

# Tendencias y Desafíos de la Computación de Alto Rendimiento en la Nube

Ezra F. Parra-González<sup>1</sup>, Uziel Jaramillo-Avila<sup>1</sup>, Pablo Salazar-Linares<sup>2</sup>,  
Carlos A. Lara-Álvarez<sup>1</sup>

**ezra.parra@cimat.mx; uziel.jaramillo@cimat.mx; salazar.pablo@ibm.com;  
carlos.lara@cimat.mx**

<sup>1</sup> Centro de Investigación en Matemáticas A. C. Unidad Zacatecas. CP 98160, Parque Quantum. Zacatecas, Zacatecas, México.

<sup>2</sup> IBM. Carretera al Castillo 2200 El Quince, CP 45680 El Salto, Jalisco, México.

**DOI: 10.17013/risti.49.131–146**

**Resumen:** En los últimos años, la demanda de aplicaciones que requieren Computación de Alto Rendimiento (HPC) ha aumentado considerablemente. Aunque tradicionalmente se ejecutan en infraestructuras privadas, cada vez es más común usar la nube como alternativa, gracias a su flexibilidad. Sin embargo, la computación en la nube también enfrenta desafíos como la gestión de recursos y virtualización, lo que afecta su desempeño. Para abordar estos desafíos, en este artículo se revisan trabajos relacionados con HPC en la nube y se analizan las principales tendencias y desafíos. Finalmente, presentamos una categorización de las tendencias identificadas.

**Palabras-clave:** Computación de Alto Rendimiento; Computación en la nube.

## *Trends and Challenges of HPC in the Cloud*

**Abstract:** In recent years, the demand for applications that require High-Performance Computing (HPC) has grown rapidly. These types of applications are traditionally executed on a company's or research center's private infrastructure; however, a different approach can also be used: the cloud. It is important to take advantage of the flexibility that decentralized computing power provides. However, like any other approach, cloud computing faces challenges such as resource management, virtualization, performance issues, etc. In this article, we review related works on HPC in the cloud to identify the main trends and challenges in this area. Finally, we propose a categorization of the identified main trends.

**Keywords:** High performance computing (HPC); Cloud Computing.

## 1. Introducción

La computación en la nube ha surgido como una alternativa al uso de recursos centralizados y se ha convertido en una solución viable para empresas y centros de

investigación debido a su aspecto económico (Rashid y Chaturvedi, 2019; Gong et al., 2010). Los proveedores de servicios en la nube ofrecen servicios de renta de recursos, como capacidad de almacenamiento y ciclos de almacenamiento (George y Pramila, 2021; Zhou et al., 2010), y una amplia variedad de servicios, tales como infraestructura como servicio (IaaS), software como servicio (SaaS), plataformas como servicios (PaaS), almacenamiento, copias de seguridad, bases de datos, análisis de datos, Internet de las cosas (IoT), Inteligencia artificial (IA) y aprendizaje automático (machine learning) (Bokhari et al., 2018; Meade et al., 2013; Mell y Grance, 2011).

Existen cuatro principales modelos de implementación de servicios en la nube: privados, públicos, comunitarios e híbridos (Bokhari et al., 2018; Dillon et al., 2010). En el contexto de la computación en la nube, existen diferentes paradigmas, siendo el más conocido el de Infraestructura como servicio (IaaS) (Chhabra et al., 2021), que está relacionado con la noción de cómputo en la nube (Malla y Christensen, 2020; Sotomayor et al., 2009). Existe una cantidad considerable de servicios de cómputo en la nube (Hoefler y Karagiannis, 2010), y casi todas las implementaciones se pueden plantear en términos de un servicio SaaS (Rimal et al., 2009).

Dado que cada tipo de aplicación requiere una cantidad diferente de recursos para desempeñarse de forma efectiva, resulta esencial revisar los estudios más relevantes relacionados con la computación de alto rendimiento (HPC) en la nube e identificar las principales tendencias y desafíos en su adopción. La revisión permitirá generar una categorización de los trabajos revisados según las principales tendencias detectadas y así identificar los tipos de implementaciones más relevantes según su ámbito de aplicación.

Con el fin de orientar el estudio, se plantearon las siguientes preguntas de investigación:

1. ¿Qué tendencias pueden ser identificadas en la adopción de HPC en la nube?
2. ¿Cuáles son los principales desafíos en la adopción de HPC en la nube?
3. ¿Cómo se pueden categorizar las principales aportaciones científicas en el área de la HPC en la nube?

La revisión de la literatura se llevó a cabo mediante una metodología narrativa, que permitió responder de manera puntual a las preguntas de investigación (Wiles, 2011). Para ello, se utilizaron bases de datos como Web of Science, Scopus y Google Scholar, utilizando términos clave relacionados con la computación en la nube y la HPC, tales como “high performance computing”, “cloud computing”, “IaaS”, entre otros. Posteriormente, se realizó una selección de los estudios relevantes y se procedió a su análisis crítico. Una vez identificadas las principales tendencias y desafíos en la adopción de HPC en la nube, se presenta una propuesta de categorización de los trabajos revisados según las principales tendencias detectadas y se expondrán las conclusiones obtenidas a partir de la revisión.

### 1.1. Aplicaciones de HPC

Aquellas aplicaciones que requieren una gran cantidad de recursos computacionales son conocidas como *Computación de Alto Rendimiento* (HPC). HPC es una disciplina de las tecnologías de la información con una gran cantidad de aplicaciones, con grandes requerimientos de recursos, en diferentes campos como la ciencia, economía, ingeniería

aeroespacial, etc. (Netto et al., 2019 ; Mauch et al., 2012); sin embargo, los sistemas HPC tradicionales son costosos y requieren una gran cantidad de recursos (Lynn et al., 2020), lo que limita su accesibilidad para muchos usuarios. En diversas aplicaciones, se requiere de modelos matemáticos para obtener un resultado satisfactorio, dichos modelos frecuentemente requieren grandes cantidades de recursos computacionales para llevar a cabo experimentos a gran escala o reducir la complejidad computacional en un momento dado (Vecchiola et al., 2009). Tradicionalmente, las aplicaciones de desempeño alto están basadas en sistemas de tipo redes o grupos (Duran-Limon et al., 2011). Ya que el cómputo en la nube ha experimentado un gran crecimiento en el modelo de tipo “*paga sobre la marcha*”, la comunidad científica ha mostrado considerable interés en su uso, dada la capacidad de sólo gastar el dinero correspondiente a los recursos computacionales requeridos en una aplicación dada (Jackson et al., 2010).

