**UNIVERSIDAD MARIANO GALVEZ DE GUATEMALA**

**Centro de Datos (supervisión)**

**Jonathan Benjamín  
Herrera Rodríguez**

**7690-13-1131**

**Alta Disponibilidad**

La alta disponibilidad es la capacidad que tiene un sistema de TI para ser accesible y confiable casi todo el tiempo, lo cual elimina o disminuye el tiempo de inactividad. Combina dos conceptos para determinar si un sistema cumple con su nivel de rendimiento operativo: el primero tiene que ver con la accesibilidad o la disponibilidad prácticamente permanentes de un servicio o servidor sin tiempo de inactividad; y el segundo se refiere a su funcionamiento según las expectativas razonables y durante un período establecido. No se trata solo de cumplir con el tiempo de actividad indicado en el acuerdo de nivel del servicio (SLA) o con las expectativas establecidas entre el proveedor y el cliente, sino que implica que el sistema sea resistente, confiable y eficaz.

**La importancia de la alta disponibilidad**

La adopción de los servicios en línea y las cargas de trabajo híbridas genera la necesidad de que las infraestructuras gestionen el aumento de las cargas en los sistemas mientras mantienen los estándares operativos.No es suficiente con que las infraestructuras, comúnmente conocidas como sistemas de alta disponibilidad, "funcionen mejor", sino que deben generar resultados definidos y cuantificables para lograr la alta disponibilidad.

Uno de los objetivos de estas soluciones o servicios es ofrecer un tiempo de actividad del 99,999 %, es decir, que el sistema funcione correctamente el 99,999 % del tiempo. Por lo general, solo los sectores más importantes, como la [salud](https://www.redhat.com/es/solutions/healthcare), el [Gobierno](https://www.redhat.com/es/solutions/public-sector) y los [servicios financieros](https://www.redhat.com/es/technologies/industries/financial), requieren este nivel de disponibilidad por razones de cumplimiento normativo o de competencia. Sin embargo, también hay otras empresas y sectores que necesitan que los sistemas mantengan el tiempo de actividad del 99,9 % o incluso del 99,99 % para ofrecer acceso digital constante a los clientes o permitir que los empleados trabajen desde sus hogares.

**El funcionamiento de la alta disponibilidad**

La infraestructura de alta disponibilidad depende de la detección y eliminación de los puntos únicos de fallas que podrían aumentar el tiempo de inactividad de los sistemas e impedir que las empresas alcancen sus objetivos de rendimiento. Los puntos únicos de fallas son aspectos de la infraestructura que podrían desconectar todo el sistema, y puede haber muchos de ellos en los sistemas complejos.

Las empresas también deben considerar los distintos tipos de fallas que se pueden producir en las [infraestructuras de TI](https://www.redhat.com/es/topics/cloud-computing/what-is-it-infrastructure) modernas y complejas, como las del hardware; del software (tanto en el sistema operativo como en las aplicaciones en ejecución); del servicio (como las que se producen por la falta de accesibilidad de la red y la latencia en los servicios de nube, o el deterioro del rendimiento), y las fallas externas, como por un corte de la energía eléctrica.

La primera medida que pueden tomar las empresas para lograr la alta disponibilidad es determinar los resultados específicos más importantes que esperan para sus servicios esenciales, los requisitos de cumplimiento normativo y de las cargas de trabajo, los parámetros de rendimiento, las aplicaciones fundamentales y las prioridades operativas:

¿Cuáles son sus requisitos en cuanto al tiempo de actividad, ya sea para el cumplimiento normativo o para la experiencia del usuario?

¿Qué tan distribuido está su entorno? ¿Cuáles son los puntos de fallas más importantes?

¿Cuáles son los requisitos de rendimiento para la aplicación? ¿Cuáles son los riesgos que corre el rendimiento de las aplicaciones? (Por ejemplo, el tráfico de usuarios o la carga de escritura).

¿Qué tipo de [almacenamiento](https://www.redhat.com/es/topics/data-storage/why-choose-red-hat-storage) utiliza?

¿Cuáles son los requisitos respecto a la pérdida de datos o el acceso a ellos?

Teniendo en cuenta los recursos informáticos actuales, ¿cuáles son los acuerdos de nivel de servicio (SLA) que se pueden cumplir en caso de que se produzca una interrupción del servicio? ¿Cuáles son los períodos de mantenimiento programados en la actualidad? ¿De qué manera afectarán el tiempo de actividad?

¿Tiene planes para los diferentes casos de recuperación ante desastres o los cambios en las operaciones de la empresa?

