Prácticas recomendadas para optimizar la observabilidad de Amazon EKS

AWS Orientación prescriptiva

AWS Orientación prescriptiva: Prácticas recomendadas para optimizar la observabilidad de Amazon EKS

Table of Contents

Introducción 1
Objetivos 2
Registro 4
Tipos de registro 4
Registros del sistema 5
Registros de componentes de Kubernetes 6
Registros de ejecución del contenedor 7
Registros de aplicaciones 8
Prácticas recomendadas 9
Consideraciones importantes 10
Supervisión 13
Tipos de monitoreo 13
Monitoreo de infraestructuras 13
Supervisión de aplicaciones 14
Monitorización de la seguridad 16
Tools (Herramientas) 17
AWS servicios 17
Soluciones de código abierto o patentadas 18
Herramientas especializadas 20
Implementación de alta disponibilidad 20
Redundancia y escalabilidad arquitectónicas 20
Estrategia de almacenamiento de datos resiliente 21
Gestión de alertas redundante 21
Equilibrio de carga y descubrimiento de servicios 21
Consideraciones adicionales sobre alta disponibilidad 22
Prácticas recomendadas 23
Enfoque de implementación estratégica 23
Gestión eficaz de los datos 23
Configuración y administración de alertas 24
Optimización de recursos 25
Seguridad 16
Consideraciones avanzadas 25
Rastreo 27
Herramientas 29

Servicios de AWS 29
Soluciones de código abierto 30
Prácticas recomendadas 30
Alertas 33
Tools (Herramientas) 33
Prácticas recomendadas 34
Pasos siguientes 39
Recursos 40
AWS documentación 40
AWS publicaciones de blog 40
Otros recursos 40
Historial de documentos 41
Glosario 42
..... 42
A 43
B 46
C 48
D 52
E 56
F 58
G 60
H 62
I 63
L 66
M 67
O 71
P 74
Q 77
R 78
S 81
T 85
U 87
V 87
W 88
Z 89
..... xc

Prácticas recomendadas para optimizar la observabilidad de Amazon EKS

Ishwar Chaauthaiwale, Naveen Suthar y Pratap Kumar Nanda, Amazon Web Services ()AWS

Marzo de 2026 (historia del documento)

Amazon Elastic Kubernetes Service (Amazon EKS) requiere soluciones de observabilidad integrales para supervisar y solucionar problemas de las cargas de trabajo en contenedores de forma eficaz. Los sistemas distribuidos y los microservicios tienen arquitecturas complejas en los entornos de Amazon EKS, por lo que la implementación de prácticas de observabilidad adecuadas es crucial para mantener la fiabilidad de las operaciones. La observabilidad efectiva en los entornos de Amazon EKS permite a los equipos obtener información detallada sobre el rendimiento de las aplicaciones, solucionar problemas de manera eficiente y mantener un estado óptimo del clúster.

El desafío consiste en navegar por el vasto ecosistema de herramientas y técnicas disponibles para la observabilidad de Amazon EKS y, al mismo tiempo, seguir las mejores prácticas que se ajusten a los objetivos de la organización y a los estándares del sector. Las estrategias de observabilidad eficaces deben equilibrar la recopilación integral de datos con las consideraciones de rendimiento, rentabilidad y escalabilidad.

Esta guía está diseñada para ayudar a las organizaciones a optimizar la observabilidad de Amazon EKS en las siguientes áreas:

- Establecer mecanismos de registro eficientes
- Implementación de soluciones de monitoreo sólidas
- Uso del rastreo distribuido para arquitecturas complejas
- Implementación de estrategias de alerta y respuesta a incidentes

Al adoptar estas prácticas recomendadas, su organización puede mejorar su capacidad de obtener información detallada sobre su entorno Amazon EKS, lo que mejora la confiabilidad, el rendimiento y la eficiencia operativa. Este enfoque simplificado de la observabilidad contribuye a la solución de problemas y al mantenimiento, y apoya la toma de decisiones basada en los datos para la mejora continua de las aplicaciones y la infraestructura basadas en Kubernetes. (Para obtener información detallada sobre Amazon EKS, consulte la documentación del servicio).

Esta guía profundiza en cada aspecto de la observabilidad de Amazon EKS y explora las herramientas y estrategias que puede personalizar para satisfacer las necesidades específicas de sus implementaciones de Amazon EKS, desde aplicaciones a pequeña escala hasta arquitecturas de microservicios grandes y complejas.

En esta guía:

- Inicio de sesión en Amazon EKS
- Supervisión en Amazon EKS
- Rastreo en Amazon EKS
- Alertas en Amazon EKS
- Pasos siguientes
- Recursos

Objetivos

Esta guía puede ayudarle a usted y a su organización a alcanzar los siguientes objetivos empresariales:

- Visibilidad operativa mejorada: obtenga una visión integral de sus clústeres y aplicaciones de Amazon EKS mediante prácticas de observabilidad eficaces.

  Este objetivo hace hincapié en la importancia de mantener una visibilidad completa en todo el entorno de Amazon EKS. Herramientas como AWS X-RayAmazon CloudWatch Container Insights y AWS Distro le OpenTelemetry ayudan a comprender el comportamiento del sistema, identificar problemas rápidamente y mantener un rendimiento óptimo.
- Mejora de la eficiencia de la solución de problemas: reduzca el tiempo medio de detección (MTTD) y el tiempo medio de resolución (MTTR) mediante estrategias eficaces de seguimiento y supervisión.

  Este objetivo se centra en implementar prácticas de observabilidad que permitan identificar y resolver problemas rápidamente. Técnicas como el rastreo distribuido, el registro efectivo y la recopilación integral de métricas son fundamentales para lograr este objetivo.
- Gestión proactiva del rendimiento: permita la detección temprana de posibles problemas antes de que afecten a los usuarios finales.

La supervisión proactiva es fundamental para mantener una alta disponibilidad y rendimiento del servicio. Este objetivo aborda la importancia de implementar alertas, análisis de tendencias y monitoreo predictivo adecuados para evitar interrupciones en el servicio.

- Observabilidad rentable: optimice los costos de observabilidad y, al mismo tiempo, mantenga una visibilidad integral del sistema.

  La optimización de costos abarca la implementación de estrategias de muestreo eficientes, políticas de retención de datos adecuadas y enfoques de instrumentación óptimos. El objetivo es equilibrar las necesidades de observabilidad con las consideraciones de costo y, al mismo tiempo, garantizar una supervisión eficaz del sistema.
- Arquitectura de monitoreo escalable: asegúrese de que sus soluciones de observabilidad se escalen sin problemas con su entorno Amazon EKS.

  Este objetivo se centra en implementar soluciones de monitoreo que puedan crecer con su aplicación. Ya sea que ejecute un solo clúster o una implementación de varios clústeres y varias regiones, su estrategia de observabilidad debe ampliarse en consecuencia

Inicio de sesión en Amazon EKS

El registro es un aspecto fundamental de la administración y el mantenimiento de las aplicaciones que se ejecutan en Amazon EKS. Las prácticas de registro eficaces en los entornos de Amazon EKS ayudan a los desarrolladores, los equipos de operaciones y los administradores de sistemas a obtener información valiosa sobre el comportamiento, el rendimiento y el estado de sus aplicaciones en contenedores y su infraestructura subyacente.

La implementación de una estrategia de registro sólida en Amazon EKS es esencial por varios motivos:

- Solución de problemas: los registros ayudan a identificar y diagnosticar los problemas rápidamente, lo que reduce el tiempo de inactividad y mejora la confiabilidad general del sistema.
- Cumplimiento: muchos sectores requieren un registro exhaustivo para fines de auditoría y regulación.
- Seguridad: el análisis de los registros puede ayudarle a detectar e investigar posibles amenazas o infracciones de seguridad.
- Optimización del rendimiento: los registros proporcionan información sobre el rendimiento de las aplicaciones y los sistemas, para que pueda identificar los cuellos de botella y optimizar la utilización de los recursos.
- Supervisión y alertas: los datos de registro se pueden utilizar para configurar sistemas de supervisión y activar alertas en caso de eventos o condiciones específicos.

En esta sección:

- Tipos de inicio de sesión en Amazon EKS
- Prácticas recomendadas para iniciar sesión en Amazon EKS
- Consideraciones importantes para iniciar sesión en Amazon EKS

Tipos de inicio de sesión en Amazon EKS

En Amazon EKS, el registro implica la captura, el almacenamiento y el análisis de varios tipos de datos de registro que generan los distintos componentes del clúster de Kubernetes, entre los que se incluyen:

- Registros del sistema: información sobre las instancias o nodos subyacentes de Amazon Elastic Compute Cloud (Amazon EC2) AWS Fargate
- Registros de componentes de Kubernetes: datos de los componentes principales de Kubernetes, como el servidor de API, el programador y el administrador de controladores
- Registros del tiempo de ejecución del contenedor: información del tiempo de ejecución del contenedor, como Docker o containerd
- Registros de aplicaciones: resultados de aplicaciones en contenedores

Para gestionar los registros en su entorno de Amazon EKS de forma eficaz, suele emplear una combinación de Servicios de AWS herramientas de terceros y prácticas recomendadas. Esto podría incluir el uso de Amazon CloudWatch, Fluent Bit, Elasticsearch, Kibana y otras herramientas de registro y análisis para recopilar, almacenar y visualizar datos de registro.

En las siguientes secciones se analizan varios aspectos del registro en Amazon EKS, incluidas las prácticas recomendadas, las herramientas y las técnicas para implementar una estrategia de registro integral en sus clústeres de Kubernetes. AWS

Registros del sistema

El registro de instancias EC2 subyacentes o nodos de Fargate en Amazon EKS implica diferentes enfoques según el tipo de nodo.

Para implementar el registro de instancias EC2 en Amazon EKS, puede usar las siguientes herramientas:

- CloudWatch agente: instale y configure el CloudWatch agente en sus instancias EC2. Configúrelo para recopilar registros del sistema, como `/var/log/messages` y`/var/log/secure`. Puede utilizar scripts de datos de usuario o herramientas de administración de la configuración para automatizar este proceso.
- Fluent Bit: Implemente Fluent Bit DaemonSet para recopilar registros de todos los nodos. Configúrelo para reenviar los CloudWatch registros a Logs u otros sistemas de registro centralizados.
- Container Insights: habilite Container Insights en su clúster de EKS para recopilar automáticamente métricas y registros de las instancias de EC2.
- Secuencias de comandos personalizadas: desarrolle secuencias de comandos personalizadas para recopilar registros específicos y enviarlos al destino de registro que prefiera.

- Agente SSM: utilice el AWS Systems Manager agente (agente SSM) para recopilar y reenviar los registros a CloudWatch Logs.

Para implementar el registro para los nodos de Fargate en Amazon EKS, utilice estas herramientas:

- Registro de Fargate: Fargate recopila `stdout` y `stderr` registra automáticamente en sus contenedores. Configure su perfil de Fargate para enviar estos registros a CloudWatch Logs.
- Fluent Bit para Fargate: AWS proporciona una imagen de Fluent Bit específica para el registro de Fargate. Colócalo como un contenedor con sidecar en tus cápsulas Fargate para recolectar y reenviar troncos.
- Container Insights para Fargate: habilite Container Insights para recopilar métricas y registros de los nodos de Fargate.

Registros de componentes de Kubernetes

La recopilación de registros de los componentes de Kubernetes, como el servidor de API, el programador y el administrador de controladores de Amazon EKS, requiere un enfoque ligeramente diferente al del registro de aplicaciones. Estos componentes se ejecutan como parte del plano de control de Amazon EKS, que administra AWS. A continuación, le indicamos cómo puede recopilar estos registros y acceder a ellos:

- Habilite el registro del plano de control: puede habilitar el registro del plano de control para su clúster de EKS mediante las herramientas Consola de administración de AWS, AWS Command Line Interface (AWS CLI) o de infraestructura como código (IaC), como AWS CloudFormationTerraform. Cuando habilitas el registro en el plano de control, los registros se envían a Amazon CloudWatch Logs. Puede verlos en la CloudWatch consola, en el grupo de `/aws/eks/<cluster-name>/cluster` registros. Dentro de este grupo de registros, cada componente del plano de control tiene su propio flujo de registro, de la siguiente manera:

Nombre de flujo	Description (Descripción)
servidor kube-api	Registros del servidor API de Kubernetes
kube-scheduler	Registros de decisiones del programador
kube-controller-manager	Registros del administrador del controlador

Nombre de flujo	Description (Descripción)
autenticador	Registros del autenticador de IAM
audit	Registros de auditoría de Kubernetes (deben estar habilitados de forma explícita)

Para ver los registros de un componente específico, navegue hasta el grupo de registros del clúster y filtre por el nombre del flujo de registro de destino.

- Use CloudWatch Logs Insights: puede usar CloudWatch Logs Insights para realizar consultas complejas en sus registros.
- Exportación de registros a Amazon S3: para almacenarlos a largo plazo o analizarlos más a fondo, puede exportar los registros a Amazon Simple Storage Service (Amazon S3).
- Utilice herramientas de terceros: puede utilizar herramientas como Fluent Bit para recopilar estos registros y reenviarlos a otros sistemas de registro, como Elasticsearch o Splunk.
- Uso AWS CloudTrail: El AWS CloudTrailservicio puede proporcionar información adicional sobre las llamadas a la API realizadas a su clúster de EKS.

Registros de ejecución del contenedor

El registro de los registros del tiempo de ejecución del contenedor en Amazon EKS implica capturar y administrar los registros del tiempo de ejecución del contenedor, que suele ser el `containerd` caso de Amazon EKS. A continuación, le explicamos cómo puede abordar el registro de los registros de tiempo de ejecución de los contenedores en Amazon EKS:

- Acceda directamente a los registros de los nodos de Amazon EC2. En el caso de los nodos EC2 autogestionados, puede acceder directamente a los registros de tiempo de ejecución del contenedor en el host desde las siguientes ubicaciones:
  - `containerd`registros: `/var/log/containers/`
  - Registros de Docker (si está utilizando el tiempo de ejecución de Docker): `/var/log/docker.log`
- Usa un DaemonSet para la recopilación de registros.
- Implemente un agente de recopilación de registros (como Fluent Bit) DaemonSet para recopilar registros de todos los nodos.

- Configure el CloudWatch agente para recopilar los registros de tiempo de ejecución del contenedor.
- Habilite Container Insights para recopilar métricas y registros del tiempo de ejecución de los contenedores.
- Usa Fargate. En el caso de los nodos de Fargate, los registros de tiempo de ejecución de los contenedores se recopilan automáticamente y se puede acceder a ellos a través CloudWatch de los registros.
- Implemente soluciones de registro personalizadas mediante herramientas como Fluent Bit o Logstash. Configure CloudWatchalarmas o utilice herramientas como Prometheus para monitorear patrones o problemas específicos en los registros de tiempo de ejecución de los contenedores. Considere la posibilidad de utilizar soluciones de registro de terceros que se integren bien con Kubernetes y Amazon EKS, como Datadog, Splunk o Elastic Stack (ELK Stack). Utilice herramientas de agregación de registros para recopilar registros de varias fuentes y reenviarlos a un sistema de registro centralizado.

Registros de aplicaciones

Los registros de aplicaciones en Amazon EKS son una parte fundamental del mantenimiento y la solución de problemas de las aplicaciones. Para implementar el registro de aplicaciones en Amazon EKS, puede elegir entre estas opciones:

- Escribir los registros en`stdout`/`stderr`: La forma más sencilla y nativa de Kubernetes de gestionar los registros de las aplicaciones consiste en escribirlos en y. `stdout` `stderr` Kubernetes captura automáticamente estas transmisiones.
- Implemente la agregación de registros: utilice un agregador de registros como Fluent Bit para recopilar los registros de todos sus pods.
- Configure el enrutamiento de registros: configure su agregador de registros para enrutar los registros al destino que desee (como CloudWatch Logs o Elasticsearch).
- Utilice CloudWatch Container Insights: habilite Container Insights para un registro y una supervisión completos.

Prácticas recomendadas para iniciar sesión en Amazon EKS

Las siguientes prácticas recomendadas ayudan a crear un sistema de registro sólido, escalable y eficiente para su entorno Amazon EKS y proporcionan una mejor solución de problemas, supervisión y administración general de sus clústeres de Kubernetes.

- Centralice la recopilación de registros: utilice una solución de registro centralizada, como CloudWatch Logs, Elasticsearch o un servicio de terceros, para agregar los registros de todos los componentes. Esto proporciona un punto de acceso único para el análisis de registros y simplifica la administración.
- Implemente un registro estructurado: utilice formatos de registro estructurados, como JSON, para que los registros se puedan analizar y buscar con mayor facilidad. Incluya los metadatos relevantes, como las marcas de tiempo, los niveles de registro y los identificadores de origen.
- Utilice los niveles de registro de forma adecuada: Implemente los niveles de registro adecuados (como `DEBUG` `INFO`WARN, yERROR) en sus aplicaciones. Configure los entornos de producción para que registren en los niveles adecuados a fin de evitar un registro excesivo.
- Habilite el registro de contenedores: configure sus contenedores para que se registren en `stdout` y`stderr`. Esto permite a Kubernetes capturar y reenviar estos registros a la solución de registro que elija.
- Habilite el registro de aplicaciones: configure las aplicaciones para que escriban registros en los archivos de registro `stdout` y `stderr` en lugar de escribirlos en ellos. Esto sigue la metodología de aplicaciones de 12 factores y se alinea con las mejores prácticas nativas de la nube.
- Usa Kubernetes DaemonSets para la recopilación de registros: implementa agentes de recopilación de registros (como Fluent Bit) DaemonSets para asegurarte de que se ejecuten en todos los nodos del clúster.
- Implemente políticas de retención: defina y aplique políticas de retención de registros para cumplir con las regulaciones y administrar los costos de almacenamiento.
- Proteja los datos de registro: cifre los registros en tránsito y en reposo. Implemente controles de acceso para restringir quién puede ver y administrar los registros.
- Supervise la ingesta de registros: configure alertas en caso de errores o retrasos en la ingesta de registros para garantizar un registro continuo.
- Utilice anotaciones y etiquetas de Kubernetes: utilice las anotaciones y etiquetas de Kubernetes para añadir metadatos a sus registros, a fin de mejorar la capacidad de búsqueda y el filtrado.

- Implemente el rastreo distribuido: utilice herramientas de rastreo distribuido, como Jaeger, para correlacionar los registros entre los microservicios. AWS X-Ray
- Optimice el volumen de los registros: Sea selectivo con lo que registra para evitar costes innecesarios y problemas de rendimiento. Utilice el muestreo para registros de gran volumen y bajo valor.
- Implemente la agregación de registros: utilice herramientas como Logstash para agregar registros de varias fuentes antes de enviarlos a su sistema de registro central.
- Úselo Servicios de AWS siempre que sea posible: servicios como CloudWatch Logs y Container Insights proporcionan una integración perfecta con otros servicios. Servicios de AWS
- Implemente el análisis y la visualización de CloudWatch registros: utilice herramientas como Logs Insights, Elasticsearch con Kibana o soluciones de terceros para el análisis y la visualización de registros.
- Implemente un análisis de registros automatizado: utilice herramientas de aprendizaje automático y basadas en inteligencia artificial para detectar automáticamente anomalías y patrones en sus registros.
- Documente su estrategia de registro: mantenga una documentación clara de su arquitectura, prácticas y herramientas de registro para su equipo.

