Eusko Jaurlaritzaren
Informatika Elkartea

Sociedad Informática
del Gobierno Vasco

EXPEDIENTE Nº: EJIE-019-2022

Servicios para el desarrollo del Sistema de Repositorio y Gestión de Series Temporales para la Agencia Vasca del Agua BigData.

Pliego de condiciones técnicas

Febrero 2022

Mediterráneo, 14 • 01010 Vitoria-Gasteiz • Posta-kutxatila / Apartado: 809 • 01080 Vitoria-Gasteiz • Tel. 945 01 73 00* • Fax. 945 01 73 01 • www.ejie.eus

Este documento es propiedad de Eusko Jaurlaritzaren Informatika Elkartea – Sociedad Informática del Gobierno Vasco, S.A. (EJIE) y su contenido es CONFIDENCIAL. Este documento no puede ser reproducido, en su totalidad o parcialmente, ni mostrado a otros, ni utilizado para otros propósitos que los que han originado su entrega, sin el previo permiso escrito de EJIE. En el caso de ser entregado en virtud de un contrato, su utilización estará limitada a lo expresamente autorizado en dicho contrato. EJIE no podrá ser considerada responsable de eventuales errores u omisiones en la edición del documento.

Versión	Fecha	Resumen de cambios	Elaborado por:	Aprobado por:
1	12/01/22		Marina Baratta	Gurutze Marijuan

Contenido

1 Introducción 5

2 Objeto del contrato 6

2.1 Objeto y alcance del servicio 6
2.2 Descripción del servicio 6

3 Requisitos del servicio 7

3.1 Situación del problema 7
3.1.1 Estado Inicial 7
3.1.2 Estado final 7
3.1.3 Estrategia de desarrollo 8

4 Metodología aplicable y Entorno Tecnológico 11

4.1 Metodología de desarrollo, Normativa y Guía de Estilo 11
4.2 Herramientas del ciclo de vida de las aplicaciones 13
4.3 Entorno Tecnológico 14
4.3.1 BigData en Batera 14
4.3.1.1 Servicios 14
4.3.1.2 Seguridad 14
4.3.2 Catálogo de servicios 15
4.3.2.1 File Service 15
4.3.2.2 Table Service 15
4.3.2.3 Event Service 15
4.3.2.4 Search Service 15
4.3.3 Versiones 16
4.4 Contenedores 17

5 Descripción de entregables 18

5.1 Normativa para la puesta en producción y posterior explotación del sistema 18

6 Planificación y organización 20

6.1 Plan de proyecto 20
6.1.1 Plan de Entregas e Hitos de Control 20
6.1.2 Diagrama de Gantt 20
6.2 Organización del Equipo de Trabajo 20
6.2.1 Asignación de recursos a fases del proyecto 20
6.2.2 Equipo de Trabajo 21
6.2.2.1 Constitución inicial del equipo de trabajo 21
6.2.2.2 Modificaciones en la composición del equipo de trabajo 21

6.2.2.3 Horario y lugar de realización de los trabajos. *22*

6.3 Transferencia tecnológica 22

7 Mecanismos de Seguimiento, Control y Supervisión 23

7.1 Mecanismos de Seguimiento y Control. 23

7.2 Supervisión del proyecto 23

8 ANEXO 1 - Series Temporales Inundaciones 24

9 ANEXO 2 – Resumen Consultoría área de Vertidos y Aprovechamientos.....25

1 Introducción

EJIE, Eusko Jaurlaritzaren Informatika Elkartea – Sociedad Informática del Gobierno Vasco, es la Empresa pública de servicios de las tecnologías de la información y las comunicaciones (TIC), cuya razón de existir es contribuir a la consecución de un Sector Público Vasco, moderno y eficiente, en el Marco Legal establecido por el Gobierno, con la seguridad y calidad necesarias y con el debido respeto al medio ambiente.

EJIE tiene como meta final la consecución de la satisfacción de sus clientes, siendo el instrumento común de prestación de servicios TIC en el Sector Público Vasco, y comprometiéndose en:

- Construir y mantener con eficiencia y calidad la infraestructura de los Sistemas de Información, posibilitando su continuidad y seguridad.
- Garantizar la interoperabilidad entre las distintas administraciones.
- Servir de apoyo a las necesidades de planificación y realización de la función informática de los Departamentos y Organismos Autónomos del Gobierno, asegurando la cobertura de sus demandas con el compromiso y profesionalidad adecuados a las relaciones contractuales que se establezcan.

Por tanto, EJIE debe ser un instrumento común de referencia para la prestación de servicios TIC en el Sector Público Vasco:

- Aportando valor añadido.
- Proporcionando soluciones competitivas.
- Transmitiendo confianza a sus clientes.
- Contando con personas cualificadas y comprometidas.

Se puede obtener información más detallada y extensa en nuestra dirección de Internet http://www.ejie.eus

2 Objeto del contrato

2.1 Objeto y alcance del servicio

Ejie, empresa pública del Gobierno Vasco, impulsada por el objetivo primordial de lograr una Administración Pública Vasca moderna y eficiente, tiene como labor principal la prestación de servicios informáticos de diversa índole.

La Comunidad Autónoma del País Vasco cuenta con una amplia red de puntos de observación hidrometeorológica y de calidad del agua en tiempo real. Esta red cuenta con una trayectoria prolongada en el tiempo y, además, se sigue trabajando en su mejora y ampliación. Por ello, se dispone de una ingente cantidad de series de datos temporales.

Las diferentes áreas de la Agencia Vasca del Agua – Uraren Euskal Agentzia (en adelante URA) utilizan esta información temporal para realizar diversos trabajos y proyectos, y se prevé que en el futuro la disponibilidad de estos datos para toda la organización será cada vez más importante. Es por ello que resulta necesario disponer de repositorios de series temporales para almacenar la información y trabajar con ella de forma ágil y eficiente, y, además, que sea abierta para poder ser interoperable con los distintos sistemas IT de URA.

A través de este pliego se pretende establecer las condiciones que han de regir los trabajos de la creación y gestión de los repositorios de series temporales para la Agencia Vasca del Agua-Uraren Euskal Agencia, y establecer los criterios y la metodología para su desarrollo.

Por consiguiente, el objeto del presente Pliego de Condiciones Técnicas define las condiciones de la contratación de los servicios de análisis, diseño, construcción, pruebas, implantación, para el sistema de Repositorio y Gestión de Series Temporales de la Agencia Vasca del Agua.

2.2 Descripción del servicio

Se contrata los servicios de análisis, diseño, construcción, pruebas e implantación del Sistema de Repositorio y Gestión de Series Temporales para la Agencia Vasca del Agua en BigData, en las siguientes tres líneas de trabajo:

1. Evolución del actual sistema del área de Inundaciones
2. Nuevo sistema para el área de Vertidos
3. Nuevo sistema para el área de Aprovechamientos

La descripción en detalle de los servicios a prestar se describe en el apartado Requisitos del Servicio del presente Pliego de Condiciones Técnicas.

3 Requisitos del servicio

3.1 Situación del problema

3.1.1 Estado Inicial

Actualmente en Gobierno Vasco se dispone de una plataforma de BigData con las últimas versiones de Hadoop y HDFS conocida como BDaaS (BigData as a Service), cuyas características se detallan en el apartado “Entorno Tecnológico”.

