



DOI: <http://dx.doi.org/10.23857/dc.v8i2.2773>

Ciencias Técnicas y Aplicadas
Artículo de Investigación

*Análisis de un sistema de gestión integral de monitoreo de nodos de
Telecomunicaciones de la COAC Jardín Azuayo*

*Analysis of a comprehensive management system for monitoring
Telecommunications nodes of the COAC Jardín Azuayo*

*Análise de um sistema de gestão integral para monitorar os nós de
Telecomunicações do COAC Jardín Azuayo*

Paúl Andrés Valdivieso-Avilés ^I
paul.valdivieso.01@est.ucacue.edu.ec
<https://orcid.org/0000-0001-7684-649X>

Javier Bernardo Cabrera-Mejía ^{II}
jcabreram@ucacue.edu.ec
<https://orcid.org/0000-0003-2027-0211>

Andrés Sebastián Quevedo-Sacoto ^{III}
asquevedos@ucacue.edu.ec
<https://orcid.org/0000-0001-5585-0270>

Correspondencia: paul.valdivieso.01@est.ucacue.edu.ec

***Recibido:** 29 de marzo del 2022 ***Aceptado:** 21 de abril de 2022 ***Publicado:** 26 de mayo de 2022

- I. Estudiante de la Maestría en Tecnologías de la Información, Universidad Católica de Cuenca, Cuenca, Ecuador.
- II. Jefe de Investigación, Universidad Católica de Cuenca, Cuenca, Ecuador.
- III. Docente investigador, Universidad Católica de Cuenca, Cuenca, Ecuador

Resumen

Se presenta una nueva arquitectura de HDSM (Sistema de Gestión de Nodos de Sensores Distribuidos Basado en Hadoop) para la gestión de nodos de sensores distribuidos mediante el marco Hadoop MapReduce y el sistema de archivos distribuido (DFS). Ofrece varias formas eficientes de recopilar datos de sensores y administrar múltiples nodos de sensores mediante el lanzamiento de aplicaciones específicas de reducción de mapas en nodos de sensores que cargan datos de nodos de sensores en DFS y recuperan datos de sensores periódicamente de DFS. Proporciona un esquema de gestión flexible para el nodo del sensor mediante la reconfiguración del firmware o la actualización de configuraciones y formatos de datos de los nodos del sensor basados en el marco de MapReduce. Además, se logra una arquitectura transparente y tolerante a fallas al explotar varias características cruciales de Hadoop. Los resultados experimentales muestran que el sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo tiene un rendimiento eficiente y estable.

Palabras Claves: Archivo; Hadoop; MapReduce; nodo; sensor.

Abstract

A new HDSM (Hadoop-based Distributed Sensor Management) architecture is presented for managing distributed sensor nodes using the Hadoop MapReduce framework and Distributed File System (DFS). It offers several efficient ways to collect sensor data and manage multiple sensor nodes by launching specific map reduction applications on sensor nodes that upload sensor node data to DFS and periodically retrieve sensor data from DFS. In addition, it provides a flexible management scheme for the sensor node by reconfiguring the firmware or updating the configurations and data formats of the sensor nodes based on the MapReduce framework. Furthermore, a transparent and fault-tolerant architecture is achieved by exploiting several crucial features of Hadoop. The experimental results show that the comprehensive management system for monitoring Telecommunications nodes of COAC Jardín Azuayo has efficient and stable performance.

Keywords: File; Hadoop; MapReduce; node; sensor.