Consideraciones importantes para iniciar sesión en Amazon EKS

En esta sección se analizan las consideraciones importantes que se deben tener en cuenta al implementar el registro en Amazon EKS.

- Impacto en el rendimiento: el registro excesivo puede afectar al rendimiento de la aplicación. Tenga en cuenta el volumen y la frecuencia de los registros que se generan.
- Administración de costos: el almacenamiento y el procesamiento de registros pueden generar costos significativos, especialmente a gran escala. Implemente políticas de retención de registros y considere la posibilidad de utilizar la agregación de registros para reducir los costos.
- Seguridad y conformidad: asegúrese de que los registros no contengan información confidencial, como contraseñas o datos personales. Implemente el cifrado para los registros en tránsito y en reposo. Tenga en cuenta los requisitos de cumplimiento, como el Reglamento General de Protección de Datos (GDPR) o la Ley de Portabilidad y Responsabilidad de los Seguros de Salud (HIPAA) cuando maneje los registros.

- Escalabilidad: asegúrese de que su solución de registro pueda escalarse con el tamaño del clúster y el volumen de registro. Considere la posibilidad de utilizar el almacenamiento en búfer y el almacenamiento por lotes para la transmisión de registros.
- Retención de registros: defina e implemente los períodos de retención de registros adecuados. Equilibre los requisitos de cumplimiento con los costos de almacenamiento.
- Control de acceso: Implemente las funciones y políticas AWS Identity and Access Management (de IAM) adecuadas para el acceso a los registros. Siga el principio de privilegios mínimos para la administración de registros.
- Coherencia de registros: utilice formatos de registro coherentes en diferentes aplicaciones y servicios. Utilice el registro estructurado para facilitar el análisis y el análisis.
- Sincronización horaria: sincronice la hora en todos los nodos para obtener marcas de tiempo consistentes en los registros.
- Asignación de recursos: asigne los recursos adecuados (como la CPU y la memoria) a los agentes de registro. Supervise el uso de recursos de los componentes de registro.
- Consideraciones sobre Fargate: Fargate tiene mecanismos de registro específicos que difieren de los nodos basados en EC2. Comprenda las limitaciones y las capacidades del registro de Fargate.
- Clústeres con varios inquilinos: en entornos con varios inquilinos, asegúrese de que los registros estén debidamente aislados entre los inquilinos.
- Análisis y análisis de registros: tenga en cuenta las herramientas y habilidades necesarias para un análisis de registros eficaz. Implemente el análisis de registros para la extracción de datos estructurados.
- Supervisión del sistema de registro: configure la supervisión de la propia infraestructura de registro. Genere alertas para registrar las fallas o los atrasos del sistema.
- Impacto en la red: tenga en cuenta el ancho de banda de red que utiliza la transmisión de registros. Considere la posibilidad de utilizar la compresión para los datos de registro.
- Eventos de Kubernetes: no pases por alto los eventos de Kubernetes como fuente de información importante.
- Registro en el plano de control: comprenda las implicaciones y los costes de habilitar el registro en el plano de control.
- Capacidades de depuración: asegúrese de que su solución de registro permita depurar y solucionar problemas fácilmente.
- Integración con las herramientas existentes: considere cómo se integra su solución de registro Amazon EKS con las herramientas de supervisión y alerta existentes.

- Pruebas: pruebe periódicamente la configuración de registro, especialmente después de actualizar el clúster.
- Documentación: mantenga una documentación clara de su arquitectura y prácticas de registro.
- Latencia de agregación de registros: tenga en cuenta cualquier latencia en la agregación de registros y cómo podría afectar a la supervisión en tiempo real.

Supervisión en Amazon EKS

La supervisión en Amazon EKS proporciona una visibilidad fundamental del estado, el rendimiento y la seguridad de las cargas de trabajo de Kubernetes. Sin una supervisión adecuada, corre el riesgo de sufrir interrupciones en el servicio, brechas de seguridad y un uso ineficiente de los recursos, lo que puede afectar a las operaciones comerciales y aumentar los costos. Una supervisión eficaz le permite identificar y resolver problemas de forma proactiva, optimizar el uso de los recursos y mantener los requisitos de conformidad en todas sus aplicaciones contenerizadas. Al implementar soluciones de monitoreo integrales, puede garantizar una alta disponibilidad, detectar anomalías de manera temprana y tomar decisiones basadas en datos para escalar y mejorar su infraestructura de Amazon EKS.

En esta sección se analizan los diversos aspectos de la supervisión de Amazon EKS, incluidos los diferentes tipos de supervisión, las herramientas disponibles y las prácticas recomendadas para ayudarle a crear una estrategia de supervisión sólida para su entorno de Kubernetes.

En esta sección:

- Tipos de supervisión en Amazon EKS
- Herramientas de supervisión para Amazon EKS
- Implementación de alta disponibilidad para las soluciones de monitoreo Amazon EKS
- Prácticas recomendadas para la supervisión en Amazon EKS
- Consideraciones de supervisión avanzada en Amazon EKS

Tipos de supervisión en Amazon EKS

La observabilidad efectiva en Amazon EKS implica actividades de supervisión de la infraestructura, las aplicaciones y la seguridad.

Monitoreo de infraestructuras

La supervisión de la infraestructura es un componente fundamental de la observabilidad de Amazon EKS que proporciona información detallada sobre el estado y el rendimiento de los elementos fundamentales de su clúster de Kubernetes. En esencia, implica hacer un seguimiento de los signos vitales tanto de los componentes del plano de control como de los nodos de trabajo, y asegurarse de que la plataforma subyacente se mantenga estable y eficiente.

- La supervisión del plano de control es crucial porque supervisa componentes clave como el servidor API, la base de datos etcd y el programador. Al monitorear la latencia del servidor API, puede identificar rápidamente los cuellos de botella en el rendimiento que podrían afectar a la implementación de aplicaciones o a las operaciones de escalado. La supervisión del rendimiento de Etcd valida que la base de datos de estado del clúster funciona de manera eficiente y evita problemas de coherencia de los datos que podrían afectar a todo el clúster.
- La supervisión a nivel de nodo es igualmente importante porque se centra en los recursos informáticos que ejecutan las cargas de trabajo en contenedores. Esto incluye el seguimiento de la utilización de la CPU, el consumo de memoria, las E/S del disco y el rendimiento de la red en todos los nodos de trabajo. Comprender estas métricas ayuda a evitar el agotamiento de los recursos, a optimizar las decisiones de escalado de los nodos y a garantizar una planificación de la capacidad adecuada.
- La supervisión de la red desempeña un papel fundamental a la hora de mantener una comunicación fiable entre los módulos, los servicios y los recursos externos. Al monitorear el rendimiento, la latencia y los estados de conexión de la red, puede identificar los problemas de conectividad de manera temprana y garantizar una comunicación fluida entre las aplicaciones. La supervisión del almacenamiento complementa la supervisión de la red mediante el seguimiento del rendimiento del volumen, la utilización de la capacidad y I/O los patrones, a fin de evitar los cuellos de botella relacionados con los datos.

La supervisión de la infraestructura sirve como un sistema de alerta temprana de posibles problemas, permite un mantenimiento proactivo y garantiza una asignación óptima de los recursos. Sin una supervisión sólida de la infraestructura, corre el riesgo de sufrir tiempos de inactividad inesperados, reducir el rendimiento y hacer un uso ineficiente de los recursos, lo que puede repercutir considerablemente en las operaciones y los costes empresariales.

Supervisión de aplicaciones

La supervisión de las aplicaciones es esencial para mantener las aplicaciones en contenedores en buen estado, eficaces y fiables en su entorno Amazon EKS. Este nivel de monitoreo se centra en las cargas de trabajo reales que se ejecutan dentro de su clúster y proporciona información fundamental sobre el comportamiento, el rendimiento y la interacción de sus aplicaciones con otros servicios.

El monitoreo de aplicaciones incluye el monitoreo a nivel de contenedor, monitoreo a nivel de servicio y rastreo distribuido.

- A nivel de contenedor, el monitoreo de aplicaciones rastrea métricas cruciales como el estado del contenedor, el número de reinicios y los patrones de consumo de recursos. Estas métricas le ayudan a identificar los contenedores problemáticos que podrían estar consumiendo recursos excesivos o que se reinicien con frecuencia, lo que podría indicar problemas subyacentes, como pérdidas de memoria o problemas de configuración. Al monitorear los eventos del ciclo de vida de los contenedores, puede garantizar el comportamiento correcto de las aplicaciones y solucionar rápidamente los problemas de implementación.
- La supervisión a nivel de servicio proporciona visibilidad de las métricas de rendimiento y confiabilidad de las aplicaciones, como los tiempos de respuesta, las tasas de error y el rendimiento de las solicitudes. Estas métricas son fundamentales para mantener los objetivos de nivel de servicio (SLOs) y garantizar una experiencia positiva para el usuario final. Puede realizar un seguimiento de la latencia en los diferentes puntos finales del servicio, identificar los cuellos de botella en el rendimiento y supervisar los patrones de error para mantener la fiabilidad de las aplicaciones.
- El rastreo distribuido es otro aspecto fundamental de la supervisión de aplicaciones, especialmente en las arquitecturas de microservicios. Al implementar el rastreo, puede hacer un seguimiento de las solicitudes a medida que pasan por los distintos servicios, comprender las dependencias e identificar los cuellos de botella en el rendimiento. Esta end-to-end visibilidad le ayuda a optimizar las interacciones de los servicios y a solucionar problemas complejos que afectan a varios componentes.

Las métricas de las aplicaciones personalizadas desempeñan un papel crucial a la hora de proporcionar información específica de la empresa. Estas pueden incluir métricas como las tasas de procesamiento de pedidos, las frecuencias de inicio de sesión de los usuarios o las tasas de éxito de las transacciones. Puede correlacionar estas métricas personalizadas con las métricas de infraestructura y contenedores para comprender mejor cómo el rendimiento de la infraestructura afecta a las operaciones empresariales y tomar decisiones basadas en datos para el escalado y la optimización.

La importancia de la supervisión de las aplicaciones reside en su capacidad de proporcionar una visión integral del estado y el rendimiento de las aplicaciones. Esta supervisión le permite mantener una alta calidad de servicio, resolver rápidamente los problemas y optimizar continuamente sus aplicaciones para cumplir los objetivos empresariales.

Monitorización de la seguridad

La supervisión de la seguridad en Amazon EKS es una actividad fundamental que ayuda a las organizaciones a mantener la integridad, la confidencialidad y el cumplimiento de sus entornos de Kubernetes. Este enfoque de seguridad integral combina la vigilancia continua, la detección de amenazas y la supervisión del cumplimiento para proteger las cargas de trabajo en contenedores de los posibles riesgos de seguridad y del acceso no autorizado. Incluye la supervisión de la autenticación y la autorización, la supervisión de la seguridad de la red y la supervisión de la configuración y el cumplimiento.

- La supervisión de la autenticación y la autorización constituye la primera línea de defensa, ya que rastrea todos los intentos de acceso al clúster. Esto incluye la supervisión de las solicitudes de los servidores de la API, el seguimiento de los intentos de inicio de sesión correctos y fallidos y la auditoría de los cambios en el control de acceso basado en roles (RBAC). Al mantener registros de auditoría detallados sobre quién accedió a qué recursos y cuándo, puede detectar rápidamente posibles brechas de seguridad, intentos de acceso no autorizado o actividades de escalamiento de privilegios. Esto es especialmente importante en los entornos con varios inquilinos, donde es esencial mantener controles de acceso estrictos.
- La supervisión de la seguridad de la red se centra en detectar y prevenir la comunicación no autorizada entre los módulos y los servicios. Al monitorear las infracciones de las políticas de red y los patrones de tráfico inusuales, puede identificar posibles amenazas a la seguridad, como los intentos de escape de los contenedores o los movimientos laterales dentro del clúster. Esto incluye el seguimiento tanto de la comunicación interna del clúster como de los patrones de tráfico externo para garantizar que los contenedores se comuniquen solo con los puntos finales autorizados y sigan las políticas de seguridad definidas.
- La supervisión de la configuración y el cumplimiento es esencial para mantener las bases de seguridad y cumplir los requisitos normativos. Implica escanear continuamente las imágenes de los contenedores para detectar vulnerabilidades, supervisar la seguridad en tiempo de ejecución y realizar un seguimiento de los cambios de configuración que puedan afectar a la postura de seguridad. Las auditorías de conformidad periódicas garantizan el cumplimiento de los estándares del sector y las políticas de seguridad de la organización, y la detección de desviaciones en la configuración ayuda a evitar cambios no autorizados que puedan suponer riesgos para la seguridad.

La supervisión de la seguridad en Amazon EKS proporciona la visibilidad y el control necesarios para ayudar a protegerse contra las amenazas de seguridad modernas y, al mismo tiempo, garantizar el

cumplimiento de los requisitos reglamentarios. Al implementar una supervisión de seguridad integral, su organización puede mantener una postura de seguridad sólida, responder rápidamente a los incidentes de seguridad y demostrar el cumplimiento de diversas normas reglamentarias.

Herramientas de supervisión para Amazon EKS

En esta sección se analizan tres categorías de herramientas de supervisión de Amazon EKS: servicios de AWS supervisión, soluciones de código abierto o patentadas y herramientas especializadas.

AWS servicios

- Amazon CloudWatch: servicio integral de monitoreo y registro

  CloudWatch constituye la columna vertebral de las soluciones de AWS monitoreo y proporciona amplias capacidades para los entornos de Amazon EKS. Ofrece Container Insights para obtener métricas pormenorizadas de contenedores y clústeres, de forma que pueda supervisar el rendimiento, la utilización de los recursos y el estado de las aplicaciones. El servicio destaca en la agregación y el análisis de registros, y admite el registro centralizado en contenedores y nodos. CloudWatch se integra de forma natural con Servicios de AWS. Proporciona una configuración de alarmas automatizada y admite métricas y paneles personalizados, lo que la convierte en una herramienta esencial para la supervisión de Amazon EKS.
- AWS X-Ray: Plataforma avanzada de rastreo distribuido

  X-Ray mejora la observabilidad al proporcionar sofisticadas capacidades de rastreo distribuido. Su visualización del mapa de servicios ofrece información clara sobre la arquitectura y las dependencias de las aplicaciones, y el seguimiento detallado de las solicitudes ayuda a identificar los cuellos de botella en el rendimiento de los servicios. X-Ray puede rastrear las solicitudes a través de arquitecturas de microservicios complejas, lo que lo hace inestimable para la resolución de problemas y la optimización, especialmente en sistemas distribuidos que abarcan varios. Servicios de AWS
- AWS Distribución para OpenTelemetry: marco de observabilidad unificado

  Distro for OpenTelemetry proporciona capacidades de recopilación de datos unificadas con soporte multiplataforma, lo que la hace ideal para entornos híbridos. Este servicio se integra con otros Servicios de AWS, admite instrumentación personalizada y ofrece flexibilidad a la hora de

implementar soluciones de monitoreo integrales, al tiempo que mantiene la compatibilidad con los estándares de la industria.

- Grafana gestionada por Amazon: visualización de nivel empresarial

  Amazon Managed Grafana proporciona un servicio totalmente gestionado para la visualización y el análisis de datos. Ofrece una integración perfecta con otras Servicios de AWS funciones de seguridad integradas y una escalabilidad de nivel empresarial. El servicio simplifica la creación y la administración de los paneles y, al mismo tiempo, proporciona funciones avanzadas, como el acceso a las fuentes de datos entre cuentas y la integración con ellas. AWS IAM Identity Center
- Amazon Managed Service para Prometheus: monitorización gestionada, segura y de alta disponibilidad

  Amazon Managed Service for Prometheus es un servicio de monitorización totalmente gestionado y compatible con Prometheus. Proporciona escalado automatizado, alta disponibilidad e ingesta y consulta seguras de métricas. El servicio se integra perfectamente con Amazon EKS y elimina la sobrecarga operativa de la administración de los servidores Prometheus.

Soluciones de código abierto o patentadas

Las AWS herramientas descritas en la sección anterior ofrecen una integración perfecta y servicios gestionados. Las herramientas de código abierto que se enumeran en esta sección se Servicios de AWS complementan al proporcionar flexibilidad y amplias opciones de personalización. Comprender las capacidades y los casos de uso de cada herramienta le ayuda a diseñar las estrategias de monitoreo que mejor se adapten a sus requisitos específicos.

- Prometheus: kit de herramientas de recopilación de métricas

  Prometheus es una solución de código abierto para la recopilación de métricas en entornos de Kubernetes. Su base de datos de series temporales y su lenguaje de consultas ProMQL permiten realizar análisis de métricas sofisticados. Las capacidades de detección de servicios de la plataforma se adaptan automáticamente a los entornos dinámicos de Kubernetes, y su sistema de gestión de alertas lo mantiene informado de los problemas críticos. Prometheus ofrece amplias opciones de integración, lo que lo convierte en una opción versátil para el monitoreo integral de métricas.
- Grafana: motor de visualización avanzada

Grafana transforma los datos de monitoreo complejos en información procesable a través de sus capacidades de visualización. La plataforma crea paneles personalizados que combinan datos de múltiples fuentes y proporcionan una vista unificada de las métricas de la infraestructura y las aplicaciones. Su compatibilidad con diversas fuentes de datos y sus funciones de gestión de alertas proporcionan una supervisión exhaustiva. Grafana puede ayudarlo a visualizar datos históricos y en tiempo real, para que pueda identificar tendencias y tomar decisiones informadas.

- Fluent Bit: capa de registro unificada

  Esta solución de registro proporciona la recopilación y administración de registros para los entornos de Kubernetes. Su integración nativa con Kubernetes garantiza una recopilación de registros fluida desde contenedores y nodos, y su compatibilidad con varios destinos de salida ofrece flexibilidad en el almacenamiento y el análisis de los registros. Las funciones avanzadas, como el análisis y el filtrado de registros, le permiten procesar y enrutar los registros en función de requisitos específicos. La naturaleza liviana de Fluent Bit lo hace especialmente adecuado para entornos en contenedores.

- Datadog: observabilidad completa

  Datadog proporciona capacidades de monitoreo integrales con soporte nativo de Kubernetes. Ofrece monitoreo de infraestructura, monitoreo del rendimiento de las aplicaciones (APM), administración de registros y análisis en tiempo real. Puede utilizar el descubrimiento automático de servicios y el amplio catálogo de integración de la plataforma para la supervisión de Amazon EKS, así como sus capacidades de aprendizaje automático para detectar anomalías y predecir posibles problemas.

