

## INFORME TÉCNICO PREVIO DE EVALUACIÓN DE SOFTWARE

N°00/-2017-SUNAT-1U4100

### 1. NOMBRE DEL AREA:

1U4200- División de Arquitectura Tecnológica

1U4100 - División de Arquitectura de Información y de Aplicaciones

### 2. RESPONSABLE DE LA EVALUACIÓN:

- Miguel Alejandro Neyra Horna - Jefe de División de Arquitectura Tecnológica
- Johnny Antonio Valdez Arevalo - Jefe de División de Arquitectura de Información y de Aplicaciones
- Paul Santamaria Romero - Profesional de la División de Arquitectura Tecnológica
- Nilton Sulca Recharte - Profesional de la División de Arquitectura de Información y de Aplicaciones

### 3. FECHA:

09 de Agosto de 2017

### 4. JUSTIFICACIÓN:

La Entidad se encuentra en un proceso de modernización tanto de servicios y plataforma tecnológica interna como externa para acelerar el cumplimiento de las funciones de su competencia y de los objetivos trazados.

El sistema referente a la Plataforma Única de Ingreso de Información (Proceso Adjudicación de Menor Cuantía N°0073-2015-SUNAT/8B1200) se ha construido basado en una arquitectura orientada a microservicios, cuyos componentes de software para almacén de estructura de datos cache, agente de mensajes y software para monitoreo de los componentes han sido explicados y sustentados en el documento de arquitectura aprobada con carta N° 1417-2016-SUNAT/8B1300 en coordinación con el equipo técnico de la INSI; los cuales permitirán el funcionamiento óptimo del sistema en la infraestructura de la institución.

El desarrollo del sistema se conceptualizó para dar asistencia al usuario que simplifique y facilite el registro y la presentación de información ante la Entidad, organizados en un único ambiente, mediante formularios electrónicos, pagos de tributos internos, programa de declaración telemática (PDT), entre otros, en cumplimiento con el Objetivo Estratégico N° 02 que permitirán reducir los costos al administrado en cumplimiento con las obligaciones tributarias y aduaneras.

Por lo expuesto y en cumplimiento de la Ley N° 28612 – “Ley que norma el uso, adquisición y adecuación del software en la administración pública” – se ha procedido a la evaluación de software ante la necesidad imprescindible de licenciarlo, para el logro de los objetivos institucionales y la modernización de la SUNAT.

### 5. ALTERNATIVAS:

Para el análisis se han tomado los siguientes componentes:



### 5.1 Software para almacén de estructura de datos cache y agente de mensajes

Los formularios del sistema de Plataforma Única de Ingreso de Información requieren de cierta información paramétrica que debe ser consultada de base de datos cada vez que un contribuyente ingresa a un formulario. Esta información paramétrica se mantiene constante con cambios muy esporádicos en el tiempo, por lo que se propone llevar esta información a un cache en memoria que permita reducir los accesos a base de datos, y ofrecer mejores tiempos de respuesta al consultarse esta información, apoyándose también a la escalabilidad de la plataforma.

Como parte de la solución de Plataforma Única de Información se ha evaluado los productos **HazelCast** y **REDIS Pack** siendo este último la herramienta utilizada de acuerdo a la arquitectura desarrollada, documentada y aprobada por SUNAT con carta N° 1417-2016-SUNAT/8B1300 como soporte al sistema de Plataforma Única de Ingreso de Información de Declaraciones Juradas y Pagos, mediante el proceso Adjudicación de Menor Cuantía N°0073-2015-SUNAT/8B1200, el cual tiene una amplia adopción en diversas soluciones e industrias a nivel mundial, demostrando un gran desempeño y escalabilidad.

