HOWTO: Multi Disk System Tuning

Stein Gjoen, `sgjoen@nyx.net` v 0.22i, 27 maggio 1999

Questo documento descrive come utilizzare al meglio dischi e partizioni multiple per un sistema Linux. Sebbene parte di questo testo è specifico di Linux, l'approccio generale evidenziato qui può essere applicato a molti altri sistemi operativi multi tasking. Traduzione di Gaetano Paolone (*gaetano at poboxes.com*). Data traduzione: 26 ottobre 1999

Indice

1 Introduzione

Per commemorare la release 2.2. del kernel Linux a questa release nuova di zecca ho dato il nome in codice **Daniella**.

Nuovi nomi in codice appariranno in relazione alle linee guida standard dell'industria per enfatizzare lo stato dell'arte di questo documento.

Questo documento è stato scritto per due ragioni, principalmente perché possedevo 3 vecchi dischi SCSI su cui installare Linux e mi stavo domandando come utilizzare al meglio l'innata possibilità di parallelismo in un sistema SCSI. In secondo luogo ho sentito che c'è una ricompensa per chi scrive documenti...

Questo documento è concepito per essere letto insieme al Linux Filesystem Structure Standard (FSSTND). Esso non lo rimpiazza in alcun modo ma cerca di suggerire dove mettere fisicamente le directory descritte in dettaglio nel FSSTND, in termini di dischi, partizioni, tipi, RAID, file system (fs), dimensioni fisiche e altri parametri che dovrebbero essere considerati e regolati in un sistema Linux, spaziando dai singoli sistemi casalinghi a grossi server su Internet.

Il successore del FSSTND è chiamato Filesystem Hierarchy Standard (FHS) e non riguarda più solo Linux. È stata rilasciata la versione 2.0 del FHS ma ci sono ancora un po' di questioni aperte e ci vorrà molto tempo prima che questo nuovo standard abbia un impatto sulle distribuzioni attuali.

È anche una buona idea leggere approfonditamente le guide di Installazione di Linux e se state usando un PC, come immagino la maggior parte di voi faccia, potete trovare molte informazioni degne di nota ed utili nella FAQ del newsgroup comp.sys.ibm.pc.hardware specialmente riguardo i supporti di archiviazione.

Questa è anche un'esperienza istruttiva per me e spero che io possa cominciare a far girare la palla con questo HOWTO e che esso possa evolvere in un HOWTO più grande, più dettagliato, e speriamo anche più corretto.

Prima di tutto abbiamo bisogno di un po' di linguaggio legale. Recenti sviluppi dimostrano che questo è abbastanza importante.

1.1 Copyright

1.2 Liberatoria

Usate le informazioni contenute in questo documento a vostro rischio. Ripudio qualsiasi potenziale responsabilità per i contenuti di questo documento. L'utilizzo dei concetti, degli esempi e/o degli altri contenuti di questo documento è completamente a vostro rischio.

Tutti i copyright sono dei rispettivi proprietari, salvo ove diversamente specificato. L'utilizzo di un termine, in questo documento, non dovrebbe essere considerato come un attacco alla validità di qualsiasi marchio di fabbrica o di servizio.

Il nominare prodotti particolari o marche, non dovrebbe essere considerato come approvazione.

Siete fermamente invitati a fare un backup del vostro sistema prima della installazione principale e a continuare a farne ad intervalli regolari.

1.3 Notizie

Questa versione è caratterizzata dall'avere una grossa ristrutturazione e molte più aggiunte di quante ne possa elencare qui, specialmente l'aggiunta del supporto a file system.

Questo HOWTO ora utilizza l'indicizzazione ed è basato sugli SGMLtools versione 1.0.5 e la vecchia versione non conferirà a questo documento una formattazione adeguata.

Inoltre sono disponibili anche una serie di nuove traduzioni.

Per quanto concerne lo sviluppo, la gente si sta concentrando sul Linux 2.2 e fino a quando non sarà rilasciato, non ci saranno molte notizie nuove sulla tecnologia disco per Linux.

La Debian 2.1 è pronta per essere rilasciata ma visto che uso Debian per i miei sistemi di test, farò più aggiornamenti quando la aggiornerò.

L'ultimo numero di versione può essere ottenuto dall'annotazione del mio progetto se fate un

finger `http://www.cs.indiana.edu/finger/nox.nyx.net/sgjoen` al mio account su Nyx.

Inoltre, l'ultima versione sarà disponibile sul mio spazio web su nyx in numerosi formati:

- *HTML* `http://www.nyx.net/~sgjoen/disk.html` .
- *semplice testo ASCII* `http://www.nyx.net/~sgjoen/disk.txt`
- *postscript compresso* `http://www.nyx.net/~sgjoen/disk.ps.gz` .
- *sorgente SGML* `http://www.nyx.net/~sgjoen/disk.sgml` .

Un mirror Europeo del

Multi Disk HOWTO `http://home.sol.no/~gjoen/stein/disk.html`

è stato appena creato.

1.4 Crediti

In questo documento ho il piacere di riconoscere molte persone che hanno contribuito in un modo o nell'altro:

```
ronnej (at ) ucs.orst.edu
cm (at) kukuruz.ping.at
armbru (at) pond.sub.org
R.P.Blake (at) open.ac.uk
neuffer (at) goofy.zdv.Uni-Mainz.de
sjmudd (at) redestb.es
nat (at) nataa.fr.eu.org
sundbyk (at) oslo.geco-prakla.slb.com
ggjoeen (at) online.no
mike (at) i-Connect.Net
roth (at) uiuc.edu
phall (at) ilap.com
szaka (at) mirror.cc.u-szeged.hu
CMckeon (at) swcp.com
kris (at) koentopp.de
edick (at) idcomm.com
pot (at) fly.cnuce.cnr.it
earl (at) sbox.tu-graz.ac.at
ebacon (at) oanet.com
```

```
vax (at) linkdead.paranoia.com
tschenk (at) theoffice.net
pjfarley (at) dorsai.org
jean (at) stat.ubc.ca
johnf (at) whitsunday.net.au
clasen (at) unidui.uni-duisburg.de
eeslgw (at) ee.surrey.asc.uk
adam (at) onshore.com
anikolae (at) wega-fddi2.rz.uni-ulm.de
cjaeger (at) dwave.net
eperezte (at) c2i.net
```

Ringraziamenti speciali vanno a `nakano (at) apm.seikei.ac.jp` per aver fatto la

taduzione in Giapponese `http://jf.linux.or.jp/JF/JF-ftp/other-formats/Disk-HOWTO/html/Disk-HOWTO.html` , per aver dato contributi generali e anche per aver fornito un esempio di un computer in un ambiente accademico, che è incluso alla fine di questo documento.

Ci sono ora molte traduzioni disponibili e ringraziamenti speciali ora vanno ai traduttori per il lavoro e la spinta che hanno dato:

- *Traduzione in Tedesco* `http://` di `chewie (at) nuernberg.netsurf.de`
- *Traduzione in Svedese* `http://www.swe-doc.linux.nu` di `jonah (at) swipnet.se`
- *Traduzione in Francese* `http://www.lri.fr/~loisel/howto/` di `Patrick.Loiseleur (at) lri.fr`

Inoltre la DPT è riconosciuta per avermi inviato la documentazione sui propri controller così come il permesso di citare il materiale. Queste citazioni sono state approvate prima di apparire qui e saranno etichettate intelligentemente. Nessuna citazione ancora ma stanno per venire.

Non è ancora tutto, quindi per cortesia leggete questo documento, date un contributo e partecipate all'elite. Se ho dimenticato qualcuno fatemelo sapere.

Nuova in questa versione è un'appendice con un po' di tabelle che potete riempire per il vostro sistema al fine di semplificare la procedura di progettazione.

Qualsiasi commento o suggerimento può essere inviato al mio indirizzo e-mail sul nyx: *sgjoen@nyx.net* .

Quindi diamo la caccia a `swap` ed a `/tmp` dal momento che stanno correndo nel disco rigido...

2 Struttura

Dal momento che questo documento si suppone che sia più che sufficiente per imparare come se fosse un documento di riferimento tecnico, ho riarrangiato la struttura in quest'ottica. Per colui che progetta un sistema è più utile avere le informazioni presentate in termini di scopi di questo esercizio piuttosto che dal punto di vista della strutturazione logica dei dispositivi stessi. Tuttavia questo documento non sarebbe completo senza tutta quella struttura a strati di cui la struttura del computer è piena, quindi la includerò come introduzione a come funziona.

È passato molto tempo da quando il *mini* nel mini-HOWTO poteva essere considerato corretto ma sono convinto che questo documento sia lungo quanto debba esserlo, utile per prendere le giuste decisioni e niente di più.

2.1 Struttura Logica

Questa è basata su come ogni strato accede ad un altro, tradizionalmente con l'applicazione nello strato più alto e lo strato fisico in quello più basso. È abbastanza utile mostrare le interrelazioni tra gli strati usati per controllare i dischi.

```
-----------------------------------------------------------
|__     Struttura dei File      ( /usr /tmp ecc)        __|
|__     File system             (ext2fs, vfat ecc)      __|
|__     Gestione del volume     (AFS)                   __|
|__     RAID, concatenazione    (md)                    __|
|__     Driver dei dispositivi  (SCSI, IDE ecc)         __|
|__     Controller              (chip, scheda)          __|
|__     Connessione             (cavo, rete)            __|
|__     Dischi                  (magnetici, ottici ecc) __|
-----------------------------------------------------------
```

Nel diagramma qui sopra sia la gestione del volume, che il RAID, che la concatenazione, sono strati opzionali. I 3 strati più bassi riguardano l'hardware. Tutte le parti sono ampiamente discusse in seguito in questo documento.

2.2 Struttura del Documento

La maggior parte degli utenti comincia con una determinata configurazione di hardware e qualcuno progetta cosa vorrebbe ottenere e quanto grande dovrebbe essere il sistema. Questo è il punto di vista che utilizzerò in questo documento presentando il materiale, iniziando con l'hardware, continuando con le costrizioni della progettazione, prima di descrivere la strategia di progettazione che ho scoperto funzionare bene. Ho usato questo sia per il mio computer di casa, che al lavoro per un server multifunzionale e ho scoperto che funziona ragionevolmente bene. Oltre a ciò, il collaboratore giapponese di questo documento ha applicato la stessa strategia su un server di un ambiente accademico con un paragonabile successo.

Per finire, alla fine ho elencato con dettaglio alcune tavole di configurazione da usare per le vostre progettazioni. Se avete qualsiasi commento al riguardo o annotazioni del vostro lavoro di progettazione, mi piacerebbe sentirvi così che questo documento possa essere aggiornato.

2.3 Pianificazione della lettura

Sebbene non sia l'HOWTO più grande, è in ogni caso abbastanza grande e mi è stato chiesto di fare una pianificazione della lettura per poter sfoltire il volume.

Esperto

(tipo l'elite). Se siete pratici di Linux e delle tecnologie dei dischi, troverete molto di quello di cui avete bisogno nelle appendici. In aggiunta siete avvisati di leggere le FAQ e il capitolo 19 (Pezzettini e Ritagli).

Pratici

(tipo Competenti). Se siete pratici di computer in generale potete direttamente andare ai capitoli sulle 6 (tecnologie) e continuare da lì.

Principianti

(per lo più disarmati). Dovete solamente leggere tutto. Mi dispiace. In aggiunta vi consiglio di leggere anche tutti gli altri HOWTO riguardanti i dischi.

3 Tecnologie dei dischi

Una discussione ben più completa sulle tecnologie dei dischi per PC IBM può essere trovata presso l'home page delle

The Enhanced IDE/Fast-ATA FAQ `http://thef-nym.sci.kun.nl/~pieterh/storage.html`

che è anche periodicamente postata sulle News di Usenet. Qui io presenterò cosa è necessario per farsi un'idea della tecnologia e per prepararsi per il setup.

3.1 Dischi

Questo è il dispositivo fisico dove i vostri dati vivono e sebbene il sistema operativo rende abbastanza somiglianti tipi differenti, essi possono in fondo essere molto differenti. Conoscere il funzionamento può essere molto utile nella vostra fase di progettazione. I floppy non fanno parte di questo documento, sebbene nel caso ci fosse una grande richiesta, potrei essere convinto ad aggiungere una piccola parte qui.

3.2 Geometria

Fisicamente, i dischi consistono in uno o più piatti contenenti dati che sono letti all'interno e all'esterno, mediante l'utilizzo di sensori montati su testine mobili fissate tra di loro. I trasferimenti di dati dunque, avvengono attraverso tutte le superfici simultaneamente, il che definisce un cilindro di tracce. Il disco è inoltre suddiviso in settori contenenti un determinato numero di campi per i dati.

I dischi sono quindi spesso classificati in relazione alla geometria: il numero di Cilindri, Testine e Settori (CHS).

Per varie ragioni c'è un bel numero di differenze tra

- la CHS fisica del disco stesso
- la CHS logica che il disco riporta al BIOS o al S.O.
- la CHS logica usata dal S.O.

Praticamente è un gran disordine ed è causa di molta confusione. Per maggiori informazioni siete caldamente invitati a leggere il *Large Disk mini-HOWTO*

3.3 Supporti

La tecnologia dei supporti determina importanti parametri come le velocità di lettura/scrittura, tempo di accesso, capacità di memorizzazione come anche se permette la lettura/scrittura o solamente la lettura.

3.3.1 Dischi Magnetici

Questo è il tipico mezzo di memorizzazione di lettura e scrittura e come qualsiasi altra cosa nel mondo dei computer è disponibile in molti 'gusti' e con proprietà differenti. Generalmente questa è la tecnologia più veloce e offre capacità di lettura/scrittura. Il disco gira con una velocità angolare costante (CAV) con una densità del settore fisico variabile per un utilizzo più efficiente dell'area magnetica del supporto. In altre parole, il numero di bit per unità di lunghezza è mantenuto costante aumentando il numero di settori logici per le tracce esterne.

I valori tipici delle velocità di rotazione sono 4500 e 5400 RPM, sebbene sia usata anche la velocità a 7200. Molto recentemente è entrata nel mercato anche la velocità a 10000 RPM. I tempi di accesso sono intorno a 10 ms, le velocità di trasferimento sono abbastanza variabili da un tipo ad un altro ma generalmente sono nell'ordine di 4-40 MB/s. Con dischi ad alte prestazioni dovreste ricordare che le prestazioni stesse necessitano di più potenza elettrica che è poi dissipata sotto forma di calore, controllate 19.3 (Alimentazione e Riscaldamento).

Notate che esistono diversi tipi di trasferimento e che questi sono riportati con unità differenti. Prima di tutto c'è il trasferimento dal 'piatto' alla cache dell'unità disco che è generalmente riportato in Mbit/s. Un valore tipico è circa 50-250 Mbit/s. Il secondo passaggio è dalla cache del drive all'adattatore ed è generalmente indicato in MB/s, i valori riportati sono intorno ai 3-40 MB/s. Notate comunque che si assume che i dati siano già contenuti nella cache e da qui l'effettiva percentuale di trasferimento diminuirà drammaticamente a causa della massima velocità di lettura dal drive.

3.3.2 Dischi Ottici

I dischi ottici in lettura/scrittura esistono ma sono lenti e non così diffusi. Furono usati nella macchina NeXT ma la bassa velocità è stata l'origine di molte lamentele. La bassa velocità è dovuta principalmente alla natura termica del cambio di fase che rappresenta l'archiviazione dei dati. Anche quando si sono usati laser potenti per indurre cambiamenti di fase, gli effetti sono stati ancora più lenti dell'effetto magnetico usato nei dischi magnetici.

Oggi molte persone usano i dischi CD-ROM che, come il nome fa capire, sono di sola lettura. L'archiviazione è di circa 650 MB, le velocità di trasferimento sono variabili e dipendono dal disco ma possono superare 1.5 MB/s. I dati sono archiviati in una singola traccia a spirale quindi non è utile parlare di geometria per essi. La densità dei dati è costante, quindi il lettore usa una velocità lineare costante (CLV). Anche l'accesso è più lento, circa 100 ms, parzialmente dovuto alla traccia a spirale. Recentemente dischi ad alta velocità, usano un misto tra CLV e CAV al fine di aumentare le prestazioni. Questo riduce inoltre il tempo di accesso causato dalla necessità di raggiungere la corretta velocità di rotazione per la lettura.

Un nuovo tipo (DVD) è all'orizzonte, in grado di offrire fino a circa 18 GB su un disco singolo.

3.3.3 Dischi a Stato Solido

Questa è un'aggiunta relativamente recente alla tecnologia disponibile ed è stata resa popolare specialmente in computer portatili come anche in sistemi fissi. Non contenendo parti rimovibili, sono molto veloci sia in termini di velocità di accesso che di velocità di trasferimento. Il tipo più popolare è la flash RAM, ma sono usati anche altri tipi di RAM. Alcuni anni fa molti riponevano molte speranze per le memorie a bolla magnetica ma si sono dimostrate essere relativamente costose e non sono così comuni.

Generalmente l'uso di dischi RAM non è considerato una buona idea dal momento che è più efficace aggiungere più RAM alla scheda madre e permettere al sistema operativo di dividere il quantitativo di memoria in memorie tampone, cache, aree di programmi e di dati. Solo in casi molto speciali, come in sistemi in tempo reale con piccoli margini di tempo, i dischi RAM possono essere una soluzione ragionevole.

La Flash RAM è disponibile in molte decine di megabyte di capacità e si potrebbe essere tentati di usarla per un'archiviazione veloce e temporanea in un computer. C'è comunque un enorme ostacolo con questo metodo: la flash RAM ha un tempo di vita limitato in termini del numero di volte in cui potete riscrivere i dati, quindi mettere `swap`, `/tmp` o `/var/tmp` su questi dispositivi accorcerà sicuramente e drammaticamente la loro vita. Invece, usare le flash RAM per directory che vengono lette spesso ma raramente scritte, porterebbe ad un grosso guadagno in termini di prestazioni.

Per ottenere l'optimum della vita media dalle flash RAM, dovrete usare driver speciali che usano la RAM ugualmente e minimizzano il numero di blocchi cancellati.

Questo esempio illustra i vantaggi di dividere la vostra struttura delle directory su più dispositivi.

I dischi a Stato Solido non hanno un reale indirizzamento cilindri/testine/settori ma per ragioni di compatibilità questi sono simulati dal driver per avere un'interfaccia uniforme al sistema operativo.

3.4 Interfacce

C'è una pletora di interfacce da cui scegliere con diverse caratteristiche di prezzo e prestazioni. Parecchie schede madri oggi includono interfaccia IDE o interfacce migliori, che sono parte integrante dei chipset. Molte altre schede includono anche un chip per l'interfaccia SCSI fatto da NCR e connesso direttamente al bus PCI. Controllate che cosa avete e quale supporto del BIOS avete per quello.

3.4.1 MFM e RLL

Un tempo questa era la tecnologia consolidata, il tempo in cui 20 MB erano solenni, che comparati con le dimensioni di oggi ti fanno pensare che i dinosauri si aggiravano per la Terra con questi dischi. Come i dinosauri questi sono datati e sono lenti e non affidabili se confrontati a quello che abbiamo oggi. Linux li gestisce ma siete stati bene avvisati di pensare due volte riguardo cosa volete metterci. Si potrebbe discutere che una partizione di emergenza potrebbe essere adattata con un'opportuna versione datata del DOS.

3.4.2 ESDI

Effettivamente, ESDI era un adattamento dell'interfaccia SMD, largamente diffusa ed utilizzata su computer grandi, ai cavi usati con l'interfaccia ST506, cosa molto più conveniente da confezionare rispetto alla copia di connettori a 60-pin e 26-pin usati con SMD. L'ST506 era un'interfaccia silenziosa che faceva completo affidamento sul fatto che il controller e l'host facessero tutto: dal valutare la disposizione di testine/cilindri/settori al tenere traccia della disposizione delle testine, ecc. ST506 necessitava che il controller estraesse il clock dai dati recuperati e che controllasse la disposizione delle caratteristiche dettagliate sulle tracce del supporto, bit per bit. Ebbe una vita di circa 10 anni, se includete l'uso degli schemi di modulazione di MFM, RLL, e ERLL/ARLL. ESDI, d'altra parte, ebbe l'intelligenza, spesso utilizzando tre o quattro microprocessori separati su un singolo drive e comandi di alto livello per formattare una traccia, trasferire dati, effettuare ricerche e così via. Il recupero del clock dal flusso dei dati era affidato al drive, che pilotava la linea di clock e presentava i suoi dati in NRZ, sebbene la correzione dell'errore era ancora compito del controller. ESDI permetteva l'uso della densità di registrazione a bit variabile, oppure, per quella ragione, di qualsiasi altra tecnica di modulazione, dal momento che era localmente generata e risolta presso il disco. Sebbene molte delle tecniche usate in ESDI, furono in seguito incorporate nell'IDE, fu la crescente popolarità dello SCSI che portò alla fine dell'ESDI nei computer. ESDI ebbe una vita di circa 10 anni, sebbene principalmente nei server o in grossi sistemi piuttosto che nei PC.

3.4.3 IDE e ATA

Il progresso fece sì che l'elettronica dei dischi migrasse dallo slot ISA al drive stesso e nacque l'Integrated Drive Electronics. Era semplice, economica e ragionevolmente veloce così coloro che progettavano il BIOS crearono quel tipo di ostacolo di cui l'industria di computer è così piena. Una combinazione tra una limitazione IDE di 16 testine insieme alla limitazione del BIOS a 1024 cilindri ci donò l'infame limite di 504 MB. Seguendo le tradizioni dell'industria del computer, l'ostacolo fu rattoppato con un programma inaffidabile ed ottenemmo ogni tipo di schema di traduzione e instabili rattoppi del BIOS. Questo vuol dire che dovete leggere la documentazione relativa all'installazione molto attentamente e controllare che BIOS avete e che data ha, dal momento che il BIOS deve dire a Linux che dimensioni di dischi avete. Fortunatamente con Linux potete dire direttamente al kernel le dimensioni del vostro disco con i suoi parametri, controllate la documentazione su LILO e Loadlin, esaurientemente. Notate anche che IDE è equivalente ad ATA, AT Attachment. IDE usa un Input/Output Programmato (PIO) CPU-intensivo per trasferire i dati verso e dai dischi e non è in grado di gestire la tecnologia di Accesso Diretto alla Memoria (DMA) che è più efficiente. La velocità di trasferimento più elevata è di 8.3 MB/s.

3.4.4 EIDE, Fast-ATA e ATA-2

Questi 3 termini sono abbastanza equivalenti, fast-ATA è ATA-2 ma EIDE include come aggiunta ATAPI. ATA-2 è quello che si usa di più in questi giorni dato che è più veloce e con DMA. La velocità trasferimento è aumentata a 16.6 MB/s.

3.4.5 Ultra-ATA

Una nuova e più veloce modalità DMA che è circa il doppio del PIO-Mode 4 dell'EIDE (33 MB/s). I dischi con o senza Ultra-ATA possono essere messi sullo stesso cavo senza degrado della velocità per gli adattatori più veloci. L'interfaccia Ultra-ATA è elettricamente identica alla normale interfaccia Fast-ATA, inclusa la massima lunghezza del cavo.

3.4.6 ATAPI

L'ATA Packet Interface venne progettata per gestire i CD-ROM utilizzando la porta IDE e come l'IDE è economica e semplice.

3.4.7 SCSI

Lo Small Computer System Interface è un'interfaccia dai molti scopi che può essere utilizzata per collegare di tutto, dai dischi, a schiere di dischi, stampanti, scanner e molto altro ancora. Il nome ha una designazione erronea dal momento che è stato usato tradizionalmente nelle fasce alte del mercato come anche nelle workstation dal momento che è adatto a sistemi multi tasking.

L'interfaccia standard è larga 8 bit e può indirizzare 8 dispositivi. C'è una versione larga 16 bit che è veloce il doppio con lo stesso clock e può indirizzare 16 dispositivi. L'adattatore conta sempre come dispositivo ed è generalmente il numero 7. È anche possibile avere bus con un'ampiezza di 32 bit ma questo generalmente necessita di doppi cavi per gestire tutte le linee.

Il vecchio standard era di 5 MB/s e il recente fast-SCSI l'ha aumentato a 10 MB/s. Recentemente ultra-SCSI, noto anche come Fast-20, è arrivato a velocità di trasferimento di 20MB/s per un bus ampio 8 bit. Il nuovo sistema di segnalazione a basso voltaggio differenziale (LVD) permette queste velocità elevate ed anche cablaggi molto più lunghi di prima.

Più recentemente è stato anche proposto uno standard ancora più veloce: SCSI 160/m che è capace di un mostruoso 160 MB/s su un bus ampio 16 bit. Il supporto è ancora scarso ma è sostenuto per un po' di dischi a 10000 RPM che possono trasferire 40 MB/s. Mettere 6 di questi dischi su un RAID manterrebbe questo bus saturo ed inoltre saturerebbe la maggior parte dei bus PCI. Ovviamente questo vale solo per i server di altissimo livello di oggi.

Le prestazioni più elevate si traducono in un costo generalmente più alto dell'(E)IDE. L'importanza di terminazioni corrette e di una buona qualità dei cavi non può essere sovraenfatizzata. I dischi SCSI inoltre spesso tendono ad essere di qualità più elevata rispetto ai dischi IDE: spesso è solo una questione di attaccare e staccare il dispositivo; molte persone fanno questo senza spegnere il sistema. Questa caratteristica è molto utile quando avete sistemi multipli e potete solamente spostare i dispositivi da un sistema all'altro qualora uno di questi fallisse per un qualsiasi motivo.

C'è un gran numero di documenti che dovreste leggere se utilizzate lo SCSI, lo SCSI HOWTO come del resto le SCSI FAQ postate sulle News di Usenet.

Lo SCSI ha inoltre il vantaggio che vi potete connettere facilmente a lettori di nastri per fare il backup dei vostri dati, come anche ad alcune stampanti e scanner. È anche possibile usarlo come una rete molto veloce tra computer dal momento che si possono condividere dispositivi SCSI sullo stesso bus. I lavori continuano ma a causa di problemi nell'assicurare la condivisione della cache tra i diversi computer connessi, non sarà un lavoro da nulla.

I numeri SCSI sono anche usati per gestire le priorità. Se diversi dischi richiedono il servizio, viene data priorità al drive con il numero più basso.

