



Temática: Software Libre

Selección de los servidores web Apache 2 y Nginx durante la migración a código abierto

Selection of Apache 2 and Nginx web servers during the open source migration

Nurisel Palma Pérez ^{1*}

¹ Centro de Software Libre, Universidad de las Ciencias Informáticas, La Habana, Cuba.

* Autor para correspondencia: npalma@uci.cu

Resumen

La presente investigación se centró en el objetivo de desarrollar una herramienta informática para aumentar la eficiencia en la selección de los servidores web Apache 2 y Nginx durante el proceso de migración a código abierto. Se aplicó el método Analítico-Sintético para el estudio de Apache 2 y Nginx y la determinación de las particularidades de cada uno. La realización de un estudio de caso aplicando la Norma Cubana ISO/IEC 25023:2017 y el método estadístico, permitió el procesamiento de la información recopilada acerca del comportamiento de Apache 2 y Nginx, esto sirvió de base para formular la teoría relacionada con la eficiencia de ambos servidores. A partir del tipo de contenido y del total de peticiones concurrentes se definieron 60 escenarios para la selección del servidor más eficiente, teniendo en cuenta todos los indicadores o cada medida específica: Rendimiento, Utilización de recursos y Capacidad. Las tres variantes a seleccionar son: Apache 2, Nginx o Nginx como proxy inverso de Apache 2. El resultado obtenido se materializa en la Herramienta para la Migración y Administración de Servicios Telemáticos, la cual posee un componente Web que permite la selección del servidor web que más se ajusta a la institución y la activación del módulo correspondiente para su administración. La herramienta se validó a través de un estudio de caso, del criterio de expertos en su variante Delphi y la técnica de Iadov. Finalmente la triangulación metodológica permitió confirmar el resultado satisfactorio de todos los métodos aplicados y el cumplimiento del objetivo planteado.

Palabras clave: Apache 2, eficiencia, herramienta informática, Nginx.

Abstract

The present research focused on the objective of developing a computer tool to increase efficiency in the selection of Apache 2 and Nginx web servers during the open source migration process. The Analytical-Synthetic method was applied for the study of Apache 2 and Nginx and the determination of the particularities of each one. The realization of a case study applying the Cuban Standard ISO/IEC 25023:2017 and the statistical method, allowed the processing

of the information collected about the behavior of Apache 2 and Nginx, this served as the basis to formulate the theory related to the efficiency of both servers. From the content type and the total of concurrent request, 60 scenarios were defined for the most efficient server selection, taking into account all the indicators or each specific measure: Performance, Resource utilization and Capacity. The three variants to select are: Apache 2, Nginx or Nginx as a reverse proxy of Apache 2. The result obtained is materialized in the Tool for Migration and Administration of Telematic Services, which has a web component that allows the selection of the web server that more it adjusts to the institution and the activation of the corresponding module for its administration. The tool was validated through a case study, expert criteria in its Delphi variant and the iadov technique. Finally, the methodological triangulation allowed to confirm the satisfactory result of all the applied methods and the fulfillment of the proposed objective.

Keywords: Apache 2, efficiency, computer tool, Nginx.

Introducción

Actualmente son numerosos los servidores web, dentro de los que se encuentran Microsoft-IIS, Apache 2, Nginx, LiteSpeed, Google, Sun Java System, Lighttpd, IBM Servers y Yahoo Traffic Server. A partir de las estadísticas reportadas por (NETCRAFT, 2018) y (W3TECHS, 2018), se calculó la media del porcentaje de uso de cada uno de los cuatro servidores web que coinciden, determinándose Apache con 33,89%, Nginx con 29,61%, Microsoft-IIS con 24,79% y Google Servers con 1,23%, donde Apache y Nginx son los dos servidores web de código abierto más usados.

