Historial de cambios del documento

Pruebas de rendimiento BCR-Salesforce

Enero

2017

|  |  |  |  |
| --- | --- | --- | --- |
| **Versión** | **Fecha** | **Autor** | **Descripción** |
| 1.0 | 2017-01-19 | Alejandro Quesada Vega | Todas las secciones |

Tabla de Contenidos

Introducción 1

Propósito 1

Audiencia 1

Alcance 1

Documentos relacionados 1

Terminología 2

Contexto General 3

Estado de la Aplicación 3

Enfoque de las Pruebas 4

Objetivo General 4

Objetivos Específicos 4

Ambiente de Pruebas 4

Infraestructura 4

Herramienta de Monitoreo de Páginas Web 5

Herramienta de Monitoreo de Servidores 6

Flujos de Trabajo 6

Variables Medidas 8

Otras Condiciones de la Prueba 8

Estrategia de Pruebas 9

Generación de carga 9

Monitoreo de Tiempos de respuesta Web 9

Monitoreo de Servidores 9

Resultados 11

Escenario I – línea base de generación de reportes 11

Escenario II – Línea base navegación web 13

Escenario III – Cola de generación de reportes y monitoreo de navegación web 18

Comportamiento de Servidores 24

Principales Hallazgos 30

Escenario I 30

Escenario II 30

Escenario III 30

Servidores Físicos 30

Conclusiones 32

# Introducción

## Propósito

El propósito de este documento es presentar la estrategia general de pruebas seguida para la ejecución del ciclo de pruebas de rendimiento al servicio web del BCR-Salesforce, así como los hallazgos y conclusiones más importantes a las cuales se llegó luego del análisis de resultados de las mismas.

La estrategia de pruebas descritas a continuación fue definida con base en la información obtenida en las últimas semanas, en las cuales personal de Avantica y del Banco de Costa Rica se han estado comunicando constantemente a través de reuniones, emails, etc.

## Audiencia

Personal de Avantica involucrado en la creación / ejecución de scripts, además del personal del Banco de Costa Rica colaborando con el soporte al equipo de Avantica.

## Alcance

Este documento presenta de forma general la estrategia y el enfoque seleccionado para la ejecución de las pruebas de rendimiento para el servicio web del BCR-Salesforce, así como los resultados obtenidos de la ejecución y las conclusiones más importantes luego del proceso de análisis.

## Documentos relacionados

|  |  |
| --- | --- |
| **Ref.** | **Documento** |
|  |  |
|  |  |

## 

## Terminología

|  |  |
| --- | --- |
| **Término** | **Definición** |
| **QA** | Quality Assurance, aseguramiento de la calidad. |
| **RAM** | Random Access Memory, memoria de acceso aleatorio. |
| **JMeter** | Herramienta para generar la carga a los servidores. |
| **Script** | Conjunto de pasos o instrucciones que se ejecutan en una máquina. Es un programa escrito en lenguaje de programación para probar una parte de la funcionalidad de un software o aplicación. |
| **Java** | Lenguaje de programación orientado a objetos. |

# Contexto General

## Estado de la Aplicación

El sistema del BCR-Salesforce se encuentra en producción y se busca comprobar la capacidad de consultas por segundo que puede procesar el servicio web, para mejorar el rendimiento de la aplicación.

La aplicación le permite al usuario, realizar consultas sobre los clientes del Banco de Costa Rica utilizando el número de cédula, si la identificación coincide con los usuarios que forman parte de los clientes del Banco entonces el servicio responde con la información del cliente, si no, solo responde que la persona buscada no se encuentra dentro de la base de datos de clientes del BCR.

Este documento plantea la ejecución de un ciclo de pruebas de rendimiento, contemplado por siete escenarios, para determinar la capacidad de consultas por segundo que permite procesar el servicio web.

# Enfoque de las Pruebas

Esta sección describe la estrategia de pruebas seguidas para medir el comportamiento del servicio web del BCR-Salesforce mientras es sometido a una capacidad determinada.

## Objetivo General

El objetivo general del proyecto es la ejecución de pruebas de rendimiento al servicio web del BCR-Salesforce que permitan verificar que dicho servicio, así como la infraestructura tecnológica utilizada para su ejecución, tendrán el rendimiento esperado por el banco, cuando se incrementa gradualmente la cantidad de consultas que se realizan por segundo.

