

Comunicación en Grupo

30221 - Sistemas Distribuidos

Rafael Tolosana

Dpto. Informática e Ing. de Sistemas

Lectura Recomendada

- Coulouris, Dollimore, Kindberg and Blair, Distributed Systems: Concepts and Design Edn. 5, 2012. Capítulo 6 y Sección 15.4
- Tanenbaum and Van Steen, Distributed Systems: Principles and Paradigms, 3e, (c) 2017 Prentice- Hall. Secciones 4.4, 6.7, 8.4

Motivación

Motivación

Comunicación entre Procesos

<i>Communicating entities</i> (what is communicating)		<i>Communication paradigms</i> (how they communicate)		
<i>System-oriented entities</i>	<i>Problem-oriented entities</i>	<i>Interprocess communication</i>	<i>Remote invocation</i>	<i>Indirect communication</i>
Nodes	Objects	Message passing	Request-reply	Group communication
Processes	Components	Sockets	RPC	Publish-subscribe
	Web services	Multicast	RMI	Message queues
				Tuple spaces

1

¹Instructor's Guide for Coulouris, Dollimore and Kindberg Distributed Systems: Concepts and Design Edn.
4©Pearson Education 2005

Motivación

Formas de Comunicación entre Procesos

- Comunicación Directa
 - Sockets IPC: TCP / UDP
 - RPC
 - Canales síncronos / asíncronos
- Comunicación Indirecta (mediada)
 - Linda, publish-subscribe, colas de mensajes, etc.

Motivación

Formas de Comunicación entre Procesos

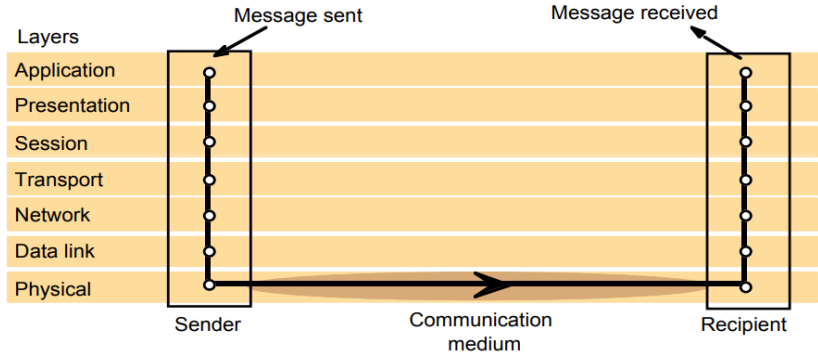
- Comunicación Directa
 - Sockets IPC: TCP / UDP
 - RPC
 - Canales síncronos / asíncronos
- Comunicación Indirecta (mediada)
 - Linda, publish-subscribe, colas de mensajes, etc.

Protocolos de Comunicación

- Formato de los Mensajes
- Inter-Process Communication (IPC)
 - Sockets IPC: TCP / UDP
 - RPC
 - Canales síncronos / asíncronos

Motivación

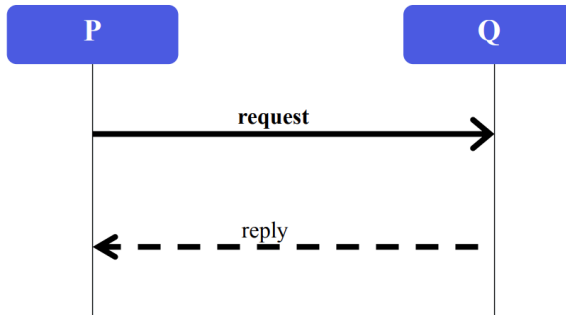
Protocolos de Comunicación



Motivación

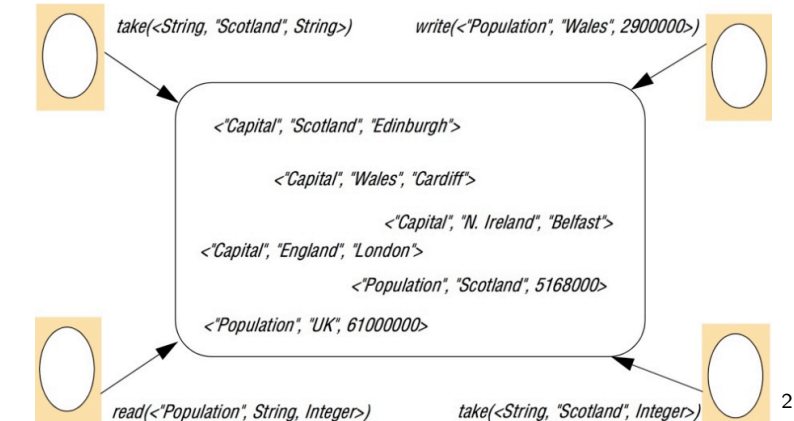
Protocolos de Interacción

- Conjunto de mensajes intercambiados entre procesos para conseguir una sincronización / coordinación



Motivación

Linda



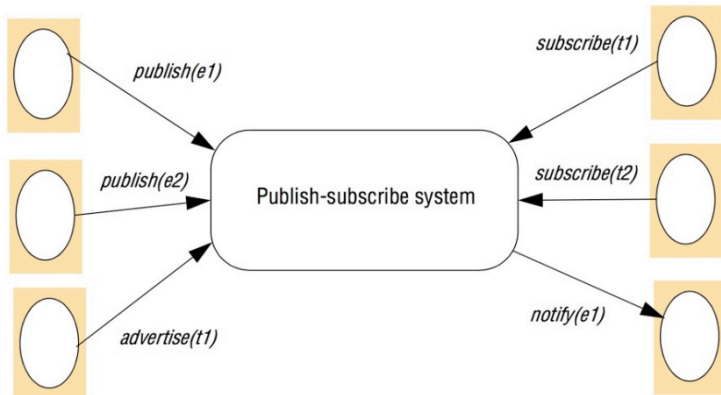
²Instructor's Guide for Coulouris, Dollimore and Kindberg Distributed Systems: Concepts and Design Edn. 4 © Pearson Education 2005

