

Esame del Sistema Operativo Linux Parte 3 - Indice

1. Caratteristiche del *File System*
2. Implementazione del *File System*
3. *File System* di rete (NFS)

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 165

Esame del Sistema Operativo Linux Caratteristiche del File System - 1

- Basato su paradigma minimalista
Small is beautiful
- *File* visto da FS come sequenza di *byte* di contenuto arbitrario
 - Significato assegnato dall'utente
- *File* regolari, *file* catalogo (*directory*) e *file* speciali (I/O di dispositivi)
- Nome inizialmente limitato a 14 caratteri (UNIX v7) poi esteso fino a 255 (UNIX BSD → Linux)
 - Estensione non richiesta, con convenzione a scelta dell'utente (e/o del programma applicativo)

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 166

Esame del Sistema Operativo Linux Caratteristiche del File System - 2

- Designazione di *file* mediante cammino (*path*) sia assoluto che relativo
 - Il cammino relativo richiede la nozione di "posizione (*directory*) corrente"
 - `pwd` per visualizzarne la designazione assoluta
 - `cd` per cambiare posizione
 - Un intero FS **B** posto su una partizione visibile può essere inglobato come parte di un FS **A** pre-esistente (`mount`)
 - La radice di **B** viene designata con un nome (cammino) specifico in **A** detto *mount point*

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 167

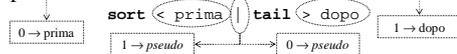
Esame del Sistema Operativo Linux Caratteristiche del File System - 3

- Controllo di accessi concorrenti (*locking*)
 - A grana grossa (per *directory* o per *file*)
 - A grana fine (per gruppi di *byte* in un *file*)
- 2 modalità
 - Accesso simultaneo condiviso (*shared lock*)
 - Consente più accessi condivisi dello stesso tipo alla stessa zona, anche in parziale sovrapposizione tra loro
 - Accesso esclusivo (*exclusive lock*)
 - Consente un solo accesso per zona selezionata

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 168

Esame del Sistema Operativo Linux Caratteristiche del File System - 4

- \forall *file* aperto l'utente ha un descrittore (`int > 0`) che indica anche la posizione corrente di lettura/scrittura
- 3 descrittori di *file* pre-assegnati dalla shell, per altrettanti *file* aperti per definizione
 - 0 per `stdin`, 1 per `stdout`, 2 per `stderr`
 - Comandi di redirectione ">" e "<" modificano tali assegnamenti
- La *pipe* "`|`" costituisce uno pseudo-file (con descrittore proprio) in lettura/scrittura per una coppia di processi



Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 169

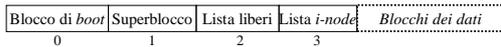
Esame del Sistema Operativo Linux Esempi di chiamate di File System

- Disponibili all'utente solo incapsulate in procedure di libreria
 - `lseek` : fissa l'indice di posizione all'interno di un file (come *offset* espresso in *byte* rispetto ad un riferimento dato) ← accesso diretto
 - `stat` : fornisce informazioni su *file* prelevandole dall'*i-node* corrispondente
 - Disponibile all'utente tramite comando `stat` di *shell*
 - Provare per esercizio (dopo aver letto "`man stat`") ☺

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 170

Esame del Sistema Operativo Linux Implementazione del File System in UNIX - 1

- Struttura di partizione secondo UNIX v7
- Il super-blocco (1) indica, tra l'altro il # di *i-node* e di blocchi nel FS, e fornisce il puntatore alla lista dei blocchi liberi (2)
- *i-node* (3) numerati 1..N, ciascuno di ampiezza ≥ 64 B
- *Directory* contiene un insieme variabile e non ordinato di unità informative (campi, *entry*) ampie 16 B
 - 14 B (= caratteri ASCII) per nome di *file*
 - 2 B per numero di *i-node*



Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 171

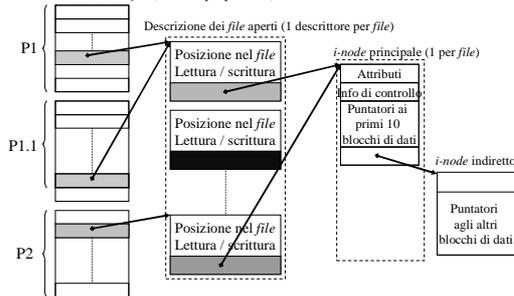
Esame del Sistema Operativo Linux Implementazione del File System in UNIX - 2

- Per ogni tipo di *file* aperto, il nucleo mantiene una tabella dei corrispondenti *i-node*
 - Un'altra tabella di nucleo, per processo, contiene i descrittori corrispondenti
 - Ad essi occorre associare la posizione corrente di lettura/scrittura nel *file*
- Quale corrispondenza tra le due tabelle?
 - Ogni processo che apre un *file* deve avere il suo proprio indicatore di posizione (più posizioni per *file*)
 - L'indicatore non può essere associato all'*i-node* (unico per *file*)
 - Sequenze ordinate di processi figli di uno stesso padre devono poter scrivere su uno stesso *file* consecutivamente
 - Lo stesso indicatore per processi di una stessa famiglia
- Una tabella intermedia, che descrive ciascun *file* aperto, associa descrittore di *file* all'*i-node* corrispondente

