

【iostat でのデバイス使用状況の確認】

と

「do file sequential read 待機イベント」、
「do file scattered read 待機イベント」が多発の対処方法

原因

「do file sequential read」、「do file scattered read」待機イベントが多発する場合は、ディスク・アクセスが長時間に亘^{わた}っているからと考えられる
以下の追加調査にて、「データ量が多い」か、「デバイス故障」からなのかを切り分ける

必要追加調査事項

1. Statspack レポートの統計情報の調査
 - ・ **アクセス I/O 量の多 or 少 の確定** (Load Profile セクションの Physical Read)
 - ・ アクセス量の多い表領域とディスクの特定 (Tablespace セクション)
 - ・ ~~単位時間アクセス I/O 量の長 or 短の確定~~
 - ・ 待機時間 (Elapsed Time - CPU Time) が長い SQL 文を特定する (SQL orderd by Reads、SQL orderd by Gets セクション)
2. IOSTAT コマンドによるディスク・デバイスの統計情報の調査
 - ・ 単位時間アクセス I/O 量の多 or 少 の調査
 - ・ **1 アクセスの処理時間の長短調査**
 - ・ ディスク稼働中 (ビジー) 率の高低調査

調査結果とその対応方針

【アクセス I/O 量の多く、ビジー率が高い場合】

アクセス量が多く過ぎてしまうと、レスポンスは悪化してしまう。

デバイスに対するアクセス量を少なくする対策を取る

- 例 実行計画の変更によるアクセス量の減少
ページング処理が発生している場合には、リソース追加
物理ディスク追加とファイル再配置

【アクセス I/O 量の少なく、レスポンスが悪い場合】

ディスク・デバイス故障が考えられる

【アクセス I/O 量の少なく、レスポンスも良い場合】

ディスク・デバイスには問題なし

他の要因が SQL 文の処理の長時間化になっている

- 例 排他ロックなどの制御ロック取得待ち
ネットワークなどの他のデバイス故障

iostat 取得コマンド

\$ iostat -xtc 取得間隔(秒) 取得回数

I/Oデバイスの実行効率と負荷率の把握

iostat 出力結果 表示

--- UNIX --- (P77)

Device	r/s	w/s	rk/s	wk/s	wait	actv	svc_t	%w	%b
sd0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0
sd1	102.2	0.0	9302.3	0.0	1.6	1.9	58.0	29	95
nfs0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0
nfs1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0

↑ 1秒あたりの読取り回数 ↑ 1秒あたりの読取り Kバイト数 ↑ I/Oにかかった時間(ms) ↑ Wait率 ↑ ディスク・ビジー率

Wait キューの長さ (待ち行列の数)

--- Linux ---

Time: 21:58:00

Linux 2.6.9-22.ELsmp (koda21.localdomain) 02/24/06

Time: 20:57:17

avg-cpu:	%user	%nice	%sys	%iowait	%idle	...	CPU 使用状態
	2.18	0.00	0.62	9.01	88.20		

↑ I/O 待機率 ↑ アイドル状態率

Device:	rrqm/s	wrqm/s	r/s	w/s	rsec/s	wsec/s	rkB/s	wkB/s
hd1	0.43	80.69	10.55	8.92	419.61	616.05	209.80	283.74
hd2	0.68	120.33	18.75	16.35	508.61	758.85	320.41	358.43
hd3	0.79	150.91	23.55	23.77	630.61	903.58	209.80	347.63

↑ 1秒あたりの読取り数

avgrq-sz	avgqu-sz	await	svctm	%util
58.37	0.44	22.33	5.33	30.37
75.53	0.84	22.33	5.01	41.77
88.21	1.93	92.33	7.43	90.37

↑ 平均待ち行列長 ↑ I/Oにかかった時間(ms) ↑ ディスク・ビジー率

実際の処理時間+待たされた時間

※ ディスク・ビジーの時は await 時間が長くなる

障害例の iostat

【ディスク負荷オーバーからのレスポンス悪化状態】

1秒あたりの読取り、書込み数が多い場合には、ディスクの作業負荷が大きくなる
このときに、1I/O 当りの処理時間が異様に長時間になったり、長時間ディスク・ビジー率が高い場合には、負荷オーバーの可能性はある

追加調査として、Disk I/O に関する待機イベントが多発していないか確認の必要がある。

Statspack レポート (Top 5 Times Events セクション)
v\$session、v\$session_wait

iostat 出力例

```
--- UNIX ---
Device      r/s    w/s    rk/s    wk/s    wait    actv    svc_t    %w    %b
sd0         0.0    0.0     0.0     0.0     0.0     0.0     0.0     0     0
sd1        102.2    0.0  9302.3     0.0     1.6     1.9     58.0    29    95

--- Linux ---
Device:     rrqm/s    wrqm/s    r/s    w/s    rsec/s    wsec/s    rkB/s    kB/s
hd1         0.43     80.69   210.55    8.92   419.61   616.05   209.80   283.74
           avgrq-sz  avgqu-sz   await   svctm    %util
           58.37    0.44    22.33    45.33   98.37
```

【ディスク障害によるレスポンス悪化状態】

1秒あたりの読取り、書込み数が少ない場合でも、1I/O 当りの処理時間が長かったり、ディスク・ビジー率が高い場合には、ディスク障害の可能性はある

iostat 出力例

```
--- UNIX ---
Device      r/s    w/s    rk/s    wk/s    wait    actv    svc_t    %w    %b
sd0         0.0    0.0     0.0     0.0     0.0     0.0     0.0     0     0
sd1         0.2    0.0     1.6     0.0     0.0     0.9   4000.0     0    80

--- Linux ---
Device:     rrqm/s    wrqm/s    r/s    w/s    rsec/s    wsec/s    rkB/s    kB/s
hd1         42.3     30.69   12.25    3.92   119.61   66.05   209.80   283.74
           avgrq-sz  avgqu-sz   await   svctm    %util
           43.37    0.66    12.33   2838.33   98.37
```