

METODOLOGÍA DE IMPLEMENTACIÓN DE CENTRO DE DATOS BASADO EN SOLUCIONES OPEN-SOURCE²³

Andy Reyes²⁴, Alex Pacheco²⁵

Pares evaluadores: Red de Investigación en Educación, Empresa y Sociedad – REDIEES.²⁶

²³ Derivado del proyecto de investigación: Implementing model applied to a virtualized data center based on open source projects

²⁴ Open Nova IT Consulting, CIO, correo electrónico: andy.reyes@opennova.pe

²⁵ Universidad Nacional de Cañete, Docente investigador, correo electrónico: apacheco@undc.edu.pe

²⁶ Red de Investigación en Educación, Empresa y Sociedad – REDIEES. www.rediees.org

METODOLOGÍA DE IMPLEMENTACIÓN DE CENTRO DE DATOS BASADO EN SOLUCIONES OPEN- SOURCE²⁷

Andy Reyes²⁸, Alex Pacheco²⁹

RESUMEN

Esta obra tiene como objetivo demostrar que el modelo de implementación llamado OSPA permite el despliegue de un centro de datos basado en cuatro soluciones de código abierto para la gestión de identidades, plataforma de virtualización, sistemas de almacenamiento y gestión de red. La principal contribución de la metodología es simplificar el proceso de implementación de los cuatro proyectos que proporcionan documentación independiente el uno del otro. Los cuatro componentes cumplen con una lista de verificación de requisitos dados por la industria, como escalabilidad, alta disponibilidad, seguridad y capacidad de administración. Una arquitectura diseñada en capas donde cada capa complementa el modelo, cada capa implementa funcionalidades clave. La división en fases permite gestionar el proyecto de forma ordenada. El resultado es un centro de datos implementado con menos complejidad, que reduce el tiempo de implementación y el costo al no incurrir en costos de licencias. Esta implementación cumple con los requisitos de las organizaciones de la misma manera que las soluciones propietarias. La metodología permite reutilizar servidores y otros componentes disponibles en la organización para implementar una arquitectura abierta, además de proponer el uso de tecnologías basadas en software y no hardware, dando paso a la integración con otros componentes fuera de esta metodología ya que esta se basa en estándares abiertos donde el nivel de integración es muy alto.

²⁷ Derivado del proyecto de investigación: Implementing model applied to a virtualized data center based on open source projects

²⁸ Open Nova IT Consulting, CIO, correo electrónico: andy.reyes@opennova.pe

²⁹ Universidad Nacional de Cañete, Docente investigador, correo electrónico: apacheco@undc.edu.pe

ABSTRACT

This work aims to demonstrate that the implementation model called OSPA allows the deployment of a data center based on four open source projects for identity management, virtualization platform, storage systems and network management. The main contribution of the methodology is to simplify the implementation process of the four projects that provide documentation independent of each other. All four components meet an industry-given checklist of requirements such as scalability, high availability, security, and manageability. An architecture designed in layers where each layer complements the model, each layer implements key functionalities. The division into phases allows the project to be managed in an orderly manner. The result is a less complex deployed data center, which reduces deployment time and cost by not incurring licensing costs. This implementation meets the requirements of organizations in the same way as proprietary solutions. The methodology allows reusing servers and other components available in the organization to implement an open architecture, in addition to proposing the use of software-based and non-hardware-based technologies, giving way to integration with other components outside of this methodology since it is based on open standards where the level of integration is very high.

PALABRAS CLAVE: Linux, open-source, metodología de TI, infraestructura como servicio.

KEYWORDS: Linux, open-source, IT methodology, infrastructure as a service.

