

Grupo de investigación de sistemas distribuidos (E.U.I.)

Sergio Arévalo y Ernesto Jiménez

Lsd

Distributed
Systems

Laboratory

Universidad Politécnica de Madrid (UPM)

<http://lsd.ls.fi.upm.es/lsd/lsd.htm>

Laboratorio de sistemas distribuidos (LSD)

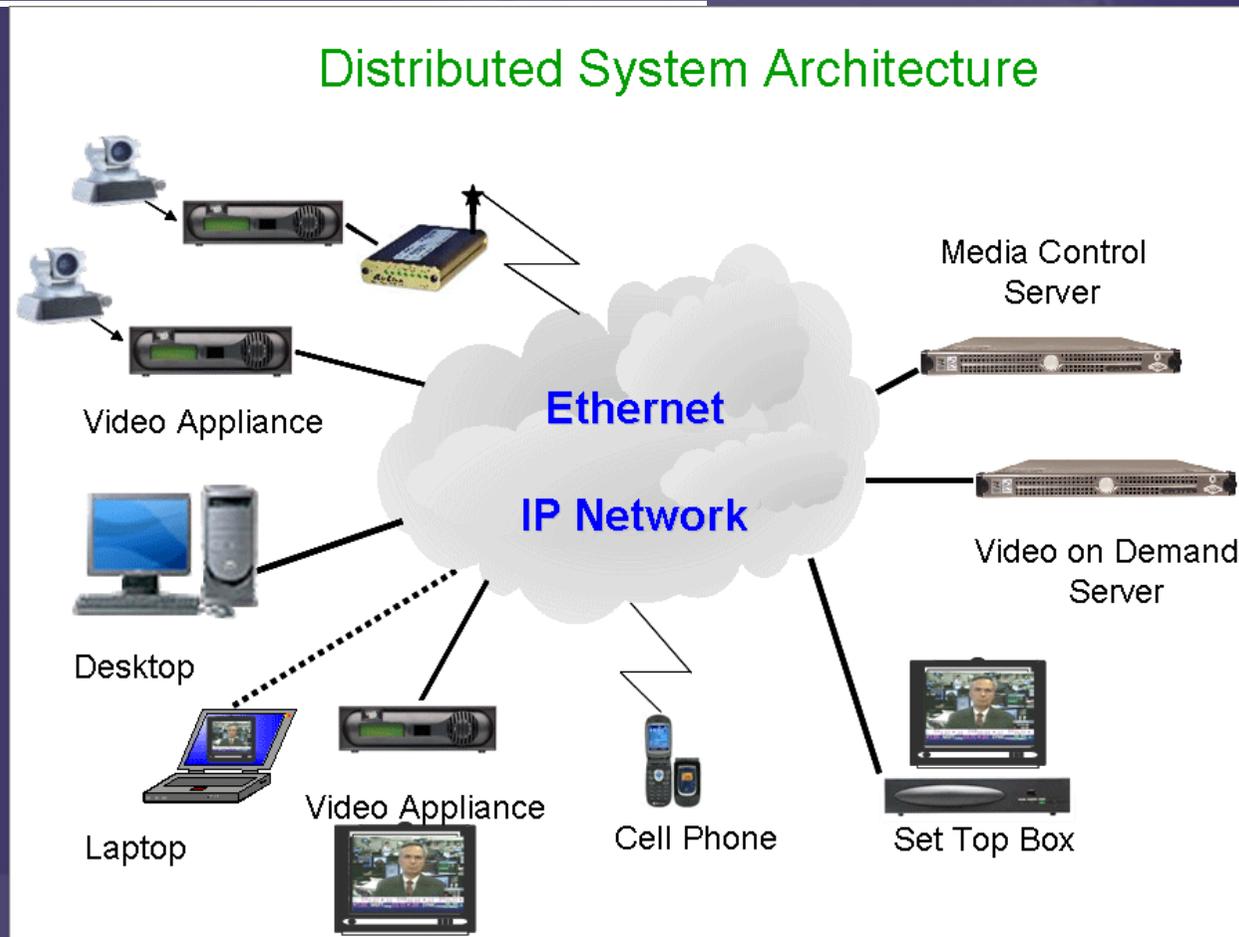
- Localización:
 - Escuela Universitaria de Informática (EUI):
 - Departamento de Informática Aplicada (I.A.).
 - Departamento de Arquitectura y Tecnología de Computadores (A.T.C.).
 - Facultad de Informática.
- Líneas de investigación:
 - Cloud-computing (computación en nube).
 - Tolerancia a fallos.
 - Coordinación y cooperación en sistemas distribuidos dinámicos.
 - Memoria compartida distribuida.
 - Bases de datos distribuidas.
 - Transacciones.

