Web Server: IIS e Apache

Nicola Gessa

In questa lezione:

- Cos'è e come funziona Internet Information Service
- Come configurare Intenet Information Service
- Cos'è e come funziona Apache
- Come configurare Apache

Internet Information Service

Con particolare riferimento alla versione 4.*

Introduzione a IIS

Definizione

IIS: Internet Information Services è il server web di Microsoft per sistemi Windows.

Una volta installato presenta la seguente struttura :

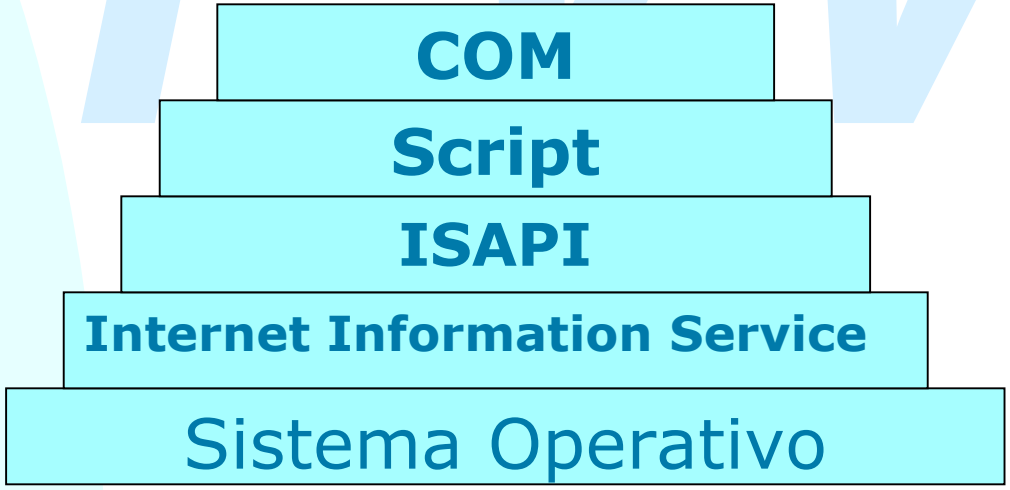
- \Inetpub
- *systemroot*\help\iishelp
- *systemroot*\system32\inetsrv

IIS crea anche un sito web di default al momento dell'installazione in \Inetpub\wwwroot dove si possono inserire delle pagine web

Architettura di IIS

IIS è parte integrante della architettura Windows DNA (Distribuited interNet Application Architetture), che definisce un framework per lo sviluppo e la distribuzione di applicazioni web e che comprende servizi quali

- Active Server Pages
- Componenti COM (Component Object Model)
- Windows security services
- Microsoft Data Acces Component



IIS request Processing

Quando IIS Riceve una richiesta, determina dall'URL se riguarda un documento statico (HTML) o dinamico

- Pagina HTML: ritorna la pagina immediatamente
- File corrispondenti a determinate estensioni ISAPI:carica la DLL appropriata con cui gestire la richiesta (ad esempio le richieste di file .asp vengono dirette al file APS.dll)
- Applicazioni CGI : viene creato un nuovo processo a cui passare tutti i parametri attraverso l'ambiente di gestione del processo

Configurare IIS

- La configurazione di ciascun sito viene registrato in un metabase, che assume valori di default al momento dell'installazione.

- Ogni sito web in IIS ha associato un insieme di "classi" per gestire le proprietà del sito.

- Esistono tre livelli gerarchici di configurazione:
  - Master: fissa una configurazione comune a tutti i siti web
  - Site by site: fissa una configurazione specifica per un sito
  - File by file: fissa una configurazione specifica per un file

- La configurazione di IIS si appoggia sulla configurazione del sistema

Organizzazione dei contenuti

I contenuti di un sito web può essere distribuito iin due modi:

1. Fissando una directory di root locale e includendo quindi tutte le sottodirectory
2. Associando al sito un certo numero di Virtual Directory. Sono directory non necessariamente contenute fisicamente in IIS ma che fanno parte dell'insieme dei contenuti associati ad un sito web. Possono essere di due tipi.
    1. Directory locali, ma non dipendenti dalla directory di root
    2. Directory remote identificate da URL