INTRODUCCIÓN

En 1991 Linus Torvalds se propuso crear un sistema operativo basado en UNIX (Aguilera, 2020) pero con la intención que este sistema se ejecute sobre cualquier plataforma, usted también sigue esa tendencia al interesarse en un centro de datos basado en estándares abiertos. La motivación principal por la cual el especialista de tecnología adopta el open-source es diversa, quizás por el ahorro de costos en licencias o suscripciones o quizás por las grandes posibilidades de integración al no ser exclusivo de un fabricante, independientemente de su motivación, el centro de datos basado en proyectos open-source es una alternativa tecnológica viable.

Un centro de datos es considerado la columna vertebral de muchos de negocios desde el punto de vista de las tecnologías de la información (Reyes, 2010), sobre él se alojan las aplicaciones que soportan al negocio, intentar diseñar, implementar, reestructurar, mantener un centro de datos es un tarea muy compleja, a pesar de que estas actividades de diseño e implementación están referenciadas en estándares internacionales y buenas practicas es correcto pensar que cada centro de datos es como una persona, no existen dos iguales.

Esta obra está dirigida a futuros administradores de sistemas que quieran innovar, no se está creando nada nuevo, se está aplicando una tecnología que existe desde el siglo pasado con un enfoque integral orientado a la infraestructura como servicio (Brandon, 2019), es decir, implementar dominios tecnológicos muy específicos como el computo, redes y almacenamiento.

Cuando implementamos un centro de datos existen requerimientos clave que la industria define como la alta disponibilidad, escalabilidad, seguridad y administrabilidad, todos los componentes del centro de datos deben alinearse a cumplir en alguna medida estos requerimientos según las necesidades de la organización buscando un equilibrio entre costo y beneficio. Esta obra nos permite cumplirlos, diseñarlos e implementarlos tomando como base la virtualización.

Aunque esta obra está dirigida a la implementación de la virtualización de componentes (cómputo, redes y almacenamiento), es la base tecnológica de otras infraestructuras como las aplicaciones, contenedores, bases de datos, etc.

La base tecnológica de este centro de datos serán proyectos open-source que proporcionan documentación extensa, sin embargo, estos proyectos requieren de un marco metodológico que guíe su implementación y sobre todo la integración entre ellos. Es así como nace esta iniciativa, si usted está interesado en la implementación se sugiere tenga una experiencia mínima en sistema operativo Linux, en el caso de la operación del día a día no es necesario ya que todas las herramientas son de administración gráfica, por lo tanto, requiere conocer solo los conceptos y operaciones.

Este centro de datos podrá implementarse sobre servidores de cualquier fabricante o arquitectura soportada sobre sistema operativo Linux incluso servidores en la nube siendo el soporte de virtualización el principal requerimiento, teniendo mayor flexibilidad y disponibilidad en el mercado con servidores X86_64.

Los procedimientos de implementación irán cambiando en el tiempo, por lo tanto, el autor de esta obra se compromete a realizar actualizaciones periódicas en los canales digitales públicos para documentar esas secciones del libro que requieran actualizaciones importantes, sin embargo, el diseño y planificación se mantendrán en el tiempo ya que se basa en la metodología propuesta la cual llamaremos OSPA – Open Source Projects Applied.

DESARROLLO DE LA METODOLOGÍA OSPA

Antes de iniciar la implementación de cualquier solución debemos diseñarla y planificarla, el principal motivo es definir alcances y conocer las limitaciones tecnológicas para estructurar un plan de implementación realista y sobre todo cubra las necesidades de la organización.

La metodología OSPA está enfocada a la implementación del centro de datos virtualizado solo en tres dominios tecnológicos dentro de la infraestructura como servicio:

- Computo
- Redes
- Almacenamiento

A fin que cumplan con cuatro requerimientos

- Alta disponibilidad

- Escalabilidad
- Seguridad
- Administrabilidad

Los proyectos open-source deben aportar al requerimiento, esto varia de empresa a empresa, por ejemplo, un proveedor de servicios de internet tiene requerimientos muy elevados en alta disponibilidad y seguridad dentro del dominio de redes, pero menores en almacenamiento, una empresa de sistemas de respaldo tiene requerimientos muy elevados en el dominio de almacenamiento y así se definen según la organización.