- New Relic: monitoreo del rendimiento de las aplicaciones

  New Relic ofrece visibilidad del rendimiento de las aplicaciones y del estado de la infraestructura. Su integración con Kubernetes proporciona información detallada sobre los contenedores, rastreo distribuido y paneles personalizados. La plataforma le ayuda a correlacionar el rendimiento de las aplicaciones con las métricas de la infraestructura, para que pueda identificar y resolver los problemas rápidamente.

- Elastic Stack (ELK Stack): análisis y búsqueda de registros

  El ELK Stack combina Elasticsearch, Logstash y Kibana para ofrecer capacidades de análisis y administración de registros. Ofrece funciones de búsqueda avanzada, herramientas de visualización y funciones de aprendizaje automático. Puede usar la pila para gestionar grandes volúmenes de datos de registro de sus entornos de Amazon EKS.

Herramientas especializadas

Puede combinar las siguientes herramientas en función de sus requisitos de supervisión específicos, la escala de las operaciones y las preferencias de la organización. La clave es crear un conjunto de monitoreo que proporcione una visibilidad completa y, al mismo tiempo, sea manejable y rentable.

- kube-state-metrics (KSM): monitoreo del estado de Kubernetes

  Este servicio complementario escucha el servidor API de Kubernetes y genera métricas sobre el estado de los objetos. Proporciona información sobre el estado de las implementaciones, los módulos y otros recursos de Kubernetes.
- Kubernetes Metrics Server: métricas de recursos

  Este servidor de métricas recopila métricas de recursos de los kubelets y las expone a través de la API de métricas de Kubernetes. Proporciona escalado automático de módulos horizontales y métricas básicas de CPU y memoria.
- Kubecost: monitoreo de costos de Kubernetes

  Herramientas como Kubecost proporcionan un análisis de costes detallado y recomendaciones de optimización para los clústeres de EKS. Le ayudan a comprender y optimizar el gasto en la nube en diferentes espacios de nombres, implementaciones y servicios.

Implementación de alta disponibilidad para las soluciones de monitoreo Amazon EKS

Una estrategia sólida de alta disponibilidad (HA) para la supervisión de Amazon EKS es fundamental para garantizar una visibilidad continua de su entorno de Kubernetes. En esta sección, se describe un enfoque integral para implementar la alta disponibilidad en diferentes aspectos de su infraestructura de monitoreo.

Redundancia y escalabilidad arquitectónicas

La creación de un sistema de monitoreo de alta disponibilidad comienza con un diseño arquitectónico adecuado. Los componentes de monitoreo deben distribuirse en varias zonas de AWS disponibilidad para protegerlos contra los errores de la zona. Esto incluye la implementación del escalado horizontal para los componentes de monitoreo críticos, como los servidores Prometheus, los recopiladores de registros y los administradores de alertas. Puede utilizar servicios AWS gestionados como

Amazon Managed Service for Prometheus y Amazon Managed Grafana para reducir los gastos operativos y, al mismo tiempo, garantizar una alta disponibilidad. Configure mecanismos automáticos de conmutación por error para mantener la continuidad del servicio durante las averías de los componentes, con controles de estado y procedimientos de recuperación automatizados.

Estrategia de almacenamiento de datos resiliente

La resiliencia del almacenamiento de datos es fundamental para mantener la confiabilidad del sistema de monitoreo. La implementación de soluciones de almacenamiento distribuido garantiza que los registros y los datos métricos permanezcan accesibles incluso si fallan los nodos de almacenamiento individuales. Esto incluye configurar la replicación de datos adecuada en varias zonas de disponibilidad y utilizar diferentes backends de almacenamiento para garantizar la redundancia. Establezca procedimientos de respaldo periódicos para los datos históricos, con procesos de recuperación documentados para diversos escenarios de falla. Para las bases de datos de series temporales, como Prometheus, la implementación de soluciones de almacenamiento remoto ayuda a separar las preocupaciones de almacenamiento de la recopilación de datos y mejora la confiabilidad general del sistema.

Gestión de alertas redundante

La gestión de alertas requiere una atención especial en una configuración de alta disponibilidad. La implementación de gestores de alertas redundantes garantiza que las notificaciones críticas lleguen a los destinatarios previstos incluso en caso de fallo del sistema. Configure varios canales de notificación, como el correo electrónico, los SMS o Slack, y PagerDuty proporcione vías de comunicación alternativas. Utiliza mecanismos de deduplicación de alertas para evitar una avalancha de alertas en caso de fallo parcial del sistema y utiliza métodos de notificación alternativos para garantizar que las alertas críticas no se pierdan nunca. La implementación de la correlación de alertas ayuda a mantener el contexto durante los escenarios de conmutación por error y evita que los sistemas redundantes envíen notificaciones duplicadas.

Equilibrio de carga y descubrimiento de servicios

El equilibrio de carga adecuado es esencial para mantener los servicios de monitoreo estables. AWS Los balanceadores de carga de aplicaciones distribuyen el tráfico de monitoreo entrante entre varios puntos finales, y las comprobaciones de estado garantizan que el tráfico se dirija solo a las instancias en buen estado. Los mecanismos de detección de servicios ayudan a que los componentes de supervisión se adapten automáticamente a los cambios del entorno, como la adición de nuevos

nodos o servicios. Implemente agentes de monitoreo de manera uniforme en todos los nodos DaemonSets para garantizar una cobertura integral a medida que el clúster se amplía.

Consideraciones adicionales sobre alta disponibilidad

Resiliencia de la red:

- Implemente rutas de red redundantes.
- Configure el diseño de subred adecuado en todas las zonas de disponibilidad.
- Úselo AWS Direct Connectcon rutas de respaldo.
- Configure los grupos de seguridad y las listas de control de acceso a la red (red ACLs) adecuados.

Supervisión de los monitores:

- Implemente sistemas de monitoreo secundarios.
- Implemente el monitoreo entre regiones.
- Configure alertas para los sistemas que no responden.
- Pruebe los procedimientos de conmutación por error con regularidad.

Planificación de la capacidad:

- Supervise las tendencias de uso de los recursos.
- Implemente el escalado predictivo.
- Pruebe el rendimiento de forma periódica.

Gestión de datos:

- Implemente políticas de retención de datos.
- Configure la agregación de métricas.
- Planifique la administración del ciclo de vida de los datos.
- Optimice el almacenamiento de forma regular.

Procedimientos de recuperación:

- Procesos de recuperación de documentos.

- Pruebe la recuperación ante desastres con regularidad.
- Implemente la recuperación automática siempre que sea posible.
- Identifique e implemente rutas de escalamiento claras.

Al implementar estas prácticas de alta disponibilidad, puede asegurarse de que su infraestructura de monitoreo de Amazon EKS siga siendo confiable y resiliente, y de que tiene una visibilidad continua de sus entornos de Kubernetes, incluso en varios escenarios de falla. Las pruebas y actualizaciones periódicas de estas configuraciones de alta disponibilidad garantizan que sigan siendo eficaces a medida que el entorno evoluciona.

Prácticas recomendadas para la supervisión en Amazon EKS

Enfoque de implementación estratégica

Una estrategia de monitoreo exitosa de Amazon EKS comienza con un enfoque de implementación gradual y bien planificado.

- Comience por identificar y monitorear las métricas críticas que afectan directamente a las operaciones empresariales y a la confiabilidad de las aplicaciones. Esta base debe incluir las métricas de infraestructura esenciales, los indicadores clave de rendimiento de las aplicaciones y las métricas de seguridad críticas. Amplíe gradualmente la cobertura de monitoreo en función de las necesidades operativas y las lecciones aprendidas, y asegúrese de que cada incorporación aporte un valor significativo.
- Implemente procesos de despliegue automatizados mediante el uso de herramientas de infraestructura como código (IaC), como Terraform, o CloudFormation para garantizar la coherencia y la repetibilidad.
- Pruebe y valide los sistemas de monitoreo para ayudar a mantener la confiabilidad y la precisión.
- Refina los parámetros de supervisión de forma continua para adaptarlos a las cambiantes necesidades empresariales.

Gestión eficaz de los datos

La gestión adecuada de los datos es fundamental para mantener una solución de supervisión eficiente y rentable.

- Implemente políticas claras de retención de datos que equilibren las necesidades de análisis histórico con los costos de almacenamiento.
- Configure las frecuencias de muestreo adecuadas para los diferentes tipos de métricas: una frecuencia más alta para las métricas críticas y una frecuencia más baja para las menos críticas.
- Utilice la agregación de métricas para reducir el volumen de datos y, al mismo tiempo, conservar información significativa, especialmente para el análisis de tendencias a largo plazo.
- Implemente procedimientos sistemáticos de conservación y archivado de registros para los sistemas de registro centralizados (como CloudWatch los registros) a fin de gestionar los costes de almacenamiento y mantener el acceso a los datos importantes de forma accesible.

> **Note**
>
> En Amazon EKS versión 1.21 o posterior, el kubelet gestiona automáticamente la rotación de registros a nivel de contenedor.

- Considere la posibilidad de implementar una hot-warm-cold arquitectura de almacenamiento de registros a fin de optimizar tanto la velocidad de acceso como la rentabilidad.

Configuración y administración de alertas

La configuración de las alertas requiere una consideración cuidadosa para mantener la eficacia sin provocar fatiga en las alertas.

- Defina umbrales claros y procesables en función de los objetivos de nivel de servicio (SLOs) y de los patrones de rendimiento históricos.
- Implemente un sistema de gravedad de las alertas por niveles que diferencie claramente entre los problemas críticos que requieren atención inmediata y los asuntos menos urgentes.
- Asegúrese de que las alertas proporcionen suficiente contexto e información procesable para facilitar la resolución rápida de los problemas.
- Establezca procedimientos de escalamiento claros con propiedad y tiempos de respuesta definidos para los diferentes niveles de gravedad de las alertas.
- Revise y perfeccione las configuraciones de alertas con regularidad para ayudar a mantener su relevancia y eficacia.

Optimización de recursos

El monitoreo continuo de la utilización de los recursos es esencial para mantener operaciones rentables.

- Implemente una supervisión integral de los recursos en todos los componentes del clúster, incluidos los nodos, los pods y los volúmenes persistentes.
- Configure el escalado automático en función de los patrones de uso reales y los requisitos de rendimiento para garantizar una utilización eficiente de los recursos y, al mismo tiempo, mantener el rendimiento.
- Utilice etiquetas de asignación de costes para realizar un seguimiento del consumo de recursos por parte de los diferentes equipos, aplicaciones o entornos.
- Analice periódicamente las métricas de eficiencia de los recursos para identificar las oportunidades de optimización e implementar mejoras.
- Considere la posibilidad de implementar herramientas de administración de costos para rastrear y optimizar el gasto en la nube.

Seguridad

Las consideraciones de seguridad deben ser parte integral de su estrategia de monitoreo.

- Implemente principios de acceso con privilegios mínimos para todos los componentes de monitoreo para garantizar que los usuarios y los servicios solo tengan los permisos que necesitan.
- Habilite un registro de auditoría integral para rastrear todos los accesos y cambios en los sistemas de monitoreo.
- Realice revisiones de seguridad periódicas de las configuraciones de monitoreo y los patrones de acceso para identificar posibles vulnerabilidades.
- Implemente el cifrado para los datos de monitoreo confidenciales, tanto en tránsito como en reposo.
- Integre la supervisión de la seguridad con los sistemas de información de seguridad y gestión de eventos (SIEM) existentes para obtener una visibilidad completa de la seguridad.

Consideraciones de supervisión avanzada en Amazon EKS

Optimización del rendimiento:

- Optimice los intervalos de recopilación de métricas.
- Configure patrones de consulta eficientes.
- Implemente la agregación previa de métricas.
- Utilice las soluciones de almacenamiento adecuadas.

Cumplimiento y gobierno:

- Mantenga los registros de auditoría.
- Implemente la supervisión del cumplimiento.
- Proporcione informes de cumplimiento periódicos.
- Documente los procedimientos de supervisión.

Recuperación ante desastres:

- Realice copias de seguridad de las configuraciones de monitoreo con regularidad.
- Procedimientos de recuperación de documentos.
- Pruebe los procesos de recuperación.

Mejora continua:

- Supervise las sesiones de revisión con regularidad.
- Optimice los ciclos de rendimiento.
- Actualice la supervisión en función de los incidentes.
- Incorpore los comentarios de los usuarios.

Estas prácticas recomendadas proporcionan un marco para implementar y mantener soluciones de monitoreo eficaces para los entornos de Amazon EKS. Revise y actualice periódicamente estas prácticas para que se ajusten a las necesidades de su organización y a los estándares del sector. La supervisión no se realiza una sola vez, sino que es un proceso continuo que requiere atención y perfeccionamiento periódicos.

Rastreo en Amazon EKS

El rastreo es un componente fundamental de la observabilidad de las aplicaciones en Amazon EKS. El rastreo proporciona una visibilidad detallada de los flujos de solicitudes y las interacciones de los servicios mediante la recopilación, el procesamiento y la visualización de la ruta de las solicitudes a medida que viajan a través de varios microservicios que se implementan en los clústeres de EKS. Esta capacidad le ayuda a comprender el comportamiento del sistema, identificar los cuellos de botella y solucionar problemas de forma eficaz en su entorno de Amazon EKS. El seguimiento eficaz elimina la complejidad de depurar los sistemas distribuidos al proporcionar visibilidad de los flujos de solicitudes. end-to-end Permite realizar un seguimiento de las transacciones a través de los límites del servicio e identificar problemas o errores de rendimiento en las cargas de trabajo de Amazon EKS.

La implementación general del rastreo en Amazon EKS le permite comprender el comportamiento del sistema, optimizar el rendimiento y mantener la confiabilidad de sus aplicaciones en contenedores. En última instancia, las capacidades de rastreo mejoran la visibilidad operativa y la capacidad de mantenimiento del sistema en los entornos de Amazon EKS.

AWS X-Ray desempeña un papel importante en el seguimiento de los datos sobre su aplicación. El rastreo implica monitorear varios aspectos de las interacciones del servicio, incluidos los siguientes:

- Las rutas y dependencias de las solicitudes proporcionan información crucial sobre el comportamiento de su sistema distribuido. Hacen un seguimiento del recorrido completo de las solicitudes a medida que atraviesan diferentes microservicios y componentes. El mapeo de las dependencias de los servicios le ayuda a comprender los patrones de comunicación e identificar las rutas críticas en la arquitectura de su aplicación. Para obtener detalles sobre la implementación, consulte Uso del mapa AWS X-Ray de rastreo del servicio en la documentación de X-Ray.
- Las latencias y los cuellos de botella del servicio son métricas esenciales para mantener un rendimiento óptimo del sistema. Al medir y analizar los tiempos de respuesta entre los servicios, puede identificar los problemas de rendimiento de forma eficaz. Estos datos le permiten identificar los servicios u operaciones específicos que están provocando retrasos en la cadena de solicitudes y realizar esfuerzos de optimización específicos. Para obtener más información sobre el análisis de latencia, consulte Interactuar con la consola de Analytics en la documentación de X-Ray.
- Los patrones de propagación de errores le ayudan a comprender la confiabilidad del sistema y la tolerancia a los errores. Al comprender cómo las fallas se propagan en cascada por el

sistema mediante el seguimiento de las rutas de error entre los servicios, podrá diseñar mejor sus aplicaciones. Esta visibilidad le ayuda a identificar la causa raíz de los errores y su impacto en los servicios dependientes, lo que se traduce en sistemas más resilientes. Para obtener detalles sobre la implementación, consulte la documentación de Traces in the X-Ray.

- La utilización de los recursos en todos los servicios proporciona información sobre la eficiencia del sistema y la optimización de los costes. Puede supervisar los patrones de uso de la CPU, la memoria y la red que están correlacionados con los datos de rastreo para comprender la demanda de recursos. Estos datos le ayudan a analizar las tendencias de consumo de recursos para optimizar el rendimiento y los costes del servicio en todo su clúster de EKS. Para ver la configuración de la supervisión, consulte Supervisar el rendimiento del clúster y ver los registros en la documentación de Amazon EKS.
- Los flujos de transacciones de los usuarios finales son fundamentales para comprender y mejorar la experiencia del usuario. Al realizar un seguimiento completo de las interacciones de los usuarios, desde los servicios de interfaz hasta los de fondo, puede garantizar un rendimiento óptimo de las aplicaciones. Puede medir y optimizar los tiempos de end-to-end respuesta para los recorridos críticos de los usuarios, lo que repercute directamente en la satisfacción del cliente. Para implementar la supervisión de los usuarios finales, utilice el AWS X-Ray SDK como lenguaje de programación.
- Las interacciones entre las pasarelas de API constituyen la primera línea del rendimiento y la seguridad de su aplicación. Puede supervisar los patrones de solicitud y el rendimiento en los puntos de entrada de la API para garantizar una prestación de servicios óptima. Esta visibilidad le ayuda a realizar un seguimiento de los efectos de la autenticación, la autorización y la limitación de la velocidad en los flujos de solicitudes, a fin de mantener los requisitos de seguridad y rendimiento. Obtenga más información sobre el rastreo de API en la documentación de Amazon API Gateway with X-Ray.

El rastreo efectivo en Amazon EKS va más allá de recopilar tramos y trazas. Requiere una estrategia bien estructurada que equilibre las necesidades de observabilidad con el rendimiento del sistema. Esta estrategia debe centrarse en:

- Implementación de las tasas de muestreo adecuadas: configure las reglas de muestreo en función de los patrones de tráfico y las prioridades comerciales para optimizar los costos y, al mismo tiempo, mantener la visibilidad de las transacciones críticas. Para obtener más información, consulte Configuración de las reglas de muestreo en la documentación de X-Ray.

- Definir las rutas y los servicios críticos que se deben rastrear: identifique y priorice los servicios esenciales y los recorridos de los usuarios que requieren un seguimiento detallado para garantizar una supervisión óptima del rendimiento. Para obtener más información, consulte Enviar datos métricos y de rastreo con ADOT Operator en la documentación de Amazon EKS.
- Establecer políticas de retención de datos adecuadas: configure reglas de administración del ciclo de vida de los datos para equilibrar las necesidades de observabilidad con los costos de almacenamiento y los requisitos de conformidad. Para ver las políticas CloudWatch de retención, consulte Trabajar con grupos de registros y flujos de registros en la documentación de CloudWatch registros.
- Configuración de herramientas de visualización y análisis eficaces: Implemente y configure herramientas de visualización como la consola de AWS X-Ray Analytics o Amazon Managed Grafana para analizar los datos de rastreo de forma eficaz. Para obtener más información, consulte Interactuar con la consola de Analytics en la documentación de X-Ray.

En esta sección:

- Herramientas de rastreo para Amazon EKS
- Prácticas recomendadas para el rastreo en Amazon EKS

Herramientas de rastreo para Amazon EKS

Amazon EKS admite varias opciones AWS y opciones de terceros para implementar el rastreo distribuido.

