Pruebas de carga del sistema

Dirección General de Sistemas de Información y Comunicaciones

Áreas de Gobierno Tecnológico de SSII y del Dato

Servicio Andaluz de Salud
Consejería de Sanidad, Presidencia y Emergencias

Contenido

- Dirección General de Sistemas de Información y Comunicaciones
  - Áreas de Gobierno Tecnológico de SSII y del Dato
- Cumplimiento normativo
- Histórico de cambios
- Introducción
- Actores
- Flujo de trabajo
- Plan de pruebas
  - Diseño
  - Planificación
  - Prerrequisitos
  - Validación de las comunicaciones
  - Línea base del entorno
  - Métricas y requisitos
    - Métricas básicas: API de Servicios
    - Métricas básicas: Aplicaciones Web
    - Métricas detalladas
- Pruebas de capacidad

Resumen

- **Versión:** v02r12
- **Fecha publicación:** 05 de diciembre de 2023
- **Entrada en vigor desde:** 05 de diciembre de 2023

    - Resultados obtenidos
- Pruebas de rendimiento
    - Resultados obtenidos
- Pruebas de estrés
    - Resultados obtenidos
- Pruebas sostenidas
        - Resultados obtenidos
    - Pruebas sostenidas con tolerancia a fallos
- Conclusiones
- Entregables tras las pruebas
    - Documentación
    - Archivos
- Anexos
    - Anexo I: Volumetría
    - Anexo II: Casos de Uso
    - Anexo III: Plantilla PPS
    - Anexo IV: Lista de revisión de documentación
    - Anexo V: Hot Deployment
        - Requisitos y recomendaciones para afrontar con garantías un hot deployment en Oracle Weblogic:

Cumplimiento normativo

Las normas expuestas son de obligado cumplimiento. La STIC podrá atender casos excepcionales que serán gestionados a través de los responsables del proyecto correspondiente y autorizados por el Área de Gobernanza de la STIC. Asimismo, cualquier aspecto no recogido en estas normas deberá regirse en primera instancia por las guías técnicas correspondientes al esquema nacional de seguridad y esquema nacional de interoperabilidad, según correspondencia, y en su defecto a los marcos normativos y de desarrollo software establecidos por la Junta de Andalucía, debiendo ser puesto de manifiesto ante la STIC.

La STIC se reserva el derecho a la modificación de la norma sin previo aviso, tras lo cual, notificará del cambio a los actores implicados para su adopción inmediata según la planificación de cada proyecto.

En el caso de que algún actor considere conveniente y/o necesario el incumplimiento de alguna de las normas y /o recomendaciones, deberá aportar previamente la correspondiente justificación fehaciente documentada de la solución alternativa propuesta, así como toda aquella documentación que le sea requerida por la STIC para proceder a su validación técnica.

Contacto Dpto: Oficina de Calidad

Histórico de cambios

Los cambios en la normativa vendrán acompañados de un registro de las modificaciones. De este modo se podrá realizar un seguimiento y consultar su evolución.

Versión	v02r12	**Fecha publicación**	5 dic 2023	**Fecha entrada en vigor**	5 dic 2023
Alcance					

- Se actualiza el apartado ***Entregables tras las pruebas,*** incluyendo información sobre la estructura y versionado de los entregables.

Versión	v02r11	**Fecha publicación**	29 may 2023	**Fecha entrada en vigor**	29 may 2023
Alcance					

- Se clarifica lo estipulado en cuanto a volumetría y casos de uso.
- Se introducen criterios de normalización en las medidas de carga para facilitar su comparación.
- Se revisan los flujos de los procesos.
- Se añade un proceso para identificar anticipadamente incongruencias entre los casos de uso del plan de pruebas y los requisitos de volumetría.

Versión	v02r10	**Fecha publicación**	10 de marzo de 2021	**Fecha entrada en vigor**	10 de marzo de 2021
Alcance					

- Se incluye el apartado ***Documentación a entregar tras las pruebas***. En él se declara la documentación a entregar en Git tras las pruebas.

Versión	v02r09	**Fecha publicación**	26 de septiembre de 2018	**Fecha entrada en vigor**	26 de septiembre de 2018
Alcance					

- Flujo de trabajo y actores que intervienen en las pruebas de carga del sistema
- Plan de pruebas y métricas
- Se incluyen varios anexos de ayuda

Introducción

La normativa que a continuación se detalla muestra la metodología a seguir para llevar a cabo las pruebas de carga del sistema sobre los diferentes aplicativos basados en los modelos de arquitectura tecnológica de referencia publicados en Confluence.

En la metodología se muestra el flujo genérico que se debe seguir para planificar y ejecutar correctamente las pruebas de carga del sistema desde el inicio del desarrollo hasta su validación.

De esta forma, se indican el conjunto de pruebas fundamentales necesarias para garantizar el correcto funcionamiento y rendimiento, previo a la entrada a producción, de una nueva aplicación o de una evolución de un producto existente. Consta de los siguientes elementos:

1. Prerrequisitos: Los elementos y herramientas necesarias para poder implementar las pruebas.
2. Pruebas de comunicación: Comprobación del entorno objetivo para las pruebas por parte del lanzador.
3. Línea base: Establecimiento del conjunto inicial de datos necesario para la restauración del punto inicial en cada prueba.
4. Pruebas de capacidad: Determinan los parámetros de concurrencia para el resto de las pruebas.
5. Pruebas de rendimiento: Determinan el correcto funcionamiento del aplicativo con la carga en los distintos niveles de concurrencia definidos.
6. Pruebas de estrés: Se ejecuta la prueba con concurrencia incremental hasta el punto de saturación del aplicativo o aplicativos con respecto al entorno objetivo. Se mide también la recuperación del sistema.
7. Pruebas sostenidas: Comprueba el correcto funcionamiento y la ausencia de degradación del aplicativo durante un período de tiempo prolongado.
8. Pruebas sostenidas con tolerancia a fallos: Comprueba el impacto sobre el funcionamiento del aplicativo o aplicativos con operativas sobre determinados componentes del entorno: caída de una instancia de base de datos, caída de una instancia del servidor de aplicaciones o réplica de la aplicación, etc.
9. Conclusiones: Conjunto de conclusiones obtenidas a partir de los resultados e interpretaciones de las pruebas realizadas.

Esta normativa **se limita exclusivamente al proceso a seguir para la ejecución de pruebas de carga del sistema**, no siendo responsabilidad de esta normativa el resto de flujos que intervienen en el proceso de creación de un nuevo software. Su propósito por tanto es que sirva de guía para el usuario que se encarga de gestionar o ejecutar las pruebas de carga del sistema, existiendo un **documento adicional que hará de plantilla**, siendo dicha plantilla el documento que realmente deberán entregar a la STIC los proveedores de los aplicativos con los resultados de las pruebas obtenidas.

Actores

A continuación, se indican los actores partícipes en la metodología y su etiqueta que determina el liderazgo en cada proceso de la metodología:

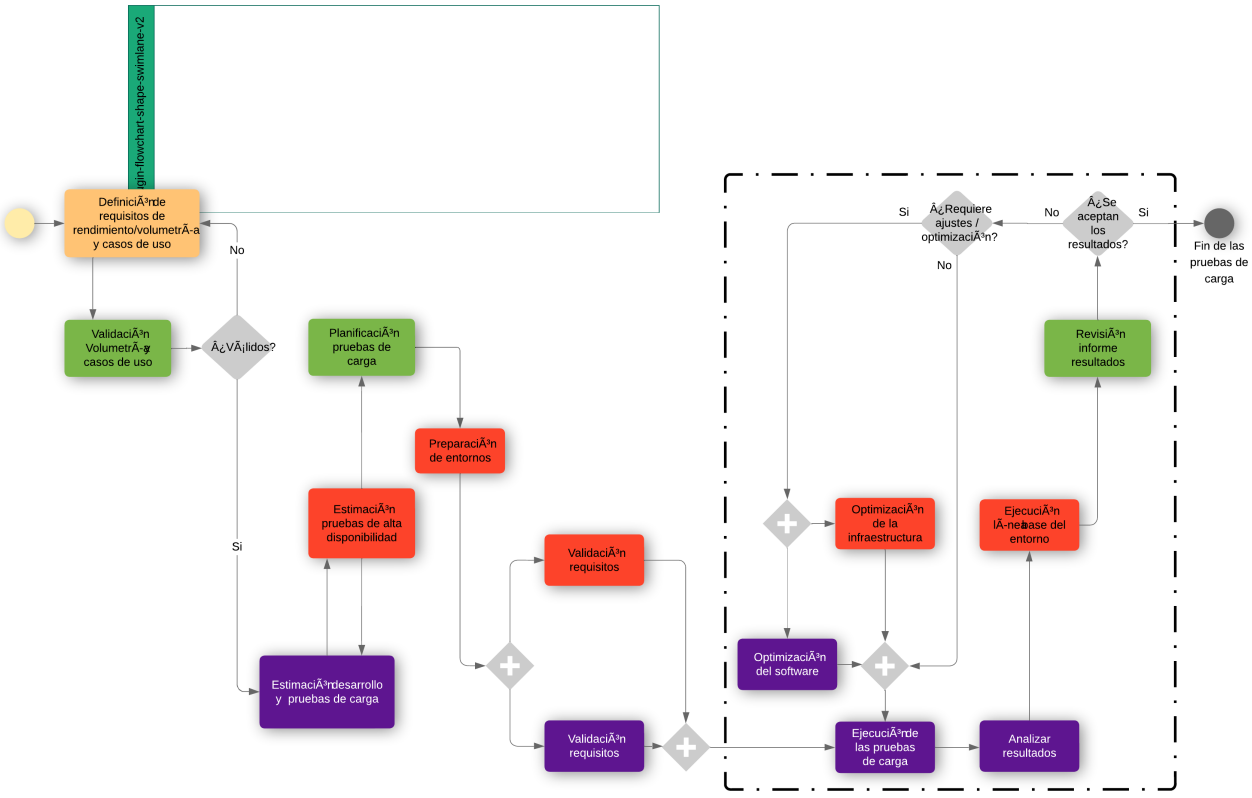
ACTOR	ETIQUETA
JEFE DE PROYECTO	NARANJA
GOBERNANZA	VERDE
SISTEMAS	ROJO
PROVEEDOR	MORADO
SISTEMAS / PROVEEDOR	AZUL

Flujo de trabajo

Es importante tener muy en cuenta las pruebas de carga de nuestros sistemas de información desde una fase temprana del desarrollo, estando seguros de los requisitos de rendimiento necesarios para la aceptación del producto por parte del SAS, y reservando el espacio de tiempo oportuno para poder llevar a cabo correctamente el conjunto de pruebas del sistema.

La responsabilidad de la coordinación del proceso de pruebas será labor de la OCA actuando como intermediario entre el proveedor de desarrollo, Sistemas y las diferentes áreas técnicas y de gestión de la STIC.

En el siguiente diagrama se puede ver una foto global del procedimiento a seguir para la planificación y ejecución de pruebas de sistemas sobre un sistema de información:



draw.io

Diagram attachment access error: cannot display diagram

Las actividades descritas en este flujo son:

- **Nuevo desarrollo**: Es el disparador del proceso, se refiere al momento, al inicio de un desarrollo, en el que surge la necesidad de realizar una estimación de un nuevo desarrollo. Ya desde el principio es **muy importante tener en cuenta las pruebas de carga del sistema**.
- **Definición de requisitos de rendimiento/volumetría y casos de uso**: Es responsabilidad del jefe de proyecto del sistema el **acordar con los responsables funcionales así como con las distintas áreas técnicas de la STIC** las necesidades y requerimientos de rendimiento que debe cumplir el sistema. Como resultado de este proceso, se obtendrán el apartado de métricas y los anexos I y II.
- **Validación volumetría y casos de uso:** El Área de Gobernanza validará la inclusión de requisitos de rendimiento/volumetría, su relevancia y que estos sean verificables con los casos de uso que se hayan incluido en el plan de pruebas.
- **Estimación del desarrollo y pruebas de carga**: El proveedor es el responsable de realizar la estimación que vea oportuna para el desarrollo del producto, **incluyendo** en dicha estimación el **tiempo necesario para realizar las pruebas de carga** aquí descritas.
- **Estimación pruebas de alta disponibilidad**: **Sistemas realizará una estimación** sobre el tiempo que necesita para hacer **pruebas de alta disponibilidad**. Esta planificación será añadida a la estimación global.
- **Planificación pruebas de carga**: La Oficina de Calidad realizará la coordinación entre las diferentes partes (Sistemas y el Proveedor) de la realización de las pruebas de carga.
- **Preparación de entornos**: Sistemas es el responsable de asegurar que el **entorno de pruebas** necesario estará correctamente **configurado de acuerdo** a las especificaciones indicadas en el **DGT** del producto. Igualmente, **Sistemas deberá asegurar la integridad del entorno** durante toda la fase de pruebas de carga.
- **Validación de requisitos previos**: Tanto el proveedor como Sistemas deberá revisar el conjunto de **requisitos previos explicados en el presente documento** con el fin de asegurar el correcto estado del entorno y poder así iniciar los ciclos de pruebas de carga del sistema.
- **Ejecución de las pruebas de carga**: Se refiere a la ejecución de forma ordenada de los diferentes tipos de pruebas de carga explicados en el presente documento.
- **Análisis de resultados**: El proveedor revisará y analizará los resultados de las pruebas de carga ejecutadas, **tra nsmitiendo al jefe de proyecto los resultados obtenidos** de cara a decidir si continuar con el proceso de pruebas o abortar el ciclo. Es necesaria una explicación clara de lo que indica cada tabla / gráfica, tanto por parte del proveedor, como de CTI, por lo que cada una de las tablas / gráficas, deberá venir personalizada incluyendo una explicación de los datos que se están arrojando.
- **Ejecución línea base del entorno**: Sistemas ejecutará el procedimiento necesario para **restaurar el entorno a su punto inicial** y poder así iniciar un nuevo ciclo de pruebas.
- **Revisión informe de resultados**: El Área de Gobernanza revisará el informe de resultados de las pruebas de carga realizadas.
- **¿Se aceptan los resultados?**: Se refiere a la **aprobación por parte del jefe de proyecto** de las pruebas ejecutadas hasta el momento, debiendo decidir entre continuar con las pruebas de carga, dar por finalizada la fase de pruebas, ya sea por haber llegado al último ciclo o por no tener resultados satisfactorios.

- **Optimización de la estructura:** Se refiere al proceso de ajustes y **optimización por parte de Sistemas** de la infraestructura ofrecida con el fin de solucionar posibles problemas de rendimiento o funcionamiento encontrados durante las pruebas de carga.
- **Optimización del software:** Se refiere al proceso de ajustes y **optimización del software por parte del proveedor** con el fin de solucionar posibles problemas de rendimiento o funcionamiento encontrados durante las pruebas de carga.

