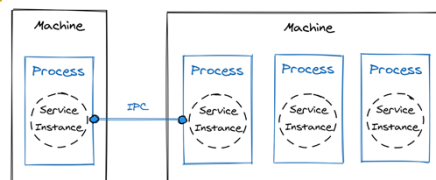


Arquitectura y Computación P2P

Resumen



- Desde la perspectiva física, el sistema distribuido se desplegará en un conjunto de máquinas que se interconectan entre sí a través de una red de datos dando la impresión de un único sistema.
- Desde la perspectiva lógica (tiempo de ejecución) es un conjunto de procesos que se comunican vía Interprocess Communication (IPCs) (p.ej., sockets, RPCs, RMI, HTTP, Mensajería, etc).
- Desde otro punto de vista se puede ver como un conjunto de componentes débilmente acoplados los cuales pueden ser desplegados y escalados de forma independiente.
- Cada uno de estos componentes, ofrece un servicio que implementa una parte específica del sistema global.
- En el corazón de cada servicio, esta la lógica de negocio. El servicio se comunica a través de interfaces muy bien definidas.

Clasificación de Sistemas de Computación y SD

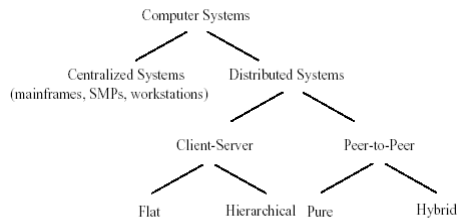
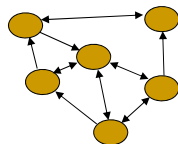


Figure 4: A Taxonomy of Computer Systems Architectures.

Peer-to-Peer (P2P)

- Computación Peer-to-Peer fue INSPIRADO como una alternativa para compartir música con servicios como Napster.
- En vez de tener todos los archivos en un mismo servidor, los archivos de música residen en las máquinas de los usuarios (peers).
- Teóricamente, se ve como un modelo muy escalable, porque NO depende de un punto central.

P2P



Not P2P

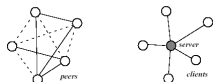
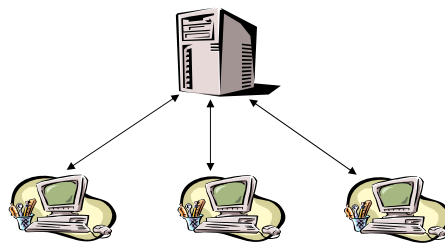
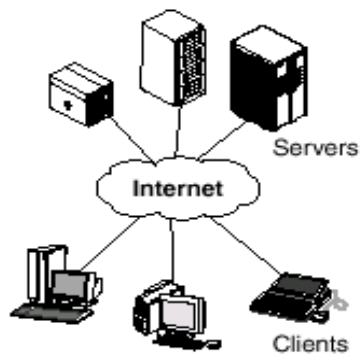


Figure 1: Simplified, High-Level View of Peer-to-Peer versus Centralized (Client-Server) Approach.

P2P como alternativa al C/S

■ Modelo Cliente/Servidor



Problemas del Modelo C/S

■ Escalabilidad

- Cuando aumenta el número de usuarios, recursos, etc, la demanda de computación, almacenamiento, ancho de banda, etc aumenta del lado del servidor?
- Solución hoy en día?
 - Cloud Computing
 - Arquitecturas C/S Escalables.

■ Confiabilidad

- Que pasa si se cae el Servidor?
 - Se va el servicio.
- Solución?
 - Esquemas de tolerancia a fallos y Cloud Computing.

¿Cómo es la Arquitectura P2P?

- Es una tecnología que habilita cualquier dispositivo de red a ofrecer servicios a otros dispositivos.
- Un Peer en una red P2P actual tanto como Cliente como Servidor, o sea que en el fondo, se presentan relaciones C/S entre peers.

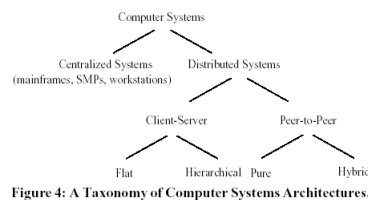
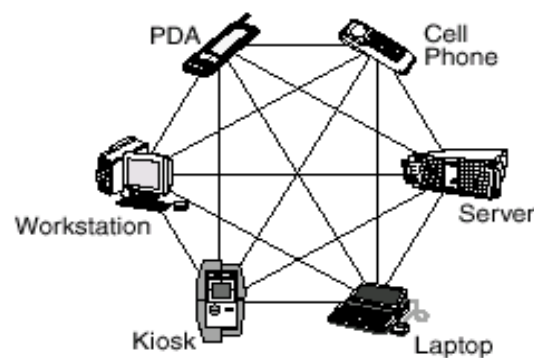


Figure 4: A Taxonomy of Computer Systems Architectures.