Motivación

Publish-Subscribe

Publishers

Subscribers



3

³Instructor's Guide for Coulouris, Dollimore and Kindberg Distributed Systems: Concepts and Design Edn.
⁴©Pearson Education 2005

Motivación

Comunicación entre Procesos

- La mayor parte de la comunicación entre procesos en SSDD se hace entre pares.
- A veces hay también necesidad de comunicar 1-N
- La comunicación 1-N puede *simularse* enviando el mismo mensaje N veces (N destinatarios).

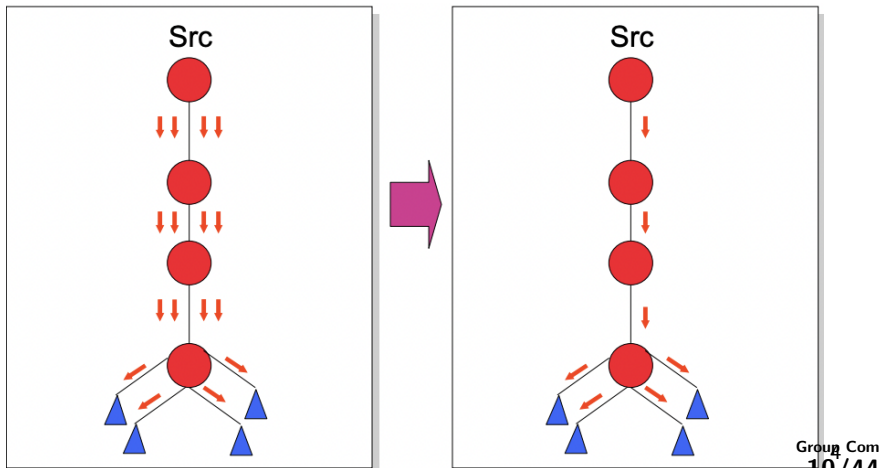
Motivación

Comunicación entre Procesos

- La mayor parte de la comunicación entre procesos en SSDD se hace entre pares.
- A veces hay también necesidad de comunicar 1-N
- La comunicación 1-N puede *simularse* enviando el mismo mensaje N veces (N destinatarios).
- **¿Es eficiente?**

Motivación

Eficiencia comunicación 1-N “simulada”



Motivación

Comunicación entre Procesos

- La mayor parte de la comunicación entre procesos en SSDD se hace entre pares.
- A veces hay también necesidad de comunicar 1-N
- La comunicación 1-N puede *simularse* enviando el mismo mensaje N veces (N destinatarios).
- **¿Es eficiente?**
- **¿Qué sucede con a tolerancia a fallos?**

Motivación

¿Necesidad de Comunicación entre múltiples procesos?

Motivación

¿Necesidad de Comunicación entre múltiples procesos?

operation acquire_mutex() **is**

- (1) $cs_state_i \leftarrow \text{trying};$
- (2) $\ell rd_i \leftarrow clock_i + 1;$
- (3) $waiting_from_i \leftarrow R_i; \quad \% R_i = \{1, \dots, n\} \setminus \{i\}$
- (4) **for each** $j \in R_i$ **do** send REQUEST($\ell rd_i, i$) to p_j **end for**;
- (5) **wait** ($waiting_from_i = \emptyset$);
- (6) $cs_state_i \leftarrow in.$

operation release_mutex() **is**

- (7) $cs_state_i \leftarrow out;$
- (8) **for each** $j \in perm_delayed_i$ **do** send PERMISSION(i) to p_j **end for**;
- (9) $perm_delayed_i \leftarrow \emptyset.$

when REQUEST(k, j) **is received do**

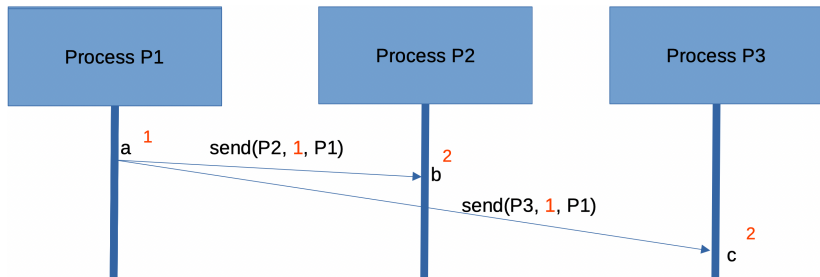
- (10) $clock_i \leftarrow \max(clock_i, k);$
- (11) $prio_i \leftarrow (cs_state_i \neq out) \wedge (\ell rd_i, i) < (k, j);$
- (12) **if** ($prio_i$) **then** $perm_delayed_i \leftarrow perm_delayed_i \cup \{j\}$
- (13) **else** send PERMISSION(i) to p_j
- (14) **end if.**

when PERMISSION(j) **is received do**

- (15) $waiting_from_i \leftarrow waiting_from_i \setminus \{j\}.$

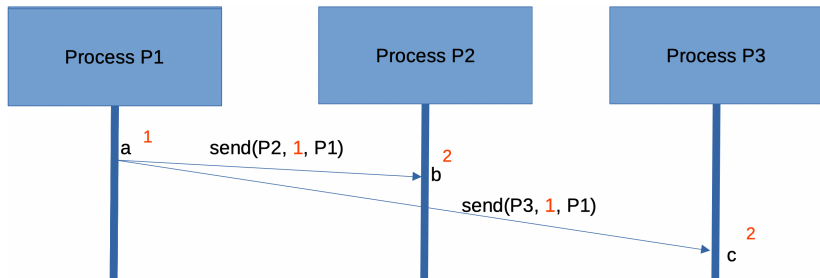
Motivación

Solicitud de Acceso SC en RA



Motivación

Solicitud de Acceso SC en RA



¡Comunicación de grupo (multicast)!