Servicios de AWS

- AWS X-Ray: Plataforma avanzada de rastreo distribuido

  X-Ray es un sistema totalmente gestionado Servicio de AWS que proporciona capacidades end-to-end de rastreo. Instrumenta Servicios de AWS y proporciona automáticamente análisis y mapas de servicios detallados para las aplicaciones que se ejecutan en Amazon EKS. X-Ray está integrado con otros Servicios de AWS, incluido Amazon CloudWatch, y ofrece una correlación automática de las trazas con Servicio de AWS las llamadas.
- AWS Distribución para OpenTelemetry: Marco de observabilidad unificado

  Distro for OpenTelemetry es una distribución segura, lista para la producción y AWS compatible de aplicaciones nativas de la nube. OpenTelemetry Ofrece capacidades de instrumentación

independientes del proveedor y, al mismo tiempo, mantiene una Servicio de AWS integración nativa, lo que la hace ideal para entornos de nube híbrida. Distro for OpenTelemetry admite múltiples backends de observabilidad y proporciona una integración perfecta con los servicios de monitoreo. AWS

Soluciones de código abierto

- OpenTelemetry: Marco de observabilidad de código abierto

  OpenTelemetry proporciona un marco de observabilidad estandarizado con bibliotecas de instrumentación completas que admiten varios lenguajes de programación. Sus opciones de backend flexibles y su enfoque independiente del proveedor lo hacen ideal para cargas de trabajo que requieren coherencia en diferentes entornos. El amplio ecosistema del marco garantiza una amplia compatibilidad con diversas soluciones de monitoreo.

- Jaeger: plataforma de rastreo distribuido de código abierto

  Jaeger ofrece capacidades de rastreo integrales con propagación de contexto distribuida en tiempo real. Proporciona un análisis de la causa raíz y una optimización del rendimiento mediante una visualización detallada de la dependencia del servicio. La arquitectura de Jaeger está diseñada para ofrecer una alta escalabilidad y admite varios backends de almacenamiento, lo que la hace adecuada para despliegues de Amazon EKS a gran escala. Consulte la configuración de Jaeger para EKS

- Grafana Tempo: rastreo distribuido

  Tempo es una solución de Grafana Labs que proporciona almacenamiento de trazas a gran escala y una integración perfecta con las métricas de Prometheus. Su rentable modelo de retención de trazas y su integración nativa con Grafana lo hacen adecuado para las organizaciones que ya utilizan Grafana para la visualización. La arquitectura de Tempo está diseñada específicamente para entornos nativos de la nube, como Amazon EKS.

Prácticas recomendadas para el rastreo en Amazon EKS

En esta sección se proporciona una lista completa de las mejores prácticas y técnicas para crear un sistema de rastreo eficaz que mejore la observabilidad y la solución de problemas de las aplicaciones basadas en Kubernetes en Amazon EKS.

- Muestreo estratégico: configure diferentes frecuencias de muestreo en función de los patrones de tráfico de su aplicación y de la importancia de los servicios que utilice. Implemente frecuencias de muestreo más altas para las rutas críticas y, al mismo tiempo, reduzca el muestreo para las rutas de gran volumen y menos críticas para optimizar los costos. Para obtener orientación, consulte Configuración de las reglas de muestreo en la AWS X-Ray documentación.
- Configuración de la instrumentación: utilice herramientas de instrumentación automáticas, como el X-Ray SDK o AWS Distro para OpenTelemetry coleccionistas, a fin de minimizar el esfuerzo de instrumentación manual. Mantenga unas convenciones de nomenclatura coherentes y una propagación del contexto en todos los servicios para mejorar la correlación de trazas. Para obtener más información, consulte la distribución para ver la documentación de los OpenTelemetry recopiladores.
- Administración de datos: implemente los períodos de retención y las estrategias de compresión adecuados para equilibrar los costos de almacenamiento con sus necesidades de observabilidad. Establezca controles de privacidad de datos y procedimientos de respaldo claros para proteger los datos de rastreo confidenciales. Para obtener más información, consulte Cambiar la retención de datos de registro en CloudWatch los registros en la documentación de CloudWatch registros.
- Optimización del rendimiento: supervise y optimice la sobrecarga de rastreo para minimizar el impacto en el rendimiento de las aplicaciones. Utilice un almacenamiento en búfer y un procesamiento asíncrono eficientes para reducir el impacto en la latencia. Para obtener más información, consulte Configuración del AWS X-Ray daemon en la documentación de X-Ray.
- Controles de seguridad: Implemente los controles de acceso y las medidas de protección de datos adecuados mediante el uso de funciones y políticas de IAM. Las auditorías de seguridad y las revisiones de conformidad periódicas ayudan a garantizar que los datos de rastreo permanezcan seguros. Para obtener más información, consulte Seguridad AWS X-Ray en la documentación de X-Ray.
- Supervisión y alertas: configure una supervisión integral del estado de la recopilación de trazas y configure las alertas en caso de problemas relacionados con la recopilación. Realice un seguimiento de las tasas de muestreo y las métricas de rendimiento del sistema para garantizar un funcionamiento óptimo. Para obtener más información, consulte Container Insights en la CloudWatch documentación.
- Alta disponibilidad: implemente recopiladores redundantes en todas las zonas de disponibilidad y configure los mecanismos de conmutación por error adecuados. Las pruebas periódicas de la configuración de alta disponibilidad garantizan una recopilación de trazas fiable. Para obtener más información, consulte Uso de AWS Distro OpenTelemetry como recopilador en la documentación de Amazon Managed Service for Prometheus.

Si sigue estas prácticas recomendadas, puede crear un sistema de rastreo sólido, eficiente y efectivo para su entorno de Amazon EKS. Esto ayudará a garantizar una observabilidad completa, una solución de problemas eficiente y un rendimiento óptimo de sus aplicaciones basadas en Kubernetes.

Alertas en Amazon EKS

Las alertas son un componente fundamental de la administración y el mantenimiento de las aplicaciones que se ejecutan en Amazon EKS. Sirve como un sistema de alerta temprana que notifica a los operadores y desarrolladores sobre posibles problemas, anomalías o degradaciones del rendimiento antes de que se conviertan en problemas graves que puedan afectar a la disponibilidad del servicio o a la experiencia del usuario. Las alertas implican la supervisión de varios aspectos del clúster de Kubernetes, entre los que se incluyen:

- Estado de la infraestructura
- Rendimiento de las aplicaciones
- Métricas de contenedores
- Métricas empresariales personalizadas

Las alertas eficaces en Amazon EKS van más allá de la simple configuración de notificaciones. Requiere una well-thought-out estrategia que equilibre la necesidad de información puntual con la posibilidad de agotar las alertas. Esta estrategia debería:

- Defina umbrales y condiciones significativos.
- Priorice las alertas en función de la gravedad y el impacto.
- Implemente los procedimientos de enrutamiento y escalamiento adecuados.
- Intégrelo con las herramientas de comunicación y gestión de incidentes.

En esta sección:

- Herramientas de alertas para Amazon EKS
- Prácticas recomendadas para la emisión de alertas en Amazon EKS

Herramientas de alertas para Amazon EKS

Amazon EKS admite varias opciones AWS y opciones de terceros para implementar alertas. Cuando elija una herramienta para las alertas de Amazon EKS, tenga en cuenta factores como las capacidades de integración, la escalabilidad, la facilidad de uso, el costo y las características específicas que se adapten a sus requisitos de monitoreo y alertas. Muchas organizaciones utilizan

una combinación de estas herramientas para crear una solución integral de supervisión y alertas para sus entornos Amazon EKS.

- Amazon CloudWatch: Servicio de AWS para monitoreo y observabilidad

  CloudWatch proporciona métricas, registros y alarmas para los clústeres de EKS y se integra bien con otros Servicios de AWS.
- Prometheus: herramienta de monitoreo y alertas de código abierto para Kubernetes

  Prometheus proporciona un potente lenguaje de consulta (ProMQL) para definir las condiciones de alerta.
- Alertmanager: el complemento de Prometheus para gestionar las alertas

  Alertmanager permite deduplicar, agrupar y enrutar las alertas. Es compatible con varios canales de notificación, incluidos el correo electrónico, Slack y. PagerDuty
- Grafana: plataforma de código abierto para monitoreo y observabilidad

  Grafana proporciona capacidades de visualización y alerta. Se puede integrar con varias fuentes de datos, incluidas CloudWatch Prometheus y.
- Elastic Stack (ELK Stack): combinación de Elasticsearch, Logstash y Kibana

  Esta herramienta es útil para la agregación, el análisis y las alertas de registros. Se puede ampliar con las funciones de observabilidad de Elastic.
- Soluciones de terceros

  Hay muchas herramientas disponibles en el mercado, incluidas Datadog, New Relic, Sysdig, Dynatrace, Zabbix, Nagios, Splunk, IBM Instana y. AppDynamics

Prácticas recomendadas para la emisión de alertas en Amazon EKS

En esta sección se describen las prácticas recomendadas para crear un sistema de alertas sólido que mejore la fiabilidad y el rendimiento de las aplicaciones basadas en Kubernetes en Amazon EKS.

Defina umbrales de alerta claros:

- Establezca umbrales significativos en función de los datos históricos y los requisitos empresariales.
- Utilice umbrales dinámicos cuando proceda para tener en cuenta las diferentes cargas de trabajo.

Implemente la priorización de alertas:

- Clasifique las alertas por gravedad (por ejemplo, críticas, altas, medias o bajas).
- Alinee las prioridades de las alertas con el impacto empresarial.

Evite la fatiga de las alertas:

- Reduzca el ruido eliminando las alertas redundantes o de bajo valor.
- Correlaciona las alertas con los problemas relacionados con el grupo.

Utilice alertas en varias etapas:

- Implemente umbrales de advertencia antes de que se alcancen los niveles críticos.
- Utilice diferentes canales de notificación para diferentes niveles de gravedad de las alertas.

Implemente un enrutamiento de alertas adecuado:

- Asegúrese de que las alertas se envíen a los equipos o personas correctos.
- Utilice los horarios y rotaciones de guardia para obtener cobertura durante todo el día y todos los días.

Aproveche las métricas nativas de Kubernetes:

- Supervise los componentes principales de Kubernetes (nodos, pods, servicios).
- Usa kube-state-metrics (KSM) para obtener métricas adicionales de objetos de Kubernetes.

Supervise tanto la infraestructura como las aplicaciones:

- Configure alertas sobre el estado del clúster, el estado de los nodos y la utilización de los recursos.
- Implemente alertas específicas de la aplicación, como las tasas de error y la latencia.

Utilice Prometheus y Alertmanager:

- Utilice Prometheus para la recopilación de métricas y ProMQL para definir las condiciones de alerta.
- Utilice Alertmanager para el enrutamiento y la deduplicación de alertas.

Integre con Amazon CloudWatch:

- Usa CloudWatchContainer Insights para obtener métricas específicas de Amazon EKS.
- Configure CloudWatchalarmas para las métricas de AWS recursos críticos.

Implemente alertas contextuales:

- Incluya información relevante en los mensajes de alerta, como el nombre del clúster, el espacio de nombres y los detalles del pod.
- Proporcione enlaces a los cuadros de mando o manuales relevantes en las alertas.

Utilice la detección de anomalías:

- Implemente la detección de anomalías basada en el aprendizaje automático para patrones complejos.
- Utilice servicios como la detección de CloudWatch anomalías o herramientas de terceros.

Implemente la supresión y el silenciamiento de alertas:

- Permita la supresión temporal de los problemas conocidos.
- Implemente períodos de mantenimiento para reducir el ruido durante los tiempos de inactividad planificados.

Supervise el rendimiento de las alertas:

- Realice un seguimiento de métricas como la frecuencia de las alertas, el tiempo de resolución y las tasas de falsos positivos.
- Revisa y refina periódicamente las reglas de alerta en función de estas métricas.

Implemente procedimientos de escalamiento:

- Defina rutas de escalamiento claras para las alertas no resueltas.
- Utilice herramientas como PagerDuty Opsgenie para las escalaciones automatizadas.

Pruebe los sistemas de alerta con regularidad:

- Realice pruebas periódicas de su canal de alertas.
- Incluya las pruebas de alerta en los simulacros de recuperación ante desastres.

Utilice plantillas para mantener la coherencia de las alertas:

- Cree plantillas de alertas estandarizadas para escenarios comunes.
- Garantice la coherencia del formato y la información en todas las alertas.

Implemente una limitación de velocidad:

- Prevenga las tormentas de alertas implementando una limitación de frecuencia en las alertas que se activan con frecuencia.

Usa métricas personalizadas:

- Implemente métricas personalizadas para el monitoreo específico de la aplicación.
- Usa la API de métricas personalizadas de Kubernetes para realizar un escalado automático en función de estas métricas.

Implemente la integración de registros:

- Correlaciona las alertas con los registros relevantes para una solución de problemas más rápida.
- Utilice herramientas como Grafana Loki o ELK Stack junto con su sistema de alertas.

Considera las alertas de costos:

- Configure alertas en caso de picos inesperados en el uso de los recursos o los costos.
- Utilice AWS Budgetsherramientas de gestión de costes de terceros.

Utilice el rastreo distribuido:

- Integre herramientas de rastreo distribuido como Jaeger o. AWS X-Ray
- Configure alertas para detectar patrones de rastreo o latencias anormales.

Documente los manuales de alertas:

- Cree manuales claros y procesables para cada tipo de alerta.
- Incluya los pasos de solución de problemas y los procedimientos de escalamiento en los manuales de instrucciones.

Si sigue estas prácticas recomendadas, puede crear un sistema de alertas sólido, eficiente y eficaz para su entorno de Amazon EKS. Esto ayudará a garantizar la alta disponibilidad, la rápida resolución de problemas y el rendimiento óptimo de sus aplicaciones basadas en Kubernetes.

Pasos siguientes

Esta guía proporciona un marco integral para implementar una observabilidad sólida en los entornos de Amazon EKS, centrándose en la recopilación de métricas, la infraestructura de registro, el rastreo distribuido y la optimización de costos. Al comprender y aplicar estos componentes principales, puede crear un entorno de contenedores rentable, altamente observable y fácil de mantener, que proporcione información detallada sobre el comportamiento de las aplicaciones y la infraestructura. La integración de Servicios de AWS Amazon CloudWatch Container Insights y AWS X-Ray, combinada con soluciones de código abierto como Prometheus, crea una base sólida para monitorear OpenTelemetry y solucionar problemas de aplicaciones en contenedores.

El éxito de la implementación se basa en un enfoque gradual, que comienza con la recopilación de las métricas principales y se amplía gradualmente hasta alcanzar capacidades integrales de registro y rastreo distribuido. Le recomendamos que comience por evaluar sus capacidades de monitoreo actuales, identificar las brechas y seleccionar las combinaciones de herramientas adecuadas que se ajusten a sus requisitos operativos y a la experiencia de su equipo. Este enfoque metódico garantiza que cada componente del conjunto de observabilidad se implemente e integre correctamente, mientras que los equipos desarrollan las habilidades y los procesos necesarios para utilizar estas herramientas de forma eficaz.

La sostenibilidad a largo plazo de la observabilidad de Amazon EKS depende de la optimización periódica de los costos, los recursos y los procesos. Debe revisar y ajustar continuamente su infraestructura de observabilidad, incluidas las políticas de retención de datos, las tasas de muestreo y la asignación de recursos, para mantener el equilibrio adecuado entre la supervisión integral y la eficiencia operativa. Este enfoque iterativo de mejora, combinado con la formación continua del equipo y las actualizaciones de la documentación, permite a su organización mantener una observabilidad efectiva y, al mismo tiempo, respaldar el crecimiento empresarial y adaptarse a las cambiantes arquitecturas de aplicaciones.

Recursos

AWS documentación

- Guía de prácticas recomendadas de Amazon EKS
- Información sobre CloudWatch contenedores de Amazon
- Servicio administrado por Amazon para Prometheus
- Amazon Managed Grafana
- AWS Distro para y OpenTelemetry AWS X-Ray
- OpenSearch Servicio Amazon

AWS publicaciones de blog

- Amazon EKS mejora la observabilidad del plano de control de Kubernetes
- Automatización de la recopilación de métricas en Amazon EKS con Amazon Managed Service para los raspadores gestionados por Prometheus
- Automatice la supervisión de su clúster de Amazon EKS mediante CloudWatch Container Insights
- Mejora de la observabilidad con una solución de monitorización gestionada para Amazon EKS

Otros recursos

- Documentación de OpenTelemetry
- Documentación de Prometheus
- Documentación de Fluent Bit
- Documentación sobre supervisión, registro y depuración en Kubernetes

Historial de documentos

En la siguiente tabla, se describen cambios significativos de esta guía. Si quiere recibir notificaciones de futuras actualizaciones, puede suscribirse a las notificaciones RSS.

Cambio	Descripción	Fecha
Actualizaciones	Hemos actualizado el capítulo Logging in Amazon EKS.	17 de marzo de 2026
Publicación inicial	—	10 de abril de 2025

AWS Glosario de orientación prescriptiva

Los siguientes son términos de uso común en las estrategias, guías y patrones proporcionados por la Guía AWS prescriptiva. Para sugerir entradas, utilice el enlace Enviar comentarios al final del glosario.

Números

Las 7 R

Siete estrategias de migración comunes para trasladar aplicaciones a la nube. Estas estrategias se basan en las 5 R que Gartner identificó en 2011 y consisten en lo siguiente:

- Refactor/re-architect — Mueva una aplicación y modifique su arquitectura aprovechando al máximo las funciones nativas de la nube para mejorar la agilidad, el rendimiento y la escalabilidad. Por lo general, esto implica trasladar el sistema operativo y la base de datos. Ejemplo: migre su base de datos Oracle local a la PostgreSQL-Compatible edición Amazon Aurora.
- Redefinir la plataforma (transportar y redefinir): traslade una aplicación a la nube e introduzca algún nivel de optimización para aprovechar las capacidades de la nube. Ejemplo: Migrar la base de datos Oracle en las instalaciones a Amazon Relational Database Service (Amazon RDS) para Oracle en la nube de Nube de AWS.
- Recomprar (readquirir): cambie a un producto diferente, lo cual se suele llevar a cabo al pasar de una licencia tradicional a un modelo SaaS. Ejemplo: migre su sistema de gestión de relaciones con los clientes (CRM) a Salesforce.com.
- Volver a alojar (migrar mediante lift-and-shift): traslade una aplicación a la nube sin hacer cambios para aprovechar las funcionalidades de la nube. Ejemplo: Migrar la base de datos de Oracle en las instalaciones a Oracle en una instancia de EC2 en la Nube de AWS.
- Reubicar: (migrar el hipervisor mediante lift and shift): traslade la infraestructura a la nube sin comprar equipo nuevo, reescribir aplicaciones o modificar las operaciones actuales. Los servidores se migran de una plataforma en las instalaciones a un servicio en la nube para la misma plataforma. Ejemplo: migrar una Microsoft Hyper-V aplicación a AWS.
- Retener (revisitar): conserve las aplicaciones en el entorno de origen. Estas pueden incluir las aplicaciones que requieren una refactorización importante, que desee posponer para más adelante, y las aplicaciones heredadas que desee retener, ya que no hay ninguna justificación empresarial para migrarlas.