3.5 Cablaggio

La mia intenzione è quella di non fare troppi commenti sull'hardware ma sento che dovrei fare una piccola nota sul cablaggio. Questo sembrerebbe essere una componente dell'insieme marcatamente di bassa rilevanza tecnologica, ma purtroppo è la causa di molti problemi di frustrazione. Viste le alte velocità di oggi, si dovrebbe pensare piuttosto ad un dispositivo RF con le sue intrinseche necessità di correttezza dell'impedenza. Se non prendete le vostre precauzioni, otterrete affidabilità altamente ridotta o un fallimento totale. Qualche adattatore SCSI è molto più sensibile a ciò rispetto ad altri.

Cavi schermati sono ovviamente migliori di quelli non schermati ma il prezzo è molto più alto. Con un po' di attenzione potrete ottenere buone prestazioni da un economico cavo non schermato.

- Per Fast-ATA e Ultra-ATA, la massima lunghezza del cavo è specificata come 45cm (18). Le linee di dati di tutti e due i canali IDE sono connessi su molte schede, sebbene, in questo modo possono rivelarsi come **un unico** cavo. In ogni caso i cavi EIDE dovrebbero essere più corti possibile. Se ci sono crash misteriosi o cambiamenti improvvisi dei dati, è opportuno verificare i vostri cavi. Provate una modalità PIO più bassa oppure disconnettete il secondo canale e controllate se il problema persiste.
- Usate cavi più corti possibile, ma non dimenticate la minima separazione di 30 cm per ultra SCSI e di 60 cm per SCSI differenziale.
- Evitate lunghi spezzoni tra il cavo ed il drive, collegate il connettore del cavo direttamente al disco senza una prolunga.
- Limiti nel cablaggio SCSI:

```
Velocita Bus (MHz)       |    Massima Lunghezza (m)
--------------------------------------------------
 5                       |       6
10  (fast)               |       3
20  (fast-20 / ultra)    |       3 (massimo 4 dispositivi),
                         |     1,5 (massimo 8 dispositivi)
xx  (differenziale)      |      25 (massimo 16 dispositivi)
--------------------------------------------------
```

- Usate una corretta terminazione per i dispositivi SCSI e le posizioni corrette ad entrambe le estremità della catena SCSI. Ricordate che lo stesso adattatore può avere il terminatore sulla scheda stessa.
- Non mischiate cavi schermati e cavi non schermati, non avvolgete i cavi attorno al metallo, provate ad evitare vicinanza a parti metalliche lungo il decorso del cablaggio. Qualsiasi discontinuità può causare imprecisioni nell'impedenza che generano riflessione dei segnali che va ad aumentare il rumore nel cavo.

> Questo problema diventa ancor più grave nel caso dei controller multi-canale. Recentemente qualcuno ha suggerito di usare la plastica con le bolle attorno ai cavi per evitare lo stretto contatto con il metallo, un problema serio nei cabinet affollati.

Più informazioni sul cablaggio SCSI e sui terminatori possono essere trovate presso *altre* `http://resource.simplenet.com/files/68_50_n.htm`

pagine sparse per il web.

3.6 Adattatori

Questo è l'altro capo dell'interfaccia dal drive, la parte che è connessa ad un bus del computer. La velocità del bus del computer e quella dei dischi dovrebbe essere vagamente similare, altrimenti avrete un collo di bottiglia nel vostro sistema. Connettere un insieme di dischi RAID 0 ad una scheda ISA è inutile. In questo periodo la maggior parte dei computer è dotata di bus PCI a 32 bit capaci di trasferimenti a 132 MB/s che non dovrebbe rappresentare un collo di bottiglia per la maggior parte della gente per il prossimo futuro.

Con il migrare dell'elettronica del disco ai dischi stessi, la parte rimanente che divenne l'interfaccia (E)IDE è così piccola che può essere facilmente inserita nel chipset PCI. L'adattatore SCSI è più complesso e spesso include una propria piccola CPU ed è quindi più costoso e non integrato nei chipset PCI a disposizione in questo periodo. L'evoluzione tecnologica potrebbe cambiare questa situazione.

Alcuni adattatori vengono con gestione della cache separata ed intelligenza ma dal momento che ciò dipende principalmente dal sistema operativo, i vantaggi sono in stretta dipendenza con il sistema operativo stesso utilizzato. Alcuni dei più primitivi, che rimarranno senza nome, ne traggono grande vantaggio. Linux, invece, ha così tanti vantaggi di per sè stesso che i guadagni sono molto più scarsi.

Mike Neuffer, che fece i driver per i controller DTP, afferma che gli stessi sono abbastanza intelligenti che avendo sufficiente memoria cache, vi porterà una grossa spinta in prestazione e ritiene che la gente che aveva ottenuto piccoli guadagni con controller efficienti non aveva usato un controller di cache intelligente.

3.7 Sistemi Multi Canali

Al fine di aumentare il flusso, è necessario identificare i colli di bottiglia più significativi e poi eliminarli. In qualche sistema, in particolare dove c'è un grande numero di dischi connessi, è vantaggioso utilizzare diversi controller che lavorano in parallelo, sia per gli adattatori SCSI che per i controller IDE che normalmente hanno 2 canali incorporati. Linux gestisce questo.

Qualche controller RAID dispone di 2 o 3 canali ed è arduo suddividere il caricamento da disco attraverso tutti i canali. In altre parole, se avete due dischi SCSI che volete porre in RAID ed un controller a due canali, dovreste mettere ogni drive su canali separati.

3.8 Sistemi Multi Scheda

Oltre ad avere sia SCSI che IDE sulla stessa macchina, è anche possibile avere più di un controller SCSI. Controllate lo SCSI-HOWTO per quali controller potete combinare. Inoltre è molto probabile che dobbiate dire al kernel di verificare l'esistenza di più di un singolo controller SCSI o IDE. Questo si può fare utilizzando i parametri del kernel quando si fa il boot, ad esempio usando LILO. Controllate gli HOWTO per lo SCSI e per il LILO per capire come farlo.

I sistemi multi scheda possono offrire significanti guadagni di velocità se configurate il vostro disco in maniera corretta, specialmente per il RAID0. Assicuratevi di segnalare i controller come anche i dischi, in modo da aggiungere i dischi al dispositivo md RAID nell'ordine corretto. Se il controller 1 è connesso ai dischi `sda` e `sdc` mentre il controller 2 ai dischi `sdb` e `sdd` otterrete più parallelismo aggiungendo in ordine `sda - sdc - sdb - sdd` piuttosto che `sda - sdb - sdc - sdd` perché la lettura o la scrittura su più di un cluster abbraccerebbe verosimilmente due controller.

Gli stessi metodi possono essere applicati anche all'IDE. La maggior parte delle schede madri si presenta con 4 porte IDE:

- `hda` controller primario
- `hdb` slave primario
- `hdc` controller secondario
- `hdd` slave secondario

in cui i due primari condividono un cavo piatto ed i secondari ne condividono un altro. I chipset moderni li mantengono indipendenti. D'altronde è meglio fare un RAID con ordine `hda - hdc - hdb - hdd` perché ciò potrebbe presumibilmente parallelizzare ambedue i canali.

3.9 Confronti di Velocità

Le tavole seguenti sono state messe solo per indicare quali velocità sono possibili. Tutte le velocità di trasferimento sono in MB al secondo e le ampiezze di bus sono misurate in bit.

3.9.1 Controller

```
IDE             :    8.3 - 16.7
Ultra-ATA       :       33

SCSI            :
                        Ampiezza Bus (bit)

Velocita Bus (MHz)      |        8      16      32
--------------------------------------------------
 5                      |        5      10      20
10  (fast)              |       10      20      40
20  (fast-20 / ultra)   |       20      40      80
40  (fast-40 / ultra-2) |       40      80      --
--------------------------------------------------
```

3.9.2 Tipi di Bus

```
ISA             :      8-12
EISA            :       33
VESA            :       40    (A volte regolato a 50)

PCI
                        Ampiezza Bus (bits)

Velocita Bus (MHz)      |       32      64
--------------------------------------------------
33                      |       132     264
66                      |       264     528
--------------------------------------------------
```

3.10 Benchmarking

Questo è un argomento molto difficile e farò solamente pochi commenti cauti su questo campo minato. Prima di tutto, è più difficile fare benchmark paragonabili che abbiano un qualsiasi significato attuale. Questo comunque non trattiene la gente dal provare.

Invece si potrebbe usare il benchmarking per diagnosticare il vostro sistema, per verificare se sta andando tanto veloce quanto dovrebbe, cioè se non sta andando più lento. Inoltre vi aspettereste un aumento significante nel cambiare da un semplice file system ad un RAID, così un mancato guadagno di prestazioni vi dirà che qualcosa non va.

Quando provate a fare un benchmark non dovreste avventurarvi da soli, ma controllare `iozone` e `bonnie` e leggere la documentazione molto bene. In particolare assicuratevi che la dimensione del vostro buffer sia più grande della dimensione della vostra RAM, altrimenti testerete la vostra RAM piuttosto che i vostri dischi, il che vi riporterà alte prestazioni non realistiche.

Un benchmark molto semplice può essere ottenuto utilizzando `hdparm -tT` il che può essere utilizzato sia sui dischi IDE che sugli SCSI.

Per maggiori informazioni riguardo il benchmarking ed il software per una serie di piattaforme, controllate la pagine dei benchmark l' *ACNC* `http://www.acnc.com/benchmarks.html` .

3.11 Confronti

SCSI offre più prestazione di un EIDE ma ad un costo elevato. La terminazione è più complessa ma l'espansione non è molto difficile. Avere più di 4 dischi IDE (in alcuni casi 2) può essere complicato, mentre con il wide SCSI potete averne fino a 15 per adattatore. Qualche adattatore SCSI ha diversi canali, il che rende ulteriormente moltiplicabile il numero di dischi possibile.

Per lo SCSI dovrete dedicare un IRQ per adattatore, il quale può controllare fino a 15 dischi. Con l'EIDE avete bisogno di un IRQ per ogni disco e ciò può causare conflitto.

RLL e MFM sono in generale troppo vecchi, lenti ed inaffidabili per essere di qualche utilità.

3.12 Sviluppo Futuro

Lo SCSI-3 è in lavorazione e speriamo venga rilasciato presto. Dispositivi più veloci sono già stati annunciati, recentemente sono state proposte le mostruose specifiche da 80 MB/s e 160 MB/s. Queste sono basate sullo standard Ultra-2 (che utilizzava un clock a 40 MHz) combinato con un cavo a 16 bit.

Alcuni produttori annunciano già dispositivi SCSI-3 ma ciò è in realtà abbastanza prematuro visto che lo standard è ancora da consolidare. Con l'aumentare delle velocità di trasferimento, il punto di saturazione del bus PCI si sta avvicinando. Attualmente la versione a 64 bit ha un limite di 264 MB/s. La velocità di trasferimento in futuro sarà aumentata dall'attuale 33MHz a 66MHz, aumentando quindi il limite a 528 MB/s.

Un altro orientamento è per dischi sempre più grandi. Ho sentito che è possibile ottenere 55 GB su un singolo disco sebbene questo sia molto costoso. Attualmente il dispositivo di memorizzazione migliore per le vostre tasche è di circa 6.4 GB ma anche questo è in continuo aumento. L'introduzione del DVD avrà nel prossimo futuro un grande impatto, con circa 20 GB su un singolo disco potete avere una copia completa anche dei maggiori siti FTP intorno al mondo. L'unica cosa di cui siamo sicuramente certi riguardo al futuro è che se non sarà migliore, sarà sicuramente più grande.

Addendum: poco dopo avere iniziato a scrivere questo documento, ho letto che la massima velocità utile per un CD-ROM era 20x dal momento che la stabilità meccanica sarebbe stata un problema molto grande a queste velocità. Dopo circa un mese furono disponibili CD-ROM a 24x... Ora potete ottenere un 40x e non c'è dubbio che velocità più alte siano in produzione.

3.13 Raccomandazioni

La mia opinione personale è che EIDE o Ultra ATA è il modo migliore per cominciare un vostro sistema, specialmente se intendete usare anche il DOS sulla vostra macchina. Se intendete espandere il vostro sistema nel corso degli anni o intendete utilizzarlo come server, vi suggererei caldamente di prendere dischi SCSI. Attualmente wide SCSI è un po' più costoso. Otterrete di più col vostro denaro mettendo uno SCSI di ampiezza standard. Ci sono inoltre versioni differenziali del bus SCSI il che aumenta la lunghezza massima del cavo. L'aumento di prezzo è ancor più sostanziale e non può essere consigliato ad utenti normali.

Oltre ai dischi, potete anche connettere qualche tipo di scanner e stampante ed anche le reti ad un bus SCSI.

Inoltre ricordatevi che espandere il vostro sistema necessiterà anche di più potenza, così assicuratevi che l'alimentazione sia adatta al lavoro e che voi abbiate raffreddamento sufficiente. Molti dischi SCSI offrono l'opzione di avviamento sequenziale, il che è una buona idea per grandi sistemi. Controllate anche 19.3 (Alimentazione e Riscaldamento).

4 Struttura del File System

Linux è stato multi tasking dall'inizio, da quando un bel numero di programmi interagivano e giravano ininterrottamente. In ogni caso è importante mantenere una struttura dei file su cui tutti potrebbero essere d'accordo, in modo che il sistema trovi i dati dove ci si aspetta di trovarli. Storicamente ci sono stati così tanti standard differenti che si è creata confusione e la compatibilità è stata mantenuta usando collegamenti simbolici, il che ha fatto ancora più confusione e la struttura è finita col sembrare un labirinto.

Nel caso di Linux, uno standard chiamato il *File Systems Standard* (FSSTND) fu fortunatamente accettato all'inizio e oggi è utilizzato dalle maggiori distribuzioni Linux.

Più tardi, venne deciso di creare un successore che supportasse anche altri sistemi operativi oltre al solo Linux e venne chiamato *Filesystem Hierarchy Standard* (FHS) attualmente alla versione 2.1. Questo standard è in continua evoluzione e sarà quanto prima adottato dalle distribuzioni Linux.

Vi consiglio di non sviluppare una vostra struttura personale visto che una marea di pensieri è servita per avere gli standard e molti pacchetti si conformano a questi standard. Invece, potete leggere di più su questo presso l'home page del

FHS `http://www.pathname.com/fhs` .

Questo HOWTO si sforza di conformarsi al FSSTND e perseguirà il FHS quando le distribuzioni saranno disponibili.

4.1 Caratteristiche del File System

Le varie parti di FSSTND hanno diverse necessità riguardo la velocità, l'affidabilità e la dimensione, ad esempio perdere root è un dolore ma può essere facilmente recuperabile. Perdere `/var/spool/mail` è una cosa abbastanza differente. Ecco un veloce sommario di alcune parti essenziali e delle loro proprietà e necessità. Badate che questa è solamente una guida e ci potrebbero essere file binari nelle directory `etc` e `lib`, librerie nelle directory `bin` e così via.

4.1.1 Swap

Velocità

Il massimo! Sebbene facciate troppo affidamento sullo swap, dovreste considerare di comprare più RAM. Notate comunque, che su molte schede madri, la cache non funzionerà se la RAM è superiore a 128 MB.

Dimensione

Stesso discorso della RAM. Un algoritmo veloce e sporco: un po' come per il tè: 16 MB per la macchina e 2 MB per ogni utente. Il kernel più piccolo gira in un mega ma sta stretto. Usate 4 MB per il lavoro in generale e applicazioni leggere, 8 MB per X11 o GCC o 16 MB per stare più comodi.

(Si sa che l'autore prende tazze di tè abbastanza potenti...)

Alcuni suggeriscono che lo spazio di swap debba essere 1-2 volte la dimensione della RAM, evidenziando che la localizzazione dei programmi determina quanto sia efficace lo spazio swap aggiunto. Notate che usare lo stesso algoritmo come per 4BSD è abbastanza scorretto dal momento che Linux non riserva spazio per pagine in core.

Un approccio più preciso è considerare lo spazio di swap più la RAM come il vostro insieme lavorativo totale, così se sapete quanto spazio, al massimo, vi serve, sottraete la RAM fisica che avete e che è lo spazio di swap di cui avrete bisogno.

C'è inoltre un'altra ragione per essere generosi quando dimensionate il vostro spazio di swap: i furti di memoria. Programmi che funzionano male, che non liberano la memoria che riservano per loro stessi si dice che provocano un furto di memoria. Questa allocazione rimane anche dopo che il programma incriminato è terminato e quindi questo è fonte di consumo di memoria. Una volta che tutta la RAM fisica e lo spazio di swap sono esaurite, l'unica soluzione è fare un reboot e cominciare da capo. Per fortuna questi programmi non sono così comuni ma se voi ne doveste incontrare qualcuno, troverete che uno spazio swap extra, vi prenderà più tempo tra i riavvii.

Ricordate inoltre di prendere in considerazione il tipo di programma che usate. Alcuni programmi che hanno grandi ambienti di lavoro, come quelli di modellazione agli elementi finiti (FEM), hanno raffinatissime strutture di dati caricate nella RAM piuttosto che lavorare esplicitamente sui file. Programmi per i dati e per il calcolo complessi come questi, causeranno swapping eccessivo se avete meno RAM del necessario.

Altri tipi di programmi possono bloccare le loro pagine nella RAM. Questo può accadere per motivi di sicurezza, prevenendo copie dei dati che raggiungono un dispositivo di swap o per motivi di prestazioni come nel modulo in tempo reale. In ogni caso, bloccare queste pagine, riduce il quantitativo di memoria swappabile rimanente e può fare in modo che il sistema swappi prima di quanto ci si possa aspettare.

Nel `man 8 mkswap` c'è spiegato che ogni partizione di swap può essere massimo 128 MB in dimensione per macchine a 32-bit e 256 MB per macchine a 64-bit.

Affidabilità

Media. Quando fallisce sapete lo fa rapidamente e il fallimento vi costerà qualche perdita di lavoro. Salvate spesso, no?

Nota 1

Linux offre la possibilità di distribuire lo swapping su più dispositivi, una caratteristica che vi può far guadagnare parecchio. Controllate `man 8 swapon` per maggiori dettagli. Tuttavia, il guadagno di utilizzare un disco RAID per lo `swap` potrebbe essere vanificato dalla maggiore pesantezza del codice di RAID.

Perciò il file `/etc/fstab` potrebbe somigliare a questo:

```
/dev/sda1       swap            swap    pri=1           0       0
/dev/sdc1       swap            swap    pri=1           0       0
```

Ricordate che il file `fstab` è *molto* sensibile alla formattazione utilizzata, leggete le pagine man attentamente e *non* fate un semplice taglia e incolla delle righe qui sopra.

Nota 2

Qualcuno per swappare utilizza un RAM disk o qualche altro file system. In ogni caso, se non avete impostazioni o necessità inusuali, difficilmente trarrete vantaggio da questo e questo porta ad usare la memoria a disposizione per le operazioni di cache e buffer.

Nota 2b

C'è un'eccezione: su un numero di schede madri mal disegnate, la memoria cache integrata su piastra madre non è in grado di porre in cache tutta la RAM che può essere indirizzata. Molte schede madri vecchie, erano in grado di accettare 128 MB di RAM ma potevano porre in cache solo le prime 64. In questi casi, aumenterebbe la prestazione se voi utilizzaste gli ultimi 64 MB di RAM come swap basato su RAMdisk o come altro supporto di memorizzazione temporaneo.

4.1.2 Archiviazione Temporanea (`/tmp` e `/var/tmp`)

Velocità

Molto alta. Su un disco/partizione separati questo generalmente ridurrà la frammentazione, sebbene `ext2fs` gestisca la frammentazione abbastanza bene.

Dimensione

Difficile a dirsi, piccoli sistemi girano facilmente con appena pochi MB ma questi sono noti posti di nascondiglio per tenere lontano dalla vista di occhi curiosi e dal rinforzo della quota e possono crescere senza controllo su macchine più grandi. Suggerito: piccole macchine casalinghe: 32 MB, piccoli server: 128 MB e grosse macchine fino a 500 MB (la macchina che l'autore usa al lavoro ha 1100 utenti e una `/tmp` di 300 MB). Controllate queste directory, non solo per file nascosti ma anche per vecchi file. Preparatevi anche al fatto che queste partizioni potrebbero essere la prima ragione che potreste avere per ridimensionare le vostre partizioni.

Affidabilità

Bassa. Quando queste aree falliscono o si riempiono, spesso i programmi emetteranno un avvertimento, oppure falliranno splendidamente. Errori di file casuali saranno ovviamente più seri, non importa quale area file sia.

File

Generalmente piccoli file ma potrebbe esservene un gran numero. Normalmente i programmi cancellano i loro file `tmp` ma se casualmente interviene un'interruzione, possono sopravvivere. Molte distribuzioni hanno una linea di condotta riguardo la pulizia dei file `tmp` all'avvio, quindi potreste voler controllare quale sia la vostra configurazione.

Nota1

In FSSTND c'è una nota riguardo al mettere `/tmp` su RAM disk. Questo comunque non è consigliato per le stesse ragioni espresse al riguardo dello swap. Inoltre, come segnalato prima, non utilizzate dispositivi flash RAM per queste directory. Ci si dovrebbe ricordare inoltre che qualche sistema è organizzato in modo da pulire automaticamente le aree `tmp` al riavvio.

Nota2

I sistemi più vecchi avevano un `/usr/tmp` ma questo non è consigliato e, per motivi storici, un link simbolico lo fa puntare ad una delle altre aree `tmp`.

(* That was 50 lines, I am home and dry! *)

4.1.3 Aree di Coda (`/var/spool/news` e `/var/spool/mail`)

Velocità

Alta, specialmente su grossi server di news. Il trasferimento e la scadenza delle news usano molto il disco e beneficieranno di dischi veloci. Code di stampa: bassa. Considerate RAID0 per le news.

Dimensione

Per i server di news/posta: quanta ve ne potete permettere. Per sistemi ad utente singolo pochi MB saranno sufficienti se leggete in continuazione. Iscriversi ad un list server e prendersi una vacanza, non è proprio una buona idea (ancora, la macchina che uso al lavoro ha 100 MB riservati per l'intera `/var/spool`).

Affidabilità

Posta: molto alta, news: media, coda di stampa: bassa. Se la vostra posta è molto importante (non lo è forse sempre?) considerate RAID per l'affidabilità.

File

Generalmente una grande quantità di file che in dimensione sono intorno a pochi KB. I file nella coda di stampa, possono d'altra parte essere pochi ma abbastanza grandi.

Nota

Qualche documentazione sulle news, suggerisce di mettere tutti i file `.overview` su un disco separato dai file delle news, controllate tutte le FAQ sulle news per più informazioni. La dimensione tipica è di circa il 3-10 per cento di tutto lo spazio di coda.

4.1.4 Directory Home (`/home`)

Velocità

Media. Sebbene molti programmi utilizzino `/tmp` per archiviare temporaneamente, altri come alcuni lettori di news, aggiornano frequentemente i file nella directory home, il che può essere notato su ampi sistemi multitenti. Per piccoli sistemi, questo non è un problema.

Dimensione

Difficile! Su qualche sistema, la gente paga per archiviare, perciò questo generalmente è una questione di soldi. Grossi sistemi come *nyx.net* `http://www.nyx.net/`

(che è un servizio di Internet gratuito con posta, news e servizi WWW) girano con successo con un limite suggerito di 100 KB per utente e 300 KB come massimo applicato. Gli ISP commerciali offrono tipicamente circa 5 MB nei loro pacchetti standard di abbonamento.

Se comunque scrivete libri o fate del lavoro grafico, le necessità si gonfiano velocemente.

Affidabilità

Variabile. Perdere `/home` su una macchina con un singolo utente è fastidioso ma quando 20000 utenti ti chiamano per dirti che le proprie directory home sono andate è più che fastidioso. Per alcuni, la propria fonte di sostentamento risiede qui. Fate backup regolari ovviamente, no?

File

Ugualmente difficile. Il setup minimo per un singolo utente è di circa una dozzina di file, di dimensione di 0.5 - 5 KB. I file relativi al progetto, possono essere sufficientemente numerosi.

Nota1

Potreste considerare RAID sia per la velocità che per l'affidabilità. Se volete velocità estremamente alta ed affidabilità, dovreste vedere altri sistemi operativi e piattaforme hardware (tolleranza ai guasti ecc.).

Nota2

I navigatori spesso utilizzano una cache locale per aumentare la velocità di navigazione e questa cache può occupare una quantità di spazio sostanziale e causare parecchia attività del disco. Ci sono molti metodi per evitare questi colpi alle prestazioni, per maggiori informazioni guardate le sezioni sulle 10.2.1 (Directory Home) e 10.2.3 (WWW).

Nota3

Gli utenti spesso tendono ad utilizzare tutto lo spazio disponibile sulla partizione `/home`. Il sottosistema Quota di Linux è in grado di limitare il numero di blocchi e il numero di inode che l'ID di un singolo utente può allocare su una base di un filesystem. Guardate il *Linux Quota mini-HOWTO* `http://metalab.unc.edu/LDP/mini` di Albert M.C. Tam `bertie (at) scn.org` per i dettagli sull'impostazione.

4.1.5 File binari principali (`/usr/bin` e `/usr/local/bin`)

Velocità

Bassa. Spesso i dati sono più grandi dei programmi che sono comunque richiesti su domanda, quindi questo non è rischioso per la velocità. Testimoni i live file system su CD ROM.

Dimensione

Il limite è il cielo ma 200 MB dovrebbero darvi molto di quello di cui avete bisogno per un sistema completo. Un grande sistema, per sviluppo software o un server dai molti fini dovrebbe forse riservare 500MB sia per l'installazione che per la crescita.

Affidabilità

Bassa. è generalmente montato sotto root dove tutti gli essenziali sono raccolti. Tuttavia perdere tutti i file binari fa male...

File

Variabile ma generalmente dell'ordine di 10 - 100 KB.

4.1.6 Librerie (`/usr/lib` e `/usr/local/lib`)

Velocità

Media. Questi sono grosse trance di dati caricati spesso, spaziano dai file oggetto ai font, tutti suscettibili di rigonfiamento. Spesso questi sono anche caricati nella loro interezza e la velocità è abbastanza utilizzata qui.

Dimensione

Variabile. Qui, ad esempio, $egrave; dove gli elaboratori di testo archiviano i loro immensi file di font. I pochi da cui ho avuto un feedback su questo, riportano circa 70 MB nelle loro varie directory `lib`. Un'installazione abbastanza completa della Debian 1.2, può arrivare a circa 250 MB, che può essere preso come un limite superiore realistico. I seguenti sono alcuni dei più grossi consumatori di spazio su disco: GCC, Emacs, TeX/LaTeX, X11 e perl.

Affidabilità

Bassa. Controllate il punto 4.1.5 (File binari principali).

