ISBN: 978-631-6557-84-1

Recibido:1 de julio de 2026 / Aceptado: 27 de agosto de 2026 / Publicado: 12 de septiembre de 2026

Área: Ingeniería

CAPÍTULO 9

GESTIÓN DE ARQUITECTURAS CLOUD Y ANALÍTICA DE DATOS: OPTIMIZACIÓN DE INFRAESTRUCTURAS TI PARA LA TOMA DE DECISIONES ESTRATÉGICAS

Management of Cloud Architectures and Data Analytics: Optimization of IT Infrastructures for Strategic Decision-Making

Blanco Reyna, Jorge Luis, Fernando Emilio Escudero Vilchez, Pedro Antonio Arévalo Martínez, Carlos Alberto Murillo Salazar, Reyno Mesías Padilla Santos
DOI: 10.55204/pmea.134.c218

Blanco Reyna, Jorge Luis[1] *0000-0002-7477-5824*
Jorge.blanco@upn.edu.pe

Fernando Emilio Escudero Vilchez[2] *0000-0002-3835-8740*
fernando.escudero@autonomadeica.edu.pe

Pedro Antonio Arévalo Martínez[3] *0009-0004-1963-7404*
pedroarevalom@outlook.com

Carlos Alberto Murillo Salazar[4] *0009-0009-7673-4813*
carlos.murillo@uppu.edu.pe

Reyno Mesías Padilla Santos[5] *0000-0001-5278-2049*
rmpadillas@hotmail.com

[1]Universidad Privada del Norte, Lima, Perú.
[2]Universidad Autónoma de Ica, Ica, Perú.
[3]Universidad Inca Garcilaso de la Vega, Lima, Perú.
[4]Universidad Privada de Puno, Puno, Perú.
[5]Universidad Tecnológica del Perú, Lima, Perú.

Resumen: Introducción: El constante crecimiento en el volumen y la heterogeneidad de los datos empresariales ha convertido a las arquitecturas en la nube en el pilar fundamental para la analítica de negocios y la toma de decisiones estratégicas. El presente capítulo examina el impacto de la gestión y optimización de las infraestructuras de tecnologías de la información (TI) en la capacidad analítica organizacional. Objetivo: Analizar las estrategias de optimización y gestión en arquitecturas cloud que permiten maximizar el rendimiento analítico e impactar positivamente en la toma de decisiones estratégicas organizacionales. Métodos: Se realizó una revisión de literatura en IEEE Xplore, ScienceDirect, Springer, O'Reilly, Gartner e IEEE Std, considerando artículos publicados entre 2020 y 2025 en inglés y español. Para el proceso de selección se siguieron las fases del método PRISMA. Se identificaron 319 registros, de los cuales se seleccionaron 16 estudios que conformaron el corpus final y fueron analizados mediante una síntesis temática. Resultados:

ISBN: 978-631-6557-84-1

Se evidencia que la adopción de infraestructuras cloud no garantiza automáticamente la efectividad analítica. La mera migración mediante esquemas directos (lift-and-shift) tiende a replicar las ineficiencias de los sistemas tradicionales, generando sobrecostos presupuestarios imprevistos. Conclusiones: Se concluye que una infraestructura cloud optimizada alinea el rendimiento técnico con la sostenibilidad financiera y la reducción de la incertidumbre en los niveles directivos.

Palabras clave: *Arquitectura Cloud, Analítica de Datos, Data Lakehouse, FinOps, Gobernanza de Datos, Toma de Decisiones Estratégicas, MLOps.*

