

Livello cinque (Livello application)

7. Generalità:

In questo livello viene effettivamente svolto il lavoro utile per l'utente, contiene al suo interno diverse tipologie di oggetti:

a) protocolli di supporto alle applicazioni :

a.1) SNMP (simple network mangement protocol) per la gestione della rete;

a.2) FTP (file transfer protocol) per il trasferimento file;

a.3) SMTP e POP3 (simple mail transfer protocol e post office protocol) per la posta elettronica;

a.4) HTTP (hyperText transfer protocol) alla base del world wide web;

b) applicazioni scritte in conformità ai protocolli di cui sopra;

c) applicazioni proprietarie basate su regole di dialogo private;

7.1 Il DNS:

Riferirsi ad una risorsa utilizzando il suo identificativo numerico (indirizzo IP) risulta quanto mai complicato, per questa ragione tutte le risorse del web (email comprese) sono identificate da un nome logico, ossia un identificativo per noi facile da memorizzare.

Questo tipo di traduzione richiede però che ci siano delle macchine che nel momento della richiesta del client convertano il nome logico nell'indirizzo IP, e questo è proprio il compito del DNS, che consiste di:

a) uno schema gerarchico di nominazione basato sul concetto di dominio;

b) un database distribuito che implementa lo schema di nominazione;

c) un protocollo per il mantenimento e la distribuzione delle informazioni;

Volendo schematizzare il funzionamento avremo che:

a) quando un'applicazione deve collegarsi ad una risorsa di cui conosce il nome logico invia una richiesta al dns server locale;

b) se il dns server locale conosce la risposta, la invia direttamente al richiedente, altrimenti interroga il dns server gerarchicamente più in alto e così via, quando alla fine arriverà una risposta il dns server locale la inoltrerà al richiedente;

c) quando l'applicazione riceve la risposta crea una connessione TCP con la destinazione usando l'indirizzo IP appena ricevuto;

I DNS come abbiamo già detto sono organizzati gerarchicamente in domini e sottodomini. Esiste un insieme di top-level domain che sono i più alti nella gerarchia.

Nel caso di un host la struttura di un possibile nome è la seguente:

host.subdomain2.subdomain1.topleveldomain dove la prima sottostringa identifica l'host, le altre i sottodomini e l'ultima il top-level domain che corrisponde ai .com, .it, .edu, .org, .net ...

Ogni dominio è responsabile della creazione dei suoi sottodomini che vanno registrati all'autorità competente.

Inoltre ogni dominio ha la responsabilità di fornire il servizio DNS per quanto di propria competenza, ossia deve poter rispondere a tutte le interrogazioni che riguardano i suoi sottodomini e host.

Le informazioni che si possono ottenere da un name server sono:

a) corrispondenza tra IP e nomi logici per ogni host del dominio

- b) l'identità dei name server responsabili dei sottodomini
- c) l'identità del nameserver del dominio di livello più alto

7.2 La posta elettronica:

La posta elettronica è sicuramente uno dei servizi più consolidati ed utilizzati su internet, quello che consente di effettuare è :

- a) comporre un messaggio;
- b) spedire il messaggio;
- c) ricevere messaggi da altri utenti;
- d) leggere i messaggi ricevuti;
- e) stampare, memorizzare, eliminare messaggi spediti e ricevuti;

Un messaggio di posta è normalmente composto da un header ed un body, separati da una linea vuota.

L'header è composto da diverse linee ciascuna relativa ad una specifica informazione e sono:

From	indirizzo mittente
To	indirizzo destinatario
Cc	indirizzi dei destinatari a cui inviare una copia per conoscenza
Bcc	Blind Carbon Copy (indirizzi segreti)
Subject	argomento del messaggio
Sender	chi materialmente effettua l'invio

Il body contiene testo in caratteri ASCII e l'ultima riga soltanto un punto.

Gli indirizzi di posta elettronica sono nel formato [username@hostname](#), dove username è una stringa che identifica il destinatario e hostname è un nome DNS oppure un indirizzo IP.

