

3 データベース設計

この章では、データベースを構成するファイルやメモリの設計方針について説明します。また、2章で確認した結果をもとに、データベースを構成するファイルやメモリの容量の見積もり方法について説明します。

本章の構成

- 3.1 リソースの設計方針
- 3.2 データベース環境の決定
- 3.3 リソース見積もりの概要
- 3.4 システムファイルの容量見積もり
- 3.5 作業表ファイルの容量見積もり
- 3.6 RD エリアの容量見積もり
- 3.7 HiRDB ファイルシステム領域の容量見積もり
- 3.8 グローバルバッファの容量見積もり
- 3.9 排他制御プールの容量見積もり
- 3.10 リソース見積もりの結果

3.1 リソースの設計方針

HiRDB を使用する業務に最適で、パフォーマンスや信頼性が高いデータベースを構築するためには、データベースを構成するファイル (HiRDB ファイル、システムファイル、作業表ファイル) やグローバルバッファを適切に設計する必要があります。

ここでは、HiRDB ファイルシステム領域、システムファイル、RD エリア、グローバルバッファの設計方針について説明します。

なお、本テキストでは、ディスクは RAID を使用し、ディスク障害^{*1}からの回復は RAID で保証されていることを前提とし、設計方針を説明します^{*2}。

3.1.1 HiRDB ファイルシステム領域の設計

HiRDB ファイルシステム領域の設計方針について、次に示します。

(1) HiRDB ファイルシステム領域は、次に示す用途ごとに作成します。

- RD エリア用
- システムファイル用
- 作業表ファイル用

(2) RD エリア用の HiRDB ファイルシステム領域は、システム RD エリア用とユーザ RD エリア用をそれぞれ別に作成します。

別々に作成することで、OS レベルでのファイルのコピーなどにおける運用容易性 (バックアップリカバリの運用容易性、HiRDB/Single Server から HiRDB/Parallel Server への移行容易性など) が向上します。

(3) システムファイル用の HiRDB ファイルシステム領域は、RD エリア用の HiRDB ファイルシステム領域とは異なる RAID グループに作成します。

RAID 障害^{*3}に備えるためです。異なる RAID グループに作成しておけば、RD エリア用の HiRDB ファイルシステム領域を作成した RAID グループに障害が発生した場合、当該 RD エリアのバックアップとシステムファイルをもとに回復できます。

さらに、異なる RAID グループに作成することで、表やインデックスデータの I/O とシステムログの I/O を分散させることができるので、パフォーマンスの向上も期待できます。

^{*1} ディスク障害とは、データベースを構築しているハードディスクの障害です。詳しくは 6 章で説明します。

^{*2} RAID を用いても、ディスク障害からの回復が必ず保証されるわけではありません。(例えば、3 本のディスクの RAID5 構成において、2 本のディスクに障害が発生した場合など。) ディスク障害に備える場合、あるいは RAID を用いない場合の設計方針は付録に示します。

^{*3} RAID 障害とは、データベースを構築している RAID グループの障害、つまり、RAID を用いてもディスク障害から回復できない場合を指します。

- (4) 作業表ファイル用の HiRDB ファイルシステム領域は任意の RAID グループに作成します。
作業表ファイルは、常に領域を確保するわけではないため、任意の RAID グループに作成します。

以上の方針を踏まえ、容量のバランスを考慮した HiRDB ファイルシステム領域の設計を次の図に示します。ここでは、HiRDB ファイルシステム領域を作成する領域として、RAID グループが 2 グループあるとします*4。