Abstract: Introduction: The constant growth in the volume and heterogeneity of business data has made cloud architectures a fundamental pillar for business analytics and strategic decision-making. This chapter examines the impact of managing and optimizing Information Technology (IT) infrastructures on organizational analytical capacity. Objective: Analyze optimization and management strategies in cloud architectures that enable the maximization of analytical performance and positively impact strategic organizational decision-making. Methods: A literature review was conducted across IEEE Xplore, ScienceDirect, Springer, O 'Reilly, Gartner, and IEEE Std, considering articles published in English and Spanish between 2020 and 2025. The selection process followed the phases of the PRISMA framework. A total of 319 records were identified, from which 16 studies were selected to form the final corpus and were analyzed through thematic synthesis. Results: It is evidenced that the adoption of cloud infrastructures does not automatically guarantee analytical effectiveness. The mere migration through direct schemes (lift-and-shift) tends to replicate the inefficiencies of traditional systems, generating unexpected budget cost overruns. Conclusions: It is concluded that an optimized cloud infrastructure aligns technical performance with financial sustainability and the reduction of uncertainty at management levels.

Keywords: *Cloud Architecture, Data Analytics, Data Lakehouse, FinOps, Data Governance, Strategic Decision-Making, MLOps.*

1. INTRODUCCIÓN:

1.1. Contexto y problema:

En la economía digital contemporánea, las organizaciones generan y acumulan volúmenes exógenos de información a velocidades sin precedentes. No obstante, la coexistencia de fuentes de datos estructuradas, semiestructuradas y no estructuradas crea un entorno de alta complejidad técnica. La gestión de esta heterogeneidad bajo infraestructuras tradicionales locales (*on-premises*) se ha vuelto inviable debido a sus rigideces estructurales, altos costos de mantenimiento fijos y limitada capacidad de escalabilidad. El problema central radica en que la infraestructura subyacente de TI con frecuencia se transforma en un cuello de botella analítico: los retrasos en la ingesta, el procesamiento ineficiente y la falta de integración impiden que la información procesada alimente oportunamente los procesos directivos de toma de decisiones estratégicas.

1.2. Formulación del problema:

A partir del contexto descrito, se formula la siguiente interrogante principal de investigación: ¿De qué manera la gestión y optimización de las arquitecturas cloud y las infraestructuras de TI potencian las capacidades analíticas de las organizaciones para sustentar la toma de decisiones estratégicas?

1.3. Justificación:

Desde la perspectiva teórica, este estudio sistematiza la evolución de los modelos de al-

ISBN: 978-631-6557-84-1

macenamiento y analítica masiva, articulando conceptos emergentes como Data Lakehouse, FinOps y MLOps en un marco interpretativo unificado. Desde el punto de vista práctico y académico, ofrece a los líderes de tecnología, arquitectos de datos y directores ejecutivos una guía metodológica para reestructurar sus ecosistemas digitales, maximizar el retorno de inversión (ROI) en tecnologías de la nube y garantizar la disponibilidad de datos de alta calidad para la toma de decisiones informadas.

1.4 Objetivos:

Analizar las estrategias de optimización y gestión en arquitecturas cloud que permiten maximizar el rendimiento analítico e impactar positivamente en la toma de decisiones estratégicas organizacionales.

1.5 Objetivos Específicos:

- Examinar la transición conceptual desde los Data Warehouses y Data Lakes tradicionales hacia la consolidación del modelo Data Lakehouse y entornos multicloud.
- Identificar las mejores prácticas en optimización financiera y operativa a través del marco FinOps, tecnologías serverless y mecanismos de gobernanza de datos.
- Evaluar el rol de la analítica en tiempo real, la democratización del dato y las metodologías MLOps como catalizadores del valor de negocio.

1.6 Estado de la cuestión y vacío de conocimiento:

La literatura reciente ha abordado de forma aislada componentes clave como la migración a la nube (Atwal & Dyché, 2020), la eficiencia financiera mediante FinOps (Storment & Fuller, 2023) o la puesta en producción de modelos de Machine Learning (Kreuzberger et al., 2023). Sin embargo, persiste un vacío de conocimiento significativo en la integración holística de estos dominios. Pocas investigaciones articulan de manera integral cómo la ingeniería de infraestructura en el nivel técnico repercute directamente en la agilidad analítica y en la certidumbre directiva al momento de tomar decisiones complejas.