En Cuba también se evidencia el gran uso de estos dos servidores web de código abierto. Se analizaron los 40 principales sitios en Cuba de interés cultural y de entretenimiento, informativo, educativo e investigativo publicados por ETECSA (2018) y las tecnologías que estos emplean. Para analizar los sitios web se utilizó la herramienta multiplataforma Wappalizer, que permite identificar la tecnología utilizada en sitios web (Rakhmawati, et al., 2018). Del análisis realizado se determinó que 33 sitios mostraron el servidor web, donde las estadísticas son Apache 2 (66,6%), Nginx (27,3%) y Microsoft-IIS (6,1%), evidenciándose que Apache 2 es notablemente el más usado.

Se puede apreciar que se hace extensivo el uso de servidores de código abierto y precisamente en Cuba el proceso de informatización tiene como pilar el empleo de software libre. Para lograr la soberanía tecnológica y desempeñar el proceso de migración, la Universidad de las Ciencias Informáticas (UCI) se destaca por desarrollar la distribución cubana de GNU/Linux Nova y realizar el proceso de migración a código abierto en diferentes instituciones cubanas (Pérez, 2015). La ejecución de este proceso se basa en la “Estrategia para la migración a aplicaciones de código abierto”, la cual plantea que la migración de servicios telemáticos es el primer paso en el proceso de ejecución de la



migración (Pérez, 2015). Dentro de estos servicios se encuentran los ofrecidos por los servidores web. Como apoyo al proceso de migración se encuentra el libro “Buenas Prácticas para la Migración a Código Abierto”, el cual establece que las alternativas libres a servidores web privativos para pequeñas y medianas empresas y grandes empresas son Apache 2 y Nginx (Pérez, García y Goñi, 2015).

Se definen las dos alternativas pero no cómo se debe realizar el proceso de selección de estos. A partir de una entrevista realizada a los especialistas en servicios telemáticos pertenecientes a la UCI, se identificó que para la selección no existe una guía que oriente qué servidor se ajusta más a la institución. No se tiene un registro del estudio de servidores web migrados en procesos anteriores. La situación existente ocasiona que: se evidencia la pérdida de tiempo y se duplican esfuerzos estudiando la alternativa a seleccionar; no se reutiliza el conocimiento para la selección del servidor web; muchas veces se selecciona el servidor web más conocido aunque sea el menos eficiente para la institución. Con vista a la solución de la problemática planteada, el objetivo general está centrado en desarrollar una herramienta informática para aumentar la eficiencia en la selección de los servidores web Apache 2 y Nginx durante el proceso de migración a código abierto.

Materiales y métodos o Metodología computacional

Para el desarrollo de la investigación el objeto de estudio está orientado a los servidores web Apache 2 y Nginx, se consultaron fuentes como Nam (2017), Nedelcu (2015) y Soni (2016). Se utilizaron los métodos científicos y técnicas que a continuación se mencionan: Analítico-Sintético, Hipotético deductivo, Modelación, Método Estadístico, Entrevista y Encuesta. La eficiencia en la selección de los servidores web Apache 2 y Nginx implica el desempeño adecuado del servidor web que se seleccione para un entorno determinado. Por tanto, se hace necesario el estudio y análisis de la eficiencia de ambos servidores mediante un estudio de caso, los cuales son utilizados para desarrollar alguna teoría (Hernández, Fernández y Baptista, 2006).

Estudio de caso

El objetivo del estudio de caso es definir la diferencia entre los servidores web Apache 2 y Nginx en cuanto a la eficiencia de desempeño, teniendo en cuenta los indicadores de la Norma Cubana ISO/IEC 25023:2017 (Oficina Nacional de Normalización, 2017): tiempo medio de conclusión de un trabajo, adecuación del tiempo de conclusión de un trabajo, rendimiento medio, la media de utilización del procesador, la media de utilización de la memoria, la

media del uso de los dispositivos de entrada salida (E/S), utilización del ancho de banda, capacidad de procesamiento de transacciones y capacidad de acceso de usuario.

