

システム性能評価法レポート

1. キャッシュ

CPU キャッシュ、ディスクキャッシュ等と比較したときに WWW キャッシュにはどのような特徴があるといえるか説明せよ。

1.1 キャッシュとは

アクセスが低速な記憶装置に記録されているデータのうち、使用頻度の高いデータを高速な記憶装置に蓄えておくことにより、低速な装置から読み出す無駄を省いて高速化すること。および、その際に使われる高速な記憶装置。メインメモリのキャッシュを CPU キャッシュ、ハードディスクのキャッシュをディスクキャッシュと呼ぶ。

例えば、ハードディスクに比べ、メモリは何百倍も高速にデータにアクセスできる。そこで、使用頻度の高いデータをメモリ内に保持しておくことで、ハードディスクにすべてのデータを置いた場合よりも処理を高速化できる。この場合、メモリがハードディスクのキャッシュになる。

同様の手法は通信においても利用でき、これが WWW キャッシュにあたる。低速な通信回線を使って読み込んだデータをハードディスクに蓄えておくことで、次の高速なデータ閲覧を可能にしている。

1.2 WWWキャッシュ

WWW キャッシュは、CPU キャッシュやディスクキャッシュと比較した際の特徴が、大きく 2 つある。

1.2.1 目的

1 つ目はその目的である。CPU キャッシュやディスクキャッシュはアクセスの高速化が目的であるが、WWW キャッシュはそれだけが目的ではない。一般的に Web 上のページを閲覧するときには、クライアントの要求がインターネットを介しサーバに届き、その応答をサーバがクライアントに返すという動作をしている。

ここで、WWW キャッシュを実現することにより、閲覧頻度の高いデータをキャッシュに保持しておくことで、クライアントの要求がインターネット上に流れることなくローカルの範囲で処理できる。こうすることで WWW キャッシュは、Web ページを公開するサーバの負荷軽減とインターネット上のトラフィックの減少という役割を果たしている。

つまり、WWW キャッシュは、クライアントの高速なデータ閲覧、サーバの負荷軽減、トラフィックの現象という 3 つの目的を持つことがわかる。

1.2.2 用途

2 つ目はその用途である。WWW キャッシュは、用途によって設置する場所が 2 通り存在する。

1 つはブラウザ上のキャッシュで、ブラウザ自身に Web ページのデータを保持しておくことでアクセスの高速化を実現している。同一のユーザが同一のページを閲覧する場合に非常に有効である。

もう 1 つはプロキシサーバ上のキャッシュで、プロキシサーバに Web ページのデータを保持しておく。こうすることで、同じプロキシサーバを利用する複数のユーザが同じページを閲覧する場合に非常に有効である。

2. RAID

耐障害性が必要なシステムにおいて RAID 0 と 1 が選択可能なとき、どちらを選択することが望ましいか、理由とともに述べよ。

2.1 RAIDとは

RAID は、Redundant Arrays of Inexpensive Disks の略語である。ディスクアレイの代表的な実装形態で、複数のディスクを組み合わせ、仮想的な 1 台のハードディスクとして運用する技術のことを指す。RAID は RAID 0 から RAID 5 までの 6 種類が定義されているが、はじめに定義された 6 種類のうち、今日では RAID 0、RAID 1、RAID 5、およびこれら 3 方式の組み合わせが用いられている。

2.2 RAIDの種類

2.2.1 RAID 0

RAID 0 はストライピングと呼ばれ、複数台のハードディスクに、データを分散して読み書きし高速化したものである(図 1)。RAID 0 はデータの読み書き速度は向上するものの、冗長性がなく耐障害性もないため信頼性は低下する。

1 台のドライブが故障しただけでアレイ内の全データが失われてしまう。そのため故障率は単体ドライブに比べ高く、2 台で RAID 0 を構成した場合、稼働率は低下する(図 2)。

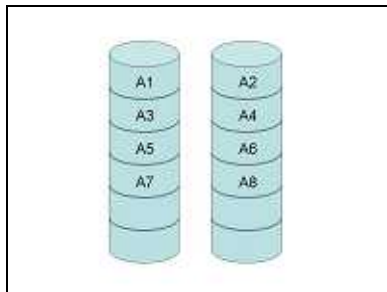


図 1 : ストライピング

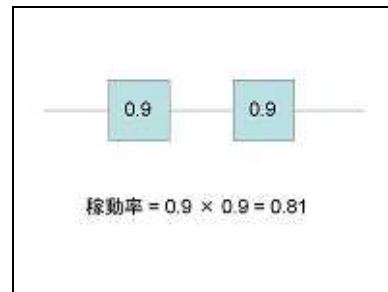


図 2 : RAID 0 の稼働率

2.2.2 RAID 1

RAID 1 はミラーリングと呼ばれ、複数台のハードディスクに同時に同じ内容を書き込むことで耐障害性を高めた RAID である(図 3)。つまり、耐障害性が必要なシステムにおいては、RAID 1 を選択することが望ましい。

一台が故障した際に、もう一方も同時に故障する可能性は低く、システムは稼働し続けることが出来るため、2 台で RAID 1 を構成した場合、稼働率は向上する(図 4)。