Motivación

El Problema Mutex Distribuido

- Un grupo de procesos se ordenan para acceder a la SC en excl mútua

El Problema Consenso Distribuido

- Un grupo de procesos acuerdan qué valores son adecuados bajo la presencia de fallos

Motivación

El Problema Mutex Distribuido

- Un grupo de procesos se ordenan para acceder a la SC en excl mútua

El Problema Consenso Distribuido

- Un grupo de procesos acuerdan qué valores son adecuados bajo la presencia de fallos

Publish-Subscribe

- N fuentes de datos generan información periódicamente
- M consumidores están interesados en algunas fuentes de datos
- Un sistema se encarga de comunicar las fuentes de datos con los consumidores

Historia

Historia

- Hasta 1993 la comunicación 1-1 funcionaba correctamente

⁵<https://tldp.org/HOWTO/Multicast-HOWTO-1.html>

Historia

- Hasta 1993 la comunicación 1-1 funcionaba correctamente
- La comunicación 1-1 como implementación para comunicar N-M

⁵<https://tldp.org/HOWTO/Multicast-HOWTO-1.html>

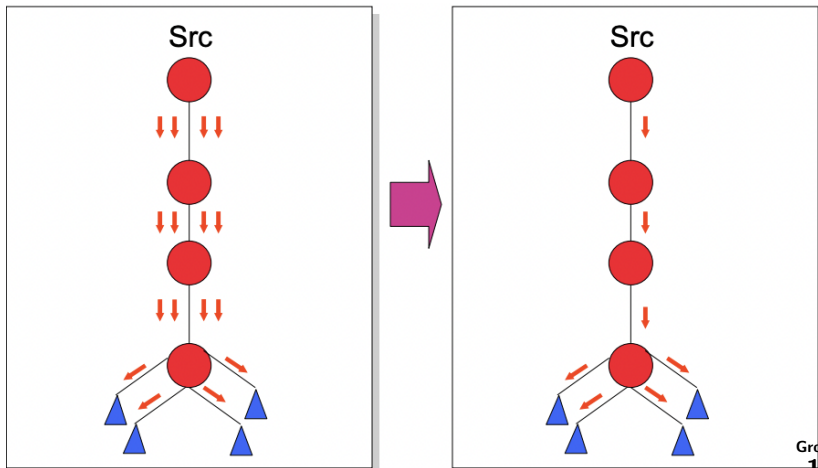
Historia

- Hasta 1993 la comunicación 1-1 funcionaba correctamente
- La comunicación 1-1 como implementación para comunicar N-M
- Algo cambió...: vídeo ⁵

⁵<https://tldp.org/HOWTO/Multicast-HOWTO-1.html>

Historia

Eficiencia comunicación 1-N “simulada”



Comunicación en Grupo

Comunicación en Grupo

¿Necesidad de Comunicación entre múltiples procesos?

Un aspecto importante de los SSDD es la capacidad de enviar mensajes a múltiples procesos a la vez:

COMUNICACIÓN MULTICAST

Comunicación en Grupo

¿Necesidad de Comunicación entre múltiples procesos?

Un aspecto importante de los SSDD es la capacidad de enviar mensajes a múltiples procesos a la vez:

COMUNICACIÓN MULTICAST

- Históricamente ha sido un aspecto de investigación exclusivo de redes de comunicación

Comunicación en Grupo

¿Necesidad de Comunicación entre múltiples procesos?

Un aspecto importante de los SSDD es la capacidad de enviar mensajes a múltiples procesos a la vez:

COMUNICACIÓN MULTICAST

- Históricamente ha sido un aspecto de investigación exclusivo de redes de comunicación
- En la práctica, involucraba enormes cantidades de gestión (requiriendo intervención humana).

Comunicación en Grupo

¿Necesidad de Comunicación entre múltiples procesos?

Un aspecto importante de los SSDD es la capacidad de enviar mensajes a múltiples procesos a la vez:

COMUNICACIÓN MULTICAST

- Históricamente ha sido un aspecto de investigación exclusivo de redes de comunicación
- En la práctica, involucraba enormes cantidades de gestión (requiriendo intervención humana).
- Con el avance de las tecnologías, su gestión se ha simplificado significativamente

Comunicación en Grupo

Requisitos Comunicación en Grupo

- Cada grupo se identifica con una dirección IP

Comunicación en Grupo

Requisitos Comunicación en Grupo

- Cada grupo se identifica con una dirección IP
- Los grupos pueden ser de tamaño variable (dinámico)

Comunicación en Grupo

Requisitos Comunicación en Grupo

- Cada grupo se identifica con una dirección IP
- Los grupos pueden ser de tamaño variable (dinámico)
- Los miembros de un grupo se pueden encontrar en cualquier parte de Internet

Comunicación en Grupo

Requisitos Comunicación en Grupo

- Cada grupo se identifica con una dirección IP
- Los grupos pueden ser de tamaño variable (dinámico)
- Los miembros de un grupo se pueden encontrar en cualquier parte de Internet
- Los miembros de un grupo pueden unirse y abandonar al grupo libremente