- Retirar: retire o elimine las aplicaciones que ya no sean necesarias en un entorno de origen.

A

A2A () Agent-to-Agent

Un protocolo completo para la colaboración entre agentes que facilita la delegación de tareas y la transferencia de estados.

ABAC

Consulte control de acceso basado en atributos.

servicios abstractos

Consulte servicios administrados.

ACID

Consulte atomicidad, consistencia, aislamiento, durabilidad.

migración activa-activa

Método de migración de bases de datos en el que las bases de datos de origen y destino se mantienen sincronizadas (mediante una herramienta de replicación bidireccional o mediante operaciones de escritura doble) y ambas bases de datos gestionan las transacciones de las aplicaciones conectadas durante la migración. Este método permite la migración en lotes pequeños y controlados, en lugar de requerir una transición única. Es más flexible, pero requiere más trabajo que una migración activa-pasiva.

migración activa-pasiva

Método de migración de bases de datos en el que las bases de datos de origen y destino se mantienen sincronizadas, pero solo la de origen gestiona las transacciones de las aplicaciones conectadas, mientras los datos se replican en la de destino. La base de datos de destino no acepta ninguna transacción durante la migración.

Agente

Un sistema de IA que puede razonar, planificar y tomar medidas de forma autónoma utilizando herramientas para alcanzar los objetivos.

Agent Ops

Prácticas operativas para crear, probar, implementar y ejecutar agentes de IA en producción a escala.

función de agregación

Función SQL que actúa en un grupo de filas y calcula un único valor de devolución para el grupo. Entre los ejemplos de funciones de agregación se incluyen `SUM` y `MAX`.

IA

Consulte inteligencia artificial.

AIOps

Consulte operaciones de inteligencia artificial

anonimización

El proceso de eliminar permanentemente la información personal de un conjunto de datos. La anonimización puede ayudar a proteger la privacidad personal. Los datos anonimizados ya no se consideran datos personales.

antipatrones

Una solución que se utiliza con frecuencia para un problema recurrente en el que la solución es contraproducente, ineficaz o menos eficaz que una alternativa.

control de aplicaciones

Enfoque de seguridad que permite usar de manera exclusiva aplicaciones aprobadas para ayudar a proteger un sistema contra el malware.

cartera de aplicaciones

Recopilación de información detallada sobre cada aplicación que utiliza una organización, incluido el costo de creación y mantenimiento de la aplicación y su valor empresarial. Esta información es clave para el proceso de detección y análisis de la cartera y ayuda a identificar y priorizar las aplicaciones que se van a migrar, modernizar y optimizar.

inteligencia artificial (IA)

El campo de la informática que se dedica al uso de tecnologías informáticas para realizar funciones cognitivas que suelen estar asociadas a los seres humanos, como el aprendizaje, la resolución de problemas y el reconocimiento de patrones. Para más información, consulte ¿Qué es la inteligencia artificial?

operaciones de inteligencia artificial (AIOps)

El proceso de utilizar técnicas de machine learning para resolver problemas operativos, reducir los incidentes operativos y la intervención humana, y mejorar la calidad del servicio. Para obtener más información sobre cómo se utiliza AIOps en la estrategia de migración de AWS , consulte la Guía de integración de operaciones.

cifrado asimétrico

Algoritmo de cifrado que utiliza un par de claves, una clave pública para el cifrado y una clave privada para el descifrado. Puede compartir la clave pública porque no se utiliza para el descifrado, pero el acceso a la clave privada debe estar sumamente restringido.

atomicidad, consistencia, aislamiento, durabilidad (ACID)

Conjunto de propiedades de software que garantizan la validez de los datos y la fiabilidad operativa de una base de datos, incluso en caso de errores, cortes de energía u otros problemas.

control de acceso basado en atributos (ABAC)

La práctica de crear permisos detallados basados en los atributos del usuario, como el departamento, el puesto de trabajo y el nombre del equipo. Para obtener más información, consulte ABAC AWS en la documentación AWS Identity and Access Management (IAM).

origen de datos fidedigno

Ubicación en la que se almacena la versión principal de los datos, que se considera la fuente de información más fiable. Puede copiar los datos del origen de datos autorizado a otras ubicaciones con el fin de procesarlos o modificarlos, por ejemplo, anonimizarlos, redactarlos o seudonimizarlos.

Zona de disponibilidad

Una ubicación distinta dentro de una Región de AWS que está aislada de los fallos en otras zonas de disponibilidad y que proporciona una conectividad de red económica y de baja latencia a otras zonas de disponibilidad de la misma región.

AWS Marco de adopción de la nube (AWS CAF)

Un marco de directrices y mejores prácticas AWS para ayudar a las organizaciones a desarrollar un plan eficiente y eficaz para migrar con éxito a la nube. AWS CAF organiza la orientación en seis áreas de enfoque denominadas perspectivas: negocios, personas, gobierno, plataforma, seguridad y operaciones. Las perspectivas empresariales, humanas y de gobernanza se centran en las habilidades y los procesos empresariales; las perspectivas de plataforma, seguridad y

operaciones se centran en las habilidades y los procesos técnicos. Por ejemplo, la perspectiva humana se dirige a las partes interesadas que se ocupan de los Recursos Humanos (RR. HH.), las funciones del personal y la administración de las personas. Desde esta perspectiva, AWS CAF proporciona orientación para el desarrollo, la formación y la comunicación de las personas a fin de preparar a la organización para una adopción exitosa de la nube. Para obtener más información, consulte la Página web de AWS CAF y el Documento técnico de AWS CAF.

AWS Marco de calificación de la carga de trabajo (AWS WQF)

Herramienta que evalúa las cargas de trabajo de migración de bases de datos, recomienda estrategias de migración y proporciona estimaciones de trabajo. AWS WQF se incluye con AWS Schema Conversion Tool ().AWS SCT Analiza los esquemas de bases de datos y los objetos de código, el código de las aplicaciones, las dependencias y las características de rendimiento y proporciona informes de evaluación.

B

bot malicioso

Bot destinado a causar interrupciones o daños a personas u organizaciones.

BCP

Consulte planificación de la continuidad del negocio.

gráfico de comportamiento

Una vista unificada e interactiva del comportamiento de los recursos y de las interacciones a lo largo del tiempo. Puede utilizar un gráfico de comportamiento con Amazon Detective para examinar los intentos de inicio de sesión fallidos, las llamadas sospechosas a la API y acciones similares. Para obtener más información, consulte Datos en un gráfico de comportamiento en la documentación de Detective.

sistema big-endian

Un sistema que almacena primero el byte más significativo. Consulte también endianidad.

clasificación binaria

Un proceso que predice un resultado binario (una de las dos clases posibles). Por ejemplo, es posible que su modelo de ML necesite predecir problemas como “¿Este correo electrónico es spam o no es spam?” o “¿Este producto es un libro o un automóvil?”.

filtro de floración

Estructura de datos probabilística y eficiente en términos de memoria que se utiliza para comprobar si un elemento es miembro de un conjunto.

blue/green despliegue

Estrategia de implementación en la que se crean dos entornos separados, pero idénticos. La versión actual de la aplicación se ejecuta en un entorno (azul) y la nueva versión de la aplicación se ejecuta en el otro entorno (verde). Esta estrategia lo ayuda a hacer reversiones rápidas con un impacto mínimo.

bot

Aplicación de software que ejecuta tareas automatizadas a través de Internet y simula la actividad o interacción humana. Algunos bots son útiles o beneficiosos, como los rastreadores web que indexan la información de Internet. Otros bots, conocidos como bots maliciosos, tienen como objetivo causar interrupciones o daños a personas u organizaciones.

botnet

Redes de bots infectadas por malware y que están bajo el control de una sola parte, conocida como pastor de bots u operador de bots. Las botnets son el mecanismo más conocido para escalar los bots y su impacto.

branch

Área contenida de un repositorio de código. La primera rama que se crea en un repositorio es la rama principal. Puede crear una rama nueva a partir de una rama existente y, a continuación, desarrollar características o corregir errores en la rama nueva. Una rama que se genera para crear una característica se denomina comúnmente rama de característica. Cuando la característica se encuentra lista para su lanzamiento, se vuelve a combinar la rama de característica con la rama principal. Para obtener más información, consulte Acerca de las sucursales (GitHub documentación).

acceso de emergencia

En circunstancias excepcionales y mediante un proceso aprobado, es una forma rápida de que un usuario pueda acceder a un Cuenta de AWS sitio al que normalmente no tiene permisos de acceso. Para obtener más información, consulte el indicador de implementación de procedimientos rompe-cristales en la AWS Well-Architected guía.

estrategia de implementación sobre infraestructura existente

La infraestructura existente en su entorno. Al adoptar una estrategia de implementación sobre infraestructura existente para una arquitectura de sistemas, se diseña la arquitectura en función de las limitaciones de los sistemas y la infraestructura actuales. Si está ampliando la infraestructura existente, puede combinar las estrategias de implementación sobre infraestructuras existentes y de implementación desde cero.

caché de búfer

El área de memoria donde se almacenan los datos a los que se accede con más frecuencia.

capacidad empresarial

Lo que hace una empresa para generar valor (por ejemplo, ventas, servicio al cliente o marketing). Las arquitecturas de microservicios y las decisiones de desarrollo pueden estar impulsadas por las capacidades empresariales. Para obtener más información, consulte la sección Organizado en torno a las capacidades empresariales del documento técnico Ejecutar microservicios en contenedores en AWS.

planificación de la continuidad del negocio (BCP)

Plan que aborda el posible impacto de un evento disruptivo, como una migración a gran escala en las operaciones y permite a la empresa reanudar las operaciones rápidamente.

C

CAF

Consulte AWS Cloud Adoption Framework.

implementación canario

Lanzamiento lento e incremental de una versión para los usuarios finales. Cuando tenga mayor confianza en la nueva versión, la implementa y reemplaza la versión actual en su totalidad.

CCoE

Consulte Centro de excelencia en la nube.

CDC

Consulte captura de datos de cambios.

captura de datos de cambio (CDC)

Proceso de seguimiento de los cambios en un origen de datos, como una tabla de base de datos, y registro de los metadatos relacionados con el cambio. Puede utilizar los CDC para diversos fines, como auditar o replicar los cambios en un sistema de destino para mantener la sincronización.

ingeniería del caos

Introducción intencionada de fallos o eventos disruptivos para poner a prueba la resiliencia de un sistema. Puedes usar AWS Fault Injection Service (AWS FIS) para realizar experimentos que estresen tus AWS cargas de trabajo y evalúen su respuesta.

CI/CD

Consulte integración continua y entrega continua.

clasificación

Un proceso de categorización que permite generar predicciones. Los modelos de ML para problemas de clasificación predicen un valor discreto. Los valores discretos siempre son distintos entre sí. Por ejemplo, es posible que un modelo necesite evaluar si hay o no un automóvil en una imagen.

Desarrollador ciudadano

Un usuario empresarial que crea aplicaciones de IA utilizando plataformas sin code/low código sin conocimientos técnicos especializados.

cifrado del cliente

Cifrado de datos localmente, antes de que el objetivo los Servicio de AWS reciba.

Centro de excelencia en la nube (CCoE)

Equipo multidisciplinario que impulsa los esfuerzos de adopción de la nube en toda la organización, incluido el desarrollo de las prácticas recomendadas en la nube, la movilización de recursos, el establecimiento de plazos de migración y la dirección de la organización durante las transformaciones a gran escala. Para obtener más información, consulte las publicaciones de CCoE en el blog de estrategia Nube de AWS empresarial.

computación en la nube

La tecnología en la nube que se utiliza normalmente para la administración de dispositivos de IoT y el almacenamiento de datos de forma remota. La computación en la nube suele estar relacionada con la tecnología de computación de periferia.

modelo operativo en la nube

En una organización de TI, el modelo operativo que se utiliza para crear, madurar y optimizar uno o más entornos de nube. Para obtener más información, consulte Creación de su modelo operativo de nube.

etapas de adopción de la nube

Las siguientes son las cuatro fases por las que suelen pasar las empresas cuando migran a la Nube de AWS:

- Proyecto: ejecución de algunos proyectos relacionados con la nube con fines de prueba de concepto y aprendizaje
- Fundamento: realización de inversiones fundamentales para escalar la adopción de la nube (p. ej., crear una zona de aterrizaje, definir un CCoE, establecer un modelo de operaciones)
- Migración: migración de aplicaciones individuales
- Re-invention — Optimizar los productos y servicios e innovar en la nube

Stephen Orban definió estas etapas en la entrada del blog The Journey Toward Cloud-First & the Stages of Adoption del blog Nube de AWS Enterprise Strategy. Para obtener información sobre su relación con la estrategia de AWS migración, consulte la guía de preparación para la migración.

CMDB

Consulte base de datos de administración de configuración.

repositorio de código

Una ubicación donde el código fuente y otros activos, como documentación, muestras y scripts, se almacenan y actualizan mediante procesos de control de versiones. Algunos repositorios en la nube comunes son GitHub o Bitbucket Cloud. Cada versión del código se denomina rama. En una estructura de microservicios, cada repositorio se encuentra dedicado a una única funcionalidad. Una sola CI/CD canalización puede utilizar varios repositorios.

caché en frío

Una caché de búfer que está vacía no está bien poblada o contiene datos obsoletos o irrelevantes. Esto afecta al rendimiento, ya que la instancia de la base de datos debe leer desde la memoria principal o el disco, lo que es más lento que leer desde la memoria caché del búfer.

datos fríos

Datos a los que se accede con poca frecuencia y que suelen ser históricos. Al consultar este tipo de datos, normalmente se aceptan consultas lentas. Trasladar estos datos a niveles o clases de almacenamiento de menor rendimiento y menos costosos puede reducir los costos.

visión artificial (CV)

Campo de la IA que utiliza el machine learning para analizar y extraer información de formatos visuales, como imágenes y videos digitales. Por ejemplo, Amazon SageMaker AI proporciona algoritmos de procesamiento de imágenes para CV.

deriva de configuración

En el caso de una carga de trabajo, un cambio en la configuración con respecto al estado esperado. Podría provocar que la carga de trabajo deje de cumplir las normas y, por lo general, es gradual e involuntaria.

base de datos de administración de configuración (CMDB)

Repositorio que almacena y administra información sobre una base de datos y su entorno de TI, incluidos los componentes de hardware y software y sus configuraciones. Por lo general, los datos de una CMDB se utilizan en la etapa de detección y análisis de la cartera de productos durante la migración.

paquete de conformidad

Un conjunto de AWS Config reglas y medidas correctivas que puede reunir para personalizar sus controles de conformidad y seguridad. Puede implementar un paquete de conformidad como una entidad única en una región Cuenta de AWS y, o en una organización, mediante una plantilla YAML. Para obtener más información, consulta los paquetes de conformidad en la documentación. AWS Config

integración y entrega continuas () CI/CD

El proceso de automatización de las etapas de origen, creación, prueba, puesta en escena y producción del proceso de publicación del software. CI/CD se describe comúnmente como una canalización. CI/CD puede ayudarlo a automatizar los procesos, mejorar la productividad, mejorar

la calidad del código y entregar más rápido. Para obtener más información, consulte Beneficios de la entrega continua. CD también puede significar implementación continua. Para obtener más información, consulte Entrega continua frente a implementación continua.

CV

Consulte visión artificial.

D

datos en reposo

Datos que están estacionarios en la red, como los datos que se encuentran almacenados.

clasificación de datos

Un proceso para identificar y clasificar los datos de su red en función de su importancia y sensibilidad. Es un componente fundamental de cualquier estrategia de administración de riesgos de ciberseguridad porque lo ayuda a determinar los controles de protección y retención adecuados para los datos. La clasificación de los datos es un componente del pilar de seguridad del AWS Well-Architected Framework. Para obtener más información, consulte Clasificación de datos.

deriva de datos

Una variación significativa entre los datos de producción y los datos que se utilizaron para entrenar un modelo de machine learning, o un cambio significativo en los datos de entrada a lo largo del tiempo. La deriva de datos puede reducir la calidad, la precisión y la imparcialidad generales de las predicciones de los modelos de machine learning.

datos en tránsito

Datos que se mueven de forma activa por la red, por ejemplo, entre los recursos de la red.

malla de datos

Marco de arquitectura que proporciona una propiedad de datos distribuida y descentralizada con una administración y una gobernanza centralizadas.

minimización de datos

El principio de recopilar y procesar solo los datos estrictamente necesarios. Practicar la minimización de los datos Nube de AWS puede reducir los riesgos de privacidad, los costos y la huella de carbono de la analítica.

perímetro de datos

Un conjunto de barreras preventivas en su AWS entorno que ayudan a garantizar que solo las identidades confiables accedan a los recursos confiables desde las redes esperadas. Para obtener más información, consulte Crear un perímetro de datos sobre. AWS

preprocesamiento de datos

Transformar los datos sin procesar en un formato que su modelo de ML pueda analizar fácilmente. El preprocesamiento de datos puede implicar eliminar determinadas columnas o filas y corregir los valores faltantes, incoherentes o duplicados.

procedencia de los datos

El proceso de rastrear el origen y el historial de los datos a lo largo de su ciclo de vida, por ejemplo, la forma en que se generaron, transmitieron y almacenaron los datos.

titular de los datos

Persona cuyos datos se recopilan y procesan.

almacenamiento de datos

Sistema de administración de datos que respalda la inteligencia empresarial, como los análisis. Los almacenes de datos suelen contener grandes cantidades de datos históricos y, por lo general, se utilizan para las consultas y los análisis.

lenguaje de definición de datos (DDL)

Instrucciones o comandos para crear o modificar la estructura de tablas y objetos de una base de datos.

lenguaje de manipulación de datos (DML)

Instrucciones o comandos para modificar (insertar, actualizar y eliminar) la información de una base de datos.

DDL

Consulte lenguaje de definición de bases de datos.

conjunto profundo

Combinar varios modelos de aprendizaje profundo para la predicción. Puede utilizar conjuntos profundos para obtener una predicción más precisa o para estimar la incertidumbre de las predicciones.

aprendizaje profundo

Un subcampo del ML que utiliza múltiples capas de redes neuronales artificiales para identificar el mapeo entre los datos de entrada y las variables objetivo de interés.

defensa en profundidad

Un enfoque de seguridad de la información en el que se distribuyen cuidadosamente una serie de mecanismos y controles de seguridad en una red informática para proteger la confidencialidad, la integridad y la disponibilidad de la red y de los datos que contiene. Al adoptar esta estrategia AWS, se añaden varios controles en diferentes capas de la AWS Organizations estructura para ayudar a proteger los recursos. Por ejemplo, un enfoque de defensa en profundidad podría combinar la autenticación multifactor, la segmentación de la red y el cifrado.

administrador delegado

En AWS Organizations, un servicio compatible puede registrar una cuenta de AWS miembro para administrar las cuentas de la organización y gestionar los permisos de ese servicio. Esta cuenta se denomina administrador delegado para ese servicio. Para obtener más información y una lista de servicios compatibles, consulte Servicios que funcionan con AWS Organizations en la documentación de AWS Organizations .

