PLIEGO DE BASES TÉCNICAS
SUMINISTRO E INSTALACIÓN DE UNA PLATAFORMA HIPERCONVERGENTE PARA LOS CENTROS DE PROCESO DE DATOS DE ERROTABURU, LOIOLA Y PETRITEGI

Donostia / San Sebastián, 2 de julio de 2021

Contenido

1. Antecedentes 3
2. Opción Tecnológica 3
3. Ventajas de la solución hiperconvergente 4
   3.1 Consolidación de recursos y múltiples funciones de TI 4
   3.2 Escalabilidad y flexibilidad 5
   3.3 Fácil automatización y mayor eficiencia 5
   3.4 Asegurar la continuidad de negocio y la recuperación ante desastres 5
   3.5 Reducción de costes 5
4. Objeto del contrato 6
5. Requisitos técnicos 6
   5.1 Numero de nodos, clúster y requerimientos HW 6
   5.2 Eficiencias y cifrado de datos 7
   5.3 Resiliencia y capacidad de almacenamiento de la solución 8
   5.4 Gestión de la solución 9
   5.5 *Backup* Integrado 9
   5.6 Distribución de datos en modo Stretched-cluster 10
   5.7 Autosoporte y uso de Inteligencia Artificial para la detección precoz de incidencias 10
   5.8 Otros requerimientos técnicos 11
   5.9 Licencias de software 11
   5.10 Servicios de instalación, configuración y formación 12
   5.11 Garantía 12

SUMINISTRO E INSTALACIÓN DE UNA PLATAFORMA HIPERCONVERGENTE PARA LOS CENTROS DE PROCESO DE DATOS DE ERROTABURU, LOIOLA Y PETRITEGI

1. Antecedentes

Durante los últimos años la incorporación y uso de las Tecnologías de la Información y las Comunicaciones, en adelante TIC, ha supuesto una medida básica para la prestación de servicios de forma óptima y un soporte esencial para las tareas que realizan todos los profesionales que trabajan en las diferentes áreas de AGASA.

Las medidas de impulso al proceso de transformación digital han motivado que el uso de las TIC se haya convertido en un factor estratégico en AGASA. Actualmente las TIC son claves para el soporte de varios servicios y sistemas críticos dentro de la organización como son los que controlan los procesos industriales, los laboratorios o la gestión administrativa y contable.

Debido a esto, la infraestructura de los centros de proceso de datos de AGASA ha ido adquiriendo una relevancia y criticidad máxima, con sistemas muy complejos y extraordinariamente relacionados entre sí. Actualmente dicha infraestructura, además de estar quedando obsoleta por la antigüedad de algunos equipos, está al límite de su capacidad en lo referente a memoria y almacenamiento. Además, en los próximos meses hemos de afrontar nuevos proyectos que requerirán de más recursos de cómputo y almacenamiento como son, el nuevo sistema de planificación de recursos empresariales (ERP) o el nuevo sistema de Supervisión, Control y Adquisición de Datos (SCADA).

AGASA precisa por tanto contratar la adquisición y puesta en marcha de una nueva infraestructura de sistemas más escalable que paulatinamente sustituya a la actual y aumente los recursos para poder albergar sus sistemas de información con los mecanismos y niveles de calidad, seguridad y disponibilidad necesarios. Se pretende no sólo mejorar el rendimiento y la capacidad de crecimiento, sino que se busca unificar la infraestructura de los diferentes centros de datos con el objetivo de simplificar su mantenimiento y gestión.

2. Opción Tecnológica

Una vez analizada la situación actual y estudiado las opciones de mercado disponibles **se ha optado por la tecnología Hiperconvergente (HCI)** buscando una plataforma de virtualización de los datos que integre almacenamiento, networking y computación, dando cobertura a todos los elementos que conforman la infraestructura de los CPD.

La virtualización de los sistemas ha multiplicado las funcionalidades del almacenamiento y mejorado el rendimiento de la red, pero implican una carga adicional sobre la capa de computación en la que se basa. Conlleva la necesidad de adquirir recursos para proceso, almacenamiento, red y equipos de comunicaciones que sean capaces de manejar las

cargas previstas para las aplicaciones, lo cual se traduce en un rápido y dinámico crecimiento del número de máquinas virtuales.

