

SISTEMI OPERATIVI

Sistemi Distribuiti
Funzioni dei Sistemi Distribuiti

Lezione 1 – Obiettivi e funzioni

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Architetture di elaborazione distribuite
- Sistemi operativi per architetture di elaborazione distribuite
- Obiettivi
- Funzionalità

Architetture di elaborazione distribuite (1)

Caratteristiche

- Processore
 - Almeno uno per sistema di elaborazione
- Memoria
 - Locale ai sistemi di elaborazione
- Periferiche
 - Locali ai sistemi di elaborazione, ma usabili in rete
 - Locali ai sistemi di elaborazione, e non condivise (orologio)
 - Globali in rete
 - Rete di comunicazioni

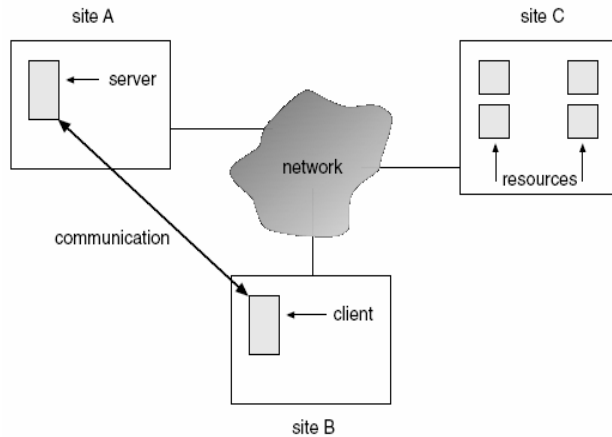
Architetture di elaborazione distribuite (2)

Caratteristiche

- Eterogeneità dei sistemi di elaborazione
- 

Architetture di elaborazione distribuite (3)

- Sito (site)
- Macchina (machine), host, computer, nodo (node)
 - Server
 - Client



Architetture di elaborazione distribuite (4)

Vantaggi

- Integrazione di sottosistemi
- Condivisione delle risorse (resource sharing)
- Parallelismo della computazione: aumento della velocità di calcolo (computation speedup)
- Interazione con utente

Architetture di elaborazione distribuite (5)

Vantaggi

- Riduzione della complessità e del costo (downsizing)
- Aumento dell'affidabilità (reliability), della tolleranza ai guasti (fault tolerance) e della disponibilità (dependability)
- Scalabilità (scalability)

Sistemi operativi per architetture di elaborazione distribuite

- Sistemi operativi di rete
network operating systems
 - Visibilità della struttura e delle caratteristiche dei componenti e della rete
- Sistemi operativi distribuiti
distributed operating systems
 - Completa trasparenza della struttura e delle caratteristiche dei componenti e della rete

Sistema operativo di rete (1)

- Accesso alle risorse condivise in rete
- Collegamento per elaborazione remota
 - Login remoto
 - telnet
 - Comunicazione tra processi remoti
 - socket
 - Attivazione di procedure remote
 - remote procedure call RPC
- Stampanti remote
 - Spooling di stampa remoto

Sistema operativo di rete (2)

- Trasferimento remoto di file
 - ftp
- File server remoto
 - Montaggio di file system remoti
- Posta elettronica

Sistema operativo distribuito (1)

- Risorse remote accessibili come risorse locali
- Migrazione dei dati e dei processi gestita dal sistema operativo in modo trasparente

Sistema operativo distribuito (2)

Migrazione dei dati

- Copia-lavoro-salvataggio
Il lavoro viene svolto su una copia locale del file; il file modificato viene ritrasmesso al server (FTP automatizzato)
- Copia di parti di file
Simile alla paginazione
- Problema: compatibilità di rappresentazione tra le varie macchine

Sistema operativo distribuito (3)

Migrazione computazionale

- Migrazione di procedura
- Migrazione di processo

Sistema operativo distribuito (4)

Migrazione di procedura

- Chiamata di procedura remota
Remote Procedure Call (RPC)
- Processo remoto con scambio messaggi
- Motivazioni
 - Preferibilità dell'hardware o del software remoti
 - Disponibilità di risorse specifiche
 - Accesso ai dati

Sistema operativo distribuito (5)

Migrazione di processo

- Processo
- Agenti mobili
- Motivazioni:
 - bilanciamento del carico
 - velocità di elaborazione
 - preferibilità dell'hardware o del software remoti
 - disponibilità di risorse specifiche
 - accesso ai dati

Sistema operativo distribuito (6)

- File server distribuito
Montaggio di file system remoti
in modo omogeneo e trasparente
- Stampanti distribuite
- Posta elettronica

Robustezza (1)

Problema

Guasti e malfunzionamenti del sistema di elaborazione distribuito

Tecniche di gestione

- Rilevazione dei guasti
- Mascheramento degli errori
- Riconfigurazione del sistema
- Ripristino del sistema

Robustezza (2)

Rilevamento dei guasti

- Monitoraggio periodico con handshaking
- Tempo massimo di attesa (time-out)
- Computazione duplicata con confronto dei risultati

Robustezza (3)

Mascheramento degli errori

- Computazione duplicata con votazione a maggioranza dei risultati
- Risorse duplicate

Robustezza (4)

Riconfigurazione

Interruzione di un collegamento o guasto ad una macchina

- Aggiornamento tabelle di instradamento per escludere il collegamento o la macchina guasti
- Sostituzione della macchina nei suoi compiti

Robustezza (5)

Ripristino

- Riconfigurazione per includere nuovamente il componente riparato
 - Informare le macchine (handshaking)
 - Aggiornamento tabelle
 - Aggiornamento siti non funzionanti
- Gestione posta non consegnata

Aspetti progettuali

- Trasparenza di allocazione delle risorse
 - Processori
 - Dispositivi di memorizzazione
 - Periferiche
 - File
- Mobilità dell'utente, della computazione e dei dati
- Tolleranza ai guasti
- Scalabilità

In sintesi

- Architetture di elaborazione distribuite
- Sistemi operativi per architetture di elaborazione distribuite
- Funzionalità

SISTEMI OPERATIVI

Sistemi Distribuiti
Comunicazione in rete

Lezione 1 – Gestione della comunicazione in rete

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

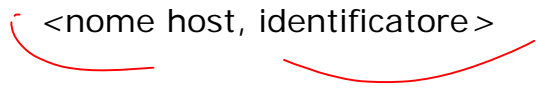
Sommario

- Nomi e risoluzione dei nomi
- Strategie di instradamento
- Strategie di pacchetto
- Strategie di connessione
- Gestione dei conflitti

Nomi e risoluzione dei nomi (1)

Identificatore di risorsa o processo

<nome host, identificatore>



- Nome host
 - Identificatore numerico 128.148.31.100
 - Identificatore logico dti.unimi.it

Nomi e risoluzione dei nomi (2)