Resumo

Uma nova arquitetura HDSM (Distributed Sensor Node Management System baseado em Hadoop) é apresentada para gerenciar nós de sensores distribuídos usando a estrutura Hadoop MapReduce e o Distributed File System (DFS). Ele oferece várias maneiras eficientes de coletar dados de sensores e gerenciar vários nós de sensores, iniciando aplicativos específicos de redução de mapa em nós de sensores que carregam dados de nós de sensores para o DFS e recuperam periodicamente dados de sensores do DFS. Ele fornece um esquema de gerenciamento flexível para o nó sensor reconfigurando o firmware ou atualizando as configurações e formatos de dados dos nós sensores com base na estrutura MapReduce. Além disso, uma arquitetura transparente e tolerante a falhas é alcançada explorando vários recursos cruciais do Hadoop. Os resultados experimentais mostram que o sistema de gestão integral de monitoramento dos nós de Telecomunicações do COAC Jardín Azuayo tem um desempenho eficiente e estável.

Palavras-chave: Arquivo; Hadoop; MapReduce; nó; sensor.

Introducción

En un entorno informático ubicuo con redes de sensores cableados/inalámbricos, la recopilación, el control y la gestión eficientes de los datos de los sensores son muy importantes para varios análisis de datos útiles (Yang, 2022). Se pueden implementar varios nodos de sensores, como CCTV, sensores de temperatura o puntos de acceso Wi-Fi con procesadores adicionales y dispositivos de almacenamiento para guardar los datos de detección localmente y conectarse entre sí a través de redes como Ethernet. Estas redes de sensores necesitan enfoques sofisticados para lograr una recopilación masiva de datos de sensores y una gestión rentable de los nodos de sensores distribuidos (Mendieta, 2022).

MapReduce es un marco útil en una amplia gama de aplicaciones que incluyen búsqueda paralela distribuida, indexación, agrupamiento y clasificación para varios tipos de datos y reemplaza el antiguo software ad hoc que realiza los diversos análisis para grandes conjuntos de datos (Abdolazimi, 2022). Además, el modelo de programación MapReduce ofrece una manera más fácil de crear aplicaciones paralelas distribuidas que se ejecutan en múltiples nodos informáticos para procesar Big data. Por lo tanto, se ha adaptado a varios entornos informáticos, como sistemas multiprocesador y entornos de nube. Hasta el momento, existen varias implementaciones para el algoritmo MapReduce y su marco de ejecución, como Hadoop. Pueden integrarse con varios sistemas de archivos distribuidos (DFS) y

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo

ofrecen una gestión general de todo el proceso para iniciar la aplicación MapReduce, integrar varios servidores y gestionar las comunicaciones y las transferencias de datos entre componentes. Proporcionan una arquitectura redundante y tolerante a fallas a través de su arquitectura distribuida, replicación de datos y mecanismo de latido.

Sin embargo, los enfoques anteriores se centran en acumular datos de registro de aplicaciones generadas desde otros servidores conectados a través de una red cableada en un repositorio basado en DFS y procesarlos mediante MapReduce (Agudelo, 2022). En esta investigación, se considera la red de sensores del sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo como un clúster de MapReduce a gran escala y se presenta una nueva plataforma de gestión de nodos de sensores, llamada HDSM (sistema de gestión de nodos de sensores distribuidos basado en Hadoop) para la gestión de nodos de sensores distribuidos mediante la explotación del marco MapReduce y DFS y el lanzamiento de varias aplicaciones para ejecutar configuraciones de carga, recuperación, vaciado, monitoreo y actualización de datos de sensores dentro del sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo. El cual permite formas eficientes de recopilar datos de sensores y administrar múltiples nodos de sensores mediante el lanzamiento de aplicaciones específicas de MapReduce en nodos de sensores que cargan datos de nodos de sensores en DFS y recuperan datos de sensores periódicamente de DFS. Además, proporciona un esquema de gestión flexible para el nodo sensor mediante la reconfiguración de un firmware o la actualización de configuraciones y formatos de datos de nodos sensores basados en el marco MapReduce. Además, logra una arquitectura transparente y tolerante a fallas al explotar varias características cruciales de Hadoop (Ma, 2022).