Proprietà delle Virtual Directory

- Flessibilità
- Scalabilità
- Prestazione ridotte

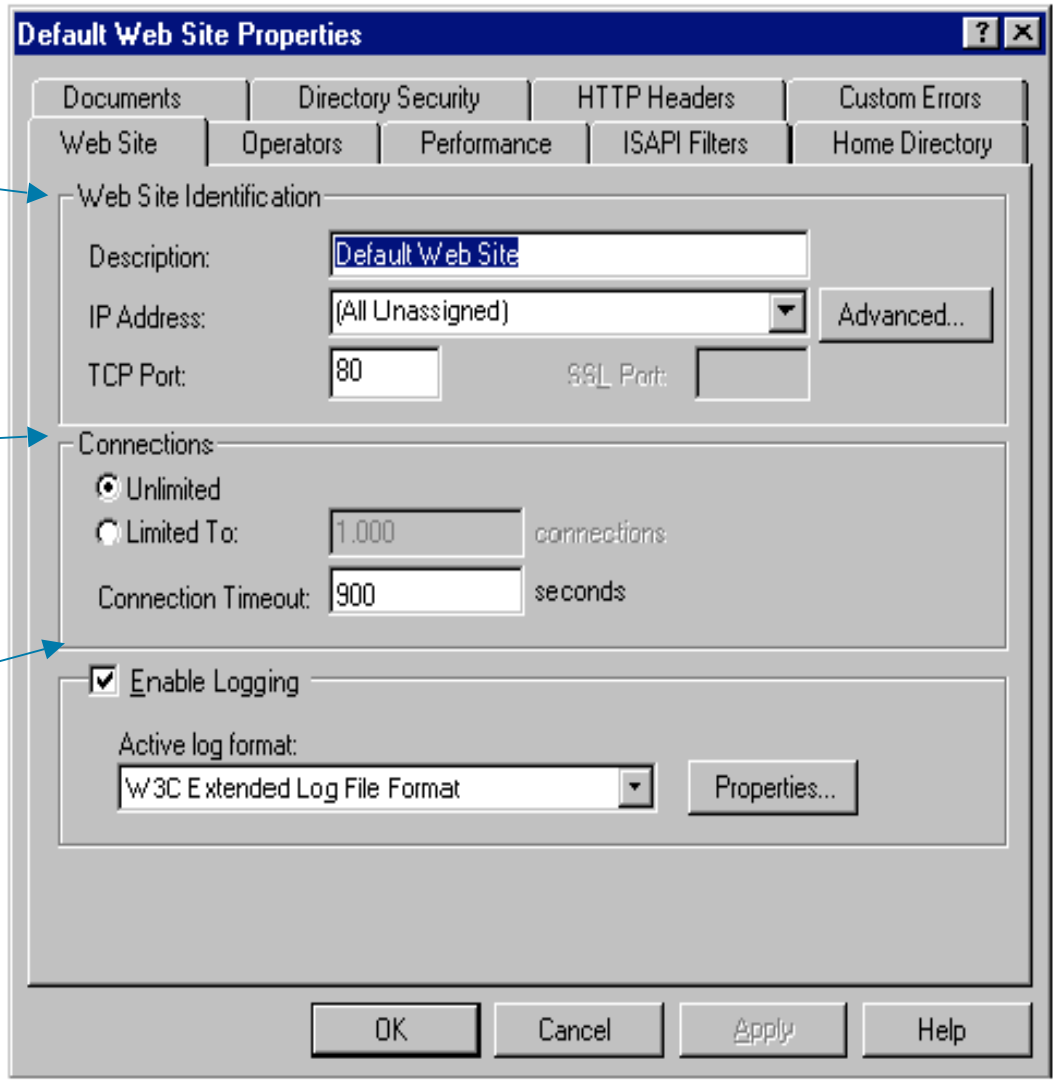
Configurare IIS

Il web server può essere configurato tramite la console Internet Service Manager che definisce 9 sezioni principali

- Web site
- Operators
- Performance
- Isapi Filter
- Home Directory
- Documents
- Directory Security
- HTTP Headers
- Custom errors

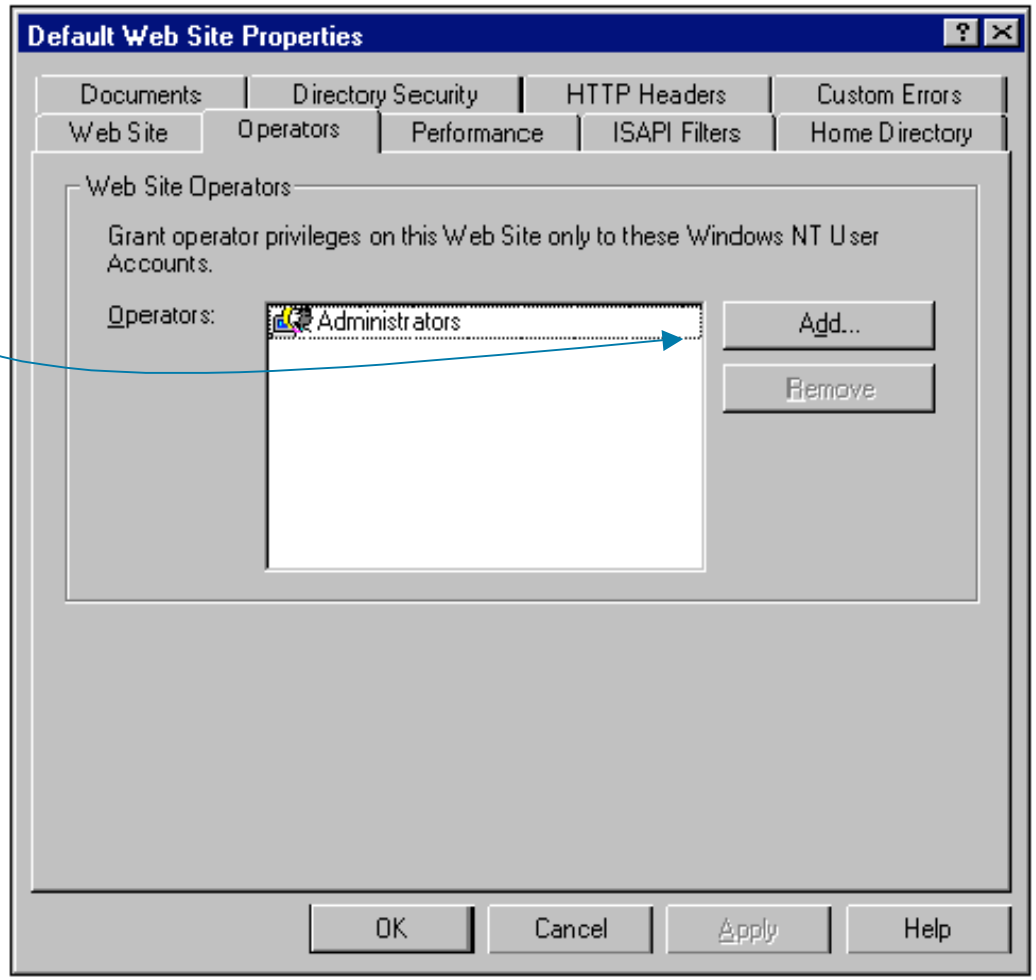
Web Site

- Identificazione del Sito Web
- Controllo sul numero delle connessioni permesse
- Controllo dei meccanismi di logging del sito
  - Come creare i log
  - Quali formati per i file di log