La posta elettronica viene implementata attraverso l'utilizzo di due sistemi distinti:

- a) Mail User Agent (MUA): che permette di comporre i messaggi, riceverli e leggerli, consegnarli ad un MTA per la trasmissione, salvarli o eliminarli;
- b) Mail Transport Agent (MTA): si occupa di trasportare i messaggi sulla rete fino alla consegna ad un MTA di destinazione, e risponde ai MUA dei vari utenti per consegnare la posta che hanno ricevuto;

Sono definiti due protocolli per l'utilizzo della posta elettronica l'SMTP per l'invio ed il trasferimento dei messaggi tra MTA e il POP3 per la ricezione dei messaggi sul MUA.

Recentemente sono stati proposti nuovi protocolli più sofisticati per la gestione come IMAP.

Le azioni compiute per inviare un messaggio sono le seguenti:

- a) Pippo compone un messaggio con il suo MUA da spedire a Minnie
- b) appena Pippo preme il pulsante send, il suo MUA:
 - b.1) interroga il DNS per sapere l'indirizzo IP del suo server di posta;
 - b.2) apre una connessione TCP ed effettua una conversazione SMTP con il suo server SMTP e gli consegna il messaggio;
 - b.3) chiude la connessione;
- c) il server SMTP di Pippo :
 - c.1) chiede al DNS l'indirizzo IP dell'host di destinazione
 - c.2) scopre che è quello di un server mail

- c.3) apre una connessione TCP con quel server e gli con un dialogo SMTP gli consegna il messaggio;
 - d) Minnie lancia il suo MUA e preme il tasto check mail:
 - d.1) il suo MUA richiede al DNS l'indirizzo IP del suo server mail
 - d.2) si connette utilizzando TCP e con un dialogo POP3 scarica la posta in arrivo;
- Bisogna tener presente che nei record DNS esistono delle voci specifiche per i server di posta, questi sono identificati dal campo MX.
- Significative estensioni delle funzionalità della posta elettronica sono abbastanza recenti, e sono:
- a) possibilità di inviare messaggi contenenti informazioni di qualunque tipo grazie alle estensioni MIME;
 - b) possibilità di inviare messaggi con firma elettronica o crittografati grazie all'uso dello standard SMIME;

7.3 Il World Wide Web:

Detto anche WWW o W3 è nato al Cern nel 1989 per consentire ai gruppi di ricerca sparsi nel mondo di comunicare tra loro.

È un'architettura software volta a fornire l'accesso e la navigazione ad un enorme insieme di documenti collegati tra loro e sparsi su milioni di elaboratori.

Tutto questo insieme di documenti forma un *ipertesto*, cioè un testo che viene percorso in modo non lineare. Tutti questi documenti che costituiscono l'ipertesto sono detti pagine web e possono contenere collegamenti ad altri documenti, immagini, video, musica, scenari tridimensionali o codice eseguibile.

Le caratteristiche principali del web, che lo hanno portato al suo grande successo sono:

- a) architettura di tipo client – server:
 - a.1) ampia scalabilità;
 - a.2) adatta ad ambienti di rete;
- b) architettura distribuita:
 - b.1) perfettamente in linea con le esigenze di gestione di un ipertesto;
- c) architettura basata su standard di pubblico dominio:
 - c.1) possibilità per chiunque di proporre una implementazione;
 - c.2) uguali possibilità di accesso per tutte le piattaforme;
- d) capacità di gestire informazioni di diverso tipo, suscitando grande interesse nel pubblico;

7.3.1. Standard:

Esistono tre standard principali, che nel loro insieme costituiscono l'architettura software del web:

- a) *sistema di indirizzamento* basato su *uniform resource locator (URL)* : è un meccanismo standard per fare riferimento alle entità indirizzabili nel web;
- b) *linguaggio HTML*: è il linguaggio con il quale si compongono le pagine web;
- c) *protocollo HTTP*: è il protocollo di comunicazione tra client e server;

7.3.2. URL:

Una URL costituisce un riferimento ad una risorsa, tale risorsa risiede ovviamente da qualche parte e proprio l'url specifica come si vuole accedere alla risorsa, dove è fisicamente localizzata e come è identificata questa risorsa.