Risoluzione dei nomi

- Processo simile alla compilazione
- Ogni host può avere un file contenente i nomi e gli indirizzi di tutti gli altri host
- Informazioni distribuite fra i computer nella rete e usare un protocollo per effettuare la risoluzione
 - server dei nomi del dominio
Domain Name Server - DNS

Nomi e risoluzione dei nomi (3)

Ottimizzazione delle prestazioni

- Caching delle tabelle dei nomi degli host
 - Aggiornamento e rinfresco
- Caching gerarchico

Strategie di instradamento (1)

Identificazione del percorso per un messaggio da A a B

- Unico percorso fisico
- Tabella di instradamento
 - Velocità percorso
 - Costo percorso
 - Aggiornamento tabelle

Strategie di instradamento (2)

Schemi di instradamento

- Instradamento statico (o fisso)
 - Percorso tra A e B definito a priori e non modificabile
 - Non richiede ricerca del percorso
 - Non adattabile a variazioni di carico
 - Non adattabile per gestire guasti
- Instradamento virtuale
 - Percorso tra A e B definito per una sessione
 - Percorso non modificabile nella sessione
 - Percorso modificabile al cambiare della sessione
 - Richiede ricerca del percorso solo all'inizio della sessione
 - Non adattabile a variazioni di carico o guasti durante la sessione

Strategie di instradamento (2)

Schemi di instradamento

- Instradamento dinamico
 - Percorso tra A e B definito per ogni messaggio
 - Percorso modificabile al cambiare del messaggio
 - Richiede ricerca del percorso all'inizio di ogni messaggio
 - Messaggi possono arrivare in ordine diverso da quello in cui sono inviati
 - Adattabile a variazioni di carico
 - Adattabile per gestire guasti

Strategie di instradamento (3)

Gateway

- Collega la rete locale ad altre reti gestendo l'eventuale cambiamento di protocollo di comunicazione
- Host → Gateway: instradamento statico
- Gateway → Host: instradamento dinamico
- Realizzazione
 - Software
 - Dispositivi hardware dedicati

Strategie di instradamento (4)

Router

- Dispositivo responsabile dell'instradamento dei messaggi tra due reti
- Realizzazione
 - Software
 - Dispositivi hardware dedicati

Strategie di pacchetto

Invio di messaggi a lunghezza variabile

- Divisione del messaggio in porzioni di dimensione fissa
 - Pacchetti **Packet**
 - Strutture **Frame**
 - Datagrammi **Datagram**
- Invio dei pacchetti e ricostruzione del messaggio
- Gestione degli errori e della perdita di pacchetti

Strategie di connessione ⁽¹⁾

Sessioni di comunicazione

Schemi per le sessioni di comunicazione

- a commutazione di circuito
circuit switching
- a commutazione dei messaggi
message switching
- a commutazione di pacchetto
packet switching

Strategie di connessione (2)

Comunicazioni a commutazione di circuito

- Due processi che devono comunicare stabiliscono un collegamento fisico fisso per tutta la sessione
- Nessun altro processo può usare tale collegamento
- Notevole tempo di attivazione
- Possibile spreco di larghezza di banda
- Sovraccarico di gestione inferiore

Strategie di connessione (3)

Comunicazioni a commutazione dei messaggi

- Due processi che devono comunicare stabiliscono un collegamento temporaneo per la durata dell'invio del messaggio
- Il collegamento è creato dinamicamente
- Sovraccarico di gestione medio

Comunicazioni a commutazione di pacchetto

- Ciascun pacchetto di un messaggio è inviato separatamente su una connessione attivata dinamicamente
- Migliore uso della larghezza di banda
- Suddivisione e riassettaggio dei pacchetti
- Sovraccarico di gestione superiore

Gestione dei conflitti (1)

Trasmissioni in

- Bus multi-accesso
- Rete ad anello

Conflitto per uso contemporaneo
del canale di comunicazione

Gestione dei conflitti (2)

Rete a bus multi-accesso

- Rilevamento delle collisioni
 - Analisi della portante ad accesso multiplo
 - Dati trasmessi diversi da quelli osservati
→ collisione
- Gestione delle collisioni
 - Ritrasmissione a tempi differenti
- Prevenzione delle collisioni
 - Limitare il numero dei nodi sulla rete

Gestione dei conflitti (3)

Rete ad anello

- Rilevamento della perdita del token
 - Tempo massimo di attesa
- Gestione della perdita del token
 - Elezione

In sintesi

- Nomi e risoluzione dei nomi
- Strategie di instradamento
- Strategie di pacchetto
- Strategie di connessione
- Gestione dei conflitti

SISTEMI OPERATIVI

Sistemi Distribuiti
Comunicazione in rete

Lezione 2 – Protocolli di comunicazione

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Obiettivi
- Modello teorico
 - Protocollo ISO/OSI
- Modelli reali
 - Protocollo TCP/IP
 - Protocollo UDP/IP
 - Protocolli applicativi

Problemi

- Comunicazione asincrone
- Probabilità di errori
- Interazione tra ambienti eterogenei

Obiettivi

- Semplificare la progettazione
- Creare un ambiente omogeneo di comunicazione tra componenti eterogenei
- Astrazione della visione delle comunicazioni in rete
- Virtualizzazione delle comunicazioni
- Gestione efficiente
- Gestione degli errori e dei guasti

Soluzione

- Protocolli di comunicazione
come driver della rete
- Sottosistema dedicato alle comunicazioni in rete
suddiviso in strati
- Comunicazione gestita tra strati equivalenti
- Protocolli specifici

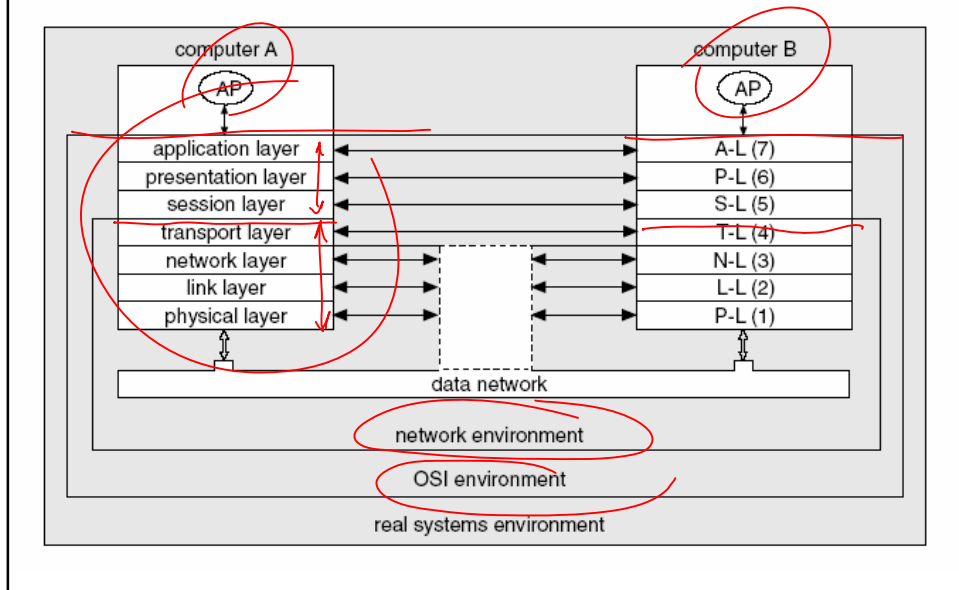
Modello teorico

Protocollo ISO / OSI

ISO: International Standards Organization

OSI: Open Systems Interconnection

Protocollo ISO/OSI (1)