El estudio se basa en los elementos de los que depende el funcionamiento del servidor web, definidos en la sección anterior: arquitectura del servidor (se evidencia en el servidor web en cuestión), la cantidad de peticiones concurrentes y el tipo de contenido publicado. La PC servidora se estudiará en tres escenarios de prueba diferentes en cuanto al tipo de contenido que puede ser: estático, dinámico con PHP o dinámico con Python. A su vez para el tipo de contenido estático existen dos escenarios en cuanto al servidor web instalado que puede ser Apache 2 o Nginx, por otra parte para tipo de contenido dinámico con ambos lenguajes de programación se encuentran tres escenarios: Apache 2, Nginx o Nginx funcionando como proxy inverso de Apache 2 (la autora de la investigación lo denominó Proxy). De lo anteriormente explicado se concluye que para cada uno de los nueve indicadores de la variable eficiencia de desempeño, se estudiarán ocho escenarios.

El estudio consistió en realizar con la herramienta ab (Apache Benchmark) un número determinado de peticiones con cierta concurrencia, desde una PC cliente al servidor web instalado en la PC servidora. En cada uno de los ocho escenarios se realizaron 10 observaciones, para todas con un total de 50 000 peticiones y cada una con concurrencias de 10, 100, 250, 500, 750, 1 000, 5 000, 10 000, 15 000 y 20 000 peticiones respectivamente.

Además al mismo tiempo se realizó el monitoreo de los recursos de la PC servidora con la herramienta dstat. Para el cálculo de la eficiencia en cada indicador se definieron cinco escenarios en cuanto a la cantidad de peticiones concurrentes que pueden existir en las instituciones (denominada con la variable p), teniendo en cuenta las respuestas de los administradores de servidores web en una entrevista realizada. A continuación se describen los escenarios.

- $p \leq 500$: se refiere a una cantidad de peticiones concurrentes menor o igual que 500. Se incluyen cuatro observaciones con concurrencias 10, 100, 250 y 500 respectivamente.
- $500 < p \leq 1000$: se refiere a una cantidad de peticiones concurrentes mayor que 500 y menor o igual que 1 000. Se incluyen dos observaciones con concurrencias 750 y 1 000 respectivamente.
- $1000 < p \leq 10000$: se refiere a una cantidad de peticiones concurrentes mayor que 1 000 y menor o igual que 10 000. Se incluyen dos observaciones con concurrencias 5 000 y 10 000 respectivamente.
- $p > 10000$: se refiere a una cantidad de peticiones concurrentes mayor que 10 000. Se incluyen dos observaciones con concurrencias 15 000 y 20 000 respectivamente.
- Todas: se refiere a todas las concurrencias de peticiones incluyendo las 10 observaciones.

A partir del análisis de la eficiencia calculada en cada indicador, con el objetivo de medir todos los indicadores en una misma escala, se asignó en cada escenario en cuanto a la concurrencia de peticiones y atendiendo al tipo de contenido, los valores 3, 2 y 1 a los servidores web en correspondencia con el orden de su eficiencia, de mayor a menor. Para casos de igual eficiencia se asignó el mismo valor. Para cada uno de los cinco escenarios en cuanto a la concurrencia de peticiones, se calculó la puntuación final para cada servidor web teniendo en cuenta todos los indicadores y además para cada conjunto de indicadores correspondientes a las tres medidas de la eficiencia que son Rendimiento, Utilización de los recursos y Capacidad. Partiendo de la teoría definida, se desarrolló la solución como parte de la Herramienta para la Migración y Administración de Servicios Telemáticos.

Resultados y discusión

La Herramienta para la Migración y Administración de Servicios Telemáticos (HMAST) en su versión 2.0 contiene varios módulos, dos de ellos corresponden a los servidores web Apache 2 y Nginx respectivamente. La herramienta permite la administración remota a través del protocolo SSH (Secure SHell) de diferentes PC servidoras. La solución informática de la presente investigación se sustenta en añadir a HMAST un componente Web que permita seleccionar el servidor web que se ajuste a la institución, a partir del resultado obtenido con la aplicación del estudio de caso del epígrafe anterior. La herramienta presenta una arquitectura N-Capas orientada al dominio, distribuida en cinco componentes o paquetes: Presentación, Aplicación, Dominio, Persistencia e Infraestructura Transversal. En las capas de Presentación, Aplicación, Dominio y Persistencia se inserta un paquete llamado Web que a su vez contiene dos paquetes, Apache2 y Nginx, con todo lo referente a cada módulo.

