

PROCEDIMIENTO PARA ESTIMAR LOS REQUERIMIENTOS DE CAPACIDAD INMINENTE Y A LARGO PLAZO DE LOS SERVICIOS A APROVISIONAR

El procedimiento para estimar la capacidad que demandan los servicios de la Nube Privada (NP)/Centro de Datos Virtualizados (CDV) a (re)diseñar, como muestra la [Figura 1](#), se divide en dos partes fundamentales: cálculo de los requerimientos de capacidad de los servicios a desplegar con la puesta a punto de la nueva NP, capacidad inminente; y cálculo de los requerimientos de capacidad de los servicios futuros a desplegar, capacidad a largo plazo. El procedimiento consta de seis actividades.

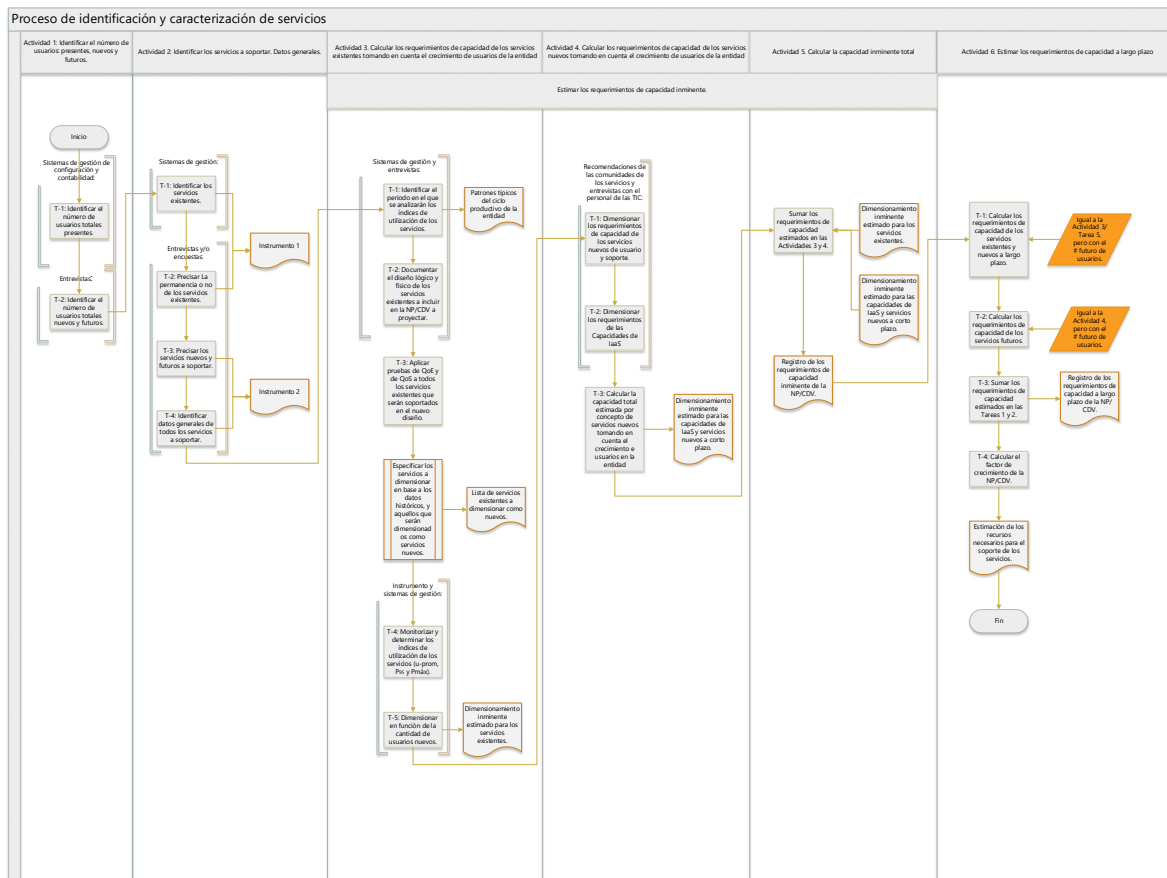


Figura 1. Procedimiento para estimar los recursos de cómputo

Actividad 1. Identificar el número de usuarios: presentes, nuevos y futuros

La Actividad 1 tiene como objetivo identificar el número de usuarios presentes, nuevos y futuros, variables indispensables en el cálculo de la estimación de los recursos que demandan los servicios ante la demanda generada por los usuarios. El índice de utilización de los servicios generada por la demanda de los usuarios presentes durante el desarrollo del proyecto, junto al número de usuarios que estarán realmente con la puesta a punto de la NP a (re)diseñar, usuarios nuevos, permite estimar la capacidad necesaria para el soporte de los servicios. El número de usuarios que estará presente en un futuro a largo plazo es considerado, para garantizar la escalabilidad horizontal de la NP/CDV. En consecuencia, la presente actividad posee dos tareas:

Tarea 1 (T-1): Identificar el número de usuarios presentes a través de los sistemas de configuración y contabilidad de la entidad. Proceder a la **Tarea 2**.

Tarea 2 (T-2): Identificar el número de usuarios nuevos y futuros a través de entrevistas con las partes interesadas de la entidad cliente. Proceder a la [Actividad 2](#).

Actividad 2. Identificar los servicios a soportar. Datos generales

La Actividad 2 tiene como objetivo identificar los servicios a aprovisionar a corto y largo plazo, tanto los tradicionales, como las categorías de servicios de la capacidad de IaaS a los usuarios finales con autoservicio y bajo demanda. Deben ser especificadas, además, características claves de los servicios que inciden directamente en el diseño de la infraestructura, como son sus niveles de criticidad

para la entidad y los RF que deben ser soportados por las categorías de la capacidad de IaaS a desplegar. Posee cuatro tareas:

Tarea 1 (T-1): Identificar los servicios existentes en la entidad y caracterizarlos de forma general, cubriendo el “[Instrumento 1](#)”. La información de esta tarea puede ser obtenida a través de los sistemas de gestión y configuración de la entidad, o mediante entrevistas con el personal de TI. Proceder a la **Tarea 2**.

Tarea 2 (T-2): Identificar de los servicios existentes los que permanecerán en la NP/CDV a (re)diseñar. Esta tarea debe ser llevada a cabo mediante entrevistas con los decisores del proyecto de la entidad y el personal de TI. Debe culminar con la finalización del cubrimiento del documento “[Instrumento 1](#)”. Proceder a la **Tarea 3**.

Tarea 3 (T-3): Identificar los servicios a nuevos y futuros a soportar. Proceder a la **Tarea 4**.

Tarea 4 (T-4): Identificar los datos generales de los servicios identificados en las **Tareas 1 y 2** haciendo uso del “[Instrumento 2](#)”. Los datos deben ser obtenidos mediante entrevistas a los decisores del proyecto de la entidad y el personal de las TIC. Proceder a la Actividad 3.

Estimación de la Capacidad Inminente:

El cálculo de la capacidad inminente se divide a su vez en dos partes: cálculo de los requerimientos de capacidad de los servicios existentes tomando en cuenta el crecimiento de usuarios de la entidad, y el cálculo de los requerimientos de capacidad de los servicios nuevos. Se divide en tres actividades:

Actividad 3. Calcular los requerimientos de capacidad de los servicios existentes tomando en cuenta el crecimiento de usuarios de la entidad

La Actividad 1 consta de cinco tareas:

Tarea 1 (T-1): Identificar el periodo en el que se analizarán los índices de utilización de los servicios:

Identificar de la entidad su(s) periodo(s) de funcionamiento típico, el cual puede ser un mes(es), 15 días, una semana(s) o días, se recomienda como máximo dos semanas; para entonces en esa etapa procesar los datos históricos de los índices de utilización de los servicios existentes, en post de estimar los recursos de cómputo, almacenamiento y red de la infraestructura de la NP/CDV a diseñar. Esta tipicidad debe ser identificada siempre que sea posible por medio del análisis de los datos históricos de los índices de utilización de los servicios de la entidad¹, y ser complementada con información obtenida mediante entrevistas a especialistas de la entidad. Se propone el empleo de la plantilla “[Patrones típicos del ciclo productivo de la entidad](#)”.