Operators

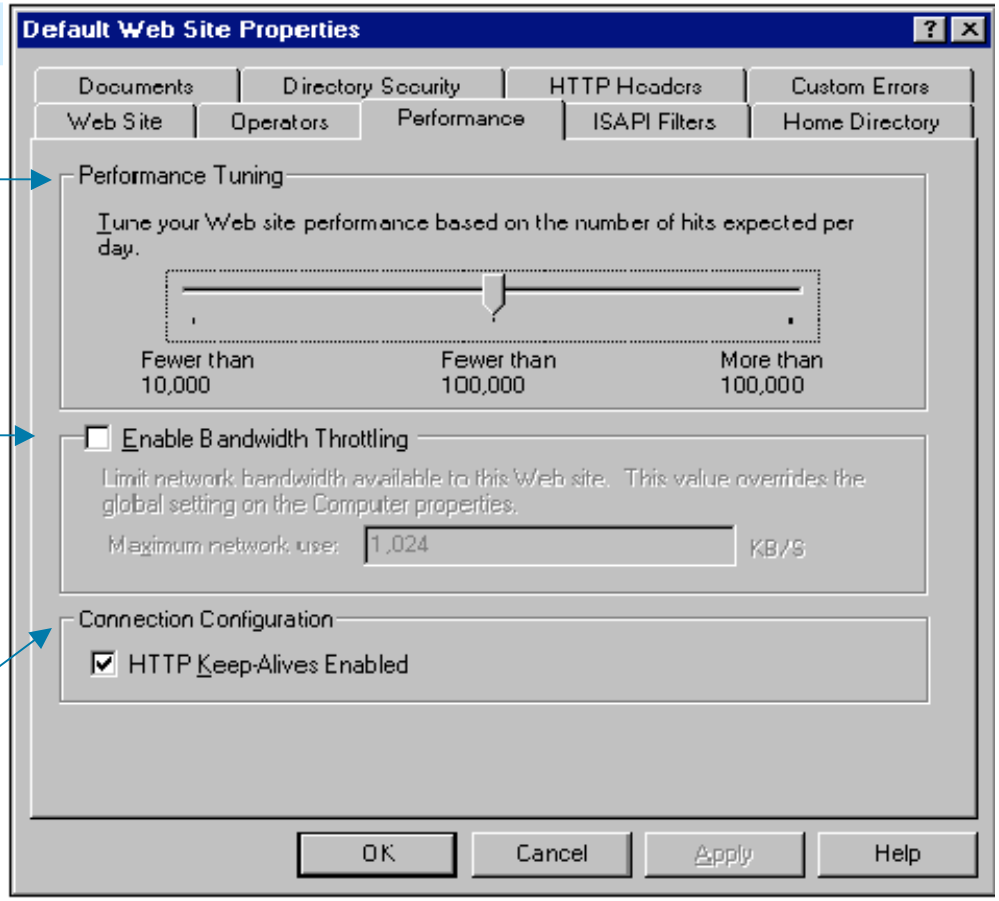
Consente di assegnare i privilegi di amministrazione ad utenti del sistema



Performance

Consente di migliorare le prestazioni del sito Web tramite la cinfigurazione di alcuni parametri

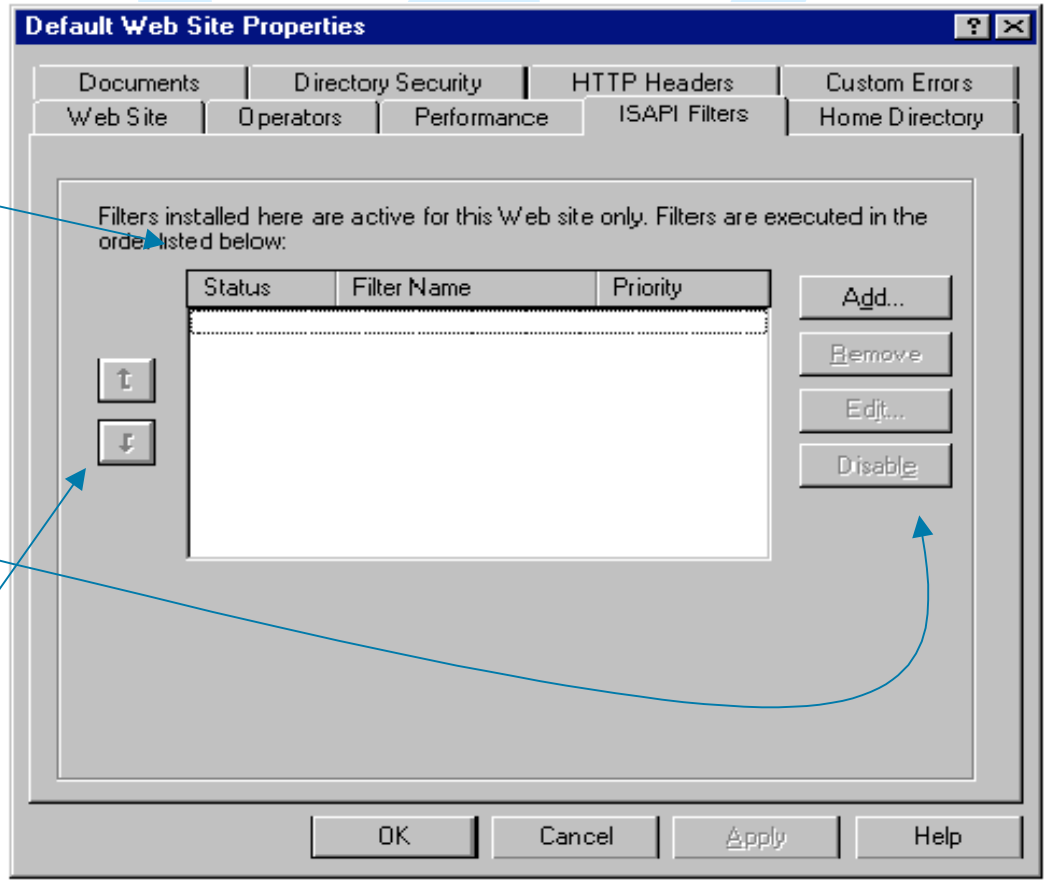
- Previsione del numero di connessioni giornaliere al sito
- Controllo dalla bandwidth disponibile alle connessioni del sito
- Gestione di connessioni HTTP persistenti



ISAPI Filters

Permette la configurazione dei filtri ISAPI (Internet Server Application Programming Interface). I filtri sono associati a determinate estensioni dei file e attivati al momento della richiesta di file con tali estensioni

