

2014

SISTEMI DISTRIBUITI



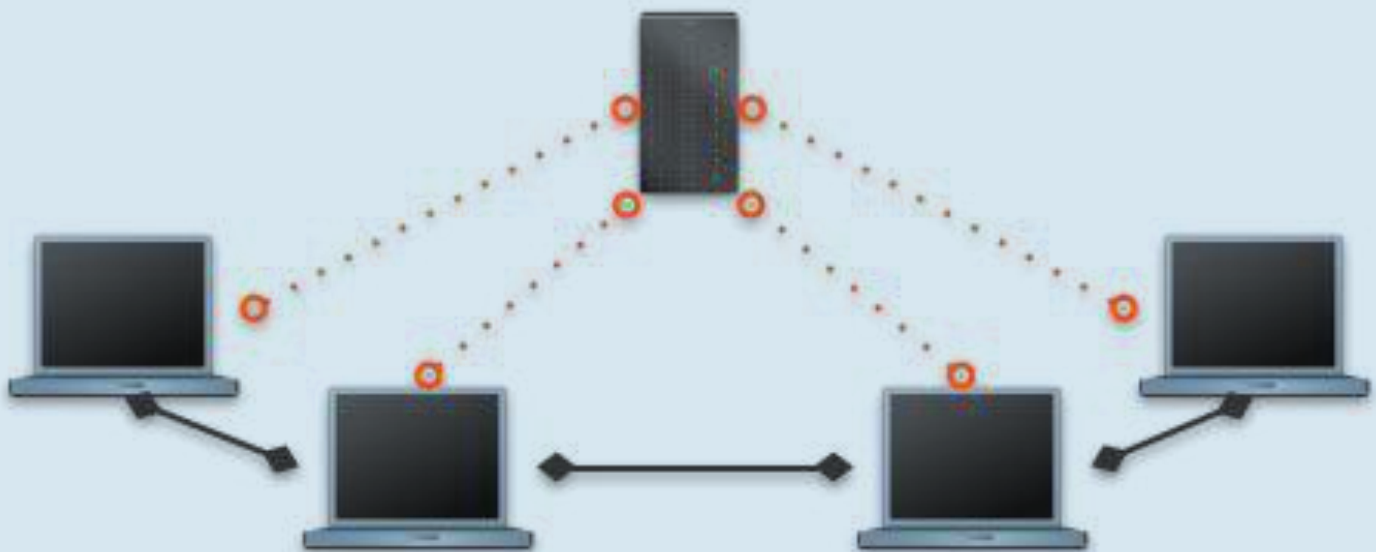
*"Tu sai di averne
uno quando il
guasto di un
computer di cui
non hai mai
sentito parlare
non ti permette di
fare il tuo lavoro."
-Lamport*

Quercioli, Pecoraro, Rando, Lucero

V AI

Sommario

Definizione	2
Caratteristiche.....	3
Pro e contro	4
Organizzazione di un sistema distribuito	8
Sitografia	9