Esta complejidad añadida puede anular las mejoras obtenidas en lo que respecta a servidores y a aplicaciones mediante una degradación del rendimiento del proceso, de la velocidad de la red, del almacenamiento o de la protección de datos y un sobreesfuerzo en las tareas de administración y gestión.

Los sistemas hiperconvergentes buscan simplificar esta complejidad. Estos sistemas integran hipervisor, servidores físicos, red y almacenamiento en una única infraestructura simplificada. El resultado es un coste de infraestructura menor, informática empresarial más rápida y una carga inferior para los administradores de los sistemas.

Una solución hiperconvergente puede ser definida como una solución llave en mano que favorece que la capa informática y la de datos coexistan en un mismo nodo. Esto permite la localización de la carga de trabajo en el nodo de almacenamiento, minimizando la necesidad de mover los datos a la capa de cómputo. El software aprovecha el almacenamiento interno de cada nodo para crear un conjunto compartido de almacenamiento al que todas las máquinas virtuales (VM) del entorno pueden acceder.

En comparación con la infraestructura tradicional, la infraestructura hiperconvergente es un software basado en nodos sobre un hardware estándar construido desde cero para entornos de alto rendimiento y alta escalabilidad. Mientras que aquella requiere de compras “escalonadas”, el diseño modular de la infraestructura hiperconvergente asegura que solo deban comprar nodos cuando realmente lo requieran. A diferencia de la infraestructura tradicional, la hiperconvergencia, en general, ofrece una relación carga de trabajo vs gestión muy ventajosa, estable y fácilmente escalable.

3. Ventajas de la solución hiperconvergente

La hiperconvergencia aprovecha lo mejor de las últimas tecnologías disponibles para los centros de datos (virtualización, almacenamiento, etc.) al tiempo que permite aunar los diferentes componentes de la infraestructura (servidores, redes, almacenamiento) en una solución integrada. La ventaja más inmediata es disponer de una tecnología unificada, lo que reduce costes y tiempos en la puesta en marcha de proyectos; es un sistema probado, previamente construido. Todos los componentes están ya instalados y conectados entre sí. Dado que todos los elementos están ya conectados, se reduce el número de incidencias.

A modo de resumen algunas de las ventajas más significativas son:

3.1 Consolidación de recursos y múltiples funciones de TI

La razón más “atractiva” para usar la HCI es sin duda que:

- Consolida los recursos informáticos, de almacenamiento y de redes.
- Consolida la carga de trabajo: Deduplicación, Compresión, Copia de seguridad y optimización de WAN.

Estando todo en la misma plataforma, facilitando la migración de máquinas virtuales entre diferentes dispositivos hiperconvergentes o incluso entre centros de datos.

3.2 Escalabilidad y flexibilidad

La HCI, al estar basada en software, permite:

- Niveles superiores de flexibilidad que la infraestructura tradicional.
- Un entorno con el con mayor escalabilidad, tanto horizontal como vertical
- Mayor facilidad para satisfacer las necesidades específicas de las aplicaciones.

Con la HCI expandir la capacidad del centro de datos se consigue de forma muy simple: integrando una o dos unidades adicionales.

3.3 Fácil automatización y mayor eficiencia

Muchas infraestructuras de TI tradicionales son tan complejas que hacen imposible la automatización, algo que se vuelve más fácil con la HCI gracias a:

- A la consolidación de recursos: cómputo, almacenamiento y redes.
- Una gestión centralizada.

La mejora de la automatización implicará un aumento en la eficiencia, eliminando los procesos manuales y la necesidad de contar con un equipo de operaciones con conocimientos aislados.

