

新 医療IT用語辞典

医療ITに関する専門書に出てくる専門用語の数々。ITベンダーがシステムを解説する際に使用する特殊な言い回し。なんとなく理解しているようにも、100%自信があるわけではない。失礼ながら、ほとんどの読者の方々にそのような経験があるわけではない。この連載は基本的なものに絞り、それらの正しい理解を目指すもの。今回は、個人のパソコン操作でも一般的に使う用語ではあるものの、医療情報システムでの深い意味についてはご存知ないのではないか。



今月のIT用語▼バックアップ

解説▼国立成育医療研究センター 山野辺 裕二

今月は「バックアップ」を取り上げます。英和辞典に載っている意味としては、支援・支援者、代役・代替物というものがあります。ITの世界で使う場合、多くはコンピューターに記憶しているプログラムやデータを別の記憶媒体に保存することをバックアップと呼びます。2011年の東日本大震災の際には、病院同士でバックアップを持ち合っていたおかげで、津波でシステムが使えなくなっても、相

手病院のバックアップから必要なデータを戻せた事例がありました。

バックアップの範囲と頻度

普段使っているパソコンや携帯電話に保存したデータが突然消えてしまったら大変困ります。そうならないように

データの保全のためのさまざまな方法を用意するわけですが、個人レベルにおいては、その熱意は人それぞれという感じがします。

バックアップには様々な分類や種類がありますが、それを決めるパラメーターの代表が、バックアップの範囲と頻度です。範囲の例としては、データ、機器、システムといったものがあります。データのバックアップは、パソコンで文書やデータを作成して保存する時に、パソコン本体の内蔵ディスクと外付けのUSBメモリの2カ所に保存するといった例があります。このようにすることで、内蔵ディスクの故障時にもデータを失うことがありません。

ただしパソコンの電源が入らなくなるような故障の場合は、データだけあっても使えません。同じパソコンをもう一台待機させておくと、仕事ができない期間を最小限に

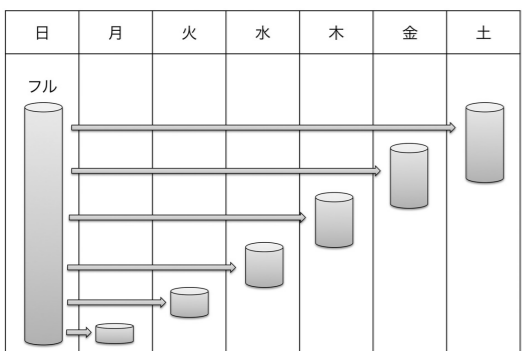
多様化するバックアップ媒体

ビジネスの世界でのバックアップのメディアとしては、磁気テープが一般的です。テープは容量あたりの単価は安いのですが、データ転送速度が低いため、書き戻しに時間がかかる欠点があります。このデータを書き戻すことをリストアと呼びます。バックアップの方は夜間など仕事が少ない時間帯に行うことができますが、リストアの場合はあらかじめ時間を決められない場合も多いので時間がかかる

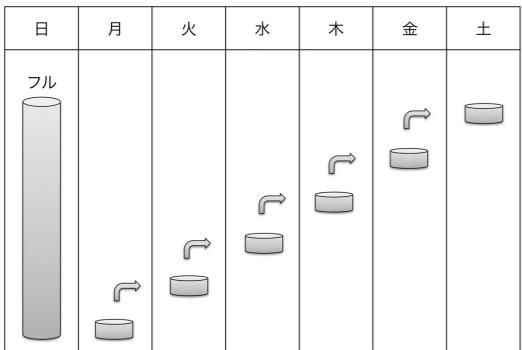
と困ります。その課題を解決するのがD to Dバックアップです。バックアップ媒体として、テープより遙かに高速なハードディスクを使います。バックアップの速度は速いものの、コストや故障率はテープの方が有利なため、D to D to Tといって、バックアップしたディスクから更にテープに移すことでディスクとテープの欠点を補い合う手法もあります。

D to Dバックアップの利点の一つは、重複排除(deduplication)という技術によりバックアップ媒体の容量を節約できることです。バックアップ対象がパソコンであればOS部分は共通ですし、医療分野でも定型文書がたくさんあります。例えば紹介状の外枠の部分は常に同じなので、例えば数千件の文書であっても、定型部分は一っだけ保存しておけば事足りるわけです。