File

Generalmente grossi, di cui molti di dimensioni dell'ordine di 1 MB.

Nota

Per motivi storici, qualche programma tiene gli eseguibili nelle aree lib. Un esempio è GCC che ha grossi file binari nella gerarchia `/usr/lib/gcc/lib`.

4.1.7 Boot

Velocità

Abbastanza bassa: dopo tutto il boot non avviene così di frequente e caricare il kernel è solo una minima frazione del tempo che si impiega per rendere operativo il sistema.

Dimensione

Abbastanza piccola, un'immagine completa con qualche extra entra in un singolo floppy, così 5 MB dovrebbero essere sufficienti.

Affidabilità

Alta. Guardate la sezione sotto su Root.

Nota 1

La parte più importante riguardo la partizione di boot è che su molti sistemi, *deve* risiedere al di sotto del cilindro 1023. Questa è una limitazione BIOS che Linux non riesce a gestire.

4.1.8 Root

Velocità

Abbastanza bassa: qui c'è solo il minimo indispensabile, la maggior parte del quale gira solo all'inizio.

Dimensione

Relativamente piccola. In ogni caso è una buona idea mantenere qualche file di recupero e utilità sulla partizione di root e qualcuno ci tiene diverse versioni del kernel.

Affidabilità

Alta. Un fallimento qui causerà probabilmente un gran dolore e potresti finire col perdere del tempo a recuperare la tua partizione di boot. Con un po' di pratica potete naturalmente farlo in un'ora o giù di lì, ma penso che anche se avete un po' di pratica nel farlo, state anche facendo qualcosa di sbagliato.

Naturalmente avete un disco di recupero no? Ovviamente questo è stato aggiornato da quando avete fatto l'installazione iniziale? Ci sono molti dischi di recupero già fatti come anche strumenti per la creazione di dischi di recupero che potreste trovare validi. Presumibilmente investire del tempo in questi vi salva dal diventare un esperto nel recupero di root.

Nota 1

Se avete molti dischi, potreste considerare di mettere una partizione boot di emergenza di ricambio su un disco fisico separato. Vi costerà un po' di spazio ma se il vostro setup è molto vasto il tempo risparmiato, nel caso qualcosa fallisse, varrebbe molto lo spazio extra.

Nota 2

Per semplicità e anche nel caso di emergenza non è consigliabile di mettere la partizione root su un sistema RAID a livello 0. Inoltre se utilizzate il RAID per la vostra partizione di boot, dovete ricordare che sia abilitata l'opzione `md` per il vostro kernel di emergenza.

Nota 3

Per semplicità è abbastanza comune mantenere Boot e Root sulla stessa partizione. Se fate ciò, allora al fine di fare il boot da LILO, è importante che i file di boot essenziali risiedano tutti entro il cilindro 1023. Questo include il kernel come anche i file che si trovano in `/boot`.

4.1.9 DOS ecc.

A rischio di apparire eretico ho incluso questa piccola sezione riguardo un qualcosa contro cui molti di quelli che leggono questo documento hanno forti sensazioni. Sfortunatamente molti componenti hardware li troviamo con setup e mezzi di mantenimento basati su questi sistemi, quindi ecco qui.

Velocità

Molto bassa. I sistemi in questione non sono famosi per la velocità, quindi c'è un piccolo appunto sull'utilizzo di dischi di prima qualità. Il multitasking o il multi-threading non sono disponibili, quindi il comando di agevolazione di accodaggio nei dischi SCSI non è una cosa di cui potrete trarre vantaggio. Se avete un vecchio disco IDE sarà sufficientemente buono. L'eccezione è in qualche modo Win95 e ancor di più NT che hanno supporto multi-threading che teoricamente dovrebbe essere in grado di trarre vantaggio delle più avanzate caratteristiche offerte dai dispositivi SCSI.

Dimensione

La compagnia che sta dietro questi sistemi operativi non è famosa per scrivere codice stringato, così dovete essere preparati a spendere poche decine di MB a seconda di quale versione installate del DOS o di Windows. Con una vecchia versione di DOS o Windows potreste fare entrare tutto in 50 MB.

Affidabilità

Ha-ha. Visto che la catena non è più forte dell'anello più debole, potete usare qualsiasi vecchio disco. Dal momento che è più facile che il SO si scombini da solo, piuttosto che il drive si autodistrugga, imparerete presto l'importanza di backup qui.

Mettetela in un altro modo: *La vostra missione, se doveste accettarla, è di fare funzionare questa partizione. La garanzia si autodistruggerà tra 10 secondi...*

Recentemente mi è stato chiesto di giustificare le mie lamentele in questa sede. Prima di tutto non sto cercando misere scuse per il Dos e Windows come sistemi operativi. Secondariamente ci sono vari articoli legali che devono essere presi in considerazione. Dire che c'è una relazione tra le due ultime frasi è solamente un vaneggiamento paranoide. Di sicuro. Invece offrirò allo stimato lettore un po' di parole chiavi: DOS 4.0, DOS 6.x e vari mezzi di compressione del disco che rimarranno senza nome.

4.2 Spiegazione dei termini

Ovviamente più veloce è meglio è, ma spesso il felice installatore di Linux ha svariati dischi di varia velocità e affidabilità, così anche se questo documento descrive la prestazione come 'veloce' o 'lenta' è solamente una rozza guida dal momento che non è determinabile alcun tipo di precisione più fine. Anche se ci sono dei dettagli che si dovrebbero ricordare:

4.2.1 Velocità

Questa è realmente una confusa commistione di svariati termini: carico della CPU, impostazioni generali del trasferimento, tempo di accesso al disco e velocità di trasferimento. È nella reale natura della regolazione che non c'è un optimum fisso e in molti casi il prezzo è il fattore che fa la differenza. Il carico della CPU è rilevante solo per i sistemi IDE dove è la CPU che esegue da sola il trasferimento, ma è generalmente bassa per lo SCSI, controllate la documentazione SCSI per i valori reali. Anche il tempo di accesso al disco è piccolo, generalmente dell'ordine del millisecondo. Questo comunque non è un problema se usate comandi di accodamento su SCSI, dove sovrapponete comandi mantenendo il bus occupato per tutto il tempo. Gli spool delle news sono un caso speciale consistente in un grande numero di file normalmente piccoli, così in questo caso il tempo di accesso può divenire molto significativo.

Ci sono due parametri principali che sono di interesse qui:

L'accesso

: è generalmente definito come il tempo che occorre alla testina di lettura/scrittura per saltare da una traccia ad un'altra. Questo parametro è importante quando si ha a che fare con un grande numero di piccoli file come visto nei file di spool. C'è inoltre l'ulteriore ritardo di accesso prima che il settore desiderato ruoti in posizione sotto la testina. Questo ritardo dipende dalla velocità angolare del disco ed ecco perché a volte questo parametro è riportato per i dischi. I valori comuni sono 4500, 5400 e 7200 RPM (rotazioni al minuto). RPM più alti riducono il tempo di accesso ma ad un costo sostanziale. Inoltre dischi che lavorano a 7200 RPM si sa che sono rumorosi e che generano un grande calore, fattore che dovrebbe essere tenuto in considerazione se state costruendo un grande insieme o una batteria di dischi. Molto di recente sono entrati nel mercato dischi che lavorano a 10000 RPM e qui le richieste di raffreddamento sono ancora più strette e vengono date pochissime indicazioni per il flusso d'aria.

Il trasferimento

: è generalmente espresso in megabyte al secondo. Questo parametro è importante quando si utilizzano grandi file che devono essere trasferiti. I file di libreria, i dizionari e i file di immagine sono un esempio di questo. I dischi che posseggono alta velocità rotazionale, normalmente hanno anche trasferimenti veloci visto che la velocità di trasferimento è proporzionale alla velocità angolare per la stessa densità di settore.

È inoltre importante leggere le specifiche per i dischi molto attentamente e noterete che la massima velocità di trasferimento è spesso riportata intendendo i trasferimenti dalla cache integrata (burst speed) e *non* direttamente dalla superficie del disco (sustained speed). Guardate anche la sezione sull' 19.3 (Alimentazione e riscaldamento).

4.2.2 Affidabilità

Naturalmente nessuno vorrebbe dischi con bassa affidabilità ma uno farebbe bene a giudicare vecchi dischi come inaffidabili. Inoltre per scopi RAID (controllate le informazioni pertinenti) è consigliato utilizzare un insieme misto di dischi così che crash simultanei di più dischi possano diventare meno probabili.

Fino ad oggi ho avuto notizia di un solo rapporto di fallimento totale del file system, ma qui l'hardware instabile è sembrato essere la causa dei problemi.

Anche se i dischi sono economici al giorno d'oggi, la gente ancora sottostima il valore dei contenuti dei dischi. Se avete bisogno di affidabilità più alta, assicuratevi di rimpiazzare i vecchi dischi e mantenete i ricambi. Non è inusuale che i dischi possano lavorare più o meno continuamente per anni ed anni ma ciò che spesso uccide i dischi alla fine sono i cicli di alimentazione.

4.2.3 File

La dimensione media dei file è importante al fine di decidere i paramentri del disco più appropriati. Un grande numero di piccoli file fa sì che il tempo di accesso divenga importante, invece per grossi file diventa importante la velocità di trasferimento. Il comando di accodamento nei dispositivi SCSI è molto comodo per maneggiare un grosso numero di piccoli file, ma per il trasferimento l'EIDE non è così lontano dallo SCSI e normalmente molto più economico dello SCSI.

5 File System

Nel corso del tempo le necessità per i file system sono aumentate e le domande per grosse strutture, grossi file, nomi lunghi e altro ancora ha generato la richiesta di file system, il sistema che accede e organizza i dati sulle unità di memorizzazione, ancora più avanzati. Oggi c'è un gran numero di file system tra cui scegliere e questa sezione li descriverà in dettaglio.

L'enfasi è su Linux ma con più richieste sarò felice di aggiungere informazioni per un'audience più ampia.

5.1 File system per scopi generali

La maggior parte dei sistemi operativi ha generalmente un file system per scopi generali per utilizzo di ogni giorno per la maggior parte dei tipi di file, mostrando caratteristiche nel SO come permessi, protezioni e recupero.

5.1.1 `minix`

Questo fu il fs originale per Linux, agli albori Linux era ospitato su macchine minix. È semplice ma limitato nelle caratteristiche e difficilmente viene utilizzato in questi giorni se non per qualche disco di recupero visto che è sufficientemente compatto.

5.1.2 `xiafs` e `extfs`

Questi sono ugualmente vecchi e sono caduti in disuso e non sono più consigliati.

5.1.3 `ext2fs`

Questo è lo standard stabilito per scopi generali nel mondo Linux. È veloce, efficiente e maturo ed è in continua evoluzione e caratteristiche quali ACL e la compressione trasparente sono prossime.

Per maggiori informazioni controllate l'home page di

ext2fs `http://web.mit.edu/tytso/www/linux/ext2.htm`

5.1.4 `ufs`

Questo è il filesystem utilizzatto da BSD e sue varianti. È maturo ma è stato anche sviluppato per tipi di dischi più vecchi dove le geometrie si conoscevano. Il fs utilizza un bel numero di trucchetti per ottimizzare le prestazioni ma dal momento che le geometrie del disco sono tradotte in un bel numero di modi, l'effetto rete non è più così ottimale.

5.1.5 efs

L'Extent File System (efs) è il giovane file system di Silicon Graphics ampiamente utilizzato su IRIX prima della versione 6.0 dopo la quale è subentrato l'xfs. Mentre viene incoraggiata la migrazione ad xfs, efs è ancora supportata e molto usata sui CD.

C'è un driver Linux in versione beta giovane, ottenibile presso l'home page

Linux extent file system `http://aeschi.ch.eu.org/efs/`

5.1.6 reiserfs

Dal 23 Luglio 1997, Hans Reiser `reiser (at) RICOCHET.NET` ha messo su web il sorgente del suo

reiserfs `http://idiom.com/~beverly/reiserfs.html`

basato su una struttura ad albero. Sebbene il suo file system abbia delle caratteristiche veramente interessanti e sia molto più veloce dell'`ext2fs`, è ancora troppo sperimentale e difficile da integrare con il kernel standard. Ci si aspetta qualche sviluppo interessante nel futuro - questo è ben differente dal vostro progetto file system medio basato su log per Linux, perché Hans ha già un codice che funziona.

5.1.7 enh-fs

Attualmente in stadio alfa, il progetto

Enhanced File System `http://www.coker.com.au/~russel/enh-fs.html`

punta a combinare su un unico livello il file system e la gestione del disco.

5.2 File System Microsoft

Questa compagnia è responsabile di molte cose, tra cui alcuni file system, tanto che alla fine ha causato confusioni.

5.2.1 fat

In realtà ci sono 2 `fat`, `fat12` e `fat16`, dipendentemente dalla dimensione della partizione utilizzata, ma fortunatamente la differenza è così piccola che l'intera questione è chiara.

Tra i fattori a favore, sono veloci e semplici e molti SO li gestiscono e possono sia leggere che scrivere su questo filesystem. E questo è quanto.

Il fattore a sfavore è la limitata sicurezza, i flag dei permessi severamente limitati e scalabilità atroce. Ad esempio con `fat` non potete avere partizioni più grandi di 2 GB.

5.2.2 fat32

Dopo circa 10 anni la Microsoft realizzò cosa fosse la `fat`. Bene, 10 anni in ritardo e creò così questo file system che scala ragionevolmente bene.

I flag dei permessi sono ancora limitati. NT 4.0 non può leggere questo file system ma Linux può.

5.2.3 `vfat`

Nello stesso periodo in cui Microsoft lanciò la `fat32`, aggiunsero anche il supporto per i nomi lunghi dei file, conosciuto come `vfat`.

Linux legge partizioni `vfat` e `fat32` mediante mount con il tipo `vfat`.

5.2.4 `ntfs`

Questo è il file system nativo di Win-NT ma dal momento che non sono disponibili informazioni complete c'è un supporto limitato per altri sistemi operativi.

5.3 File system per il Logging e il Journaling

Apportano un approccio radicalmente differente agli aggiornamenti dei file registrando le modifiche di un file in un log e successivamente controllando saltuariamente i log.

La lettura in pratica è veloce come un file system tradizionale che aggiorna sempre i file direttamente. La scrittura è invece molto più veloce, dal momento che gli aggiornamenti sono aggiunti ad un log. Tutto ciò è trasparente all'utente. È nell'affidabilità e particolarmente nel controllo dell'integrità del file system che questi file system brillano. Dal momento che dall'ultimo controllo si sa che i dati stanno bene, dovranno essere controllati solo i log e questa cosa è molto più veloce rispetto ai sistemi tradizionali.

Notate che i file system di *logging* tengono traccia dei cambiamenti fatti sia ai dati chee agli inode, mentre i file system che effettuano *journaling* tengono traccia solamente del cambiamento degli inode.

Linux ha abbastanza scelta tra questi file system ma nessuno è ancora in qualità tale da essere prodotto. Alcuni sono anche sospesi.

- Adam Richter della Yggdrasil ha postato qualche tempo fa la notizia che stavano lavorando su un file system basato su un file di log compresso e che questo progetto è attualmente sospeso. Nonostante ciò, una versione non funzionante è disponibile sui propri server FTP. Controllate

  the Yggdrasil ftp server `ftp://ftp.yggdrasil.com/private/adam`

  dove possono essere trovate speciali versioni con patch del kernel.

- Un altro progetto è il

  Linux log-structured Filesystem Project `http://collective.cpoint.net/lfs/`

  che è anch'esso tristemente sospeso. Nonostante ciò questa pagina contiene parecchie informazioni sull'argomento.

- Infine c'è il

  dtfs – A Log-Structured Filesystem For Linux `http://www.complang.tuwien.ac.at/czezatke/lfs.html`

  che sembra crescere forte. Ancora in alfa ma sufficientemente completo da far girare programmi su questo file system.

5.4 File System di sola lettura

I supporti di sola lettura non sono sfuggiti alle sempre più crescenti complessità viste nei file system più generali, quindi c'è ancora una vasta scelta con corrispondenti opportunità per errori eccitanti.

Molti di questi sono usati nei CD-ROM ma anche il nuovo DVD può utilizzarli ed è pure possibile utilizzarli attraverso il dispositivo di loopback su un file di un hard disk per verificare un'immagine prima di masterizzare una ROM.

C'è un `romfs` per Linux ma visto che non è relativo al disco, più nulla potrà essere detto relativamente a questo in questa sede.

5.4.1 `High Sierra`

Questo fu uno dei più giovani standard per i formati CD-ROM, chiamato così forse dall'albergo in cui si raggiunse l'ultimo accordo.

`High Sierra` era così limitato nelle caratteristiche che le nuove estensioni dovevano semplicemente apparire e dal momento che non c'è stata conclusione di nuovi formati, l'originale `High Sierra` rimane il comune precursore ed è quindi ancora ampiamente supportato.

5.4.2 `iso9660`

L'International Standards Organisation fece le proprie estensioni e formalizzò lo standard nel quale noi riconosciamo lo standard `iso9660`.

Il file system Linux iso9660 gestisce sia le estensioni High Sierra che quelle `Rock Ridge`.

5.4.3 `Rock Ridge`

Non tutti accettano limiti come i nomi corti e assenza di permessi, così molto presto sopraggiunsero le estensioni `Rock Ridge` per rettificare queste mancanze.

5.4.4 `Joliet`

La Microsoft, per non essere superata nel gioco delle estensioni standard, decise che avrebbe dovuto estendere i formati CD-ROM con qualche caratteristica di internazionalizzazione e l'ha chiamata `Joliet`.

Linux gestisce questi standard nei kernel 2.0.34 o superiori. Dovete abilitare l'NLS per usarlo.

5.4.5 Trivia

Joliet è una città fuori Chicago; più conosciuta per essere stata il posto della prigione dove Jake fu ingabbiato nel film Blues Brothers. Il Rock Ridge (le estensioni UNIX all'ISO 9660) è chiamato così dalla città (fittizia) nel film Blazing Saddles.

5.4.6 `UDF`

Con l'arrivo del DVD con fino a 17 GB di capacità di immagazzinamento, il mondo sembrò apparentemente volere un altro formato, questa volta fu chiamato ambiziosamente Universal Disk Format (UDF). Fu inteso come il rimpiazzo per l'`iso9660` e sarà richiesto per il DVD.

Attualmente non è nel kernel standard di Linux ma un progetto è in corso per fare un

driver UDF `http://trylinux.com/projects/udf/index/htm`

per Linux. Patch e documentazione sono disponibili.

5.5 File System di Rete

Sono disponibili un gran numero di tecnologie di rete che vi permettono di distribuire dischi attraverso reti locali o globali. Ciò è in qualche modo marginale all'argomento di questo HOWTO ma dal momento che può essere utilizzato su dischi locali, lo tratterò brevemente. Sarebbe meglio se qualcuno ne facesse un HOWTO separato.

5.5.1 `NFS`

Questo è stato uno dei primi sistemi che permettono di montare uno spazio file di una macchina,su un altra. Ci sono una serie di problemi con `NFS` che oscillano dalle prestazioni alla sicurezza, ma in ogni caso ciò si è stabilizzato.

5.5.2 `AFS`

Questo è un sistema che permette un'efficiente condivisione di file attraverso ampie reti. Iniziato come un progetto accademico, è oggi venduto da

Transarc `http://www.transarc.com`

la cui homepage vi darà più dettagli.

Derek Atkins, del MIT, ha fatto il porting di AFS per Linux ed ha inoltre organizzato per questo la mailing list del Linux AFS (*linux-afs@mit.edu*) che è aperta al pubblico. Richieste per partecipare alla lista dovrebbero essere spedite a *linux-afs-request@mit.edu*

ed infine si dovrebbero riportare i bug a *linux-afs-bugs@mit.edu* .

Importante: dal momento che AFS utilizza la cifratura, è un software che è caratterizzato da restrizione e non può essere facilmente esportato dagli Stati Uniti.

L'IBM che possiede la Transarc, ha annunciato la disponibilità dell'ultima versione del client come anche del server per Linux.

Arla è un'implementazione AFS gratuita, controllate l' *Arla homepage* `http://www.stacken.kth.se/projekt/arla`

per maggiori informazioni come anche per la documentazione.

5.5.3 Coda

È iniziato un lavoro per un sostituto gratuito di `AFS` ed è chiamato

Coda `http://coda.cs.cmu.edu/` .

5.5.4 `nbd`

Il

Dispositivo di Blocco di Rete `http://atrey.karlin.mff.cuni.cz/~pavel`

(`nbd`) è disponibile nel kernel 2.2 di Linux e successivi e offre prestazioni eccellenti accertate. La cosa interessante è che può essere combinato con RAID (vedere dopo).

5.5.5 GFS

Il

File System Globale `http://gfs.lcse.umn.edu/`

è un nuovo file system progettato per immagazzinare attraverso un'ampia area di rete. Attualmente è agli stadi iniziali e più informazioni si avranno più avanti.

5.6 File System Speciali

Oltre ai file system generali, ce ne sono altri più specifici, generalmente utilizzati per fornire prestazioni migliori o altre caratteristiche, di solito con mancanze su altri aspetti.

5.6.1 `tmpfs` e `swapfs`

Per un'archiviazione veloce di file a breve termine, Sun OS offre `tmpfs` che è praticamente la stessa cosa di `swapfs` su NeXT. Questo risolve l'intrinseca lentezza di `ufs` mettendo in cache i dati e mantenendo l'informazione di controllo in memoria. Ciò significa che i dati su un tale file system si perderanno al reboot ed è quindi adatto per l'area `/tmp` ma non per `/var/tmp` che è la sede dove vengono posti i dati temporanei che devono sopravvivere al reboot.

SunOS offre scarsa regolazione per `tmpfs` ed il numero dei file è anche limitato dalla memoria fisica totale della macchina.

Linux non ha un equivalente a questo file system e si ritiene che `ext2fs` sia sufficientemente veloce da eliminarne la necessità.

5.6.2 `userfs`

Il file system utente (`userfs`) permette una serie di estensioni verso l'utilizzo di file system tradizionali come file system basati su FTP, compressione (`arcfs`) e veloci prototipazioni e molte altre caratteristiche. Il `docfs` è basato su questo filesystem. Controllate l'

userfs homepage `http://www.goop.org/~jeremy/userfs/`

per maggiori informazioni.

5.6.3 `devfs`

Quando i dischi vengono aggiunti, rimossi o semplicemente falliscono, è probabile che i nomi dei dispositivi dei dischi rimanenti cambieranno. Ad esempio `sdb` fallisce e quindi il vecchio `sdc` diventa `sdb`, il vecchio `sdc` diventa `sdb` e così via. Notate che in questo caso `hda`, `hdb` ecc. rimarranno invariati. Allo stesso modo se un nuovo disco viene aggiunto, può accadere il contrario.

Non c'è nessuna garanzia che lo SCSI ID 0 diventi `sda` e che aggiungere dischi aumentando l'ordine degli ID aggiungerà solamente un nuovo nome di dispositivo senza rinominare le voci precedenti, dal momento che qualche driver SCSI assegna partendo dall'ID 0 in poi mentre altri invertono l'ordine di controllo. Allo stesso modo anche l'aggiunta di un adattatore SCSI può causare rinomine.

Generalmente i nomi dei dispositivi sono assegnati nell'ordine in cui vengono trovati.

L'origine dei problemi giace nel numero limitato di bit disponibili per la numerazione principale e secondaria nei file di dispositivo usati per descrivere il dispositivo stesso. Questo lo potete verificare nella directory

`/dev`, informazioni sulla numerazione e sull'allocazione possono essere trovate con il comando `man MAKEDEV`. Attualmente ci sono 2 soluzioni a questo problema a vari stadi di sviluppo:

scsidev

funziona mediante la creazione di un database dei dischi e di cosa essi fanno parte, controllate *man scsifs* per maggiori informazioni.

devfs

è un progetto a lungo termine orientato a raggirare l'intero impiccio della numerazione dei dispositivi facendo sì che la directory `/dev` diventi un file system del kernel come lo è il `/procfs`.

5.7 Raccomandazioni sui File System

C'è una marea di scelte ma generalmente è consigliabile usare il file system generale che è presente nella vostra distribuzione. Se utilizzate `ufs` e avete disponibilità di qualcosa tipo `tmpfs` dovreste inizialmente cominciare con il file system generale per avere un'idea delle necessità di spazio e se necessario, comprare più RAM per gestire la dimensione del `tmpfs` di cui avete bisogno. Altrimenti finirete con fallimenti misteriosi e tempo perso.

Se utilizzate un dual boot e dovete trasferire dati tra i due SO, uno dei modi più semplici è utilizzare una partizione opportunamente dimensionata formattata con `fat` visto che molti sistemi possono affidabilmente leggere e scrivere su questa. Ricordatevi del limite dei 2 GB per le partizioni `fat`.

Per maggiori informazioni sulle interconnettività tra file system potete controllare la pagina del

file system `http://www.ceid.upatras.gr/~gef/fs/` .

Per evitare il tracollo totale con la rinomina dei dispositivi, se un disco fallisce, controllate l'ordine di scansione del vostro sistema e provate a mantenere il vostro sistema root su `hda` o su `sda` e supporti rimovibili cone dischi ZIP alla fine dell'ordine di scansione.

6 Technologie

Al fine di decidere come ottenere il massimo dai vostri dispositivi dovete sapere quali tecnologie sono disponibili e le loro implicazioni. Come sempre ci possono essere esigenze conflittuali riguardo la velocità, l'affidabilità, l'alimentazione, la flessibilità, la facilità di utilizzo e la complessità.

Molte delle tecniche descritte qui sotto possono essere unite in maniere diverse per aumentare la prestazione e l'affidabilità, sebbene a scapito di ulteriore complessità.

6.1 RAID

Questo è un metodo per aumentare l'affidabilità, la velocità o entrambi utilizzando più dischi in parallelo così da diminuire il tempo di accesso ed aumentare la velocità di trasferimento. Per aumentare l'affidabilità può essere usato un sistema di controllo o di mirroring. Grossi server possono trarre vantaggio da questo setup ma potrebbe essere troppo per un utente singolo a meno che non abbiate un gran numero di dischi disponibili. Guardate altri documenti e FAQ per maggiori informazioni.

Si può avere RAID in Linux o mediante software (il modulo `md` nel kernel), una scheda controller compatibile Linux (PCI-to-SCSI) o un controller SCSI-to-SCSI. Controllate la documentazione per sapere per cosa possono essere utilizzati i controller. Una soluzione hardware è generalmente più veloce, e forse anche più sicura, ma costa.