Durante el año 2021 la Agencia Vasca del Agua ha realizado un proyecto de creación del sistema de repositorio y gestión de series temporales en las infraestructuras BigData del Gobierno Vasco. En este primer proyecto se ha construido una iniciativa BigData para el área de Inundaciones de la Dirección de Planificación y Obras, que ya disponía de una aplicación para almacenar series temporales desde 2012, pero se trataba de una herramienta propietaria y cerrada. Disponen ahora de una iniciativa BigData para el almacenamiento, transformación y acceso y uso para modelización de series temporales y los metadatos asociados a estas series temporales, en los volúmenes históricos y los que previsiblemente se generarán en las próximas décadas. Sobre esta iniciativa se desea seguir evolucionando para dotarla de nuevas funcionalidades. Las funcionalidades actuales se detallan en el Anexo 1 de este pliego.

Por otra parte, desde la Dirección de Gestión del Dominio Público Hidráulico, se detectó también la necesidad de gestionar datos en continuo, tanto medidas del caudal de aprovechamiento, como distintas medidas recogidas en las distintas plantas de depuración, tratamiento, u otras instalaciones. Es por este motivo que realizaron un proyecto de consultoría, para estudio de la transmisión en continuo del caudal de los aprovechamientos y la telemetría de los datos de calidad del efluente de EDARs tanto de naturaleza urbana como industrial. Además, en esa consultoría obtuvieron unos requisitos generales para el sistema de repositorio y gestión de series temporales que se resumen en el Anexo 2 de este pliego.

3.1.2 Estado final

Al finalizar la ejecución de los servicios objeto de esta contratación, la Agencia Vasca del Agua deberá contar en la infraestructura del Gobierno Vasco BigData con tres iniciativas de sistemas de repositorio, gestión y procesamiento de series temporales:

Sistema de Inundaciones, se dará continuidad a los trabajos ya realizados, y se evolucionará el actual sistema con las nuevas funcionalidades detalladas en el apartado Estrategia de Desarrollo para esta área.

Iniciativas de Aprovechamientos y Vertidos:

Además del área de Inundaciones, durante los últimos años en otras áreas de la Agencia Vasca del Agua, de Vertidos y Aprovechamientos, ha surgido también la necesidad de almacenar series temporales, y requieren sus propias iniciativas para el tratamiento y gestión de datos en continuo, que les deberán llegar de distintas fuentes.

Se trata de nuevos sistemas que permitan la recogida, almacenamiento y explotación de las series de datos temporales tanto del caudal de aprovechamientos, como de la telemetría de distintos parámetros de calidad, caudales, y otros datos generados en las Estaciones Depuradoras de Aguas Residuales (EDAR), tanto de naturaleza urbana como industrial.

Al finalizar el proyecto se contarán con dos nuevos sistemas de Repositorio y Gestión de Series Temporales para la Agencia Vasca del Agua en BigData para la gestión de datos en continuo:

1. Nuevo sistema para el área de Vertidos
2. Nuevo sistema para el área de Aprovechamientos

3.1.3 Estrategia de desarrollo

Se detallan a continuación los requisitos básicos que debe cumplir incluidos en el ámbito de la presente contratación.

Son de aplicación, total o parcialmente, al menos, todas las normativas reguladoras mencionadas anteriormente en este documento.

Todos los requisitos aquí enumerados deben considerase en cualquier caso como una primera aproximación, los cuales deberán ser refinados convenientemente durante la realización de los trabajos objeto del presente pliego de condiciones técnicas. Del mismo modo, es muy probable que éstos se vean ampliados a medida que se conozcan más en detalle las funcionalidades a implementar en el nuevo sistema.

Para la realización de los trabajos se deberá partir y reutilizar en la medida de lo posible los trabajos ya realizados durante 2021 para la gestión de series temporales del área de Inundaciones de la Agencia Vasca del Agua.

REQUISITOS DEL SERVICIO:

FASE DE INUNDACIONES:

1. Desarrollo de funciones de transformación expuestas en el API para generar formatos de salida específicos válidos para las distintas herramientas que usan en URA/Inundaciones. Los formatos que habría que implementar son (Tetis, Piformat, SimpleCSV, ColCSV)
2. Desarrollo de Funciones de transformación expuestas en el API para generar nuevas series calculadas. Función media ponderada, potencias, sum, resta etc... Con el fin de preparar las series temporales para su uso en el modelado se requiere que dichas series cumplan con la completitud de datos necesarios, por lo que será necesario aplicar interpolaciones en partes de ellas, medias, ponderaciones etc.
3. Análisis, construcción e implantación de aplicación de usuario para visualizar las series y gestionar las mismas. Altas/Bajas/Modificaciones de las entidades de negocio existentes usando el API ya desarrollado, además de dotarle de funcionalidades de visualización y de explotación y exportación que se concretarán en el análisis.
4. Integrar el API de series temporales con series calculadas históricas y predichas con los modelos de inundaciones para la automatización de la generación de informes
5. Nuevo modelo automático de alarmas basado en el forecast del histórico de las series temporales ingestadas en la plataforma.
6. Estudiar formatos raster

FASE DE VERTIDOS y FASE DE APROVECHAMIENTOS:

Se deberá desarrollar la capa de ingesta, almacenamiento, procesamiento y explotación de las series temporales de los distintos dispositivos de captura, para obtener como entregables de esta parte del proyecto los dos nuevos sistemas de Repositorio y Gestión de Series Temporales para la Agencia Vasca del Agua en la infraestructura BigData del Gobierno, para la gestión de datos en continuo del área de Vertidos y otro para la gestión de datos en continuo del área de Aprovechamientos.

El proyecto partirá de la primera consultoría ya realizada por URA en la que se ha definido que los distintos agentes que capturan esta información dejarán con una frecuencia predeterminada las variables a almacenar en un sistema de colas MQTT.

Para cada uno de los dos nuevos sistemas se realizarán los siguientes trabajos

1. **Fase de análisis**

   En esta fase se analizará y definirá teniendo en cuenta el desarrollo ya existente para el área de Inundaciones, todo lo necesario para la ejecución a realizar en fases posteriores, al menos los siguientes

   - Las fuentes de datos
   - Las entidades de negocio
   - Las particularidades de las series temporales, y su tratamiento
   - Definición las alarmas técnicas y de negocio a aplicar
   - Al menos 5 sistemas que se modelarán de Vertidos
   - Al menos 5 sistemas que se modelarán de Aprovechamientos
   - Los reports que tienen cabida en este proyecto
   - La aplicación básica piloto de funcionalidades
   - Los formatos en detalle de los mensajes enviados a las colas MQTT, y la arquitectura de comunicaciones Gateway- colas MQTT-plataforma BDaaS
   - Modelo de comunicación MQTT→ NiFi

2. **Implementación del modelo de ingesta y tratamiento de datos**

   En esta fase se analizará, desarrollará e implantará los procesos de importación de los diferentes tipos de datos

   - Ingesta de las variables de los sistemas de vertidos
   - Ingesta de las variables de los sistemas de aprovechamientos

   Durante esta fase se configurará el sistema para importar datos en tiempo real y se importarán los datos históricos. Todos los procesos de importación al almacén de datos se documentarán mediante logs
   Además, se realizará el tratamiento de los datos (validación y transformación) que se requiera en el análisis.