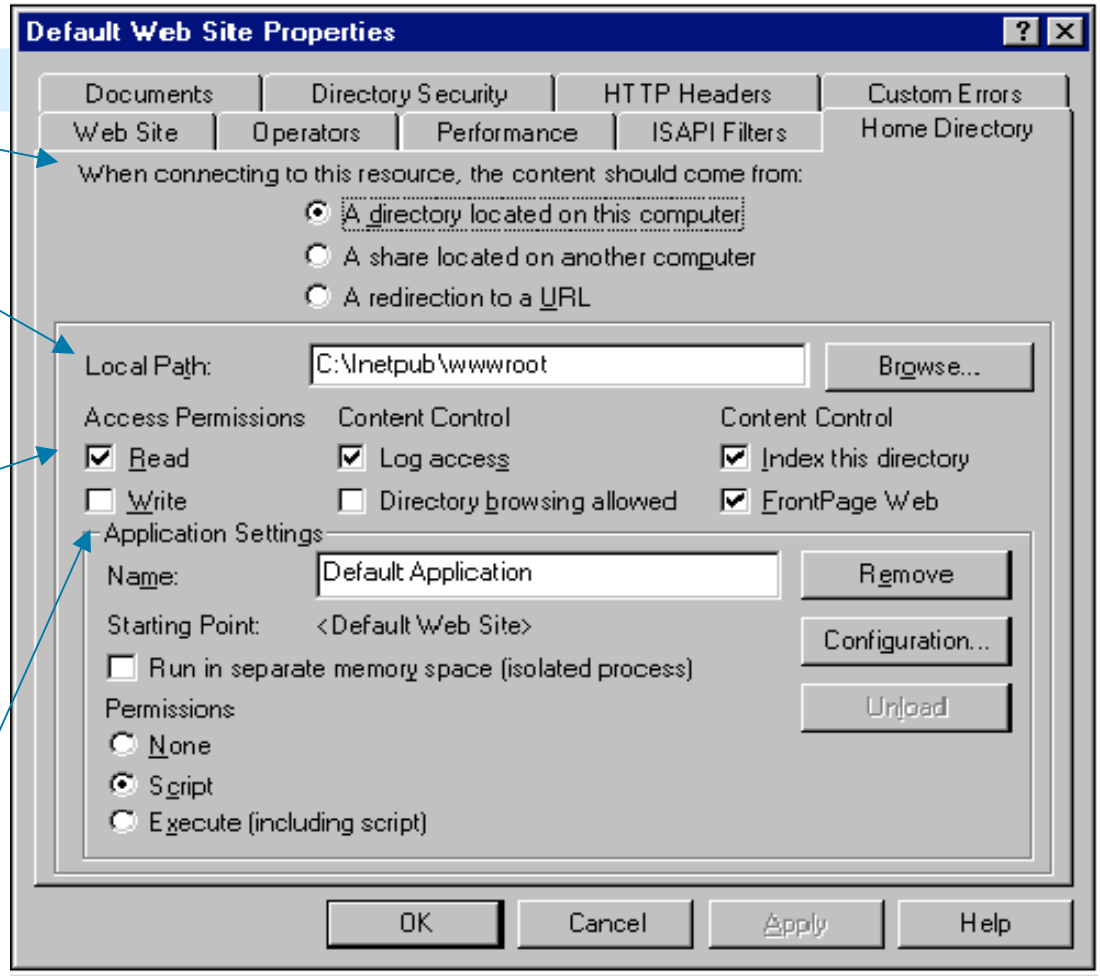
- Lista dei filtri ISAPI con stato (Loaded, Unloaded, Disabled), nome, priorità (High, Medium, Low
- Possibilità di aggiungere e rimuovere i filtri o modificarne lo stato
- Modifica delle priorità dei filtri



Home Directory

Consente di configurare le caratteristiche della directory principale del sito

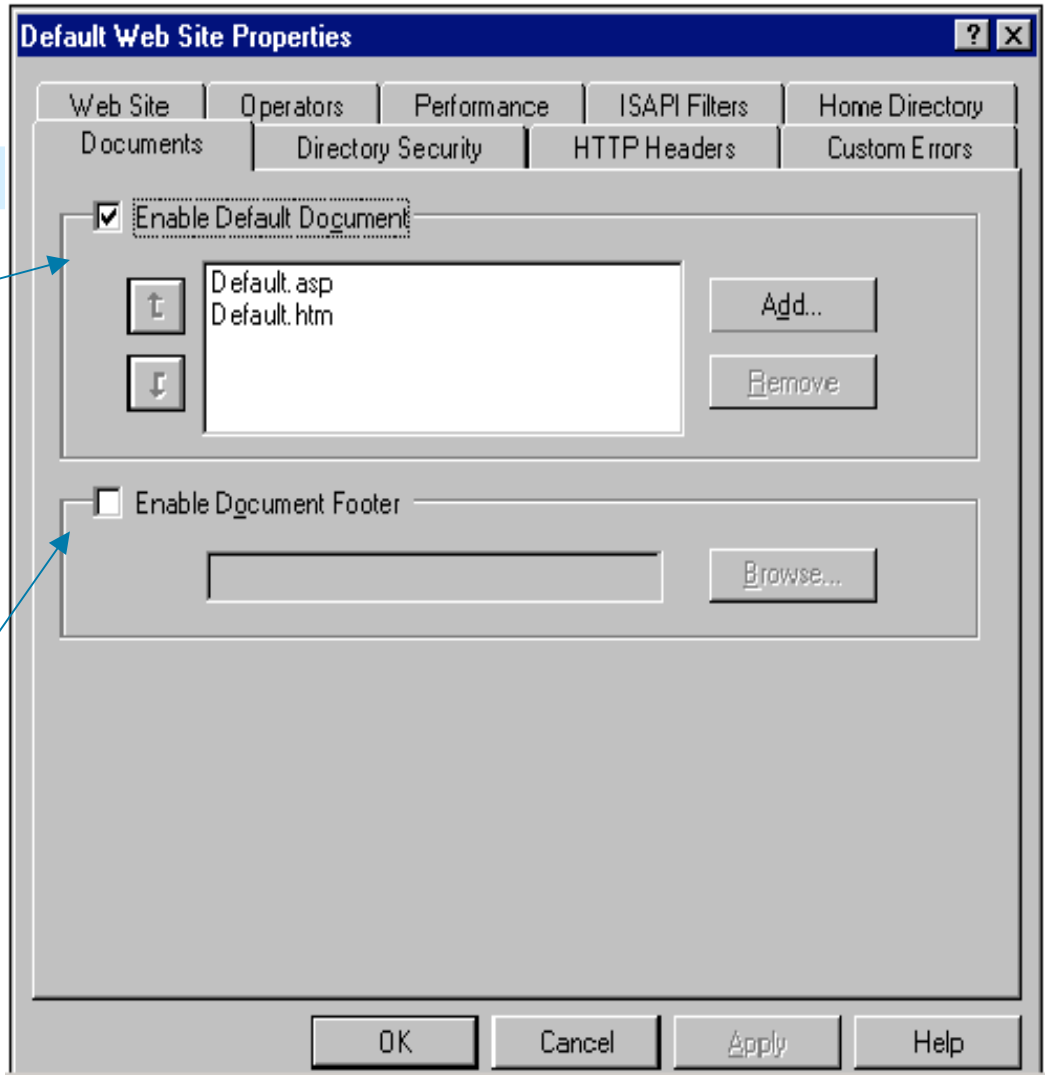
- Locazione della home directory
- Specifica del Path della home directory
- Gestione dei permessi di accesso e della navigazione dei contenuti
- Gestione delle applicazioni associate alla directory



