

Manizales, 18 de junio de 2021

IDEAM

Referencia: Desarrollo de herramienta web acompañada de una solución de High Performance Computing (HPC) como soporte para la ejecución y visualización de modelos meteorológicos descriptivos y predictivos.

AVISO DE CONFIDENCIALIDAD

Los destinatarios y receptores de la presente propuesta, reconocen y aceptan que toda la información presentada, metodologías, modelos, Know How, el plan técnico y su desarrollo, la descripción del trabajo a realizar, la información sobre las hojas de vida de los profesores investigadores, y la propuesta financiera sobre el costo del desarrollo del proyecto, tienen el carácter de confidencial, y por lo tanto la empresa, compañía, entidad, o institución receptora; así como todas y cada una de las personas que detente la calidad de subordinados, empleados o dependientes y que tengan acceso a la información con ocasión de la presente propuesta; se comprometen a conservar y mantener de manera estrictamente confidencial y no revelarla a terceros, todas las informaciones y datos suministrados en la misma. La Cláusula de confidencialidad se Regirá por la reglamentación sobre Informaciones Confidenciales y Secretos Industriales de que tratan los artículos 260 a 266 de la Decisión 486 de la Comisión del Acuerdo de Cartagena.

PROPIEDAD INTELECTUAL SOBRE LOS CONTENIDOS: Los destinatarios y receptores de la presente propuesta, reconocen y aceptan que el contenido de la misma contiene material, productos o procedimientos, contenidos e información protegida por las normas nacionales y extranjeras vigentes sobre Propiedad Intelectual. Por lo tanto cualquier utilización, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso o digital se encuentran prohibidos, y solo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito del Centro de Bioinformática y Biología Computacional. Todo lo anterior de conformidad con lo establecido en la normatividad vigente en Propiedad Intelectual, particularmente con lo establecido en el Convenio de Berna, el Convenio de París; la ley 23 de 1982, Ley 44 de 1993, ley 1450 de 2011, la Decisión Andina 351 de 1993, La Decisión Andina 345 de 1993, la Decisión Andina 486 de 2000, y en general los convenios internacionales sobre la materia de los cuales Colombia es miembro; y demás normas concordantes que las modifiquen o adicionen.

Contenido

Tabla de contenido

AVISO DE CONFIDENCIALIDAD 2

Desarrollo de herramienta web acompañada de una solución de High Performance Computing (HPC) como soporte para la ejecución y visualización de modelos meteorológicos descriptivos y predictivos. 5

Justificación 5

Pertinencia 13

Alcance 14

Objetivo general 14

Objetivos específicos 14

Metodología 15

1. Gestión del Proyecto 15
2. Diseño de la herramienta web 15
3. Disponibilidad de infraestructura HPC 16
4. Desarrollo de la plataforma 16
5. Acompañamiento científico especializado 16
6. Pruebas de rendimiento 16
7. Implementación en ambiente de producción 17
8. Soporte y mejoras 17
9. Transferencia de conocimiento 17

Equipo de trabajo base 17

Herramienta web 17

Solución de HPC 18

Aclaraciones 19

Entregables 19
Presupuesto y EDT 20
Herramienta de software 21
Solución de HPC 27
Resumen presupuestal 32
Forma de desembolso de aportes 32
Lugar de la prestación de la solución 35
Vigencia de la propuesta 35
Bibliografía 35

Desarrollo de herramienta web acompañada de una solución de High Performance Computing (HPC) como soporte para la ejecución y visualización de modelos meteorológicos descriptivos y predictivos.

Justificación

Los procesos de analítica de datos apoyados en las tecnologías emergentes y el creciente recurso tecnológico están generando una transformación a escala social y económica. Se estima que para el 2030, estas tecnologías podrían contribuir con hasta 15.7 billones de dólares a la economía global, incluidos 0.5 billones en América latina, según proyecciones de la empresa consultora PricewaterhouseCoopers (PwC). Esto presenta cifras concretas y evidencia que estas herramientas son algo más que una simple tecnología novedosa y que por el contrario pueden impactar y lograr efectos de largo alcance en el mundo, lo que plantea una creciente necesidad para las empresas y los gobiernos de crear sistemas de gobernanza al más alto nivel para ayudar a las organizaciones a involucrarse y aprovechar sus datos de forma efectiva.

Actualmente, alrededor del mundo las compañías han visto el beneficio de sus inversiones en Big Data, lo cual se evidencia en el estudio llevado a cabo por PwC en Estados Unidos. Por ejemplo, el sector que obtiene más valor es el público con un 43%, seguido del sector de la energía y los suministros con un 37%, y la industria con un 33%. Adicionalmente, el Big Data mejoraría el desarrollo de sus compañías a través de la inclusión de la Inteligencia Artificial [1].

Siguiendo estas premisas, el Gobierno de Colombia reconoce la necesidad de crear un ecosistema que permita aumentar el aprovechamiento de datos públicos y privados y la adaptación de tecnologías emergentes como las mencionadas anteriormente, esto por medio de la creación de políticas públicas como el Conpes 3920 de Política Nacional de Explotación de Datos (Big Data) y el Conpes 3975 de Política Nacional para la Transformación Digital e Inteligencia Artificial, las cuales sientan las bases para la transformación digital del país. Estas políticas y en especial el Conpes 3920 buscan superar los retos que hasta el momento han impedido la disponibilidad masiva de datos digitales de las entidades públicas, como una etapa fundamental para el aprovechamiento de estos, a través de la aceleración y fortalecimiento de la digitalización, apertura de datos y la interoperabilidad entre entidades, elementos que generan valor social y económico.

Por otro lado, con el objetivo de dar valor a la gran cantidad de datos que se han estructurado, así como los que se están generando constantemente, se hace evidente la necesidad de articular los

recursos computacionales que permitan tecnificar y fortalecer la infraestructura de procesamiento de datos del país y en especial del IDEAM. Esto es, como se mencionó, la disposición de servicios de analítica de datos apoyados en infraestructura de supercomputación al alcance de las instituciones académicas, la sociedad civil, los emprendedores y las entidades públicas que hacen uso de grandes cantidades de datos para la creación de soluciones innovadoras, optimización de procesos y creación nuevos productos y/o servicios. Es así como toma importancia considerar que en el ámbito de la analítica de datos, la interoperación de diferente fuentes de información y la carga masiva de nuevos datos genera una necesidad latente en términos de escalamiento vertical y horizontal en la infraestructura de computación, así como la incursión en los paradigmas asociados a computación distribuida y computación de alto desempeño (HPC), permitiendo de esta manera, la optimización de procesamiento requerido en soluciones que deben dar respuestas concisas en intervalos de tiempo cortos.

Teniendo en cuenta el contexto anterior, relacionado con la cuarta revolución, y a raíz de la ola invernal ocasionada por el fenómeno de la Niña 2010-2011, el Ministerio de Ambiente y Desarrollo Sostenible reportó al Fondo Adaptación (FA) la necesidad de implementar acciones para reducir los riesgos en las zonas afectadas como una estrategia integral de ordenamiento ambiental del territorio tendiente a la mitigación y prevención de riesgos y a la protección en lo sucesivo de la población de las amenazas ambientales que están sucediendo [1].