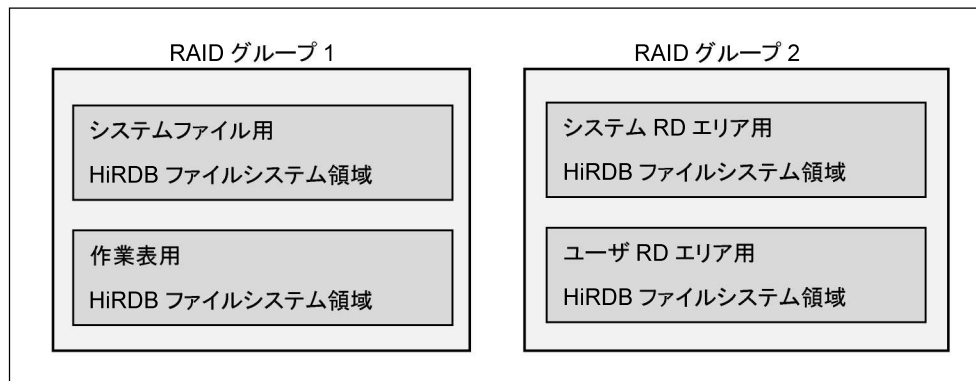


図 3.1-1 HiRDB ファイルシステム領域の設計

参考 -RAID グループ-

■RAID グループ

複数の物理ハード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し冗長性を向上させる技術を RAID といいます。このとき、RAID を構成する物理ハードディスクのグループを RAID グループといいます。

また、論理的に分けたディスクスペースのことを LU (Logical Unit = 論理ユニット) といい、LU と RAID グループは別物です。

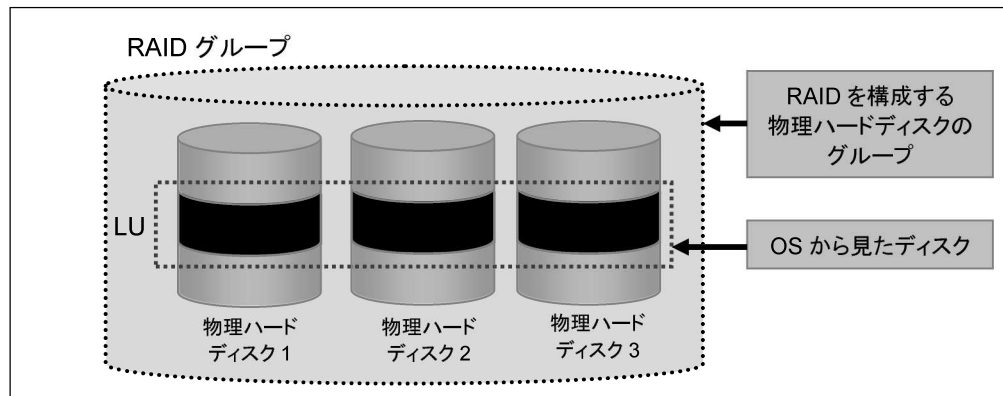


図 3.1-2 RAID グループ

*4 バックアップ、アンロードログファイルの格納先として、もう 1 つ RAID グループがあるとします。バックアップ、アンロードログファイルについては、後述します。

3.1.2 システムファイルの設計

システムログファイル、シンクポイントダンプファイル、ステータスファイルはシステムファイル用 HiRDB ファイルシステム領域に作成します。

ここでは、システムファイルの設計方針を説明します。

補足 -システムファイルのグループ-

Supplementation

システムファイルは、二重化したファイルをまとめて「グループ(世代)」といい、二重化したそれぞれのファイルをA系、B系といいます。二重化したファイルには、同じ情報が書き込まれます。

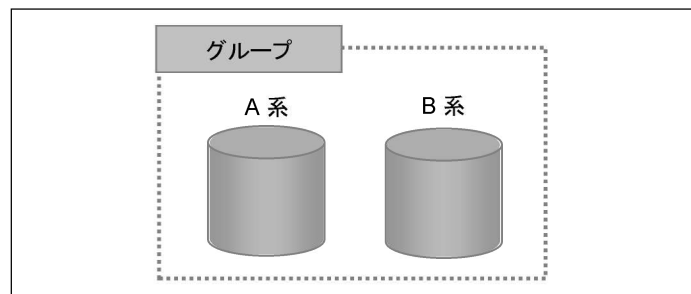


図 3.1-3 システムファイルのグループ(二重化あり)

二重化していない場合、ファイルは1つですが、1グループといいます。

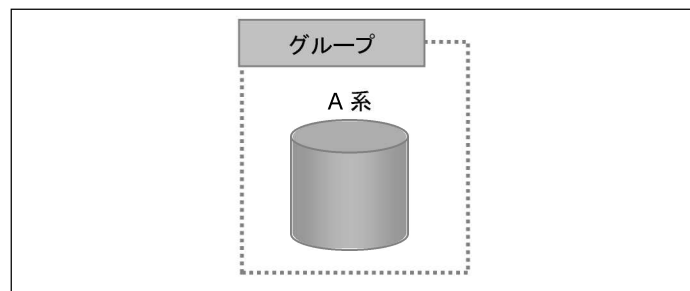


図 3.1-4 システムファイルのグループ(二重化なし)