Obras relacionadas

La aplicación MapReduce se compone de dos etapas: mapeado para procesar tareas de unidades de granularidad fina, como filtrado, conteo o clasificación, y reductor para resumir operaciones (Regalado Valenzuela, 2022). Hay varias implementaciones para el algoritmo MapReduce y su marco de ejecución, como Google MapReduce, Apache Hadoop y Twitter (Bawankule, 2022), y las entradas y salidas de MapReduce generalmente se almacenan en un sistema de archivos distribuido como GFS y HDFS. Ofrecen la orquestación de todos los procesos de administración para iniciar la aplicación MapReduce, monitorear múltiples nodos, transferir datos y administrar las comunicaciones entre componentes (Soledispa Peña, 2022).

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo

Apache Hadoop es una implementación de código abierto del marco MapReduce que admite la ejecución de aplicaciones en grandes grupos de hardware básico y un sistema de archivos distribuido denominado sistema de archivos distribuido Hadoop (HDFS) que almacena grandes datos en los nodos informáticos (Zhang, 2022). Está escrito en el lenguaje de programación Java (Lopez, 2021). Un clúster de Hadoop incluye un solo nodo maestro y varios nodos esclavos, como se describe en la Figura 1. Para HDFS, Namenode es un componente maestro que administra las ubicaciones de los archivos y los metadatos de todos los archivos en HDFS. Namenode verifica la actividad de todos los nodos esclavos mediante un mecanismo de latido basado en notificaciones periódicas con mensajes simples y ofrece una interfaz de línea de comandos y una capa API para los clientes. Los archivos de datos reales en HDFS se dividen en varios fragmentos y se administran mediante un componente esclavo denominado nodo de datos. Almacena fragmentos en el disco local de cada nodo trabajador. El motor de MapReduce para ejecutar la aplicación que se compone de un solo rastreador de trabajo en el nodo maestro como un proceso coordinador de trabajos y múltiples rastreadores de tareas en los nodos esclavos como nodos trabajadores. Después de recibir la solicitud de trabajo para iniciar una aplicación MapReduce, el rastreador de trabajos programa y asigna tareas de mapa y reducción a múltiples rastreadores de tareas con conocimiento de la ubicación de los datos y hace que se ejecuten en forma paralela. La mayoría de los resultados de los trabajos de MapReduce se guardan en el HDFS. Tanto el motor de MapReduce como el HDFS están diseñados para que el marco maneje automáticamente las fallas de los nodos. Proporcionan una arquitectura redundante y tolerante a fallas a través de su arquitectura expandible y un esquema de monitoreo único llamado mecanismo de latido. Por lo tanto, su arquitectura es útil para construir un sistema de gestión de nodos de sensores distribuidos.

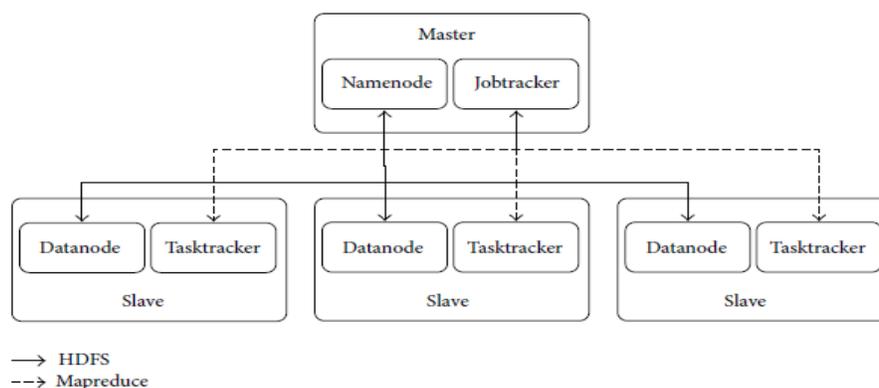


Figura 1. Arquitectura de Hadoop implementada en múltiples nodos; HDFS y MapReduce e interacción entre componentes.

Fuente. Elaboración propia.