3.4 Asegurar la continuidad de negocio y la recuperación ante desastres

La HCI facilita mucho los esfuerzos de recuperación ante desastres:

- La copia de seguridad y la recuperación de desastres están integradas como parte de la infraestructura hiperconvergente, lo que hace que la protección de datos sea más eficiente y asequible.
- El modelo de escalamiento horizontal se basa en la distribución de datos en múltiples nodos en todo el centro de datos, o incluso entre centros de datos en diferentes ubicaciones geográficas, lo que significa que cuando un dispositivo se desactiva, el rendimiento y la disponibilidad no se ven afectados.

3.5 Reducción de costes

Con la hiperconvergencia se reducen costes gracias a:

- Su escalabilidad, evitando redes complejas diseñadas ad-hoc.
- La posibilidad de ampliar capacidad fácilmente según las necesidades y sin interrupciones, evitando un aprovisionamiento excesivo, es decir, se optimiza el gasto en la infraestructura en los centros de datos.

- Se reducen costes de mantenimiento, energía, o licencias.

4. Objeto del contrato

Es objeto del presente expediente la adquisición de una plataforma hiperconvergente, incluyendo las licencias software necesarias y los servicios de configuración y puesta en marcha inicial de la misma.

La solución ofertada debe seguir los principios básicos de la tecnología hiperconvergente, donde la computación y el almacenamiento residan en el propio sistema de manera mono-bloque homogéneo, y no en alternativas de integración x86 que requieran integraciones heterogéneas.

5. Requisitos técnicos

Actualmente, AGASA dispone de tres CPDs que son los que se pretenden actualizar, en los siguientes apartados detallamos los requisitos técnicos mínimos que ha de cumplir la solución ofertada.

5.1 Numero de nodos, clúster y requerimientos HW

La solución deberá proporcionar el número de nodos necesario para cumplir los objetivos descritos en el documento.

Por razones de compatibilidad con las máquinas virtuales existentes, el hipervisor utilizado será vShpere de VMware, en sus versiones más recientes y estables.

Solo serán admitidas soluciones con nodos HCI puros. Es decir, no serán admitidas soluciones desagregadas donde el cómputo y el almacenamiento estén separados;

El número total de clúster que componen la solución global es de 3, uno por cada centro de proceso de datos:

- Errotaburu.
- Petritegi.
- Loiola.

Solo queda establecido el número mínimo en 2 nodos en cada clúster, a fin de conformar una solución 100% HA en CPU, RAM y Almacenamiento.

De forma independiente a los componentes que defina cada solución, cada nodo debe disponer de los siguientes mínimos:

- Doble fuente de alimentación en cada Nodo.
- Puerto de gestión remota (iLO, iDRAC, IPMI, etc.).
- Almacenamiento de datos en SSD.
- Almacenamiento para sistema de inicio (boot) en RAID1, no inferior a 250GB.

- Controladoras de discos independientes para Datos y Sistema.
- 16 núcleos físicos (cores) de al menos 3,0GHz cada uno.
- 256 GB de memoria RAM.
- 2 puertos de 10/25 GbE SFP+.
- 4 puertos de 1/10 GbE Base-T.

Todos los clúster que proponga la solución presentada deben permitir ampliación en caliente (con nodos adicionales), es decir, sin necesidad de reconfigurar el clúster ya instalado o realizar una apagado completo, o necesidad de cambiar el licenciamiento (del software HCI) existente.

5.2 Eficiencias y cifrado de datos

La solución propuesta debe tener activados todos sus motores de eficiencia (no solo disponer de las funcionalidades).

Los motores de eficiencia exigidos son compresión, deduplicación y optimización de escrituras en los discos de datos. La solución podrá proveer de más motores de eficiencia, siempre que no afecten a los tres exigidos.

La solución debe tener habilitada la funcionalidad de cifrado de datos. En concreto aquella funcionalidad que permita cifrar directamente el contenido de los discos de datos. Es necesario que si un disco de datos queda expuesto (por retirada, sustitución, robo, etc...) no pueda ser ‘montado’ o ‘accedido’ por ningún otro sistema.

Esta funcionalidad de cifrado debe ser ofrecida por la capa HCI, al margen de las opciones que pueda ofrecer el nivel hipervisor.