(1) システムログファイルの設計

■システムログファイルの動作

システムログファイルは、循環して使用します。例えば、次の図に示すようにシステムログファイルを3グループで運用している場合、第1グループが満杯になると第2グループに書き込み、第2グループが満杯になると第3グループに書き込み、第3グループが満杯になると第1グループに上書きします。

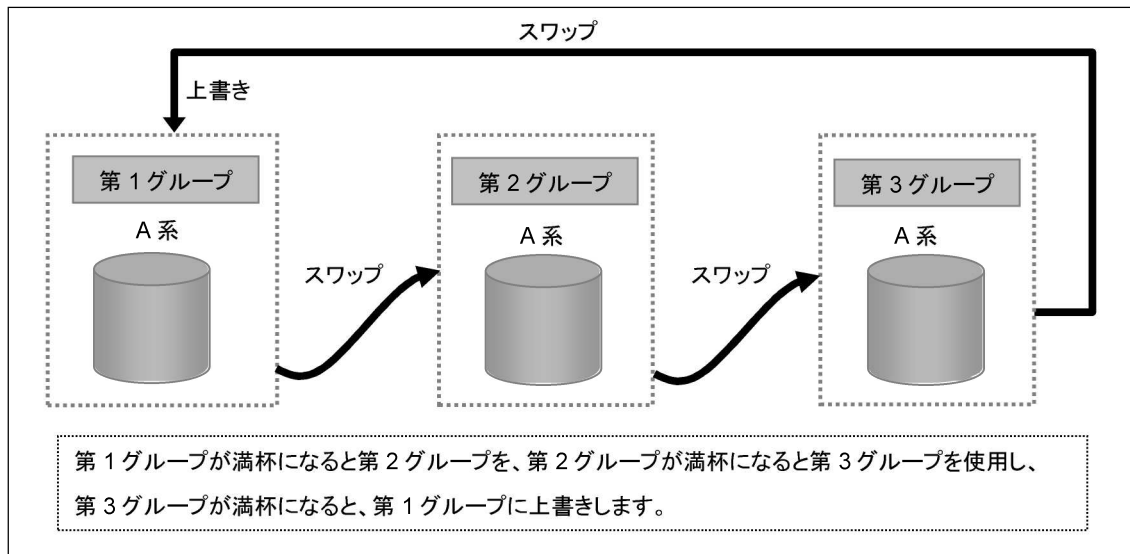


図 3.1-5 システムログファイルの循環利用

■システムログファイルのグループ数

システムログファイルが1グループのみの場合、システムログファイルが満杯になると、次に書き込むシステムログファイルがないため、HiRDBは異常終了します。

システムログファイルが2グループの場合、運用は可能ですが、一巡して上書きされる前に余裕を持ってアンロード処理(後述)ができない恐れがあります。そのため、システムログファイルは3グループ以上作成します。

ユーザのオペレーションミスなどによるシステムログファイルの損失に備える場合は、余裕を持ってアンロード処理ができるように、10グループ以上の作成を検討してください。

補足 -システムログファイルのスワップ-

Supplementation

システムログファイルが満杯になり、次のシステムログファイルを使用する(ログを出力先となるシステムログファイルを変更する)ことを「システムログファイルのスワップ」といいます。

■システムログファイルの二重化

システムログファイルは、二重化できます。システム障害*⁵によるシステムログファイルの破損*⁶に備える場合は、システムログファイルの二重化を検討してください。二重化しておくことで、システム障害によるシステムログファイルの損失が発生した場合、正常な系のシステムログファイルを用いて回復できます。

二重化したシステムログファイルの設計については、付録を参照してください。

*⁵ システム障害とは、サーバマシンの電源停止などによる HIRDB の異常終了です。詳しくは、6 章で説明します。

*⁶ システムログファイルへの書き込み中にシステム障害が発生した場合、書き込み中のシステムログファイルが破損する恐れがあります。ファイルの破損かつ二重化していない場合、システム障害から回復できません。

参考 -システムログファイルの自動拡張機能-

■システムログファイルの自動拡張機能

システムログファイルの容量不足が発生した際に、自動的に HiRDB ファイルシステム領域^{*7}とシステムログファイルを拡張する機能を、システムログファイルの自動拡張機能といいます。