6.1.1 SCSI-to-SCSI

I controller SCSI-to-SCSI sono generalmente implementati come cabinet completi con dischi ed un controller che si connette al computer con un secondo bus SCSI. Questo fa sì che l'intero cabinet sembri un singolo grosso, veloce disco SCSI e non richiede nessun driver RAID particolare. Lo svantaggio è che il bus SCSI che connette il cabinet al computer diventa un collo di bottiglia.

Uno svantaggio significativo per la gente con grosse quantità di dischi è che c'è un limite a quante voci SCSI possono esserci nella directory `/dev`. In questi casi utilizzare lo SCSI-to-SCSI conserverà le voci.

Generalmente sono configurati attraverso il pannello frontale o con un terminale connesso alle loro interfacce seriali.

Tra i produttori di questi sistemi ci sono

CMD `http://www.cmd.com`

e

Syred `http://www.syred.com`

nelle cui pagine web sono descritti diversi sistemi.

6.1.2 PCI-to-SCSI

I controller PCI-to-SCSI sono, come suggerisce il nome, connessi al bus PCI ad alta velocità e non soffrono dello stesso collo di bottiglia come i controller SCSI-to-SCSI. Questi controller richiedono driver speciali ma potete anche capire cosa significa controllare la configurazione RAID attraverso la rete il che semplifica la gestione.

Attualmente solo poche famiglie di adattatori PCI-to-SCSI sono gestiti sotto Linux.

DPT

I più vecchi e più maturi sono una gamma di controller della *DPT* `http://www.dpt.com`

incluse le famiglie di controller SmartCache I/III/IV e SmartRAID I/III/IV. Questi controller sono supportati dai driver EATA-DMA presenti nel kernel standard. Questa compagnia ha inoltre una

home page `http://www.dpt.com`

informativa che descrive i vari aspetti generali del RAID e dello SCSI oltre alle informazioni dei prodotti.

Maggiori informazioni dall'autore dei driver dei controller DPT (driver EATA* possono essere trovate sulla sua pagina presso

SCSI `http://www.uni-mainz.de/~neuffer/scsi`

e

DPT `http://www.uni-mainz.de/~neuffer/scsi/dpt` .

Questi non sono i più veloci ma hanno una affidabilità più che provata.

Notate che gli strumenti di gestione per i controller DTP attualmente girano solamente sotto DOS/Win, così avrete bisogno di una piccola partizione Dos/Win per un po' di software. Ciò significa pure che dovrete lanciare il sistema con Windows per fare manutenzione del vostro sistema RAID.

ICP-Vortex

Una recentissima aggiunta è un set di controller della

ICP-Vortex `http://www.icp-vortex.com` con la caratteristica di avere fino a 5 canali indipendenti e hardware molto veloce basati sul chip i960. Il driver per Linux è stato scritto dalla compagnia stessa il che dimostra che supportano Linux.

Visto che ICP-Vortex fornisce il software di manutenzione per Linux, non è necessario fare un boot verso altri sistemi operativi per il setup e la manutenzione del vostro sistema RAID. Questo vi fa risparmiare tempo.

Mylex DAC-960

Questo è una delle ultime uscite ed è in beta. Maggiori informazioni come anche i driver sono disponibili presso la

Dandelion Digital's Linux DAC960 Page `http://www.dandelion.com/Linux/DAC960.html` .

Compaq Smart-2 PCI Disk Array Controllers

Un'altra entrata recentissima e attualmente in versione beta è il driver

Smart-2 `http://www.insync.net~frantze/cpqarray.html` .

6.1.3 Software RAID

Qualche sistema operativo offre software RAID su dischi comuni e controller. Il costo è basso e le prestazioni per il crudo IO su disco possono essere molto alte.

Visto che può richiedere molte risorse della CPU, aumenta il carico evidente quindi se la macchina è limitata nelle prestazioni della CPU piuttosto che nelle prestazioni dell'IO, sarebbe meglio per voi risolvere con un controller hardware PCI-to-RAID.

Il costo reale, le prestazioni e specialmente l'affidabilità del RAID software contro quello hardware sono un argomento molto controverso. L'affidabilità su sistemi Linux è stata finora molto buona.

L'attuale progetto del software RAID per Linux, è il sistema `md` (multiple devices) che offre molto più del RAID, quindi è descritto maggiormente più avanti.

6.1.4 Livelli di RAID

Il RAID lo troviamo con molti livelli e sapori su cui farò una breve descrizione qui. Molto è stato scritto riguardo questo ed il lettore interesato è invitato a leggere di più sulle FAQ del RAID.

- RAID *0* non è affatto ridondante ma offre la più alta velocità di trasferimento di tutti i livelli. I dati sfilano attraverso un bel numero di dischi quindi le operazioni di lettura e scrittura avvengono in parallelo attraverso tutti i dischi. D'altra parte se un singolo disco fallisce, allora tutto è perduto. Avevo parlato di backup?
- RAID *1* è il metodo più primitivo per ottenere ridondanza duplicando i dati attraverso tutti i drive. Naturalmente questo è un metodo che spreca molto ma ottenete un vantaggio che è l'accesso rapido. Il disco che accede ai dati per primo vince. I trasferimenti non sono più veloci che per dischi singoli, sebbene potreste ottenere trasferimenti di lettura maggiori utilizzando una lettura di traccia per disco.

  Inoltre se avete solo due dischi, questo è l'unico modo per raggiungere la ridondanza.
- RAID *2* e *4* non sono così comuni e non ci sono qui.
- RAID *3* utilizza un certo numeri di dischi (almeno 2) per immagazzinare i dati come nella modalità RAID 0. Esso utilizza inoltre un disco di ridondanza addizionale per archiviare la somma XOR dei dati dai dischi contenenti i dati. Dovesse fallire il disco di ridondanza, il disco può continuare ad operare come se niente fosse accaduto. Se un disco contenente i dati dovesse fallire, il sistema può determinare

i dati su questo disco dall'informazione sul disco di ridondanza e sui dischi rimanenti. Qualsiasi doppio fallo porterà alla disattivazione del RAID.

RAID 3 ha senso solo con almeno 2 dischi di dati (3 dischi incluso il disco di ridondanza). Teoricamente non c'è limite per il numero di dischi nell'insieme, ma la probabilità di un errore aumenta con il numero di dischi nell'insieme RAID. Generalmente il limite superiore è dai 5 ai 7 dischi in un singolo insieme RAID.

Visto che RAID 3 immagazzina tutte le informazioni ridondanti su un disco dedicato e visto che questa informazione deve essere aggiornata qualora ci sia un'operazione di scrittura su qualsiasi disco di dati, la velocità generale di scrittura dell'insieme RAID 3 è limitata dalla velocità di scrittura del disco di ridondanza. Anche questo è un limite per il numero di dischi in un insieme RAID. La velocità generale di lettura di un insieme RAID 3 con tutti i dischi di dati attivati, è quella di un insieme RAID 0 con quel numero di dischi. Se l'insieme deve ricostruire i dati immagazzinati su un disco fallito da un'informazione ridondante, la prestazione sarà severamente limitata: tutti i dischi nell'insieme devono essere letti e dev'essere fatto l'XOR dai dati per calcolare le informazioni rimanenti.

- RAID *5* è come il RAID 3, ma l'informazione ridondante è sparsa su tutti i dischi dell'insieme RAID. Questo aumenta la prestazione in scrittura, perché il carico è distribuito più omogeneamente tra tutti i dsichi disponibili.

Ci sono anche ibridi basati su RAID 0 o 1 e un altro livello. Molte combinazioni sono possibili ma ne ho viste solamente poche. Questi sono più complessi dei livelli RAID sopra menzionati.

RAID *0/1* combina lo spandimento con la duplicazione, il che conferisce trasferimenti molto veloci insieme ad accessi veloci come anche ridondanza. Lo svantaggio è il grosso consumo di disco come anche la complessità sopra descritta.

RAID *1/5* combina la velocità e i benefici della ridondanza del RAID 5 con l'accesso veloce del RAID 1. La ridondanza è migliorata se paragonata al RAID 0/1 ma il consumo di disco è ancora importante. Implementare questo sistema vuol dire utilizzare in genere più di 6 dischi, forse anche diversi controller o canali SCSI.

6.2 Gestione dei Volumi

La gestione dei volumi è un modo per superare le costrizioni delle partizioni e dei dischi a dimensione fissa mantenendo sempre il controllo di dove le varie parti dei file risiedono. Con un sistema del genere potete aggiungere nuovi dischi al vostro sistema e aggiungere spazio da questo disco alle parti dello spazio file dove è necessario, come anche migrare i dati fuori da un disco che ha dei problemi ad altri dischi prima che avvenga un fallimento catastrofico.

Il sistema sviluppato da

Veritas `http://www.veritas.com`

è diventato lo standard di fatto per la gestione logica dei volumi.

La gestione dei volumi è una delle cose in cui attulamente Linux è carente.

Uno è il progetto del sistema a partizioni virtuali

VPS `http://www.uiuc.edu/ph/www/roth`

che reimplementerà molte delle funzioni di gestione trovate nel sistema AIX di IBM. Sfortunatamente questo progetto è fermo.

Un altro progetto è il

Logical Volume Manager `http://linux.msede.com/lvm/`

che è simile ad un progetto della HP.

6.3 La patch `md` per il kernel Linux

Il Linux Multi Disk (md) fornisce un certo numero di caratteristiche di livelli di blocco in vari stadi di sviluppo.

RAID 0 (spargimento) e concatenamento sono molto solidi anche nella qualità di produzione e anche i RAID 4 e 5 sono molto maturi.

È inoltre possibile ammucchiare qualche livello, ad esempio fare il mirror (RAID 1) di due paia di dischi, ogni paio costruito come dischi sparsi (RAID 0), il che offre la velocità del RAID 0 con l'affidabilità del RAID 1.

Oltre al RAID questo sistema offre (allo stadio alfa) la gestione dei volumi a livelli di blocco e presto anche spazio file traslucido. Dal momento che questo è fatto a livello di blocco, esso può essere utilizzato in combinazione con qualsiasi file system, anche per la `fat` utilizzando Wine.

Pensate attentamente quali dischi combinate così potete agire su tutti i dischi in parallelo, il che vi fornisce migliori prestazioni e meno usura. Leggete di più riguardo a ciò nella documentazione che accompagna `md`.

Sfortunatamente la documentazione è piuttosto vecchia ed in parte fuorviante e si riferisce solamente alla versione 0.35 di `md` che usa un setup di vecchio stile. Il nuovo sistema è molto differente e presto sarà rilasciato come versione 1.0 ma è attualmente non documentato. Se voleste provarlo, dovreste seguire la mailing list `linux-raid`.

La documentazione sta migliorando e il

Software RAID HOWTO `http://ostenfeld.dk/~jakob/Software-RAID.HOWTO/`

è in lavorazione.

Suggerimento: se non riuscite a farlo funzionare correttamente, avete dimenticato di impostare il flag `persistent-block`. La vostra documentazione migliore è attualmente il codice sorgente.

6.4 Compressione

La compressione del disco contro la compressione dei file è un dibattito caldo specialmente riguardo il rischio aggiunto di corruzione dei file. Nonostante ciò ci sono diverse opzioni disponibili per gli amministratori avventurosi. Questi intraprendono molte forme, dai moduli del kernel ed le patch alle librerie extra, ma notate che molti soffrono di diverse forme di limitazione quali essere di sola lettura. Visto che lo sviluppo va così rapidamente, sicuramente le specifiche saranno cambiate nel momento in cui leggete questo. Come sempre: controllate da soli gli aggiornamenti. Qui sono elencati solo pochi riferimenti.

- Compressione file dalle doppie caratteristiche con alcune limitazioni.
- Zlibc aggiunge una decompressione in tempo reale e trasparente dei file mentre li carica.
- ci sono molti moduli disponibili per leggere i file compressi o le partizioni che sono native ad altri svariati sistemi operativi sebbene attualmente la maggior parte di questi è di sola lettura.
- *dmsdos* `http://bf9nt.uni-duisburg.de/mitarbeiter/gockel/software/dmsdos/`

  (attualmente alla versione 0.9.2.0) offre molto della compressione disponibile per il DOS e per Windows. Non è ancora completa ma il lavoro sta andando avanti e nuove caratteristiche aggiunte regolarmente.
- `e2compr` è un pacchetto che implementa l'`ext2fs` con le capacità di compressione. È ancora in fase di test e sarà principalmente di interesse per gli hacker del kernel ma dovrebbe acquisire ben presto la stabilità necessaria per un uso più vasto. Controllate l'

homepage e2compr `http://netspace.net.au/~reiter/e2compr.html`

per maggiori informazioni. Ho saputo della velocità e della buona stabilità, ecco perché è menzionato qui.

6.5 ACL

L'Access Control List (ACL) offre un controllo più fine sull'accesso ai file di un utente sulle basi dell'utente stesso piuttosto che il tradizionale owner, group and others, come si vede nel listato della directory (`drwxr-xr-x`). Ciò non è ancora disponibile in Linux ma ci si aspetta che lo sarà nel kernel 2.3 visto che gli agganci sono già posizionati nell'`ext2fs`.

6.6 `cachefs`

Questo utilizza una parte del disco rigido per mettere in cache supporti più lenti come i CD-ROM. è disponibile per SunOS ma non ancora per Linux.

6.7 File System Traslucidi o Nascosti

Questo è un sistema copy-on-write dove le scritture finiscono su un sistema differente dalla reale origine facendolo sembrare uno spazio file ordinario. Quindi visto che lo spazio file nasconde i dati originali ed il traslucente li riscrive indietro, il buffer può essere privato per ogni utente.

C'è un certo numero di applicazioni:

- aggiornare un file system vivo su un CD-ROM, rendendolo flessibile, veloce e che risparmi anche spazio,
- uno scheletro dei file originale per ciascun nuovo utente, risparmiando spazio visto che i dati originali sono mantenuti in uno spazio singolo e condivisi,
- progetto di sviluppo parallelo prototipizzando il fatto che ogni utente possa modificare il sistema globalmente non intaccando gli altri utenti.

SunOS offre questa caratteristica ed è in fase di sviluppo per Linux. C'era un vecchio progetto chiamato Inheriting File Systems (`ifs`) ma questo progetto si è fermato. Un progetto attuale è parte del sistema `md` e offre traslucenza del livello di blocco così può essere applicato a qualsiasi file system.

La Sun ha una

pagina `http://www.sun.ca/white-papers/tfs.html`

informativa sul file system traslucido.

6.8 Posizionamento Fisico delle Tracce

Questo trucco era molto importante quando i dischi erano lenti e piccoli e qualche file system era solito considerare le caratteristiche variabili nel posizionare i file. Oltre alla velocità generale più alta, la cache integrata nei controller e nei dischi ha ridotto questo effetto.

Comunque c'è ancora qualcosa da guadagnare anche oggi. Da quel che sappiamo, *il dominio del mondo* è a portata di mano, ma per raggiungerlo *velocemente* dobbiamo impiegare tutti i trucchi che possiamo .

Per capire la strategia dobbiamo richiamare questo vecchio pezzo di conoscenza e le proprietà delle varie localizzazioni delle tracce. Questo è basato sul fatto che le velocità di trasferimento generalmente aumentano per le tracce più ci si allontana dal pignone, come anche per il fatto che è più veloce accedere verso o dalle tracce centrali che verso o dalle tracce più interne o più esterne.

La maggior parte delle unità utilizzano dischi che girano con velocità angolare costante ma utilizzano una densità di dati (abbastanza) costante attraverso tutte le tracce. Questo vuol dire che otterrete una più alta velocità di trasferimento sulle tracce esterne che su quelle più interne; una carateristica che è molto buona per le necessità di grosse librerie.

I dischi più recenti utilizzano una geometria logica di mappatura che differisce dall'attuale mappatura fisica con la quale sono mappati trasparentemente dal disco stesso. Questo fa sì che la stima delle tracce di mezzo sia più difficile.

Nella maggior parte dei casi la traccia 0 è nella traccia più esterna e questo è lo standard per molte persone. Comunque, si dovrebbe ricordare che non ci sono garanzie che sia sempre così.

Le tracce più interne

sono generalmente lente in trasferimento e anche il fatto di giacere alla fine della posizione di accesso ne rende lento l'accesso.

Sono quindi più adatte per le directory low end come DOS, root e gli spool di stampa.

Le tracce intermedie

sono mediamente più veloci riguardo al trasferimento rispetto alle tracce più interne e essendo in mezzo vi si accede più velocemente.

Questa caratteristica è ideale per le parti che richiedono di più quali `swap`, `/tmp` e `/var/tmp`.

Le tracce esterne

hanno di media caratteristiche di trasferimento più veloci ma, come le tracce interne, sono collocate alla fine dell'accesso; quindi, statisticamente, accedervi è ugualmente lento che per le tracce interne.

Grandi file come le librerie beneficerebbero del posizionamento in questa sede.

Quindi la riduzione del tempo di accesso può essere raggiunta posizionando le tracce con accesso frequente in mezzo così che la distanza media di accesso e di conseguenza il tempo di accesso possa essere breve. Questo può essere fatto sia utilizzando `fdisk` o `cfdisk` per fare una partizione nelle tracce di mezzo o facendo prima un file (utilizzando `dd`) uguale alla metà della grandezza dell'intero disco prima di creare i file cui si ha accesso di frequente, dopo di che il file fittizio può essere cancellato. In entrambe i casi si assume che si inizi da un disco vuoto.

L'ultimo trucco è adatto per gli spool delle news dove la struttura delle directory vuote può essere posizionata nel mezzo prima di metterci i file di dati. Questo aiuta inoltre un po' la frammentazione.

Questo piccolo trucco può essere utilizzato sia sui dischi ordinari sia per i sistemi RAID. Nell'ultimo caso il calcolo per mettere al centro le tracce sarà differente, se possibile. Consultate l'ultimo manuale RAID.

La differenza di velocità sta al drive ma un 50 per cento di miglioramento è un valore comune.

6.8.1 Valori di velocità del disco

Lo stesso assemblaggo della testata (HDA) è spesso disponibile su un bel numero di interfacce (IDE, SCSI, ecc.) ed i parametri meccanici sono quindi paragonabili. Oggigiorno la meccanica è spesso il fattore limitante ma lo sviluppo sta migliorando le cose rapidamente. Ci sono due parametri principali, generalmente riportati in millisecondi (ms):

- Movimento delle testine - la velocità alla quale la testina di lettura e scrittura è capace di muoversi da una traccia a quella successiva, chiamata tempo di accesso. Se fate i calcoli e integrate l'accesso prima attraverso tutte le possibili tracce di partenza e poi attraverso tutte le possibili tracce di destinazione, troverete che ciò è uguale ad un colpo (stroke) attraverso un terzo di tutte le tracce.
- Velocità rotazionale - che determina il tempo necessario per arrivare al settore giusto, chiamato latenza.

Dopo che le bobine audio hanno rimpiazzato i motori passo passo per il controllo del movimento delle testine, i miglioramenti sembrarono essersi livellati e maggiore energia è oggi impiegata (letteralmente) per migliorare la velocità rotazionale. Questo ha il beneficio secondario di aumentare anche le velocità di trasferimento.

Qualche valore tipico:

Tipo di Disco

Tempo di acceso (ms)	Veloce	Tipico	Vecchio
Traccia-a-traccia	<1	2	8
Tempo medio di accesso	10	15	30
Fine-a-fine	10	30	70

Ciò dimostra che gli ultimissimi dischi, offrono solo marginalmente un miglior tempo di accesso della media dei dischi ma che i vecchi dischi basati su motori passo passo sono decisamente peggiori.

Velocita di rotazione	3600	4500	4800	5400	7200	10000
Latenza (ms)	17	13	12.5	11.1	8.3	6.0

Dal momento che la latenza è il tempo medio per raggiungere un settore particolare, la formula è sufficientemente semplice

```
latenza (ms) = 60000 / velocita (RPM)
```

Ovviamente anche questo è un esempio della diminuzione dei compensi per gli sforzi messi nello sviluppo. Comunque, ciò che veramente importa qui è il consumo di elettricità, il calore ed il rumore.

6.9 Sovrapposizione di strati RAID

Uno dei vantaggi di un progetto a strati di un sistema operativo è che si ha a disposizione la flessibilità di mettere insieme i pezzi in un gran numero di modi. Ad esempio potete mettere in cache un CD-ROM con `cachefs` che è un volume suddiviso tra due dischi. Questo a turno può essere organizzato in maniera traslucente con un volume che è montato da un'altra macchina via NFS. RAID può essere raggruppato in svariati strati per offrire accesso molto veloce e trasferimento in una maniera tale che potrebbe funzionare se anche fallissero 3 dischi. Le scelte sono tante, limitate solamente dall'immaginazione e, cosa probabilmente più importante, dal denaro.

6.10 Raccomandazioni

C'è un numero praticamente infinito di combinazioni disponibili ma la mia raccomandazione è di iniziare con un'installazione semplice senza alcuna aggiunta immaginosa. Capite cosa vi serve, dove è richiesta la massima prestazione, se il collo di bottiglia è il tempo di accesso o la velocità di trasferimento, e così via. Poi affrontate ogni componente a turno. Se voi potete raggruppare abbastanza liberamente, dovreste essere capaci di rimpiazzare la maggior parte dei componenti con poche difficoltà.

RAID è generalmente una buona idea ma assicuratevi di avere una buona padronanza della tecnologia ed un solido sistema di back up.

7 Altri Sistemi Operativi

Molti utenti Linux hanno diversi sistemi operativi installati, spesso necessari a causa di sistemi di predisposizione dell'hardware che girano sotto altri sistemi operativi, In particolare il DOS o qualche tipo di Windows. Una piccola sezione su come avere a che fare al meglio con questi è qui inclusa.

7.1 DOS

Mantenendo in disparte il dibattito sul fatto che il DOS si qualifichi o meno come sistema operativo, uno potrebbe dire che è molto poco sofisticato riguardo alle operazioni del disco. La più grave conseguenza di questo fatto è che ci possono essere serie difficoltà facendo girare varie versioni del DOS su dischi larghi, e voi siete quindi fortemente invitati a leggere il *Large Drives mini-HOWTO*.

Un effetto collaterale di ciò è che spesso si propende a mettere il DOS su numeri di traccia bassi.

Essendo stato progettato per piccoli dischi, ha un file system alquanto poco sofisticato (`fat`) e quando usato su dischi grossi alloca blocchi di dimensioni enormi. Inoltre, causa la frammentazione dei blocchi, che dopo un po' causerà accessi spropositati e rallenterà i trasferimenti effettivi.

Una soluzione a questo è fare una deframmentazione regolarmente ma è fortemente raccomandato fare un backup prima di deframmentare. Tutte le versioni del DOS hanno il `chkdsk` che può fare un po' di controllo del disco, le nuove versioni hanno anche lo `scandisk` che è in qualche modo migliore. Ci sono molti programmi di deframmentazione disponibili, qualche versione ne ha uno chiamato `defrag`. Le Norton Utilities hanno un grosso insieme di attrezzi per il disco e ce ne sono anche molti altri.

Come sempre ci sono imprevisti, e questo particolare serpente nel nostro paradiso del disco è chimato *file nascosti*. Qualche rivenditore cominciò ad usarli per schemi di protezione delle copie perché non avrebbero reagito bene nell'essere spostati in un'altra parte sul disco, anche se rimanevano nello stesso posto nella struttura della directory. Il risultato di questo fu che i programmi di deframmentazione non toccavano alcun file nascosto, che a lungo andare riduceva l'effetto della deframmentazione.

Essendo un sistema operativo mono tasking, mono threading e mono molte altre cose, ci sono molti pochi vantaggi nell'usare dischi multipli se almeno non utilizzate un controller con un qualsiasi supporto RAID integrato.

Ci sono un po' di utilità chiamate `join` e `subst` che possono eseguire una configurazione di dischi multipli ma c'è un guadagno molto piccolo da questo per l'enorme lavoro da fare. Qualcuno di questi programmi è stato rimosso nelle nuove versioni.

Alla fine c'è molto poco che voi potete fare, ma niente è perduto. Molti programmi necessitano di archiviazione veloce, temporanea e quelli che si comportano bene cercheranno variabili d'ambiente chiamate `TMPDIR` o `TEMPDIR` che potete predisporre per farle puntare ad un altro disco. Questo è di solito fatto nell'`autoexec.bat`.

```
SET TMPDIR=E:/TMP
SET TEMPDIR=E:/TEMP
```

Non solo vi potrà far guadagnare più velocità ma può anche ridurre la frammentazione.

Ci sono state delle affermazioni riguardo la difficoltà nel rimuovere partizioni primarie multiple utilizzando il programma `fdisk` che accompagna il DOS. Se dovesse capitare potete utilizzare un disco di recupero Linux con l'`fdisk` di Linux per riparare il sistema.

Non dimenticate che ci sono altre alternative al DOS, le più note sono il

DR-DOS `http://www.caldera/dos/`

della

Caldera `http://www.caldera/` . Questo è il diretto discendente del DR-DOS della Digital Research. Esso offre molte caratteristiche non trovate nel più comune DOS, come il multi tasking ed i nomi di file lunghi.

Un'altra alternativa, che è pure libera, è

Free DOS `http://www.freedos.org/`

che è un progetto in sviluppo. Un numero di utilità libere sono disponibili.

7.2 Windows

La maggior parte dei punti espressi qui sopra sono validi anche per Windows, con l'eccezione di Windows95 che apparentemente ha una migliore gestione del disco, che trarrà migliori prestazioni dai dischi SCSI.

Una cosa utile è l'introduzione dei nomi di file lunghi, per leggerli da Linux dovrete avere il file system `vfat` per montare queste partizioni.

La frammentazione del disco è ancora un problema. Un po' di questa può essere evitata facendo una deframmentazione immediatamente prima e immediatamente dopo l'installazione di grossi programmi o sistemi. Uso questa procedura al lavoro e mi sono accorto che funziona abbastanza bene. Eliminare i file inutilizzati e svuotare il cestino prima di tutto può migliorare la deframmentazione ancora di più.

Anche Windows utilizza i dischi swap, e reindirizzare ciò verso un altro disco può concedervi guadagni di prestazione. Ci sono diversi mini-HOWTO che vi dicono come condividere al meglio lo spazio di swap tra vari sistemi operativi.

Il trucco di organizzare la `TEMPDIR` può ancora essere utilizzato ma non tutti i programmi soddisferanno questa disposizione. Alcuni comunque lo fanno. Per avere una buona visione delle disposizioni nei file di controllo, potete lanciare `sysedit` che aprirà un bel numero di file da editare, uno dei quali è l'`autoexec` dove potete aggiungere i settaggi della `TEMPDIR`.

