

vSphere のリソース管理

Update 3

VMware vSphere 7.0

VMware ESXi 7.0

vCenter Server 7.0

最新の技術ドキュメントは、VMware の Web サイト (<https://docs.vmware.com/jp/>)

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

VMware株式会社
〒108-0023 東京都港区芝浦 3-1-1
田町ステーションタワー N 18 階
www.vmware.com/jp

Copyright © 2006-2021 VMware, Inc. All rights reserved. 著作権および商標情報。

目次

vSphere のリソース管理について 10

1 リソース管理について 11

- リソースのタイプ 11
- リソースの提供者 11
- リソースの消費者 12
- リソース管理の目的 12

2 リソース割り当て設定の構成 13

- リソース割り当てのシェア 13
- リソース割り当ての予約 14
- リソース割り当て制限 15
- リソース割り当て設定の推奨事項 15
- 設定の編集 15
- リソース割り当て設定の変更：例 16
- アドミッション コントロール 17

3 CPU 仮想化の基礎 18

- ソフトウェア ベースの CPU 仮想化 18
- ハードウェア アシストによる CPU 仮想化 19
- 仮想化とプロセッサ固有の動作 19
- CPU 仮想化によるパフォーマンスの影響 19

4 CPU リソースの管理 21

- プロセッサ情報の表示 21
- CPU 構成の指定 22
- マルチコア プロセッサ 22
- ハイパースレッド 22
 - ハイパースレッドと ESXi ホスト 23
 - ハイパースレッドの有効化 24
- CPU アフィニティの使用 24
 - 特定のプロセッサへの仮想マシンの割り当て 25
 - CPU アフィニティでの潜在的な問題 25
- ホストの電源管理ポリシー 26
 - CPU 電源管理ポリシーの選択 26
 - ホストの電源管理用のカスタム ポリシー パラメータの構成 27

5 メモリ仮想化の基礎 29

- 仮想マシンのメモリ 29
- メモリのオーバーコミット 30
- メモリの共有 31
- メモリ仮想化 31
 - ハードウェア アシストによるメモリ仮想化 32
- ラージ ページ サイズのサポート 33

6 メモリ リソースの管理 34

- メモリ オーバーヘッドについて 34
 - 仮想マシン上のオーバーヘッド メモリ 35
- ESXi ホストによるメモリ割り当ての方法 35
 - アイドル仮想マシンのメモリ税 36
 - VMX スワップ ファイル 36
- メモリの解放 37
 - メモリ バルーン ドライバ 37
- スワップ ファイルの使用 38
 - スワップ ファイルの場所 38
 - DRS クラスタでのホストのローカル スワップの有効化 39
 - スタンドアロン ホストでのホストのローカル スワップの有効化 39
 - スワップ スペースとメモリのオーバーコミット 40
 - ホストの仮想マシン スワップファイル プロパティの設定 41
 - クラスタ用の仮想マシンのスワップ ファイルのための場所の設定 42
 - スワップ ファイルの削除 42
- 仮想マシン間でのメモリの共有 43
- メモリ圧縮 43
 - メモリ圧縮キャッシュの有効化または無効化 43
 - メモリ圧縮キャッシュの最大サイズの設定 44
- メモリ使用量の測定とタイプの識別 44
- メモリの信頼性 46
 - エラー隔離通知の訂正 46
- システム スワップについて 46
 - システム スワップの構成 47

7 永続的なメモリ 48

- PMEM 仮想マシンに対する vSphere HA の構成 50
- vSphere HA アドミッション コントロールの PMEM 予約 51
- vSphere メモリの監視と修正 52

8 仮想グラフィックの構成 54

- GPU 統計の表示 54
- 仮想マシンへの NVIDIA GRID vGPU の追加 55

- ホスト グラフィックの設定 55
- グラフィック デバイスの設定 56

9 ストレージ I/O リソースの管理 57

- 仮想マシンのストレージ ポリシーについて 58
- I/O フィルタについて 58
- ストレージ I/O コントロールの要件 59
- ストレージ I/O コントロールのリソースのシェアおよび制限 59
 - ストレージ I/O コントロールのシェアおよび制限の表示 59
 - ストレージ I/O コントロールのシェアの監視 60
- Storage I/O Control のリソースのシェアおよび制限の設定 60
- Storage I/O Control を有効にする 61
- ストレージ I/O コントロールのしきい値の設定 61
- Storage DRS とストレージ プロファイルの統合 63

10 リソース プールの管理 64

- リソース プールを使用する理由 65
- リソース プールの作成 67
- リソース プールの編集 68
- リソース プールへの仮想マシンの追加 69
- リソース プールからの仮想マシンの削除 70
- リソース プールの削除 70
- リソース プールのアドミッション コントロール 71
 - 拡張可能な予約の例 1 71
 - 拡張可能な予約の例 2 71

11 DRS クラスタの作成 74

- vSphere クラスタ サービス 75
 - vCLS データストアの配置 76
 - vSphere クラスタ サービスの監視 77
 - vSphere クラスタ サービスの健全性の維持 77
 - クラスタの退避モードへの切り替え 79
 - vCLS 仮想マシンのパスワードの取得 80
 - vCLS 仮想マシン非アフィニティ ポリシー 80
- アドミッション コントロールと初期配置 81
 - 1台の仮想マシンのパワーオン 82
 - グループ パワーオン 82
- 仮想マシンの移行 83
 - DRS 移行のしきい値 84
 - 移行の推奨 84
- DRS クラスタの要件 85

共有ストレージの要件	85
共有 VMFS ボリュームの要件	85
プロセッサの互換性要件	86
DRS クラスタの vMotion 要件	86
仮想フラッシュによる DRS の構成	87
クラスタの作成	87
クラスタ設定の編集	88
仮想マシンのカスタム自動化レベルの設定	90
DRS の無効化	91
リソース プール ツリーのリストア	92
vSAN ストレッチ クラスタの DRS 認識	92

12 ROBO Enterprise ライセンスを保持している場合の DRS メンテナンス モードの機能 94

ROBO Enterprise ライセンスでの DRS メンテナンス モードの制限事項	94
ROBO Enterprise ライセンスでの DRS メンテナンス モードの使用	94
ROBO Enterprise ライセンスでの DRS メンテナンス モードのトラブルシューティング	95

13 DRS クラスタを使用したリソース管理 97

クラスタへのホストの追加	97
クラスタへの管理対象ホストの追加	97
管理されていないホストのクラスタへの追加	98
クラスタへの仮想マシンの追加	99
仮想マシンのクラスタへの移動	99
クラスタからの仮想マシンの削除	100
クラスタからの仮想マシンの移動	100
クラスタからのホストの削除	101
ホストをメンテナンス モードに切り替える	101
クラスタからのホストの削除	102
スタンバイ モードの使用	102
DRS クラスタの妥当性	103
有効な DRS クラスタ	103
オーバーコミットされた DRS クラスタ	105
無効な DRS クラスタ	106
電力リソースの管理	107
vSphere DPM 用の IPMI 設定または iLO 設定の構成	108
vSphere DPM 用の Wake-On-LAN のテスト	109
DRS クラスタ用の vSphere DPM の有効化	110
vSphere DPM の監視	112
DRS アフィニティ ルールの使用	113
ホストの DRS グループの作成	113
仮想マシンの DRS グループの作成	114

- 仮想マシン間のアフィニティ ルール 114
- 仮想マシンとホスト間のアフィニティ ルール 115

14 データストア クラスタの作成 118

- 初期配置と実行中のバランシング 119
- ストレージの移行の推奨 119
- データストア クラスタの作成 120
- Storage DRS の有効化および無効化 120
- データストア クラスタの自動化レベルの設定 121
- Storage DRS の積極性レベルの設定 121
 - ストレージ DRS ランタイム ルールの設定 122
- データストア クラスタの要件 123
- データストアのデータストア クラスタへの追加と削除 123

15 データストア クラスタを使用したストレージ リソースの管理 124

- ストレージ DRS メンテナンス モードの使用 124
 - データストアをメンテナンス モードに切り替える 124
 - Storage DRS アフィニティ ルールを無視したメンテナンス モードへの切り替え 125
- Storage DRS 推奨の適用 126
 - Storage DRS 推奨の更新 126
- 仮想マシンの Storage DRS 自動レベルの変更 127
- オフピークの Storage DRS スケジュールの設定 127
- Storage DRS 非アフィニティ ルール 129
 - 仮想マシン非アフィニティ ルールの作成 130
 - VMDK 非アフィニティ ルールの作成 130
 - VMDK アフィニティ ルールのオーバーライド 131
- Storage DRS 統計の削除 132
- ストレージ vMotion のデータストア クラスタとの互換性 133

16 NUMA システムと ESXi の併用 134

- NUMA とは 134
 - オペレーティング システムについての課題 135
- ESXi の NUMA スケジュール設定の機能 135
- VMware NUMA 最適化のアルゴリズムと設定 136
 - ホーム ノードと初期配置 136
 - 動的ロード バランシングとページの移行 137
 - NUMA 用に最適化された透過的なページ共有 137
- NUMA アーキテクチャでのリソース管理 138
- 仮想 NUMA の使用 138
 - 仮想 NUMA 制御 139
- NUMA 制御の指定 139

- 仮想マシンと特定のプロセッサとの関連付け 140
- メモリ アフィニティを使用した、メモリ割り当てと特定の NUMA ノードとの関連付け 140
- 仮想マシンと指定 NUMA ノードの関連づけ 141

17 詳細属性 143

- ホストの詳細属性の設定 143
 - メモリの詳細属性 143
 - NUMA の詳細属性 145
- 仮想マシンの詳細属性の設定 146
 - 仮想マシンの詳細属性 146
 - 仮想 NUMA の詳細属性 147
- 遅延感度 148
 - 遅延感度の調整 148
- 信頼性の高いメモリについて 149
 - 信頼性の高いメモリの表示 149
- 1 GB ページでのゲスト vRAM のバッキング 149

18 障害の定義 150

- 仮想マシンが固定されている 151
- 仮想マシンがいずれのホストとも互換性がない 151
- 異なるホストへの移動時に仮想マシン間の DRS ルール違反が発生する 151
- ホストに仮想マシンとの互換性がない 151
- 仮想マシン間の DRS ルールに違反する仮想マシンがホストにある 151
- 仮想マシン用の十分なキャパシティがホストにない 152
- ホストの状態が正しくない 152
- 仮想マシン用の物理 CPU の数がホストで不足している 152
- 各仮想マシン CPU 用の十分なキャパシティがホストにない 152
- 仮想マシンが vMotion 処理中である 152
- クラスタにアクティブなホストがない 152
- リソース不足 152
- HA 用として構成されたフェイルオーバー レベルを満たすための十分なリソースがない 153
- 互換性のあるハード アフィニティ ホストがない 153
- 互換性のあるソフト アフィニティ ホストがない 153
- ソフト ルール違反の修正が許可されていない 153
- ソフト ルール違反の修正による影響 153

19 DRS トラブルシューティング情報 154

- クラスタの問題 154
 - クラスタでの負荷不均衡 154
 - クラスタが黄色で表示される 155
 - リソース プールの不整合によりクラスタが赤色で表示される 155

- フェイルオーバー キャパシティに違反しているためクラスタが赤色で表示される 156
- クラスタの総負荷が低い場合、ホストがパワーオフされない 156
- クラスタの総負荷が高い場合、ホストがパワーオフされる 157
- DRS で vMotion の移行が実行されることがほとんど、またはまったくない 157
- ホストの問題 158
 - クラスタの総負荷が低い場合に、ホストをパワーオンしてキャパシティを増やすことを DRS が推奨する 158
 - クラスタの総負荷が高い 158
 - クラスタの総負荷が低い 159
 - 要求されたホストを DRS が回避してメンテナンス モードまたはスタンバイ モードに切り替えることができない 159
 - DRS によって仮想マシンがホストに移動されない 160
 - DRS によってホストから仮想マシンが移動されない 160
- 仮想マシンの問題 161
 - CPU またはメモリ リソース不足 161
 - 仮想マシン間の DRS ルール違反または仮想マシンとホスト間の DRS ルール違反が発生する 162
 - 仮想マシンのパワーオン操作が失敗する 163
 - DRS によって仮想マシンが移動されない 163

vSphere のリソース管理について

『vSphere リソース管理』では、VMware[®] ESXi および vCenter[®] Server 環境のリソース管理について説明します。

本書では次の内容を中心に説明しています。

- リソース割り当ておよびリソース管理の概念
- 仮想マシンの属性とアドミッション コントロール
- リソース プールとその管理方法
- クラスタ、vSphere[®] DRS (Distributed Resource Scheduler)、vSphere DPM (Distributed Power Management) の操作方法
- データストア クラスタ、Storage DRS、Storage I/O Control の操作方法
- リソース管理の詳細オプション
- パフォーマンスについての考慮事項

VMware では、多様性の受け入れを尊重しています。お客様、パートナー企業、社内コミュニティとともにこの原則を推進することを目的として、多様性に配慮した言葉遣いでコンテンツを作成します。

対象読者

本書は、システムによるリソースの管理方法およびデフォルト動作のカスタマイズ方法を理解する必要があるシステム管理者を対象としています。また、リソース プール、クラスタ、DRS、データストア クラスタ、Storage DRS、Storage I/O Control、または vSphere DPM を理解および使用するユーザーにも有用です。

本書は、読者が VMware ESXi および vCenter Server を使用した経験があることを前提としています。

注： ここで「メモリ」は、物理メモリまたは永続的なメモリを指します。

リソース管理について

1

リソース管理について理解するには、その構成要素、目的、およびクラスタ設定における最良の実装方法について認識しておく必要があります。

仮想マシンのリソース割り当て設定（シェア、予約、制限）について、その設定方法と表示方法も含めて説明します。また、アドミッション コントロールのほか、既存リソースに対するリソース割り当て設定の検証プロセスについても説明します。

リソース管理とは、リソースの提供者からリソースの利用者へリソースを割り当てることです。

リソース管理の必要性は、リソースのオーバーコミット、つまり容量を上回る需要から生じ、また、需要と容量が時間とともに変化するという事実からも発生します。リソース管理を使用すると、使用可能な容量をより効率的に使用できるよう、リソースを動的に割り当て直すことができます。

注： この章では、「メモリ」は物理 RAM を指します。

この章には、次のトピックが含まれています。

- リソースのタイプ
- リソースの提供者
- リソースの消費者
- リソース管理の目的

リソースのタイプ

リソースには、CPU、メモリ、電力、ストレージ、およびネットワーク リソースなどがあります。

注： ESXi はホストごとにネットワーク バンド幅とディスク リソースを管理し、その際、ネットワーク トラフィックシェーピングと比例的なシェア メカニズムが、それぞれ使用されます。

リソースの提供者

ホストとクラスタは、データストア クラスタも含め、物理リソースの提供者です。

ホストの場合、使用可能なリソースとは、仮想化ソフトウェアによって使用されるリソースをホストのハードウェア仕様から差し引いた分です。

クラスタとは、ホストのグループです。vSphere Client を使用してクラスタを作成し、そのクラスタに複数のホストを追加できます。vCenter Server は、それらのホストのリソースをまとめて管理します。クラスタは、すべてのホストの CPU とメモリをすべて所有します。共通のロード バランシングとフェイルオーバーについて、クラスタを有効にすることができます。詳細については [11 章 DRS クラスタの作成](#) を参照してください。

データストア クラスタは、データストアのグループです。DRS クラスタ同様、vSphere Client を使用してデータストア クラスタを作成し、そのクラスタに複数のデータストアを追加できます。vCenter Server は、データストア リソースをまとめて管理します。Storage DRS を有効にして、I/O 負荷と容量の使用率の負荷を分散できます。 [14 章 データストア クラスタの作成](#) を参照してください。

リソースの消費者

仮想マシンは、リソースの消費者です。

作成時に割り当てるデフォルトのリソース設定で、ほとんどのマシンは十分に機能します。仮想マシンの設定をあとで編集して、リソースの提供者の CPU、メモリ、およびストレージ I/O の合計、または保証された CPU とメモリの予約に対する、シェア ベースの割合を割り当てることができます。仮想マシンをパワーオンする場合、サーバは、未予約の十分なリソースが使用可能かどうかを確認して、十分なリソースがある場合のみパワーオンすることを許可します。このプロセスをアドミッション コントロールと呼びます。

リソース プールは、リソースを柔軟に管理するための論理上の抽象概念です。リソース プールは階層にグループ化することができ、使用可能な CPU リソースとメモリ リソースを階層状にパーティショニングするために使用できます。したがって、リソース プールはリソースの提供者と消費者の両方であるとみなすことができます。リソース プールは子リソース プールと仮想マシンにリソースを提供しますが、親のリソースを消費するので、リソースの消費者でもあります。 [10 章 リソース プールの管理](#) を参照してください。

ESXi ホストは、次のようないくつかの要素に基づいて、基盤となるハードウェア リソースの一部を各仮想マシンに割り当てます。

- ユーザーが定義するリソースの制限。
- ESXi ホスト（またはクラスタ）で使用可能なリソースの合計。
- パワーオンする仮想マシンの数、およびそれらの仮想マシンによるリソースの使用率。
- 仮想化の管理に必要なオーバーヘッド。

リソース管理の目的

リソースを管理する場合、その目的が何であるかを認識しておく必要があります。

リソース管理は、リソースのオーバーコミットを解決できるだけでなく、次のことの実現にも役立ちます。

- パフォーマンスの分離：仮想マシンがリソースを独占するのを防ぎ、予測可能なサービス率を確保する。
- 効率的な利用：アンダーコミットされているリソースを活用し、性能低下を抑えたオーバーコミットを行う。
- 容易な管理：仮想マシンの相対的な重要度を制御し、柔軟な動的パーティショニングを提供し、絶対的なサービス レベル アグリーメントを達成する。

リソース割り当て設定の構成

2

使用可能なリソース容量がリソースの消費者（および仮想化オーバーヘッド）の需要を満たさない場合、管理者は、仮想マシンまたは仮想マシンが配置されているリソース プールに割り当てるリソースの量をカスタマイズする必要があります。

リソース割り当て設定（シェア、予約、制限）は、仮想マシンに提供する CPU、メモリ、およびストレージ リソースの量を決定するために使用します。特に、システム管理者にはリソース割り当てのオプションがいくつかあります。

- ホストまたはクラスタの物理リソースを予約する。
- 仮想マシンに割り当てることができるリソースの上限を設定する。
- 特定の仮想マシンに、常にほかの仮想マシンより高い割合の物理リソースが割り当てられることを保証する。

注： この章では、「メモリ」は物理 RAM を指します。

この章には、次のトピックが含まれています。

- [リソース割り当てのシェア](#)
- [リソース割り当ての予約](#)
- [リソース割り当て制限](#)
- [リソース割り当て設定の推奨事項](#)
- [設定の編集](#)
- [リソース割り当て設定の変更：例](#)
- [アドミッション コントロール](#)

リソース割り当てのシェア

シェアは、仮想マシン（またはリソース プール）の相対的な重要度を指定します。ある仮想マシンのリソースのシェアが別の仮想マシンの 2 倍である場合、その仮想マシンは、別の仮想マシンの 2 倍のリソースを消費できます（2 台の仮想マシンがリソースを獲得するために競合する場合）。

シェアは、通常、[高]、[標準]、または [低] で指定され、これらの値によって、それぞれのシェア値が 4:2:1 の割合で指定されます。特定のシェア数（比重を表す）を各仮想マシンに割り当てる場合は、[カスタム] を選択することもできます。

シェアの指定が意味を持つのは、同レベルの仮想マシンまたはリソース プール（つまり、リソース プール階層内で親が同じである仮想マシンまたはリソース プール）に関してだけです。同レベルの場合、予約および制限によって制限される、相対的なシェア値に応じてリソースを共有します。仮想マシンにシェアを割り当てる場合は、その仮想マシンの、パワーオンされたそれ以外の仮想マシンに対する相対的な優先順位を常に指定します。

次の表は、仮想マシンについて、CPU とメモリのデフォルトのシェア値を示しています。リソース プールの場合、CPU とメモリのデフォルトのシェア値は同じですが、そのリソース プールを、4 つの仮想 CPU と 16 GB のメモリを備えた 1 台の仮想マシンと見なして乗算する必要があります。

表 2-1. シェア値

設定	CPU シェア値	メモリ シェア値
高	仮想 CPU あたり 2000 シェア	構成された仮想マシン メモリ 1MB あたり 20 シェア
標準	仮想 CPU あたり 1,000 シェア	構成された仮想マシン メモリ 1 MB あたり 10 シェア
低	仮想 CPU あたり 500 シェア	構成された仮想マシン メモリ 1 MB あたり 5 シェア

たとえば、2 つの仮想 CPU と 1GB の RAM を持ち、CPU とメモリのシェアが [標準] に設定されている SMP 仮想マシンは、 $2 \times 1,000 = 2,000$ シェアの CPU と、 $10 \times 1,024 = 10,240$ シェアのメモリを持ちます。

複数の仮想 CPU を持つ仮想マシンを、SMP（対称型マルチプロセッシング）仮想マシンと呼びます。

各シェアによって表される相対的な優先順位は、新しい仮想マシンがパワーオンされると変化します。これは、同じリソース プールのすべての仮想マシンに影響を与えます。すべての仮想マシンは、同じ数の仮想 CPU を持ちます。次の例について考えます。

- 2 台の CPU バインドの仮想マシンが、全体の CPU 容量が 8GHz であるホスト上で稼働しています。CPU シェアは [標準] に設定されていて、それぞれ 4GHz を獲得します。
- 3 台目の CPU バインドの仮想マシンがパワーオンされます。その仮想マシンの CPU シェア値は [高] に設定されています。これは、その仮想マシンが、[標準] に設定されているマシンの 2 倍のシェアを持つことを意味します。この新しい仮想マシンは 4GHz を獲得し、ほかの 2 台のマシンはそれぞれ 2GHz だけを獲得します。ユーザーが 3 台目の仮想マシンで、カスタム シェア値を 2,000 に指定した場合も、同じ結果になります。

リソース割り当ての予約

予約は、仮想マシンに保証される最小割り当てを指定します。

vCenter Server または ESXi では、仮想マシンの予約を達成するのに十分な未予約リソースが存在する場合にのみ、仮想マシンをパワーオンできます。サーバは、物理サーバの負荷が大きくても、その量を保証します。予約は、具体的な単位（メガヘルツまたはメガバイト）で表されます。

たとえば、2GHz が使用可能であり、VM1 に 1GHz の予約を指定し、VM2 に 1GHz の予約を指定するとします。これで、必要な場合に各仮想マシンが 1GHz を獲得できることが保証されます。ただし、VM1 が 500MHz しか使用していない場合、VM2 は 1.5GHz を使用できます。

予約は、デフォルトでは 0 です。仮想マシンが最低限必要な CPU またはメモリを常に使用できることを保証する必要がある場合に予約を指定してください。

リソース割り当て制限

制限は、仮想マシンに割り当てることができる CPU、メモリ、またはストレージ I/O リソースの上限を指定します。

サーバは、予約より多くの量を仮想マシンに割り当てることができますが、システムに未使用のリソースがあっても、制限を超える割り当ては行いません。制限は、具体的な単位（メガヘルツ、メガバイト、または 1 秒あたりの I/O 処理数）で表されます。

CPU、メモリ、およびストレージ I/O リソースの制限は、デフォルトでは制限なしです。メモリの制限が制限なしのときは、仮想マシンを作成するときに構成されたメモリの量が事実上の制限になります。

ほとんどの場合、制限を指定する必要はありません。制限には、メリットとデメリットがあります。

- **メリット**：少数の仮想マシンから使用をはじめて、ユーザーの要望に対応する場合に、制限を割り当てると役立ちます。仮想マシンを追加するごとに、パフォーマンスは低下します。制限を指定することによって、使用可能なリソースが少なくなることをシミュレートできます。
- **デメリット**：制限を指定すると、アイドルリソースが無駄になることがあります。システムが活用されていない状態で、アイドルリソースが使用可能な場合でも、システムは、制限以上のリソースを仮想マシンが使用することを許可しません。適切な理由がある場合にのみ制限を指定します。

リソース割り当て設定の推奨事項

使用する ESXi 環境に適切なリソース割り当て設定（予約、制限、シェア）を選択します。

次のガイドラインは、仮想マシンのパフォーマンスを向上させるのに役立ちます。

- [予約] では、ユーザーが使用可能にしたい量ではなく、条件に合った最小の CPU またはメモリの量を指定します。予約によって表される具体的なリソースの量は、仮想マシンの追加や削除など、環境を変更しても変化しません。ホストは、仮想マシンの制限、シェア数、および需要予測数に基づいて、使用可能な追加のリソースを割り当てます。
- 仮想マシンの予約を指定する場合、すべてのリソースをコミットしないでください（10% 以上を未予約にしてください）。システム内のすべての容量が完全に予約された状態に近づくほど、アドミッションコントロールに違反せずに予約とリソースプール階層に変更を加えることが困難になっていきます。DRS の有効なクラスタでは、クラスタの容量またはクラスタ内の個々のホストの容量を完全にコミットする予約によって、DRS が仮想マシンをホスト間で移行できなくなることがあります。
- 合計使用可能リソースが頻繁に変化することが予想される場合は、[シェア] を使用して、仮想マシン間で適正にリソースを割り当てます。[シェア] を使用していて、ホストをアップグレードする場合、各シェアがより多くのメモリ、CPU、またはストレージ I/O リソースを表していても、各仮想マシンの優先順位は変わりません（同じシェア数のままです）。

設定の編集

[設定の編集] ダイアログボックスを使用して、メモリおよび CPU リソースの割り当てを変更します。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
- 2 右クリックして [設定の編集] を選択します。
- 3 CPU リソースを編集します。

オプション	説明
シェア	親の合計リソースに対するこのリソース プールの CPU シェア。兄弟のリソース プールは、予約と制限の範囲内で、相対的シェア値に従ってリソースを共有します。[低]、[標準]、または [高] を選択します。これらの値はそれぞれ 1:2:4 の割合でシェア値を指定します。各仮想マシンに対して、比重に見合う特定のシェア値を指定するには、[カスタム] を選択します。
予約	このリソース プールで確保されている CPU の割り当て。
制限	このリソース プールに割り当てる CPU の上限。上限を指定しない場合は、[制限なし] を選択します。

- 4 メモリ リソースを編集します。

オプション	説明
シェア	親の合計リソースに対するこのリソース プールのメモリ シェア。兄弟のリソース プールは、予約と制限の範囲内で、相対的シェア値に従ってリソースを共有します。[低]、[標準]、または [高] を選択します。これらの値はそれぞれ 1:2:4 の割合でシェア値を指定します。各仮想マシンに対して、比重に見合う特定のシェア値を指定するには、[カスタム] を選択します。
予約	このリソース プール用に確保されているメモリの割り当て。
制限	このリソース プールに割り当てるメモリの上限。制限を指定しない場合は、[制限なし] を選択します。

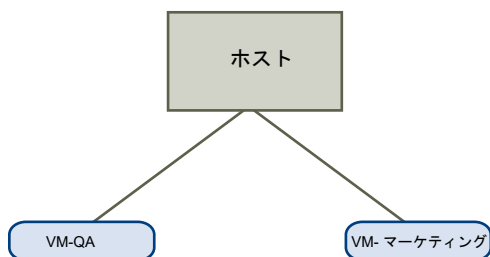
- 5 [OK] をクリックします。

リソース割り当て設定の変更：例

次の例は、リソース割り当て設定を変更して、仮想マシンのパフォーマンスを向上させる方法を示しています。

ESXi ホストで、2 台の新しい仮想マシンを作成したとします。この 2 台はそれぞれ、QA (VM-QA) 部門とマーケティング (VM- マーケティング) 部門用です。

図 2-1. 2 台の仮想マシンを持つ単一ホスト



次の例では、VM-QA が多くのメモリを必要とし、そのために 2 台の仮想マシンのリソース割り当て設定を変更して、次のことを行うものとします。

- システム メモリがオーバーコミットされる場合に、VM-QA が VM- マーケティングの 2 倍の CPU リソースとメモリ リソースを使用できるように指定する。VM-QA の CPU シェアとメモリ シェアを [高] に設定し、VM- マーケティングでは [標準] に設定する。
- VM- マーケティングが、一定量の保証された CPU リソースを所有するようにする。これは、予約設定を使用して実行できます。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
- 2 [VM-QA] (シェアを変更する仮想マシン) を右クリックして、[設定の編集] を選択します。
- 3 [仮想ハードウェア] で、[CPU] を展開し、[シェア] ドロップダウン メニューから [高] を選択します。
- 4 [仮想ハードウェア] で、[メモリ] を展開し、[シェア] ドロップダウン メニューから [高] を選択します。
- 5 [OK] をクリックします。
- 6 マーケティング仮想マシン ([VM- マーケティング]) を右クリックし、[設定の編集] を選択します。
- 7 [仮想ハードウェア] で、[CPU] を展開し、[予約] の値を目的の数値に変更します。
- 8 [OK] をクリックします。

アドミッション コントロール

仮想マシンをパワーオンすると、まだ予約されていない CPU リソースとメモリ リソースの量をシステムが確認します。システムは、未予約の使用可能なリソースに基づいて、仮想マシンが構成されている予約を保証できるかどうかを判定します (予約がある場合)。このプロセスをアドミッション コントロールと呼びます。

未予約の CPU とメモリが十分に使用可能な場合、または予約がない場合、仮想マシンはパワーオンされます。それ以外の場合は、リソース不足です という警告が表示されます。

注： ユーザー指定のメモリ予約のほか、仮想マシンごとにオーバーヘッド メモリがあります。この余分なメモリコミットは、アドミッション コントロールの計算に含まれます。

vSphere DPM 機能が有効になっている場合、電力の消費を減らすために、ホストがスタンバイ モード (パワーオフ状態) になることがあります。これらのホストが提供する未予約のリソースは、アドミッション コントロールに使用できると見なされます。これらのリソースがないと仮想マシンをパワーオンできない場合、十分なスタンバイ ホストをパワーオンする推奨が作成されます。詳細については、『[電力リソースの管理](#)』を参照してください。

CPU 仮想化はパフォーマンスを重視し、可能なかぎりプロセッサ上で直接実行されます。基盤となる物理リソースが可能なかぎり使用され、必要な場合にだけ仮想化レイヤーが命令を実行して、あたかも物理マシン上で直接実行されているかのように仮想マシンを動作させます。

CPU 仮想化は、エミュレーションと同じものではありません。ESXi は、仮想 CPU の実行にエミュレーションを使用しません。エミュレーションでは、すべての処理がソフトウェア内でエミュレータによって実行されます。ソフトウェア エミュレータにより、プログラムが元々記述されたコンピュータ システム以外のシステム上でそのプログラムを実行できるようになります。エミュレータは、これを行うために、同じデータまたは入力を受け入れて、同じ結果を達成することで、元のコンピュータの動作をエミュレート（再生）します。エミュレーションは移植性を提供し、1つのプラットフォーム用に設計されたソフトウェアをいくつかのプラットフォームをまたいで実行します。

CPU リソースがオーバーコミットされている場合、ESXi ホストは、すべての仮想マシンにわたって物理プロセッサを時分割するため、指定された数の仮想プロセッサが備わっているかのように各仮想マシンが実行されます。ESXi ホストが複数の仮想マシンを実行する場合、ESXi ホストは物理リソースのシェアを各仮想マシンに割り当てます。デフォルトのリソース割り当て設定では、同じホストに関連付けられているすべての仮想マシンは、1つの仮想 CPU につき同等の CPU シェアを取得します。つまり、単一プロセッサ仮想マシンには、デュアル プロセッサ仮想マシンのリソースの半分だけが割り当てられます。

この章には、次のトピックが含まれています。

- ソフトウェア ベースの CPU 仮想化
- ハードウェア アシストによる CPU 仮想化
- 仮想化とプロセッサ固有の動作
- CPU 仮想化によるパフォーマンスの影響

ソフトウェア ベースの CPU 仮想化

ソフトウェア ベースの CPU 仮想化では、ゲスト アプリケーション コードはプロセッサ上で直接実行されます。このとき、ゲスト特権コードは変換され、変換されたコードがプロセッサ上で実行されます。

変換されたコードはわずかに大きくなり、通常はネイティブ バージョンより実行速度が遅くなります。結果として、小さな特権コード コンポーネントを持つゲスト アプリケーションは、ネイティブに非常に近い速度で実行されます。大きな特権コード コンポーネント（システム呼び出し、トラップ、ページ テーブル アップデートなど）を持つアプリケーションは、仮想化環境では実行速度が遅くなる場合があります。

ハードウェア アシストによる CPU 仮想化

一部のプロセッサでは、CPU を仮想化するハードウェア アシストが提供されます。

このアシストを使用すると、ゲストではゲスト モードと呼ばれる個別の実行モードを使用できます。ゲスト コードは、アプリケーション コードか特権コードかにかかわらず、ゲスト モードで実行されます。場合によっては、プロセッサがゲスト モードを終了し、ルート モードを開始します。この場合、ハイパーバイザーがルート モードで実行され、ゲスト モードを終了した理由を判断し、必要なアクションを実行し、ゲスト モードでゲストを再起動します。

仮想化にハードウェア アシストを使用すると、コードを変換する必要がなくなります。その結果、システム呼び出しやトラップを多用するワークロードが、ネイティブに近い速度で実行されます。一部のワークロード（ページテーブルのアップデートを行うものなど）では、ゲスト モードからルート モードへの移行が頻繁に発生します。この移行の回数および移行で費やされる合計時間によっては、ハードウェア アシストによる CPU 仮想化の実行速度は大幅に速くなることがあります。

仮想化とプロセッサ固有の動作

VMware ソフトウェアは CPU を仮想化しますが、仮想マシンは実行しているプロセッサの特定のモデルを認識します。

各プロセッサ モデルは、提供する CPU 機能が異なる場合があります。仮想マシン内で実行されるアプリケーションは、それらの機能を使用できます。したがって、異なる機能セットを備えたプロセッサ上で実行されているシステム間で、vMotion[®] を使用して仮想マシンを移行させることはできません。この機能をサポートするプロセッサで Enhanced vMotion Compatibility (EVC) を使用すると、この制限を回避できることもあります。詳細は、『vCenter Server and Host Management』ドキュメントを参照してください。

CPU 仮想化によるパフォーマンスの影響

CPU 仮想化によって追加されるオーバーヘッドの量は、ワークロードと使用する仮想化のタイプによって異なります。

アプリケーションの時間のほとんどが、外部イベント（ユーザーの対話、デバイス入力、データ検索など）の待機のためではなく命令の実行に費やされる場合、そのアプリケーションは CPU バインドです。そのようなアプリケーションの場合、CPU 仮想化オーバーヘッドには、実行する必要がある追加命令が含まれます。このオーバーヘッドは、アプリケーション自体が使用できる CPU 処理時間を奪います。通常、CPU 仮想化オーバーヘッドによって、全体的にパフォーマンスが低下します。

CPU バインドでないアプリケーションの場合、CPU 仮想化は CPU 使用率の増加を意味することが多くなります。オーバーヘッドを吸収するのに予備の CPU 容量を使用できる場合、CPU は、全体的なスループットに関して同等のパフォーマンスを提供できます。

ESXi は 1 台の仮想マシンにつき最大 128 の仮想プロセッサ（CPU）をサポートします。

注： 最高のパフォーマンスとリソース使用率を得るためには、単一スレッドアプリケーションを（複数の CPU が搭載された SMP 仮想マシンではなく）ユニプロセッサ仮想マシンに導入してください。

単一スレッドアプリケーションは、単一 CPU のみを活用できます。そのようなアプリケーションをデュアルプロセッサ仮想マシンに導入しても、アプリケーションの速度は向上しません。それどころか、ほかの仮想マシンが使用可能な物理リソースを 2 番目の仮想 CPU が使用してしまいます。

CPU リソースの管理

4

1つ以上の仮想プロセッサを使用して、それぞれ独自のレジスタ セットおよび制御構造を備える仮想マシンを構成できます。

仮想マシンをスケジュール設定すると、その仮想プロセッサが物理プロセッサ上で実行するようにスケジュール設定されます。VMkernel リソース マネージャは、物理 CPU 上に仮想 CPU をスケジュール設定することによって、仮想マシンの物理 CPU リソースへのアクセスを管理します。ESXi は最大 128 個の仮想 CPU を持つ仮想マシンをサポートします。

注： この章では、「メモリ」は物理 RAM または永続的なメモリを指します。

この章には、次のトピックが含まれています。

- プロセッサ情報の表示
- CPU 構成の指定
- マルチコア プロセッサ
- ハイパースレッド
- CPU アフィニティの使用
- ホストの電源管理ポリシー

プロセッサ情報の表示

vSphere Client では、現在の CPU 構成に関する情報にアクセスできます。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [ハードウェア] で、[CPU] を展開して、物理プロセッサの数とタイプ、および論理プロセッサの数に関する情報を参照します。

注： ハイパースレッド システムでは、各ハードウェア スレッドが論理プロセッサになります。たとえば、ハイパースレッドが有効なデュアル コア プロセッサでは、コアが 2 つで論理プロセッサが 4 つになります。

CPU 構成の指定

CPU 構成を指定して、リソース管理を向上できます。ただし、CPU 構成をカスタマイズしない場合、ESXi ホストは、ほとんどの状況で良好に機能するデフォルトを使用します。

次の方法で CPU 構成を指定できます。

- vSphere Client を通じて使用可能な属性と特殊な機能を使用します。vSphere Client を使用して ESXi ホストまたは vCenter Server システムに接続できます。
- 特定の状況下で高度な設定を使用します。
- スクリプト化した CPU 割り当てに vSphere SDK を使用します。
- ハイパースレッドを使用します。

マルチコア プロセッサ

ホストが仮想マシンのマルチタスクを実行する場合、マルチコア プロセッサにはさまざまな利点があります。

注： このトピックでは、「メモリ」は物理 RAM または永続的なメモリを指します。

Intel および AMD は、複数のプロセッサ コアを 1 つの集積回路（通常はパッケージまたはソケットと呼ばれます）に組み合わせたプロセッサを開発しました。VMware では、1 つ以上のプロセッサ コアを装備できる 1 つのまとまりを表すのにソケットという用語を使用します。

各プロセッサ コア内には、1 つ以上の論理プロセッサが含まれます。たとえば、デュアルコア プロセッサは、同時に 2 つの仮想 CPU を実行できるため、シングルコア プロセッサのほぼ 2 倍のパフォーマンスを発揮します。同じプロセッサ内のコアは、通常、すべてのコアが使用する最終レベルの共有キャッシュを装備して構成されており、より低速のメイン メモリにアクセスする必要性は軽減されます。実行中の仮想マシンが、同じメモリ バス リソースで競合するようなメモリを多用するワークロードを実行している場合、物理プロセッサをメイン メモリに接続する共有メモリ バスによって論理プロセッサのパフォーマンスが制限されることがあります。

仮想マシンを実行する ESXiCPU スケジューラによって各プロセッサ コアの各論理プロセッサが独立して使用されるため、SMP システムと同様のパフォーマンスが得られます。たとえば、2 方向の仮想マシンは、仮想プロセッサを同じコアに属する論理プロセッサ上で実行することも、異なる物理コア上の論理プロセッサで実行することもできます。

ESXiCPU スケジューラは、プロセッサ トポロジと、そのプロセッサ コアと論理プロセッサとの関係を検出します。この情報を使用して、仮想マシンをスケジュール設定し、パフォーマンスを最適化します。

ESXiCPU スケジューラは、ソケット、コア、および論理プロセッサ間の関係も含め、プロセッサ トポロジを認識できます。スケジューラはトポロジ情報を使用して、さまざまなソケットへの仮想 CPU の配置を最適化します。この最適化により、キャッシュ全体の使用率を最大化し、仮想 CPU の移行を最小化してキャッシュのアフィニティを高めることができます。

ハイパースレッド

ハイパースレッド テクノロジーを使用すると、単一の物理プロセッサ コアを 2 つの論理プロセッサのように機能させることができます。プロセッサは、同時に 2 つの異なるアプリケーションを実行できます。論理プロセッサと物理

プロセッサの混同を避けるために、Intel は物理プロセッサをソケットと呼んでおり、このセクションの説明でもこの用語を使用します。

Intel Corporation は、Pentium IV および Xeon のプロセッサ ラインのパフォーマンスを向上させるハイパースレッドテクノロジーを開発しました。ハイパースレッドテクノロジーにより、単一のプロセッサ コアが 2 つの独立したスレッドを同時に実行できます。

ハイパースレッドはシステムのパフォーマンスを 2 倍にするわけではありませんが、アイドル リソースを有効に活用することによってパフォーマンスを向上でき、特定の重要なワークロード タイプについてスループットを向上させます。ビジーなコアの 1 つの論理プロセッサ上で実行中のアプリケーションは、非ハイパースレッド プロセッサ上で単独で実行される場合に取得するスループットの半分よりわずかに多いスループットを期待できます。ハイパースレッドのパフォーマンス向上はアプリケーションに強く依存し、アプリケーションによっては、ハイパースレッドを使用するとパフォーマンスが低下する場合があります。これは、両方の論理プロセッサ間で多くのプロセッサ リソース（キャッシュなど）が共有されるためです。

注： Intel ハイパースレッドテクノロジーを使用するプロセッサでは、メモリ キャッシュや機能単位など、コアのリソースの大部分を共有する 2 つの論理プロセッサを、1 つのコアに担当させることができます。このような論理プロセッサを、通常、スレッドと呼びます。

多くのプロセッサはハイパースレッドをサポートしておらず、その結果、1 つのコアにつき 1 つのスレッドだけを備えています。それらのプロセッサの場合、コアの数も論理プロセッサの数に一致します。次の各プロセッサはハイパースレッドをサポートしており、1 つのコアにつき 2 つのスレッドを備えています。

- Intel Xeon 5500 プロセッサ マイクロアーキテクチャに基づいたプロセッサ
- Intel Pentium 4 (HT 有効)
- Intel Pentium EE 840 (HT 有効)

ハイパースレッドと ESXi ホスト

ハイパースレッドが有効になっているホストは、ハイパースレッドがないホストと同様な動作をする必要があります。しかし、ハイパースレッドを有効にする場合は、特定の要素について検討する必要があることがあります。

ESXi ホストは、プロセッサ時間をインテリジェントに管理して、システム内のプロセッサ コア間で負荷が円滑に分散されるように保証します。CPU 0 と 1 はともに第 1 のコア上にあり、CPU 2 と 3 は第 2 のコア上にあるといったように、同じコア上の論理プロセッサは、連続した CPU 番号を持ちます。仮想マシンは、同じコアの 2 つの論理プロセッサ上よりも、2 つの異なるコア上に優先的にスケジュール設定されます。

論理プロセッサの作業がない場合、論理プロセッサは停止状態になり、論理プロセッサの実行リソースが解放されて、同じコア上のほかの論理プロセッサ上で実行中の仮想マシンが、コアの実行リソース全体を使用できます。VMware スケジューラは、この停止時間を適切に捕捉し、コアのリソース全体を使用して実行中の仮想マシンを、半分のコア上で実行されている仮想マシンよりも多くチャージします。このプロセッサ管理方法により、標準の ESXi リソース割り当てルールにサーバが違反しないようにします。