想定以上のログが出力され、システムログファイルが全て満杯になり、上書き可能なシステムログファイルがなくなってしまう場合、ログの書き込み先がないため HiRDB は異常終了します。本機能を適用することで、システムログファイルの容量不足による HiRDB の異常終了の頻度を低減できます。

なお、本機能は HiRDB Version 08・05 から提供されています。

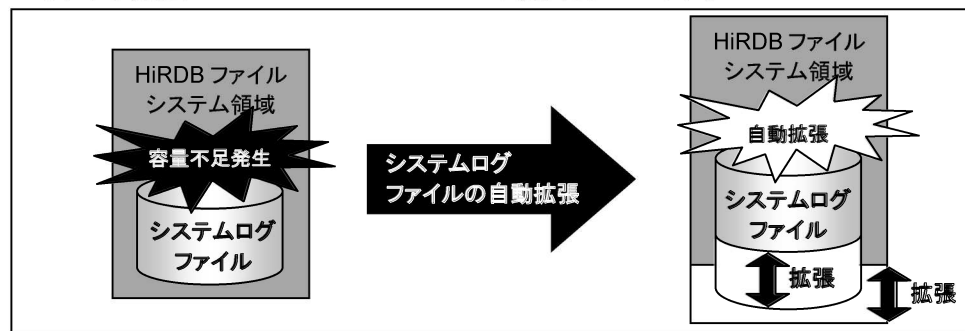


図 3.1-6 システムログファイルの自動拡張

■適用基準

基本的には、本機能は適用せず、HiRDB ファイルシステム領域とシステムログファイルは固定サイズとし、容量不足が発生しないよう自動ログアンロード機能(4章)と空き容量監視機能^{*8}を適用して運用することを推奨します。

しかし、システムログファイルの容量の見積もりが難しい場合や、試験的に HiRDB を使用する場合、あるいはリラン用のシステムログファイル^{*9}にのみ本機能の適用を推奨します。

なお、適用する場合は HiRDB ファイルシステム領域が作成されているディスクに自動拡張できる余裕が十分にあることと、AP による不当な大量のデータ更新を防止する仕組み(実行時間監視^{*10}、使用リソース数制限など)を実装することが前提です。

^{*7} HiRDB ファイルシステム領域を作成する際に、自動拡張を可能とするオプションを指定する必要があります。(4章参照)

^{*8} pd_log_remain_space_check パラメータに safe を指定します。(付録参照)

^{*9} pd_log_rerun_reserved_file_open パラメータに Y を指定すると、HiRDB 再開始時に「スワップ先にできる状態」のシステムログファイルがなくても、「予約状態(HiRDB 開始時にオープンせず、出力対象外)」のシステムログファイルがあれば、予約状態のファイルをオープンして HiRDB を再開始できるため、運用が容易になります。このように障害に備えてあらかじめ作成しておく予約状態のファイルのことを、リラン用のシステムログファイルといいます。

^{*10} クライアント環境定義の PDCWAITTIME パラメータもしくは、システム共通定義の pd_utl_exec_time パラメータで設定します。

(2) シンクポイントダンプファイルの設計

■シンクポイントダンプファイルの動作

シンクポイントダンプファイルは、循環して使用します。例えば、次の図に示すように、シンクポイントダンプファイルを2グループで運用している場合、t1時点のシンクポイント情報は第1グループに書き込み、t2時点のシンクポイント情報は第2グループに書き込み、t3時点のシンクポイント情報は第1グループに上書きします。