Plan de pruebas

Diseño

Uno de los aspectos más importante a la hora de plantear unas pruebas de carga del sistema es hacer un buen diseño de las pruebas a realizar, ya que estas deben ser representativas de la situación real del producto, esto es:

- **Casos de Usos**: Se deben identificar los casos de usos más representativos del sistema, aquellos usados con más frecuencia así como los casos de uso de nueva incorporación.
- **Volumetría**: Es imprescindible la definición de una volumetría semejante a la que se espera recibir en producción, ya que establecer una carga muy diferente (ya sea por encima o por debajo) no será representativo del estrés real que sufrirá la aplicación, haciendo que las pruebas realizadas no sean representativas.
- **Diseño de lanzadores de pruebas**: En la medida de lo posible, se debe añadir en las herramientas de pruebas aquellos controles necesarios para validar que el resultado obtenido en cada petición es el esperado (Por ejemplo, aserciones en jMeter).
- **Pruebas conectadas o desconectadas**: Un aspecto clave a la hora de diseñar un buen plan de pruebas de carga del sistema es decidir qué situación es más representativa con respecto a la situación real que vamos a encontrarnos en producción. En general, puede tener más sentido hacer pruebas de carga desconectadas (Donde la aplicación testeada no se conectará con terceras aplicaciones o lo hará mediante simuladores), pero la STIC junto al proveedor deberá decidir si tiene más sentido hacer pruebas de carga conectadas (Por ejemplo, cuando se está probando una aplicación en la que la mayor parte de sus funcionalidades hacen uso de servicios ofrecidos por terceras aplicaciones).

Planificación

Como un aspecto fundamental a la hora de diseñar un plan de pruebas de carga del sistema, debemos tener en cuenta los distintos tipos de pruebas necesarios para llevar a cabo el ciclo de pruebas completo. Por ello, en la siguiente captura se muestra una plantilla a modo de ejemplo con los **ítems que deben ser recogidos en la planificación de cualquier producto** que vaya a realizar pruebas de carga.

Modo de	Nombre de tarea	Duración	Comienzo	Fin	Predecesoras
	Pruebas de carga del sistema	**30 días**	**mar 03/04/18**	**lun 14/05/18**	
	Fase previa (Preparación y validación de entornos y lanzadores)	10 días	mar 03/04/18	lun 16/04/18	
	Pruebas de capacidad	3 días	mar 17/04/18	jue 19/04/18	2
	Pruebas de rendimiento	2 días	vie 20/04/18	lun 23/04/18	3
	Pruebas de estrés	2 días	mar 24/04/18	mié 25/04/18	4
	Pruebas sostenidas	3 días	jue 26/04/18	lun 30/04/18	5
	Pruebas de alta disponibilidad	10 días	mar 01/05/18	lun 14/05/18	6

Para iniciar el proceso de lanzamiento de pruebas de carga del sistema será necesario enviar correo electrónico a I-pruebasderendimiento.stic.sspa@juntadeandalucia.es con la siguiente información:

- **Planificación detallada por tipo de prueba a realizar** (se puede hacer referencia al apartado planificación del PPS). Debe incluir las pruebas de alta disponibilidad (sostenidas con tolerancia a fallos).
- **Horario en el que se realizarán las pruebas**, es importante indicar la franja horaria en la que se realizan las pruebas para poder contar con el soporte de Sistemas y sobre todo sí estas se realizan fuera de horario laboral (por ejemplo las cargas sostenidas se suelen hacer los fines de semana).
- **Nº de máquinas lanzadoras** (controladora + agentes) necesarias para la realización de las pruebas.
- **Usuario DMSAS-TEST** que realizará las pruebas, se utilizará para conceder el acceso a las máquinas.
- **Software y versión** necesario para realizar la pruebas.
- **Aplicaciones afectadas** por las pruebas de carga de sistemas, es necesario indicar las integraciones con aplicaciones que se realizan durante las pruebas.
- **PPS**, adjuntar en el correo el enlace al artículo Confluence del plan de pruebas de carga de sistemas actualizado.

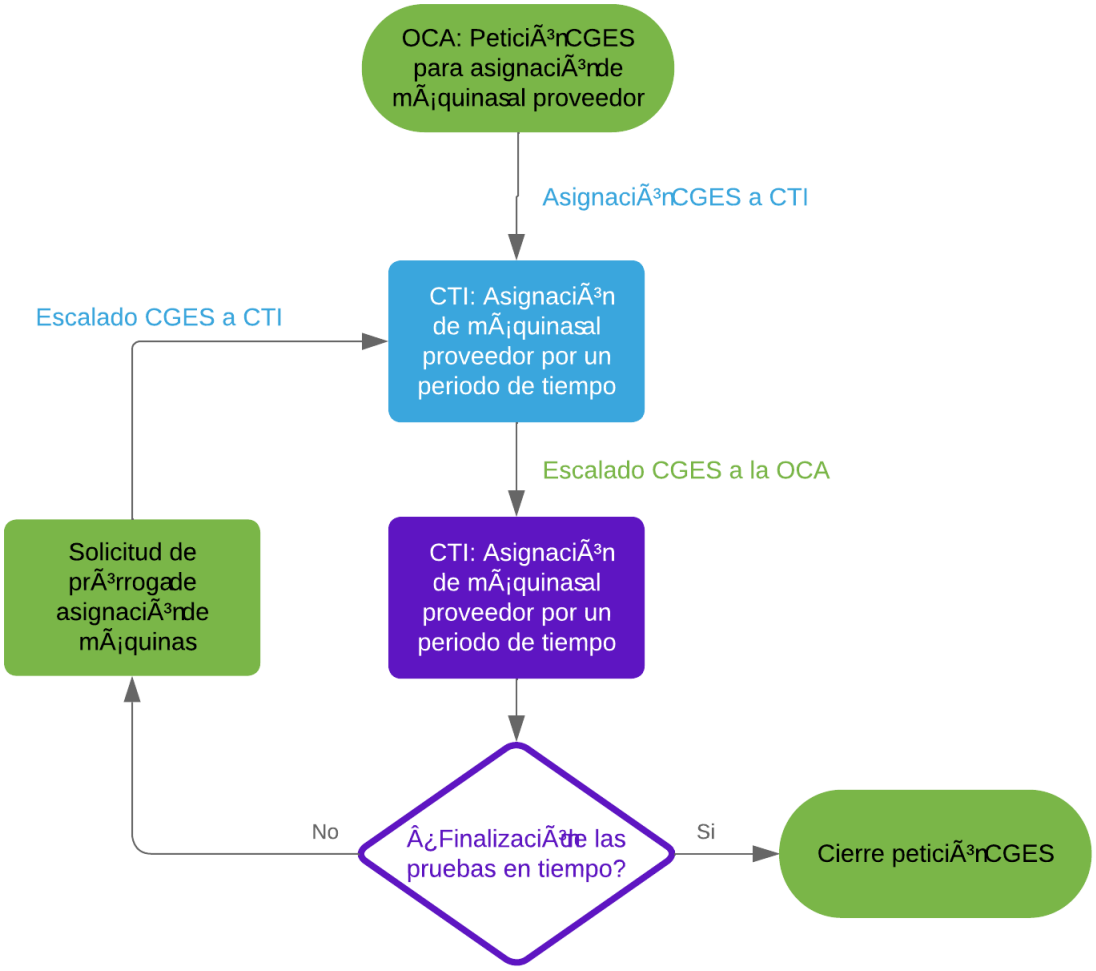
El correo se debe enviar con al menos un mes de antelación a la fecha prevista de inicio de las pruebas.

Una vez el Área de Gobernanza reciba la solicitud de ejecución de pruebas de rendimiento y estudie su viabilidad, realizará petición CGES a CTI para la asignación de las máquinas lanzadoras a los proveedores para así ejercer el control de la asignación de éstas durante todo el periodo de pruebas. **Esta asignación será por un periodo de tiempo determinado y de manera nominal** (preferiblemente del usuario que vaya a realizar las pruebas) y será informado mediante correo electrónico al proveedor encargado de realizar las pruebas. (Calendario)

La solicitud CGES contendrá la siguiente información:

- Motivo de la reserva.
- Usuario DMSAS-TEST que realizará las pruebas.
- Fecha inicio y fecha fin de la reserva.
- Número de máquinas.
- Nombre de las maquinas a reservar.
- Software y versión necesario para realizar la pruebas.

A continuación se muestra un diagrama con el flujo a seguir de dicha petición CGES:



draw.io

Diagram attachment access error: cannot display diagram

Si por algún motivo justificado fuese necesaria una prórroga del tiempo establecido inicialmente en la petición CGES, el usuario encargado de realizar las pruebas y que tiene asignadas las máquinas lanzadoras deberá solicitar mediante un correo electrónico a "l-pruebasderendimiento.stic.sspa@juntadeandalucia.es" una ampliación del tiempo de asignación. Dicha solicitud de prórroga será tramitada por la OCA según el flujo anterior, informando al ejecutor de las pruebas sobre la resolución de la misma.

Una vez finalizado el periodo de asignación de las máquinas lanzadoras y tras confirmación por parte del proveedor que las pruebas han finalizado, la OCA devolverá la petición CGES a CTI para que proceda a la desasignación de las mismas:

- **Se quitan los permisos al usuario DMSAS.**
- **Se elimina su perfil y se desinstala el software solicitado.**

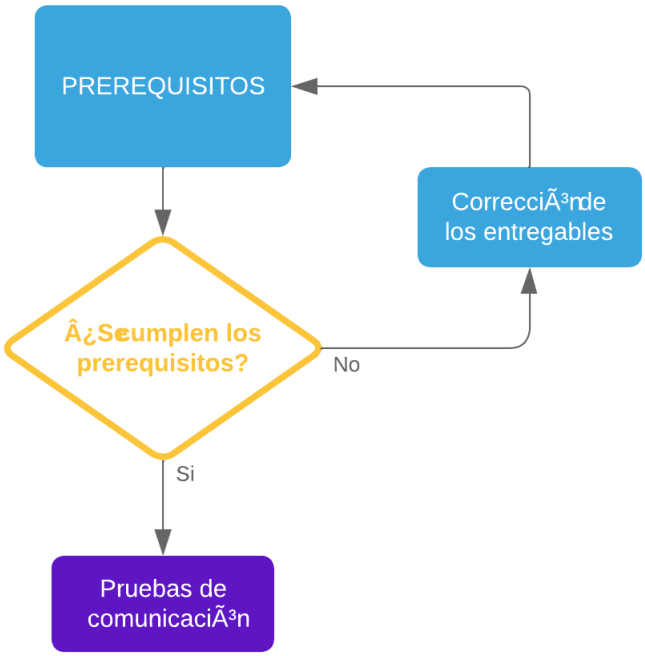
Todas las máquinas se encuentran alojadas en el entorno de preproducción.

El usuario que realiza las pruebas no tendrá en ningún momento permisos de administración sobre las máquinas lanzadoras.

En caso de ser necesario el reinicio de las máquinas lanzadoras, el proveedor generará una petición CGES la cual será asignada a CTI para que resuelva lo antes posible dicha petición y no produzca ninguna demora en la realización de las pruebas. Dicha petición CGES debe ser informada mediante correo electrónico a la lista l-pruebasderendimiento.stic.sspa@juntadeandalucia.es

Prerrequisitos

Los prerrequisitos son los elementos necesarios para una correcta ejecución de las pruebas. Es un proceso iterativo hasta que todos los actores acuerden los mismos y sean aceptados por la STIC.



draw.io

Diagram attachment access error: cannot display diagram

Aunque no todo el conjunto de prerrequisitos dependen exclusivamente del proveedor, si será **responsabilidad del proveedor** las tareas de gestión necesarias para desbloquear los posibles impedimentos que pueda encontrarse al iniciar el proceso de pruebas del sistema.

En la siguiente tabla se muestra el conjunto de prerrequisitos que el proveedor deberá validar y marcar como revisado. En el caso de no poder validar algunos de los requisitos deberá contactar con el coordinador de las pruebas:

REQUISITO	DESCRIPCIÓN	RESPONSABLE
Construcción del entorno	El entorno objetivo para la ejecución de las pruebas. Habitualmente el entorno de PREPRODUCCIÓN	SISTEMAS
Despliegue inicial del aplicativo	Los aplicativos receptores de las pruebas deben estar desplegados y los elementos de sistema necesarios de los que depende (JMS, Datasources, Vistas Materializadas, etc.)	SISTEMAS
Monitorización del entorno	Los elementos constituyentes del entorno deben presentar la monitorización necesaria para la obtención de métricas de sistema. Esto incluye SCOM (BeanSpy), grabaciones de JVM y estadísticas de BD.	SISTEMAS
Validación de la aplicación o aplicaciones desplegadas	El proveedor tiene que validar el aplicativo o aplicativos desplegados en el entorno objetivo.	PROVEEDOR
Indicadores de negocio	Métricas que determinan la productividad de la aplicación o aplicaciones a probar.	PROVEEDOR / SISTEMAS
Conjunto inicial de datos	El entorno objetivo debe presentar los datos para las pruebas. Habitualmente se implementa con la importación de datos desde producción y su enmascarado. El procedimiento debe ser definido por el PROVEEDOR y ejecutado por Sistemas.	PROVEEDOR / SISTEMAS
Herramienta lanzadora	JMETER instalado y disponible para que el proveedor pueda ejecutar las pruebas.	SISTEMAS
Conjunto de casos de usos	Paquetes que implementan los casos de usos protagonistas en las pruebas. Paquete a cargar en JMETER. (Ver Anexo I)	PROVEEDOR / STIC
Usuarios nominativos	Creación de usuarios nominativos (solo lectura) para los servidores (p. ej. WebLogic) y la base de datos (p. ej. Oracle). El PROVEEDOR debe indicar los usuarios necesarios.	SISTEMAS
Acceso a entorno objetivo	Los participantes del PROVEEDOR deben tener acceso al entorno objetivo. El PROVEEDOR debe indicar el mapa de comunicación necesario.	SISTEMAS
Configuración de réplica de sesiones	Si el entorno objetivo consta de clúster es necesaria la réplica de sesiones para garantizar la alta disponibilidad	PROVEEDOR
Ausencia de conflictos de clases	Un despliegue con conflictos de clases genera incertidumbre en las pruebas. En cada momento la versión de librerías utilizadas puede no ser la adecuada. Ejecución de informe Class Loader Analysis Tool (CAT) con 0	PROVEEDOR / SISTEMAS

Validación de las comunicaciones

Las pruebas de comunicación son necesarias para que el PROVEEDOR pueda interactuar con el entorno y pueda probar las comunicaciones necesarias para las pruebas. Es un proceso iterativo hasta que el PROVEEDOR considere necesario.

Es importante **no confundir** con las comunicaciones entre aplicativos que los sistemas podrían realizar, estas pruebas se refieren **SOLO** a las pruebas que el proveedor necesita realizar para poder acceder a la herramienta y poder así hacer un seguimiento de las pruebas de carga realizadas.