Los equipos de TI utilizan varios indicadores comunes para determinar si la arquitectura de los entornos de alta disponibilidad cumple con sus objetivos. Es posible que algunos sean más útiles que otros para su [arquitectura](https://www.redhat.com/es/topics/cloud-native-apps/what-is-an-application-architecture), pero vale la pena evaluarlos a todos para establecer las expectativas de referencia en cuanto al rendimiento:

Tiempo medio entre fallas (MTBF): determina el tiempo de funcionamiento del entorno entre una falla y otra del sistema.

Tiempo medio de inactividad: determina el tiempo de inactividad del sistema, en minutos, antes de que se recupere o sustituya.

Objetivo de tiempo de recuperación (RTO): determina el tiempo total que toma completar una reparación y volver a conectar el sistema.

Objetivo de punto de recuperación (RPO): determina el período en el que se deben poder recuperar los datos. Es el período en que se pierden los datos. Por ejemplo, si un sistema depende de la recuperación de otro cuyos backups se ejecutan una vez al día, podría haber una pérdida de datos de hasta 24 horas. Por el contrario, si cuenta con el almacenamiento replicado o compartido, la pérdida puede ser de tan solo unos minutos o menos.

**La alta disponibilidad y la recuperación ante desastres**

La arquitectura de alta disponibilidad incorpora los principios de cada aspecto del plan de continuidad, como la supervisión y la [automatización](https://www.redhat.com/es/topics/automation). Esto permite que todo el sistema sea resistente a cualquier tipo de fallas, desde las locales y específicas hasta las que se producen por los cortes generales de energía eléctrica. Además, permite mantenerlo en ejecución incluso durante los períodos de mantenimiento programados u otras interrupciones.

El plan de recuperación ante desastres o de continuidad permite abordar cada tipo de falla:

Previsión de fallas específicas: los arquitectos de TI se aseguran de que los sistemas sean redundantes y cuenten con backups en caso de que se produzcan fallas. Luego, se encargan de automatizar los procesos de conmutación por error y de detección de fallas para identificar automáticamente la desconexión de los sistemas y trasladar los servicios al sistema de backup

Gestión preventiva del rendimiento: la tolerancia a los errores puede ocuparse de las interrupciones, pero no necesariamente del deterioro del rendimiento. En este caso, resulta muy útil contar con el equilibrio de carga y la capacidad de ajuste. Los arquitectos de TI supervisan el rendimiento del sistema y utilizan varios sistemas para gestionar las operaciones y solicitudes de los usuarios. La gestión del tráfico y los equilibradores de carga dirigen de forma inteligente el tráfico en tiempo real en función del ancho de banda, el rendimiento del sistema y el tipo de usuario o de solicitud.

Gestión de las fallas generalizadas: si bien son poco frecuentes, como la desconexión de un proveedor de nube o un desastre natural en el centro de datos, requieren un enfoque más integral que las fallas en el hardware o el software. No solo es necesario recuperar la conexión de la infraestructura, sino también disponer de los datos actualizados mediante su replicación (la cual se ejecuta de forma sincrónica, pero puede afectar el rendimiento), o bien mediante los backups (los cuales se ejecutan de forma asíncrona, pero pueden implicar la pérdida de datos).

[**Aumente la continuidad empresarial**](https://www.redhat.com/en/resources/maximize-business-continuity-datasheet)

**Los clústeres de alta disponibilidad**

Las arquitecturas de alta disponibilidad ejecutan clústeres activos de conmutación por error, por lo que la conmutación y la redundancia están integradas, y, con suerte, no habrá tiempos de inactividad.Dentro del clúster, se supervisa la disponibilidad de los nodos y el rendimiento general de las aplicaciones, los servicios y la red. Dado que el almacenamiento es compartido, es decir que todos los nodos del clúster utilizan la misma fuente de datos, si un nodo deja de funcionar, no se pierde ningún dato. Se puede utilizar el equilibrio de carga para gestionar el tráfico y obtener el mejor rendimiento.

Además de ofrecer estas características generales, también se pueden diseñar los clústeres de alta disponibilidad para tareas más específicas en función de las prioridades y las actividades de la infraestructura de TI. Por ejemplo, [Red Hat Enterprise Linux High Availability Add-On](https://www.redhat.com/en/store/high-availability-add) tiene cuatro configuraciones predeterminadas:

Alta disponibilidad: se centra en el tiempo de actividad y la disponibilidad.