Financiación del LSD-EUI

- Principales proyectos de investigación abiertos:
 - CLOUDS: Cloud computing para servicios escalables, confiables y ubicuos. Comunidad de Madrid. 2010-2013.
 - CloudStorm: Scalable and Dependable Cloud Service Platforms. Ministerio de Ciencia e Innovación. 2010-2013.
- Proyectos en evaluación:
 - Sistemas distribuidos dinámicos y ambientes inteligentes. Ministerio de Ciencia e Innovación. 2011-2014
- Becas:
 - China Scholarship Council + UPM

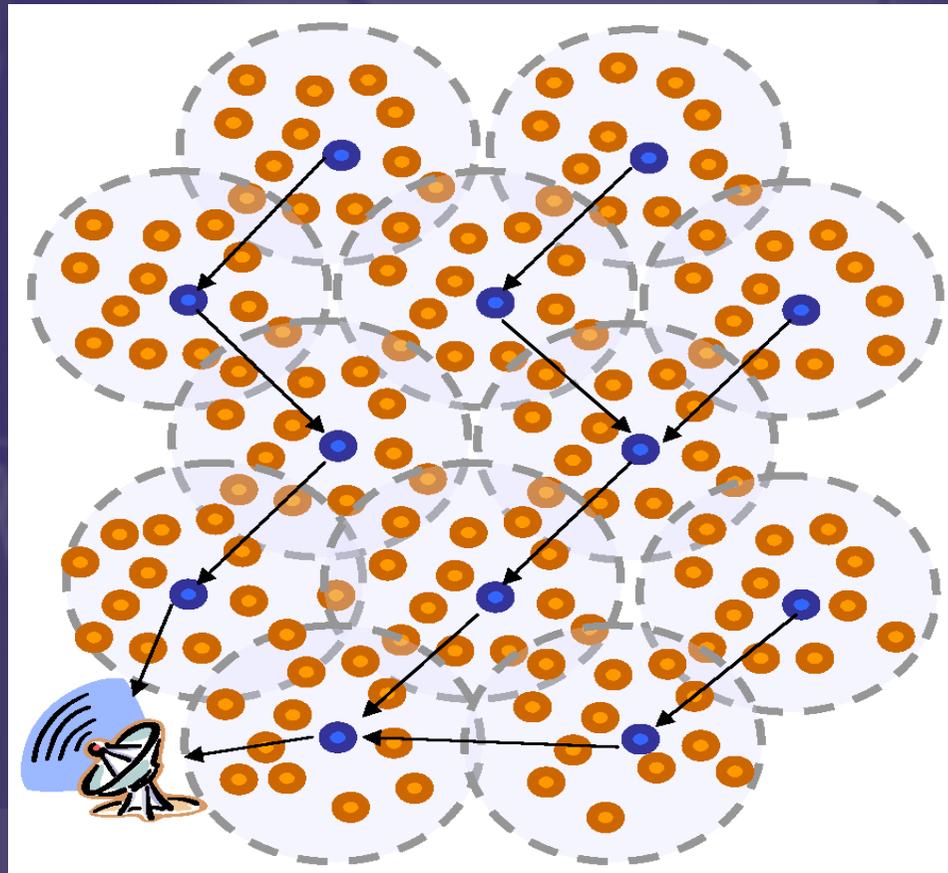
Líneas de investigación del LSD-EUI

- Sistemas Distribuidos:



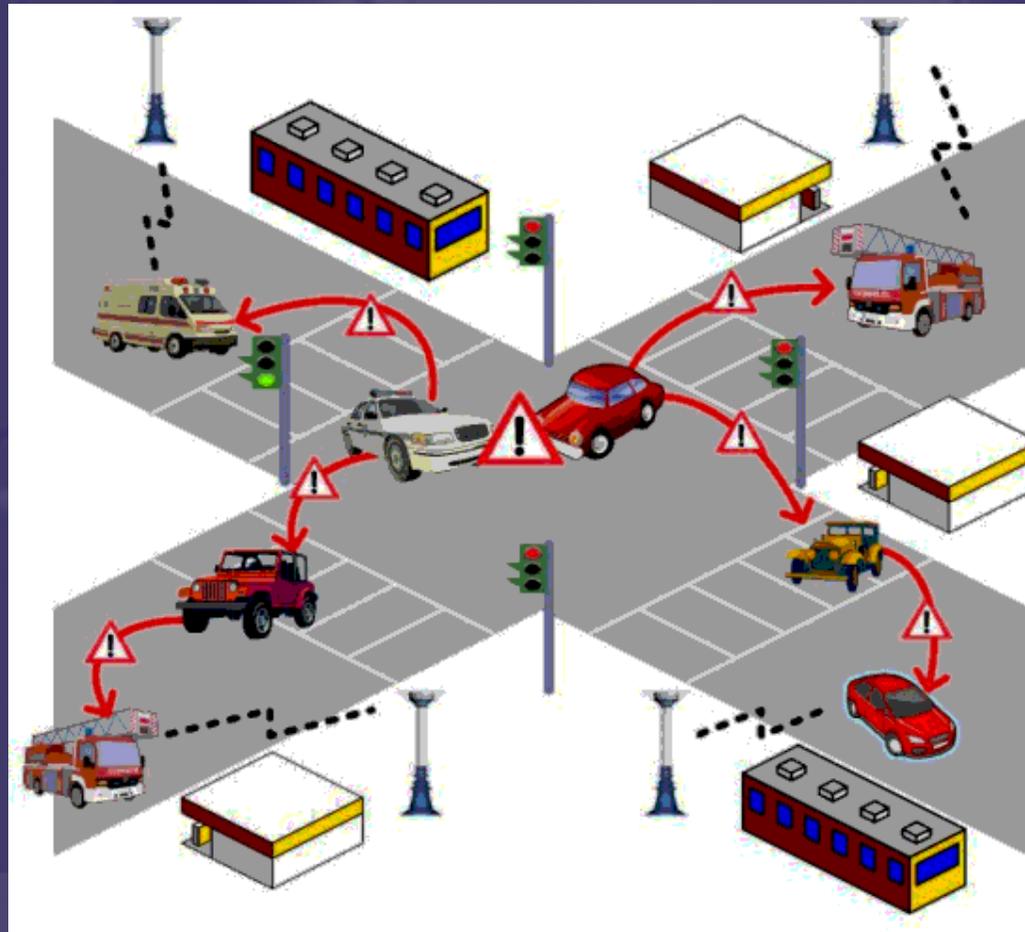
Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos: redes de sensores.