Como solución de la presente investigación, en lugar de tener los módulos Apache 2 y Nginx aisladamente, se inserta en HMAST un componente Web que contiene dentro ambos módulos, que inicialmente están desactivados. Cuando se accede por primera vez al componente, se deben completar los campos en la opción Seleccionar servidor web: Cantidad total de usuarios que acceden al contenido (número entero positivo); Tipo de contenido (Estático/Dinámico); En caso de ser dinámico especificar lenguaje de programación (PHP/Python); Medida de la eficiencia que desea priorizar (Todos/Rendimiento/Utilización de los recursos/Capacidad).

Posteriormente se presiona el botón Mostrar resultado, la herramienta tomando como referencia los 60 escenarios obtenidos en la Tabla 1, visualiza el nombre del servidor web más indicado según los elementos especificados anteriormente. Se muestra la opción Instalar servidor web, que permite instalar el servidor seleccionado o ambos servidores en caso de ser Proxy la opción resultante. En este último caso, el sistema instala primero Apache 2 y se

habilita solo el puerto 8080 escuchando por 127.0.0.1 para posteriormente instalar Nginx y que no existan conflictos. El usuario puede desinstalar el servidor web en el momento deseado.

Tabla 1. Escenarios definidos para la selección del servidor web

(Fuente: elaboración propia).

Escenario	Indicadores	Tipo de contenido		
		Estático	Dinámico	
			PHP	Python
p≤500	Rendimiento	Nginx	Nginx	Nginx
	Recursos	Nginx	Nginx	Nginx
	Capacidad	Apache	Nginx	Proxy
	Todos	Nginx	Nginx	Nginx
500<p≤1000	Rendimiento	Nginx	Nginx	Proxy
	Recursos	Nginx	Nginx	Nginx
	Capacidad	Apache	Nginx	Nginx/Proxy
	Todos	Nginx	Nginx	Nginx
1000<p≤10000	Rendimiento	Apache	Proxy	Proxy
	Recursos	Nginx	Nginx	Nginx
	Capacidad	Nginx	Proxy	Nginx/Proxy
	Todos	Nginx	Proxy	Nginx
p>10000	Rendimiento	Nginx	Proxy	Proxy
	Recursos	Nginx	Nginx	Nginx
	Capacidad	Apache/Nginx	Nginx	Nginx/Proxy
	Todos	Nginx	Nginx	Nginx
Todas	Rendimiento	Nginx	Proxy	Proxy
	Recursos	Nginx	Nginx	Nginx
	Capacidad	Apache/Nginx	Nginx	Nginx/Proxy
	Todos	Nginx	Nginx	Nginx

Como se puede apreciar, teniendo en cuenta todos los indicadores, el servidor web seleccionado en todos los escenarios es Nginx excepto para contenido dinámico con PHP y concurrencia de peticiones entre 1 000 y 10 000, que el seleccionado es Proxy. En cuanto al consumo de recursos el más eficiente en todos los casos es Nginx. Teniendo en

cuenta el rendimiento y la capacidad, los tres servidores son elegidos respectivamente en algunos escenarios. En el módulo Apache 2 se pueden administrar las tres secciones de configuración: entorno global, servidor principal y hosts virtuales. El módulo Nginx cuenta con las funcionalidades necesarias para configurar las conexiones en el servidor, controlar el tráfico HTTP, gestionar los hosts virtuales, configurar los permisos de acceso y alias para cada host virtual.

Validación de la solución

La validación de la herramienta informática tiene la finalidad de verificar el cumplimiento del objetivo propuesto, que es aumentar la eficiencia en la selección de los servidores web Apache 2 y Nginx durante el proceso de migración a código abierto. Primeramente se evalúa la herramienta informática con la realización de un estudio de caso midiendo cada uno de los indicadores de la eficiencia de desempeño. Además se aplica la técnica de Iadov para determinar el nivel de satisfacción de los especialistas en servicios telemáticos y se evalúa la aplicabilidad de la solución con el criterio de expertos a través del método Delphi.