Documents

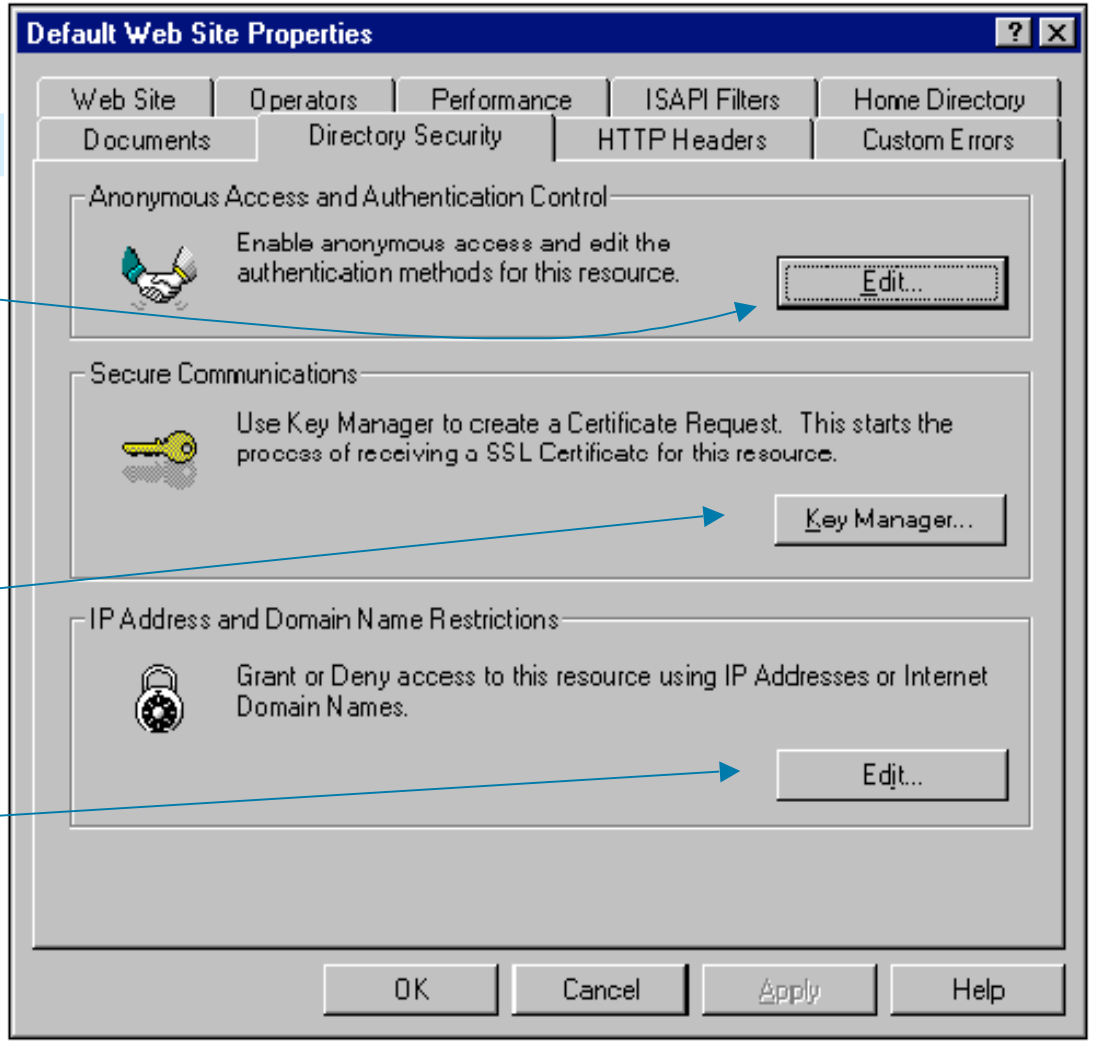
Consente di specificare i documenti di default e gli attach delle pagine web

- Definisce i documenti di default da inviare
- Specifica il footer da inserire automaticamente nella pagina



Directory Security

- Configurazione dell'accesso anonimo e e i metodi per il controllo dell'autenticazione
- Creazione della richiesta per ricevere un certificato digitale SSL
- Gestione delle politiche di accesso
  - granted access
  - denied access



Anonymous Access and Authentication Control

Metodi di configurazione per l'accesso al server

- Allow Anonymous Access: consente l'accesso al server a utenti non riconosciuti (anonymous user), creando una connessione anonima e assegnando loro un account anonimo. L'account usato è un account valido per il sistema.

- Basic Authentication: esegue un riconoscimento dell'utente che vuole accedere al server tramite la richiesta di login e password trasmessi in chiaro sulla connessione.

- Windows NT Challenge/Response: meccanismo di autenticazione basato sullo scambio non in chiaro di login e password che richiede l'uso di Microsoft Internet Explorer versione 2.0 o successiva

Cosa è un certificato?