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo

Existe una serie de investigaciones para el procesamiento de datos de sensores recopilados mediante computación paralela distribuida (Moreno Segura, 2022). Además, existen varios enfoques para construir grandes sistemas de recopilación de datos en entornos de cuadrícula y nube (Narro Mestanza, 2022). En particular, hay algunos enfoques que están integrados con el marco MapReduce de Hadoop. Apache Flume (Rashid, 2022) es un servicio distribuido para recopilar, agregar y transferir de manera eficiente grandes cantidades de datos de registro de muchas fuentes diferentes a HDFS. Ofrece una arquitectura flexible basada en flujos de datos de transmisión y utiliza un modelo de datos extensible simple que permite la aplicación analítica en línea. Honu es una canalización de recopilación y procesamiento de datos de transmisión a gran escala creada con Hadoop. Recopila y procesa todos los eventos de registro estructurados y no estructurados generados por varias aplicaciones que se ejecutan en Amazon EC2 o EMR y los pone a disposición del usuario a través de Hive (Yusoff, 2022).

En el presente análisis se considera a la red de sensores como un clúster de MapReduce a gran escala y se presenta una plataforma de gestión de nodos de sensores mediante la explotación del marco de MapReduce y DFS y el lanzamiento de varias aplicaciones de MapReduce para ejecutar configuraciones de carga, recuperación, vaciado, monitoreo y actualización de datos de sensores.

Sistema de gestión de nodos de sensores distribuidos basado en MapReduce

Características clave de HDSM.

- Ofrece una manera más fácil de construir un sistema de gestión de nodos de sensores distribuidos. HDSM utiliza la arquitectura y los componentes básicos del marco Hadoop sin corregir sus funciones ni desarrollar componentes específicos para manipular transferencias de datos o comunicaciones. Mediante la integración de múltiples nodos de sensores con recursos de hardware limitados a través de la red y la tecnología Hadoop, HDSM crea fácilmente un enorme grupo de nodos de sensores lógicos en una red de sensores distribuidos.
- Garantiza una gestión automatizada de todo el proceso para integrar múltiples nodos de sensores, iniciar aplicaciones MapReduce y gestionar comunicaciones y transferencias de datos entre DFS y otros componentes a través del marco Hadoop MapReduce. Por lo tanto, puede ahorrar costos de construcción y mantenimiento del sistema de gestión para nodos de sensores distribuidos.

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo

- Se enfoca en analizar los datos de detección recopilados sin preocuparse por fallas, ya que Hadoop tiene un mecanismo interno para manejar las fallas del sistema.
- Permite actualizar fácilmente el software de los nodos de sensores, como los controladores de dispositivos de sensores, el firmware, el formato de datos y las configuraciones de nodos mediante el lanzamiento de la aplicación MapReduce exclusiva sin mucho esfuerzo para manipular cada uno de los nodos de sensores en un proceso complejo.

Arquitectura

HDSM consta del administrador de control de nodos (NCM), el nodo maestro de Hadoop y el esclavo. El NCM y el nodo maestro de Hadoop funcionan como un administrador de HDSM que controla las operaciones generales de HDSM. Una red de sensores se considera como un clúster MapReduce a gran escala. El nodo esclavo Hadoop se implementa en cada uno de los nodos sensores. Consiste en el rastreador de tareas y el nodo de datos e interactúa con el nodo maestro de Hadoop, que está compuesto por el rastreador de trabajos y el nodo de nombres. NCM administra y controla el maestro de Hadoop mediante el uso de comandos de Hadoop y media la interacción entre HDFS y sistemas externos, como repositorios o servicios de consumidores de datos. La recuperación de datos de detección acumulados en HDFS y el vaciado de HDFS se realizan mediante la ejecución de comandos de Hadoop. En particular, ofrece una forma eficaz de recuperar grandes datos acumulados en el grupo de nodos de sensores, ya que el tráfico de datos se distribuye entre todos los nodos de sensores (nodos de datos) en la operación de lectura de HDFS. NCM solicita al nodo maestro de Hadoop que ejecute varias operaciones, incluida la carga de datos de detección en HDFS, el vaciado de HDFS y la aplicación de parches al firmware o las configuraciones del nodo del sensor mediante el lanzamiento de las aplicaciones adecuadas. En cada operación, se asigna una sola tarea de mapa en cada nodo de sensor, respectivamente, mediante la asignación de una cantidad de tareas de mapa como número total de nodos de sensor, ya que el programador de Hadoop tiende a asignar tareas a nodos de trabajo inactivos que pueden permitirse que se les asignen nuevas tareas. Por lo tanto, todas las tareas del mapa procesan su función de forma independiente en los nodos sensores en paralelo (Sanabria, 2022).