### 1.2. HPC en la nube

El uso de recursos computacionales bajo demanda proporciona una mejora al obtener resultados a problemas grandes en un tiempo aceptable (Wang et al., 2017). El interés en el uso de servicios de HPC en la nube crece conforme a su disponibilidad, poder de cómputo y desempeño incrementan (Expósito et al., 2012). Algunas de las compañías que ofrecen estos servicios lo hacen con tecnología patentada, aunque en algunos casos también se hace a través de código abierto (Dukaric y Juric, 2012; Wind, 2011; Cordeiro et al., 2010). Una parte fundamental del cómputo en la nube es el proporcionar infraestructura como un servicio, esto es logrado típicamente proporcionando máquinas virtuales a los usuarios para completar sus tareas (von Laszewski et al., 2012; Vila et al., 2019).

A pesar de que existen un gran número de iniciativas para el despliegue de HPC en la nube, hay un número de problemas asociados a las aplicaciones de gran desempeño (Benedict, 2013), siendo las principales; virtualización, desempeño, manejo de memoria, seguridad, dominio de aplicación e impactos económicos.

## 2. Tendencias

La revisión de trabajos realizada permitió de manera general identificar las siguientes tendencias en estudios referentes al HPC en la nube que de manera general se clasificaron en: aportaciones relacionadas a la virtualización, desempeño de los sistemas, aportaciones relacionadas al dominio de la aplicación, gestión de recursos y finalmente los relacionados a los aspectos económicos.

*Virtualización*: Es una técnica que permite la creación de múltiples instancias de sistemas operativos y aplicaciones en una misma infraestructura de hardware, lo que permite el uso eficiente de recursos y la mejora del desempeño de los sistemas. En el ámbito de la HPC en la nube, la virtualización ha sido objeto de un gran interés en la investigación, con un enfoque particular en el desarrollo de sistemas de peso ligero (Ren et al., 2017; Duran-Limon et al., 2011; Xavier et al., 2013; Lee et al., 2013; Dai et al., 2013; Chaudhary et al., 2008), peso pesado (Regola y Ducom, 2010; Youseff et al., 2006), hipervisores (Younge et al., 2011; Regola y Ducom, 2010; Perera et al., 2011; Aladyshev et al., 2018) y enfoques híbridos (Zhang y Panda, 2017; Stewart et al., 2012; Lee et al., 2013).

La virtualización de peso ligero y los contenedores son enfoques importantes para reducir los costos generales del uso de HPC en la nube (Ren et al., 2017; Colonnelli et al., 2020). Xavier et al. (2013) evaluaron el desempeño de la virtualización basada en contenedores para HPC, mientras que Lee et al. (2013) investigaron técnicas de optimización como la virtualización ligera y la afinidad de CPU. Guidi et al. (2021) demostraron que dar acceso directo a las I/O a las máquinas virtuales puede reducir la brecha en el desempeño entre implementaciones en la nube y nativas.

El hipervisor es un factor crítico en el desempeño experimentado por el sistema operativo invitado y las aplicaciones en máquinas de alto desempeño. Younge et al. (2011) evaluaron la factibilidad de entornos de virtualización HPC y determinaron que los hipervisores deben cumplir ciertas condiciones, siendo la máquina virtual basada en el núcleo (KVM) la mejor opción para implementaciones en la nube. Perera et al. (2011) identificaron características importantes para los clientes al seleccionar un hipervisor.

La virtualización híbrida puede minimizar los problemas causados por un solo enfoque de virtualización. Zhang y Panda (2017) propusieron una virtualización híbrida que utiliza procesos MPI almacenados en diferentes contenedores y máquinas virtuales, mientras que Stewart et al. (2012) presentaron un enfoque híbrido que ofrece un desempeño superior al de la virtualización completa, aunque no tan bueno como el de la para-virtualización o la virtualización del sistema operativo, con la capacidad adicional de virtualizar sistemas operativos invitados arbitrariamente.

*Desempeño:* En la investigación del HPC en la nube, se ha detectado una tendencia general hacia la mejora del rendimiento de los sistemas. Uno de los mayores desafíos en esta área es la necesidad de mejorar la eficiencia del almacenamiento, la entrada/salida (I/O), la resolución de problemas relacionados con la comunicación, como el ancho de banda, la tolerancia a fallos y las latencias.

Para mejorar la eficiencia del almacenamiento, se están investigando nuevos enfoques para reducir los tiempos de acceso a los datos (Kundu et al., 2010; Agarwal y Prasad, 2012; Chen et al., 2011). Ekanayake y Fox (2010) presentan experimentos en los que se comparan diversas aplicaciones con altos requerimientos de almacenamiento, indicando que muchas de ellas utilizan tareas paralelas para acceder a los datos almacenados en discos locales, aunque no está claro cómo funcionaría este enfoque usando máquinas virtuales.

Los costos de virtualización son especialmente visibles en la virtualización de I/O, donde el host o el monitor de la máquina virtual (VMM) debe revisar cada operación I/O. Esto no es favorable para aplicaciones HPC donde el desempeño de la comunicación es crítico (Maliszewski et al., 2019).

Ren et al. (2017) proponen una infraestructura de nube para aplicaciones HPC basada en la máquina virtual liviana “nOSV” que ejecuta aplicaciones HPC básicas y preserva la flexibilidad de la nube al permitir la virtualización anidada. Zhang y Panda (2017) proponen enfoques de virtualización para evitar cuellos de botella en cargas intensivas de E/S y sugieren una solución compuesta por una localidad de dos capas de alto rendimiento y una biblioteca MPI compatible con NUMA. Además, Expósito et al. (2013) muestran que el uso de discos efímeros proporciona un mejor rendimiento que los volúmenes EBS y que el rendimiento general puede verse limitado por un bajo rendimiento en la red.