### 5.2 Software para monitoreo de los componentes de la arquitectura

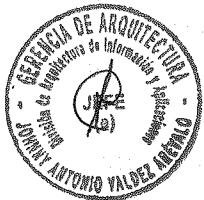
El software de monitoreo permitirá recolectar la información del consumo de los recursos de cada nodo que conforma la arquitectura de plataforma única, así como de las métricas de desempeño de los diversos microservicios que se ejecutan; permitiendo generar a partir de esta información tableros de monitoreo del desempeño de la plataforma y posibles problemas que se presenten, generándose alertas a los operadores de la plataforma cuando se alcancen determinadas condiciones.

Como parte de la solución de Plataforma Única de Información se ha evaluado los productos **Splunk** y **ELK / X-Pack (Elasticsearch, Logstash, Kibana, Beats y X-Pack)** siendo este último la herramienta utilizada de acuerdo a la arquitectura desarrollada, documentada y aprobada por SUNAT con carta N° 1417-2016-SUNAT/8B1300 como soporte al sistema de Plataforma Única de Ingreso de Información de Declaraciones Juradas y Pagos, mediante el proceso Adjudicación de Menor Cuantía N°0073-2015-SUNAT/8B1200.

## 6. ANÁLISIS COMPARATIVO TÉCNICO

Se precisa que para las evaluaciones se ha establecido un puntaje técnico mínimo que debe alcanzar cada producto evaluado para poder ser considerado como apto para su implementación, si el producto no supera dicho puntaje técnico no será aceptado por no contar con la funcionalidad básica requerida. La referencia al puntaje mínimo se basa en el siguiente cuadro:

Id	Descripción	Puntaje mínimo
1	Software para almacén de estructura de datos cache y agente de mensajes.	84
2	Software para monitoreo de los componentes de la arquitectura.	98



El presente análisis se realizó aplicando el ítem 3 de la Guía de Evaluación de Software para la Administración Pública. A continuación se detalla el análisis comparativo por el software base planteado para cada producto.

### 6.1 Propósito de la evaluación

Validar que las alternativas seleccionadas sean las más convenientes para el uso en la SUNAT.

### 6.2 Tipo de producto

- Software para almacén de estructura de datos cache y agente de mensajes.
- Software para monitoreo de los componentes de la arquitectura.

### 6.3 Modelo de calidad

De acuerdo a lo establecido en la Guía de Evaluación del Software para la Administración Pública aprobada mediante resolución ministerial N° 139-2004-PCM, se aplicara el modelo de calidad de software establecido en el mismo.

### 6.4 Selección de Métricas

Las métricas fueron seleccionadas en base al análisis de las características técnicas del software seleccionado en el punto 5 (alternativas de software), a través de Internet y el documento de arquitectura de la Plataforma Única de Ingreso de Información de Declaraciones Juradas y Pagos (proceso Adjudicación de Menor Cuantía N°0073-2015-SUNAT/8B1200).

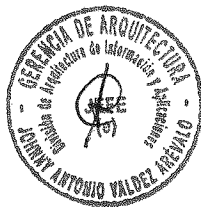
#### 6.4.1 Software para almacén de estructura de datos cache y agente de mensajes

Las características y la métrica de evaluación del Software está basado en el siguiente cuadro:

Ítem	Características	Puntaje Máximo	REDIS Pack	HazelCast
1	Debe ser un motor de base de datos NoSQL en memoria, basado en el almacenamiento en tablas de hashes (clave/valor) y pueda ser usada como una base de datos durable o persistente	10	10	10
2	Debe permitir realizar el caché de páginas HTML, o fragmentos de estas, acelerando el acceso a las mismas, evitando llegar a los servidores web o de aplicaciones.	10	10	10
3	Debe permitir el almacén de sesiones de muy rápido acceso, en el que se mantenga un identificador de sesión junto con toda la información asociada a la misma.	10	10	8
4	Reducir los tiempos de latencia, evitando accesos a otras bases de datos innecesarios.	10	10	8