Una URL è quindi composta da tre parti:

- a) il metodo d'accesso;
- b) l'host di destinazione;
- c) l'identità della risorsa;

Il metodo d'accesso indica il protocollo che va utilizzato per poter colloquiare con il server che mantiene la risorsa, i metodi d'accesso più comuni sono:

http	protocollo nativo Web
ftp	file transfer protocol
gopher	vecchio protocollo per il riempimento di informazioni
mailto	usato per spedire posta
telnet	terminale virtuale
file	accesso a documenti locali
news	accesso ai gruppi di discussione

Il nome dell'host, può invece essere o l'indirizzo IP o più comunemente il nome DNS del server, dopo il nome può essere incluso anche il numero di porta, ma se viene omissso si intende la porta standard 80.

L'identità della risorsa consiste nel path e nel nome del file da raggiungere, la sintassi è di tipo Unix.

7.3.3. HTML:

E' lo standard per la formattazione del testo nelle pagine web, è una specializzazione del linguaggio SGML.

E' un Markup Language, nel senso che la sua sintassi è composta di marcatori che modificano il testo che contengono, tipicamente la struttura di una pagina è composta da due elementi principali, l'header e il corpo della pagina.

Quando un client richiede una pagina, il suo browser interpreta questi marcatori riformattando il testo, e mostrando a video il risultato.

7.3.4. HTTP:

E' il protocollo che sta alla base di tutte le comunicazioni sul web, ne esistono due versioni, la seconda delle quali ancora non è uno standard ma è in fase di proposta, la sua struttura è molto semplice, usa una comunicazione di tipo ASCII quindi convoglia solo caratteri, ed è stateless, ossia ogni richiesta è una connessione a parte e non si tiene traccia di quelle precedenti, quindi volendo schematizzare il suo comportamento abbiamo:

- a) apertura connessione tra client e server;
- b) invio richiesta singola per l'oggetto;
- c) invio da parte del server di una risposta;
- d) chiusura connessione;

Analizzando in dettaglio le fasi di invio della richiesta e di risposta possiamo notare che il funzionamento è il seguente:

a) Invio richiesta:

- a.1) si specifica il metodo della richiesta (GET, POST, PUT, DELETE, HEAD...);
- a.2) si specifica il numero di versione del protocollo;
- a.3) si indica l'oggetto che si vuole richiedere;
- a.4) varie altre informazioni tra le quali:
 - a.4.1) il tipo di client;
 - a.4.2) i tipi di dati che il client può accettare

b) Risposta del server:

La risposta del server è articolata in più parti, poichè esiste un problema di fondo, ossia come fa il client a sapere in che modo dovrà gestire le informazioni che gli verranno inviate?

- b.1) una riga di stato che indica l'esito della richiesta (200 OK, 301 Moved permanently, 404, 500..), composta a sua volta da:
 - b.1.1) versione del protocollo
 - b.1.2) codice di stato
 - b.1.3) specifica testuale dello stato
- b.2) le metainformazioni che descrivono la natura del messaggio;
 - b.2.1) uno dei campi più significativi è il Content-type dove viene indicato il tipo MIME (multipurpose internet mail exchange) del messaggio, i tipi mime sono identificati dalla coppia *Mime type/Mime subtype*.
- b.3) le informazioni vere e proprie;

7.3.5. I problemi:

I problemi principali di cui soffre il web, che ricordiamo è nato come mezzo di comunicazione per gruppi di ricercatori, sono essenzialmente dovuti all'ampio sviluppo che ha conseguito negli ultimi tempi, e al fatto che si sta cercando di sfruttarlo sempre più per transazioni di tipo commerciale.

Questi due aspetti hanno messo in luce tutte le sue limitazioni come l'efficienza nell'uso delle risorse di rete, l'assenza del concetto di sessione che comporta grandi sprechi per instaurare per ogni oggetto una connessione, e la carenza di meccanismi per lo scambio di dati sensibili.