **Las plataformas de nube** brindan a los ingenieros y analistas de datos, los ingenieros de ML y los desarrolladores de aplicaciones acceso rápido a los recursos que necesitan para agilizar su trabajo.
- ▶ **Los aceleradores de red, almacenamiento e informática** agilizan la preparación de los datos, el desarrollo de modelos y las tareas de inferencia.
- ▶ **Los extremos de la infraestructura** brindan recursos para todas las etapas de las operaciones de IA/ML, tanto en los entornos locales, virtuales y del extremo como en los de nube privada, pública e híbrida.
- ▶ **Las implementaciones en el extremo de la red** son opcionales y proporcionan grandes cantidades de datos obtenidos de dispositivos y sensores con los que se puede entrenar a los modelos y obtener conocimientos de manera inmediata.

En este ebook, se analizan los aspectos fundamentales para diseñar una infraestructura de IA/ML eficaz.

Contenedores

Los **contenedores** son unidades básicas de software que empaquetan las aplicaciones con todas sus dependencias.

Para que los analistas de datos, los ingenieros de ML y los desarrolladores de aplicaciones puedan aumentar al máximo la productividad, deben tener acceso a sus herramientas y recursos preferidos. A la vez, los equipos de operaciones de TI deben garantizar que los recursos estén actualizados, cumplan con los estándares del sector y se utilicen de forma segura. Los contenedores simplifican los procesos de diseño de las aplicaciones y permiten que se implementen en los distintos entornos sin necesidad de realizar cambios. También posibilitan la implementación uniforme de una amplia selección de herramientas de IA/ML en los entornos híbridos. Para favorecer la transparencia, los equipos pueden modificar y compartir imágenes en contenedores de forma constante a través de las funciones de control de versiones, las cuales permiten hacer un seguimiento de los cambios. Mientras tanto, el aislamiento de los procesos y el control de los recursos mejoran la protección ante las amenazas.

Recomendaciones sobre las soluciones de contenedores

Busque una plataforma de contenedores sólida y con alta disponibilidad que incluya funciones de seguridad integradas y reduzca la complejidad de la implementación, la gestión y el traslado de los contenedores en su entorno. Elija una plataforma open source que pueda integrarse con una gran variedad de tecnologías para adquirir más flexibilidad y acceder a más posibilidades.

Organización en contenedores

La organización en contenedores implica gestionar su creación, implementación y ciclo de vida en todo el entorno.

La adopción de los contenedores es solo el primer paso; luego, necesita una forma de implementarlos, gestionarlos y ajustarlos de manera eficiente. Puede utilizar herramientas de organización en contenedores para administrar su ciclo de vida de manera uniforme. Estas herramientas brindan acceso centralizado a los recursos informáticos, de almacenamiento y de redes en los diferentes entornos en las instalaciones, en el extremo de la red y en la nube. También ofrecen programación unificada de las cargas de trabajo, controles de la arquitectura multiempresa y la aplicación de cuotas.

Recomendaciones sobre la organización en contenedores

Seleccione una herramienta de organización basada en **Kubernetes** para obtener los beneficios de una tecnología de open source líder.

Gestión del ciclo de vida de las aplicaciones

La gestión del ciclo de vida de las aplicaciones supone implementar, ajustar y administrar aquellas que se ejecutan dentro de los contenedores.

Los entornos de IA/ML son complejos por naturaleza. Los elementos para la gestión del ciclo de vida que se integran a su herramienta de organización le permiten administrar las aplicaciones en contenedores en forma directa, incluidas las herramientas de desarrollo de IA/ML. Los equipos de operaciones de TI pueden automatizar las tareas frecuentes de la gestión del ciclo de vida, como la configuración, la implementación y la actualización, para adquirir eficiencia, velocidad y precisión. Los analistas de datos, los ingenieros de ML y los desarrolladores de aplicaciones pueden utilizar las herramientas y las aplicaciones contenidas en un catálogo de servicios aprobados previamente, sin que sea necesario involucrar a los equipos de TI. La automatización también permite que el personal deje de lado algunas tareas tediosas para enfocarse en actividades estratégicas más interesantes.