## Objetivos Específicos

* Identificar un conjunto de escenarios que permita simular un incremento gradual de consultas por segundo.
* Crear un script que permita la ejecución de los escenarios a probar, para observar y analizar el comportamiento del servicio web al manejar múltiples consultas por segundo.
* Crear documentos con los resultados de prueba y entregar los mismos al BCR.

## Ambiente de Pruebas\*

### Infraestructura\*

El servicio web del BCR-Salesforce está alojado en una infraestructura compuesta por múltiples servidores, entre los que podemos destacar:

* Servidor de Base de datos
* Servidores de aplicación
* Servidores de presentación
* Balanceador de carga

### Herramienta de Monitoreo de Servidores

El monitoreo de los recursos en los servidores de aplicación durante el periodo de carga es clave para determinar patrones de comportamiento y posibles problemas de rendimiento.

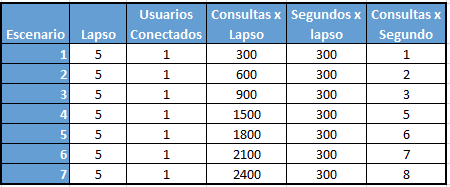
Debido a la complejidad de la infraestructura y las limitantes de acceso por temas de seguridad, para monitorear los servidores que alojan al servicio web de Salesforce y a sus bases de datos se utiliza el sistema Foglight, un sistema controlado por el Banco de Costa Rica para el monitoreo de su infraestructura.

Al finalizar el ciclo de ejecución, el Banco de Costa Rica le hace llegar a Avantica los datos recolectados luego del monitoreo realizado a los servidores durante el tiempo de la prueba.

## Flujos de Trabajo

Para el detalle de los flujos de trabajo utilizados en la prueba, por favor referirse al correo electrónico “Solicitud #021 - Salesforce Pruebas de Capacidad”.

Para las pruebas de rendimiento, se plantearon siete escenarios diferentes.



## Variables Medidas

Para efectos de recabar información que permita el análisis posterior, las siguientes variables han sido medidas en los servidores físicos y las páginas de los sitios web a prueba:

* Tiempo de respuesta del servicio web
* Niveles de uso de memoria RAM (Servidores)
* Niveles de uso de CPU (Servidores)

## Otras Condiciones de la Prueba

Para el análisis de la prueba de capacidad, es importante tomar en cuenta otros aspectos relacionados con la ejecución de la prueba:

* Las pruebas se realizaron en las instalaciones del Banco de Costa Rica, Aranjuez, desde la máquina de Carlos Cueva ya que era la única que contaba con permisos de acceso al servicio web.
* Las pruebas se realizaron contra el ambiente de producción, mediante el acceso a un URL definido por el banco.
* Las pruebas de capacidad se realizaron el día miércoles 18 de enero 2017, la carga de trabajo en los servidores de producción, adicional a la generada por la prueba, fue mínima.
* Se contó con el apoyo de personal del BCR para realizar labores adicionales de monitoreo y control que están fuera del alcance del personal de Avantica.

## Estrategia de Pruebas

Se realizaron las pruebas de rendimiento de acuerdo a los escenarios descritos en el correo electrónico de “Solicitud #021 - Salesforce Pruebas de Capacidad”.

Durante el período de carga, se ejecutaron pruebas de capacidad al servicio web del BCR-Salesforce, midiendo los tiempos de respuesta para cada uno de los flujos mencionados en dicho correo.

Finalmente se monitoreó el comportamiento de los servidores físicos relacionados con la aplicación a prueba, para medir el uso de recursos respectivo a lo largo de la ejecución.

## Generación de carga

A nivel general, la estrategia de carga para los flujos del BCR-Salesforce, es la siguiente:

* Se incrementó la carga de forma constante, iniciando con trecientas consultas durante cinco minutos para el primer escenario y aumentando en trecientas para el siguiente manteniendo los cinco minutos de duración de las pruebas

## Monitoreo de Servidores

A nivel de servidores, la estrategia a seguir se basa en lo siguiente:

* Ejecución del monitoreo mediante la herramienta Foglight, para realizar las respectivas mediciones de las variables a tomar en cuenta: Memoria RAM usada, así como el porcentaje de uso del CPU de cada uno de los servidores.
* El equipo de soporte del Banco de Costa Rica, recolecta la información necesaria para cada servidor y la envía a Avantica para el análisis respectivo.