Protocollo ISO/OSI (2)

- Strato fisico **physical layer**
- Strato di collegamento tra i dati **data-link layer**
- Strato di rete **network layer**
- Strato di trasporto **transport layer**
- Strato di sessione **session layer**
- Strato di presentazione **presentation layer**
- Strato di applicazione **application layer**

Protocollo ISO/OSI (3)

Strato fisico

- Dettagli della trasmissione dei bit
 - Meccanici
 - Elettrici

Protocollo ISO/OSI (4)

Strato del collegamento tra i dati

- Gestione dell'invio e della ricezione del singolo pacchetto
- Rilevazione e correzione degli errori

Protocollo ISO/OSI (5)

Strato di rete

- Connessione
- Instradamento dei pacchetti
- Indirizzo pacchetti in uscita
- Decodifica indirizzo pacchetti in entrata

Protocollo ISO/OSI (6)

Strato di trasporto

- Partizionamento dei messaggi in pacchetti
- Mantenimento dell'ordine dei pacchetti
- Controllo del flusso
- Gestione degli errori a livello di messaggio

Protocollo ISO/OSI (7)

Strato di sessione

- Realizzazione delle sessioni
- Protocolli di comunicazione tra processi

Protocollo ISO/OSI (8)

Strato di presentazione

- Risolvere le differenze di formato tra le varie macchine
 - conversione dei formati di rappresentazione dei dati
 - modalità semi-duplex
 - modalità full-duplex

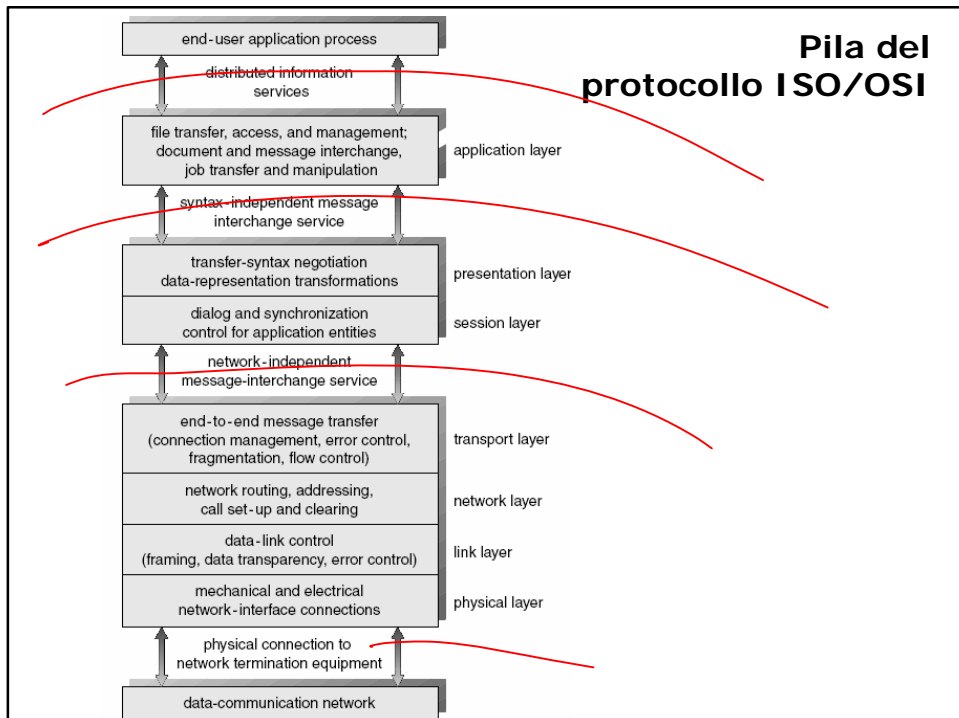
Protocollo ISO/OSI (9)

Strato di applicazione

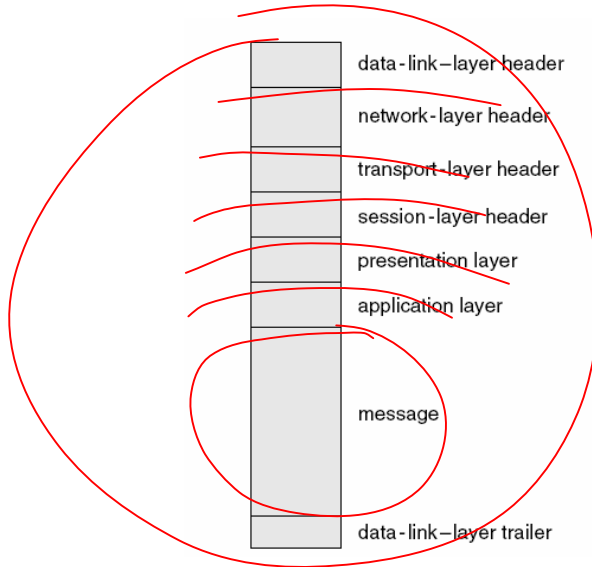
- Interazione a livello di applicazione

Esempi:

- trasferimento di file
- protocolli di connessione remoti
- posta elettronica
- basi di dati distribuite



Messaggio nel protocollo ISO/OSI



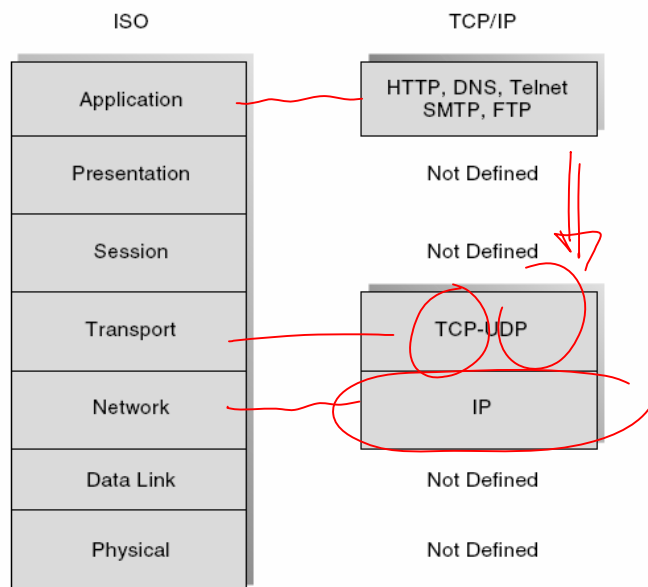
Modelli reali (1)