Molti dei file temporanei sono messi nella directory `/windows/temp` e cambiare questo è più arduo. Per fare questo potete utilizzare `regedit` che è abbastanza potente e capace di mettere il vostro sistema in uno stato che non gradireste o, più precisamente, in uno stato meno gradevole di Windows in generale. Registry database error è un messaggio che significa seriamente cattive notizie. Inoltre vedrete che molti programmi hanno la propria directory temporanea sparsa nel sistema.

Predisporre il file di swap su una partizione separata è un'idea migliore e molto meno rischiosa. Ricordatevi che questa partizione non può essere utilizzata per niente altro, anche se dovesse sembrare che c'è spazio residuo.

Ora è possibile leggere le partizioni `ext2fs` da Windows, anche montando le partizioni utilizzando

FSDEXT2 `http://www.yipton.demon.co.uk/`

o utilizzando un esploratore di file chiamato

Explore2fs `http://uranus.it.swin.edu.au/~jn/linux/Explore2fs.html` .

7.3 OS/2

L'unica nota speciale qui è che potete ottenere un driver del file system per OS/2 che può leggere una partizione `ext2fs`.

7.4 NT

Questo è un sistema più serio caratterizzato da termini di gran moda noti al marketing. È bene notare che può fare lo striping e molte altre sofisticate predisposizioni. Notate il drive manager nel pannello di controllo. Non ho facile accesso ad NT, maggiori dettagli su questo possono necessitare un po' di tempo.

Un intoppo importante è stato riportato recentemente da acahalan at cs.uml.edu : (riformattato da un messaggio Usenet)

Il DiskManager di NT ha un bug serio che può corrompere il vostro disco quando avete diverse (più di una?) partizioni estese. La Microsoft rilascia un programma per correggere questo problema sul proprio sito. Vedete la

knowledge base `http://www.microsoft.com/kb/`

per saperne di più (questo tocca gli utenti Linux, perché hanno partizioni extra).

Ora potete leggere le partizioni `ext2fs` da NT mediante

Explore2fs `http://uranus.it.swin.edu.au/~jn/linux/Explore2fs.html` .

7.5 Sun OS

C'è un po' di confusione in quest'area tra Sun OS contro Solaris. In maniera molto concisa, Solaris non è altro che Sun OS 5.x confezionato con Openwindows e poche altre cose. Se eseguite Solaris, vi basta scrivere `uname -a` per vedere la vostra versione. Parte della ragione di questa confusione è che la Sun Microsystems soleva utilizzare un SO proveniente dalla famiglia BSD, sebbene con un po' di pezzi da altre parti come anche con cose fatte da loro. Questa è stata la situazione fino al Sun OS 4.x.y quando presero una decisione strategica della pianificazione e decisero di passare oltre lo Unix ufficiale, il System V, la Versione 4 (aka SVR5), e fu creato il SO Sun 5. Questo rese scontenti molti. Inoltre venne unito ad altre cose e commercializzato sotto il nome di Solaris, che attualmente è alla versione 7 che proprio da poco ha rimpiazzato la versione 2.6 essendo l'ultima e la migliore. A differenza del grosso salto nel numero della versione c'è stato attualmente un piccolo miglioramento ma un salto enorme per il marketing.

7.5.1 Sun OS 4

Questo è abbastanza familiare alla maggior parte degli utenti Linux. L'ultima versione è la 4.1.4 più varie patch. Notate comunque che la struttura del file system è abbastanza differente e non è conforme al FSSTND quindi ogni pianificazione deve essere fatta sulla struttura standard. Potete ottenere qualche informazione su questo dalle pagine man: `man hier`. Questo è, come molte pagine man, abbastanza conciso ma dovrebbe essere un buon inizio. Se siete ancora confusi dalla struttura almeno sarà ad un livello più alto.

7.5.2 Sun OS 5 (aka Solaris)

Si presenta con un elegante sistema di installazione che gira sotto Openwindows, che vi aiuterà a partizionare ed a formattare i dischi prima dell'installazione del sistema da CD-ROM. Fallirà anche se il vostro setup dei dischi è troppo vasto e dal momento che serve un'intera sessione di installazione da un CD-ROM pieno, in un'unità 1x questo fallimento vi piomberà addosso dopo troppo tempo. Questa è l'esperienza che abbiamo fatto quando lo usavo al lavoro. Quindi installavamo tutto su un solo disco e poi spostavamo le directory tra i dischi.

Le impostazioni abituali sono sensibili alla maggior parte delle cose, ma rimane ancora una piccola controversia: i dischi swap. Sebbene il manuale ufficiale raccomanda dischi di swap multipli (che sono utilizzati in maniera simile a Linux) abitualmente si utilizza un disco solo. Si raccomanda di cambiare ciò il prima possibile.

Sun OS 5 offre inoltre un file system progettato in maniera specifica per i file temporanei, `tmpfs`. Offre un miglioramento significativo della velocità rispetto a `ufs` ma non c'è sopravvivenza al riavvio.

L'unico commento quindi è: state attenti! Sotto Solaris 2.0 sembra che creare file troppo grandi in `/tmp` possa causare una trappola con kernel panic da saturazione dello spazio swap. Ciò che risulta da ciò che accade è la perdita di ogni dato su un RAMdisk dopo avere spento e quindi risulta difficile capire ciò che accade dopo avere spento. Ciò che è peggio, sembra che i processi dello spazio utente possano causare questo kernel panic e sebbene questo problema sia molto a cuore, è meglio non utilizzare `tmpfs` in ambienti potenzialmente ostili.

Controllate anche le note su 5.6.1 (tmpfs).

Trivia: C'è anche un film chiamato Solaris, un film di fantascienza che è molto, molto lungo, lento ed incomprensibile. Questo è spesso stato segnalato quando Solaris (il SO) apparve...

7.5.3 BeOS

Questo sistema operativo è uno dei più recenti ad arrivare e ha la caratteristica di avere un file system che ha uno stampo ad archivio.

C'è un driver per il file system BFS che stanno sviluppando per Linux ed è disponibile in versione alpha. Per maggiori informazioni controllate la

Linux BFS page `http://hp.vector.co.jp/authors/VA008030/bfs`

dove sono disponibili anche le patch.

8 Cluster

In questa sezione accennerò brevemente al modo in cui le macchine possono essere connesse tra loro ma questo è un argomento molto ampio che potrebbe essere oggetto di un HOWTO separato, suggerimento, suggerimento. Inoltre parlando concisamente, questa sezione giace fuori dallo scopo di questo HOWTO, quindi se avete voglia di gloria, *voi* potreste contattarmi e accollarvi questa parte e riversarla su un nuovo documento.

In questi giorni i computer diventano datati con una frequenza eccezionale. Non c'è comunque ragione per non fare un buon uso del vecchio hardware con Linux. Utilizzando un computer vecchio e fuori moda come un server di rete può essere sia utile nella propria essenza che come un buon esercizio didattico. Questo gruppo di computer collegati in rete può assumere diverse forme ma per rimanere nell'argomento di questo HOWTO mi limiterò alle strategie dei dischi. Nonostante ciò spererei che qualcun altro si accollasse questo argomento e lo riversasse in un proprio documento.

Questa è un'area eccitante di attività oggi, e molte forme di raggruppamento sono disponibili oggi, spaziando dal bilanciamento del carico di lavoro automatico della rete locale ad un hardware più esotico come la Scalable Coherent Interface (SCI) che fornisce una stretta integrazione di macchine, covertendole effettivamente in una macchina unica. Vari tipi di raggruppamenti sono stati disponibili per macchine più grandi per qualche tempo e il gruppo VAXcluster è forse un esempio ben conosciuto di ciò. Il raggruppamento è fatto di solito al fine di condividere le risorse quali dischi, stampanti, terminali, ecc., ma anche per elaborare risorse ugualmente in maniera trasparente tra i nodi computazionali.

Non c'è una definizione universale di raggruppamento, qui è inteso per definire una rete di macchine che combina le proprie risorse per servire gli utenti. Ammetto che questa è una definizione poco valida ma ciò cambierà più tardi.

In questi giorni inoltre Linux offre caratteristiche di raggruppamento ma all'inizio descriverò solamente una semplice rete locale. È un buon modo per riutilizzare il vecchio ed altrimenti inutilizzabile hardware, fino a che possono permettere a Linux o a qualcosa del genere di girare.

Uno dei modi migliori di utilizzare una vecchia macchina è un server di rete nel caso in cui la velocità effettiva sia verosimilmente limitata dall'ampiezza di banda piuttosto che dalla pura prestazione computazionale. Per utilizzo domestico potreste spostare la seguente funzionalità presso una macchina più vecchia utilizzata come server:

- news
- posta
- proxy web
- server di stampa
- modem server (PPP, SLIP, FAX, posta Vocale)

Potete anche montare con `NFS` i dischi sulla vostra stazione operativa riducendo le necessità di spazio disco. Comunque leggete il FSSTND per vedere quali directory *non* dovrebbero essere esportate. I miglior candidati per esportare tutte le macchine sono `/usr` e `/var/spool` e possibilmente `/usr/local` ma probabilmente non `/var/spool/lpd`.

La maggior parte delle volte anche i dischi lenti, fornirebbero una prestazione sufficiente. D'altro canto, se processate direttamente sui dischi del server o avete un sistema di rete molto veloce, potreste volere ripensare alla vostra strategia ed utilizzare dischi più veloci. Le caratteristiche di ricerca su un server web o ricerca su un archivio di news ne sono esempi.

Una rete del genere può essere un modo eccellente di imparare l'amministrazione di rete e mettere su la propria rete tostapane, come è spesso chiamata. Potete ottenere maggiori informazioni su questo in altri HOWTO ma ci sono due cose importanti che dovreste tenere a mente:

- Non tirate fuori dal nulla i numeri IP. Configurate la vostra rete interna utilizzando numeri IP riservati per utilizzo privato, es. utilizzate il vostro server di rete come router che gestisca questo mascheramento di IP.
- Ricordate che se configurate in aggiunta il router come un firewall, potreste non essere in grado di accedere ai vostri dati dal di fuori, a seconda della configurazione del firewall.

La rete *nyx* fornisce un esempio di gruppo nel senso definito qui. Esso consiste delle seguenti macchine:

nyx

è una delle due macchine per il login utente e fornisce qualche servizio di rete.

nox

(aka nyx10) è la macchina di login utente principale ed è anche il server di posta.

noc

è un server dedicato alle news. Lo spool delle news è reso accessibile attraverso NFS montando verso nyx e nox.

arachne

(aka www) è il server web. Le pagine Web sono scritte montando NFS su nox.

Ci sono inoltre altri progetti di raggruppamento più avanzato in atto

- *Il progetto Beowulf* `http://cesdis.gsfc.nasa.gov/linux/beowulf/beowulf.html`
- *The Genoa Active Message Machine (GAMMA)* `http://www.disi.unige.it/project/gamma/`

Il raggruppamento di alta tecnologia richiede interconnesioni ad alta tecnologia e SCI è uno di essi. Per scoprirne di più potete sia cercare sull'home page della

Dolphin Interconnect Solutions `http://www.dolphinics.no/`

che è uno degli attori principali in questo campo, o potete vedere

scizzl `http://www.scizzl.com/` .

I server di posta centralizzati che utilizzano IMAP stanno diventando sempre più popolari visto che i dischi diventano grandi abbastanza da contenere tutta la posta archiviata indefinitamente ed inoltre sufficientemente economici da renderla un'opzione realizzabile. Sfortunatamente è diventato chiaro che montare attraverso `NFS` la posta da un'altra macchina può causare la corruzione del database IMAP visto che il software del server non gestisce le interruzioni dell'NFS troppo bene e le interruzioni di NFS sono un'evenienza abbastanza comune. Mantenete in ogni caso l'archivio della posta localmente rispetto al server IMAP.

9 Punti di Montaggio

Nel progettare la struttura del disco è importante non suddividere in maniera errata la struttura dell'albero delle directory, ecco il perché di questa sezione. Dal momento che è altamente dipendente dal FSSTND è stato messo da parte in una sezione separata, e probabilmente andrà forse completamente riscritta quando FHS sarà adottato in una distribuzione Linux. Nel frattempo lo farà questa.

Ricordate che questa è una lista di dove una separazione *può* avvenire, non dove *deve* essere. Come sempre, è sempre richiesto un giudizio coscienzioso.

Ancora una volta qui sarà data un'indicazione rozza. I valori indicano

```
0=non separare qui
1=sconsigliato
 ...
4=utile
5=consigliato
```

Al fine di mantenere ancora la lista, le parti non interessanti sono state rimosse.

```
Directory   Appropriatezza
/
|
+-bin       0
```

```
+-boot      0
+-dev       0
+-etc       0
+-home      5
+-lib       0
+-mnt       0
+-proc      0
+-root      0
+-sbin      0
+-tmp       5
+-usr       5
| \
| +-X11R6     3
| +-bin       3
| +-lib       4
| +-local     4
| | \
| | +bin        2
| | +lib        4
| +-src       3
|
+-var       5
  \
  +-adm       0
  +-lib       2
  +-lock      1
  +-log       0
  +-preserve  1
  +-run       1
  +-spool     4
  | \
  | +-mail      3
  | +-mqueue    3
  | +-news      5
  | +-smail     3
  | +-uucp      3
  +-tmp       5
```

C'è ovviamente una marea di aggiustamenti possibili, ad esempio ad un utente casalingo non importerà la separazione della gerarchia `/var/spool` ma ad un ISP serio dovrebbe. La soluzione qui è *l'utilizzo*.

QUIZ! Perché `/etc` non dovrebbe mai risiedere su una partizione separata? Risposta: Le istruzioni di montaggio durante l'avvio si trovano nel file `/etc/fstab` quindi se questo è su una partizione separata e non montata è come se una chiave di un cassetto chiuso sia nel cassetto stesso (sì, non faccio praticamente nulla per ravvivare questo HOWTO).

10 Considerazioni e Dimensionamento

Il punto di inizio in questo, sarà considerare dove siete e cosa volete fare. Un tipico sistema casalingo inizia con hardware esistente e l'utente Linux convertito da poco vorrà ottenere il massimo dall'hardware esistente. Qualcuno che mette su un nuovo sistema per uno scopo specifico (come un ISP) dovrà considerare invece quale è lo scopo e comprare in relazione ad esso. Essendo ambizioso, cercherò di ricoprire l'intero ambito.

Vari scopi avranno anche necessità differenti riguardanti il posizionamento del file system sui dischi, una grande macchina multiutente sarà migliore con la directory `/home` su un disco separato, solo per dare un esempio.

In generale, per prestazione è vantaggioso dividere la maggior parte delle cose su più dischi possibili ma c'è un numero limitato di dispositivi che possono vivere su un bus SCSI ed il costo è naturalmente un altro fattore. Ugualmente importante, la manutenzione del file system diventa più complicata con l'aumentare del numero delle partizioni e dei dischi fisici.

10.1 Sistemi casalinghi

Con l'hardware economico che si può comprare oggi, è possibile avere un sistema grande a casa che è ancora economico, sistemi che battono i maggiori server del passato. Mentre molti hanno iniziato a mettere su un server Linux con vecchi dischi scartati (che è il motivo per il quale è nato questo HOWTO), molti possono permettersi oggi di comprare dischi da 20 GB.

La dimensione rimane importante per alcuni, e qui ci sono un po' di linee guida:

Testare

Linux è semplice e non avete nemmeno bisogno di un disco rigido per provarlo, se potete fare il boot dai floppy, probabilmente riuscirete a farlo funzionare sul vostro hardware. Se il kernel standard non vi funziona, non dimenticate che spesso ci possono essere versioni speciali dei dischi di boot per combinazioni inusuali di hardware che possono risolvere i vostri problemi iniziali fino a che non compilate il vostro kernel personale.

Conoscere

il sistema operativo è qualcosa in cui Linux eccelle, c'è una marea di documentazione ed i sorgenti sono disponibili. Un disco singolo con 50 MB è sufficiente per farvi iniziare con una shell e una cerchia ristretta dei comandi e delle utilità più frequentemente utilizzate.

Un utilizzo per hobby

o per un apprendimento più serio richiede più comandi ed utilità ma un disco singolo è ancora ciò che è necessario, 500 MB saranno spazio sufficiente, sia per i sorgenti che per la documentazione.

Serio

sviluppo software o semplicemente serio lavoro richiede anche molto altro spazio. A questo stadio, probabilmente avrete entrate di posta e news che richiedono file di coda e molto spazio. Dischi separati per compiti di vario genere cominceranno a mostrare un beneficio. A questo stadio probabilmente avrete anche un po' di dischi. Le necessità di dischi diventano più dure da stimare ma mi aspetterei che 2-4 GB siano sufficienti, anche per un piccolo server.

I server

sono di molti tipi, variando da server di posta fino a server ISP di piena grandezza. Una base di 2 GB per il sistema principale dovrebbe essere sufficiente, poi aggiungete spazio e forse anche dischi per caratteristiche separate che offrirete. Il costo è il maggior fattore limitante qui ma siate preparati a spendere un po' se volete giustificare la S dell'ISP. In verità non tutti lo fanno.

Praticamente un server è dimensionato come ogni macchina per utilizzo serio con spazio aggiunto per i servizi offerti, e tende ad essere limitato dall'IO piuttosto che dalla CPU.

Con tecnologia economica sia per linee di terra come anche per reti radio, è molto probabile che molto presto, gli utenti casalinghi avranno i propri server più o meno permenentemente agganciati alla rete.

10.2 Server

Grossi compiti richiedono grossi dischi ed una sezione separata qui. Se possibile mantenete quanto più possibile su dischi separati. Qualcuna delle appendici descrivono il setup di un piccolo server dipartimentale per 10-100 utenti. Qui presenterò un po' di considerazioni per i server limite. In generale non dovreste avere paura di utilizzare RAID, non solo perché è veloce e sicuro ma anche perché rende la crescita un po' meno dolorosa. Tutte le note qui sotto sono aggiunte ai punti menzionati in precedenza.

I server popolari raramente sono lì per caso, piuttosto, crescono nel tempo e questo richiede quantitativi generosi di spazio disco come anche una buona connessione di rete. In molti di questi casi potrebbe essere una buona idea riservare ad ogni compito interi dischi SCSI, da soli o in fila. In questo modo potrete spostare i dati nel caso il computer fallisse. Notate che trasferire i dischi attraverso i computer non è semplicissimo e potrebbe anche non funzionare sempre, specialmente nel caso di dischi IDE. Gli insiemi di dischi, richiedono setup attenti al fine di ricostruire i dati correttamente, quindi potreste voler mantenere una copia cartacea del vostro file `fstab` come anche una nota degli ID degli SCSI.

10.2.1 Directory Home

Stimate di quanti dischi avete bisogno, se sono più di 2, raccomanderei RAID, fortemente. Altrimenti dovreste separare gli utenti attraverso i vostri dischi dedicati agli utenti basati su una qualche specie di semplice algoritmo di hash. Ad esempio potreste utilizzare le prime due lettere del nome utente, quindi `jbloggs` viene situato in `/u/j/b/jbloggs` dove `/u/j` è un link simbolico ad un disco fisico quindi potete ottenere un carico bilanciato sui vostri dischi.

10.2.2 FTP Anonimo

Questo è un servizio essenziale se siete seri riguardo al servizio. I server buoni sono ben mantenuti, documentati, aggiornati e immensamente popolari, non importa dove sono localizzati nel mondo. Il grosso server

ftp.funet.fi `ftp://ftp.funet.fi`

è un eccellente esempio di ciò.

In generale questo non è una questione di CPU ma di ampiezza di banda di rete. La dimensione è difficile da calcolare, principalmente è una questione di ambizione e attitudini del server. Credo che il grosso archivio presente presso

ftp.cdrom.com `ftp://ftp.cdrom.com`

sia una macchina *BSD con un disco da 50 GB. Inoltre anche la memoria è importante per un server FTP dedicato, circa 256 MB di RAM sarebbero sufficienti per un server molto grande, anche se server più piccoli possono farcela bene anche con 64 MB di RAM. Le connessioni di rete sarebbero comunque sempre il fattore più importante.

10.2.3 WWW

Per molti questa è la ragione principale per andare in Internet, infatti ora sembra che molti li considerino la stessa cosa. Inoltre ad essere intensi in rete consegue il fatto di avere un bel po' di attività nei dischi per questo motivo, principalmente riguardante le cache. Mantenere le cache su un disco separato e veloce, porterebbe beneficio. Anche meglio sarebbe installare un server proxy di cache. In questo modo potreste ridurre la dimensione della cache per ogni utente ed aumentare la velocità del servizio riducendo nello stesso tempo le necessità di ampiezza di banda.

Con un server proxy di cache, avrete bisogno di un insieme di dischi veloci; RAID0 sarebbe l'ideale visto che l'affidabilità qui non è importante. Una capacità più alta è importante ma circa 2 GB sarebbero sufficienti per la maggior parte. Ricordatevi di far coincidere il periodo della cache con la capacità e la domanda. Periodi troppo lunghi sarebbero d'altro canto uno svantaggio, se possibile provate a regolare a seconda dell'URL. Per maggiori informazioni, controllate i server più utilizzati come `Harvest`,

Squid `http://www.nlanr.net/Squid`

e quello della

Netscape `http://www.netscape.com` .

10.2.4 Posta

Gestire la posta è qualcosa che molte macchine fanno fino ad un certo punto. I grandi server di posta, comunque, fanno gruppo a parte. Questo è un compito su richiesta e un grande server può essere lento anche se connesso a dischi veloci e ad una rete molto efficiente. Nel mondo Linux, il grande server come `vger.rutgers.edu` è un esempio ben noto. Diversamente da un servizio di news che è distribuito e che può parzialmente ricostruire lo spool utilizzando altre macchine come meccanismo di alimentazione, i server di posta, sono centralizzati. Questo rende la sicurezza molto più importante, quindi per un server principale, potreste considerare una soluzione RAID con enfasi sull'affidabilità. La dimensione è difficile da stabilire, dipende tutto da quante liste fate girare e da quanti iscritti avete.

Notate che in questi giorni si sta passando dall'utilizzare `POP` per prelevare la posta sulla macchiana locale dal server di posta all'utilizzo di `IMAP` per servire la posta mantenendo gli archivi di posta centralizzati. Questo vuol dire che la posta non è più accodata nel senso originale ma spesso cresce, richiedendo un'enormità di spazio disco. Inoltre sempre più si (ab)usano i messaggi con allegati per spedire ogni sorta di roba, anche un piccolo documento di un elaboratore testi può facilmente finire sopra il MB. Dimensionate i vostri dischi generosamente e controllate quanto spazio resta.

10.2.5 News

Questo è sicuramente un compito di grande volume e molto dipendente dal gruppo a cui vi iscrivete. Sul Nyx c'è un meccanismo di alimentazione molto completo ed i file di coda occupano circa 17 GB. I gruppi più grandi sono senza dubbio nella gerarchia `alt.binary.*`, quindi se per qualche ragione decidete di non averli, potete avere un buon servizio forse con 12 GB. In ogni caso altri, che rimarranno senza nome, pensano che 2 GB siano sufficienti per conferirsi il titolo di ISP. In questo caso le news scadono molto velocemente che penso che chiamarli IsP è abbastanza giustificato. Un completo meccanismo di alimentazione per le news significa un traffico di qualche GB ogni giorno e questo è un numero sempre crescente.

10.2.6 Altri

Ci sono molti servizi disponibili sulla rete e nonostante ciò molti sono stati messi nell'ombra dalla rete. Nonostante ciò servizi quali *archie*, *gopher* e *wais* esistono ancora e rimangono strumenti di valore sulla rete. Se state pensando in maniera coscienziosa di iniziare a creare un server principale, dovreste anche considerare questo servizio. Determinare lo spazio necessario è difficile, dipende tutto dalla popolarità e dalla domanda. Fornire un buon servizio ha inevitabilmente i suoi costi, lo spazio disco è solamente uno di essi.

10.2.7 Raccomandazioni sul Server

I server oggi richiedono molti dischi per funzionare in maniera soddisfacente per le impostazioni commerciali. Visto che il tempo medio tra i fallimenti (MTBF) diminuisce rapidamente con l'aumentare dei componenti, è consigliabile utilizzare RAID per protezione ed utilizzare un numero di dischi di media grandezza piuttosto che uno singolo ed enorme. Inoltre guardate anche nell progetto High Availability (HA) per maggiori informazioni.

10.3 Trappole

I pericoli di dividere ogni cosa su partizioni separate sono brevemente menzionati nella sezione riguardante la gestione del volume. Comunque, molte persone mi hanno chiesto di enfatizzare questo punto più fermamente: quando una partizione si riempie, non può crescere di più, non importa se c'è un mare di spazio su altre partizioni.

In particolare tenete d'occhio la crescita esplosiva nella coda delle news (`/var/spool/news`). Per macchine multi utente con le quote tenete sotto controllo `/tmp` e `/var/tmp` visto che c'è qualcuno che cerca di nascondere i propri file lì, basta cercare i file che finiscono per gif o jpeg...

Praticamente, per singoli dischi fisici questo schema offre guadagni molto piccoli, piuttosto che rendere più facile il controllo della crescita dei file (utilizzando '`df`') e del posizionamento fisico delle tracce. Più importante, non c'è possibilità per accesso di disco parallelo. La disponibilità di avere un sistema per la gestione di un volume potrebbe risolvere ciò, ma ciò accadrà in futuro. Comunque, quando file system più specializzati saranno disponibili, anche un disco singolo potrà beneficiare dall'essere diviso in diverse partizioni.

11 Struttura del Disco

Con tutto questo in testa siamo pronti per iniziare la struttura. Ho basato ciò sul mio metodo sviluppato quando ho preso 3 vecchi dischi SCSI e ho provato tutte le possibilità.

Le tavole nelle appendici sono disegnate per semplificare il processo di mappatura. Sono state progettate per aiutarvi attraverso il processo di ottimizzazione come anche per fare un log utile nel caso di riparazione del sistema. Sono dati anche alcuni esempi.

11.1 Selezione per il Partizionamento

Determinate le vostre necessità ed organizzate una lista di tutte le parti del file system che volete posizionare su partizioni separate ed ordinatele in ordine decrescente di richiesta di velocità e quanto spazio volete dare ad ogni partizione.

La tabella presente nella sezione 20 (Appendice A) è uno strumento utile per selezionare quali directory dovreste mettere su partizioni differenti. È ordinata in ordine logico con lo spazio per le vostre aggiunte e annotazioni riguardo i punti di montaggio e sistemi addizionali. Inoltre esso non è ordinato per velocità, invece le necessità di velocità sono indicate da pallini ('ò).