Estudio de caso en ECOAIND3

Se realizó un estudio de caso en la Empresa Constructora de Obras de Arquitectura e Industriales No. 3, que tuvo como objetivo comparar la eficiencia del servidor web seleccionado en dos proyectos de migración a código abierto: el proyecto de migración a aplicaciones de código abierto y el proyecto de actualización a la migración de los servidores. En el primero sin aplicar la solución se seleccionó Apache2 y en el segundo aplicándola, el servidor web seleccionado fue Nginx debido a que la empresa posee un total de 200 usuarios y el contenido publicado es estático (repositorio de Nova y la actualización del antivirus). Para la medición de la eficiencia se empleó la Norma Cubana ISO/IEC 25023:2017. Se realizaron cuatro observaciones, utilizando la herramienta ab mediante la cual se crearon un total de 200 peticiones con concurrencias 10, 100, 150 y 200 respectivamente, desde una PC cliente al servidor web instalado en la PC servidora. Al mismo tiempo se realizó el monitoreo de los recursos de la PC servidora con la herramienta dstat. Finalmente se realizó la comparación de la eficiencia de ambos servidores web, ver Tabla 2.

En la tabla se muestra en negrita el valor resultante del servidor más eficiente, donde en siete indicadores Nginx supera a Apache 2, se concluye que con la aplicación de la herramienta en ECOAIND3 se aumentó la eficiencia en la selección del servidor web.

Tabla 2. Cálculo de la eficiencia en ECOAIND3

(Fuente: elaboración propia).

Indicadores	Apache 2	Nginx
Tiempo medio de conclusión de un trabajo (ms)	562,250	190,250
Adecuación del tiempo de conclusión de un trabajo (ms)	2,811	0,951
Rendimiento medio	1,783	5,264
La media de utilización del procesador	0,012	0,005
La media de utilización de la memoria	0,275	0,126
La media del uso de los dispositivos de E/S	0,090	0,090
Utilización del ancho de banda	0,647	0,543
Capacidad de procesamiento de transacciones	2,864	2,921
Capacidad de acceso de usuario	200	200

Criterio de expertos

Para la selección de los expertos se aplicó una encuesta a 10 especialistas en servicios telemáticos, a partir de la información expresada en la misma se obtuvo el nivel de competencia de cada experto, donde el 90% posee un nivel de competencia Alto y un 10% un nivel Medio. Para la evaluación de la solución, se realizó el envío a los expertos de un documento que describe la solución desarrollada y una encuesta, donde se solicitó que evaluaran cinco aspectos relacionados con la misma en cuanto a cinco categorías. Posteriormente se emplea el Modelo de Torgerson, que, a partir de las valoraciones de los expertos se calculó la frecuencia absoluta por aspecto, la frecuencia acumulada por aspecto, la frecuencia acumulada relativa y los Puntos de Corte, que su ubicaron en una recta numérica, clasificando de esta forma cada aspecto sometido a evaluación. Se concluye que de los cinco aspectos, tres se evalúan de Muy Adecuado (60%) y dos de Bastante Adecuado (40%), sin mostrarse ninguno de los casos como Adecuado, Poco adecuado o Nada adecuado. Se demostró la existencia de un 100% de concordancia con la solución.

Técnica de Iadov

Se realizó una encuesta a 10 especialistas en servicios telemáticos pertenecientes al Centro de Software Libre de la UCI, para determinar su grado de satisfacción con la herramienta propuesta a partir de su experiencia en la migración y administración de servicios telemáticos. De los 10 especialistas encuestados, 7 respondieron clara satisfacción, 2

más satisfechos que insatisfechos y 1 indefinido, resultando el índice de satisfacción grupal (ISG) igual a 0,8, indica que los especialistas en servicios telemáticos presentan satisfacción con la herramienta.