En consecuencia, el Ministerio de Ambiente y Desarrollo Sostenible presentó una postulación al FA, de carácter estratégica y visionaria para la reducción de las futuras condiciones de riesgo del país y que permitiera emitir con mayor oportunidad alertas tempranas ante eventos externos de origen hidrometeorológicos en pro de salvaguardar la vida y bienes de los colombianos en el territorio nacional. Dicha postulación se definió en 4 componentes ante el consejo directivo del FA, quien las aprobó en la sesión del 2 de febrero de 2012, y se describen a continuación:

1. Incorporación del componente de gestión del riesgo como determinante ambiental del ordenamiento territorial en los procesos de formulación y/o actualización de 130 Planes de Ordenación y Manejo de Cuencas Hidrográficas afectadas por el fenómeno de La Niña 2010-2011.
2. Delimitación de ecosistemas estratégicos en las cuencas priorizadas para su protección y para la regulación hídrica como medida de prevención del riesgo: Humedales y Páramos.
3. Diseño e implementación de una red hidrometeorológica para alertas tempranas como medida de reducción del riesgo en las cuencas priorizadas.
4. Diseño e implementación de una herramienta de información para la toma de decisiones en materia de ordenamiento ambiental del territorio.

De las anteriores postulaciones, la número 3 y 4 fueron dispuestas para ser ejecutadas por el Instituto de Hidrología, Meteorología y Estudios Ambientales – IDEAM mediante el fortalecimiento del sistema de alertas tempranas y desde ese momento se han realizado todas aquellas acciones para su complimiento, como lo es el caso de la postulación 3 que fue entregada en año 2018 y la número 4 que se encuentra en estructuración. Esta última es fundamental para los tomadores de decisión en los territorios, ya que ni siquiera han transcurrido los cien primeros días de 2021 y ya se han vivido fenómenos hidrometeorológicos realmente sorprendentes en diversas latitudes. Uno de los ejemplos más próximo en Colombia fue la devastación dejada por el huracán Iota en las islas de San Andres y Providencia, y en el caso de España, la borrasca Filomena que a principios de enero se convirtió en la mayor tormenta de nieve del país de los últimos cincuenta años, y hace también pocas semanas, el estado norteamericano de Texas experimentó una ola de frío que incluyó la mayor nevada de su historia.

Estos acontecimientos, tan excepcionales e impactantes, se derivan de la variabilidad climática, un concepto que hace referencia a condiciones meteorológicas extremas pero que son temporales y transitorias. Variabilidad climática no es lo mismo que cambio climático, esta última noción, por el contrario, indica variaciones en la media de la temperatura terrestre y las condiciones ambientales a lo largo de una escala temporal más prolongada y constante. Sin embargo, el cambio climático implica una mayor frecuencia e intensidad de algunos fenómenos extremos, como olas de calor, sequías o lluvias torrenciales.

Carlo Buontempo, director del Servicio de Cambio Climático de Copernicus (C3S, por sus siglas en inglés), afirma: “En este contexto de alteraciones, la capacidad de recopilar e interpretar la enorme cantidad de datos sobre el clima a largo plazo y de forma continua representa una herramienta fundamental para entender si los fenómenos meteorológicos excepcionales como los referidos pueden considerarse eventos aislados o ser indicadores de importantes cambios en las condiciones climatológicas del planeta”. Este organismo al igual que el IDEAM trabajan para poner a disposición de la sociedad información clave sobre la situación climática pasada, presente y futura en áreas determinadas (Países, regiones, lugares focalizados) o del planeta.

También, como lo describe la revistas Nature, el impacto humano en las concentraciones de gases de efecto invernadero en la atmósfera y los efectos en el sistema climático han sido documentados y explicados por un vasto recurso de publicaciones científicas, y la conclusión —de que las emisiones antropogénicas de gases de efecto invernadero deben reducirse drásticamente en unas pocas décadas para evitar una catástrofe climática— es aceptada por más del 97% de la comunidad científica del sistema terrestre hoy en día [2].

La presión para proporcionar predicciones en un clima cambiante, por ejemplo, el número y la intensidad de los ciclones tropicales y la probabilidad de que las olas de calor y la sequía coexistan es particularmente alta porque el impacto actual de los peligros naturales a nivel mundial es asombroso. En el período 1998-2017, se han producido más de 1 millón de muertes y varios billones de dólares en pérdidas económicas [3]. Los años entre 2010 y 2019 han sido la década más costosa registrada, con un daño económico que alcanzó los US$2,98 billones, US$1,19 billones más que entre 2000 y 2009 [4]. Tanto el clima extremo como la posible falta de acción sobre el cambio climático se clasifican como los principales riesgos que combinan la máxima probabilidad e impacto para nuestro futuro [5].

Estas pérdidas no invalidan el progreso constante logrado en la predicción meteorológica en las últimas décadas, es decir, el resultado combinado de la mejora de los sistemas de observación, una mejor comprensión de los procesos físicos relevantes que ocurren e interactúan en el sistema de la Tierra y el crecimiento exponencial del rendimiento de la tecnología informática de propósito general a un costo casi constante [6]. Sin embargo, continuar a este ritmo está siendo cuestionado, debido a los efectos aparentes del cambio climático en nuestro medio ambiente— en particular en la frecuencia de ocurrencia y la intensidad de los extremos ambientales— los cuales requieren una respuesta política urgente y progresos mucho más rápidos en la entrega de predicciones de cambios futuros [7] [8]. Los modelos del sistema terrestre necesitan un progreso más que constante y dar un salto a una resolución muy alta, una representación más realista de los procesos a todas las escalas y su interacción entre la atmósfera, el océano, la criosfera, las superficies terrestres y la biosfera. Este movimiento inevitablemente se traducirá en un salto en nuestras necesidades computacionales y de capacidad de manejo de datos.

Para algunos científicos y teóricos, la predicción meteorológica ha sido una aplicación pionera de simulaciones numéricas por computador desde el 'Proyecto de Meteorología' de John von Neuman a finales de la década de 1940 [9] . Mucho se ha logrado desde entonces y las predicciones globales operativas de hoy se completan en una hora para modelos con alrededor de 10 millones de puntos de cuadrícula, 100 capas verticales y 10 variables pronósticos, inicializadas utilizando 100 millones de observaciones por día.

Estos cálculos se ejecutan en cientos de nodos de unidades de procesamiento central de uso general (CPU) ofrecidas por los proveedores en los EE. UU., Asia y Europa. La necesidad de ejecutar conjuntos de simulación para predecir tanto el estado como la incertidumbre [10] multiplica tanto la carga informática como la carga de datos, pero ha demostrado ser enormemente beneficiosa para la toma de decisiones [11].

Reconociendo las limitaciones computacionales y sus resultados sobre los modelos meteorológicos y modelos climáticos, los cuales han demostrado notables divergencias en las últimas décadas, se puede afirmar en términos generales que los modelos climáticos deben representar ciclos estables a expensas de los detalles de los procesos a una pequeña escala, mientras que los modelos meteorológicos, necesitan un gran nivel de detalle para los pronósticos locales, a expensas de excluir aquellos procesos del sistema terrestre que son menos relevantes para el clima. Por ejemplo, la descripción precisa de los procesos del ciclo del agua es altamente relevante tanto para los modelos meteorológicos como climáticos, mientras que la representación del ciclo del carbono sólo es importante en las escalas de tiempo climático. Sin embargo, cada vez más, el reconocimiento de que las pequeñas escalas importan para las predicciones climáticas [12] y que la complejidad del sistema terrestre importa para la predicción meteorológica [13] está llevando a desarrollos convergentes. Como consecuencia, la necesidad de ambos de una resolución muy alta y un completo proceso y definición del sistema terrestre. Por lo tanto, utilizando la arquitectura de superordenadores (HPC) se podrá medir cuánto más realistas se vuelven las simulaciones de muy alta resolución [14] [15]

Esto ha dado lugar a códigos heredados muy grandes, impulsados principalmente por preocupaciones científicas, es decir, los modelos del sistema terrestre discretizan el conjunto de ecuaciones físicas para los procesos resueltos en el espacio y el tiempo [16] y utilizan parametrizaciones para procesos no resueltos como microfísica en la nube y turbulencias, que afectan a las variables de pronóstico en las escalas resueltas [17]. Lo mismo se aplica a la asimilación de datos, cuyo rendimiento informático está impulsado principalmente por el modelo de pronóstico y los componentes acoplados que representan procesos oceánicos, olas superficiales, hielo marino, superficies terrestres, incluida la vegetación, etc. en el sistema terrestre [18]. Diferentes opciones de discretización implican diferentes solucionadores con patrones específicos para el acceso a la memoria y la comunicación de datos por paso de tiempo. El paso del tiempo en sí es un factor de costo importante y depende de la elección de la discretización [19] [20], pero también está limitado por la no linealidad del problema y el tipo y la velocidad de los movimientos que se deben resolver [21].