# Resultados

Esta sección presenta un detalle de los resultados más relevantes observados durante la ejecución de las pruebas, los cuales serán utilizados como base para la definición de los principales hallazgos.

El proceso de generación de la carga siguió el siguiente patrón de comportamiento:

* Se dio inicio a las pruebas ejecutando el primer escenario sin problemas.
* Para el segundo y tercer escenario se encontraron errores de timeout para el bus de datos, se ejecutaron.
* Luego se ejecutó el cuarto escenario el cual presentó problemas tanto en producción como en clientes

Inicialmente se planteó el objetivo de realizar un proceso en donde el paso de un escenario al siguiente se diera de manera ininterrumpida. Sin embargo, eventos puntuales que se dieron durante la ejecución de la prueba provocaron el tener que realizar interrupciones forzadas:

* En el escenario II y III, durante la prueba se detectaron problemas de timeout los cuales fueron investigados y solucionados por parte del personal del Banco.
* En el escenario IV, se presentaron problemas de timeout, se procedió a ejecutar el escenario en BCR-Clientes donde también se detectaron timeouts.

## Escenario I – Una Consulta por segundo

Durante el escenario I se busca establecer una línea base de tiempos de respuesta para que cada uno de los escenarios siguientes.

Para este escenario se configura trecientas consultas durante cinco minutos, esto quiere decir que el servicio recibió una consulta cada segundo durante cinco minutos.

Durante la ejecución de este escenario no se detectaron problemas.

## Escenario II – Dos Consultas por segundo

Durante la ejecución del escenario II se configura seiscientas consultas durante cinco minutos, generando dos consultas por segundo.

En este escenario se descubrió que el servicio estaba respondiendo algunas peticiones con timeouts del bus de datos.

Gráfico 1- Tiempos de respuesta total ejecuciones del escenario II

## Escenario III – Tres Consultas por segundo

En el Escenario III se buscaba cargar con tres consultas por segundo al servicio web ejecutando novecientas consultas durante cinco minutos.

Este escenario se ejecutó sin problemas con una media de tiempos de respuesta de cuatrocientos once milisegundos.

## Escenario IV – Cinco Consultas por segundo

Para este cuarto escenario se ejecutó mil quinientas consultas en un período de cinco minutos, para un total de cinco consultas realizadas por segundo.

Durante la ejecución de este escenario se identificó problemas de bus nuevamente, el personal del Banco confirmó que se trataba de timeouts. Como medida final, se tomó la decisión de ejecutar el escenario III y IV en el servicio de BCR-Clientes.

Gráfico 2 - Clientes vs Producción

## Comportamiento de Servidores

Durante la ejecución del ciclo de pruebas de rendimiento, se monitoreó el comportamiento del uso de recursos en la infraestructura del sitio del BCR-Salesforce, particularmente los servidores de Producción y BCR-Clientes.

Esta sección presenta una serie de gráficos que muestran las tendencias de comportamiento de cada uno de los servidores para las variables monitoreadas durante el tiempo de ejecución de los diferentes escenarios del ciclo de pruebas.

Cabe resaltar que las pruebas se realizaron en ambiente de producción por lo que algunos picos de consumo pueden ser provocados por factores externos a las pruebas. En este documento únicamente vamos a hablar de los relacionados con las pruebas realizadas.

**Uso de Memoria RAM**

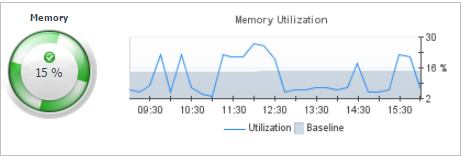
En el caso de los servidores de **producción**, el uso de memoria se vio impactado en pocas ocasiones a lo largo de la prueba. Se dieron variaciones en los valores observados, tal y como lo muestra el siguiente gráfico:

Gráfico 3 – Servidor Producción - BCR0683AWCI41 - Memoria

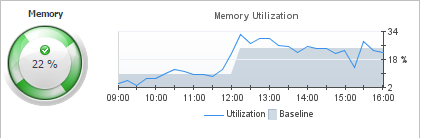
* Consumo promedio a lo largo de la prueba fue de un 15%
* El pico más alto de consumo lo presenta alrededor de las 12:00pm sobrepasando el 25% de uso, durante este momento se estaba ejecutando el cuarto escenario.