- I certificati sono una forma di identificazione digitale che consente ai Web server a ai client l'autenticazione prima di stabilire una connessione e comprende le chiavi pubbliche
- I certificati sono parte del protocollo SSL per l'utilizzo di connessioni sicure: per sfruttare connessioni SSL un Web server **deve** ottenere ed installare un certificato
- I certificati contengono chiavi utilizzate per criptare le connessioni
- I certificati possono essere essere adottati sia dal client che dal server solo con SSL 3.0
- I certificati sono rilasciati da organizzazioni per la certificazione che garantiscono sulle informazioni contenute nel certificato
- I certificati sono utilizzabili solo dal sito a cui sono associati

Gestione dei certificati

Per installare una connessione SSL sul un sito, si deve creare un certificato che garantisca l'identità del sito. IIS mette a disposizione una procedura per **la creazione della richiesta di certificato**

- si lancia il tool di IIS che richiede l'inserimento delle informazioni da includere nel certificato.
- questo tool genera un codice criptato che rappresenta la richiesta da inoltrare ad un ente certificatore per ottenere il certificato vero e proprio.
- ottenuta questa richiesta, ci si collega ad un sito che fornisca i certificati (www.instantssl.com)
- si spedisce via form la richiesta criptata
- si riceve per posta il certificato vero e proprio
- In seguito IIS permete di installare il certificato ricevuto, di modificarlo, cambiarlo ecc. e abilitare la connessione SSL

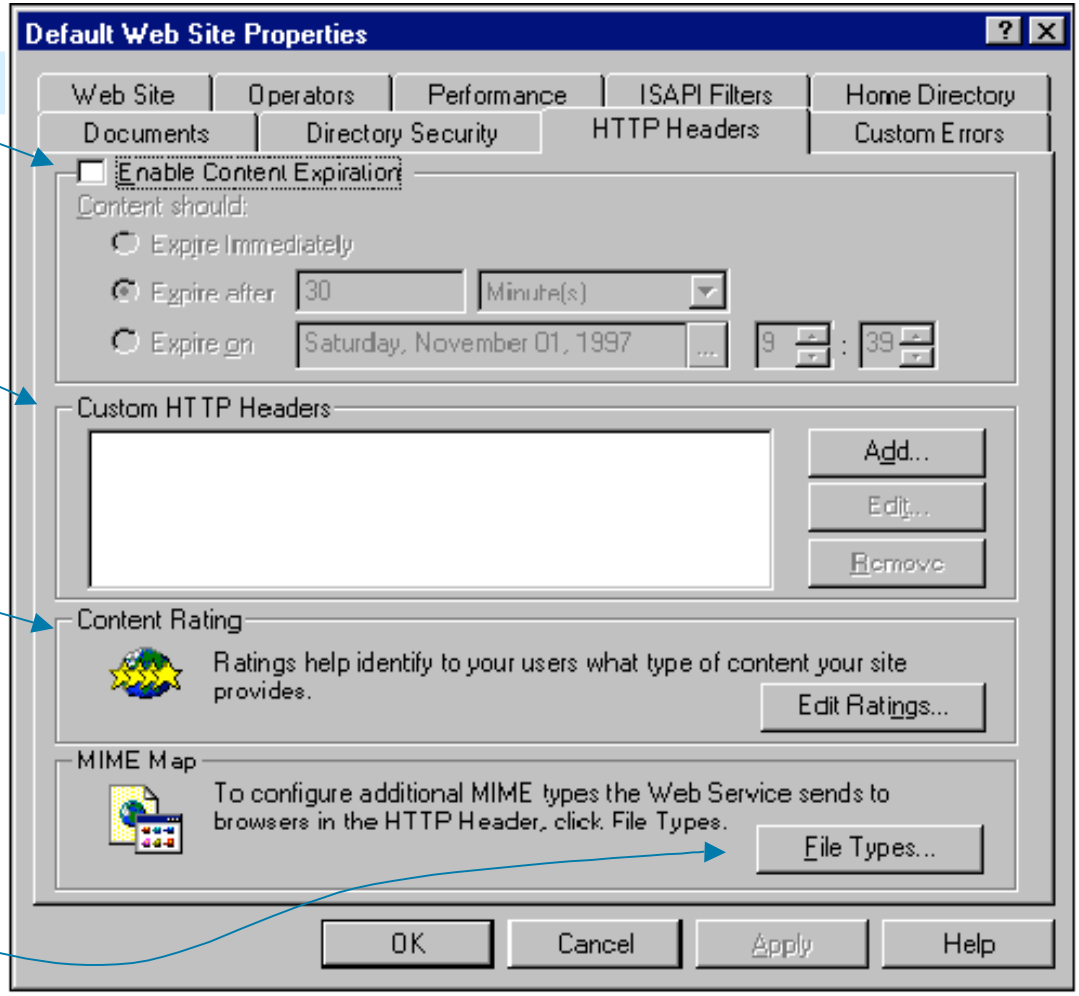
Utilizzo dei certificati

- Il browser cerca di stabilire una connessione sicura (le pagine si richiamano con https://)
- Il server invia il suo certificato con la chiave pubblica
- Il browser genera una chiave di sessione e la cifra utilizzando la chiave pubblica del server e gliela invia
- Utilizzando la sua chiave privata il server decodifica la chiave di sessione e il canale è così stabilito.
- Server e browser comunicano usando la chiave di sessione