Recomendaciones sobre la gestión de aplicaciones

Elija herramientas para la gestión del ciclo de vida de las aplicaciones en contenedores que ofrezcan la integración a sus herramientas de IA/ML preferidas y la posibilidad de realizar automatizaciones de manera sencilla. Entre las opciones más conocidas se encuentran los **operadores de Kubernetes** y los **charts de Helm**.

Prácticas de MLOps

Las prácticas de MLOps combinan las herramientas, las plataformas y los procesos que hacen falta para poner en funcionamiento los procesos de IA/ML según sea necesario.

Las empresas necesitan desarrollar e implementar los modelos de IA/ML y las aplicaciones que los utilizan de manera rápida y eficiente. Para poder hacerlo con éxito, es esencial que los equipos trabajen juntos. Como en los enfoques **DevOps**, en MLOps se fomenta la colaboración entre los equipos de IA/ML, los de operaciones de TI y los desarrolladores de aplicaciones para agilizar la creación, el entrenamiento, la implementación y la gestión de los modelos de ML y las aplicaciones impulsadas por ML. La automatización, generalmente en forma de canales de **integración y distribución continuas (CI/CD)**, favorece la implementación de cambios rápidos, graduales y repetitivos para acelerar los ciclos de vida del desarrollo de los modelos y las aplicaciones.

Prácticas recomendadas de MLOps

El enfoque de MLOps no se trata solo de la tecnología; el personal y los procesos también son elementos fundamentales. Incorpore estas **prácticas** en todo el ciclo de vida de IA/ML. Para crear flujos de trabajo y canales de CI/CD, utilice la automatización en sus plataformas y herramientas, junto con tecnologías de open source como Argo, Kubeflow, Tekton y Jenkins.

Plataforma de nube híbrida

Las plataformas de nube híbrida ofrecen una base para el desarrollo, la implementación y la gestión de los modelos y las aplicaciones inteligentes en los entornos locales, de la nube y del extremo de la red.

Para desarrollar e implementar las aplicaciones y los modelos de IA/ML, es necesario contar con una infraestructura. Cuando adopta una plataforma de nube híbrida uniforme, puede desarrollar, probar, implementar y gestionar las aplicaciones y los modelos de la misma forma en todas las partes de su infraestructura. Esto le otorga la portabilidad, la flexibilidad y la capacidad de ajuste necesarias para implementar los entornos de IA/ML según lo requiera. También puede proporcionarle funciones de autoservicio para que agilice la distribución de los recursos sin perder el control de la TI. Por último, una plataforma de este tipo ofrece la base que necesita para integrar las tecnologías de terceros y de las comunidades open source, así como cualquier herramienta personalizada que utilice.

Recomendaciones sobre la plataforma de nube híbrida

Seleccione una plataforma centrada en la seguridad que admita la aceleración de los sistemas de hardware, un ecosistema diverso de herramientas para el desarrollo de aplicaciones e IA/ML y funciones integradas para la gestión de DevOps. Las plataformas open source pueden otorgarle más flexibilidad y oportunidades de integración.

Implementaciones en el extremo de la red

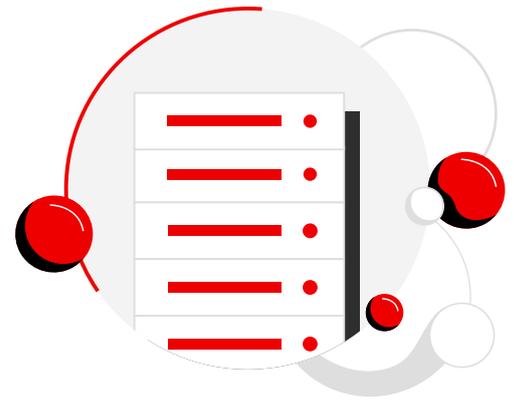
Las **implementaciones en el extremo de la red** son entornos descentralizados con dispositivos que recolectan datos y ejecutan funciones de manera local, pero fuera de los centros de datos principales.