Sobre el sistema operativo CentOS se implementarán cuatro soluciones open-source:

- FreeIPA (FreeIPA, 2020) el cual será el sistema principal de gestión de usuarios, grupos y autorización a recursos
- Katello/Foreman (Foreman, 2020) para la gestión de operaciones y actualización de paquetería de software
- GlusterFS (Gluster, 2020) para el sistema de almacenamiento
- OVirt (oVirt, 2020) para el sistema de virtualización

Identificando los cuatro proyectos open-source cruzamos la información con los cuatro requerimientos para estimar el porcentaje que aporta cada proyecto, donde se busca alcanzar el 100% de las expectativas de la empresa como se muestra en la Tabla 1.

Tabla 1. Contribución de cada componente a la solución general

Proyecto	Alta Disponibilidad	Escalabilidad	Seguridad	Administrabilidad
FreeIPA	5	5	70	25
Katello/Foreman	5	5	20	45
GlusterFS	45	45	5	5
OVirt	45	45	5	25
			10	
Total	100	100	0	100

A fin de elaborar el plan de proyecto, se definen cuatro fases en la metodología OSPA la cual facilita la elaboración de la documentación y los responsables de elaborar, revisar y aplicar dicha documentación, las fases son las siguientes:

- Análisis
- Diseño
- Implementación
- Operación

En la fase de análisis, el especialista en infraestructura analiza los requerimientos de la organización y propone una arquitectura tentativa sin mayor detalle, siendo su enfoque identificar las limitaciones de las aplicaciones y determinando si están soportadas sobre la solución basada en open-source.

En la fase de diseño, se crean las arquitecturas de servidores de virtualización, almacenamiento, redes virtuales, así como los perfiles de afinamiento de estos tres dominios tecnológicos con funcionalidades claves como la calidad de servicio, cuotas de recursos, gestión de acceso a recursos por perfiles de grupos de usuarios, etc.

En la fase de implementación se despliegan los diseños de la fase anterior en la plataforma seleccionada (nube o local), certificando la solución implementada con pruebas que la organización determine que garanticen los niveles de servicio acordados.

En la fase de operación se elaboran y aprueban los procedimientos para los operadores como la creación de sistemas virtualizados, despliegue de nuevas redes de servicios o migración de volúmenes de almacenamiento, siendo el procedimiento más importante los de respaldos y restauración de componentes seleccionados lo cual garantiza la continuidad de los sistemas seleccionados.

Las fases las desarrollaremos de forma secuencial a fin de elaborar un cronograma de proyecto ordenado, sin embargo, es importante recalcar que no todas las actividades son obligatorias ya que una organización puede contar con infraestructura vigente y no desea implementar todos los dominios tecnológicos, en estos se sugiere adaptar los planes y omitir lo que considere innecesario.

Finalmente, el plan de proyecto puede ser similar a la Figura 1 considerando una implementación de todos los dominios tecnológicos.