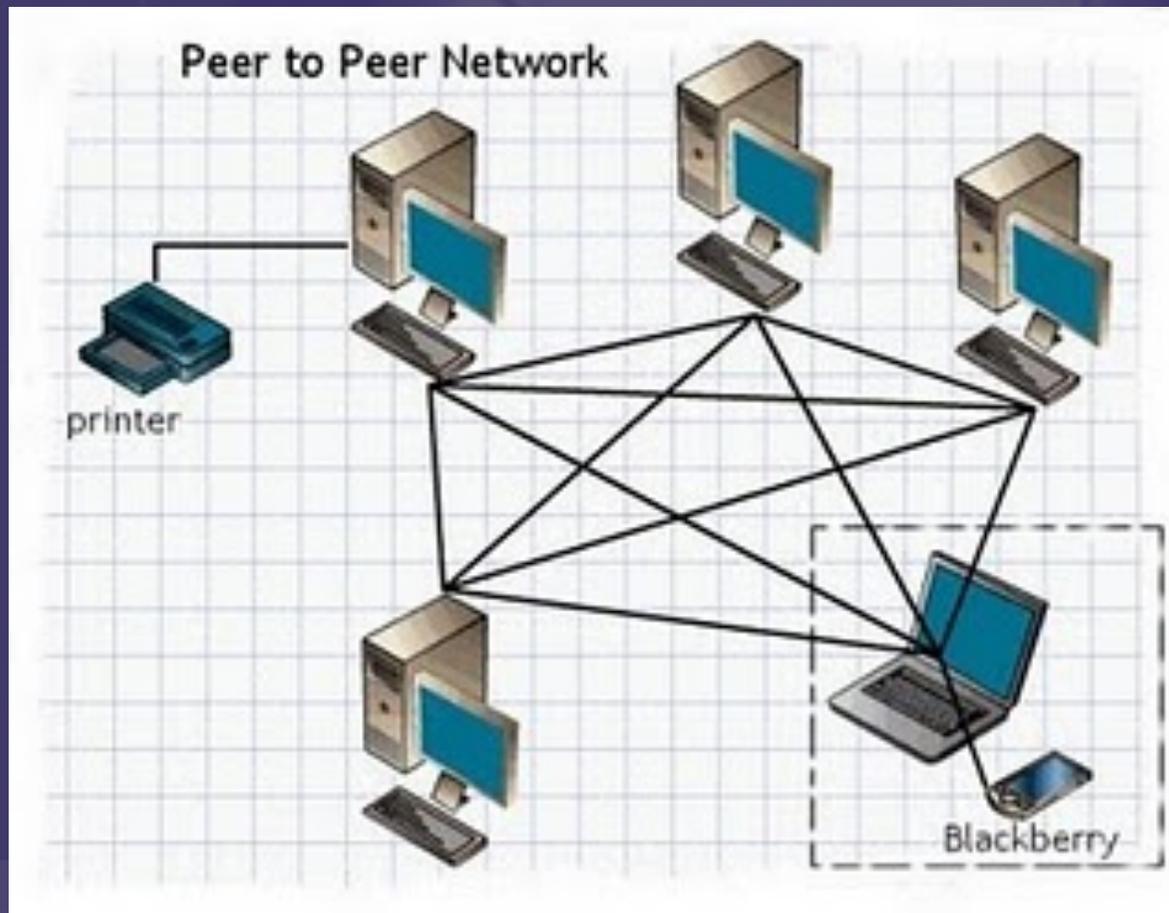
Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos: redes adhoc móviles



Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos: redes entre pares



Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Problemas a resolver:
 - Elección presente y futura de líder (*lider election and eventual leader election*).
 - Consenso distribuido. Zookeeper (Yahoo), Megastore (Google)
 - Memoria Compartida distribuida.
 - Radiado atómico y fiable. (Bases de datos replicadas)
 - Escalabilidad en radiado.

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Incertidumbres:
 - Fallos de nodos y de red.
 - Asincronía de nodos y de red.
 - Imposibilidad de conocer la membresía.
 - No existencia de identificadores: anonimato.
 - Múltiples identificadores: homonimia.

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Incertidumbres (cont.): Fallos de nodos y de red.
 - Si los fallos son maliciosos/arbitrarios resulta complejo resolver problemas de coordinación.
 - $N > 3f + 1$ (N número de procesos, f de fallos) para resolver consenso [Lamport, Pease y Shostack 1980].
 - Los procesos maliciosos modelan a intrusos en problemas de seguridad.