Ha habido varios programas con el objetivo de acelerar sustancialmente las infraestructuras climáticas y de código de predicción climática en la última década. Sin embargo, estas mejoras "tradicionales" se abstienen de tocar los conceptos algorítmicos básicos y trabajar a lo largo de los desarrollos de software científico conocidos. Por ejemplo, el código para realizar los análisis (Domino y Downscaling) está escrito principalmente por científicos hidrometeorológicos y luego científicos informáticos extraen el rendimiento refactorizando gradualmente el código, normalmente mejorando el manejo de la memoria y la comunicación, e introduciendo directivas de código para explotar el paralelismo y la vectorización basados en modelos de programación estándar.

En términos más generales, la superposición de informática y la transferencia de datos pueden acelerar algoritmos numéricos individuales que dependen en gran medida de la comunicación de datos [22] [23], o acelerar los flujos de trabajo en los que el análisis de datos y el postprocesamiento se ejecutan simultáneamente con el modelo [24].

Si bien una implementación a medida en una arquitectura (HPC) específica puede devolver aceleraciones sustanciales [25], lograr el rendimiento sin sacrificar la portabilidad se considera crucial para evitar una solución en la que se tiene que reescribir continuamente software complejo para una opción de hardware en particular. Hoy en día, la mayoría de los modelos y sistemas de asimilación de datos todavía se basan en millones de líneas de código Fortran. Además, adoptan una estructura de bloques bastante rígida en el contexto de un esquema de paralelización híbrida mediante la interfaz de paso de mensajes (MPI) y el multiprocesamiento abierto (OpenMP) combinado con un flujo de datos inspirado en el conocimiento del dominio o basado en la conveniencia dentro de la aplicación de modelo [26].

Por otro lado, el factor tecnológico en sí mismo no representa una solución a un problema de análisis de datos, dado que es necesario disponer del personal técnico y científico idóneo con experiencia en diseño, desarrollo y ejecución de modelos necesarios para modelar los complejos comportamientos meteorológicos. De igual forma, dicho personal deberá estar en la capacidad de proponer activamente mejoras, avances y trabajos futuros, orientados al escalamiento y optimización de los modelos desarrollados. La base para revisar por completo este enfoque son de nuevo estructuras de datos genéricas y conceptos de software específicos del dominio que separan el código científico de las capas de software dependientes del hardware, distinguiendo entre la preocupación de flexibilidad algorítmica del front-end y la preocupación por la flexibilidad del hardware del back-end [27]. Idealmente, esto da como resultado una vista de estructura de datos única de toda la compleja aplicación acoplada en una amplia gama de arquitecturas, que se utiliza en todo el flujo de trabajo de manejo de observaciones, simulación, asimilación, E/S, postprocesamiento y archivo de datos [28] [29] [30] [31]

El desarrollo y la innovación de los sistemas informáticos nunca se detienen. El mejor precio-punto de rendimiento cambiará rápidamente y en tres años, un diseño del sistema probablemente se verá muy diferente, no obstante, los desafíos sociales derivados del cambio climático exigen una permanente evolución y cambio gradual en las habilidades y capacidades predictivas, por lo tanto, es sobre las anteriores afirmaciones y en particular la última conclusión que el IDEAM sostiene la necesidad de realizar el fortalecimiento de la red de origen hidrometeorológico al Sistema Nacional para la Prevención y Atención de Desastres.

Paralelamente, el panorama en avances tecnológicos trae consigo mejoras sustanciales a nivel de desarrollo de software, ya que es necesario escalar los algoritmos de inteligencia artificial de tal manera que Aproveche al máximo los recursos de supercomputación disponibles y las nuevas técnicas en ciencia de datos. En este orden de ideas, y teniendo en cuenta las pruebas de concepto realizada con los modelos del IDEAM, se han descubierto puntos de mejora que permiten escalar los modelos actuales hacia resultados más precisos y tiempos de ejecución más cortos, donde se espera afrontar retos tales como la paralelización de procesamiento utilizando el modelo actual, dado que el mismo está construido con restricciones de procesamiento iterativas que dificultan la optimización en los tiempo de entrega de resultados.

Tal fortalecimiento se constituye en desarrollo, adquisición e implementación de una herramienta tecnológica con sus diferentes componentes tecnológicos. Bajo este contexto, la Oficina de Informática y la Oficina de Pronósticos y Alertas - OSPA del IDEAM han interpretado de manera eficaz las cadenas de valor al interior de Entidad, en las cuales las diferentes dependencias y grupos de trabajo que la componen en torno a la "Prevención y Atención de Desastres" han validado y reconocido las necesidades expuestas y en un trabajo conjunto, se han identificado proyectos que pueden atender de manera transversal y específica los requerimientos de los procesos para cada área involucrada.

Con base en lo anterior, nace el alcance a un alto nivel de construir una Herramienta Web diseñada sobre tres (3) componentes estructurales, los cuales son:

1. Fábrica de Software. Cuyo objetivo propuesto sería: Desarrollar, adquirir, probar, documentar e implementar una interfaz gráfica de interacción o portal web, para el uso y apropiación al servicio de los tomadores de decisión, ciudadanos del común y actores claves, para que, de manera orgánica, centralizada y andragógica, en torno a los productos y servicios de información relacionados con el monitoreo de variables ambientales en Colombia se pueda tener un mayor control de tipo hidrometeorológico del territorio. Lo anterior bajo los preceptos de la política pública del gobierno digital dispuesta por MinTIC.
2. Solución de Computación de alto desempeño – HPC y Análisis de datos de alto rendimiento – HPDA. Cuyo objetivo propuesto sería: Adquirir una solución de servicios de tipo HPC y HPDA para ejecutar los diferentes modelos meteorológicos, optimizando los tiempos de ejecución, los niveles de precisión, las áreas de análisis y la calidad de los pronósticos actuales del IDEAM.
3. Adquisición de equipos. Cuyo objetivo propuesto sería: Adquirir una solución de arquitectura de hardware propia con su correspondiente licenciamiento de software, escalable e hiperconvergente; requerida para que la Herramienta Web, desarrollada en el componente

1 y 2, pueda ser instalada, configurada, parametrizada, soportada, almacenada, respaldada y publicada sobre Internet.

Una vez planteados los tres ejes articuladores de la solución, es importante entender que el hardware y el software informáticos avanzan en escalas de tiempo muy diferentes. La vida útil del software puede ser de décadas, mientras que el hardware de alto rendimiento se utiliza generalmente durante menos años. Así, las inversiones algorítmicas y de software propuestas deben proporcionar la máxima flexibilidad y apertura a las nuevas tecnologías de hardware en evolución constante. En ese sentido, adquirir una solución de tipo HPC para el IDEAM pone de manifiesto:

1. Invertir en infraestructura de alto costo y con obsolescencia en menos de 5 años.
2. Cargar con costos y gastos operativos a la oficina de informática para mantener y soportar una infraestructura tan especializada.
3. Considerar una reconfiguración y ampliación del espacio físico del Datacenter actual del IDEAM.
4. Considerar una adaptación del suministro eléctrico y sus redundancias en el Datacenter actual del IDEAM.
5. Considerar una reestructuración de cableado y refrigeración para el Datacenter actual del IDEAM.
6. Evaluar y mitigar los impactos ocasionados por la migración de los servicios actuales que presta el IDEAM desde su Datacenter, a otro Datacenter alterno y temporal.