Para mejorar el desempeño de I/O en aplicaciones HPC, Shan et al. (2008) proponen caracterizar y predecir el desempeño de I/O seleccionando requerimientos de I/O que correspondan a los patrones de aplicaciones específicas, mientras que Liu et al. (2011) sugieren que la configurabilidad de los servicios de nube pública puede proporcionar oportunidades para mejorar el rendimiento y/o reducir los costos de las aplicaciones de HPC. En particular, los clusters basados en la nube pueden permitir al usuario construir soluciones de I/O paralelas que satisfagan todas las necesidades de la aplicación, según lo analizado por Yu y Vetter (2008).

La variación en los retardos y el rendimiento inestable de la red tienen un fuerte impacto negativo en las aplicaciones científicas (Llopis et al., 2019). Incluso una granularidad de tiempo de decenas de milisegundos puede causar fluctuaciones en el rendimiento de una aplicación rápida de 1 Gb/s a cero (Wang et al., 2010). La configuración de parámetros, la pérdida de paquetes y la virtualización de la máquina son factores que afectan el rendimiento TPC.

El ancho de banda es una condición determinante para los entornos de computación en la nube, y la disponibilidad de interconexiones Infiniband en máquinas virtuales mejora el rendimiento para aplicaciones científicas (Brunetta y Borin, 2019; Ramakrishnan et al., 2012).

La escalabilidad de las aplicaciones HPC depende del soporte de las redes de comunicación en entornos virtualizados. Las latencias de arranque altas y los anchos de banda limitados limitan severamente la escalabilidad de programas intensivos en comunicación (Expósito et al., 2012).

La tolerancia a fallos permite a sistemas HPC con múltiples nodos en la nube completar la ejecución de aplicaciones intensivas, como el enfoque de redundancia en el nivel de proceso (PLR) presentado por Egwutuoha et al. (2012, 2012b, 2013).

***Dominio de aplicación:*** Otra forma de clasificar las aportaciones científicas en HPC en la nube agrupando es según el dominio de aplicación.

En el ámbito científico, la HPC en la nube se ha utilizado para acelerar la investigación en diversos campos, como la física, la química, la biología y la medicina. De manera general, los científicos han utilizado la HPC en la nube para modelar y simular sistemas complejos (Sharma y Jadhao, 2021; Vecchiola et al., 2009; Changchun et al., 2011; Hoffa et al., 2008; Jackson et al., 2010).

En el dominio de datos intensivos, la HPC en la nube se utiliza para manejar grandes conjuntos de datos y realizar análisis complejos en ellos. La HPC en la nube permite el procesamiento paralelo y distribuido, lo que acelera el tiempo de respuesta y reduce el costo total del procesamiento de datos (Abadi, 2009; Cooper et al., 2010; Shamsi et al., 2013).

En áreas tan particulares como la física de altas energías (Sobie et al., 2011), la HPC en la nube se ha utilizado para modelar y simular fenómenos altamente complejos, lo que ayuda a comprender mejor las problemáticas.

En el campo de la bioinformática, la HPC en la nube se enfoca a demostrar que la computación en la nube y la virtualización son tecnologías que pueden ser utilizadas

para abordar los desafíos en el análisis de datos biológicos, especialmente en términos de escalabilidad y eficiencia de los recursos de cómputo (Yin et al., 2019). Además, se discuten las posibles soluciones para abordar los desafíos y limitaciones actuales de estas tecnologías emergentes, incluyendo la necesidad de desarrollar algoritmos más eficientes y optimizados para su uso en plataformas de cómputo en la nube (Ekanayake et al., 2011; Youseff et al., 2006b; Arrais y Oliveira, 2010; Benedict, 2013; Calabrese y Cannataro, 2015).

En el ámbito de los videojuegos, la HPC en la nube se utiliza para el procesamiento de gráficos y la simulación en tiempo real. Investigaciones como la propuesta por Chen et al., (2014) realizan una evaluación de la calidad de servicio en los sistemas de juego en la nube y con sus experimentos demuestran que aplicaciones de alto rendimiento como los videojuegos se benefician del uso de la computación en la nube y el HPC. En general los desarrolladores de videojuegos pueden utilizar la HPC en la nube para crear mundos virtuales más realistas y detallados, lo que mejora la experiencia del usuario.

La simulación a gran escala es otro dominio de aplicación que fue detectado, en donde la HPC en la nube se utiliza para simular eventos complejos en una amplia gama de áreas, como la ingeniería, la física y la medicina. La HPC en la nube permite a los científicos y a los ingenieros realizar simulaciones complejas que antes no eran posibles debido a las limitaciones de hardware y software.

*Gestión:* Como ya se ha establecido en éste documento, la evolución de la HPC en la nube ha generado interés por ofrecer recursos bajo demanda y a escala. Sin embargo, la gestión de estos algunos recursos en la nube presenta una serie de desafíos.

En la revisión realizada se identificaron tres tendencias principales relacionadas a la gestión de HPC en la nube: seguridad, memoria y energía. En lo que respecta a la seguridad, es esencial garantizar la privacidad y la integridad de los datos y la infraestructura. El artículo de Grobauer et al., (2011) reporta detalladamente las vulnerabilidades en sistemas HPC en la nube y realizan un análisis de los riesgos de seguridad en los modelos de servicio (IaaS, PaaS y SaaS), identificando las principales amenazas y debilidades en cada uno de ellos. Vaquero et al., (2010) aborda otros aspectos relacionados a la seguridad como es la confidencialidad de los datos, la integridad de los sistemas y la disponibilidad del servicio. Los trabajos proponen estrategias de seguridad para la nube en diferentes áreas, desde la arquitectura y la gestión de identidad hasta la seguridad de la red y la virtualización.

En cuanto a la memoria, es crucial garantizar una gestión eficiente para maximizar el rendimiento en HPC en la nube. Los sistemas de HPC en la nube deben garantizar un acceso rápido y eficiente a la memoria de alta velocidad y proporcionar herramientas para optimizar el uso de la memoria y minimizar la latencia de acceso (Dong et al., 2013; Kundu et al., 2010; Wu et al., 2011).