3. **Implementación de API**

   Desarrollo del API que permita el acceso a los datos o la modificación de los datos, de forma transparente, con el objetivo de que todo cambio en la estructura del almacén no suponga ningún cambio para la integración con otras aplicaciones que vayan a trabajar sobre este sistema de repositorio de datos. El consumo de ese API debería ser posible desde cualquier lenguaje de programación. Las funcionalidades que expondrá el API serán por una parte para la obtención de los datos de las series temporales, y por otra parte para posibilitar la integración con los procesos de negocio.

4. **Monitorización y alertas**

   En esta fase se analizará, desarrollará e implantará el modelo de calidad del dato y de alarmas a varios niveles

   - Reglas
   - KIPs de calidad
   - Alarmas técnicas
   - Alarmas funcionales

   Con este módulo de monitorización y alertas, se podrá informar a diferentes agentes según eventos que puedan generarse en el sistema. Por ejemplo, en el caso de interrupción del flujo de datos, de superación de umbrales o de superación de pendientes críticas en la evolución de una variable.

5. **Explotación, Reporting y aplicación básica**

En esta fase se analizará, desarrollará e implantarán los informes y reportes acordados en el diseño y una pequeña aplicación básica piloto de funcionalidades que sirva para ilustrar el uso del API en una aplicación web.
Como resultado de esta fase de los trabajos se deberá poder realizar procesos sobre los datos de forma periódica-programada o a demanda por los usuarios del sistema.

En la explotación habrá que tener en cuenta la posibilidad de interconexión entre áreas a la hora de explotar los datos, que se posibilite el acceso a la información contenida en cualquiera de los tres repositorios pudiendo ser relacionada y comparada los datos contenidos en uno de los repositorios con los de otros, mediante tablas, gráficos, … independientemente al área de URA al que pertenezca el usuario que realice dicha explotación

6. Soporte y trasferencia de conocimientos

Esta última fase previa al cierre del proyecto incluye las tareas de entrega del sistema a los distintos actores de URA y de EJIE que se encargarán de su continuidad, como son las siguientes;

Definición del procedimiento organizativo de uso y gobierno del sistema

Formación a usuarios de URA y técnicos de EJIE y documentación

Soporte post-implantación y tutorización para el arranque tras la puesta en servicio.

Los productos a entregar serán aquellos que define la metodología de desarrollo Arinbide en referencia a las fases definidas dentro del alcance del proyecto: análisis, diseño, construcción e implantación del sistema.

Una vez completado este proyecto, los productos obtenidos serán los entregables, tanto software como documentales, resultantes de la ejecución de las anteriores líneas de trabajo.

4 Metodología aplicable y Entorno Tecnológico

La organización del trabajo y la prestación del servicio estará basada en ARINbide. No obstante, el nivel de cumplimiento de la misma dentro de los servicios a prestar deberá concretarse en las fases iniciales del proyecto. Aun así, las ofertas presentadas deberán tomar como referencia dicha metodología, que se explica brevemente a continuación.

4.1 Metodología de desarrollo, Normativa y Guía de Estilo

ARINbide es la metodología para el desarrollo y mantenimiento de sistemas de información promovida y utilizada por EJIE S.A. en el ámbito de su prestación de servicios al Gobierno Vasco. ARINbide procura sistematizar y estandarizar las actividades del ciclo de vida de los proyectos software, así como establecer los elementos que componen los productos finales a obtener (código fuente, ejecutables, documentación, etc.)

Para la presente contratación se opta por la opción de **ARINbide-Predictivo**.

ARINbide-Predictivo se concibe como una metodología práctica para el ciclo de vida completo del software, basada en Métrica 3, y adaptada a las necesidades y directrices de EJIE Además consta de un apartado para el establecimiento de una metodología de Gestión de Proyectos. Como Plan de Calidad la propia metodología, en sus apartados de trabajo habitual, genera los registros de calidad necesarios para el sistema de calidad de EJIE.

Para todo el ciclo de vida del proyecto, ARINbide-Predictivo define los siguientes procesos:

- Ingeniería del Software (ISW)
- Gestión del Proyecto (GPR)
- Gestión de Riesgos del Proyecto (GRP)
- Gestión de la configuración (GCO)
- Mantenimiento del Sistema de Información (MSI)

El proceso principal de ingeniería del software (ISW) describe y normaliza la secuencia de fases y actividades a realizar en el proyecto de desarrollo de un sistema de información, así como los entregables a obtener en cada una de ellas. En este ámbito, ARINbide-Predictivo contempla el desarrollo de las siguientes fases metodológicas:

- Catálogo de Requisitos de Usuario (CRU) (proceso diferenciado dentro del Análisis del Sistema de Información, ASI)
- Análisis del sistema (ASI)
- Diseño del sistema (DSI)
- Construcción del sistema (CSI)
- Implantación y aceptación del sistema (IAS)

El ciclo de vida descrito en la metodología de desarrollo ARINbide-Predictivo se alinea y complementa a lo largo del tiempo con la metodología de pruebas corporativa PROBAMET, centrándose ésta última en la especificación de todas las actividades relacionadas con la planificación, definición y ejecución de los diferentes tipos de pruebas. Existe una correspondencia entre las fases de ejecución de ARINbide-Predictivo y las fases de ejecución de PROBAMET.

En la sección "Perfil de contratante>Información técnica>Metodología de desarrollo" de la página web de EJIE se puede obtener información detallada sobre ARINbide-Predictivo, PROBAMET y el modelo SQA de EJIE.

En las distintas fases del proyecto y, teniendo en cuenta las mencionadas metodologías la empresa adjudicataria deberá basarse, entre otros, en los siguientes elementos:

- Catálogo de Requisitos de Usuario detallado en el documento CRU. Lo realizará en su totalidad el adjudicatario y en sus dependencias, que incluye:
  - ✓ La descripción del sistema
  - ✓ Los requisitos de usuario
  - ✓ Los casos de uso
  - ✓ La definición, análisis y validación del catálogo de requisitos