1.7 Marco conceptual:

El marco conceptual de esta investigación se apoya en tres ejes axiales:

- **Arquitectura Nube (Cloud Architecture):** Diseño de componentes de computación, almacenamiento y redes desacoplados y elásticos (IaaS/PaaS).
- **Gobierno y Calidad de Datos:** Conjunto de políticas y controles (incluyendo el linaje del dato) que aseguran la consistencia, integridad y seguridad de la información.
- **Analítica Estratégica:** Transformación sistemática de datos mediante procesamiento batch o streaming para respaldar la ventaja competitiva organizacional.

ISBN: 978-631-6557-84-1

2. METODOLOGÍA:

Para abordar los objetivos planteados, se desarrolló una investigación cualitativa de carácter descriptivo-analítico basada en una revisión sistemática de la literatura y una síntesis conceptual de la disciplina. El procedimiento metodológico se estructuró en cuatro fases principales:

1. **Estrategia de Búsqueda y Fuentes de Información:** Se consultaron bases de datos científicas de alto impacto (IEEE Xplore, ScienceDirect, SpringerLink) y literatura especializada del sector de ingeniería de software (O'Reilly Media, Gartner Research) publicadas predominantemente entre 2017 y 2024.

2. **Criterios de Selección:** Se incluyeron publicaciones arbitradas, libros de texto de referencia y marcos estándares de la industria (IEEE Std 2301-2023) centrados en arquitectura cloud, analítica de datos, FinOps, gobernanza e Inteligencia Artificial.

3. **Criterios de Inclusión y Exclusión:** Se seleccionaron artículos con rigor metodológico demostrado y aportación conceptual relevante sobre infraestructuras cloud modernas. Se excluyeron publicaciones con un enfoque puramente comercial o sin sustento arquitectónico.

4. **Análisis y Sintesis de Datos:** El material recopilado se organizó temáticamente conforme al esquema del capítulo, utilizando análisis de contenido cualitativo para identificar convergencias técnicas, patrones de diseño y marcos de mejores prácticas.

Figura 1 Diagrama de flujo PRISMA del corpus

Nota. Elaboración propia.

3. EVOLUCIÓN DE LAS ARQUITECTURAS DE DATOS EN ENTORNOS CLOUD:

3.1. Data Warehouse:

Históricamente, el

3.1. Data Warehouse:

(DWH) ha constituido el pilar analítico central de la empresa (Inmon, 2005; Kimball & Ross, 2013). Operando bajo esquemas relacionales rígidos (estrella o copo de nieve) y procesos ETL (Extracción, Transformación y Carga), el DWH impone una filosofía de Schema-on-Write. Si bien garantiza un rendimiento excepcional en consultas SQL sobre datos altamente estructurados, presenta limitaciones severas en escalabilidad vertical, altos costos de almacenamiento y falta de flexibilidad para procesar datos no estructurados.

3.2. Data Lake:

Como respuesta a la explosión del Big Data y fuentes no estructuradas, surgió el paradigma del **Data Lake** (Lago de Datos). Fundamentado en sistemas de archivos distribuidos y almacenamiento de objetos de bajo costo (como Amazon S3 o Google Cloud Storage), adopta la metodología ELT (Extracción, Carga y Transformación) y el enfoque Schema-on-Read.