Gráfico 4 - Servidor Producción - BCR0683AWCI42 -Memoria

* Consumo promedio a lo largo de la prueba fue de un 22%
* Aproximadamente a las 12:00pm se presenta el pico más elevado de consumo alcanzando el 34% de uso.

En el caso de los servidores de **clientes** el uso de memoria se vio afectado mayoritariamente en horas de la tarde que se hicieron las pruebas sobre el mismo.

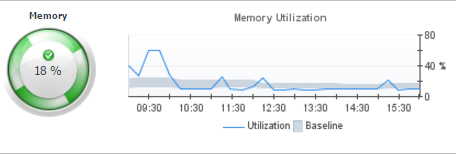
* El consumo promedio a lo largo de la prueba fue de un 18%

Gráfico 5 - Servidor Clientes - BCR0106WIT25 - Memoria

* El pico de consumo más alto es alrededor de las 9:55am, durante la ejecución del primer escenario.

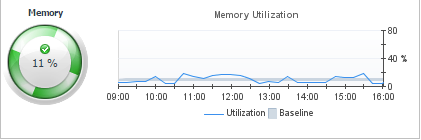


Gráfico 6 - Servidor Clientes - BCR0106WIT26 - Memoria

* El consumo promedio de memoria durante la prueba fue de un 11%
* El consumo general de memoria durante la prueba es muy estable sin altos picos de uso, se mantiene por debajo del 25%

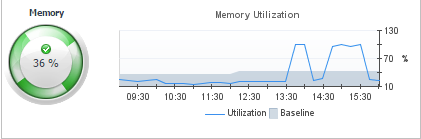


Gráfico 7 - Servidor Clientes - BCR0106ACL41 - Memoria

* Consumo promedio para la memoria de este servidor fue de un 36% de uso.
* Presenta picos de consumo de hasta 80% de uso pasadas las 2:30pm hasta las 3:30pm, en este periodo de tiempo se ejecutó el escenario tres y cuatro

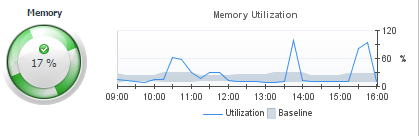


Gráfico 8 - Servidor Clientes - BCR0106ACL42 - Memoria.

* Promedia un 17% de uso de memoria con tres altas elevaciones de uso.
* A las 10:30am alcanza la marca del 60% de consumo de memoria
* Alrededor de las 3:30pm y 3:45pm supera el 60% alcanzando un máximo de un 90% de uso de memoria

**Uso de CPU**

Durante la ejecución del ciclo de pruebas también se monitoreó el nivel de uso de CPU en los distintos servidores de la infraestructura del servicio del BCR-Salesforce.

El uso de CPU en los servidores de **producción y clientes** presenta valores promedio elevados para unos servidores y más bajos en los otros.

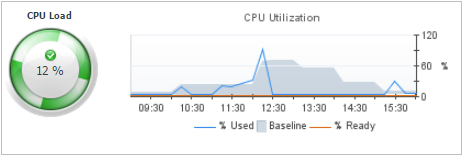
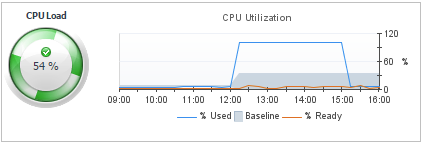
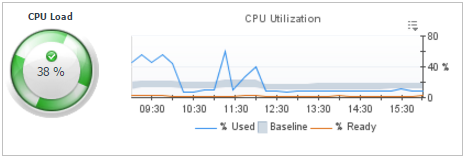
****Ambos comportamientos se muestran en los siguientes gráficos:

Gráfico 9 - Servidor Producción - BCR0683AWCI41 – CPU

* El promedio de consumo de procesador para este servidor fue de 12%
* El pico más alto lo presenta cercano a la marca del 90% aproximadamente a las 12:00pm
* 