ハイパースレッドを使用するホストでは、CPU アフィニティを有効にする前にリソース管理の必要性を検討してください。たとえば、優先順位の高い仮想マシンを CPU 0 にバインドしていて、別の優先順位の高い仮想マシンを CPU 1 にバインドしている場合、この 2 台の仮想マシンは同じ物理コアを共有する必要があります。この場合、これらの仮想マシンのリソース要求を満たせないことがあります。カスタム アフィニティ設定が、ハイパースレッドシステムにとって意味があることを確認してください。

ハイパースレッドの有効化

ハイパースレッドを有効にするには、まず、使用するシステムの BIOS 設定でハイパースレッドを有効にし、vSphere Client でオンにする必要があります。ハイパースレッドは、デフォルトで有効になっています。

使用している CPU がハイパースレッドをサポートしているかどうかを確認するには、システムのドキュメントを参照してください。

手順

- 1 システムがハイパースレッド テクノロジーをサポートしていることを確認します。
- 2 システムの BIOS でハイパースレッドを有効にします。
メーカーによって、このオプションは [論理プロセッサ] と呼ばれることもあれば、[ハイパースレッディングの有効化] と呼ばれることもあります。
- 3 ESXi のハイパースレッドが有効なことを確認します。
 - a vSphere Client で、ホストを参照して移動します。
 - b [構成] をクリックします。
 - c [システム] の [システムの詳細設定] をクリックし、[VMkernel.Boot.hyperthreading] を選択します。
設定を有効にするには、ホストを再起動する必要があります。値が **true** ならば、ハイパースレッドは有効です。
- 4 [ハードウェア] の [プロセッサ] をクリックして、論理プロセッサ数を表示します。

結果

ハイパースレッドが有効になりました。

CPU アフィニティの使用

各仮想マシンに CPU アフィニティ設定を指定することにより、仮想マシンの割り当てをマルチプロセッサ システム内の使用可能なプロセッサのサブセットに制限できます。この機能を使用することで、各仮想マシンを指定したアフィニティ セットのプロセッサに割り当てることができます。

CPU アフィニティは、仮想マシンのプロセッサへの配置の制約を指定します。これは、仮想マシンの仮想マシン ホストへの配置の制約を指定する、仮想マシン間または仮想マシンとホスト間のアフィニティ ルールによって作成された関係とは異なります。

ここでは、CPU という用語は、ハイパースレッド システム上の論理プロセッサを指し、また、非ハイパースレッド システム上のコアのことを指します。

仮想マシンの CPU アフィニティ設定は、その仮想マシンに関連付けられているすべての仮想 CPU と、その仮想マシンに関連付けられているその他のすべてのスレッド（ワールドとも呼ばれる）に適用されます。これらの仮想マシン スレッドは、マウス、キーボード、画面、CD-ROM、およびその他のレガシー デバイスのエミュレーションに必要な処理を実行します。

表示を多用するワークロードなど、一部のケースでは、仮想 CPU とその他の仮想マシン スレッドの間で大量の通信が発生することがあります。仮想マシンのアフィニティ設定によって、これらのその他のスレッドを仮想マシンの仮想 CPU と同時にスケジュール設定できない場合、パフォーマンスが低下する可能性があります。たとえば、ユニプロセッサ仮想マシンに 1 つの CPU がアフィニティ設定されている場合や、2 方向の SMP 仮想マシンに 2 つの CPU がアフィニティ設定されている場合などです。

最大のパフォーマンスを得るために、手動でアフィニティを設定する場合は、少なくとも 1 つの仮想マシン スレッドをその仮想 CPU と同時にスケジュール設定できるように、アフィニティ設定に少なくとも 1 つの物理 CPU を追加して含めておくことをお勧めします。たとえば、ユニプロセッサ仮想マシンに少なくとも 2 つの CPU をアフィニティ設定したり、2 方向の SMP 仮想マシンに少なくとも 3 つの CPU をアフィニティ設定します。

特定のプロセッサへの仮想マシンの割り当て

CPU アフィニティを使用すると、特定のプロセッサに仮想マシンを割り当てることができます。これにより、仮想マシンの割り当てをマルチプロセッサ システム内で使用可能な特定のプロセッサだけに制限できます。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
 - a データセンター、フォルダ、クラスタ、リソース プール、またはホストを選択して、仮想マシンを検索します。
 - b [仮想マシン] を選択します。
- 2 仮想マシンを右クリックし、[設定の編集] をクリックします。
- 3 [仮想ハードウェア] の [CPU] を展開します。
- 4 [スケジュール設定のアフィニティ] で仮想マシンの物理プロセッサ アフィニティを選択します。

範囲を指定する場合は「-」を使用します。また、値を区切るときは、「,」を使用します。

たとえば、「0, 2, 4-7」と入力すると、プロセッサ 0、2、4、5、6 および 7 が指定されます。
- 5 仮想マシンを実行するプロセッサを選択し、[OK] をクリックします。

CPU アフィニティでの潜在的な問題

CPU アフィニティを使用する前に、いくつかの問題について検討する必要がある場合があります。

CPU アフィニティでの潜在的な問題には、次のようなものがあります。

- マルチプロセッサ システムの場合、ESXi システムは自動ロード バランシングを実行します。仮想マシン アフィニティを手動で指定しないようにすると、プロセッサ間の負荷を均衡させるためのスケジューラの機能が向上します。
- アフィニティは、仮想マシンに指定された予約とシェアを満たす ESXi ホストの機能を妨害する場合があります。
- CPU アドミッション コントロールはアフィニティを考慮しないので、手動のアフィニティ設定を持つ仮想マシンは、予約を完全に受け取るとはかぎりません。

手動のアフィニティ設定がない仮想マシンは、手動のアフィニティ設定がある仮想マシンから不都合な影響を受けません。

- あるホストから別のホストへ仮想マシンを移動する場合、新しいホストの持つプロセッサの数が異なる場合があるので、アフィニティが適用しなくなることがあります。
- NUMA スケジューラは、アフィニティを使用してすでに特定のプロセッサに割り当て済みの仮想マシンを管理できない場合があります。詳細については、『16 章 NUMA システムと ESXi の併用』を参照してください。
- アフィニティは、マルチコア プロセッサまたはハイパースレッド プロセッサ上の仮想マシンをスケジュール設定するホストの機能に影響を及ぼすため、そういったプロセッサで共有されるリソースを最大限に活用できます。

ホストの電源管理ポリシー

ESXi で、ホスト ハードウェアが提供するいくつかの電源管理機能を適用して、パフォーマンスと消費電力のバランスを調整できます。電源管理ポリシーを選択することにより、これらの機能を ESXi でどのように使用するかを制御できます。

高パフォーマンス ポリシーを選択すると、絶対的なパフォーマンスは高まりますが、1 ワットあたりの電力の使用効率とパフォーマンスは低下します。省電力ポリシーを使用すると、絶対的なパフォーマンスは低くなりますが、電力の使用効率は向上します。

管理対象ホストのポリシーは、VMware Host Client を使用して選択できます。ポリシーを選択しない場合、ESXi ではデフォルトで [balancing] が使用されます。

表 4-1. CPU 電源管理ポリシー

電源管理ポリシー	説明
高パフォーマンス	電源管理機能は使用しないでください。
balancing (デフォルト)	パフォーマンスへの影響を最小限に抑えてエネルギー消費を削減します
省電力	パフォーマンスが低下するおそれがありますがエネルギー消費を削減します
カスタム	ユーザー定義の電源管理ポリシーです。高度な設定が可能になります。

CPU が低い周波数で実行されると、電圧も低くなるため、電力消費を削減できます。このタイプの電源管理は、通常、Dynamic Voltage and Frequency Scaling (DVFS) と呼ばれます。ESXi では、仮想マシンのパフォーマンスに影響を与えないように、CPU 周波数を調整します。

CPU がアイドル状態の場合、ESXi は C ステートと呼ばれるさまざまなレベル (深度) の停止状態を適用できます。C ステートの深度が深いほど、CPU での電力消費は少なくなります。CPU が稼働を再開するまでに時間がかかります。CPU がアイドル状態になると、ESXi はアイドル状態の時間を予想するアルゴリズムを適用し、適切な C ステートを選択します。深い C ステートに移動しない電源管理ポリシーでは、ESXi は、アイドル状態の CPU に対して深度が最も浅い停止状態 (C1) を使用します。

CPU 電源管理ポリシーの選択

ホストの CPU 電源管理ポリシーは、vSphere Client を使用して設定します。

前提条件

ホスト システムの BIOS 設定で、オペレーティング システムによる電源管理の制御が許可されていることを確認します ([OS 制御] など)。ホスト ハードウェアでオペレーティング システムによる電源管理が許可されていない場合、使用できるのは "サポート対象外" ポリシーのみです。(一部のシステムでは、"高パフォーマンス" ポリシーのみ使用できます)。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] をクリックします。
- 3 [ハードウェア] の [概要] で [電源管理] を選択し、[編集] ボタンをクリックします。
- 4 ホストの電源管理ポリシーを選択し、[OK] をクリックします。

この選択されたポリシーはホストの構成に保存され、起動時に再度使用されます。このポリシーはいつでも変更可能で、サーバを再起動する必要はありません。

ホストの電源管理用のカスタム ポリシー パラメータの構成

ホストの電源管理用にカスタム ポリシーを使用する場合、ESXi では、電源管理ポリシーはいくつかの詳細設定パラメータの値に基づきます。

前提条件

電源管理ポリシーに [カスタム] を選択します。これは [CPU 電源管理ポリシーの選択](#) に記載されているとおりです。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] をクリックします。
- 3 [システム] の [システムの詳細設定] を選択します。
- 4 右側のペインで、カスタム ポリシーに影響する電源管理パラメータを編集できます。

カスタム ポリシーに影響する電源管理パラメータには、[(カスタム ポリシー項目)] で始まる説明が付けられています。その他すべての電源パラメータが、電源管理ポリシーのすべてに影響します。

- 5 パラメータを選択して [編集] ボタンをクリックします。

注： デフォルトの電源管理パラメータの値は、[バランシング済みポリシー](#)と一致します。

パラメータ	説明
Power.UsePStates	プロセッサが高負荷の状態のときに電力消費を抑えるため、ACPI P ステートを使用します。
Power.MaxCpuLoad	指定した CPU の使用率をリアルタイムで下回ると、CPU の電力消費を抑えるため、P ステートを使用します。
Power.MinFreqPct	指定した CPU 周波数の割合を下回る場合、P ステートを使用しません。

パラメータ	説明
Power.UseStallCtr	キャッシュ ミスなどのイベントを待機して CPU が頻繁にアイドル状態になる場合に、深度が深い P ステートを使用します。
Power.TimerHz	各 CPU がどの P ステートにあるべきかを ESXi が 1 秒間に再評価する回数を制御します。
Power.UseCStates	プロセッサがアイドル状態の場合に、深度が深い ACPI C ステート (C2 以下) を使用します。
Power.CStateMaxLatency	遅延がこの値を超えた場合、C ステートを使用しません。
Power.CStateResidencyCoef	CPU がアイドル状態になったときに、ホストが CPU のアイドル状態の継続時間を予測し、この時間が指定した値の遅延を下回る場合、最大深度の C ステートを選択します。大きな値を指定するほど、ESXi は深い深度の C ステートの使用を控え、低い値を指定するほど、C ステートを積極的に使用します。
Power.CStatePredictionCoef	CPU のアイドル状態の継続時間予測に使用する ESXi アルゴリズムのパラメータです。この値の変更は推奨されません。
Power.PerfBias	電力調整のパフォーマンスのバイアスを指定します (Intel のみ)。Intel プロセッサ上の MSR を、Intel 推奨値に設定します。Intel の推奨値は、パフォーマンスを高くする場合に 0、バランスをとる場合に 6、省電力の場合に 15 です。その他の値は定義されていません。

6 [OK] をクリックします。

メモリ仮想化の基礎

5

メモリ リソースを管理する前に、メモリ リソースが ESXi によってどのように仮想化され、使用されるかを理解しておく必要があります。

VMkernel は、ホスト上のすべての物理 RAM を管理します。VMkernel は、この管理対象物理 RAM の一部を自身専用に使います。残りが、仮想マシン用に使用可能になります。

仮想および物理メモリ スペースは、ページと呼ばれるブロックに区分されます。物理メモリが一杯になった場合、物理メモリ内に存在しない仮想ページのデータは、ディスクに格納されます。プロセッサ アーキテクチャに応じて、ページは通常 4 KB または 2 MB です。[メモリの詳細属性](#)を参照してください。

この章には、次のトピックが含まれています。

- [仮想マシンのメモリ](#)
- [メモリのオーバーコミット](#)
- [メモリの共有](#)
- [メモリ仮想化](#)
- [ラージ ページ サイズのサポート](#)

仮想マシンのメモリ

各仮想マシンは、構成されたサイズに基づいてメモリを消費し、加えて仮想化のためのオーバーヘッド メモリを消費します。

構成されたサイズは、ゲスト OS に提供されるメモリ容量です。これは、仮想マシンに割り当てられる物理 RAM 容量とは異なります。後者は、リソース設定（共有、予約、制限）やホストのメモリ プレッシャーのレベルに依存します。

たとえば、構成済みサイズが 1GB の仮想マシンについて考えます。ゲスト OS は起動するときに、物理メモリが 1GB の専用マシンで実行していると考えています。場合によっては、仮想マシンに 1GB が完全に割り当てられます。別の場合には、より小さな割り当てを受けることがあります。実際の割り当てにかかわらず、ゲスト OS は、物理メモリが 1GB の専用マシンで実行している場合と同じ動作を続けます。

シェア

予約以外に指定可能な場合、ここで仮想マシンの相対的な優先順位を指定します。

予約

メモリがオーバーコミットされる場合でも、ホストが仮想マシン用に予約する物理 RAM 容量の、保証された下限です。過大なページングがなく効率的に動作するのに十分なメモリを仮想マシンが確実に利用できるよう、予約のレベルを設定します。

仮想マシンは、その予約内のすべてのメモリを消費した後、たとえアイドル状態になった場合でもその量のメモリを保持することができ、そのメモリは解放されません。一部のゲスト OS (Linux など) は、起動直後に構成済みメモリの全部にはアクセスできない場合があります。仮想マシンがその予約内のすべてのメモリを消費するまで、VMkernel はその予約の任意の未使用部分を他の仮想マシンに割り当てることができます。しかし、ゲストのワークロードが増加し、仮想マシンがその予約分全体を消費した場合、仮想マシンはそのメモリを保持できません。

制限

ホストが仮想マシンに割り当てることができる物理 RAM 容量の上限です。仮想マシンのメモリ割り当ては、構成済みサイズでも暗黙的に制限されます。

メモリのオーバーコミット

システムは実行中の仮想マシンごとに、仮想マシンの予約 (ある場合) 用と、仮想マシンの仮想化オーバーヘッド用に物理 RAM を予約します。

すべての仮想マシンの構成済み総メモリ サイズが、ホストで使用可能な物理メモリ容量を超える可能性があります。ただし、必ずしもメモリがオーバーコミットされることを意味するわけではありません。すべての仮想マシンの動作中のメモリ フットプリントの合計がホストのメモリ サイズを超える場合に、メモリはオーバーコミットされます。

ESXi ホストのメモリ管理手法のおかげで、仮想マシンは、ホストで利用可能な物理 RAM より多い仮想 RAM を使用できます。たとえば、2GB のメモリを持つホストがある場合に、それぞれ 1GB のメモリを持つ 4 台の仮想マシンを実行できます。この場合、メモリがオーバーコミットされます。たとえば、4 つの仮想マシンすべてがアイドル状態の場合、消費メモリの合計は 2 GB を大幅に下回っている可能性があります。ただし、4 GB の仮想マシンすべてがアクティブにメモリを消費している場合は、メモリ フットプリントが 2 GB を超えて、ESXi ホストがオーバーコミットされる可能性があります。

オーバーコミットが有用なのは、一般に、仮想マシンには負荷の軽いものと負荷の重いものがあり、相対的な動作レベルが時間とともに変動するためです。

メモリの使用率を向上させるために、ESXi ホストは、アイドル状態の仮想マシンから、より多くのメモリを必要とする仮想マシンにメモリを移動します。予約またはシェアのパラメータを使用して、重要な仮想マシンにメモリを優先的に割り当てます。このメモリは、使用されていない場合にほかの仮想マシンで利用できます。ESXi は、ホストのメモリが大きくオーバーコミットされていない場合でも、バルーン、メモリ共有、メモリ圧縮、スワップなど、さまざまなメカニズムを実装して適切なパフォーマンスを実現します。

仮想マシンのメモリがオーバーコミットされた環境で予約可能なメモリをすべて消費すると、ESXi ホストはメモリ不足になる可能性があります。パワーオンされている仮想マシンは影響を受けませんが、新しい仮想マシンはメモリが不足しているためにパワーオンに失敗する場合があります。

注： すべての仮想マシンのメモリ オーバーヘッドも予約済みとして認識されます。

また、[メモリ圧縮](#) で説明するように、メモリのオーバーコミット時に仮想マシンのパフォーマンスを向上させるため、ESXi ホストではメモリ圧縮がデフォルトで有効になっています。

メモリの共有

メモリ共有は、ホスト上でメモリ密度を高めるための独自の ESXi 技法です。

メモリ共有は、いくつかの仮想マシンが同じゲスト OS のインスタンスを実行している状態を利用する技術です。これらの仮想マシンは、同じアプリケーションまたはコンポーネントをロードする場合や、共通のデータを含む場合があります。このような場合、ホストは、独自の透過的なページ共有 (TPS) 技法を使用してメモリ ページの冗長コピーを除去します。メモリの共有により、仮想マシンで実行しているワークロードが使用するメモリが、物理マシン上で実行する場合の使用量よりも少なくなることがよくあります。結果的に、より高いレベルのオーバーコミットを効率的にサポートできます。メモリ共有によって節約されるメモリ容量は、ワークロードがほぼ同じマシンに集中しているために、より多くのメモリが解放されるかどうかによって異なります。ワークロードが分散するほど、メモリ節約の度合いが低くなる可能性があります。

注： セキュリティ上の懸念があるため、仮想マシン間の透過的なページ共有はデフォルトで無効になっており、ページ共有は仮想マシン内のメモリ共有に制限されます。ページ共有は、複数の仮想マシン間で行われることはなく、特定の仮想マシン内でのみ行われます。詳細については[仮想マシン間でのメモリの共有](#)を参照してください。

メモリ仮想化

仮想化によって導入される特別なレベルのメモリ マッピングのため、ESXi は、すべての仮想マシン間でメモリを効率的に管理できます。

仮想マシンの一部の物理メモリが共有ページにマッピングされるか、マッピングされていないページまたはスワップアウトされているページにマッピングされます。

ホストは仮想メモリ管理を行います。ゲスト OS が認識したり、ゲスト OS のメモリ管理サブシステムの干渉を受けたりすることはありません。

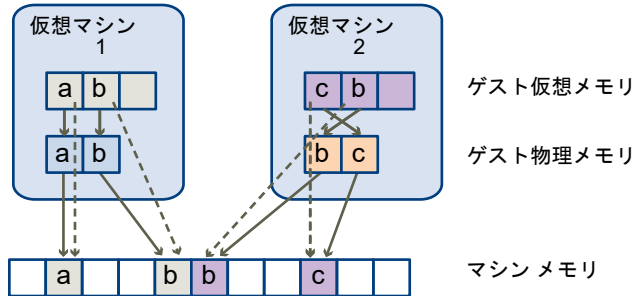
各仮想マシンの VMM は、ゲスト OS の物理メモリ ページから、基盤となるマシン上の物理メモリ ページへのマッピングを保持します (VMware は、基盤となるホスト物理ページを「マシン」ページとして参照し、ゲスト OS の物理ページを「物理」ページとして参照します)。

各仮想マシンは、ゼロ ベースの連続したアドレス可能物理メモリ スペースを参照します。各仮想マシンによって使用される、サーバの基盤となるマシン メモリは、必ずしも連続しません。

ゲストの仮想アドレスとゲストの物理アドレスとのマッピングは、ゲスト OS によって管理されます。ハイパーバイザーの役割は、ゲストの物理アドレスをマシン アドレスに変換することだけです。ハードウェア アシストによるメモリ仮想化では、ハードウェア機器を活用し、ゲストのページ テーブルとハイパーバイザーによって管理されるネストしたページ テーブルを使用して統合されたマッピングを生成します。

図は、ESXi でのメモリ仮想化の実装を示しています。

図 5-1. ESXi のメモリ マッピング



- ボックスはページを表し、矢印はそれぞれメモリ マッピングを表します。
- ゲスト仮想メモリからゲスト物理メモリへの矢印は、ゲスト OS 内のページ テーブルが保持するマッピングを表します (仮想メモリから x86 アーキテクチャ プロセッサのリニア メモリへのマッピングは示していません)。
- ゲスト物理メモリからマシン メモリへの矢印は、VMM が保持するマッピングを表します。
- 点線の矢印は、ゲストの仮想メモリから、シャドウ ページ テーブル (これも VMM が保持) のマシン メモリへのマッピングを表します。仮想マシンを実行している基盤となるプロセッサは、シャドウ ページ テーブルのマッピングを使用します。

ハードウェア アシストによるメモリ仮想化

AMD SVM-V や Intel Xeon 5500 シリーズなどの一部の CPU は、ページ テーブルの 2 つのレイヤーを使用することで、ハードウェアからメモリ仮想化をサポートする機能を提供しています。

注： このトピックでは、「メモリ」は物理 RAM または永続的なメモリを参照できます。[]

ページ テーブルの 1 番目のレイヤーには、ゲストの仮想メモリから物理メモリへの変換が格納されています。ページ テーブルの 2 番目のレイヤーには、ゲストの物理メモリからマシンへの変換が格納されています。TLB (変換索引バッファ) は、プロセッサの MMU (メモリ管理ユニット) ハードウェアが維持する変換のキャッシュです。TLB ミスはこのキャッシュのミスであり、ハードウェアはメモリに (おそらく複数回) アクセスして必要な変換を検索する必要があります。TLB でゲストの仮想アドレスが参照できない場合、ハードウェアが両方のページ テーブルを検索し、ゲストの仮想アドレスをマシン アドレスに変換します。ページ テーブルの 1 番目のレイヤーは、ゲスト OS によって維持されます。VMM は、ページ テーブルの 2 番目のレイヤーのみを維持します。

パフォーマンスについての考慮事項

ハードウェア アシストを使用すると、ソフトウェア メモリ仮想化のオーバーヘッドをなくすことができます。特に、ハードウェア アシストによって、シャドウ ページ テーブルをゲスト ページ テーブルと同期させるために必要なオーバーヘッドがなくなります。ただし、ハードウェア アシストを使用すると、TLB ミスの遅延が大幅に増えます。デフォルトでは、ハイパーバイザーは、ハードウェア アシスト モードで大きなページを使用して、TLB ミスのコストを削減します。結果として、ハードウェア アシストを使用することによってワークロードにメリットがあるかどうかは、ソフトウェアによるメモリ仮想化を使用した場合に発生するメモリ仮想化のオーバーヘッドに大きく依存します。ワークロードに含まれるページ テーブル アクティビティ (プロセス作成、メモリのマッピング、コンテキスト スイッチなど) が少ない場合、ソフトウェアの仮想化では、大きなオーバーヘッドは発生しません。逆に、ワークロードに大量のページ テーブル アクティビティが含まれている場合は、ハードウェア アシストを使用すると有益です。

デフォルトでは、ハイパーバイザーは、ハードウェア アシスト モードで大きなページを使用して、TLB ミスのコストを削減します。最適なパフォーマンスは、ゲストの仮想アドレスからゲストの物理アドレスへの変換とゲストの物理アドレスからマシン アドレスへの変換の両方で大きなページを使用することで実現されます。

オプション **LPage.LPageAlwaysTryForNPT** によって、ゲストの物理アドレスからマシン アドレスへの変換で大きなページを使用するようにポリシーを変更できます。詳細については、[メモリの詳細属性](#)を参照してください。

ラージ ページ サイズのサポート

ESXi は、ラージ ページ サイズに対して限定的なサポートを提供しています。

x86 アーキテクチャにより、システム ソフトウェアで 4 KB、2 MB、および 1 GB のページを使用できます。4 KB のページはスモール ページと呼び、2 MB と 1 GB のページはラージ ページと呼びます。ラージ ページは、トランスレーション ルックアサイド バッファ (TLB) の負荷を軽減し、ページ テーブル ウォークのコストを削減します。これにより、ワークロードのパフォーマンスが向上します。

仮想化環境では、ラージ ページは、ハイパーバイザーとゲスト OS によって個別に使用できます。ラージ ページがゲストとハイパーバイザーによって使用された場合、パフォーマンスに著しい影響が及びますが、通常、ラージ ページがハイパーバイザー レベルでのみ使用された場合でも、パフォーマンスには影響が及びます。

ESXi ハイパーバイザーは、デフォルトでゲスト vRAM のバッキングに 2 MB のページを使用します。vSphere 6.7 ESXi は、1 GB ページでのゲスト vRAM のバッキングに対して限定的なサポートを提供しています。詳細については、「1 GB ページでのゲスト vRAM のバッキング」を参照してください。

メモリ リソースの管理

6

vSphere Client を使用して、メモリ割り当ての設定に関する情報を表示し、それらの設定を変更できます。メモリ リソースを効果的に管理するには、メモリ オーバーヘッド、アイドル メモリ税、および ESXi ホストがメモリをどのように解放するかについて、よく理解しておく必要もあります。

メモリ リソースを管理する場合、メモリ割り当てを指定できます。メモリ割り当てをカスタマイズしない場合、ESXi ホストは、ほとんどの状況で良好に機能するデフォルトを使用します。

メモリ割り当てを指定する方法は、いくつかあります。

- vSphere Client []を通じて使用可能な属性と特殊な機能を使用します。vSphere Client を使用して ESXi ホストまたは vCenter Server システムに接続できます。
- 詳細設定を使用します。
- スクリプト化したメモリ割り当てに vSphere SDK を使用します。

注： この章では、「メモリ」は物理 RAM または永続的なメモリを指します。

この章には、次のトピックが含まれています。

- [メモリ オーバーヘッドについて](#)
- [ESXi ホストによるメモリ割り当ての方法](#)
- [メモリの解放](#)
- [スワップ ファイルの使用](#)
- [仮想マシン間でのメモリの共有](#)
- [メモリ圧縮](#)
- [メモリ使用量の測定とタイプの識別](#)
- [メモリの信頼性](#)
- [システム スワップについて](#)

メモリ オーバーヘッドについて

メモリ リソースの仮想化には、関連するいくつかのオーバーヘッドがあります。

ESXi 仮想マシンは、2 種類のメモリ オーバーヘッドを被ることがあります。

- 仮想マシン内のメモリにアクセスするための追加の時間。
- ESXi ホストが自身のコードとデータ構造のために必要とする、各仮想マシンに割り当てられたメモリを超える余分のスペース。

ESXi のメモリ仮想化では、メモリ アクセスへの時間のオーバーヘッドはほとんど追加されません。プロセッサのページングハードウェアがページテーブル（ソフトウェアベースのアプローチ用のシャドウ ページ テーブルまたはハードウェア アシストによるアプローチ用の 2 レベルのページ テーブル）を直接使用するので、仮想マシン内のほとんどのメモリ アクセスは、アドレス変換のオーバーヘッドをせずに実行できます。

メモリ スペース オーバーヘッドには、2 つの構成要素があります。

- VMkernel が必要とする、システム全体の固定されたオーバーヘッド。
- 各仮想マシンの追加のオーバーヘッド

オーバーヘッド メモリには、仮想マシンのフレーム バッファと、シャドウ ページ テーブルなどの各種の仮想化データ構造のために予約されたスペースが含まれます。オーバーヘッド メモリは、仮想 CPU の数とゲスト OS 用に構成されたメモリに依存します。

仮想マシン上のオーバーヘッド メモリ

仮想マシンをパワーオンするには、特定の量の使用可能なオーバーヘッド メモリが必要です。このオーバーヘッドの量を認識しておく必要があります。

仮想マシンに必要なオーバーヘッド メモリの量は、vCPU の数、メモリ サイズ、デバイスの数とタイプ、モニタで使用している実行モード、仮想マシンのハードウェア バージョンなどの多数の要素によって決まります。また、使用している vSphere のバージョンも、必要なメモリの量に影響する可能性があります。仮想マシンに必要なオーバーヘッド メモリの量は、VMX によって自動的に計算されます。

特定の構成に必要なオーバーヘッド メモリの量を確認するには、まず対象となる仮想マシンをパワーオンします。vmware.log ファイルを調べます。仮想マシンをパワーオンすると、必要なオーバーヘッド メモリの量がログに出力されます。ログ内で VMMEM を検索して、仮想マシン用に予約されているオーバーヘッド メモリの最初の正確な量を確認します。

ESXi ホストによるメモリ割り当ての方法

ホストは、メモリがオーバーコミットされていないかぎり、制限 パラメータによって指定されたメモリを各仮想マシンに割り当てます。ESXi ホストは、指定された物理メモリ サイズを上回るメモリを仮想マシンに割り当てることはしません。

たとえば、1GB の仮想マシンには、デフォルトの制限（制限なし）またはユーザー指定の制限（たとえば 2GB）がある場合があります。このどちらの場合も、ESXi ホストは、1GB（仮想マシン用に指定された物理メモリ サイズ）を上回るメモリは割り当てません。

メモリがオーバーコミットされる場合、各仮想マシンには、[予約] によって指定された量と [制限] によって指定された量の間のメモリを割り当てられます。仮想マシンの予約を超えて仮想マシンに与えられるメモリの量は、通常、現在のメモリ負荷とともに変動します。

ホストは、仮想マシンに割り当てられたシェアの数と最新の作業セット サイズの概算に基づいて、各仮想マシンへの割り当てを決定します。

- シェア：ESXi ホストは、修正された比例的なシェアのメモリ割り当てポリシーを使用します。メモリ シェアは、使用可能な物理メモリの一部を仮想マシンに割り当てます。
- 作業セット サイズ：ESXi ホストは、仮想マシン実行時間の連続的な期間にわたってメモリの動作を監視することによって、仮想マシンの作業セットを見積もります。見積もりは、いくつかの期間にわたって平滑化されます。これには、作業セット サイズの拡大にすばやく対応し、作業セット サイズの縮小にゆっくりと対応する技法が使用されます。

この方法により、アイドル メモリを解放した仮想マシンがメモリをより有効に使用し始めた場合、共有ベースの割り当てすべてを短時間のうちに確実に獲得できます。

作業セットのサイズを見積もるために、デフォルトで 60 秒間、メモリの動作が監視されます。このデフォルトを変更するには、Mem.SamplePeriod 詳細設定を調整します。[ホストの詳細属性の設定](#) を参照してください。

アイドル仮想マシンのメモリ税

現在割り当てられているメモリのすべてを仮想マシンが有効に使用していない場合、ESXi は、使用中のメモリに対してより、アイドル メモリに対してより多くのコストを課します。これは、仮想マシンがアイドル メモリを蓄えないうようにするのに役立ちます。

アイドル メモリ税は、累進的に課税されます。仮想マシンのアクティブなメモリに対するアイドル メモリの割合が大きくなるにつれて、実効税率も増加します（以前のバージョンの ESXi は階層リソース プールをサポートしておらず、仮想マシンのすべてのアイドル メモリには均一に課税されていました）。

アイドル メモリの税率は、Mem.IdleTax オプションで変更できます。このオプションを Mem.SamplePeriod 詳細属性とともに使用して、仮想マシンのターゲット メモリ割り当ての決定方法を制御します。[ホストの詳細属性の設定](#) を参照してください。

注： ほとんどの場合、Mem.IdleTax の変更は必要なく、適切でもありません。

VMX スワップ ファイル

仮想マシン実行可能（VMX）スワップ ファイルを使用すると、VMX 処理のために予約されるホストのオーバーヘッド メモリの量を大幅に削減できます。

注： VMX スワップ ファイルは、ホスト スワップ キャッシュ機能へのスワップや通常のホスト レベルのスワップ ファイルとは関係がありません。

ESXi は、さまざまな目的で仮想マシンごとにメモリを予約します。仮想マシン モニタ（VMM）や仮想デバイスなどの特定のコンポーネントに必要なメモリは、仮想マシンのパワーオン時にすべて予約されます。しかし、VMX 処理のために予約されるオーバーヘッド メモリの一部は、スワップすることができます。VMX スワップ機能によって、VMX メモリの予約が大幅に削減されます（たとえば、仮想マシンごとに 50MB 以上だったのが、仮想マシンごとに 10MB 程度など）。これにより、ホストのメモリが割り当てられすぎているときに、残りのメモリをスワップアウトすることが可能になり、仮想マシンごとのオーバーヘッド メモリの予約量を削減できます。

VMX スワップ メモリは、仮想マシンのパワーオン時に空きディスク領域が十分にあれば、ホストで自動的に作成されます。

メモリの解放

ESXi ホストは、仮想マシンからメモリを解放できます。

ホストは、予約によって指定された量のメモリを、仮想マシンに直接割り当てます。予約を超えるメモリは、ホストの物理リソースを使用して割り当てられるか、物理リソースを使用できない場合は、バルーニングやスワップなどの特別な方法を使用して処理されます。ホストは、2つの方法を使用して、仮想マシンに割り当てられたメモリの量を動的に拡大または縮小します。

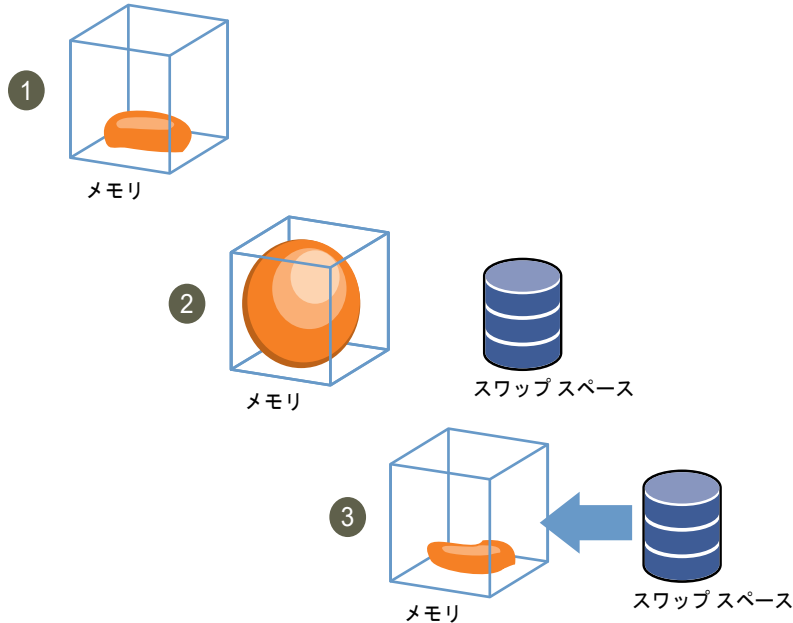
- ESXi システムは、仮想マシン内で実行中のゲスト OS にロードされたメモリ バルーン ドライバ (`vmmemctl`) を使用します。[メモリ バルーン ドライバ](#) を参照してください。
- ESXi システムは、ゲスト OS による関与なしに仮想マシンからサーバ スワップ ファイルにページをスワップアウトします。各仮想マシンには、独自のスワップ ファイルがあります。

メモリ バルーン ドライバ

メモリ バルーン ドライバ (`vmmemctl`) は、サーバと協同して、ゲスト OS によって最も価値がないと見なされたページを解放します。

このドライバは、同様のメモリ制約下のネイティブ システムの動作と緊密に一致する、予測可能なパフォーマンスを提供する専用のバルーニング技法を使用します。この技法は、ゲスト OS に対するメモリ プレッシャーを増減して、ゲスト自身のネイティブ メモリ管理アルゴリズムをゲストに使用させます。メモリが十分ではない場合、ゲスト OS は、どのページを解放するかを決定し、必要に応じて、それらのページを自身の仮想ディスクにスワップします。

図 6-1. ゲスト OS のメモリ バルーン化



注： 十分なスワップスペースを持つゲスト OS を構成する必要があります。ゲスト OS によっては、追加の制限事項があります。

必要な場合は、特定の仮想マシンについて [sched.mem.maxmemctl] パラメータを設定することによって、vmmemctl が解放するメモリの量を制限できます。このオプションは、仮想マシンから解放できるメモリの最大量をメガバイト（MB）単位で指定します。[仮想マシンの詳細属性の設定](#) を参照してください。

スワップファイルの使用

ゲスト スワップファイルの場所の指定、メモリがオーバーコミットされている場合のスワップスペースの予約、およびスワップファイルの削除が可能です。

ESXi ホストは、次のように vmmemctl ドライバが使用できない場合、または反応しない場合、スワップを使用して仮想マシンからメモリを強制的に解放します。

- インストールされていない。
- 明示的に無効にされている。
- 実行されていない（たとえば、ゲスト OS の起動中）。
- 現在のシステム要求を十分に満たすほど迅速にメモリを解放することが、一時的にできない。
- 適切に機能しているが、最大バルーンサイズに達している。

標準の要求ページング技法は、仮想マシンにページが必要な場合、ページをスワップバックします。

スワップファイルの場所

デフォルトでは、スワップファイルは仮想マシンの構成ファイルと同じ場所（VMFS データストア、vSAN データストア、または VMware vSphere® Virtual Volumes™ データストアのいずれか）に作成されます。vSAN デ

ータストアまたは vVol データストアでは、スワップ ファイルは個別の vSAN オブジェクトまたは vVol オブジェクトとして作成されます。

仮想マシンがパワーオンされると、ESXi ホストはスワップ ファイルを作成します。スワップ ファイルを作成できない場合、仮想マシンをパワーオンできません。デフォルトを使用せずに、次のように操作することもできます。

- 仮想マシンごとの構成オプションを使用して、データストアを別の共有ストレージの場所に変更する。
- ホストのローカル スワップを使用して、データストアをホスト上にローカルで格納するように指定する。このように指定すると、ホスト レベルでスワップするため、SAN の容量を節約できます。ただし、スワップ元ホスト上のローカル スワップ ファイルへスワップされたページは、ネットワークを介してスワップ先ホストへ転送されるため、vSphere vMotion のパフォーマンスがわずかに低下する場合があります。現在、vSAN データストアおよび vVol データストアをホスト ローカル スワップに指定することはできません。

DRS クラスタでのホストのローカル スワップの有効化

ホストのローカル スワップを使用すると、ホスト上にローカルに格納されているデータストアをスワップ ファイルの場所として指定できます。DRS クラスタで、ホストのローカル スワップを有効にすることができます。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 [構成] をクリックします。
- 3 [構成] の [全般] を選択して、スワップ ファイルの場所を表示します。次に、[編集] をクリックしてスワップ ファイルを変更します。
- 4 [ホストが指定するデータストア] オプションを選択し、[OK] をクリックします。
- 5 vSphere Client で、クラスタ内のホスト 1 台を参照して移動します。
- 6 [構成] をクリックします。
- 7 [仮想マシン] で、[スワップ ファイルの場所] を選択します。
- 8 [編集] をクリックし、使用するローカル データストアを選択して、[OK] をクリックします。
- 9 クラスタ内の各ホストに対して、**手順 5** から **手順 8** の手順を繰り返します。

結果

これで、ホストのローカル スワップが DRS クラスタで有効になりました。

スタンドアロン ホストでのホストのローカル スワップの有効化

ホストのローカル スワップを使用すると、ホスト上にローカルに格納されているデータストアをスワップ ファイルの場所として指定できます。スタンドアロン ホストで、ホストのローカル スワップを有効にすることができます。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] をクリックします。
- 3 [仮想マシン] で、[スワップ ファイルの場所] を選択します。

- 4 [編集] をクリックし、[選択したデータストア] を選択します。
- 5 リストからローカル データストアを選択し、[OK] をクリックします。

結果

これで、ホストのローカル スワップがスタンドアローン ホストで有効になりました。

スワップ スペースとメモリのオーバーコミット

仮想マシンごとのスワップ ファイルで、予約されていない仮想マシン メモリ（これは予約されたメモリ サイズと構成済みメモリ サイズの差です）に対して、スワップ スペースを予約する必要があります。

このスワップ予約は、どのような状況下でも ESXi ホストが仮想マシン メモリを維持できるようにするのに必要です。実際には、ホスト レベルのスワップ スペースのごく一部分だけが使用されます。

バルーニングによって生じるゲスト内スワップをサポートするために、ESXi でメモリをオーバーコミットする場合は、必ず、ゲスト OS にも十分なスワップ スペースが存在することを確認してください。このゲスト レベルのスワップ スペースは、仮想マシンの構成済みメモリ サイズと仮想マシンの [予約] の差以上である必要があります。

注意： メモリがオーバーコミットされ、ゲスト OS に構成されたスワップ スペースが不十分な場合、仮想マシン内のゲスト OS は起動に失敗します。

仮想マシンの障害を防ぐには、仮想マシン内のスワップ スペースのサイズを拡大します。

- Windows ゲスト OS：Windows オペレーティング システムでは、スワップ スペースをページング ファイルと呼びます。一部の Windows オペレーティング システムは、十分な空きディスク領域がある場合、ページング ファイルのサイズを自動的に拡大しようとします。

Microsoft Windows のドキュメントを参照するか、Windows のヘルプ ファイルで「ページング ファイル」を検索します。仮想メモリ ページング ファイルのサイズを変更する手順に従ってください。

- Linux ゲスト OS：Linux オペレーティング システムでは、スワップ スペースをスワップ ファイルと呼びます。スワップ ファイルの拡大については、次の Linux の man ページを参照してください。
 - mkswap：Linux スワップ領域を設定します。
 - swapon：ページングとスワップのためのデバイスとファイルを有効にします。

メモリが大量にあり、仮想ディスクが小さいゲスト OS（たとえば、8GB の RAM と 2GB の仮想ディスクを持つ仮想マシン）は、スワップ スペースが不十分な場合の影響を受けやすくなります。

注： スワップ ファイルは、シン プロビジョニングされた LUN 上に保存しないでください。シン プロビジョニングされた LUN 上に保存されたスワップ ファイルを使用する仮想マシンを実行した場合、スワップ ファイルの増大にエラーが起きる可能性があり、仮想マシンが終了する恐れがあります。

大きなスワップ ファイルを作成する場合（100GB を超える場合など）、仮想マシンのパワーオンに要する時間が大幅に長くなる可能性があります。これを回避するには、大規模な仮想マシン用に予約を増やして設定します。