De no contarse con los datos históricos el periodo debe ser identificado mediante entrevistas a especialistas de la entidad. Como resultado de la investigación debe ser precisado la variabilidad en el año fiscal de los procesos fundamentales de la entidad, aquellos que garantizan su objeto social, en vistas a determinar los periodos de alta y baja demanda / comportamiento, junto a sus causas. El instrumento se aplicará en el periodo correspondiente a la mayor demanda para garantizar la disponibilidad y el desempeño de los servicios de las Tecnologías de

¹ Empleando los sistemas de gestión de la entidad.

la Información y las Comunicaciones (TIC), y con esto la continuidad de las operaciones básicas en todo el año.

Proceder a la **Tarea 2**.

Tarea 2 (T-2): Documentar el diseño lógico y físico de los servicios existentes a incluir en la NP a proyectar:

Deben concebirse los mapas de los diseños lógico y físico de las aplicaciones/servicios, los que deben ser anexados al documento resultante del proceso, [Informe de la capacidad estimada](#). Como resultado de este paso deben ser completadas las Tablas 1, 2², 3, 4³ y 5. Solo en el caso que la entidad incluya entre sus servicios existentes el aprovisionamiento de categorías de servicios de la capacidad de Infraestructura como Servicio (IaaS⁴) a subentidades deben ser identificados los recursos asignados en las Tablas 3, 4 y 5.

Tabla 1. Servicios, sub-servicios, aprovisionados de forma tradicional sobre nodos físicos, Bare-Metal (BM)

Servicio	Sub- servicios ⁵	Solución	Nodo

Tabla 2. Servicios, sub-servicios, aprovisionados de forma tradicional, virtualizados con “hipervisor⁶”

Servicio	Sub- servicios ⁷	Solución	IV	Nodo

Tabla 3. Aprovisionamiento de la categoría de IaaS a sub-entidades con BM.

Sub-entidad	Nodo

² Deben cubrirse tantas tablas de este tipo como soluciones de hipervisores hayan sido desplegadas en el CD.

³ Deben cubrirse tantas tablas de este tipo como soluciones de hipervisores sean empleadas para brindar IaaS.

⁴ Siglas correspondientes al término en inglés: Infrastructure as a Service.

⁵ Elemento funcional que forma parte de la topología de un servicio.

⁶ Debe ser sustituido el término “hipervisor” con el nombre del hipervisor empleado para virtualizar los servicios en cuestión.

⁷ Elemento funcional que forma parte de la topología de un servicio.

Tabla 4. Aprovisionamiento de la categoría de IaaS a sub-entidades sobre “hipervisor⁸”.

Sub-entidad	Instancia Virtual (IV)	Nodo

Tabla 5. Aprovisionamiento de la categoría de Almacenamiento como servicio (DSaaS⁹) a subentidades.

Sub-entidad	Capacidad asignada (GB)	
	Almacenamiento Objetos	Salvas

Proceder a la **Tarea 3**.

Tarea 3 (T-3): Aplicar pruebas de Calidad de Experiencia (QoE¹⁰) y de Calidad de Servicio (QoS¹¹) a todos los servicios existentes que serán soportados en el nuevo diseño:

El objetivo de estas pruebas es garantizar la fiabilidad de los índices de utilización monitorizados para proyectar los requerimientos de capacidad. No se deben tomar los datos históricos necesarios para dimensionar como válidos si los servicios en cuestión no tienen los índices de QoE y QoS esperados. En ese caso deberá recurrirse a la información de buenas prácticas y datos técnicos del fabricante del servicio, es decir, deben ser dimensionados como “servicios nuevos”, Actividad 4.

En el [Anexo A](#) se proponen pruebas de QoS y QoE para desarrollar esta tarea.

Proceder a la **Tarea 4**.

⁸ Debe ser sustituido el término “hipervisor” con el nombre del hipervisor empleado para virtualizar los servicios en cuestión.

⁹ Siglas correspondientes al término en inglés: [Data Storage as a Service](#).

¹⁰ Siglas correspondientes al término en inglés: [Quality of Experience](#).

¹¹ Siglas correspondientes al término en inglés: [Quality of Service](#).

Tarea 4 (T-4): Monitorizar y determinar los índices de utilización de los servicios (se propone el empleo a partir de este momento y hasta la **Tarea 5** de la aplicación LiliPad, su enlace en GitHub: <https://github.com/Echenique0702/Lilipad>):

Ejecutar la monitorización de los índices de utilización promedio (**u-prom**) y máximo (**u-máx**) de cada IV que soporte los elementos funcionales de los servicios, sub-servicios, en el periodo definido en la **Tarea 1**, durante la jornada laboral, con un tiempo de muestreo no superior a los cinco minutos, así como el uso de los recursos virtuales aprovisionados como capacidades de IaaS. Se debe mantener la agrupación de los servicios en función de si están virtualizados o no, y la solución de virtualización, como fue identificado en la Tarea 2¹². Las métricas a obtener por IV y nodo físico no virtualizado que soporte servicios, se indican en las Tablas 6 y 7, mientras la Tabla 8 muestra las del DSaaS.

Tabla 6. Métricas de utilización por nodo no virtualizado.

Nodo	Unidad Central de Procesamiento (CPU ¹³)			Memoria de Acceso Aleatorio (RAM ¹⁴) (MB)		Almacenamiento								Red, Ancho de Banda (BW ¹⁵) (Mbps)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
						Capacidad (GB)		Throughput						Transmisión (TX)			Recepción (RX)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
	Operaciones de Entrada/Salida por Segundo (IOPS ¹⁶)			Throughput (Mbps)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
							Capacidad	u-prom	u-máx	Capacidad	u-prom		u-máx		Capacidad	u-prom	u-máx	Capacidad	u-prom	u-máx																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
	Lectura (L)	Escritura (E)	L	E																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Días																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																

¹² A partir de este momento se les referirá como Agrupación de Recursos de Cómputo (ARC).

¹³ Siglas correspondientes al término en inglés: Central Processing Unit.

¹⁴ Siglas correspondientes al término en inglés: Random Access Memory.

¹⁵ Siglas correspondientes al término en inglés: Bandwidth.

¹⁶ Siglas correspondientes al término en inglés: Input/Output Operations Per Second.

Índices	n	...
u-prom		
P95; Pmáx		
u-prom		
P95; Pmáx		
Pmáx		
u-prom		
P95; Pmáx		
u-prom		
u-prom		
P95; Pmáx		
P95; Pmáx		
u-prom		
P95; Pmáx		
u-prom		
P95; Pmáx		

[illegible]

Sub-entidades	Capacidad asignada (GB)			
	Almacenamiento Objetos		Salvas	
	Asignado	Uso	Asignado	Uso
...				
Índices	Totales			

¹⁸ La unidad de medida depende del CMP y/o solución de virtualización empleada, generalmente número de CPU virtuales o número de núcleos físicos.

Los índices totales P_{95} y $P_{\text{máx}}$ obtenidos del periodo de monitorización de cada instancia virtual y nodo físico no virtualizado que soporte servicios, serán empleados para¹⁹: el P_{95} para estimar la capacidad que se requiere para soportar los servicios con QoS durante la jornada laboral de cada tipo de Agrupación de Recursos de Cómputo (ARC) como lo indica la Fórmula 1, Capacidad.ARC ; mientras que $P_{\text{máx}}$, empleando la Fórmula 2, será empleado para calcular la capacidad extra necesaria para la tolerancia a fallos, y a su vez satisfacer la demanda en los horarios picos de cada tipo de agrupación de recursos, $\text{Capacidad}_{\text{disp_picos.ARC}}$, como muestra la Fórmula 3. De no ser necesarios estos recursos, deberán permanecer en un estado de bajo consumo eléctrico o apagados.

$$\text{Capacidad.ARC} = \sum_{k=1}^n P_{95} \text{ de la IVk} \quad (1)$$

$$\text{Capacidad}_{\text{disp_picos.ARC}} = \text{Capacidad}_{\text{pico.ARC}} - \text{Capacidad.ARC} \quad (2)$$

En donde:

$$\text{Capacidad}_{\text{pico.ARC}} = \sum_{k=1}^n P_{\text{máx}} \text{ de la IVk} \quad (3)$$

Se empleó el P_{95} para el dimensionamiento debido a que el promedio indica el uso promedio de los recursos virtuales asignados a la instancia virtual/nodo durante el periodo típico, pero esta métrica no muestra los índices de utilización disgregados en el tiempo; así como que se considera por la autora de la presente investigación que dimensionar tomando en cuenta los picos puede conllevar a un sobre-dimensionamiento, que se ve agravado por la adquisición necesaria de hardware adicional para lograr una buena tolerancia ante fallos. El P_{95} permite eliminar datos que ocurren de forma aislada, y tomar en cuenta los valores más presentes en las

¹⁹ Las fórmulas propuestas sirven tanto para estimar las capacidades de las ARC que emplean virtualización, como aquellas que no, solo debe ser sustituida la IVk por NODOK.

muestras tomadas [2], [3], [4], [5]. Además, esta métrica está siendo introducida en la mayoría de los sistemas de gestión de servicios telemáticos y de infraestructuras de nube [2], [3], [4], [5].