Se avete intenzione di impostare il RAID annotatevi i dischi che volete utilizzare e quali partizione volete usare con RAID. Ricordatevi che le varie soluzioni RAID offrono differenti velocità e gradi di affidabilità.

(Solo per farla più facile assumerò che abbiamo un insieme identico di dischi SCSI e nessun tipo di RAID).

11.2 Organizzare le Partizioni sui Dischi

Quindi ora vogliamo mettere le partizioni sui dischi fisici. Lo scopo del seguente algoritmo è di aumentare il parallelismo e la capacità del bus. In questo esempio i dischi sono A, B e C e le partizioni sono 987654321 dove 9 è la partizione con la più alta necessità di velocità. Partendo da un drive noi 'avremo' la linea della partizione attraverso i dischi in questo modo:

```
A : 9 4 3
B : 8 5 2
C : 7 6 1
```

Questo fa sì che la 'somma delle necessità di velocità' sia il più possibile omogenea per tutti i dischi.

Utilizzate la tabella nella sezione 21 (Appendice B) per selezionare quale disco utilizzare per ogni partizione al fine di ottimizzare il parallelismo.

Annotate le caratteristiche di velocità dei vostri dischi e annotate ogni directory nell'apposita colonna. Siate preparati a mischiare directory, partizioni e dischi prima di essere soddisfatti.

11.3 Ordinare le Partizioni sui Dischi

Dopo di ciò è consigliato selezionare la numerazione delle partizioni per ciascun disco.

Utilizzate la tabella nella sezione 22 (Appendice C) per selezionare i numeri di partizione al fine di ottimizzare in base alle caratteristiche di traccia. Alla fine dovreste avere una tabella ordinata in ordine crescente per numero di partizione. Riempite poi questi numeri nelle tabelle presenti nell'appendice A e B.

Troverete utili queste tabelle quando eseguirete il programma di partizionamento (`fdisk` or `cfdisk`) e al momento di fare l'installazione.

11.4 Ottimizzazione

Dopo di ciò, ci sono generalmente poche partizioni che devono essere 'mischiate' nei dischi sia per farcele entrare sia se ci fossero considerazioni speciali riguardanti la velocità, l'affidabilità, file system speciali ecc. In ogni caso, questo fornisce quello che questo autore crede sia un buon punto di inizio per un setup completo dei dischi e delle partizioni. Alla fine è l'utilizzo del momento che determinerà le necessità reali dopo aver fatto così tante premesse. Dopo le operazioni preliminari si dovrebbe assumere che arriva il momento in cui ripartizionare porterebbe benefici.

Ad esempio se uno dei 3 dischi nell'esempio menzionato sopra è molto lento in confronto agli altri due, un miglior progetto potrebbe essere il seguente:

```
A : 9 6 5
B : 8 7 4
C : 3 2 1
```

11.4.1 Ottimizzare per Caratteristica

Spesso i dischi possono essere simili nell'apparente velocità globale ma qualche vantaggio può essere ottenuto accoppiando i dischi alla dimensione, alla distribuzione ed alla frequenza di accesso. Quindi i file binari sono fatti per dischi con accesso rapido che offre possibilità di accodamento dei comandi, e le librerie sono fatte per dischi con velocità di trasferimento più ampie dove IDE offre una buona prestazione rispetto al prezzo.

11.4.2 Ottimizzare mediante Parallelizzazione del Disco

Evitate la contenzione del disco guardando ai task: ad esempio se state accedendo a `/usr/local/bin` ci sono possibilità anche che avrete molto presto bisogno di file da `/usr/local/lib` quindi posizionarli in dischi separati permette meno ricerca e possibili azioni in parallelo e caching del drive. È abbastanza probabile che scegliendo ciò che può apparire inferiore alle caratteristiche del disco, sarà sicuramente vantaggioso se potete ottenere operazioni in parallelo. Identificate compiti comuni, che partizioni utilizzano e cercate di mantenerle su dischi separati.

Giusto per illustrare il mio punto di vista, fornirò un po' di esempi dell'analisi dei task.

Software da Ufficio

come l'editing, l'elaborazione testi ed i fogli elettronici, sono tipici esempi di software di bassa intensità sia per quanto riguarda la CPU che l'intensità di disco. In ogni caso, se doveste avere un singolo server per un grande numero di utilizzatori non dovreste dimenticare che molto di questo software ha opportunità di auto salvataggi il che causa traffico extra, generalmente nelle directory home. Dividere gli utenti su più dischi potrebbe ridurre la contenzione.

I news reader

fanno anche loro auto salvataggi nella directory home quindi gli ISP dovrebbero considerare di separare le directory home.

Le code delle News sono note per avere directory profondamente ramificate e gran numero di file molto piccoli. La perdita di una partizione con le code delle News non è per molti un grande problema, quindi sono buone candidate per organizzare un RAID 0 con molti dischi piccoli per distribuire i vari accessi attraverso alberini multipli. Si raccomanda nei manuali e nelle FAQ del server delle news INN di mettere le cose delle news e i file `.overview` su dischi separati per installazioni più grandi.

C'è anche una pagina web dedicata all' *ottimizzazione di INN* `http://www.spinne.com/usenet/inn-perf.html` che sarebbe bene leggere.

Le applicazioni database

possono essere pretenziose sia in termini di utilizzo del disco che in termini di richiesta di velocità. I dettagli sono naturalmente specifici alle applicazioni, leggete la documentazione attentamente tenendo ben presente la richiesta di spazio su disco. Inoltre considerate il RAID sia per le prestazioni che per l'affidabilità.

Leggere ed inviare e-mail

coinvolge directory home come anche i file delle code in entrata ed in uscita. Se possibile mantenete le directory home e i file delle code su dischi separati. Se siete un server di posta o un hub di posta considerate di mettere le directory home e i file delle code su dischi separati.

Perdere la posta è una cosa molto brutta, se state gestendo un ISP o un hub principale. Riflettete sul fatto di fare RAID della vostra posta e considerate backup frequenti.

Lo sviluppo di software

può richiedere un gran numero di directory per file binari, librerie, file include, nonchśorgenti e file di progetto. Se possibile dividete il più possibile su dischi separati. Su piccoli sistemi potete posizionare `/usr/src` e i file di progetto sullo stesso disco delle directory home.

Navigare il Web

sta diventando sempre più popolare. Molti browser hanno una cache locale che può espandersi in volumi abbastanza grandi. Dal momento che questa è utilizzata per richiamare le pagine e nel ritornare alle pagine precedenti, la velocità è abbastanza importante in questi casi. Se invece siete connessi attraverso un proxy server ben configurato, non avete bisogno di più di pochi megabyte per utente per sessione. Controllate anche le sezioni sulle 10.2.1 (Directory Home) e 10.2.3 (WWW).

11.5 Compromessi

Un modo per evitare le trappole menzionate è semplicemente di dedicare le partizioni fisse a directory di dimensione ben nota come swap, `/tmp` e `/var/tmp` e raggruppare insieme i richiami nelle partizioni rimanenti utilizzando link simbolici.

Esempio: un disco lento (`discolento`), un disco veloce (`discoveloce`) ed un assortimento di file. Avendo organizzato `swap` e `tmp` sul `discoveloce`; e `/home` e root sul discolento abbiamo le rimanenti directory fittizie `/a/lento`, `/a/veloce`, `/b/lento` e `/b/veloce` da allocare sulle partizioni `/mnt.discolento` e `/mnt.discoveloce` che rappresentano le rimanenti partizioni dei due dischi.

Mettere `/a` o `/b` direttamente su entrambe i dischi, conferisce le stesse proprietà alle sottodirectory. Potevamo fare tutte e 4 le directory come partizioni separate ma perderebbe un po' di flessibilità nella gestione della dimensione di ciascuna directory. Una soluzione migliore è quella di fare link simbolici delle 4 directory a directory appropriate sui rispettivi dischi.

Quindi facciamo

```
/a/veloce punta a /mnt.discoveloce/a/veloce o   /mnt.discoveloce/a.veloce
/a/lento  punta a /mnt.discolento/a/lento   o   /mnt.discolento/a.lento
/b/veloce punta a /mnt.discoveloce/b/veloce o   /mnt.discoveloce/b.veloce
/b/lento  punta a /mnt.discolento/b/lento   o   /mnt.discolento/b.lento
```

e abbiamo tutte le directory veloci sul disco veloce senza dovere organizzare una partizione per le 4 directory. La seconda alternativa ci da un file system più piatto che in questo caso rende più semplice tenere sotto controllo la struttura.

Lo svantaggio è che all'inizio è uno schema complicato da mettere su e da pianificare e che tutti i punti di montaggio e le partizioni devono essere definiti prima dell'installazione del file system.

12 Implementazione

Avendo fatto la struttura, dovreste avere una descrizione dettagliata di cosa va dove. Probabilmente questo sarà su carta ma si spera che qualcuno farà un sistema più automatizzato che possa gestire qualsiasi cosa, a partire dal progetto, attraverso il partizionamento e la formattazione, fino all'installazione. Questa è la strada che uno dovrà seguire per realizzare il progetto.

Le distribuzioni moderne hanno mezzi di installazione che vi guideranno attraverso il partizionamento e la formattazione e anche ad organizzare per voi il file `/etc/fstab` automaticamente. Per modifiche posteriori, dovrete capire il meccanismo che sta dietro.

12.1 Dischi e Partizioni

Quando fate partire il DOS o qualcosa del genere troverete che tutte le partizioni sono nominate da `C:` in avanti, con nessuna differenza tra IDE, SCSI, rete o qualsiasi tipo di supporto che avete. Nel mondo di Linux, questo è abbastanza differente. Durante l'avvio, vedrete le partizioni descritte in questo modo:

```
Dec  6 23:45:18 demos kernel: Partition check:
Dec  6 23:45:18 demos kernel:  sda: sda1
Dec  6 23:45:18 demos kernel:  hda: hda1 hda2
```

I dischi SCSI sono etichettati `sda`, `sdb`, `sdc` ecc., e i dischi (E)IDE sono etichettati `hda`, `hdb`, `hdc` ecc. Ci sono inoltre nomi standard per tutti i dispositivi, tutte le informazioni possono essere trovate in `/dev/MAKEDEV` e in `/usr/src/linux/Documentation/devices.txt`.

Le partizioni sono etichettate numericamente per ciascun disco `hda1`, `hda2` e così via. Sui dischi SCSI ci possono essere 15 partizioni per disco, sui dischi EIDE ce ne possono essere 63. Tutti e due questi limiti eccedono ciò che è attualmente utile per molti dischi.

Questi sono poi montati attenendosi al file `/etc/fstab` prima di entrare a far parte del sistema.

12.2 Partizionamento

Prima di tutto, voi dovete partizionare ogni disco in un bel numero di partizioni separate. Sotto Linux ci sono due metodi principali, `fdisk` e quello più grafico `cfdisk`. Questi sono programmi complessi, leggetene il manuale *molto* attentamente.

Le partizioni sono di 3 tipi, `primarie`, `estese` e `logiche`. Dovete utilizzare le partizioni `primarie` per avviare il sistema, ma c'è un massimo di 4 partizioni primarie. Se volete di più, dovete definire una partizione `estesa` nell'ambito della quale voi definite le vostre partizioni logiche.

Ogni partizione ha un numero identificativo che dice al sistema operativo che cosa è, riguardo Linux i tipi che dovete conoscere sono `swap(82)` e `ext2fs(83)`.

C'è un file readme che accompagna `fdisk` che fornisce maggiori informazioni sul partizionamento.

Qualcuno ha appena fatto un *Partitioning HOWTO* che contiene informazioni eccellenti ed approfondite sui grattacapi del partizionamento. Piuttosto che ripeterlo qui e occupare oltre questo documento, farò invece riferimento ad esso per voi.

La Redhat ha scritto un'utility visuale chiamata *Disk Druid* che si suppone sia un'alternativa di facile utilizzo di `fdisk` e `cfdisk` ed inoltre automatizza un po' di altre cose. Sfortunatamente questo prodotto non è ancora sufficientemente maturo, quindi se lo utilizzate e non riuscite a farlo funzionare, siete ben avvisati di provare `fdisk` o `cfdisk`.

Il

Ranish Partition Manager `http://www.users.intercom.com/~ranish/part/`

è un'altra alternativa libera, mentre

Partition Magic `http://www.powerquest.com`

è un'alternativa commerciale che inoltre offre qualche supporto per ridimensionare le partizioni `ext2fs`.

Notate che Windows si lamenterà se trova più di una partizione primaria su un disco. Inoltre sembra assegnare le lettere del disco alle partizioni primarie mentre trova i dischi, prima di reiniziare dal primo per assegnare nomi di dischi consecutivi alle partizioni logiche.

Se volete DOS/Windows sul vostro sistema, prima di tutto dovreste fare quella partizione, una primaria su cui avviare il boot, fatta con l'`fdisk` del DOS. Poi, se volete NT, mettete quello. Alla fine, per Linux, create quelle partizioni con l'`fdisk` di Linux o con programmi equivalenti. Linux è sufficientemente flessibile da avviarsi sia dalle partizioni primarie che da quelle logiche.

12.3 Ripartizionamento

Alcune volte è necessario cambiare le dimensioni delle partizioni esistenti mantenendone intatto il contenuto. Un modo è ovviamente quello di fare il backup di tutto, ricreare le nuove partizioni e poi ripristinare i vecchi contenuti, e sebbene questo vi fornisca un ottimo modo di testare il vostro sistema di back up, vi porta via un bel po' di tempo.

Ridimensionare le partizioni è una semplice alternativa in cui un file system è dapprima contratto nel volume desiderato e poi la tavola di partizione è aggiornata per riflettere la nuova fine della posizione della partizione. Questo processo è comunque molto sensibile al file system.

Ripartizionare necessita che ci sia spazio libero alla fine dello spazio file, quindi per assicurarvi di essere in grado di contrarre la dimensione, dovreste innanzitutto deframmentare il vostro disco e svuotare qualsiasi cestino.

Utilizzando

fips `http://www.igd.fgh.de/~aschaefe/fips/`

potete ridimensionare una partizione `fat`, e l'ultima versione 1.6 di `fips` o di `fips 2.0` sono in grado anche di ridimensionare le partizioni `fat32`. Notate che questi programmi attualmente girano sotto DOS.

Ridimensionare altri file system è molto più complesso ma un sistema commerciale abbastanza popolare è

Partition Magic `http://www.powerquest.com`

che è in grado di ridimensionare più tipi di file system, inclusi `ext2fs` mediante il programma `resize2fs`.

Al fine di ottenere il massimo da `fips` dovreste innanzitutto cancellare i file non necessari, svuotare i cestini ecc., prima di deframmentare il vostro disco. In questo modo potete allocare più spazio per altre partizioni. Se il programma si lamenta che ci sono ancora file alla fine del disco, ciò forse è dovuto a file nascosti generati da Microsoft Mirror o da Norton Image. Questi sono probabilmente chiamati `image.idx` e `image.dat` e contengono backup di qualche file di sistema.

Ci sono prove che in alcuni programmi di deframmentazione di Windows è necessario che la casella permetti a Windows di spostare i file *non* sia contrassegnata, altrimenti finirete con l'avere qualche file nell'ultimo cilindro della partizione il che eviterà a FIPS di richiedere spazio.

Se avete ancora file non spostabili alla fine della vostra partizione DOS, dovreste ottenere il programma DOS showfat versione 3.0 o superiore. Questo vi mostra quali file sono dove così potete avere a che fare con loro direttamente.

Ripartizionare è un processo pericoloso come qualsiasi altro partizionamento quindi siete avvisati di avere a portata di mano un backup fresco.

12.4 Il Bug della Partizione Microsoft

Nei prodotti Microsoft fino a Win 98 c'è un bug che vi può causare un po' di problemi: se avete diverse partizioni `fat` primarie e l'ultima partizione estesa non è una partizione `fat`, il system Microsoft cercherà di montare l'ultima partizione al posto dell'ultima partizione FAT primaria.

Maggiori

informazioni `http://www.v-com.com/95Notes.html`

su questo sono disponibili sulla rete.

Per evitare ciò, potete sistemare una piccola parizione logica `fat` proprio alla fine del disco.

Visto che qualche componente hardware è accompagnato da setup software che è disponibile sotto DOS, questo potrebbe essere a portata di mano in ogni caso. Esempi degni di nota sono i controller RAID da DPT ed un bel numero di schede di rete.

12.5 Dispositivi multipli (md)

Essendo in uno stato di sviluppo dovreste assicurarvi di leggere la documentazione più recente su questa caratteristica del kernel.

Spiegata brevemente, funziona aggiungendo partizioni insieme in nuovi dispositivi `md0`, `md1` ecc. utilizzando `mdadd` prima che voi li attiviate mediante `mdrun`. Questo processo può essere automatizzato mediante il file `/etc/mdtab`.

Il sistema `md` più recente utilizza `/etc/raidtab` ed una sintassi differente. Assicuratevi che il pacchetto di strumenti RAID combaci con la versione `md` dal momento che il protocollo interno è cambiato.

Dopo di che li trattate come qualsiasi altra partizione su disco. Procedete con la formattazione ecc. come descritto qui sotto utilizzando questi nuovi dispositivi.

C'è ora anche un HOWTO in sviluppo per RAID che utilizza `md` che dovreste leggere.

12.6 Formattazione

Dopo viene la formattazione della partizione, in cui viene predisposta la struttura dei dati che descriverà i file e dove saranno posti. Se questa è la prima volta si raccomanda la formattazione con verifica. Parlando stringatamente, non dovrebbe essere necessario ma questo stimola l'I/O così tanto da scoprire problemi potenziali come incorretta terminazione, prima che voi archiviate i vostri preziosi dati. Guardate il comando `mkfs` per maggiori dettagli.

Linux è in grado di gestire un gran numero di file system, piuttosto che ripetere i dettagli, potete leggere le pagine di manuale di `fs` che li descrive in maggior dettaglio. Notate che il kernel, per poter utilizzare queste funzioni, deve avere i driver compilati internamente o come moduli. Quando è ora di compilare il kernel, dovreste leggere attentamente la lista delle caratteristiche del file system. Se utilizzate `make menuconfig` potete avere un aiuto in linea per ogni tipo di file system.

Notate che qualche sistema di disco di recupero ha bisogno che `minix`, `msdos` e `ext2fs` vengano compilati nel kernel.

Anche le partizioni di swap devono essere preparate e per far questo utilizzate `mkswap`.

12.7 Montaggio

I dati presenti su una partizione non sono disponibili in un file system fino a che essa non è montata in un punto di montaggio. Questo può essere fatto manualmente utilizzando `mount` o automaticamente durante l'avvio, aggiungendo linee appropriate all'`/etc/fstab`. Leggete il manuale di `mount` e prestate molta attenzione alle tabulazioni.

12.8 fstab

Durante il processo di boot il sistema monta tutte le partizioni come elencate nel file `fstab` che può apparire così:

```
# <file system>    <punto mont.>    <tipo>  <opzioni>    <dump>  <pass>
/dev/hda2           /                ext2    defaults     0       1
None                none             swap    sw           0       0
proc                /proc            proc    defaults     0       0
/dev/hda1           /dosc            vfat    defaults     0       1
```

Questo file è in qualche modo sensibile alla formattazione utilizzata quindi è meglio e più conveniente modificarlo utilizzando uno degli strumenti di modifica creati per questo scopo.

Brevemente, i campi sono: nome della partizione, dove montare la partizione, tipo di file system, quando montarla, quando effettuare il dump per il backup e quando fare `fsck`.

Linux offre la possibilità di fare un controllo di file parallelo (`fsck`) ma per essere efficiente è importante non fare `fsck` su più di una partizione di un disco alla volta.

Per maggiori informazioni fate riferimento alle pagine di manuale di `mount` e `fstab`.

12.9 Raccomandazioni

Avendo costruito ed implementato il vostro intelligente schema, siete ben avvisati di annotarvelo tutto, su carta. È inutile avere le informazioni necessarie su disco, se la macchina è andata.

Le tabelle di partizione possono essere danneggiate o perse, nel qual caso è estremamente importante che voi mettiate gli stessi, esatti, numeri in `fdisk`, così da recuperare il vostro sistema. Potete utilizzare il programma `printpar` per fare una chiara annotazione delle tabelle. Scrivetevi anche i numeri SCSI o i nomi IDE per ogni disco, così potete rimettere insieme il sistema nell'ordine giusto.

13 Manutenzione

È dovere dell'amministratore di sistema tenere d'occhio i dischi e le partizioni. Se dovesse riempirsi una delle partizioni, è probabile che il sistema possa smettere di funzionare correttamente, non importa quanto spazio sia disponibile su altre partizioni, fino a che lo spazio è richiesto.

Le partizioni e i dischi possono essere facilmente monitorate utilizzando `df` e dovrebbe essere fatto frequentemente, forse mediante l'utilizzo di un job di cron o con altri strumenti di gestione del sistema.

Non dimenticate le partizioni di swap, queste sono monitorate meglio utilizzando uno dei programmi di statistica della memoria come `free`, `procinfo` o `top`.

Monitorare l'utilizzo dei dischi è più difficile ma è importante nell'interesse delle prestazioni per evitare la disputa - ponendo troppa domanda su un singolo disco se altri sono disponibili e inattivi.

È importante, quando si installano i pacchetti software, avere una chiara idea di dove i vari file devono andare. Come menzionato precedentemente GCC mantiene i file binari in una directory di libreria e ci sono anche altri programmi che per ragioni storiche sono difficili da tenere presente, X11 per esempio ha una struttura inusualmente complessa.

Quando il vostro sistema è quasi pieno, è ora di controllare e sfoltire i vecchi messaggi di log ed anche eliminare i file core. L'uso corretto di `ulimit`, nelle impostazioni globali della shell, può essere utile a salvarvi dall'avere i file core sparsi per tutto il sistema.

13.1 Backup

Il lettore attento avrà notato un po' di consigli sulla utilità di fare backup. Ci sono storie horror riguardo incidenti e cosa è accaduto al responsabile quando il backup si è rivelato non essere funzionante o anche non esistente. Potreste pensare che sia più semplice investire in backup corretti che in una seconda identità segreta.

Ci sono molte opzioni ed anche un mini-HOWTO (`Backup-With-MSDOS`) che spiega in dettaglio cosa dovete sapere. In aggiunta alle specifiche DOS, esso contiene inoltre informazioni generali e gli orientamenti futuri.

Oltre a fare questi backup, dovreste assicurarvi che siate in grado di ripristinare i dati. Non tutti i sistemi verificano che i dati scritti siano corretti e molti amministratori hanno iniziato felici a ripristinare i dati dopo un incidente, nella speranza che tutto fosse funzionante, solo per scoprire con orrore che i backup erano inutili. Fate attenzione.

13.2 Deframmentazione

Questa dipende molto dalla struttura del file system, alcuni soffrono di una frammentazione veloce e abbastanza debilitante. Fortunatamente per noi, `ext2fs` non appartiene a questo gruppo e inoltre c'è stato molto poco da dire riguardo gli strumenti di deframmentazione. Infatti esistono, ma difficilmente se ne ha bisogno.

Se per qualche ragione ritenete che ciò sia necessario, la soluzione veloce e facile è di fare un backup e ripristinare. Se solo una piccola area ne è affetta, ad esempio le directory home, potreste fare un `tar` di essa su una area temporanea di un'altra partizione, *verificare* l'archivio, cancellare l'originale e poi rifarne l'estrazione di nuovo con tar.

13.3 Cancellazioni

Abbastanza spesso le carenze di disco possono essere rimediate semplicemente cancellando i file non necessari accumulati nel sistema. Molto spesso i programmi che terminano di funzionare in maniera anormale, causano confusioni di tutti i tipi nei posti peggiori. Normalmente un core dump si ha dopo una situazione del genere e se non avete intenzione di fare il debug, potete cancellarlo. Questi li possiamo trovare dappertutto quindi siete avvisati di farne una ricerca completa ora e sempre.

L'interruzione improvvisa di un programma può inoltre causare tutta una serie di file temporanei che rimangono in posti quali `/tmp` o `/var/tmp`, file che sono automaticamente rimossi quando il programma finisce normalmente. Il riavvio pulisce qualcuna di queste aree ma non necessariamente tutte e se passa molto tempo potreste finire con avere un sacco di roba. Se lo spazio è poco, dovete cancellare con attenzione, assicurandovi che il file non sia in uso in quel momento. Utilità quali `file` possono spesso dirvi che tipo di file state guardando.

Molte cose vengono messe nei log quando il sistema gira, per lo più nell'area `/var/log`. In particolare il file `/var/log/messages` tende a crescere fino a che viene cancellato. È una buona idea mantenere un piccolo archivio di file log per paragonarli se il sistema dovesse cominciare a comportarsi in modo non corretto.

Se il sistema di posta o delle news non funziona correttamente, potreste avere crescita eccessiva nelle aree di coda, `/var/spool/mail` e `/var/spool/news` rispettivamente. Fate attenzione ai file di overview visto che hanno un punto che li rende invisibili all'`ls -l`, generalmente è meglio utilizzare `ls -Al` che ve li svelerà.

La saturazione dello spazio dell'utente è un argomento particolarmente complesso. Guerre sono state fatte tra amministratori di sistemi e utenti. Tatto, diplomazia e un budget generoso per i nuovi dischi è ciò di cui si ha bisogno. Utilizzate la caratteristica message-of-the-day, visualizzato durante il login dal file `/etc/motd` per dire agli utenti quando lo spazio è poco. Predisporre la configurazione standard della shell per prevenire la formazione di file core può far risparmiare molto spazio.

Alcuni tipi di persone cercano di mascherare i file nel sistema, generalmente cercando di trarre vantaggio dal fatto che i file con un punto davanti al nome sono invisibili al comando `ls`. Un esempio comune sono i file che appaiono come `...` che anche normalmente non sono visibili, o quando si usa `ls -al` scompaiono nel rumore di tutti i file come `.` o `..` che sono presenti in ogni directory. C'è comunque una contromisura a ciò, utilizzate `ls -Al` che sopprime `.` o `..` ma mostra tutti gli altri file con il punto.

13.4 Aggiornamenti.

Non importa quanto grandi siano i vostri dischi, verrà il tempo in cui scoprirete che ne avete bisogno di altri. Con l'avanzare della tecnologia potete ottenere sempre di più col vostro denaro. Al momento di scrivere ciò, sembra che i dischi da 6.4 GB siano quanto di meglio si possa avere coi vostri soldi.