Bajo el panorama anterior, son evidentes las desventajas intrínsecas que conlleva la instalación y puesta a punto de una infraestructura HPC propia, cuando la misma va a ser utilizada para un servicio específico y no se extenderá su uso a otros servicios que ocupen y sostengan económicamente este tipo de tecnología, siendo viable considerar accesos alternativos al servicio que permitan alcanzar los objetivos propuestos.

En consecuencia, el IDEAM enfocará sus esfuerzos estratégicos y misionales para el aprovechamiento de la información resultante de manera confiable, consistente y oportuna, entre otras, sobre el origen y el estado de las variables ambientales en Colombia, de esta manera lograr el fortalecimiento del Sistema Nacional para la prevención y Atención de Desastres.

Pertinencia

El Centro de Bioinformática y Biología Computacional de Colombia - BIOS es una iniciativa de Ciencia, Tecnología e Innovación (CTeI) que nació como una respuesta del gobierno nacional (MinTIC y Colciencias) y la empresa privada, para incentivar el desarrollo de la investigación científica de alto nivel en Colombia, apoyada en infraestructura computacional de alto rendimiento. El aprovechamiento de esta infraestructura está ligada a la prestación de servicios en bioinformática, al desarrollo de proyectos de investigación y en generar soluciones de bioingeniería. En esta última área y como parte de sus objetivos está formular y ejecutar proyectos de investigación y desarrollo tecnológico ofreciendo soluciones ajustadas a las necesidades particulares de la academia y la industria.

Entre las líneas de investigación y desarrollo de BIOS se encuentran el despliegue de aplicaciones en salud, ambiente, procesamiento inteligente de datos y arquitecturas avanzadas de cómputo de alto desempeño. De esta manera, BIOS ha definido una estrategia para consolidarse como un nodo de articulación para el desarrollo científico en el área de ciencia de datos, trabajando de manera conjunta con la academia, la industria y el gobierno para promover el establecimiento de sinergias que impulse la innovación y la transferencia tecnológica en estas áreas.

Por otro lado, desde el año 2017 BIOS ha venido trabajando en proyectos relacionados con gestión del riesgo ambiental, principalmente en el monitoreo de laderas y calidad del aire. Con base en lo anterior, y teniendo en cuenta la importancia de la gestión del riesgo en una ciudad como Manizales, BIOS consideró enfocar una de sus líneas estratégicas en el análisis de datos ambientales, por lo cual, proyectos de este tipo representan un aprendizaje fundamental para sofisticar las herramientas ya desarrolladas y generar nuevas que permitan futuros desarrollos que basados en el análisis de datos faciliten la toma de decisiones estratégicas enfocadas en evitar perdidas humanas, ambientales y materiales. Adicionalmente y como parte de su rol como Centro de Desarrollo Tecnológico, este proyecto permitirá fomentar la articulación entre las instituciones y la generación de conocimiento de valor.

Alcance

El alcance del presente proyecto incluye dos líneas importantes a saber: i) La solución HPC (Infraestructura supercomputación) y ii) La fábrica de software, las cuales convergen de manera simultánea con el fin de lograr la difusión científica y técnica de la información resultante luego de su procesamiento.

Adicionalmente la solución que se pretende implementar, requiere de la realización de análisis iterativos, que permitan identificar cuál de las configuraciones para correr los modelos en HPC, resultan ser los más apropiados con relación al tiempo estimado mínimo (6 horas) para que el IDEAM realice los pronósticos necesarios.

En conclusión se desarrollará e implementará una herramienta web con sus diferentes componentes tecnológicos, entre ellos, una arquitectura e infraestructura de referencia HPC y HPDA, y, capacidades científicas y técnicas especializadas para la ejecución de modelos de pronósticos y asimilación de datos, que se utilizarán tanto para la optimización de las capacidades tecnológicas, tiempos de procesamiento de datos, comprensión e investigación, como en la predicción de las condiciones meteorológicas mesoescalares, entre otras, para Colombia. La herramienta web contará con soporte durante los doce (12) meses siguientes a la finalización del proyecto.

Objetivo general

Desarrollar e implementar una herramienta tecnológica a través de una arquitectura e infraestructura de referencia HPC y HPDA, y, capacidades científicas y técnicas especializadas, para el fortalecimiento de la red de alertas de origen hidrometeorológico.

En este contexto, se busca:

Objetivos específicos

1. Desarrollar adquirir e implementar una plataforma de visualización (herramienta Web) sobre los paradigmas de UX/UI24, de tal manera que permita una navegación fácil e intuitiva de cara a la comunicación de resultados de los modelos analíticos descriptivos y predictivos generados por el IDEAM, para los diferentes usuarios y tomadores de decisiones, en torno a los productos y servicios relacionados con el monitoreo de variables ambientales en Colombia.

2. Adquirir una solución de computación de alto desempeño (HPC y HPDA), soporte científico y técnico como servicio, en función de las necesidades institucionales y su alineación con los objetivos estratégicos de la entidad.

Metodología

Dado la magnitud del proyecto se plantea el uso de una metodología PMI a través de la cual sea posible planear a detalle la ejecución de las diferentes fases necesarias para alcanzar los objetivos planteados. De esta manera será posible asegurar que el producto y resultado del proyecto, esté claramente definido y acotado por todas las partes implicadas.

El desarrollo de esta propuesta se basa en los ejes fundamentales que se describen a continuación:

1. Gestión del Proyecto

Se refiere a las actividades que son necesarias para articular a todos los participantes en la construcción de la solución, la asignación de tareas y el seguimiento de estas para asegurar un correcto proceso de desarrollo del proyecto. En esta fase se busca establecer las bases y normas que se deben respetar en el proyecto para una correcta ejecución. De esta manera se espera construir el plan de trabajo, detallando tiempos, hitos, responsables, así como los formatos necesarios para reportar avances y entregables asociado al proyecto.

2. Diseño de la herramienta web

Basado en el proceso tradicional de ingeniería de software, esta etapa permite abordar en primer lugar el proceso de levantamiento de requerimientos funcionales y no funcionales, a través del cual es posible conocer las características generales y restricciones que deben ser consideradas en la solución, haciendo uso de los respectivos formatos previamente definidos, así como el enfoque utilizado para un levantamiento mencionado. Para lograrlo se destaca el uso de historias de usuarios y casos de uso.

Por otro lado, y posterior el levantamiento de requerimientos, se realiza la consolidación de dichos requerimientos en una propuesta que defina la estructura visual y los lineamientos estéticos de la misma teniendo en cuenta paradigmas UX/UI. Para esto se emplea el

desarrollo de mockups, guías de estilo y assets de la aplicación para finalmente construir con esto los mapas de navegación.

De forma similar, se definirá el patrón arquitectónico que mejor se ajuste a la solución planteada, para lograrlo se utilizan metodologías tales como Domain Driven Design, API's, REST o GraphQL, CORS, Event-Driven y Event Sourcing.

3. Disponibilidad de infraestructura HPC

Esta etapa comprende la configuración y puesta a punto de la infraestructura HPC necesaria para el entrenamiento y ejecución de los modelos meteorológicos. Es importante aclarar que para lograrlo es necesario disponer de personal especializado en temas de computación distribuida y expertos en HPC capaces de disponer la infraestructura computacional de acuerdo con los lineamientos de capacidad y velocidad esperada.