- Análisis del Sistema de Información, ASI, recogido en documento ERS. Lo realizará en su totalidad el adjudicatario y en sus dependencias, que incluye:
  - ✓ La especificación de requisitos del sistema
  - ✓ La relación de los módulos del sistema.
  - ✓ Las clases del sistema
  - ✓ El prototipo de interfaz
  - ✓ Especificación del Plan de Pruebas, según las directrices establecidas en PROBAMET
- Diseño del sistema. Lo realizará en su totalidad el adjudicatario y en sus dependencias. Debe incluir:
  - ✓ Diseño de las aplicaciones, DSI. (EDS)
  - ✓ Especificar el detalle del Plan de Pruebas del sistema por cada uno de los niveles de prueba: unitarias, de integración, de sistema, de implantación y de aceptación, según las directrices establecidas en PROBAMET
- Construcción del sistema.
  - ✓ Lo realizará en su totalidad el adjudicatario en sus dependencias, teniendo en cuenta las directrices de EJIE en lo que se refiere a los módulos estándar, las normas de funcionamiento y albergue de las aplicaciones en Internet/Intranet, el Manual de Identidad Corporativa del Gobierno Vasco y las directrices de la Dirección de Atención a la Ciudadanía en lo que respecta a estética, diseño y funcionalidades de las páginas Web.
  - ✓ Incluirá la ejecución del Plan de Pruebas, para verificar el cumplimiento de los requisitos establecidos en el mismo, abarcando pruebas unitarias, de integración y del sistema, siguiendo el modelo y las actividades descritas en la metodología PROBAMET
  - ✓ Los entregables de esta fase incluyen la Especificación de Construcción del Sistema (ECS) y el sistema construido.
  - ✓ Asimismo, deberán respetarse las convenciones adoptadas para un desarrollo de aplicaciones homogéneas recogidas en el Libro de Estilo de la Agencia Vasca del Agua
  - ✓ El diseño del sistema tendrá en cuenta el nivel de accesibilidad AA (WAI-AA), para aquellos módulos que sean accesibles desde Internet.
  - ✓ Todos los sistemas desarrollados deberán soportar al menos los siguientes idiomas: euskera, castellano.
  - ✓ La aplicación, las páginas, los textos, los mensajes de error/aviso y/o cualquier otro componente (tool-tip, textos en imagen) dirigidos al usuario final deberán estar al menos en ambos idiomas: euskera y castellano.
- Implantación en desarrollo. La empresa adjudicataria realizará los siguientes puntos
  - ✓ La entrega de la aplicación en el sistema de corporativo de control de versiones Subversion, tanto de la entrega inicial como de las sucesivas
  - ✓ Llevará a cabo la instalación de la aplicación en el entorno de desarrollo de EJIE, habiendo realizado previamente las pruebas necesarias durante la fase de construcción
  - ✓ Ejecución de las Pruebas del sistema, acorde con lo establecido en la metodología PROBAMET y el Modelo SQA, evaluando los resultados
- Implantación en entorno de test o preexplotación. La empresa adjudicataria efectuará los siguientes puntos
  - ✓ Preparará los paquetes y dará soporte a la realización por EJIE de la correspondiente implantación en el entorno de test o preexplotación
  - ✓ Ejecución de las Pruebas del sistema, y en el caso de que se requiera, estas pruebas se realizarán en las dependencias de E.J.I.E en sus distintos niveles de prueba, incluidas las Pruebas de Aceptación, acorde con lo establecido en la metodología PROBAMET y el Modelo SQA, evaluando los resultados
- Pruebas de prestaciones:
  - ✓ Conforme a lo establecido en el Modelo SQA, realización de pruebas de rendimiento y carga, que incluyen la verificación de los niveles de respuesta de la aplicación ante las previsiones de

carga del sistema, así mismo se verificará el comportamiento global del sistema en cuanto a consumo de memoria y CPU de sus componentes, así como según se considere de capacidad, stress, escalabilidad

- Implantación en producción y puesta en marcha del sistema. Incluirá:
  - ✓ Fuentes de la aplicación.
  - ✓ BBDD: scripts y carga inicial de datos.
  - ✓ Sistema implantado en entorno pruebas de EJIE
  - ✓ Informe de los resultados de los distintos niveles de pruebas, acorde con el Modelo SQA.
  - ✓ Manuales requeridos para la prestación del servicio en producción y definidos por los procesos ITIL de EJIE.
  - ✓ Manuales de Usuario y la Ayuda On-line deberán estar en ambos idiomas: euskera y castellano.
  - ✓ Formación a los usuarios y administradores del sistema.
  - ✓ Sistema implantado en entorno de producción.
  - ✓

En todas las fases del proyecto, así mismo será de referencia el documento de Estándares de desarrollo de sistemas software, que establece requisitos obligatorios y recomendaciones a seguir en todo el proceso de ejecución del proyecto en cuando a la elaboración de los productos y la entrega de los mismos a EJIE Cabe destacar que para el caso de desarrollo de aplicativos con interfaz web basado en navegador que será desplegado en el contexto de internet, las pruebas del sistema deberán contemplar los ensayos necesarios para garantizar que el sistema se comporta correctamente en los navegadores más utilizados (Internet Explorer, Firefox, Chrome, Opera, y Safari). Mientras que si el contexto de despliegue es intranet (con PCs gestionados), el sistema deberá soportar los navegadores homologados por los estándares tecnológicos publicados por la Dirección de informática y telecomunicaciones del GV para dicho contexto. Para el caso de extranet, la lista de navegadores a soportar se deberá decidir en función de los colectivos usuarios de la aplicación y del uso o no de PCs gestionados.

Igualmente será de referencia el **Documento de Estándares Tecnológicos** de Gobierno Vasco, publicado en: www.euskadi.eus/informatika

4.2 Herramientas del ciclo de vida de las aplicaciones

Como soporte e instrumento necesario en la ejecución de todas las fases del proyecto, existe un conjunto de Herramientas homologadas por EJIE, que abarcan todo el ciclo de vida de las aplicaciones, y que facilitan la realización de distintas tareas y normalizan la obtención de entregables.

Estas herramientas homologadas son las que se utilizan en el entorno de trabajo de EJIE, no pudiendo utilizarse en el mismo otras herramientas similares o equivalentes.

Para los trabajos a realizar en las dependencias del proveedor, su uso es recomendado frente a otros productos o herramientas del mercado, para dar cobertura a los cometidos para los que están destinadas. No obstante, en los casos en los que el resultado de uso de las herramientas sea un entregable con un formato específico y normado, su uso será obligatorio frente a otras herramientas de mercado, o bien en cualquier caso deberá proporcionarse un formato compatible.

En el documento PCT-Anexo Herramientas se detallan las herramientas homologadas.

4.3 Entorno Tecnológico

El entorno tecnológico de referencia será el establecido en los estándares del Gobierno Vasco junto con los específicos de la Agencia Vasca del Agua. Se deberán aplicar por tanto aquellos que se consideren necesarios para la ejecución del servicio, conforme a las características específicas de cada una de las aplicaciones incluidas en el alcance del presente pliego de condiciones técnicas.

El adjudicatario se compromete a adoptar y soportar en todo momento cualquier modificación o adaptación del entorno tecnológico establecido por el Gobierno Vasco e incluido dentro de sus estándares, así como todas aquellas otras actualizaciones específicas impuestas por la Agencia Vasca del Agua.

En este caso el entorno tecnológico de aplicación principal es el Entorno Tecnológico BigData de BATERA.

4.3.1 BigData en Batera

El entorno BigData de Batera es una plataforma gestionada (BDaaS), a disposición de las entidades clientes en el ámbito de la Convergencia, que se organiza sobre un catálogo de servicios estandarizados que se pueden consumir de forma ágil, pero con flexibilidad para adaptarse y evolucionar según las necesidades.

4.3.1.1 Servicios

Los servicios que realiza el equipo de BDaaS consisten en la prestación de consultoría especializada, diseño, soporte y mantenimiento asociado a la difusión, utilización y evolución de la plataforma BDaaS implantada.
Los distintos equipos implicados en el servicio BDaaS se encargan de promover y prestar un servicio eficiente y confiable, de manera que se garantice la integridad, su correcto dimensionamiento y utilización por los diversos proyectos e iniciativas de estas entidades.
En concreto, los equipos anteriormente indicados cubren los siguientes servicios:

- Soporte y mantenimiento de la Plataforma
- Formación
- Consultoría Especializada
- Apoyo en la Difusión y Promoción
- Seguimiento de best practices

Aportando servicios de control, monitorización, soporte, formación y evolución de la plataforma BDaaS en sí misma, tanto a nivel estratégico, como táctico (Proyectos y Servicios) y operativo.