【図1】差分バックアップ



【図2】増分バックアップ



もっとも注意すべきことは人為ミス

コンピューターに保存したデータが失われる要因としては、次のようなものがあるとされています。

- 1 ストレージ機器のハードウェア障害
- 2 操作ミス
- 3 地震・火災・水害等の災害
- 4 ウィルスなど悪意のプログラム
- 5 アプリケーションのバグ等

このうちで最も多いのは人による操作ミスと言われ、私も次のような経験があります。その日は機器のトラブルで、バックアップの記憶装置からデータを戻そうとしていました。その時に「リストア」でなく、誤って「バックアップ」のボタンを押してしまい、戻そうとしていたデータを上書き消去してしまっただけでなく、バックアップ用機器に

昨年の6月には、あるサーバー請負会社が預かった企業のデータを大規模に消去してしまった事件がありました。これも人為的なミスで、作業中に誤ってファイル削除の命令を出してしまっただけでなく、バックアップ用機器に

きるでしょう。これが機器のバックアップですが、余分な費用がかかります。

更に高度なところでは、待機系のシステムを普段から動かしておき、通常使っているシステムに異常が起ると自動的に切り替えて運用を続けるシステムもあります。このようなものを、システムの二重化とかクラスター構成などと呼びます。

病院の電子カルテシステムなど、複数の機器からなるシステム全体のデータをバックアップすることもあります。ただ、以前取り上げたアーカイブと異なるのは、通常バックアップといえば保存する対象の選別をしないということです。電子カルテなどの病院情報システムは、物理的にはデータを患者別に保存しているわけではないので、「内科の患者のみバックアップ」といったことはできません。アーカイブは保存対象を選別して行うので、「死亡患者のカルテのみアーカイブ」といったことが可能なのです。

バックアップの頻度としては、毎月、毎日、リアルタイムといった例があります。頻度が高いほど必要な資源は大きくなりますが、障害などでデータを失ったときの被害が少ないこととなります。

バックアップの手法には大きく分けてフルバックアップ、差分バックアップ、増分バックアップがあります。フルバックアップは対象すべてをバックアップします。それに対し、差分バックアップは最後のフルバックアップとの差を保存します(図1)。増分バックアップは、前回のバックアップとの差を保存します(図2)。この両者は一見どこが違うのかという感じを受けます。増分バックアップは、毎回のバックアップ量が少なく短い時間で済むという利点があります。図2の木曜のデータを作るには、日月火水木の5つを順番にリストアする必要があります。逆に差分バックアップは、毎回のバックアップ量がだんだん増えるため次第に時間を要するようになりますが、リストア時には常に2つのデータで済むという利点があります。対象データの変化量に合わせて、最適な組み合わせを選ぶことになります。

まで同じ削除コマンドを送ってしまったというものでした。普通考えられないようなミスでも現実には起こりうることを実感させられました。

以上のことから、バックアップにはさまざまな仕組みと頻度を複数組み合わせることが重要です。

結論

コンピューターに記憶しているプログラムやデータを別の記憶媒体に保存することをバックアップと呼びます。手段や媒体が多様化しており、目的に合致した手法を多重に用いることが必要です。

今月のおまけはレプリケーションです。これは同じデータを2つのところに同時に書き込むというもので、リアルタイムのバックアップに似ており、RAIDという仕組みでのディスク装置の冗長化もその一種と言えるでしょう。理屈は分かりやすいですが、遠隔地への同時書き込みは回線速度などの問題があり、やや複雑になります。即時同時に書き込むということは、操作ミスによるデータ消去などに対しては無力であること、「一週間前」といった過去の時点に戻ることが難しいという欠点もあります。



山野辺 裕二
(やまのべ・ゆうじ)

国立成育医療研究センター
情報管理部 情報解析室長

62年長崎県生まれ。86年長崎大医卒。91年同大大学院修了。愛媛県立中央病院、国立佐賀病院、長崎大病院、米国マウントサイナイメディカルセンター(医療情報学客員研究員)等を経て、05年国立成育医療センター臨床研究開発部医療情報室長、11年5月より現職。