Otro aspecto que requiere una gestión adecuada en el HPC en la nube es la energética. Administrar adecuadamente la energía permite reducir los costos y la huella de carbono de los centros de datos. Los sistemas de gestión de energía en la nube de HPC deben optimizar el uso de la energía y garantizar la disponibilidad de recursos para las aplicaciones de HPC de manera eficiente (Rodero et al., 2012; Anton et al., 2012).

***Economía:*** La rentabilidad de las soluciones de HPC en la nube es otra área de interés para los investigadores. El objetivo es estudiar los desafíos relacionados al tema económico, incluyendo la gestión de costos y la garantía de un acuerdo de nivel de servicio (SLA) satisfactorio.

En cuanto a la gestión de costos, el análisis de costo-beneficio es una tendencia clave en la HPC en la nube. El análisis de costo-beneficio es una herramienta importante para determinar la viabilidad económica de una solución de HPC en la nube (Carlyle et al., 2010; Zhai et al., 2011). Además, la garantía de un acuerdo de nivel de servicio (SLA) satisfactorio es una tendencia clave en la HPC en la nube. Un SLA es un contrato entre el proveedor de la nube y el cliente que establece los niveles de servicio que se deben cumplir, como la disponibilidad del servicio, el tiempo de respuesta y la capacidad de procesamiento. Los proveedores deben asegurarse de que sus SLA sean claros, específicos y medibles, para que los clientes puedan confiar en la calidad y la consistencia del servicio de HPC en la nube. Asimismo, deben desarrollar herramientas y procesos para monitorear y medir el rendimiento de la nube y garantizar que se cumplan los SLA acordados (Carlyle et al., 2010; Zhai et al., 2011). Por ejemplo, Nasr et al., (2019) presentan un algoritmo para un algoritmo para la programación de tareas de HPC en la nube de forma rentable.

### **3. Desafíos**

Las capacidades de la HPC en la nube han generado un especial interés en las industrias y organizaciones. Sin embargo, su implementación también plantea desafíos que deben ser abordados para su correcto funcionamiento. En este sentido, se han identificado cinco áreas principales en las que se presentan los desafíos más importantes en la HPC en la nube.

***Problemas de seguridad y privacidad:*** Uno de los mayores desafíos de la computación en la nube es la seguridad y la privacidad. Cuando las empresas trasladan sus datos a la nube, le están confiando a un tercero su información confidencial, y eso puede generar inquietudes sobre violaciones de datos, acceso no autorizado y privacidad de datos (Chakraborti, 2022).

***Dependencia de la conectividad a Internet:*** Los servicios de computación en la nube dependen de Internet para la comunicación entre el proveedor de la nube y el usuario. Esto significa que, si hay una interrupción en la conectividad a Internet, es posible que el usuario no pueda acceder a sus datos o aplicaciones. Esto puede ser un desafío importante para las empresas que requieren un acceso continuo a sus datos. Esto puede ser particularmente desafiante en países subdesarrollados ya que las velocidades de transferencia de datos pueden ser lentas (Razi y Batan, 2023). En esos casos es posible que no se pueda servicios como Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Software as a Service (SaaS).

***Bloqueo del proveedor (Vendor lock-in):*** Cuando las empresas eligen un proveedor de la nube, están bloqueadas en la pila de tecnología de ese proveedor, lo que puede dificultar el cambio a un proveedor diferente en el futuro. Algunas de las soluciones propuestas es seleccionar el mejor proveedor de servicios de acuerdo con diferentes criterios (Mohamed

y Abdelsalam, 2020), o seleccionar al proveedor que logre un equilibrio entre el costo y la disponibilidad (Hari Kumar y Mala, 2022).

*Problemas normativos y de cumplimiento:* La computación en la nube puede generar problemas normativos y de cumplimiento para las empresas. Estas regulaciones pueden describir los estándares requeridos para tomar medidas de precaución para prevenir ataques (Parast et al. 2022).

*Gestión de costos:* La computación en la nube puede ser rentable, ya que las organizaciones pueden reducir gastos y aumentar la eficiencia compartiendo recursos y pagando solo por lo que necesitan (Razi y Batan, 2023). Sin embargo, también puede ser costosa si no se administra adecuadamente; por ejemplo, las empresas pueden incurrir en costo de tarifas de transferencia de datos, de almacenamiento u otros cargos.

#### **4. Categorización de trabajos analizados**

En la Figura 1 se muestran las categorías más relevantes obtenidas del análisis realizado. Cada categoría agrupa los trabajos revisados según su área de aplicación, lo que permite identificar las principales tendencias en el uso de HPC en la nube en diferentes campos. Las categorías identificadas fueron: virtualización, desempeño, dominio de aplicación, gestión y los trabajos relacionados al tema económico. Las categorías reflejan el amplio espectro de aplicaciones que pueden beneficiarse de la utilización de HPC en la nube.

#### **5. Conclusiones**

La computación en la nube se presenta como una solución para el procesamiento de cargas de trabajo de alto rendimiento. Aunque tanto las aplicaciones de HPC como las convencionales tienen problemas similares y diferentes, es tarea de los investigadores identificar los problemas que siguen siendo desafíos en el campo y mejorar el estado del arte. Este estudio propuso identificar las tendencias más importantes sobre la computación en la nube de alto rendimiento y destacar las preocupaciones abiertas y cerradas empleando un sistema de clasificación. Los resultados de esta revisión sugieren que, aunque la HPC basada en la nube ha logrado importantes avances, todavía hay áreas en las que se requiere más investigación. Al abordar estos problemas, se pueden mejorar aún más la eficiencia, la escalabilidad y la robustez de los sistemas de HPC basados en la nube. Además de identificar las áreas de oportunidad en el campo de la HPC basada en la nube, es importante considerar cómo estos vacíos pueden abordarse. Una posible estrategia es mediante el desarrollo de herramientas y recursos para la HPC en la nube. Por ejemplo, se podría crear una plataforma para evaluar el rendimiento y la eficiencia de diferentes proveedores de la nube y plataformas de virtualización para aplicaciones de HPC. También se podrían desarrollar herramientas de monitoreo y análisis para ayudar a los usuarios a comprender y optimizar el rendimiento de sus aplicaciones. En última instancia, el objetivo es mejorar la eficiencia y la accesibilidad de la HPC basada en la nube para una amplia gama de usuarios.