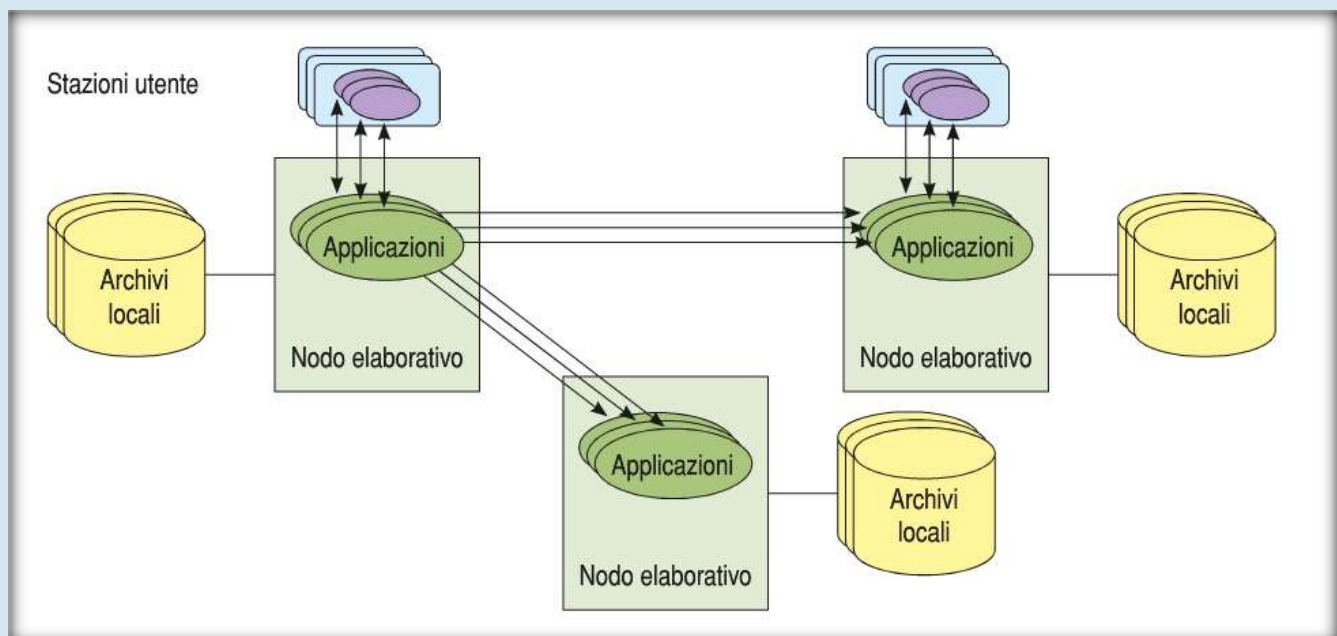


Definizione

Cos'è un sistema distribuito?

A questa domanda si possono dare diverse risposte poiché esistono diverse definizioni:

1. Un sistema distribuito è costituito da un insieme di **entità autonome** (componenti software e hardware) spazialmente separate che comunicano e coordinano tra loro le loro azioni attraverso scambio di **messaggi**.
2. Secondo **Tanenbaum** un sistema distribuito consiste in un insieme di **calcolatori** che all'utente vengono mostrati come un singolo calcolatore.
3. Indica genericamente una tipologia di sistema informatico costituito da un insieme di **processi interconnessi** tra loro in cui le comunicazioni avvengono solo esclusivamente tramite lo scambio di opportuni messaggi. Un sistema distribuito è una **collezione** di computer indipendenti che appare all'utente come un solo **sistema coerente**.



Caratteristiche

Un sistema distribuito ha determinate caratteristiche:

- È un sistema in cui l'SE (sistema di elaborazione delle informazioni) non è centralizzata su una singola macchina ma distribuito su diverse.
- È un sistema di elaborazione in cui il numero di componenti coopera comunicando in rete.
- È un sistema in cui i componenti hardware o software comunicano in rete solo tramite messaggi.
- Non ha un clock globale poiché impossibile sincronizzare i clock di tutti i processi, questo comporta l'impossibilità di ordinare, in modo preciso ed univoco, tutti gli eventi che occorrono all'interno del sistema.
- L'essere sincrono o asincrono. Questa distinzione è essenziale poiché alcune problematiche nell'ambito dei sistemi distribuiti possono essere risolte o meno proprio in base a queste caratteristiche.
- Un sistema distribuito si dice sincrono quando è possibile calcolare le seguenti proprietà, altrimenti si dice asincrono:
 - L'intervallo di tempo massimo e minimo per l'esecuzione di un'istruzione da parte di un processo.
 - L'intervallo di tempo massimo per la trasmissione di un messaggio dalla sorgente alla destinazione.
 - E la massima deviazione del valore di ciascun orologio locale (clock drift rate) rispetto al tempo reale.

Vantaggi

- Consentire facilmente la connessione tra utenti e risorse
- Essere trasparente, cioè nascondere che le risorse sono distribuite
- Essere aperto
- Essere flessibile
- Essere scalabile

Inoltre:

- Le macchine sono autonome (**hardware**)
- L'utente pensa di lavorare su una sola macchina (**software**)

Pro e contro

In un sistema distribuito si possono individuare dei vantaggi e degli svantaggi

I **pro** sono:

- L'**affidabilità** è il principale vantaggio dei sistemi distribuiti: grazie alla sua ridondanza intrinseca un sistema distribuito è in grado di “sopravvivere” a un guasto di un suo componente.
- **Eterogeneità**, infatti i vari processi possono essere fisicamente diversi. Infatti questi possono utilizzare diversi sistemi operativi software scritti con differenti linguaggi di programmazione o utilizzare molteplici dispositivi hardware.
- La **scalabilità**, cioè la capacità di erogare le medesime prestazioni, in termini di throughput e latenza, rispetto agli utilizzatori nonostante l'aumento del carico operativo sul sistema.
- La **trasparenza**, come trasparenza si intende il concetto di vedere il sistema distribuito non come un insieme di componenti ma come un unico sistema di elaborazione: l'utente non deve accorgersi che di interagire con un sistema distribuito ma deve avere la percezione di utilizzare un singolo elaboratore.
- **Economicità**, i sistemi distribuiti offrono spesso un miglior rapporto prezzo/qualità dei sistemi centralizzati basati su mainframe: una rete di PC connessi ha un prezzo di alcuni ordini di grandezza inferiore rispetto a quello di un mainframe e con le tecnologie odierne la capacità computazionale è paragonabile.