Comunicación en Grupo

Requisitos Comunicación en Grupo

- Cada grupo se identifica con una dirección IP
- Los grupos pueden ser de tamaño variable (dinámico)
- Los miembros de un grupo se pueden encontrar en cualquier parte de Internet
- Los miembros de un grupo pueden unirse y abandonar al grupo libremente
- La pertenencia a un grupo no se conoce *a priori*

Comunicación en Grupo

Requisitos Comunicación en Grupo

- Cada grupo se identifica con una dirección IP
- Los grupos pueden ser de tamaño variable (dinámico)
- Los miembros de un grupo se pueden encontrar en cualquier parte de Internet
- Los miembros de un grupo pueden unirse y abandonar al grupo libremente
- La pertenencia a un grupo no se conoce *a priori*
- Los emisores no tienen por qué ser miembros del grupo

Comunicación en Grupo

Requisitos Comunicación en Grupo

- Cada grupo se identifica con una dirección IP
- Los grupos pueden ser de tamaño variable (dinámico)
- Los miembros de un grupo se pueden encontrar en cualquier parte de Internet
- Los miembros de un grupo pueden unirse y abandonar al grupo libremente
- La pertenencia a un grupo no se conoce *a priori*
- Los emisores no tienen por qué ser miembros del grupo

Analogía:

- Radio frecuencia: cualquiera puede transmitir y cualquiera puede conectarse y escuchar

Comunicación en Grupo

Pertenencia al Grupo

- Los mensajes al grupo se reciben por todos sus miembros
- Si los procesos se añaden / borran de un grupo (fallo de proceso, cambio en la red, cambio en las preferencias de usuario, etc.), se tiene que notificar el cambio a todos los miembros activos del grupo, manteniendo la consistencia en todos ellos.
- Todo mensaje se entrega bajo ciertas configuraciones de precisión. No obstante, siempre se quiere garantizar atomicidad, uniformidad y terminación.

Comunicación en Grupo

Retos de Comunicación en grupo

- Ordenación de mensajes para los miembros del grupo
- Garantías de Entrega
- Gestión de la pertenencia a grupo
- Tolerancia a fallos

Comunicación en Grupo

Retos de Comunicación en Grupo

- Vivacidad y Tolerancia a fallos
 - Los mensajes terminan llegando (fiabilidad en entrega)

Comunicación en Grupo

Retos de Comunicación en Grupo

- Vivacidad y Tolerancia a fallos
 - Los mensajes terminan llegando (fiabilidad en entrega)
- Atomic Multicast
 - Un mensaje enviado a un grupo, o se entrega a todos o nadie lo recibe

Comunicación en Grupo

Retos de Comunicación en Grupo

- Vivacidad y Tolerancia a fallos
 - Los mensajes terminan llegando (fiabilidad en entrega)
- Atomic Multicast
 - Un mensaje enviado a un grupo, o se entrega a todos o nadie lo recibe
- Orden de Llegada de los mensajes
 - ¿Es el orden de llegada como el de emisión?

Comunicación en Grupo

Retos de Comunicación en Grupo

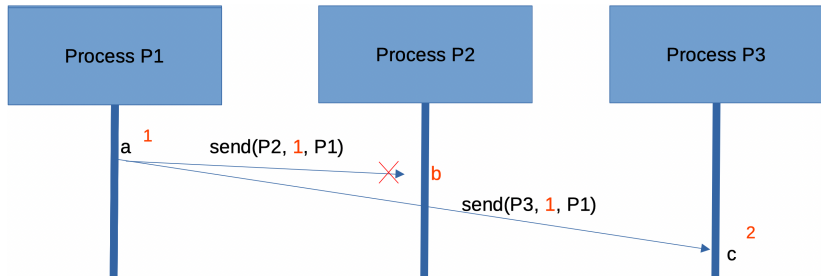
- Vivacidad y Tolerancia a fallos
 - Los mensajes terminan llegando (fiabilidad en entrega)
- Atomic Multicast
 - Un mensaje enviado a un grupo, o se entrega a todos o nadie lo recibe
- Orden de llegada de los mensajes
 - ¿Es el orden de llegada como el de emisión?
- Gestión de la pertenencia al grupo
 - Dinamismo y garantías de identidad anónima

Ordenación de Mensajes

- Best-Effort broadcast (sin control alguno)
- Reliable broadcast
- FIFO broadcast
- Causal broadcast
- Total Order broadcast

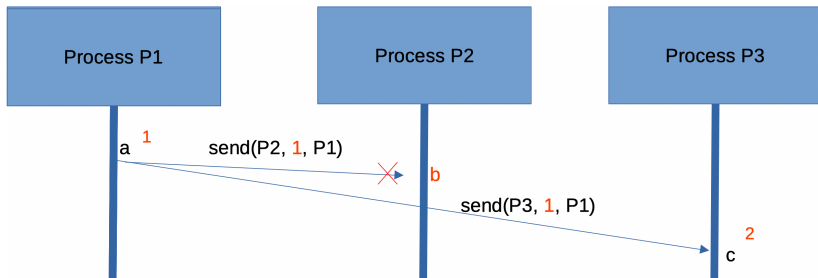
Ordenación de Mensajes

Best-Effort (sin control alguno)