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Incertidumbres (cont.): Asincronía de red y/o nodos.
 - Si hay un solo fallo de nodo no se puede resolver consenso [Fischer, Lynch, Paterson, 1985]:

No se distingue un nodo fallido de un nodo lento o del que no llegan mensajes porque son lentos.
 - No se puede resolver elección de líder.
 - Se pueden resolver diversos problemas si se utilizan detectores de fallos (Chandra, Toueg, 1990)
 - Los detectores de fallos se pueden implementar en sistemas parcialmente síncronos (Dolev, Dwork, Stockmeyer, 1988)

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Incertidumbres (cont.): No se conoce la membresía.
 - En los sistemas distribuidos dinámicos los procesos entran y salen del sistema a lo largo del tiempo.
 - Es complejo coordinar acciones si no se sabe con quién hay que hacerlo.
 - Comunicación mediante radiado hardware.

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Incertidumbres (cont.): Los nodos carecen de identificador (anónimos).
 - Puede darse este caso en las redes de sensores (sin id. MAC).
 - Puede darse el caso en redes entre iguales por privacidad.
 - El anonimato crea problema de simetría en problemas de coordinación distribuida.
 - Imposibilidad de elección de líder.
 - Problemas de solución de consenso con detectores de fallos realistas (implementables).

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Incertidumbres (cont.): Algunos nodos pueden tener el mismo identificador (homónimos).
 - Es un modelo intermedio entre el modelo normal y el modelo con anónimos que incluye a ambos.
 - Se presenta cuando los nodos eligen su identificador al azar.
 - Se presenta cuando nodos replicados tienen el mismo id.
 - Se puede resolver el problema del consenso con detectores de fallos realistas (implementables).

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Soluciones: Algoritmos indulgentes con detectores de fallos no fiables.
- Un detector de fallos no fiable:
 - Cada proceso tiene un módulo del detector de fallos que le indica de forma no fiable los procesos sospechosos de fallo.
 - Los módulos del detector de fallos intercambian mensajes de *heartbeat* y van adaptando sus temporizadores con el tiempo.
 - Tarde o temprano los detectores de fallos dejan de sospechar de los procesos correctos.

Líneas de investigación del LSD-EUI (cont)

- Sistemas distribuidos dinámicos (cont.)
- Soluciones: Algoritmos indulgentes con detectores de fallos no fiables (cont.).
- Un algoritmo indulgente:
 - Un proceso intercambia mensajes de manera asíncrona con otros procesos del sistema.
 - Consulta el detector de fallos para decidir por cuántos mensajes espera.
 - La corrección del algoritmo no depende de la salida del detector de fallos. (Es indulgente)
 - La viveza (*liveness*) del algoritmo sí depende del detector de fallos.

Líneas de investigación del LSD-EUI

- Sistemas distribuidos dinámicos (cont.)
- Definición y análisis de los problemas.
- Implementaciones de los problemas:
 - Algoritmos de consenso indulgentes.
 - Algoritmos de detectores de fallos.
 - Algoritmos sin membresía.
 - Algoritmos con anonimia.

Líneas de investigación del LSD-EUI

- Memoria distribuida compartida.
 - Modelo de memoria: los procesos sólo se comunican entre sí mediante operaciones de lectura ($r_i(x)v$) y escritura ($w_i(y)s$).
 - Semánticas:
 - Atómico
 - Secuencial
 - Causal
 - pRAM
 - Implementaciones de los modelos con paso de mensajes, y réplica parcial/total de los datos en los procesos.

Líneas de investigación del LSD-EUI

- Radiado con calidad de servicio escalable.
 - Calidad de servicio (QoS) significa garantías de entrega de mensaje.
 - Significa también garantías de orden de entrega.
 - El radiado con orden causal implica información de control de $O(n^2)$ en el mensaje.
 - No permite escalar.
 - Se usan técnicas de dividir el sistema en subsistemas, y luego interconectarlos con pasarelas (*gateways*).
 - Hay que evitar el *efecto convoy* en los nodos pasarela.