4.3.1.2 Seguridad

Cada usuario entidad dispone de unas credenciales para acceder a los servicios.

Todos los datos en este servicio deberán estar cifrados, de forma que, si alguien accede a la máquina, ya sea de forma física o a través del sistema operativo, no pueda leer los datos almacenados en HDFS.

Todas las comunicaciones con este servicio están cifradas con algoritmos al menos tan seguros como AES-256.

Todos los accesos están monitorizados y quedan registrados durante al menos un mes de forma que se puedan hacer auditorías de seguridad si fuese necesario.

La seguridad del servicio está basada en Kerberos, para la autenticación, y en Ranger, para la autorización de los usuarios.

Mediante el ticket de Kerberos, el servicio sabrá con que usuario estamos *logueados*, y mediante las políticas de Ranger, el servicio sabrá que permisos tiene cada usuario, y que ficheros y carpetas puede ver dentro de HDFS.

4.3.2 Catálogo de servicios

4.3.2.1 File Service

Almacenamiento de documentos para su posterior tratamiento. Dicho servicio, permite a un usuario y/o a las aplicaciones, almacenar sus ficheros y recuperar datos en cualquier momento, independiente de cuál sea su tamaño.

Se cubren los siguientes escenarios:

- Transferencia y almacenamiento de ficheros de manera distribuida
- Procesamiento de datos de forma distribuida
- Explotación y visualización de datos
- Descarga de ficheros

Este servicio es accesible desde diversas fuentes (HUE, navegador y consola web…) para gestionar los ficheros, así como sus aplicaciones.

Cada usuario/iniciativa puede contratar una capacidad determinada de File Service, de forma que se le asignará un directorio dentro del sistema de ficheros HDFS. Dentro de esta carpeta el usuario tendrá todos los permisos CRUD.

Este servicio se apoya en las siguientes tecnologías principalmente → HDFS, NiFi, HIVE, Spark, DBeaver, Zeppelin, …

4.3.2.2 Table Service

Es un servicio muy parecido al anterior, con el añadido de que se pueden consultar los datos almacenados en tablas mediante diferentes herramientas. Este servicio es accesible desde diversas herramientas (HUE, HIVE, navegador y consola web…) para gestionar los datos.

Se cubren los siguientes escenarios:

- Transferencia de datos y almacenamiento en tablas
- Procesamiento de datos
- Explotación y visualización de datos
- Consulta de datos

Cada usuario/iniciativa puede contratar servicio de consultas de forma que se le asignará un directorio dentro del sistema de ficheros HDFS. Dentro de esta carpeta el usuario tendrá todos los permisos CRUD.

Este servicio se apoya en las siguientes tecnologías principalmente → HBase, NiFi, Spark, …

4.3.2.3 Event Service

Este servicio permite al usuario y/o aplicaciones, publicar y consumir mensajes de colas de eventos en tiempo real.

Se cubren presentan los siguientes escenarios:

- Transferencia y almacenamiento de eventos de manera distribuida
- Procesamiento de datos de forma distribuida
- Explotación y visualización de datos

En este servicio tenemos las siguientes tecnologías principalmente → Kafka, Spark Streaming, …

4.3.2.4 Search Service

Es un servicio de un motor de búsqueda que permite al usuario realizar búsquedas de cualquier tipo de información en grandes volúmenes de datos, gestionando todo tipo de ficheros/documentos independientemente de su formato.

Dispone de un sistema de indexación de información, de forma que guarda y procesa de forma distribuida toda la información ingestada en el sistema.

De este modo, se cubren los siguientes escenarios:

- Transferencia y almacenamiento de ficheros de manera distribuida
- Explotación y visualización de datos

En este servicio tenemos las siguientes tecnologías principalmente → Elasticsearch, Kibana, …

4.3.3 Versiones

A continuación, se detallan las versiones de los distintos servicios que componen el nuevo entorno:

Servicio	Versión
HDFS	3.1.1
YARN	3.1.1
Tez	0.9.1
Hive	3.1.0
Hbase	2.0.2
Oozie	4.3.1
Zookeeper	3.4.6
Infra Solr	0.1.0
NiFi	1.9.0
Atlas	1.1.0
Kafka	2.0.0
Knox	1.0.0
Log Search	0.5.0
Ranger	1.2.0
Ranger KMS	1.2.0
Spark2	2.3.2
Zeppelin	0.8.0
Kerberos	1.10.3-30
ECE	2.8.1
ElasticSearch	7.11.1

4.4 Contenedores

La solución de contenedores corporativa se basa en Red Hat Openshift v4.4 y se encuentra desplegada sobre servidores virtuales con soporte de alta disponibilidad/contingencia ofrecida por VMWare.

5 Descripción de entregables

Los productos a entregar, tanto productos software como la documentación, serán aquellos que define la metodología de desarrollo ARINbide-Predictivo en referencia a las fases definidas dentro del alcance del proyecto.

Igualmente, dentro del alcance definido del proyecto, la Metodología de Pruebas PROBAMET, así como el modelo de Aseguramiento de la Calidad de Sistemas Software, detallan los entregables a proporcionar en el ámbito de la calidad.

Con independencia de lo descrito y requerido a nivel metodológico (metodología de desarrollo ARINbide-Predictivo, y Metodología de Pruebas PROBAMET), a continuación se detallan las tareas y entregables a realizar/suministrar por el adjudicatario:

- Implantación en el Entorno de Desarrollo de EJIE.
- Definición, ejecución y gestión del Plan de Pruebas
- Preparación de componentes y soporte a la implantación en el entorno de Test o Preproducción.
- Soporte a la implantación en el entorno de Producción.
- La documentación de usuario deberá estar en ambos idiomas: euskera y castellano.
- La aplicación, las páginas, los textos, tablas de códigos, los mensajes de error/aviso y/o cualquier otro componente (tool-tip, textos en imagen, …) dirigidos al usuario final deberán estar en ambos idiomas: euskera y castellano.
- Entrega de fuentes de todos los componentes, ficheros de configuración, elementos gráficos utilizados, … y cualquier otro elemento que se considere necesario para el normal y adecuado mantenimiento/evolución del sistema, en soporte magnético y/o electrónico.
- Confección y entrega de la documentación (funcional, técnica, de explotación y de usuario) correspondiente a cada una de las fases del proyecto tanto en formato papel, como en formato electrónico.

Todos los entregables deberán proporcionarse utilizando las herramientas homologadas en EJIE para el ciclo de vida de las aplicaciones, o en su caso proporcionando formatos compatibles.

5.1 Normativa para la puesta en producción y posterior explotación del sistema

EJIE cuenta con actividades formalizadas para la gestión e implantación de los sistemas de información generados en los distintos entornos, establecidas en base a los procesos de Gestión de Cambios y Gestión de Despliegues.