ISBN: 978-631-6557-84-1

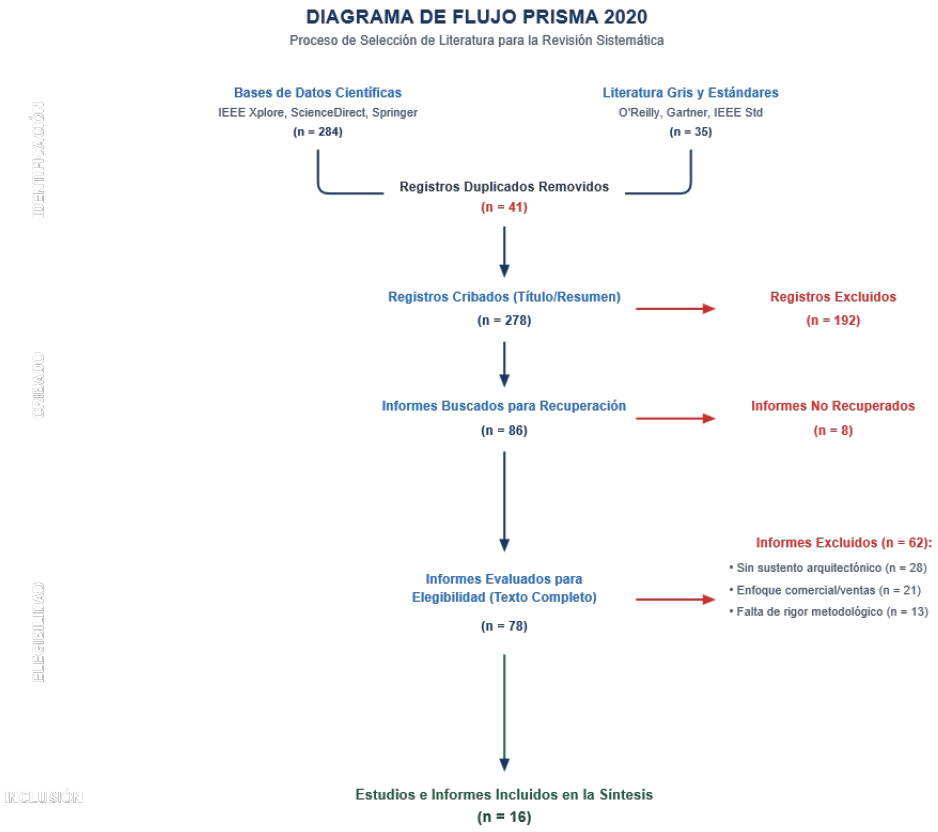


Figura 1: Diagrama El contenido generado por IA puede ser incorrecto.

ISBN: 978-631-6557-84-1

Aunque resolvió el almacenamiento masivo, los Data Lakes carecen de soporte nativo para transacciones ACID, derivando frecuentemente en problemas de calidad e integridad de datos denominados "pantanos de datos"(Data Swamps).

3.3 Data Lakehouse:

El paradigma contemporáneo del **Data Lakehouse** surge para unificar la estructura y fiabilidad transaccional del DWH con la flexibilidad y costo reducido del Data Lake (Armbrust et al., 2021). Mediante capas de metadatos abiertos (como Apache Iceberg, Delta Lake o Apache Hudi), esta arquitectura ofrece:

- Soporte transaccional ACID en almacenamiento de objetos (Zaharia et al., 2020).
- Desacoplamiento total entre los recursos de cómputo y almacenamiento.
- Capacidad para ejecutar analítica descriptiva (BI) y avanzada (Machine Learning) sobre una única copia de los datos.

3.4. Arquitecturas híbridas y Multicloud:

Para mitigar el riesgo de dependencia de un solo proveedor (vendor lock-in) y responder a exigencias de soberanía de datos, las empresas han migrado hacia soluciones **Multicloud e Híbridas** (Rajan & Sridhar, 2022). No obstante, estas arquitecturas imponen retos considerables en la gestión de la latencia de red, la reducción de costos por transferencia o egreso de datos (Data Egress Fees) y la homogeneización de las políticas de seguridad inter-plataforma.

4. ESTRATEGIAS DE OPTIMIZACIÓN Y GOBERNANZA:

4.1. FinOps:

La transición del modelo fijos de gastos de capital (CapEx) a gastos operacionales dinámicos (OpEx) exige la adopción del marco FinOps (Operaciones Financieras en la Nube) (Storment & Fuller, 2023). **FinOps** promueve una responsabilidad financiera compartida mediante un ciclo continuo de tres fases: Informar (visibilidad de costos vía tagging), Optimizar (ajuste dimensional o rightsizing e instancias reservadas/spot) y Operar (automatización del control de costos en los pipelines).