また、ホストのローカルのスワップ ファイルを使用することで、コストの低いストレージにスワップ ファイルを配置することもできます。

ホストの仮想マシン スワップファイル プロパティの設定

ホストでスワップ ファイルの場所を構成し、vSphere Client 内で仮想マシン スワップ ファイルのデフォルトの場所を指定します。

デフォルトで仮想マシンのスワップファイルは、データストア上のほかの仮想マシン ファイルを含むフォルダにあります。ただし、他のデータストアに仮想マシンのスワップ ファイルを配置するように、ホストを設定できます。

この方法で、低コストまたは高パフォーマンスのストレージに仮想マシンのスワップ ファイルを配置することができます。このホスト レベルの設定は、各仮想マシンにオーバーライドできます。

スワップファイルにデフォルトとは異なる配置場所を設定すると、vMotion での移行時間がより長くなる可能性があります。vMotion のパフォーマンスを高めるには、仮想マシンを仮想マシンのスワップ ファイルと同じディレクトリではなく、ローカル データストアに格納してください。仮想マシンがローカル データストアに格納されている場合は、その他の仮想マシン ファイルとともにスワップ ファイルを格納しても、vMotion のパフォーマンスは向上しません。

前提条件

必要な権限 : ホスト マシン.構成.ストレージ パーティション構成

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] をクリックします。
- 3 [仮想マシン] の [スワップ ファイルの場所] をクリックします。

選択したスワップ ファイルの場所が表示されます。選択したホストでスワップ ファイルの場所を構成する機能がサポートされていない場合は、このタブにそのように表示されます。

ホストがクラスタに属しており、クラスタ設定でスワップファイルを仮想マシンと同じディレクトリに格納するように指定されている場合は、ホストの [構成] タブではスワップファイルの場所を編集できません。この場合、ホストのスワップ ファイルの場所を変更するには、クラスタ設定を編集します。

- 4 [編集] をクリックします。
- 5 スワップ ファイルを格納する場所を選択します。

オプション	説明
仮想マシン ディレクトリ	仮想マシン構成ファイルと同じディレクトリにスワップファイルを格納します。
特定のデータストアを使用する	指定した場所にスワップ ファイルを格納します。 ホストが指定するデータストアにスワップファイルを格納できない場合、スワップファイルは仮想マシンと同じフォルダに格納されます。

- 6 (オプション) [特定のデータストアを使用] を選択した場合は、データストアをリストから選択します。
- 7 [OK] をクリックします。

結果

仮想マシンのスワップ ファイルが、選択された場所に格納されます。

クラスタ用の仮想マシンのスワップ ファイルのための場所の設定

デフォルトで仮想マシンのスワップ ファイルは、データストア上のほかの仮想マシン ファイルを含むフォルダにあります。ただし、代わりにクラスタでホストを構成し、別のデータストアを選択して仮想マシンのスワップ ファイルを配置することもできます。

低コストまたは高パフォーマンスのストレージに仮想マシンのスワップ ファイルを配置するため、必要に応じて別のスワップ ファイルの場所を設定できます。

前提条件

クラスタで仮想マシンのスワップ ファイルの場所を設定する前に、[ホストの仮想マシン スワップファイル プロパティの設定](#)に記載されているように、クラスタ内のホストで仮想マシンのスワップ ファイルの場所を設定する必要があります。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 [構成] をクリックします。
- 3 [構成] - [全般] を選択します。
- 4 スワップ ファイル ロケーションの横にある [編集] をクリックします。
- 5 スワップ ファイルを格納する場所を選択します。

オプション	説明
仮想マシン ディレクトリ	仮想マシン構成ファイルと同じディレクトリにスワップ ファイルを格納します。
ホストが指定するデータストア	ホストの設定で指定された場所にスワップ ファイルを格納します。 ホストが指定するデータストアにスワップ ファイルを格納できない場合、スワップ ファイルは仮想マシンと同じフォルダに格納されます。

- 6 [OK] をクリックします。

スワップ ファイルの削除

ホストに障害が起きた場合、そのホストにスワップ ファイルを使用している実行中の仮想マシンがあるとスワップ ファイルは存在し続け、ディスク スペースを何ギガバイトも消費します。スワップ ファイルを削除すると、この問題を解消できます。

手順

- 1 障害のあったホスト上の仮想マシンを再起動します。
- 2 仮想マシンを停止します。

結果

仮想マシンのスワップ ファイルが削除されます。

仮想マシン間でのメモリの共有

多くの ESXi ワークロードでは、複数の仮想マシン間で（単一の仮想マシン内でも）メモリを共有する機会があります。

ESXi のメモリ共有は、時間の経過とともに共有の機会をスキャンするバックグラウンド動作として実行されます。節約されるメモリの量は、時間とともに変動します。ある程度一定のワークロードの場合、共有のすべての機会が利用されるまで、量は一般にゆっくりと増加します。

特定のワークロードについてのメモリ共有の有効性を判定するには、ワークロードを実行してみて、`resxtop` または `esxtop` を使用して実際の節約を観察します。その情報は、メモリ ページの対話モードの `PSHARE` フィールドにあります。

Mem.ShareScanTime と **Mem.ShareScanGHz** の詳細設定を使用すると、メモリ共有の機会を特定するためにシステムがメモリをスキャンする速度を制御できます。

sched.mem.pshare.enable オプションを設定することで、個々の仮想マシンの共有を構成することもできます。

セキュリティ上の懸念があるため、仮想マシン間の透過的なページ共有はデフォルトで無効になっており、ページ共有は仮想マシン内のメモリ共有に制限されます。このため、ページ共有は複数の仮想マシン間で行われることはなく、特定の仮想マシン内でのみ行われます。ソルティングの概念は、システム管理者が透過的なページ共有のセキュリティへの影響について抱く懸念に対処するため、導入されました。ソルティングを使用すると、透過的なページ共有に参加する仮想マシンを、これまでと比べてより高精度で管理できます。新しいソルティング設定では、ページのソルト値と内容が一致する場合に限り、複数の仮想マシンでページを共有できます。ソルティングの有効化または無効化を設定するには、新しいホスト構成オプション **Mem.ShareForceSalting** を使用します。

詳細オプションを設定する方法については、[17 章 詳細属性](#)を参照してください。

メモリ圧縮

ESXi はメモリ圧縮によるキャッシュ機能を備えているため、メモリのオーバーコミットメントの使用時に仮想マシンのパフォーマンスを向上させます。メモリ圧縮はデフォルトで有効になっています。ホストのメモリがオーバーコミット状態になると、ESXi は仮想ページを圧縮し、メモリに保存します。

圧縮メモリへのアクセスは、ディスクにスワップされたメモリへのアクセスよりも早いため、ESXi でメモリ圧縮を使用すると、パフォーマンスに大きな影響を与えることなくメモリをオーバーコミットできます。仮想ページでスワップが必要な場合、ESXi はまずページを圧縮しようとします。2KB 以下に圧縮できたページは仮想マシンの圧縮用キャッシュに保存されるため、ホストの容量が増加します。

vSphere Client の詳細設定ダイアログ ボックスで、圧縮キャッシュの最大サイズの設定、およびメモリ圧縮の無効化を行うことができます。

メモリ圧縮キャッシュの有効化または無効化

メモリ圧縮はデフォルトで有効になっています。vSphere Client の [システムの詳細設定] を使用して、ホストのメモリ圧縮を有効または無効にできます。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] をクリックします。
- 3 [システム] の [システムの詳細設定] を選択します。
- 4 Mem.MemZipEnable を探して、[編集] ボタンをクリックします。
- 5 メモリ圧縮キャッシュを有効にするには 1 を、無効にするには 0 を入力します。
- 6 [OK] をクリックします。

メモリ圧縮キャッシュの最大サイズの設定

ホストの仮想マシン用に、メモリ圧縮キャッシュの最大サイズを設定できます。

圧縮キャッシュのサイズは、仮想マシンのメモリ サイズの割合 (%) で設定します。たとえば、仮想マシンのメモリ サイズが 1,000 MB のときに値 20 を入力した場合、ESXi では、仮想マシンの圧縮されたページを格納するのに、最大で 200 MB のホスト メモリを使用できます。

圧縮キャッシュのサイズを設定しない場合、ESXi ではデフォルト値の 10 % を使用します。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] をクリックします。
- 3 [システム] の [システムの詳細設定] を選択します。
- 4 M Mem.MemZipMaxPct を探して、[編集] ボタンをクリックします。
この属性の値によって、仮想マシンの圧縮キャッシュの最大サイズが決まります。
- 5 圧縮キャッシュの最大サイズを入力します。
この値は仮想マシンのサイズの割合で、5 ~ 100 % である必要があります。
- 6 [OK] をクリックします。

メモリ使用量の測定とタイプの識別

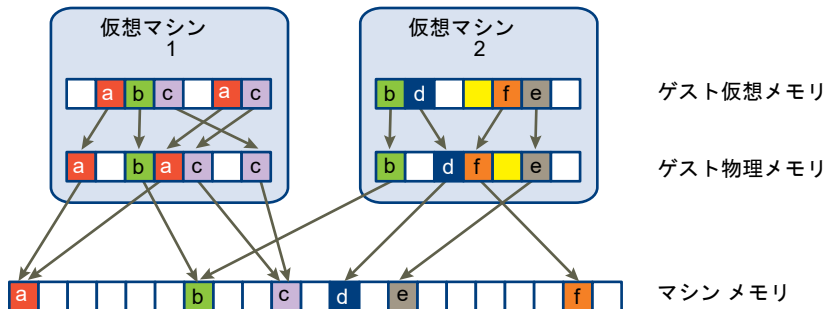
vSphere Client の [パフォーマンス] タブに、メモリ使用量の分析に使用できる、いくつかのメトリックが表示されます。

これらのメモリ メトリックには、ゲストの物理メモリを測定するものと、マシン メモリを測定するものがあります。たとえば、パフォーマンス メトリックを使用して検査できるメモリ使用量の 2 つのタイプは、ゲスト物理メモリとマシンメモリです。ゲスト物理メモリを測定するには、与えられたメモリ メトリック（仮想マシンの場合）または共有メモリ（ホストの場合）を使用します。しかし、マシンメモリを測定するには、消費されたメモリ（仮想マシンの場合）または一般共有メモリ（ホストの場合）を使用します。これらのメトリックが何を測定し、これらのメトリックをどのように解釈すべきかを知るためには、これらのタイプのメモリ使用量について、概念上の違いを理解することが重要です。

VMkernel はゲスト物理メモリをマシン メモリにマッピングしますが、必ずしも常に 1 対 1 でマッピングされるわけではありません。ゲスト物理メモリの複数の領域が、マシン メモリと同じ領域にマッピングされることもあれば (メモリ共有の場合)、ゲスト物理メモリの特定の領域がマシン メモリにマッピングされないこともあります (VMkernel がゲスト物理メモリをスワップアウトするかバレーニングする場合)。そうした状況では、個々の仮想マシンやホストについてのゲスト物理メモリの使用量およびマシン メモリの使用量の計算が異なります。

次図の例を考えてみましょう。これは、1つのホストで実行中の 2 つの仮想マシンを示しています。各ブロックは 4KB のメモリを表しており、それぞれの色または文字は、ブロック上の異なるデータ セットを表しています。

図 6-2. メモリ使用量の例



仮想マシンのパフォーマンス メトリックは、次のようにして判別できます。

- 仮想マシン 1 の与えられたメモリ (マシン メモリにマッピングされたゲスト物理メモリの量) は、マシン メモリへ向かう矢印を持つ、仮想マシン 1 のゲスト物理メモリのブロック数をカウントし、それに 4KB を乗算します。矢印を持つブロックは 5 つあるため、与えられたメモリは 20KB です。
- 使用されたメモリは、仮想マシンに割り当てられたマシン メモリの量で、共有メモリから差し引かれる量を示しています。最初に、仮想マシン 1 のゲスト物理メモリからの矢印を持つ、マシン メモリ内のブロックの数をカウントします。そのようなブロックは 3 つありますが、1 ブロックは仮想マシン 2 と共有されています。したがって、2 つの完全なブロックと、3 番目のブロックの半分をカウントし、それに 4KB を乗算し、合計 10KB が消費されたメモリになります。

これら 2 つのメトリックの重要な違いは、与えられたメモリがゲスト物理メモリ レベルで矢印を持つブロックの数をカウントし、使用されたメモリはマシン メモリ レベルで矢印を持つブロックの数をカウントすることです。ブロックの数はメモリ共有のために 2 つのレベル間で異なり、そのため与えられたメモリと使用されたメモリに違いが出ます。メモリが共有またはその他の解放の技法によって節約されています。

ホストの共有メモリおよび一般共有メモリの判別でも、同様な結果が得られます。

- ホストの共有メモリは、各仮想マシンの共有メモリの合計です。共有メモリを計算するには、各仮想マシンのゲスト物理メモリに注目し、マシン メモリ ブロックを指す矢印を持ち、それ自体も複数の矢印で指されているブロックの数を数えます。例では、そのようなブロックが 6 つあるので、このホストの共有メモリは 24KB です。
- 一般共有メモリは、複数の仮想マシンによって共有されているマシン メモリの量です。一般メモリを判別するには、マシン メモリに注目し、複数の矢印で指されているブロックの数をカウントします。そのようなブロックは 3 つあるので、一般共有メモリは 12KB です。

共有メモリは、ゲスト物理メモリに関するもので、矢印の起点に注目します。一方、一般共有メモリはマシン メモリに関するもので、矢印の宛先に注目します。

ゲスト物理メモリとマシンメモリを測定するメモリメトリックは、矛盾して見える場合があります。これらのメトリックは、実際には仮想マシンのメモリ使用量の異なる局面を測定しています。これらのメトリックの違いを理解することにより、これらのメトリックをより上手に利用して、パフォーマンスの問題を診断できます。

メモリの信頼性

メモリの信頼性は、error insolation と呼ばれ、ESXi により、実際に障害が発生した場合と同様、障害が発生している可能性があるとして判断した場合にメモリの一部の使用を停止できる機能です。

特定のアドレスで、訂正済みエラーの報告が十分集まったら、ESXi はこのアドレスの使用を中止し、訂正済みエラーが未訂正エラーになることを防ぎます。

メモリの信頼性は、RAM 内に訂正済みエラーや未訂正エラーがあっても、VMkernel の信頼性を高めます。また、システムが、エラーを含む可能性のあるメモリページを使用することを防ぎます。

エラー隔離通知の訂正

メモリの信頼性を使用して、VMkernel はエラー隔離通知を受信するページの使用を中止します。

ユーザーが vSphere Client でイベントを受信するのは、VMkernel が修正不可のメモリエラーから復旧したとき、修正可能なエラーが多数あるために VMkernel がシステムメモリのかなりの割合を廃棄するとき、または廃棄できないページが多数あるときです。

手順

- 1 ホストとしての使用を中止します。
- 2 仮想マシンを移行します。
- 3 メモリー関連のハードウェアテストを実行します。

システムスワップについて

システムスワップとは、システム全体の未使用のメモリリソースを有効に活用するために、メモリを再利用するプロセスです。

システムスワップを使用すると、システムは仮想マシン以外のメモリの利用者からメモリを再要求できます。システムスワップを有効にすると、別のプロセスによるメモリの再利用と、メモリを必要としている仮想マシンにメモリを割り当てのバランスを調整できます。システムスワップに必要なハードディスク容量は 1GB です。

メモリからデータを取り出し、バックグラウンドストレージに書き込むことで、メモリを再利用できます。バックグラウンドストレージからデータへアクセスする速度は、メモリからデータへアクセスするより遅いため、スワップしたデータの格納場所を慎重に選ぶ必要があります。

ESXi では、システムスワップの格納先が自動的に決定されます。これが、[優先スワップファイルの場所] です。格納場所の決定に、特定のオプションセットを指定することができます。この場合、システムは、有効なオプションの中から最適なものを選択します。利用可能なオプションがない場合、システムスワップは有効になりません。

次のオプションを使用できます。

- データストア： 指定したデータストアを使用できるようにします。vSAN データストアまたは VMware vSphere® Virtual Volumes™ データストアをシステム スワップ ファイル用に指定することはできません。
- ホスト スワップ キャッシュ： ホスト スワップ キャッシュの一部を使用できるようにします。
- 優先スワップ ファイルの場所： ホストに構成されている優先スワップ ファイルの場所を使用できるようにします。

システム スワップの構成

システム スワップの場所を決定するオプションをカスタマイズできます。

前提条件

[システム スワップの設定] ダイアログ ボックスの [有効] チェック ボックスをオンにします。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] をクリックします。
- 3 [システム] で [システム スワップ] を選択します。
- 4 [編集] をクリックします。
- 5 有効にする各オプションのチェック ボックスを選択します。
- 6 [データストア] オプションを選択するには、ドロップダウン メニューからデータストアを選択します。
- 7 [OK] をクリックします。

永続的なメモリ

7

不揮発性メモリ (NVM) と呼ばれる永続的なメモリ (PMEM) は、電源が停止した後もデータを保持することができます。PMEM は、ダウンタイムの影響を受けやすく高いパフォーマンスが必要なアプリケーションで使用できます。

仮想マシンは、スタンドアロン ホスト上またはクラスタ内で PMEM を使用するように構成できます。PMEM は、ローカル データストアとして扱われます。永続的なメモリは、ストレージの遅延を大幅に軽減します。ESXi では、PMEM で構成されている仮想マシンを作成することができ、これらの仮想マシン内のアプリケーションはこの高速化を利用できます。仮想マシンが最初にパワーオンされたときに、仮想マシンがパワーオンされているかパワーオフされているかに関係なく、PMEM が仮想マシン用に予約されます。この PMEM は、仮想マシンが移行または削除されるまで予約済みのままです。

永続的なメモリは、2 つの異なるモードで仮想マシンによって使用できます。レガシーのゲスト OS では、仮想永続的なメモリ ディスクの機能も利用できます。

■ [仮想永続的なメモリ (vPMem)]

vPMem を使用すると、メモリは仮想 NVDIMM としてゲスト OS に公開されます。これにより、ゲスト OS は PMEM をバイト アドレスが指定可能なランダム モードで使用することができます。

注： 仮想マシン ハードウェア バージョン 14 と NVM テクノロジーをサポートするゲスト OS を使用する必要があります。

注： PMEM 仮想マシン用に vSphere HA を構成する場合は、仮想マシン ハードウェア バージョン 19 を使用する必要があります。詳細については、『[PMEM 仮想マシンに対する vSphere HA の構成](#)』を参照してください。

■ [仮想永続的なメモリ ディスク (vPMemDisk)]

vPMemDisk を使用すると、ゲスト OS は仮想 SCSI デバイスとしてメモリにアクセスできますが、仮想ディスクは PMEM データストアに格納されます。

PMEM で仮想マシンを作成すると、ハード ディスクの作成時にメモリが仮想マシン用に予約されます。アドミッション コントロールもハード ディスクの作成時に実行されます。詳細については、『[vSphere HA アドミッション コントロールの PMEM 予約](#)』を参照してください。

クラスタでは、各仮想マシンに PMEM 用の一部の容量があります。PMEM の合計量は、クラスタで使用可能な合計量以下にする必要があります。PMEM の使用量には、パワーオン状態の仮想マシンとパワーオフ状態の仮想マシンの両方が含まれます。仮想マシンが PMEM を使用するように構成されており、DRS を使用しない場合は、仮想マシンを配置するのに十分な PMEM があるホストを手動で選択する必要があります。

[NVDIMM と従来のストレージ]

NVDIMM はメモリとしてアクセスされます。従来のストレージを使用すると、アプリケーションとストレージ デバイスの間にソフトウェアが配置されるため、処理時間が遅延することがあります。PMEM を使用すると、アプリケーションはストレージを直接使用します。つまり、PMEM のパフォーマンスは従来のストレージよりも優れています。ストレージがホストに対してローカルです。ただし、システム ソフトウェアは変更を追跡できないため、バックアップなどのソリューションは現在 PMEM では動作しません。

非 PMEM データストアへのライトスルー以外のモードで vPMEM を使用する場合、vSphere HA などのソリューションの範囲は制限されます。フェイルオーバーが有効な vPMEM 仮想マシンに対して vSphere HA を有効にすると、仮想マシンを別のホストにフェイルオーバーできます。この場合、仮想マシンは新しいホストの PMEM リソースを使用します。古いホストのリソースを解放するために、ガベージ コレクタはこのようリソースを定期的に特定して解放し、他の仮想マシンが使用できるようにします。

[名前空間]

PMEM の名前空間は、ESXi が起動する前に構成されます。名前空間は、システム上のディスクと似ています。ESXi は名前空間を読み取り、GPT ヘッダーを書き込むことで複数の名前空間を 1 つの論理ボリュームにまとめます。以前にこれを設定していない場合、デフォルトで自動的にフォーマットされます。すでにフォーマットされている場合は、ESXi は PMEM をマウントします。

[GPT テーブル]

PMEM ストレージ内のデータが破損していると、ESXi が動作しなくなることがあります。これを回避するため、ESXi は PMEM のマウント時にメタデータのエラーをチェックします。

[PMEM 領域]

PMEM 領域とは、単一の vNVDimm または vPMemDisk を表す連続したバイト ストリームです。各 PMEM ボリュームは 1 台のホストに属しています。これは、管理者が多数のホストを持つクラスタ内の各ホストを管理する必要がある場合には、管理が難しいことがあります。ただし、個々のデータストアを管理する必要はありません。代わりに、クラスタ内の PMEM 容量全体を 1 つのデータストアと見なすことができます。

vCenter Server および DRS は PMEM データストアの初期配置を自動化します。仮想マシンを作成したとき、またはデバイスを仮想マシンに追加したときに、ローカルの PMEM ストレージ プロファイルを選択します。残りの設定は自動化されています。ESXi では仮想マシン ホームを PMEM データストアに配置することができないという制限があります。これは、仮想マシン ログと統計情報ファイルを保存するのに貴重な容量を必要とするためです。これらの領域は、仮想マシン データを表すために使用され、バイト アドレス指定可能な nvDimms または VpMem ディスクとして公開できます。

[移行]

PMEM はローカル データストアであるため、仮想マシンを移動する場合は Storage vMotion を使用する必要があります。vPMem を持つ仮想マシンは、PMEM リソースを持つ ESX ホストにのみ移行することができます。vPMemDisk を持つ仮想マシンは、PMEM リソースなしで ESX ホストに移行できます。

[エラー処理と NVDimm の管理]

ホスト障害により、ライトスルー モードではない vPMEM 仮想マシンの可用性が失われる可能性があります。致命的なエラーが発生した場合は、すべてのデータが失われる可能性があるため、手動で PMEM を再フォーマットする必要があります。

vSphere Client での vSphere の永続的なメモリ

永続的なメモリの簡単な概要については、以下を参照してください。



(vSphere Client での vSphere の永続的なメモリ)

vSphere Client と PMEM の併用に対する機能向上

PMEM と併用する際の HTML5 ベース vSphere Client の機能向上の概要については、以下を参照してください：



(vSphere Client と PMEM の併用に対する機能向上)

PMEM を使用した vSphere Client での仮想マシンの移行およびクローン作成

PMem を使用する仮想マシンの移行およびクローン作成の詳細については、以下を参照してください：



(PMEM を使用した vSphere Client での仮想マシンの移行およびクローン作成)

この章には、次のトピックが含まれています。

- PMEM 仮想マシンに対する vSphere HA の構成
- vSphere HA アドミッション コントロールの PMEM 予約
- vSphere メモリの監視と修正

PMEM 仮想マシンに対する vSphere HA の構成

vSphere HA を PMEM 仮想マシンに対しライトスルー モードで構成できます。これにより、ホストで障害が発生した場合に仮想マシンを機能している別のホストでリストアできるようになります。

前提条件

- ハードウェア バージョン 19 を選択する必要があります。
- vPMemDisks を持つ PMEM 仮想マシンはサポートされていません。

手順

- 1 [新規仮想マシン] ウィザードで新しい仮想マシンを作成する場合は、[ハードウェアのカスタマイズ] を選択します。

- a [新規デバイスを追加] をクリックし、ドロップダウンメニューから [NVDIMM を追加] を選択します。
- b チェックボックス [すべての NVDIMM デバイスに対し別のホストでのフェイルオーバーを許可] をクリックします。
- c [次へ] をクリックし、[新規仮想マシン] ウィザードを完了します。

ホスト障害が発生した場合、NVDIMM PMEM データはリカバリできません。デフォルトでは、HA でこの仮想マシンは別のホストで再起動されません。ホスト障害が発生した場合に HA で仮想マシンのフェイルオーバーを許可すると、新しい空の NVDIMM が搭載されている別のホストで仮想マシンが再起動されます。

- 2 既存の仮想マシンで HA を有効にするには、仮想マシンを参照します。

- a [仮想マシンのハードウェア] で、[編集] をクリックします。
- b NVDIMM を選択します。
- c チェックボックス [すべての NVDIMM デバイスに対し別のホストでのフェイルオーバーを許可] をクリックします。
- d [OK] をクリックします。

ホスト障害が発生した場合、HA は新しい空の NVDIMM が搭載されている別のホスト上でこの仮想マシンを再起動します。

vSphere HA アドミッション コントロールの PMEM 予約

アドミッション コントロールは vSphere HA が使用するポリシーです。これにより、クラスタ内のフェイルオーバーのキャパシティが確保されます。

許容されるホスト障害の数を増やすと、可用性の制約や予約キャパシティが増えます。ホストのフェイルオーバー キャパシティ用に一定割合の永続的なメモリを予約できます。これはブロックされる実際のストレージ容量であり、ホストのパワーオフを考慮する必要があります。

[クラスタ設定の編集] で [アドミッション コントロール] を選択して、ホストで許容する障害の数を指定できます。

次で定義された CPU/メモリ予約を選択した場合について以下に示します。

- [クラスタ リソースの割合 (%)]。クラスタ内の仮想マシンで永続的なメモリを使用していない場合でも、クラスタ内の永続的なメモリ容量の一部はフェイルオーバー専用となります。この割合は、オーバーライドで指定できるほか、[許容するホスト障害] 設定に従って自動的に計算されます。PMEM アドミッション コントロールを有効にすると、PMEM をディスクとして使用している仮想マシンがある場合でも、PMEM キャパシティがクラスタ全体で予約されます。

- [スロット ポリシー (パワーオン状態の仮想マシン)]。永続的なメモリのアドミッション コントロールでは、永続的なメモリ リソースに対してのみ、クラスタ リソースの割合 (%) ポリシーを使用してスロット ポリシーをオーバーライドします。割合の値は、[クラスタで許容するホスト障害] 設定から自動的に計算され、オーバーライドすることはできません。
- [専用フェイルオーバー ホスト]。専用のフェイルオーバー ホスト上の永続的なメモリはフェイルオーバー専用であり、これらのホスト上に永続的なメモリを持つ仮想マシンをプロビジョニングすることはできません。

注： アドミッション コントロール ポリシーを選択した後、[永続的なメモリのフェイルオーバー キャパシティの予約] チェックボックスをクリックして PMEM アドミッション コントロールを有効にする必要もあります。

vSphere メモリの監視と修正

vMMR では、アプリケーションのワークロードがメモリ モードが原因で回帰されるかどうかを判断できるよう、データが収集され、パフォーマンス統計が表示されます。

Intel Optane パーシステント メモリ は、アプリケーション直接モードまたはメモリ モードで BIOS 設定で構成できます。アプリケーション直接モードでは、パーシステント メモリに DRAM とともにバイト アドレスが指定可能なパーシステント メモリとしてアクセス可能です。メモリ モードでは、DRAM がハードウェア キャッシュになり、より大きな PMEM が揮発性になり、システム メモリとして表示されます。

メモリ モードは、仮想マシンに認識されず、透過的です。メモリ モードでシステムを構成すると、システムは DRAM が搭載された従来のシステムとして表示されます。クラスタには、構成が異なるホストを混在させることができます。vSphere では、メモリ モードになっているシステムに関する追加情報が表示されます。ESXi では、ホスト レベルと仮想マシン レベルの統計に関する情報を収集するパフォーマンス カウンタがプログラムされます。これらのパフォーマンス統計は、アラームの作成に使用されます。統計情報は、パフォーマンス チャートでも追跡されます。

システムがメモリ モードになっているかどうかは、ホストの [サマリ] タブの [メモリ階層化: ハードウェアおよび追加詳細] で確認できます。

The screenshot shows the vSphere Summary tab for a host. The 'Monitor' section displays the following configuration:

Logical Processors:	96
NICs:	3
Virtual Machines:	1
Memory Tiering:	Hardware
State:	Connected
Uptime:	6 hours

A notification box is overlaid on the 'Memory Tiering' row, stating: "Intel Optane™ Persistent Memory configured in Memory Mode." The notification has a close button (X) in the top right corner.


DRAM と PMEM のサイズを [構成] - [ハードウェア] - [概要] - [メモリ] で表示することもできます。

Summary	Monitor	Configure	Permissions	VMs	Datastores	Networks	Updates												
System Resource Reservation		Memory																	
Firewall		<table border="1"> <tr> <td>Total</td> <td>503.68 GB</td> </tr> <tr> <td>System</td> <td>385.17 MB</td> </tr> <tr> <td>Virtual machines</td> <td>503.3 GB</td> </tr> <tr> <td>Memory Tiering</td> <td>Hardware ⓘ</td> </tr> <tr> <td>Tier 0</td> <td>256 GB DRAM (Cache)</td> </tr> <tr> <td>Tier 1</td> <td>503.67 GB PMem (Memory)</td> </tr> </table>						Total	503.68 GB	System	385.17 MB	Virtual machines	503.3 GB	Memory Tiering	Hardware ⓘ	Tier 0	256 GB DRAM (Cache)	Tier 1	503.67 GB PMem (Memory)
Total	503.68 GB																		
System	385.17 MB																		
Virtual machines	503.3 GB																		
Memory Tiering	Hardware ⓘ																		
Tier 0	256 GB DRAM (Cache)																		
Tier 1	503.67 GB PMem (Memory)																		
Services																			
Security Profile																			
System Swap																			
Packages																			
Hardware ▼																			
Overview																			
Graphics																			

ESXi では、次の 2 種類のメモリ統計情報が収集および公開されます。

- ホスト レベルの統計情報：メモリのサブコンポーネントでは、パフォーマンスカウンタのプログラムによって DRAM と PMEM のパフォーマンスが測定されます。ホスト レベルの統計情報は、さまざまなメモリ タイプ (DRAM、PMEM) の合計、読み取り/書き込みのバンド幅、読み取り/書き込み遅延、およびミス率です。
- 仮想マシン レベルの統計情報：vSphere では、パフォーマンスカウンタが監視され、仮想マシンの DRAM および PMEM 読み取りバンド幅に関するデータが取得されます。

ホストと仮想マシンの両方には、[パフォーマンス チャート] の下に新しい [メモリ] ペインがあります。ここに、メモリ使用率やメモリの解放などのメモリ詳細が新しい統計情報とともに表示されます。ESXi ホスト レベルでは、メモリのバンド幅とメモリのミス率のチャートを監視できます。仮想マシン レベルでは、PMEM と DRAM の読み取りバンド幅を表示できます。

ESXi ホストの [仮想マシン] タブで、ホストに配置されているすべての仮想マシンに関するパフォーマンス情報を含むリストを表示できます。仮想マシンに対するメモリモードの影響に関する情報を表示するには、列表示  アイコンをクリックして、アクティブなメモリ、DRAM 読み取りバンド幅、および PMEM 読み取りバンド幅のメトリックを選択します。

事前構成済みのデフォルト アラームは 2 つあります。ホスト レベル (メモリモードのホストのアクティブな DRAM の使用率が高い) と仮想マシン レベル (仮想マシンの PMEM バンド幅の使用率が高い) のアラームです。アラーム条件が満たされると、イベントが発行され、対応するアラームがトリガされます。パフォーマンスメトリックに基づき、カスタムアラームを作成することもできます。vMMR アラームは、メモリモードで構成されたホストでのみ機能します。

DRS がクラスタで有効で、完全自動化されている場合、ホストのアクティブなメモリの使用率が DRAM キャッシュのサイズの一定割合を超えると、負荷をより適切に分散させるために、DRS で一部の仮想マシンがホスト外に移動されることがあります。

詳細については、『vSphere の監視とパフォーマンス』を参照してください。

注： vMMR は、Intel Broadwell、Skylake、Cascade Lake、Ice Lake プラットフォームでサポートされています。ホストレベルの DRAM 統計情報は、これらのプラットフォームで使用できます。ホストレベルおよび仮想マシンレベルの PMEM 統計情報は、メモリモードで構成された Cascade Lake ホストおよび Ice Lake ホストでのみ使用できます。

仮想グラフィックの構成

8

実装されたサポート対象のグラフィックのグラフィック設定を編集できます。

vSphere では、複数のグラフィックの実装がサポートされます。

- VMware では、AMD、Intel、および NVIDIA の 3D グラフィック ソリューションがサポートされます。
- NVIDIA GRID のサポート。
- 単一の NVIDIA vib で、vSGA と vGPU の両方の実装をサポートできるようにします。
- vCenter Server に Intel と NVIDIA の GPU パフォーマンス チャートを提供します。
- Horizon View VDI デスクトップのグラフィックを有効にします。

ホストのグラフィックを設定し、仮想マシンごとに vGPU グラフィック設定をカスタマイズできます。

注： この章では、「メモリ」は物理 RAM を指します。

この章には、次のトピックが含まれています。

- GPU 統計の表示
- 仮想マシンへの NVIDIA GRID vGPU の追加
- ホスト グラフィックの設定
- グラフィック デバイスの設定

GPU 統計の表示

ホストのグラフィック カードの詳細情報を表示できます。

GPU の温度、使用率、およびメモリ使用量を確認できます。

注： これらの統計は、GPU ドライバがホストにインストールされている場合にのみ表示されます。

手順

- 1 vSphere Client で、ホストに移動します。
- 2 [監視] > [パフォーマンス] の順にクリックします。
- 3 [詳細] をクリックし、ドロップダウン メニューから [GPU] を選択します。

仮想マシンへの NVIDIA GRID vGPU の追加

ESXi ホストに NVIDIA GRID GPU グラフィック デバイスが搭載されている場合は、NVIDIA GRID の仮想 GPU (vGPU) テクノロジーを使用するように仮想マシンを設定できます。

NVIDIA GRID GPU グラフィック デバイスは、複雑なグラフィック操作を最適化し、CPU に過大な負荷をかけずに高パフォーマンスで動作するように設計されています。

前提条件

- NVIDIA GRID GPU グラフィック デバイスおよび適切なドライバがホストにインストールされていることを確認します。『vSphere のアップグレード』ドキュメントを参照してください。
- 仮想マシンが ESXi 6.0 以降と互換性があることを確認します。

手順

- 1 仮想マシンを右クリックし、[設定の編集] を選択します。
- 2 [仮想ハードウェア] タブで、[新規デバイスを追加] を選択し、ドロップダウン メニューから [新しい PCI デバイス] を選択します。
- 3 [新しい PCI デバイス] を展開し、仮想マシンに接続する NVIDIA GRID vGPU パススルー デバイスを選択します。

注： 全面的なメモリ予約が自動的に適用されます。これは、PCI デバイスに必要です。

- 4 GPU プロファイルを選択します。
GPU プロファイルで、vGPU のタイプが指定されます。
- 5 [OK] をクリックします。

結果

仮想マシンがデバイスにアクセスできるようになります。

ホスト グラフィックの設定

グラフィック オプションは、ホストごとにカスタマイズできます。

前提条件

仮想マシンの電源はオフにする必要があります。

手順

- 1 ホストを選択し、[設定] - [グラフィック] を選択します。
- 2 [ホスト グラフィック] の [編集] を選択します。

- 3 [ホスト グラフィック設定の編集] ウィンドウで、次を選択します。

オプション	説明
共有	VMware の共有仮想グラフィック
直接共有	ベンダーの共有パススルー グラフィック

- 4 共有パススルー GPU の割り当てポリシーを選択します。
- a 複数の GPU にわたって仮想マシンを分散 (ベスト パフォーマンス)
 - b GPU の上限まで仮想マシンをグループ化 (GPU 統合)
- 5 [OK] をクリックします。

次のステップ

[OK] をクリックした後は、ホストで Xorg を再起動する必要があります。

グラフィック デバイスの設定

ビデオ カードのグラフィックの種類を編集できます。

前提条件

仮想マシンの電源はオフにする必要があります。

手順

- 1 [グラフィック デバイス] でグラフィック カードを選択し、[編集] をクリックします。
 - a VMware 共有仮想グラフィックには [共有] を選択します。
 - b ベンダーの共有パススルー グラフィックには [直接共有] を選択します。
- 2 [OK] をクリックします。

結果

デバイスを選択すると、そのデバイスを使用しているアクティブな仮想マシンが表示されます。

次のステップ

[OK] をクリックした後は、ホストで Xorg を再起動する必要があります。

ストレージ I/O リソースの管理

9

vSphere Storage I/O Control によって、クラスタ全体でストレージ I/O の優先順位付けが可能となり、ワークロードの統合が強化されるため、過剰なプロビジョニングに伴うコストの削減につながります。

Storage I/O Control では、ストレージ I/O リソースを処理するためのシェアと制限の構造を拡張できます。I/O が輻輳状態の間に仮想マシンに割り当てるストレージ I/O の量を制御できます。こうすると重要度の高い仮想マシンに対し、重要度の低い仮想マシンよりも多くの I/O リソースを確実に割り当てることができます。

データストアで Storage I/O Control を有効にすると、ESXi は、そのデータストアとの通信時にホストが監視するデバイスの待ち時間の監視を始めます。デバイスの待ち時間がしきい値を上回ると、そのデータストアは輻輳状態であると見なされ、そのデータストアにアクセスする各仮想マシンに、それぞれのシェアに応じた I/O リソースが割り当てられます。仮想マシンごとにシェアを設定します。必要に基づいて、それぞれの数を調整できます。

I/O フィルタ フレームワークである Sphere API for I/O Filtering (VAIO) により、VMware およびそのパートナーは各 VMDK の I/O を遮断し、VMDK の詳細レベルに必要な機能を提供するフィルタを開発できます。VAIO はストレージ ポリシー ベース管理 (SPBM) とともに動作し、VMDK に関連付けられているストレージ ポリシーを介してフィルタの優先順位を設定できます。

Storage I/O Control は、2 つの手順で設定します。

- 1 データストアに対して Storage I/O Control を有効にします。
- 2 各仮想マシンに許可される、ストレージ I/O シェアの数と 1 秒あたりの I/O 処理数 (IOPS) の上限を設定します。

デフォルトでは、すべての仮想マシンのシェアは標準 (1000) に、IOPS は制限なしに設定されます。

注： ストレージ DRS 対応のデータストア クラスタでは、Storage I/O Control がデフォルトで有効となっています。

注： この章では、「メモリ」は物理 RAM を指します。

この章には、次のトピックが含まれています。

- 仮想マシンのストレージ ポリシーについて
- I/O フィルタについて
- ストレージ I/O コントロールの要件
- ストレージ I/O コントロールのリソースのシェアおよび制限
- Storage I/O Control のリソースのシェアおよび制限の設定

- Storage I/O Control を有効にする
- ストレージ I/O コントロールのしきい値の設定
- Storage DRS とストレージ プロファイルの統合

仮想マシンのストレージ ポリシーについて

仮想マシンのストレージ ポリシーは、仮想マシンのプロビジョニングに不可欠です。ポリシーは、仮想マシンに提供するストレージの種類、仮想マシンをストレージ内に配置する方法、および仮想マシンに提供するデータ サービスを制御します。

vSphere は、デフォルトのストレージ ポリシーを提供しています。ただし、新しいポリシーを定義して割り当てることもできます。

ストレージ ポリシーを作成するには、仮想マシン ストレージ ポリシー インターフェイスを使用します。ポリシーを定義する際は、仮想マシンで実行するアプリケーション用に、さまざまなストレージ要件を指定します。ストレージ ポリシーは、キャッシュ、レプリケーションなどの仮想ディスクの特定のデータ サービスを要求するために使用することもできます。

仮想マシンの作成、クローン作成、または移行の際に、ストレージ ポリシーを適用します。ストレージ ポリシーを適用すると、ストレージ ポリシー ベースの管理 (SPBM) メカニズムは、仮想マシンを対応するデータストアに配置します。また、特定のストレージ環境では、必要なサービス レベルを保証するため、仮想マシンのストレージ オブジェクトをストレージ リソース内にどのようにプロビジョニングして割り当てるかを決定します。また、SPBM は要求された仮想マシンのデータ サービスを有効にします。vCenter Server は、ポリシーのコンプライアンスを監視し、割り当てられているストレージ ポリシーに仮想マシンが違反するとアラートを送信します。

詳細については、『vSphere のストレージ』を参照してください。

I/O フィルタについて

基盤となるストレージ トポロジに関わらず、仮想ディスクに関連付けられている I/O フィルタは、仮想マシンの I/O バスに直接アクセスできます。

VMware は、I/O フィルタの特定のカテゴリを提供します。さらに、サードパーティ ベンダーは、I/O フィルタを作成することができます。通常、I/O フィルタはパッケージとして配布され、インストーラを使用して、vCenter Server と ESXi ホスト クラスタにフィルタ コンポーネントをデプロイします。

I/O フィルタを ESXi クラスタにデプロイすると、vCenter Server は VASA プロバイダとも呼ばれる I/O フィルタ ストレージ プロバイダを、クラスタ内の各ホストに自動的に構成して登録します。ストレージ プロバイダは、vCenter Server と通信し、I/O フィルタによって提供されるデータ サービスを仮想マシン ストレージ ポリシー インターフェイスに表示できるようにします。これらのデータ サービスは、仮想マシン ポリシーの共通ルールを作成する際に参照できます。仮想ディスクをこのポリシーに関連付けると、仮想ディスクで I/O フィルタが有効になります。

詳細については、『vSphere のストレージ』を参照してください。

ストレージ I/O コントロールの要件

ストレージ I/O コントロールには、要件と制限事項がいくつかあります。

- ストレージ I/O コントロールが有効になっているデータストアは、単一の vCenter Server システムで管理される必要があります。
- ストレージ I/O コントロールは、ファイバ チャンネルに接続されたストレージ、iSCSI に接続されたストレージ、および NFS に接続されたストレージでサポートされます。Raw デバイス マッピング (RDM) はサポートされていません。
- ストレージ I/O コントロールは、複数のエクステントを持つデータストアはサポートしません。
- ストレージの自動階層化機能を持つアレイにバックアップされているデータストアでストレージ I/O コントロールを使用する前に、『VMware ストレージ/SAN 互換性ガイド』を参照して、自動的に階層化されたストレージアレイがストレージ I/O コントロールと互換性があると確認されているかどうかを調べてください。

ストレージの自動階層化機能とは、ユーザーが設定したポリシーおよび現在の I/O パターンに基づいて、LUN/ボリューム、または LUN/ボリュームの一部を別のタイプのストレージメディア (SSD、FC、SAS、SATA) に移行する、アレイ (またはアレイ グループ) の機能です。これらの自動的な移行/階層化機能を持たないアレイについては、特別な証明は不要です。これには、異なるタイプのストレージメディア間で手動でデータを移行する機能を提供するアレイも含まれます。