Implementación

El proceso de hacer que una aplicación, características nuevas o correcciones de código se encuentren disponibles en el entorno de destino. La implementación abarca implementar cambios en una base de código y, a continuación, crear y ejecutar esa base en los entornos de la aplicación.

entorno de desarrollo

Consulte entorno.

control de detección

Un control de seguridad que se ha diseñado para detectar, registrar y alertar después de que se produzca un evento. Estos controles son una segunda línea de defensa, ya que lo advierten sobre los eventos de seguridad que han eludido los controles preventivos establecidos. Para obtener más información, consulte Controles de detección en Implementación de controles de seguridad en AWS.

asignación de flujos de valor para el desarrollo (DVSM)

Proceso que se utiliza para identificar y priorizar las restricciones que afectan negativamente a la velocidad y la calidad en el ciclo de vida del desarrollo de software. DVSM amplía el proceso de asignación del flujo de valor diseñado originalmente para las prácticas de fabricación ajustada. Se centra en los pasos y los equipos necesarios para crear y transferir valor a través del proceso de desarrollo de software.

gemelo digital

Representación virtual de un sistema del mundo real, como un edificio, una fábrica, un equipo industrial o una línea de producción. Los gemelos digitales son compatibles con el mantenimiento predictivo, la supervisión remota y la optimización de la producción.

tabla de dimensiones

En un esquema en estrella, tabla más pequeña que contiene los atributos de datos sobre los datos cuantitativos en una tabla de hechos. Los atributos de la tabla de dimensiones suelen ser campos de texto o números discretos que se comportan como texto. Estos atributos se suelen utilizar para restringir consultas, filtrarlas y etiquetar los conjuntos de resultados.

desastre

Un evento que impide que una carga de trabajo o un sistema cumplan sus objetivos empresariales en su ubicación principal de implementación. Estos eventos pueden ser desastres naturales, fallos técnicos o el resultado de acciones humanas, como una configuración incorrecta involuntaria o un ataque de malware.

recuperación de desastres (DR)

Estrategia y proceso que utiliza para minimizar el tiempo de inactividad y la pérdida de datos a causa de un desastre. Para obtener más información, consulte Recuperación de cargas de trabajo ante desastres en AWS: Recuperación en la nube en el AWS Well-Architected marco.

DML

Consulte lenguaje de manipulación de bases de datos.

diseño basado en el dominio

Un enfoque para desarrollar un sistema de software complejo mediante la conexión de sus componentes a dominios en evolución, o a los objetivos empresariales principales, a los que sirve cada componente. Eric Evans introdujo este concepto en su libro Domain-Driven Design: Tackling Complexity in the Heart of Software (Boston: Addison-Wesley Professional, 2003). Para

obtener información sobre cómo utilizar el diseño basado en dominios con el patrón de higos estranguladores, consulte Modernización gradual de los servicios web antiguos de ASP.NET Microsoft (ASMX) mediante contenedores y Amazon API Gateway.

DR

Consulte recuperación ante desastres.

Detección de desviaciones

Seguimiento de las desviaciones con respecto a una configuración con línea de base. Por ejemplo, puedes usarlo AWS CloudFormation para detectar desviaciones en los recursos del sistema o puedes usarlo AWS Control Tower para detectar cambios en tu landing zone que puedan afectar al cumplimiento de los requisitos de gobierno.

DVSM

Consulte asignación de flujos de valor para el desarrollo.

E

EDA

Consulte análisis de datos de tipo exploratorio.

EDI

Consulte intercambio electrónico de datos.

computación en la periferia

La tecnología que aumenta la potencia de cálculo de los dispositivos inteligentes en la periferia de una red de IoT. En comparación con la computación en la nube, la computación de periferia puede reducir la latencia de la comunicación y mejorar el tiempo de respuesta.

intercambio electrónico de datos (EDI)

Intercambio automatizado de documentos comerciales entre organizaciones. Para más información, consulte ¿Qué es el intercambio electrónico de datos?

cifrado

Proceso de computación que transforma datos de texto plano, que son legibles por humanos, en texto cifrado.

clave de cifrado

Cadena criptográfica de bits aleatorios que se genera mediante un algoritmo de cifrado. Las claves pueden variar en longitud y cada una se ha diseñado para ser impredecible y única.

endianidad

El orden en el que se almacenan los bytes en la memoria del ordenador. Big-endian los sistemas almacenan primero el byte más significativo. Little-endian los sistemas almacenan primero el byte menos significativo.

punto de conexión

Consulte punto de conexión de servicio.

servicio de punto de conexión

Servicio que puede alojar en una nube privada virtual (VPC) para compartir con otros usuarios. Puede crear un servicio de punto final con AWS PrivateLink entidades principales Cuentas de AWS o AWS Identity and Access Management (de IAM) y conceder permisos a ellas. Estas cuentas o entidades principales pueden conectarse a su servicio de punto de conexión de forma privada mediante la creación de puntos de conexión de VPC de interfaz. Para obtener más información, consulte Creación de un servicio de punto de conexión en la documentación de Amazon Virtual Private Cloud (Amazon VPC).

planificación de recursos empresariales (ERP)

Sistema que automatiza y administra los procesos empresariales clave (como la contabilidad, MES y la administración de proyectos) de una empresa.

cifrado de sobre

El proceso de cifrar una clave de cifrado con otra clave de cifrado. Para obtener más información, consulte el cifrado de sobres en la documentación de AWS Key Management Service (AWS KMS).

entorno

Una instancia de una aplicación en ejecución. Los siguientes son los tipos de entornos más comunes en la computación en la nube:

- entorno de desarrollo: instancia de una aplicación en ejecución que solo se encuentra disponible para el equipo principal responsable del mantenimiento de la aplicación. Los

entornos de desarrollo se utilizan para probar los cambios antes de promocionarlos a los entornos superiores. Este tipo de entorno a veces se denomina entorno de prueba.

- entornos inferiores: todos los entornos de desarrollo de una aplicación, como los que se utilizan para las compilaciones y pruebas iniciales.
- entorno de producción: instancia de una aplicación en ejecución a la que pueden acceder los usuarios finales. En un CI/CD proceso, el entorno de producción es el último entorno de implementación.
- entornos superiores: todos los entornos a los que pueden acceder usuarios que no sean del equipo de desarrollo principal. Esto puede incluir un entorno de producción, entornos de preproducción y entornos para las pruebas de aceptación por parte de los usuarios.

epopeya

En las metodologías ágiles, son categorías funcionales que ayudan a organizar y priorizar el trabajo. Las epopeyas brindan una descripción detallada de los requisitos y las tareas de implementación. Por ejemplo, las epopeyas AWS de seguridad de CAF incluyen la gestión de identidades y accesos, los controles de detección, la seguridad de la infraestructura, la protección de datos y la respuesta a incidentes. Para obtener más información sobre las epopeyas en la estrategia de migración de AWS , consulte la Guía de implementación del programa.

ERP

Consulte planificación de recursos empresariales.

análisis de datos de tipo exploratorio (EDA)

El proceso de analizar un conjunto de datos para comprender sus características principales. Se recopilan o agregan datos y, a continuación, se realizan las investigaciones iniciales para encontrar patrones, detectar anomalías y comprobar las suposiciones. El EDA se realiza mediante el cálculo de estadísticas resumidas y la creación de visualizaciones de datos.

F

tabla de hechos

Tabla central de un esquema en estrella. Almacena datos cuantitativos sobre operaciones empresariales. Por lo general, una tabla de hechos contiene dos tipos de columnas: las que contienen medidas y las que contienen una clave externa para una tabla de dimensiones.

Fail Fast

Filosofía que utiliza pruebas frecuentes e incrementales para reducir el ciclo de vida del desarrollo. Es una parte fundamental de los enfoques ágiles.

límite de aislamiento de errores

En el Nube de AWS, un límite, como una zona de disponibilidad Región de AWS, un plano de control o un plano de datos, que limita el efecto de una falla y ayuda a mejorar la resiliencia de las cargas de trabajo. Para más información, consulte AWS Fault Isolation Boundaries.

rama de característica

Consulte rama.

características

Los datos de entrada que se utilizan para hacer una predicción. Por ejemplo, en un contexto de fabricación, las características pueden ser imágenes que se capturan periódicamente desde la línea de fabricación.

importancia de las características

La importancia que tiene una característica para las predicciones de un modelo. Por lo general, esto se expresa como una puntuación numérica que se puede calcular mediante diversas técnicas, como las explicaciones aditivas de Shapley (SHAP) y los gradientes integrados. Para obtener más información, consulte Interpretabilidad del modelo de aprendizaje automático con. AWS

transformación de funciones

Optimizar los datos para el proceso de ML, lo que incluye enriquecer los datos con fuentes adicionales, escalar los valores o extraer varios conjuntos de información de un solo campo de datos. Esto permite que el modelo de ML se beneficie de los datos. Por ejemplo, si divide la fecha del “27 de mayo de 2021 00:15:37” en “jueves”, “mayo”, “2021” y “15”, puede ayudar al algoritmo de aprendizaje a aprender patrones matizados asociados a los diferentes componentes de los datos.

peticiones con pocos pasos

Proporcionar a un LLM una pequeña cantidad de ejemplos que demuestren la tarea y el resultado deseado antes de pedirle que lleve a cabo una tarea similar. Esta técnica es una aplicación del aprendizaje contextual, en el que los modelos aprenden a partir de ejemplos (tomas) integrados en las instrucciones. Few-shot Las indicaciones pueden ser eficaces para tareas que requieren

un formato, un razonamiento o un conocimiento del dominio específicos. Consulte también peticiones desde cero.

FGAC

Consulte control de acceso detallado.

control de acceso preciso (FGAC)

El uso de varias condiciones que tienen por objetivo permitir o denegar una solicitud de acceso.

migración relámpago

Método de migración de bases de datos que utiliza la replicación continua de datos mediante la captura de datos de cambio para migrar los datos en el menor tiempo posible, en lugar de utilizar un enfoque gradual. El objetivo es reducir al mínimo el tiempo de inactividad.

FM

Consulte modelo fundacional.

Modelo fundacional (FM)

Gran red neuronal de aprendizaje profundo que se entrenó con conjuntos de datos masivos de datos generalizados y no etiquetados. Los FM pueden hacer una amplia variedad de tareas generales, como comprender el lenguaje, generar texto e imágenes y conversar en lenguaje natural. Para más información, consulte ¿Qué son los modelos fundacionales?

Puerta de enlace FM

Un intermediario centralizado que controla y normaliza el acceso a los modelos básicos. También se conoce como puerta de enlace LLM.

G

IA generativa

Subconjunto de modelos de IA que se entrenaron con grandes cantidades de datos y que pueden utilizar una simple petición de texto para crear contenido y artefactos nuevos, como imágenes, videos, texto y audio. Para más información, consulte ¿Qué es la IA generativa?

bloqueo geográfico

Consulte restricciones geográficas.

restricciones geográficas (bloqueo geográfico)

En Amazon CloudFront, una opción para impedir que los usuarios de países específicos accedan a las distribuciones de contenido. Puede utilizar una lista de permitidos o bloqueados para especificar los países aprobados y prohibidos. Para obtener más información, consulta Restringir la distribución geográfica del contenido en la CloudFront documentación.

Flujo de trabajo de Gitflow

Un enfoque en el que los entornos inferiores y superiores utilizan diferentes ramas en un repositorio de código fuente. El flujo de trabajo de Gitflow se considera heredado, mientras que el flujo de trabajo basado en enlaces troncales es el enfoque moderno preferido.

imagen dorada

Instantánea de un sistema o software que se usa como plantilla para implementar nuevas instancias de ese sistema o software. Por ejemplo, en la fabricación, una imagen dorada se puede utilizar para aprovisionar software en varios dispositivos y ayuda a mejorar la velocidad, la escalabilidad y la productividad de las operaciones de fabricación de dispositivos.

estrategia de implementación desde cero

La ausencia de infraestructura existente en un entorno nuevo. Al adoptar una estrategia de implementación desde cero para una arquitectura de sistemas, puede seleccionar todas las tecnologías nuevas sin que estas deban ser compatibles con una infraestructura existente, lo que también se conoce como implementación sobre infraestructura existente. Si está ampliando la infraestructura existente, puede combinar las estrategias de implementación sobre infraestructuras existentes y de implementación desde cero.

barrera de protección

Una regla de alto nivel que ayuda a regular los recursos, las políticas y la conformidad en todas las unidades organizativas (OU). Las barreras de protección preventivas aplican políticas para garantizar la alineación con los estándares de conformidad. Se implementan mediante políticas de control de servicios y límites de permisos de IAM. Las barreras de protección de detección detectan las vulneraciones de las políticas y los problemas de conformidad, y generan alertas para su corrección. Se implementan mediante Amazon AWS Config AWS Security Hub CSPM GuardDuty AWS Trusted Advisor, Amazon Inspector y AWS Lambda cheques personalizados.

barandas (AI)

Mecanismos de seguridad que filtran, validan y restringen las entradas y salidas de los agentes para ayudar a garantizar un comportamiento responsable y seguro de la IA.

H

HA

Consulte alta disponibilidad.

migración heterogénea de bases de datos

Migración de la base de datos de origen a una base de datos de destino que utilice un motor de base de datos diferente (por ejemplo, de Oracle a Amazon Aurora). La migración heterogénea suele ser parte de un esfuerzo de rediseño de la arquitectura y convertir el esquema puede ser una tarea compleja. AWS ofrece AWS SCT, lo cual ayuda con las conversiones de esquemas.

alta disponibilidad (HA)

La capacidad de una carga de trabajo para funcionar de forma continua, sin intervención, en caso de desafíos o desastres. Los sistemas de alta disponibilidad están diseñados para realizar una conmutación por error automática, ofrecer un rendimiento de alta calidad de forma constante y gestionar diferentes cargas y fallos con un impacto mínimo en el rendimiento.

modernización histórica

Un enfoque utilizado para modernizar y actualizar los sistemas de tecnología operativa (TO) a fin de satisfacer mejor las necesidades de la industria manufacturera. Un histórico es un tipo de base de datos que se utiliza para recopilar y almacenar datos de diversas fuentes en una fábrica.

datos de reserva

Parte de los datos históricos etiquetados que se ocultan de un conjunto de datos que se utiliza para entrenar un modelo de machine learning. Puede utilizar los datos de reserva para evaluar el rendimiento del modelo mediante la comparación de las predicciones del modelo con los datos de reserva.

human-in-the-loop (HiTL)

Un patrón de flujo de trabajo en el que la ejecución de los agentes se detiene para su revisión y aprobación por parte de una persona en los puntos de decisión críticos.

migración homogénea de bases de datos

Migración de la base de datos de origen a una base de datos de destino que comparte el mismo motor de base de datos (por ejemplo, Microsoft SQL Server a Amazon RDS para SQL Server).

La migración homogénea suele formar parte de un esfuerzo para volver a alojar o redefinir la plataforma. Puede utilizar las utilidades de bases de datos nativas para migrar el esquema.

datos recientes

Datos a los que se accede con frecuencia, como datos en tiempo real o datos traslacionales recientes. Por lo general, estos datos requieren un nivel o una clase de almacenamiento de alto rendimiento para proporcionar respuestas rápidas a las consultas.

hotfix

Una solución urgente para un problema crítico en un entorno de producción. Debido a su urgencia, una revisión suele realizarse fuera del flujo de trabajo habitual de las DevOps versiones.

periodo de hiperatención

Periodo, inmediatamente después de la transición, durante el cual un equipo de migración administra y monitorea las aplicaciones migradas en la nube para solucionar cualquier problema. Por lo general, este periodo dura de 1 a 4 días. Al final del periodo de hiperatención, el equipo de migración suele transferir la responsabilidad de las aplicaciones al equipo de operaciones en la nube.

I

IaC

Consulte infraestructura como código.

políticas basadas en identidades

Política asociada a uno o más directores de IAM que define sus permisos en el entorno. Nube de AWS

aplicación inactiva

Aplicación que utiliza un promedio de CPU y memoria de entre 5 y 20 por ciento durante un periodo de 90 días. En un proyecto de migración, es habitual retirar estas aplicaciones o mantenerlas en las instalaciones.

IIoT

Consulte Internet de las cosas industrial.

infraestructura inmutable

Modelo que implementa una nueva infraestructura para las cargas de trabajo de producción en lugar de actualizar o modificar la infraestructura existente o aplicarle revisiones. Las infraestructuras inmutables son de manera intrínseca más coherentes, fiables y predecibles que las infraestructuras mutables. Para obtener más información, consulte las mejores prácticas del Framework para implementar con una infraestructura inmutable. AWS Well-Architected

VPC entrante (de entrada)

En una arquitectura de AWS cuentas múltiples, una VPC que acepta, inspecciona y enruta las conexiones de red desde fuera de una aplicación. La Arquitectura de referencia de seguridad de AWS recomienda configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

migración gradual

Estrategia de transición en la que se migra la aplicación en partes pequeñas en lugar de realizar una transición única y completa. Por ejemplo, puede trasladar inicialmente solo unos pocos microservicios o usuarios al nuevo sistema. Tras comprobar que todo funciona correctamente, puede trasladar microservicios o usuarios adicionales de forma gradual hasta que pueda retirar su sistema heredado. Esta estrategia reduce los riesgos asociados a las grandes migraciones.

Industria 4.0

Un término que Klaus Schwab introdujo en 2016 para referirse a la modernización de los procesos de fabricación mediante avances en la conectividad, los datos en tiempo real, la automatización, el análisis y. AI/ML

infraestructura

Todos los recursos y activos que se encuentran en el entorno de una aplicación.

infraestructura como código (IaC)

Proceso de aprovisionamiento y administración de la infraestructura de una aplicación mediante un conjunto de archivos de configuración. La IaC se ha diseñado para ayudarlo a centralizar la administración de la infraestructura, estandarizar los recursos y escalar con rapidez a fin de que los entornos nuevos sean repetibles, fiables y consistentes.

Internet de las cosas industrial (IIoT)

El uso de sensores y dispositivos conectados a Internet en los sectores industriales, como el productivo, el eléctrico, el automotriz, el sanitario, el de las ciencias de la vida y el de la

agricultura. Para obtener más información, consulte Creación de una estrategia de transformación digital del Internet de las cosas industrial (IIoT).

VPC de inspección

En una arquitectura de AWS cuentas múltiples, una VPC centralizada que gestiona las inspecciones del tráfico de red entre las VPC (iguales o Regiones de AWS diferentes), Internet y las redes locales. La Arquitectura de referencia de seguridad de AWS recomienda configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

Internet de las cosas (IoT)

Red de objetos físicos conectados con sensores o procesadores integrados que se comunican con otros dispositivos y sistemas a través de Internet o de una red de comunicación local. Para obtener más información, consulte ¿Qué es IoT?.

interpretabilidad

Característica de un modelo de machine learning que describe el grado en que un ser humano puede entender cómo las predicciones del modelo dependen de sus entradas. Para obtener más información, consulte Interpretabilidad del modelo de aprendizaje automático con. AWS

IoT

Consulte Internet de las cosas.

biblioteca de información de TI (ITIL)

Conjunto de prácticas recomendadas para ofrecer servicios de TI y alinearlos con los requisitos empresariales. La ITIL proporciona la base para la ITSM.

administración de servicios de TI (ITSM)

Actividades asociadas con el diseño, la implementación, la administración y el soporte de los servicios de TI para una organización. Para obtener información sobre la integración de las operaciones en la nube con las herramientas de ITSM, consulte la Guía de integración de operaciones.