Ítem	Características	Puntaje Máximo	REDIS Pack	HazelCast
5	Permitir asociar un tiempo de expiración a cada clave, con lo que las sesiones deben finalizar automáticamente sin tener que gestionarlo en el código de la aplicación.	10	10	10
6	De forma muy similar al almacén de sesiones de usuario, permita almacenar de forma dinámica en el software la información que se va adquiriendo en las diferentes páginas web y centralizarlas para realizar operaciones posteriores (carrito de compra).	10	10	10
7	Permita almacenar en el software el resultado de determinadas consultas que se ejecuten con mucha frecuencia a nivel de base de datos, y cuya información no cambia a menudo, o no es crítico mantener actualizada al instante.	10	9	9
8	Permita manejar contadores y estadísticas en tiempo real, y cuente con soporte para la gestión concurrente y atómica de los mismos. Algunos ejemplos posibles serían el contador de visualización de un producto, votos de usuarios, o contadores de acceso a un recurso para poder limitar su uso.	10	9	9
9	Permita acceder a las listas de elementos recientes que aparecen en las últimas actualizaciones de algún elemento hechas por los usuarios. Por ejemplo, los últimos comentarios sobre un post, las últimas imágenes subidas, los artículos de un catálogo vistos recientemente, etc. El software debe ser capaz de evitar el uso de las bases de datos relacionales, sobre todo cuando el volumen de información se va haciendo mayor, siendo capaz de resolver esta operación con independencia del volumen.	10	9	9
10	Permita usarse como almacenamiento principal, realizar replicaciones, clustering de su base datos, replicación asíncrona, no bloqueante, Failover automático, entre otros.	10	10	5
11	Permitir el uso en aplicaciones basadas en microservicios, en los que se pueda aprovechar su velocidad para construir soluciones especializadas, simples de implementar y mantener, que a la vez ofrezcan un alto rendimiento.	10	10	7



Ítem	Características	Puntaje Máximo	REDIS Pack	HazelCast
12	Debe funcionar en la mayoría de los sistemas operativos POSIX como Linux, BSD, OS X, entre otros.	10	10	10

El software se utilizará como cache de datos para mantener en memoria la información de parametría de los formularios. Si el cache no está disponible, o la información no se encuentra en cache, los microservicios consultarán la información directamente en la base de datos.

El cache se utiliza también para mantener información asociada al contribuyente que ha iniciado sesión para interactuar con la plataforma. Esta información se utilizará para verificar la validez de cada invocación que realizan las aplicaciones cliente a los microservicios. Dado que las peticiones de las aplicaciones cliente pueden ser atendidas en ambos sites indistintamente, es necesario que esta información se encuentre disponible en el cache de ambos SITES, para lo cual se realizará una replicación de datos del cache entre ambos SITES.

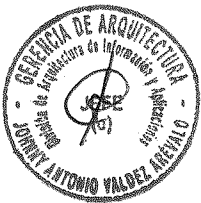
También se hará uso del software como cache al realizarse el UPLOAD de archivos grandes de PDT, facilitando que la escritura de los archivos se pueda segmentar en distintos llamados, sin importar la instancia del microservicio que atiende cada llamado, usando el cache para mantener los distintos segmentos del archivo y posteriormente escribir el archivo completo en FileSystem una vez que se envía el último segmento del archivo.

Dada la relevancia de la información que se mantendrá en cache, se ha apreció que REDIS Pack funciona mejor en un esquema de cluster para contar con un esquema de alta disponibilidad, permitiendo que la información sobreviva aún ante caídas de algún nodo del cluster de REDIS.