Gráfico 10 - Servidor Producción - BCR0683AWCI42 – CPU

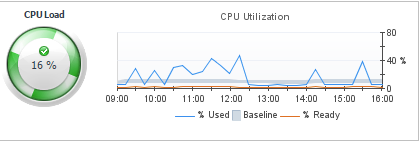
Gráfico 11 - Servidor Clientes - BCR0106WIT25 – CPU

Gráfico 12 - Servidor Clientes - BCR0106WIT26 – CPU

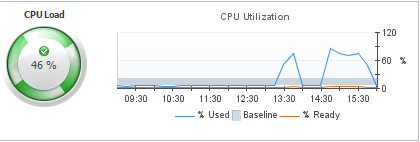


Gráfico 13 - Servidor Clientes - BCR0106ACL41 – CPU

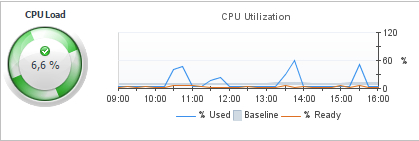
* 

Gráfico 14 - Servidor Clientes - BCR0106ACL42 - CPU

# Principales Hallazgos

Durante la ejecución de este ciclo de pruebas de rendimiento, se han dado los siguientes hallazgos:

## Escenario I

* Todos los reportes se ejecutan en menos de 28 minutos, lo cual se encuentra dentro del rango de tiempo esperado que toma un reporte en procesarse.
* Los reportes que mayor tiempo toman en procesamiento son “Activos y Pasivos” y “Pasivos”, que se también es parte de los supuestos proporcionados por el banco.
* Para los reportes de tipo “Backtesting” y “Efectivo” los tiempos de procesamiento son muy rápidos, tomando menos de 1 minuto para cada uno.

## Escenario II

* Se identificaron 3 acciones que, en comparación con el resto, toman tiempos de respuesta elevados, estas acciones son: Iniciar Sesión, Cerrar Sesión, Cargar Archivo, todas con valores por encima de los 10 segundos.
* En general, la mayoría de acciones se mantienen por debajo de 5 segundos.

## Escenario III

* En algunos casos como Iniciar y Cerrar sesión, al momento de inducir carga de trabajo en la cola que procesa los reportes, los tiempos de respuesta se duplicaron en comparación con los resultados de la línea base generada en el escenario II.
* Cargar archivo triplicó su valor máximo alcanzado en la línea base, pasando de 11.9 a 45 segundos con carga de trabajo.
* Acciones que se mostraban estables en la línea base, se vieron afectadas por la inducción de carga en la cola de procesamiento de reportes, Ir a página plantillas paso de un pico máximo de 2.5 segundos a 16 segundos.
* Para el resto de las acciones se presentan valores menores a 10 segundos

## Servidores Físicos

* El uso de CPU muestra tendencias de valores promedio bajos para todos los servidores.
  + Los Servidores de aplicación son los más afectados, con valores que alcanzan el 45% en el peor escenario.
  + Los servidores de presentación y base de datos presentan valores promedio menores a un 10%.
* En el caso del consumo de memoria RAM
  + Los valores mostrados a lo largo de la prueba no muestran alteraciones importantes en los Servidores de base de datos manteniendo un constante de 8.5GB de uso.
  + En los Servidores de aplicación y presentación, se mantienen niveles casi invariables de uso de 4.2GB para ambos servidores

# Conclusiones

Luego de la primera ejecución de pruebas de rendimiento y tomando como referencia los hallazgos antes mencionados, se puede concluir lo siguiente:

* Los tiempos de procesamiento de reportes no se ven afectados significativamente en el escenario de carga.
* Los flujos de “Activos y Pasivos” y “Pasivos” con niveles altos de confianza y horizonte son los que generan mayor carga de trabajo en el sistema, sin embargo aún en estos casos la carga de trabajo no llegan a niveles que comprometan los tiempos de procesamiento en el sistema.
* Sí se puede determinar un impacto en la experiencia de usuario, sobre todo en las acciones de cargar archivo, ir a página solicitudes, iniciar y cerrar sesión.

En general podemos concluir que el sistema responde bien ante la carga de trabajo proporcionada, sin embargo se recomienda una revisión de los procesos de iniciar y cerrar sesión así como del proceso de ir a página solicitudes que fueron los más afectados durante la prueba.