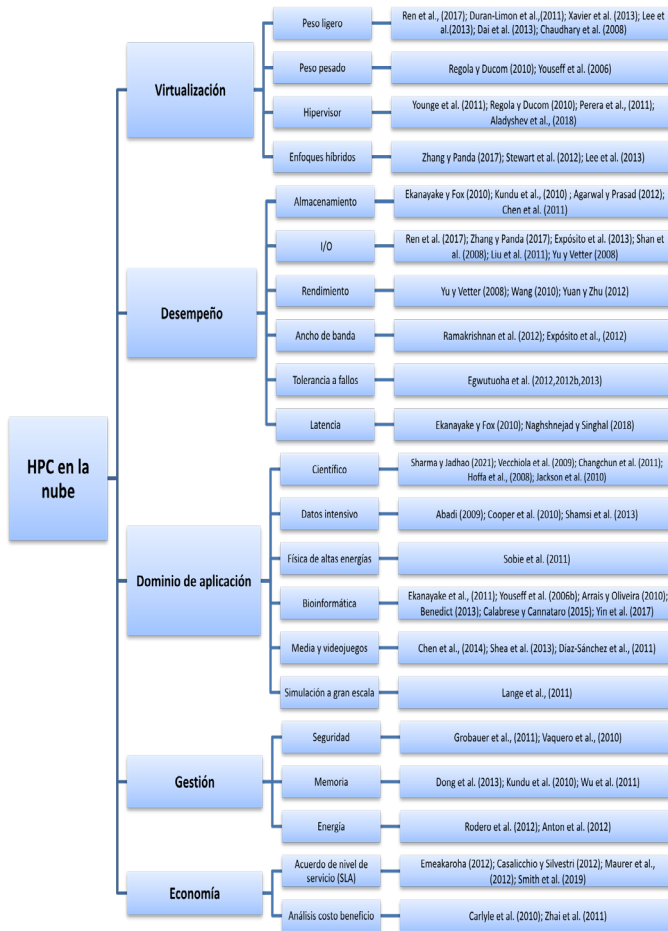


Figura 1 – Categorías de problemáticas específicas a diferentes tipos de HPC en la nube

## Referencias

Abadi, D. J. (2009). *Data management in the cloud: Limitations and opportunities*. IEEE Data Eng. Bull, 32(1), 3-12.

Agarwal, D., & Prasad, S. K. (2012). *AzureBench: Benchmarking the Storage Services of the Azure Cloud Platform*. In Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW), 2012 IEEE 26th International (pp. 1048-1057).

Aladyshev, O. S., Baranov, A. V., Ionin, R. P., Kiselev, E. A., & Shabanov, B. M. (2018). Variants of deployment the high performance computing in clouds. En 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus) (pp. 1453-1457) Russia.

- Anton, B., Rajkumar, B., Young, C. L., & Albert, Z. (2012). *A taxonomy and survey of energy-efficient data centers and cloud computing systems*. Technical Report CLOUDS-TR-2010-3. arXiv preprint arXiv:1007.0066v2.
- Rashid A., Chaturvedi A., (2019). Cloud Computing Characteristics and Services: A Brief Review. *International Journal of Computer Sciences and Engineering*, 7(2), 421-426.
- Arrais, J.P. & Oliveira, J.L. (2010). *On the exploitation of cloud computing in bioinformatics*. Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on, 1-4.
- Benedict, S. (2013). *Performance issues and performance analysis tools for HPC cloud applications: a survey*. *Computing*, 95(2), 89-108.
- Bokhari, M.U., Makki, Q., & Tamandani, Y.K. (2018). A Survey on Cloud Computing. In V. Aggarwal, V. Bhatnagar, & D. Mishra (Eds.), *Big Data Analytics . Advances in Intelligent Systems and Computing*, vol 654 (pp. 197-212). Springer, Singapore.
- Brunetta, J. R., & Borin, E. (2019). *Selecting efficient cloud resources for HPC workloads*. In Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing (pp. 155-164).
- Calabrese, B., & Cannataro, M. (2015). *Cloud computing in healthcare and biomedicine*. *Scalable computing: practice and experience*, 16 (1), 1-18. <https://doi.org/10.12694/scpe.v16i1.1057>
- Carlyle, A. G., Harrell, S. L., & Smith, P. M. (2010). *Cost-effective HPC: The community or the Cloud?*. In *Cloud Computing Technology and Science (CloudCom)*, 2010 IEEE Second International Conference on (pp. 169-176).
- Casalicchio, E., & Silvestri, L. (2012). *Mechanisms for SLA provisioning in cloud-based service providers*. *Computer Networks*, 56(11), 2671-2686.
- Changchun S., Hiroshi N., Susan J., Kai S., Krishna M., & Yong Q. (2011). *HPC Cloud Applied To Lattice Optimization*. In Proceedings of 2011 Particle Accelerator Conference (pp. 3813-3815). IEEE.
- Chaudhary, V., Cha, M., Walters, J. P., Guercio, S., & Gallo, S. (2008). *A comparison of virtualization technologies for HPC*. In *Advanced Information Networking and Applications*, 2008. AINA 2008. 22nd International Conference on (pp. 861-868). IEEE.
- Chhabra, A., Singh, G., & Kahlon, K. S. (2021). *Multi-criteria HPC task scheduling on IaaS cloud infrastructures using meta-heuristics*. *Cluster Computing*, 24, 885-918. <https://doi.org/10.1007/s10586-021-03387-2>
- Chen, H., Cheng, K., Jheng, Z., Pan, Y., Wu, C. & Yu, H. (2011). *Easy Virtual Cluster with Cloud Storage for HPC Numerical Simulation*. *Commerce and Enterprise Computing (CEC)*, 2011 IEEE 13th Conference on , vol., no., pp.197,200, 5-7 Sept. 2011. <https://doi.org/10.1109/CEC.2011.36>