A partir de estos procesos se garantiza que los sistemas que serán instalados en la infraestructura de EJIE cumplen con los estándares y normativas técnicas de la organización, que no impactarán en el correcto funcionamiento del resto de productos y que se conocen los procedimientos y particularidades relacionadas con su operación y soporte.

Para ello, es necesario que, durante el ciclo de vida del proyecto, el equipo de trabajo genere una serie de entregables, que serán validados a través del proceso de Gestión de Cambios:

- **Ficha de proyecto y Plan de Implantación**: primera aproximación a Gestión de Cambios donde se informa qué se va a hacer y calendario aproximado.
- **Documento de albergue:** recoge el diseño lógico e información de cualquier elemento no estándar o que pueda impactar tanto en necesidades de infraestructura como de usuario final.
- **Manual de implantación**: recoge los aspectos técnicos y configuraciones necesarias que deben ser realizadas para implantar la aplicación en un entorno tecnológico de EJIE.

- **Manual de producción**: contiene toda información que necesita tanto operación como los procesos para poder dar correctamente el servicio de explotación a partir del día de Puesta en Producción del proyecto.

- **Plantilla de monitorización**: recoge los elementos no estándar cuya monitorización permite conocer el estado de los sistemas.

Para cada uno de estos documentos, EJIE dispone de plantillas estándar que el proveedor deberá utilizar y entregar según los procedimientos establecidos.

Asimismo, en la cumplimentación de las plantillas se podrá solicitar el asesoramiento de los distintos Responsables de Proceso y Coordinadores de Infraestructura, para asegurar que la solución construida e implantada cumple con la normativa de EJIE y es posible operarla y soportarla con garantías.

6 Planificación y organización

6.1 Plan de proyecto

6.1.1 Plan de Entregas e Hitos de Control

El licitador deberá incluir en su Documento de Propuesta Técnica una descripción de las principales:

- Fases y tareas del proyecto,
- Entregas resultantes de cada fase,
- Hitos de control y/o verificación,

Junto con sus correspondientes fechas orientativas de finalización.

Así mismo, el licitador podrá establecer propuestas de entregas y/o certificaciones parciales acordes con el planteamiento del proyecto establecido, siempre que:

- Lo considere más adecuado para facilitar la implantación y puesta en marcha del objeto del contrato.
- No se establezca nada en sentido contrario en los apartados de: *Objeto, Alcance y Descripción del Proyecto*.

6.1.2 Diagrama de Gantt

El licitador deberá presentar un Diagrama de Gantt o Cronograma de las Fases y Tareas propuestas para la ejecución de las realizaciones objeto del presente pliego.

6.2 Organización del Equipo de Trabajo

El licitador deberá describir en su Documento de Propuesta Técnica:

- La organización (perfiles) del equipo de proyecto asignado a la realización de las actividades resultantes del presente pliego
- Las funciones de los mismos.

6.2.1 Asignación de recursos a fases del proyecto

El licitador deberá incluir en su Documento de Propuesta Técnica un desglose de horas y % de dedicación total por perfil y actividades del proyecto, siguiendo exactamente el siguiente modelo:

Descripción Perfil	Gestión Proyecto		Análisis		Diseño		Construcción		Implantación		Total Horas
	Horas	%	Horas	%	Horas	%	Horas	%	Horas	%	
Jefe Proyecto											
Analista Funcional											
Analista Técnico											
Analista-Programador											
Programador											
Diseñador											
TOTAL											

- Este desglose de horas se considerará como orientativo y será tenido en consideración en el momento de valorar el grado de aproximación a la planificación del proyecto según la estimación del licitador,

permitiendo, de esta forma, valorar la idoneidad del dimensionamiento del equipo de trabajo propuesto y su adecuación a la consecución de los objetivos. No obstante, este desglose de horas no se considera vinculante, al no tratarse de una contratación de horas de desarrollo, sino un proyecto "llave en mano" según el importe total ofertado.

- Se deberán ofertar los perfiles indicados en la tabla anterior y no otros.
- Para cada uno de los perfiles, deberá detallarse el número de horas y % de dedicación de cada uno de los miembros del equipo propuesto en función de la categoría asignada

6.2.2 Equipo de Trabajo

El equipo de trabajo propuesto estará formado por personal técnico con categoría profesional y nivel de especialización adecuados a las necesidades planteadas en cada momento, de acuerdo con las actividades que se vayan desarrollando.

El licitador se compromete, en caso de ser adjudicatario, a mantener el equipo, según lo establecido en la oferta, y durante el periodo fijado en cada actividad específica. Los recursos asignados a la ejecución del contrato deberán estar asignados desde el inicio del mismo. En el caso de modificación o sustitución de recursos, deberá comunicarlo previamente a EJIE, con la antelación mínima necesaria en función de la criticidad del servicio, y en cualquier caso no menor de 15 días naturales, y EJIE aprobará o rechazará dicha modificación o sustitución. En todo caso el sustituto o sustitutos tendrán un perfil igual o superior al ofertado, o recurso a sustituir.

EJIE podrá solicitar el cambio de los recursos asignado a los servicios, por inadecuación de los recursos o por incumplir los requisitos tanto de solvencia como técnicos exigidos, o haber declarado el contratista unas circunstancias que no son veraces, en cuyo caso podrá conllevar la rescisión del contrato y adjudicación al licitador siguiente en las calificaciones.

6.2.2.1 Constitución inicial del equipo de trabajo

El equipo técnico a incorporar tras la formalización del contrato para la ejecución de los trabajos deberá estar formado por componentes y perfiles relacionados en la oferta adjudicataria y consecuentemente valorados.

Si tras la adjudicación se observara que el equipo de proyecto no se corresponde con el establecido en la oferta y:

- Caso que el adjudicatario presente justificación escrita, detallada y suficiente, explicando el motivo que suscita el cambio, se procederá a:
    - La presentación por el adjudicatario de sustituto/s con un perfil de cualificación técnica igual o superior al de la persona que se pretende sustituir.
    - Aceptación del sustituto/s por parte del responsable del servicio de EJIE
- Caso de que se demostrase que el cambio no se corresponde con causa justificada, EJIE se reserva el derecho no solo a la aprobación del sustituto/s sino incluso a la revisión de la adjudicación y en su caso la rescisión del pedido/contrato, si este hecho fuera elemento determinante en la mencionada adjudicación.

6.2.2.2 Modificaciones en la composición del equipo de trabajo

La valoración de la calidad de los servicios objeto del presente pliego corresponde a la Dirección del Proyecto de EJIE, siendo potestad suya solicitar el cambio de cualquiera de los componentes del equipo de trabajo, con un preaviso de quince días, por otro de igual categoría o superior, si existen razones justificadas que lo aconsejen.

Si el adjudicatario propusiera el cambio de alguna de las personas del equipo de trabajo, se deberá solicitar por escrito con quince días naturales de antelación, y requerirá de las siguientes condiciones:

- Justificación escrita, detallada y suficiente, explicando el motivo que suscita el cambio.
- Presentación de sustituto/s con un perfil de cualificación técnica igual o superior al de la persona que se pretende sustituir.