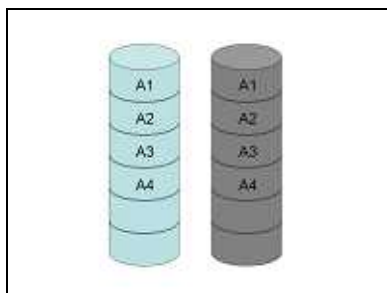


図 3 : ミラーリング

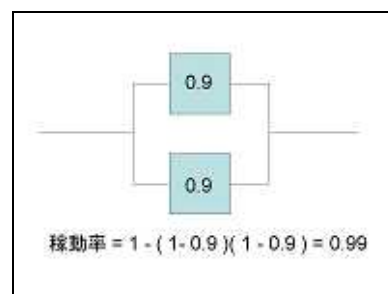


図 4 : RAID 1 の稼働率

3. ターンアラウンドタイムとレスポンスタイム

ターンアラウンドタイムとレスポンスタイムの違いについて述べよ。

3.1 ターンアラウンドタイム

ターンアラウンドタイム (Turn Around Time) とは、ジョブをコンピュータシステムに投入してから、完全な処理結果を出力するまでの時間のことである。バッチ処理の処理能力の指標に使われ、以下の式で表すことができる。

$$\text{ターンアラウンドタイム} = \text{処理待ち時間} + \text{CPU 時間} + \text{入出力時間}$$

3.2 レスポンスタイム

レスポンスタイム (Response Time) とは、応答時間とも呼ばれ、処理要求を送ってから出力が始まるまでの時間のことである。オンライントランザクション処理の処理能力の指標に使われる。ターンアラウンドタイムとの違いを以下に示す (図 5)。

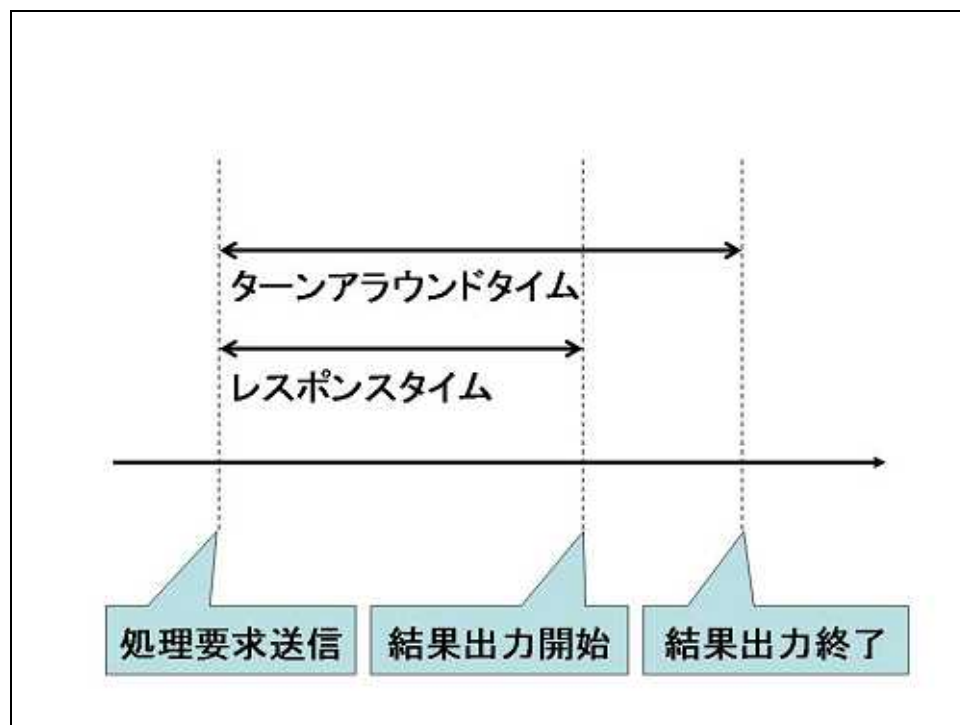


図 5 : 処理の流れと時間

4. 信頼性

MTBF、MTTR、稼働率という指標と RASIS との関係を説明せよ。また、この3つの指標の関係について説明せよ。

4.1 RASIS

RASIS とは、コンピュータシステムの信頼性を評価する基準であり、以下の5つの指標の頭文字を取ったものである。

- ・ Reliability【信頼性】 後述の『MTBF』に当たる。
- ・ Availability【可用性】 後述の『稼働率』に当たる。
- ・ Serviceability【保守性】 後述の『MTTR』に当たる。
- ・ Integrity【保全性】
- ・ Security【機密性】

4.2 MTBF

Mean Time Between Failures の略で、平均故障間隔を示す。

システムが故障せずに連続動作する時間の平均値。

RASIS の R (Reliability : 信頼性) の指標で、値が大きいほど信頼性が高い。

MTBF = システムの総稼働時間 / 故障回数

4.3 MTTR

Mean Time To Repair の略で、平均修理時間を示す。

故障が発生したときに要する時間の平均値。

RASIS の S (Serviceability : 保守性) の指標で、値が小さいほど保守性が高い。

MTTR = システムの総故障時間 / 故障回数

4.4 稼働率

ある時点においてシステムが稼働している確率を示す。

RASIS の A (Availability : 可用性) の指標で、値が1に近いほど可用性が高い。

稼働率は、以下のように MTBF と MTTR を用いて算出することができる (図6)。

$$\text{稼働率} = \frac{\text{総稼働時間}}{\text{全体時間}} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

図6 : 稼働率、MTBF、MTTR の関係式