- Chen, K.-T., Chang, Y.-C., Hsu, H.-J., Chen, D.-Y., Huang, C.-Y., & Hsu, C.-H. (2014). *On the quality of service of cloud gaming systems*. *IEEE Transactions on Multimedia*, 16(2), 480-495.
- Chakraborti, A., Curtmola, R., Katz, J., Nieh, J., Sadeghi, A. R., Sion, R., & Zhang, Y. (2022). *Cloud computing security: foundations and research directions*. *Foundations and Trends in Privacy and Security*, 3(2), 103-213.
- Colonnelli, I., Cantalupo, B., Merelli, I., & Aldinucci, M. (2020). *StreamFlow: cross-breeding cloud with HPC*. *IEEE Transactions on Emerging Topics in Computing*, 9(4), 1723-1737.
- Cooper, B.F., Silberstein, A., Tam, E., Ramakrishnan, R., & Sears, R. (2010). *Benchmarking cloud serving systems with YCSB*. *Proceedings of the 1st ACM symposium on Cloud computing*, 2010.
- Cordeiro, T., Damalio, D., Pereira, N., Endo, P., Palhares, A., Gonçalves, G., ... & Mångs, J. E. (2010). *Open source cloud computing platforms*. In *Grid and Cooperative Computing (GCC)*, 2010 9th International Conference on (pp. 366-371). IEEE.
- Dai, Y., Qi, Y., Ren, J., Shi, Y., Wang, X., & Yu, X. (2013). *A lightweight VMM on many core for high performance computing*. In *Proceedings of the 9th ACM SIGPLAN/SIGOPS international conference on Virtual execution environments (VEE '13)* (pp. 111-120). ACM.
- Díaz-Sánchez D., Almenarez F., Marín A., Proserpio, D. & Cabarcos PA. (2011). *Media cloud: an open cloud computing middleware for content management*. *IEEE Trans Consumer Electron* 57(2):970–978
- Dillon, T., Wu, C., & Chang, E. (2010). *Cloud computing: Issues and challenges*. In *Advanced Information Networking and Applications (AINA)*, 2010 24th IEEE International Conference on (pp. 27-33). IEEE.
- Dong O.S., Hong J.C., Jae H.P. & Cheol H. K. (2013). *Analysis of Memory Management Policies for Heterogeneous Cloud Computing*. *Information Science and Applications (ICISA)*, 2013 International Conference on, pp.1,3, 24-26 June 2013 <https://doi.org/10.1109/ICISA.2013.6579403>
- Dukaric, R., & Juric, M. B. (2012). *Towards a unified taxonomy and architecture of cloud frameworks*. *Future Generation Computer Systems*.
- Duran-Limon, H. A., Silva-Bañuelos, L. A., Tellez-Valdez, V. H., Parlavantzas, N., & Zhao, M. (2011). *Using Lightweight Virtual Machines to Run High Performance Computing Applications: The Case of the Weather Research and Forecasting Model*. In *Utility and Cloud Computing (UCC)*, Fourth IEEE International Conference on (pp. 146-153). IEEE.
- Egwutuoha, I.P., Chen, S., Levy, D., Selic, B., & Calvo, R. (2012). *A Proactive Fault Tolerance Approach to High Performance Computing (HPC) in the Cloud*. In *Cloud and Green Computing (CGC)*, Second International Conference on (pp. 268-273). IEEE.

- Egwutuoha, I.P., Shiping C., Levy, D. & Selic, B. (2012). *A Fault Tolerance Framework for High Performance Computing in Cloud*. Cluster, Cloud and Grid Computing (CCGrid), 12th IEEE/ACM International Symposium on, 13-16 May 2012, (pp.709,710). <https://doi.org/10.1109/CCGrid.2012.80>
- Egwutuoha, I.P., Shiping, C., Levy, D., Selic, B. & Calvo, R. (2013). *Energy Efficient Fault Tolerance for High Performance Computing (HPC) in the Cloud*. Cloud Computing (CLOUD), IEEE Sixth International, June 28 2013-July 3 2013, (pp.762,769). <https://doi.org/10.1109/CLOUD.2013.69>
- Ekanayake, J., & Fox, G. (2010). *High performance parallel computing with clouds and cloud technologies*. In Cloud Computing (pp. 20-38). Springer Berlin Heidelberg.
- Ekanayake, J., Gunarathne, T. & Qiu J. (2011). *Cloud Technologies for Bioinformatics Applications*. IEEE Transactions on Parallel and Distributed Systems, 22(6).
- Expósito, R. R., Taboada, G. L., Ramos, S., González-Domínguez, J., Touriño, J., & Doallo, R. (2013). *Analysis of I/O performance on an amazon EC2 cluster compute and high I/O platform*. Journal of Grid Computing, 1-19.
- Expósito, R. R., Taboada, G. L., Ramos, S., Touriño, J., & Doallo, R. (2012). *Performance analysis of HPC applications in the cloud*. Future Generation Computer Systems.
- George, S. S., & Pramila, R. S. (2021). A review of different techniques in cloud computing. Materials Today: Proceedings, 46(Part 17), (pp. 8002-8008).
- Gong, C., Liu, J., Zhang, Q., Chen, H., & Gong, Z. (2010). *The characteristics of cloud computing*. In Parallel Processing Workshops (ICPPW), 2010 39th International Conference on (pp. 275-279). IEEE.
- Grobauer, B., Walloschek, T. & Stocker, E. (2011) *Understanding cloud computing vulnerabilities*. Secur Privacy IEEE 9(2),50-57
- Guidi, G., Ellis, M., Buluç, A., Yelick, K., & Culler, D. (2021). *10 years later: Cloud5computing is closing the performance gap*. In Companion of the ACM/SPEC International Conference on Performance Engineering (41-48).
- Hari Kumar, P., & Mala, G. A. (2022). *H2RUN: An efficient vendor lock-in solution for multi-cloud environment using horse herd Runge Kutta based data placement optimization*. Transactions on Emerging Telecommunications Technologies, 33(9), e4541.
- Hofer, C. N., & Karagiannis, G. (2010). *Taxonomy of cloud computing services*. In GLOBECOM Workshops (GC Wkshps), 2010 IEEE (pp. 1345-1350).
- Hoffa, C., G. Mehta, T. Freeman, E. Deelman, K. Keahey, B. Berriman & Good J. (2008). *On the Use of Cloud Computing for Scientific Workflows*. SWBES.
- Jackson, K., Ramakrishnan, L., Muriki, K., Canon, S., Cholia, S., Shalf, J., Wasserman, H., & Wright, N. (2010). *Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud*. In 2nd IEEE International Conference on Cloud Computing Technology and Science. IEEE, 2010, (pp. 159-168).