La siguiente tabla muestra el listado de comprobaciones realizadas:

[Ejemplo. Reemplazar por tabla real]

SISTEMA	COMPONENTE	ACEPTACIÓN
CITACIÓN	WEB	¿
CITACIÓN	BASE DE DATOS	¿
PRESCRIPCIÓN	WEB	¿
¿		

Línea base del entorno

En este proceso se recogen las tareas necesarias para poder reproducir el entorno original antes de cualquier prueba de carga. El responsable de **ejecutar** las tareas de este proceso es Sistemas, pero **será responsabilidad del PROVEEDOR el definir los pasos** que necesita ejecutar para la restauración del entorno.

Las tareas son las siguientes:

TAREA	DESCRIPCIÓN	RESPONSABLE	ACEPTACIÓN
Copia de seguridad en frío de los servidores	Copia de seguridad de todos los servidores del entorno objetivo. Lo habitual es que los servidores sean independientes del negocio y no requieran su restauración. El backup puede ser por NetBackup o a través de Hyper-V	SISTEMAS	¿
Copia de seguridad de la base de datos	Backup RMAN de la base de datos del entorno objetivo. Si la base de datos está sin archivado será un backup offline.	SISTEMAS	¿
Snapshot de los volúmenes de la base de datos	Snapshot a nivel de Veritas de los volúmenes de la base de datos.	SISTEMAS	¿
Creación de procedimiento	Definición del procedimiento de restauración del entorno entre pruebas. Será necesaria una prueba de este procedimiento. (Este proceso lo definirá el proveedor, pudiendo ir desde una restauración de un backup hasta la ejecución de unos scripts que sitúe el sistema en una situación válida para lanzar un nuevo ciclo de pruebas)	PROVEEDOR	¿

[En caso de necesitar tareas adicionales, estas deben ser indicadas aquí

Ejemplo: Si la prueba va a generar cambios en otras aplicaciones, se debe hacer backup de esa otra aplicación para poder restaurarla en caso de ser necesario.]

Métricas y requisitos

Uno de los primeros pasos a realizar a la hora de abordar un conjunto de pruebas de carga de un sistema es identificar correctamente el conjunto de pruebas y métricas, de forma que sea lo suficientemente representativas para reflejar fielmente el funcionamiento del sistema ante una elevada carga de usuarios.

Por ello, uno de los requisitos obligatorios será que el nuevo desarrollo sea capaz de soportar la carga de usuarios a los que deberá dar servicio. Los umbrales a superar serán:

- En el caso de productos que ya están en producción o que reemplacen a alguno existente conservando, en su mayor parte, su funcionalidad, el desarrollo deberá soportar una carga de transacciones que sea un **25% superior a la carga máxima** recibida en horario de trabajo indicada en el Anexo I. Como norma general, se considerará el periodo comprendido entre las 9:00 y las 14:00 de lunes a viernes como horario de trabajo o los 6 intervalos de 1 hora con mayor carga si estos no coinciden con el horario de trabajo habitual.
- Para los productos que no estén en producción también se establecerá un **25% superior** a la estimación inicial realizada por el proveedor indicada en el Anexo I.

Se proponen dos conjuntos de posibles métricas a realizar dependiendo del tipo de prueba a realizar, definiendo en función del caso un conjunto de métricas básicas a medir.

Las métricas básicas se refieren al conjunto de métricas dirigidas a medir los tiempos de respuestas y buen funcionamiento del sistema. Estas métricas deben ser medidas en las diferentes tipos de pruebas que se proponen en el presente documento, debiéndose incluir en el informe final generado los resultados obtenidos en las mediciones de cada una de las pruebas.

IMPORTANTE: Es especialmente importante destacar la necesidad de diseñar indicadores de negocio específicos de cada aplicación y que serán de utilidad para conocer el buen funcionamiento del sistema y la validez de las pruebas (Por ejemplo, si estamos probando un sistema en el que sabemos que en producción se crean 1000 episodios por hora, y nuestro indicativo de "Episodios creados por hora" nos da un resultado de 100 episodios, sabemos que los resultados no son correctos, puesto que no está simulando una situación real.

Métricas básicas: API de Servicios

Las pruebas de carga del sistema orientadas a testear herramientas que principalmente ofrecen servicios web mediante API (Como por ejemplo es Maco, Estructura o BDU) deben ir dirigidas a testear cada servicio web que se ofrece.

Este tipo de servicios se caracterizan por ser rápidos, ágiles y concretos, dedicándose a un conjunto reducido de funcionalidades específicas y obteniendo tiempos de respuesta reducidos, por lo que las exigencias en cuanto a requisitos de respuesta son mayores que el caso de aplicaciones web.

Para este tipo de aplicaciones se propone medir las siguientes métricas básicas:

Para REST:

MÉTRICA	PROPORCIONADO POR	Valor esperado(*)
Indicador de negocio 1 cada 10 mins	Monitorización	A definir
Indicador de negocio 2 cada 10 mins	Monitorización	A definir
¿	Monitorización	A definir
Indicador de negocio n cada 10 mins	Monitorización	A definir
Tiempo medio de respuesta por cada llamada a la API de servicios	Herramienta lanzadora	< 300ms
Tiempo de respuesta del percentil 90 por cada llamada a la API de servicios	Herramienta lanzadora	< 600ms
Otras mediciones básicas dependiendo del caso de uso	N/A	A definir

Para Soap:

MÉTRICA	PROPORCIONADO POR	Valor esperado(*)
Indicador de negocio 1 cada 10 mins	Monitorización	A definir
Indicador de negocio 2 cada 10 mins	Monitorización	A definir
¿	Monitorización	A definir
Indicador de negocio n cada 10 mins	Monitorización	A definir
Tiempo medio de respuesta por cada llamada a un servicio Asíncrono	Herramienta lanzadora	< 500ms
Tiempo medio de respuesta por cada llamada a un servicio Síncrono	Herramienta lanzadora	< 1seg
Tiempo de respuesta del percentil 90 por cada llamada a un servicio Síncrono	Herramienta lanzadora	< 2seg
Otras mediciones básicas dependiendo del caso de uso	N/A	A definir

(*): Es importante tener en cuenta que los valores marcados en la presente tabla como valores objetivos se refiere al conjunto de peticiones REST o Soap de operaciones sobre datos "típicas". Si la aplicación a testear cuenta con alguna funcionalidad específica de la que se espere un tiempo de respuesta mayor, debe indicarse en la siguiente tabla, aportando en cada caso una justificación de porqué se espera un tiempo de respuesta mayor al indicado como referencia en este documento:

Petición	Valor esperado	Justificación
(Ejemplo) /MiAplicacion /autenticación/login	*< 2seg*	*El proceso de autenticación depende de múltiples servicios externos, cuyos tiempos de respuesta están en torno a 1 segundo.*

Métricas básicas: Aplicaciones Web

Las aplicaciones web se suelen caracterizar por contar con una división entre la capa visual y otras capas que gestionan el negocio de la aplicación. Los frameworks de desarrollo web ofrecen un conjunto de funcionalidades (gestión de la sesión, páginas cacheadas en servidor, etc.) que hacen que los tiempos de respuestas en peticiones web suelan ser superiores a las aplicaciones que ofrecen API de servicios.

En el caso de aplicaciones web, las peticiones al servidor suelen dar como resultado respuestas de mayor tamaño que las API de servicios, puesto que en la misma respuesta se debe ofrecer toda la información necesaria para representar visualmente el contenido (HTML).

Para este tipo de aplicaciones se propone medir las siguientes métricas básicas:

MÉTRICA	PROPORCIONADO POR	Valor esperado(*)
Indicador de negocio 1 cada 10 mins	Monitorización	A definir
Indicador de negocio 2 cada 10 mins	Monitorización	A definir
¿	Monitorización	A definir
Indicador de negocio n cada 10 mins	Monitorización	A definir
Tiempo medio de respuesta por cada llamada al servidor Web	Herramienta lanzadora	< 1seg
Tiempo de respuesta del percentil 90 por cada llamada al servidor Web	Herramienta lanzadora	< 3seg
Otras mediciones básicas dependiendo del caso de uso	N/A	A definir

(*): Es importante tener en cuenta que los valores marcados en la presente tabla como valores objetivos se refiere al conjunto de peticiones de navegación típicas de cualquier aplicación web. Si la aplicación a testear cuenta con alguna funcionalidad específica de la que se espere un tiempo de respuesta mayor, debe indicarse en la siguiente tabla, aportando en cada caso una justificación de porqué se espera un tiempo de respuesta mayor al indicado como referencia en este documento:

Petición	Valor esperado	Justificación
(Ejemplo) /MiAplicacion /listado.jsf	*< 4seg*	*La pantalla de listados muestra resultados masivos, debiendo hacer consultas a base de datos de alto coste, por lo que los tiempos de respuesta estarán por encima de lo indicado como referencia.*

Métricas detalladas

Este conjunto de métricas van destinadas a profundizar en el comportamiento del sistema desde un punto de vista de sistemas, donde se pretende analizar que los diferentes componentes internos que forman parte de la arquitectura del sistema (Servidor, bases de datos, conexiones, etc.) funcionan correctamente.

El resultado de las mediciones obtenidas debe analizarse en las diferentes pruebas, **debiéndose incluir en este informe las anomalías que se encuentren en dicho análisis**. Igualmente, el **proveedor podrá incluir en cualquiera de las pruebas el resultado de una o varias de estas métricas** si lo considera necesario para comprender el informe generado.

Las métricas establecidas para esta prueba son:

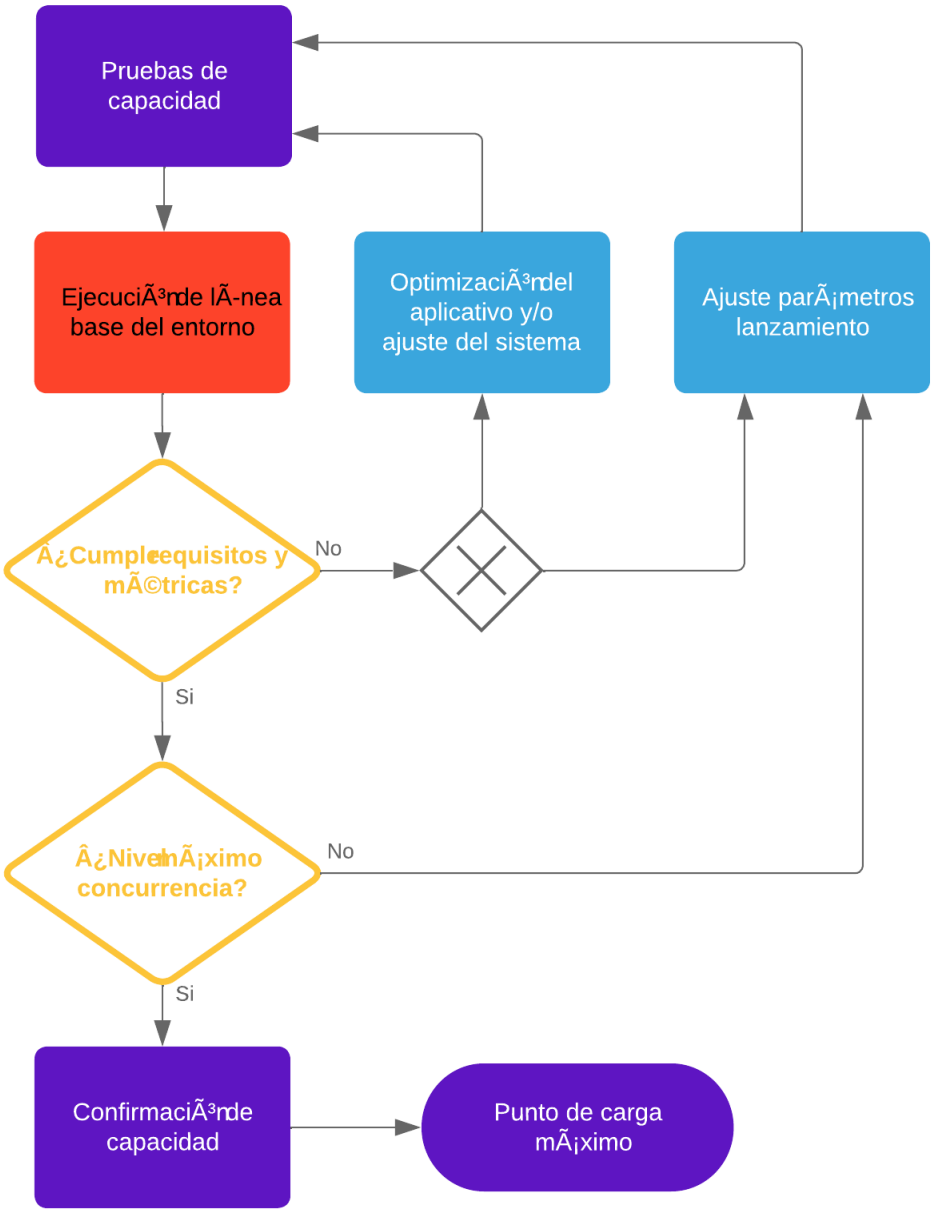
MÉTRICA	PROPORCIONADO POR
Estado de los destinos JMS cada 10 mins	Monitorización
Estado y consumo CPU de los hilos de instancia WebLogic/contenedor cada 5 mins	Thread dumps de JVM
Consumo medio de CPU de instancia WebLogic/contenedor cada 10 mins	Grabación de JVM
Consumo máximo de CPU de instancia WebLogic/contenedor cada 10 mins	Grabación de JVM
Gráfico de consumo medio de Java Heap cada 10 mins	Grabación de JVM
Conexiones activas media de orígenes de datos cada 10 minutos	Monitorización
Conexiones activas máxima de orígenes de datos cada 10 minutos	Monitorización
Fuga de conexiones de orígenes de datos cada 10 minutos	Monitorización
Peticiones de conexiones de orígenes de datos en espera cada 10 minutos	Monitorización
Top consultas por tiempo de ejecución cada 10 minutos	Estadísticas de base de datos
Top consultas por lecturas de memoria cada 10 minutos	Estadísticas de base de datos
Top consultas por lecturas físicas cada 10 minutos	Estadísticas de base de datos
Top consultas por tiempo de clúster cada 10 minutos	Estadísticas de base de datos
Load Profile cada 10 minutos	Estadísticas de base de datos
Instance Efficiency cada 10 minutos	Estadísticas de base de datos
E/S de Datafiles cada 10 minutos	Estadísticas de base de datos
E/S de Segmentos cada 10 minutos	Estadísticas de base de datos
Time Model cada 10 minutos	Estadísticas de base de datos
Eventos de base de datos cada 10 minutos	Estadísticas de base de datos
Gráfico con actividad de GC cada 10 minutos	Grabación de JVM
Mensajería en servidor JMS cada 10 minutos	Monitorización
Bytes en servidor JMS cada 10 minutos	Monitorización
Tiempo de peticiones sobre WebLogic/servidor integrado	Access log format W3C Extended
Evolución de los pooles de la memoria HEAP (TLAB) cada 10 minutos	Grabación de JVM

Pruebas de capacidad

Se trata de la ejecución de casos de uso conjuntos y con concurrencia escalable de usuarios y con incremento paulatino de forma iterativa en un escenario de duración determinada a petición del PROVEEDOR. Si se dispone de datos de carga (Volumetría) de un sistema preexistente, se recomienda que la duración del lanzamiento sea la misma para facilitar la comparación de los resultados obtenidos. Por ejemplo, si los datos de volumetría de los que se dispone vienen agrupados en periodos de una hora (accesos por hora), la duración de la ejecución debiera ser de una hora.