El **edge computing** permite obtener información valiosa y experiencias en el momento y en la ubicación exacta en que se las necesita. Los sensores y los dispositivos suelen generar grandes cantidades de datos que se utilizan en los flujos de trabajo de IA/ML para entrenar los modelos y hacer inferencias sobre los tiempos de ejecución, pero puede resultar difícil y costoso transmitir estos datos a una nube central de manera inmediata. Por ejemplo, los algoritmos de reconocimiento de imágenes se ejecutan de forma más eficiente en las ubicaciones más cercanas al origen de los datos, lo cual elimina la necesidad de transferir grandes cantidades de datos para procesarlos.

Prácticas recomendadas para el extremo de la red

La capacidad de ajuste, la conectividad y la gestión de dispositivos son elementos fundamentales al realizar implementaciones en el extremo de la red. Elija soluciones que puedan gestionarse con las mismas herramientas y procesos que su infraestructura de nube y su centro de datos. Es esencial contar con plataformas que puedan lidiar con las interrupciones de la comunicación y la desconexión de los entornos. Por último, es importante utilizar soluciones que sean compatibles con diferentes entornos de hardware y dispositivos, ya que ofrecen mayor flexibilidad y más posibilidades de personalización.

Comience con una base abierta y flexible



Red Hat ofrece una cartera completa de productos tecnológicos, experiencia comprobada y asociaciones estratégicas para ayudarlo a alcanzar sus metas en el ámbito de la inteligencia artificial y el aprendizaje automático. No solo proporcionamos una base para el diseño de entornos de IA/ML listos para la producción, sino también los servicios y la capacitación que necesita para una adopción rápida.

Red Hat® OpenShift® es una plataforma de aplicaciones unificada y lista para el uso empresarial que permite generar innovaciones en la nube. Gracias a que ofrece recursos informáticos en función de las necesidades, posibilita la aceleración de los sistemas de hardware y las unidades de procesamiento de gráficos (GPU) y brinda uniformidad en los entornos locales, de nube pública y del extremo de la red, los equipos obtienen la velocidad y la flexibilidad que necesitan para tener éxito. Por ejemplo, puede crear una plataforma de autoservicio para las operaciones de MLOps en la que los desarrolladores, los analistas de datos y los ingenieros de datos diseñen modelos, los incorporen en las aplicaciones y lleven a cabo tareas de inferencia con rapidez. Con las funciones de colaboración, los equipos pueden crear resultados de modelado en contenedores y compartirlos con sus colegas y con los desarrolladores de manera uniforme.

Red Hat OpenShift AI es una cartera de productos que le permite entrenar, distribuir, supervisar y gestionar el ciclo de vida de las aplicaciones y los modelos de IA/ML. Esta oferta incluye **Red Hat OpenShift Data Science**, que ofrece a los analistas de datos y los desarrolladores una plataforma de IA/ML eficiente para recopilar información valiosa y diseñar aplicaciones inteligentes. Los equipos pueden pasar de la etapa de experimentación a la de producción en un entorno uniforme y colaborativo que incorpora importantes ofertas certificadas de partners como NVIDIA, Intel, Starburst, Anaconda, IBM y Pachyderm, entre muchos otros.

La cartera de productos de **Red Hat Application Services** le permite crear un entorno unificado para desarrollar, distribuir, integrar y automatizar las aplicaciones. Los servicios de integración de datos posibilitan el diseño de canales de datos efectivos, mientras que los de tiempos de ejecución permiten simplificar el desarrollo de las aplicaciones. Debido a que los servicios y las herramientas de automatización de los procesos pueden acceder a los modelos de ML y las aplicaciones inteligentes, es posible automatizar las decisiones empresariales.

Por último, obtiene una infraestructura con capacidad de ajuste y definida por el software gracias a los productos de plataforma de Red Hat, como **Red Hat Enterprise Linux®**, **Red Hat OpenStack® Platform** y **Red Hat OpenShift Platform Plus**.