En el subsistema de CPU en los nodos físicos las métricas de utilización finales a emplear en el dimensionamiento se definen como lo indica la Fórmula 4:

$$CPU_{util_{nodo}} = CPU_{util\%} * (f_{nucleo}) * (\#_{nucleos}) \quad (4)$$

En donde:

$CPU_{util_{nodo}}$: son los ciclos de CPU consumidos del nodo en GHz, puede corresponderse con el índice de utilización promedio, el 95 percentil o el valor máximo en función de los valores que tome la variable $CPU_{util\%}$.

$CPU_{util\%}$: es el porcentaje del procesamiento utilizado en el nodo. Puede tomar los valores de u-prom, P95 o Pmáx.

f_{nucleo} : frecuencia de un núcleo del CPU.

$\#_{nucleos}$: número de núcleos físicos totales del nodo.

En el caso de las IV, dado que sus procesos son planificados en los diferentes CPU virtuales (vCPU²⁰), los que a su vez son distribuidos por la solución de virtualización

²⁰ Siglas correspondientes al término en inglés: virtual CPU.

entre los CPU físicos del nodo, las métricas de utilización del subsistema de CPU a emplear en el dimensionamiento se definen como lo indica la Fórmula 5:

$$CPU_{util_{IV}} = vCPU_{util\%} * (f_{núcleo}) * (\#_{vCPU}) \quad (5)$$

En donde:

$CPU_{util_{IV}}$: son los ciclos de CPU consumidos del nodo en GHz, puede corresponderse con el índice de utilización promedio, el 95 percentil o el valor máximo en función de los valores que tome la variable $vCPU_{util\%}$.

$vCPU_{util\%}$: es el porcentaje del procesamiento utilizado en la instancia virtual. Puede tomar los valores de u-prom, P95 o Pmáx.

$f_{núcleo}$: frecuencia de un núcleo del CPU.

$\#_{vCPU}$: número de vCPU de la IV.

Se recomienda obtener:

- 1- La asignación de recursos: a partir de las plantillas definidas en el gestor del CD/CMP/hipervisor empleado.
- 2- Las métricas de utilización, **u-prom** y **u-máx**, de cada IV o nodo físico: a través del gestor del CD/CMP o de herramientas de monitorización como Zabbix.
- 3- Los índices: **u-prom**, **P95** y **Pmáx**: deben ser calculados por el equipo de diseño.

La Tabla 9 indica los valores de capacidad que deben ser especificados para una ARC y el CD en general. La capacidad del CD en general es la sumatoria de las capacidades de cada ARC. Los valores de capacidad para el DSaaS se encuentran reflejados en la Tabla 8.

Tabla 9. Métricas de capacidad requerida para una ARC y CD general

CPU (GHz)		RAM (GB)		Almacenamiento							Red BW (Mbps)			
				Capacida d (GB)	Throughput							TX	RX	TX
Capacidad	Cpacidad_disp_picos	Capacidad	Cpacidad_disp_picos	Capacidad	IOPS	Cpacidad_disp_picos	Throughput (Mbps)				Capacidad	Capacidad	Cpacidad_disp_picos	Cpacidad_disp_picos
							Capacidad		Cpacidad_disp_picos					
							L	E	L	E				

Proceder a la **Tarea 5**.

Tarea 5: Dimensionar en función de la cantidad de usuarios nuevos. Se propone aplicar la Fórmula 6 a los valores de las métricas de capacidad obtenidas en la **Tarea 4**, por ARC y a nivel de CD, así como a los del DSaaS. Los valores deben ser consultados con el personal de las TIC de la entidad y en caso de modificaciones, estas deben quedar registradas.

$$Utilizaciónnueva = Utilizaciónactual * \frac{Usuariosnuevos}{Usuariosactuales} * \beta \quad (6)$$

β constituye el factor de comportamiento de los usuarios nuevos respecto a los usuarios existentes, en relación a su patrón de uso de los servicios. Si el comportamiento de los usuarios nuevos es igual al de los existentes su valor numérico debe ser uno.

Actividad 4. Calcular los requerimientos de capacidad de los servicios nuevos tomando en cuenta el crecimiento de usuarios de la entidad

La Actividad 4 consta de tres tareas:

Tarea 1 (T-1): Dimensionar los requerimientos de capacidad de los servicios de usuario y soporte:

Consultar las recomendaciones y/o indicaciones del ecosistema entorno del servicio en cuanto a la tecnología de virtualización a emplear y su dimensionamiento; así como encuestar a la entidad cliente en cuanto a sus proyecciones. Se aboga por el empleo en primer lugar de la virtualización OSLV, y de no ser posible emplear la HVM.

Deben ser estimadas las demandas para el número de usuarios que se deben tomar en cuenta a la hora de la puesta a punto del sistema. En post de tributar a la tolerancia a fallos y al soporte de los horarios pico, se propone identificar el porcentaje que representa la **Capacidad_disp_picos** de la **Capacidad** general del CD, procedente de los resultados de la **Actividad 3 / Tarea 5**, correspondiente a cada subsistema: CPU, RAM, almacenamiento y red; y entonces aplicarles ese porcentaje a los valores resultantes del estudio de las recomendaciones de las comunidades y/o proveedores. La Tabla 10, 11 y 12 muestran los datos a obtener, por ARC, agrupando las aplicaciones/servicios por el tipo de virtualización de servidores que se estima emplear.

Tabla 10. Demanda de los servicios de usuario y soporte nuevos de una ARC virtualizada con una solución de tipo OSLV

Servicio	Sub-servicio	Solución	IV	CPU (GHz)		RAM (GB)	Almacenamiento				Red BW (Mbps)			
							Capacidad (GB)	Throughput			T X	R X	T X	R X
				Capacidad	IOPS	Throughput (Mbps)		Capacidad	Capacidad_disp_picos	Capacidad	Capacidad_disp_picos	Capacidad	Capacidad_disp_picos	
						Capacidad								Capacidad_disp_picos

[illegible]

Tabla 11. Demanda de los servicios de usuario y soporte nuevos de una ARC virtualizada con una solución de tipo HVM

[illegible]

Tabla 12. Demanda de los servicios de usuario y soporte nuevos de una ARC no virtualizada

[illegible]

[illegible]

Σ Totales													
------------------	--	--	--	--	--	--	--	--	--	--	--	--	--

Tabla 14. Demanda de los servicios de IaaS nuevos de una ARC virtualizada con una solución de virtualización de tipo HVM

Sub-entidades	CPU (GHz)		RAM (GB)		Almacenamiento						Red BW (Mbps)			
					Capacidad (GB)	<u>Throughput</u>				TX	RX	TX	RX	
	IOPS	<u>Throughput</u> (Mbps)												
		Capacidad	Capacidad			Cpacidad_disp_picos								
			L	E		L	E							
	Capacidad	Cpacidad_disp_picos	Capacidad	Cpacidad_disp_picos	Capacidad	IOPS	L	E	L	E	Capacidad	Capacidad	Cpacidad_disp_picos	Cpacidad_disp_picos
ΣTotales														

Tabla 15. Métricas de utilización en los servicios DSaaS aprovisionados.

Sub-entidades	Capacidad estimada (GB)	
	Almacenamiento Objetos	Salvas
Σ Totales		

Tarea 3 (T-3): Calcular la capacidad total estimada por concepto de servicios nuevos tomando en cuenta el crecimiento de la entidad:

Deben sumarse los valores de capacidad estimados en las **Tareas 1 y 2**. Debe ser sumados los valores totales de cada subsistema: vCPU, RAM, almacenamiento y red, de las capacidades correspondientes a una misma ARC: ARC virtualizada con “hipervisor²¹” o ARC-BM, obteniéndose para cada ARC una Tabla igual a la 9. Posteriormente sumar las capacidades estimadas de cada ARC para identificar el total de la capacidad que demandan los servicios nuevos a nivel del CD, tomando

²¹ Debe ser sustituido el término “hipervisor” con el nombre del hipervisor empleado para virtualizar los servicios en cuestión.

en consideración el crecimiento de la entidad. La demanda de la capacidad de almacenamiento para el servicio DSaaS ya se encuentra reflejado en la Tabla 15.

