

Architetture dei WIS

Prof.ssa E. Gentile
a.a. 2011-2012

Definizione di WIS

- Un WIS può essere definito come un insieme di applicazioni in grado di reperire, cooperare e fornire informazioni utilizzando il Web come canale di comunicazione

Benefici dei WIS

- **Interoperabilità in ambienti eterogenei**
 - Piattaforme differenti
- **Servizi di Business attraverso il Web**
 - Migliorare la gestione della supply chain
- **Libertà di scelta nelle configurazioni**
 - Disponibilità di configurazioni diverse
- **Supporto a vari tipi di client**
 - Indipendentemente dalla piattaforma del client

Supply Chain

- La **Supply Chain** è "Una rete globale usata per consegnare prodotti e servizi dalla materia prima fino al consumatore finale attraverso un flusso ingegneristicamente impostato di informazione, distribuzione fisica, e denaro"
- La **Supply Chain Management** è "La progettazione, pianificazione, esecuzione, controllo, e monitoraggio delle attività della catena di fornitura con l'obiettivo di creare valore netto, creando una infrastruttura competitiva, prendendo vantaggio dalla logistica a livello mondiale, sincronizzando l'offerta con la domanda, e misurando globalmente i risultati"

Struttura di un WIS

- I WIS si strutturano come le architetture distribuite in differenti livelli:
 - Livelli applicativi (**Layer**)
 - Livelli di architettura fisica (**Tier**)

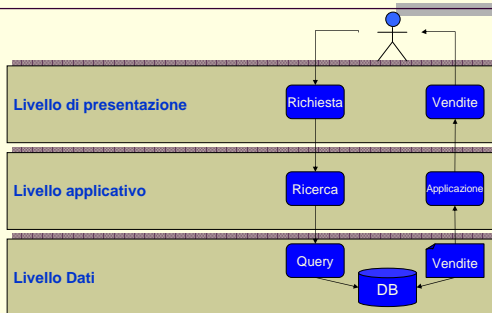
Layer Applicativi

- **Layer di presentazione** (front end)
 - Modalità di interazione con l'utente
 - Modalità di interfacciamento grafico
 - Modalità di rendering delle informazioni
- **Layer di logica applicativa** (o Business)
 - Funzioni da mettere a disposizione degli utenti
- **Layer di accesso ai dati**
 - Gestione dell'informazione
 - Accesso alle basi di dati, ai sistemi esterni e ai sistemi legacy

Livelli Fisici (Tier)

- **Single Tiered**
 - I layer applicativi sono su una singola macchina
- **Two Tiered**
 - layer di presentazione sulla macchina utente
 - layer di accesso ai dati sulla macchina server
 - layer di logica applicativa su una o sull'altra macchina o su entrambe
- **Tree Tiered**
 - I layer risiedono su macchine dedicate

Architettura logica a tre livelli



Classificazione dei WIS

- **Servizi di informazione**
 - Accesso ad informazioni strutturate e classificate
 - Pagine statiche o dinamiche
 - L'utente non inserisce contenuto informativo nel sistema
- **Servizi di comunicazione**
 - Supportano la comunicazione tra gruppi di utenti
 - L'utente è identificato
 - Le informazioni possono essere memorizzate nel sistema
- **Servizi transazionali**
 - Supporto all'acquisto di beni o servizi
 - L'utente identificato modifica i dati nel sistema
 - Problemi di sicurezza e riservatezza dei dati

Livelli di interazione

Livello 1: Sono presenti solo informazioni per l'avvio della procedura del servizio

- Trova informazioni solo sui contatti e sull'organizzazione

Livello 2: Scaricare e stampare moduli per l'avvio della procedura del servizio

- Bisogna poi far pervenire la documentazione con modalità tradizionali

Livello 3: Interazione avanzata

- Possibilità di avviare on line la procedura e la compilazione dei moduli

Livello 4: Esecuzione on line della procedura

- Assenza di moduli cartacei e gestione on line dell'intera procedura

Modalità di accesso ai dati

■ Sito Internet

- Accessibile da tutta la rete e da qualsiasi utente
- Problemi di accessibilità

■ Sito Intranet

- Accesso limitato all'interno dell'organizzazione
- Knowledge Management
- Uso di architetture hardware dedicate