4. Desarrollo de la plataforma

Esta etapa contempla la construcción de los elementos levantados en el diseño. Para lograrlo se discriminan dos grandes desarrollos, uno a nivel de Back-end, el cual comprende el desarrollo de la lógica necesaria para la atención de las peticiones, acceso a bases de datos y disponibilizar la información a través de servicios web. Por otro lado, a nivel de Front-end se busca desarrollar la capa de presentación con la cual interactúa el usuario para hacer uso de las funcionalidades de la plataforma. Cabe mencionar que en este punto es importante contratar, configurar e integrar los servicios en nube necesarios para interactuar con la herramienta web, así como la puesta a punto de bases de datos y servicios necesarios para el correcto funcionamiento de la solución.

5. Acompañamiento científico especializado

Esta etapa permite poner a disposición el personal técnico necesario para atender las necesidades de intervención y/o escalamiento de los modelos meteorológicos, haciendo uso de la infraestructura HPC, como estrategia para disminuir el tiempo de procesamiento actual (6 horas).

6. Pruebas de rendimiento

Esta etapa contempla la construcción de scripts de testeo necesario para la realización de las pruebas pertinentes tales como: pruebas unitarias, pruebas de estrés, pruebas de componentes, pruebas de humo, pruebas de integración, pruebas de regresión, pruebas de confiabilidad, detección y corrección de fallos. Es importante aclarar que el debido desarrollo de esta etapa definirá el óptimo funcionamiento de la solución.

7. Implementación en ambiente de producción

Una vez se tiene una versión estable de la solución, se procede a realizar el respectivo despliegue en servidores de producción donde el usuario final podrá interactuar libremente con la herramienta web. Para lograrlo es necesario llevar a cabo la configuración, puesta a punto y documentación de todo el proceso para el despliegue de la solución.

8. Soporte y mejoras

El soporte que se dará para la herramienta web es de 12 meses a partir de la finalización del proyecto, estimada para el 30 de noviembre del 2021. Dicho soporte incluye atención a fallos y posibles eventos en la plataforma web que no se comporten de manera esperada ya sea a nivel de Front-end, Back-end o servicios en la nube.

Por otro lado, el soporte estimado para la solución de HPC se limita a configuraciones y atención de incidentes que puedan presentarse en el proyecto durante la ejecución del mismo en la fecha estimada anteriormente.

9. Transferencia de conocimiento

Una vez se tiene la solución funcionando de manera correcta, es necesario la generación del material necesario para abordar sesiones en las cuales BIOS realiza unos procesos de enseñanza al equipo destinado por el IDEAM sobre el uso y entendimiento de la plataforma. Esto con el objetivo que la herramienta web pueda ser explotada al máximo en temas de monitoreo y seguimiento climatológico.

Equipo de trabajo base

Herramienta web

EQUIPO BASE DE OPERACIÓN	CANTIDAD	DEDICACIÓN	DESCRIPCIÓN
DIRECTOR INGENIERO ESPECIALISTA	1	100%	Cinco (5) años como director o gerente de consultoría de proyectos relacionados con el desarrollo e implementación de software. Estar certificado en Project Management Institute.
INGENIERO TEMÁTICO	3	100%	Cinco (5) años en desarrollo y/o planeación, e implementación de proyectos de meteorología/climatología o el desarrollo o implementación de herramientas de software relacionados con la meteorología/climatología

INGENIERO DE ARQUITECTURA DE SOFTWARE	1	100%	Cinco (5) años en arquitecturas técnicas de soluciones complejas de tecnologías de información, desarrollo de arquitectura de microservicios haciendo uso de plataformas Openshift de Redhat, desarrollo de sistema de información geográficos.
ANALISTA EN SISTEMAS	3	100%	Dos (2) años en: procesos de gestión de bases de datos geográficas y de información geográfica, administración de servidores de mapas y geoservicios, y estándares de información geográficos
EXPERTO EN SEGURIDAD INFORMÁTICA	1	100%	Tres (3) años en procesos de diseño, implementación y mantenimiento de políticas de seguridad integral en la Fábrica de Software.
DISEÑADOR WEB UX/UI	1	100%	Tres (3) años en procesos de diseño de interacción UX/UI, con Wireframes. Estar certificado en Project Management Institute.
DISEÑADOR VISUAL	1	100%	Tres (3) años generando material audiovisual para realizar trasferencia de conocimiento de todas las funciones por parte de los usuarios finales.
COMUNICADOR	1	100%	Tres (3) años generando contenidos para realizar trasferencia de conocimiento de todas las funciones por parte de los usuarios finales.
QA TESTER	1	100%	Tres (3) años realizando procesos de calidad centrada en la prevención y la detección de fallos en la fábrica de software.

Solución de HPC

EQUIPO BASE DE OPERACIÓN	CANTIDAD	DEDICACIÓN
PERFIL 1: Un (1) EXPERTO EN COMPUTACIÓN DISTRIBUIDA	1	100%
PERFIL 2: Un (1) EXPERTO EN HPC	1	100%
PERFIL 3: Un (1) EXPERTO EN MODELOS DE PREDICCIÓN METEOROLÓGICOS	1	100%
PERFIL 4: Un (1) CIENTÍFICO DE DATOS	1	100%

PERFIL 5: Un (1) EXPERTO EN VISUAL ANALYTICS	1	100%

Aclaraciones

El soporte mencionado en la propuesta cubre la atención en horarios de lunes a viernes de 8am - 12m y de 2pm-6pm.

Entregables

- Plan de trabajo que contenga los requerimientos técnicos para cada proyecto a ejecutar, que incluya todas las actividades consideradas, cronograma para llevar el control y los hitos y responsables.
- Definición de la Estructura de Desglose del Trabajo (EDT).
- Formatos definidos para establecer el modelo de gobernanza del proyecto: Gestión de cambios, canales de comunicación y formatos para entrega de documentación asociada al proyecto.
- Documento de levantamiento de requerimientos funcionales y no funcionales.
- Documento de diseño de la solución web incluyendo requerimientos, casos de uso detallando el flujo de información que seguirá la solución, mapas de navegación y propuesta visual de la solución.
- Documento de arquitectura de la solución.
- Manual de configuración del ambiente de pre-producción.
- Documentación y código fuente de la solución tanto a nivel de Back-end como a nivel de Front-end, describiendo la funcionalidad de cada clase, función y atributo
- Documentación de configuraciones de bases de datos, diccionarios de datos y configuración de servicios contratados en la nube.
- Manual de usuario de la plataforma.
- Scripts de testeo debidamente comentados.
- Documento de resultados de pruebas unitarias y QA, detallando hallazgos y puntos de mejora de la solución.
- Documento con respuesta a los hallazgos encontrados en las pruebas unitarias.
- Documento final (incluyendo posibles cambios) con configuraciones necesarias en ambiente de producción para despliegue de la plataforma.
- Actas de reunión de sesiones de acompañamiento al equipo del IDEAM en relación al manejo y funcionamiento de la solución.
- Bitácora de fallos y atención a fallos de la solución web.
- Informe de cambios y mejoras de la solución.
- Documento con plan para sesiones de transferencia de conocimiento. al equipo del IDEAM.

- Acta de asistencia a las sesiones de transferencia de conocimiento y capacitaciones técnicas y funcionales al equipo del IDEAM.
- Informe mensual de uso y disponibilidad de HPC.
- Reporte de características del cluster.
- Hojas de vida del personal científico asociado. tanto para herramienta web como para la solución de HPC.
- Informe de parametrización genérica de la infraestructura HPC (SO, Librerías, SW conexo).
- Bitácora de fallos y atención en fallos HPC.
- Canales y protocolos de comunicación para acceso a soporte técnicos HPC.