- Obiettivo:
 - più efficiente
 - più semplice
- Problema:
 - più difficile da realizzare
 - meno astratto

Modelli reali (2)

- Internet Protocol
IP
- Transmission Control Protocol
TCP
- User Datagram Protocol
UDP
- Protocolli applicativi

Relazione tra TCP/IP, UDP/IP e ISO/OSI



Internet Protocol

Strato di rete

- Gestione pacchetti
- Instradamento pacchetti

Protocolli di trasporto

Strato di trasporto

- User Datagram Protocol - UDP
 - Inaffidabile
 - Senza gestione delle connessioni
- Transmission Control Protocol - TCP
 - Affidabile
 - Orientato alle connessioni

Strati TCP/IP

Strato di applicazione

- Telnet, SSH *comunicazione remota*
- FTP, SFTP *file transfer*
- HTTP, HTTPS *web*
- SMTP *posta elettronica*
- DNS *risoluzione nomi*

In sintesi

- Gestione della comunicazione in rete
driver di rete
=
protocolli di comunicazione in rete
- Modello teorico
 - Protocollo ISO/OSI
- Modelli reali
 - Protocollo TCP/IP
 - Protocollo UDP/IP
 - Protocolli applicativi

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 1 – Distribuzione della computazione

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Motivazioni
- Obiettivi
- Tecniche di distribuzione della computazione

Motivazioni

- Aumento della velocità di elaborazione
Parallelelismo
- Accesso efficiente a risorse informative o fisiche
- Elaborazione su grandi quantità di dati localizzate
- Tolleranza ai guasti

Obiettivi

Spostare la computazione
sulla macchina
in cui si trovano le risorse adatte
a raggiungere
l'obiettivo applicativo

Tecniche per la distribuzione della computazione

- Chiamate di procedure remote
- Allocazione di processi
- Agenti

Tecniche per il supporto alla computazione distribuita

- Comunicazione tra processi distribuiti
- Sincronizzazione tra processi distribuiti

In sintesi

- Parallelismo
 - Diminuzione latenza
 - Aumento throughput
 - Diminuzione tempo risposta
- Uso di risorse fisiche o informative locali
- Tolleranza ai guasti
 - Rilevamento e correzione errori
 - Sopravvivenza
- Tecniche per la distribuzione della computazione e il suo supporto

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 2 – Chiamata di procedura remota

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Obiettivo
- Realizzazione
- Esecuzione
- Valutazione
- Invocazione di metodo remoto

Obiettivo

- Eseguire una procedura sulla macchina su cui sono disponibili le risorse informative o fisiche necessarie
- Lasciare il resto del processo sulla macchina su cui è stato attivato

Chiamata di procedura remota

Remote Procedure Call

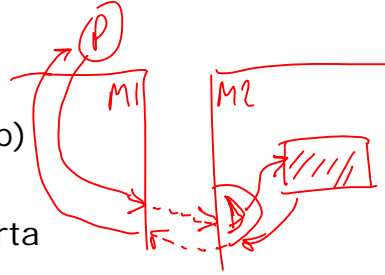
RPC

Chiamata di procedura fra macchine in rete

- Simile alla chiamata di procedura all'interno del processo oppure al sistema operativo
- Processo chiamante è entità attiva, procedura chiamata è entità passiva
- RPC è realizzata mediante comunicazione tra processi basata su messaggi strutturati

Realizzazione (1)

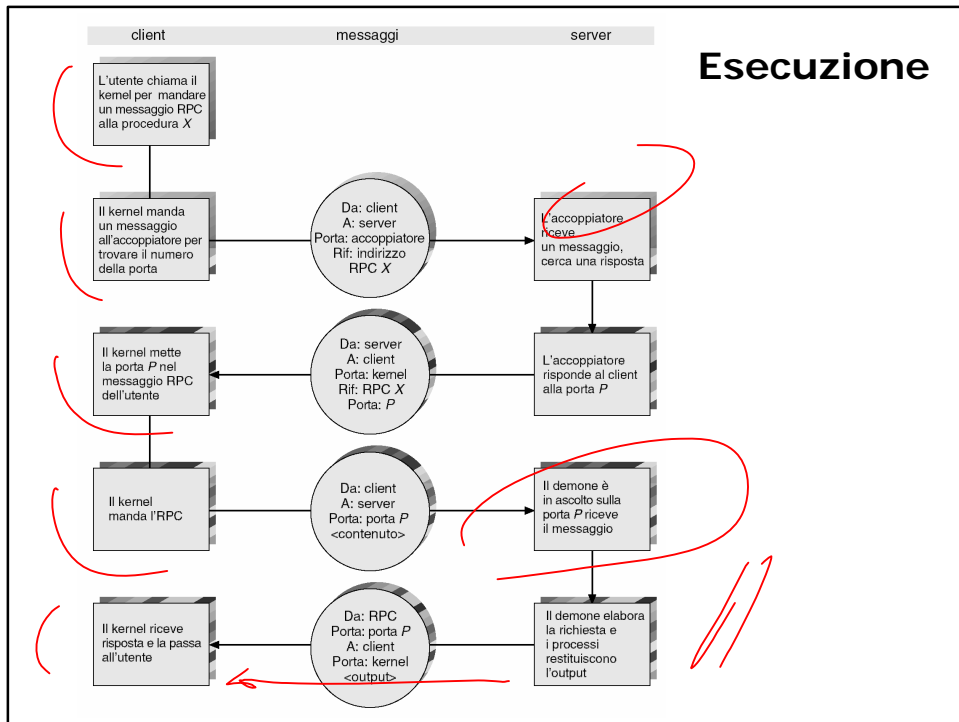
- Processo chiamante
- Procedura chiamata (stub)
- Demone in ascolto su porta
- Attivazione mediante messaggio strutturato
 - Identificatore della funzione richiesta
 - Parametri
- Risultato restituito al mittente in messaggio separato



Realizzazione (2)

Stub (terminale remoto della procedura)

- Uno stub per ogni procedura
- Parametri
- Traduzione parametri
- Scambio messaggi



Vantaggi

- Procedura remota come se fosse locale
- Dettagli nascosti

Problemi (1)

- Diverse rappresentazioni dei dati
- Rappresentazione esterna dei dati
external data representation, XDR

Problemi (2)

- Semantica della chiamata
 - Al più una volta
 - Marche di tempo
 - Storico delle richieste
 - Esattamente una volta
 - Unicità delle richieste
 - Esecuzione garantita

Esempio di uso delle RPC

Realizzazione di un file system distribuito

- Demoni RPC
- Client
- Messaggi sulla porta DFS del server dove risiede il file
- Messaggio di ritorno col risultato