El objetivo de estas pruebas es buscar los parámetros de configuración adecuados para los lanzadores de pruebas (jMeter), realizándose una carga progresiva sobre los sistemas a probar en busca del **punto de carga máximo sin degradar el servicio**, es decir, sin dejar de cumplir los requisitos particulares del sistema y las métricas establecidas. De nuevo, para facilitar la comparación de los datos obtenidos, si los datos de volumetría de los que se dispone vienen expresados en base a una métrica (p. ej.: peticiones por segundo), los datos del punto de carga máximo deberían expresarse en la misma métrica. Con el objetivo de armonizar los resultados y, teniendo en cuenta la información que recopila JMeter en sus informes, se propone la P*roductividad (Transacciones/s)* como métrica para la volumetría y el punto de carga máximo.

Debido a que los niveles de concurrencia encontrados en estas pruebas serán la **base para las siguientes pruebas**, una vez encontrado el punto máximo de carga sobre un sistema, es necesario **ejecutar el plan de pruebas una segunda vez** para confirmar que los datos observados son correctos.



draw.io

Diagram attachment access error: cannot display diagram

La iteración del proceso dependerá de la autorización de la STIC, los problemas encontrados que necesiten algún ajuste, o que el PROVEEDOR no haya encontrado aún los niveles de concurrencia.

Resultados obtenidos

*[El presente apartado se muestra únicamente a **modo de ejemplo**, debiendo ser cumplimentado al COMPLETO por el PROVEEDOR con los resultados, tanto de métricas básicas como de métricas detalladas, obtenidos en sus pruebas de carga.]*

La aplicación X tiene como requisito el tener que soportar una carga máxima de 200 usuarios concurrentes, por ello se ha configurado la herramienta lanzadora para generar 100 hilos por cada uno de los 2 casos de usos programados para ser lanzado.

Es necesaria una explicación clara de lo que indica cada tabla / gráfica, tanto por parte del proveedor, como de CTI, por lo que cada una de las tablas / gráficas, deberá venir customizada incluyendo una explicación de los datos que se están arrojando.

Tras la ejecución de la lanzadera durante 30 minutos, se ha obtenido un resultado estable de tiempos de respuestas que puede visualizarse en las siguientes gráficas (ordenada por tiempo medio de respuesta de mayor a menor):

Requests	Executions			Response Times (ms)					
Label	#Samples	KO	Error %	Average	Min	Max	90th pct	95th pct	99th pct
Total	**5066**	**0**	**0.00%**	**767.57**	**16**	**7344**	**1743.60**	**2362.50**	**3957.31**
Login	400	0	0.00%	1714.20	625	6171	2836.00	3438.70	4936.69
Modificar concepto - Enviar datos	382	0	0.00%	1076.74	219	7344	2229.60	3420.45	6000.96
Cambio de perfil	1568	0	0.00%	900.57	127	5113	1752.00	2254.55	3717.70
Modificar concepto - Cerrar panel	381	0	0.00%	783.07	160	5438	1683.00	2303.80	4439.32
Modificar concepto - Carga panel	382	0	0.00%	758.11	148	6913	1700.70	2379.55	4559.42
Listado Formularios (Pantalla bienvenida)	398	0	0.00%	691.26	136	4769	1355.00	2008.50	3040.70
Listado Conceptos Fragmento	386	0	0.00%	653.17	117	6289	1398.40	2307.45	3973.86
Listado Conceptos Pagina	386	0	0.00%	451.40	88	3715	946.80	1283.55	2872.78
Logout	379	0	0.00%	155.04	23	2650	409.00	653.00	1315.00
Acceso	404	0	0.00%	77.27	16	1584	163.50	393.75	824.65

Tras el lanzamiento de estas pruebas se realizan las siguientes comprobaciones:

- *No se observan errores en el log de la aplicación*
- *El funcionamiento de la base de datos ha sido correcto, no detectándose cargas elevadas en CPU y Memoria.*

Host CPU (CPUs: 8 Cores: 8 Sockets: 1)

Load Average Begin	Load Average End	%User	%System	%WIO	%Idle
2.55	2.10	28.5	12.8	3.1	58.4

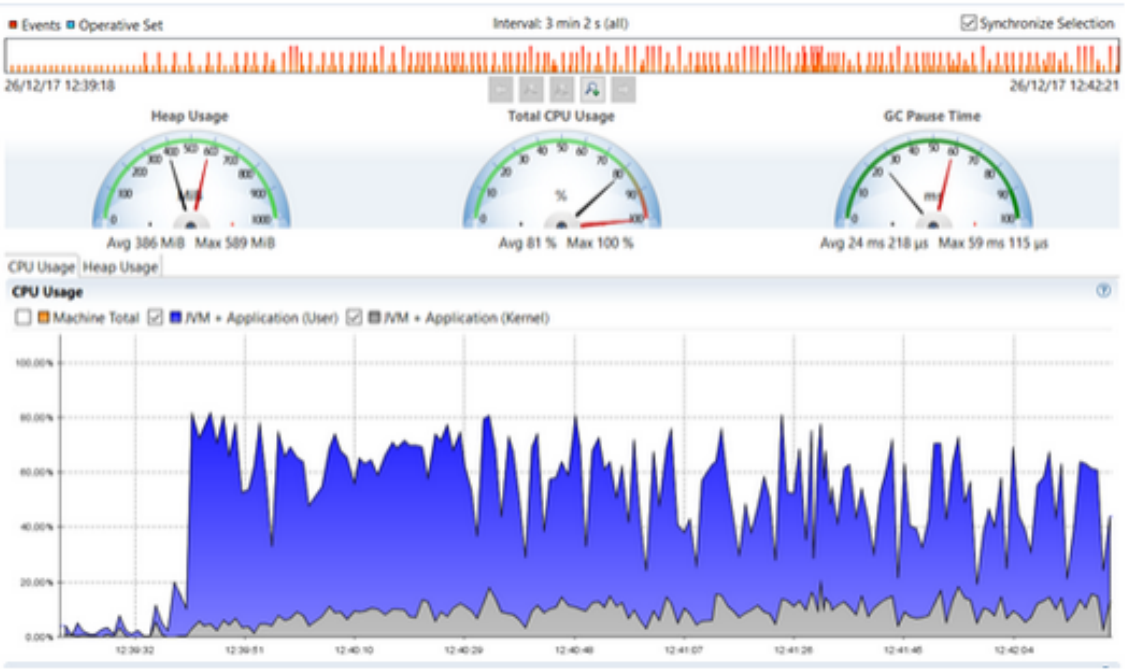
Instance CPU

%Total CPU	%Busy CPU	%DB time waiting for CPU (Resource Manager)
41.2	99.0	0.0

Memory Statistics

	Begin	End
Host Mem (MB):	16,048.8	16,048.8
SGA use (MB):	5,097.1	5,097.1
PGA use (MB):	1,038.0	1,035.3
% Host Mem used for SGA+PGA:	38.23	38.23

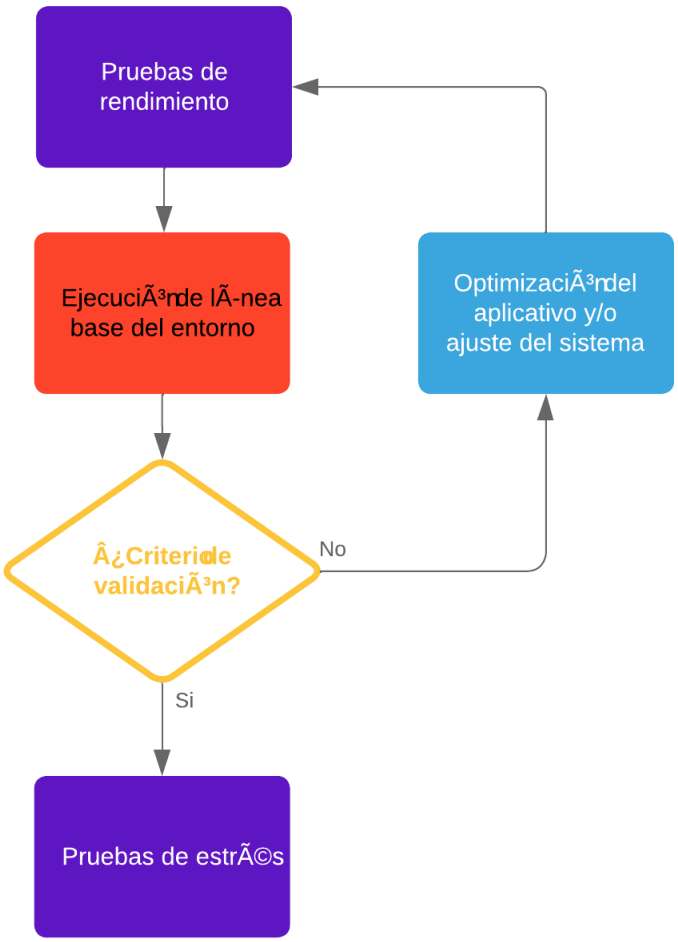
- *En los servidores de aplicaciones no se detectan cargas significativas.*



Dados los datos obtenidos, se dan por válidas las pruebas de capacidad, pudiendo afrontar las siguientes pruebas definidas en el presente documento.

Pruebas de rendimiento

Se trata de la ejecución de casos de uso conjuntos y con concurrencia de hilos determinada por el PROVEEDOR de forma iterativa en un escenario de duración determinada por las pruebas de capacidad. El objetivo es **observar el comportamiento del producto con diferentes niveles de concurrencia**. Las pruebas de rendimiento son fundamentales.



draw.io

Diagram attachment access error: cannot display diagram

La iteración del proceso dependerá de la autorización de la STIC y los problemas encontrados que necesiten algún ajuste y/u optimización del aplicativo o aplicativos.

Previo acuerdo con el PROVEEDOR y autorización de la STIC, los niveles de concurrencia recomendables serían los obtenidos por la prueba de capacidad. Lo habitual es que la prueba consista en tres iteraciones: una con el nivel de concurrencia máximo, otra con el nivel de concurrencia óptimo y una última con un nivel bajo.

En la siguiente tabla puede verse el conjunto de pruebas a realizar:

PRUEBA	CONCURRENCIA	DURACIÓN (HORAS)	JUSTIFICACIÓN
Concurrencia alta	80% de la concurrencia máxima determinada por las pruebas de capacidad	Determinado por el PROVEEDOR	Justificación de la duración establecida
Concurrencia media	60% de la concurrencia máxima determinada por las pruebas de capacidad	Determinado por el PROVEEDOR	Justificación de la duración establecida
Concurrencia baja	40% de la concurrencia máxima determinada por las pruebas de capacidad	Determinado por el PROVEEDOR	Justificación de la duración establecida

Resultados obtenidos

*[El presente apartado se muestra únicamente a **modo de ejemplo**, debiendo ser cumplimentado al COMPLETO por el PROVEEDOR con los resultados, tanto de métricas básicas como de métricas detalladas, obtenidos en sus pruebas de carga.]*

Es necesaria una explicación clara de lo que indica cada tabla / gráfica, tanto por parte del proveedor, como de CTI, por lo que cada una de las tablas / gráficas, deberá venir customizada incluyendo una explicación de los datos que se están arrojando.

Partiendo de los niveles de concurrencia definidos en las pruebas de capacidad (200 hilos concurrentes), se definen las siguientes pruebas de rendimiento a realizar:

PRUEBA	*CONCURRENCIA*	*DURACIÓN (HORAS)*
Concurrencia alta	*160 hilos*	*1*
Concurrencia media	*120 hilos*	*1*
Concurrencia baja	*80 hilos*	*1*

Obteniendo los siguientes resultados:

Concurrencia alta (160 hilos):

Requests	Executions			Response Times (ms)					
Label	#Samples	KO	Error %	Average	Min	Max	90th pct	95th pct	99th pct
Total	**5066**	**0**	**0.00%**	**767.57**	**16**	**7344**	**1743.60**	**2362.50**	**3957.31**
Login	400	0	0.00%	1714.20	625	6171	2836.00	3438.70	4936.69
Modificar concepto - Enviar datos	382	0	0.00%	1076.74	219	7344	2229.60	3420.45	6000.96
Cambio de perfil	1568	0	0.00%	900.57	127	5113	1752.00	2254.55	3717.70
Modificar concepto - Cerrar panel	381	0	0.00%	783.07	160	5438	1683.00	2303.80	4439.32
Modificar concepto - Carga panel	382	0	0.00%	758.11	148	6913	1700.70	2379.55	4559.42
Listado Formularios (Pantalla bienvenida)	398	0	0.00%	691.26	136	4769	1355.00	2008.50	3040.70
Listado Conceptos Fragmento	386	0	0.00%	653.17	117	6289	1398.40	2307.45	3973.86
Listado Conceptos Pagina	386	0	0.00%	451.40	88	3715	946.80	1283.55	2872.78
Logout	379	0	0.00%	155.04	23	2650	409.00	653.00	1315.00
Acceso	404	0	0.00%	77.27	16	1584	163.50	393.75	824.65

Concurrencia media (120 hilos):

Label	#Samples	KO	Error %	Average response time	90th pct	95th pct	99th pct
Total	**910**	**0**	**0.00%**	**394.44**	**779.90**	**1033.65**	**1827.89**
Login	70	0	0.00%	1024.96	1625.10	1935.45	3699.00
Modificar concepto - Enviar datos	70	0	0.00%	501.76	776.50	1784.55	2586.00
Cambio de perfil	280	0	0.00%	456.73	766.70	1003.45	1492.50
Listado Formularios (Pantalla bienvenida)	70	0	0.00%	379.79	609.40	984.60	2425.00
Modificar concepto - Carga panel	70	0	0.00%	355.16	590.00	710.25	1258.00
Modificar concepto - Cerrar panel	70	0	0.00%	346.89	537.90	822.95	1163.00
Listado Conceptos Fragmento	70	0	0.00%	287.80	446.00	577.10	1111.00
Listado Conceptos Pagina	70	0	0.00%	240.61	369.10	866.90	1177.00
Logout	70	0	0.00%	108.13	217.00	342.70	1827.00
Acceso	70	0	0.00%	55.70	151.70	186.90	495.00