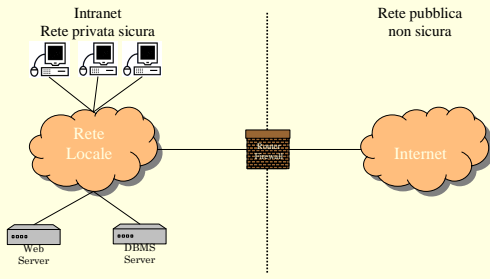
■ Sito Extranet

- Accesso a utenti ben definiti
- Fondamentale la standardizzazione del sistema

Confronto tra i Sistemi

	Internet	Intranet	Extranet
Utenti	Noti con approssimazione	Dipendenti dall'organizzazione	Dipendenti di più organizzazioni
Compiti	Servizi informativi o di vendita	Supporto al lavoro dei dipendenti Gestione dei processi	Integrazione dei processi delle organizzazioni
Larghezza di Banda	Bassa	Alta, affidabile	Alta, affidabile
Compatibilità	Diversi browser Massima accessibilità del sito	Scelta a priori del browser	Scelta a priori della tecnologia - Conoscenza della tecnologia adottata dai partner
Quantità Informazioni	Poche e in tempi rapidissimi	Quantità elevate, per la gestione dei processi operativi	Quantità elevate, integrazione della supply chain

Struttura Fisica



Prof.ssa E. Gentile

Sistemi Informativi su Web

13

Architetture dei WIS

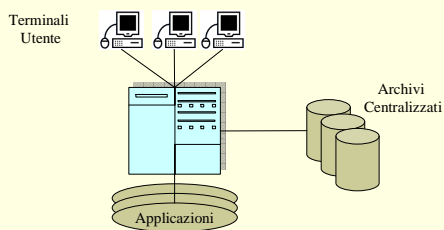
- **Sistema centralizzato**
 - Quando i dati e le applicazioni risiedono in un unico nodo elaborativo
- **Sistema distribuito**
 - Quando almeno una delle seguenti condizioni è verificata:
 - Le applicazioni, fra loro cooperanti, risiedono su più nodi elaborativi (elaborazione distribuita)
 - Il patrimonio informativo, unitario, è ospitato su più nodi elaborativi (basi di dati distribuite)

Prof.ssa E. Gentile

Sistemi Informativi su Web

14

Sistema centralizzato

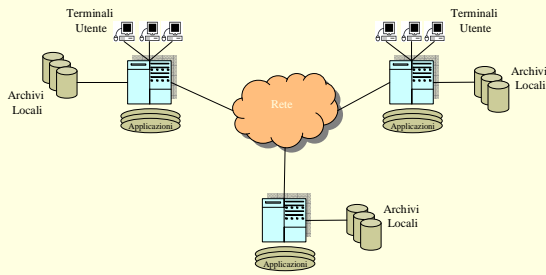


Prof.ssa E. Gentile

Sistemi Informativi su Web

15

Sistema distribuito

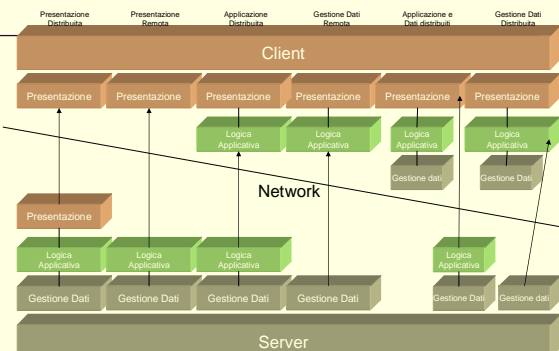


Prof.ssa E. Gentile

Sistemi Informativi su Web

16

Architetture 2-tier

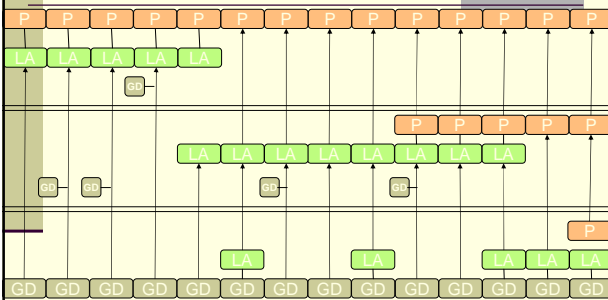


Prof.ssa E. Gentile

Sistemi Informativi su Web

17

Architetture 3-tier



Prof.ssa E. Gentile

Sistemi Informativi su Web

18

Server Farm

- Un **Server Farm** è pensato come un insieme di elaboratori che condividono il carico elaborativo, le applicazioni e, a seconda delle configurazioni, i dati
- Il vantaggio della **Server Farm** rispetto ad una singola macchina consiste in un minor costo di acquisto dell'hardware e nella maggiore affidabilità e scalabilità del sistema