ストレージ I/O コントロールのリソースのシェアおよび制限

各仮想マシンに許可されるストレージ I/O シェアの数と、1 秒あたりの I/O 処理数 (IOPS) の上限を割り当てます。あるデータストアでストレージ I/O の輻輳状態が検出されると、そのデータストアにアクセスする仮想マシンの I/O ワークロードは、各仮想マシンが持つシェアの比率に基づいて調整されます。

ストレージ I/O シェアは、メモリおよび CPU リソースの割り当てに使用されるシェアと似ています。これについては、[リソース割り当てのシェア](#) に記述されています。これらのシェアは、ストレージ I/O リソースの配分に関する仮想マシンの相対的な重要性を表します。リソースの競合がある場合、シェア値が高い仮想マシンにストレージアレイへのアクセスが優先的に割り当てられます。ストレージ I/O リソースを割り当てるときに、仮想マシンに許可する IOPS を制限できます。デフォルトでは、IOPS は制限なしです。

リソースの制限を設定することのメリットとデメリットについては、[リソース割り当て制限](#) に記述してあります。仮想マシンに設定する制限を IOPS ではなく 1 秒あたりのメガバイト (Mbps) で考えている場合は、その仮想マシンの一般的な I/O サイズに基づいて Mbps を IOPS に変換できます。たとえば、64 KB I/O のバックアップアプリケーションを 10Mbps に制限する場合は、160 IOPS に設定します。

ストレージ I/O コントロールのシェアおよび制限の表示

データストアで実行しているすべての仮想マシンのシェアと制限を表示できます。この情報を表示すると、実行しているクラスタに関係なく、そのデータストアにアクセスしているすべての仮想マシンの設定を比較できます。

手順

- 1 vSphere Client で、データストアを参照して移動します。

- 2 [仮想マシン] タブをクリックします。

このタブには、そのデータストアで実行している各仮想マシンとそれに関連するシェアの値、およびデータストアシェアのパーセンテージが表示されます。

ストレージ I/O コントロールのシェアの監視

データストアにアクセスしている仮想マシンの I/O ワークロードを、Storage I/O Control がそれらのシェアに基づいてどのように処理しているかを監視するには、データストアの [パフォーマンス] タブを使用します。

データストアのパフォーマンス チャートでは、次の情報を監視できます。

- データストアの平均遅延および IOPS の合計
- ホスト間の遅延
- ホスト間のキューの深さ
- ホスト間の読み取り/書き込み IOPS
- 仮想マシンのディスク間の読み取り/書き込み遅延
- 仮想マシンのディスク間の読み取り/書き込み IOPS

手順

- 1 vSphere Client で、データストアを参照して移動します。
- 2 [監視] タブの [パフォーマンス] をクリックします。
- 3 [詳細] を選択します。

Storage I/O Control のリソースのシェアおよび制限の設定

仮想マシンに相対的なシェアの量を割り当てることで、重要度に基づいて仮想マシンにストレージ I/O リソースを割り当てることができます。

仮想マシンのワークロードが非常に似ている場合を除き、シェアは、1 秒あたりの I/O 処理数や MB 数で割り当てを指定するとはかぎりません。高いシェアの仮想マシンは、低いシェアの仮想マシンと比べて、ストレージ デバイスやデータストアでより多くの同時 I/O 処理を保留状態に保つことができます。ワークロードに応じて、2 台の仮想マシンのスループットがそれぞれ異なることがあります。

前提条件

仮想マシン ストレージ ポリシーの作成および仮想マシン ストレージ ポリシーの共通ルールの定義については、『vSphere のストレージ』を参照してください。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
 - a データセンター、フォルダ、クラスター、リソース プール、またはホストを選択して、仮想マシンを検索します。
 - b [仮想マシン] タブをクリックします。

- 2 仮想マシンを右クリックし、[設定の編集] をクリックします。
- 3 [仮想ハードウェア] タブをクリックし、リストから仮想ハード ディスクを選択します。[ハード ディスク] を展開します。
- 4 ドロップダウン メニューから [仮想マシン ストレージ ポリシー] を選択します。
ストレージ ポリシーを選択する場合は、手動で [シェア] と [制限 - IOPS] を構成しないでください。
- 5 [シェア] のドロップダウン メニューをクリックし、仮想マシンに割り当てるシェアの量 (低、標準、または高) を選択します。
[カスタム] を選択するとユーザー定義のシェア値を入力できます。
- 6 [制限 - IOPS] のドロップダウン メニューをクリックし、仮想マシンに割り当てるストレージ リソースの上限を入力します。
IOPS は、1 秒あたりの I/O 処理数です。デフォルトでは、IOPS は制限なしです。低 (500)、標準 (1000)、または高 (2000) を選択します。[カスタム] を選択してユーザーが定義するシェアの値を入力することもできます。
- 7 [OK] をクリックします。

Storage I/O Control を有効にする

Storage I/O Control を有効にすると、ESXi はデータストアの遅延を監視し、平均遅延がしきい値を超えると、I/O 負荷を調整します。

手順

- 1 vSphere Client で、データストアを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [設定] をクリックし、[全般] をクリックします。
- 4 [データストア機能] の [編集] をクリックします。
- 5 [Storage I/O Control を有効にする] チェック ボックスをオンにします。
- 6 [OK] をクリックします。

結果

[データストア機能] の Storage I/O Control がデータストアで有効になります。

ストレージ I/O コントロールのしきい値の設定

データストアの輻輳しきい値とは、データストアに許容された遅延の上限で、この上限に達すると、Storage I/O Control はシェア値に応じて仮想マシンのワークロードに重み付けをし始めます。

ほとんどの環境では、しきい値の設定を調整する必要はありません。

注意: 2つの異なるデータストアで同じスピンドルを共有する場合、Storage I/O Control は正常に機能しないことがあります。

輻輳しきい値の設定を変更する場合は、次の考慮事項に従って値を設定します。

- 通常、値が高いほど、スループットの合計が高くなり、隔離レベルが低くなります。全体の平均遅延がしきい値より高い場合を除いて、スロットルは発生しません。
- スループットの方が遅延より重要な場合、値を低くし過ぎないでください。たとえば、ファイバチャネルディスクの場合、20 ミリ秒より低い値にすると、最大のディスクスループットが小さくなる可能性があります。50 ミリ秒を超える非常に高い値を設定すると、スループット全体で大きなメリットは得られず、遅延が極めて長くなる可能性があります。
- 値が低いほどデバイスの遅延は短くなり、仮想マシンの I/O のパフォーマンスの隔離レベルが高くなります。隔離レベルが高いということは、より頻繁にシェアコントロールが実行されることを意味します。デバイスの遅延が短くなると、最も高いシェア値を持つ仮想マシンの I/O の遅延が短くなります。一方、シェアの少ない仮想マシンでは I/O の遅延は長くなります。
- 非常に低い値（20 ミリ秒よりも低い）を設定すると、デバイスの遅延は短くなり、I/O 間の隔離レベルは向上しますが、データストアのスループットの合計が減少する可能性があります。
- 極端に高い値または極端に低い値に設定すると、隔離レベルが低下します。

前提条件

Storage I/O Control が有効になっていることを確認します。

手順

- 1 vSphere Client で、データストアを参照して移動します。
- 2 [構成] タブをクリックし、[設定] をクリックします。
- 3 [全般] をクリックします。
- 4 [データストア機能] の [編集] をクリックします。
- 5 [Storage I/O Control を有効にする] チェックボックスをオンにします。

Storage I/O Control は、データストアが最大スループットの 90% で稼働した場合の遅延を推定して、それに対応するしきい値を自動的に設定します。

- 6 (オプション) [輻輳しきい値] を調整します。

- ◆ [最大スループットのパーセント] ドロップダウンメニューから値を選択します。

最大スループット値の比率は、データストアを推定最大スループットの比率で稼働する際に推定される遅延のしきい値を意味します。

- ◆ [手動] ドロップダウンメニューから値を選択します。

有効値は 5 ミリ秒～100 ミリ秒です。不適切な輻輳しきい値を設定すると、データストア上の仮想マシンのパフォーマンスが低下します。

- 7 (オプション) [デフォルトに戻す] をクリックして、輻輳しきい値をデフォルト値 (30 ミリ秒) に戻します。

- 8 [OK] をクリックします。

Storage DRS とストレージ プロファイルの統合

ストレージ ポリシー ベース管理 (SPBM) を使用すると、Storage DRS によって適用される仮想マシン ポリシーを指定できます。データストア クラスタには、異なる機能プロファイルが指定された一連のデータストアを使用できます。仮想マシンにストレージ プロファイルが関連付けられている場合、Storage DRS は基盤となるデータストアの機能に基づいて配置を適用できます。 []

Storage DRS のストレージ プロファイルとの統合の一環として、Storage DRS クラスタ レベルの詳細オプション `EnforceStorageProfiles` が導入されています。詳細オプション `EnforceStorageProfiles` には、0、1、または 2 のいずれかの整数値を設定します。デフォルト値は 0 です。このオプションを 0 に設定すると、Storage DRS クラスタにストレージ プロファイルまたはポリシーが適用されません。このオプションを 1 に設定すると、Storage DRS クラスタにストレージ プロファイルまたはポリシーがソフト適用されます。これは、DRS ソフト ルールに似ています。Storage DRS は、最適なレベルでストレージ プロファイルまたはポリシーに従いません。Storage DRS は、必要な場合、ストレージ プロファイルに従いません。Storage DRS アフィニティ ルールは、ストレージ プロファイルの適用が 1 に設定されている場合のみ、ストレージ プロファイルより優先されます。このオプションを 2 に設定すると、Storage DRS クラスタにストレージ プロファイルまたはポリシーがハード適用されます。これは、DRS ハード ルールに似ています。Storage DRS は、ストレージ プロファイルまたはポリシーに違反しません。ストレージ プロファイルは、アフィニティ ルールより優先されます。Storage DRS は、非アフィニティ ルール違反を解決できませんでしたというエラーを生成します。

前提条件

デフォルトでは、Storage DRS は仮想マシンに関連付けられているストレージ ポリシーを適用しません。要件に応じて `EnforceStorageProfiles` オプションを設定してください。オプションは、デフォルト (0)、ソフト (1)、ハード (2) です。

手順

- 1 vSphere Client に管理者としてログインします。
- 2 vSphere Client で、Storage DRS クラスタをクリックし、[管理] - [設定] - [Storage DRS] の順に選択します。
- 3 [編集] - [詳細オプション] - [構成パラメータ] をクリックし、[追加] の順に選択します。
- 4 [オプション] 見出しの下のフィールドをクリックし、「`EnforceStorageProfiles`」と入力します。
- 5 前に入力した詳細オプション名の右の [値] 見出しの下のフィールドをクリックし、0、1、または 2 のいずれかの値を入力します。
- 6 [OK] をクリックします。

リソース プールの管理

10

リソース プールは、リソースを柔軟に管理するための論理上の抽象概念です。リソース プールは階層にグループ化することができ、使用可能な CPU リソースとメモリ リソースを階層状にパーティショニングするために使用できません。

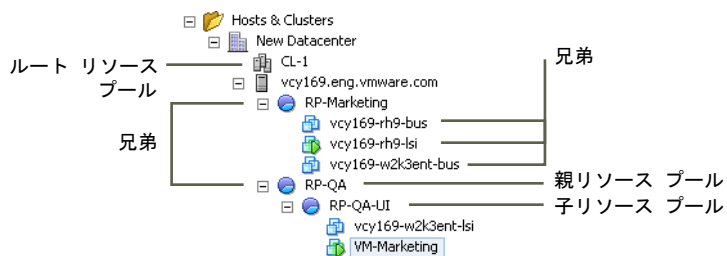
各スタンドアロン ホストと各 DRS クラスタには、(非表示の) ルート リソース プールがあり、そのホストまたはクラスタのリソースがグループ分けされています。ホスト (またはクラスタ) のリソースとルート リソース プールのリソースは常に同じであるため、ルート リソース プールは表示されません。

ユーザーは、ルート リソース プールの子リソース プール、またはユーザーが作成した任意の子リソース プールの子リソース プールを作成できます。それぞれの子リソース プールは、親のリソースの一部を保持し、これによって、連続した小さな単位のコンピュータ機能を表す子リソース プールの階層構造が形成されます。

リソース プールには、子リソース プール、仮想マシン、またはその両方を格納できます。共有リソースの階層を形成できます。より上位にあるリソース プールを親リソース プールと呼びます。同じレベルにあるリソース プールと仮想マシンを兄弟と呼びます。クラスタ自体は、ルート リソース プールを表します。子リソース プールを作成しない場合は、ルート リソース プールだけが存在します。

次の例の RP-QA は、RP-QA-UI の親リソース プールです。RP- マーケティングと RP-QA は兄弟です。RP- マーケティングのすぐ下にある 3 台の仮想マシンも兄弟です。

図 10-1. リソース プール階層での親、子、および兄弟



各リソース プールに対し、予約、制限、シェア、予約拡張の可否を指定します。リソース プールのリソースは、その子リソース プールと仮想マシンからも使用できます。

注： この章では、「メモリ」は物理 RAM を指します。

この章には、次のトピックが含まれています。

- リソース プールを使用する理由
- リソース プールの作成

- リソース プールの編集
- リソース プールへの仮想マシンの追加
- リソース プールからの仮想マシンの削除
- リソース プールの削除
- リソース プールのアドミSSION コントロール

リソース プールを使用する理由

リソース プールを使用すると、ホスト（またはクラスタ）のリソース管理を委任できます。リソース プールを使用して、クラスタ内のすべてのリソースを区分化する場合にメリットがあります。ホストまたはクラスタの直接の子として、複数のリソース プールを作成し、構成できます。その後、リソース プールの管理をほかの個人または組織に委任できます。

リソース プールを使用すると、次のメリットを得られる場合があります。

- 柔軟な階層型編成：必要に応じて、リソース プールを追加、削除、再編成し、リソース割り当てを変更します。
- プール間での分離とプール内での共有：最上位レベルの管理者は、リソース プールを部門レベルの管理者が使用できるようにすることができます。1つの部門のリソース プールの内部的な割り当ての変更が、ほかの無関係なリソース プールに不当な影響を及ぼすことはありません。
- アクセス コントロールと委任：最上位レベルの管理者が、リソース プールを部門レベルの管理者が使用できるようにした場合、部門レベルの管理者は、現在の設定（シェア、予約、制限）によってリソース プールに割り当てられるリソースの限度内で、すべての仮想マシンの作成と管理を実行できます。委任は、通常、権限の設定と結び付けて行われます。
- リソースとハードウェアとの分離：DRS が有効に設定されているクラスタを使用している場合、すべてのホストのリソースは、常にそのクラスタに割り当てられます。これは、リソースを提供する実際のホストと無関係に管理者がリソース管理を実行できることを意味します。たとえば 3 台の 2GB ホストを 2 台の 3GB ホストに置き換える場合、リソース割り当てを変更する必要はありません。

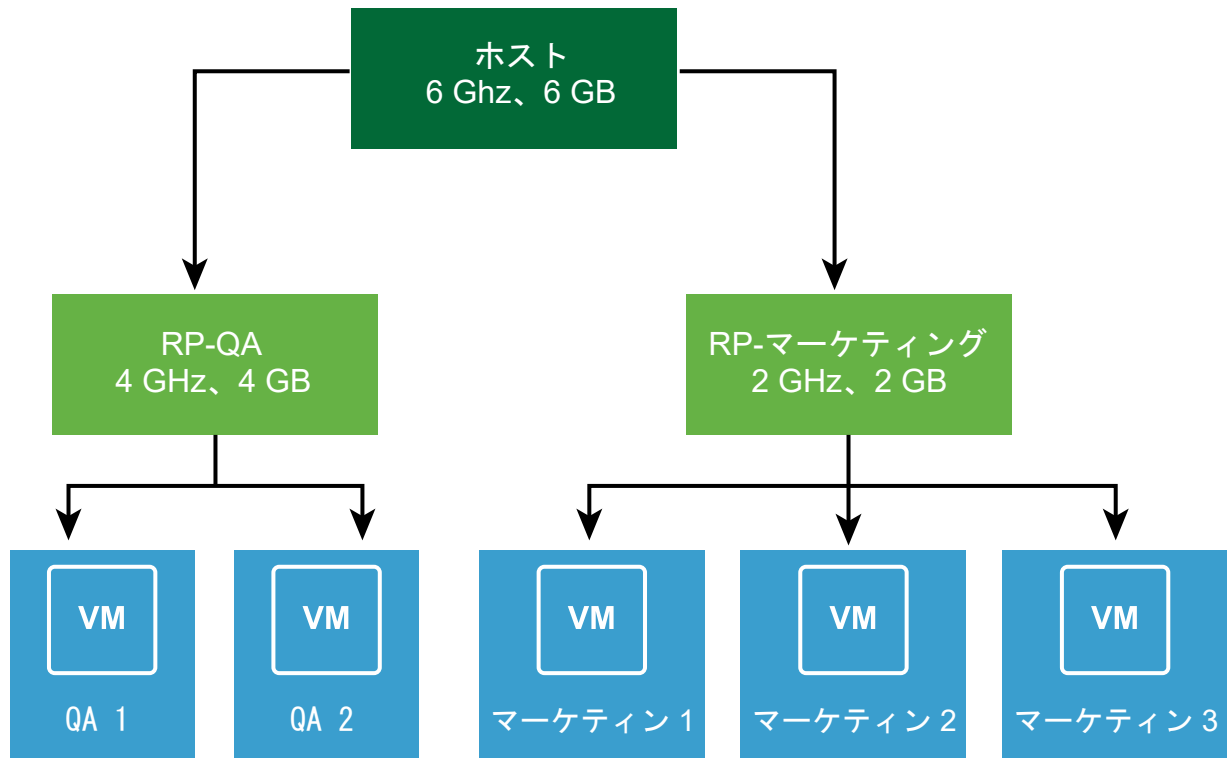
このように切り離して考えることで、管理者は個々のホストというより、コンピューティング能力全体に注意を払えるようになります。

- 多層サービスを実行する一連の仮想マシンの管理：リソース プール内で多層サービス用に仮想マシンをグループ化します。それぞれの仮想マシン上でリソースを設定する必要はありません。代わりに、それらの仮想マシンを含むリソース プール上で設定を変更することで、仮想マシンのセットに対するリソースの全体的な割り当てを制御できます。

たとえば、ホストにいくつかの仮想マシンがあるとします。マーケティング部門は 3 台の仮想マシンを使用し、QA 部門は 2 台の仮想マシンを使用します。QA 部門は、より多くの CPU とメモリを必要とするため、管理者は、グループごとに 1 つのリソース プールを作成します。QA 部門のユーザーは自動化されたテストを実行するため、管理者は QA 部門のプールの [CPU シェア] を [高] に設定し、マーケティング部門のプールを [標準] に設定します。CPU とメモリのリソースがより少ない 2 番目のリソース プールでも、マーケティングスタッフの負荷は軽いので十分です。QA 部門が割り当て分のすべてを使用していない場合、マーケティング部門はいつでも使用可能リソースを使用できます。

次の図の数値は、リソース プールに対する実際の割り当てを示しています。

図 10-2. リソース プールへのリソースの割り当て



リソース プールの作成

任意の ESXi ホスト、リソース プール、または DRS クラスタの子リソース プールを作成できます。

注： ホストがクラスタに追加された場合、そのホストの子リソース プールは作成できません。クラスタが DRS に対して有効な場合、クラスタの子リソース プールを作成できます。

子リソース プールを作成する際、リソース プールの属性情報を入力するよう求められます。システムは、アドミッション コントロールを使用して、使用できないリソースが割り当てられないようにします。仮想マシンを追加または削除するときにシェアが動的にスケールされるようにするには、スケラブル シェアを選択します。

注： シェアは親レベルでスケールされます。スケラブル シェアを持つ親から作成されたすべての子孫リソース プールは、デフォルトでスケラブル シェアを持ちます。

前提条件

vSphere Client は、vCenter Server システムに接続されます。

手順

- 1 vSphere Client で、リソース プールの親オブジェクト（ホスト、別のリソース プール、または DRS クラスタ）を選択します。
- 2 オブジェクトを右クリックし、[新規リソース プール] を選択します。
- 3 リソース プールを識別するための名前を入力します。
- 4 スケラブル シェアを有効にする場合は、チェックボックスを選択します。
- 5 CPU リソースおよびメモリ リソースの割り当て方法を指定します。

リソース プールの CPU リソースは、ホストがリソース プール用に予約する、保証された物理リソースです。通常、デフォルト値を受け入れ、ホストにリソース割り当てを任せます。

オプション	説明
シェア	<p>親の合計リソースに対するこのリソース プールのシェアを指定します。兄弟のリソース プールは、予約と制限の範囲内で、相対的シェア値に従ってリソースを共有します。</p> <ul style="list-style-type: none"> ■ [低]、[標準]、または [高] を選択して、それぞれ 1:2:4 の比率でシェア値を指定します。 ■ 各仮想マシンに対して、比重に見合う特定のシェア値を指定するには、[カスタム] を選択します。
予約	<p>リソース プールで確保されている CPU またはメモリの割り当てを指定します。デフォルトは 0 です。</p> <p>0 以外の予約は、親（ホストまたはリソース プール）の未予約のリソースから差し引かれます。リソースは、仮想マシンがリソース プールに関連付けられているかどうかに関係なく、予約済みとみなされます。</p>

オプション	説明
拡張可能な予約	このチェック ボックスを選択すると (デフォルト)、アドミSSION コントロール中に拡張可能な予約が考慮されます。 このリソース プール内の仮想マシンをパワーオンすると仮想マシンの予約の合計がリソース プールの予約よりも大きくなる場合、リソース プールは親または先祖のリソースを使用できません。
制限	リソース プールに割り当てる CPU またはメモリの上限を指定します。通常はデフォルト値 ([制限なし]) を使用します。 制限を指定するには、[制限なし] チェック ボックスを選択解除します。

6 [OK] をクリックします。

結果

リソース プールを作成したら、仮想マシンをリソース プールに追加できます。仮想マシンのシェアは、同じ親リソース プールを持つほかの仮想マシン (またはリソース プール) に対して相対的です。

例：リソース プールの作成

マーケティング部門と QA 部門の間で共有する必要がある、6GHz の CPU と 3GB のメモリを備えたホストがあると仮定します。また、リソースを均一に共有するのではなく、一方の部門 (QA) に高い優先順位を与えたいものとします。これは、各部門でリソース プールを作成し、[シェア] 属性を使用してリソース割り当ての優先順位を設定することで実現できます。

この例では、ESXi ホストを親リソースとするリソース プールの作成方法を説明します。

- 1 [新規リソース プール] ダイアログ ボックスで、QA 部門のリソース プールの名前を入力します (例: RP-QA)。
- 2 RP-QA の CPU リソースとメモリ リソースについて、[シェア] を [高] に指定します。
- 3 2 番目のリソース プールの RP- マーケティングを作成します。

CPU とメモリのシェアは、[標準] のままにします。

4 [OK] をクリックします。

リソースの競合がある場合、RP-QA は 4GHz および 2GB のメモリを受け取り、RP- マーケティングは 2GHz および 1GB を受け取ります。競合がない場合は、どちらも前述の値を上回る割り当てを受け取ります。その後、これらのリソースはそれぞれのリソース プール内の仮想マシンで使用できます。

リソース プールの編集

リソース プールの作成後、CPU およびメモリのリソース設定を編集できます。

手順

- 1 vSphere Client で、リソース プールを参照して移動します。
- 2 [アクション] ドロップダウン メニューから [リソース設定の編集] を選択します。

3 (オプション) **リソース プールの作成**で説明されているように、選択したリソース プールの属性はすべて変更できます。

- スケーラブル シェアを有効にする場合は、チェックボックスを選択します。

注： シェアは親レベルでスケーリングされます。スケーラブル シェアを持つ親から作成されたすべての子孫リソース プールは、デフォルトでスケーラブル シェアを持ちます。

- [CPU] で、CPU リソースの設定を選択します。
- ◆ [メモリ] で、メモリ リソース設定を選択します。

4 [OK] をクリックして、変更内容を保存します。

リソース プールへの仮想マシンの追加

仮想マシンを作成する場合、作成プロセスの一環として、リソース プールの場所を指定できます。既存の仮想マシンをリソース プールに追加することもできます。

新しいリソース プールに移動した仮想マシンは、次のような状態になります。

- 仮想マシンの予約と制限は変更されません。
- 仮想マシンのシェアが高、標準、または低の場合、シェア率は、新規リソース プールで使用されているシェア数の合計に合わせて調整されます。
- 仮想マシンにカスタム シェアが割り当てられている場合は、そのシェア値が維持されます。

注： シェアの割り当てはリソース プールに対して相対的なもので、仮想マシンをリソース プールに移動する場合は、新規リソース プール内の相対的な値と仮想マシンのシェアが一致するように、仮想マシンのシェアを手動で変更しなければならない場合があります。合計シェアに対して非常に大きい（または非常に小さい）割合が仮想マシンに割り当てられると、警告が表示されます。

- [監視] では、[リソース予約] タブに表示される、リソース プールの予約済みおよび未予約の CPU とメモリのリソースに関する情報は、仮想マシンに関連付けられた予約（ある場合）に合わせて変更されます。

注： 仮想マシンがパワーオフまたはサスペンド状態の場合、その仮想マシンを移動することはできますが、リソース プール全体で使用可能なリソース（予約済みおよび未予約の CPU とメモリなど）は変更されません。

手順

1 vSphere Client で、仮想マシンを参照して移動します。

- a データセンター、フォルダ、クラスター、リソース プール、またはホストを選択して、仮想マシンを検索します。
- b [仮想マシン] タブをクリックします。

2 仮想マシンを右クリックして、[移行] を選択します。

- 仮想マシンは別のホストに移動できます。
- 仮想マシンのストレージは別のデータストアに移動できます。

- 仮想マシンを別のホストに移動し、そのストレージを別のデータストアに移動できます。
- 3 仮想マシンを配置するリソース プールを選択します。
 - 4 選択内容を確認し、[終了] をクリックします。

結果

仮想マシンがパワーオン状態で、ターゲットのリソース プールに仮想マシンの予約を確保するだけの十分な CPU またはメモリがない場合、仮想マシンの移動はアドミッション コントロールによって許可されないため、失敗します。エラー ダイアログ ボックスには、使用可能なリソースと要求されたリソースが表示されるので、調整することで問題が解決されるかどうかを判断できます。

リソース プールからの仮想マシンの削除

仮想マシンを別のリソース プールへ移動するか、仮想マシンを削除することにより、仮想マシンをリソース プールから削除できます。

仮想マシンをリソース プールから削除すると、リソース プールに関連付けられたシェアの総数が減り、残りの各シェアが、より多くのリソースを表すようになります。たとえば、シェアを [標準] に設定した 3 台の仮想マシンを含み、6GHz を割り当てられたプールがあるとします。仮想マシンが CPU バインドであったとすると、各仮想マシンは、均等に 2GHz を割り当てられます。仮想マシンの 1 台を別のリソース プールに移動した場合、残りの 2 台の仮想マシンは、それぞれ均等に 3GHz を割り当てられます。

手順

- 1 vSphere Client で、リソース プールを参照して移動します。
- 2 次のいずれかの方法で、仮想マシンをリソース プールから削除します。
 - 仮想マシンを右クリックし、[移動先...] を選択して、仮想マシンを別のリソース プールに移動します。
仮想マシンを移動する前に仮想マシンをパワーオフする必要はありません。
 - 仮想マシンを右クリックし、[ディスクから削除] を選択します。
仮想マシンを完全に削除するには、事前に仮想マシンをパワーオフする必要があります。

リソース プールの削除

リソース プールをインベントリから削除できます。

手順

- 1 vSphere Client でリソース プールを右クリックし、[削除] を選択します。
確認のダイアログ ボックスが表示されます。
- 2 [はい] をクリックしてリソース プールを削除します。

リソース プールのアドミッション コントロール

リソース プール内の仮想マシンをパワーオンする場合、または子リソース プールの作成を試みる場合、システムは追加のアドミッション コントロールを実行して、リソース プールの制限に違反しないことを確認します。

仮想マシンをパワーオンしたり、リソース プールを作成したりする前に、vSphere Client の [リソース予約] タブを使用して、リソースが十分にあることを確認します。CPU とメモリの [使用可能な予約] 値で、未予約のリソースが表示されます。

使用可能な CPU とメモリ リソースの計算法、およびアクションが実行されるかどうかは、[予約タイプ] によって決まります。

表 10-1. 予約タイプ

予約タイプ	説明
[固定]	選択されたリソース プールに予約のないリソースが十分にあるかどうかをシステムが確認します。ある場合、アクションを実行できます。ない場合、メッセージが表示され、アクションを実行することはできません。
[拡張可能] (デフォルト)	システムは、選択されたリソース プールおよび直接の親のリソース プール内で使用可能なリソースを考慮します。親のリソース プールでも [拡張可能な予約] オプションが選択されている場合は、その親のリソース プールからリソースを借りることができます。[拡張可能な予約] オプションが選択されているかぎり、リソースの借用は、現在のリソース プールの先祖から繰り返し発生します。このオプションを選択しておく、柔軟性が高まりますが、同時に保護が弱くなります。子リソース プールの所有者は、予想以上のリソースを予約することがあります。

事前構成された [予約] または [制限] の設定に違反することはできません。リソース プールを再構成するごとに、または仮想マシンをパワーオンするごとに、システムはすべてのパラメータを検証して、すべてのサービス レベルを保証できるようにします。

拡張可能な予約の例 1

この例では、拡張可能な予約を持つリソース プールの機能を示します。

管理者が、プール P を管理し、2 人の異なるユーザー（または 2 つの異なるグループ）用に 2 つの子リソース プール S1 と S2 を定義するとします。

管理者は、予約を持つ仮想マシンをユーザーがパワーオンしたいことを知っていますが、各ユーザーがどれだけの量を予約する必要があるかは知りません。S1 と S2 の予約を拡張可能にすると、管理者は、プール P の共通の予約をより柔軟に共有および継承できます。

拡張可能な予約を使用しない場合、管理者は、特定の量を明示的に S1 と S2 に割り当てる必要があります。このような特定の割り当ては、特に深いリソース プール階層において柔軟性がなくなる場合があり、リソース プール階層での予約設定が複雑になる場合があります。

拡張可能な予約では、厳密な分離がなくなります。つまり、S1 は、P の予約をすべて使い始めることができ、S2 が直接使用できるメモリまたは CPU がなくなります。

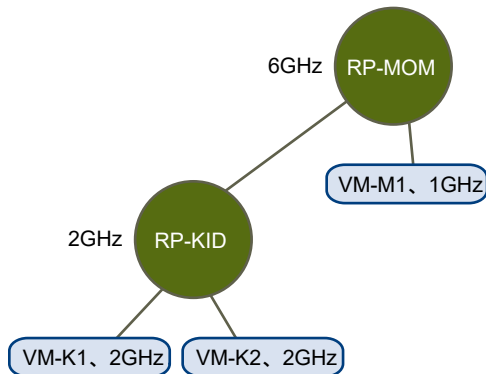
拡張可能な予約の例 2

この例では、拡張可能な予約を持つリソース プールの機能を示します。

図に示すように、次のシナリオを想定しています。

- 親プールの RP-MOM に、6GHz の予約と、1GHz を予約する実行中の 1 台の仮想マシン VM-M1 があります。
- 2GHz の予約を持ち、[拡張可能な予約] が選択された子リソース プール RP-KID を作成します。
- それぞれ 2GHz の予約を持つ 2 台の仮想マシン VM-K1 と VM-K2 を子リソース プールに追加し、これらの仮想マシンのパワーオンを試みます。
- VM-K1 は、(2GHz を持つ) RP-KID からリソースを直接予約できます。
- VM-K2 が使用できるローカル リソースはないので、VM-K2 は親リソース プール RP-MOM からリソースを借ります。RP-MOM は、6GHz マイナス 1GHz (仮想マシンによる予約) マイナス 2GHz (RP-KID による予約) を持ち、3GHz が未予約として残ります。3GHz が使用可能なので、2GHz の仮想マシンをパワーオンすることができます。

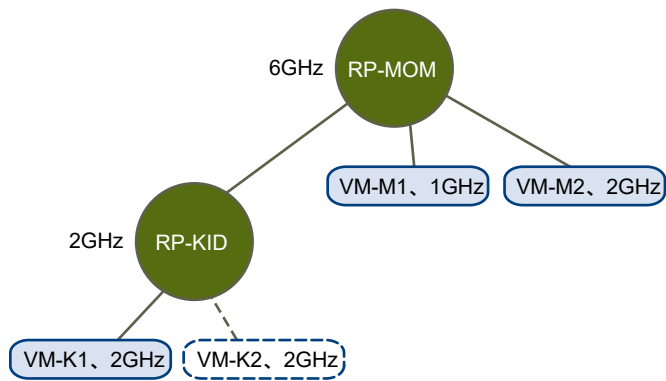
図 10-3. 拡張可能なリソース プールを使用したアドミッション コントロール：正常にパワーオンされた場合



次に、VM-M1 と VM-M2 を使用した別のシナリオについて考えます。

- 予約の合計が 3GHz である、RP-MOM 内の 2 台の仮想マシンをパワーオンします。
- この場合でも、2GHz はローカルで使用可能なので、RP-KID 内の VM-K1 をパワーオンすることができます。
- VM-K2 をパワーオンするとき、RP-KID には未予約の CPU 容量がないため、親を確認します。RP-MOM には未予約で使用可能な容量が 1GHz しかありません (RP-MOM 用に 5GHz が使用済みで、ローカル仮想マシンが 3GHz を予約し、RP-KID が 2GHz を予約しています)。その結果、2GHz の予約が必要な VM-K2 をパワーオンできません。

図 10-4. 拡張可能なリソース プールを使用したアドミッション コントロール：正常にパワーオンされなかった場合



DRS クラスタの作成

11

クラスタは、共有リソースと共有管理インターフェイスを持つ複数の ESXi ホストとそれに関連する仮想マシンの集合体です。クラスタ レベルのリソース管理の利点を得る前に、クラスタを作成し、DRS を有効にする必要があります。

Enhanced vMotion Compatibility (EVC) が有効であるかどうかに応じて、クラスタ内の vSphere フォールトトレランス (vSphere FT) 仮想マシンを使用するときの DRS 動作は異なります。

表 11-1. vSphere FT 仮想マシンおよび EVC との DRS 動作

EVC	DRS (ロード バランシング)	DRS (初期配置)
有効にする	有効 (プライマリ仮想マシンとセカンダリ仮想マシン)	有効 (プライマリ仮想マシンとセカンダリ仮想マシン)
無効	無効 (プライマリ仮想マシンとセカンダリ仮想マシン)	無効 (プライマリ仮想マシン) 完全自動化 (セカンダリ仮想マシン)

この章には、次のトピックが含まれています。

- vSphere クラスタ サービス
- アドミッション コントロールと初期配置
- 仮想マシンの移行
- DRS クラスタの要件
- 仮想フラッシュによる DRS の構成
- クラスタの作成
- クラスタ設定の編集
- 仮想マシンのカスタム自動化レベルの設定
- DRS の無効化
- リソース プール ツリーのリストア
- vSAN ストレッチ クラスタの DRS 認識

vSphere クラスタ サービス

vSphere クラスタ サービス (vCLS) はデフォルトで有効で、すべての vSphere クラスタで実行されます。vCLS は、vCenter Server が利用できなくなった場合に、クラスタ内で実行されるワークロードのリソースと健全性を維持するために、クラスタ サービスを引き続き使用できるようにします。DRS と HA を実行するには、vCenter Server が引き続き必要です。

vCLS は、vSphere 7.0 Update 3 にアップグレードするか、vSphere 7.0 Update 3 を新規にデプロイすると有効になります。vCLS は、vCenter Server アップグレードの一部としてアップグレードされます。

vCLS は、エージェント仮想マシンを使用してクラスタ サービスの健全性を維持します。ホストをクラスタに追加すると、vCLS エージェント仮想マシン (vCLS 仮想マシン) が作成されます。各 vSphere クラスタでは、最大で 3 台の vCLS 仮想マシンをクラスタ内に分散して実行する必要があります。vCLS は、ホストが 1 台または 2 台のみのクラスタでも有効になります。これらのクラスタでは、vCLS 仮想マシンはそれぞれ 1 台、2 台です。

新しい非アフィニティ ルールが自動的に適用されます。3 分ごとにチェックが実行され、複数の vCLS 仮想マシンが 1 台のホストにある場合、それらの仮想マシンは異なるホストに自動的に再配分されます。

表 11-2. クラスタ内の vCLS エージェント仮想マシンの数

クラスタ内のホストの数	vCLS エージェント仮想マシンの数
1	1
2	2
3 以上	3

vSphere DRS や vSphere HA などのクラスタ サービスがクラスタで有効になっていない場合でも、すべてのクラスタで vCLS 仮想マシンが実行されます。vCLS 仮想マシンのライフサイクル処理は、ESX Agent Manager やワークロード制御プレーンなどの vCenter Server サービスによって管理されます。vCLS 仮想マシンでは NIC はサポートされていません。

vCLS が有効なクラスタには、vCenter Server と互換性がある ESXi バージョンであれば、異なるバージョンの ESXi ホストを含めることができます。vCLS は、vSphere Lifecycle Manager クラスタと連携します。

vSphere DRS

vSphere DRS は、vSphere の重要な機能で、vSphere クラスタ内で実行されるワークロードの健全性の維持に必要です。DRS は vCLS 仮想マシンの可用性に依存します。

注： vCLS 仮想マシンに問題があるクラスタで DRS を有効にすると、[クラスタのサマリ] 画面に警告メッセージが表示されます。

注： DRS が有効でも vCLS 仮想マシンに問題がある場合、DRS が動作するにはその問題を解決する必要があります。[クラスタのサマリ] 画面に警告メッセージが表示されます。

DRS が機能しない場合でも、必ずしも DRS が無効になっているわけではありません。既存の DRS 設定とリソースプールは、vCLS 仮想マシンのクォーラムが失われても存続します。vCLS の健全性が [非健全] に変更されるのは、DRS が有効なクラスタで vCLS 仮想マシンが実行されておらず、そのために DRS の最初のインスタンスがスキップされた場合に限られます。1 台以上の vCLS 仮想マシンが実行されていない場合、DRS が有効でないクラスタでの vCLS の健全性は [低下] のままです。

vCLS 仮想マシンでのデータストア選択

vCLS 仮想マシン用のデータストアは、クラスタ内のホストに接続されているすべてのデータストアのランクに基づいて自動的に選択されます。データストアに接続されているクラスタ内のホストに予約済みの空き DRS スロットがある場合、そのデータストアが選択される可能性は高くなります。このアルゴリズムでは、可能であれば vCLS 仮想マシンをローカル データストアではなく共有データストアに配置することを試行します。空き容量の大きいデータストアが優先されます。また、このアルゴリズムでは同じデータストアに複数の vCLS 仮想マシンを配置しないようにします。vCLS 仮想マシンのデータストアの変更は、デプロイしてパワーオンした後のみ可能です。

vCLS 仮想マシンの VMDK を別のデータストアに移動する場合、または別のストレージ ポリシーを適用する場合は、vCLS 仮想マシンを再構成します。この操作を実行すると、警告メッセージが表示されます。

Storage vMotion を実行すると、vCLS 仮想マシンを別のデータストアに移行できます。1 つのデータセンターで実行されているすべての仮想マシンに適用される特定のメタデータ戦略がある場合など、vCLS 仮想マシンをワークロード仮想マシンから切り離してグループ化する必要があるときは、タグを付けたり、カスタム属性を割り当てたりすることができます。

注： データストアがメンテナンス モードのとき、データストアで vCLS 仮想マシンをホストする場合は、手動で Storage vMotion を適用して vCLS 仮想マシンを新しい場所に移動するか、クラスタを退避モードにする必要があります。警告メッセージが表示されます。

メンテナンス モードへの切り替えタスクは開始されますが、終了できません。これは、データストアに 1 台の仮想マシンが配置されているためです。続行する場合でも、[最近のタスク] でタスクを常にキャンセルできます。

選択したデータストアには、パワーオフできない vSphere クラスタ サービス仮想マシンが保存されている可能性があります。vSphere クラスタ サービスの健全性を確保するには、このデータストアをメンテナンスのために停止する前に、これらの仮想マシンを vMotion によりクラスタ内の別のデータストアに手動で移動する必要があります。ナレッジベースの記事 KB79892 を参照してください。

続行するには、[すべての仮想マシンのストレージを移行し、移行後もメンテナンス モードへの切り替えを継続します。] チェックボックスをオンにします。

vCLS データストアの配置

デフォルトの vCLS 仮想マシン データストアの配置をオーバーライドできます。

vSphere クラスタ サービス (vCLS) 仮想マシン データストアの場所は、デフォルトのデータストア選択ロジックによって選択されます。クラスタのデフォルトの vCLS 仮想マシン データストア配置をオーバーライドするには、クラスタを参照し、[構成] - [vSphere クラスタ サービス] - [データストア] で [追加] をクリックして、許可されるデータストアのセットを指定します。一部のデータストアは、vCLS を構成できない SRM などのソリューションまたは vSAN メンテナンス モードによってブロックされるため、一部のデータストアは、vCLS に対して選択できません。ユーザーは、vCLS 仮想マシンに対してソリューションにブロックされたデータストアを追加または削除できません。

vSphere クラスタ サービスの監視

vCLS 仮想マシンによって使用されたリソースと、それらの健全性ステータスを監視できます。

vCLS 仮想マシンは、[ホストおよびクラスタ] タブのインベントリ ツリーに表示されません。データセンター内のすべてのクラスタの vCLS 仮想マシンは、[vCLS] という名前の別の仮想マシンおよびテンプレート フォルダ内に配置されています。このフォルダと vCLS 仮想マシンは、vSphere Client の [仮想マシンおよびテンプレート] タブにのみ表示されます。これらの仮想マシンは、通常のワークロード仮想マシンとは異なるアイコンで識別されます。vCLS 仮想マシンの [サマリ] タブで、vCLS 仮想マシンの目的に関する情報を確認できます。

[監視] タブで、vCLS 仮想マシンによって使用されるリソースを監視できます。

表 11-3. vCLS 仮想マシンのリソース割り当て

プロパティ	サイズ
VMDK サイズ	245 MB (シン ディスク)
メモリ	128 MB
CPU	1 vCPU
ハード ディスク	2 GB
データストア上のストレージ	480 MB (シン ディスク)

注： vCLS 仮想マシンごとに、クラスタに 100 MHz、100 MB のキャパシティが予約されています。クラスタ内で実行されている vCLS 仮想マシンの数に応じて、これらの仮想マシンに対して最大 400 MHz、400 MB のキャパシティを予約できます。