Los motores de eficiencia y el cifrado de datos deben quedar habilitados para el 100% de las cargas de trabajo y sin restricciones, ni limitaciones, entre ellos.

En caso de que la solución propuesta no pueda cumplir de forma intrínseca con este requerimiento, será necesario que las limitaciones técnicas queden reflejadas en la respuesta al pliego.

Dado que las funcionalidades solicitadas utilizarán los mismos recursos (CPU/RAM) que las cargas de trabajo productivas, las propuestas deberán identificar de forma clara la cantidad de vCPU (o Ghz) y vRAM que el software HCI consumirá en cada nodo.

En caso de que dichos consumos no sean fijos, o estén vinculados a las diferentes funcionalidades habilitadas, o no estén debidamente informadas en la documentación del fabricante, sus propuestas deberán añadir al menos un 20% extra sobre los recursos (CPU/RAM) solicitados. La documentación entregada con la propuesta deberá identificar claramente este incremento de recursos (CPU/RAM).

Aquellas propuestas que no identifiquen estos consumos y/o el incremento de recursos solicitado, serán automáticamente descartadas.

5.3 Resiliencia y capacidad de almacenamiento de la solución

Cada clúster que compone el proyecto debe proporcionar un almacenamiento de alta resiliencia, similar al obtenido en una cabina de discos con doble controladora activa-activa y configuración RAID5 o superior.

No es obligatorio el uso de una tecnología especifica en la gestión de discos/datos, cada propuesta podrá ofrecer diferentes métodos de gestión (RAIN, RAID, grupos de discos, Erasure codes, o combinaciones de estas).

Cada clúster debe proporcionar los mecanismos necesarios para permitir la perdida simultánea de 1 nodo (controlador) y al menos 1 disco de datos de cada nodo superviviente, sin que esta condición afecte a la disponibilidad de la información almacenada.

Si la solución ofertada precisa de reconstrucciones lógicas ante la pérdida de algún disco, deben proporcionarse nodos con discos pequeños (no mayores de 2TB) a fin de minimizar el tiempo de reconstrucción y su impacto en el rendimiento (Soluciones que no incorporan RAID).

En caso de que la solución solo tenga que reconstruir datos en el cambio de un disco y no en la perdida, deberá disponer de mecanismos que limiten el impacto en el rendimiento de dicha reconstrucción (Soluciones que incorporan RAID en cada nodo).

La solución debe proporcionar 6,9 TiB (7,5 TB) Netos Usables en cada clúster.

El concepto ‘Neto Usable’ hace referencia al espacio disponible en el clúster después de aplicar formato, mecanismos de resiliencia/replica/tolerancia a fallo de disco, espacios reservados para índices y operativas internas de las soluciones; pero antes de aplicar los motores de eficiencia. Ejemplo: Si la solución proporciona 30TB de espacio total en el clúster, reserva un 20% para índices y reservas para post-dedup o post-compresión, y almacena 3 copias (para permitir un doble fallo, como el solicitado), el espacio Neto Usable es de 8TB

La respuesta técnica deberá reflejar obligatoriamente las fórmulas, cálculos, funcionalidades o ejercicios de dimensionamiento realizados para llegar a los ‘Netos Usables’ presentados. Aquellas propuestas que solo proporcionen un listado de materiales físicos, o solo la suma de estos, serán descartadas.

5.4 Gestión de la solución

La solución presentada debe estar totalmente integrada con el hypervisor (VMware), es decir que no sea necesaria ninguna otra consola de gestión.

En caso de existir o requerir de alguna otra consola de gestión, la solución deberá proporcionar sus propios recursos CPU/RAM/Disco, de forma independiente a los solicitados y aportar las licencias necesarias en la edición, o versión, más completa y rica en funcionalidades.

En caso de caída, perdida o apagado del vCenter, la solución debe seguir trabajando con toda normalidad, e incluso permitir acciones de “backup/restore”.

Si la solución presentada considera el componente vCenter como crítico, o indispensable, para mantener y operar la solución, deberá proporcionar una solución HA para el propio vCenter.