Affidabilità del sistema

- Probabilità che il sistema sia operativo ad uno specifico istante temporale t

$$\frac{MTTF}{MTTF + MTTR}$$

- MTTF (Mean Time To Failure) indica il tempo medio di guasto
- MTTR (Mean Time To Repair) indica il tempo medio di riparazione del sistema

Scalabilità

- La capacità dell'infrastruttura di soddisfare richieste crescenti da parte degli utenti con aggiornamenti adeguati
- Un sistema si dice **non scalabile** se il costo aggiuntivo per far fronte a un dato aumento di traffico o dimensione è eccessivo o non attuabile

Server Farm

- **Cloning (clonazione)**
 - Su ogni nodo vengono installate le stesse applicazioni software ed i medesimi dati
- **Partitioning (partizionamento)**
 - Prevede la duplicazione dell'hardware e del software ma non dei dati, che invece vengono ripartiti tra i nodi

Cloning

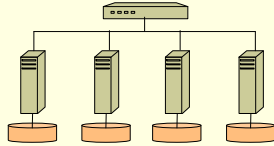
- La clonazione permette di raggiungere dei vantaggi sia di scalabilità che di tolleranza dei guasti e di disponibilità del servizio stesso
- Se un clone subisce un guasto un altro clone può continuare a erogare il servizio

Reliable Array of Cloned Services

- Un insieme di cloni dedicati ad un particolare servizio si dicono RACS
- Si possono presentare in due configurazioni:
 - Shared-nothing
 - Shared-disk

Shared-nothing

- I dati memorizzati sono replicati su ogni clone e risiedono in un disco fisso locale ad ogni clone
- Un aggiornamento di dati deve essere applicato ad ognuno dei cloni



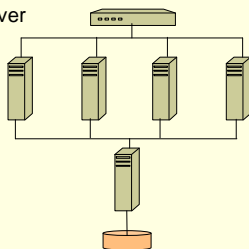
Prof.ssa E. Gentile

Sistemi Informativi su Web

25

Shared-disk o Cluster

- I cloni condividono un server di memorizzazione che gestisce i dischi fissi



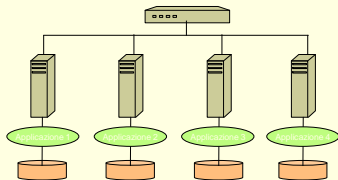
Prof.ssa E. Gentile

Sistemi Informativi su Web

26

Partizionamento

- Il partizionamento è completamente trasparente alle applicazioni e le richieste vengono inviate alla partizione che possiede i dati rilevanti.



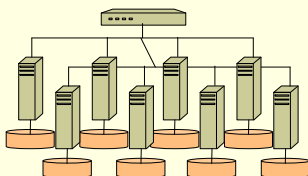
Prof.ssa E. Gentile

Sistemi Informativi su Web

27

Reliable Array of Partitioned Service (RAPS)

- Si risolve il problema dei guasti utilizzando la clonazione dei singoli server che costituiscono la partizione, creando dei *pack*



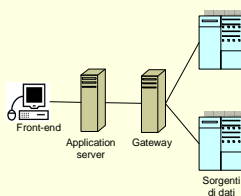
Prof.ssa E. Gentile

Sistemi Informativi su Web

28

Sistemi n tier

- Su ogni tier è possibile implementare una server farm
- I layer sono sempre 3
- Però l'uso di più tier riduce l'affidabilità del sistema



$$\text{Affidabilità} = \prod_i a_i$$

dove $a_i < 1$ per ogni i è l'affidabilità di ogni singolo tier

Prof.ssa E. Gentile

Sistemi Informativi su Web

29

Architetture dei sistemi Web

- **Web server**
 - Si occupa della gestione delle richieste HTTP provenienti da internet o dalla intranet aziendale
- **Script engine**
 - È un processo che esegue script per la generazione di pagine HTML dinamiche
- **Application server**
 - Assume il ruolo di middle tier ed implementa la logica di business dell'applicazione web
- **DBMS server**
 - Si occupa della gestione dei dati
- **Componenti di sicurezza**
 - Firewall e Intrusion Detection System (IDS)

Prof.ssa E. Gentile

Sistemi Informativi su Web

30

Elementi dell'architettura

- Il **Web Server** si colloca fisicamente tra l'utente finale ed il sistema informativo. Si occupa della presentazione di pagine HTML statiche.
- Lo **Script Engine** genera pagine dinamiche interagendo con l'Application server o con il DBMS server. Le pagine dinamiche vengono restituite al Web Server che le inoltrerà all'utente.
- L'**Application Server** implementa la logica di business e in alcuni casi svolge il ruolo di contenitore di oggetti
- Il **DBMS** si occupa dell'accesso e della gestione dei dati aziendali