クラスタの [サマリ] タブに表示されている [クラスタ サービス] ポートレット内で、vCLS の健全性ステータスを監視できます。

表 11-4. vCLS の健全性ステータス

ステータス	色分け	サマリ
健全	緑	1 台以上の vCLS 仮想マシンが実行されている場合、クラスタ内のホストの数に関係なく、ステータスは健全なままになります。
低下しました	黄色	vCLS 仮想マシンが実行されていない状態が 3 分未満続いた場合、ステータスは低下になります。
非健全	赤	vCLS 仮想マシンが実行されていない状態が 3 分以上続いた場合、DRS 対応クラスタのステータスは非健全になります。

vSphere クラスタ サービスの健全性の維持

vSphere DRS はこれらの仮想マシンの可用性に依存しているため、vCLS 仮想マシンは常にパワーオン状態になります。これらの仮想マシンはシステム仮想マシンとして扱う必要があります。管理者のみが vCLS 仮想マシンで選

択的な操作を実行できます。クラスタ サービスの障害を回避するために、vCLS 仮想マシンでは構成や操作を実行しないでください。

vCLS 仮想マシンは、誤って削除されないように保護されています。クラスタ仮想マシンとフォルダは、管理者を含むユーザーが変更できないように保護されています。

管理者 SSO グループのメンバーであるユーザーだけが次の管理操作を実行できます。

- vCLS 仮想マシンに対する読み取り専用アクセス
- vCLS 仮想マシンに対するコンソール アクセス
- コールド移行またはホット移行を使用して、vCLS 仮想マシンを新しいストレージ、コンピューティング リソース、またはその両方に再配置します
- vCLS 仮想マシンに対するタグとカスタム属性の使用

vCLS 仮想マシンの健全な機能を妨げる可能性のある操作は、次のとおりです。

- vCLS 仮想マシンの電源状態の変更
- CPU、メモリ、ディスク サイズ、ディスクの配置など、vCLS 仮想マシンのリソースの再構成
- 仮想マシンの暗号化
- vCLS 仮想マシンの vMotion のトリガ
- BIOS の変更
- インベントリからの vCLS 仮想マシンの削除
- ディスクからの vCLS 仮想マシンの削除
- vCLS 仮想マシンの FT の有効化
- vCLS 仮想マシンのクローン作成
- PMem の構成
- vCLS 仮想マシンを別のフォルダへの移動
- vCLS 仮想マシンの名前変更
- vCLS フォルダの名前変更
- vCLS 仮想マシンでの DRS ルールとオーバーライドの有効化
- vCLS 仮想マシンでの HA アドミッション コントロール ポリシーの有効化
- vCLS 仮想マシンでの HA オーバーライドの有効化
- vCLS 仮想マシンのリソース プールへの移動
- スナップショットからの vCLS 仮想マシンのリカバリ

vCLS 仮想マシンで中断操作を実行すると、警告ダイアログ ボックスが表示されます。

トラブルシューティング：

電源状態を含む vCLS 仮想マシンの健全性は、EAM および WCP サービスによって管理されます。vCLS 仮想マシンのパワーオンに失敗した場合、または vCLS 仮想マシンのクォーラム不足のためにクラスタの DRS の最初のインスタンスがスキップされた場合は、クラスタのサマリ ページにバナーが表示され、エラー状態のトラブルシューティングに役立つナレッジベースの記事へのリンクが示されます。

vCLS 仮想マシンはシステム仮想マシンとして扱われるため、これらの仮想マシンをバックアップしたり、スナップショットを作成したりする必要はありません。これらの仮想マシンの健全性の状態は、vCenter Server サービスによって管理されます。

クラスタの退避モードへの切り替え

データストアがメンテナンス モードのとき、データストアで vCLS 仮想マシンをホストする場合は、Storage vMotion によって vCLS 仮想マシンを新しい場所に手動で移動するか、クラスタを退避モードにする必要があります。

このタスクでは、クラスタを退避モードにする方法について説明します。

手順

- 1 vSphere Client にログインします。
- 2 vCLS を無効にするクラスタに移動します。
- 3 ブラウザの URL フィールドからクラスタのドメイン ID をコピーします。これは、**domain-c(数字)** のような形式です。

注： URL のコロンの左側にある数字のみをコピーします。

- 4 vCenter Server の [構成] タブに移動します。
- 5 [詳細設定] で、[設定の編集] ボタンをクリックします。
- 6 新しいエントリ **config.vcls.clusters.domain-c(数字).enabled** を追加します。手順 3 でコピーしたドメイン ID を使用します。
- 7 [値] を **False** に設定します。
- 8 [保存] をクリックします。

結果

vCLS 監視サービスは 30 秒ごとに実行されます。1 分以内に、クラスタ内のすべての vCLS 仮想マシンがクリーンアップされ、[クラスタ サービス] の健全性が [低下] に設定されます。クラスタで DRS が有効になっている場合、DRS 機能は停止し、[クラスタのサマリ] に追加の警告が表示されます。DRS が有効になっていても、退避モードから取り除くことで vCLS を再構成するまで DRS は機能しません。

vSphere HA による配置は、ホスト障害が発生している状況では最適になりません。HA は、配置の推奨について DRS に依存しています。それでも HA は仮想マシンをパワーオンしますが、これらの仮想マシンは最適でないホスト上でパワーオンされている可能性があります。

クラスタの退避モードを解除するには、手順 7 の値を **True** に変更します。

vCLS 仮想マシンのパスワードの取得

vCLS 仮想マシンにログインするためのパスワードを取得できます。

クラスタ サービスの健全性を確保するために、vCLS 仮想マシンへのアクセスは避けてください。このドキュメントは、vCLS 仮想マシンの明示的な診断を目的としています。

手順

- 1 SSH を使用して vCenter Server Appliance にログインします。
- 2 次の Python スクリプトを実行します。

```
/usr/lib/vmware-wcp/decrypt_clustervm_pw.py
```

- 3 パスワードの出力を読み取ります。

```
pwd-script-output

Read key from file

Connected to PSQL

PWD: (password displayed here)
```

結果

取得したパスワードを使用して、vCLS 仮想マシンにログインできます。

vCLS 仮想マシン非アフィニティ ポリシー

vSphere では、vCLS 仮想マシンとワークロード仮想マシンの別のグループ間の非アフィニティがサポートされません。

コンピューティング ポリシーを使用すると、vSphere Distributed Resource Scheduler (DRS) でリソース プール内のホストに仮想マシンが配置される方法を指定できます。vSphere コンピューティング ポリシー エディタを使用して、コンピューティング ポリシーを作成および削除します。コンピューティング ポリシーは作成または削除できますが、変更することはできません。ポリシーの定義で使用されているカテゴリ タグを削除すると、ポリシーも削除されます。vSphere の [仮想マシンのサマリ] ページを開き、仮想マシンに適用されるコンピューティング ポリシーと、各ポリシーとのコンプライアンス ステータスを表示します。vCLS 仮想マシンのグループに非アフィニティであるワークロード仮想マシンのグループのコンピューティング ポリシーを作成できます。vCLS 非アフィニティ ポリシーは、ワークロード仮想マシンのグループに対して単一のユーザー表示タグを持つことができ、vCLS 仮想マシンの他のグループは内部的に認識されます。

vCLS 仮想マシン非アフィニティ ポリシーの作成または削除

vCLS 仮想マシン非アフィニティ ポリシーでは、仮想マシンと vCLS システム仮想マシンのカテゴリ間の関係を説明します。

vCLS 仮想マシン非アフィニティ ポリシーでは、vCLS 仮想マシンとアプリケーション仮想マシンを同じホストに配置することは推奨されません。このタイプのポリシーは、vCLS 仮想マシンと重要なワークロードを実行している仮想マシンを同じホスト上で実行しない場合に有用です。SAP HANA などの重要なワークロードを実行するためのいくつかのベスト プラクティスには、専用ホストが必要です。ポリシーが作成されると、配置エンジンは、ポリシー仮想マシンが実行されていないホストに vCLS 仮想マシンを配置しようとします。

vCLS 仮想マシン非アフィニティ ポリシーの実施は、いくつかの点で影響を受ける可能性があります。

- さまざまなホスト上の複数の仮想マシンにポリシーが適用され、vCLS 仮想マシンを分散するための十分なホストを確保できない場合、vCLS 仮想マシンはポリシー仮想マシンのないホストに統合されます。
- プロビジョニング操作でターゲット ホストが指定された場合、その指定はポリシーに違反しても常に考慮されません。DRS は、後続の修正サイクルで vCLS 仮想マシンを準拠ホストに移動しようとします。

手順

- 1 vCLS 仮想マシン非アフィニティ ポリシーに含める仮想マシンのグループごとに、カテゴリを作成してタグを付けます。
- 2 含める仮想マシンにタグを付けます。
- 3 vCLS 仮想マシン非アフィニティ ポリシーを作成します。
 - a vSphere で、[ポリシーおよびプロファイル] - [コンピューティング ポリシー] の順にクリックします。
 - b [追加] をクリックし、[新しいコンピューティング ポリシー] ウィザードを開きます。
 - c ポリシーの [名前] を入力し、[ポリシー タイプ] ドロップダウン コントロールから [vCLS 仮想マシンの非アフィニティ] を選択します。
 ポリシーの [名前] は一意である必要があります。
 - d ポリシーの [説明] を入力してから、[仮想マシン タグ] を使用して、ポリシーを適用する [カテゴリ] と [タグ] を選択します。
 複数の仮想マシン タグが1つのカテゴリに関連付けられている場合を除き、タグ [カテゴリ] を選択すると、ウィザードで仮想マシン タグが入力されます。
 - e [作成] をクリックしてポリシーを作成します。
- 4 (オプション) コンピューティング ポリシーを削除するには、vSphere を開き、[ポリシーおよびプロファイル] - [コンピューティング ポリシー] の順にクリックして、各ポリシーをカードとして表示します。[削除] をクリックして、ポリシーを削除します。

アドミッション コントロールと初期配置

DRS が有効なクラスタ内の単一の仮想マシンまたは仮想マシンのグループをパワーオンする際、vCenter Server がアドミッション コントロールを実行します。アドミッション コントロールによって、クラスタ内に仮想マシンをサポートする十分なリソースがあるかどうかを確認されます。

単一の仮想マシン、または仮想マシン グループ内のいずれかの仮想マシンのパワーオンに十分なリソースがクラスタにない場合、メッセージが表示されます。十分なリソースがある場合、DRS は各仮想マシンを実行するホストの配置を推奨し、次のアクションのいずれかを行います。

- 配置の推奨を自動実行する。
- 配置の推奨を表示し、ユーザーは承諾するか変更するかを選択できる。

注： スタンドアロン ホストまたは非 DRS クラスタの仮想マシンには、初期配置の推奨は表示されません。パワーオン時に、現在格納されているホストに配置されます。

- DRS では、ネットワーク バンド幅が考慮されます。ホスト ネットワークの利用状況を計算することで、DRS は配置についてさらに適正な判断を下すことができます。これにより、環境をさらに包括的に把握し、仮想マシンのパフォーマンスの低下を回避できます。

1 台の仮想マシンのパワーオン

DRS クラスタで、単一仮想マシンをパワーオンし、初期配置の推奨を取得します。

単一仮想マシンをパワーオンする場合、2 種類の初期配置の推奨があります。

- パワーオンに必要な前処理はなく、1 台の仮想マシンがパワーオンされる。
ユーザーには、相互に排他的な仮想マシンの初期配置の推奨のリストが表示されます。1 つだけを選択できます。
- パワーオンに必要な前処理を行ってから、1 台の仮想マシンがパワーオンされる。
これらの操作には、スタンバイ モードのホストのパワーオン、別の仮想マシンのホスト間の移行が含まれます。この場合、推奨は複数行で表示され、必要な各操作が示されます。ユーザーは、この推奨全体を受け入れるか、仮想マシンのパワーオンをキャンセルできます。

グループ パワーオン

複数の仮想マシンを同時にパワーオンできます（グループ パワーオン）。

グループのパワーオンの試行に選択する仮想マシンは、同じ DRS クラスタにある必要はありません。クラスタをまたがって選択できますが、同じデータセンター内にある必要があります。また、DRS 以外のクラスタまたはスタンドアロン ホストにある仮想マシンを含めることもできます。これらの仮想マシンは自動的にパワーオンされますが、初期配置の推奨には含まれません。

グループ パワーオンの初期配置に関する推奨は、クラスタ単位で提供されます。グループのパワーオンの試行に関する配置関連のすべてのアクションが自動モードの場合、初期配置の推奨は表示されずに、仮想マシンがパワーオンされます。いずれかの仮想マシンの配置関連のアクションが手動モードである場合は、すべての仮想マシン（自動モードである仮想マシンを含む）のパワーオンは、手動で行われます。そのようなアクションは、初期配置の推奨に含まれています。

パワーオンする仮想マシンが属する DRS クラスタごとに、必要なすべての前提条件が含まれた単一の推奨が作成されます（または、推奨は作成されません）。このようなクラスタ固有の推奨は、[パワーオン推奨] タブの下にまとめて表示されます。

自動モードでないグループのパワーオンが試行され、初期配置に関する推奨の対象外である仮想マシン（スタンドアロン ホストまたは DRS 以外のクラスタにある仮想マシン）が含まれている場合、vCenter Server はこれらの仮想マシンを自動的にパワーオンしようとします。このパワーオンが成功すると、[パワーオン開始] タブの下に仮想マシンが一覧表示されます。パワーオンに失敗した仮想マシンは、[パワーオン失敗] タブに一覧表示されます。

例：グループ パワーオン

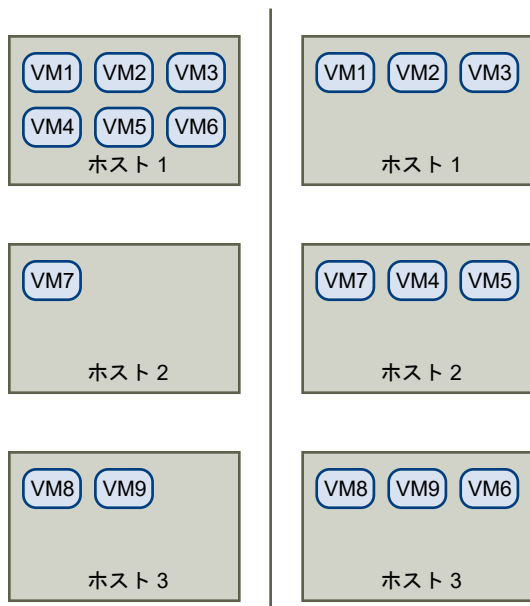
ユーザーは、グループのパワーオンを試行するために、同じデータセンターにある 3 台の仮想マシンを選択します。最初の 2 台の仮想マシン（VM1 と VM2）は、同じ DRS クラスタ（Cluster1）にあり、3 番目の仮想マシン（VM3）はスタンドアロン ホストにあります。VM1 は自動モードで、VM2 は手動モードです。このシナリオでは、VM1 と VM2 をパワーオンするアクションで構成された Cluster1 の初期配置の推奨が（[パワーオン推奨] タブに）表示されます。VM3 のパワーオン試行は自動的に行われ、成功した場合、[パワーオン開始] タブに表示されます。この試行が失敗した場合は、[パワーオン失敗] タブに一覧表示されます。

仮想マシンの移行

DRS が初期配置を実行するため、クラスタ間の負荷は均衡になりますが、仮想マシンの負荷やリソースの可用性が変化すると、クラスタがアンバランスになります。このようなアンバランスを調整するには、DRS が移行の推奨を作成します。

クラスタで DRS が有効な場合は、負荷をより均一に分散し、前述の不均衡を軽減できます。たとえば、この図の左側にある 3 台のホストは、不均衡です。ホスト 1、ホスト 2、ホスト 3 の容量が同一であり、すべての仮想マシンの構成と負荷（予約が設定されている場合はそれも含む）が同じであるとしてします。しかし、ホスト 1 には 6 台の仮想マシンがあるので、ホスト 1 のリソースは使用過剰になる場合があるのに対し、ホスト 2 とホスト 3 では、使用可能なリソースが十分にあります。DRS は、仮想マシンをホスト 1 からホスト 2 とホスト 3 に移行します（または移行を推奨します）。図の右側に、適切にロード バランシングされたあとのホストの構成が表示されます。

図 11-1. ロード バランシング



クラスタが不均衡になった場合、DRS はデフォルトの自動化レベルに応じて、推奨を行うか仮想マシンを移行します。

- クラスタまたはいずれかの仮想マシンが手動または一部自動化の場合、vCenter Server はリソースのバランスをとるための自動的なアクションを実行しません。代わりに、移行の推奨が使用可能であることがサマリ ページに表示され、クラスタ内で最も効率的にリソースを使用できる変更の推奨が DRS 推奨ページに表示されます。
- 関連するクラスタおよび仮想マシンが完全自動化の場合、vCenter Server は、実行中の仮想マシンを必要に応じてホスト間で移行して、クラスタ リソースが効率的に使用できるようにします。

注： 自動移行設定であっても、ユーザーが個々の仮想マシンを明示的に移行できますが、vCenter Server は、クラスタ リソースを最適化するためにそれらの仮想マシンをほかのホストに移動する場合があります。

デフォルトでは、自動化レベルはクラスタ全体に対して指定されます。個々の仮想マシンについてカスタム自動化レベルを指定することもできます。

DRS 移行のしきい値

DRS 移行のしきい値を使用すると、どの推奨を生成して適用するか（推奨に関与する仮想マシンが完全自動化モードの場合）、または表示するか（手動モードの場合）を指定できます。このしきい値は、仮想マシンの健全性を高めるために移行を推奨する上で、DRS の積極性をどの程度にするかを示す尺度です。

しきい値スライダを移動して、[保守的] から [積極的] までの 5 つの設定のうち、1 つを使用できます。積極性設定を高くすると、仮想マシンの健全性を高めるために DRS が移行を推奨する頻度が高くなります。[保守的] に設定すると、優先順位が最上位の推奨（必須の推奨）のみが生成されます。

推奨が優先順位レベルを取得したあと、そのレベルは設定した移行のしきい値と比較されます。優先順位レベルがしきい値の設定以下の場合、推奨が適用される（関連する仮想マシンが完全自動化モードの場合）か、ユーザーの確認のために表示されます（手動モードまたは一部自動化モードの場合）。

DRS スコア

個々の移行の推奨は、実行効率を測定する仮想マシン健全性メトリックを使用して計算されます。このメトリックは、vSphere Client にあるクラスタの [サマリ] タブに DRS スコアとして表示されます。DRS ロード バランシングの推奨は、仮想マシンの DRS スコアの改善を意図しています。クラスタの DRS スコアは、クラスタ内でパワーオンされているすべての仮想マシンの仮想マシン DRS スコアの加重平均です。クラスタの DRS スコアはゲージコンポーネントに表示されます。値表示部分の色は、仮想マシンの DRS スコアのヒストグラムの対応するバーと一致するように変化します。ヒストグラムのバーには、DRS スコアがその範囲内に該当する仮想マシンの割合が表示されます。クラスタの [監視] タブを選択して [vSphere DRS] を選択すると、サーバ側でソートおよびフィルタリングしたリストを表示できます。これにより、クラスタ内の仮想マシンが DRS スコアの昇順にソートされて一覧表示されます。

移行の推奨

デフォルトが手動または一部自動化モードのクラスタを作成した場合、vCenter Server は DRS 推奨ページに移行の推奨を表示します。

システムは、ルールを実施してクラスタのリソースのバランスをとるのに必要なだけの推奨を行います。各推奨には、移動する仮想マシン、現在の（ソース）ホストとターゲットホスト、および推奨の理由が含まれます。理由は次のいずれかです。

- CPU 負荷または予約を均衡化する。
- メモリ負荷または予約を均衡化する。
- リソース プールの予約に従う。
- アフィニティ ルールに従う。
- ホストがメンテナンス モードまたはスタンバイ モードに切り替えている最中である。

注： vSphere Distributed Power Management の機能を使用している場合、DRS は移行の推奨のほかに、ホストの電源状態の推奨を提供します。

DRS クラスタの要件

クラスタ機能を正しく使用するには、DRS クラスタに追加するホストが、特定の要件を満たしている必要があります。

注： vSphere DRS は、vSphere の重要な機能で、vSphere クラスタ内で実行されるワークロードの健全性の維持に必要です。vSphere 7.0 Update 1 以降では、DRS は vCLS 仮想マシンの可用性に依存します。詳細については [vSphere クラスタ サービス](#) を参照してください。

共有ストレージの要件

DRS クラスタには、一定の共有ストレージ要件があります。

管理対象ホストが共有ストレージを使用している必要があります。共有ストレージは、SAN 上にあるのが一般的ですが、NAS 共有ストレージを使用して実装することもできます。

他の共有ストレージの情報については、『vSphere のストレージ』ドキュメントを参照してください。

共有 VMFS ボリュームの要件

DRS クラスタには、一定の共有 VMFS ボリュームの要件があります。

共有 VMFS ボリュームを使用するように、すべての管理対象ホストを構成します。

- ソースホストとターゲットホストからアクセスできる VMFS ボリューム上に、全仮想マシンのディスクを設置します。
- VMFS ボリュームが、仮想マシンのすべての仮想ディスクを保存するのに十分なサイズであることを確認します。

- ソース ホストとターゲット ホスト上のすべての VMFS ボリュームがボリューム名を使用すること、およびすべての仮想マシンがそれらのボリューム名を使用して仮想ディスクを指定することを確認します。

注： 仮想マシン スワップ ファイルも、.vmdk 仮想ディスクと同様に、ソース ホストとターゲット ホストがアクセスできる VMFS 上に配置される必要があります。すべてのソース ホストとターゲット ホストが ESX Server 3.5 以降で、ホストのローカル スワップを使用している場合、この要件は適用されません。この場合は、非共有ストレージのスワップ ファイルを使用する vMotion がサポートされます。スワップ ファイルはデフォルトで VMFS 上に配置されますが、管理者は、仮想マシンの詳細構成オプションを使用してファイルの場所をオーバーライドできません。

プロセッサの互換性要件

DRS クラスタには、一定のプロセッサ互換性要件があります。

DRS の機能が制限されないようにするためには、クラスタ内のソース ホストとターゲット ホストのプロセッサの互換性を最大にする必要があります。

vMotion は、実行中の仮想マシンのアーキテクチャ状態を、基盤となる ESXi ホスト間で移行します。vMotion の互換性とは、ソース ホストのプロセッサのサスペンドに用いたのと同等の命令を使用して、ターゲット ホストのプロセッサが実行をレジュームできなければならないことを意味します。vMotion での移行に必要な互換性を保つためには、プロセッサのクロック速度とキャッシュ サイズは異なってもかまいませんが、プロセッサが同じクラスのベンダー（Intel と AMD）製で、同じプロセッサ ファミリーでなければなりません。

プロセッサ ファミリーはプロセッサ ベンダーによって決められています。同じファミリー内の異なるプロセッサ バージョンは、プロセッサのモデル、ステッピング レベル、拡張機能を比較することで区別できます。

プロセッサ ベンダーが、同じプロセッサ ファミリー内で大幅なアーキテクチャ変更（64 ビット拡張や SSE3 など）を行うことがあります。vMotion での正常な移行を保証できない場合、これらの例外を確認します。

vCenter Server には、vMotion で移行した仮想マシンがプロセッサの互換性要件を確実に満たすようにする機能が備わっています。これには、次の機能が含まれます。

- Enhanced vMotion Compatibility (EVC)：EVC を使用して、クラスタ内のホストの vMotion 互換性を確保できます。ホスト上の実際の CPU が異なる場合でも、EVC によって 1 つのクラスタ内のすべてのホストが確実に同じ CPU 機能セットを仮想マシンに提供するようになります。これにより、互換性のない CPU が原因で vMotion での移行が失敗することがなくなります。

EVC は、クラスタ設定ダイアログ ボックスで構成します。クラスタで EVC を使用するためには、クラスタ内のホストが特定の要件を満たす必要があります。EVC および EVC の要件については、『vCenter Server およびホストの管理』ドキュメントを参照してください。

- CPU 互換性マスク：vCenter Server は仮想マシンで使用できる CPU 機能をターゲット ホストの CPU 機能と比較して、vMotion での移行が可能かどうかを判断します。個々の仮想マシンに CPU 互換性マスクを適用することにより、特定の CPU 機能を仮想マシンから隠し、互換性のない CPU が原因で vMotion での移行が失敗する可能性を減らすことができます。

DRS クラスタの vMotion 要件

DRS クラスタには、一定の vMotion 要件があります。

DRS 移行の推奨を使用できるようにする場合は、クラスタ内のホストが vMotion ネットワークの一部である必要があります。ホストが vMotion ネットワーク内にない場合でも、DRS は初期配置の推奨を行うことができます。

vMotion 用に構成するためには、クラスタ内の各ホストが次の要件を満たしている必要があります。

- vMotion は、Raw ディスク、または MSCS (Microsoft Cluster Service) でクラスタリングしたアプリケーションの移行をサポートしていません。
- vMotion では、vMotion に対応したすべての管理対象ホスト間に、プライベートなギガビット イーサネット移行ネットワークが必要です。管理対象ホスト上で vMotion を有効にする場合は、管理対象ホストに固有のネットワーク ID オブジェクトを構成し、これをプライベートな移行用ネットワークに接続します。

仮想フラッシュによる DRS の構成

DRS は、仮想フラッシュ予約を持つ仮想マシンを管理できます。

仮想フラッシュ容量は、ホストから vSphere Client に定期的に報告される統計情報として表されます。DRS が実行されるたびに、報告された最新の容量値が使用されます。

ホストあたり 1 つの仮想フラッシュ リソースを構成できます。そのため、仮想マシンがパワーオンになっている間、DRS は特定のホスト上のさまざまな仮想フラッシュ リソースの中から選択する必要がありません。

DRS は、仮想マシンの起動に十分な仮想フラッシュ容量のあるホストを選択します。DRS が仮想マシンの仮想フラッシュ予約を満たせない場合、パワーオンにすることはできません。DRS は、仮想フラッシュ予約のあるパワーオン状態の仮想マシンを、その現在のホストとソフト アフィニティがあるものとして扱います。DRS では vMotion にそのような仮想マシンを使用することは推奨されません。ただし、ホストをメンテナンス モードにしたり、過剰に使用されているホストの負荷を減らしたりするなどの強制的な理由がある場合はこの限りではありません。

クラスタの作成

クラスタとは、ホストのグループです。あるクラスタに 1 台のホストが追加されると、そのホストのリソースはそのクラスタのリソースの一部になります。クラスタは、そのクラスタ内のすべてのホストのリソースを管理します。

クラスタは、vSphere High Availability (HA) および vSphere Distributed Resource Scheduler (DRS) ソリューションを有効化します。

注： vSphere DRS は、vSphere の重要な機能で、vSphere クラスタ内で実行されるワークロードの健全性の維持に必要です。vSphere 7.0 Update 1 以降では、DRS は vCLS 仮想マシンの可用性に依存します。詳細については [vSphere クラスタ サービス](#) を参照してください。

前提条件

- クラスタ オブジェクトの作成に必要な権限を持っていることを確認します。
- データセンターがインベントリに存在することを確認します。
- vSAN を使用する場合は、vSphere HA を構成する前に有効にする必要があります。

手順

- 1 vSphere Client で、データセンターを参照します。

- 2 データセンターを右クリックし、[新規クラスタ] を選択します。
- 3 クラスタの名前を入力します。
- 4 DRS と vSphere HA クラスタ機能を選択します。

オプション	説明
クラスタで DRS を使用するには、次の手順を実行します。	<ul style="list-style-type: none"> a [DRS をオンにする] チェック ボックスを選択します。 b 自動化レベルと移行のしきい値を選択します。
クラスタで HA を使用するには、次の手順を実行します。	<ul style="list-style-type: none"> a [vSphere HA をオンにする] チェック ボックスを選択します。 b ホスト監視およびアドミッション コントロールを有効にするかどうかを選択します。 c アドミッション コントロールを有効にする場合は、ポリシーを指定します。 d 仮想マシン監視オプションを選択します。 e 仮想マシンの監視感度を指定します。

- 5 Enhanced vMotion Compatibility (EVC) の設定を選択します。

実際の CPU が異なる場合でも、クラスタ内のすべてのホストが同じ CPU 機能セットを仮想マシンに確実に提供します。このため、互換性のない CPU が原因で vMotion での移行が失敗することはありません。

- 6 [OK] をクリックします。

結果

クラスタがインベントリに追加されます。

次のステップ

クラスタにホストとリソース プールを追加します。

注： [クラスタのサマリ] ページに、vSphere Cluster サービスの健全性のステータスを示す [クラスタ サービス] が表示されます。

クラスタ設定の編集

ある DRS クラスタにホストを追加すると、ホストのリソースはクラスタのリソースの一部になります。このリソースの統合に加えて、DRS クラスタを使用すると、クラスタ全体にわたるリソース プールがサポートされ、リソース割り当てポリシーをクラスタ レベルで適用できます。

次のクラスタ レベルのリソース管理機能も使用できます。

ロード バランシング

クラスタ内にあるすべてのホストおよび仮想マシンの CPU およびメモリ リソースの配分と使用率を継続的に監視します。DRS は、これらのメトリックを、クラスタのリソース プールと仮想マシンの属性に設定された理想的なリソース使用率、現在の需要、および負荷が調整されていないホストと比較します。次に DRS は推奨を行い、それに応じて仮想マシン移行を実行します。[仮想マシンの移行](#)を参照してください。クラスタ内で仮想マシンをパワーオンしたときに、DRS は仮想マシンを適切なホストに配置するか、推奨を提示して、適切なロード バランシングを維持しようとします。[アドミSSION コントロールと初期配置](#)を参照してください。

電源管理

vSphere Distributed Power Management (DPM) の機能が有効になっている場合、DRS はクラスタ レベルおよびホスト レベルのキャパシティと、クラスタの仮想マシンに必要なキャパシティ（最近の需要履歴を含む）とを比較します。次に DRS は、ホストをスタンバイ状態にすることを推奨するか、または十分なキャパシティがあることがわかったなら、ホストをスタンバイ電源モードにします。DRS は、キャパシティが必要になった場合にはホストをパワーオンします。ホストの電源状態に関する推奨に応じて、ホスト間で仮想マシンの移行が必要になる場合があります。[電力リソースの管理](#)を参照してください。

アフィニティ ルール

アフィニティ ルールを割り当てることで、クラスタ内のホスト上の仮想マシンの配置を制御できます。[DRS アフィニティ ルールの使用](#)を参照してください。

前提条件

クラスタの作成には特別なライセンスは必要ありませんが、クラスタで vSphere DRS または vSphere HA を有効にするにはライセンスが必要です。

注： vSphere DRS は、vSphere の重要な機能で、vSphere クラスタ内で実行されるワークロードの健全性の維持に必要です。vSphere 7.0 Update 1 以降では、DRS は vCLS 仮想マシンの可用性に依存します。詳細については [vSphere クラスタ サービス](#)を参照してください。

手順

- 1 vSphere Client 内のクラスタに移動します。
- 2 [構成] タブをクリックし、[サービス] をクリックします。
- 3 [vSphere DRS] で [編集] をクリックします。
- 4 [DRS 自動化] で、DRS のデフォルト自動化レベルを選択します。

自動化レベル	操作
手動	<ul style="list-style-type: none"> ■ 初期配置： 推奨されるホストが表示されます。 ■ 移行： 推奨が表示されます。
一部自動化	<ul style="list-style-type: none"> ■ 初期配置： 自動的に実行されます。 ■ 移行： 推奨が表示されます。
完全自動化	<ul style="list-style-type: none"> ■ 初期配置： 自動的に実行されます。 ■ 移行： 推奨が自動的に実行されます。

- 5 DRS の [移行のしきい値] を設定します。

- 6 [Predictive DRS] チェックボックスを選択します。リアルタイム メトリックに加え、DRS は vRealize Operations サーバが提供する予測メトリックにも対応します。この機能をサポートするバージョンの vRealize Operations でも、[Predictive DRS] を構成する必要があります。
- 7 [仮想マシンの自動化] チェック ボックスを選択して、個々の仮想マシンの自動化レベルを有効にします。各仮想マシンのオーバーライドは「仮想マシンのオーバーライド」ページから設定できます。
- 8 [その他のオプション] で、チェック ボックスを選択してデフォルト ポリシーの1つを適用します。

オプション	説明
仮想マシンの分散	可用性のために、複数のホスト間で仮想マシンをより均等に分散させます。これは、DRS のロード バランシングに次いで実行されます。
ロード バランシングのメモリ メトリック	アクティブなメモリではなく仮想マシンの使用メモリに基づいたロード バランシングを行います。この設定は、ホスト メモリがオーバーコミットしていないクラスタにのみ推奨されません。 注： この設定はサポートが終了しました。vCenter 7.0 では表示されません。
CPU オーバーコミットメント	クラスタ内の CPU オーバーコミットメントを制御します。
スケーラブル シェア	このクラスタのリソース プールに対して、スケーラブル シェアを有効にします。

- 9 [電源管理] で、自動化レベルを選択します。
- 10 DPM を有効にする場合は、[DPM しきい値] を選択します。
- 11 [OK] をクリックします。

次のステップ

注： [クラスタのサマリ] ページに、vSphere Cluster サービスの健全性のステータスを示す [クラスタ サービス] が表示されます。

vSphere Client で、DRS のメモリ使用率を確認できます。詳細については、以下を参照してください：



(Distributed Resource Scheduler (DRS) のメモリ使用率の表示)

仮想マシンのカスタム自動化レベルの設定

DRS クラスタを作成したあとで、個々の仮想マシンの自動化レベルをカスタマイズして、クラスタのデフォルト自動化レベルをオーバーライドできます。

たとえば、完全に自動化されたクラスタ内の特定の仮想マシンに対して [手動] を選択したり、手動のクラスタ内の特定の仮想マシンに対して [一部自動化] を選択したりできます。

仮想マシンを [無効] に設定すると、vCenter Server でその仮想マシンの移行や、移行に関する推奨を行うことがなくなります。

手順

- 1 vSphere Client のクラスタを参照して移動します。

- 2 [構成] タブをクリックし、[サービス] をクリックします。
- 3 [サービス] で [vSphere DRS] を選択し、[編集] をクリックします。[DRS 自動化] を展開します。
- 4 [各仮想マシンの自動化レベルを有効にする] チェック ボックスを選択します。
- 5 個々の仮想マシンのオーバーライドを一時的に無効にするには、[各仮想マシンの自動化レベルを有効にする] チェック ボックスを選択解除します。
チェック ボックスをもう一度選択すると、仮想マシンの設定は元に戻ります。
- 6 クラスタ内のすべての vMotion アクティビティを一時的にサスペンドするには、クラスタを手動モードに切り替え、[各仮想マシンの自動化レベルを有効にする] チェック ボックスを選択解除します。
- 7 1 つ以上の仮想マシンを選択します。
- 8 [自動化レベル] 列をクリックし、ドロップダウン メニューから自動化レベルを選択します。

オプション	説明
手動	配置と移行の推奨が表示されますが、推奨が手動で適用されるまで実行はされません。
完全自動化	配置と移行の推奨が自動的に実行されます。
一部自動化	初期配置が自動的に実行されます。移行の推奨は表示されますが、実行はされません。
無効	vCenter Server では、仮想マシンの移行や、移行のための推奨は行われません。

- 9 [OK] をクリックします。

結果

注： ほかの当社製品または機能（vSphere vApp および vSphere Fault Tolerance など）が、DRS クラスタ内の仮想マシンの自動化レベルをオーバーライドする場合があります。詳細については、個別の製品ドキュメントを参照してください。

DRS の無効化

クラスタの DRS をオフにできます。

DRS を無効にすると、DRS を再度オンにしたときに、クラスタのリソース プールの階層およびアフィニティ ルールが再構築されません。DRS を無効にすると、リソース プールはクラスタから削除されます。リソース プールの消失を回避するには、リソース プール ツリーのスナップショットをローカル マシンに保存します。このスナップショットを使用すれば、DRS を有効にしたときに、リソース プールをリストアできます。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 [構成] タブをクリックし、[サービス] をクリックします。
- 3 [vSphere DRS] の [編集] をクリックします。
- 4 [vSphere DRS をオンにする] チェック ボックスを選択解除します。
- 5 [OK] をクリックして DRS をオフにします。

- 6 (オプション) リソース プールを保存するオプションを選択します。
- リソース プール ツリーをローカル マシンに保存する場合は、[はい] をクリックします。
 - リソース プール ツリーのスナップショットを保存せずに、DRS をオフにする場合は、[いいえ] をクリックします。

結果

DRS はオフになりました。

注： vSphere DRS は、vSphere の重要な機能で、vSphere クラスタ内で実行されるワークロードの健全性の維持に必要です。vSphere 7.0 Update 1 以降では、DRS は vCLS 仮想マシンの可用性に依存します。詳細については [vSphere クラスタ サービス](#) を参照してください。

リソース プール ツリーのリストア

保存したリソース プール ツリーのスナップショットはリストアできます。

前提条件

- vSphere DRS はオンにする必要があります。
- スナップショットは、それを取得したクラスタだけにリストアできます。
- それ以外のリソース プールはクラスタにありません。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 クラスタを右クリックし、[リソース プール ツリーのリストア] を選択します。
- 3 [参照] をクリックし、ローカル マシンにスナップショット ファイルを格納した場所を選択します。
- 4 [開く] をクリックします。
- 5 [OK] をクリックして、リソース プール ツリーをリストアします。

vSAN ストレッチ クラスタの DRS 認識

vSphere 7.0 U2 を使用して DRS を有効にした場合は、ストレッチ クラスタで vSAN ストレッチ クラスタの DRS 認識機能を使用できます。vSAN ストレッチ クラスタには読み取りの局所性があり、仮想マシンはローカル サイトからデータを読み取ります。リモート サイトからデータを読み取ると、仮想マシンのパフォーマンスが影響を受ける可能性があります。vSphere 7.0 U2 より前のリリースでは、DRS は vSAN ストレッチ クラスタの読み取り局所性を認識せず、読み取り局所性がないリモート サイトに仮想マシンを誤って配置することがありました。vSAN ストレッチ クラスタの DRS 認識機能により、DRS は仮想マシンの読み取り局所性を完全に認識するようになったため、仮想マシンは読み取り局所性に完全に対応するサイトに配置されます。これは自動で行われます。構成可能なオプションはありません。vSAN ストレッチ クラスタの DRS 認識は、既存のアフィニティ ルールと連携します。この機能は vSphere 7.0 U2 および VMware Cloud on AWS で動作します。

vSphere HA および vSphere DRS を使用する vSAN ストレッチ クラスタでは、回復性を実現するために、2 つのフォルト ドメインに 2 つのデータ コピーを分散し、障害に備えて 3 番目のフォルト ドメインに監視ノードを配置します。2 つのアクティブなフォルト ドメインでデータのレプリケーションが提供されるため、両方のフォルト ドメインに現在のデータ コピーを保持できます。

vSAN ストレッチ クラスタには、2 つのフォルト ドメイン内でのワークロードの移動を自動化する方法が用意されています。サイト全体で障害が発生した場合、vSphere HA は仮想マシンをセカンダリ サイトで再起動します。これにより、重要な本番ワークロードのダウンタイムを確実に回避できます。プライマリ サイトがオンラインに戻ると、DRS は直ちに、ソフト アフィニティ ホストが配置されたプライマリ サイトに仮想マシンをリバランスします。このプロセスにより、仮想マシンは、仮想マシン データ コンポーネントがまだ再構築中である場合もセカンダリ サイトから読み取りおよび書き込みを行うため、仮想マシンのパフォーマンスが低下する可能性があります。

vSphere 7.0 U2 より前のリリースでは、プライマリ サイトへの再同期中に仮想マシンが移行されるのを回避するために、DRS を完全自動化モードから部分的な自動化モードに変更することを推奨します。再同期が完了した後にのみ、DRS を完全自動化に戻してください。

vSphere 7.0 U2 では、vSAN ストレッチ クラスタの DRS 認識機能により、vSAN ストレッチ クラスタの障害からリカバリするための読み取り局所性ソリューションが完全に自動化されています。読み取り局所性の情報は、仮想マシンが完全にアクセスできるホストを示します。DRS は、仮想マシンを vSAN ストレッチ クラスタのホストに配置するときに、この情報を使用します。サイト リカバリ フェーズ中も vSAN の再同期が進行している場合、DRS は仮想マシンがプライマリ サイトにフェイルバックするのを防止します。データ コンポーネントが完全な読み取り局所性を実現している場合、DRS は仮想マシンをプライマリ アフィニティ サイトに自動的に移行して戻します。これにより、サイト全体で障害が発生した場合に、DRS を完全自動モードで操作できます。

サイトの一部に障害が発生した場合に、許容される障害の数以上の数のデータ コンポーネントが失われたために仮想マシンの読み取り局所性が失われると、vSphere DRS は、読み取りバンド幅を大量に使用する仮想マシンを特定して、セカンダリ サイトにリバランスします。これにより、サイトの一部に障害が発生した場合に、読み取り負荷が高いワークロードを処理する仮想マシンの負担が軽減されることはありません。プライマリ サイトがオンラインに戻り、データ コンポーネントの再同期が完了すると、仮想マシンはアフィニティ サイトに戻されます。

ROBO Enterprise ライセンスを保持している場合の DRS メンテナンスモードの機能

12

vSphere 6.7 U2 で VMware の大規模な Remote Office Branch Office (ROBO) Enterprise ライセンスを使用している場合は、ホストがメンテナンスモードになったときに仮想マシンの自動退避がサポートされます。

ROBO Enterprise クラスタでは DRS がデフォルトで無効になっているため、DRS 設定を変更できません。ROBO Enterprise クラスタ内のホストがメンテナンスモードになると、仮想マシンは DRS によってホストから自動的に退避されます。ホストから仮想マシンを退避する前に、DRS は仮想マシンとホスト間のアフィニティ マッピングを作成して、仮想マシンが配置されていた場所を追跡します。ホストのメンテナンスモードが終了すると、そのホストで実行されていた仮想マシンはホストに移行されます。ホストが移行されると、仮想マシンとホスト間のアフィニティ マッピングはクリアされます。

この章には、次のトピックが含まれています。

- ROBO Enterprise ライセンスでの DRS メンテナンスモードの制限事項
- ROBO Enterprise ライセンスでの DRS メンテナンスモードの使用
- ROBO Enterprise ライセンスでの DRS メンテナンスモードのトラブルシューティング

ROBO Enterprise ライセンスでの DRS メンテナンスモードの制限事項

ROBO Enterprise ライセンスでの DRS 機能は、DRS のフル機能ではありません。

ROBO Enterprise クラスタでメンテナンスモードを開始する場合は、事前にいくつかの制限事項を把握しておく必要があります。ROBO Enterprise クラスタでは、DRS はデフォルトでは無効です。DRS 対応ライセンスから ROBO Enterprise ライセンスに移行した場合、システムにはアフィニティ ルールまたは非アフィニティ ルールを伴う仮想マシンがあることがあります。アフィニティ ルールまたは非アフィニティ ルールを伴う仮想マシンは無効にするか、削除する必要があります。そうしないと、ROBO Enterprise のメンテナンスモード操作は無効になります。DRS が完全自動化モードに設定されていない場合、ROBO Enterprise のメンテナンスモード操作は無効です。ホストのメンテナンスワークフローを介して仮想マシンを自動的に退避させるには、DRS 自動化レベルを完全自動化モードに設定する必要があります。仮想マシンが DRS 完全自動化モードをオーバーライドした場合は、仮想マシンを手動で退避させる必要があります。