- NCM solicita al nodo maestro de Hadoop que inicie la aplicación "Uploader" para hacer que todos los nodos sensores carguen sus datos de detección locales en HDFS. La aplicación de carga solo tiene una clase de mapeador sin reductor y crea una cantidad de tareas de mapa idénticas a la cantidad de nodos sensores. Cada tarea de

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín
Azuayo

mapa se asigna y ejecuta en cada nodo y copia sus datos de detección locales en HDFS con nombres de archivos distinguibles. Cada dato de detección local se divide en uno o varios fragmentos de datos y nodo de datos los guarda en HDFS. De esta manera, cada nodo sensor carga sus datos locales en su nodo de datos y nodo de nombres identifica la ubicación de todos los datos. Después de cargarlos en HDFS, los datos de detección pueden estar listos para que mediante el nodo maestro Hadoop y sus clientes los recuperen.

- NCM solicita al nodo maestro de Hadoop que recupere los datos de detección cargados de HDFS y los almacene en el disco local, otros servidores de repositorio o servicios de consumo de datos. Todos los datos de detección acumulados se leen de varios nodos de datos a la vez periódicamente mediante la lectura de varios fragmentos de varios nodos de datos mediante el comando Hadoop.
- NCM solicita al nodo maestro de Hadoop que vacíe el almacenamiento HDFS para retener el espacio disponible en el disco de los nodos del sensor debido al tamaño de almacenamiento limitado del nodo del sensor.
- NCM solicita al nodo maestro Hadoop que actualice el controlador del dispositivo del sensor avanzado o las configuraciones del nodo del sensor y ejecute la aplicación "Updater". La aplicación de actualización no tiene etapa de reducción y la cantidad de tareas de mapas es idéntica a la de los nodos de sensores. NCM proporciona varias líneas de comandos para asignar tareas, que a su vez las ejecuta como tareas secundarias, como actualizar el controlador del sensor o los códigos de firmware a una versión más nueva, cambiar los formatos de datos, editar las configuraciones de los nodos del sensor o incluso ejecutar comandos como reiniciar, calibrar e inicializar de dispositivos sensores (Zhang, 2022).

HDSM realiza las operaciones anteriores periódicamente de acuerdo con la configuración de NCM. Además, ofrece una forma flexible para que los usuarios agreguen más operaciones mediante el desarrollo de sus aplicaciones MapReduce correspondientes, como analizadores de datos simples o convertidores de datos. Si el poder de cómputo de los nodos sensores no es suficiente, el usuario puede ejecutar la migración de datos a otros repositorios para realizar copias de seguridad o sistemas de procesamiento de datos para un análisis de datos de alto rendimiento.

Metodología

Es importante señalar que, tomando en cuenta los parámetros éticos y de confidencialidad de los datos recabados dentro de la investigación, se manejó con cautela y privacidad los direccionamientos de los protocolos de internet (IP), nombres de equipos, identificadores de servicio (SSID) de las redes analizadas y cualquier dato que pueda ser utilizado por terceros para exponer la seguridad de la información.

Para el análisis se procedió a definir los sistemas operativos en los que se va a trabajar dentro del pentest, definiendo los sistemas Kali Linux, Windows 10 y Windows 7 por la estabilidad actual y las herramientas incorporadas en los mismos.