Diseño en comunidad

Red Hat participa activamente en las comunidades open source **Kubeflow** y **Open Data Hub**. Open Data Hub es un proyecto de la comunidad que ofrece un plan para integrar las herramientas de IA/ML open source de uso más frecuente a un entorno de OpenShift. Estas herramientas comunes de aprendizaje automático y análisis de datos, como Ray, Ceph®, Apache Kafka, Kubeflow, TensorFlow y Jupyter Notebooks, se integran a la arquitectura de referencia.

Obtenga flexibilidad con un ecosistema de partners certificados

Gracias al **ecosistema de partners certificados de Red Hat**, puede integrar a esta arquitectura las herramientas más conocidas de IA/ML, análisis de datos, gestión, almacenamiento, seguridad y desarrollo. Trabajamos en estrecha colaboración con los partners para certificar el funcionamiento de sus sistemas de software en nuestras plataformas con el fin de mejorar la capacidad de gestión, la seguridad y el soporte. Muchos partners también ofrecen **operadores de Red Hat OpenShift** certificados para simplificar la gestión del ciclo de vida del software.

Elija las tecnologías y los productos que prefiera

Red Hat promueve el crecimiento de un ecosistema de partners certificados de IA/ML para que pueda incorporar las tecnologías y los productos conocidos a su entorno.

NVIDIA y Red Hat ofrecen soluciones para agilizar la distribución de las aplicaciones inteligentes que funcionan con IA a través de diferentes entornos. La solución **NVIDIA AI Enterprise con Red Hat OpenShift** es un conjunto de software de IA y análisis de datos completo, optimizado y desarrollado en la nube. Los sistemas NVIDIA DGX, Red Hat Enterprise Linux y Red Hat OpenShift le permiten gestionar fácilmente la TI de la infraestructura de IA. El operador **NVIDIA GPU Operator** automatiza la gestión de todos los elementos de software de NVIDIA necesarios para implementar las GPU.

Intel y Red Hat trabajan en conjunto para ofrecer una infraestructura definida por el software y plataformas estándar del sector que mejoran la flexibilidad y la agilidad del centro de datos. La distribución de Intel del **kit de herramientas OpenVINO** optimiza los modelos de DL y los convierte en motores de inferencia de alto rendimiento que pueden adaptarse a miles de nodos de Red Hat OpenShift automáticamente. **Intel AI Analytics Toolkit con la tecnología de oneAPI** ofrece un conjunto completo de herramientas interoperables de software de IA que permiten agilizar y ajustar los flujos de trabajo de ML.

Starburst y Red Hat le permiten obtener información valiosa de diferentes fuentes de datos distribuidas. **Starburst Enterprise** funciona en conjunto con Red Hat OpenShift para que pueda realizar análisis rápidos en varias plataformas. Esta combinación le otorga funciones de supervisión, flexibilidad, alta disponibilidad y automatización de nivel empresarial. Con esta solución, puede modernizar sus sistemas de datos; ejecutar las cargas de trabajo de extracción, transformación y carga (ETL) con eficiencia; realizar análisis de datos interactivos, y proporcionar información a las herramientas de inteligencia comercial.

SAS y Red Hat colaboran para crear tecnologías abiertas de nube híbrida y funciones de análisis que generen inteligencia de grado empresarial. **SAS Viya** en Red Hat OpenShift combina las funciones líderes de análisis, IA y ML con una plataforma de nube híbrida para que pueda diseñar sistemas solo una vez e implementarlos en todas partes. La gestión uniforme en todas las infraestructuras une a los equipos y promueve la colaboración entre ellos. Gracias a esta plataforma unificada, puede utilizar las interfaces, los lenguajes y las infraestructuras que prefiera para desarrollar e implementar los modelos.

Casos de éxito



Gracias al trabajo con Red Hat Consulting, **Banco Galicia** diseñó una solución de procesamiento del lenguaje natural (NLP) basada en la inteligencia artificial con la tecnología de Red Hat OpenShift, Red Hat Integration y Red Hat Single Sign-On.

El tiempo de incorporación de clientes nuevos se redujo de

20 días a minutos,

y se obtuvo una precisión del 90 % en el análisis de los datos.

Lea el [caso de éxito](#).



Nippon Telegraph and Telephone East Corporation (NTT East) utilizó Red Hat OpenShift para diseñar un servicio de análisis de datos de edge computing.