ROBO Enterprise ライセンスでの DRS メンテナンスモードの使用

vSphere 6.7 U2 では、ROBO Enterprise ライセンスで一部の DRS メンテナンスモード機能を利用できます。

前提条件

- クラスタ内のすべてのホストに ROBO Enterprise ライセンスがインストールされているかどうかを確認します。インストールされていない場合は、ライセンスをインストールする必要があります。
- DRS ルールを設定して有効にしているかどうかを確認します。有効な場合は無効にするか削除して、ROBO Enterprise メンテナンス モードで操作を行う必要があります。

手順

- 1 ROBO Enterprise ライセンスで DRS メンテナンス モードを使用するには、クラスタの各ホストに ROBO Enterprise ライセンスがインストールされていることを確認します。
 - ライセンスがインストールされていない場合は、手順 2 に進みます。
 - ライセンスがインストールされている場合は、手順 3 に進みます。
- 2 ROBO Enterprise ライセンスをインストールします。
 - a vSphere Client で、ホストを参照して移動します。
 - b [構成] タブで [ライセンス] を選択します。
 - c [ライセンスの割り当て] をクリックします。
 - d ROBO Enterprise ライセンス キーを入力して、[OK] をクリックします。

クラスタ内のすべてのホストで、この手順を繰り返す必要があります。
- 3 クラスタ内のホストを選択し、右クリックして [メンテナンス モードへの切り替え] を選択し、[OK] をクリックします。

ホスト上の仮想マシンを自動的に退避させます。

結果

ホストでメンテナンス モードを終了すると、仮想マシンはホストに自動的に移行されます。ホストは、元の状態にリストアされます。ただし、ホストに負荷がかかりすぎる場合、DRS は仮想マシンを元のホストに移行できません。DRS はホストを元の状態に復旧しますが、ホストを過負荷にすることはありません。

次のステップ

ROBO Enterprise ライセンスを使用して DRS メンテナンス モードを無効にする必要がある場合は、vpxd.cfg ファイルを編集します。vpxd.cfg ファイルを開きます。<クラスタ> オプションの <roboMMEnabled>true</roboMMEnabled> を <roboMMEnabled>>false</roboMMEnabled> に変更します。これは、ランタイムの設定であるため、設定の更新後に vpxd を再起動する必要はありません。

ROBO Enterprise ライセンスでの DRS メンテナンス モードのトラブルシューティング

ROBO Enterprise クラスタでメンテナンス モードを使用しているときに問題が発生した場合は、以下を確認してください。

以下のことを確認して、ROBO Enterprise クラスタでメンテナンス モードが正しく機能するようにします。

- クラスタ内のすべてのホストに ROBO Enterprise ライセンスがインストールされているかどうかを確認します。インストールされていない場合は、ライセンスをインストールする必要があります。
- DRS ルールを設定して有効にしているかどうかを確認します。有効な場合は無効にするか削除して、ROBO Enterprise メンテナンス モードで操作を行う必要があります。
- 互換性チェックに失敗した場合は、他のホストが仮想マシンと互換性があるかどうかを確認します。

DRS クラスタを作成した後で、それをカスタマイズできます。また、それを使用してリソースを管理できます。

DRS クラスタおよびそれに含まれるリソースをカスタマイズするには、アフィニティ ルールを構成し、ホストおよび仮想マシンを追加または削除します。クラスタの設定とリソースを定義したら、そのクラスタが有効なクラスタであり、有効であり続けるようにする必要があります。また、有効な DRS クラスタを使用すると、電力リソースを管理し、vSphere HA と相互運用することもできます。

注： この章では、「メモリ」は物理 RAM または永続的なメモリを指します。

この章には、次のトピックが含まれています。

- クラスタへのホストの追加
- クラスタへの仮想マシンの追加
- クラスタからの仮想マシンの削除
- クラスタからのホストの削除
- DRS クラスタの妥当性
- 電力リソースの管理
- DRS アフィニティ ルールの使用

クラスタへのホストの追加

クラスタにホストを追加する手順は、同じ vCenter Server によって管理されているホスト（管理対象ホスト）と、そのサーバによって管理されていないホストとで異なります。

ホストを追加すると、そのホストにデプロイされている仮想マシンはクラスタの一部になり、DRS は、一部の仮想マシンをクラスタ内の別のホストへ移行するよう推奨する場合があります。

注： vSphere DRS は、vSphere の重要な機能で、vSphere クラスタ内で実行されるワークロードの健全性の維持に必要です。vSphere 7.0 Update 1 以降では、DRS は vCLS 仮想マシンの可用性に依存します。詳細については [vSphere クラスタ サービス](#) を参照してください。

クラスタへの管理対象ホストの追加

すでに vCenter Server によって管理されているスタンドアロン ホストを DRS クラスタに追加すると、そのホストのリソースがクラスタに関連付けられます。

ユーザーは、既存の仮想マシンとリソース プールをクラスタのルート リソース プールと関連付けるか、それともリソース プール階層を移動するかを決定できます。

注： ホストに子のリソース プールまたは仮想マシンがない場合、ホストのリソースがクラスタに追加されますが、最上位レベルのリソース プールを持つリソース プール階層は作成されません。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 ホストを右クリックして [移動先...] を選択します。
- 3 クラスタを選択します。
- 4 [OK] をクリックして変更内容を保存します。
- 5 ホストの仮想マシンとリソース プールをどうするかを選択します。
 - [このホストの仮想マシンをクラスタのルート リソース プールに入れます]

vCenter Server はホストの既存のリソース プールをすべて削除し、ホストの階層内にあるすべての仮想マシンは、ルートに接続されます。シェア割り当てはリソース プールに対して相対的なもので、リソース プール階層を破棄するこのオプションを選択したあとで、手動で仮想マシンのシェアを変更する必要が生じることがあります。

- [このホストの仮想マシンとリソース プール用のリソース プールを作成します]

vCenter Server は、クラスタの直接の子になる最上位レベルのリソース プールを作成し、ホストのすべての子とその新規リソース プールに追加します。この新しい最上位レベルのリソース プールの名前を指定できます。デフォルトは、[<host_name> から移動] です。

結果

ホストがクラスタに追加されます。

管理されていないホストのクラスタへの追加

管理されていないホストをクラスタに追加できます。現在、そのクラスタと同じ vCenter Server システムによって管理されておらず、vSphere Client に認識されていないホストを追加できます。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 クラスタを右クリックし、[ホストの追加] を選択します。
- 3 ホスト名、ユーザー名、およびパスワードを入力して、[次へ] をクリックします。
- 4 概要情報を確認して、[次へ] をクリックします。
- 5 既存のライセンス キーまたは新しいライセンス キーを割り当てて、[次へ] をクリックします。

- 6 (オプション) ロックダウン モードを有効にして、リモート ユーザーがホストに直接ログインすることを防止できます。

ロックダウン モードを有効にしない場合、ホスト設定のセキュリティ プロファイルを編集して、このオプションを設定できます。

- 7 ホストの仮想マシンとリソース プールをどうするかを選択します。

- [このホストの仮想マシンをクラスタのルート リソース プールに入れます]

vCenter Server はホストの既存のリソース プールをすべて削除し、ホストの階層内にあるすべての仮想マシンは、ルートに接続されます。シェア割り当てはリソース プールに対して相対的なので、リソース プール階層を破棄するこのオプションを選択したあとで、手動で仮想マシンのシェアを変更する必要が生じることがあります。

- [このホストの仮想マシンとリソース プール用のリソース プールを作成します]

vCenter Server は、クラスタの直接の子になる最上位レベルのリソース プールを作成し、ホストのすべての子とその新規リソース プールに追加します。この新しい最上位レベルのリソース プールの名前を指定できます。デフォルトは、[<host_name> から移動] です。

- 8 設定内容を確認して、[終了] をクリックします。

結果

ホストがクラスタに追加されます。

クラスタへの仮想マシンの追加

仮想マシンは多数の方法でクラスタに追加できます。

- クラスタにホストを追加すると、そのホスト上のすべての仮想マシンがクラスタに追加されます。
- 仮想マシンを作成すると、[新規仮想マシン] ウィザードは、その仮想マシンを配置する場所を尋ねてきます。スタンドアロン ホストまたはクラスタを選択でき、そのホストまたはクラスタの内部にある任意のリソース プールを選択できます。
- [仮想マシンの移行] ウィザードを使用して、仮想マシンをスタンドアロン ホストからクラスタに、またはあるクラスタから別のクラスタに移行できます。このウィザードを起動するには、仮想マシン名を右クリックして、[移行] を選択します。

仮想マシンのクラスタへの移動

仮想マシンはクラスタに移動できます。

手順

- 1 vSphere Client で仮想マシンを検索します。
 - a データセンター、フォルダ、クラスタ、リソース プール、またはホストを選択して、仮想マシンを検索します。
 - b [仮想マシン] タブをクリックします。

- 2 仮想マシンを右クリックし、[移動先...] を選択します。
- 3 クラスタを選択します。
- 4 [OK] をクリックします。

クラスタからの仮想マシンの削除

仮想マシンをクラスタから削除できます。

次の 2 つの方法で、仮想マシンをクラスタから削除できます。

- ホストをクラスタから削除すると、ほかのホストに移行していないパワーオフ状態の仮想マシンもすべて削除されます。ホストを削除できるのは、ホストがメンテナンス モードになっているか、切断されている場合だけです。DRS クラスタからホストを削除した場合、クラスタは、オーバーコミットされるために黄色で表示されることがあります。
- [移行]ウィザードを使用して、仮想マシンをクラスタからスタンドアロン ホストに、またはあるクラスタから別のクラスタに移行できます。ウィザードを起動するには、仮想マシン名を右クリックして、[[移行]] を選択します。

クラスタからの仮想マシンの移動

仮想マシンはクラスタから移動できます。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
 - a データセンター、フォルダ、クラスタ、リソース プール、またはホストを選択して、仮想マシンを検索します。
 - b [仮想マシン] タブをクリックします。
- 2 仮想マシンを右クリックして [移行] を選択します。
- 3 [データストアの変更] を選択し、[次へ] をクリックします。
- 4 データストアを選択して、[次へ] をクリックします。
- 5 [終了] をクリックします。

仮想マシンが DRS クラスタのルール グループのメンバーである場合、vCenter Server は、移行の続行を許可する前に警告を表示します。この警告は、従属仮想マシンの移行が自動的に実行されないことを示します。移行を続行するには、警告を了承する必要があります。

クラスタからのホストの削除

DRS クラスタからホストを削除すると、リソース プールの階層や仮想マシンに影響を与え、また無効なクラスタを作成する恐れがあります。影響を受けるオブジェクトを考慮してから、ホストを削除します。

- リソース プールの階層：ホストをクラスタから削除すると、ホストをクラスタに追加した時点で DRS クラスタを使用してホストのリソース プールを移動した場合でも、ホストが保持するのはルート リソース プールだけになります。この場合、階層はクラスタで保持されます。ホスト固有のリソース プール階層を作成できます。

注：最初にホストをメンテナンス モードにすることで、クラスタからホストを確実に削除できます。メンテナンス モードにせずに、クラスタからホストを削除する前にホストを切断すると、ホストはクラスタ階層を反映したリソース プールを保持します。

- 仮想マシン：クラスタから削除する前にホストをメンテナンス モードにしておく必要があります。ホストをメンテナンス モードにするには、パワーオン状態の仮想マシンをすべてそのホストから移行する必要があります。また、ホストをメンテナンス モードに切り替えることを要求する場合、そのホストにあるパワーオフ状態のすべての仮想マシンを、クラスタ内の別のホストに移行するか尋ねられます。
- 無効なクラスタ：クラスタからホストを削除すると、クラスタで使用可能なリソースが減少します。クラスタ内のすべての仮想マシンとリソース プールの予約を満たすのに十分なリソースがまだクラスタにある場合、クラスタは、リソース割り当てを調整して、減少後のリソース量を反映させます。クラスタに、すべてのリソース プールの予約を満たすのに十分なリソースはないが、すべての仮想マシンの予約を満たす十分なリソースはある場合は、アラームが発行され、クラスタが黄色で表示されます。DRS は、実行を続けます。

ホストをメンテナンス モードに切り替える

メモリの増設など、ホストの保守作業を行う必要がある場合は、ホストをメンテナンス モードにします。ホストは、ユーザーの要求があった場合のみ、メンテナンス モードを開始または終了します。

メンテナンス モードに切り替えるホストで実行中の仮想マシンは、別のホストに（手動で、または DRS によって自動的に）移行するか、シャットダウンする必要があります。ホストは、稼働中のすべての仮想マシンがパワーオフされるか別のホストへ移行されるまで、[メンテナンス モードに切り替えています] と表示された状態になります。メンテナンス モードのホストで仮想マシンをパワーオンしたり、そのホストに仮想マシンを移行したりすることはできません。

ホスト上に実行中の仮想マシンがなくなると、ホストのアイコンが変化して [メンテナンス中] が組み込まれ、ホストのサマリ パネルに新しい状態が表示されます。メンテナンス モードの間は、そのホストでは、仮想マシンをデプロイしたり、パワーオンしたりすることはできません。

注：DRS は、要求されたモードにホストが切り替わると vSphere HA のフェイルオーバー レベルに違反することになる場合、メンテナンス モードまたはスタンバイ モードに切り替わるホストから仮想マシンが移行することを推奨（完全自動モードの場合は実行）しません。

手順

- 1 vSphere Client で、ホストを参照して移動します。

- 2 ホストを右クリックして [メンテナンス モード] - [メンテナンス モードへの切り替え] を選択します。
 - ホストが一部自動化された、または手動の DRS クラスタの一部である場合は、[クラスタ] - [監視] - [DRS] - [推奨]の順に移動して、[推奨事項の適用] をクリックします。
 - ホストが、自動化された DRS クラスタの一部になっている場合、メンテナンス モードに切り替わると、仮想マシンは別のホストに移行します。
- 3 該当する場合は、[はい] をクリックします。

結果

[メンテナンス モード] - [メンテナンス モードの終了] を選択するまで、ホストはメンテナンス モードです。

クラスタからのホストの削除

ホストをクラスタから削除できます。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 ホストを右クリックして [メンテナンス モード] - [メンテナンス モードへの切り替え] を選択します。

ホストがメンテナンス モードのとき、そのホストを別のインベントリの場所（最上位レベルのデータセンター、または別のクラスタ）に移動します。
- 3 ホストを右クリックして [移動先...] を選択します。
- 4 ホストの新しい移動先を選択して、[OK] をクリックします。

結果

ホストを移動すると、そのホストのリソースがクラスタから削除されます。ホストのリソース プール階層をクラスタに移動した場合、その階層はクラスタで保持されます。

次のステップ

ホストをクラスタから削除後、次のタスクを実行できます。

- vCenter Server から、そのホストを削除します。
- そのホストをスタンドアロン ホストとして vCenter Server の下で稼働させます。
- ホストを別のクラスタに移動できます。

スタンバイ モードの使用

ホスト マシンがスタンバイ モードになると、パワーオフされます。

通常、電力の使用を最適にするために、ホストは vSphere DPM 機能によってスタンバイ モードになります。手動でホストをスタンバイ モードにすることもできます。ただし、次に実行するときに、DRS が変更を取り消す（または取り消すように推奨する）ことがあります。ホストを強制的にオフのままにするには、メンテナンス モードにして、パワーオフします。

DRS クラスタの妥当性

vSphere Client には、DRS クラスタが有効か、オーバーコミットされているか（黄色）、または無効（赤）であるかが示されます。

DRS クラスタは、いくつかの理由でオーバーコミットまたは無効になります。

- クラスタは、ホストに障害が起きた場合にオーバーコミットされる可能性があります。
- vCenter Server を使用できない場合に、vSphere Client を使用して仮想マシンをパワーオンにすると、クラスタは無効になります。
- クラスタは、仮想マシンのフェイルオーバーの最中にユーザーが親リソース プール上で予約を減らした場合に、無効になります。
- vCenter Server を使用できないときに、vSphere Client を使用してホストまたは仮想マシンに変更を行う場合、それらの変更は有効になります。しかし、vCenter Server が再び使用可能になると、クラスタ要件がもはや満たされていないためにクラスタが赤または黄色で表示されることがあります。

クラスタの妥当性のシナリオについて考える場合、次の用語を理解する必要があります。

予約

ユーザーが入力した、固定で保証されたリソース プールの割り当て。

使用済みの予約

それぞれの子リソース プールの予約または使用済みの予約（いずれか大きな方）の合計。再帰的に加算されます。

予約なし

負でない値で、リソース プールのタイプによって異なります。

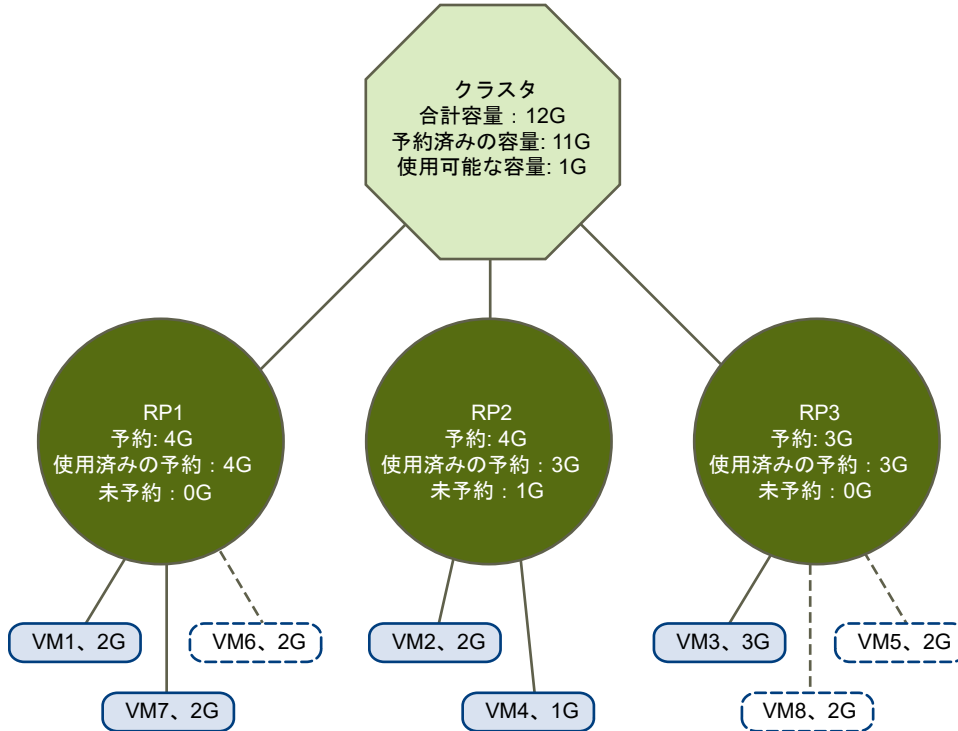
- 拡張不能なリソース プール： 予約 - 使用済み予約
- 拡張可能なリソース プール（予約 - 使用済み予約） + 先祖のリソース プールから借用できる未予約リソース

有効な DRS クラスタ

有効なクラスタには、すべての予約を満たして、実行中のすべての仮想マシンをサポートするのに十分なリソースがあります。

次図に、固定リソース プールを持つ有効なクラスタの例と、その CPU およびメモリ リソースの計算方法を示します。

図 13-1. 固定リソース プールを持つ有効なクラスタ

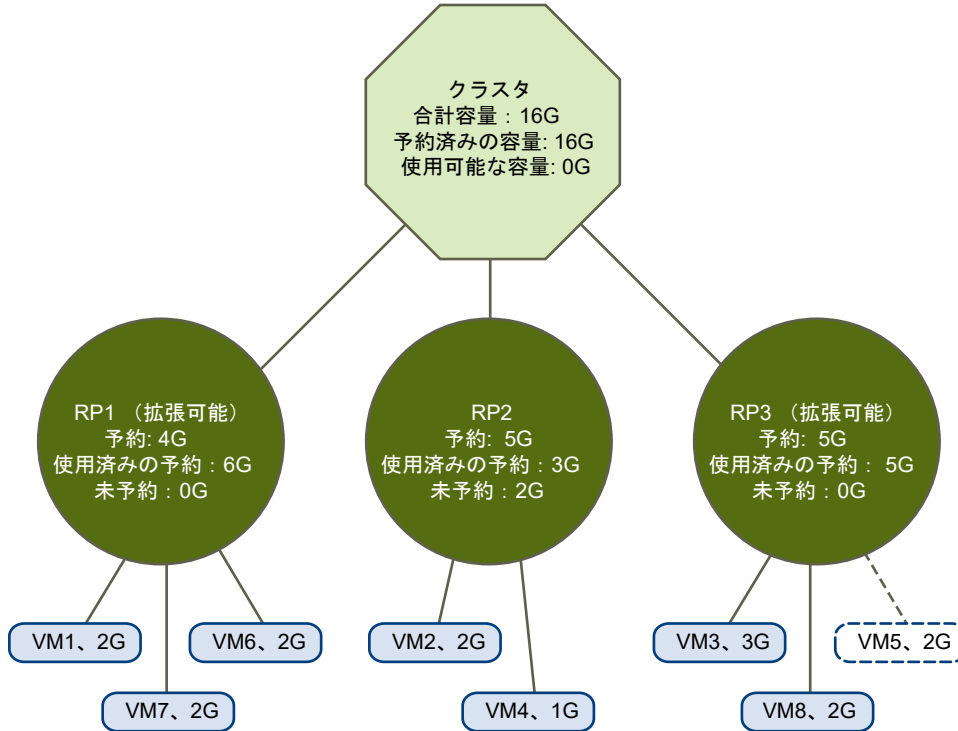


このクラスタの特性は次のとおりです。

- クラスタのリソースは、合計 12GHz です。
- それぞれタイプが [固定] ([拡張可能な予約] が選択されていない) のリソース プールが 3 つあります。
- 3 つのリソース プールをあわせた予約の合計は、11GHz (4+4+3GHz) です。合計は、クラスタの [予約済みの容量] フィールドに示されます。
- RP1 は、4GHz の予約を使用して作成されています。それぞれ 2GHz の 2 台の仮想マシン (VM1 および VM7) がパワーオンされます ([使用済みの予約] : 4GHz)。さらに仮想マシンをパワーオンするリソースは、残っていません。VM6 は、パワーオンされていないものとして表示されます。VM6 は予約を消費しません。
- RP2 は、4GHz の予約を使用して作成されています。1GHz と 2GHz の 2 台の仮想マシンがパワーオンされています ([使用済みの予約] : 3GHz)。1GHz が未予約で残っています。
- RP3 は、3GHz の予約を使用して作成されています。3GHz を持つ 1 台の仮想マシンがパワーオンされています。さらに仮想マシンをパワーオンするのに使用可能なリソースはありません。

次図に、予約タイプ [拡張可能]を使用した一部のリソース プール (RP1 および RP3) を持つ有効なクラスタの例を示します。

図 13-2. 拡張可能なリソース プールを持つ有効なクラスタ



有効なクラスタを次のように構成できます。

- クラスタのリソースは、合計 16GHz です。
- RP1 および RP3 はタイプ [拡張可能] で、RP2 はタイプ固定です。
- 3 つのリソース プールを統合した使用済みの予約の合計は、16GHz です (RP1 の 6GHz、RP2 の 5GHz、および RP3 の 5GHz)。16GHz は、最上位レベルのクラスタの [予約済みの容量] として表示されます。
- RP1 は、4GHz の予約を使用して作成されています。それぞれ 2GHz の 3 台の仮想マシンがパワーオンされています。これらの仮想マシンのうちの 2 台 (たとえば、VM1 と VM7) は、RP1 の予約を使用でき、3 台目の仮想マシン (VM6) は、クラスタのリソース プールから予約を使用できます。(このリソース プールのタイプが [固定] の場合、追加の仮想マシンのパワーオンはできません)。
- RP2 は、5GHz の予約を使用して作成されています。1GHz と 2GHz の 2 台の仮想マシンがパワーオンされています ([使用済みの予約]: 3GHz)。2GHz が未予約で残っています。

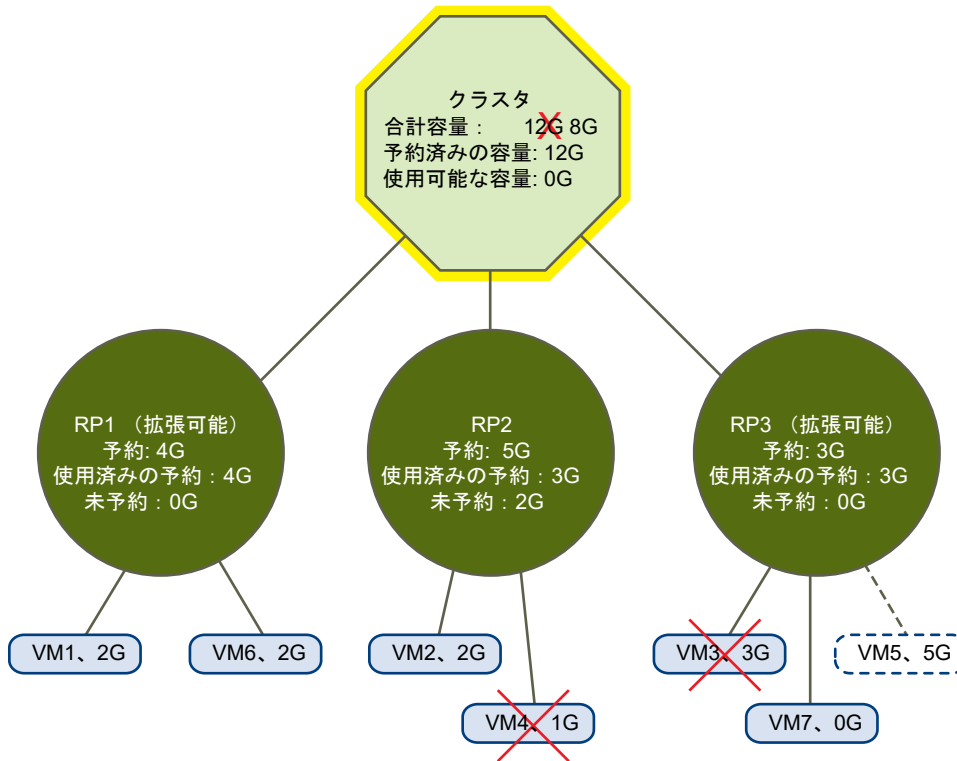
RP3 は、5GHz の予約を使用して作成されています。3GHz と 2GHz の 2 台の仮想マシンがパワーオンされています。このリソース プールのタイプが [拡張可能] であっても、親の余分なリソースがすでに RP1 によって使用されているので、追加の 2GHz の仮想マシンをパワーオンすることはできません。

オーバーコミットされた DRS クラスタ

クラスタがオーバーコミット (黄色) で表示されるのは、リソース プールと仮想マシンのツリーが内部的に整合性があるが、子リソース プールによって予約されるすべてのリソースをサポートする容量がクラスタにない場合です。

ホストが使用不能になると、そのすべての仮想マシンが使用不能になるので、実行中のすべての仮想マシンをサポートするのに十分なリソースが常に存在するようになります。通常、クラスタは、クラスタの容量が突然減少した場合、たとえば、クラスタ内のホストが使用不能になった場合に黄色になります。クラスタが黄色にならないように、十分な追加クラスタ リソースを残すことを推奨します。

図 13-3. 黄色のクラスタ



この例では、次のようになります。

- クラスタのリソースは、それぞれ 4GHz の 3 台のホストから合計 12GHz です。
- 3 つのリソース プールが合計で 12GHz を予約します。
- 統合した 3 つのリソース プールによって使用される予約の合計は、12GHz (4+5+3 GHz) です。これは、クラスタ内の [予約済みの容量] として表示されます。
- 4GHz ホストの 1 台が使用不能になり、合計リソースが 8GHz に減ります。
- 同時に、障害の起きたホスト上で実行されていた VM4 (1GHz) および VM3 (3GHz) は、すでに実行されていません。
- クラスタは、合計で 6GHz を必要とする仮想マシンを実行中です。クラスタではまだ 8GHz が使用可能であり、これは、仮想マシンの要件を満たすのに十分です。

リソース プールの 12GHz の予約を満たすことはすでにできないので、クラスタは黄色で表示されます。

無効な DRS クラスタ

DRS が有効に設定されているクラスタは、ツリーが内部的に整合性がない、すなわち、リソース制約が遵守されていない場合に無効 (赤) になります。

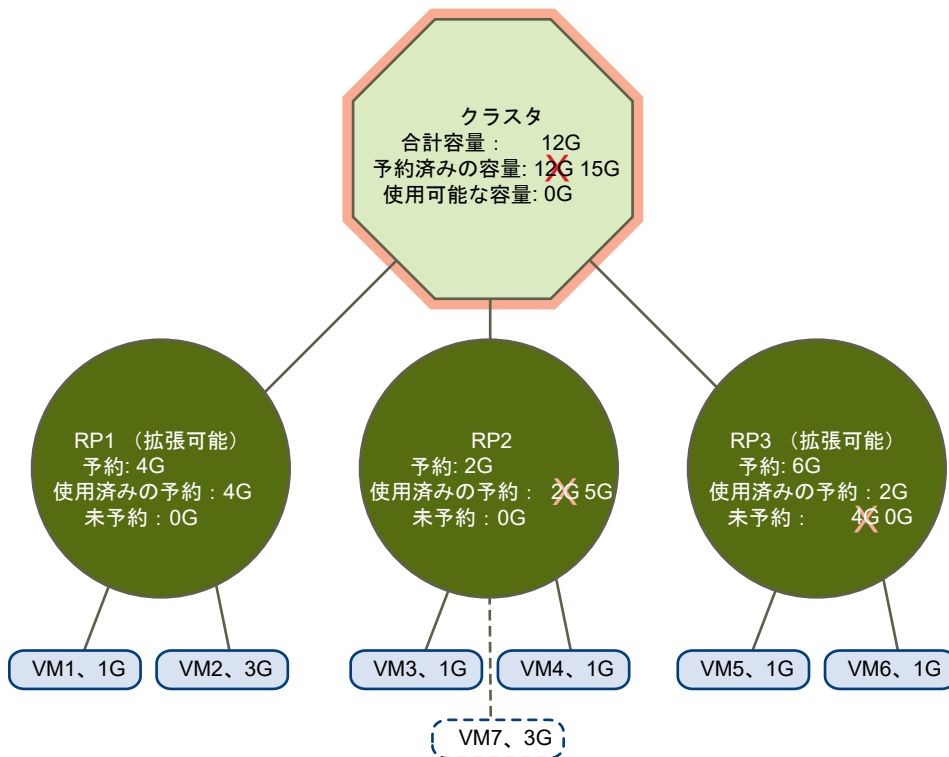
クラスタ内のリソースの合計量は、クラスタが赤になるかどうかに影響を与えません。クラスタは、ルートレベルに十分なリソースが存在しても、子レベルで不整合があると、赤で表示される場合があります。

赤の DRS クラスタの問題は、1 台以上の仮想マシンをパワーオフするか、十分なリソースがあるツリーの一部に仮想マシンを移動するか、あるいは赤い部分のリソース プール設定を編集することによって解決できます。一般的に、リソースの追加は、黄色の状態のときにのみ役立ちます。

仮想マシンのフェイルオーバーの最中にリソース プールを再構成した場合も、クラスタが赤で表示されることがあります。フェイルオーバー中の仮想マシンは、切断され、親リソース プールによって使用される予約の計算に含まれません。フェイルオーバーが完了する前に、親リソース プールの予約を減らすことができます。フェイルオーバーが完了したら、仮想マシンのリソースが親リソース プールに再びチャージされます。プールの使用量が新しい予約よりも多くなった場合、クラスタは赤になります。

ユーザーがリソース プール 2 の下で、3GHz の予約で仮想マシンを（サポートされていない方法で）起動した場合、次の図に示すようにクラスタは赤になります。

図 13-4. 赤のクラスタ



電力リソースの管理

vSphere DPM (Distributed Power Management) の機能を使用すると、クラスタのリソース使用状況に応じてホストをパワーオンおよびパワーオフするため、DRS クラスタの電力消費量を削減できます。

vSphere DPM は、クラスタにあるすべての仮想マシンのメモリ リソース、および CPU リソースの度重なる需要を監視し、この需要と、クラスタのすべてのホストで使用可能なリソース キャパシティの総量とを比較します。余剰キャパシティが十分にある場合、vSphere DPM は 1 台以上のホストをスタンバイ モードにし、それらの仮想マシンを別のホストに移行したあとでパワーオフします。逆に、容量が不足しそうな場合、DRS はホストのスタンバイモードを終了し（パワーオンし）、vMotion を使用して仮想マシンをそれらのホストに移行します。これらの計算を行う場合、vSphere DPM は現在の需要だけではなく、ユーザーが指定した仮想マシンのリソース予約についても考慮します。

DRS クラスタを作成するときに [予測メトリック] を有効にすると、選択するローリング予測ウィンドウに応じて、DPM が事前にプロポーザルを発行します。

注： ESXi ホストは、vCenter Server が管理するクラスタ内で実行されている場合を除き、自動的にスタンバイモードを終了することはできません。

vSphere DPM では、次の 3 つの電源管理プロトコルを使用して、ホストのスタンバイ モードを終了させることができます。IPMI (Intelligent Platform Management Interface)、Hewlett-Packard iLO (Integrated Lights-Out)、または WOL (Wake-On-LAN) です。各プロトコルで独自のハードウェア サポートおよび構成が必要です。ホストがこれらのプロトコルのいずれもサポートしていない場合、vSphere DPM はそのホストをスタンバイ モードにできません。ホストが複数のプロトコルをサポートしている場合に使用される順番は、IPMI、iLO、WOL です。

注： 初回のパワーオンを行わずに、スタンバイ モードのホストを切断したり、DRS クラスタの外に移動したりしないでください。vCenter Server はそのホストを再びパワーオンできなくなります。

vSphere DPM 用の IPMI 設定または iLO 設定の構成

IPMI はハードウェア レベルの仕様であり、Hewlett-Packard iLO は組み込みのサーバ管理テクノロジーです。どちらも、コンピュータをリモートで監視および制御するためのインターフェイスを記述および提供します。

各ホストで次の手順を行う必要があります。

前提条件

IPMI および iLO は、ハードウェア制御機能にアクセスするためのゲートウェイを提供するためにハードウェア Baseboard Management Controller (BMC) を必要とします。また、これらはそのインターフェイスにシリアル接続または LAN 接続を使用してリモート システムからアクセスできるようにします。ホスト自体がパワーオフされている場合でも、BMC はパワーオンされた状態です。適切に有効にすると、BMC はリモート パワーオン コマンドに応答できます。

IPMI または iLO を Wake プロトコルとして使用する場合は、BMC を構成する必要があります。BMC の構成手順は、モデルによって異なります。詳細については、ベンダーのドキュメントを参照してください。IPMI を使用する場合は、BMC LAN チャネルが常時使用可能で、オペレータ特権コマンドを許可するように構成する必要もあります。一部の IPMI システムでは、「IPMI over LAN」を有効にする場合は、BIOS 内でこれを構成し、特定の IPMI アカウントを指定する必要があります。

IPMI のみを使用している vSphere DPM は、MD5 認証およびプレーンテキスト ベースの認証をサポートしていますが、MD2 ベースの認証はサポートされていません。MD5 がサポートされており、オペレータ ロールでそれを使用可能であることをホストの BMC が報告している場合、vCenter Server は MD5 を使用します。または、プレーンテキスト ベースの認証がサポートされており、使用可能であることを BMC が報告している場合、プレーンテキスト ベースの認証を使用します。MD5 認証もプレーンテキスト認証も使用できない場合、ホストで IPMI を使用できません。vCenter Server は Wake-on-LAN を使用しようとします。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [システム] の [電源管理] をクリックします。
- 4 [編集] をクリックします。
- 5 以下の情報を入力します。
 - BMC アカウントのユーザー名とパスワード（このユーザー名は、リモート側からホストをパワーオンできる必要があります）。
 - BMC に関連付けられている NIC の IP アドレス（ホストの IP アドレスとは異なるため）。この IP アドレスは、静的アドレスまたはリース期間が無制限の DHCP アドレスにします。
 - BMC に関連付けられている NIC の MAC アドレス。
- 6 [OK] をクリックします。

vSphere DPM 用の Wake-On-LAN のテスト

vSphere DPM 機能に対する Wake-On-LAN (WOL) の使用は、この機能を VMware ガイドラインに従って構成し、テストが成功した場合、完全にサポートされます。クラスタの vSphere DPM をはじめて有効にする前に、または、vSphere DPM を使用しているクラスタに追加された任意のホストで、次の手順を実行する必要があります。

前提条件

WOL をテストする前に、クラスタが前提条件を満たしていることを確認します。

- クラスタに少なくとも 2 台の ESX 3.5（または ESX 3i バージョン 3.5）以降のホストが含まれている必要があります。
- 各ホストの vMotion ネットワーク リンクが正しく機能している必要があります。vMotion ネットワークは、ルータによって分離された複数のサブネットではなく、単一の IP サブネットであることが必要です。
- 各ホスト上の vMotion NIC は、WOL をサポートしている必要があります。WOL のサポートを確認するには、はじめに VMkernel ポートに対応する物理ネットワーク アダプタの名前を突き止めます。vSphere Client のインベントリ パネルでホストを選択し、[構成] タブを選択して、[ネットワーク] をクリックします。この情報を取得したあと、[ネットワーク アダプタ] をクリックし、そのネットワーク アダプタに対応する登録情報を探します。関連するアダプタの [Wake on LAN 対応] 列が「はい」になっているはずです。

- ホスト上の各 NIC の WOL 互換ステータスを表示するには、vSphere Client のインベントリ パネルでホストを選択し、[構成] タブを選択してから、[ネットワーク アダプタ] をクリックします。[Wake on LAN 対応] 列で、NIC が「はい」になっている必要があります。
- WOL 対応の各 vMotion NIC が接続されているスイッチ ポートは、リンク速度が固定速度（1,000 Mb/s など）ではなく、自動ネゴシエートするように設定されている必要があります。多くの NIC は、ホストがパワーオフ状態のときに 100 Mb/s 以下に切り替えることができる場合にのみ、WOL をサポートします。

これらの前提条件を確認したあと、vSphere DPM のサポートに WOL を使用する予定の各 ESXi ホストをテストします。これらのホストをテストするときは、vSphere DPM 機能がクラスタに対して無効になっていることを確認します。

注意: WOL を Wake プロトコルとして使用する vSphere DPM クラスタに追加されるすべてのホストをテストし、テストに合格しないホストは電源管理の使用を無効にしてください。これをしなかった場合、vSphere DPM はあとでパワーオンに戻ることができないホストをパワーオフする可能性があります。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 ホストを右クリックし、[電源] - [スタンバイ モードへの切り替え] を選択します。
この操作により、ホストはパワーオフされます。
- 3 ホストを右クリックし、[電源] - [パワーオン] を選択して、スタンバイ モードから切り替えます。
- 4 ホストが正常にパワーオンしたかどうかを観察します。
- 5 スタンバイ モードから移行できないホストについては、次の手順を実行します。
 - a vSphere Client でホストを選択し、[設定] タブを選択します。
 - b [ハードウェア] - [電源管理] で [編集] をクリックして、電源管理ポリシーを調整します。
これを実行すると、vSphere DPM はそのホストをパワーオフする候補と見なさなくなります。

DRS クラスタ用の vSphere DPM の有効化

各ホスト上で使用している Wake プロトコルに必要な構成手順またはテスト手順を実行した後、vSphere DPM を有効にすることができます。

電源管理自動化レベル、しきい値、およびホスト レベルのオーバーライドを構成します。これらの設定は、クラスタの [設定] ダイアログ ボックスの [電源管理] で構成されます。

[スケジュール設定タスク: クラスタ電源設定の変更]ウィザードを使用して、クラスタの DPM を有効または無効にするスケジュール設定タスクを作成することもできます。

注: DRS クラスタ内のホストに USB デバイスが接続されている場合は、そのホストについて DPM を無効にする必要があります。無効にしないと、DPM によって、そのホストとサーバがオフになり、そのデバイスとそのデバイスを使用していた仮想マシンとの接続が切れることがあります。

自動化レベル

vSphere DPM によって生成されたホストの電源状態および移行の推奨を自動的に実行するかどうかは、各機能に指定された電源管理の自動化レベルによって決まります。

自動化レベルは、クラスタの [設定] ダイアログ ボックスの [電源管理] で構成します。

注： 電源管理の自動化レベルは、DRS の自動化レベルと同じではありません。

表 13-1. 電源管理の自動化レベル

オプション	説明
オフ	機能が無効になり、推奨は生成されません。
手動	ホストの電源操作および関連する仮想マシンの移行に関する推奨が行われますが、それが自動的に実行されることはありません。
自動	関連する仮想マシンの移行をすべて自動的に実行できる場合、ホストの電源操作は自動的に実行されます。

vSphere DPM しきい値

vSphere DPM 機能によって生成される電源状態（ホストのパワーオンまたはオフ）の推奨には、優先順位 1 の推奨から優先順位 5 の推奨までの優先順位が割り当てられます。

これらの優先順位の評価は、DRS クラスタ内で検出された高いまたは低い使用率と、ホストの電源状態の意図的な変更から予想される改善の程度に基づいています。優先順位 1 の推奨は必須である一方、優先順位 5 の推奨はわずかな改善のみをもたらします。

しきい値は、クラスタの設定ダイアログ ボックスの [電力管理] で構成します。vSphere DPM しきい値のスライダを右へ 1 レベル移動するごとに、1 つ下のレベルの優先順位を、自動的に実行される推奨セットに含めたり、または手動で実行する推奨として表示したりできます。保守的設定では、vSphere DPM は優先順位 1 の推奨のみを作成し、その右にある次のレベルは、優先順位 2 以上の推奨を作成します。以後、同様に作成し、積極的レベルは優先順位 5 以上の推奨（つまり、すべての推奨）を作成します。

注： DRS しきい値と vSphere DPM しきい値は、基本的に独立しています。ユーザーは移行の積極性と、そのそれぞれが提供するホストの電源状態の推奨を区別できます。

ホスト レベルのオーバーライド

DRS クラスタで vSphere DPM を有効にすると、デフォルトでクラスタのすべてのホストが vSphere DPM の自動化レベルを継承します。

クラスタの [設定] ダイアログ ボックスの [ホスト オプション] ページを選択し、[電源管理] 設定をクリックすることで、各ホストにあわせてこのデフォルト値をオーバーライドできます。この設定を次のオプションに変更できます。