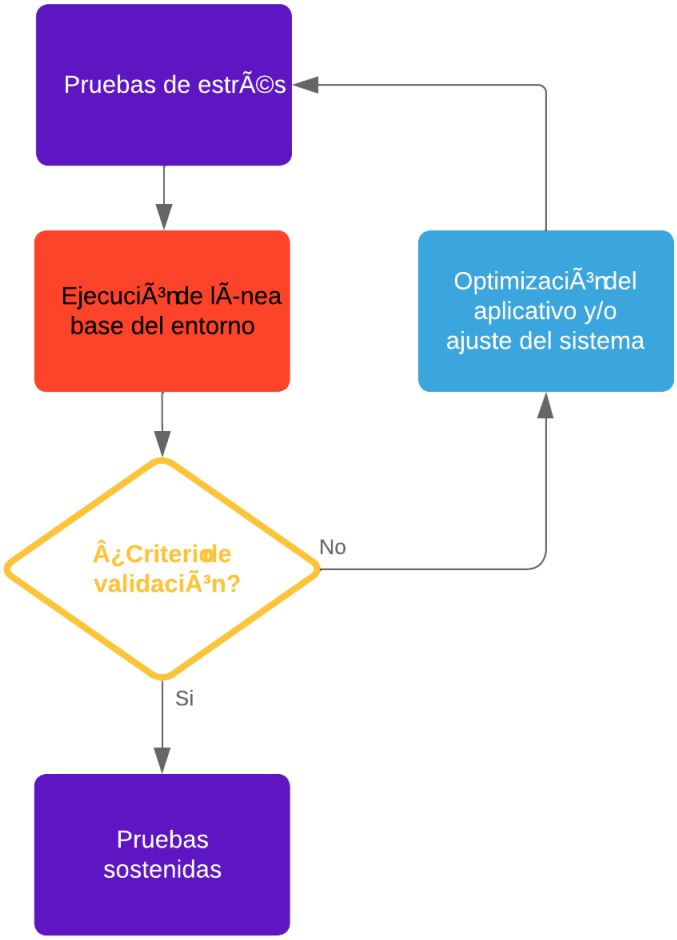
Concurrencia baja (80 hilos):

Label	#Samples	KO	Error %	Average response time	90th pct	95th pct	99th pct
Total	**650**	**0**	**0.00%**	**390.08**	**854.70**	**1186.75**	**2182.33**
Login	50	0	0.00%	1170.50	2140.60	2714.45	3767.00
Cambio de perfil	200	0	0.00%	444.58	852.70	1194.25	2007.69
Modificar concepto - Enviar datos	50	0	0.00%	415.18	734.70	1002.30	1375.00
Modificar concepto - Carga panel	50	0	0.00%	380.36	648.80	1514.55	3070.00
Listado Formularios (Pantalla bienvenida)	50	0	0.00%	363.26	608.20	866.40	1058.00
Modificar concepto - Cerrar panel	50	0	0.00%	340.24	624.50	778.90	1875.00
Listado Conceptos Fragmento	50	0	0.00%	256.14	541.70	744.45	821.00
Listado Conceptos Pagina	50	0	0.00%	247.40	483.90	776.80	967.00
Logout	50	0	0.00%	83.28	128.90	398.35	866.00
Acceso	50	0	0.00%	36.28	78.00	129.05	134.00

Se aprecia una bajada importante de los tiempos de respuesta entre la concurrencia alta y la concurrencia media, lo cual nos hace ver que la concurrencia alta definida en las pruebas de capacidad está muy cerca del límite de los servidores actuales. Se recomienda asignar mayor cantidad de RAM a los servidores de weblogic, ya que el gráfico de carga muestra que se llegan a picos del 100%.

Pruebas de estrés

Se trata de la ejecución de casos de uso conjuntos y con concurrencia de hilos determinada por el PROVEEDOR y con máximos y mínimos locales determinados por el PROVEEDOR de forma iterativa en un escenario de duración determinada por el PROVEEDOR. La finalidad de estas pruebas es **determinar la capacidad de recuperación del entorno** frente a un período de carga superior a la máxima admitida.



draw.io

Diagram attachment access error: cannot display diagram

La iteración del proceso dependerá de la autorización de la STIC y los problemas encontrados que necesiten algún ajuste y/u optimización del aplicativo o aplicativos.

Las pruebas deben ejecutarse de forma incremental, de forma que cada una de las pruebas que se muestran a continuación se deben lanzar sobre la anterior (Por ejemplo, se lanza el proceso inicial y se inicia la monitorización. A los 10 minutos se incrementa la carga, a los 20 se vuelve a incrementar, etc.):

PRUEBA	CONCURRENCIA	DURACIÓN (HORAS)
Concurrencia inicial	Concurrencia aproximada al 80% del límite de la máquina (Determinado por las pruebas de capacidad)	Determinado por el PROVEEDOR
Incremento 1: Situar la concurrencia al límite	Aumentar la concurrencia en un 20%	Determinado por el PROVEEDOR
Incremento 2: Situar la concurrencia por encima del límite	Aumentar la concurrencia entre un 20% y un 40%	Determinado por el PROVEEDOR
Decremento 1: Situar la concurrencia al límite	Reducir la concurrencia en un 20%	Determinado por el PROVEEDOR
Decremento a nivel inicial	Dejar en nivel de concurrencia inicial	Determinado por el PROVEEDOR

Desde un punto de vista teórico, al sobrepasar el límite admitido por el servidor, algunas peticiones deberían comenzar a fallar, obteniendo respuestas erróneas durante el periodo de carga superior al máximo obtenido. El objetivo de estas pruebas es observar cómo, después de haber estado durante un periodo de tiempo por encima del límite, la aplicación vuelve a su estado normal y las nuevas peticiones enviadas siguen funcionando correctamente.

Por ello, como resultado de este proceso se debe observar los resultados obtenidos **ANTES y DESPUÉS** de la ejecución de estas pruebas, para lo cual existen varias alternativas:

- **Incremento paulatino**: Programar en la aplicación lanzadora un número elevado de hilos a lanzar y hacer que el incremento se realice durante un periodo de subida muy largo
    - Por ejemplo, si el límite de nuestra máquina está en 200 hilos, se puede programar una prueba que lance 300 hilos y que tenga un periodo de subida de 30 minutos.
- **Programar diferentes lanzaderas**: Programar una lanzadera inicial que ejecute el 80% de la capacidad máxima y posteriormente se van lanzando otras lanzaderas en las que cada una de ellas van aumentando la carga en un 20%. De esta forma se pueden ir parando manualmente cuando se vea necesario para volver a la carga inicial.

Resultados obtenidos

*[El presente apartado se muestra únicamente a **modo de ejemplo**, debiendo ser cumplimentado al COMPLETO por el PROVEEDOR con los resultados, tanto de métricas básicas como de métricas detalladas, obtenidos en sus pruebas de carga.]*

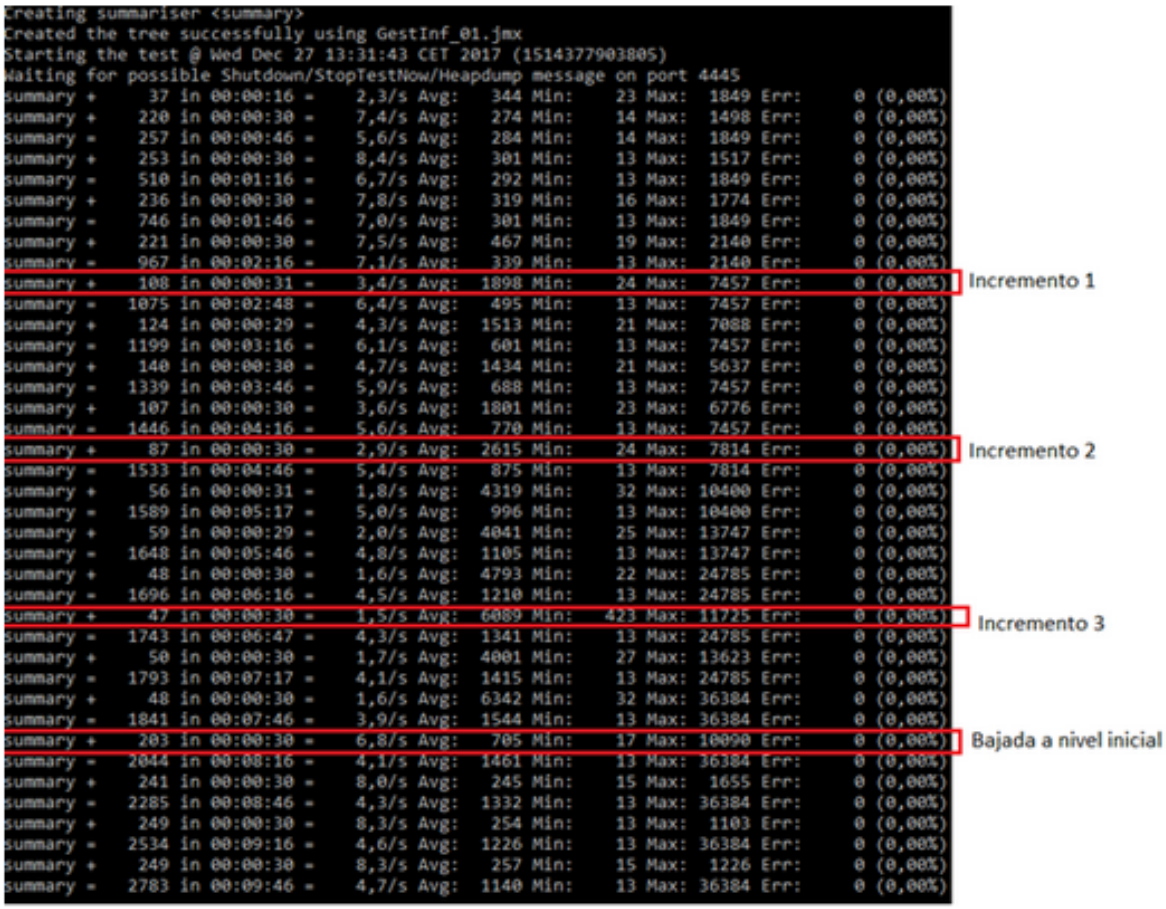
Es necesaria una explicación clara de lo que indica cada tabla / gráfica, tanto por parte del proveedor, como de CTI, por lo que cada una de las tablas / gráficas, deberá venir customizada incluyendo una explicación de los datos que se están arrojando.

Para la realización de estas pruebas se ha partido de la concurrencia definida en las pruebas de capacidad, realizándose las siguientes pruebas:

PRUEBA	***CONCURRENCIA***	***DURACIÓN (HORAS)***
Concurrencia inicial	*160 hilos*	*10 minutos**
Incremento 1	*Aumentar la concurrencia en 200 hilos*	*2 minutos**
Incremento 2	*Aumentar la concurrencia en 240 hilos*	*2 minutos**
Incremento 3	*Aumentar la concurrencia en 260 hilos*	*2 minutos**
Decremento 1	*Aumentar la concurrencia en 200 hilos*	*2 minutos**
Decremento a nivel inicial	*160 hilos*	*Hasta final de grabación*

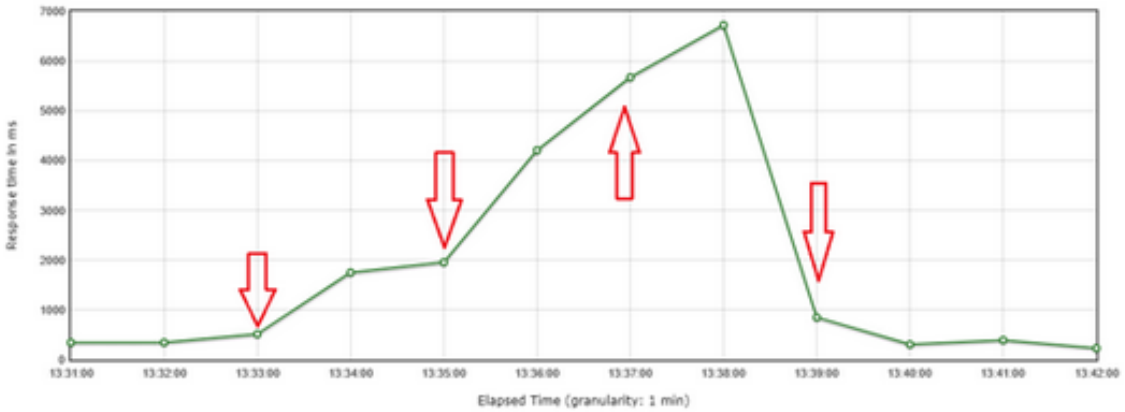
[(*): A modo de ejemplo se ha lanzado el test durante solo 10 minutos. Se recomienda que en un caso real se ejecute al menos 1 hora y se vayan realizando incrementos de concurrencia cada 10 o 20 minutos]

Ejecutadas las pruebas, se observa que a partir del primer incremento los tiempos de respuesta ya se elevan a valores no admisibles.

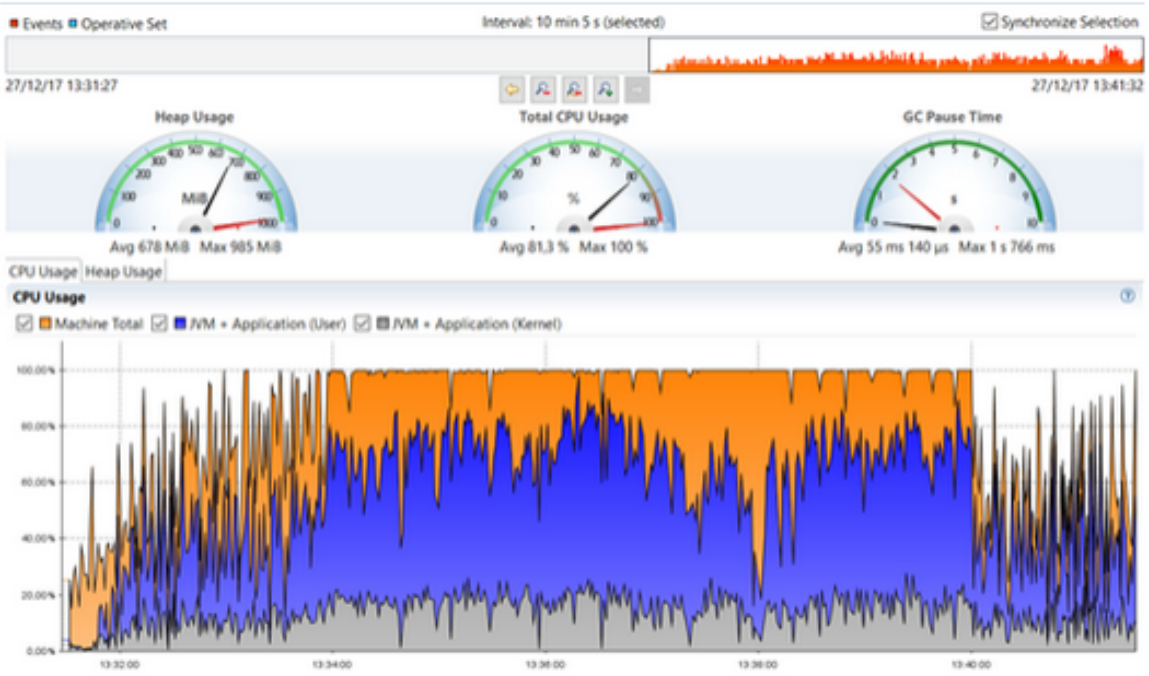


Esto nos lleva a pensar que, como se había mencionado anteriormente, la máquina ya estaba cerca de su límite máximo.