- Kundu, A., Banerjee, C., Guha, S.K., Mitra, A., Chakraborty, S., Pal, C. & Roy, R. (2010). *Memory utilization in cloud computing using transparency*. Computer Sciences and Convergence Information Technology (ICCIT), 2010 5th International Conference on, 22-27, <https://doi.org/10.1109/ICCIT.2010.5711023>.
- Lange, J., Pedretti, K., Dinda, P., Bridges, P., Bae, C., Soltero, P. & Merritt, A. (2011). *Minimal Overhead Virtualization of a Large Scale Supercomputer*. Proceedings of the 2011 ACM SIGPLAN/SIGOPS. VEE Conference, 2011.
- Lee, P. F., Kaufmann, R., & Milojicic, D. (2013). *The Who, What, Why and How of High Performance Computing Applications in the Cloud*. IEEE Cloud Computing, 1(2), 24-31.
- Llopis, P., Lindqvist, C., Høimyr, N., van der Ster, D., & Ganz, P. (2019). *Integrating HPC into an agile and cloud-focused environment at CERN*. In EPJ Web of Conferences (Vol. 214, p. 07025). EDP Sciences.
- Liu, M., Zhai, J., Zhai, Y., Ma, X., & Chen, W. (2011). *One optimized I/O configuration per HPC application: leveraging the configurability of cloud*. In Proceedings of the 2nd ACM SIGOPS Asia-Pacific Workshop on Systems (APSys'11), Shanghai, China (pp. 1-5).
- Lynn, T., Fox, G., Gourinovitch, A., & Rosati, P. (2020). *Understanding the determinants and future challenges of cloud computing adoption for high performance computing*. Future Internet, 12(8), 135.
- Malla, S., & Christensen, K. (2020). *HPC in the cloud: Performance comparison of function as a service (FaaS) vs infrastructure as a service (IaaS)*. Internet Technology Letters, 3(1), e137.
- Maliszewski, A. M., Vogel, A., Griebler, D., Roloff, E., Fernandes, L. G., & OA, N. P. (2019, June). *Minimizing communication overheads in container-based clouds for HPC applications*. In 2019 IEEE Symposium on Computers and Communications (ISCC) (pp. 1-6). IEEE.
- Mauch, V., Kunze, M., & Hillenbrand, M. (2012). *High performance cloud computing*. Future Generation Computer Systems, 28(1), 76-90. <https://doi.org/10.1016/j.future.2011.02.005>
- Maurer, M., Emeakaroha, V. C., Brandic, I., & Altmann, J. (2012). *Cost-benefit analysis of an SLA mapping approach for defining standardized Cloud computing goods*. Future Generation Computer Systems, 28(1), 39-47.
- Meade, B., Manos, S., Sinnott, R., Fluke, C., van der Knijff, D., & Tseng, A. (2013). *Research Cloud Data Communities*. Journal of Grid Computing, 11(4), 601-618. <https://doi.org/10.1007/s10723-013-9262-8>
- Mell, P., & Grance, T. (2011). *The NIST definition of cloud computing*. National Institute of Standards and Technology, 53(6), 50.
- Mohamed, A. M., & Abdelsalam, H. M. (2020). *A multicriteria optimization model for cloud service provider selection in multicloud environments*. Software: Practice and Experience, 50(6), 925-947.

- Naghshnejad, M., & Singhal, M. (2018). Adaptive Online Runtime Prediction to Improve HPC Applications Latency in Cloud. En 2018 IEEE 11th International Conference on Cloud Computing (CLOUD) (pp. 762-769).
- Nasr, A. A., El-Bahnasawy, N. A., Attiya, G., & El-Sayed, A. (2019). *Cost-effective algorithm for workflow scheduling in cloud computing under deadline constraint*. Arabian Journal for Science and Engineering, 44, 3765-3780.
- Netto, M. A. S., Calheiros, R. N., Rodrigues, E. R., Cunha, R. L. F., & Buyya, R. (2019). HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges. ACM Computing Surveys, 51(1), Article 8.
- Parast, F. K., Sindhav, C., Nikam, S., Yekta, H. I., Kent, K. B., & Hakak, S. (2022). *Cloud computing security: A survey of service-based models*. Computers & Security, 114, 102580.
- Perera, P. Muditha, & Keppitiyagama, C. (2011). *A performance comparison of hypervisors*. In 2011 International Conference on Advances in ICT for Emerging Regions (ICTer) (pp. 107-112). IEEE.
- Ramakrishnan, L., Canon, R.S., Muriki, K., Sakrejda, I., & Wright, N.J. (2012). *Evaluating interconnect and virtualization performance for high performance computing*. SIGMETRICS Perform. Eval. Rev, 40(2), 55-60.
- Razi, M., & Batan, A. (2023). *Opportunities and Challenges of Cloud Computing in Developing Countries*. Artificial Intelligence in Society, 3(1), 1-8.
- Ren, J., Qi, Y., Dai, Y., Xuan, Y., & Shi, Y. (2017). nOSV: A lightweight nested-virtualization VMM for hosting high performance computing on cloud. Journal of Systems and Software, 124, 137-152.
- Regola, N., & Ducom, J. C. (2010). *Recommendations for virtualization technologies in high performance computing*. In Proceedings of the 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom'10) (pp. 409-416). Indianapolis, IN, USA.
- Rimal, B. P., Choi, E., & Lumb, I. (2009). *A taxonomy and survey of cloud computing systems*. In INC, IMS and IDC, 2009. NCM'09. Fifth International Joint Conference on (pp. 44-51). IEEE.
- Rodero, I., Viswanathan, H., Lee, E.K., Gamell, M., Pompili, D., & Parashar, M. (2012). *Energy-efficient thermal aware autonomic management of virtualized HPC cloud infrastructure*. J. Grid Comput., 10(3), 447-473.
- Shamsi, J., Khojaye, M. A., & Qasmi, M. A. (2013). *Data-Intensive Cloud Computing: Requirements, Expectations, Challenges, and Solutions*. Journal of Grid Computing, 1-30.
- Shan, H., Antypas, K., & Shalf, J. (2008). *Characterizing and predicting the I/O performance of HPC applications using a parameterized synthetic benchmark*. In Proceedings of the 20th ACM/IEEE Supercomputing Conference (SC'08) (pp. 42:1-42:12). Austin, TX, USA.