Ordenación de Mensajes

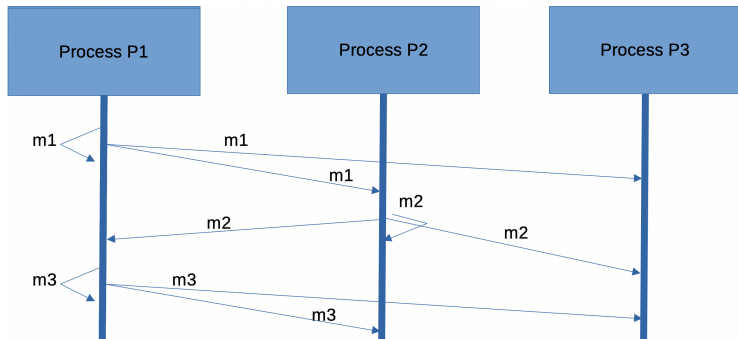
Reliable Multicast



Se producen retransmisiones hasta que llega el mensaje

Ordenación de Mensajes

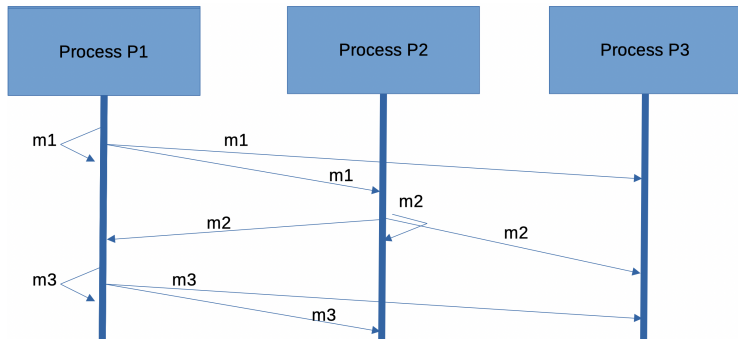
FIFO Multicast



Los mensajes enviados por un nodo tienen que llegar en el mismo orden

Ordenación de Mensajes

FIFO Multicast

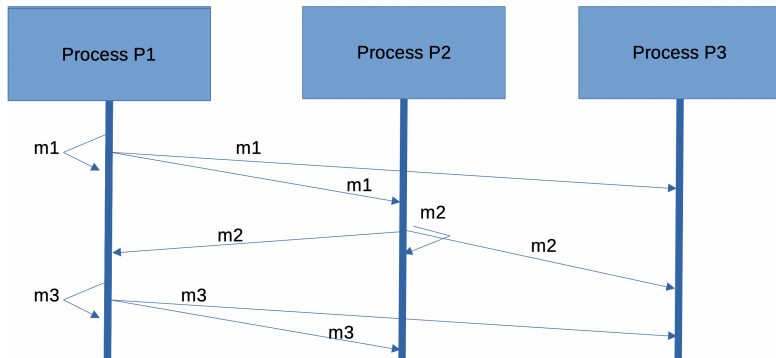


Los mensajes enviados por un nodo tienen que llegar en el mismo orden

Órdenes válidos: (m1, m2, m3), (m2, m1, m3), (m1, m3, m2)