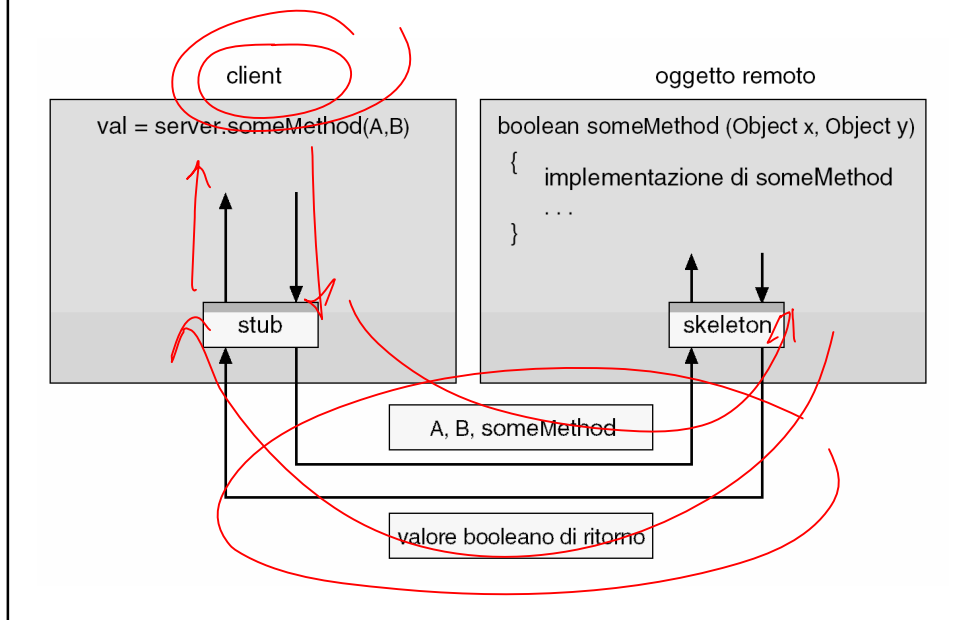
Invocazione di metodo remoto ⁽¹⁾

Remote Method Invocation

RMI

- Specifica di linguaggi ad oggetti (ad esempio Java)
- Simile a RPC
- Thread invoca un metodo su un oggetto remoto
 - differente JVM sullo stesso computer
 - computer remoto connesso tramite rete

Invocazione di metodo remoto (2)



RPC rispetto RMI

- **RPC**

- Gestione della connessione è responsabilità del programmatore

- **RMI**

- Gestione trasparente al programmatore
- Gestita dall'ambiente di programmazione

In sintesi

- Chiamate di procedure remote
 - Realizzazione
 - Esecuzione
 - Vantaggi e problemi
- Invocazione di Metodi Remoti

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 3 – Allocazione dei processi

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Obiettivi
- Allocazione statica
- Allocazione dinamica

Obiettivi

- **Attivare** i processi sulla macchina più adatta ad effettuare la computazione
tenendo conto del carico computazionale e delle risorse informative o fisiche necessarie
- **Migliorare** lo sfruttamento dei processori:
bilanciamento del carico
- **Migliorare** lo sfruttamento delle risorse
- **Gestire** in modo efficiente l'uso delle risorse minimizzando il carico computazionale di gestione
- **Realizzare** la tolleranza ai guasti

Allocazione statica (1)

- L'allocazione dei processi è definita alla loro attivazione ottimizzando uno o più obiettivi
- Effettuata l'allocazione dei processi, tale allocazione è permanente e i processi vengono attivati

Allocazione statica (2)

- **Allocazione completa**

Tutti i processi che devono essere attivati
sono allocati contemporaneamente
e poi sono attivati

- **Allocazione incrementale**

Quando un gruppo di processi deve essere
attivato,
viene prima allocato tenendo fissa
l'allocazione dei processi già attivati
e poi viene attivato

Allocazione statica (3)

- Funzione obiettivo per l'ottimizzazione

- Sfruttamento dei processori

- Tempo di idle totale dei processori
- Distribuzione del tempo di idle dei processori
- Latenza media dei processi
- Throughput dei processi
- Tempo di risposta in sistemi in tempo reale
- ...

- Efficienza della gestione

- Minimizzazione del tempo di accesso alle informazioni
- Minimizzazione del tempo di accesso alle periferiche
- Minimizzazione del tempo di accesso ai servizi del sistema operativo
- ...

- ...

Allocazione statica (4)

- Vincoli per l'ottimizzazione
 - Locazione di processi
 - Locazione delle risorse informative o fisiche necessarie con accesso solo locale
 - Interazione con utente
 - Basi dati
 - Sensori e attuatori
 - Sicurezza e autenticazione
 - ...
 - Incompatibilità con locazione (hw, sw, sistema operativo)
 - Incompatibilità tra processi
 - ...

Allocazione statica (5)

- Caratteristiche degli algoritmi di allocazione
 - Modalità di ricerca della soluzione
 - Algoritmi deterministici
 - Algoritmi euristici
 - Modalità di esecuzione
 - Algoritmi centralizzati
 - Algoritmi distribuiti
 - Qualità della soluzione
 - Soluzione ottima
 - Soluzione sub-ottima

Allocazione statica (6)

- Caratteristiche degli algoritmi di allocazione
 - Allocazione
 - Globale
 - Locale
 - Attivazione dell'algoritmo di allocazione
 - Eseguita dal processore mittente
 - Eseguita dal processore ricevente

Allocazione dinamica (1)

- L'allocazione dei processi è definita durante tutta la loro vita ottimizzando uno o più obiettivi
- L'allocazione effettuata all'attivazione non è permanente

Allocazione dinamica (2)

- **Allocazione totale**
L'allocazione è definita considerando tutti i processi contemporaneamente
- **Allocazione parziale**
L'allocazione è definita considerando un sottoinsieme di processi che soddisfa una regola di candidatura alla riallocazione

Allocazione dinamica (3)

- **Allocazione periodica**
L'allocazione è effettuata a intervalli regolari
- **Allocazione reattiva**
L'allocazione è effettuata quando si verifica una condizione di riallocazione
- **Riallocazione volontaria**
La riallocazione di uno o più processi è richiesta da un processo

Allocazione dinamica (4)

- Funzione obiettivo per l'ottimizzazione
- Vincoli per l'ottimizzazione
- Caratteristiche degli algoritmi di allocazione

Allocazione dinamica (5)

- Migrazione dei processi
 - Stato di evoluzione della computazione del processo
 - Trasferimento del processo
 - Riattivazione del processo
 - Compatibilità e traduzione della rappresentazione dei dati e del codice

Allocazione dinamica (6)

- Costo della migrazione dei processi
 - Tempo di gestione dell'algoritmo di allocazione
 - Tempo di gestione della migrazione

In sintesi

- Allocazione statica
- Allocazione dinamica

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 4 – Agenti mobili

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Obiettivi
- Modello computazionale ad agenti
- Agenti mobili

Obiettivi

- Innalzare il livello di astrazione della computazione secondo i principi dell'ingegneria del software
- Descrivere la computazione mediante tecnologie ad oggetti con capacità di esecuzione di azioni

Modello della computazione ad oggetti

- Oggetti
collezione di dati e
procedure (metodi) per la loro gestione
- Incapsulamento delle caratteristiche dei dati e delle procedure
- Programmi → Processi/Thread