Modelo P2P



Ventajas del P2P

- ♦ La comunicación es directa entre los peer, elimina retardos (por ejemplo en aplicaciones como Skype)
- ♦ El sistema puede sobrevivir a cambios extremos de la red.
- ♦ Crecimiento del número de clientes.
- ♦ Tolerancia a fallos

Retos

- ¿Cómo localizar los recursos de la red P2P?
 - Indexar, Descubrir y Buscar RECURSOS.
- Indexación
- Búsqueda (local vs distributed – flooding, supernodes, DHT)
- Descubrimiento de Peers & Recursos
- Recuperación
- Seguridad

Propuestas de Arquitecturas P2P

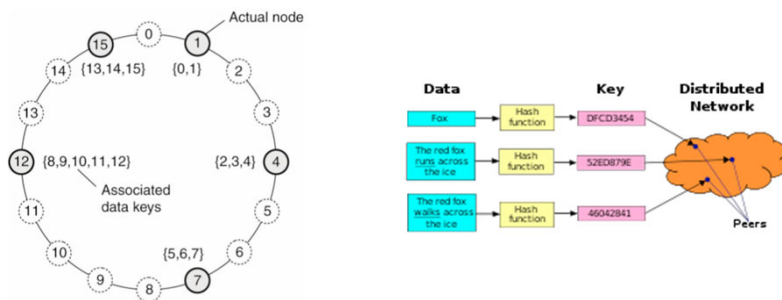
- Redes P2P Estructuradas y P2P NO estructuradas.
- Muchas variantes para las Estructuradas
- Para las NO estructuradas:
 - Redes P2P puras, Híbridas basadas en SuperPeers, Híbridas Basadas en “Servidor”
 - Que hace este “Servidor”?
 - Muchas cosas, menos tener los recursos

Arquitectura General P2P

- **Network Communication:** conectividad IP
- **Overlay Nodes Management:** Descubrimiento de Peers y Algoritmos de enrutamiento
- **Feature Management:** Seguridad, Confiabilidad, Tolerancia a fallos, Disponibilidad de recursos, etc
- **Services Specific:** Soporta la infraestructura subyacente P2P y los componentes específicos de aplicación tales como Planificación, contenido y gestión de archivos.
- **Metadata:** Descripción de los recursos y su localización
- **Application:** Herramientas, Aplicaciones y servicios sobre la P2P

P2P Estructurado

- Conexiones 'fijas' entre peers
- Usa algoritmos como DHT(Distributed Hash Table) para mantener la red. i.e. Chord (MIT)



P2P Estructurado

- Emplea un protocolo global consistente para asegurar que cualquier nodo pueda eficientemente enrutar una búsqueda a algún otro peer a un destino deseado
- DHT
 - ❑ Servicio de búsqueda similar a una: (*key*, *value*)
 - ❑ Escala muy bien (add, remove or failure of peers)

P2P NO Estructurado

- No hay algoritmo para la conexión entre peers.
- 3 modelos
 - Pure P2P
 - Hybrid P2P? 2 variantes: Serverless & Server-based.
 - Centralized P2P (server-based)

Ejemplos de aplicaciones

P2P Computing

- Distributed computing (e.g., SETI@home, Search for Extraterrestrial Intelligence)
- File sharing (e.g., Gnutella, Freenet, Publius, Free Haven)
- Collaboration (e.g., Magi, Groove, Jabber)
- Platforms (e.g., JXTA)

Distributed Computing: SETI@home

- Search for Extraterrestrial Intelligence that has over two million computers crunching away and downloading data gathered from the Arecibo radio telescope in Puerto Rico
- The SETI@Home project is widely regarded as the fastest computer in the world
- In fact, the project has already performed **the single largest cumulative computation to date**
- From the architecture point of view Seti@Home is based upon client-server
- The centralised servers hold enormous amounts of data gathered from the Arecibo radio telescope "listening" to the skies
- That data needs to be analysed for distinct or unusual radio waves that might suggest extraterrestrial communications
- <http://setiathome.ssl.berkeley.edu>

File Sharing: Napster

- Here is how Napster used to work:
 - It is first necessary to download and install Napster software and to sign up for a free account at Napster
 - By clicking the Napster icon, a connection is established with Napster Web site, where the name of the music searched can be entered in a search window
 - The Napster software tracks all users who are online at that particular time and provides access to tracks stored on users' hard drives
 - When the song requested is found, Napster establishes a connection between the two computers so that it could be downloaded
 - However, after the Recording Industry Association of America sued Napster for copyright infringements, the court ruled that Napster was indeed violating copyrights and it is shut down

File Sharing: Gnutella

- Gnutella is a file sharing protocol
- **Applications that implement the Gnutella protocol allow users to search for and download files from other users connected to the Internet**
- **Its architecture is completely decentralised and distributed**, meaning that there are no central servers and that all computations and interactions happen between clients
- **When a client wishes to connect to the network they run through a list of nodes that are most likely to be up or take a list from a website and then connect to how ever many nodes they want**

File Sharing: Gnutella

- Routing in the network is accomplished through broadcasting
- When a search request arrives into a client that client searches itself for the file and broadcasts the request to all its other connections
- Broadcasts are cut off by a time to live that specifies how many hops they may cover before clients should drop them rather than broadcast them

Collaborative Application: Groove

- Groove is a collaborative P2P system (<http://www.groove.net>)
- It is mainly targeted to Internet and intranet users, although it can also be used on mobile devices, such as PDAs, mobile phones, and tablets
- It is intended to enable communication, content sharing, and tools for joint activities