- Sharma, P., & Jadhao, V. (2021). *Molecular Dynamics Simulations on Cloud Computing and Machine Learning Platforms*. In 2021 IEEE 14th International Conference on Cloud Computing (CLOUD) (pp. 751-753).
- Shea, R., Jiangchuan Liu, Ngai, E. C.-H., & Yong Cui. (2013). *Cloud gaming: architecture and performance*. IEEE Network, 27(4), 16-21.
- Smith, P., Harrell, S. L., Younts, A., & Zhu, X. (2019). *Community Clusters or the Cloud: Continuing cost assessment of on-premises and cloud HPC in Higher Education*. In Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning) (pp. 1-4)
- Sobie, R. J., Agarwal, A., Anderson, M., Armstrong, P., Fransham, K., Gable, I., & Podaima, W. (2011). *Data intensive high energy physics analysis in a distributed cloud*. arXiv preprint arXiv:1101.0357.
- Sotomayor, B., Montero, R. S., Llorente, I. M., & Foster, I. (2009). *Virtual infrastructure management in private and hybrid clouds*. IEEE Internet Computing, 13(5), 14-22.
- Stewart, K. E., Humphries, J. W., & Andel, T. R. (2012). *An Automated Virtualization Performance Analysis Platform*. The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 9(3).
- Vaquero, L., Rodero-Merino, L., & Morán, D. (2010). *Locking the sky: A survey on IaaS cloud security*. Computing, 91(1), 93-118.
- Vecchiola, C., Pandey, S., & Buyya, R. (2009). *High-performance cloud computing: A view of scientific applications*. In Pervasive Systems, Algorithms, and Networks (ISPAN), 10th International Symposium on (pp. 4-16). IEEE.
- Von Laszewski, G., Diaz, J., Wang, F., & Fox, G. C. (2012). *Comparison of multiple cloud frameworks*. In Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on (pp. 734-741). IEEE.
- Vila, S., Guirado, F., Lerida, J. L., & Cores, F. (2019). *Energy-saving scheduling on IaaS HPC cloud environments based on a multi-objective genetic algorithm*. The Journal of Supercomputing, 75(3), 1483-1495
- Wang, G. & Ng, T.E. (2010). *The Impact of Virtualization on Network Performance of Amazon EC2 Data Center*". INFOCOM 2010.
- Wang, L., Ranjan, R., Chen, J., & Benatallah, B. (2017). *Cloud computing: methodology, systems, and applications*. CRC press.
- Wind, S. (2011). *Open source cloud computing management platforms: Introduction, comparison, and recommendations for implementation*. In Open Systems (ICOS), 2011 IEEE Conference on (pp. 175-179). IEEE.
- Wiles, R., Crow, G., & Pain, H. (2011). *Innovation in qualitative research methods: A narrative review*. Qualitative Research, 11(5), 587-604.

- Wu, H., Tantawi, A. N., Diao, Y., & Wang, W. (2011). *Adaptive memory load management in cloud data centers*. IBM Journal of Research and Development, 55(6), 5:1-5:10. <https://doi.org/10.1147/JRD.2011.2170869>
- Xavier, M.G., Neves, M.V., Rossi, F.D., Ferreto, T.C., Lange, T., & De Rose, C.A.F. (2013). *Performance Evaluation of Container Based Virtualization for High Performance Computing Environments*. Euromicro International Conference on Parallel, Distributed and Network-Based Processing, (pp. 233-240).
- Yin, Z., Lan, H., Tan, G., Lu, M., Vasilakos, A. V., & Liu, W. (2017). *Computing Platforms for Big Biological Data Analytics: Perspectives and Challenges*. Computational and Structural Biotechnology Journal, 15, (pp. 403-411).
- Younge, A. J., Henschel, R., Brown, J. T., von Laszewski, G., Qiu, J., & Fox, G. C. (2011). *Analysis of virtualization technologies for high performance computing environments*. In *Cloud Computing (CLOUD)*. 2011 IEEE International Conference on (pp. 9-16). IEEE.
- Youseff, L., Wolski, R., Gorda, B., & Krintz, C. (2006). *Paravirtualization for HPC systems*. In: Proc. International Workshop on XEN in HPC Cluster and Grid Computing Environments (XHPC'06), Sorrento, Italy, (pp. 474-486).
- Youseff, L., Wolski, R., Gorda, B., & Krintz, C. (2006b). *Evaluating the performance impact of Xen on MPI and process execution for HPC systems*. In First International Workshop on Virtualization Technology in Distributed Computing (VTDC 2006). IEEE.
- Yu, W., & Vetter, J. S. (2008). *Xen-based HPC: A parallel I/O perspective*. In: Proc. 8th IEEE International Symposium on Cluster Computing and the Grid (CCGRID'08), Lyon, France, pp. 154-161.
- Yuan, W., & Zhu, Y. (2012). *An autonomic capacity management approach with cloud insight towards cost-efficient throughput optimization for high performance computing*. In Communication Systems and Network Technologies (CSNT), 2012 International Conference on (740-744). IEEE.
- Zhang, J., Lu, X., & Panda, D. K. (2017). *Designing Locality and NUMA Aware MPI Runtime for Nested Virtualization based HPC Cloud with SR-IOV Enabled InfiniBand*. In Proceedings of the 13th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments (187-200).
- Zhai, Y., Liu, M., Zhai, J., Ma, X., & Chen, W. (2011). *Cloud versus in-house cluster: evaluating Amazon cluster compute instances for running MPI applications*. In State of the Practice Reports (SC '11) (pp. 11, 10). ACM.
- Zhou, M., Zhang, R., Zeng, D., & Qian, W. (2010). *Services in the cloud computing era: A survey*. In UCS 4th International (pp. 40-46). IEEE.