- 無効
- 手動

- 自動

注： スタンバイ モードの終了テストに失敗したためにホストの電源管理設定が無効になっている場合は、この設定を変更しないでください。

vSphere DPM を有効にして実行したあと、それが正しく機能しているかどうかを確認できます。確認するには、クラスタの設定ダイアログ ボックスの [ホスト オプション] ページ、および各クラスタの [ホスト] タブに表示される、各ホストの [前回のスタンバイ終了] 情報を参照します。このフィールドには、vCenter Server が前回、ホストのスタンバイ モードを終了させようと試みたときのタイム スタンプと、それが成功または失敗のどちらであったかが表示されます。そのような試みが行われなかった場合、このフィールドには実行しないと表示されます。

注： [前回のスタンバイ終了] フィールドの時刻は、vCenter Server イベント ログから生成されます。このログが消去された場合、時刻は実行しないにリセットされます。

vSphere DPM の監視

vCenter Server でイベント ベースのアラームを使用し、vSphere DPM を監視できます。

vSphere DPM の使用時に発生する可能性がある最も深刻なエラーは、DRS クラスタがホストの容量を必要としたときに、ホストがスタンバイ モードを終了できないことです。このエラーが発生していないかどうかを監視するために、事前構成された [スタンバイの終了エラー] アラームを vCenter Server で使用できます。vSphere DPM がホストのスタンバイ モードを終了できない場合（vCenter Server イベント DrsExitStandbyModeFailedEvent）、このアラームを構成して、管理者にアラート E メールを送信するか、SNMP トラップを使用して通知を送信できます。デフォルトでは、このアラームは vCenter Server がそのホストに正常に接続できるようになると消去されます。

vSphere DPM のアクティビティを監視するために、次に示す vCenter Server イベントのアラームも作成できます。

表 13-2. vCenter Server イベント

イベント タイプ	イベント名
スタンバイ モードに切り替えています (ホストをパワーオフしようとしている)	DrsEnteringStandbyModeEvent
正常にスタンバイ モードに入りました (ホストのパワーオフが成功した)	DrsEnteredStandbyModeEvent
スタンバイ モードを終了しています (ホストをパワーオンしようとしている)	DrsExitingStandbyModeEvent
正常にスタンバイ モードを終了しました (パワーオンが成功した)	DrsExitedStandbyModeEvent

アラームの作成と編集の詳細については、『vSphere の監視とパフォーマンス』ドキュメントを参照してください。

vCenter Server 以外の監視ソフトウェアを使用しており、そのソフトウェアが物理ホストの予想外のパワーオフ時にアラームを起動する場合は、vSphere DPM がホストをスタンバイ モードにするときに、誤ったアラームが生成される可能性があります。そのようなアラームを受信しないようにする場合は、ベンダーと協力し、vCenter Server に統合したバージョンの監視ソフトウェアを導入してください。また、vCenter Server 自体を監視ソリューションとして使用することもできます。vSphere 4.x 以降では、vCenter Server は vSphere DPM を本質的に認識でき、そうした誤ったアラームを起動しないからです。

DRS アフィニティ ルールの使用

アフィニティ ルールを使用することで、クラスタ内のホストへの仮想マシンの配置を制御できます。

次の 2 つのタイプのルールを作成できます。

- 仮想マシンのグループとホストのグループとの間にアフィニティまたは非アフィニティを指定するために使用します。アフィニティ ルールは、選択した仮想マシンの DRS グループのメンバーが、指定したホストの DRS グループのメンバー上で実行できる、または実行する必要があることを指定します。非アフィニティ ルールは、選択した仮想マシンの DRS グループのメンバーが、指定したホストの DRS グループのメンバー上で実行できないことを指定します。

このタイプのルールの作成と使用については、[仮想マシンとホスト間のアフィニティ ルール](#) を参照してください。

- 個々の仮想マシン間のアフィニティまたは非アフィニティを指定するために使用します。アフィニティを指定したルールにより、DRS は指定された仮想マシンを、パフォーマンス上の理由などから同一ホスト上に保持しようとしています。非アフィニティ ルールを使用すると、DRS は指定された仮想マシンを別々に配置しようとしています。たとえば、1 台のホストに問題が発生しても、両方の仮想マシンを失うことはありません。

このタイプのルールの作成と使用については、[仮想マシン間のアフィニティ ルール](#) を参照してください。

アフィニティ ルールの追加または編集時に、クラスタの現在の状態がルールに違反している場合、システムは操作を続行して違反を修正しようとしています。手動、または部分的に自動化された DRS クラスタの場合、ルールの達成とロード バランシングに基づいた移行の推奨が表示され、許可が求められます。ルールを達成する必要はありませんが、ルールを達成するまで、それに対応する推奨が表示されます。

有効なアフィニティ ルールに、DRS で修正できない違反があるかどうかを確認するには、クラスタの [DRS] タブを選択し、[障害] をクリックします。現在違反があるすべてのルールについて、それに対応する障害がこのページに表示されます。障害を読んで、DRS がこの特定のルールに応じることができない理由を判断してください。ルール違反があると、ログ イベントも生成されます。

注： 仮想マシン間、および仮想マシンとホスト間のアフィニティ ルールは、個々のホストの CPU アフィニティ ルールとは異なります。

ホストの DRS グループの作成

仮想マシンとホスト間のアフィニティ ルールは、仮想マシンの DRS グループとホストの DRS グループとの間のアフィニティ（または非アフィニティ）関係を定義します。リンクするためのルールを作成する前に、これら両方のグループを作成する必要があります。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [構成] で [仮想マシン/ホスト グループ] を選択し、[追加] をクリックします。
- 4 [仮想マシン/ホスト グループの作成] ダイアログ ボックスで、グループの名前を入力します。
- 5 [タイプ] ドロップダウン ボックスから [ホスト グループ] を選択し、[追加] をクリックします。

- 6 ホストの横にあるチェック ボックスをクリックしてホストを追加します。必要なホストがすべて追加されるまで、この手順を繰り返します。
- 7 [OK] をクリックします。

次のステップ

ホストの DRS グループを使用して、仮想マシンとホスト間のアフィニティ ルールを作成できます。これにより、適切な仮想マシンの DRS グループとのアフィニティ（または非アフィニティ）関係を定義します。

[仮想マシンの DRS グループの作成](#)

[仮想マシンとホスト間のアフィニティ ルールの作成](#)

仮想マシンの DRS グループの作成

アフィニティ ルールは、DRS グループ間のアフィニティ（またはアンチアフィニティ）関係を規定します。リンクするためのルールを作成する前に、DRS グループを作成する必要があります。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [構成] で [仮想マシン/ホスト グループ] を選択し、[追加] をクリックします。
- 4 [仮想マシン/ホスト グループの作成] ダイアログ ボックスで、グループの名前を入力します。
- 5 [タイプ] ドロップダウン ボックスから [仮想マシン グループ] を選択し、[追加] をクリックします。
- 6 仮想マシンの横にあるチェック ボックスをクリックして仮想マシンを追加します。必要な仮想マシンがすべて追加されるまで、この手順を繰り返します。
- 7 [OK] をクリックします。

次のステップ

[ホストの DRS グループの作成](#)

[仮想マシンとホスト間のアフィニティ ルールの作成](#)

[仮想マシン間のアフィニティ ルールの作成](#)

仮想マシン間のアフィニティ ルール

仮想マシン間のアフィニティ ルールは、選択した個々の仮想マシンを同じホスト上で実行するか、別のホスト上に維持するかを指定します。このタイプのルールは、選択した個々の仮想マシン間のアフィニティまたは非アフィニティ関係を作成するのに使用します。

アフィニティ ルールが作成されると、DRS は、指定された仮想マシンを同じホスト上にまとめて維持します。パフォーマンス上の理由などで、この処理が必要になります。

非アフィニティ ルールの場合、DRS は、指定された仮想マシンを分離します。特定の仮想マシンをそれぞれ異なる物理ホスト上に常に維持する必要がある場合に、このルールを使用できます。その場合、1 台のホストに問題が発生しても、すべての仮想マシンにリスクが及ぶことはありません。

仮想マシン間のアフィニティ ルールの作成

仮想マシン間のアフィニティ ルールを作成して、選択した個々の仮想マシンを同じホスト上で実行するか、別のホスト上に維持するかを指定できます。

注： vSphere HA のフェイルオーバー ホストの指定アドミSSION コントロール ポリシーを使用して、複数のフェイルオーバー ホストを指定する場合、仮想マシン間のアフィニティ ルールはサポートされません。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [構成] で [仮想マシン/ホスト ルール] をクリックします。
- 4 [追加] をクリックします。
- 5 [仮想マシン/ホスト ルールの作成] ダイアログ ボックスで、ルールの名前を入力します。
- 6 [タイプ] ドロップダウン メニューから、[仮想マシンの包括] または [仮想マシンの分割] を選択します。
- 7 [追加] をクリックします。
- 8 ルールを適用する仮想マシンを少なくとも 2 台選択し、[OK] をクリックします。
- 9 [OK] をクリックします。

仮想マシン間のアフィニティ ルールの競合

仮想マシン間のアフィニティ ルールは複数作成して使用できますが、その場合、ルール同士が競合する可能性があります。

仮想マシン間の 2 つのアフィニティ ルールが競合する場合、両方のルールを有効にすることはできません。たとえば、特定のルールで 2 台の仮想マシンをまとめて扱う一方で、別のルールでこれらの 2 台の仮想マシンを個別に扱うようにした場合、両方のルールを有効にすることはできません。適用するルールを 1 つ選択し、それと競合するルールは無効にするか、削除します。

仮想マシン間の 2 つのアフィニティ ルールが競合する場合、古い方のルールが優先され、新しいルールは無効になります。DRS は有効なルールのみ準拠し、無効なルールは無視します。DRS は、アフィニティ ルールの違反よりも、非アフィニティ ルールの違反を防ぐことを優先させます。

仮想マシンとホスト間のアフィニティ ルール

仮想マシンとホスト間のアフィニティ ルールは、選択した仮想マシン DRS グループのメンバーが、特定のホスト DRS グループのメンバー上で実行できるかどうかを指定します。

個々の仮想マシンの間にアフィニティ (または非アフィニティ) を指定する仮想マシン間のアフィニティ ルールとは異なり、仮想マシンとホスト間のアフィニティ ルールは、仮想マシンのグループとホストのグループとの間にアフィニティ 関係を指定します。アフィニティ ルールには、「必須」ルール (「must」で指定) と「必須でない」ルール (「should」で指定) があります。

仮想マシンとホスト間のアフィニティ ルールには、次のコンポーネントがあります。

- 1 つの仮想マシン DRS グループ

- 1つのホスト DRS グループ
- ルールが必須（「must」）か必須でない（「should」）か、アフィニティ（「run on」）か非アフィニティ（「not run on」）かの指定。

仮想マシンとホスト間のアフィニティ ルールはクラスタ ベースであるため、ルールに含まれている仮想マシンとホストが、すべて同じクラスタ内にある必要があります。仮想マシンがクラスタから削除された場合、あとでそのクラスタに戻したとしても、その仮想マシンは DRS グループの関連付けを失います。

仮想マシンとホスト間のアフィニティ ルールの作成

仮想マシンとホスト間のアフィニティ ルールを作成して、選択した仮想マシン DRS グループのメンバーが、特定のホスト DRS グループのメンバー上で実行できるかどうかを指定します。

前提条件

仮想マシンとホスト間のアフィニティ ルールを適用する、仮想マシンの DRS グループおよびホストの DRS グループを作成します。

手順

- 1 vSphere Client のクラスタを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [構成] で [仮想マシン/ホスト ルール] をクリックします。
- 4 [追加] をクリックします。
- 5 [仮想マシン/ホスト ルールの作成] ダイアログ ボックスで、ルールの名前を入力します。
- 6 [タイプ] ドロップダウン メニューで、[仮想マシンからホストへ] を選択します。
- 7 ルールを適用する仮想マシンの DRS グループおよびホストの DRS グループを選択します。
- 8 ルールの仕様を選択します。
 - [グループ内のホスト上で実行する必要があります]。仮想マシン グループ 1 の仮想マシンはホスト グループ A のホスト上で実行する必要があります。
 - [グループ内のホスト上で実行します]。仮想マシン グループ 1 の仮想マシンはホスト グループ A のホスト上で実行しますが、それが必須ではありません。
 - [グループ内のホスト上では実行しない必要があります]。仮想マシン グループ 1 の仮想マシンはホスト グループ A のホスト上では実行しない必要があります。
 - [グループ内のホスト上では実行しません]。仮想マシン グループ 1 の仮想マシンはホスト グループ A のホスト上で実行しませんが、実行する場合があります。
- 9 [OK] をクリックします。

仮想マシンとホスト間のアフィニティ ルールの使用

仮想マシンとホスト間のアフィニティ ルールを使用して、仮想マシンのグループとホストのグループの間にアフィニティ関係を指定します。仮想マシンとホスト間のアフィニティ ルールを使用する際には、そのルールが最も役立つタ

イミング、ルール間の競合を解決する方法、および必須アフィニティ ルール設定時の注意事項の重要性を認識しておく必要があります。

仮想マシンとホスト間のアフィニティ ルールを複数作成する場合、それらのルールのランク付けは行われず、同等に適用されます。これがルール間の相互作用に与える影響に注意してください。たとえば、1 台の仮想マシンが 2 つの DRS グループに属し、DRS グループのそれぞれに異なる必須ルールが適用されている場合、その仮想マシンは、各ルールに表されるホスト DRS グループの両方に属するホスト上でのみ実行できます。

仮想マシンとホスト間のアフィニティ ルールを作成するときに、そのルールがほかのルールに関連して機能できるかどうかはチェックされません。このため、使用しているほかのルールと競合するルールを作成してしまう可能性があります。仮想マシンとホスト間の 2 つのアフィニティ ルールが競合する場合、古い方のルールが優先され、新しいルールは無効になります。DRS は有効なルールのみを準拠し、無効なルールは無視します。

DRS、vSphere HA、および vSphere DPM は、必須アフィニティ ルール（仮想マシン DRS グループのホスト DRS グループに対する「実行必須」または「実行禁止」ルール）の違反につながるアクションは実行しません。クラスタの機能に悪影響が生じる可能性があるため、このタイプのルールを使用するときには注意が必要です。仮想マシンとホスト間のアフィニティ ルールを不適切に使用した場合、クラスタが断片化され、DRS、vSphere HA、および vSphere DPM の機能が阻害される可能性があります。

必須アフィニティ ルールの違反につながる場合、いくつかのクラスタ機能は実行されません。

- DRS は、ホストをメンテナンス モードにするために仮想マシンを退避させません。
- DRS は、仮想マシンのパワーオンや仮想マシンのロード バランシングを行いません。
- vSphere HA は、フェイルオーバーを実行しません。
- vSphere DPM は、ホストをスタンバイ モードにして電源管理を最適化することをしません。

これらの問題を回避するため、必須アフィニティ ルールを複数作成するときには注意してください。または、仮想マシンとホスト間の、必須ではないアフィニティ ルール（仮想マシン DRS グループのホスト DRS グループに対する「実行」または「不実行」ルール）を使用することを検討します。各仮想マシンが関連付けられているクラスタ内のホストの数が十分であり、1 台のホストが失われても、仮想マシンを実行できるホストがなくなることを確認します。DRS、vSphere HA、および vSphere DPM が適切に機能するように、必須でないルールに違反することは可能です。

注： 仮想マシンとホスト間のアフィニティ ルールに仮想マシンが違反したときに起動される、イベント ベースのアラームを作成できます。仮想マシン用の新しいアラームを追加し、イベント トリガとして [仮想マシンが仮想マシン/ホスト アフィニティ ルールに違反しています] を選択します。アラームの作成と編集の詳細については、『vSphere Monitoring and Performance』ドキュメントを参照してください。

データストア クラスタの作成

14

データストア クラスタは、リソースと管理インターフェイスが共有されたデータストアの集まりです。データストア クラスタとデータストアの関係は、クラスタとホストの関係と似ています。データストア クラスタを作成すると、vSphere ストレージ DRS を使用してストレージ リソースを管理できます。

注： データストア クラスタは、vSphere API でストレージ ポッドとも呼ばれます。

データストア クラスタにデータストアを追加すると、データストアのリソースはそのデータストア クラスタのリソースの一部になります。ホストのクラスタと同様に、データストア クラスタを使用してストレージ リソースを集約することで、データストア クラスタ レベルでリソース割り当てポリシーをサポートすることができます。次のリソース管理機能も各データストア クラスタで使用できます。

領域使用率のロード バランシング

領域使用のしきい値を設定することができます。データストアの領域使用量がしきい値を超えると、ストレージ DRS によって推奨が生成されるか、Storage vMotion 移行が実行されて、データストア クラスタ全体で領域使用量のバランスが調整されます。

I/O 待ち時間のロード バランシング

I/O 待ち時間のしきい値を設定して、ポトルネックを回避できます。データストア上の I/O 待ち時間がしきい値を超えると、ストレージ DRS によって推奨が生成されるか、Storage vMotion 移行が実行されて、I/O の高負荷が軽減されます。

非アフィニティ ルール

仮想マシン ディスク用の非アフィニティ ルールを作成できます。たとえば、特定の仮想マシンの仮想ディスクは別々のデータストアに配置する必要があります。デフォルトでは、仮想マシンのすべての仮想ディスクが同じデータストアに配置されます。

この章には、次のトピックが含まれています。

- [初期配置と実行中のバランシング](#)
- [ストレージの移行の推奨](#)
- [データストア クラスタの作成](#)
- [Storage DRS の有効化および無効化](#)
- [データストア クラスタの自動化レベルの設定](#)
- [Storage DRS の積極性レベルの設定](#)

- データストア クラスタの要件
- データストアのデータストア クラスタへの追加と削除

初期配置と実行中のバランシング

ストレージ DRS では、初期配置と実行中のバランシングの推奨が、ストレージ DRS 対応のデータストア クラスタ内のデータストアに対して提供されます。

初期配置は、仮想マシン ディスクを配置するデータストア クラスタ内のデータストアが、ストレージ DRS で選択されるときに発生します。これは、仮想マシンが作成またはクローン作成される時、仮想マシン ディスクが別のデータストア クラスタに移行される時、または既存の仮想マシンにディスクを追加するときに発生します。

初期配置の推奨は、領域の制約に従って、領域および I/O ロード バランシングの目標を基準に作成されます。これらの目標は、1つのデータストアのオーバー プロビジョニング、ストレージ I/O のボトルネック、および仮想マシンのパフォーマンス上の影響のリスクを最低限にすることを意図しています。

ストレージ DRS は、構成された頻度（デフォルトでは 8 時間ごと）で起動する、またはデータストア クラスタ内の 1つ以上のデータストアがユーザー構成可能な領域使用率のしきい値を超えたときに起動します。ストレージ DRS が起動すると、各データストアの領域使用率および I/O 待ち時間の値が、しきい値と照合して確認されます。I/O 待ち時間について、ストレージ DRS では、1日の過程で測定された I/O 待ち時間の 90 パーセントの値が、しきい値との照合に使用されます。

ストレージの移行の推奨

vCenter Server では、手動での自動化モードがあるデータストア クラスタのストレージ DRS 推奨ページに、移行の推奨が表示されます。

このシステムでは、ストレージ DRS ルールの適用と、データストア クラスタの領域と I/O リソースのバランスの調整に必要なだけの推奨が提供されます。各推奨には、仮想マシン名、仮想ディスク名、データストア クラスタ名、ソース データストア、ターゲット データストア、および推奨の理由が含まれます。

- データストアの領域使用量のバランス調整
- データストア I/O ロードのバランス調整

ストレージ DRS は、次の状況で移行に必須の推奨を行います。

- データストアの空き領域がなくなった。
- 非アフィニティまたはアフィニティ ルールに違反している。
- データストアがメンテナンス モードを開始しようとしており、退避させる必要がある。

さらに、データストアの領域がなくなりかかっている場合や、領域や I/O ロード バランシングのために調整する必要がある場合に、オプションの推奨が行われます。

ストレージ DRS では、領域のバランシングのためにパワーオフ状態またはパワーオン状態の仮想マシンの移動が検討されます。ストレージ DRS では、スナップショットがあるパワーオフ状態の仮想マシンもこれらの検討に含まれます。

データストア クラスタの作成

データストア クラスタ リソースは Storage DRS を使用して管理できます。

手順

- 1 vSphere Client で、データセンターを参照します。
- 2 データセンター オブジェクトを右クリックし、[新規データストア クラスタ] を選択します。
- 3 [新規データストア クラスタ] ウィザードを実行するには、画面の指示どおりに操作します。
- 4 [終了] をクリックします。

Storage DRS の有効化および無効化

Storage DRS を使用することで、データストア クラスタの統合されたリソースを管理できます。Storage DRS が有効な場合、データストア クラスタ内のデータストア全体で容量と I/O リソースの負荷を分散するための仮想マシンのディスク配置および移行に関する推奨が提供されます。

Storage DRS を有効にすると、以下の機能が有効になります。

- データストア クラスタ内のデータストア間における容量のロード バランシング。
- データストア クラスタ内のデータストア間における I/O のロード バランシング。
- 容量および I/O ワークロードに基づく仮想ディスクの初期配置。

データストア クラスタ設定ダイアログ ボックスの Storage DRS を有効にするためのチェック ボックスで、これらすべてのコンポーネントが一度に有効または無効になります。必要に応じて、Storage DRS の I/O 関連の機能を、容量のバランシング機能とは別に無効にすることができます。

データストア クラスタ上で Storage DRS を無効にした場合、Storage DRS 設定は保存されます。Storage DRS を有効にしたときに、データストア クラスタに対する設定は、Storage DRS が無効にされた時点の設定に戻ります。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[サービス] をクリックします。
- 3 [Storage DRS] を選択し、[編集] をクリックします。
- 4 [vSphere DRS をオンにする] を選択し、[OK] をクリックします。
- 5 (オプション) Storage DRS の I/O 関連の機能のみを無効にするには、容量関連のコントロールを有効にしたままで、次の手順を実行します。
 - a [Storage DRS] の [編集] を選択します。
 - b [SDRS 推奨に対して I/O メトリックを有効にします] チェック ボックスを選択解除して、[OK] をクリックします。

データストア クラスタの自動化レベルの設定

データストア クラスタの自動化レベルによって、Storage DRS からの配置および移行の推奨が自動的に適用されるかどうかが決まります。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[サービス] をクリックします。
- 3 [DRS] を選択し、[編集] をクリックします。
- 4 [DRS 自動化] を展開して、自動化レベルを選択します。

デフォルトの自動化レベルは手動です。

オプション	説明
自動化なし (手動モード)	配置と移行の推奨が表示されますが、推奨が手動で適用されるまで実行はされません。
完全自動化	配置と移行の推奨が自動的に実行されます。

- 5 [OK] をクリックします。

Storage DRS の積極性レベルの設定

使用容量と I/O 遅延のしきい値を指定することで、Storage DRS の積極性が決まります。

Storage DRS は、データストア クラスタ内のデータストアのリソース使用情報を収集します。vCenter Server では、データストア上の仮想ディスクの配置用の推奨を生成するため、この情報が使用されます。

データストア クラスタに低い積極性レベルを設定した場合、Storage DRS では、明らかに必要な場合にのみ、Storage vMotion の移行が推奨されます。たとえば、I/O ロード、容量使用率、またはそれらの負荷分散の不均衡が高い場合です。データストア クラスタに高い積極性レベルを設定した場合、Storage DRS では、データストア クラスタが容量または I/O ロードの分散によって利益を受けるときはいつでも、移行が推奨されます。

vSphere Client では、Storage DRS に対する積極性レベルの設定に、次のしきい値を使用できます。

容量使用率

vSphere Client に設定したしきい値をデータストアの容量使用率の割合が上回ったときに、Storage DRS で推奨が生成される、または移行が実行されます。

I/O 遅延

データストアについて 1 日に測定された I/O 遅延の 90 パーセントの値が、しきい値を上回ったときに、Storage DRS で推奨が生成される、または移行が実行されます。

また、詳細オプションを設定し、Storage DRS の積極性レベルをさらに構成することもできます。

容量使用率差

このしきい値は、ソースとターゲットで使用されている容量の差が最低限になるようにします。たとえば、データストア A で使用された容量が 82% で、データストア B が 79% の場合、差は 3 になります。しきい値が 5 の場合、Storage DRS はデータストア A からデータストア B への移行推奨を行いません。

I/O ロード バランシングの起動間隔

この時間の経過後、Storage DRS で I/O ロードの調整が実行されます。

I/O 不均衡のしきい値

この値を下げると、I/O ロード バランシングの積極性が低下します。Storage DRS は、0 と 1 の間の I/O 公平性のメトリックを算出します。1 は、最も公平な分散を表します。I/O ロード バランシングは、算出されたメトリックが 1 より小さい場合 (I/O 不均衡のしきい値 / 100) のみ実行されます。

ストレージ DRS ランタイム ルールの設定

ストレージ DRS トリガーを設定し、データストア クラスタの詳細オプションを構成します。

手順

- 1 (オプション) [Storage DRS 推奨に対する I/O メトリックの有効化] チェック ボックスを選択または選択解除し、I/O メトリックを含めるかどうかを決めます。

このオプションを無効にすると、ストレージ DRS 推奨を行うときに vCenter Server は I/O メトリックを考慮しなくなります。このオプションを無効にすると、ストレージ DRS の以下の要素が無効になります。

- データストア クラスタ内のデータストア間における I/O のロード バランシング。
- I/O ワークロードに基づく仮想ディスクの初期配置。初期配置は、領域のみに基づきます。

- 2 (オプション) ストレージ DRS のしきい値を設定します。

使用領域と I/O 待ち時間のしきい値を指定することで、ストレージ DRS の積極性レベルを設定します。

- 使用されている領域のスライダを使用して、ストレージ DRS が起動されるまでに許容される消費領域の最大のパーセンテージを示します。データストアでの領域使用量がしきい値を上回ると、ストレージ DRS は、推奨を行い、移行を実行します。
- I/O 待機時間のスライダを使用して、ストレージ DRS が起動されるまでに許容される最大 I/O 待機時間を示します。待機時間がしきい値を上回ると、ストレージ DRS は、推奨を行い、移行を実行します。

注： データストア クラスタのストレージ DRS I/O 待ち時間のしきい値は、ストレージ I/O コントロールの輻輳のしきい値以下である必要があります。

- 3 (オプション) 詳細オプションを構成します。

- ソースとターゲットの使用率の差が次の場合、推奨が行われます。スライダを使用して、領域使用率差のしきい値を指定します。使用率は、使用量 × 100 / キャパシティです。

このしきい値は、ソースとターゲットで使用されている容量の差が最低限になるようにします。たとえば、データストア A で使用された容量が 82% で、データストア B が 79% の場合、差は 3 になります。しきい値が 5 の場合、Storage DRS はデータストア A からデータストア B への移行推奨を行いません。

- 不均衡の確認間隔：ストレージ DRS が領域と I/O ロード バランシングを評価する頻度を指定します。

- I/O 不均衡のしきい値：スライダを使用して、I/O ロード バランシングの積極性を示します。この値を下げると、I/O ロード バランシングの積極性が低下します。Storage DRS は、0 と 1 の間の I/O 公平性のメトリックを算出します。1 は、最も公平な分散を表します。I/O ロード バランシングは、算出されたメトリックが 1 より小さい場合（I/O 不均衡のしきい値 / 100）のみ実行されます。

4 [OK] をクリックします。

データストア クラスタの要件

データストア クラスタ機能を適切に使用するには、データストア クラスタと関連付けられたデータストアおよびホストが、特定の要件を満たしている必要があります。

データストア クラスタを作成するときは、次のガイドラインに従ってください。

- データストア クラスタに含めるデータストアは、相互に類似していて、交換可能なものにする必要があります。1 つのデータストア クラスタには、サイズおよび I/O キャパシティが異なるデータストアや、アレイおよびベンダーの異なるデータストアを混在させることができます。しかし、1 つのデータストア クラスタ内に、次のタイプのデータストアを共存させることはできません。
 - 同一のデータストア クラスタで、NFS データストアと VMFS データストアを組み合わせることはできません。
 - レプリケートされたデータストアは、同一のストレージ DRS 対応データストア クラスタに属するレプリケートされていないデータストアと組み合わせることはできません。
- データストア クラスタ内のデータストアに接続されるすべてのホストは、ESXi 5.0 以降である必要があります。データストア クラスタ内のデータストアが ESX/ESXi 4.x 以前のホストと接続されている場合、ストレージ DRS は実行されません。
- 複数のデータ センターで共有されているデータストアは、データストア クラスタに含めることはできません。
- ベスト プラクティスとして、ハードウェア アクセラレーション機能が有効なデータストアは、ハードウェア アクセラレーション機能が有効でないデータストアと同じデータストア クラスタ内に含めないようにします。ハードウェア アクセラレーションでサポートされる動作を確保するには、データストア クラスタ内のデータストアを同種のクラスタにする必要があります。

データストアのデータストア クラスタへの追加と削除

既存のデータストア クラスタに対して、データストアを追加または削除します。

次の場合を除き、データストア クラスタには vSphere Client インベントリ内のホストにマウントされたどのデータストアも追加できます。

- データストアに接続されるすべてのホストは、ESXi 5.0 以降である必要があります。
- データストアは、vSphere Client の同じインスタンス内にある複数のデータセンターに含めることはできません。

データストア クラスタからデータストアを削除しても、そのデータストアは vSphere Client インベントリ内に残り、ホストからアンマウントされません。

データストア クラスタを使用したストレージ リソースの管理

15

データストア クラスタを作成したあとで、それをカスタマイズしたり、それを使用してストレージ I/O および領域使用率リソースを管理したりできます。

この章には、次のトピックが含まれています。

- ストレージ DRS メンテナンス モードの使用
- Storage DRS 推奨の適用
- 仮想マシンの Storage DRS 自動レベルの変更
- オフピークの Storage DRS スケジュールの設定
- Storage DRS 非アフィニティ ルール
- Storage DRS 統計の削除
- ストレージ vMotion のデータストア クラスタとの互換性

ストレージ DRS メンテナンス モードの使用

データストアを使用停止にして保守作業を行うために、メンテナンス モードに切り替えます。データストアは、ユーザーの要求があった場合のみ、メンテナンス モードを開始または終了します。

ストレージ DRS 対応データストア クラスタ内のデータストアは、メンテナンス モードを使用できます。スタンドアロン データストアはメンテナンス モードにすることはできません。

メンテナンス モードになっているデータストア上の仮想ディスクは、手動またはストレージ DRS を使用して、別のデータストアに移行する必要があります。データストアをメンテナンス モードに切り替えようとする、[配置の推奨] タブに移行の推奨リスト、仮想ディスクを移行できる同じデータストア クラスタ内のデータストアが表示されません。vCenter Server では、[障害] タブに、移行できないディスクのリストとその理由が表示されます。ストレージ DRS のアフィニティまたは非アフィニティのルールにより、ディスクを移行できない場合は、[メンテナンスで関係ルールを無視する] オプションを有効にできます。

すべての仮想ディスクが移行されるまで、データストアは [メンテナンス モードに切り替えています] の状態のままになります。

データストアをメンテナンス モードに切り替える

データストアのサービスを停止する必要がある場合、データストアを Storage DRS メンテナンス モードに切り替えることができます。

前提条件

メンテナンス モードに切り替えるデータストアを含むデータストア クラスタで、Storage DRS が有効になっている。

データストアに、CD-ROM のイメージ ファイルが保存されていない。

データストア クラスタに少なくとも 2 つのデータストアがある。

手順

- 1 vSphere Client で、データストアを参照して移動します。
- 2 データストアを右クリックして、[メンテナンス モード] - [メンテナンス モードへの切り替え]を選択します。
データストアのメンテナンス モード移行に関する推奨リストが表示されます。
- 3 (オプション) [配置の推奨] タブで、適用しない推奨を選択解除します。

注： ディスクをすべて取り外さないと、データストアをメンテナンス モードに切り替えることはできません。推奨を選択解除した場合は、影響を受ける仮想マシンを手動で移動する必要があります。

- 4 必要に応じて、[推奨の適用] をクリックします。

vCenter Server は Storage vMotion を使用して、ソース データストアからターゲット データストアに仮想ディスクを移行します。データストアはメンテナンス モードに切り替わります。

結果

データストア アイコンがすぐに更新されず、データストアの最新の状態が反映されないことがあります。アイコンを強制的に更新するには、[更新] をクリックします。

Storage DRS アフィニティ ルールを無視したメンテナンス モードへの切り替え

Storage DRS アフィニティまたは非アフィニティ ルールにより、データストアがメンテナンス モードを開始できなくなる可能性があります。データストアをメンテナンス モードに切り替えるときに、これらのルールを無視できません。

データストア クラスタに対し、メンテナンスでアフィニティ ルールを無視するオプションを有効にした場合、データストアのメンテナンス モードへの切り替えを妨げる Storage DRS のアフィニティ ルールまたは非アフィニティ ルールが、vCenter Server で無視されます。

Storage DRS ルールは、退避の推奨についてのみ無効になります。容量およびロード バランシングの推奨や初期配置の推奨を作成するときのルールが、vCenter Server によって阻害されることはありません。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[サービス] をクリックします。
- 3 [DRS] を選択し、[編集] をクリックします。
- 4 [詳細オプション] を展開して、[追加] をクリックします。

- 5 「オプション」列に「IgnoreAffinityRulesForMaintenance」と入力します。
- 6 オプションを有効にするには、値の列に「1」と入力します。
オプションを無効にするには、「0」と入力します。
- 7 [OK] をクリックします。

結果

メンテナンスでアフィニティ ルールを無視するオプションがデータストア クラスタに適用されます。

Storage DRS 推奨の適用

Storage DRS は、データストア クラスタ内のすべてのデータストアにおけるリソースの使用情報を収集します。Storage DRS はこの情報を使用して、データストア クラスタ内のデータストア上の仮想マシン ディスク配置に関する推奨を生成します。

Storage DRS の推奨は、vSphere Client データストア ビューの [Storage DRS] タブに表示されます。また、データストアを Storage DRS メンテナンス モードにする際、推奨が表示されます。Storage DRS 推奨を適用すると、vCenter Server は Storage vMotion を使用して仮想マシンのディスクをデータストア クラスタ内の他のデータストアに移行して、リソースを調整します。

推奨のサブセットを適用するには、[提案された DRS 推奨のオーバーライド] チェック ボックスを選択し、各推奨を選択して適用します。

表 15-1. Storage DRS 推奨

ラベル	説明
優先順位	推奨の優先順位レベル (1 ~ 5)。(デフォルトでは非表示。)
推奨	Storage DRS が推奨するアクション。
理由	アクションが必要な理由。
移行前の容量使用率 % (ソース) および (ターゲット)	移行前にソースとターゲットのデータストアで使用される容量の割合。
移行後の容量使用率 % (ソース) および (ターゲット)	移行後にソースとターゲットのデータストアで使用される容量の割合。
移行前の I/O 遅延 (ソース)	移行前のソース データストア上の I/O 遅延の値。
移行前の I/O 遅延 (ターゲット)	移行前のターゲット データストア上の I/O 遅延の値。

Storage DRS 推奨の更新

Storage DRS の移行に関する推奨は、vSphere Client の [Storage DRS] タブに表示されます。これらの推奨は、Storage DRS を実行することで更新できます。

前提条件

1 つ以上のデータストア クラスタが vSphere Client インベントリに含まれている必要があります。

データストア クラスタ用の Storage DRS を有効にします。[Storage DRS] タブは、Storage DRS が有効になっている場合のみ表示されます。

手順

- 1 vSphere Client データストア ビューで、データストア クラスタを選択して、[Storage DRS] タブをクリックします。
- 2 [推奨] ビューを選択し、右上隅の [Storage DRS の実行] リンクをクリックします。

結果

推奨が更新されます。最終更新日タイムスタンプに、Storage DRS の推奨の更新時が表示されます。

仮想マシンの Storage DRS 自動レベルの変更

データストア クラスタ全体に渡る自動化レベルは、仮想マシンごとにオーバーライドできます。デフォルトの仮想ディスク アフィニティ ルールをオーバーライドすることもできます。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[構成] をクリックします。
- 3 [仮想マシンのオーバーライド] で、[追加] を選択します。
- 4 仮想マシンを選択します。
- 5 [自動化レベル] ドロップダウン メニューをクリックし、仮想マシンの自動化レベルを選択します。

オプション	説明
デフォルト (手動)	配置と移行の推奨が表示されますが、推奨が手動で適用されるまで実行はされません。
完全自動化	配置と移行の推奨が自動的に実行されます。
無効	vCenter Server では、仮想マシンの移行や、移行のための推奨は行われません。

- 6 [VMDK の包括] ドロップダウン メニューをクリックし、デフォルトの VMDK アフィニティをオーバーライドします。

[VMDK アフィニティ ルールのオーバーライド](#)を参照してください。

- 7 [OK] をクリックします。

オフピークの Storage DRS スケジュールの設定

スケジュール設定タスクを作成して、データストア クラスタの Storage DRS 設定を変更できます。これにより、完全に自動化されたデータストア クラスタの移行をピーク時以外の時間帯に実行できるようになります。

スケジュール設定タスクを作成することで、データストア クラスタの自動化レベルおよび積極性レベルを変更することができます。たとえば、パフォーマンスが重視される場合には、ピーク時に Storage DRS の実行を抑制するように構成して、ストレージ移行の発生を最小限にすることができます。ピーク時以外には、Storage DRS をより積極的に実行し、より頻繁に移行させることができます。

前提条件

Storage DRS を有効にします。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[サービス] をクリックします。
- 3 [vSphere DRS] で [DRS をスケジュール] ボタンをクリックします。
- 4 データストア クラスタの編集ダイアログ ボックスで、[SDRS スケジュール設定] をクリックします。
- 5 [DRS 自動化] を展開します。
 - a 自動化レベルを選択します。
 - b 移行のしきい値を設定します。

移行スライドを使用して、クラスタのロード バランシングを行う vCenter Server の推奨の優先順位レベルを選択します。
 - c 仮想マシンの自動化を有効にするかどうかを選択します。

各仮想マシンのオーバーライドは「仮想マシンのオーバーライド」ページから設定できます。
- 6 [電源管理] を展開します。
 - a 自動化レベルを選択します。
 - b DPM のしきい値を設定します。

DPM スライダーを使用して、vCenter Server に適用する 電源の推奨事項を選択します。
- 7 タスク名を入力します。
- 8 作成したタスクの説明を入力します。
- 9 [構成済みスケジューラー] の [変更] をクリックし、タスクを実行する時間を選択して、[OK] をクリックします。
- 10 タスク完了時に送信する通知メールの送信先アドレスを入力します。
- 11 [OK] をクリックします。

結果

指定した時間に、スケジュール設定タスクが実行されます。

Storage DRS 非アフィニティ ルール

ストレージ DRS 非アフィニティ ルールを作成して、データストア クラスタ内の同じデータストアに配置すべきではない仮想ディスクを指定することができます。デフォルトでは、仮想マシンの仮想ディスクはすべて、同じデータストア内で格納されます。

非アフィニティ ルールを作成すると、データストア クラスタ内の関連する仮想ディスクに適用されます。非アフィニティ ルールは初期配置および Storage DRS 推奨による移行の際に適用されますが、移行がユーザーによって開始される場合は適用されません。

注： 非アフィニティ ルールは、データストア クラスタ内のデータストアに保存された CD-ROM ISO イメージ ファイルには適用されず、ユーザー定義の場所に保存されたスワップファイルにも適用されません。

仮想マシン非アフィニティ ルール

どの仮想マシンを同じデータストアに格納できないようにするかを指定します。[仮想マシン非アフィニティ ルールの作成](#) を参照してください。

VMDK 非アフィニティ ルール

特定の仮想マシンに関連付けられた仮想ディスクのうち、どの仮想ディスクを別のデータストアに格納すべきかを指定します。[VMDK 非アフィニティ ルールの作成](#) を参照してください。

仮想ディスクをデータストア クラスタから除外すると、アフィニティ ルールまたは非アフィニティ ルールはそのディスクに適用されなくなります。

既存のアフィニティ ルールまたは非アフィニティ ルールのあるデータストア クラスタに仮想ディスク ファイルを移動する場合、以下の動作が適用されます。

- データストア クラスタ B に仮想マシン内のアフィニティ ルールがある場合。仮想ディスクをデータストア クラスタ A からデータストア クラスタ B に移動すると、データストア クラスタ A 内の特定の仮想マシンの仮想ディスクに適用されていたルールは適用されなくなります。仮想ディスクは、データストア クラスタ B 内の仮想マシン内のアフィニティ ルールの対象となります。
- データストア クラスタ B に仮想マシン非アフィニティ ルールがある場合。仮想ディスクをデータストア クラスタ A からデータストア クラスタ B に移動すると、データストア クラスタ A 内の特定の仮想マシンの仮想ディスクに適用されていたルールは適用されなくなります。仮想ディスクは、データストア クラスタ B 内の仮想マシン非アフィニティ ルールの対象となります。
- データストア クラスタ B に VMDK 非アフィニティ ルールがある場合。仮想ディスクをデータストア クラスタ A からデータストア クラスタ B に移動すると、ルールがデータストア クラスタ B 内の指定の仮想ディスクのみを対象とするため、データストア クラスタ B 内の VMDK 非アフィニティ ルールは特定の仮想マシンの仮想ディスクに適用されません。

注： Storage DRS ルールにより、データストアがメンテナンス モードを開始できなくなる可能性があります。[メンテナンスするアフィニティ ルールが無視します] オプションを有効にすると、メンテナンス モードのために Storage DRS ルールを無視できます。

仮想マシン非アフィニティ ルールの作成

非アフィニティ ルールを作成して、特定の仮想マシンのすべての仮想ディスクは別々のデータストアに存在する必要があることを示すことができます。このルールは個々のデータストア クラスタに適用されます。

データストア クラスタ内の仮想マシン非アフィニティ ルールに参加する仮想マシンを、データストア クラスタの仮想マシン内アフィニティ ルールと関連付ける必要があります。仮想マシンは、仮想マシン内アフィニティ ルールにも準拠する必要があります。

仮想マシンが仮想マシン非アフィニティ ルールの対象となる場合は、次の動作が適用されます。

- Storage DRS は、仮想マシンの仮想ディスクをルールに従って配置します。
- Storage DRS は、データストアをメンテナンス モードにするなどの強制的な理由で移行する場合でも、ルールに従って vMotion を使用して仮想ディスクを移行します。
- 仮想マシンの仮想ディスクがルールに違反している場合、Storage DRS は移行推奨を行ってエラーを修正します。推奨を行ってエラーを修正できない場合は、その違反をエラーとして報告します。

デフォルトでは、仮想マシン非アフィニティ ルールは定義されません。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[構成] をクリックします。
- 3 [仮想マシン/ホスト ルール] を選択します。
- 4 [追加] をクリックします。
- 5 ルールの名前を入力します。
- 6 [タイプ] メニューで [仮想マシン非アフィニティ] を選択します。
- 7 [追加] をクリックします。
- 8 [仮想マシンの選択] をクリックします。
- 9 少なくとも 2 台の仮想マシンを選択し、[OK] をクリックします。
- 10 [OK] をクリックしてルールを保存します。