Alto rendimiento: se centra en las operaciones simultáneas de alta velocidad.

Equilibrio de carga: se centra en la capacidad de ajuste rentable.

Almacenamiento: se centra en la gestión flexible de los datos.

En los entornos reales, los sistemas de alta disponibilidad podrían incorporar características de esos elementos.

[El sistema operativo puede marcar la diferencia](https://www.redhat.com/es/engage/operating-system-matters-s-202203180234)

**Red Hat y la alta disponibilidad**

La alta disponibilidad abarca toda la infraestructura y debe contemplar la [gestión de los datos y el almacenamiento](https://www.redhat.com/es/topics/data-services/what-is-data-management) en entornos distintos, tanto físicos como en la [nube](https://www.redhat.com/es/topics/cloud-native-apps), así como las diferentes ubicaciones de los servicios y las aplicaciones. Por eso es tan útil contar con una plataforma común y un [entorno operativo estándar](https://www.redhat.com/es/topics/management/what-is-an-soe): crean uniformidad independientemente del entorno de implementación.

[Red Hat Enterprise Linux](https://www.redhat.com/es/technologies/linux-platforms/enterprise-linux) tiene funciones y servicios adicionales que pueden incluirse mediante paquetes de complementos. [Red Hat Enterprise Linux High Availability Add-On](https://www.redhat.com/en/store/high-availability-add) se ocupa de los aspectos de la red, la agrupación en clústeres y el almacenamiento.

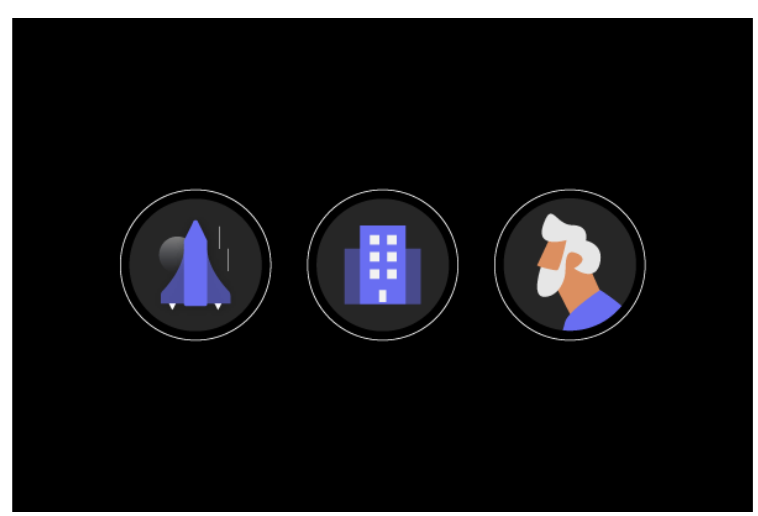
Las implementaciones de Red Hat Enterprise Linux para [Microsoft SQL Server](https://www.redhat.com/es/technologies/linux-platforms/enterprise-linux/sql-server) y [SAP](https://www.redhat.com/es/technologies/linux-platforms/enterprise-linux/sap)también incluyen Red Hat Enterprise Linux High Availability Add-On, dado que la alta disponibilidad está muy relacionada con la [gestión de los datos](https://www.redhat.com/es/topics/data-services/what-is-data-management).

**Misión Crítica**

Los Sistemas de Misión Crítica son las aplicaciones que son esenciales para el funcionamiento de una actividad en una empresa, organización o gobierno. Si estas soluciones se dañan o fallan, pueden provocar el colapso de todos los procesos y operaciones que dependen de esta.

**Ejemplos de Sistemas de Misión Crítica**

Sistemas de Misión  
Son las soluciones que fueron concebidas para cumplir con un propósito en la organización. Si algo falla, ese objetivo se deja de cumplir.  
  
Sistemas de Negocio  
Son las soluciones que, si fallan, inciden directamente en la economía de la organización que usa ese sistema.  
  
Sistemas de la Gente  
Son las soluciones que, en el caso de fallar, pueden poner en riesgo la vida o la salud de las personas.