Proceder a la [Actividad 5](#).

Actividad 5. Calcular la capacidad inminente total

El cálculo de la capacidad inminente total consiste en sumar las capacidades totales estimadas en:

- El cálculo de los requerimientos de capacidad de los servicios existentes tomando en cuenta el crecimiento de usuarios de la entidad, [Actividad 3](#); y en
- El cálculo de los requerimientos de capacidad de los servicios nuevos tomando en cuenta el crecimiento de usuarios de la entidad, [Actividad 4](#).

El resultado debe ser registrado en el “[Informe de la capacidad estimada](#)”.

Estimación de la capacidad a largo plazo

Actividad 6. Calcular los requerimientos de capacidad a largo plazo

La Actividad 6 consta de tres tareas:

Tarea 1 (T-1): Calcular los requerimientos de capacidad de los servicios existentes y nuevos a largo plazo:

Tomar los resultados obtenidos en “Cálculo de la capacidad inminente total” y aplicar la **Actividad 3 / Tarea 5**, pero empleando el número de usuarios futuros.

Proceder a la **Tarea 2**.

Tarea 2 (T-2): Calcular los requerimientos de capacidad de los servicios futuros:

Se propone el mismo procedimiento descrito en la **Actividad 4**, pero tomando como número de usuarios, el total de usuarios futuros. Proceder a la **Tarea 3**.

Tarea 3 (T-3): Calcular la capacidad a largo plazo total:

El cálculo de la capacidad a largo plazo total consiste en sumar las capacidades totales estimadas en:

- El cálculo de los requerimientos de capacidad de los servicios existentes y nuevos a largo plazo, **Tarea 1**; y en
- El cálculo de los requerimientos de capacidad de los servicios futuros, **Tarea 2**.

El resultado debe ser registrado en el "[Informe de la capacidad estimada](#)". Proceder a la **Tarea 4**.

Tarea 4 (T-4): Calcular el "Factor de Crecimiento de la NP/CDV a Largo Plazo"

El "Factor de Crecimiento de la NP/CDV a Largo Plazo (FCLP)" contribuye a garantizar la EH de la NP/CDV, sin afectar: los niveles de desempeño requeridos; los presupuestos anuales correspondientes a las Inversiones de Capital (CAPEX²²) y a los Gastos de Operaciones (OPEX²³), específicamente a los gastos de consumo de energía eléctrica y a la gestión de los activos; la selección de la Plataforma de Gestión de Nube (CMP²⁴) y/o plataforma de virtualización de NP; y los recursos facilitadores.

Su cálculo requiere de las siguientes acciones:

²² Siglas correspondientes al término en inglés: Capital Expenditure.

²³ Siglas correspondientes al término en inglés: Operational Expenditures.

²⁴ Siglas correspondientes al término en inglés: Cloud Management Platform.

Acción 1: sumar los requerimientos de capacidad inminentes de CPU, RAM y almacenamiento de las ARC, incluyendo la demanda para los horarios picos, como indican las Fórmulas 7-11. Solo son empleadas para el cálculo las métricas de capacidad de CPU, de RAM y de almacenamiento porque son las que de manera general definen el número de nodos de cómputo, de almacenamiento y/o de nodos convergentes²⁵.

$$CPU_{total} = \sum_{i=1}^n CPU_{ARC_i} \quad (7)$$

$$RAM_{total} = \sum_{i=1}^n RAM_{ARC_i} \quad (8)$$

$$SA_{total} = \sum_{i=1}^n SA_{ARC_i} \quad (9)$$

$$CPU_{pico} = \sum_{i=1}^n CPU_{pico_{ARC_i}} \quad (10)$$

$$RAM_{pico} = \sum_{i=1}^n RAM_{pico_{ARC_i}} \quad (11)$$

Acción 2: sumar los requerimientos de capacidad a largo plazo de CPU, RAM y almacenamiento de las ARC, incluyendo la demanda para los horarios picos, como indican las Fórmulas 12-16.

$$CPU_{LP_{total}} = \sum_{i=1}^n CPU_{LP_{ARC_i}} \quad (12)$$

$$RAM_{LP_{total}} = \sum_{i=1}^n RAM_{LP_{ARC_i}} \quad (13)$$

$$SA_{LP_{total}} = \sum_{i=1}^n SA_{LP_{ARC_i}} \quad (14)$$

$$CPU_{LP_{pico}} = \sum_{i=1}^n CPU_{pico_{LP_{ARC_i}}} \quad (15)$$

$$RAM_{LP_{pico}} = \sum_{i=1}^n RAM_{pico_{LP_{ARC_i}}} \quad (16)$$

Acción 3: Calcular los factores de crecimiento por métricas de capacidad por subsistema como indican las Fórmulas 17-21.

²⁵ Nodos convergentes: nodos que juegan el role de nodos de cómputo y nodos de almacenamiento.

$$FC_{CPU} = \frac{CPU_LP_{total}}{CPU_{total}} \quad (17)$$

$$FC_{RAM} = \frac{RAM_LP_{total}}{RAM_{total}} \quad (18)$$

$$FC_{SA} = \frac{SA_LP_{total}}{SA_{total}} \quad (19)$$

$$FC_{CPU_{pico}} = \frac{CPU_LP_{pico}}{CPU_{pico}} \quad (20)$$

$$FC_{RAM_{pico}} = \frac{RAM_LP_{pico}}{RAM_{pico}} \quad (21)$$

Acción 4: Calcular el FCLP, para una infraestructura convergente emplear la Fórmula 22, para una infraestructura no convergente emplear la Fórmula 23.

$$FCLP = \max\{FC_{CPU}; FC_{RAM}; FC_{SA}\} + \max\{FC_{CPU_{pico}}; FC_{RAM_{pico}}\} \quad (22)$$

$$FCLP = \max\{FC_{CPU}; FC_{RAM}\} + FC_{SA} + \max\{FC_{CPU_{pico}}; FC_{RAM_{pico}}\} \quad (23)$$

Documentar los resultados del procedimiento para estimar los requerimientos de capacidad

El procedimiento debe ser documentado empleando la plantilla “[Informe de la capacidad estimada](#)”.

Anexos

Anexo A. Propuesta de proyecto de pruebas para evaluar la QoE y QoS de servicios de usuario y soporte

La Figura A.1 muestra la propuesta del proyecto de pruebas para evaluar de forma general la QoE y QoS de los servicios de usuario brindados en la entidad para entonces realizar el estudio de los datos históricos. Puede observarse que consta de tres partes elementales: pruebas de QoS en la red de la entidad, pruebas de

QoS de los servicios y pruebas de QoE. El proyecto de pruebas se encuentra basado en [6].