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 172

Esame del Sistema Operativo Linux Implementazione del File System in UNIX - 3

Tabella dei descrittori di *file* (1 tabella per processo)



Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 173

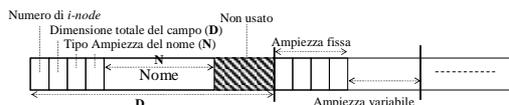
Esame del Sistema Operativo Linux Implementazione del File System in UNIX - 4

- L'*i-node* principale contiene l'indirizzo dei primi 10 blocchi dati del *file*
- Per *file* di maggiore dimensione, un campo dell'*i-node* principale punta al blocco (*single-indirect*) di un *i-node* secondario che contiene i puntatori agli altri blocchi dati
 - Esempio: blocco da 1 KB; indirizzi di blocco su 4 B \rightarrow max dimensione di *file* = $(10 + 1 \text{ KB} / 4 \text{ B}) * 1 \text{ KB} = 266 \text{ KB}$
- Per *file* più grandi, l'*i-node* secondario (*double-indirect*) contiene puntatori a nodi *single-indirect*
 - Esempio: stesse ipotesi \rightarrow max = $(10 + [1 \text{ KB} / 4 \text{ B}]^2) * 1 \text{ KB} = 10 \text{ KB} + 64 \text{ MB}$

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 174

Esame del Sistema Operativo Linux Implementazione del File System in UNIX - 5

- La versione BSD introduce alcune migliorie significative
 - Estensione del nome di *file* fino a 255 caratteri
 - *Directory* di dimensione multipla di blocco
 - Facilita e velocizza la scrittura su disco
 - La dimensione così fissata induce frammentazione interna



Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 175

Esame del Sistema Operativo Linux Implementazione del File System in UNIX - 6

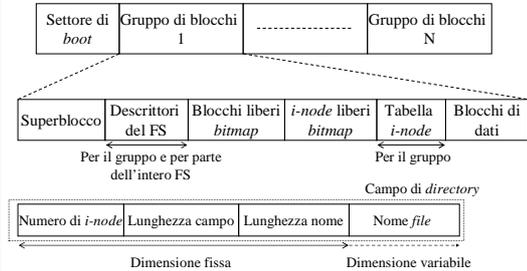
- Migliorie BSD (segue)
 - Uso di *cache* dei nomi per evitare la costosa ricerca lineare su *directory*
 - Suddivisione del disco in "gruppi di cilindri" (~ sotto-partizioni virtuali) per avvicinare gli *i-node* ai relativi blocchi dati
 - Oggi di scarso interesse perché i dischi moderni tendono a nascondere la loro geometria interna al S/O
 - Uso di 2 ampiezze di blocco
 - Blocchi grandi per *file* molto grandi
 - Blocchi piccoli per *file* piccoli e medi (la norma)
 - Maggior complessità di gestione

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 176

Esame del Sistema Operativo Linux Implementazione del File System Linux - 1

- Inizialmente basato sul FS di MINIX, però subito abbandonato per le eccessive limitazioni
 - Nomi fino a 14 caratteri
 - Indirizzi di blocco limitati a 16 bit per blocchi ampi 1 KB
→ ampiezza di *file* ≤ 64 MB
- **ext2** diviene presto la versione di riferimento
 - Basata sulle scelte BSD, ma con diversa struttura fisica
 - La maggiore innovazione è stata la suddivisione della partizione in “gruppi di blocchi”
 - Stessa località tra *i-node* e blocchi di dati
 - Maggior robustezza replicando su ciascun gruppo le informazioni di controllo del superblocco

Esame del Sistema Operativo Linux Implementazione del File System Linux - 2



Esame del Sistema Operativo Linux Implementazione del File System Linux - 3

- Dimensione di *i-node* estesa a 128 B
 - Indirizzi di blocco ampi 4 B → partizione ≥ 4 GB (blocchi di dimensione 1,2, 4 KB, scelta a configurazione del FS)
 - 12 indirizzi diretti di blocco dati
 - 3 indirizzi indiretti (*single, double, ...*)
 - Più informazioni di controllo
 - Una parte riservata per uso futuro
- Ogni aggiunta a *file* regolare e *directory* viene resa quanto più locale possibile entro lo stesso gruppo
 - Preallocazione di fino ad 8 blocchi dati adiacenti per localizzare scritture su *file*
 - Località tra *file* correlati tramite gruppi
 - Località entro *file* tramite preallocazione