Ordenación de Mensajes

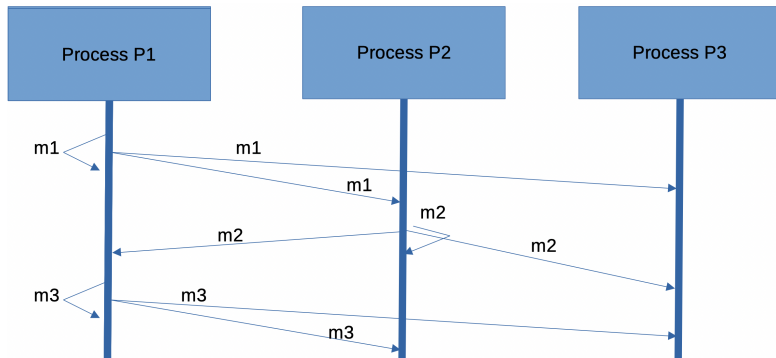
Causal Broadcast



Los mensajes con relaciones de causalidad tienen que llegar en orden causal

Ordenación de Mensajes

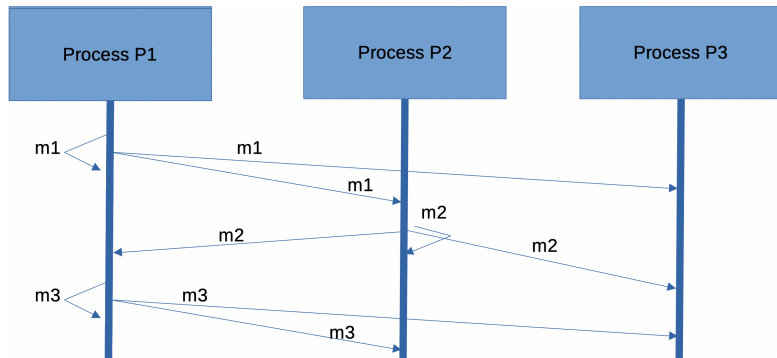
Causal Broadcast



Los mensajes con relaciones de causalidad tienen que llegar en orden causal

Ordenación de Mensajes

Causal Broadcast

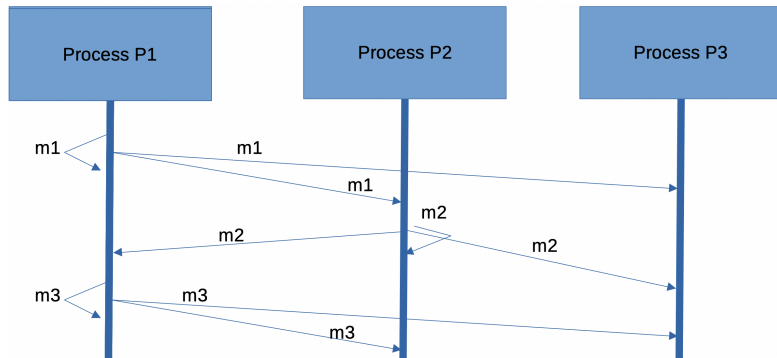


$\text{broadcast}(m1) \rightarrow \text{broadcast}(m2)$

$\text{broadcast}(m1) \rightarrow \text{broadcast}(m3)$

Ordenación de Mensajes

Causal Broadcast



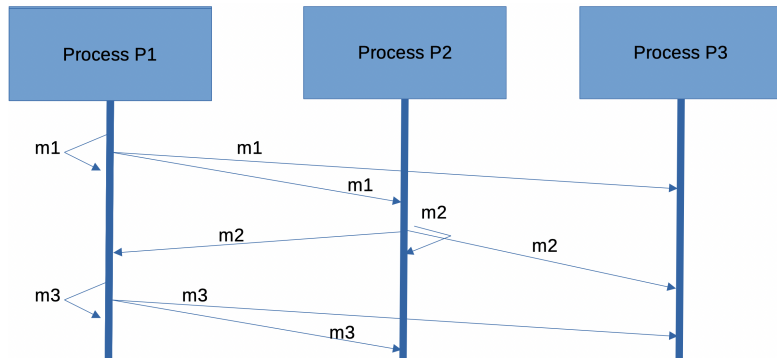
$\text{broadcast}(m1) \rightarrow \text{broadcast}(m2)$

$\text{broadcast}(m1) \rightarrow \text{broadcast}(m3)$

Órdenes válidos: (m1, m2, m3), (m1, m3, m2)

Ordenación de Mensajes

Causal Broadcast



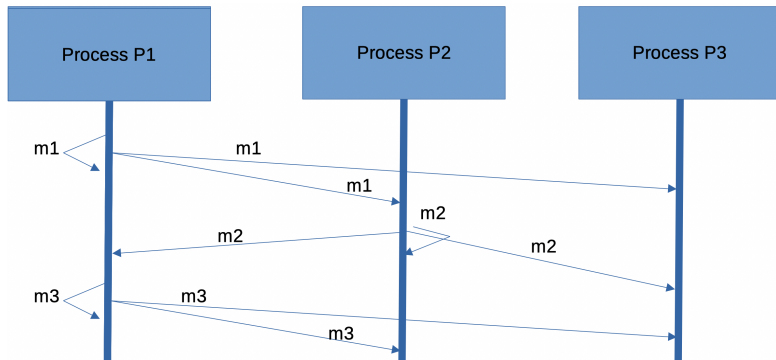
$\text{broadcast}(m1) \rightarrow \text{broadcast}(m2)$

$\text{broadcast}(m1) \rightarrow \text{broadcast}(m3)$

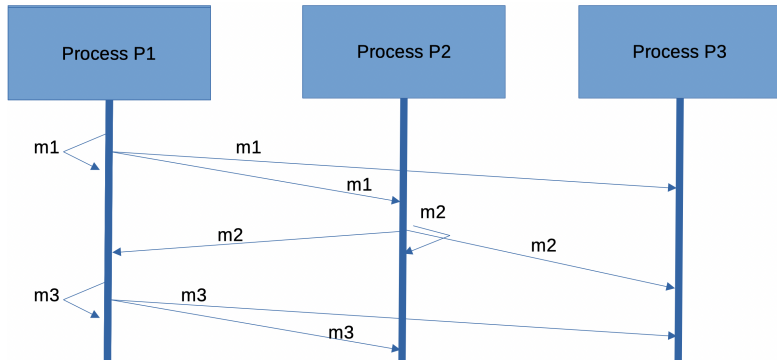
Órdenes válidos: (m1, m2, m3), (m1, m3, m2)

Ordenación de Mensajes

Total Order Broadcast



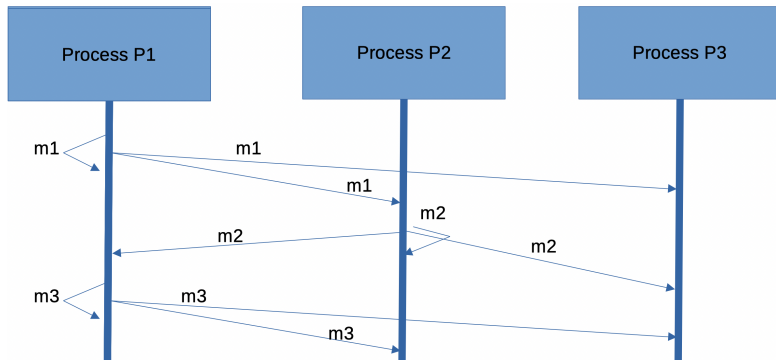
Todos los nodos tienen el mismo orden de recepción.



Todos los nodos tienen el mismo orden de recepción.
El orden que se determine

Ordenación de Mensajes

Total Order Broadcast



Todos los nodos tienen el mismo orden de recepción.
El orden que se determine

Diseño e Implementación de Comunicación en Grupo

Diseño e Implementación de Comunicación en Grupo

¿Qué alternativas hay para la comunicación en grupo?

- El emisor envía el mensaje repetidamente punto a punto: unicast
 - Poco eficiente en recursos

Diseño e Implementación de Comunicación en Grupo

¿Qué alternativas hay para la comunicación en grupo?

- El emisor envía el mensaje repetidamente punto a punto: unicast
 - Poco eficiente en recursos
- IP Multicast
 - Soporte hardware

Diseño e Implementación de Comunicación en Grupo

¿Qué alternativas hay para la comunicación en grupo?