Componenti di sicurezza: Firewall

- Il **Firewall** controlla il traffico in entrata e in uscita dalla rete da proteggere con lo scopo di prevedere, rilevare ed annullare eventuali attacchi e richieste non autorizzate.
- Esistono tre tipologie di Firewall:
 - Packet filtering
 - Application proxy
 - Statefull inspection packet filter

Firewall

- Un firewall packet filtering effettua la ritrasmissione di un pacchetto basandosi su regole che fanno riferimento all'header di pacchetto (indirizzo IP o numero dei porta)
- Un firewall application proxy è un programma applicativo eseguito da un host che connette due reti

Confronto tra Firewall

- Nei firewall packet filtering le comunicazioni vengono istaurate direttamente tra i processi client e server
- Nei firewall application proxy i client devono stabilire una connessione con il firewall e richiedere al proxy di stabilire una connessione con il servizio di destinazione

Firewall stateful inspection packet filter

- Effettua la funzione di filtraggio basandosi sul contenuto dei pacchetti ricevuti in precedenza e sulla base dello stato complessivo della connessione esistente tra client e server
- È un soluzione intermedia tra i firewall packet filtering e application proxy

Intrusion Detection System

- È il processo di monitoraggio degli eventi di un sistema, o di una rete di calcolatori, il cui obiettivo è individuare le intrusioni definite come il tentativo di compromettere la confidenzialità, l'integrità o la disponibilità del sistema
- Esistono due tipologie di IDS
 - Network based
 - Host based

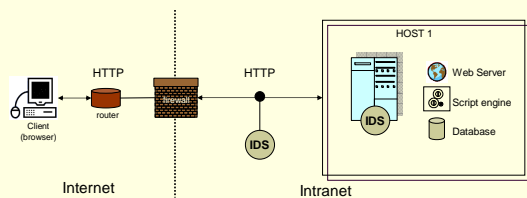
Progettazione architetturale

- Consiste nello stabilire:
 - Il numero di tier fisici del sistema
 - Quante macchine devono essere introdotte ad ogni tier fisico
 - Come collegare tra di loro le diverse macchine
 - Come introdurre i componenti di sicurezza

Soluzioni architetturali

- Configurazione 2 tier single host
- Configurazione 3 tier dual host
- Configurazione 3 tier con server farm
- Configurazione 5 tier con server farm

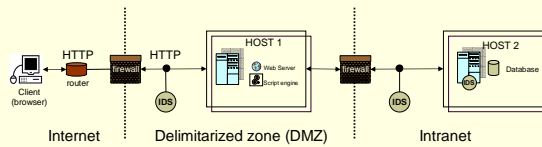
Configurazione 2 tier single host



Configurazione 2 tier single host

- Vantaggi:
 - semplicità nella installazione e nella manutenzione si deve gestire una sola macchina
- Svantaggi:
 - bassa affidabilità: il malfunzionamento di un componente blocca l'accesso all'intero sistema
 - scarsa sicurezza: se un intruso supera il firewall, ha accesso alla macchina con i dati aziendali
 - la scalabilità è limitata alla possibilità di upgrade della singola macchina

Configurazione 3 tier dual host



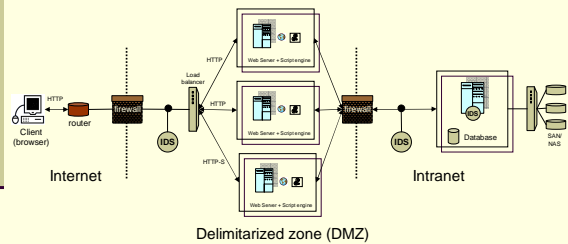
Delimitarized Zone (DMZ)

- consiste in un segmento, o insieme di segmenti, della rete localizzati tra reti protette e non protette
- è necessario un tempo aggiuntivo per violare il secondo firewall e il DBMS, tempo che può essere utile al sistema di *Intrusion Detection (IDS)* e agli amministratori per rilevare e bloccare l'attacco

Configurazione 3 tier dual host

- Migliora la *scalabilità*
 - si può intervenire separatamente su tier intermedio e data tier
- Critica la *disponibilità*
 - il guasto di un componente blocca l'intero sistema