Notate che con i dischi IDE potreste dover rimuovere un vecchio disco, visto che il massimo numero gestito dalla vostra scheda madre è generalmente di 2 o qualche volta di 4. Con lo SCSI, ne potete avere fino a 7 per canale per lo SCSI stretto (8-bit) e fino a 15 per il wide SCSI (15-bit). Qualche adattatore può gestire più di un singolo canale e in ogni caso potete avere più di un adattatore per sistema. La mia raccomandazione personale è che a lungo andare vi troverete meglio con lo SCSI.

La domanda che viene è, dove dovrei mettere questo nuovo disco? In molti casi, la ragione per l'espansione è che volete un'area per le code più ampia e in quel caso la soluzione più semplice e veloce è di montare il disco da qualche parte sotto `/var/spool`. D'altra parte, i dischi più nuovi sono generalmente più veloci dei vecchi quindi a lungo andare vi potrebbe sembrare opportuno di fare una riorganizzazione generale, possibilmente utilizzando i vostri vecchi schemi di pianificazione.

Se l'aggiornamento è forzato dal fatto che non c'è più spazio in partizioni utilizzate per cose come `/usr` o `/var`, l'aggiornamento è un po' più complesso. Dovreste considerare la possibilità di una riorganizzazione completa dalla vostra distribuzione preferita (e si spera più aggiornata). In questi casi dovreste stare attenti a non sovrascrivere le vostre configurazioni essenziali. Generalmente queste cose sono nella directory `/etc`. Procedete con cura, backup recenti e dischi di recupero che funzionano. L'altra possibilità è di copiare semplicemente la vecchia directory sopra quella nuova che è montata su un punto di montaggio temporaneo, editate il file `/etc/fstab`, fate ripartire il sistema con la nuova partizione al suo posto e controllate che funzioni. Qualora dovesse fallire, potete riavviare con il disco di recupero, rieditare il file `/etc/fstab` e provare di nuovo.

Fino a che la gestione dei volumi non diventerà disponibile per Linux, questo è sia complicato che pericoloso. Non vi meravigliate più di tanto se scoprite che avete bisogno di ripristinare il sistema da un backup.

Il Tips-HOWTO fornisce l'esempio seguente su come spostare un'intera struttura della directory.

```
(cd /directory/di/origine; tar cf - . ) | (cd /directory/di/destinazione; tar xvfp -)
```

Sebbene questo approccio per muovere alberi di directory è portabile per molti sistemi Unix, è difficile da ricordare. Inoltre esso fallisce con directory troppo annidate quando i nomi dei percorsi diventano troppo lunghi per essere maneggiati da tar (il tar GNU ha caratteristiche speciali per gestire nomi di file lunghi).

Se potete utilizzare il cp di GNU (il che accade sempre nel caso di sistemi linux), potete anche utilizzare

```
cp -av /directory/di/origine /directory/di/destinazione
```

Il cp di GNU ha una conoscenza specifica di link simbolici, FIFO e file di dispositivo e li copierà correttamente.

Ricordatevi che non è una buona idea provare a trasferire `/dev` o `/proc`.

13.5 Recupero

I crash di sistema avvengono in molti e divertenti modi e la corruzione della tabella di partizione garantisce sempre il massimo dell'eccitazione. Uno strumento recente e senza dubbio utile per quelli di noi che sono contenti con il normale livello di eccitazione è

gpart `http://www.stud.uni-hannover.de/user/76201/gpart`

che significa Guess PC-Type hard disk partitions (indovina il tipo delle partizioni dell'hard disk. ndt). Utile.

14 Questioni Avanzate

Linux ed i sistemi correlati, offrono una marea di possibilità per una distruzione veloce, efficiente e devastante. Questo documento non fa eccezione. Con la potenza c'è il pericolo e le sezioni seguenti descrivono poche altre imprese esoteriche che non dovrebbero essere tentate prima di aver letto e capito la documentazione, le questioni ed i pericoli. Dovreste anche fare backup. Ricordatevi inoltre di provare a ripristinare il sistema da zero dal vostro backup almeno una volta. Altrimenti potreste scoprire di non essere il primo ad avere un backup perfetto e nessuno strumento disponibile per reinstallarlo (o, ancora più imbarazzante, qualche file critico mancante su nastro).

Le tecniche qui descritte sono raramente necessarie, ma possonoe essere utilizzate per setup molto specifici. Pensate bene cosa volete ottenere prima di averci a che fare.

14.1 Regolazione del Disco Rigido

I paramentri del disco rigido possono essere regolati utilizzando l'utilità `hdparm`. Qui, il parametro più interessante è probabilmente il parametro di lettura che determina quanto prefetch debba essere fatto nella lettura sequenziale. Se volete provarlo, ha più senso regolare per la dimensione di file caratteristica sul vostro disco ma ricordate che questa regolazione è per il disco *intero* il che lo rende più difficile. Probabilmente questo è usato solamente sui grandi server, utilizzando dischi dedicati alle news ecc.

Per motivi di sicurezza, i settaggi convenzionali di hdparm sono abbastanza conservativi. Lo svantaggio è che ciò significa che potete ottenere interrupt persi se avete alta frequenza di IRQ, come avreste utilizzando la porta seriale e un disco IDE, visto che gli IRQ degli ultimi maschererebbero altri IRQ. Questo si potrebbe notare da una prestazione inferiore all'ideale nello scaricare i dati dalla rete al disco. Fare `hdparm -u1 dispositivo` dovrebbe prevenire il mascheramento e aumentare anche le prestazioni, o a seconda dell'hardware, corrompere i dati dell'hard disk. Sperimentazioni con cautela e backup recenti.

14.2 Regolazione del File System

Molti file system sono disponibili con un'utilità di regolazione e per `ext2fs` c'è l'utilità `tune2fs`. Diversi parametri possono essere modificati ma forse il parametro più importante è che dimensione dovrebbe essere preservata e chi dovrebbe trarre vantaggio da ciò, il che vi potrà aiutare ad ottenere più spazio utile dai vostri dischi possibilmente al costo di un minor spazio per riparare un sistema qualora dovesse piantarsi.

14.3 Sincronizzazione dell'Alberino

Questo non dovrebbe essere pericoloso, se non per il fatto che i dettagli esatti delle connessioni rimangono poco chiare per molti dischi. La teoria è semplice: mantenere una prefissata differenza di fase tra i dischi diversi in un setup RAID fa sì che si debba aspettare di meno che la traccia giusta arrivi in posizione per la testina di lettura/scrittura. In pratica sembra ora che con grandi buffer di lettura nei dischi, l'effetto sia da trascurare.

La sincronizzazione dell'alberino non dovrebbe essere utilizzata su RAID0 o RAID 0/1 visto che perdereste il beneficio di avere le testine di lettura sopra aree differenti dei settori rispecchiati.

15 Ulteriori Informationi

C'è un'enormità di informazioni che si dovrebbero conoscere nel metter su un sistema più grande, ad esempio per un Fornitore di Accessi News o un generico Fornitore di Accessi Internet. Le FAQ nei gruppi seguenti sono utili:

15.1 News group

Alcuni dei news group più interessanti sono:

- *Archiviazione* `news:comp.arch.storage` .
- *Archiviazione su PC* `news:comp.sys.ibm.pc.hardware.storage` .
- *AFS* `news:alt.filesystems.afs` .
- *SCSI* `news:comp.periphs.scsi` .
- *Linux setup* `news:comp.os.linux.setup` .

La maggior parte dei newsgroup ha le proprie FAQ che sono organizzate per rispondere a molte delle vostre domande, come indica il nome Frequently Asked Question (domande poste frequentemente). Versioni recenti dovrebbero essere postate regolarmente sui newsgroup rilevanti. Se non potete trovarle nelle vostre code di stampa, potete andare direttamente al

sito FTP del maggior archivio di FAQ `ftp://rtfm.mit.edu` . Le versioni WWW posso essere lette nel

sito WWW del maggior archivio di FAQ `http://www.cis.ohio-state.edu/hypertext/faq/usenet/FAQ-List.html` .

Qualche FAQ ha il proprio sito, di particolare interesse qui abbiamo

- *SCSI FAQ* `http://www.paranoia.com/~filipg/HTML/LINK/F_SCSI.html` e
- *comp.arch.storage FAQ* `http://alumni.caltech.edu/~rdv/comp_arch_storage/FAQ-1.html` .

15.2 Mailing List

Questi sono canali di bassa confusione principalmente dedicati agli sviluppatori. Pensate due volte prima di chiedere lì visto che il chiasso rallenta lo sviluppo. Tra le liste rilevanti troviamo `linux-raid`, `linux-scsi` e `linux-ext2fs`. Molte delle mailing list più utili girano sul server `vger.rutgers.edu` ma questo è notoriamente sovraccarico, quindi cercate di trovare un mirror. C'è qualche lista mirrorata presso

The Redhat Home Page `http://www.redhat.com` . Molte liste sono inoltre accessibili presso

linuxhq `http://www.linuxhq.com/lnxlists` , ed il resto del sito è una miniera di utili informazioni.

Se volete trovare di più riguardo le liste disponibili potete mandare un messaggio con la riga `lists` al list server presso vger.rutgers.edu (*majordomo@vger.rutgers.edu*). Se avete bisogno di aiuto su come utilizzare il mail server, mandate semplicemente la riga `help` allo stesso indirizzo. Vista la popolarità di questo server, è probabile che passerà un po' di tempo prima che otteniate una risposta, o anche i messaggi, dopo aver spedito il comando `subscribe`.

C'è inoltre un bel numero di list server majordomo che possono essere di qualche interesse come la lista dei driver EATA (*linux-eata@mail.uni-mainz.de*) e la lista intelligente di IO *linux-i2o@dpt.com* .

Le mailing list sono in uno stato di flux ma potete trovare link ad un bel numero di liste interessanti dalla

Linux Documentation Homepage `http://metalab.unc.edu/LDP/` .

15.3 HOWTO

Questi vanno intesi come punti di partenza per ottenere l'informazione di supporto o anche per mostrarvi come risolvere uno specifico problema. Tra gli HOWTO rilevanti troviamo `Bootdisk`, `Installation`, `SCSI` e `UMSDOS`. Il sito principale per questi è l' *archivio LDP* `http://metalab.unc.edu/LDP/`

presso Metalab (formalmente conosciuto come Sunsite).

C'è un nuovo HOWTO che ha a che fare con il mettere su un sistema RAID DPT, controllate la

DPT RAID HOWTO homepage `http://www.ram.org/computing/linux/dpt_raid.html` .

15.4 Mini-HOWTO

Questi sono testi liberi più piccoli rispetto agli HOWTO. Tra i mini-HOWTO rilevanti troviamo `Backup-With-MSDOS`, `Diskless`, `LILO`, `Large Disk`, `Linux+DOS+Win95+OS2`, `Linux+OS2+DOS`, `Linux+Win95`, `NFS-Root`, `Win95+Win+Linux`, `ZIP Drive`. Li potete trovare allo stesso posto degli HOWTO, generalmente in una sottodirectory chiamata `mini`. Notate che questi sono destinati ad essere convertiti in SGML ed a diventare veri e propri HOWTO in un futuro prossimo.

Il vecchio `Linux Large IDE mini-HOWTO` non è più valido, leggete invece `/usr/src/linux/drivers/block/README.ide` o `/usr/src/linux/Documentation/ide.txt`.

15.5 Risorse locali

In molte distribuzioni di Linux, c'è installata una directory di documenti, controllate nella directory */usr/doc* in cui la maggior parte dei pacchetti pone la propria documentazione principale, i file README, ecc. Inoltre troverete qui l'archivio HOWTO (*/usr/doc/HOWTO*) di HOWTO già formattati ed inoltre l'archivio di mini-HOWTO (*/usr/doc/HOWTO/mini*) con documenti in testo semplice.

La maggior parte dei file di configurazione citati in precedenza può essere trovata nella directory */etc*

In particolare vorrete lavorare con il file */etc/fstab*

che predispone il montaggio delle partizioni e probabilmente anche col file */etc/mdtab*

che è utilizzato dal sistema `md` per organizzare il RAID.

I sorgenti del kernel */usr/src/linux*

sono, ovviamente, il documento finale. In altre parole, *usa i sorgenti, Luke*. Dovrebbe anche essere puntualizzato che il kernel non ̀dato solo con i codici sorgenti, che sono anche commentati (beh, almeno una parte), ma anche con una

directory di documentazione `file:///usr/src/linux/Documentation` informativa. Se state per chiedere qualcosa riguardo al kernel, dovreste prima leggerla; eviterà a voi e a molti altri di perdere un sacco di tempo e anche possibili imbarazzi.

Controllate anche nel file log del sistema (*/var/log/messages*) per vedere cosa sta succedendo e in particolare come è andato l'avvio, se vi è sfuggito via dallo schermo. Utilizzando `tail -f /var/log/messages` in una finestra o in uno schermo separati, otterrete un continuo aggiornamento su cosa succede nel vostro sistema.

Potete anche trarre vantaggio del file system */proc*

che è una finestra dell'intrinseco funzionamento del vostro sistema. Utilizzate `cat` piuttosto che `more` per vedere i file nel momento in cui sono dichiarati essere di lunghezza zero.

15.6 Pagine Web

C'è un grande numero di pagine web informative lì fuori e per loro stessa natura cambiano velocemente, quindi non siate troppo sorpresi se questi link diventano facilmente datati.

Un buon punto di inizio è ovviamente l' *archivio LDP* `http://metalab.unc.edu/LDP/` di Metalab che è un centro di informazioni sulla documentazione, sulle pagine di progetti e molto altro.

- Mike Neuffer, l'autore dei driver dei controller RAID con caching DPT, ha qualche pagina interessante sullo

  SCSI `http://www.uni-mainz.de/~neuffer/scsi`

  e

  DPT `http://www.uni-mainz.de/~neuffer/scsi/dpt` .
- Informazioni sullo sviluppo del Software RAID possono essere trovate presso il

  sito del kernel Linux `http://www.kernel.org/`

  con patch ed utility.
- Informazioni riguardanti i dischi, sul benchmarking, sul RAID, sull'affidabilità e molto altro, possono essere trovate presso la pagina del progetto

  Linas Vepstas `http://linas.org` .
- Ci sono inoltre anche altre informazioni disponibili su come

  fare il RAID della partizione root `ftp://ftp.bizsystems.com/pub/raid/Root-RAID-HOWTO.html`

  e quali pacchetti software sono necessari per farlo.
- È disponibile anche una documentazione approfondita su

  ext2fs `http://step.polymtl.ca/~ldd/ext2fs/ext2fs_toc.html` .
- Gente che cerca informazioni su VFAT, FAT32 e Joliet può dare un'occhiata alla

  pagina di sviluppo `http://bmrc.berkeley.edu/people/chaffee/index.html` . Questi driver sono nella serie di sviluppo del kernel 2.1.x come anche nel kernel 2.0.34 e successivi.
- Per maggiori informazioni sull'avvio e anche qualche informazione su BSD, date un'occhiata alla pagina con le

  informazioni sull'avvio `http://www.paranoia.com/~vax/boot.html`

Per i diagrammi e le informazioni su tutti i tipi di dischi, controller ecc. Sia per linee continue che discontinue

The Ref `http://theref.c3d.rl.af.mil`

è il sito di cui avete bisogno. Ci sono un bel po' di informazioni utili qui, un vero tesoro. Potete anche scaricarvi il database presso l'

FTP `ftp://theref.c3d.rl.af.mil/public` .

Fatemi sapere se avete qualche altra idea che può essere utile.

15.7 Motori di Ricerca

Ricordatevi che potete anche utilizzare i motori di ricerca e che alcuni, come

- *Altavista* `http://www.altavista.digital.com`

- *Excite* `http://www.excite.com`
- *Hotbot* `http://www.hotbot.com`

permettono anche ricerche usenet.

Ricordatevi inoltre che

Dejanews `http://www.dejanews.com`

è un motore dedicato alle news che mantiene una coda di news dal 1995 in poi.

Se dovete chiedere aiuto potrete forse trovarne nel newsgroup

Linux Setup `news:comp.os.linux.setup`

A causa di un grande carico di lavoro e una lenta connessione di rete non sono in grado di seguire quel newsgroup quindi se volete contattarmi dovete farlo via e-mail.

16 Ottenere Aiuto

Alla fine vi potrete trovare nella condizione di non poter risolvere i problemi ed avere bisogno di aiuto da qualcuno. La maniera più efficiente è quella di chiedere localmente o nel più vicino LUG, cercate sul web il più vicino.

Un'altra possibilità è di chiedere su Usenet in uno dei moltissimi newsgroup disponibili. Il problema è che questi hanno un traffico e un rumore così elevati (chiamato basso rapporto segnale/rumore), che la vostra domanda può finire col non ricevere risposta.

Non importa dove lo chiedete ma è importante che lo chiediate correttamente o non sarete presi sul serio. Dire solamente *il mio disco non funziona* non vi aiuterà ed aumenterà il brusio ancora di più e se siete fortunati qualcuno vi chiederà delucidazioni.

Invece descrivete i vostri problemi in dettaglio, il che permetterà alla gente di aiutarvi. Il problema potrebbe giacere da qualche parte che non sospettavate. Quindi siete avvisati di elencare le seguenti caratteristiche del vostro sistema:

Hardware

- Processore
- DMA
- IRQ
- Chip set (LX, BX ecc)
- Bus (ISA, VESA, PCI ecc)
- Schede di espansione utilizzate (Adattatori, video, IO ecc.)

Software

- BIOS (Su scheda madre e possibilmente gli adattatori SCSI)
- LILO, se utilizzato
- Versione del kernel Linux ed anche possibili patch e modifiche
- Parametri del kernel, se ce ne sono
- Software che mostra l'errore (con il numero di versione o la data)

Periferiche

- Tipo di lettori con la marca, la versione ed il tipo
- Altre periferiche rilevanti connesse allo stesso bus

Come esempio su come queste cose sono relazionate tra loro: un vecchio chip set ha causato problemi con una certa combinazione di controller video e adattatori SCSI.

Ricordatevi che il testo del boot è presente in `/var/log/messages` e può rispondere a molte delle domande qui sopra. Ovviamente, se i dischi falliscono non sarete in grado di avere i file di log salvati ma potete almeno tornare indietro nello schermo utilizzando i tasti `SHIFT` e `PAGE UP`. Potrebbe anche essere utile includere parte di ciò nella vostra richiesta di aiuto ma non esagerate, *capite* che un completo file di log inserito in Usenet è più che un semplice fastidio.

17 Annotazioni conclusive

La messa a punto del disco e le decisioni sulle partizioni sono difficili da intraprendere e non ci sono grosse regole. Ciò nonostante è una buona idea lavorare di più su queste cose visto che i vantaggi possono essere considerevoli. Aumentare al massimo l'utilizzo su un disco solo quando gli altri sono inutilizzati è tutt'altro che una buona idea, guardate le luci dei dischi, non stanno lì solo per decorazione. Per un sistema organizzato bene le luci dovrebbero sembrare come Natale in una discoteca. Linux offre il software RAID ma gestisce anche qualche controller SCSI RAID hardware. Controllate cosa è disponibile. Con il progredire del sistema e delle esperienze, potreste ripartizionare e controllare questo documento ancora una volta. Le aggiunte sono sempre benvenute.

Per finire mi piacerebbe riassumere qualche raccomandazione

- I dischi sono economici ma i dati che essi contengono potrebbero valere molto di più, utilizzate e testate il vostro sistema di backup.
- Anche il lavoro è costoso, assicuratevi di ottenere dischi sufficientemente grandi visto che installarne altri, o ripartizionare i vecchi, richiede tempo.
- Pensate all'affidabilità, rimpiazzate i vecchi dischi prima che falliscano.
- Mantenete una copia cartacea della vostra configurazione, avere tutto su disco quando la macchina è andata non aiuta molto.
- Cominciate con un semplice progetto con un minimo di tecnologia fantasiosa e piuttosto mettetecela dopo. In generale aggiungere è più facile che rimpiazzare, che siano dischi, tecnologia o altre caratteristiche.

17.1 Presto disponibili

Ci sono un po' di cose importanti che stanno per apparire qui. In particolare aggiungerò più tabelle di esempi visto che sto per mettere su due sistemi grossi e generici, uno al lavoro e uno a casa. Questo dovrebbe dare una sensazione generale su come un sistema può essere organizzato per uno di questi due scopi. Esempi su sistemi che girano bene sono benvenuti.

C'è anche un po' di lavoro da fare sui vari tipi di file system e sulle utility.

Ci sarà una grande aggiunta sulle tecnologie dei dischi molto presto come anche una descrizione più dettagliata sull'utilizzo di `fdisk`, `cfdisk` e `sfdisk`. I file system saranno aggiornati con l'avvento di nuove caratteristiche come anche più sul RAID e quali directory possono beneficiare da quale livello di RAID.

C'è qualche piccola interferenza tra il Linux Filesystem Structure Standard e il FHS che spero di integrare meglio presto, il che probabilmente vuol dire un grande rimaneggiamento di tutte le tavole alla fine di questo documento.

Con la lettura di questo documento da parte di più persone dovrei ottenere qualche ulteriore commento e feedback. Sto anche pensando ad un programma in grado di automatizzare un bel po' di queste decisioni e sebbene non sia certo l'optimum, dovrebbe fornire un punto di inizio più semplice e completo.

17.2 Richieste ed Informazioni

C'è voluto un bel po' di tempo per ottenere questo documento e sebbene molti pezzi stanno congiungendosi ci sono ancora delle informazioni di cui abbiamo bisogno prima di uscire dalla versione beta.

- Sono necessarie maggiori informazioni sulle politiche di dimensionamento dello swap come anche informazioni sulla massima dimensione di swap possibile con le varie versioni del kernel.
- Quanto è diffusa la corruzione del disco o del file system? Fino ad ora ho sentito solo di problemi causati da hardware critico.
- Sono necessarie notizie sulla velocità e sui dischi.
- Ci sono altri controller RAID compatibili con Linux?
- Quali strumenti rilevanti sono disponibili per il monitoraggio, la gestione ed il mantenimento?
- Sono necessari riferimenti generali sulle fonti di informazione, dovrebbe forse essere un documento separato?
- L'utilizzo di `/tmp` e `/var/tmp` è stato difficile da determinare, infatti quali programmi utilizzano quali directory, non è ben definito e maggiori informazioni sono richieste qui. Ancora, alla fine sembra essere chiaro che queste dovrebbero risiedere su drive fisici differenti al fine di aumentare il parallelismo.

17.3 Progetti di Lavoro Suggeriti

Ora e sempre la gente posta su comp.os.linux.*, cercando buone idee per i progetti. Ora ne elencherò alcune che mi vengono in mente e che sono pertinenti a questo documento. Dovrebbero essere postati anche piani riguardanti enormi progetti, come nuovi file system, al fine di trovare collaboratori o di vedere se qualcuno ci sta già lavorando.

I mezzi per pianificare

che possono automatizzare i contorni del progetto più velocemente, farebbero probabilmente un progetto di media grandezza, forse come un esercizio in programmazione di base.

I mezzi per partizionare

che prendono l'output del programma precedentemente menzionato e formattano i dischi in parallelo e applicano i collegamenti simbolici appropriati alla struttura della directory. Sarebbe generalmente meglio se questo fosse integrato nel software per l'installazione del sistema esistente. Il setup del partizionamento dei dischi utilizzato in Solaris è un esempio di cosa può servire.

Strumenti di sorveglianza

che controllano le dimensioni delle partizioni e avvertono prima che una partizione si riempia.

Strumenti di migrazione

che vi permettono di spostare vecchie strutture verso nuovi sistemi (ad esempio RAID). Questo potrebbe essere fatto probabilmente con script shell che controllano programmi di backup e sarebbe abbastanza semplice. Ancora una volta, accertatevi che sia sicuro e che i cambiamenti possano essere ripristinati.

18 Domande e risposte

Questa è solamente una collezione di quello che io penso siano le domande più comuni che la gente può fare. Contattatemi e muterò questa sezione in una vera e propria FAQ.

- D: Di quanti dischi fisici (alberini) ha bisogno un sistema Linux? R: Linux può girare bene su un unico disco (alberino). Avere abbastanza RAM (attorno a 32 MB e fino a 64 MB) per gestire lo swap è una scelta di prezzo/prestazione migliore che prendere un disco secondario. Un disco (E)IDE è generalmente più economico (ma un po' più lento) di uno SCSI.
- D: Ho un singolo disco, mi aiuterà questo HOWTO? R: Si, sebbene solo ad un grado inferiore. La sezione 6.8 (Posizionamento Fisico delle Tracce) vi fornirà del guadagno.
- D: Ci sono svantaggi in questo schema? R: C'è solo una piccola rogna: se anche solo una singola partizione si riempie, il sistema potrebbe smettere di lavorare correttamente. La gravità dipende ovviamente da quale partizione è interessata. Comunque questo non è difficile da tenere sotto controllo, il comando `df` vi da una buona visione della situazione. Controllate inoltre le partizioni di swap mediante `free` per assicurarvi che non state per saturare la memoria virtuale.
- D: OK, quindi dovrei dividere il sistema in quante più partizioni possibili per un disco solo? R: No, ci sono diversi svantaggi nel fare ciò. Prima di tutto la manutenzione diventa inutilmente complessa e ottenete molto poco. Infatti se le vostre partizioni sono troppo grandi, accederete su aree più larghe del necessario. Questo è un bilancio dipendente dal numero di dischi fisici che avete.
- D: Questo significa che più dischi permettono più partizioni? R: In qualche modo sì. Comunque, qualche directory non dovrebbe essere separata da root, controllate il file system standard per maggiori dettagli.
- D: Che devo fare se ho molti dischi che voglio utilizzare? R: Se avete più di 3-4 dischi dovreste considerare l'utilizzo di un qualche tipo di RAID. Comunque è una buona idea di mantenere root su una singola partizione senza RAID, controllate la sezione 6.1 (RAID) per maggiori dettagli.
- D: Ho installato l'ultimo Windows95 ma non posso accedere a questa partizione dal sistema Linux, cosa è sbagliato? R: È presumibile che tu stia utilizzando `FAT32` nella partizione windows. Sembra che Microsoft abbia deciso che noi abbiamo bisogno di ancora un altro sistema e questo è stato introdotto nell'ultima versione di Windows95, chiamata OSR2. Il vantaggio è che questo formato è meglio adatto a dischi larghi.

  Può interessare sapere che NT 4.0 della Microsoft non lo gestisce ancora.
- D: Non riesco a far coincidere la dimensione del disco e quella della partizione, manca qualcosa. Cosa è successo? R: È possibile che abbiate montato una partizione su un punto di montaggio che non era una directory vuota. I punti di montaggio sono directory e se non sono vuote il montaggio maschererà i contenuti. Se sommate vedrete che la quantità di spazio disco utilizzato in questa directory manca dal totale osservato.