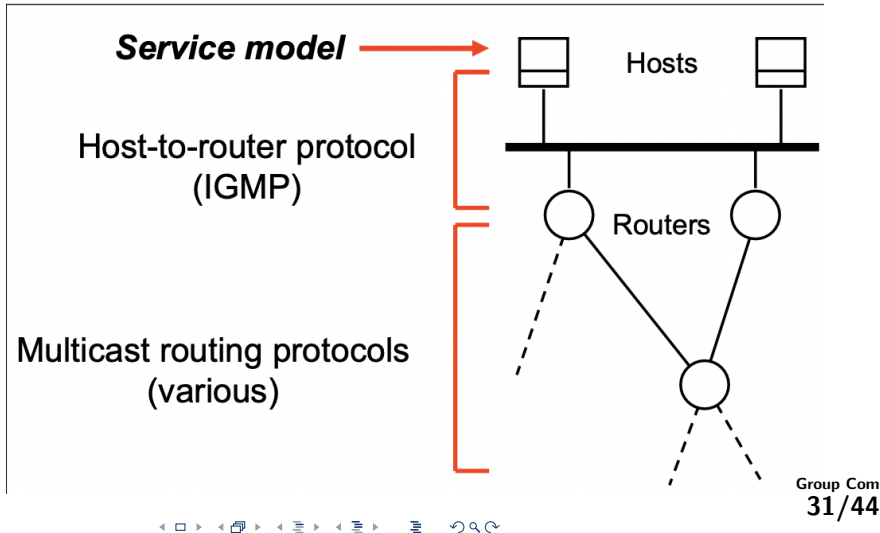
- El emisor envía el mensaje repetidamente punto a punto: unicast
 - Poco eficiente en recursos
- IP Multicast
 - Soporte hardware
- Overlay Multicast: Comunicación en Grupo en Internet

IP Multicast

- IP multicast service basics
- Multicast routing

IP Multicast

IP multicast service basics



IP Multicast

IP multicast service basics

- IP Multicast típicamente se construye sobre UDP: no fiable

IP Multicast

IP multicast service basics

- IP Multicast típicamente se construye sobre UDP: no fiable
- Operaciones de recepción / envío
- Operaciones de gestión de pertenencia al grupo

IP Multicast

IP multicast service basics

	<i>octet 1</i>	<i>octet 2</i>	<i>octet 3</i>	<i>Range of addresses</i>
	<i>Network ID</i>		<i>Host ID</i>	
Class A:	1 to 127	0 to 255	0 to 255	0 to 255
	<i>Network ID</i>		<i>Host ID</i>	
Class B:	128 to 191	0 to 255	0 to 255	0 to 255
	<i>Network ID</i>		<i>Host ID</i>	
Class C:	192 to 223	0 to 255	0 to 255	1 to 254
	<i>Multicast address</i>			
Class D (multicast):	224 to 239	0 to 255	0 to 255	1 to 254
Class E (reserved):	240 to 255	0 to 255	0 to 255	1 to 254

IP Multicast

Multicast routing

- Para el encaminamiento de los paquetes, los routers tienen que gestionar estados grandes en tamaño
- Presenta problemas potenciales de denegación de servicio
- Es un servicio costoso económicamente

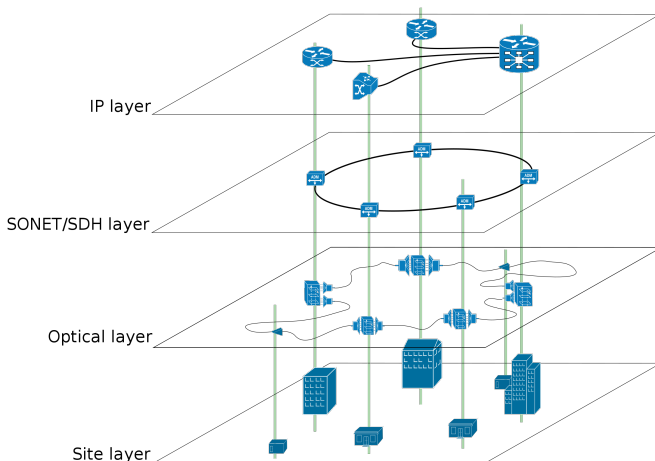
IP Multicast

Multicast routing

- Para el encaminamiento de los paquetes, los routers tienen que gestionar estados grandes en tamaño
- Presenta problemas potenciales de denegación de servicio
- Es un servicio costoso económicamente
- Los proveedores de Internet no suelen dar servicio IP Multicast
- IP Multicast es una buena solución para una LAN que uno controla, pero EN GENERAL NO PARA INTERNET

Network Overlay

Comunicación en Grupo en Internet



Network Overlay

Overlay network

- Puede entenderse como una red de computadores que se construye sobre otra red. Todos los nodos de la *overlay network* se conectan mediante enlaces lógicos o virtuales, y esos enlaces constituyen un camino en la red subyacente.

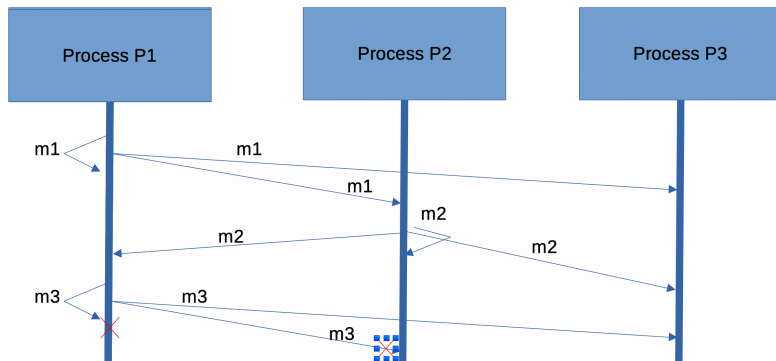
Network Overlay

Overlay network

- Puede entenderse como una red de computadores que se construye sobre otra red. Todos los nodos de la *overlay network* se conectan mediante enlaces lógicos o virtuales, y esos enlaces constituyen un camino en la red subyacente.
- Un ejemplo de overlay network pueden ser los sistemas distribuidos cliente servidor o peer-to-peer. Esos sistemas y sus redes pueden verse como una overlay network puesto que se construyen sobre Internet.