HTTP Headers

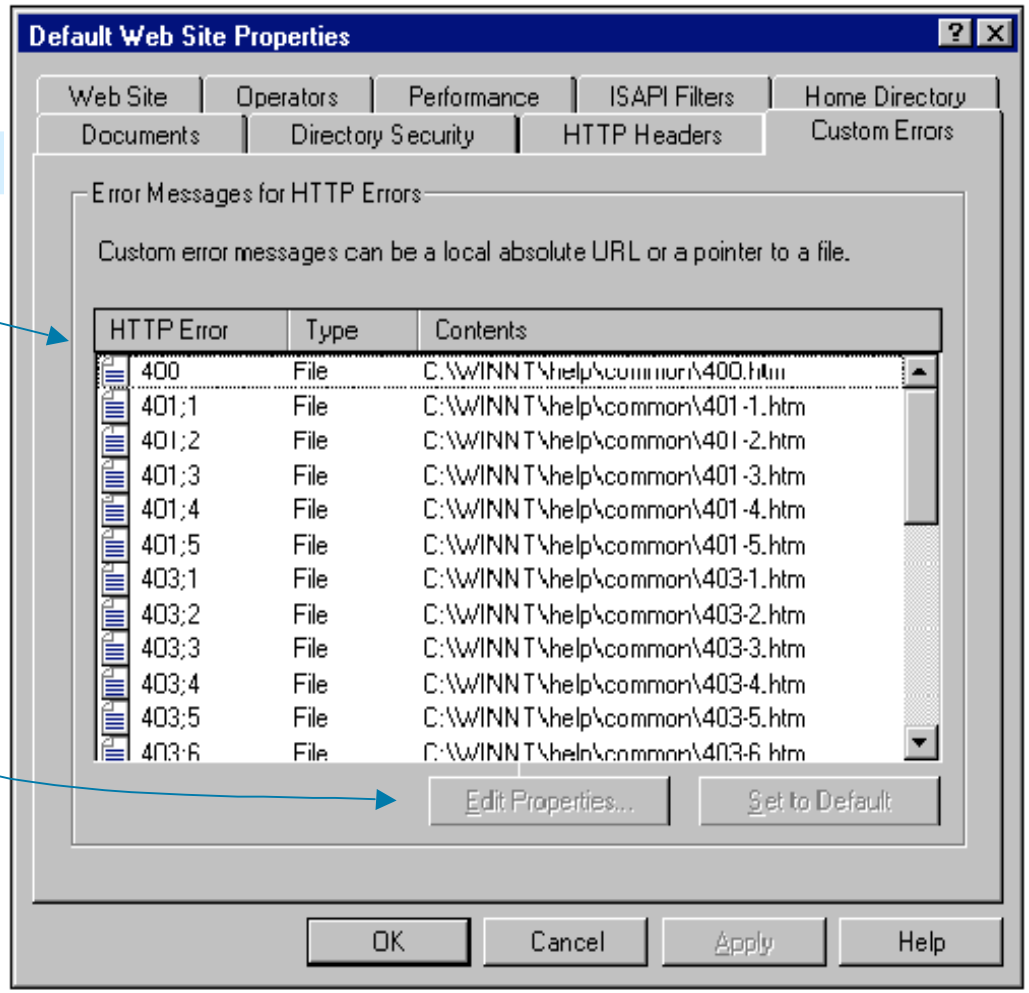
Consente di creare specifici header HTTP da spedire al browser

- Gestione della scadenza delle pagine
- Creazione di “custom HTTP Header”
- Classificazione via PICS (secondo le categorie ICRA) dei contenuti del sito
- Gestione dell’associazione ai file dei tipi MIME (Multipurpose Internet Mail Extension)



Custom Error

- Lista dei messaggi restituiti al browser in caso di errore HTTP
- Personalizzazione o redirezione degli errori HTTP



Apache

Con particolare riferimento alle versioni 1.3 (attualmente la più diffusa) e 2.0 (recentemente rilasciata)

Apache HTTP Server

- Apache HTTP Server è il web server HTTP della Apache Software Foundation
- Apache viene sviluppato e aggiornato per sistemi operativi come UNIX e Windows
- Apache è open-source
- Apache è dal 1996 il web server più diffuso in Internet (nel 2002 il 63% dei siti web usavano Apache)

Compilazione e installazione per sistemi Unix-like

- Download:
  - http://www.apache.org/dist/httpd/httpd-2_0_NN.tar.gz
  - http://httpd.apache.org/download.cgi: lista di siti mirror per il download
- Estrazione
  - gzip –d httpd-2_0_NN.tar.gz
  - tar xvf httpd-2_0_NN.tar
- Configurazione del sistema
  - environment variables
  - pathnames
  - moduli
- Compilazione – necessario un compilatore ansi C (es. GCC)
- Installazione
- Configurazione parametri di Apache
- Esecuzione

Starting Apache

- In Unix il programma httpd viene eseguito in background come demone che gestisce le richieste

- Il programma httpd può essere eseguito tramite lo script apachectl, che configura alcune variabili di ambiente dipendenti dal sistema operativo necessarie al funzionamento del server.

- Lo script apachectl permette il controllo del ciclo di vita del web server
  - start, stop,restart,graceful restart

- Httpd legge il file di configurazione httpd.conf

- Per lanciare il demone httpd all'avvio del sistema si deve aggiungere una chiamata allo script apachectl ai file di startup. Questo fa eseguire Apache come root.

Modularità di Apache

Apache è un server modulare: le sue funzionalità sono rese disponibili tramite dei moduli che possono essere aggiunti o sottratti al server sfruttando il modulo mod_so

- Solo le funzionalità di base sono incluse nel core Server
- Nel processo di configurazione si possono scegliere i moduli da aggiungere al server
- Alcuni moduli sono inclusi per default, gli altri devono essere esplicitamente aggiunti
- I moduli possono essere :
  - Compilati staticamente nel web server
  - Aggiunti dinamicamente nel file httpd.conf sfruttando il supporto DSO (Dynamic Shared Object)
- Esempi: mod_auth (autenticazione utenti usando file di testo), mod_cgi (esecuzion di moduli cgi), mod_ISAPI (estensione ISAPi per Apache sotto Windows)

Multi – Processing Modules (MPMs)

- Apache è stato progettato per essere flessibile su ogni tipo di piattaforma e con “ogni” configurazione d’ambiente
- Apache fornisce un insieme di moduli per la gestione di operazioni quali il binding, accept delle richieste, gestione dei processi / threads per gestire le connessioni e consente a
  - sistemi operativi differenti di fornire moduli MPM appropriati per ottenere maggiore efficienza
  - agli amministratori, in base alle esigenza del proprio web server di applicare politiche diverse
- Il modulo MPM deve essere compilato nel server (sfruttando l’ottimizzazione di alcuni compilatori nella gestione dei threads, se sono usati)
- Solo un modulo MPM per volta può esserecaricato nel server
- modili MPM sono: mpm_winnt, prefork,mpmt_os2

Esempio di configurazione

```
$>CC=$"pgcc" \
.configure --prefix="/sw/pkg/apache \
--enable-cgi
--enable-rewrite=shared

$>make

$>make install

$>apachectl start
```

File di configurazione

- Apache viene configurato tramite delle direttive riportate i file di testo
- Il file di configurazione principale è httpd.conf, ma se ne possono creare di nuovi
- La configurazione può essere aggiornata solo riavviando Apache
- La gestione dei tipi MIME è gestita nel file mime.types
- Le direttive si applicano all'intero server, a meno di specificare lo scope, se consentito, in cui sono valide (es. usando <Directory> per definire la directory in cui applicarla)
- Apache consente di decentralizzare la configurazione dell'albero delle directory usando i file .htaccess, le cui direttive vengono applicate solo al sotto albero in cui il file .htaccess è presente.