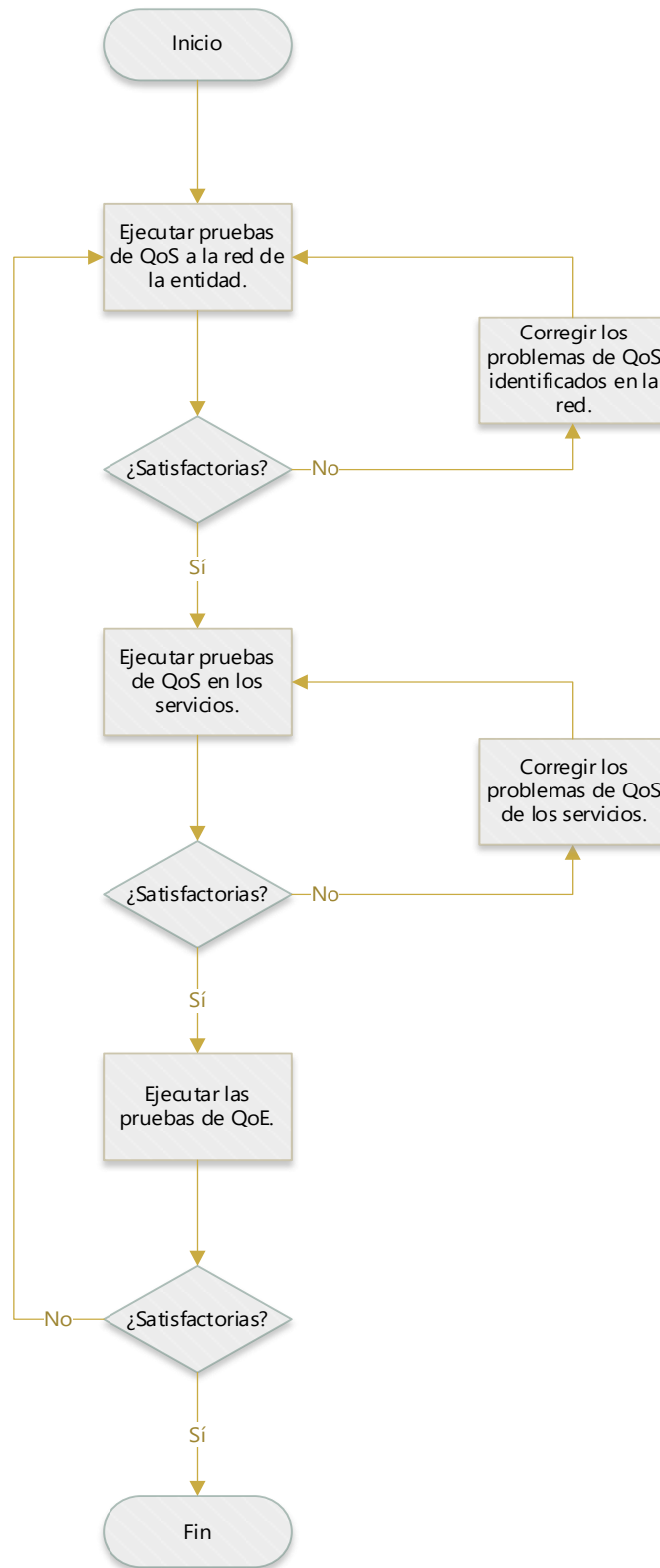


Figura A.1. Procedimiento para evaluar la QoE y QoS de servicios de usuarios

Pruebas para evaluar el desempeño de la red

Las métricas a evaluar para identificar el desempeño de la red del cliente e identificar los cuellos de botella son: throughput, tiempo de respuesta, pérdida de paquetes y demoras. Se propone dividir la red en segmentos lógicos, o sea entre elementos lógicos, partiendo desde el usuario final hasta el cortafuegos que se encuentra de “cara a Internet”, tratando de abarcar la mayor extensión posible de la red. La Figura A.2 muestra un escenario típico para la realización de la prueba, el tiempo de procesamiento de los elementos de red no se tomará en cuenta debido a que ronda el orden de los nanosegundos [7], [8].

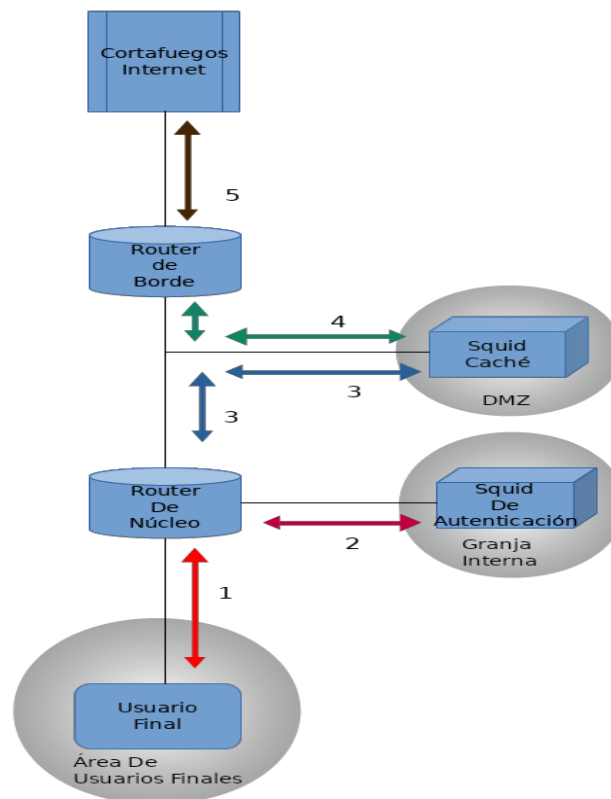


Figura A.2. Escenario típico del diagrama lógico de una red de una entidad

Las métricas de desempeño deben cumplir con los umbrales:

- Demora: preferiblemente menor que 2s, aceptable menor que 4s.
- Pérdida de Paquetes: 0%.
- Tiempo de respuesta: hasta 100ms.

De no ser favorables los resultados es necesario identificar y resolver los problemas para proseguir con el proyecto de pruebas de QoE y QoS. Si los resultados cumplen con los umbrales propuestos se deben realizar las pruebas de QoS a los servicios de usuario y soporte pertinentes.

Prueba para evaluar el throughput de la red:

Nombre de la Prueba: Throughput de la red.

Tipo de prueba: Micro-benchmark.

Objetivo de la prueba: Obtener los valores de throughput que se pueden alcanzar en la red.

Número de iteraciones: Cinco.

Parámetros a ser medidos y medios a emplear: El parámetro a ser medido es el throughput y la herramienta empleada es **Iperf**.

Descripción del escenario y las pruebas: El escenario aproximado es el que se muestra en la Figura A.2. Se propone identificar el horario pico del consumo del ancho de banda de la red, y ejecutar las pruebas en ese periodo.

Instalación y configuración:

- 1- Instalar Iperf en el nodo físico o IV a emplear como origen y destino.
- 2- En el servidor se ejecutará **iperf -s**.

3- En el cliente se ejecutará **iperf -c “IP del servidor” -i 0 -t 60** (cliente al que los paquetes se enviarán seguidos, sin intervalo de tiempo entre ellos y por un tiempo de 60 segundos).²⁶

Desarrollo de la prueba:

- 1- Ejecutar **iperf** en modo cliente y servidor para enviar paquetes TCP durante 60 segundos y medir el throughput.
- 2- Limpieza del escenario: Desinstalar la herramienta **iperf**.

Cálculo y agregación de métricas:

- 1- Calcular el throughput promedio de las iteraciones para cada escenario de la forma que se propone en la Tabla A.1.

Tabla A.1. Cálculo de throughput

	Iteración #1	Iteración #2	Iteración #3	Iteración #4	Iteración #5	Promedio
Segmento 1						
Segmento 2						
Segmento 3						
Segmento 4						
Segmento 5						
Segmento 6						

- 2- Identificar el segmento con menor throughput, ya que este será el que definirá el throughput del tramo de red analizado.

Prueba de tiempo de respuesta

Nombre de la Prueba: Tiempo de respuesta de la red.

²⁶ Iperf utiliza tramas TCP por defecto.

Tipo de prueba: Micro-benchmark.

Objetivo de la prueba: Obtener los valores de los tiempos de respuestas propiciados por la red.

Número de iteraciones: Cinco.

Parámetros a ser medidos y medios a emplear: Los parámetros a ser medidos son el tiempo de respuesta mínimo, promedio y máximo. Se propone utilizar la herramienta **Ping**.

Descripción del escenario y las pruebas: Se recomienda el mismo escenario e indicaciones que en la prueba de throughput. Se enviarán 100 000 peticiones ICMP de forma que se inunde la red.

Instalación y configuración: Instalar **Ping** en cada IV o nodo físico identificado como origen y destino.

Desarrollo de la prueba:

- 1- Ejecutar desde la terminal **sudo ping -f -c 100 000 "IP-Destino"**.
- 2- Recoger las métricas especificadas en la descripción.

Limpieza del escenario: Dejar la herramienta **Ping** instalada ya que puede ser útil en otro momento.

Cálculo y agregación de métricas:

- 1- Completar la Tabla A.2 con las métricas obtenidas.
- 2- Calcular el tiempo de respuesta total de la red sumando los valores promedios de los diferentes escenarios.

Tabla A.2. Valores de tiempos de respuestas

	Iteración # 1	Iteración # 2	Iteración # 3	Iteración # 4	Iteración # 5	Promedio
--	---------------	---------------	---------------	---------------	---------------	----------

	Promedio	Máximo	Mínimo	Promedio	Máximo	Mínimo	Promedio	Máximo	Mínimo	Promedio	Máximo	Mínimo	Promedio	Máximo	Mínimo	Promedio	Máximo	Mínimo
Segment o 1																		
Segment o 2																		
Segment o 3																		
Segment o 4																		
Segment o 5																		
Total (Sumatoria de los valores finales)																		

- 3- Evaluar la red en función del valor promedio máximo obtenido: si supera los 100ms hay demoras, la red presenta problemas.

Prueba de demoras:

Nombre de la Prueba: Demoras de la red.

Objetivo: Obtener métricas de demoras de la red.

Duración: Relativa. La duración de esta prueba depende de las habilidades del personal de TI y del conjunto de datos recogido.