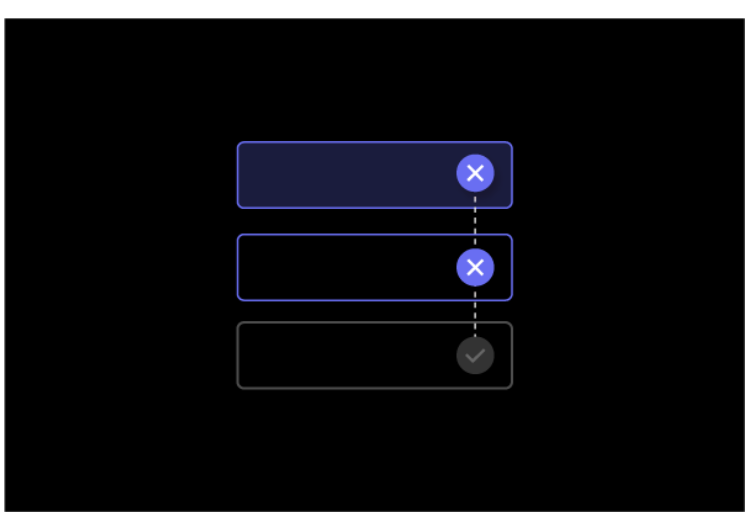


**¿Por qué fracasan los Sistemas de Misión Crítica?**

Por pretender resolver todos los problemas y funcionalidades en una primera etapa.

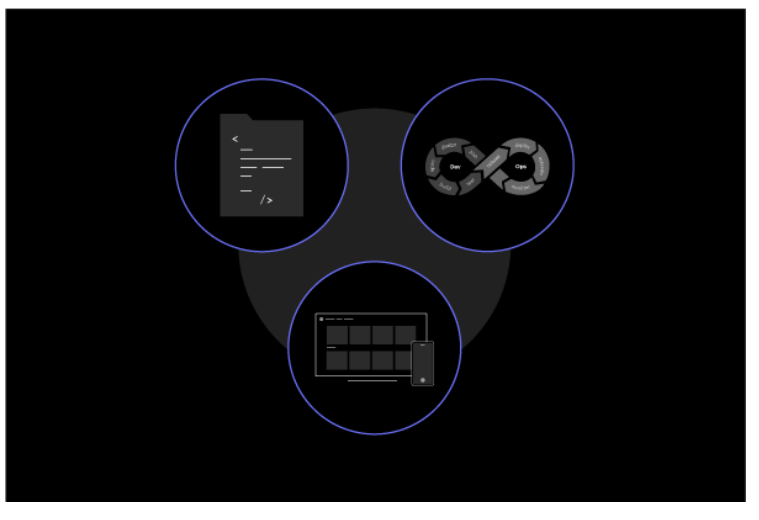
Por enfocarse solo en las necesidades técnicas y no en las necesidades del usuario.

Por restarle importancia al Modelado de Datos y a la Observabilidad.



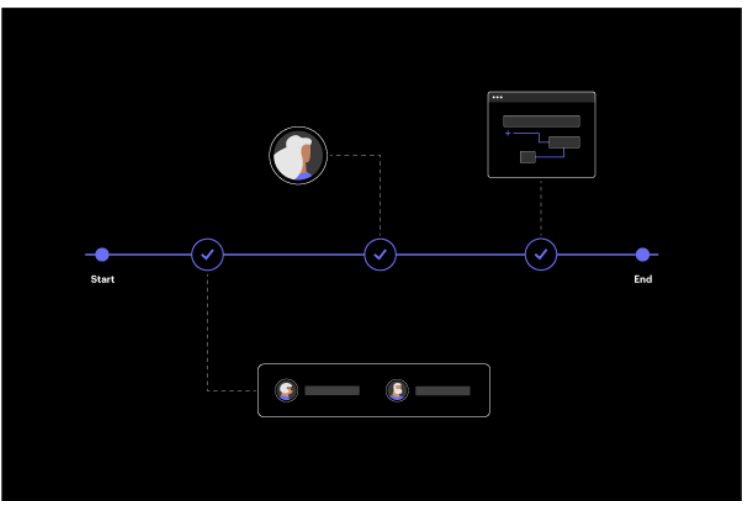
**¿Cuál es la fórmula para construir eficazmente un Sistema de Misión Crítica?**

Modelado de Datos  
Son la base para que ese desarrollo pueda hacerse de forma rápida y evolutiva.  
  
Integración  
La solución debe poder integrarse con el resto del ecosistema: APIs, datos y más.  
  
Multi-experiencia  
La clave es centrarse en los problemas que deben resolverles a los usuarios y en la experiencia que tendrán cuando usen la solución.



**¿Cuáles son los pasos para desarrollar un Sistema de Misión Crítica?**

Lo primero es crear un equipo de trabajo y delegar roles y responsabilidades. Lo siguiente es comprender cuáles son las historias, necesidades y problemas que deben resolverles a los usuarios que usarán el sistema. El Modelado de datos y Procesos, constituyen los otros puntos relevantes que deben considerarse para iniciar con buen pie la construcción de un Sistema de Misión Crítica.