- Aceptación del sustituto/s por parte de la Dirección del Proyecto de EJIE

Los posibles inconvenientes de adaptación al entorno de trabajo y al proyecto debidos a las sustituciones de personal, deberán subsanarse mediante periodos de solapamiento sin coste adicional, durante el tiempo necesario. Si a criterio de la Dirección del Proyecto de EJIE, esto no fuera posible, las dos primeras semanas de trabajo del sustituto/s no serán facturables corriendo a cargo del adjudicatario.

6.2.2.3 Horario y lugar de realización de los trabajos.

Los trabajos de desarrollo se realizarán normalmente en las dependencias del adjudicatario en cuyo caso:

- La jornada de trabajo estará de acuerdo a la establecida por el adjudicatario,
- Los componentes del grupo de trabajo deberán estar en una única ubicación y desarrollarán su labor con hardware y software propiedad del adjudicatario. Se deberá garantizar una presencia rápida en EJIE, ante cualquier eventualidad que pudiera surgir.

Las pruebas de integración, de carga, aceptación y la puesta a punto de los productos se realizarán en los locales de EJIE.

En aquellos casos en que los trabajos deban ser realizados en las dependencias de EJIE, estos se realizarán en las siguientes condiciones:

- La jornada de trabajo estará de acuerdo a la establecida por EJIE,
- Con carácter general los componentes del grupo de trabajo deberán desarrollar su labor con hardware y software propiedad del adjudicatario, salvo para labores de instalación, implantación, y puesta en marcha, que se realizará mediante los puestos asignados por la Dirección del Proyecto de EJIE
- Si por circunstancias excepcionales y cuando la realización efectiva de los trabajos no se ajuste a la planificación o así se requiera por las necesidades del servicio, el adjudicatario deberá comprometerse a una plena disponibilidad incluso fuera del horario habitual (salvo acuerdo previo por la Dirección del Proyecto de EJIE), sin que la realización del trabajo tenga una consideración especial a efectos de cómputo de horas o tarifa aplicable a las mismas

6.3 Transferencia tecnológica

Durante la ejecución de los trabajos objeto del contrato el adjudicatario se compromete, en todo momento, a facilitar a las personas designadas por la Dirección del proyecto de EJIE, y a tales efectos, la información y documentación que ésta solicite para disponer de un pleno conocimiento de los trabajos desarrollados, así como de los eventuales problemas que puedan plantearse y de las tecnologías, métodos, y herramientas utilizados para resolverlos.

7 Mecanismos de Seguimiento, Control y Supervisión

La gestión del proyecto se realizará según las normas procedimentales y de calidad de la Gestión de Proyectos de EJIE y de la Agencia Vasca del Agua.

El control y seguimiento de la calidad de servicio, será realizado por la Dirección del Proyecto de EJIE

Durante el desarrollo y ejecución del proyecto, EJIE podrá establecer controles de calidad sobre la actividad desarrollada.

7.1 Mecanismos de Seguimiento y Control.

El equipo de trabajo estará dirigido por un responsable del proyecto con dedicación estable y suficiente, quien coordinará las tareas implicadas en el proyecto y será el interlocutor principal con la Dirección del Proyecto de EJIE

Deberán establecerse todos los mecanismos de control que permitan verificar de forma continuada el cumplimiento de los objetivos especificados, tanto en plazos como en términos de calidad final. Para ello se contempla:

- Reuniones de seguimiento: a partir del plan de entregas requerido se celebrarán reuniones periódicas o coincidiendo con los puntos de decisión, destinadas a revisar el grado de cumplimiento de las tareas planificadas, le entrega de productos, las reasignaciones de personal, la validación de las programaciones de actividades a realizar, etc.
- Reuniones de control económico: de control económico del proyecto para llevar a cabo el plan de pagos establecido.
- Informes de situación y progreso: con los que la empresa adjudicataria comunicará a la Dirección del Proyecto de EJIE, del estado puntual de los trabajos, de los objetivos alcanzados, de las incidencias ocurridas, etc.
- Certificaciones parciales: si procede.
- Hojas de control de tareas: si procede.
- Documentos de gestión de proyectos: a solicitud del responsable de EJIE, la empresa adjudicataria cumplimentará los documentos de gestión de proyectos, según un modelo definido: actas de reuniones, registro de inspección y ensayo, informes de progreso, control de configuración (módulos-componentes instalados), etc.

7.2 Supervisión del proyecto

Tanto la dirección como la supervisión del Proyecto serán llevadas a cabo por el personal de EJIE, S.A. asignado al proyecto.

Los productos del trabajo deberán ser aprobados tanto por el responsable del proyecto por parte del adjudicatario, como por la Dirección del Proyecto de EJIE, si los encuentran conformes, emitirán su informe favorable, empezándose a computar desde ese momento el período de garantía que todo trabajo de realización externa debe aportar en previsión de defectos no detectados en las pruebas realizadas.

El incumplimiento de los plazos pactados sin causa que lo justifique dará lugar a las penalizaciones que se acuerden en el presente pliego (si existieran).

8 ANEXO 1 - Series Temporales Inundaciones

Se resume en este apartado la situación del sistema de repositorio de series temporales del área de Inundaciones en BDaaS infraestructura de Gobierno Vasco.

Se dispone de un sistema de captura, almacenamiento y explotación de series temporales para el área de Inundaciones que consiste en:

1. Captura o ingesta: Basada en el componente NIFI y se encarga de preprocesar las series temporales, enviarlas a unas colas de Kafka donde son procesadas y almacenadas por medio de unos Jobs de Flink. Además, se aplican unas reglas de negocio de validación del dato para detectar alarmas de valores anómalos.
2. El almacenamiento se basa en una estructura sobre la base de datos HBASE en la que se incluye por una parte el dato de la propia serie, y por otra parte los metadatos asociados. Se considera serie temporal a los valores de una medida obtenida de un sensor de una estación de control. Los metadatos son los valores distintos del dato, de la propia variable, sensor y estación.
3. Para la explotación se dispone de una API Rest, que permite exponer las distintas series temporales de manera individual o agrupadas en conceptos que denominamos clusters y son agrupaciones lógicas de estaciones. Además de la explotación del dato con este API se puede gestionar las altas/Bajas/modificaciones de: agrupaciones o clusters, estaciones, variables, y correcciones de los propios valores de la serie. Permite también duplicar o borrar una serie.
4. El uso de este API se ha validado con el ejemplo del cálculo de la evapotranspiración, y se ha integrado con procesos PYTHON de los que previamente se disponía en la agencia.

9 ANEXO 2 – Resumen Consultoría área de Vertidos y Aprovechamientos

A modo orientativo, se proporciona un resumen de una consultoría realizada por la Agencia Vasca del Agua, debiendo tenerse en cuenta que la volumetría proporcionada estaba pensada para una primera fase, es una aproximación inicial que podría servir para el arranque, pero la solución deberá ser escalable y no limitante, de manera que a futuro pueda contemplar todas las fuentes de datos de estas áreas.

REQUISITOS FUNCIONALES

El sistema ha de ser capaz de recoger, almacenar y procesar tanto medidas del caudal de aprovechamiento, como distintas medidas recogidas en las distintas plantas de depuración, tratamiento, u otras instalaciones.