Presupuesto y EDT

Herramienta de software

COMPONENTES		ÍTEM	DESCRIPCIÓN	APORTES BIOS	APORTES IDEAM	VALOR TOTAL
1	***Definición de plan de operativo, herramientas y estrategias para seguimiento y gobernabilidad del proyecto***	1,1	Definición de equipo de trabajo y jerarquía para escalamiento de solicitudes	$ 0	$ 57.782.258	$ 57.782.258
		1,2	Modelo operativo y plan de trabajo			
		1.2.1	Desarrollo de EDT con objetivos, procesos, actividades, hitos, entregables, tiempos de entrega y responsables asociados			
		1,3	Identificación de stakeholders en el desarrollo de la herramienta web y definición de roles.			
		1,4	Definición del modelo de gobernanza para la herramienta web enmarcado en un modelo de ingeniería como RUP y o un marco de desarrollo ágil.			
		1,5	Definición de canales de comunicación, recursos para la gestión y seguimiento al diseño y desarrollo de la herramienta web.			
		1,6	Definición de plantillas para la entrega de la documentación técnica y administrativa a lo largo del proyecto.			
		1,7	Documentación relacionada con la definición del plan de trabajo, modelos de gobernanza, EDT y demás procesos desarrollados en el marco de los lineamientos para seguimiento técnico y administrativo del proyecto.			

2	***Diseño de la Plataforma***	2,1	Definición de las necesidades, recursos y datos disponibles para el desarrollo de la herramienta.	$ 0	$ 140.756.134	$ 140.756.134
		2,2	Levantamiento de requerimientos funcionales y no funcionales			
		2,3	Creación de casos de uso/historias de usuarios.			
		2,4	Creación de mockups y mapas de navegación basado en los paradigmas UX/UI.			
		2,5	Definición de arquitectura de la solución.			
		2.5.1	Definición de estrategia para integración de las capas de la solución sobre openshift			
		2.5.2	Definición de estrategia para la integración de servicios tales como FUSE y MQ para la gestión de colas de mensajería.			
		2.5.3	Definición de arquitectura necesaria para despliegue de los microservicios requeridos, bases de datos y geoserver.			
		2.6	Documento que contenga el levantamiento de requerimientos técnicos para el desarrollo de la plataforma, el cual debe incluir el tipo y roles de usuario, así como diseños y mapas de navegación de la solución			

		2.7	Documentación de la arquitectura con el detalle de integración con otros sistemas y/o soluciones basadas en la nube			
3	*Desarrollo de la plataforma*	3.1	Aprovisionamiento de ambientes de desarrollo necesarios para el desarrollo de la solución	$ 0	$ 288.252.748	$ 288.252.748
		3.2	Maqueteado de la solución basado en la capa de colores, diseños y navegación definida en la fase de "Diseño de la plataforma"			
		3.3	Ajustes e implementaciones a nivel de bases de datos para la creación de procesos, paquetes, funciones, triggers y demás recursos necesarios para cumplir con los requerimientos de la solución			
		3.4	Desarrollo de API's de servicios para comunicación entre bases de datos y herramienta web			
		3.5	Acople de desarrollo web con diferentes navegadores de acuerdo a los requerimiento de la entidad contratante			
		3.6	Herramienta web con capa de administración exclusiva para los roles permitidos			
		3.7	Integración de vistas asociadas a la migración y manejo de informes asociados a integraciones con otros módulos o sistema definidos en la arquitectura			
		3.8	Generación de módulos para administración y visualización de logs almacenados en base de datos.			

		3.9	Documentación de la solución			
		3.9.1	Diccionarios de datos, motores de base de datos implementados y configuraciones realizadas.			
		3.9.2	Documentación de configuración de las soluciones externas contratadas para integrarse con la herramienta web.			
		3.9.3	Referenciación y comentario del código fuentes de cada sección desarrollada tanto a nivel de back-end como de Front-end. Describiendo la funcionalidad de cada clase, función y atributo			
		3.9.4	Creación de manuales técnicos y manuales de usuario detallando la construcción, funcionalidad y bondades de la solución			
4	***Pruebas de rendimiento***	4.1	Aislamiento de la solución en entorno controlado (entorno de pruebas) para el testeo de la solución	$ 0	$ 109.261.385	$ 109.261.385
		4.1.1	Generación de scripts para el testeo de las diferentes secciones de la solución			
		4.1.2	Pruebas de estrés			
		4.1.3	Pruebas QA			
		4.1.4	Documento con informe detallado de las pruebas unitarias diseñadas, resultados y hallazgos generados			

		4.2	Depuración y corrección de bugs presentes en la herramienta web			
		4.3	Acompañamiento técnico al equipo del IDEAM en el proceso de despliegue de la solución en ambientes de pre-producción			
		4.4	Informe de acciones tomadas para la corrección de eventos encontrados en las pruebas unitarias.			
5	*Implementación en ambiente de producción*	5.1	Aprovisionamiento de ambientes de producción y escalamiento de recursos de las diferentes soluciones de nube contratadas	$ 0	$ 102.139.143	$ 102.139.143
		5.2	Acompañamiento al equipo del IDEAM para el despliegue final de la solución			
		5.3	Apoyo en el tiempo acordado, a la operación de la solución			
		5.4	Monitorear, en el tiempo acordado, el funcionamiento de la solución y atender incidentes que puedan generarse a nivel técnico respetando las siguientes reglas de intervención: 1. Incidente crítico: Debe estar solucionado en máximo 4 horas. La solución debe permitir operar aunque no sea la solución definitiva. 2. Incidente funcional: Debe estar solucionado en máximo 2 días hábiles. 3. Incidente menor: Debe estar solucionado en máximo 5 días hábiles.			

		5.5	Actualización de toda la documentación de la solución de acuerdo con los cambios/actualizaciones realizadas para el correcto funcionamiento de la solución			
		5.6	Documentación de despliegue de la solución			
6	*Soporte y mejoras*	6.1	Garantizar soporte por al menos un (1) año, a través de medio telefónico, en sitio, correo electrónico, o mesa de ayuda.	$ 0	$ 212.769.151	$ 212.769.151
		6.1.1	Garantizar la disponibilidad de licencias y servicios externos necesarios para el correcto funcionamiento de la plataforma			
		6.1.2	Atención a fallos y mejoras acordes a los lineamiento estipulados en los levantamiento de requerimientos iniciales			
		6.2	Establecer actividades de mantenimiento preventivos sobre la solución tales como actualización de librerías y demás recursos necesarios para salvaguardar la seguridad de datos, accesos y consumo de información.			
7	*Transferencia de conocimiento*	7.1	Implementar estrategia de transferencia de conocimiento sobre la solución desarrollada, al equipo dispuesto por el IDEAM	$ 0	$ 116.555.437	$ 116.555.437
		7.1.1	Capacitaciones técnicas al equipo dispuesto por IDEAM para cada sección desarrollada en la herramienta web			

		7.1.2	Capacitaciones funcionales al equipo dispuesto por IDEAM para cada sección desarrollada en la herramienta web			
		7.1.3	Sesiones para transferencia de conocimiento sobre procesos de aprovisionamiento, acople, despliegue y atención de fallos de la herramienta web			
			Total	$ 0	$ 1.027.516.255	$ 1.027.516.255

Solución de HPC

COMPONENTES			APORTES BIOS	APORTES IDEAM	VALOR TOTAL
1	Disponibilidad infraestructura HPC - hora/Core	Los componentes del clúster del cómputo científico ofertados para el servicio como: procesadores, memoria RAM, discos duros, red, NAS, deben hacer parte de una arquitectura de computación de alto desempeño o HPC.	$ 494.962.650	$ 89.360.376	$ 584.323.026
		Capacidad de procesamiento en GPU mínimo con 10 tarjetas NVIDIA P100 o superiores.			