Agenti

Entità software

- Autonoma
- Inserita in un ambiente
- Pro-attiva
- Cooperante

Agenti mobili

- Agenti
con capacità di
muoversi nel sistema distribuito
interagire con le singole macchine
scoprire le risorse e i servizi
delle singole macchine

In sintesi

- Agenti
- Agenti mobili

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 5 – Coordinamento distribuito tra processi (parte 1)

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Ordinamento degli eventi
- Sincronizzazione
 - Mutua esclusione
 - Atomicità
 - Gestione concorrenza

Sincronizzazione dei processi

Ordinamento degli eventi in sistema distribuito

- Soluzione ideale:
ordinamento totale
- Problema:
in sistemi distribuiti non ci sono
orologio e memoria comune
per realizzare ordinamento totale
- Soluzione pratica:
ordinamento parziale

Relazione “accaduto prima” ⁽¹⁾

Definizione

- Se A e B sono eventi dello stesso processo
ed A avviene prima di B,
allora $A \rightarrow B$
- Se A è l'evento di trasmissione del messaggio
in un processo,
e B è l'evento di ricezione di quel messaggio
da parte di un altro processo,
allora $A \rightarrow B$
- Se $A \rightarrow B$ e $B \rightarrow C$,
allora $A \rightarrow C$

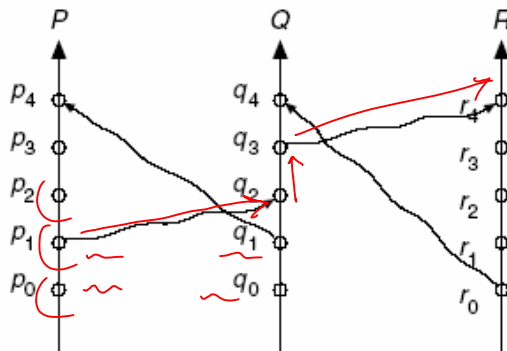
Relazione "accaduto prima" (2)

Caratteristiche

- Non riflessiva
- Eventi non in relazione sono concorrenti e non si influenzano
- Se due eventi sono in relazione possono influenzarsi

Relazione "accaduto prima" (3)

Tre processi concorrenti



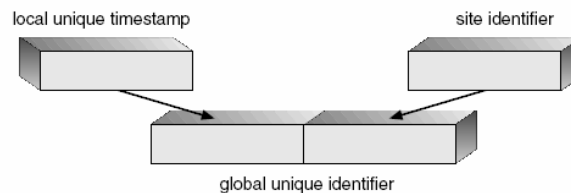
Ordinamento globale

- Marca di tempo
- Orologio logico
- Incremento monotono
- Avanzamento forzato per processi diversi che comunicano

Marca di tempo

Generazione delle marche di tempo uniche a livello di sistema distribuito

- Soluzione centralizzata: unico distributore
- Soluzione distribuita:
marca di tempo locale + identificatore del sito



Marca di tempo locale

- Orologio di sistema
- Orologio logico locale

Orologi locali

- Velocità diverse
- Orologi logici
- Sincronizzazione degli orologi logici
 - Aggiornamento dell'orologio a $n+1$
qualora un sito riceva una marca di valore n

Mutua esclusione (1)

- Metodo centralizzato
- Metodo distribuito

Mutua esclusione (2)

Metodo centralizzato

- Processo coordinatore centralizzato per la gestione dell'accesso alle sezioni critiche
- Coda dei processi in attesa
- Prestazioni: limitate per centralizzazione
- Tolleranza ai guasti: vulnerabile

Mutua esclusione (3)

Metodo distribuito

- Quando P vuole entrare in sezione critica:
 - genera marca tempo
 - invia richiesta di entrata in sezione critica a tutti i processi
- Quando un processo Q riceve richiesta di entrare in sezione critica:
 - ritarda la risposta se è in sezione critica
 - risponde immediatamente se non intende entrare in sezione critica
 - se desidera entrare nella propria sezione critica ma non vi è entrato, compara la propria marca di tempo con quella di P: se la propria è più grande, allora risponde immediatamente, altrimenti ritarda la risposta per entrare prima

Mutua esclusione (4)

Metodo distribuito

- No starvation
- No deadlock
- Tolleranza ai guasti

Mutua esclusione (5)

Metodo a passaggio di token

- Processi organizzati ad anello logico
- Token indica autorizzazione ad accedere

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 5 – Coordinamento distribuito tra processi (parte 2)

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Atomicità (1)

- Coordinatore delle transazioni in ciascuna macchina
 - Iniziare la transazione
 - Dividere la transazione in sotto-transazioni
 - Distribuzione delle sotto-transazioni sulle macchine adatte
 - Coordinamento della fine della transazione

Atomicità (2)

Protocollo di commit a due fasi

- Transazione inizia su macchina S_i
- Quando tutti i siti comunicano a S_i che è stata completata la porzione di T ad essi assegnata si attiva il protocollo a due fasi per definire il commit della transazione
- Fase 1
 - C_i aggiunge $\langle \text{prepare } T \rangle$ e manda messaggio "prepare T " a tutte le macchine coinvolte
 - Quando tale messaggio è ricevuto da una macchina S_j , C_j decide se commit sulla macchina è possibile
 - Se non possibile, aggiunge $\langle \text{no } T \rangle$ e risponde a C_i con abort
 - Se possibile, aggiunge $\langle \text{ready } T \rangle$ e risponde a C_i con "ready T "

Atomicità (3)

Protocollo di commit a due fasi

- Fase 2
 - Quando C_i riceve le risposte oppure dopo che è trascorso time-out, C_i stabilisce se può chiudere la transazione
 - T è chiusa con commit se sono ricevuti $\langle \text{ready } T \rangle$ da tutte le macchine, aggiungendo $\langle \text{commit } T \rangle$ al log
 - Altrimenti T è abortita aggiungendo $\langle \text{abort } T \rangle$
 - Il risultato è trasmesso a tutte le macchine
 - Quando un coordinatore locale riceve "commit T " o "abort T ", registra $\langle \text{commit } T \rangle$ o $\langle \text{abort } T \rangle$ rispettivamente

Atomicità (4)

Protocollo di commit a due fasi

- Tolleranza ai guasti delle macchine
- Tolleranza ai guasti della rete

Concorrenza

Ambiente distribuito

- Gestore delle transazioni
 - Transazioni locali
 - Transazioni globali
- File di log
- Protocolli
 - bloccanti
 - non bloccanti

Protocolli bloccanti (1)

- Realizzazione del gestore del blocco in ambiente distribuito
- Replicazione dei dati
- Modalità di gestione del blocco condiviso o esclusivo

Protocolli bloccanti (2)

Coordinatore centralizzato dei lock

- Dati non replicati
- Unico responsabile centralizzato dei lock
- Realizzazione semplice: richiesta di lock mediante scambio di messaggi (2 messaggi)
- Gestione stallo: complicata
- Prestazioni: limiti dovuti a coordinatore unico
- Tolleranza ai guasti: critica dovuta alla centralizzazione