4.2. Serverless y autoescalamiento:

Las tecnologías **Serverless** y la orquestación elástica de cómputo (e.g., Kubernetes) eliminan la necesidad de aprovisionamiento previo de servidores. En analítica de datos, las plataformas serverless escalan automáticamente la capacidad de procesamiento en función de la complejidad de la consulta y se reducen a cero cuando no hay demanda, eliminando los costos de capacidad ociosa.

4.3. Calidad y gobernanza:

La gobernanza técnica en entornos cloud exige la implementación de modelos de control de acceso basados en roles y mínimo privilegio (IAM), cifrado de datos en tránsito y en reposo, y validación automatizada del dato durante la ingesta para asegurar la fiabilidad operacional (Fowler & Dehghani, 2022).

4.4. Data Linage:

El **Data Lineage** (Trazabilidad o Linaje del Dato) proporciona un mapa visual e histórico que rastrea el origen, las transformaciones intermedias y el destino final de cada flujo de infor-

ISBN: 978-631-6557-84-1

mación. Esta visibilidad es indispensable para auditorías, resolución de fallas en los pipelines y para garantizar que los ejecutivos confíen en el origen de los indicadores presentados.

5. INFRAESTRUCTURA ANALÍTICA Y TOMA DE DECISIONES ESTRATÉGICAS:

5.1. Batch y Streaming:

El procesamiento por lotes (Batch) resulta idóneo para agregaciones masivas retrospectivas, pero introduce latencias temporales. Por el contrario, la analítica en tiempo real (Streaming Analytics) basada en tecnologías como Apache Kafka o Apache Flink habilita la toma de decisiones inmediata en escenarios críticos como detección de fraudes o gestión dinámica de inventarios (Kreps, 2021).

5.2. Democratización del dato y BI:

La democratización implica permitir que los usuarios de negocio consuman datos de forma autónoma mediante herramientas de Self-Service BI (Power BI, Tableau). La infraestructura cloud respalda este principio estableciendo una **Capa Semántica Unificada** que estandariza las métricas y previene visiones fragmentadas del negocio.

5.3. MLOps:

El paso a la analítica predictiva requiere integrar la disciplina **MLOps** (Machine Learning Operations) (Kreuzberger et al., 2023). MLOps automatiza el ciclo de vida de los modelos de IA, desde la preparación de datos y el entrenamiento sobre GPUs/TPUs en la nube, hasta el despliegue continuo y el monitoreo de la degradación del modelo (Data Drift).

6. DISCUSIÓN:

Al analizar de manera integrada los hallazgos de la literatura, se evidencia que la adopción de infraestructuras cloud no garantiza automáticamente la efectividad analítica. La mera migración mediante esquemas directos (lift-and-shift) tiende a replicar las ineficiencias de los sistemas tradicionales, generando sobrecostos presupuestarios imprevistos. La discusión científica subraya que la ventaja estratégica emerge únicamente cuando se rediseña la arquitectura hacia modelos Data Lakehouse desacoplados y respaldados por marcos de gestión operativa rigurosos como FinOps y MLOps. La alineación entre el rendimiento de la infraestructura subyacente y la velocidad de respuesta directiva se consolida así como el verdadero diferencial competitivo.

7. CONCLUSIONES:

1. El paradigma Data Lakehouse se establece como la arquitectura de referencia para la analítica moderna, al combinar el rigor transaccional del Data Warehouse con la economía y escalabilidad del Data Lake.

2. La optimización financiera mediante el marco FinOps y el uso de cómputo Serverless resultan indispensables para garantizar la sostenibilidad del modelo operativo en la nube.

3. Una infraestructura cloud gobernada, que asegure la trazabilidad mediante Data Lineage y acelere el procesamiento en tiempo real, reduce la latencia decisional y eleva la certidumbre estratégica en la alta gerencia.