Esame del Sistema Operativo Linux Implementazione del File System Linux - 4

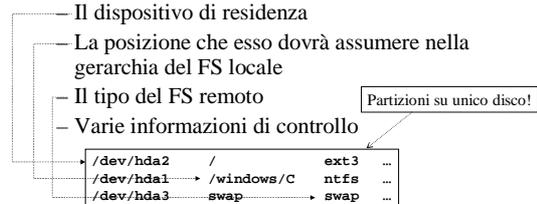
- Introduzione di una *directory /proc virtuale* (non esistente su disco) contenente una *directory* √ processo presente nel sistema
 - Il nome della *directory* è il PID del processo
 - Il contenuto della *directory* è un insieme di *file* che descrivono il processo ed il suo ambiente
 - L'informazione originale resta nel nucleo, da dove essa viene estratta alla lettura del *file* virtuale corrispondente
- L'accesso di utente al FS viene filtrato da un FS virtuale che consente la coesistenza di più FS di tipo diverso (p.es.: FAT-32 ed **ext2**)
 - Modalità sostanzialmente analoga ad NFS (vedi seguito ...)

Esame del Sistema Operativo Linux File System di rete (NFS) - 1

- L'idea è consentire ad un insieme arbitrario di utenti remoti di condividere un FS comune
- Il FS comune viene ospitato su un *server* al quale più clienti possono fare accesso
 - Sia clienti remoti (posti su rete locale o geografica ed eterogenea) che locali (posti sullo stesso nodo)
 - Il *server* esporta un FS designandone come radice una *directory* del proprio FS
 - La lista delle “radici” esportate viene posta nel *file* di configurazione **/etc/exports**
 - Il cliente importa il FS esportato mediante **mount**, assegnandogli una posizione (un cammino) nel suo FS

Esame del Sistema Operativo Linux File System di rete (NFS) - 2

- Il file **/etc/fstab** di ciascun cliente fornisce la lista dei FS da importare mediante **mount**
- Per ciascun FS remoto si indicano
 - Il dispositivo di residenza
 - La posizione che esso dovrà assumere nella gerarchia del FS locale
 - Il tipo del FS remoto
 - Varie informazioni di controllo



Esame del Sistema Operativo Linux File System di rete (NFS) - 3

- NFS definisce 2 protocolli che regolano il dialogo tra il *server* ed i suoi clienti
 - Un protocollo regola l'importazione di FS (**mount**)
 - Il cliente invia al *server* il cammino della radice del FS esportato
 - Se la richiesta ha successo, il *server* invia al cliente un descrittore unicamente associato al FS
 - Tipo di FS, disco di residenza, informazioni di controllo, numero di *i-node* della *directory* radice
 - Ogni successivo accesso ad un *file* del FS importato userà tale descrittore
 - 2 modalità di importazione
 - Esplicita, all'inizializzazione (eseguita da **/etc/rc**)
 - Automatica, al riferimento a *file* residenti in FS importato

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 183

Esame del Sistema Operativo Linux File System di rete (NFS) - 4

- L'altro protocollo NFS regola l'accesso a *file* e *directory*
 - La sua realizzazione non richiede al *server* di mantenere informazioni di stato (*stateless*)
 - Il cliente emette richieste in forma di messaggio contenente il descrittore del *file* remoto ed i parametri dell'operazione richiesta su di esso
 - Questa scelta semplifica il lavoro del *server*, ma si presta ad inconsistenze (lo stato di un *file* può cambiare tra due accessi remoti successivi)
- Il controllo di accesso a *file* usa diritti di privilegio *owner*, *group*, *others* facilmente falsificabili se non autenticati
- Nessun supporto per *lock*

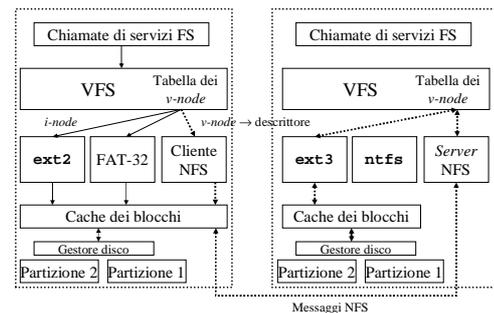
Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 184

Esame del Sistema Operativo Linux File System di rete (NFS) - 5

- Presso ogni cliente, un livello di FS virtuale (VFS) filtra ogni chiamata a servizi dei FS localmente supportati
- VFS mantiene un *v-node* per ogni *file* aperto, al quale associa un *i-node*, se il *file* è locale, od un *r-node*, se remoto
 - All'*r-node* viene associato il descrittore di *file* remoto fornito dal *server* che ne esporta il FS
 - Per ridurre il traffico di rete, trasferimenti in unità da 8 KB tra cliente e *server* ed uso di *cache* dati e riferimenti presso il cliente
 - Modalità *read-ahead* in lettura
 - Invio differito al riempimento dell'unità di trasferimento per scrittura
 - Scadenza associata alla validità dell'informazione in *cache*

Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 185

Esame del Sistema Operativo Linux File System di rete (NFS) - 6



Esame del Sistema Operativo Linux Architettura degli elaboratori 2 - T. Vardanega Pagina 186