Elección futura de Líder.

Clase de detector de fallos Ω (*eventual leader election*) [Chandra, Hadzilacos, Toueg 1996]:

- Cada proceso p tiene una variable $leader_p$ con un id. de proceso.
- Terminación y acuerdo: A partir de un determinado momento, todo proceso correcto p tiene permanentemente $leader_p = \ell$, siendo ℓ un proceso correcto.
- Es el detector de fallos más débil que permite resolver consenso en sistemas **asíncronos** ($\Omega \equiv \langle \diamond W \rangle$)

Elección futura de Líder (cont)

- Elección de líder y Ω no se pueden resolver en un sistema asíncrono con $f > 0$: Por el resultado de imposibilidad FLP.

Hay que imponer alguna restricción adicional al sistema: **sistema parcialmente síncrono.**

Elección futura de Líder. Algoritmo con canales síncronos futuros

Sistema distribuido **parcialmente síncrono** con paso de mensajes:

- Conjunto finito de n procesos $\Pi = \{p_1, \dots, p_n\}$.
- Fallos de parada (crash)
- Los procesos envían y reciben mensajes usando canales bidireccionales que unen cada par de procesos: *send(m)* y *receive(m)*
- La mínima y máxima velocidad de avance de los procesos están acotadas, pero las cotas no son conocidas.

Elección futura de Líder. Algoritmo con canales síncronos futuros (cont.)

- Canal síncrono futuro (*eventually timely*): tiene dos parámetros T y Δ desconocidos tal que:
 - Hasta el instante T el canal no da garantías. Todos los mensajes se pueden perder o llegar tarde.
 - A partir de T , un mensaje enviado en un instante t se entrega como muy tarde en instante $t + \Delta$.
- Todos los canales son síncronos futuros.
- $send(m)$: si m es enviado por un proceso correcto a un proceso correcto a través de un canal síncrono futuro en un instante $t > T$, m es recibido como tarde en instante $t + \Delta$ (antes de t no hay garantías).
- Fiable es un caso particular de síncrono futuro (cuando $T=0$ y Δ es conocido).

Elección futura de Líder. Algoritmo con canales síncronos futuros (cont.)

[Chandra & Toueg 1996]

Initially:

$leader_p \leftarrow 0$

$tout_p \leftarrow \eta$

set timer to $tout_p$

start T1 y T2

Task T1:

repeat each η time

send (ALIVE) a todos

Task T2:

when recibido mensaje ALIVE
from q and $q = leader_p$:

reset timer to $tout_p$

when recibido mensaje ALIVE
from q and $q < leader_p$:

$leader_p \leftarrow q$

$tout_p \leftarrow tout_p + 1$

reset timer to $tout_p$

when timer expira:

$leader_p \leftarrow leader_p + 1$

set timer to $tout_p$

Elección futura de Líder. Algoritmo con canales síncronos futuros (cont.)

Consideremos instante $t > T$ en que todos los procesos que van a fallar lo han hecho:

- Los latidos del proceso correcto con identificador más pequeño, ℓ , llegan como mucho cada $\eta + \Delta$.
- Si la variable $tout_p$ no alcanza ese valor, saltará el temporizador e incrementará $tout_p$. Entonces, tras recibir un ALIVE de ℓ , $leader_p = \ell$.
- Si la variable alcanza ese valor, no se incrementa más, y tras recibir un ALIVE de ℓ , $leader_p = \ell$.

Máster y doctorado

- **Máster de investigación:**
 - Ciencias y Tecnologías de la Computación.
- **Asignatura:**
 - Simulación de Redes de Comunicaciones:
 - Especialidad en Ciencias de la Computación.
 - 6 créditos ECTS.
 - 3 horas/semana.
- **Tesis:**
 - Duración estimada: 3/4 años (4 horas/día).