ITIL

Consulte biblioteca de información de TI.

ITSM

Consulte administración de servicios de TI.

L

control de acceso basado en etiquetas (LBAC)

Una implementación del control de acceso obligatorio (MAC) en la que a los usuarios y a los propios datos se les asigna explícitamente un valor de etiqueta de seguridad. La intersección entre la etiqueta de seguridad del usuario y la etiqueta de seguridad de los datos determina qué filas y columnas puede ver el usuario.

zona de aterrizaje

Una landing zone es un AWS entorno multicuenta bien diseñado, escalable y seguro. Este es un punto de partida desde el cual las empresas pueden lanzar e implementar rápidamente cargas de trabajo y aplicaciones con confianza en su entorno de seguridad e infraestructura. Para obtener más información sobre las zonas de aterrizaje, consulte Configuración de un entorno de AWS seguro y escalable con varias cuentas.

modelo de lenguaje de gran tamaño (LLM)

Modelo de IA de aprendizaje profundo que se entrenó previamente con una gran cantidad de datos. Un LLM puede llevar a cabo varias tareas, como responder preguntas, resumir documentos, traducir textos a otros idiomas y completar oraciones. Para más información, consulte ¿Qué es un LLM (modelo de lenguaje de gran tamaño)?

migración grande

Migración de 300 servidores o más.

LBAC

Consulte control de acceso basado en etiquetas.

privilegio mínimo

La práctica recomendada de seguridad que consiste en conceder los permisos mínimos necesarios para realizar una tarea. Para obtener más información, consulte Aplicar permisos de privilegio mínimo en la documentación de IAM.

migrar mediante lift-and-shift

Consulte Las 7 R.

sistema little-endian

Un sistema que almacena primero el byte menos significativo. Consulte también endianidad.

LLM

Consulte modelo de lenguaje de gran tamaño.

entornos inferiores

Consulte entorno.

M

machine learning (ML)

Un tipo de inteligencia artificial que utiliza algoritmos y técnicas para el reconocimiento y el aprendizaje de patrones. El ML analiza y aprende de los datos registrados, como los datos del Internet de las cosas (IoT), para generar un modelo estadístico basado en patrones. Para más información, consulte Machine learning.

rama principal

Consulte rama.

malware

Software diseñado para comprometer la seguridad o la privacidad de la computadora. El malware podría interrumpir los sistemas informáticos, filtrar información confidencial u obtener acceso no autorizado. Algunos ejemplos de malware son los virus, los gusanos, el ransomware, los troyanos, el spyware y los registradores de pulsaciones de teclas.

Servicios administrados

Servicios de AWS en el que AWS opera la capa de infraestructura, el sistema operativo y las plataformas, y se accede a los puntos finales para almacenar y recuperar datos. Amazon Simple Storage Service (Amazon S3) y Amazon DynamoDB son ejemplos de servicios administrados. También se conocen como servicios abstractos.

sistema de ejecución de fabricación (MES)

Sistema de software para seguir, supervisar, documentar y controlar los procesos de producción que convierten las materias primas en productos acabados en la zona de producción.

MAP

Consulte Programa de aceleración de la migración.

MCP

Consulte Model Context Protocol.

Protocolo de contexto para modelos (MCP)

Un protocolo sin estado para la comunicación entre el agente y la herramienta.

Servidor MCP

Un servicio que expone una o más herramientas a través del protocolo Model Context.

mecanismo

Proceso completo mediante el que se crea una herramienta, se impulsa su adopción y, a continuación, se inspeccionan los resultados para hacer ajustes. Un mecanismo es un ciclo que se refuerza y mejora por sí mismo a medida que funciona. Para obtener más información, consulte Creación de mecanismos en el AWS Well-Architected marco.

cuenta de miembro

Todas las Cuentas de AWS demás cuentas, excepto la de administración, que forman parte de una organización AWS Organizations. Una cuenta no puede pertenecer a más de una organización a la vez.

MES

Consulte sistema de ejecución de fabricación.

Message Queuing Telemetry Transport (MQTT)

Un protocolo de comunicación ligero de máquina a máquina (M2M), basado en el publish/subscribepatrón, para dispositivos de IoT con recursos limitados.

microservicio

Un servicio pequeño e independiente que se comunica a través de API bien definidas y que, por lo general, es propiedad de equipos pequeños e independientes. Por ejemplo, un sistema de seguros puede incluir microservicios que se adapten a las capacidades empresariales, como las de ventas o marketing, o a subdominios, como las de compras, reclamaciones o análisis. Los beneficios de los microservicios incluyen la agilidad, la escalabilidad flexible, la facilidad de implementación, el código reutilizable y la resiliencia. Para obtener más información, consulte Integrar microservicios mediante servicios sin servidor. AWS

arquitectura de microservicios

Un enfoque para crear una aplicación con componentes independientes que ejecutan cada proceso de la aplicación como un microservicio. Estos microservicios se comunican a través de una interfaz bien definida mediante API ligeras. Cada microservicio de esta arquitectura se puede actualizar, implementar y escalar para satisfacer la demanda de funciones específicas de una aplicación. Para obtener más información, consulte Implementación de microservicios en. AWS

Programa de aceleración de la migración (MAP)

Un AWS programa que proporciona soporte de consultoría, formación y servicios para ayudar a las organizaciones a crear una base operativa sólida para migrar a la nube y para ayudar a compensar el costo inicial de las migraciones. El MAP incluye una metodología de migración para ejecutar las migraciones antiguas de forma metódica y un conjunto de herramientas para automatizar y acelerar los escenarios de migración más comunes.

migración a escala

Proceso de transferencia de la mayoría de la cartera de aplicaciones a la nube en oleadas, con más aplicaciones desplazadas a un ritmo más rápido en cada oleada. En esta fase, se utilizan las prácticas recomendadas y las lecciones aprendidas en las fases anteriores para implementar una fábrica de migración de equipos, herramientas y procesos con el fin de agilizar la migración de las cargas de trabajo mediante la automatización y la entrega ágil. Esta es la tercera fase de la estrategia de migración de AWS.

fábrica de migración

Cross-functional equipos que agilizan la migración de las cargas de trabajo mediante enfoques ágiles y automatizados. Los equipos de las fábricas de migración suelen estar compuestos por analistas y propietarios de operaciones, ingenieros de migración, desarrolladores y DevOps profesionales que trabajan a pasos agigantados. Entre el 20 y el 50 por ciento de la cartera de aplicaciones empresariales se compone de patrones repetidos que pueden optimizarse mediante un enfoque de fábrica. Para obtener más información, consulte la discusión sobre las fábricas de migración y la Guía de fábricas de migración a la nube en este contenido.

metadatos de migración

Información sobre la aplicación y el servidor que se necesita para completar la migración. Cada patrón de migración requiere un conjunto diferente de metadatos de migración. Algunos ejemplos de metadatos de migración son la subred de destino, el grupo de seguridad y AWS la cuenta.

patrón de migración

Tarea de migración repetible que detalla la estrategia de migración, el destino de la migración y la aplicación o el servicio de migración utilizados. Ejemplo: rehospede la migración a Amazon EC2 AWS con Application Migration Service.

Migration Portfolio Assessment (MPA)

Herramienta en línea que proporciona información a fin de validar los argumentos comerciales necesarios para migrar a la Nube de AWS. La MPA ofrece una evaluación detallada de la cartera (adecuación del tamaño de los servidores, precios, comparaciones del costo total de propiedad, análisis de los costos de migración), así como una planificación de la migración (análisis y recopilación de datos de aplicaciones, agrupación de aplicaciones, priorización de la migración y planificación de oleadas). La herramienta MPA (requiere iniciar sesión) está disponible de forma gratuita para todos los AWS consultores y consultores de los socios de APN.

Evaluación de la preparación para la migración (MRA)

Proceso que consiste en obtener información sobre el estado de preparación de una organización para la nube, identificar sus puntos fuertes y débiles y elaborar un plan de acción para cerrar las brechas identificadas mediante el AWS CAF. Para obtener más información, consulte la Guía de preparación para la migración. La MRA es la primera fase de la estrategia de migración de AWS.

estrategia de migración

Enfoque utilizado para migrar una carga de trabajo a la Nube de AWS. Para más información, consulte la entrada Las 7 R de este glosario y también Mobilize your organization to accelerate large-scale migrations.

ML

Consulte machine learning.

modernización

Transformar una aplicación obsoleta (antigua o monolítica) y su infraestructura en un sistema ágil, elástico y de alta disponibilidad en la nube para reducir los gastos, aumentar la eficiencia y aprovechar las innovaciones. Para más información, consulte Strategy for modernizing applications in the Nube de AWS.

evaluación de la preparación para la modernización

Evaluación que ayuda a determinar la preparación para la modernización de las aplicaciones de una organización; identifica los beneficios, los riesgos y las dependencias; y determina qué

tan bien la organización puede soportar el estado futuro de esas aplicaciones. El resultado de la evaluación es un esquema de la arquitectura objetivo, una hoja de ruta que detalla las fases de desarrollo y los hitos del proceso de modernización y un plan de acción para abordar las brechas identificadas. Para más información, consulte Evaluating modernization readiness for applications in the Nube de AWS.

aplicaciones monolíticas (monolitos)

Aplicaciones que se ejecutan como un único servicio con procesos estrechamente acoplados. Las aplicaciones monolíticas presentan varios inconvenientes. Si una característica de la aplicación experimenta un aumento en la demanda, se debe escalar toda la arquitectura. Agregar o mejorar las características de una aplicación monolítica también se vuelve más complejo a medida que crece la base de código. Para solucionar problemas con la aplicación, puede utilizar una arquitectura de microservicios. Para obtener más información, consulte Descomposición de monolitos en microservicios.

MPA

Consulte Migration Portfolio Assessment.

MQTT

Consulte Message Queuing Telemetry Transport.

clasificación multiclase

Un proceso que ayuda a generar predicciones para varias clases (predice uno de más de dos resultados). Por ejemplo, un modelo de ML podría preguntar “¿Este producto es un libro, un automóvil o un teléfono?” o “¿Qué categoría de productos es más interesante para este cliente?”.

infraestructura mutable

Modelo que actualiza y modifica la infraestructura actual para las cargas de trabajo de producción. Para mejorar la coherencia, la confiabilidad y la previsibilidad, el AWS Well-Architected Marco recomienda el uso de una infraestructura inmutable como práctica recomendada.

O

OAC

Consulte control de acceso de origen.

OAI

Consulte identidad de acceso de origen.

OCM

Consulte administración del cambio organizacional.

migración fuera de línea

Método de migración en el que la carga de trabajo de origen se elimina durante el proceso de migración. Este método implica un tiempo de inactividad prolongado y, por lo general, se utiliza para cargas de trabajo pequeñas y no críticas.

OI

Consulte integración de operaciones.

OLA

Consulte acuerdo de nivel operativo.

migración en línea

Método de migración en el que la carga de trabajo de origen se copia al sistema de destino sin que se desconecte. Las aplicaciones que están conectadas a la carga de trabajo pueden seguir funcionando durante la migración. Este método implica un tiempo de inactividad nulo o mínimo y, por lo general, se utiliza para cargas de trabajo de producción críticas.

OPC-UA

Consulte Open Process Communications: arquitectura unificada.

Comunicaciones de proceso abierto: arquitectura unificada () OPC-UA

Un protocolo de comunicación de máquina a máquina (M2M) para la automatización industrial. OPC-UA proporciona un estándar de interoperabilidad con esquemas de cifrado, autenticación y autorización de datos.

acuerdo de nivel operativo (OLA)

Acuerdo que aclara lo que los grupos de TI operativos se comprometen a ofrecerse entre sí, para respaldar un acuerdo de nivel de servicio (SLA).

revisión de la preparación operativa (ORR)

Lista de comprobación de preguntas y prácticas recomendadas asociadas que son útiles para comprender, evaluar, prevenir o reducir el alcance de los incidentes y posibles errores. Para

obtener más información, consulte las revisiones de preparación operativa (ORR) en el AWS Well-Architected marco.

tecnología operativa (TO)

Sistemas de hardware y software que funcionan con el entorno físico para controlar las operaciones, los equipos y la infraestructura industriales. En el sector de la fabricación, la integración de los sistemas de TO y tecnología de la información (TI) es un enfoque clave para las transformaciones de la industria 4.0.

integración de operaciones (OI)

Proceso de modernización de las operaciones en la nube, que implica la planificación de la preparación, la automatización y la integración. Para obtener más información, consulte la Guía de integración de las operaciones.

registro de seguimiento organizativo

Un registro creado por y AWS CloudTrail que registra todos los eventos Cuentas de AWS de una organización AWS Organizations. Este registro de seguimiento se crea en cada Cuenta de AWS que forma parte de la organización y realiza un seguimiento de la actividad en cada cuenta. Para obtener más información, consulte Crear un registro para una organización en la CloudTrail documentación.

administración del cambio organizacional (OCM)

Marco para administrar las transformaciones empresariales importantes y disruptivas desde la perspectiva de las personas, la cultura y el liderazgo. La OCM ayuda a las empresas a prepararse para nuevos sistemas y estrategias y a realizar la transición a ellos, al acelerar la adopción de cambios, abordar los problemas de transición e impulsar cambios culturales y organizacionales. En la estrategia de AWS migración, este marco se denomina aceleración de personal, debido a la velocidad de cambio que requieren los proyectos de adopción de la nube. Para obtener más información, consulte la Guía de OCM.

control de acceso de origen (OAC)

En CloudFront, una opción mejorada para restringir el acceso y proteger el contenido del Amazon Simple Storage Service (Amazon S3). El OAC admite todos los buckets de S3 Regiones de AWS, el cifrado del lado del servidor con AWS KMS (SSE-KMS) y `DELETE` las solicitudes `PUT` y dinámicas al bucket de S3.

identidad de acceso de origen (OAI)

En CloudFront, una opción para restringir el acceso y proteger el contenido de Amazon S3. Cuando utiliza OAI, CloudFront crea un principal con el que Amazon S3 puede autenticarse. Los directores autenticados solo pueden acceder al contenido de un bucket de S3 a través de una distribución específica. CloudFront Consulte también el OAC, que proporciona un control de acceso más detallado y mejorado.

ORR

Consulte revisión de la preparación operativa.

OT

Consulte tecnología operativa.

VPC saliente (de salida)

En una arquitectura de AWS cuentas múltiples, una VPC que gestiona las conexiones de red que se inician desde una aplicación. La Arquitectura de referencia de seguridad de AWS recomienda configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

P

límite de permisos

Una política de administración de IAM que se adjunta a las entidades principales de IAM para establecer los permisos máximos que puede tener el usuario o el rol. Para obtener más información, consulte Límites de permisos en la documentación de IAM.

información de identificación personal (PII)

Información que, vista directamente o combinada con otros datos relacionados, puede utilizarse para deducir de manera razonable la identidad de una persona. Algunos ejemplos de información de identificación personal son los nombres, las direcciones y la información de contacto.

PII

Consulte información de identificación personal.

manual de estrategias

Conjunto de pasos predefinidos que capturan el trabajo asociado a las migraciones, como la entrega de las funciones de operaciones principales en la nube. Un manual puede adoptar la forma de scripts, manuales de procedimientos automatizados o resúmenes de los procesos o pasos necesarios para operar un entorno modernizado.

PLC

Consulte controlador lógico programable.

PLM

Consulte administración del ciclo de vida del producto.

policy

Objeto que puede definir permisos (consulte política basada en identidad), especificar las condiciones de acceso (consulte política basada en recursos) o definir los permisos máximos para todas las cuentas de una organización de AWS Organizations (consulte política de control de servicio).

persistencia políglota

Elegir de forma independiente la tecnología de almacenamiento de datos de un microservicio en función de los patrones de acceso a los datos y otros requisitos. Si sus microservicios tienen la misma tecnología de almacenamiento de datos, pueden enfrentarse a desafíos de implementación o experimentar un rendimiento deficiente. Los microservicios se implementan más fácilmente y logran un mejor rendimiento y escalabilidad si utilizan el almacén de datos que mejor se adapte a sus necesidades.

evaluación de cartera

Proceso de detección, análisis y priorización de la cartera de aplicaciones para planificar la migración. Para obtener más información, consulte la Evaluación de la preparación para la migración.

predicate

Condición de consulta que devuelve `true` o `false`. En general, se encuentra en una cláusula `WHERE`.

inserción de predicados

Técnica de optimización de consultas en bases de datos que filtra los datos de la consulta antes de transferirlos. Esta técnica reduce la cantidad de datos de la base de datos relacional que se tienen que recuperar y procesar. Además, mejora el rendimiento de las consultas.

control preventivo

Un control de seguridad diseñado para evitar que ocurra un evento. Estos controles son la primera línea de defensa para evitar el acceso no autorizado o los cambios no deseados en la red. Para obtener más información, consulte Controles preventivos en Implementación de controles de seguridad en AWS.

entidad principal

Una entidad AWS que puede realizar acciones y acceder a los recursos. Esta entidad suele ser un usuario raíz para un Cuenta de AWS rol de IAM o un usuario. Para obtener más información, consulte Entidad principal en Términos y conceptos de roles en la documentación de IAM.

Privacidad desde el diseño

Enfoque de ingeniería de sistemas que tiene en cuenta la privacidad durante todo el proceso de desarrollo.

zonas alojadas privadas

Contenedor que aloja información acerca de cómo desea que responda Amazon Route 53 a las consultas de DNS de un dominio y sus subdominios en una o varias VPC. Para obtener más información, consulte Uso de zonas alojadas privadas en la documentación de Route 53.

control proactivo

Control de seguridad que se diseñó para evitar la implementación de recursos que no cumplan con la normativa. Estos controles analizan los recursos antes de aprovisionarlos. Si el recurso no cumple con los requisitos del control, no se aprovisiona. Para obtener más información, consulte la guía de referencia de controles en la AWS Control Tower documentación y consulte Controles proactivos en Implementación de controles de seguridad en AWS.

administración del ciclo de vida del producto (PLM)

Administración de los datos y los procesos de un producto a lo largo de todo su ciclo de vida, desde el diseño, el desarrollo y el lanzamiento, pasando por el crecimiento y la madurez, hasta la reducción de su uso y su retirada.

entorno de producción

Consulte entorno.

controlador lógico programable (PLC)

En el sector de la fabricación, computadora adaptable y altamente fiable que supervisa las máquinas y automatiza los procesos de fabricación.

encadenamiento de peticiones

Uso de la salida de una petición de LLM como entrada para la siguiente petición a fin de generar mejores respuestas. Esta técnica se utiliza para dividir una tarea compleja en tareas secundarias o para refinar o ampliar de forma iterativa una respuesta preliminar. Ayuda a mejorar la precisión y la relevancia de las respuestas de un modelo y permite obtener resultados más detallados y personalizados.

seudonimización

El proceso de reemplazar los identificadores personales de un conjunto de datos por valores de marcadores de posición. La seudonimización puede ayudar a proteger la privacidad personal. Los datos seudonimizados siguen considerándose datos personales.

publish/subscribe (pub/sub)

Patrón que permite establecer comunicaciones asíncronas entre microservicios para mejorar la escalabilidad y la capacidad de respuesta. Por ejemplo, en un MES basado en microservicios, un microservicio puede publicar mensajes de eventos en un canal al que se pueden suscribir otros microservicios. El sistema puede agregar nuevos microservicios sin cambiar el servicio de publicación.