**Claves para escoger una plataforma de desarrollo de Sistemas de Misión Crítica**

Adaptabilidad a lo impredecible  
Para eso se requiere de una herramienta ágil, que permita entregar y desarrollar software velozmente.

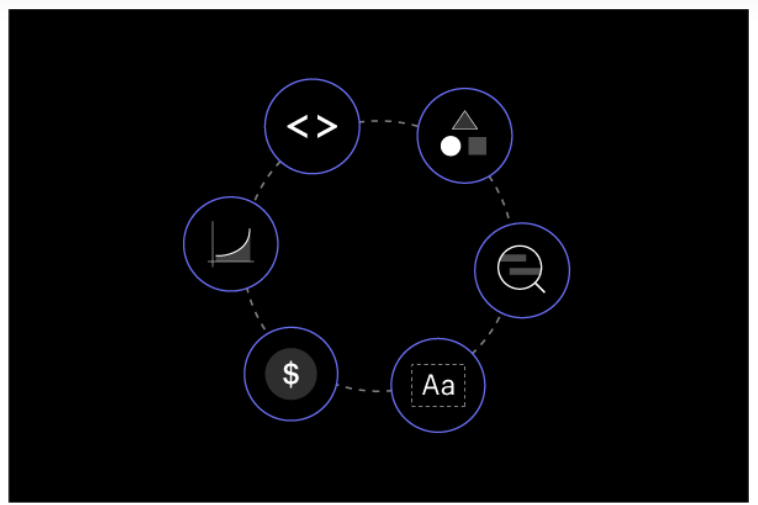
Flexibilidad  
Debe poder adaptarse a nuevas situaciones e integrarse con otras tecnologías y sistemas.

Metodología incremental  
Debe permitir hacer testing continuo.

Diseño  
Debe brindar una buena experiencia visual, donde fluyan las interacciones entre el mundo real y la tecnología.

Bajo costo  
Para que los proyectos sean viables, debe permitir desarrollar software a un costo razonable.

Fácil de aprender  
La curva de aprendizaje debe ser mínima, así los desarrolladores podrán invertir tiempo en lo que realmente importa: aportar valor y calidad al negocio.



**Redundancia N, N+1, 2N, 2(N+1)**

**N**: Es la cantidad necesaria de un componente o sistema para cumplir con los requisitos operativos. Por ejemplo, si se necesitan tres servidores para manejar la carga de trabajo, N sería 3.

**N+1**: Se refiere a la provisión de un componente adicional al número necesario para garantizar la operación en caso de fallo de uno de ellos. En el ejemplo anterior, N+1 sería 4 servidores.

**2N**: Implica la duplicación completa de todos los componentes o sistemas necesarios, creando un sistema completamente redundante.

**2(N+1)**: Se refiere a la duplicación del esquema N+1, proporcionando un nivel de redundancia extremadamente alto, donde cada componente principal y de respaldo tiene su duplicado.

**PUE (Power Usage Effectiveness)**

El PUE es una métrica utilizada para determinar la eficiencia energética de un centro de datos. Se calcula dividiendo la cantidad total de energía consumida por el centro de datos entre la energía utilizada únicamente por los equipos de TI. Un PUE ideal es 1.0, lo que significa que toda la energía se utiliza para la carga de TI, sin desperdicio en sistemas auxiliares como la refrigeración o la iluminación.

**BMS (Building Management System)**

El BMS es un sistema de gestión de edificios que controla y monitorea la infraestructura del centro de datos, incluyendo la iluminación, la climatización (HVAC), la energía y los sistemas de seguridad. El BMS es crucial para asegurar que las condiciones operativas del centro de datos se mantengan dentro de los parámetros establecidos para evitar fallos y asegurar la eficiencia operativa.

**DCIM (Data Center Infrastructure Management)**

El DCIM es una solución de software que permite a los administradores de centros de datos monitorizar, medir y gestionar la utilización de los recursos del centro de datos y el consumo de energía. Ayuda a optimizar la disposición de los equipos, la distribución de energía, y la gestión térmica, lo que contribuye a mejorar la eficiencia y reducir los costes operativos.

Estos conceptos son esenciales para garantizar que un centro de datos funcione de manera eficiente, segura y con el mínimo tiempo de inactividad, cumpliendo con los requisitos operativos de las organizaciones que dependen de ellos.