Uso delle direttive nei file di configurazione

- Le direttive nei file di configurazione possono essere applicate all'intero server, a directory, a file, host o URL
- Possono essere valutate solo avvio del server (es. <If Module> oppure ad ogni richiesta ricevuta
- Possono riguardare la struttura del filesystem (es. <Directory>) oppure la struttura del sito (Es. location)
- Alcune direttive consentono l'uso di espressioni regolari secondo la sintassi Perl (es. DirectoryMatch)

ES:

```
<Directory /home/Rossi/public_html
Options Indexes
</Directory>
```

Autenticazione

- Può essere gestita nel file httpd.conf all'interno di una sezione <Directory> o nel file .htaccess
- Richiede la creazione di un file delle password (htpasswd -c /usr/locl/apache/passwd/password) utente. Utilizzando il modulo mod_auth_dbm le utenze possono essere gestite tramite un database
- Configurazione degli accessi inserendo le seguenti direttive nei file di configurazione:

```
AuthType Basic
AuthName "Restricted Area"
AuthUserFile /usr/locl/apahce/passwd/password
require user nomeUtente
```

- L'abilitazione all'accesso può essere anche assegnato in base all'host name o all'host address della macchina che richiede il documento

```
allow from 205.252.46.165
```

Autenticazione

- AuthType Basic
  Seleziona il metodo di autenticazione utilizzato
  Basic è implementato dal modolo mod_auth
  Digest è implementato dal modulomod_auth_digest

- AuthName “Restricted Area”
  identifica un nome per l’area a cui accedere per consentire il riutilizzo delle passord al client

- AuthUserFile /usr/locl/apache/passwd/password
  specifica il file contenente le password

- require user nomeUtente
  specifica il nome dell’utente abilitato all’accesso. Si possono anche creare dei gruppi per consentire l’accesso a più utenti

SSL con APACHE

- la Apache Software Foundation non include nel progetto Apache Httpd il modulo per il supporto di SSL

- esistono due progetti open source che si propongono di garantire il supporto di SSL :mod ssl e Apache-SSL , ed entrambi gestiscono le connessioni sicure appoggiandosi alle librerie del progetto OpenSSL

- esistono altre implementazioni commerciali di moduli SSL per Apache:"Covalent Raven SSL Module For Apache", la distribuzione IBM di Apache "IBM Http Server".

Direttive per l'uso di SSL

- **SSLSessionCache**: imposta il tipo di memorizzazione della cache
- **SSLEngine** è la direttiva che abilita o disabilita le connessioni SSL
- **SSLProtocol** controlla la versione del protocollo che verrà usata durante le transazioni sicure. Es. SSLProtocol All -SSLv2
- **SSLCertificateFile** e **SSLCertificateKeyFile** contengono banalmente il path ai file certificato e chiave del server
- **SSLRequireSSL** :usata nelle sezioni "<directory>" ci consente di rendere accessibili alcune aree del nostro server solo attraverso connessioni sicure
- **SSLCipherSuite** permette di configurare gli algoritmi crittografici che il client può usare quando viene stabilita la connessione

Variabili d'ambiente

- Apache utilizza variabili d'ambiente per gestire operazione come il log o il controllo degli accessi e per comunicare con programmi esterni come script CGI
- Le direttive SetEnv, SetEnvIf permettono di definire le variabili d'ambiente, specificando eventualmente determinate condizioni
- Le variabili possono essere utilizzate all'interno delle direttive per regolare operazioni del server

Utilizzo di script CGI con Apache

- Il CGI (Common Gateway Interface) fornisce al web server un meccanismo per interagire con programmi esterni nella creazione dei contenuti attraverso l'uso di programmi (script) CGI.
- Ci sono due metodi per configurare Apache nell'utilizzi di CGI.
  - Definire la directory che contiene gli script, i quali devono avere i diritti di esecuzione

    ScriptAlias /cgi-bin /usr/local/apache/cgi-bin

  - Dichiarare il permesso di eseguire gli script all'interno di determinate directory

    <Directory /usr/local/apache/htdocs/somedir>

    Options +ExecCGI

    </Directory>

    e specificare l'estensione dei file da eseguire con

    AddHandler cgi-script cgi pl

Gestione degli utenti

- Apache consente in vari modi di definire per ogni utente del sistema una directory nel quale rendere disponibili sul web i propri documenti, immagini ecc.

UserDir public_html

UserDir /var/html

UserDir /var/www/*/public_html

- Si possono selezionare gli utenti che possono usare tale funzionalità

- Utilizzando la direttiva

<Directory /home/*/cgi-bin/>

Options ExecCGI

SetHandler cgi-script

</Directory>

si abilitano gli utenti all'uso di script CGI personali

URL Rewriting

Apache fornisce con il modulo mod_rewrite una tecnica per manipolare le URL richieste in modo tale da consentirne una elaborazione interna o una redirezione esterna

Tale tecnica

- utilizza le espressioni regolari
- permette di specificare un numero illimitato di regole di riscrittura
- consente di specificare un numero illimitato di condizioni per la riscrittura
- può essere adattata a parametri come variabili d'ambiente, header HTTP, time stamp
- può operare sia nel contesto del web-server (nel file httpd.conf) che in quello di una directory (nel .htaccess), ma con differenti prestazioni

URL Rewriting:esempi

- Spostare le home su un webserver differente

```
RewriteEngine on
RewriteRule     ^/~(.+)   http://newserver/~$1      [R,L]
```

- Cambiare la struttura delle directory

```
RewriteEngine on
RewriteRule     ^/~(([a-z][a-z0-9]+)(.*)     /home/$2/$1/.www$3
```

- Modificare i contenuti i base all'ora del giorno

```
RewriteEngine on
RewriteCond     %{TIME_HOUR}%{TIME_MIN}     >0700
RewriteCond     %{TIME_HOUR}%{TIME_MIN}     <1900
RewriteRule     ^home\.html$          home_day.html
RewriteRule     ^home\.html$          home_night.html
```

Riferimenti

- http://www.microsoft.com/technet/default.asp
- http://msdn.microsoft.com/
- http://httpd.apache.org
- www.openssl.org