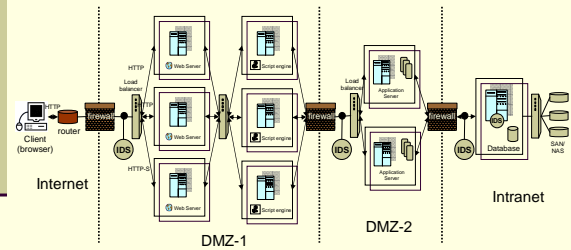
Configurazione 3 tier con Server Farm



Configurazione 3 tier con Server Farm

- *Migliore disponibilità*
 - se cade uno dei processi, il suo carico di lavoro viene distribuito sugli altri processi in esecuzione
- *Migliore scalabilità*
 - collo di bottiglia a qualsiasi livello: è possibile introdurre nuove istanze di processi e nuove macchine server dove necessario, senza il limite della possibilità di upgrade di una singola macchina)
 - replicazione migliora le prestazioni poiché il carico viene distribuito da *processi di load balancing* in modo bilanciato sui server attivi.

Configurazione 5 tier con Server Farm



Vantaggi e Svantaggi delle architetture Web

	Vantaggi	Svantaggi
2 tier single host	<p>Costo: basso, se non serve hardware ad alte prestazioni</p> <p>Complessità: soluzione semplice da installare e mantenere</p> <p>Mantenimento dello stato: semplice memorizzato su una singola macchina</p>	<p>Prestazioni: legate alle caratteristiche della macchina, database e web server competono per le risorse</p> <p>Scalabilità: limitata dalla possibilità di upgrade dell'host</p> <p>Disponibilità: se cade un componente, il sistema non è più accessibile</p> <p>Sicurezza: dati non difesi se il firewall viene superato</p>

Vantaggi e Svantaggi delle architetture Web

	Vantaggi	Svantaggi
3 tier dual host	<p>Prestazioni: dimensionamento più efficace di Web Server e DBMS server</p> <p>Scalabilità: possibilità di intervenire separatamente su middle tier e data tier che hanno requisiti prestazionali differenti</p> <p>Sicurezza: dati su macchine distinti sono più sicuri</p>	<p>Scalabilità: limitata dalla possibilità di upgrade dell'host</p> <p>Disponibilità: un componente fermo blocca ancora il sistema</p>

Vantaggi e Svantaggi delle architetture Web

	Vantaggi	Svantaggi
3 tier con server farm	<p>Prestazioni: distribuzione del carico di lavoro sui server processi in modo bilanciato</p> <p>Scalabilità: se necessario, è possibile aggiungere nuove macchine server</p> <p>Disponibilità: fail-over, se cade uno dei processi, il suo carico di lavoro viene distribuito sui processi funzionanti e il sistema continua a fornire il servizio</p>	<p>Soluzione architetturale complessa sia in termini di gestione e configurazione sia per le problematiche di implementazione del load balancing (per il mantenimento delle sessioni) che di gestione dello stato della sessione (la memorizzazione nel DBMS è onerosa dal punto di vista delle prestazioni)</p>

Prof.ssa E. Gentile

Sistemi Informativi su Web

49

Vantaggi e Svantaggi delle architetture Web

	Vantaggi	Svantaggi
5 tier con server farm	<p>Sicurezza: maggiore livello di sicurezza dovuto al terzo livello di firewall</p> <p>Prestazioni: load balancing dinamico</p> <p>Scalabilità: se necessario, è possibile aggiungere nuove macchine server</p> <p>Disponibilità: capacità di fail-over a livello dei singoli oggetti</p>	<p>Complessità: ambienti generalmente complessi da mantenere</p>

Prof.ssa E. Gentile

Sistemi Informativi su Web

50

Architettura THIN lato Client

- Se l'applicazione web è accessibile attraverso Internet
- Scarso controllo sulla configurazione dei client
- Il client implementa solo l'interfaccia utente grafica
- Il server implementa la logica dell'applicazione e la gestione dei dati

Prof.ssa E. Gentile

Sistemi Informativi su Web

51

Architettura THICK lato Client

- Interfaccia utente sofisticata
- Il client implementa sia l'interfaccia grafica che la logica dell'applicazione
- Il server implementa la gestione dei dati

Svantaggi del Thick Client

- Nessun luogo centralizzato per aggiornare la logica dell'applicazione
- Problemi di sicurezza: il server deve fidarsi dei client
 - Il controllo di accesso e l'autenticazione devono essere gestiti dal server
 - I client devono lasciare la base di dati del server in uno stato consistente
 - Una possibilità: incapsulare tutti gli accessi alla base di dati in stored procedure
- Non scalabile a più di un centinaio di client
 - Grossi trasferimenti di dati tra server e client
 - Più di un server crea un problema, se ci sono x client e y server sono necessarie $x \cdot y$ connessioni