Se aplicó las dos primeras fases del Hacking Ético para reconocer el estado de la infraestructura y las vulnerabilidades existentes en el sistema, para luego realizar un análisis descriptivo del estado de la seguridad de la información dentro de esta organización. Las acciones realizadas se señalan en la Figura 1.



Figura 2. Fases de Análisis de Vulnerabilidades.
Fuente. Elaboración propia.

Resultados

Se ha implementado un prototipo de HDSM y MapReduce dentro del sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo, para operaciones clave utilizando el lenguaje de programación Java y la API de Hadoop. El HDSM se implementó en varios nodos de guardado de información, cada uno de los cuales registra datos de la comunicación interna dentro de la entidad financiera y la comunicación dentro del sistema de gestión integral de monitoreo

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo

de nodos de Telecomunicaciones de la COAC Jardín Azuayo los cuales guardan periódicamente en su disco de almacenamiento local. Los archivos (correos con imágenes incorporadas, grabaciones de llamadas telefónicas internas o fuera de la entidad financiera, solicitudes de acceso al sistema de alta prioridad) tienen varios tamaños, desde 50 MB hasta 800 MB. Se asignó una dirección IP única a cada nodo sensor y todos los nodos sensores estuvieron conectados al administrador HDSM a través de Fast Ethernet.

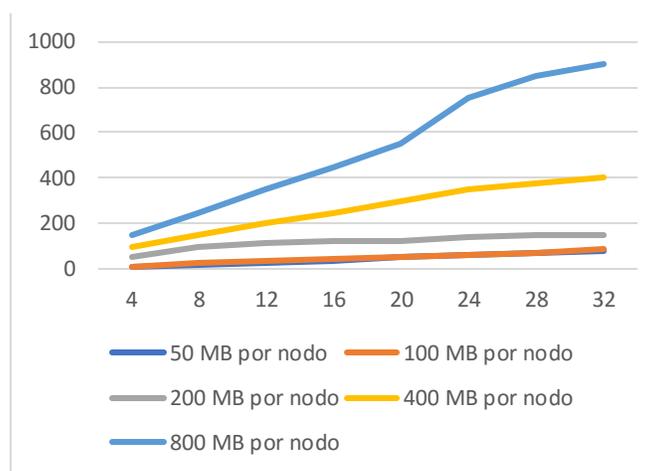


Figura 2. Tiempo promedio de carga de datos de detección con respecto a la cantidad de nodos sensores con varios tamaños de datos.

Fuente. Elaboración propia.

Se evaluó el rendimiento de HDSM para recopilar y recuperar datos de sensores midiendo su tiempo de procesamiento y rendimiento, en la Figura 2 se muestra el tiempo promedio de carga de datos de detección con varios tamaños a HDFS en varios nodos de sensores. El tamaño del bloque del fragmento de datos es de 64 MB y no hay replicación en este experimento. El aumento del tiempo de carga depende de la cantidad de nodos sensores, ya que la cantidad de datos de sensores recopilables en un período es proporcional a la cantidad de nodos sensores. Sin embargo, los rendimientos promedio están relativamente estabilizados, independientemente de la cantidad de nodos sensores o del tamaño de los datos, como se muestra en la Figura 3, la media de los rendimientos en todos los casos es de 26,230 MB/s y la desviación estándar es de 3,80.

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín
 Azuayo

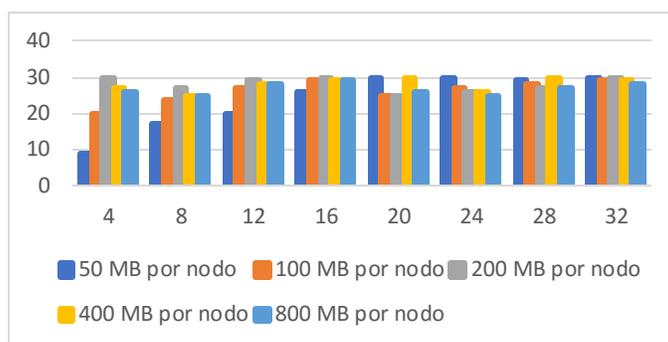


Figura 3. Rendimiento promedio de carga de datos de detección con respecto a la cantidad de nodos de sensores con varios tamaños de datos.
Fuente. Elaboración propia.