Di contro abbiamo le seguenti caratteristiche:

- Produzione di software i programmatori del secolo scorso hanno dovuto modificare il proprio stile di programmazione e aggiornarsi con lo studio dei nuovi linguaggi e dei nuovi strumenti di sviluppo per poter realizzare applicazioni distribuite.
- Proprio per la struttura hardware i sistemi distribuiti sono più complessi di quelli centralizzati: richiedono strumenti per l'interconnessione degli host e tecniche per l'instradamento corretto dei messaggi e dei dati.
- Sicurezza nei vecchi sistemi per lo più bastava proteggere il sistema dall'accesso fisico delle persone ai locali dove erano presenti i dispositivi da proteggere (hard disk e supporti di memorizzazione). Oggi l'accesso avviene via etere e via cavo e anche le trasmissioni sono soggette a rischio di intercettazione (sniffing) e quindi richiedono l'applicazione di appositi accorgimenti per tutelare tutti gli utenti e garantire sicurezza e riservatezza nei dati, sia memorizzati sui propri computer personali, sia trasmessi per transazioni commerciali o semplicemente personali (email).
- Comunicazione Il trasferimento a distanza delle informazioni richiede nuove tipologie di sistemi di telecomunicazione, sia cablati che wireless, e l'aumento esponenziale degli utenti fa sì che giornalmente aumenti la richiesta di bande trasmissive, anche per migliorare la qualità del servizio offerto e offrire nuove tipologie di applicazioni sempre più performanti (alta velocità, alta definizione, video streaming ecc.).
- La possibilità di fallimenti indipendenti ai processi. I fallimenti che possono affliggere i processi possono essere di varia tipologia, ma è possibile raggrupparli in due categorie: fallimenti di tipo crash e fallimenti bizantini. Nel primo caso abbiamo che il processo che va in crash smette improvvisamente di funzionare mentre nel secondo caso è impossibile, in genere, fare qualsiasi tipo di assunzione sulla causa o sugli effetti del fallimento. In quest'ultimo caso infatti il comportamento del processo che fallisce in modo bizantino è tipicamente arbitrario.



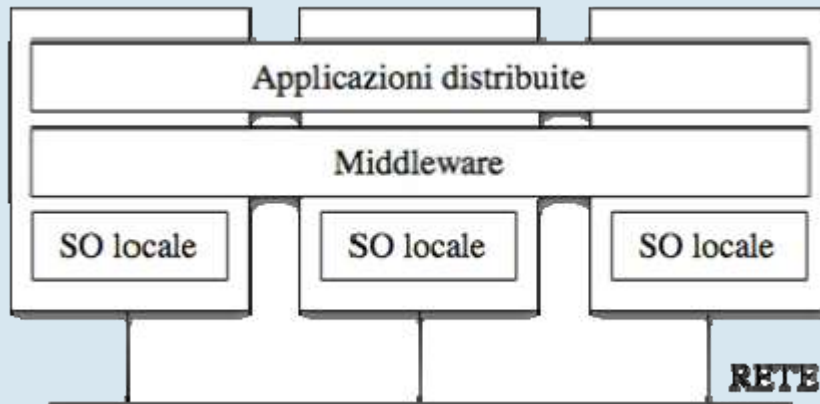
Vantaggi	Motivo
Affidabilità	Grazie alla ridondanza il sistema è in grado di sopravvivere ad un guasto.
Integrazione	La capacità di un sistema di integrare componenti eterogenei.
Trasparenza	<p>Identificare il sistema distribuito non come un insieme di macchine ma come un unico sistema che gestisce informazioni.</p> <p>Ci sono otto tipi di trasparenze:</p> <ul style="list-style-type: none"> • Accesso • Locazione • Concorrenza • Replicazione • Guasti • Migrazione • Prestazione • Scalabilità
Economicità	Ha un buon rapporto qualità/prezzo poiché permette di utilizzare vecchie tecnologie insieme a quelle recenti.
Apertura	<p>Il sistema utilizza dei protocolli standard favorendo l'accesso ad HW e SW di fornitori diversi.</p> <p>Vi sono all'interno del sistema le seguenti caratteristiche:</p> <ul style="list-style-type: none"> • Interoperabilità • Portabilità • Incrementabilità
Connettività e Collaborazione	Sistema che ha la possibilità di condividere delle risorse di tipo HW avvantaggiando la parte economica.
Tolleranza ai Guasti	Quando si presenta un guasto parziale del sistema, quest'ultimo ha la possibilità di copiare le risorse in modo tale che un componente danneggiato non ne impedisca il funzionamento.