#### 6.4.2 Software para monitoreo de los componentes de la arquitectura

Las características y la métrica de evaluación del Software está basado en el siguiente cuadro:

Ítem	Características	Puntaje Máximo	(ELK / X-Pack)	Splunk
1	Deberá contar con un almacén de estructura de datos, utilizado como base de datos, caché y agente de mensajes	10	10	9
2	Los componentes deben realizar las siguientes acciones: Procesar los logs que llegan, Almacenar todos los logs y visualizar mediante Interfaz Web para buscar y visualizar los logs	10	10	10
3	El software debe ser escalable hasta petabytes de datos estructurados y no estructurados.	10	10	9



4	Pueda ser utilizado como un sustituto de almacenamiento de documentos como MongoDB y RavenDB	10	10	7
5	Debe utilizar des-normalización de datos para mejorar el rendimiento de la búsqueda	10	10	9
6	Su respuesta de búsqueda debe ser en tiempo real, es decir, después de un segundo de añadido algún documento o trabajado en otra aplicación estos puedan ser buscados mediante el software de monitoreo de componentes	10	10	10
7	Deberá recibir los logs y hacer un análisis sintáctico de la estructura del log, para ello se deberá utilizar las siguientes configuraciones: INPUT (PLUGINS) datos y que estructura de datos queremos procesar, FILTER (PLUGINS) debe utilizar una expresión regular con un patrón (grok) para dar forma o analizar el dato, OUTPUT (PLUGINS) se remite en formato JSON	10	10	9
8	Permitir recabar la información de uso y salud de los diversos nodos de Plataforma Única y generar a partir de esta información tableros (dashboards) para el monitoreo del sistema en cuanto a cantidad total de presentaciones, cantidad de presentaciones por minuto y operatividad del sistema	10	10	9
9	Permitir alertas y visualización de la información mediante histogramas, gráficos de líneas, gráficos circulares, entre otros	10	10	10
10	Permitir administrar o configurar las funciones de seguridad mediante APIs para navegar de forma intuitiva, evitando la dependencia de la línea de comandos	10	9	9
11	Soporte para extraer eventos de fuentes comunes, todo al mismo tiempo. Permita ingresar fácilmente a sus registros, métricas, aplicaciones web, almacenes de datos y varios servicios, todos de forma continua y en streaming	10	9	7
12	Como parte de la Seguridad de la herramienta esta debe incluir: autenticación nativa, control de accesos basado en perfiles y roles, así como la sincronización e integración con el directorio activo	10	10	10
13	Permita detectar una caída o comportamiento inusual como parte de las operaciones de aplicación y registrar las acciones comunes que permita un aprendizaje continuo	10	10	5



14	Permita especificar de forma gráfica los datos que le interesan analizar (solicitudes por segundo) y qué otras propiedades pueden influir en ella (servidor, IP, nombre de usuario), especificar umbrales o construir manualmente modelos estadísticos	10	10	5
----	--	----	----	---

La información que se registrará en la herramienta de monitorio de acuerdo al proyecto de Plataforma Única de Información es la siguiente:

- Información sobre uso de CPU, Memoria, Procesos, Espacio en disco obtenida mediante MetricBeats. Se obtendrá la información de cada uno de los nodos de plataforma única.
- Información generada por metrics en los micro servicios, misma que contiene el detalle del desempeño de cada instancia de micro servicio como lo es el número de peticiones atendidas, errores presentados, uso de recursos de la JVM, así como su salud.
- Logs generados por las distintas instancias de los micro servicios.
- Logs de operación generados por el software base de infraestructura de NGINX, IBM MQ y RLEC.
- Información de negocio sobre cada presentación procesada correctamente, así como incidencias presentadas al realizarse el pago por alguno de los medios disponibles.

Basado en la información que se desea gestionar, la decisión de utilizar el **(ELK / X-Pack)** está motivada porque el esquema de escalamiento que provee esta herramienta contribuye de mejor manera a asegurar la alta disponibilidad que requiere la arquitectura planteada en el proyecto.