"[...] Red Hat OpenShift posibilitó el desarrollo y el funcionamiento estable de los servicios innovadores de IA de video a través de la colaboración con desarrolladores de IA".

Masashi Toyama

Gerente de Tecnología de Infraestructura de Servidores para el departamento de Ingeniería de Servidores en la Nube, División de Promoción Avanzada, sede comercial de la red NTT East

Lea el [caso de éxito](#).

Departamento de Asuntos de los Veteranos de los Estados Unidos

El **Team Guidehouse del departamento de Asuntos de los Veteranos de los Estados Unidos** implementó Red Hat OpenShift y Red Hat OpenShift Data Science con el fin de utilizar técnicas de aprendizaje automático en una solución prototipo para prevenir el suicidio entre los veteranos.

Ganador de la Fase 2

del desafío Mission Daybreak

Lea la [publicación del blog](#).



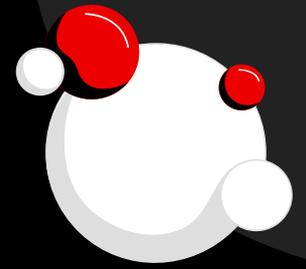
La **Universidad de Boston** utiliza Red Hat OpenShift Data Science como plataforma principal en el aula para los cursos de ciencias de la computación y sistemas de ingeniería informática.

"Esta iniciativa ofrece a mis estudiantes una experiencia completa de Linux que no oculta ningún aspecto, pero que resulta accesible y se integra bien a mis metodologías y materiales didácticos".

Jonathan Appavoo

Profesor adjunto de la Universidad de Boston

Lea la [publicación del blog](#).



Empiece a utilizar IA/ML

Las tecnologías de IA/ML y las MLOps están transformando el mundo empresarial casi por completo.

De la mano de Red Hat, puede diseñar un entorno de IA/ML listo para la producción que agilice el desarrollo y la distribución de las aplicaciones inteligentes para acompañar sus objetivos empresariales.



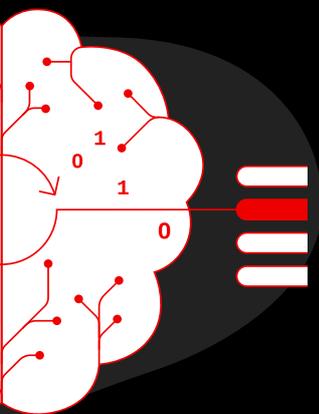
Red Hat OpenShift AI

Con Red Hat OpenShift, puede agilizar los flujos de trabajo de IA/ML y la distribución de las aplicaciones que usan la IA. Obtenga más información al respecto en red.ht/OpenShiftAI.



Red Hat OpenShift Data Science

Utilice Red Hat OpenShift Data Science para establecer las prácticas recomendadas de MLOps. Conozca la manera de hacerlo en red.ht/datascience.



Comience ahora mismo con Red Hat Consulting

Trabaje con los especialistas de Red Hat para impulsar sus proyectos de IA/ML. Red Hat ofrece servicios de capacitación y consultoría para que su empresa adopte esta tecnología más rápido.

- ▶ Conozca nuestros servicios de IA/ML en red.ht/aiml-consulting.
- ▶ Programe una sesión de exploración gratuita en redhat.com/consulting

Copyright © 2023 Red Hat, Inc. Red Hat, el logotipo de Red Hat, OpenShift y Ceph son marcas comerciales o marcas comerciales registradas de Red Hat, Inc. o sus filiales en Estados Unidos y en otros países. Linux® es la marca comercial registrada de Linus Torvalds en Estados Unidos y en otros países. La marca denominativa de OpenStack y el logotipo de Square O Design, ya sea juntos o separados, son marcas comerciales o marcas comerciales registradas de OpenStack Foundation en Estados Unidos y en otros países, y se utilizan con el permiso de OpenStack Foundation. Red Hat, Inc. no está afiliada a OpenStack Foundation ni a la comunidad de OpenStack, y tampoco goza de su respaldo ni de su patrocinio.

479615_0823_KVM



Red Hat