En la siguiente gráfica se muestra como se deterioran las peticiones con cada incremento de concurrencia:

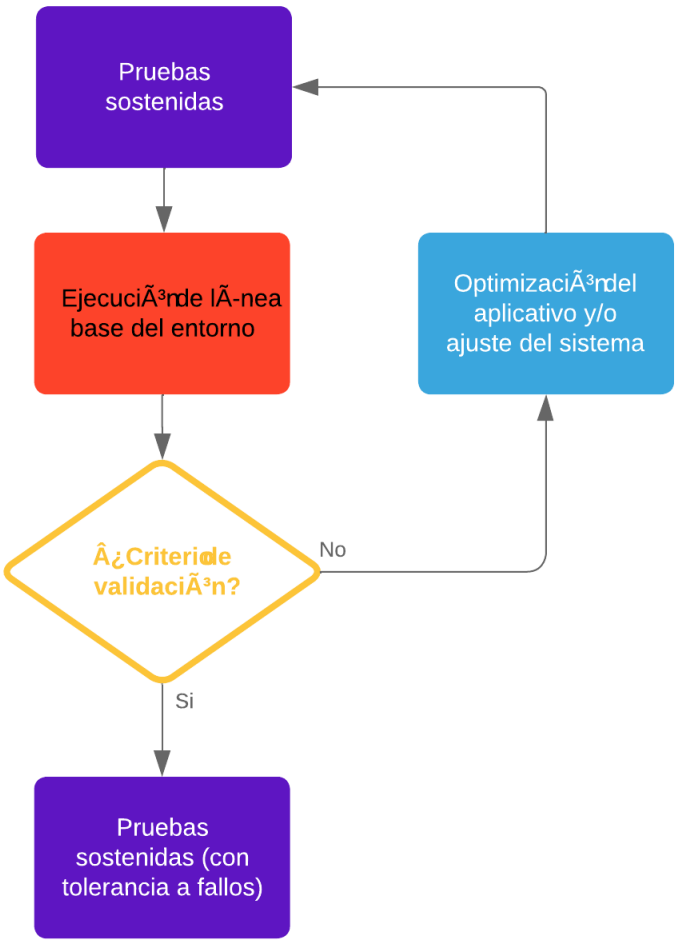


En este caso concreto, el problema se debe a una limitación de la máquina en la que está desplegada Weblogic, ya que se está quedando sin memoria con demasiada facilidad:


Events
Operative Set
Interval: 10 min 5 s (selected)
Synchronize Selection
27/12/17 13:31:27
27/12/17 13:41:32
Heap Usage
Total CPU Usage
GC Pause Time
MiB
%
s
Avg 678 MiB Max 985 MiB
Avg 81,3 % Max 100 %
Avg 55 ms 140 µs Max 1 s 766 ms
CPU Usage
Heap Usage
CPU Usage
Machine Total
JVM + Application (User)
JVM + Application (Kernel)
100,00 %
80,00 %
60,00 %
40,00 %
20,00 %
0,00 %
13:32:00
13:34:00
13:36:00
13:38:00
13:40:00

Pruebas sostenidas

Se trata de la ejecución de casos de uso conjuntos y con concurrencia de usuarios determinada por el PROVEEDOR, **no rmalmente medio/bajo**, de forma iterativa en **un escenario de larga duración** determinada por el PROVEEDOR. Estas pruebas **determinan la degradación del servicio** y el ciclo de vida de las instancias de WebLogic/réplicas del servicio y las instancias de bases de datos.



draw.io

Diagram attachment access error: cannot display diagram

La iteración del proceso dependerá de la autorización de la STIC y los problemas encontrados que necesiten algún ajuste y/u optimización del aplicativo o aplicativos.

A modo general, **se propone lanzar estas pruebas con un 40% del nivel de concurrencia máximo** determinado en las pruebas de capacidad **durante al menos 48 horas** (siendo el PROVEEDOR quien determine el periodo exacto que ve más apropiado).

Resultados obtenidos

*[El presente apartado se muestra únicamente a **modo de ejemplo**, debiendo ser cumplimentado al COMPLETO por el PROVEEDOR con los resultados, tanto de métricas básicas como de métricas detalladas, obtenidos en sus pruebas de carga.]*

Es necesaria una explicación clara de lo que indica cada tabla / gráfica, tanto por parte del proveedor, como de CTI, por lo que cada una de las tablas / gráficas, deberá venir customizada incluyendo una explicación de los datos que se están arrojando.

Tras ejecutar todos los casos de pruebas durante 48 horas con un nivel de concurrencia definido durante las pruebas de capacidad (200 hilos) se obtienen los siguientes resultados:

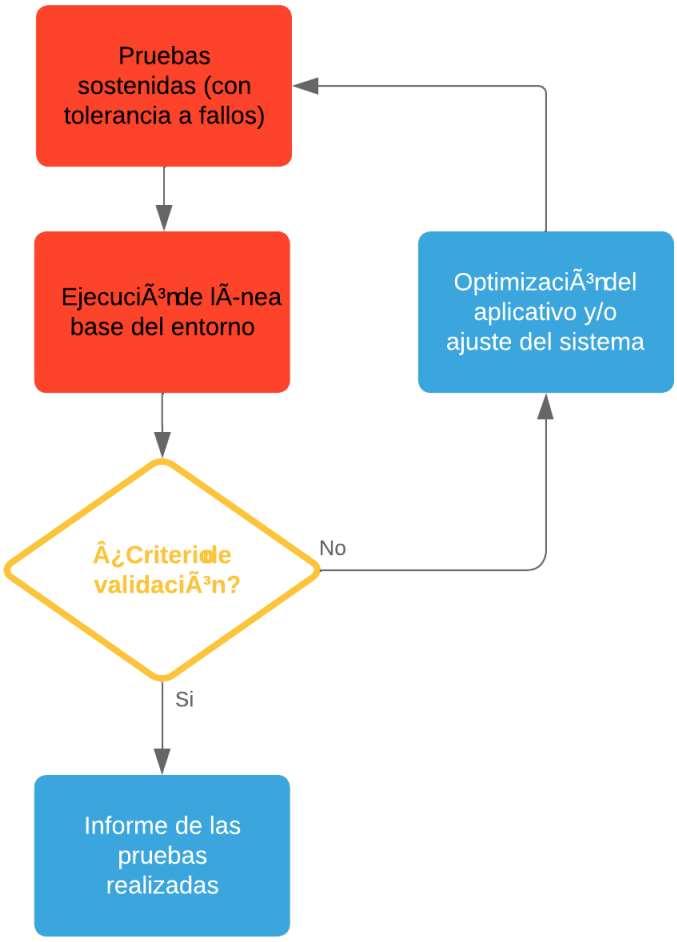
Prueba ejecutada entre el 26/12/2017 08:00 y el 28/12/2017 08:00

Requests	Executions			Response Times (ms)					
Label	#Samples	KO	Error %	Average	Min	Max	90th pct	95th pct	99th pct
Total	**5066**	**0**	**0.00%**	**767.57**	**16**	**7344**	**1743.60**	**2362.50**	**3957.31**
Login	400	0	0.00%	1714.20	625	6171	2836.00	3438.70	4936.69
Modificar concepto - Enviar datos	382	0	0.00%	1076.74	219	7344	2229.60	3420.45	6000.96
Cambio de perfil	1568	0	0.00%	900.57	127	5113	1752.00	2254.55	3717.70
Modificar concepto - Cerrar panel	381	0	0.00%	783.07	160	5438	1683.00	2303.80	4439.32
Modificar concepto - Carga panel	382	0	0.00%	758.11	148	6913	1700.70	2379.55	4559.42
Listado Formularios (Pantalla bienvenida)	398	0	0.00%	691.26	136	4769	1355.00	2008.50	3040.70
Listado Conceptos Fragmento	386	0	0.00%	653.17	117	6289	1398.40	2307.45	3973.86
Listado Conceptos Pagina	386	0	0.00%	451.40	88	3715	946.80	1283.55	2872.78
Logout	379	0	0.00%	155.04	23	2650	409.00	653.00	1315.00
Acceso	404	0	0.00%	77.27	16	1584	163.50	393.75	824.65

Pruebas sostenidas con tolerancia a fallos

Se trata de pruebas sostenidas pero **dirigidas por Sistemas** y con soporte del PROVEEDOR en el lanzamiento. **La idea subyacente es la de realizar operativas sobre las instancias de WebLogic/réplicas del servicio e instancias de bases de datos** correspondiente al entorno objetivo, ya sean administrativas o no, y estudiar el comportamiento de los casos de uso e indicadores de negocio.

La duración de la prueba estará acotada por lo acordado con la STIC y el tiempo disponible al tratarse de las pruebas realizadas en último lugar.



draw.io

Diagram attachment access error: cannot display diagram

La iteración del proceso dependerá de la autorización de la STIC y los problemas encontrados que necesiten algún ajuste y/u optimización del aplicativo o aplicativos.

El diseño de las pruebas se muestra en la siguiente tabla:

PRUEBA	CONCURRENCIA	DURACIÓN (HORAS)	FUNCIONAMIENTO ESPERADO
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia 1 de WebLogic o réplica 1 del servicio	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia 2 de WebLogic o réplica 2 del servicio	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
...	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia n de WebLogic o réplica n del servicio	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia 1 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia 2 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
¿	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia n de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia 1 de WebLogic o réplica 1 del servicio	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia 2 de WebLogic o réplica 2 del servicio	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
...	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia n de WebLogic o réplica n del servicio	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia 1 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia 2 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
¿	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia n de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR

Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia 1 de WebLogic o réplica 1 del servicio e instancia 1 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia 2 de WebLogic o réplica 2 del servicio e instancia 2 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
...	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada administrativa de la instancia n de WebLogic o réplica n del servicio e instancia n de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia 1 de WebLogic o réplica 1 del servicio e instancia 1 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia 2 de WebLogic o réplica 2 del servicio e instancia 2 de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
...	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con parada inesperada de la instancia n de WebLogic o réplica n del servicio e instancia n de base de datos	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con failover de base de datos si es Single Instance	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR
Ejecución de prueba sostenida sobre vip con switchover de base de datos si es Single Instance	Nivel de concurrencia óptima	Por determinar Sistemas	Por determinar por el PROVEEDOR

Conclusiones

En el presente apartado se hace un resumen de los resultados obtenidos, su interpretación y si los resultados obtenidos cubren las necesidades definidas en los requisitos. Los resultados **deben satisfacer** las necesidades definidas, por lo que la información que se refleje en las conclusiones **Capacidad máxima, Margen actual** y **Nodos recomendados** deb erá mostrar de forma clara que la aplicación, junto con los recursos hardware recomendados, será capaz de cubrir la carga definida en los requisitos a partir de las necesidades detectadas en Volumetría . Por ello, será necesario que los resultados obtenidos en los análisis y las necesidades definidas sean comparables. Idealmente, el caso de uso que se haya utilizado como fuente de información para la Volumetría debería formar parte de los Casos de Uso del plan de pruebas o que este contenga algún caso de uso lo más similar posible (Número de peticiones, peticiones de complejidad similar, similar proporción de peticiones a servicios externos, etc.)

Aunque la cumplimentación de este apartado es responsabilidad del PROVEEDOR, los datos aquí expuestos deben ser acordados con los diferentes grupos de la STIC dependiendo del caso (Sistemas, Arquitectura, Interoperabilidad, etc.).

En estas conclusiones se deben aportar al menos los siguientes datos, teniendo en cuenta que se refieren en cada momento a la infraestructura actual existente:

CONCLUSIÓN	DESCRIPCIÓN	TIPO
Capacidad máxima	Número de hilos por nodo que se estima como máximo antes de empezar a deteriorarse el servicio.	Obligatorio
Margen actual	Margen existente en la actualidad con la infraestructura existente (Comparando con la estimación de usuarios activos actualmente en producción. En el caso de que se trate de una nueva aplicación que aún no está en producción la comparación se realizará con la estimación realizada al comienzo del desarrollo de la pruebas).	Obligatorio
Nodos recomendados	Dada la capacidad por nodo existente, así como una previsión de usuarios del sistema. Qué estimación de nodos se aconseja ideal para ofrecer un buen servicio.	Obligatorio
Incidentes	Registro de incidentes y soluciones encontradas durante las pruebas en el caso de haberse producido.	Opcional
Mejoras detectadas	En caso de que proceda, se mostrarán también el conjunto de mejoras detectadas tanto en el software probado como en la infraestructura hardware y que se estima que mejorará el rendimiento del sistema.	Opcional

*[A partir de este punto la información mostrada es únicamente a **modo de ejemplo**, debiendo ser cumplimentado al completo por el PROVEEDOR con los resultados obtenidos durante sus pruebas]*

Dados los resultados de las pruebas realizadas, podemos aportar la siguiente información:

- ***Capacidad máxima**: Una vez analizadas las pruebas de estrés, a partir de 300 usuarios concurrentes por nodo los tiempos de respuestas son demasiado elevados (1,8 segundos). Por lo que se considera que 300 es el máximo de usuarios por nodos que soporta la aplicación.*
- ***Margen actual**: Dado que 200 usuarios (Pruebas de capacidad) ya supera aproximadamente un 25% el número de usuarios en picos de trabajo, aún hay un margen de 100 usuarios por nodo.*
- ***Nodos recomendados**: Dado que la previsión es contar con 500 usuarios concurrentes en momentos de máximo trabajo, creemos que la arquitectura actual (3 nodos) es suficiente.*

Adicionalmente, se desea destacar las siguientes conclusiones extraídas de los resultados obtenidos:

- ***Memoria RAM**: La memoria actualmente asignada a Weblogic (1GB) puede considerarse demasiado justa para el sistema. Se recomienda aumentar la memoria asignada a 1.5GB.*
- ***Lentitud de peticiones**: Se detecta un tiempo medio de respuesta de más de 2 segundos en la acción de logar. Se analizará para futuras versiones cómo optimizar este proceso para reducir el tiempo de respuesta.*

Finalmente, con la información recogida a lo largo del proceso de pruebas de carga del sistema, junto a la experiencia del proveedor como conocedor del producto a nivel técnico, el PROVEEDOR establece que las pruebas han sido:

CONCLUSIÓN	JUSTIFICACIÓN
¿ FAVORABLE	*[Razón por la que el **PROVEEDOR** en el momento de la entrega del informe, considera que la aplicación **SI** es apta para su puesta en producción]*
¿ NO FAVORABLE	*[Razón por la que el **PROVEEDOR** considera que la aplicación **NO** es apta para su puesta en producción]*

CONCLUSIÓN	JUSTIFICACIÓN
¿ FAVORABLE	*[Razón por la que el **CTI** en el momento de la entrega del informe, considera que la aplicación **SI** es apta para su puesta en producción]*
¿ NO FAVORABLE	*[Razón por la que el **CTI** considera que la aplicación **NO** es apta para su puesta en producción]*

CONCLUSIÓN	JUSTIFICACIÓN
¿ FAVORABLE	*[Razón por la que el **GOBERNANZA** en el momento de la entrega del informe, considera que la aplicación **SI** es apta para su puesta en producción]*
¿ NO FAVORABLE	*[Razón por la que el **GOBERNANZA** considera que la aplicación **NO** es apta para su puesta en producción]*

El documento debe finalizar con las conclusiones globales por parte del proveedor y CTI, indicando si del resultado de las pruebas realizadas se desprende que la aplicación es apta o no para su puesta en producción.