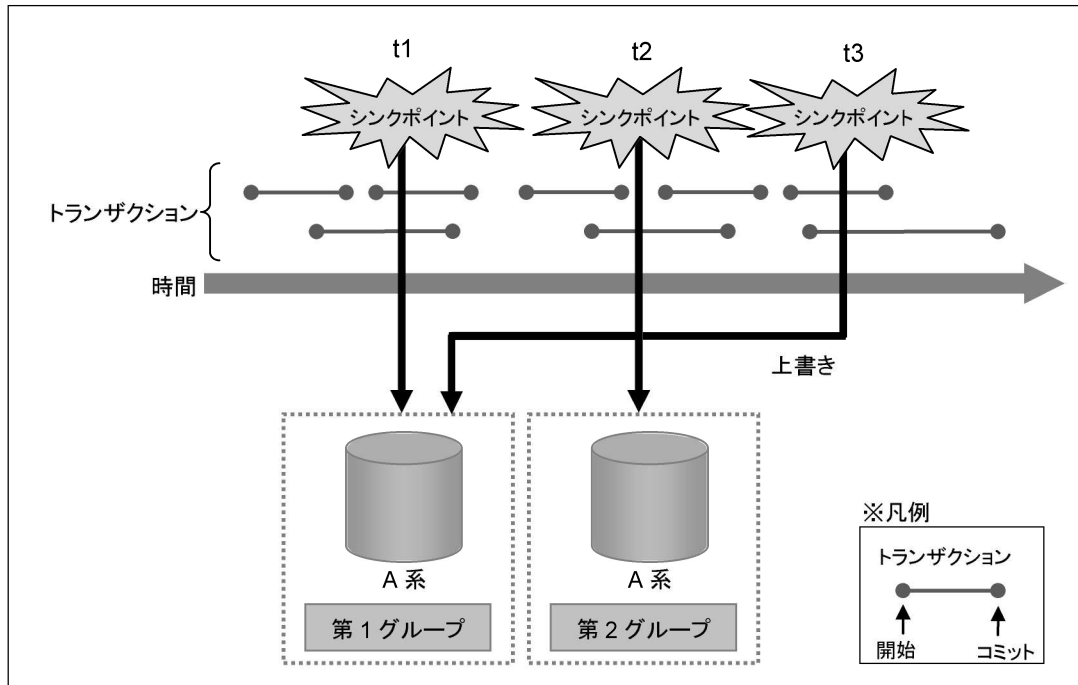


図 3.1-7 シンクポイントダンプファイルの利用

■シンクポイントダンプファイルのグループ数

シンクポイント情報の書き込み時点で、実行中の更新系トランザクションがある場合、トランザクションがコミットするまで、シンクポイントダンプファイルは無効です。このトランザクションがコミットすると、シンクポイントダンプファイルは有効になります*11。

システム障害が発生した時点で有効なシンクポイントダンプファイルを確保するために、シンクポイントダンプファイルは2グループ以上作成します。

ユーザのオペレーションミスなどによるシンクポイントダンプファイルの損失に備える場合は、有効なシンクポイントダンプファイルが確保できるように、4グループ以上の作成を検討してください。

■シンクポイントダンプファイルの二重化

シンクポイントダンプファイルは、二重化できます。二重化したシンクポイントダンプファイルには、同じ情報が書き込まれます。

*11 デフォルトでは、シンクポイント取得時に1グループ前のシンクポイントダンプファイルが有効になっていない場合、シンクポイントはスキップされ、実行されません。

(3) ステータスファイルの設計

■ステータスファイルの動作

ステータスファイルは循環して使用しません。システムの状態が変化すると、情報を上書きします。

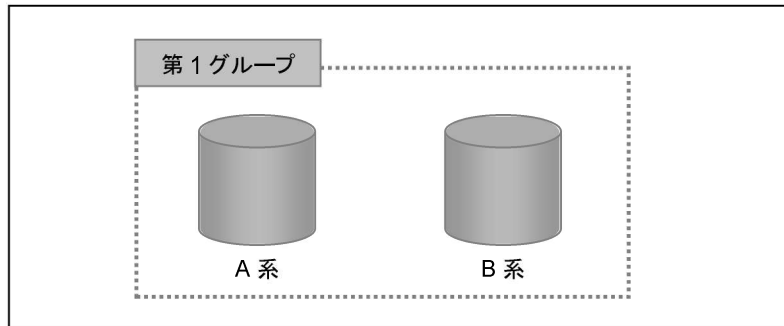


図 3.1-8 ステータスファイルの利用

■ステータスファイルのグループ数

ステータスファイルは1グループ以上作成します。

ユーザのオペレーションミスなどによるステータスファイルの損失に備える場合は、他グループのステータスファイルを使用して運用が継続できるよう、2グループ以上の作成を検討してください。

■ステータスファイルの二重化

ステータスファイルは、二重化が必須です。

3. データベース設計

(4) システムファイルの設計方針のまとめ

システムファイルの設計方針を次にまとめます。

- (1) システムログファイルは、3 グループ以上作成します。
- (2) シンクポイントダンプファイルは、2 グループ以上作成します。
- (3) ステータスファイルは、1 グループ以上作成し、二重化します。