Recogida de información

▪ La plataforma ha de ser capaz de recibir los datos de telelectura de los distintos servicios, así como de los caudales de aprovechamiento de manera continua, independientemente del momento en que se realice su procesamiento. La legislación actual establece, respecto a la obligación de reportar las medidas de caudal de aprovechamiento a URA, unas periodicidades anuales, mensuales o diarias/horarias, según sea el caudal máximo de la cuenca considerada. Esto nos plantea un escenario donde en principio no es necesaria una transmisión continua de las lecturas de caudal, aunque parece que una modificación en la que se está trabajando actualmente de esta orden incluirá el reporte de los datos en continuo.

▪ La legislación actual permite para parte de los caudales de aprovechamiento (tipo uno y dos), la toma de datos sobre papel y su posterior envío. Además, puede ser necesaria la incorporación de datos históricos de estas mismas lecturas. Por otra parte, de cara a un posterior procesamiento y explotación de estos datos será necesario incorporar otros datos (climáticos, económicos, …).

▪ Se deberá permitir por tanto también, la incorporación de datos por lotes.

▪ En principio la ingesta y procesado de los datos de aprovechamiento requerirá de una periodicidad al menos diaria, mientras que para los datos de vertidos se requiere que aun admitiéndose una cierta discretización (del orden de 10 a 30 minutos) el procesado sea near-real-time.

▪ La solución debe dar soporte a unas 100 estaciones de trabajo, cada una emitiendo unos 10 parámetros, además del volumen de datos emitidos por la parte de aprovechamiento, que deberán dar soporte a unas 108 estaciones de trabajo con unas 200 tomas.

Almacenamiento de la información

▪ La plataforma ha de ser capaz de mantener los datos originales recibidos (caudales y telemetría de los distintos servicios), de cara a habilitar escenarios de reprocesamiento frente a errores o nuevas necesidades futuras.

▪ La plataforma ha de permitir un almacenamiento de los datos refinados, así como de distintas agregaciones sobre los mismos: el volumen queda dependiente de los distintos escenarios a evaluar. En principio podemos consideran un volumen de 3 o 4 veces el volumen de los datos iniciales.

▪ La plataforma ha de permitir almacenar los datos recibidos, durante un tiempo en principio indeterminado del orden de decenas de años.

▪ La plataforma ha de permitir almacenar el resultado de los procesos, durante un tiempo dependiente del escenario de explotación, pero que en general será de unos pocos años.

▪ El almacenamiento ha de permitir tanto la lectura de datos secuencial (tipo fichero), como la realización de consultas con un tipo de acceso aleatorio (tipo SQL o similar). Este segundo tipo de lectura se ha de mantener como potencialidad, ante la posibilidad futura de una explotación tipo aplicación Web (la explotación inicial será de tipo reporting).

▪ El almacenamiento ha de permitir el borrado de los datos.

▪ Sin ser un requisito, sería altamente valorable que la plataforma admita la modificación de los datos almacenados.

Tratamiento y cálculos de la información

▪ El procesamiento ha de tener la expresividad suficiente para permitir todo tipo de escenarios de ETL: limpieza, filtrado, agregaciones, … que involucren distintos elementos de la entrada de datos, y distintos cálculos generados a partir de ellos.

▪ El procesamiento ha de poder incluir algoritmos y modelos de Machine Learning aunque inicialmente no se explotará esta faceta.

▪ El procesamiento ha de poder acceder a fuentes de datos externas a la plataforma en sí (Servicios de APIs de terceros), aunque en un escenario inicial no sea de aplicación.

▪ El procesamiento debe ofrecer una monitorización de negocio de cara a permitir su mejor operación.

Explotación

▪ La plataforma ha de permitir una explotación de tipo informe (reporting) sobre los datos y la creación de cuadros de mando.

▪ La explotación de los datos debe admitir incluir en un futuro la interacción con los datos y los resultados del procesamiento desde un interfaz web desarrollado a medida para los distintos casos de uso posibles.

▪ Se ha de incluir la posibilidad de levantar alertas ante determinadas situaciones de negocio.

Elementos transversales

▪ El sistema ha de evitar accesos no autorizados a los datos.

▪ Se ha de poder acceder a los datos desde los distintos elementos de proceso.

▪ El sistema de procesamiento debe permitir su testeo ante modificaciones y el paso entre entornos de nuevas versiones del proceso.

▪ El sistema ha de emitir unos logs y ha de permitir una monitorización del funcionamiento de los procesos involucrados; idealmente estos logs han de presentar una interface de explotación única.

REQUISITOS TÉCNICOS:

Ingesta

▪ El sistema desarrollado ha de permitir la ingesta de datos tanto en Batch como en Streaming.

▪ La comunicación de los datos hacia el sistema vía Streaming ha de poder emplear los protocolos MQTT y HTTP (API Rest).

▪ La herramienta que provee BDaaS como elemento de recepción de datos en Streaming es Kafka 2.0.

Almacenamiento

▪ Volumen de datos a almacenar: en la plataforma se guardarán 350GB/año. Para mantener 15 años de datos el volumen será de 3,5 TB.

▪ Consideraciones para el cálculo:

➢ 100 estaciones de vertidos con 10 parámetros cada una, es decir 1000 mensajes, quince minútales.

➢ Para aprovechamiento, 200 tomas aproximadamente, con 5 parámetros distintos por aprovechamiento cada una con una recogida 15 minutal.

➢ Cada mensaje se ha estimado con un tamaño de 500 B (tamaño estimado con un mensaje de tipo JSON muy poco “optimizado”).

➢ Esto da unos 35 GB de datos de entrada (raw) al año. Consideramos un multiplicador de x 10 en concepto de almacenamiento de las transformaciones intermedias, generación de índices, etc.

▪ Throughput de entrada de los datos. Una media de 4 MB/hora con picos de 100KB/seg (Considerando la entrada 15 minutal de todas las lecturas concurrente, y dando un tiempo de 10 segundos para la ingesta).
▪ BDaaS provee como elementos de almacenamiento de datos Hadoop HDFS 3.1.1 y Hbase 2.1.6.
▪ Dotaría de mayor confiabilidad y ahorraría costes asociados a procesos de consolidación si el almacenamiento provee de cierta transaccionalidad como la ofrecida por el formato Delta.

Procesamiento
▪ La plataforma propuesta ha de ser capaz de escalar horizontalmente su capacidad de procesamiento y adaptarse a una carga creciente de datos, así como de realizar una distribución del trabajo a realizar.
▪ Se hace necesario proveer de un mecanismo de procesamiento en Streaming, así como de mantener procesamientos de tipo Batch.
▪ Para mantener la posible explotación futura vía aplicación web (inicialmente no se contempla), se hace necesario que el procesamiento tenga la posibilidad de acceder a alguna base de datos externa para almacenar los datos que mostrará la aplicación (agregados, predicciones, …).
▪ La plataforma ha de permitir acceder a los datos almacenados desde todos los elementos de procesamiento, en una forma automatizable, con una identidad de procesos “múltiple” .

Explotación
▪ La plataforma en primera instancia solo proveerá de una explotación de tipo reporte. BDaaS ofrece para esto el stack de ELK (Elasticsearch, Beats,y Kibana – versión 2.8.1).
▪ La plataforma debe permitir incluir sobre este mismo reporting alertas ante determinadas condiciones de negocio.
▪ La plataforma debe contemplar la posibilidad futura de una implementación de determinados procesos vía web application.