Q

plan de consulta

Serie de pasos, como instrucciones, que se utilizan para acceder a los datos de un sistema de base de datos relacional SQL.

regresión del plan de consulta

El optimizador de servicios de la base de datos elige un plan menos óptimo que antes de un cambio determinado en el entorno de la base de datos. Los cambios en estadísticas,

restricciones, configuración del entorno, enlaces de parámetros de consultas y actualizaciones del motor de base de datos PostgreSQL pueden provocar una regresión del plan.

R

Matriz RACI

Consulte responsable, fiable, consultada e informada (RACI).

RAG

Consulte generación aumentada por recuperación.

ransomware

Software malicioso que se ha diseñado para bloquear el acceso a un sistema informático o a los datos hasta que se efectúe un pago.

Matriz RASCI

Consulte responsable, fiable, consultada e informada (RACI).

RCAC

Consulte control de acceso por filas y columnas.

réplica de lectura

Una copia de una base de datos que se utiliza con fines de solo lectura. Puede enrutar las consultas a la réplica de lectura para reducir la carga en la base de datos principal.

rediseñar

Consulte Las 7 R.

objetivo de punto de recuperación (RPO)

La cantidad de tiempo máximo aceptable desde el último punto de recuperación de datos. Esto determina qué se considera una pérdida de datos aceptable entre el último punto de recuperación y la interrupción del servicio.

objetivo de tiempo de recuperación (RTO)

La demora máxima aceptable entre la interrupción del servicio y el restablecimiento del servicio.

refactorizar

Consulte Las 7 R.

Region

Conjunto de AWS recursos en un área geográfica. Cada uno Región de AWS está aislado e independiente de los demás para proporcionar tolerancia a las fallas, estabilidad y resiliencia. Para más información, consulte Specify which Regiones de AWS your account can use.

regresión

Una técnica de ML que predice un valor numérico. Por ejemplo, para resolver el problema de "¿A qué precio se venderá esta casa?", un modelo de ML podría utilizar un modelo de regresión lineal para predecir el precio de venta de una vivienda en función de datos conocidos sobre ella (por ejemplo, los metros cuadrados).

volver a alojar

Consulte Las 7 R.

versión

En un proceso de implementación, el acto de promover cambios en un entorno de producción.

reubicar

Consulte Las 7 R.

redefinir la plataforma

Consulte Las 7 R.

recomprar

Consulte Las 7 R.

resiliencia

Capacidad de una aplicación para resistir interrupciones o recuperarse de ellas. Al planificar la resiliencia en la Nube de AWS, la alta disponibilidad y la recuperación ante desastres son consideraciones comunes. Para más información, consulte Resiliencia en la Nube de AWS.

política basada en recursos

Una política asociada a un recurso, como un bucket de Amazon S3, un punto de conexión o una clave de cifrado. Este tipo de política especifica a qué entidades principales se les permite el acceso, las acciones compatibles y cualquier otra condición que deba cumplirse.

matriz responsable, confiable, consultada e informada (RACI)

Una matriz que define las funciones y responsabilidades de todas las partes involucradas en las actividades de migración y las operaciones de la nube. El nombre de la matriz se deriva de los tipos de responsabilidad definidos en la matriz: responsable (R), contable (A), consultado (C) e informado (I). El tipo de soporte (S) es opcional. Si incluye el soporte, la matriz se denomina matriz RASCI y, si la excluye, se denomina matriz RACI.

control receptivo

Un control de seguridad que se ha diseñado para corregir los eventos adversos o las desviaciones con respecto a su base de seguridad. Para obtener más información, consulte Controles receptivos en Implementación de controles de seguridad en AWS.

retain

Consulte Las 7 R.

retirar

Consulte Las 7 R.

Generación aumentada de recuperación (RAG)

Tecnología de IA generativa mediante la que un LLM hace referencia a un origen de datos autorizado que se encuentra fuera de sus orígenes de datos de entrenamiento antes de generar una respuesta. Por ejemplo, un modelo de RAG podría hacer una búsqueda semántica en la base de conocimientos o en los datos personalizados de una organización. Para más información, consulte ¿Qué es RAG (generación aumentada por recuperación)?

rotación

Proceso mediante el que periódicamente se actualiza un secreto para que resulte más difícil que un atacante pueda acceder a las credenciales.

control de acceso por filas y columnas (RCAC)

El uso de expresiones SQL básicas y flexibles que tienen reglas de acceso definidas. El RCAC consta de permisos de fila y máscaras de columnas.

RPO

Consulte objetivo de punto de recuperación.

RTO

Consulte objetivo de tiempo de recuperación.

manual de procedimientos

Conjunto de procedimientos manuales o automatizados necesarios para realizar una tarea específica. Por lo general, se diseñan para agilizar las operaciones o los procedimientos repetitivos con altas tasas de error.

S

SAML 2.0

Un estándar abierto que utilizan muchos proveedores de identidad (IdPs). Esta función permite el inicio de sesión único (SSO) federado, de modo que los usuarios pueden iniciar sesión Consola de administración de AWS o llamar a las operaciones de la AWS API sin tener que crear un usuario en IAM para todos los miembros de la organización. Para obtener más información sobre la federación basada en SAML 2.0, consulte Acerca de la federación basada en SAML 2.0 en la documentación de IAM.

SCADA

Consulte control de supervisión y adquisición de datos.

SCP

Consulte política de control de servicio.

secreta

En AWS Secrets Manager, información confidencial o restringida, como una contraseña o credenciales de usuario, que se almacena de forma cifrada. Se compone del valor del secreto y de sus metadatos. El valor del secreto puede ser binario, una sola cadena o varias cadenas. Para más información, consulte What's in a Secrets Manager secret? en la documentación de Secrets Manager.

seguridad desde el diseño

Enfoque de ingeniería de sistemas que tiene en cuenta la seguridad durante todo el proceso de desarrollo.

control de seguridad

Barrera de protección técnica o administrativa que impide, detecta o reduce la capacidad de un agente de amenazas para aprovechar una vulnerabilidad de seguridad. Existen cuatro tipos de controles de seguridad principales: preventivos, de detección, de respuesta y proactivos.

refuerzo de la seguridad

Proceso de reducir la superficie expuesta a ataques para hacerla más resistente a los ataques. Esto puede incluir acciones, como la eliminación de los recursos que ya no se necesitan, la implementación de prácticas recomendadas de seguridad consistente en conceder privilegios mínimos o la desactivación de características innecesarias en los archivos de configuración.

sistema de información sobre seguridad y administración de eventos (SIEM)

Herramientas y servicios que combinan sistemas de administración de información sobre seguridad (SIM) y de administración de eventos de seguridad (SEM). Un sistema de SIEM recopila, monitorea y analiza los datos de servidores, redes, dispositivos y otras fuentes para detectar amenazas y brechas de seguridad y generar alertas.

automatización de la respuesta de seguridad

Acción predefinida y programada que está diseñada para responder automáticamente a un evento de seguridad o corregirlo. Estas automatizaciones sirven como controles de seguridad preventivos o adaptables que le ayudan a implementar las mejores prácticas AWS de seguridad. La modificación de un grupo de seguridad de VPC, la aplicación de revisiones a una instancia de Amazon EC2 o la rotación de credenciales son algunos ejemplos de acciones de respuesta automatizadas.

cifrado del servidor

Cifrado de los datos en su destino, por parte de Servicio de AWS quien los recibe.

política de control de servicio (SCP)

Una política que proporciona un control centralizado de los permisos de todas las cuentas de una organización en AWS Organizations. Las SCP definen barreras de protección o establecen límites a las acciones que un administrador puede delegar en los usuarios o roles. Puede utilizar las SCP como listas de permitidos o rechazados, para especificar qué servicios o acciones se encuentra permitidos o prohibidos. Para obtener más información, consulte las políticas de control del servicio en la AWS Organizations documentación.

punto de enlace de servicio

La URL del punto de entrada de un Servicio de AWS. Para conectarse mediante programación a un servicio de destino, puede utilizar un punto de conexión. Para obtener más información, consulte Puntos de conexión de Servicio de AWS en Referencia general de AWS.

acuerdo de nivel de servicio (SLA)

Acuerdo que aclara lo que un equipo de TI se compromete a ofrecer a los clientes, como el tiempo de actividad y el rendimiento del servicio.

indicador de nivel de servicio (SLI)

Medición de un aspecto del rendimiento de un servicio, como la tasa de errores, la disponibilidad o el rendimiento.

objetivo de nivel de servicio (SLO)

Métrica objetivo que representa el estado de un servicio medido mediante un indicador de nivel de servicio.

modelo de responsabilidad compartida

Un modelo que describe la responsabilidad con AWS la que compartes la seguridad y el cumplimiento de la nube. AWS es responsable de la seguridad de la nube, mientras que usted es responsable de la seguridad en la nube. Para obtener más información, consulte el Modelo de responsabilidad compartida.

Shadow AI

Aplicaciones de IA no autorizadas creadas o utilizadas fuera de los canales regulados dentro de una organización.

SIEM

Consulte sistema de administración de eventos e información de seguridad.

único punto de error (SPOF)

Error en un único componente crítico de una aplicación que puede interrumpir el sistema.

SLA

Consulte acuerdo de nivel de servicio.

SLI

Consulte indicador de nivel de servicio.

SLO

Consulte objetivo de nivel de servicio.

modelo de dividir y sembrar

Un patrón para escalar y acelerar los proyectos de modernización. A medida que se definen las nuevas funciones y los lanzamientos de los productos, el equipo principal se divide para crear nuevos equipos de productos. Esto ayuda a ampliar las capacidades y los servicios de su organización, mejora la productividad de los desarrolladores y apoya la innovación rápida. Para más información, consulte Phased approach to modernizing applications in the Nube de AWS.

SPOF

Consulte único punto de error.

esquema en estrella

Estructura organizativa de una base de datos que utiliza una tabla de hechos de gran tamaño para almacenar datos transaccionales o medidos y una o varias tablas dimensionales más pequeñas para almacenar los atributos de los datos. Esta estructura está diseñada para utilizarse en un almacén de datos o con fines de inteligencia empresarial.

patrón de higo estrangulador

Un enfoque para modernizar los sistemas monolíticos mediante la reescritura y el reemplazo gradual de las funciones del sistema hasta que se pueda desmantelar el sistema heredado. Este patrón utiliza la analogía de una higuera que crece hasta convertirse en un árbol estable y, finalmente, se apodera y reemplaza a su host. El patrón fue presentado por Martin Fowler como una forma de gestionar el riesgo al reescribir sistemas monolíticos. Para ver un ejemplo de cómo aplicar este patrón, consulte Modernización gradual de los servicios web antiguos de Microsoft ASP.NET (ASMX) mediante contenedores y Amazon API Gateway.

subred

Un intervalo de direcciones IP en la VPC. Una subred debe residir en una sola zona de disponibilidad.

control de supervisión y adquisición de datos (SCADA)

En el sector de la fabricación, sistema que utiliza hardware y software para supervisar los activos físicos y las operaciones de producción.

cifrado simétrico

Un algoritmo de cifrado que utiliza la misma clave para cifrar y descifrar los datos.

pruebas sintéticas

Prueba de un sistema de manera que simule las interacciones de los usuarios para detectar posibles problemas o supervisar el rendimiento. Puede usar Amazon CloudWatch Synthetics para crear estas pruebas.

petición del sistema

Técnica para proporcionar contexto, instrucciones o pautas a un LLM para dirigir su comportamiento. Las peticiones del sistema ayudan a establecer el contexto y las reglas para las interacciones con los usuarios.

T

etiquetas

Key-value pares que actúan como metadatos para organizar sus AWS recursos. Las etiquetas pueden ayudar a administrar, identificar, organizar, buscar y filtrar recursos de . Para obtener más información, consulte Etiquetado de los recursos de AWS.

variable de destino

El valor que intenta predecir en el ML supervisado. Esto también se conoce como variable de resultado. Por ejemplo, en un entorno de fabricación, la variable objetivo podría ser un defecto del producto.

lista de tareas

Herramienta que se utiliza para hacer un seguimiento del progreso mediante un manual de procedimientos. La lista de tareas contiene una descripción general del manual de procedimientos y una lista de las tareas generales que deben completarse. Para cada tarea general, se incluye la cantidad estimada de tiempo necesario, el propietario y el progreso.

entorno de prueba

Consulte entorno.

entrenamiento

Proporcionar datos de los que pueda aprender su modelo de ML. Los datos de entrenamiento deben contener la respuesta correcta. El algoritmo de aprendizaje encuentra patrones en los

datos de entrenamiento que asignan los atributos de los datos de entrada al destino (la respuesta que desea predecir). Genera un modelo de ML que captura estos patrones. Luego, el modelo de ML se puede utilizar para obtener predicciones sobre datos nuevos para los que no se conoce el destino.

herramienta

Una función o API que un agente puede invocar para realizar operaciones en sistemas externos.

puerta de enlace de tránsito

Centro de tránsito de red que puede utilizar para interconectar las VPC y las redes en las instalaciones. Para obtener más información, consulte Qué es una pasarela de tránsito en la AWS Transit Gateway documentación.

flujo de trabajo basado en enlaces troncales

Un enfoque en el que los desarrolladores crean y prueban características de forma local en una rama de característica y, a continuación, combinan esos cambios en la rama principal. Luego, la rama principal se adapta a los entornos de desarrollo, preproducción y producción, de forma secuencial.

acceso de confianza

Otorgar permisos a un servicio que especifique para realizar tareas en su organización AWS Organizations y en sus cuentas en su nombre. El servicio de confianza crea un rol vinculado al servicio en cada cuenta, cuando ese rol es necesario, para realizar las tareas de administración por usted. Para obtener más información, consulte AWS Organizations Utilización con otros AWS servicios en la AWS Organizations documentación.

ajuste

Cambiar aspectos de su proceso de formación a fin de mejorar la precisión del modelo de ML. Por ejemplo, puede entrenar el modelo de ML al generar un conjunto de etiquetas, incorporar etiquetas y, luego, repetir estos pasos varias veces con diferentes ajustes para optimizar el modelo.

equipo de dos pizzas

Un DevOps equipo pequeño al que puedes alimentar con dos pizzas. Un equipo formado por dos integrantes garantiza la mejor oportunidad posible de colaboración en el desarrollo de software.

U

incertidumbre

Un concepto que hace referencia a información imprecisa, incompleta o desconocida que puede socavar la fiabilidad de los modelos predictivos de ML. Hay dos tipos de incertidumbre: la incertidumbre epistémica se debe a datos limitados e incompletos, mientras que la incertidumbre aleatoria se debe al ruido y la aleatoriedad inherentes a los datos.

tareas indiferenciadas

También conocido como tareas arduas, es el trabajo que es necesario para crear y operar una aplicación, pero que no proporciona un valor directo al usuario final ni proporciona una ventaja competitiva. Algunos ejemplos de tareas indiferenciadas son la adquisición, el mantenimiento y la planificación de la capacidad.

entornos superiores

Consulte entorno.

V

succión

Una operación de mantenimiento de bases de datos que implica limpiar después de las actualizaciones incrementales para recuperar espacio de almacenamiento y mejorar el rendimiento.

control de versión

Procesos y herramientas que realizan un seguimiento de los cambios, como los cambios en el código fuente de un repositorio.

Emparejamiento de VPC

Conexión entre dos VPC que permite enrutar el tráfico mediante direcciones IP privadas. Para obtener más información, consulte ¿Qué es una interconexión de VPC? en la documentación de Amazon VPC.

vulnerabilidad

Defecto de software o hardware que pone en peligro la seguridad del sistema.

W

caché caliente

Un búfer caché que contiene datos actuales y relevantes a los que se accede con frecuencia. La instancia de base de datos puede leer desde la caché del búfer, lo que es más rápido que leer desde la memoria principal o el disco.

datos templados

Datos a los que el acceso es infrecuente. Al consultar este tipo de datos, normalmente se aceptan consultas moderadamente lentas.

función de ventana

Función SQL que hace un cálculo en un grupo de filas que se relacionan de alguna manera con el registro actual. Las funciones de ventana son útiles para las tareas de procesamiento, como calcular una media móvil o acceder al valor de las filas en función de la posición relativa de la fila actual.

carga de trabajo

Conjunto de recursos y código que ofrece valor comercial, como una aplicación orientada al cliente o un proceso de backend.

flujo de trabajo

Grupos funcionales de un proyecto de migración que son responsables de un conjunto específico de tareas. Cada flujo de trabajo es independiente, pero respalda a los demás flujos de trabajo del proyecto. Por ejemplo, el flujo de trabajo de la cartera es responsable de priorizar las aplicaciones, planificar las oleadas y recopilar los metadatos de migración. El flujo de trabajo de la cartera entrega estos recursos al flujo de trabajo de migración, que luego migra los servidores y las aplicaciones.

WORM

Consulte escritura única y lectura múltiple.

WQF

Consulte AWS Workload Qualification Framework.

escritura única y lectura múltiple (WORM)

Modelo de almacenamiento que escribe los datos una sola vez y evita que se eliminen o modifiquen. Los usuarios autorizados pueden leer los datos tantas veces como sea necesario, pero no los pueden cambiar. Esta infraestructura de almacenamiento de datos se considera inmutable.

Z

ataque de día cero

Ataque, normalmente de malware, que se aprovecha de una vulnerabilidad de día cero.

vulnerabilidad de día cero

Un defecto o una vulnerabilidad sin mitigación en un sistema de producción. Los agentes de amenazas pueden usar este tipo de vulnerabilidad para atacar el sistema. Los desarrolladores suelen darse cuenta de la vulnerabilidad a raíz del ataque.

peticiones desde cero

Proporcionar a un LLM instrucciones para llevar a cabo una tarea, pero sin ejemplos (pasos) que puedan ayudar a guiarlo. El LLM debe usar los conocimientos del entrenamiento previo para llevar a cabo la tarea. La eficacia de la petición desde cero depende de la complejidad de la tarea y de la calidad de la petición. Consulte también peticiones con pocos pasos.

aplicación zombi

Aplicación que utiliza un promedio de CPU y memoria menor al 5 por ciento. En un proyecto de migración, es habitual retirar estas aplicaciones.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la version original de inglés, prevalecerá la version en inglés.