### 6.5 Niveles, escala para las métricas y comparación de los criterios

Se muestran las escalas consideradas para cada una de las métricas señaladas:

La siguiente tabla muestra un resumen de los puntajes obtenidos por cada software considerando el Análisis Técnico Comparativo:

#### 6.5.1 Software para almacén de estructura de datos cache y agente de mensajes:

REDIS Pack	HazelCast
117 (Puntos)	105 (Puntos)

A nivel técnico REDIS Pack permite mantener mejor la información en memoria de uso frecuente (no variable), minimizando accesos de lectura a base de datos o disco, apoyando a la escalabilidad y desempeño de la plataforma. Asimismo permitirá mantener en memoria la información asociada a un usuario que ha iniciado sesión en la plataforma, facilitando que las peticiones provenientes de un



mismo usuario puedan ser atendidas por cualquier instancia de un microservicio, independientemente del SITE al que sea enviada la petición.

El uso del cache permite que las peticiones de upload de archivos puedan ser segmentadas en pequeños fragmentos y direccionadas indistintamente a cualquiera de los SITES.

### 6.5.2 Software para monitoreo de los componentes de la arquitectura

(ELK / X-Pack)	Splunk
138 (Puntos)	118 (Puntos)

A nivel técnico (ELK / X-Pack) permitirá indexar mejor cualquier información que se le entregue, facilitando búsquedas sobre la misma, podrá escalar muy fácilmente agregándose nuevos nodos de elasticsearch al cluster.

La recolección de la información de monitoreo y su envío al ElasticSearch server se ejecutará de manera asíncrona respecto al resto de los procesos que se ejecutan de la solución, evitando generar algún tipo de contención en estos elementos por el procesamiento de los datos de monitoreo.

## 7. ANÁLISIS COMPARATIVO COSTO - BENEFICIO

A continuación se ha realizado un análisis de costos referencial (en dólares americanos, no incluye IGV, de los productos y complementos necesarios para cumplir con la evaluación técnica realizada.

### 7.1 Software para almacén de estructura de datos cache y agente de mensajes

REDIS	HazelCast
(*) \$9,000.00 (01 shard)	(*) \$7,000.00 (Por JVM)
Total requerido para el proyecto de acuerdo a su equivalencia	
\$36,000.00 (04 shard)	\$42,000.00 (06 JVM)

(\*) Precio de lista obtenida del fabricante, que puede variar de acuerdo a los costos donde se adquiere el producto y los distribuidores autorizados por cada país.

### 7.2 Software para monitoreo de los componentes de la arquitectura

(ELK / X-Pack)	Splunk
(*) \$6,660.00 (Por nodo)	(*) \$600.00 (100 GB/Day)
Total requerido para el proyecto de acuerdo a su equivalencia	
\$33,300.00 (05 nodos)	Bajo este esquema no se pudo proyectar el costo

(\*) Precio de lista obtenida del fabricante, que puede variar de acuerdo a los costos donde se adquiere el producto y los distribuidores autorizados por cada país.






## 8. CONCLUSIONES

- a) El software Redis Pack para almacén de estructura de datos cache y agente de mensajes se adapta a las necesidades planteadas en el documento de arquitectura del sistema de Plataforma Única de Ingreso de Declaraciones Juradas y Pagos (Adjudicación de Menor Cuantía N°0073-2015-SUNAT/8B1200).
- b) El Software ELK / X-Pack para monitoreo de los componentes de la arquitectura se adapta a las necesidades planteadas en el documento de arquitectura del sistema de Plataforma Única de Ingreso de Declaraciones Juradas y Pagos (Adjudicación de Menor Cuantía N°0073-2015-SUNAT/8B1200).



.....  
MIGUEL ALEJANDRO NEYRA HORNA  
Jefe Div. de Arquitectura Tecnológica (e)  
GERENCIA DE ARQUITECTURA  
Miguel Alejandro Neyra Horna

Reg: 1737



.....  
JOHNNY ANTONIO VALDEZ AREVALO  
Jefe Div. de Arquitectura de Información  
y Aplicaciones (a)  
GERENCIA DE ARQUITECTURA  
Johnny Antonio Valdez Arevalo

Reg: 1245



Paul Santamaria Romero

Reg:AG24



Nilton Sullca Recharte

Reg: NJ69