Posteriormente el Área de Gobernanza revisará toda la información entregada emitiendo también sus conlusiones.

Entregables tras las pruebas

Una vez finalizadas las pruebas se entregará tanto el plan de pruebas con los resultados obtenidos, como cualquier información, documentación y artefactos que se haya utilizado en la realización de las mismas. El Área de Gobernanza revisará y dará como FAVORABLE o planteará una serie de objeciones al plan de pruebas una vez se le informe de la disponibilidad del informe del plan de pruebas.

Documentación

Con respecto al plan de pruebas, se entregará en Confluence y deberá contener toda la información de las pruebas realizadas y resultados obtenidos siguiendo las indicaciones que aparecen en el *Anexo III: Plantilla PPS* de este documento.

En el apartado **Entregables** del PPS deberán adjuntarse la exportación de los informes HTML (https://jmeter.apache.org/usermanual/generating-dashboard.html) obtenidos en las pruebas:

1. Confirmación de las pruebas de capacidad
2. Pruebas de rendimiento en concurrencia alta
3. Pruebas de estrés
4. Pruebas sostenidas

Archivos

Los archivos relacionados con las pruebas de carga del sistema se entregarán a través de Git, siguiendo un versionado específico y organizanándose de acuerdo a los siguientes criterios:

- **HELM:** Los archivos se ubicarán dentro del proyecto HELM, en el directorio denominado *"pruebasdecarga".*
  - **Versionado:** Este directorio seguirá el versionado del proyecto HELM, existiendo un único directorio asociado a cada versión/etiqueta del repositorio helm.

- **Aplicación/componente (sin componente HELM):** Se creará un directorio al mismo nivel que los componentes existentes, con el nombre <CALM>-pruebasdecarga, donde CALM corresponde al código ALM de la aplicación.
  - **Versionado:** Se creará una rama a partir de develop con el contenido necesario para las pruebas de carga de la versión a lanzar. Posteriormente, se mergeará esta rama en develop y se versionará /etiquetará desde develop con la **misma versión que la aplicación padre**

- El resto, en el directorio *"pruebasdecarga".*
  - **Versionado:** Dentro del directorio de pruebas de carga se entregará una capeta por cada versión del producto sobre la que se realizan pruebas de carga siguiendo la nomenclatura de 3 dígitos o 4 cuando sea necesario por motivos de reentrega.

Estos directorios se creará en el repositorio del proyecto no siendo necesaria su entrega previa a la versión.

Contenido, se entregará:

- Archivos de JMeter utilizados para las pruebas (*archivo *.jmx*).
- Scripts para la preparación de datos siguiendo la normativa (si aplica).
- Archivo Readme.md que contendrá, entre otros apartados, un enlace al plan de pruebas de Confluence, así como las características técnicas que sean necesarias para la ejecución de las pruebas. La plantilla está disponible en Git: Plantilla del Readme.md para las pruebas de carga del sistema.

Anexos

Anexo I: Volumetría

El objetivo de las pruebas de carga del sistema es garantizar el comportamiento de la aplicación en su puesta en producción evitando caídas del sistema, lentitud, etc. Por ello, resulta necesario establecer la carga del sistema en producción y simular el comportamiento de la aplicación en dichas circunstancias. En sistemas que ya se encuentren en producción, preferentemente se utilizará el pico de carga del sistema del que se conserven registros. En caso de que el sistema no se encuentre en producción o no se conserven registros, se deberá establecer una estimación razonada de la carga que deberá soportar el sistema.

Como norma general, el caso de uso que origine el pico de carga identificado en Volumetría debe formar parte de los Casos de Uso del plan de pruebas. En su defecto, el plan de pruebas deberá contener algún caso de uso lo más similar posible (Número de peticiones, peticiones de complejidad similar, similar proporción de peticiones a servicios externos, etc.)

Preferentemente, las métricas a utilizar en Volumetría vendrán expresadas en términos que permitan una correspondencia con los resultados que proporciona la herramienta utilizada (JMeter) en sus informes. Con caracter general, se recomienda utilizar Productividad (Transacciones/s) como unidad de medida realizando la conversion de la información de carga máxima de la que se disponga a esta unidad de medida.

Anexo II: Casos de Uso

El conjunto de casos de usos debe ser acordado entre el proveedor y la STIC a través de los responsables funcionales de la aplicación.

*A continuación se muestra un **ejemplo** de casos de uso a probar:*

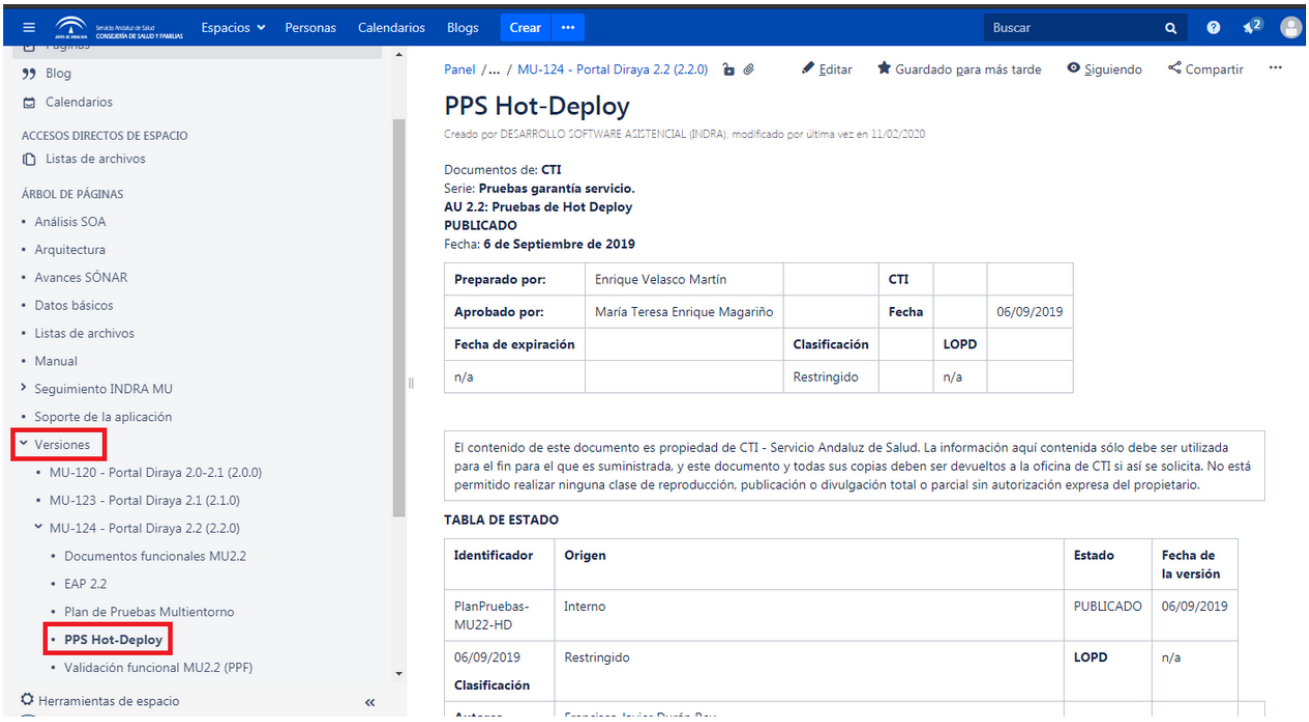
ID	CASO DE USO	DESCRIPCIÓN
CU-001	Modificar concepto	1. Acceso a la aplicación 2. Mostrar listado de conceptos existentes 3. Editar un concepto existente 4. Salir de la aplicación
CU-002	Listar formularios	1. Acceso a la aplicación 2. Mostrar listado de formularios existentes 3. Salir de la aplicación

En la descripción se incluirá la serie de actividades/pasos a realizar para llevar a cabo el proceso. Normalmente, cada actividad/paso coincidirá con alguna de las peticiones que proporciona la funcionalidad propia de la aplicación.

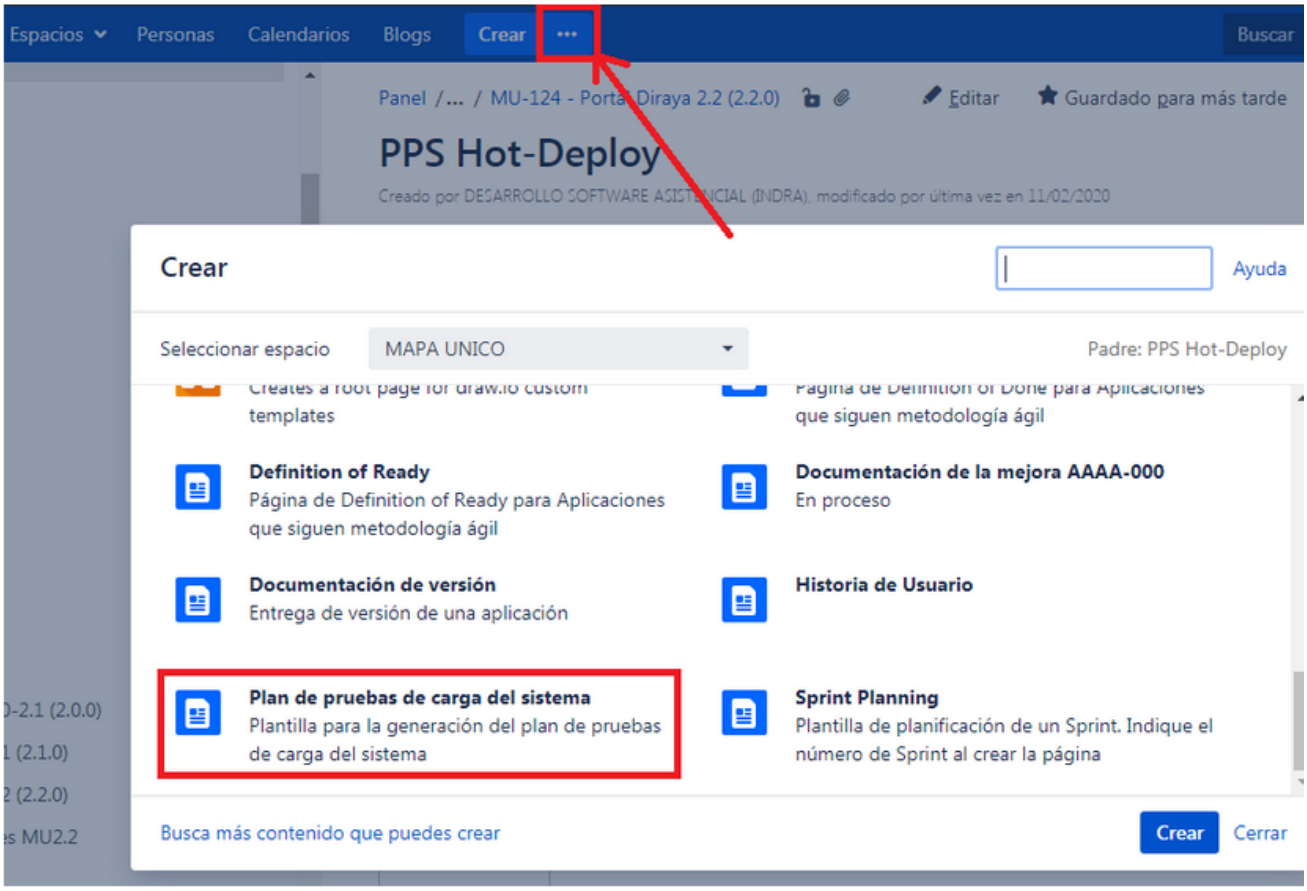
Anexo III: Plantilla PPS

Siguiendo la estrategia de migración a Confluence adoptada por la STIC, se ha creado una plantilla en dicha herramienta con el fin de abandonar el documento que se utilizaba hasta ahora.

El informe se ubicará en el espacio del proyecto correspondiente, en el apartado "versiones", dentro de la versión sobre la que se hayan ejecutado las pruebas.



Una vez situados en el lugar en el que queremos generar el informe, sólo tenemos que crearlo utilizando la plantilla llamada "Plan de pruebas de carga del sistema".



Si surgiese alguna cuestión al respecto, la Oficina de Calidad está siempre dispuesta a resolver cualquier duda que surja.

Anexo IV: Lista de revisión de documentación

A continuación se presentan los puntos de revisión que se aplicarán al documento PPS entregado:

Tipo	Código	PPS /Normativa	Descripción

Definición de pruebas	PPS-OBJ-01	Objeto	En el objeto del documento debe aparecer la aplicación y la versión sobre la que se van a realizar las pruebas
Definición de pruebas	PPS-ACT-01	Actores	Deben aparecer al menos los actores y roles indicados en normativa
Definición de pruebas	PPS-DIS-001	Diseño del plan de pruebas	Se deben identificar los casos de usos más representativos del sistema, aquellos usados con más frecuencia así como los casos de uso de nueva incorporación. Los casos de uso deben contener al menos la siguiente información: identificador, descripción y secuencia de pasos
Definición de pruebas	PPS-DIS-002	Diseño del plan de pruebas	Es imprescindible simular una volumetría semejante a la que se espera recibir en producción, ya que simular una carga muy diferente (ya sea por encima o por debajo) no será representativo del estrés real que sufrirá la aplicación, haciendo que las pruebas realizadas no sean representativas.