	i Nombre de tarea	Duración	Comienzo	Fin	Predecesoras
1	▲ Implementación de centro de datos de OpenNova	43 días	jue 25/06/20	mar 25/08/20	
2	▲ Fase de análisis	7 días	jue 25/06/20	lun 6/07/20	
3	Recolectar información de requerimientos de plataforma	3 días	jue 25/06/20	lun 29/06/20	
4	Analizar la información de la infraestructura a implementar o migrar	1 día	mar 30/06/20	mar 30/06/20	3
5	Evaluar los niveles de servicio de la organización	1 día	mié 1/07/20	mié 1/07/20	4
6	Elaborar el informe de evaluación de infraestructura propuesta	2 días	jue 2/07/20	vie 3/07/20	5
7	Presentar el informe de evaluación de infraestructura propuesta	0 horas	lun 6/07/20	lun 6/07/20	6
8	▲ Fase de diseño	13 días	lun 6/07/20	jue 23/07/20	
9	Validar requerimientos mínimos de la plataforma base	1 día	lun 6/07/20	lun 6/07/20	7
10	Elaborar el diseño de alta disponibilidad de la plataforma	2 días	mar 7/07/20	mié 8/07/20	9
11	Elaborar el diseño de integración con otras plataformas pre-existentes	2 días	mar 7/07/20	mié 8/07/20	9
12	Definir los componentes para cumplir con los niveles de servicio.	2 días	jue 9/07/20	vie 10/07/20	11;10
13	Elaborar el diseño de redes	3 días	lun 13/07/20	mié 15/07/20	12
14	Elaborar el diseño de almacenamiento	3 días	lun 13/07/20	mié 15/07/20	12
15	Elaborar el diseño de computo	3 días	lun 13/07/20	mié 15/07/20	12
16	Elaborar el diseño integral de la plataforma	3 días	jue 16/07/20	lun 20/07/20	13;14;15
17	Elaborar el plan de despliegue	2 días	mar 21/07/20	mié 22/07/20	16
18	Elaborar el plan de migración	2 días	mar 21/07/20	mié 22/07/20	16
19	Presentar los diseños de la plataforma	0 días	jue 23/07/20	jue 23/07/20	17;18
20	Presentar el plan de implementación	0 días	jue 23/07/20	jue 23/07/20	17;18
21	Presentar el plan de migración	0 horas	jue 23/07/20	jue 23/07/20	17;18
22	▲ Fase de implementación	20 días	jue 23/07/20	jue 20/08/20	
23	Instalación de la plataforma	8 días	jue 23/07/20	lun 3/08/20	19;20;21
24	Despliegue de perfiles de redes	2 días	mar 4/08/20	mié 5/08/20	23
25	Despliegue de perfiles de almacenamiento	2 días	mar 4/08/20	mié 5/08/20	23
26	Despliegue de perfiles de computo	2 días	mar 4/08/20	mié 5/08/20	23
27	Elaborar el informe general de implementación.	2 días	jue 6/08/20	vie 7/08/20	24;25;26
28	Ejecutar el plan de migración	8 días	lun 10/08/20	mié 19/08/20	27
29	Elaborar el informe general de implementación.	0 horas	lun 10/08/20	lun 10/08/20	27
30	Elaborar el informe de recursos migrados	0 horas	jue 20/08/20	jue 20/08/20	28
31	▲ Fase de operación	3 días	jue 20/08/20	mar 25/08/20	
32	Ejecutar procedimientos de aseguramiento y lineamientos de seguridad	3 días	jue 20/08/20	lun 24/08/20	22
33	Ejecutar procedimientos de respaldo y monitoreo de plataforma	3 días	jue 20/08/20	lun 24/08/20	22
34	Entregar manuales de operación continua	0 horas	mar 25/08/20	mar 25/08/20	32;33

Figura 1. Plan de proyecto

Con la ejecución de las actividades de la metodología OSPa el equipo de proyecto desarrollara perfiles muy especializados en cada fase. Dependiendo del tamaño del proyecto se recomienda que el equipo sea de al menos 3 personas incluyendo al gestor de proyecto. Mientras más aplicaciones tengan el centro de datos a implementar mayor será la cantidad de personas en el equipo. En la Tabla 2 se muestra los perfiles definidos por fases, se deja a criterio del líder asignar esos roles a las mismas personas.

Tabla 2. Roles de la metodología OSPA por fases

Fase	Rol
Análisis	Analista de infraestructura
Diseño	Especialista en infraestructura Especialista en aplicaciones
Implementación	Especialista en virtualización de servidores Especialista en almacenamiento y redes
Operación	Operador de infraestructura

Considerando que se planifiquen todas las actividades de la metodología, existen situaciones donde la migración se dificulta principalmente por limitaciones tecnológicas como, por ejemplo:

RESULTADOS

Considerando que se planifiquen todas las actividades de la metodología, existen situaciones donde la migración se dificulta principalmente por limitaciones tecnológicas como, por ejemplo:

El sistema operativo a migrar no sea compatible con la plataforma oVirt porque es muy antiguo y no certifica los drivers. En este caso se sugiere implementarlo nuevamente bajo un sistema operativo soportado, pero en caso no se pueda se sugiere clonarlo y realizar las pruebas para certificarlo, pero con la intención de migrarlo en un futuro.