Protocolli bloccanti (3)

Coordinatori multipli dei lock

- Dati non replicati
- Responsabile locale dei lock
- Realizzazione semplice:
richiesta di lock mediante scambio di messaggi
(2 messaggi)
- Gestione stallo: complicata

Protocolli bloccanti (4)

Coordinatore dei lock a maggioranza

- Dati replicati
- Responsabile dei lock per ogni sito
- Richiesta di almeno $n/2+1$ lock locali;
lock globale ottenuto quanto almeno $n/2+1$ lock
locali sono ottenuti
- Realizzazione complicata
 - $2(n/2 + 1)$ messaggi per gestire il blocco
 - $(n/2 + 1)$ messaggi per gestire lo sblocco
- Bisogna modificare gli algoritmi per la gestione
dello stallo

Protocolli bloccanti (5)

Protocollo polarizzato

- Dati replicati
- Responsabile dei lock per ogni sito
- Blocchi condivisi: richiesta locale
- Blocchi esclusivi: richiesta globale come in lock a maggioranza
- Minore sovraccarico nella lettura rispetto ai protocolli di lock a maggioranza
- Ulteriore sovraccarico in scrittura
- Difficile gestire lo stallo

Coordinatore (1)

Compito

- Gestione del coordinamento tra processi in ambiente distribuito
 - Mutua esclusione
 - Rilevazione degli stalli
 - Sostituzione token persi
 - Controllo input/output

Coordinatore (2)

Algoritmi di elezione

- Processo con identificatore di priorità più alto
- Algoritmo del bullo
- Algoritmo dell'anello

Coordinatore (3)

Algoritmo del bullo

P cerca di eleggere se stesso come coordinatore.

Quando P si accorge che il coordinatore non funziona

- manda un messaggio di inizio elezione ai processi con priorità più alta e attende risposta
- se nessuna risposta entro timeout T, P elegge se stesso come coordinatore e informa tutti i processi che ha assunto tale ruolo
- se P riceve una risposta, allora attende l'identificatore del nuovo coordinatore
- se P non riceve l'identificatore del nuovo coordinatore, deve far ripartire l'elezione

Coordinatore (4)

Algoritmo dell'anello

- Collegamento logico ad anello unidirezionale
- Lista attiva

Quando P si accorge che il coordinatore non funziona

- genera una nuova lista attiva vuota
- attiva elezione inviando la lista attiva al processo successivo nell'anello

Se P riceve un messaggio di elezione

- se è primo messaggio ricevuto o inviato, crea una lista attiva includendo se stesso e il processo che lo precede nell'anello e invia un messaggio di elezione indicando se stesso e il processo predecessore
- se il messaggio ricevuto non contiene P, aggiunge il predecessore alla propria lista attiva ed inoltra il messaggio di elezione al processo successivo
- se il messaggio ricevuto contiene P, la lista attiva contiene tutti i processi: P può identificare il coordinatore nella lista come il processo con priorità più elevata

In sintesi

- Ordinamento degli eventi
- Tecniche per la sincronizzazione
 - Mutua esclusione
 - Atomicità
 - Gestione concorrenza

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 6 – Deadlock in ambiente distribuito

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Estensione ai sistemi di elaborazione distribuiti
- Prevenzione in ambiente distribuito
- Rilevamento e gestione in ambiente distribuito

Prevenzione dello stallo

- Estensione degli algoritmi per la macchina singola al caso distribuito
- Ordinamento globale delle risorse nel sistema distribuito
 - Sovraccarico minimo
- Algoritmo del banchiere
 - Sovraccarico elevato
 - Prestazioni basse per centralizzazione

Marche di tempo con rilascio della risorsa ⁽¹⁾

- Identificatore di priorità per ogni processo
- Rilascio anticipato
Se P possiede risorsa, e Q ha priorità più alta di P
 - P deve rilasciare la risorsa
 - Rollback di P
 - Risorsa assegnata a Q
- Possibile starvation
Soluzione: marche di tempo

Marche di tempo con rilascio della risorsa (2)

Schema wait-die

- Senza rilascio anticipato
- Se la risorsa è occupata dal processo P
 - Se processo richiedente Q ha marca più piccola, attende
 - Se processo richiedente Q ha marca più grande, fa rollback e muore
- Si evita starvation se non si assegna una nuova marca al rollback

Marche di tempo con rilascio della risorsa (3)

Schema wound-wait

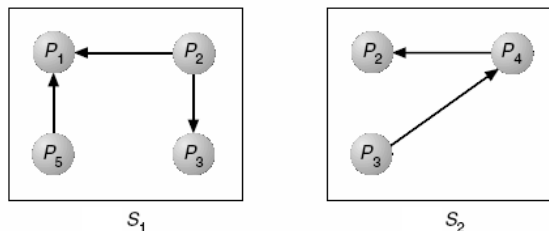
- Con rilascio anticipato delle risorse
- Se la risorsa è occupata dal processo P
 - Se processo richiedente Q ha marca più grande, attende
 - Se processo richiedente Q ha marca più piccola, fa rollback
- Si evita la starvation se non si assegna una nuova marca al rollback

Rilevamento dello stallo

- Grafo di allocazione delle risorse
→ grafo di attesa
- Ciclo = Stallo
- Come realizzare e mantenere
il grafo di attesa in ambiente distribuito?

Grafo di attesa in ambiente distribuito

- Ogni macchina ha un grafo di attesa locale
- Se non ci sono cicli nel grafo di attesa locale,
non è garantito che non ci siano stalli



- Unione dei grafi non ha cicli → non ci sono stalli

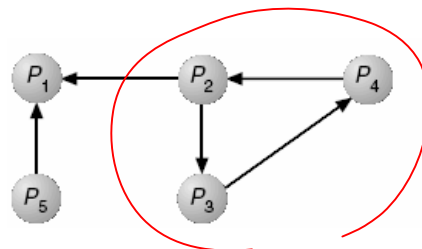
Grafo di attesa centralizzato (1)

- Coordinatore centralizzato del rilevamento degli stalli
- Grafo di attesa globale = unione dei grafi di attesa locali
 - Grafo reale
 - Grafo costruito dall'algoritmo
 - se esiste uno stallo, questo deve essere segnalato
 - se viene rilevato uno stallo, il sistema deve essere effettivamente in stallo

Grafo di attesa centralizzato (2)

- Aggiornamento del grafo di attesa
 - Inserimento o rimozione di un arco
 - La macchina manda un messaggio per avvisare della modifica del grafo locale
 - Dopo un certo numero di cambiamenti
 - La macchina manda un messaggio dopo un certo numero di cambiamenti locali