上記方針を踏まえたシステムファイルの配置を次の図に示します。

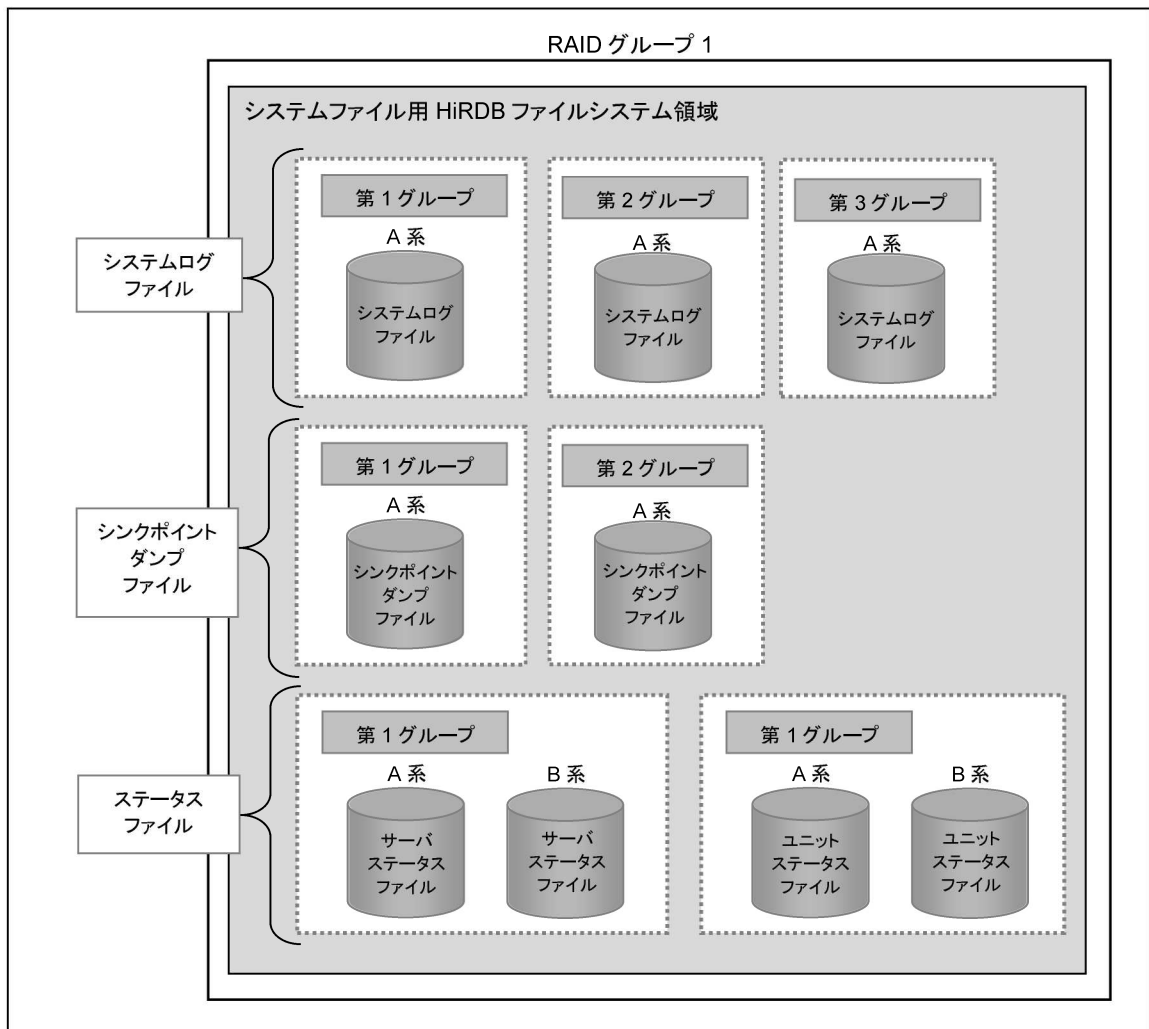


図 3.1-9 システムファイルの構成

補足 -サーバステータスファイルとユニットステータスファイル-

Supplementation

ステータスファイルには、サーバステータスファイルとユニットステータスファイルの 2 種類があり、それぞれ二重化が必須です。