La aplicación primaria del sistema migrado tiene restricciones que impidan su migración por un componente de hardware o licencia. En este caso se sugiere identificar este componente primero y evaluar su migración específica con una especie de pasarela de hardware directa al sistema virtual.

El sistema operativo a implementar no está disponible en plataforma x86_64 o usa drivers propietarios no disponibles. En este caso se sugiere implementar un nodo de virtualización en la arquitectura requerida y registrarlo como cluster alternativo al cluster principal.

Considerando todas las limitaciones y sugerencias descrita, recuerde que la tecnología de virtualización le ofrecerá ventajas y beneficios que ya son conocidos y esta

demás explicarlas, sin embargo, el principal aporte de esta obra es ofrecer una alternativa libre la cual permitirá reutilizar el hardware existente en la empresa y obtener todos los beneficios a un costo cero (0) de suscripciones o licencias, dejando solo el costo de aplicación de ingeniería.

DISCUSIÓN Y CONCLUSIONES

Todos los proyectos open-source tienen la ventaja de ser desarrollados con estándares abiertos por lo tanto el nivel de integración con otras tecnologías ya sean abiertas o propietarios es muy alto. Se recomienda que revise la documentación de referencias de Rest APIs de los 4 proyectos y vera que pueden ser consumidos por otros sistemas como los de monitoreo, respaldo, reportes, etc.

No todas las actividades de la metodología OSPA son obligatorias, siéntase libre en modificarla acorde a los requerimientos de su organización.

REFERENCIAS BIBLIOGRÁFICAS

- Aguilera, L. (2020). El Manifiesto de Linux y la Historia de Linus Torvalds. Retrieved 26 June 2020, from <http://www.maestrosdelweb.com/linus/>
- Reyes, A. (2020). Implementing Model Applied to a Virtualized Data Center based on Open Source Projects. Retrieved 26 June 2020, from <http://www.testmagzine.biz/index.php/testmagzine/article/view/5280/4276>
- Brandon, J. (2019). What is Infrastructure-as-a-Service? Everything you need to know about IaaS. Retrieved 26 June 2020, from <https://www.techradar.com/news/what-is-infrastructure-as-a-service>
- FreeIPA. (2020). FreeIPA. Retrieved 26 June 2020, from <https://www.freeipa.org/>
- Foreman. (2020). Foreman. Retrieved 26 June 2020, from <https://theforeman.org/>
- Gluster. (2020). Gluster | Storage for your Cloud. Retrieved 26 June 2020, from <https://www.gluster.org/>
- oVirt. (2020). oVirt. Retrieved 26 June 2020, from <https://ovirt.org/>

Semblanza

METODOLOGÍA DE IMPLEMENTACIÓN DE CENTRO DE DATOS BASADO EN SOLUCIONES OPEN-SOURCE

Andy Reyes

Open Nova IT Consulting – CIO – Lima - Perú

ORCID: <https://orcid.org/0000-0002-4193-1069>

PUBLONS: <https://publons.com/researcher/3580360/andy-reyes/>

Correo electrónico Institucional: andy.reyes@opennova.pe

Correo electrónico Personal: controlandy@gmail.com

Alex Pacheco

Universidad Nacional de Cañete – Docente investigador – San Vicente de Cañete - Peru

ORCID: <https://orcid.org/0000-0001-9721-0730>

PUBLONS: <https://publons.com/researcher/3293185/alex-pacheco/>

Correo electrónico Institucional: apacheco@undc.edu.pe

Correo electrónico Personal: alex01_p@yahoo.es