		Infraestructura HPC con mínimo de 22 nodos de disponibilidad, con discos duros de estado sólido – SSD y/o discos mecánicos de 10Krpm SAS o superiores.			
		Capacidad de almacenamiento de 240 TB o superior después de arreglos de discos.			
		Infraestructura de computación basada en un modelo de computación híbrida de distintas tecnologías (Xeon, Phi) o superior.			
		Total de Cores equivalente a 1.100 o superior.			
		Total de Cores con Hyper-threading equivalente a 2.200 o superior			
		Enlace de acceso a internet a través de un canal simétrico con 250MB en alta disponibilidad, y, conexión y acceso a la red RENATA.			
		Utilización de gestores de tareas tales como SLURM para la administración de procesos HPC.			
		Protocolo ligero de acceso a directorios (LDAP) como gestor de acceso a servicios de directorio ordenado y distribuido.			
		Contar con planta eléctrica que permita tener autonomía hasta por dos días sin intervención (según capacidad del tanque de combustible) con respecto a alimentación eléctrica o superior.			

		La infraestructura HPC del clúster debe permitir la comunicación entre nodos de procesamiento científico a una velocidad de 56 Gbps o superior.			
2	Acompañamiento Científico Especializado	Contar con un equipo científico con formación en universidades acreditadas en alta calidad, internacional y en diferentes áreas de computación y ciencias de investigación aplicada para el análisis, interpretación, procesamiento y visualización de datos científicos. Contar con personal especializado para el acompañamiento y administración de recursos HPC para la ejecución de modelos predictivos y paralelización de procesos. Contar con personal técnico con conocimiento en el uso de diferentes librerías para el manejo y gestión de procesos de HPC como lo son SLUM y acceso a través de SSH. Contar con un equipo de desarrollo de software para fortalecer los proyectos con especialidad en Back-end, Front-end y diseño de interfaces en múltiples lenguajes de programación (C, C++, Python, Java, R, Matlab, Fortran, D3JS, Angular, …). Contar con un grupo de investigación categoría B o superior, de acuerdo con la clasificación de Minciencias.	$ -	$ 592.908.359	$ 592.908.359

		identificar y definir las plataformas de software, librerías, compiladores, modos de compilación, versiones, tipo de máquinas virtuales, cantidad de procesadores y número de nodos del clúster HPC, requeridos para el procesamiento de modelos mesoescalares de pronósticos, asimilación de datos para el análisis óptimo y eficiente de datos meteorológicos, y en todo caso, reduciendo los tiempos de procesamiento actuales según los tiempos obtenidos en el Datacenter del IDEAM para la corrida de mayor duración (Bogotá con 3 dominios y un tiempo de procesamiento de 5 horas y media).			
		Contar con personal científico certificado por una entidad nacional o internacional en ciencia de datos con experiencia superior a 2 años en la ejecución de proyectos relacionados con tecnologías emergentes.			
		Contar con personal técnico en desarrollo de portales de visualización enfocado en la transmisión de resultados que utilizan grandes fuentes de información.			
		Garantizar un PERFIL 1: Un (1) EXPERTO EN COMPUTACIÓN DISTRIBUIDA			
		Garantizar un PERFIL 2: Un (1) EXPERTO EN HPC			
		Garantizar un PERFIL 3: Un (1) EXPERTO EN MODELOS DE PREDICCIÓN METEOROLÓGICOS : experiencia en los temas de simulación meteorológica, modelado del clima, uso del modelo WRF y experiencia de investigación			

		Garantizar un PERFIL 4: Un (1) CIENTÍFICO DE DATOS			
		Garantizar un PERFIL 5: Un (1) EXPERTO EN VISUAL ANALYTICS			
3	**Brindar soporte de servicio con disponibilidad 24/7**	Disponibilidad de uso y capacidad total de la infraestructura HPC y HPDA de por lo menos 10 horas diarias o superior, los 7 días de la semana, en diferentes franjas horarias (Una vez se estabilice la parametrización y/o configuración del software que corre el modelo WRF).	$ -	$ 186.002.541	$ 186.002.541
		Disponer de una línea telefónica que será el punto de contacto con el IDEAM sin límites horarios y los siete días de la semana.			
		El horario de soporte técnico será mínimo de 7 a.m. a 7:00 p.m, durante los 7 días de la semana y con horarios flexibles y extendidos bajo demanda en el caso de requerirse corridas de modelos en horario extendido. Acceso de soporte a través de atención telefónica y/o correo electrónico, y acceso remoto.			
		En el caso de que el IDEAM detecte fallas, el contratista debe llevar a cabo un diagnóstico inicial para determinar la(s) causa(s), y los tiempos de respuesta estarán discriminados en los ANS del contrato.			
		Corregir las fallas en el servicio dentro del plazo que le señale el supervisor del contrato.			
		Total	**$ 494.962.650**	**$ 868.271.276**	**$ 1.363.233.926**

NOTA: La contrapartida de BIOS se efectuará en especie con disponibilidad de infraestructura HPC medida en horas/core.

Resumen presupuestal

CONCEPTO	CONTRAPARTIDA BIOS (ESPECIE)	CONTRAPARTIDA IDEAM (EFECTIVO)	TOTAL
Fábrica de Software	$0	$1.027.516.255	$1.027.516.255
Solución HPC	$494.962.650	$868.271.276	$1.363.233.926
TOTAL	**$494.962.650**	**$1.895.787.531**	**$2.390.750.181**

Forma de desembolso de aportes

El pago de los aportes se realizará con cada entregable así:

PLAN DE DESEMBOLSOS				
MES	PORCENTAJE	VALOR	ENTREGABLES	LEGALIZACION

MES 1	40%	$758.315.012,4	- Firma de Acta de Inicio - Plan operativo del proyecto (Herramienta PMI), donde se incluya todas las actividades consideradas, cronograma para llevar el control y los hitos y responsables. - Documento formal detallando la Estructura de Desglose de Trabajo (EDT). - Documento con la definición de canales y protocolos de comunicación para acceso a soporte técnicos HPC. - Formatos definidos para establecer modelo de gobernanza del proyecto, donde se incluyan formatos para gestión de cambios, canales de comunicación y formatos para entrega de documentación asociada al proyecto. - Hojas de vida del personal científico asociado y del personal de desarrollo contemplado para el desarrollo de la herramienta de software.	
MES 3	30%	$568.736.259,3	- Documento de levantamiento de requerimiento funcionales y no funcionales. - Documento de diseño de la solución web incluyendo requerimientos, casos de uso, flujo de información, mapas de navegación y propuesta visual. - Documento de arquitectura de la solución. - Informe mensual de uso y disponibilidad de HPC. - Bitácora de fallos y atención en fallos HPC.	- Al finalizar el mes 3 se legalizará el 100% del desembolso mes 1.
MES 5	30%	$568.736.259,3	- Documentación y código fuente de la solución. - Manual de configuración del ambiente de preproducción. - Documentación de configuraciones de bases de datos, diccionarios de datos y configuración de servicios contratados en la nube privada. - Manual de usuario de la plataforma. - Scripts de testeo debidamente comentados. - Documento de resultados de pruebas unitarias y QA, detallando hallazgos y puntos de mejora de la solución.	- Al finalizar el mes 5 se legalizará 100% del mes 3 y 5