- Cercare un ciclo



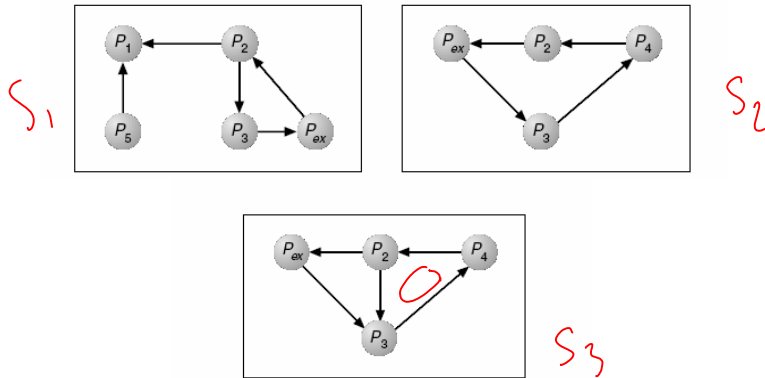
Algoritmo di rilevamento centralizzato

- Trasmissione dei grafi di attesa locali
- Costruzione del grafo di attesa globale
 - Un nodo per ogni processo
 - Riporta gli archi dei grafi di attesa locali
- C'è un ciclo nel grafo di attesa globale?
 - Sì: il sistema è in stallo
 - No: la verifica non è significativa

Algoritmo di rilevamento distribuito (1)

- Ogni macchina costruisce una parte del grafo di attesa globale
- In ogni grafo di attesa locale si inserisce un nodo P_{ex} per identificare attesa coinvolgente risorse di altre macchine
- Se esiste uno stallo, apparirà almeno un ciclo in uno dei grafi di attesa locale
 - Se un grafo di attesa locale contiene un ciclo che non coinvolge il nodo P_{ex} , il sistema è in stallo
 - Se esiste un ciclo che coinvolge P_{ex} , allora si ha la possibilità di stallo
 - contattare macchine per verificare esistenza stallo

Algoritmo di rilevamento distribuito (2)



Algoritmo di rilevamento distribuito (3)

Problema:

rilevamento contemporaneo di cicli in grafi di attesa locali provoca sovraccarico di gestione e messaggi ridondanti

- Assegnare un unico identificatore ad ogni processo P_i
- Quando la macchina S_i scopre un ciclo che coinvolge il nodo P_{ex} nel proprio grafo locale, manda un messaggio di rilevamento dello stallo ad altre macchine
solo se il processo precedente P_{ex} nel ciclo ha identificatore minore del processo successivo a P_{ex}
- Altrimenti, la macchina S_i continua la propria attività, lasciando il compito di iniziare la procedura di rilevazione dello stallo ad altre macchine.

Gestione dello stallo (1)

- Tra processi in stallo, il coordinatore sceglie una vittima a cui applicare il rollback
- Tutte le macchine vengono informate
 - Chi stava interagendo con la vittima fa rollback

Gestione dello stallo (2)

Rollback inutili

- Falsi cicli nel grafo globale dovuti a tempi di trasmissione dei messaggi di acquisizione e rilascio delle risorse
- Le richieste, provenienti da macchine differenti, devono essere individuate da identificatori o marche di tempo unici
- Nel grafo globale vanno solo le richieste non immediatamente soddisfacenti

In sintesi

- Tecniche di prevenzione dello stallo
- Tecniche di rilevamento e gestione dello stallo

SISTEMI OPERATIVI

Sistemi Distribuiti
Computazione Distribuita

Lezione 7 – Comunicazione tra processi in rete

Vincenzo Piuri

Università degli Studi di Milano

Sommario

- Messaggi
- Mailbox
- File
- Socket

Scambio di messaggi

- Analogo allo scambio di messaggi tra processi su una singola macchina
- Buffer in ambiente distribuito

Mailbox

- Analogo allo scambio di messaggi attraverso mailbox in una singola macchina
- Mailbox in ambiente distribuito

File

- Analogo all'uso di file in una singola macchina
- File in ambiente distribuito

Socket

- Analogo all'uso di socket in una singola macchina
- Porte e messaggi in ambiente distribuito

Realizzazione

- Il sistema operativo remotizza le operazioni di comunicazione

In sintesi

- Abbiamo visto come estendere le tecniche per la comunicazione tra processi in una macchina ad architetture distribuite
 - Messaggi
 - Mailbox
 - File
 - Socket

SISTEMI OPERATIVI

Sistemi Distribuiti
File System Distribuiti

Lezione 1 – Struttura e funzioni

Vincenzo Piuri

Università degli Studi di Milano - SSRI - CDL ONLINE

Sommario

- Obiettivi
- File system di rete
- File system distribuito

Obiettivi

- Accedere ai file system
(risorse informative e fisiche)
delle macchine connesse in una architettura
di elaborazione distribuita
- Trasparenza dell'accesso al file system
- Gestione efficiente dell'accesso ai file system

File system in sistemi di elaborazione distribuiti

File system di rete

Network File System - **NFS**

File system distribuito

Distributed File System - **DFS**

Condivisione di file delocalizzati

File system di rete

- Collezione di file system delle macchine della rete
- Montaggio locale di file system remoti
- Problema:
visibilità della struttura della rete
e dell'allocazione delle risorse

File system distribuito

- Integrazione dei file system delle macchine in rete in un unico file system globale
- Omogeneità di visione del file system su tutte le macchine della rete
- Trasparenza dell'allocazione delle risorse

Nomi dei file (1)

- Identificatore unico del file nel sistema di elaborazione distribuito
- Identificatore del file in file system di rete
 - nome macchina + nome file nel file system della macchina
 - montaggio file system remoto su direttorio locale con nome file nel file system della macchina (incluso il file system remoto montato)
- File system distribuito
 - nome unico nel file system globale

Nomi dei file (2)

- Nome del file mappato automaticamente
 - ad un indirizzo locale
 - ad un indirizzo remoto
- Locazione del file
 - visibile agli utenti
 - invisibile agli utenti
 - trasparenza della posizione
 - trasparenza (indipendenza) della locazione
- Repliche dei file
- Migrazione dei file

Accesso ai file (1)

- File system di rete
 - Servizi remoti mediante RPC
 - Copiatura locale, aggiornamento locale, salvataggio remoto
- File system distribuito
 - Servizi locali, eventualmente eseguiti remotamente in modo trasparente

Accesso ai file (2)

- Cache
 - Locazione
 - Politiche di aggiornamento
 - Write-through
 - Delayed-write
 - Write-on-close
 - Consistenza
 - Verifica iniziata dal client
 - Verifica iniziata dal server
 - Scelta della dimensione

Stato del file server

- Definizione
Insieme delle informazioni che caratterizzano l'uso di un file aperto
- File server senza stato
 - Richieste soddisfatte in modo indipendente
 - Più semplice
- File server con stato
 - Operazione di open()
 - Identificatore di connessione
 - Più efficiente

Replica dei file

- Ridondanza
- Migliori prestazioni
- Replicazione invisibile agli utenti
- Aggiornamento

In sintesi

- File system di rete
- File system distribuito