Número de iteraciones: Cinco.

Parámetros a ser medidos y medios a emplear: Se emplea para esta prueba los datos de las pruebas de tiempo de respuesta y herramientas de análisis estadístico como Microsoft Excel.

Descripción de la Prueba: La realización de esta prueba se hace en base a los datos de Tiempo de Respuesta, donde se define un Umbral de demora por el cliente. Si el cliente no define umbrales, en [7], [8] proponen que sea 100 ms para que el usuario no perciba demoras. De los valores que exceden el umbral de tiempo de respuesta seleccionar el mínimo, máximo y el promedio.

Pérdidas de paquetes:

Esta prueba se debe realizar al mismo tiempo que la prueba de tiempo de respuesta recolectando los valores de pérdida de paquetes para cubrir la Tabla A.3.

Tabla A.3 Cálculo de las pérdidas de paquetes

	Iteración # 1	Iteración # 2	Iteración # 3	Iteración # 4	Iteración # 5	Promedio
Segmento 1						
Segmento 2						
Segmento 3						
Segmento 4						
Segmento 5						
Total (Σ)						

Cálculo y agregación de métricas: Cubrir la Tabla A.3 con los valores obtenidos en las cinco iteraciones. Si la pérdida de paquetes supera el 0,1% la red presenta problemas.

Identificación del cuello de botella

Luego de verificar los valores obtenidos se analizará en cada segmento de red la relación de parámetros más pobres de throughput, tiempo de respuesta y demoras, para así detectar en qué segmento lógico de la red existe el cuello de botella.

Pruebas de QoS a servicios de usuario y/o soporte

Para la evaluación de la QoS a servicios de usuario y/o soporte se propone la realización de pruebas de carga y monitorización en entornos de desarrollo. La Figura A.3 ilustra el escenario a desplegar para la realización de las pruebas. En él se propone la ubicación de los elementos a emplear. Los generadores de carga se deben colocar en clientes comunes de la infraestructura, las herramientas de monitorización deben estar configuradas para monitorizar los servicios a probar, así como todos los elementos de la red y la infraestructura presentes en la prueba.

Las herramientas de benchmark se deben colocar dentro de IV desplegadas en el CD. La Figura A.4 muestra el procedimiento a realizar para desplegar las herramientas de benchmark dentro del CD.

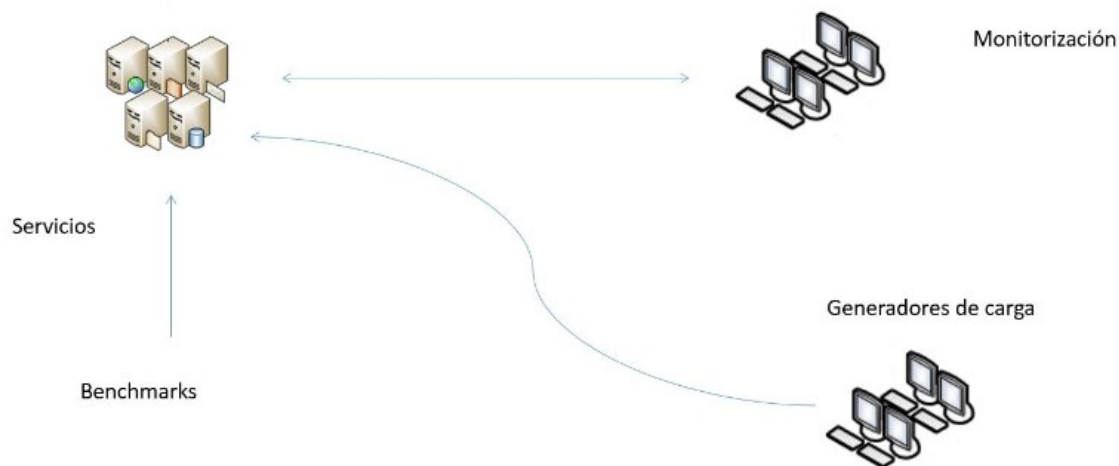


Figura A.3. Escenario de pruebas

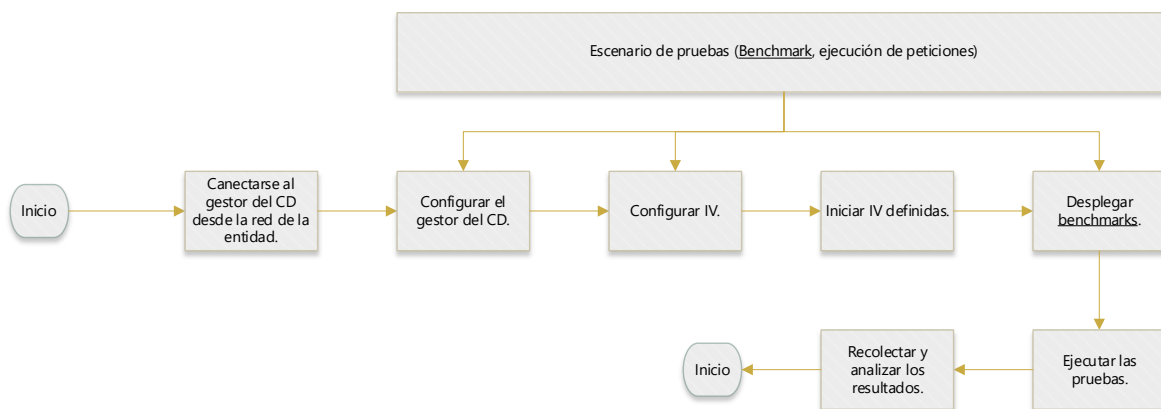


Figura A.4. Procedimiento a seguir para desplegar los benchmarks

Pruebas de throughput

Nombre de la prueba: Prueba de throughput de aplicaciones por servicio.

Tipo de prueba: Carga y monitorización.

Objetivo de las pruebas: Especificar el throughput de las aplicaciones para perfiles de carga definidos de forma individual.

Duración: 3 horas, descartando los resultados de la primera y última media hora.

Número de iteraciones: 10.

Parámetros a ser medidos y medios a emplear: La Tabla A.4 muestra las métricas de throughput por servicio y las herramientas recomendadas para obtenerlas.

Tabla A.4. Métricas de throughput por servicio

Servicio	KPI	Herramientas
DNS	Peticiones/segundo	Namebench generador de carga específico para DNS. Sistemas de gestión como Zenoss, Munin, Zabbix para monitorizar las métricas de la generación de carga. JMeter para generar carga.
	Actualizaciones dinámicas por segundo	
	Peticiones resolutivas por segundo	
	Peticiones cacheadas por segundo	
Web	Peticiones/segundo recibidas, servidas y generadas	JMeter Zenoss Zabbix
Proxy	Peticiones/segundo	JMeter Zenoss Zabbix
FTP	<u>Throughput</u> de salida máximo del servidor MB/s	JMeter Zenoss Zabbix
BD	Consultas/segundo	Sysbench, <u>benchmark</u> para sistemas de BD. Zenoss Zabbix
<u>Big Data</u>	Consultas/segundo <u>Throughput</u> de la operación ²⁷	YCSB, <u>benchmark</u> de <u>Big Data</u> . Zenoss Zabbix
Servidor de aplicaciones	Métricas propias de la aplicación.	Herramientas propias de la aplicación.
HPC	<u>Throughput</u> total en GFlops/s	HPCC, <u>benchmark</u> de computación de alto rendimiento. Zenoss Zabbix
Servidor de correo	peticiones/segundo peticiones servidas/peticiones realizadas peticiones fallidas/peticiones realizadas longitud de la cola	JMeter Zenoss Zabbix

Descripción de las pruebas:

²⁷ Se refiere al throughput de las operaciones realizadas en el clúster de Big Data.

En cada prueba se realizan las pruebas a un servicio, y se procede a iniciar la siguiente.

Instalación y configuración:

- 1- Instalar JMeter para HTTP, FTP, Proxy, POP, IMAP y SMTP; Namebench para DNS, HPCC para HPC, YCSB para Big Data y Sysbench para BD. Los generadores de carga deben instalarse en distintos nodos, uno por cada IV bajo pruebas, excepto HPCC y Sysbench que se instalan directamente en los servidores a probar. El número de hilos de cada JMeter se determinará en base a los datos propios de cada aplicación/servicio. Las configuraciones de Sysbench para BD vienen dadas por las tareas y tecnología de la BD a emplear. Los tiempos de ascenso de las cargas y el número de estas serán definidas por los perfiles de carga a aplicar. Definir recuperadores de datos en el JMeter de tipo Resumen de operaciones y Gráfico de Resultados.
- 2- Es recomendable la utilización de la herramienta de monitorización ya existente en la entidad. En escenarios donde no sea posible, emplear Zabbix, instalando el frontend en un equipo no sometido a pruebas, instalando y configurando agentes en todos los elementos de la red asociados a las pruebas. Las plantillas, elementos y disparadores pertinentes deben ser instalados.