Triangulación metodológica

La triangulación es un procedimiento de validez en el que los investigadores buscan la convergencia entre múltiples y diferentes fuentes de información para formar temas o categorías en un estudio. La triangulación metodológica se define como el uso de más de dos métodos para estudiar el mismo fenómeno bajo investigación. Se basa en el uso de métodos y análisis de recolección de datos tanto cualitativos como cuantitativos en el estudio del mismo fenómeno (HUSSEIN, 2009). Se concluye que los resultados obtenidos en los métodos cuantitativos y cualitativos coinciden, por lo que se puede afirmar que la herramienta informática aumenta la eficiencia en la selección de los servidores web Apache 2 y Nginx durante el proceso de migración a código abierto.

Conclusiones

Al finalizar la presente investigación se concluye lo siguiente:

- El análisis de las fuentes bibliográficas relacionadas con los servidores web Apache 2 y Nginx, permitió incluir como alternativa para las instituciones cubanas la configuración de Nginx como proxy inverso de Apache 2 y determinar que el correcto funcionamiento del servidor web depende de la arquitectura del servidor, la cantidad de peticiones concurrentes y el tipo de contenido publicado.
- El empleo de un estudio de caso aplicando la Norma Cubana ISO/IEC 25023:2017 y calculando la eficiencia de cada servidor web a partir del tipo de contenido y la cantidad de peticiones concurrentes, permitió el desarrollo de un componente Web en la Herramienta para la Migración y Administración de Servicios Telemáticos, que selecciona el servidor web más eficiente en cuanto a 60 escenarios.
- La validación de la herramienta informática mediante un estudio de caso en ECOAIND3 analizando dos proyectos de migración; el empleo del criterio de expertos a través del método Delphi y el modelo de Torgerson; la aplicación de la técnica de Iadov para medir el grado de satisfacción; y la triangulación metodológica; permitió evidenciar el cumplimiento del objetivo de la presente investigación.

Agradecimientos

Al Centro de Software Libre (CESOL) de la Universidad de las Ciencias Informáticas.



Referencias

- ETECSA. (2018). Sitios web y portales nacionales de interés informativo, cultural e investigativo. [En línea]. [Consultado el: 5 de diciembre de 2018]. Disponible en: [http://www.etecsa.cu/inicio/sitios_web_nacionales/].
- Hernández, R., Fernández, C., Baptista, P. (2006). *Metodología de la investigación. Cuarta Edición*, McGraw-Hill, México, 223-224.
- Hussein, A. (2009). The use of triangulation in social sciences research: Can qualitative and quantitative methods be combined? *Journal of comparative social work*, 4 (1), 1-8.
- Kabir, M. (2003). *La Biblia del Servidor Apache 2*, Anaya Multimedia, España, 44-45.
- Nam, V. (2017). Comparative Performance Evaluation of Web Servers. *VNU Journal of Science: Computer Science and Communication Engineering*, 31 (3), 28–34.
- Nedelcu, C. (2015). *Nginx HTTP Server. Third Edition*, Packt Publishing, Reino Unido, 276-310.
- NETCRAFT. (2018). Most Reliable Hosting Company Sites in November 2018. [En línea]. [Consultado el: 6 de diciembre de 2018]. Disponible en: [<https://news.netcraft.com/archives/2018/>].
- Oficina Nacional de Normalización. Norma Nubana NC ISO/IEC 25023:2017. (2017). *Ingeniería de Software y Sistemas – Requisitos de la Calidad y Evaluación de Software y Sistemas (SQUARE) – Medición de la Calidad del Producto de Software y Sistema*.
- Pérez, Y., García, A., Goñi, A. (2015). *Buenas Prácticas para la Migración a Código Abierto*, Ediciones Futuro, Cuba, 1-106.
- Pérez, Y. (2015). *Estrategia para la migración a aplicaciones de código abierto*. Tesis para optar por el título de Máster en Informática Aplicada, Universidad de las Ciencias Informáticas, La Habana.
- Rakhmawati, N. A., et al. (2018). A survey of web Technologies in Indonesia Local Governments. *Jurnal Sisfo*, 7 (3), 213-222.
- Soni, R. (2016). *Nginx: From Beginner to Pro*, Apress, Estados Unidos, 1-255.
- W3TECHS. (2018). Usage of web servers for websites. [En línea]. [Consultado el: 6 de diciembre de 2018]. Disponible en: [https://w3techs.com/technologies/overview/web_server/all].