Network Overlay

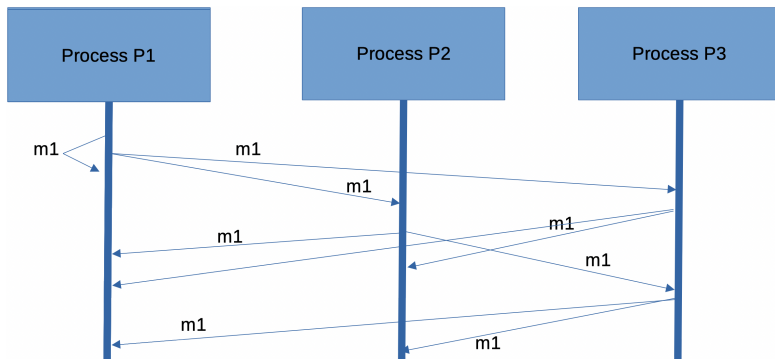
Primer intento: punto a punto



Incluso teniendo IP multicast fiable, si un proceso falla, no se retransmiten sus mensajes

Network Overlay

Segundo intento: retransmisiones exhaustivas



Se asegura la retransmisión pero a un coste altísimo

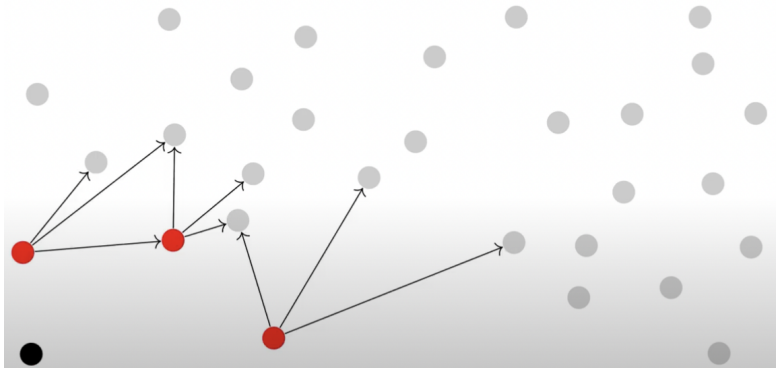
Network Overlay

Tercer intento: gossiping



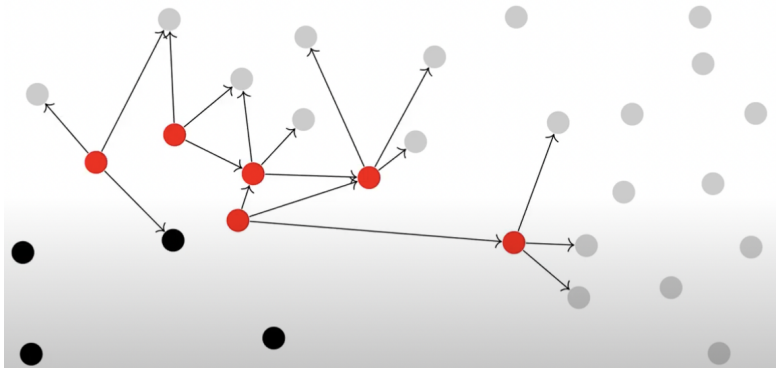
Network Overlay

Tercer intento: gossiping



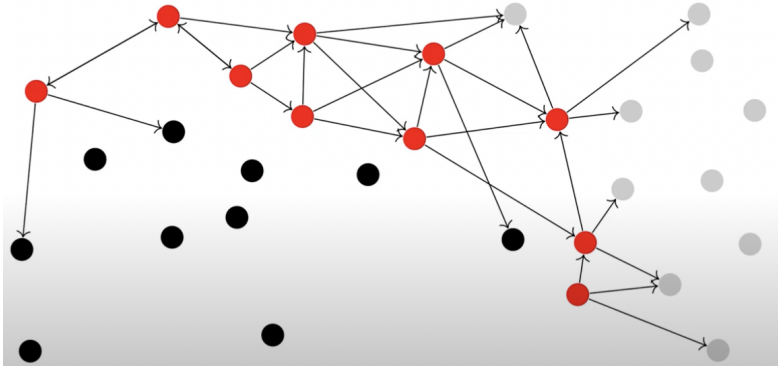
Network Overlay

Tercer intento: gossiping



Network Overlay

Tercer intento: gossiping



Network Overlay

Tercer intento: gossiping



Conclusiones

Conclusiones

Comunicación en grupo

- Los sistemas distribuidos requieren de comunicación en grupo
- Muchas son las aplicaciones que requieren enviar mensajes a un conjunto de procesos.
- La comunicación punto a punto (unicast) es la más extendida pero es ineficiente respecto de la red
- La comunicación en grupo plantea muchos retos: dinamismo de los grupos, privacidad, seguridad, orden de los mensajes y carencias en los proveedores de internet para implementarlo
- En las LAN puede usarse IP Multicast, en Internet no en general
- Para implementar la comunicación en grupo hay otras alternativas a IP Multicast, pero de menor eficiencia

Comunicación en Grupo

30221 - Sistemas Distribuidos

Rafael Tolosana

Dpto. Informática e Ing. de Sistemas