Desarrollo de las pruebas:

- 1- Aplicar perfiles de carga ascendente durante 30 minutos y posteriormente detener las pruebas por un intervalo de 10 minutos. Aplicar perfiles de carga descendente durante 30 minutos y posteriormente detener las pruebas durante 10 minutos.

- 2- Aplicar perfiles de carga por picos durante 30 minutos en intervalos de 5 minutos cada uno de carga y descanso.
- 3- Esperar un intervalo de 10 minutos y emplear perfiles de sobrecarga, durante 10 minutos con descansos de 1 minuto.
- 4- Recoger las métricas definidas de los generadores de carga y las herramientas de gestión.

Limpieza del escenario: Detener todos los procesos de generación de carga dispuestos.

Cálculo y agregación de métricas:

- 1- Calcular media, máximo, desviación estándar y P_{95} para cada valor de las métricas en cada iteración.
- 2- Se promediarán los valores P_{95} obtenidos de las 10 iteraciones para cada servicio/aplicación.
- 3- Se debe evaluar si el valor obtenido es satisfactorio, de no serlo el servicio debe ser rediseñado lógica y/o físicamente.

Pruebas de tiempo de respuesta

Nombre de la prueba: Prueba de tiempo de respuesta de aplicaciones por servicio.

Tipo de prueba: Carga y monitorización.

Objetivo de las pruebas: Obtener los tiempos de respuesta de los servicios individualmente.

Duración: tres horas, descartando los resultados de la primera y última media hora.

Número de iteraciones: 10.

Parámetros a ser medidos y medios a emplear: La Tabla A.5 muestra las métricas de tiempo de respuesta por servicio y las herramientas recomendadas para obtenerlas.

Tabla A.5. Métricas de tiempo de respuesta por servicio

Servicio	KPI	Herramientas
DNS	Demoras de la resolución de nombres de dominios (lookup) con y sin la caché activa.	Namebench Zenoss Munin Zabbix JMeter
Web	Tiempo de respuesta promedio, máximo y 95 percentil (de una petición y de un grupo de ellas).	JMeter Zenoss Zabbix
Proxy		JMeter Zenoss Zabbix
FTP		JMeter Zenoss Zabbix
BD		Sysbench Zenoss Zabbix
<u>Big Data</u>		YCSB Zenoss Zabbix
Servidor de aplicaciones		Herramientas propias de la aplicación.
HPC		HPCC Zenoss Zabbix
Servidor de correo		JMeter Zenoss Zabbix

Descripción de las pruebas:

En cada prueba se realizan las pruebas a un servicio, y se procede a iniciar la siguiente.

Instalación y configuración: similar a la prueba de throughput. Definir en el JMeter además el Gráfico de Tiempo de Respuesta.

Desarrollo de las pruebas:

- 1- Aplicar perfiles de carga ascendente durante 30 minutos y posteriormente detener las pruebas por un intervalo de 10 minutos. Aplicar perfiles de carga

descendente durante 30 minutos y posteriormente detener las pruebas durante 10 minutos.

- 2- Aplicar perfiles de carga por picos durante 30 minutos en intervalos de 5 minutos cada uno de carga y descanso de 5 minutos entre cargas.
- 3- Esperar un intervalo de 10 minutos y emplear perfiles de sobrecarga, durante 10 minutos con descansos de 1 minuto.
- 4- Recoger las métricas definidas de los generadores de carga y las herramientas de gestión.

Limpieza del escenario: Detener todos los procesos de generación de carga dispuestos.

Cálculo y agregación de métricas:

- 1- Calcular media, máximo, desviación estándar y P_{95} para cada valor de las métricas en cada iteración.
- 2- Se promediarán los valores P_{95} obtenidos de las 10 iteraciones para cada servicio/aplicación.
- 3- Evaluar los valores obtenidos, de no ser satisfactorios deben ser corregidos los problemas, de lo contrario se debe proceder con las pruebas de QoE.

Pruebas de QoE

Nombre de la prueba: Prueba de QoE para servicios de usuario y soporte que no sean de voz y/o video sobre redes IP.

Tipo de prueba: monitorización.

Objetivo de las pruebas: Identificar la calidad de los servicios percibida por los usuarios finales.

Duración: tres horas.

Número de iteraciones: 3.

Parámetros a ser medidos y medios a emplear:

Por parte de los usuarios: debe ser aplicada una encuesta para conocer la QoE experimentada por los usuarios con el servicio objeto de estudio. En la presente propuesta se toma como propuesta y ejemplo el servicio de navegación a Internet, por lo que se propone la plantilla en el [Anexo B](#). Se emplea la escala Mean Opinion Scale (MOS) para evaluar la QoE de la navegación a Internet: Excelente (5), Bien (4), Regular (3), Pobre (2) y Mal (1).

Por parte de la administración de la infraestructura y servicios de red:

- Medio a emplear: sistemas de gestión de desempeño de los servicios, **logs** de los servicios objeto de estudio. En el caso del servicio de navegación a Internet se propone la aplicación descrita en el [Anexo C](#), la que tiene como objetivo evaluar la QoS brindada al usuario durante su interacción con cada sitio.
- Métricas: deben ser evaluadas las KPI del servicio en relación a la QoE respecto al usuario. En el caso del servicio de navegación a Internet, se evaluará el throughput y las demoras del servicio.

Descripción de las pruebas:

La prueba de la QoE debe realizarse en un periodo de trabajo típico de la entidad, en donde los usuarios deben acceder a 20 sitios pertinentes a la entidad, y evaluar el desempeño percibido empleando la escala MOS. Una vez obtenida la QoE de los usuarios, estos deben ser agrupados atendiendo a la evaluación brindada según se percepción de la calidad del servicio, promedio de las evaluaciones otorgadas a

cada acápite de la encuesta, de manera que existan cinco grupos: Excelente, Bien, Regular, Pobre y Mal, y descartar las respuestas consideradas incoherentes.

Una vez con los criterios consistentes, son promediadas las evaluaciones para obtener una opinión final, y los porcentajes por categorías. Los usuarios deben ser como mínimo seis de cada comunidad de usuarios de la entidad.

Instalación y configuración:

- Es recomendable la utilización de la herramienta de monitorización ya existente en la entidad. En escenarios donde no sea posible, emplear Zabbix, instalando el frontend en un equipo no sometido a pruebas, instalando y configurando agentes en todos los elementos de la red asociados a las pruebas. Las plantillas, elementos y disparadores pertinentes deben ser instalados.
- Instalación y uso del **script** para obtener la QoS brindada a los usuarios, como se indica en el [Anexo C](#).

Desarrollo de las pruebas:

1- Seleccionar el periodo de ejecución de la prueba: la realización de la prueba debe corresponderse con un período típico de alta demanda por parte de los procesos fundamentales de la entidad, aquellos que garantizan su objeto social, en post de garantizar la disponibilidad y el desempeño de los servicios TIC, y con esto la continuidad de las operaciones básicas en todo el año. En el caso del servicio de navegación a Internet, el día de la ejecución deben ser identificados los tres momentos del día en que ocurren los horarios picos. En esos horarios deben ser ejecutadas las pruebas.

2- Conformar la encuesta: para conformar la encuesta se hace necesario seguir las siguientes tareas:

2.1- Identificar las comunidades de usuario de la entidad. Una comunidad de usuario debe caracterizarse por su Área Objetivo Social dentro de la entidad.

2.2- Identificar las redes lógicas que sirven a las diferentes comunidades de usuario.

2.3- Identificar los servicios de usuario objeto de estudio por comunidad de usuario. En el caso específico del servicio de navegación a Internet:

2.3-1. Identificar, haciendo uso de los datos históricos de los **logs** de la entidad, generalmente de los **logs** de los proxys, **servicio squid** generalmente, los sitios más visitados a nivel de entidad y por comunidades de usuario.

2.3-2. Presentar una lista con los 10 sitios más visitados a nivel de entidad y de comunidad de usuario a los decisores del proyecto, para que se concilie la pertinencia para la entidad de los sitios a evaluar, y en consecuencia se modifique la lista de sitios en función de los intereses de la organización cliente. Debe obtenerse como máximo diez sitios a nivel de entidad, y diez por comunidad de usuarios.