  Per risolvere ciò potete fare il boot da un disco di recupero e controllare cosa si sta nascondendo nei punti di montaggio e rimuovere o spostarne i contenuti montando la partizione in questione su un punto di montaggio provvisorio. Potreste trovare utile avere punti di montaggio di emergenza sparsi già fatti.

- D: Non mi sembra che la partizione di swap sia in uso, come è successo? R: È possibile che non sia stato necessario fare swap, specialmente se avete molta RAM. Controllate i vostri file di log per controllare se avete saturato la memoria su di un punto o su un altro. In quel caso il vostro spazio di swap dovrebbe essere stato messo in uso. Se così non è, è possibile che o alla partizione di swap non è stato assegnato il numero giusto, o che non l'avete preparata con `mkswap` o che non avete fatto lo `swapon` o che non lo avete aggiunto al vostro `fstab`.

- D: Cos'è qusto nyx cui si accenna molte volte qui? R: È un sistema Unix gratuito con attualmente circa 10000 utenti. Lo utilizzo per le mie pagine web, per questo HOWTO come anche per una sorgente di idee per un setup su grossi sistemi UNIX. Sono parecchi anni che gira ed ha un setup abbastanza stabile. Per maggiori informazioni controllate la

  Nyx homepage `http://www.nyx.net`

  che vi dà anche informazioni su come ottenere il vostro account gratuito.

19 Pezzettini e Ritagli

Questa è praticamente una sezione dove ammucchio tutti i pezzetti che ancora non ho deciso dove mettere, ma che sento sia necessario saperne qualcosa. È una specie di area di transizione.

19.1 Partizione Swap: usarla o non usarla

In molti casi non avete bisogno di una partizione di swap, ad esempio se avete molta RAM, tipo più di 64 MB, e siete l'unico utente della macchina. In questo caso potete provare a procedere senza una partizione di swap e controllare in ogni momento se avete mai saturato la memoria virtuale.

Rimuovere le partizioni di swap ha due vantaggi:

- risparmiate spazio disco (in realtà abbastanza ovvio)
- risparmiate tempo di accesso alle partizioni swap altrimenti giacerebbero al centro del vostro spazio disco.

In fine, avere una partizione di swap è come avere un bagno riscaldato: non lo usate molto spesso, ma sicuramente lo apprezzerete quando ne avete bisogno.

19.2 Punto di montaggio e `/mnt`

Nelle versioni più recenti di questo documento ho proposto di mettere tutte le partizioni montate in maniera permanente sotto `/mnt`. Il che, comunque, non è che sia proprio una buona idea visto che essa può essere utilizzata come un punto di montaggio, che guida a tutte le partizioni montate che diventano disponibili. Invece proporrò di montare direttamente da root mediante un nome significativo come `/mnt.nome-descrittivo`.

Recentemente sono diventato conscio del fatto che qualche distribuzione Linux utilizza punti di montaggio su subdirectory *sotto* `/mnt`, come `/mnt/floppy` e `/mnt/cdrom`, che mostra solamente quanto confusa sia l'intera faccenda. Speriamo che FHS chiarisca tutto ciò.

19.3 Alimentazione e Riscaldamento

Non molti anni fa, una macchina con la potenza di un moderno PC richiedeva alimentazione e raffreddamento a 3 fasi, generalmente mediante aria condizionata nella stanza della macchina, alcune volte anche raffreddando ad acqua. La tecnologia si è evoluta velocemente portando non solo alta velocità ma anche componenti a basso consumo. Comunque, c'è un limite definito dalla tecnologia, qualcosa che uno dovrebbe tenere a mente visto che il sistema si espande con ancora un altro disco o un'altra scheda PCI. Quando l'alimentazione gira a pieno regime tenete presente che tutta questa energia sta andando da qualche parte, per lo più in calore. Se questo non viene dissipato attraverso ventole, avrete un serio riscaldamento nel cabinet seguito da un'affidabilità ridotta ed anche da una riduzione della vita delle elettroniche. I produttori stabiliscono un minimo di necessità di raffreddamento per i loro dischi, generalmente in termini di piedi cubi al minuto (CFM). Siete tutti avvisati di prenderlo seriamente in considerazione.

Mantenete i passaggi di aria aperti, pulite la polvere e controllate la temperatura del vostro sistema in funzione. Se è troppo caldo da toccare, probabilmente gira in condizioni di calore eccessivo.

Se possibile utilizzate spin up sequenziali per i dischi. È durante l'avvio, quando le piastre dei dischi accelerano fino a velocità normale che un disco consuma il massimo dell'energia e se tutti i dischi partono simultaneamente potreste andare oltre il massimo della vostra alimentazione.

19.4 Dejanews

Questo è un sistema Internet che indubbiamente è familiare a molti di voi. Esso cerca e fornisce articoli *Usenet* dal 1995 fino ai più recenti messaggi e offre inoltre un'interfaccia web di lettura e scrittura. C'è molto altro, controllate

Dejanews `http://www.dejanews.com`

per maggiori informazioni.

Ciò che forse è meno risaputo, è che utilizzano circa 120 computer Linux SMP molti dei quali utilizzano per questo servizio il modulo `md` per gestire dai 4 ai 24 Giga di spazio disco (più di 1200 Giga tutti insieme). Il sistema cresce in continuazione ma al momento di scrivere essi utilizzano generalmente dual Pentium Pro 200MHz e sistemi Pentium II 300 MHz con 256 MB di RAM o più.

Una macchina per production database ha generalmente un disco per il sistema operativo e tra i 4 ed i 6 gestiti dal modulo `md` dove gli articoli sono archiviati. I dischi sono connessi ad adattatori PCI SCSI BusLogic Model BT-946C e BT-958

Per i sistemi di produzione (che sono accesi 365 giorni l'anno) il tempo di interruzione dovuto ad errori di disco è inferiore allo 0.25% (che è un quarto di 1%, non di 25%).

Solo una cosa: questa non è pubblicità, è inserito come un esempio di quanto è richiesto per quel che è un servizio Internet principale.

20 Appendice A: Tabella della Struttura del Partizionamento: Montare e Linkare

La tabella seguente è progettata per fare della struttura un semplice esercizio di carta e matita. È probabilmente meglio stamparlo (utilizzando caratteri non scalabili) ed aggiustarne i numeri fino a che non ne siete soddisfatti.

I punti di montaggio sono le directory dove vorreste montare le partizioni o il dispositivo reale. Questo è anche il posto giusto dove annotare come avrete intenzione di utilizzare i link simbolici.

La dimensione data corrisponde ad una installazione di Debian 1.2.6 abbastanza grande. Altri esempi seguiranno più avanti.

Principalmente utilizzate questa tabella per selezionare quali strutture e dischi utilizzerete, i numeri di partizione e le lettere arriveranno dalle prossime 2 partizioni.

```
Directory        Punto Montag.    veloc.   accesso trasferimento    dim.     DIM


swap             __________       ooooo    ooooo    ooooo           32       ____

/                __________       o        o        o               20       ____

/tmp             __________       oooo     oooo     oooo                     ____

/var             __________       oo       oo       oo              25       ____
/var/tmp         __________       oooo     oooo     oooo                     ____
/var/spool       __________                                                  ____
/var/spool/mail  __________       o        o        o                        ____
/var/spool/news  __________       ooo      ooo      oo                       ____
/var/spool/____  __________       ____     ____     ____                     ____

/home            __________       oo       oo       oo                       ____

/usr             __________                                         500      ____
/usr/bin         __________       o        oo       o               250      ____
/usr/lib         __________       oo       oo       ooo             200      ____
/usr/local       __________                                                  ____
/usr/local/bin   __________       o        oo       o                        ____
/usr/local/lib   __________       oo       oo       ooo                      ____
/usr/local/____  __________                                                  ____
/usr/src         __________       o        oo       o               50       ____

DOS              __________       o        o        o                        ____
Win              __________       oo       oo       oo                       ____
NT               __________       ooo      ooo      ooo                      ____

/mnt._________   __________       ____     ____     ____                     ____
/mnt._________   __________       ____     ____     ____                     ____
/mnt._________   __________       ____     ____     ____                     ____
/_____________   __________       ____     ____     ____                     ____
/_____________   __________       ____     ____     ____                     ____
/_____________   __________       ____     ____     ____                     ____



Capacita totale:
```

21 Appendice B: Tabella della Struttura del Partizionamento. Numerazione e dimensionamento

Questa tabella segue la stessa struttura logica della tabella precedente in cui avete deciso che disco utilizzare. Qui voi selezionate il physical tracking, ricordando l'effetto del posizionamento delle tracce menzionato in precedenza in 6.8 (Posizionamento Fisico delle Tracce).

Il numero della partizione finale verrà fuori dalla tabella successiva.

```
  Disco           sda     sdb     sdc     hda     hdb     hdc     ___

SCSI ID         |  __   |  __   |  __   |

Directory
swap            |       |       |       |       |       |       |

/               |       |       |       |       |       |       |

/tmp            |       |       |       |       |       |       |

/var            :       :       :       :       :       :       :
/var/tmp        |       |       |       |       |       |       |
/var/spool      :       :       :       :       :       :       :
/var/spool/mail |       |       |       |       |       |       |
/var/spool/news :       :       :       :       :       :       :
/var/spool/____ |       |       |       |       |       |       |

/home           |       |       |       |       |       |       |

/usr            |       |       |       |       |       |       |
/usr/bin        :       :       :       :       :       :       :
/usr/lib        |       |       |       |       |       |       |
/usr/local      :       :       :       :       :       :       :
/usr/local/bin  |       |       |       |       |       |       |
/usr/local/lib  :       :       :       :       :       :       :
/usr/local/____ |       |       |       |       |       |       |
/usr/src        :       :       :       :

DOS             |       |       |       |       |       |       |
Win             :       :       :       :       :       :       :
NT              |       |       |       |       |       |       |

/mnt.___/_____  |       |       |       |       |       |       |
/mnt.___/_____  :       :       :       :       :       :       :
/mnt.___/_____  |       |       |       |       |       |       |
/_____________  :       :       :       :       :       :       :
/_____________  |       |       |       |       |       |       |
/_____________  :       :       :       :       :       :       :


Capacita Totale:
```

22 Appendice C: Tabella della Struttura del Partizionamento: Posizionamento delle Partizioni

Questa serve solamente per ordinare i numeri delle partizioni in ordine crescente, pronti per essere immessi in fdisk o cfdisk. Qui prendete in considerazione il posizionamento fisico delle tracce quando finalizzate il vostro progetto. Se non avete informazioni specifiche, potete assumere che la traccia 0 sia la traccia più esterna.

Questi numeri e lettere sono poi utilizzati per aggiornare le tabelle precedenti, che vi risulteranno tutte molto utili nella manutenzione futura.

In caso di crash del disco, potreste trovare utile sapere quale SCSI id appartiene a quale drive, considerate di tenere una copia di ciò.

```
              Disco :   sda     sdb     sdc     hda     hdb     hdc     ___

     Capac.  totale: |  ___  |  ___  |  ___  |  ___  |  ___  |  ___  |  ___
     SCSI ID         |  __   |  __   |  __   |

     Partizione

     1               |       |       |       |       |       |       |
     2               :       :       :       :       :       :       :
     3               |       |       |       |       |       |       |
     4               :       :       :       :       :       :       :
     5               |       |       |       |       |       |       |
     6               :       :       :       :       :       :       :
     7               |       |       |       |       |       |       |
     8               :       :       :       :       :       :       :
     9               |       |       |       |       |       |       |
     10              :       :       :       :       :       :       :
     11              |       |       |       |       |       |       |
     12              :       :       :       :       :       :       :
     13              |       |       |       |       |       |       |
     14              :       :       :       :       :       :       :
     15              |       |       |       |       |       |       |
     16              :       :       :       :       :       :       :
```

23 Appendice D: Esempio: Server Multifunzionale

La tabella seguente viene da un setup di un server multifunzionale di media grandezza dove lavoro. Oltre che essere una macchina generale Linux, esso sarà anche un server di rete (DNS, posta, FTP, news, stampanti ecc.), un server X per vari programmi CAD, masterizzatori e molte altre cose. I file risiedono su 3 dischi SCSI con capacità di 600, 1000 e 1300 MB.

Un po' di ulteriore velocità potrebbe essere ottenuta separando `/usr/local` dal resto del sistema `/usr` ma abbiamo ritenuto che l'ulteriore complessità aggiunta, non sarebbe valsa lo sforzo. Con un altro paio di dischi sarebbe stato più vantaggioso. In questa configurazione, sda è vecchio e lento e potrebbe essere benissimo sostituito da un disco IDE. Gli altri due dischi sono entrambi sufficientemente veloci. Praticamente abbiamo diviso la maggior parte del carico tra questi due. Per ridurre i pericoli di sbilanciamento nel dimensionamento delle partizioni abbiamo deciso di mantenere `/usr/bin` e `/usr/local/bin` in un solo disco e `/usr/lib` e `/usr/local/lib` su un altro disco separato, il che inoltre ci permette un po' di parallelizzazione.

Potrebbe essere guadagnato anche di più utilizzando RAID, ma abbiamo avuto la sensazione che come server avevamo bisogno di più affidabilità di quella che permetteva la patch `md` e un controller RAID dedicato era lontano dalla nostra portata.

24 Appendice E: Esempio: Montare e Linkare

```
Directory       Punto di mont.  veloc.  accesso trasferimento   dim.    DIM.


swap            sdb2, sdc2      ooooo   ooooo   ooooo           32      2x64

/               sda2            o       o       o               20       100

/tmp            sdb3            oooo    oooo    oooo                     300

/var            __________      oo      oo      oo                      ____
/var/tmp        sdc3            oooo    oooo    oooo                     300
/var/spool      sdb1                                                     436
/var/spool/mail __________      o       o       o                       ____
/var/spool/news __________      ooo     ooo     oo                      ____
/var/spool/____ __________      ____    ____    ____                    ____

/home           sda3            oo      oo      oo                       400

/usr            sdb4                                            230      200
/usr/bin        __________      o       oo      o               30      ____
/usr/lib        -> libdisk      oo      oo      ooo             70      ____
/usr/local      __________                                              ____
/usr/local/bin  __________      o       oo      o                       ____
/usr/local/lib  -> libdisk      oo      oo      ooo                     ____
/usr/local/____ __________                                              ____
/usr/src        ->/home/usr.src o       oo      o               10      ____

DOS             sda1            o       o       o                        100
Win             __________      oo      oo      oo                      ____
NT              __________      ooo     ooo     ooo                     ____

/mnt.libdisk    sdc4            oo      oo      ooo                      226
/mnt.cd         sdc1            o       o       oo                       710


Capacita Totale: 2900 MB
```

25 Appendice F: Esempio: Numerazione e Dimensionamento

Qui facciamo gli aggiustamenti delle dimensioni e del posizionamento.

```
Directory         sda     sdb     sdc


swap            |       |  64   |  64   |

/               |  100  |       |       |

/tmp            |       |  300  |       |

/var            :       :       :       :
```

```
/var/tmp        |       |       | 300   |
/var/spool      :       :  436  :       :
/var/spool/mail |       |       |       |
/var/spool/news :       :       :       :
/var/spool/____ |       |       |       |

/home           |  400  |       |       |

/usr            |       |  200  |       |
/usr/bin        :       :       :       :
/usr/lib        |       |       |       |
/usr/local      :       :       :       :
/usr/local/bin  |       |       |       |
/usr/local/lib  :       :       :       :
/usr/local/____ |       |       |       |
/usr/src        :       :       :       :

DOS             |  100  |       |       |
Win             :       :       :       :
NT              |       |       |       |

/mnt.libdisk    |       |       | 226   |
/mnt.cd         :       :       : 710   :
/mnt.___/_____  |       |       |       |


Capac. Totale:  |  600  | 1000  | 1300  |
```

26 Appendice G: Esempio: Posizionamento delle Partizioni

Questo è solo per ordinare i numeri delle partizioni in ordine ascendente pronti per essere immessi in fdisk o cfdisk. Ricordate di ottimizzare il posizionamento fisico delle tracce (non fatto qui).

```
        Disco :    sda     sdb     sdc

Capac. Totale : |   600 |   1000 |   1300 |

Partizione

1               |   100 |    436 |    710 |
2               :   100 :     64 :     64 :
3               |   400 |    300 |    300 |
4               :       :    200 :    226 :
```

27 Appendice H: Esempio II

Il seguente è un esempio di una predisposizione di un server con impostazioni accademiche ed è un contributo di `nakano (at) apm.seikei.ac.jp`. Ho fatto piccolissime modifiche a questa sezione.

`/var/spool/delegate` è una directory per archiviare i file di log e i file di cache di un programma per server proxy WWW, delegato. Dal momento che non lo noto ampiamente, attualmente ci sono 1000–1500 richieste al giorno e l'utilizzo medio del disco è del 15–30% con scadenza giornaliera delle cache.

`/mnt.archive` è utilizzato per i file che sono grandi e cui non si fa frequentemente riferimento come i dati sperimentali (specialmente quelli grafici), vari sorgenti di archivi e backup di Win95 (che crescono molto velocemente).

`/mnt.root` è un file system di backup che contiene utilità di recupero. Un boot floppy è inoltre preparato per fare un boot su questa partizione.

```
=================================================
Directory              sda      sdb     hda

swap                   |   64  |   64  |       |
/                      |       |       |   20  |
/tmp                   |       |       |  180  |

/var                   :  300  :       :       :
/var/tmp               |       |  300  |       |
/var/spool/delegate    |  300  |       |       |

/home                  |       |       |  850  |
/usr                   |  360  |       |       |
/usr/lib               -> /mnt.lib/usr.lib
/usr/local/lib         -> /mnt.lib/usr.local.lib

/mnt.lib               |       |  350  |       |
/mnt.archive           :       : 1300  :       :
/mnt.root              |       |   20  |       |

Capac. Totale:           1024    2034    1050


=================================================
        Disco :           sda     sdb     hda
Capac. Totale :        | 1024  | 2034  | 1050  |

Partizione
1                      |  300  |   20  |   20  |
2                      :   64  : 1300  :  180  :
3                      |  300  |   64  |  850  |
4                      :  360  :  ext  :       :
5                      |       |  300  |       |
6                      :       :  350  :       :


Filesystem         1024-blocks  Used Available Capacity Mounted on
/dev/hda1             19485   10534      7945     57%   /
/dev/hda2            178598      13    169362      0%   /tmp
/dev/hda3            826640  440814    343138     56%   /home
/dev/sda1            306088   33580    256700     12%   /var
/dev/sda3            297925   47730    234807     17%   /var/spool/delegate
/dev/sda4            363272  170872    173640     50%   /usr
/dev/sdb5            297598       2    282228      0%   /var/tmp
/dev/sdb2           1339248  302564    967520     24%   /mnt.archive
/dev/sdb6            323716   78792    228208     26%   /mnt.lib
```

Apparentemente `/tmp` e `/var/tmp` sono troppo grandi. Queste directory dovrebbero essere unite insieme in una partizione quando lo spazio disco diminuisce.

Anche `/mnt.lib` dovrebbe esistere, ma ho pianificato di installare archivi TeX e ghostscript più nuovi, quindi `/usr/local/lib` potrebbe crescere di circa 100 MB o giù di lì (visto che dobbiamo utilizzare caratteri giapponesi!).

Viene fatto un backup dell'intero sistema da un Seagate Tapestore 8000 (Travan TR-4, 4G/8G).

28 Appendice I: Esempio III: SPARC Solaris

La sezione seguente è il progetto base utilizzato al lavoro per una serie di server Sun SPARC su cui gira Solaris 2.5.1 in un ambiente di sviluppo industriale. Esso serve una serie di archivi e applicazioni cad oltre ai normali servizi come la posta.

La semplicità è qui enfatizzata quindi `/usr/lib` non è stata separata da `/usr`.

Questa è la struttura di base, pianificata per circa 100 utenti.

```
Disco:        SCSI 0                    SCSI 1

Partizione    Dim. (MB)   Punto di Mount Dim. (MB)   Punto di Mount

  0           160         swap           160         swap
  1           100         /tmp           100         /var/tmp
  2           400         /usr
  3           100         /
  4            50         /var
  5
  6           remainder   /local0        remainder   /local1
```

A causa di specifiche necessità in questa sede è a volte necessario avere la disponibilità di grandi partizioni con un breve preavviso. Quindi al disco 0 vengono conferiti più compiti possibili, mantenendo una grande partizione `/local1`.

Questa configurazione è stata in uso per qualche tempo ed ora la riteniamo soddisfacente.

Per un sistema più generale e bilanciato, sarebbe meglio dividere `/tmp` e `/var/tmp` e spostare `/var` al disco 1.

29 Appendice J: Esempio IV: Server con 4 Dischi

Questo fornisce un esempio di utilizzo di tutte le tecniche descritte prima, con un po' di RAID. Si deve riconoscere che è abbastanza complicato ma offre in compenso alta prestazione da un hardware modesto. Il dimensionamento è saltato ma descrizioni ragionevoli possono essere trovate in esempi precedenti.

```
Partizione      sda            sdb            sdc            sdd
                ----           ----           ----           ----
         1      root           overview       lib            news
         2      swap           swap           swap           swap
         3      home           /usr           /var/tmp       /tmp
         4                     spare root     mail           /var
```

La configurazione è ottimizzata nel rispetto del posizionamento delle tracce ma anche per mancanza di accessi ai dischi.

Se inoltre volete DOS o Windows dovrete utilizzare `sda1` per questo e spostare durante queste sessioni le altre partizioni su `sdb2`, `sdc2` e `sdd2` per lo swap di Windows, `TEMPDIR` e la directory temporanea di Windows. Una serie di altri HOWTO descrive come potete far coesistere diversi tipi di sistemi operativi sulla vostra macchina.

Per completezza viene fornito un esempio di 4 dischi con l'utilizzo di diversi tipi di RAID il che è anche molto più complesso dell'esempio sopra.

```
Partizione      sda             sdb             sdc             sdd
                ----            ----            ----            ----
        1       boot            overview        news            news
        2       overview        swap            swap            swap
        3       swap            lib             lib             lib
        4       lib             overview        /tmp            /tmp
        5       /var/tmp        /var/tmp        mail            /usr
        6       /home           /usr            /usr            mail
        7       /usr            /home           /var
        8       / (root)        spare root
```

Qui tutti i duplicati sono parti di un set RAID 0 con due eccezioni, swap che è lasciato da parte e home e mail che sono implementati come RAID 1 per sicurezza.

Notate che boot e root sono separati: solo il file di boot con il kernel deve risiedere entro il limite del cilindro 1023. Il resto della root può essere dovunque e qui sono piazzati sulla partizione più lenta e lontana. Per semplicità e sicurezza la partizione root non è su un sistema RAID.

Con una complessità tale ne consegue che avremo un file `fstab` egualmente complicato. Il grande numero di partizioni rende importante fare passaggi di `fsck` nell'ordine esatto, altrimenti il processo può metterci dieci volte il tempo per concludersi come la soluzione migliore.

```
/dev/sda8       /               ?       ?               1 1 (a)
/dev/sdb8       /               ?       noauto          1 2 (b)
/dev/sda1       boot            ?       ?               1 2 (a)
/dev/sdc7       /var            ?       ?               1 2 (c)
/dev/md1        news            ?       ?               1 3 (c+d)
/dev/md2        /var/tmp        ?       ?               1 3 (a+b)
/dev/md3        mail            ?       ?               1 4 (c+d)
/dev/md4        /home           ?       ?               1 4 (a+b)
/dev/md5        /tmp            ?       ?               1 5 (c+d)
/dev/md6        /usr            ?       ?               1 6 (a+b+c+d)
/dev/md7        /lib            ?       ?               1 7 (a+b+c+d)
```

Le lettere tra parentesi indicano quali dischi saranno attivi per ogni voce e passaggio di `fsck`. Queste lettere *non* sono presenti in un vero file `fstab`. In tutto ci sono 7 passaggi.

30 Appendice K: Esempio V: Sistema con Doppio Disco

Un sistema con doppio disco offre meno opportunità per schemi intelligenti ma il seguente dovrebbe fornire un semplice punto di partenza.

```
Partizione      sda             sdb
                ----            ----
        1       boot            lib
        2       swap            news
        3       /tmp            swap
        4       /usr            /var/tmp
        5       /var            /home
        6       / (root)
```

Se utilizzate un sistema con due SO dovete ricordarvi che molti altri sistemi devono fare il boot dalla prima partizione del primo disco. Un semplice sistema DOS / Linux potrebbe apparire così:

```
Partizione      sda             sdb
                ----            ----
        1       DOS             lib
        2       boot            news
        3       swap            swap
        4       /tmp            /var/tmp
        5       /usr            /home
        6       /var            DOSTEMP
        7       / (root)
```

Inoltre ricordate che DOS e Windows preferiscono essere la sola partizione che deve anche essere la prima da dove si fa il boot. Visto che Linux può esistere felicemente in partizioni logiche, questo non è un grande problema.

31 Appendice L: Esempio VI: Sistema con un Singolo Disco

Sebbene questo cada in qualche modo fuori da quello che è lo scopo di questo HOWTO, non può essere negato che recentemente qualche disco abbastanza grande è diventato molto accessibile. Dischi con 10-20 GB stanno diventando comuni ed è frequente la domanda su come partizionare questi mostri. È abbastanza interessante il fatto che molto pochi sembrano avere problemi nel riempire questi dischi ed il futuro sembra generalmente abbastanza roseo per i produttori che puntano su dischi ancora più grandi.

Opportunità per ottimizzazioni sono ovviamente anche più piccole che per i sistemi a due dischi ma qualche trucco può ancora essere utilizzato per ottimizzare il posizionamento delle tracce nel minimizzare i movimenti delle testine.

```
Partizione      hda             Dimensione Stimata (MB)
                ----            ------------------
         1      DOS             500
         2      boot            20
         3      Winswap         200
         4      data            La massa del disco
         5      lib             50 - 500
         6      news            300+
         7      swap            128     (Dimensione massima per una CPU a 32-bit)
         8      tmp             300+    (/tmp e /var/tmp)
         9      /usr            50 - 500
        10      /home           300+
        11      /var            50 - 300
```

```
12      mail        300+
13      dosdata     10      ( Si aggirano i bug di Windows!)
```

Ricordate che la partizione `dosdata` è un filesystem DOS che deve essere l'ultima partizione sul disco, altrimenti Windows si confonde.