Definición de pruebas	PPS-DIS-003	Diseño del plan de pruebas	Se debe indicar la herramienta con la que se lanzarán las pruebas.
Definición de pruebas	PPS-DIS-004	Diseño del plan de pruebas	En la medida de lo posible, se debe añadir en las herramientas de pruebas aquellos controles necesarios para validar que el resultado obtenido en cada petición es el esperado (Por ejemplo, aserciones en jMeter). En el caso de no poder establecer los controles necesarios para validar el resultado obtenido se debe justificar el motivo.
Definición de pruebas	PPS-DIS-005	Diseño del plan de pruebas	Se debe indicar si las pruebas se realizan conectadas o desconectadas junto con su justificación.
Definición de pruebas	PPS-PLA-001	Planificación	En la planificación debe aparecer desglosado los distintos tipos de pruebas necesarios para llevar a cabo el ciclo de pruebas completo.

Definición de pruebas	PPS-PLA-002	Planificación	Aunque las realice CTI esta planificación debe añadirse a la estimación global. En el caso de que no sea necesario realizar este tipo de pruebas se debe incluir la justificación oportuna.
Definición de pruebas	PPS-PLA-003	Planificación	Se revisará que no haya solape entre las fechas y que los tiempos de realización sean acordes a la duración de cada prueba individual.
Previo ejecución	PPS-PRE-001	Prerrequisitos	Se revisará que la tabla la información de: check, resultado y descripción.
Previo ejecución	PPS-PRE-002	Prerrequisitos	En el caso de marcar "No Verificado" se deben indicar cuales son los requisitos que no se han podido verificar y el motivo. Esta opción se marcará cuando no sea posible realizar las pruebas debido a que no se cumpla alguno de los prerrequisitos
Previo ejecución	PPS-PRE-003	Prerrequisitos	En el caso de marcar "Verificado parcialmente" se deben indicar cuales son los requisitos que no se han podido verificar y el motivo. Esta opción se marcará cuando algunos prerrequisitos no han podido ser validados, pero se obtiene el visto bueno por parte del Jefe de Proyecto STIC para continuar con las pruebas de carga

Previo ejecución	PS-PRE-004	Prerrequisitos	Se deben incluir todos los prerrequisitos adicionales en su tabla correspondiente completando todos los campos
Previo ejecución	PS-COM-001	Pruebas de comunicación	Se revisará que la tabla contenga la información: sistema, componente y aceptación. Las pruebas de comunicación se refieren SOLO a las pruebas que el proveedor necesita realizar para poder acceder a la herramienta y poder así, hacer un seguimiento de las pruebas de carga realizadas. No se trata de pruebas que se realizan entre aplicativos.
Previo ejecución	PS-LBA-001	Línea base del entorno	En caso afirmativa se debe incluir una tabla con la siguiente información: tarea, descripción, responsable y aceptación. En caso negativo se debe justificar el motivo por el cual no es necesario restaurar el entorno.
Definición de pruebas	PS-MET-001	Métricas	Se deben incluir el conjunto de métricas básicas dirigidas a medir los tiempos de respuestas y buen funcionamiento del sistema. Las métricas NO deben definirse a posteriori, es decir, después de haber finalizado las pruebas sino antes de realizarlas.
Definición de pruebas	PS-MET-002	Métricas	Se deben incluir indicadores de negocio específicos de cada aplicación y que sean de utilidad para conocer el buen funcionamiento del sistema y la validez de las pruebas. Se refieren a los casos de uso más representativos de la aplicación y al número de usuarios que la aplicación debe soportar. Para el caso de aplicaciones que ya se encuentran en producción las métricas suelen ser fácilmente extraibles de este entorno. Para el caso de nuevas aplicaciones las métricas son más difíciles de extraer, pero el proveedor debe encontrar la manera de definir un mínimo antes del comienzo del desarrollo de las pruebas. Los indicadores de negocio NO deben definirse a posteriori, es decir, después de haber finalizado las pruebas sino antes de realizarlas.

Definición de pruebas	PPS-MET-003	Métricas	Se debe incluir la siguiente información: petición/funcionalidad, valor esperado y justificación. Esta información NO debe definirse a posteriori, es decir, después de haber finalizado las pruebas sino antes de realizarlas
Definición de pruebas	PPS-MET-004	Métricas	En caso negativo se debe indicar la justificación oportuna. En caso positivo hay que incluir la siguiente información: métrica, herramienta que proporciona el valor y valor esperado. Estas métricas NO debe definirse a posteriori, es decir, después de haber finalizado las pruebas sino antes de realizarlas
Informe resultados	PPS-CAP-001	Pruebas de capacidad	En base a la volumetría definida es conveniente mencionar en este apartado el requisito concreto que se persigue abordar con las pruebas de carga.
Informe resultados	PPS-CAP-002	Pruebas de capacidad	Se debe indicar tanto la duración de la prueba como la justificación de la duración establecida.
Informe resultados	PPS-CAP-003	Pruebas de capacidad	El objetivo de una prueba de capacidad es obtener punto de carga máximo sin degradar la aplicación.

Informe resultados	PPS-CAP-004	Pruebas de capacidad	El proveedor debe justificar su conclusión sobre el número máximo de hilos que soporta la aplicación sin degradarse, para ello se puede basar en los datos aportados por las herramientas lanzadoras de las pruebas.
Informe resultados	PPS-CAP-005	Pruebas de capacidad	Para productos que están en producción se establece este porcentaje (25%) para la realización de las pruebas de capacidad. Para los productos que no estén en producción también se establecerá el 25% superior a la estimación inicial realizada por el proveedor.
Informe resultados	PPS-CAP-006	Pruebas de capacidad	Al menos se debe mostrar la siguiente información: Tabla (peticiones, ejecuciones y tiempos de respuesta), consumo de CPU y memoria al empezar la prueba, durante y al finalizarla.
Informe resultados	PPS-CAP-007	Pruebas de capacidad	El proveedor debe indicar las comprobaciones que ha realizado tras la ejecución de las pruebas y el análisis de los resultados obtenidos.
Informe resultados	PPS-CAP-008	Pruebas de capacidad	El proveedor debe indicar si el resultado de la prueba ha sido satisfactoria o no. En ambos casos debe dar los argumentos en los que se basa su decisión.

Informe resultados	PS-REN-001	Pruebas de rendimiento	Los niveles de concurrencia deben ser los siguientes: Concurrencia alta: 80% de la concurrencia máxima determinada por las pruebas de capacidad Concurrencia media: 60% de la concurrencia máxima determinada por las pruebas de capacidad Concurrencia baja: 40% de la concurrencia máxima determinada por las pruebas de capacidad. Por cada nivel de concurrencia se debe indicar la siguiente información: prueba, concurrencia, duración, justificación de la duración
Informe resultados	PS-REN-002	Pruebas de rendimiento	Al menos se debe mostrar la siguiente información para cada uno de los niveles de concurrencia: Tabla (peticiones, ejecuciones y tiempos de respuesta), consumo de CPU y memoria al empezar la prueba, durante y al finalizarla.
Informe resultados	PS-REN-003	Pruebas de rendimiento	El proveedor debe indicar las comprobaciones que ha realizado tras la ejecución de las pruebas y el análisis de los resultados obtenidos.
Informe resultados	PS-REN-004	Pruebas de rendimiento	El proveedor debe indicar si el resultado de la prueba ha sido satisfactoria o no. En ambos casos debe dar los argumentos en los que se basa su decisión.
Informe resultados	PS-EST-001	Pruebas de estrés	Los niveles de concurrencia deben ser los siguientes: Concurrencia inicial: 80% de la concurrencia máxima determinada por las pruebas de capacidad Incremento 1: aumentar la concurrencia en un 20% Incremento 2: aumentar la concurrencia entre un 20% y un 40% Decremento 1: reducir la concurrencia en un 20% Decremento 2: reducir la concurrencia hasta el nivel inicial Por cada nivel de concurrencia se debe indicar la siguiente información: prueba, concurrencia, duración, justificación de la duración

Informe resultados	PPS-EST-002	Pruebas de estrés	Al menos se debe mostrar la siguiente información para cada uno de los niveles de concurrencia: Tabla (peticiones, ejecuciones y tiempos de respuesta), consumo de CPU y memoria al empezar la prueba, durante y al finalizarla.
Informe resultados	PPS-EST-003	Pruebas de estrés	El proveedor debe indicar las comprobaciones que ha realizado tras la ejecución de las pruebas y el análisis de los resultados obtenidos.
Informe resultados	PPS-EST-004	Pruebas de estrés	El proveedor debe indicar si el resultado de la prueba ha sido satisfactoria o no. En ambos casos debe dar los argumentos en los que se basa su decisión.
Informe resultados	PPS-SOS-001	Pruebas sostenidas	El nivel de concurrencia debe ser el 40% de la concurrencia máxima determinada por las pruebas de capacidad. Por cada nivel de concurrencia se debe indicar la siguiente información: prueba, concurrencia, duración, justificación de la duración
Informe resultados	PPS-SOS-002	Pruebas sostenidas	Se debe indicar tanto la duración de la prueba como la justificación de la duración establecida.

Informe resultados	PPS-SOS-003	Pruebas sostenidas	Al menos se debe mostrar la siguiente información: Tabla (peticiones, ejecuciones y tiempos de respuesta), consumo de CPU y memoria al empezar la prueba, durante y al finalizarla.
Informe resultados	PPS-SOS-004	Pruebas sostenidas	El proveedor debe indicar las comprobaciones que ha realizado tras la ejecución de las pruebas y el análisis de los resultados obtenidos.
Informe resultados	PPS-SOS-005	Pruebas sostenidas	El proveedor debe indicar si el resultado de la prueba ha sido satisfactoria o no. En ambos casos debe dar los argumentos en los que se basa su decisión.
Informe resultados	PPS-STF-001	Pruebas sostenidas con tolerancia a fallos	Por cada prueba se debe indicar la siguiente información: prueba, concurrencia, duración, funcionamiento esperado
Informe resultados	PPS-STF-002	Pruebas sostenidas con tolerancia a fallos	Al menos se debe mostrar la siguiente información: Tabla (peticiones, ejecuciones y tiempos de respuesta), consumo de CPU y memoria al empezar la prueba, durante y al finalizarla.

Informe resultados	PS-STF-003	Pruebas sostenidas con tolerancia a fallos	El proveedor debe indicar las comprobaciones que ha realizado tras la ejecución de las pruebas y el análisis de los resultados obtenidos.
Informe resultados	PS-STF-004	Pruebas sostenidas con tolerancia a fallos	El proveedor debe indicar si el resultado de la prueba ha sido satisfactoria o no. En ambos casos debe dar los argumentos en los que se basa su decisión.
Informe resultados	PS-CON-001	Conclusiones	Se debe indicar al menos la siguiente información: Capacidad máxima, Margen actual, Nodos recomendados.
Informe resultados	PS-CON-002	Conclusiones	El proveedor debe indicar si el resultado global de las pruebas han sido favorable o no. En ambos casos debe dar los argumentos en los que se basa su decisión.
Informe resultados	PS-CON-003	Conclusiones	El Área de Gobernanza, una vez revisado el informe, indicará si el plan de pruebas se considera favorable o no

Anexo V: Hot Deployment

Requisitos y recomendaciones para afrontar con garantías un hot deployment en Oracle Weblogic:

1. **Aplicación del workaround relativo a los conflictos de clases**. Nos hemos encontrado 3 tipos de conflictos en los entornos del SAS:

- Duplicidad en clases aportadas en el desplegable del proveedor y el software base de WebLogic. En el war/ear /jar tenemos las clases que tiene Weblogic en el software base existe con una versión distinta. En este caso podemos tratarlo con informes de conflicto de clases desde WebLogic (informes CAT).
- Duplicidad en clases aportadas en el desplegable. En el mismo war/ear/jar hay clases duplicadas con distintas versiones. El informe CAT no detecta estos casos, luego tendrá que realizarse mediante una herramienta de auditoría de código. En este caso hemos acordado que el aplicativo esté Mavenizado y compruebe que no haya "clases" en las librerías duplicadas (ojo, librerías con distinto nombre pueden tener clases repetidas con distintas versiones, luego no es sólo comprobar que haya librerías con distinto nombre).
- Duplicidad en clases aportadas por librerías copiadas manualmente en el rutas accesibles (CLASSPATH) como si fuera software base (por ejemplo $DOMAIN/lib). Este caso tampoco es comprobable con el informe CAT. Se ha acordado que el aplicativo candidato a despliegue caliente no puede tener más allá del software base de weblogic en el directorio de Middlware. Cualquier clase distinta al software base debe estar en el war/ear/jar e indicada la preferencia para weblogic en el descriptor de despliegue weblogic.xml.

2. **No se deben incluir a nivel de máquina virtual de java parametrización que pueda provocar un comportamiento irregular del servidor de WebLogic**. Por ejemplo, en el caso del "nodo mudo" se indica un parser a nivel de JVM distinto al utilizado por defecto en weblogic, cuando se debe utilizar como recursos dentro de los paquetes EAR. El ejemplo está documentado por Oracle en el siguiente enlace: https://docs.oracle.com/cd/E23943_01/web.1111/e13724/appscope.htm#XMLPG224 para WebLogic 10.3.6. En él se ejemplifica el cómo cambiar el Parser desde el EAR y no a nivel de JVM.

3. **El cambio a desplegar debe ser pilotable en BBDD y otros elementos compartidos como JMS, JTA, EJB, etc¿** Es decir, si el desplegable conlleva un cambio en base de datos y hay que parar para ese cambio, no se puede desplegar en caliente. Si el aplicativo necesita un reinicio un timeout excesivo en JTA, parar JMS, cambio en paramétricas en base de datos, etc¿ Puede ser contraproducente el despliegue en caliente y que dos versiones estén simultáneamente en ejecución.

4. **Redimensionamiento de los pooles de java relacionados con la carga de clases: PermGen en java 7 y MetaSpace en java 8.**
En Java 8 se suele indicar un máximo, aunque sea memoria nativa, suele estar restringido el máximo al igual que con Java 7 y el PermGen. Se debe, al menos, tener el doble para que no estamos ante excepciones de tipo OutOfMemory en los despliegues en caliente.

Todos estos requisitos son añadidos a las recomendaciones indicadas por Oracle en la documentación oficial:

https://docs.oracle.com/cd/E13222_01/wls/docs100/deployment/redeploy.html (apartado "Requirements and Restrictions for Production Redeployment")

Los requisitos aquí indicados son los mínimos, pero evidentemente hay que probarlo por cada plataforma y de forma exhaustiva ya que el modo de programación es ajeno a sistemas.

Para porder comprobar que se ha realizado correctamente el despliegue en caliente, proponemos realizar la una prueba de una versión (la ya desplegada) a otra (la que queremos desplegar en caliente) con una carga sostenida, lo que significa tener preparado dos versiones (versión A y B) con cambios en uno o varios casos de uso que se estén probando en dicha carga:

1. Despliegue versión A con versionado y reinicio/limpieza de entorno.
2. Lanzamiento de prueba sostenida.
3. Redespliegue online de versión B en la mitad de la prueba sostenida con un retire timeout de 5 minutos (el tiempo de expiración de sesión se consensuará entre las partes interesadas).
4. Espera a final de prueba sostenida de al menos 48 horas.