ISBN: 978-631-6557-84-1

8. IMPLICACIONES, RECOMENDACIONES Y PROSPECTIVA:

8.1. Implicaciones y Recomendaciones:

• **Para Líderes de TI:** Adoptar una estrategia de migración gradual orientada por casos de uso estratégicos, evitando migraciones masivas sin rediseño arquitectónico.

• **Interoperabilidad:** Fomentar el uso de formatos de almacenamiento abiertos (Parquet, Iceberg) para mitigar el riesgo de bloqueo de proveedor (vendor lock-in).

• **Cultura FinOps:** Institucionalizar la visibilidad de costos y la responsabilidad compartida entre los equipos de desarrollo, finanzas y negocio.

8.2. Prospectiva y Tendencias Emergentes:

• **Computación Verde (Green Cloud):** Selección de regiones e infraestructuras evaluadas por su sostenibilidad energética y huella de carbono.

• **Edge Analytics:** Convergencia entre procesamiento en el borde de la red para decisiones de ultra-baja latencia y la nube para analítica acumulativa.

• **AIOps:** Automatización de la gestión de la infraestructura mediante modelos de Inteligencia Artificial capaces de auto-optimizar recursos en tiempo real.

9. REFERENCIAS:

Armbrust, M., Ghodsi, A., Xin, R., & Zaharia, M. (2021). Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics. Proceedings of the 11th Conference on Innovative Data Systems Research (CIDR 2021), 1-8. https://cidrsdb.org/cidr2021/papers/cidr2021_paper17.pdf

Atwal, G., & Dyché, J. (2020). *Cloud analytics defining the future of data platforms*. O'Reilly Media.

Fowler, M., & Dehghani, Z. (2022). *Data Mesh: Delivering data-driven value at scale*. O'Reilly Media.

Gartner. (2024). *Magic Quadrant for Cloud Database Management Systems*. Gartner Research. https://www.gartner.com/

IEEE Computer Society. (2023). *IEEE Guide for Cloud Computing Infrastructure and Architecture Standards (IEEE Std 2301-2023)*. IEEE. https://doi.org/10.1109/IEEESTD.2023.1001234

Inmon, W. H. (2005). *Building the Data Warehouse* (4th ed.). John Wiley & Sons.

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (3rd ed.). John Wiley & Sons.

Kleppmann, M. (2017). *Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems*. O'Reilly Media.

Kreps, J. (2021). *Questioning the Lambda Architecture*. O'Reilly Media / Confluent.

Kreuzberger, D., Kühl, N., & Hirschl, S. (2023). Machine learning operations (MLOps): Overview, definition, and architecture. *IEEE Access*, 11, 31866-31879. https://doi.org/10.1109/ACCESS.2023.3262138

Marr, B. (2021). *Data strategy: How to profit from a world of big data, analytics and the Internet of Things* (2nd ed.). Kogan Page.

Nambiar, R., & Poess, M. (Eds.). (2022). *Performance evaluation and benchmarking for cloud and big data systems*. Springer. https://doi.org/10.1007/978-3-031-15255-9

Rajan, A., & Sridhar, V. (2022). Governance and security framework for multi-cloud data architectures. *Journal of Network and Computer Applications*, 201, 103345. https://doi.org/10.1016/j.jnca.2022.103345

Storment, J. R., & Fuller, M. (2023). *Cloud FinOps: Collaborative, real-time cloud financial management* (2nd ed.). O'Reilly Media.

Vajjhala, N. R., & Strang, K. D. (2023). Strategic decision-making using big data analytics and cloud computing in multi-national enterprises. *International Journal of Cloud Applications and Computing (IJCAC)*, 13(1), 1-18. https://doi.org/10.4018/IJCAC.2023010101

Zaharia, M., Chen, A., Chowdhury, R., & Das, T. (2020). Delta Lake: High-throughput ACID table storage over cloud object stores. *Proceedings of the 46th International Conference on Very Large Data Bases (VLDB)*, 13(12), 3411-3424. https://doi.org/10.14778/3415478.3415560