# Sistemi 12

Nebula sono una serie di esercizi sulla sicurezza informatica, questa lezione sarà di esercizi su programmi fallaci.

Andiamo a vedere l’ultimo argomento, la persistenza. Cosa sono i dispositivi a blocchi/i file system.

Chiamiamo comunemente “dischi” (magnetici/ottici/pennette usb/SSD) i dispositivi a blocchi di memoria secondaria. Questi dischi lavorano a blocchi: mentre in RAM possiamo toccare singoli byte a piacere, in questi dispositivi non possiamo leggere o scrivere un solo byte ma solo un intero blocco. Tipicamente, quindi, la granularità dei dischi è di circa 512 byte.

Se un disco è magnetico i blocchi sono logici, creati dalla formattazione a basso livello (da decenni questa formattazione a basso livello è già fatta).

Una volta i dischi erano veramente dischi, cioè oggetti piatti e di forma circolare. Dunque i dati sono letti come su un giradischi e su questo i dati subito dopo la testina di lettura sono molto più veloci da leggere di quelli subito prima, poiché per prendere quelli c’è bisogno di attendere un ritardo di rotazione. Ha senso leggere un intero blocco, perché se dovessimo leggere un byte dopo l’altro e poi ce ne servisse uno prima, ciò ci farebbe perdere parecchio tempo, è più conveniente leggere pezzi di disco in una volta.

Per questo motivo, prima dell’SSD, i file non erano frammentati nel disco, visto che se si accedeva a un file era probabile che poi si sarebbe acceduto ad altri punti del file e quindi era conveniente di tenere i pezzi dei file attaccati.

Sulle SSD si usano sempre i blocchi logici, ma su di essa, per questioni fisiche, essi vengono spostati in giro nel disco perché ogni sezione ha un numero di scritture limitato.

Un altro motivo per cui si indirizza la memoria secondaria a blocchi è per risparmiare sul numero di byte necessari per indirizzare i dati sul disco tramite RAM (il disco sarà ovviamente più grande della RAM quindi si rischierebbe di usare troppi bit per indirizzarlo se si cercassero singoli byte).

Un moderno SSD non è rotondo, non ruota, ma l’interfaccia hardware (a blocchi) è la stessa: un sistema operativo potrebbe usare un SSD pensando di star usando un hard disk, anche se ovviamente sarebbe conveniente che lo notasse visto che così potrebbe ottimizzare le operazioni.

L’interfaccia che considereremo sono dischi di n blocchi indirizzabili da 0 a n-1.

L’I/O su disco è ordini di grandezza più costoso dell’accesso in RAM (questo include anche i veloci SSD); quindi è furbo tenere i blocchi del disco letti in una cache, in caso quella porzione di file serva a un altro processo (analogo discorso per le scritture, meglio “bufferizzare” la scrittura in RAM/Cache e poi fare un’unica scrittura dopo un po’ di tempo). Questo è il motivo per cui i sistemi Unix usano una (unified) buffed cache. Una volta i blocchi di questa buffer-cache erano usati per i blocchi del disco acceduti tramite read/write e una page-cache per i file mappati in memoria (che si appoggiava sulla buffer-cache).

Su Linux per Forzare la scrittura di dati in cache si usa sync.

La ragione di memorie cache è la ragione per cui è importante espellere i drive prima di scollegarli (es. rimozione sicura dell’usb).