2.3-3. Confeccionar la encuesta empleando la plantilla propuesta en el [Anexo B](#).

3- Ejecutar la encuesta en las diferentes comunidades de usuarios, con no menos de seis usuarios por comunidad, en los tres horarios del día seleccionados. En el caso del servicio de navegación a Internet, cada usuario

no debe dedicarle más de tres minutos a cada sitio, por lo que en cada momento del día la prueba no debe excederse de una hora. Los usuarios deberán estar conectados a través de redes de acceso con las mismas prestaciones. Los terminales deberán tener similares propiedades ya que estas influyen en la experiencia de navegación.

- 4- Colectar las encuestas y calcular por usuario el promedio de la QoE experimentada. En el caso del servicio de navegación a Internet sería el promedio de la QoE percibida en los 20 sitios indicados, de esta forma se obtiene una opinión general del usuario en el momento de la prueba de navegación.
- 5- Identificar el throughput experimentado por cada usuario durante la prueba de QoE. En el caso del servicio de navegación a Internet se propone el empleo de la aplicación descrita en el [Anexo C](#).
- 6- Agrupar a los usuarios de acuerdo al grupo que le corresponde en función de la QoE resultante.
- 7- Identificar en cada grupo el valor promedio de throughput.
- 8- Eliminar en cada grupo a aquellos usuarios cuyo throughput es menor al máximo valor de los grupos inferiores en QoE, sus opiniones no serán tomadas en cuenta para el procesamiento final de los datos.

Cálculo y agregación de métricas:

Calcular las estadísticas de los porcentajes de usuarios que emitieron criterios en cada una de las posibles categorías: Excelente, Bien, Regular, Pobre y Mal. Para que la prueba sea considerada exitosa como promedio debe obtenerse un criterio superior a Bien.

Anexo B. Plantilla para aplicación de encuesta de QoE en el servicio de navegación a Internet

Cuestionario para la realización de pruebas de QoE para el servicio de navegación a Internet

Nombre y apellidos: _____

Área en que se aplica el instrumento: _____

Estudiante: _____, año: _____

Profesor: _____

Usuario: _____

Carnet de identidad: _____

Fecha: _____ Hora de inicio _____ y fin _____

Indicaciones: el usuario debe acceder a los sitios especificados y en estos de ser posible:

- Consultar documentos online
- Descargar documentos
- Acceder a archivos multimedia (videos)
- Descargar archivos multimedia (videos)
- No debe demorar más de tres minutos en cada sitio
- Evaluar la experiencia en cada sitio de: Excelente, Bien, Regular, Pobre o Mal

Sitios pertinentes a nivel de entidad objetos de estudio:

Sitio 1: _____

Calidad de experiencia percibida: __ Excelente __ Bien __ Regular __ Pobre __ Mal

...

Sitio 10: _____

Calidad de experiencia percibida: __ Excelente __ Bien __ Regular __ Pobre __ Mal

Sitios pertinentes a nivel de comunidad de usuario objetos de estudio:

Sitio 1: _____

Calidad de experiencia percibida: __ Excelente __ Bien __ Regular __ Pobre __ Mal

...

Sitio 10: _____

Calidad de experiencia percibida: __ Excelente __ Bien __ Regular __ Pobre __ Mal

Firma: _____

Gracias por su colaboración

Anexo C. Aplicación para evaluar la QoS brindada a los usuarios durante la prueba de QoE

La aplicación tiene como objetivo calcular el throughput, definido por la Fórmula C.1, y la demora promedio por usuario extraída de los **logs** del **squid** como muestra la Figura C.1, así como también contabilizar los errores obtenidos por estos usuarios durante la prueba, a modo de evaluar el servicio de navegación por Internet.

$$Throughput = \frac{bytes * 8}{demora / 10^3} \quad (C.1)$$

```
1522341451.387    951 10.8.88.133 TCP_MISS/200 590 GET
http://detectportal.firefox.com/success.txt ftavelazqueza
ROUNDROBIN_PARENT/10.8.7.221 text/plain
```

Figura C.1. Entrada del **log** del **squid**

La estructura del **log** es:

**Hora – demora - host remoto – código/estado – bytes – método – URL - rfc931
-peerstatus/peerhost tipo**

El **script** propuesto se encuentra en [“scripts QoS”](#). Su forma de uso se explica a continuación:

Requerimientos:

- Instalar **python3** en el sistema.
- Es fundamental que los **logs** a procesar estén situados en el mismo directorio desde donde se ejecuta el **script**.

Forma de ejecución:

- Para usuarios Linux abrir una terminal en la dirección donde se encuentra el **script** y escribir en la terminal: **sudo python3 QoE.py**.
- Para usuarios Windows una vez instalado **Python** en el sistema, existen varias formas:
 1. Ejecutar **click** derecho sobre el **script** a ejecutar, y seleccionar “**edit with IDLE**”, intérprete de **Python** que se instala conjuntamente con **Python**, posteriormente presionar **F5** para ejecutar el **script**.
 2. Abrir una terminal como administrador y ubicarse en la dirección donde se encuentra el **script** y escribir en la terminal: **python QoE.py**.
 3. Simplemente ejecutar doble **click** sobre el **script**.

Pasos de uso:

1. Introducir los usuarios a evaluar dentro del **script**, en el sitio indicado.
2. Introducir el horario de inicio de la encuesta para filtrar en el archivo **log** en el formato **HH:MM**.

3. Introducir el horario final de la encuesta para filtrar en el archivo **log** en el formato **HH:MM**.

El programa habrá terminado una vez se observe **Done!!!**, en la pantalla. El archivo con la información extraída, estará ubicado en la misma ubicación de los scripts, el nombre de ese archivo será un **timestamp.txt**, donde **timestamp** es el tiempo en **unix time** correspondiente a la fecha de inicio de la encuesta.

En caso de que el archivo **log** no tenga el nombre por defecto, **access.log**, o no se encuentre ubicado en el mismo directorio que el **script**, es posible especificar el nuevo camino hacia el **log** de la siguiente manera:

Linux: **python3 QoE.py camino-al-archivo/archivo.extensión**

Windows, desde la terminal: **python QoE.py camino-al-archivo/archivo.extensión**

Referencias

- [1] Z. Mahmood, Ed., *Continued Rise of the Cloud: Advances and Trends in Cloud Computing*. London: Springer-Verlag, 2014.
- [2] «Application Performance Percentiles and Request Tracing for AWS Application Load Balancer | AWS Blog». <https://aws.amazon.com/es/blogs/aws/application-performance-percentiles-and-request-tracing-for-aws-application-load-balancer/> (accedido jun. 20, 2017).
- [3] «Amazon CloudWatch Update – Percentile Statistics and New Dashboard Widgets | AWS Blog». <https://aws.amazon.com/es/blogs/aws/amazon-cloudwatch-update-percentile-statistics-and-new-dashboard-widgets/> (accedido jun. 20, 2017).
- [4] «Calculating Percentiles | Elasticsearch: The Definitive Guide [2.x] | Elastic». <https://www.elastic.co/guide/en/elasticsearch/guide/current/percentiles.html> (accedido jun. 19, 2017).
- [5] «Percentiles Aggregation | Elasticsearch Reference [5.4] | Elastic». <https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations-metrics-percentile-aggregation.html> (accedido jun. 20, 2017).
- [6] «Estimating end-to-end performance in IP networks for data applications», ITU-T, Recommendation ITU-T G.1030 provides a framework of tools to obtain

IP network performance, estimate the performance of user applications, and apply perceptual models to gauge user satisfaction with the end-to-end performance. The user-perceived performance of data applications on packet networks is dependent on many factors, including the end-to-end performance of the packet network, the application's dependency on the communications network, the performance of the terminals and other devices beyond the purview of the network operator(s), and the user's task and the extent of user interaction with the application. Network designers take these factors into account to assure user satisfaction. Once the application performance has been estimated, perceptual models can be applied to interpret the level of end-to-end performance attained. This Recommendation assumes that the reader will be able to provide at least some level of detail about each of the key factors above, and will then use the framework of tools to estimate end-to-end performance., feb. 2014.

- [7] J. F. Kurose y K. W. Ross, *Computer networking : a top-down approach*, 7ma. United States of America: Pearson Education, Inc., 2017.
- [8] P. Oppenheimer, *Top-down network design*, 3rd ed. Indianapolis, IN: Cisco Press, 2011.