Svantaggi	Motivo
Produzione di Software	Nel corso degli anni i programmatori hanno dovuto: <ul style="list-style-type: none">• Definire il protocollo TCP/IP.• Sviluppare architetture sia lato client che lato server.
Complessità	I sistemi distribuiti sono costruiti in maniera più complessa rispetto ad un sistema centralizzato poiché necessitano di operazioni quali: <ul style="list-style-type: none">• L'interconnessione degli host.• L'instradamento dei dati e messaggi.
Sicurezza	Un sistema distribuito ha maggiori problemi di sicurezza rispetto ad uno centralizzato perché hai dati ci possono accedere anche quelli non autorizzati.
Comunicazione	La comunicazione ha varie problematiche: <ul style="list-style-type: none">• Mancanza di prevedibilità perché le richieste generano casualità e poiché non hanno una risposta standardizzata.• Ha bisogno sempre di nuove vie di comunicazione (cablate o wireless).
Mancanza di Clock	La mancanza di un clock globale all'interno del sistema rende del tutto impossibile la sincronizzazione delle componenti.



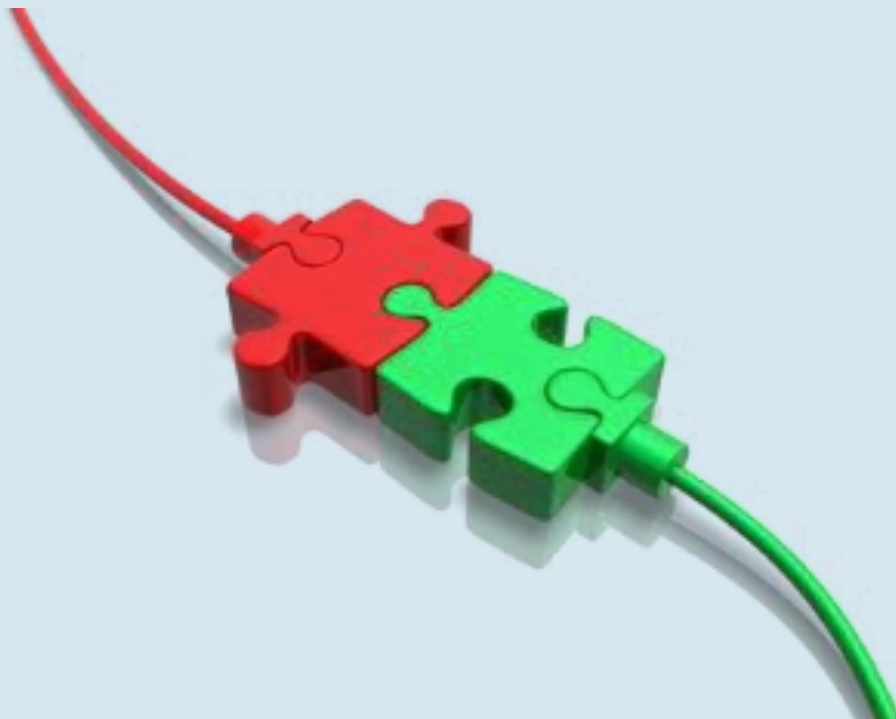
Organizzazione di un sistema distribuito

- **Obiettivo:** offrire una visione unica del sistema che in realtà è composto da computer e reti eterogenei



- **Soluzione:** organizzazione a strati (layer)
 - *Livello superiore:* utenti e applicazioni
 - *Livello intermedio:* strato software
 - *Livello basso:* sistema operativo

N.B Il livello intermedio, interfaccia tra piattaforma e applicazione, è detto **middleware**.



Sitografia

http://it.wikipedia.org/wiki/Sistema_distribuito

http://home.deib.polimi.it/loiacono/uploads/Teaching/InfoB0809/14_Sistemi_Distribuiti.pdf

http://www.hoepliscuola.it/media/file/sfogliolibro/Tecnologie_progettazione3/00396203617501.pdf/index.html

