La evolución de los sistemas de computación distribuida: desde los fundamentos hasta las nuevas fronteras

Dominic Lindsay, Sukhpal Singh Gill, Daria Smirnova, Peter Garraghan

March 9, 2025

Abstract

Este artículo proporciona una visión general de la evolución de los sistemas de computación distribuida. Desde los fundamentos de los primeros sistemas mainframe, hasta los paradigmas modernos como la computación en la nube, la computación en la niebla (fog computing) y el Internet de las cosas (IoT), examinamos los factores clave que han impulsado esta evolución. Además, consideramos los desafíos actuales y futuros en la investigación de sistemas distribuidos, incluyendo la gestión de grandes volúmenes de datos, la seguridad y la privacidad.

Keywords: Distributed systems, Cloud computing, Fog computing, Internet of Things (IoT), Edge computing, Data management, Security, Privacy.

Introducción

La evolución de los sistemas distribuidos, impulsada por la creciente dependencia social de Internet, ha transformado la infraestructura de computación desde los mainframes centralizados hasta paradigmas descentralizados como loT y la computación en la niebla (Fog Computing). Este cambio ha sido catalizado por avances tecnológicos como la agrupación (clustering), la conmutación de paquetes, las interfaces gráficas de usuario (GUIs) y la estandarización de protocolos de red, permitiendo la distribución de servicios y la expansión de la infraestructura más allá de las redes centrales.

Introducción

Los paradigmas de sistemas distribuidos han evolucionado para distribuir y facilitar el servicio desde clusters centralizados, extendiendo la infraestructura más allá de los límites de las redes centrales, formando paradigmas como loT y computación en la niebla (Fog computing)

Fundamentos de los sistemas distribuidos DEFINICIÓN

Los sistemas distribuidos describen una clase de sistema de computación en la cual los componentes de hardware y software están conectados por medio de una red, y coordinan sus acciones a través del envío de mensajes para cumplir un objetivo compartido.

Fundamentos de los sistemas distribuidos CARACTERÍSTICAS

Concurrencia transparente. Los sistemas distribuidos son inherentemente concurrentes, con cualquier recurso participante accesible a través de cualquier número de procesos locales o remotos. La capacidad y disponibilidad de tal sistema se puede aumentar agregando recursos que requieren mecanismos para la contabilidad e identificación. Tal sistema es vulnerable a comportamientos volátiles entre actores y debe ser resistente a fallas de nodos, así como a mensajes perdidos y retrasados. La gestión y el acceso de objetos, hardware o datos en un entorno de red distribuida también son de particular importancia debido al potencial de contienda de recursos físicos.

Fundamentos de los sistemas distribuidos CARACTERÍSTICAS

Falta de reloj compartido. Los sistemas mantienen su propio tiempo independiente, interpretado de diversas fuentes, y como tales, los sistemas operativos (SO) son susceptibles a la desviación y la deriva del reloj. Además, detectar cuándo se envió o recibió un mensaje es importante para asegurar el comportamiento correcto del sistema. Por lo tanto, los eventos se rastrean por medio de relojes lógicos y vectoriales conceptuales; al secuenciar los mensajes, los procesos distribuidos a través de una red pueden asegurar el ordenamiento total de los eventos.

Fundamentos de los sistemas distribuidos CARACTERÍSTICAS

Operación confiable y segura. Los componentes de un sistema distribuido son autónomos, y las solicitudes de servicio dependen de la correcta transacción de operación entre subsistemas. La falla de cualquier subsistema puede afectar el resultado de las solicitudes de servicio y puede manifestarse de maneras que son difíciles de mitigar eficazmente. La tolerancia a fallos y la confiabilidad son características clave para asegurar la supervivencia de los sistemas distribuidos y permitir que los servicios se recuperen de fallos mientras mantienen un servicio correcto.

La arquitectura del sistema físico identifica los dispositivos físicos que intercambian mensajes en un sistema distribuido y a través de qué medio se comunican. Los primeros sistemas distribuidos, como los mainframes, estaban conectados físicamente a los clientes. Posteriormente, la conmutación de paquetes permitió la comunicación de largo alcance con múltiples saltos. Las redes celulares incorporan sistemas de computación móvil, mientras que los sistemas modernos alojan servicios en hardware especializado entre proveedores de servicios y consumidores.

Entidades. Una perspectiva lógica de un sistema distribuido describe varios procesos intercambiando mensajes para lograr un objetivo común. Los sistemas contemporáneos extienden esta definición considerando entidades lógicas y agregadas, como Objetos y Componentes, utilizados para abstraer recursos y funcionalidad. Los sistemas más recientes aprovechan los servicios web y los microservicios, que consideran su despliegue en hardware físico, así como restricciones que incluyen localidad, utilización y políticas de los interesados. La computación Grid y la computación en la nube habilitan la computación distribuida al abstraer la agregación de procesamiento, memoria y espacio en disco, mientras que la computación en la niebla (Fog) y la computación en el borde (Edge) enfatizan la integración de dispositivos móviles v embebidos.

Modelos de comunicación en los sistemas distribuidos: (1) Comunicación entre procesos: Habilitando a dos procesos diferentes para comunicarse entre sí por medio de primitivas del sistema operativo como tuberías, flujos y datagramas en una arquitectura cliente-servidor; (2) Invocación remota: Mecanismos y conceptos que permiten a un proceso en un espacio de direcciones afectar la ejecución de operaciones, procedimientos y métodos en otro espacio de direcciones; y (3) Comunicación indirecta: Mecanismos que permiten el intercambio de mensajes entre uno y muchos procesos a través de un intermediario. En los modelos de comunicación, los procesos emisores y receptores están desacoplados, y la responsabilidad de facilitar el intercambio de mensajes se transfiere al intermediario.

Los sistemas distribuidos toman decisiones entre grupos de procesos cooperantes, cada uno poseyendo estados posiblemente inconsistentes. Los algoritmos de consenso son un mecanismo en el que un subconjunto mayoritario de nodos o 'quorum' puede cumplir con una solicitud de cliente, negociar una verdad y cumplir con una solicitud de cliente. La replicación y la partición son técnicas comunes utilizadas para mejorar la escalabilidad, la confiabilidad y la disponibilidad del sistema.

Referencias



Dominic Lindsay, Sukhpal Singh Gill, Daria Smirnova, Peter Garraghan. "The evolution of distributed computing systems: from fundamental to new frontiers". Computing (2022).