Además, se compiló el rendimiento de la operación de carga con varios niveles de replicación en la Figura 4, donde se muestra el rendimiento promedio de la carga de datos en 8 nodos con varios niveles de replicación. El rendimiento de la carga de datos disminuye con el aumento del nivel de replicación, ya que, aunque la replicación de datos de Hadoop se realiza para brindar tolerancia a fallas y disponibilidad, puede causar más tráfico de datos en la red y un menor rendimiento de E/S de HDFS.

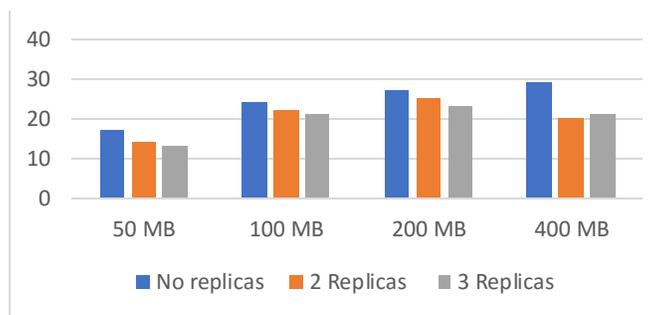


Figura 4. Rendimiento promedio de carga de datos en 8 nodos con respecto a diferentes tamaños de datos con varios niveles de replicación.
Fuente. Elaboración propia.

Se evaluó el rendimiento de recuperación de los datos del sensor acumulados en HDFS, en las Figuras 5 y 6, donde el tamaño del bloque del fragmento de datos es de 64 MB y no hay replicación en este experimento. La Figura 5 muestra que el tiempo promedio de recuperación disminuye a medida que aumenta el número de nodos sensores. La Figura 6 muestra que el rendimiento promedio de la recuperación de datos es más alto que el de la operación de carga y está relacionado con la cantidad

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo

de tamaño de los datos. Además, muestra que el rendimiento medio en todos los casos es de 37,79 MB/s, pero el rendimiento disminuye con el aumento del tamaño de los datos.



Figura 5. Tiempo promedio de recuperación de datos de sensores de HDFS con respecto a la cantidad de nodos de sensores con varios tamaños de datos

Fuente. Elaboración propia.



Figura 6. Rendimiento promedio de recuperación de datos de sensores de HDFS con respecto a la cantidad de nodos de sensores con varios tamaños de datos.

Fuente. Elaboración propia.

En un entorno real, el nodo sensor puede tener un rendimiento inferior al entorno experimental. Por lo tanto, se evaluó el tiempo de carga en los nodos de sensores de bajo rendimiento mediante el uso de máquinas virtuales para compararlas con las de alto rendimiento. La máquina virtual tiene un procesador de un solo núcleo de 2,13 GHz, 512 MB de memoria y 150 GB de disco de bajo rendimiento más económico siendo dispositivos inferiores a \$200 dólares. Las condiciones experimentales son idénticas a las de la Figura 3.

Conclusiones

En este documento, se analiza un sistema de gestión de nodos de sensores distribuidos basado en Hadoop para el esquema de gestión de nodos de sensores distribuidos en el sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín Azuayo, explotando el marco Hadoop MapReduce y el sistema de archivos distribuidos y lanzando varias aplicaciones MapReduce para ejecutar la carga, recuperación, vaciado, monitoreo, y actualizar las configuraciones. Ofrece varias formas eficientes de recopilar datos de sensores y administrar múltiples nodos de sensores mediante el lanzamiento de aplicaciones específicas en nodos de sensores que cargan datos de nodos de sensores en DFS y recuperan datos de sensores periódicamente de DFS.