Al margen de la gestión integrada, la solución debe disponer de un interface de comandos y un interface Rest-API para cualquier integración, desarrollo de scripts o acciones en caída de vCenter. Estos interfaces deben estar totalmente licenciados y documentados. Deberán entregarse las guías y/o accesos a URL con dicha documentación.

5.5 *Backup* Integrado

Al margen de la solución de *Backup* externa existente, la solución propuesta debe ser capaz de auto proteger los datos, mediante funcionalidades de *Backup* integradas, así como de funcionalidades para realizar replicas remotas a nodos o clúster externos.

Estas funcionalidades no deben ser únicamente Snapshot, sino funcionalidades de backup real, es decir, sin vínculos directos con las imágenes productivas, que permitan recuperar dichas copias de seguridad incluso cuando la máquina virtual original haya sido borrada, además de permitir recuperar cualquiera de dichas copias sin afectación en la secuencia de protección. En resumen, una funcionalidad del tipo ‘Full Backup’.

Las funcionalidades mínimas exigidas serán:

- Proteger y recuperar máquinas virtuales independientes (no *Datastores*) con sobre escritura o generando una máquina virtual nueva (GUID nuevo, etc.).
- Recuperación de ficheros y carpetas.
- Integración con VMware Tools e integración directa con VSS de las máquinas virtuales para conseguir consistencia de aplicaciones.
- Modificación de la retención de una copia especifica.
- Eliminación de cualquier copia sin afectar a las secuencias de backup.
- RTO/ RPO garantizados o, al menos, identificar tiempos asumibles.
- No debe limitar el uso de la solución de *backup* existente.

En caso de que la solución propuesta no disponga de estas características de forma nativa, deberá proporcionar algún software adicional (incluyendo licenciamiento), certificado por el fabricante, que permita realizarlas. Dicho software adicional deberá ser compatible con el ya existente.

5.6 Distribución de datos en modo Stretched-cluster

De forma independiente a como sean instalados, inicialmente, los diferentes nodos que compongan la solución, será necesario que la propuesta integre las licencias/funcionalidades necesarias para establecer un modelo Stretched-Cluster para cada uno de los clústeres presentados.

A fin de que las ubicaciones de los nodos, y su modo de funcionamiento, puedan ser modificadas, sin incurrir en costes de licenciamientos adicionales durante el ciclo de vida de la solución propuesta.

5.7 Autosoporte y uso de Inteligencia Artificial para la detección precoz de incidencias

La solución ofertada deberá disponer de un servicio, o solución basada en Inteligencia Artificial que permita beneficiarse de todas las experiencias compartidas, de forma anónima, por otros clientes con entornos similares al ofertado.

Dicho servicio, o solución, deberá aportar informes, predicciones, recomendaciones e incluso que pueda llegar a abrir casos de soporte de forma automática y directa con el fabricante.

En caso de no disponer de una solución 'IA', la propuesta debe incluir el software adicional necesario (con licencias para todo el ciclo de vida) para, al menos cumplir los siguientes objetivos:

- Informes sobre el consumo de almacenamiento de cada máquina virtual.
- Informes de rendimiento del almacenamiento para cada máquina virtual.
- Monitorización del estado de salud de cada clúster presentado.
- Informes de predicción sobre el almacenamiento de la solución.
- Detección de fallo en los componentes físicos de la solución y apertura automática de incidencias con el fabricante de dichas piezas.

Estos objetivos deben ser proporcionados por la solución 'IA', o software adicional, de forma independiente a la información ofrecida por la propia consola del hipervisor (vCenter)

5.8 Otros requerimientos técnicos

La solución presentada debe cumplir los siguientes requerimientos generales sobre cualquiera de los clúster propuestos:

- Gestión unificada de toda la solución HCI, que permita una gestión y operación completa desde la consola del hipervisor (vCenter de VMware)
- Deduplicación, compresión y optimización de datos In-line, para el 100% de los mismos.
- Almacenamiento cifrado para el 100% de los datos.
- Sistema de inicio y datos ubicados en diferentes juegos de discos, cada uno de ellos gestionado por su propia controladora de discos.
- Disponer de motores que permitan a la solución balancear datos entre nodos para mejorar el rendimiento y no solo para el balanceo de datos por su volumen.
- Reconstrucción local (dentro del nodo) en la sustitución de un disco averiado.
- Proporcionar herramientas de protección de datos (Backup) y mantenimiento de copias en centro de contingencia.
- Conectores para copia directa a dispositivos de backup (ejem. CloudBank, HPE StoreOnce, etc.)
- Configuración de nodos heterogéneos dentro del mismo clúster.
- Configuraciones asimétricas entre clúster principal y centros de contingencia. por ejemplo 5 nodos en clúster principal y solo 2 en contingencia.
- Permitir el añadido de nodos de solo computación, que permitan aumentar los recursos CPU/RAM del clúster al que sean conectados.
- Permitir el añadido de almacenamiento externo (no-HCI) y que las máquinas virtuales puedan hacer uso del mismo.
- Permitir la configuración de almacenes de datos con menor protección (tolerancia a fallos) que la requerida, de forma simultánea a la exigida y en el mismo clúster
- Control o limitación de ancho de banda para las comunicaciones WAN.
- Conexión directa entre nodos (sin switch) en escenarios con solo 2 equipos.
- La solución debe ofrecer un soporte unificado, desde las piezas hardware, hasta el software HCI y el propio hipervisor.
- Proporcionar un sistema NAS gratuito, definido por software, de al menos de 4TB.
- El fabricante de la solución Hiperconvergente propuesta debe aportar por escrito algún tipo de GARANTIA que asegure van a cumplirse ciertos hitos, tales como: el % de ahorro de capacidad de almacenamiento, el tiempo en realizar un backup (ie.- 1 minuto para 1TB…), el número de pasos para hacer una copia de seguridad, restauración o clonación de una VM, el tiempo en crear políticas de copia de seguridad, o el tiempo de inactividad producido al quitar o añadir un nuevo nodo en la solución, etc.

5.9 Licencias de software

5.9.1 Licencias para HCI

- Tantas licencias sean necesarias para cubrir todos los nodos propuestos.

- Si la solución HCI dispone de diferentes ediciones o versiones, a fin de evitar posibles costes ocultos posteriores, deberán presentar la edición, o versión, más completa y rica en funcionalidades.

5.9.2 Licencias para virtualización

- Al menos 1 licencia de vCenter Standard
- Tantas licencias para ESXi como procesadores aporte cada solución ofertada.

Nota: software VMware vCenter Server y VMware vSphere debe contar con un periodo de licenciamiento de al menos 5 años.

5.9.3 Licencias para Sistemas Operativos

- Microsoft Windows Server 2019 DataCenter

Nota: Se deberán licenciar el 100% de los cores de cada nodo

5.10 Servicios de instalación, configuración y formación

Se deberán ofertar servicios cuyo alcance mínimo será el siguiente:

- Transporte a los CPD de Errotaburu, Loiola y Petritegi.
- Instalación física de los elementos en los CPD y del cableado intra-rack que conecte entre sí los diferentes elementos.
- Configuración básica de la solución para su puesta en marcha, siguiendo la documentación aportada por el fabricante de buenas prácticas con con vSphere como hipervisor.
- Instalación y configuración inicial del vCenter.
- Formación básica del fabricante en la plataforma ofertada que incluya gestión del entorno y procedimientos simples como:
    - Crear, borrar almacenamiento
    - Operaciones centradas en VMs (clonar, copiar, recuperar, rendimiento, …).

5.11 Soporte y garantía

La plataforma ofertada deberá contar con un contrato de soporte por parte del fabricante, con una duración de 5 años, donde se incluya resolución de incidencias, actualizaciones y, reposición de piezas averiadas y sustitución de las mismas in-situ.

El Proveedor deberá incluir el soporte de software y servicios de mantenimiento de licencias que aseguren que se dispone de las actualizaciones y las últimas versiones del firmware o software utilizado, durante los 5 años del período de vigencia de la garantía.