			- Documento con respuesta a los hallazgos encontrados en las pruebas unitarias - Documento final (incluyendo posibles cambios) con configuraciones necesarias en ambiente de producción para despliegue de la plataforma. - Actas de reunión de sesiones de acompañamiento al equipo del IDEAM y actores locales regionales en relación con manejo y funcionamiento de la solución. - Bitácora de fallos y atención a fallos de la solución web. - Informe de cambios y mejoras de la solución - Documento con plan para sesiones de transferencia de conocimiento. al equipo del IDEAM. - Acta de asistencia a las sesiones de transferencia de conocimiento y capacitaciones técnicas y funcionales al equipo del IDEAM. - Informe mensual de uso y disponibilidad de HPC. - Reporte de características del clúster. - Bitácora de fallos y atención en fallos HPC. - Informe de parametrización genérica de la infraestructura HPC (SO, Librerías, SW conexo). - Informe describiendo los canales y protocolos de comunicación implementados para el acceso a soporte técnico.	
TOTAL	**100%**	**$1.895.787.531**		

Lugar de la prestación de la solución

Manizales, Caldas. Colombia

Vigencia de la propuesta

10 (diez) días a partir de la fecha de la propuesta

Bibliografía

[1] Telefónica, F. (2019). Informe "Sociedad digital en España 2018". Recuperado de: https://www. fundaciontelefonica. com/arte_cultura/sociedad-de-la- informacion/sdie-2017.

[2] Cook, J. et al. Cuantificar el consenso sobre el calentamiento global antropogénico en la literatura científica. Environ. Lett. 8, 024024 (2013).

[3] Wallemacq, P., Abajo, I. & McLean, D. Pérdidas Económicas, Pobreza y Desastres: 1998-2017 (UNISDR, CRED, 2018).

[4] Informe de análisis meteorológico, climático y catastrófico GDM05083 (AON, 2019).

[5] The Global Risks Report 2020 (Foro Económico Mundial, 2020).

[6] Bauer, P., Thorpe, A. y Brunet, G. La revolución silenciosa de la predicción meteorológica numérica. Naturaleza 525, 47–55 (2015).

[7] Hausfather, Z., Drake, H. F., Abbott, T. & Schmidt, G. A. Evaluando el desempeño de proyecciones de modelos climáticos pasados. Geofísica. Lett. 47, e2019GL085378 (2020).

[8] Sillmann, J. et al. Comprensión, modelado y predicción de los extremos climáticos y climáticos: desafíos y oportunidades. Tiempo Clim. Extremos 18,65-74 (2017).

[9] Platzman, G. W. Los cálculos de LAIAC de 1950— puerta de entrada a la predicción meteorológica numérica. El toro. Amer. Meteorol. Soc. 60, 302–312 (1979).

[10] Leutbecher, M. & Palmer, T. N. Ensemble pronosticando. J. Comput. Phys. 227, 3515-3539 (2008).

[11] Zhu, Y., Toth, Z., Wobus, R., Richardson, D. & Mylne, K. El valor económico de las previsiones meteorológicas basadas en conjuntos. El toro. Amer. Meteorol. Soc. 83, 73-84 (2002).

[12] Palmer, T. & Stevens, B. El desafío científico de entender y estimar el cambio climático. Pr. Natl Acad. Ciencia EE.UU. 116, 24390-24395 (2019).

[13] Brunet, G. et al. Colaboración de las comunidades meteorológicas y climáticas para avanzar en la predicción subseasonal a estacional. El toro. Amer. Meteorol. Soc. 91, 1397-1406 (2010).

[14] DYAMOND: la dinámica de la circulación general atmosférica modelada en dominios no hidrostáticos. Planeta Tierra Prog. La ciencia. 6, 61 (2019).

[15] Wedi, N.P. et al. Una línea de base para simulaciones meteorológicas y climáticas globales a resolución de 1 km. Modelo J. Adv. Earth Syst. 12, e2020MS002192 (2020)

[16] Williamson, D. L. La evolución de los núcleos dinámicos para los modelos atmosféricos globales. J. Meteorol. Soc. Jpn Ser. II 85, 241-269 (2007).

[17] McFarlane, N. Parametrizaciones: representar procesos clave en los modelos climáticos sin resolverlos. Wiley Interdiscip. Reverendo Clim. Cambio 2, 482–497 (2011).

[18] Flato, G.M. Modelos del sistema de la Tierra: una visión general. Wiley Interdiscip. Reverendo Clim. Cambio 2,783–800 (2011).

[19] Steppeler, J., Hess, R., Schättler, U. & Bonaventura, L. Revisión de métodos numéricos para modelos de predicción meteorológica no hidrostática. Meteorol. Phys. 82, 287-301 (2003).

[20] Mengaldo, G. et al. Estrategias actuales y emergentes de integración del tiempo en la predicción numérica global del clima y el clima. Comput. Metanfetamina. Eng. 26, 663-684 (2019).

[21] Teixeira, J., Reynolds, C. A. & Judd, K. Time step sensitivity of nonlinear atmospheric models: convergencia numérica, crecimiento de errores de truncamiento y diseño de conjuntos. J. Atmos. 64, 175-189 (2007).

[22] Mozdzynski, G., Hamrud, M., Wedi, N., Doleschal, J. & Richardson, H. 2012 SC Companion: High Performance Computing, Networking Storage and Analysis 652–661 (2012).

[23] Sanan, P., Schnepp, S.M. & May, D. A. Métodos subespaciales Krylov flexibles. Siam J. Sci. Comput. 38, C441–C470 (2016). [24] CDI-pio & XIOS I/O Servidores compatibilidad con modelos de clima HR TR/CMGC/17/52 (CERFACS, 2017).

[25] 10M-core solucionador totalmente implícito escalable para dinámicas atmosféricas no hidrostáticas. En SC16: Conferencia Internacional proc. para computación de alto rendimiento, redes, almacenamiento y análisis 57–68 (2016).
[26] Mozdzynski, G., Hamrud, M. & Wedi, N. Un espacio de dirección global dividido que implementa el Centro Europeo de Pronósticos Meteorológicos de Medio Alcance Sistema Integrado de Previsión. Int. J. Alto rendimiento. Comput. Appl. 29, 261-273 (2015).

[27] Schulthess, T.C. Programación revisitada. Nat. 11, 369-373 (2015).

[28] Deconinck, W. et al. Atlas: una biblioteca para la predicción meteorológica numérica y el modelado climático. Comput. Comm. 220, 188-204 (2017).

[29] Trèmolet, Y. El esfuerzo conjunto para la integración de la asimilación de datos (JEDI). JCSDA Q. 66, 1–5 (2020).

[30] Hill, C., DeLuca, C., Balaji, V., Suárez, M. & da Silva, A. La arquitectura del marco de modelado del sistema de la Tierra. Comput. Eng. 6, 18–28 (2004).

[31] Inteligente, S., Quintino, T. & Raoult, B. Un almacén de objetos distribuidos de alto rendimiento para la predicción meteorológica numérica a gran escala y el clima. En la Conferencia 1–11 (2019) de la Plataforma para la Computación Científica Avanzada.

CENTRO DE BIOINFORMÁTICA
Y BIOLOGÍA COMPUTACIONAL
Conocimiento e innovación para el desarrollo

Otros servicios BIOS

BIOS prospection.	BIOS consulting.	BIOS data.
BIOS industrial.	BIOS sensorial.	

Contacto
Carolina Ramírez Corrales
Carolina.ramirez@bios.co
Líder Mercadeo y Comunicaciones BIOS

MIGUEL TRUJILLO LONDOÑO
DIRECTOR EJECUTIVO
BIOS