Además, se proporciona un esquema de gestión flexible para el nodo sensor mediante la reconfiguración de un firmware o la actualización de configuraciones y formatos de datos de nodos sensores basados en el marco MapReduce. Logrando una arquitectura transparente y tolerante a fallas al explotar varias características cruciales de Hadoop. Los resultados experimentales muestran que tiene un rendimiento eficiente y estable. Como trabajo futuro, se plantea se realicen más investigaciones para mejoras e incorporar funcionalidades adicionales con tolerancia a fallas y alto rendimiento.

Referencias

1. Abdolazimi, R. H. (2022). Mapreduce preprocess of big graphs for rapid connected components detection. *In 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*, 112-118.
2. Agudelo, G. F.-G. (2022). Estudio de plataformas de monitoreo para seleccionar la pila tecnológica base de un sistema de analíticas especializado para pruebas de software. *Ingeniería y Competitividad*, 24(1).
3. Bawankule, L. D. (2022). Historical data based approach to mitigate stragglers from the Reduce phase of MapReduce in a heterogeneous Hadoop cluster. *Cluster Computing*, 1-19.
4. Lopez, D., y Marce, C. (2021). Herramienta de soporte para la inspección de la calidad del código de las aplicaciones HPC. *Ciencia y Educación-Revista Científica*, 2(3), 6-14.
5. Ma, J. R. (2022). Improved Hadoop-based cloud for complex model simulation optimization: Calibration of SWAT as an example. *Environmental Modelling & Software*, 105330.

Análisis de un sistema de gestión integral de monitoreo de nodos de Telecomunicaciones de la COAC Jardín
Azúayo

6. Mendieta, V. Z. (2022). Análisis de rendimiento en red de sensores inalámbrica con distintas topologías y cantidades de nodos. *Informática y Sistemas: Revista de Tecnologías de la Informática y las Comunicaciones*, 6(1), 11-21.
7. Moreno Segura, G. J. (2022). Programación paralela y distribuida. arquitecturas paralelas y distribuidas. programación de aplicaciones multiproceso. *TAUJA: Repositorio de Trabajos Académicos de la Universidad de Jaén*.
8. Narro Mestanza, M. (2022). El sistema de gestión de seguridad de la información y la gestión de riesgos en el área informática de una universidad pública, región Cajamarca 2020.
9. Rashid, M. S. (2022). Efficient big data-based storage and processing model in Internet of Things for improving accuracy fault detection in industrial processes. *In Research Anthology on Big Data Analytics, Architectures, and Applications*, 945-957.
10. Regalado Valenzuela, H. (2022). Aplicación cloud para la carga de información crediticia de clientes en el área comercial de una entidad financiera.
11. Sanabria, M., y Méndez-Romero, A. (2022). Tecnologías clave para la transformación digital en las organizaciones. *Transformación digital en las organizaciones*, 31.
12. Soledispa Peña, W. (2022). Implementación de un laboratorio virtual para realizar simulaciones de sistemas de comunicaciones digitales ASK Y FSK. (*Doctoral dissertation, Universidad de Guayaquil. Facultad de Ingeniería Industrial. Carrera de Ingeniería en Teleinformática*).
13. Yang, S. Y. (2022). SoftVideo: Improving the Learning Experience of Software Tutorial Videos with Collective Interaction Data. *In 27th International Conference on Intelligent User Interfaces*, 646-660.
14. Yusoff, H. K. (2022). Big Data-Based Image Handling—A Review of Implementation Using Amazon Web Services. *Multimedia Technologies in the Internet of Things Environment, Volume 3*, 95-106.
15. Zhang, J., y Lin, M. (2022). A comprehensive bibliometric analysis of Apache Hadoop from 2008 to 2020. *International Journal of Intelligent Computing and Cybernetics, (ahead-of-print)*.