VMDK 非アフィニティ ルールの作成

仮想マシンの VMDK 非アフィニティ ルールを作成して、どの仮想ディスクを別のデータストアに格納する必要があるかを示すことができます。

VMDK 非アフィニティ ルールは、ルールが定義された仮想マシンに適用されます。すべての仮想マシンに適用されるわけではありません。ルールは、互いに分離する仮想ディスクのリストとして表示されます。

1 台の仮想マシンに対して VMDK 非アフィニティ ルールと仮想マシン内アフィニティ ルールを設定しようとすると、最後に定義されたルールが vCenter Server で拒否されます。

仮想マシンが VMDK 非アフィニティ ルールの対象となる場合は、次の動作が適用されます。

- Storage DRS は、仮想マシンの仮想ディスクをルールに従って配置します。

- Storage DRS は、データストアをメンテナンス モードにするなどの強制的な理由で移行する場合でも、ルールに従って vMotion を使用して仮想ディスクを移行します。
- 仮想マシンの仮想ディスクがルールに違反している場合、Storage DRS は移行推奨を行ってエラーを修正します。推奨を行ってエラーを修正できない場合は、その違反をエラーとして報告します。

デフォルトでは、VMDK 非アフィニティ ルールは定義されません。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[構成] をクリックします。
- 3 [仮想マシン/ホスト ルール] を選択します。
- 4 [追加] をクリックします。
- 5 ルールの名前を入力します。
- 6 [タイプ] メニューで [VMDK 非アフィニティ] を選択します。
- 7 [追加] をクリックします。
- 8 [仮想マシンの選択] をクリックします。
- 9 仮想マシンを選択し、[OK] をクリックします。
- 10 ルールが適用される仮想ディスクを少なくとも 2 つ選択し、[OK] をクリックします。
- 11 [OK] をクリックしてルールを保存します。

VMDK アフィニティ ルールのオーバーライド

VMDK アフィニティ ルールにより、特定の仮想マシンに関連付けられているデータストア クラスタ内のすべての仮想ディスクが、データストア クラスタ内の同じデータストアに配置されます。このルールは個々のデータストア クラスタに適用されます。

VMDK アフィニティ ルールは、データストア クラスタ内のすべての仮想マシンに対し、デフォルトで有効になっています。データストア クラスタや個々の仮想マシンでは、このデフォルト設定をオーバーライドできます。

VMDK アフィニティ ルールが適用される仮想マシンは、次のように動作します。

- Storage DRS は、仮想マシンの仮想ディスクをルールに従って配置します。
- Storage DRS は、データストアをメンテナンス モードにするなどの強制的な理由で移行する場合でも、ルールに従って vMotion を使用して仮想ディスクを移行します。
- 仮想マシンの仮想ディスクがルールに違反している場合、Storage DRS は移行推奨を行ってエラーを修正します。推奨を行ってエラーを修正できない場合は、その違反をエラーとして報告します。

Storage DRS が有効なデータストア クラスタにデータストアを追加し、仮想マシンがそのデータストア上の仮想ディスクを使用するとします。この仮想マシンが他のデータストアに配置された仮想ディスクを使用している場合は、VMDK アフィニティ ルールが無効になります。

手順

- 1 vSphere Client で、データストア クラスタを参照して移動します。
- 2 [構成] タブをクリックし、[構成] をクリックします。
- 3 [仮想マシンのオーバーライド] を選択します。
- 4 [追加] をクリックします。
- 5 [+] ボタンを使用して、仮想マシンを選択します。
- 6 [VMDK の包括] ドロップダウン メンテナンスをクリックし、[いいえ] を選択します。
- 7 [OK] をクリックします。

Storage DRS 統計の削除

Storage DRS の問題を診断するには、手動で Storage DRS を実行する前に、Storage DRS 統計情報を削除できます。

重要： Storage DRS 統計情報を削除するオプションを有効にした場合、このオプションを無効にしないかぎり、統計情報は Storage DRS が実行されるたびに削除されます。Storage DRS の問題の診断後は、常にこのオプションを無効にしてください。

前提条件

データストア クラスタ用の Storage DRS を有効にします。

手順

- 1 **ClearIoStatsOnSdrsRun** オプションを有効にします。
 - a vSphere Client で、データストア クラスタを参照して移動します。
 - b [構成] タブをクリックし、[サービス] をクリックします。
 - c [vSphere DRS] を選択し、[編集] をクリックします。
 - d [詳細オプション] を展開して、[追加] をクリックします。
 - e 「オプション」列に「**ClearIoStatsOnSdrsRun**」と入力します。
 - f 対応する値テキスト ボックスに、「**1**」 と入力します。
 - g [OK] をクリックします。
- 2 データストア クラスタ用の Storage DRS を実行します。

vSphere Client インベントリ内のすべてのデータストア クラスタにあるすべてのデータストアおよび仮想ディスクについての現在の Storage DRS 統計情報が削除されますが、新しい統計情報は収集されません。
- 3 **ClearIoStatsOnSdrsRun** フラグの値を **0** に変更して無効にします。
- 4 Storage DRS を再度実行します。

Storage DRS が通常どおりに実行されます。新しい設定が有効になるまで、数時間かかります。

ストレージ vMotion のデータストア クラスタとの互換性

データストア クラスタには、特定の vSphere Storage vMotion[®] 要件があります。

- ホストは Storage vMotion をサポートしているバージョンの ESXi を実行する必要があります。
- ホストには、ソース データストアとターゲット データストアの両方に書き込みアクセスがある必要があります。
- ホストに Storage vMotion に対応するだけの十分な空きメモリ リソースがある必要があります。
- ターゲット データストアに十分なディスク領域がある必要があります。
- ターゲット データストアがメンテナンス モードである、またはメンテナンス モードを開始するところであってはなりません。

NUMA システムと ESXi の併用

16

ESXi は、NUMA (Non-Uniform Memory Access) をサポートするサーバ アーキテクチャで、Intel と AMD Opteron のプロセッサのメモリ アクセス最適化をサポートします。

ESXi における NUMA のスケジュール設定の実行方法と、VMware NUMA アルゴリズムの機能を理解すれば、NUMA 制御を指定して、仮想マシンのパフォーマンスを最適化できます。

この章には、次のトピックが含まれています。

- NUMA とは
- ESXi の NUMA スケジュール設定の機能
- VMware NUMA 最適化のアルゴリズムと設定
- NUMA アーキテクチャでのリソース管理
- 仮想 NUMA の使用
- NUMA 制御の指定

NUMA とは

NUMA システムは、複数のシステム バスを持つ高度なサーバ プラットフォームです。単一のシステム イメージ内で多数のプロセッサを利用でき、コスト パフォーマンスに優れています。

この 10 年間、プロセッサのクロック速度は劇的に上昇しました。しかし、マルチギガヘルツ CPU は、その処理能力を効果的に使用するために、大量のメモリ バンド幅とともに提供されなければなりません。科学計算アプリケーションなどのメモリ集約型のワークロードを実行する単一 CPU でさえ、メモリ バンド幅によって制約される場合があります。

この問題は、対称型マルチプロセッシング (SMP) システムで深刻になり、同じシステム バス上で多くのプロセッサがバンド幅を奪い合うこととなります。一部のハイエンド システムでは、高速データ バスの構築によってこの問題を解決しようとします。しかし、そのような解決策は高価であり、スケーラビリティが制限されます。

NUMA は、高性能接続を使用してコスト効率の高いいくつかの小型のノードをリンクする別の方法です。各ノードには、小型の SMP システムのように、プロセッサとメモリが含まれます。ただし、高度なメモリ コントローラにより、ノードはほかのすべてのノード上のメモリを使用でき、単一のシステム イメージを作成します。プロセッサ自身のノード内にはないメモリ (リモート メモリ) にプロセッサがアクセスする場合、ローカル メモリのアクセスよりも低速な NUMA 接続を通じてデータを転送する必要があります。テクノロジーの名前が示唆するように、メモリ アクセスの時間は一定ではなく、メモリの場所およびメモリにアクセスするノードによって異なります。

オペレーティング システムについての課題

NUMA アーキテクチャは、単一のシステム イメージを提供するので、特別な最適化を行わずにオペレーティング システムを実行できることがよくあります。

リモート メモリ アクセスの遅延が大きいと、データがローカル ノードに転送されるのを常に待つため、プロセッサの使用率が低いままになることがあり、メモリ バンド幅の要求の高いアプリケーションにとって NUMA 接続がボトルネックになることがあります。

さらに、そのようなシステムでのパフォーマンスは、変動が大きくなる場合があります。たとえば、1 回のベンチマーク実行においてローカルに配置されたメモリがアプリケーションにあるが、以後の実行ではそのすべてのメモリがリモート ノード上に配置される場合、パフォーマンスが変動します。この現象により、キャパシティ プランニングが困難になることがあります。

一部のハイエンド UNIX システムは、コンパイラとプログラミング ライブラリ内で NUMA 最適化のサポートを提供します。このサポートでは、ソフトウェア開発者が、最適なパフォーマンスのためにプログラムをチューニングおよび再コンパイルする必要があります。1 つのシステムを最適化しても、同じシステムの次の世代において適切に機能することは保証されません。ほかのシステムでは、アプリケーションを実行するノードを管理者が明示的に決定することが許可されています。これは、メモリの 100 パーセントがローカルであることを要求する特定のアプリケーションにとっては許容可能かもしれませんが、管理の負担が発生し、ワークロードが変化した場合にノード間で不均衡が発生することがあります。

理想的なのは、修正を行うことなくただちにアプリケーションが利益を得られるように、システム ソフトウェアが透過的な NUMA サポートを提供することです。システムは、管理者が常に介入する必要がなく、ローカル メモリを最大限に使用し、プログラムをインテリジェントにスケジュール設定する必要があります。最後に、公平性またはパフォーマンスを阻害することなく、状態の変化に適切に対応する必要があります。

ESXi の NUMA スケジュール設定の機能

ESXi は、高度な NUMA スケジューラを使用して、プロセッサの負荷とメモリの局所性またはプロセッサの負荷バランスを動的に均衡させます。

- 1 NUMA スケジューラによって管理される各仮想マシンは、ホーム ノードに割り当てられます。ホーム ノードは、システム リソース割り当てテーブル (SRAT) に示されたとおりにプロセッサとローカル メモリを含む、システムの NUMA ノードの 1 つです。
- 2 メモリが仮想マシンに割り当てられる場合、ESXi ホストは、ホーム ノードから優先的にメモリを割り当てます。仮想マシンの仮想 CPU は、メモリの局所性を最大限にするため、ホーム ノード上での実行に制限されません。
- 3 NUMA スケジューラは、システム負荷の変化に対応するために、仮想マシンのホーム ノードを動的に変更できます。NUMA スケジューラは、プロセッサの負荷不均衡を減らすために、新しいホーム ノードに仮想マシンを移行することがあります。このためにより多くのメモリがリモートになる場合があるので、NUMA スケジューラは、仮想マシンのメモリを動的に新しいホーム ノードに移行して、メモリの局所性を向上させることがあります。NUMA スケジューラは、ノード間で仮想マシンをスワップすると全体的なメモリの局所性が高まる場合、そのようなスワップを行うこともあります。

一部の仮想マシンは、ESXi の NUMA スケジューラによって管理されません。たとえば、仮想マシンのプロセッサまたはメモリ アフィニティを手動で設定する場合、NUMA スケジューラがその仮想マシンを管理できないことがあります。NUMA スケジューラによって管理されない仮想マシンであっても、正常に動作します。ただし、それらの仮想マシンは ESXi の NUMA 最適化の利点を得られません。

ESXi における NUMA のスケジューラ設定とメモリ配置のポリシーは、すべての仮想マシンを透過的に管理できるため、管理者は、ノード間で仮想マシンのバランスをとることの複雑さを明示的に処理する必要はありません。

最適化は、ゲスト OS のタイプにかかわらずにシームレスに機能します。ESXi は、NUMA ハードウェアをサポートしない仮想マシン（Windows NT 4.0 など）にも、NUMA サポートを提供します。結果的に、レガシー オペレーティング システムであっても新しいハードウェアを利用できます。

1つのハードウェア ノード上で使用可能な物理プロセッサ コアの数よりも多くの仮想プロセッサを持つ仮想マシンは、自動的に管理できます。NUMA スケジューラは、複数の NUMA ノードを利用することで、そのような仮想マシンに対応します。つまり、そのような仮想マシンを複数の NUMA クライアントに分割し、それぞれの NUMA クライアントに1つのノードに割り当ててから、通常の、分割されていないクライアントとしてスケジューラが管理します。これにより、局所性の高い、メモリを多用するワークロードのパフォーマンスを強化できます。この機能の動作の構成については、[仮想マシンの詳細属性](#) を参照してください。

ESXi 5.0 以降では、仮想 NUMA トポロジのゲスト OS への公開のサポートが含まれています。仮想 NUMA 制御の詳細については、[仮想 NUMA の使用](#) を参照してください。

VMware NUMA 最適化のアルゴリズムと設定

このセクションでは、アプリケーションのパフォーマンスを最大化しながらリソース保証を維持するために ESXi が使用するアルゴリズムと設定について説明します。

ホーム ノードと初期配置

仮想マシンがパワーオンされると、ESXi がその仮想マシンにホーム ノードを割り当てます。仮想マシンはホーム ノード内のプロセッサ上でのみ実行され、新しく割り当てられるメモリもホーム ノードから得られます。

仮想マシンのホーム ノードが変化しないかぎり、仮想マシンはローカル メモリだけを使用して、ほかの NUMA ノードへのリモート メモリ アクセスに伴うパフォーマンス上のペナルティを回避します。

仮想マシンがパワーオンされると初期ホーム ノードが割り当てられ、NUMA ノード間の CPU とメモリ負荷全体のバランスが保たれます。大規模な NUMA システムではノード間の待ち時間に大きな幅があるので、ESXi は起動時にノード間の待ち時間を確認し、単一の NUMA ノードよりワイドな仮想マシンを初期配置するときにこの情報を使用します。これらのワイド仮想マシンは、メモリ アクセスの待ち時間を最小限にするために、たがいに近い NUMA ノードに配置されます。

初期配置だけを考慮する方法でも、ベンチマーク構成などのように、システムが実行中のあいだ変化しない、1つのワークロードのみを実行するシステムには通常、十分です。しかし、この方法では、ワークロードが変化するデータセンタークラスのシステムに対して、良好なパフォーマンスと公平さを保証できません。そのため、初期配置に加え、ESXi 5.0 では、NUMA ノード間での仮想 CPU とメモリの動的移行を行い、CPU バランスの向上とメモリの局所性増加を図っています。

動的ロード バランシングとページの移行

ESXi は、従来の初期配置方法を動的な再バランシング アルゴリズムと統合します。システムは、定期的に（デフォルトでは 2 秒ごとに）さまざまなノードの負荷を調べて、仮想マシンをあるノードから別のノードに移動することによって負荷を再バランシングするべきかどうかを判定します。

公平性またはリソース資格に違反しないでパフォーマンスを向上させるために、この計算では、仮想マシンとリソース プールのリソース設定が考慮されます。

リバランサは、適切な仮想マシンを選択して、そのホーム ノードを負荷が最小のノードに変更します。可能な場合、リバランサは、すでに一部のメモリがターゲット ノード上に配置されている仮想マシンを移動します。そこから先は（仮想マシンが再び移動されないかぎり）、仮想マシンがその新しいホーム ノード上にメモリを割り当てて、新しいホーム ノード内のプロセッサ上でのみ実行されます。

再バランシングは、公平性を維持して、すべてのノードが完全に使用されるようにするための効果的な解決策です。リバランサは、メモリをほとんどまたはまったく割り当てなかったノードに仮想マシンを移動しなければならない場合があります。その場合、仮想マシンは、大量のリモート メモリ アクセスに伴うパフォーマンス上のペナルティを被ります。ESXi は、仮想マシンの元のノードから新しいホーム ノードへメモリを透過に移行することによって、このペナルティを排除できます。

- 1 システムは、元のノード上のページ（連続するメモリの 4KB）を選択し、そのデータをターゲット ノード内のページにコピーします。
- 2 システムは、ターゲット ノード上のページを以後のすべての参照に使用してリモート メモリ アクセスのペナルティを排除するように、仮想マシン監視レイヤーとプロセッサのメモリ管理ハードウェアを使用して仮想マシンのメモリ ビューをシームレスに再マッピングします。

仮想マシンが新しいノードに移動するとき、ESXi ホストはすぐに、この方法によるメモリの移行を開始します。仮想マシンのリモート メモリの残りがほとんどない場合やターゲット ノードの使用可能な空きメモリがほとんどない場合は特に、ESX/ESXi ホストは、この速度を管理してシステムに過度な負担をかけないようにします。このメモリ移行アルゴリズムは、仮想マシンが短期間だけ新しいノードに移動される場合は ESXi ホストがメモリを不必要に移動しないことも保証します。

初期配置、動的再バランシング、およびインテリジェントなメモリ移行が共同で動作する場合、ワークロードの変化があっても、NUMA システム上での良好なメモリ パフォーマンスがそれらの動作によって保証されます。主要なワークロードの変化が発生した場合、たとえば、新しい仮想マシンが起動された場合、システムは再調整に時間をかけて、仮想マシンとメモリを新しい場所に移行します。短期間（通常は数秒または数分）でシステムは再調整を完了し、安定状態に達します。

NUMA 用に最適化された透過的なページ共有

仮想マシン間でメモリを共有する機会を提供する ESXi ワークロードは、いくつもあります。

たとえば、いくつかの仮想マシンが、同じゲスト OS のインスタンスを実行していたり、同じアプリケーションまたはコンポーネントをロードしていたり、共通のデータを含んでいたりとすることがあります。このような場合、ESXi システムは、専用の透過的なページ共有技法を使用してメモリ ページの冗長コピーを安全に削除します。メモリの共有により、仮想マシンで実行しているワークロードが消費するメモリが、物理マシン上で実行する場合の消費量よりも少なくなることがよくあります。結果的に、より高いレベルのオーバーコミットを効率的にサポートできます。

ESXi システムの透過的なページ共有も、NUMA システム上で使用するために最適化されています。NUMA システムでは、ページはノードごとに共有されるので、各 NUMA ノードは、数多く共有されたページの独自のローカルコピーを持ちます。仮想マシンは、共有ページを使用する場合、リモートメモリにアクセスする必要がありません。

注： デフォルトの動作は、以前の ESX および ESXi のすべてのバージョンと同じです。

NUMA アーキテクチャでのリソース管理

さまざまなタイプの NUMA アーキテクチャでリソース管理を行うことができます。

マルチコア システムが広く普及したことにより、NUMA アーキテクチャの人気は一層高まりつつあります。このアーキテクチャが、メモリを多用するワークロードのパフォーマンスを強化できることがその理由です。最近の Intel および AMD のシステムには、いずれも NUMA サポートがプロセッサに組み込まれています。さらに、IBM Enterprise X アーキテクチャなどの従来の NUMA システムもあり、これらでは Intel および AMD のプロセッサが、専用のチップセット サポートを備えた NUMA の動作で拡張されます。

通常、NUMA の動作は BIOS 設定を使用して有効または無効にできます。たとえば、AMD Opteron ベースの HP ProLiant サーバでは、BIOS のノード インタリービングを有効にすることによって、NUMA を無効にできます。NUMA が有効な場合、BIOS でシステム リソース割り当てテーブル (SRAT) が構築され、このテーブルは、最適化に使用される NUMA 情報を ESXi が生成するために使用されます。スケジュール設定の公平性を確保するため、NUMA の最適化は、NUMA ノードあたりのコア数が少なすぎる、または全体のコア数が少なすぎるシステムに対しては有効になりません。この動作は、`numa.rebalancecorestotal` と `numa.rebalancecoresnode` の各オプションを使用して変更できます。

仮想 NUMA の使用

vSphere では、仮想 NUMA トポロジのゲスト OS への公開のサポートが含まれています。これにより、ゲスト OS の活用と NUMA の最適化の適用を介してパフォーマンスを改善できます。

仮想 NUMA トポロジは、ハードウェア バージョン 8 仮想マシンに使用でき、仮想 CPU の数が 8 を超えた場合にデフォルトで有効になります。また、詳細構成オプションを使用することで、仮想 NUMA トポロジを手動で調整できます。

仮想 NUMA が有効な仮想マシンを初めてパワーオンするとき、その仮想 NUMA トポロジは、基盤となる物理ホストの NUMA トポロジに基づいて決定されます。仮想マシンの仮想 NUMA トポロジは初期化されると、その仮想マシンの vCPU の数が変更されるまで変更されません。

仮想 NUMA トポロジでは、仮想マシンに対して構成されたメモリを考慮しません。また、仮想ソケットの数や仮想マシンのソケット当たりのコア数に影響されません。

仮想 NUMA トポロジをオーバーライドする必要がある場合は、[仮想 NUMA 制御](#)を参照してください。

注： CPU HotAdd を有効にすると、仮想 NUMA が無効になります。<https://kb.vmware.com/kb/2040375>を参照してください。

仮想 NUMA 制御

非常にメモリ使用量が多い仮想マシンに対し、詳細オプションを使用してデフォルトの仮想 CPU 設定をオーバーライドできます。

これらの詳細オプションは、仮想マシン構成ファイルに追加できます。

表 16-1. 仮想 NUMA 制御のための詳細オプション

オプション	説明	デフォルト値
<code>cpuid.coresPerSocket</code>	仮想 CPU ソケットあたりの仮想コア数の決定。 <code>numa.vcpu.followcorespersocket</code> が設定されている場合を除き、このオプションは仮想 NUMA トポロジには影響しません。	1
<code>numa.vcpu.maxPerVirtualNode</code>	この値を除数として vCPU の合計数を均等に分割して、仮想 NUMA ノードの数を決定します。	8
<code>numa.autosize.once</code>	これらの設定で仮想マシン テンプレートを作成した場合、その後デフォルト値 TRUE で仮想マシンをパワーオンするたびに、設定は同じになります。値が FALSE に設定されている場合、パワーオンするたびに仮想 NUMA トポロジは更新されます。仮想マシンの仮想 CPU の構成数が任意の時点で変更されると、仮想 NUMA トポロジは再評価されます。	TRUE
<code>numa.vcpu.min</code>	仮想 NUMA トポロジを生成するために必要な最小限の、仮想マシン内の仮想 CPU 数。サイズが <code>numa.vcpu.min</code> より小さい場合、仮想マシンは常に UMA です。	9
<code>numa.vcpu.followcorespersocket</code>	1 に設定した場合、 <code>cpuid.coresPerSocket</code> に関連付けられている仮想 NUMA ノードのサイジングの以前の動作に戻ります。	0

NUMA 制御の指定

メモリを大量に使用するアプリケーションがある場合、または仮想マシンの台数が少ない場合、仮想マシンの CPU とメモリの配置を明示的に指定することによってパフォーマンスを最適化できます。

コントロールを指定することは、仮想マシンがメモリ集約型のワークロード（メモリ内データベース、または大きいデータセットを持つ科学計算アプリケーションなど）を実行する場合に役に立ちます。システムのワークロードが単純で変化しないことがわかっている場合も、NUMA の配置を手動で最適化できます。たとえば、同様のワークロードを持つ 8 台の仮想マシンを実行する 8 プロセッサ システムは、明示的に簡単に最適化できます。

注： ほとんどの状況において、ESXi ホストの自動 NUMA 最適化で良好なパフォーマンスが得られます。

ESXi には NUMA 配置用のコントロールが 3 セット備わっているため、管理者が仮想マシンのメモリとプロセッサの配置を操作できます。

次のオプションを指定できます。

NUMA ノードのアフィニティ

このオプションを設定した場合、アフィニティに指定したノード上でのみ、NUMA は仮想マシンをスケジューリング設定できます。

CPU アフィニティ

このオプションを設定した場合、仮想マシンは、アフィニティに指定したプロセッサのみを使用します。

メモリ アフィニティ

このオプションを設定した場合、サーバは指定したノード上でのみメモリを割り当てます。

NUMA ノード アフィニティを指定した場合にも、仮想マシンは NUMA で管理されますが、そのマシンの仮想 CPU は、NUMA ノード アフィニティに指定したノード上でのみスケジューリング設定できます。同様に、NUMA ノード アフィニティに指定したノードからのみ、メモリを取得できます。CPU またはメモリのアフィニティを指定した場合、仮想マシンは、NUMA で管理されなくなります。これらの仮想マシンの NUMA 管理は、CPU およびメモリのアフィニティの制約を削除した場合は有効です。

手動での NUMA 配置は、プロセッサ リソースをシステム全体で均等に配布する ESXi のリソース管理アルゴリズムに干渉する可能性があります。たとえば、プロセッサ集約型のワークロードを持つ 10 台の仮想マシンを 1 つのノードに手動で配置し、2 台の仮想マシンだけを手動で別のノードに配置した場合、システムは、12 台の仮想マシンすべてに均等にシステム リソースのシェアを与えることはできません。

仮想マシンと特定のプロセッサとの関連付け

仮想マシンの仮想 CPU を固定のプロセッサに関連付けることによって、仮想マシン上のアプリケーションのパフォーマンスを向上できることがあります。これにより、仮想 CPU が NUMA ノード間で移行されないようにできます。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
 - a データセンター、フォルダ、クラスター、リソース プール、またはホストを選択して、仮想マシンを検索します。
 - b [仮想マシン] タブをクリックします。
- 2 仮想マシンを右クリックし、[設定の編集] をクリックします。
- 3 [仮想ハードウェア] タブをクリックし、[CPU] を展開します。
- 4 [スケジューリング設定のアフィニティ] で CPU アフィニティを優先するプロセッサに設定します。

注： NUMA ノード内のすべてのプロセッサを、手動で選択する必要があります。CPU アフィニティは、ノード単位ではなくプロセッサ単位で指定します。

メモリ アフィニティを使用した、メモリ割り当てと特定の NUMA ノードとの関連付け

仮想マシン上の将来のすべてのメモリ割り当てが、特定の NUMA ノードに関連付けられたページを使用することを指定できます（手動メモリ アフィニティとも呼ばれます）。

注： 将来のメモリ割り当てに使用するノードを指定するのは、CPU アフィニティも指定した場合だけにしてください。メモリ アフィニティの設定だけを手動で変更すると、自動 NUMA 再バランシングが適切に動作しなくなります。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [設定] をクリックし、[仮想マシンのハードウェア] をクリックします。
- 4 [編集] をクリックします。
- 5 [仮想ハードウェア] タブを選択し、[メモリ] を展開します。
- 6 [NUMA メモリのアフィニティ] で、メモリのアフィニティを設定します。

例：単一の NUMA ノードへの仮想マシンのバインド

次の例は、8 方向サーバ上の 2 方向の仮想マシン用に最後の 4 つの物理 CPU を単一の NUMA ノードに手動バインドする場合を示しています。

CPU（たとえば、4、5、6、7）は、物理 CPU 番号です。

- 1 vSphere Client で、仮想マシンを右クリックし、[設定の編集] を選択します。
- 2 [オプション] を選択し、[詳細] をクリックします。
- 3 [構成パラメータ] ボタンをクリックします。
- 4 vSphere Client で、プロセッサ 4、5、6、および 7 の CPU のアフィニティをオンにします。

次に、この仮想マシンをノード 1 上でのみ実行するとします。

- 1 vSphere Client のインベントリ パネルで仮想マシンを選択し、[設定の編集] を選択します。
- 2 [オプション] を選択し、[詳細] をクリックします。
- 3 [構成パラメータ] ボタンをクリックします。
- 4 vSphere Client で、NUMA ノードのメモリ アフィニティを 1 に設定します。

これらの 2 つのタスクを完了すると、仮想マシンが NUMA ノード 1 でのみ実行されるようになり、可能な場合、同じノードからメモリを割り当てるようになります。

仮想マシンと指定 NUMA ノードの関連づけ

NUMA ノード アフィニティを指定するために NUMA ノードを仮想マシンに関連付けた場合、ESXi が仮想マシンの仮想 CPU およびメモリのスケジュールを設定できる NUMA ノードを特定のノード セットに制限することができます。

注： NUMA ノードのアフィニティを制限すると、NUMA ノード全体の仮想マシンの負荷が公平になるように再分散するための ESXi NUMA スケジューラの機能に影響を与えることがあります。NUMA ノードのアフィニティを指定する前には必ず、再分散の問題を検討してください。

手順

- 1 vSphere Client で仮想マシンを参照して移動します。
- 2 [構成] タブをクリックし、[設定] をクリックします。

- 3 [仮想マシン オプション] の [編集] ボタンをクリックします。
- 4 [仮想マシン オプション] タブをクリックし、[詳細] を展開します。
- 5 [構成パラメータ] の [構成の編集] ボタンをクリックします。
- 6 [行の追加] をクリックし、新しいオプションを追加します。
- 7
 - 仮想マシンの NUMA ノードを指定するには、[名前] 列に「**numa.nodeAffinity**」と入力します。
 - 仮想マシンの特定の仮想 NUMA ノードに対して NUMA ノードを指定するには、[名前] 列に「**sched.nodeX.affinity**」と入力します。ここで、X は仮想 NUMA ノードの番号です。たとえば、「**sched.node0.affinity**」は、仮想マシン上の仮想 NUMA ノード 0 を指定します。
- 8 値の列に、仮想マシンまたは仮想 NUMA ノードをスケジュール設定できる NUMA ノードを入力します。
複数のノードを指定するには、コンマで区切ったリスト形式を使用します。たとえば、仮想マシンのリソース スケジュール設定を NUMA ノード 0 および 1 に制限するには、「**0,1**」 と入力します。
- 9 [OK] をクリックします。
- 10 [OK] をクリックして、[仮想マシンの編集] ダイアログ ボックスを閉じます。

ホストまたは個々の仮想マシンの詳細属性を設定すると、リソース管理をカスタマイズできます。

ほとんどの場合、基本的なリソース割り当て設定（[予約]、[制限]、[シェア]）を変更するか、デフォルト設定をそのまま使用することで、適切なリソース割り当てが行われます。ただし、詳細属性を使用すると、ホストまたは特定の仮想マシンのリソース管理をカスタマイズできます。

この章には、次のトピックが含まれています。

- ホストの詳細属性の設定
- 仮想マシンの詳細属性の設定
- 遅延感度
- 信頼性の高いメモリについて
- 1 GB ページでのゲスト vRAM のバッキング

ホストの詳細属性の設定

ホストの詳細属性を設定できます。

注意： 詳細オプションの変更はサポート外とみなされます。通常、デフォルトの設定で最適な結果になります。詳細オプションの変更は、VMware テクニカル サポートまたはナレッジベースの記事から具体的な手順を確認した場合にのみ行います。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [設定] タブをクリックします。
- 3 [システム] メニューの [システムの詳細設定] をクリックします。
- 4 [システムの詳細設定] から適切な項目を選択します。
- 5 [編集] ボタンをクリックして、値を変更します。
- 6 [OK] をクリックします。

メモリの詳細属性

メモリの詳細属性を使用して、メモリのリソース使用率をカスタマイズできます。

表 17-1. メモリの詳細属性

属性	説明	デフォルト
Mem.ShareForceSalting	<p>Mem.ShareForceSalting 0 : 仮想マシン間の透過的なページ共有 (TPS) の動作は引き続き保持されます。VMX オプション <code>sched.mem.pshare.salt</code> の値は存在していても無視されます。</p> <p>Mem.ShareForceSalting 1 : デフォルトでは、ソルトの値は <code>sched.mem.pshare.salt</code> から取得されます。指定しない場合、仮想マシンのソルトの値が 0 とみなされて古い TPS (仮想マシン間) の動作にフォールバックします。</p> <p>Mem.ShareForceSalting 2 : デフォルトでは、ソルトの値は <code>sched.mem.pshare.salt</code> (存在する場合) または <code>vc.uuid</code> から取得されます。存在しない場合、ページ共有アルゴリズムによって、仮想マシンごとにソルティングのためのランダムな一意の値が生成されます。この値をユーザーが構成することはできません。</p>	2
Mem.SamplePeriod	作業セット サイズを見積もるためにメモリ動作が監視される、定期的な時間間隔を指定します。これは仮想マシン実行時間の秒数で測定されます。	60
Mem.BalancePeriod	メモリの自動再割り当ての定期的な時間間隔を秒単位で指定します。空きメモリ量の大幅な変更によっても、再割り当てが開始されます。	15
Mem.IdleTax	アイドル メモリ税率をパーセンテージで指定します。この税は、事実上、仮想マシンが有効に使用しているメモリよりも多くのアイドル メモリを仮想マシンに請求します。税率が 0 パーセントの場合は、作業セットを無視する割り当てポリシーを定義し、シェアに基づいて厳密にメモリを割り当てます。税率を高くすると、非生産的にアイドル メモリを蓄えている仮想マシンからアイドル メモリを再割り当てできる割り当てポリシーになります。	75
Mem.ShareScanGHz	使用可能なホストの CPU リソース 1 GHz あたりについて、ページを共有するタイミングを調べるためにスキャンするメモリ ページの最大量 (1 秒あたり) を指定します。たとえば、デフォルトを 1 GHz あたり 4MB/秒にします。	4
Mem.ShareScanTime	ページを共有するタイミングを調べるために、仮想マシン全体をスキャンする時間を分単位で指定します。デフォルトは 60 分です。	60
Mem.CtlMaxPercent	メモリ バルーン ドライバ (<code>vmmemctl</code>) を使用して、任意の仮想マシンから解放されるメモリの最大量を、構成されたメモリ サイズのパーセンテージに基づいて制限します。0 を指定すると、すべての仮想マシンで解放が無効になります。	65
Mem.AllocGuestLargePage	ホストの大きなページでのゲストの大きなページのバッキングを可能にします。TLB ミスを減らし、ゲストの大きなページを使用するサーバ ワークロードのパフォーマンスを向上させます。0 = 無効。	1
Mem.AllocUsePSharePool および Mem.AllocUseGuestPool	ゲストの大きなページをホストの大きなページでバッキングできる可能性を高めることで、メモリの断片化を低減します。ホスト メモリが断片化されていると、ホストの大きなページの利用可能率が低下します。0 = 無効。	15
Mem.MemZipEnable	ホストに対してメモリ圧縮を有効にします。0 = 無効。	1
Mem.MemZipMaxPct	圧縮メモリとして格納できる各仮想マシンのメモリの最大のパーセンテージとして、圧縮キャッシュの最大サイズを指定します。	10
LPage.LPageDefragEnable	大容量ページのデフラグを有効にします。0 = 無効。	1
LPage.LPageDefragRateVM	1 台の仮想マシンが 1 秒間に試みる大容量ページのデフラグの最大回数。可能な値の範囲は 1 ~ 1024 です。	32

表 17-1. メモリの詳細属性 (続き)

属性	説明	デフォルト
LPage.LPageDefragRateTotal	1 秒間に試みる大容量ページのデフラグの最大回数。可能な値の範囲は 1 ~ 10240 です。	256
LPage.LPageAlwaysTryForNPT	ネストしたページ テーブル (AMD 社では「RVI」、Intel 社では「EPT」という名称) に大きなページの割り当てを試みます。このオプションを有効にすると、すべてのゲスト メモリは、ネストしたページ テーブルを使用するマシン (たとえば、AMD Barcelona など) で大きなページによって補助されます。NPT が使用可能でない場合は、ゲスト メモリの一部のみが大きなページで補助されます。0 = 無効。	1

NUMA の詳細属性

NUMA の詳細属性を使用して、NUMA の使用率をカスタマイズできます。

表 17-2. NUMA の詳細属性

属性	説明	デフォルト
Numa.RebalancePeriod	ミリ秒単位で指定される、再バランス期間の頻度を制御します。再バランスの頻度を上げると、実行中の仮想マシンを多数持つマシンにおいて特に、CPU オーバーヘッドが増大することがあります。再バランスの頻度を上げると、より適正になる場合もあります。	2000
Numa.MigImbalanceThreshold	NUMA リバランサは、各仮想マシンの CPU 時間資格と実際の消費との間の差を考慮して、ノード間の CPU 不均衡を計算します。このオプションは、仮想マシンの移行を起動するのに必要なノード間の最小負荷不均衡をパーセントで制御します。	10
Numa.RebalanceEnable	NUMA 再バランスおよびスケジュール設定を有効にします。すべての NUMA 再バランスと仮想マシンの初期配置を無効にして、NUMA スケジュール設定システムを実質的に無効にするには、このオプションを 0 に設定します。	1
Numa.RebalanceCoresTotal	NUMA リバランサを有効にするために必要な、ホスト上の合計プロセッサ コアの最小数を指定します。	4
Numa.RebalanceCoresNode	NUMA リバランサを有効にするために必要な、ノードあたりのプロセッサ コアの最小数を指定します。 このオプションと Numa.RebalanceCoresTotal は、NUMA 再バランスが有効な場合に合計のまたはノードあたりの少数のプロセッサがスケジュール設定の公平性を解決できる、小型の NUMA 構成 (たとえば、2 方向 Opteron ホスト) 上で NUMA 再バランスを無効にする場合に役に立ちます。	2
Numa.AutoMemAffinity	CPU アフィニティ セットを持つ仮想マシンのメモリ アフィニティを自動的に設定します。	1
Numa.PageMigEnable	メモリの局所性を改善するための NUMA ノード間での自動的なページ移行を行います。その場合も、手動で設定されたページ移行速度は有効です。	1

仮想マシンの詳細属性の設定

仮想マシンの詳細属性を設定できます。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
 - a データセンター、フォルダ、クラスタ、リソース プール、またはホストを選択して、仮想マシンを検索します。
 - b [仮想マシン] タブをクリックします。
- 2 仮想マシンを右クリックし、[設定の編集] を選択します。
- 3 [仮想マシン オプション] をクリックします。
- 4 [詳細] を展開します。
- 5 [構成パラメータ] の [構成の編集] をクリックします。
- 6 表示されるダイアログ ボックスで [行の追加] をクリックして、新しいパラメータとその値を入力します。
- 7 [OK] をクリックします。

仮想マシンの詳細属性

仮想マシンの詳細属性を使用して、仮想マシンの構成をカスタマイズできます。

表 17-3. 仮想マシンの詳細属性

属性	説明	デフォルト
sched.mem.maxmem ctl	選択した仮想マシンからバルーニングによって解放される、メガバイト (MB) 単位での最大メモリ量です。ESXi ホストが追加メモリを解放しなければならない場合、ESXi ホストはスワップを強制されます。スワップはバルーニングよりも好ましくありません。	-1 (制限なし)
sched.mem.pshare.en able	選択した仮想マシンのメモリ共有を有効にします。 このブール値のデフォルトは [True] です。仮想マシンに対して [False] に設定すると、メモリ共有がオフになります。	True
sched.mem.pshare.sa lit	ソルト値は、仮想マシンごとに構成可能な VMX オプションです。仮想マシンの VMX ファイルにこのオプションが存在しない場合は、 vc.uuid vmx オプションの値がデフォルト値として使用されます。 vc.uuid は各仮想マシンに一意的な値であるため、透過的なページ共有は、デフォルトでは、特定の仮想マシンに属するページ間 (仮想マシン内) でのみ行われます。仮想マシンのグループが信頼できるとみなされる場合は、グループ内のすべての仮想マシン (仮想マシン間) に共通のソルト値を設定することによって、グループ内でページを共有できます。	ユーザーが構成可能

表 17-3. 仮想マシンの詳細属性（続き）

属性	説明	デフォルト
sched.swap.persist	仮想マシンがパワーオフになった場合に仮想マシンのスワップ ファイルを残すか、削除するかを指定します。デフォルトでは、システムは、仮想マシンがパワーオンになった時点でその仮想マシンのスワップ ファイルを作成し、仮想マシンがパワーオフになった時点でスワップ ファイルを削除します。	False
sched.swap.dir	仮想マシンのスワップ ファイルのディレクトリの場所です。デフォルトでは、仮想マシンの作業ディレクトリになります。つまり、仮想マシンの構成ファイルを格納するディレクトリです。このディレクトリは、仮想マシンにアクセス可能なホスト上にある必要があります。仮想マシン（または仮想マシンから作成されたクローン）を移動する場合は、この属性をリセットする必要があります。	workingDir に等しい

仮想 NUMA の詳細属性

仮想 NUMA の詳細属性を使用して、仮想 NUMA の使用率をカスタマイズできます。

表 17-4. NUMA の詳細属性

属性	説明	デフォルト
cpuid.coresPerSocket	仮想 CPU ソケットあたりの仮想コア数の決定。この値が 2 以上の場合、仮想マシンに仮想 NUMA トポロジがあれば、仮想 NUMA ノードのサイズも決定します。各物理ホストの仮想 NUMA トポロジを正確に知っている場合、このオプションを設定できます。	1
numa.autosize	このオプションを設定するとき、仮想 NUMA トポロジには、仮想ノードあたり、各物理ノードのコア数と同数の仮想 CPU があります。	FALSE
numa.autosize.once	これらの設定で仮想マシンのテンプレートを作成した場合、その後仮想マシンをパワーオンするたびに、設定は同じであることが保証されます。仮想 NUMA トポロジは、仮想マシンの仮想 CPU の構成数に変更されると再評価されます。	TRUE
numa.vcpu.maxPerVirtualNode	cpuid.coresPerSocket は 2 のべき乗で制約が強すぎる場合、直接、numa.vcpu.maxPerVirtualNode を設定できます。このケースでは、cpuid.coresPerSocket は設定しません。	8
numa.vcpu.min	仮想 NUMA トポロジを生成するために必要な最小限の、仮想マシン内の仮想 CPU 数。	9
numa.vcpu.maxPerMachineNode	NUMA ノードで同時にスケジュール可能な、同一の仮想マシンに属する最大仮想 CPU 数。この属性を使用することにより、別の NUMA ノードの別の NUMA クライアントに強制して、確実に最大の帯域幅を得ます。	仮想マシンが実行中の物理ホストの、ノードあたりのコア数。
numa.vcpu.maxPerClient	NUMA クライアント中の最大仮想 CPU 数。1つのクライアントは、単一エンティティとして NUMA に管理された仮想 CPU のグループです。デフォルトでは、各仮想 NUMA ノードは NUMA クライアントです。ただし、ある仮想 NUMA ノードが物理的な NUMA ノードより大きい場合、単一の仮想 NUMA ノードを複数の NUMA クライアントでバックアップできます。	numa.vcpu.maxPerMachineNode に等しい

表 17-4. NUMA の詳細属性 (続き)

属性	説明	デフォルト
numa.nodeAffinity	<p>仮想マシンの仮想 CPU とメモリがスケジュールできる NUMA ノードのセットに制約を加えます。</p> <p>注: NUMA ノードのアフィニティを制限すると、NUMA ノード全体の仮想マシンを再バランスする NUMA スケジューラの機能を妨害している場合があります。NUMA ノードのアフィニティを指定する前には必ず、再バランスの問題を検討してください。</p>	
numa.mem.interleave	1 台の仮想マシンに割り当てられたメモリが、それが構成する NUMA クライアントが実行されているすべての NUMA ノードで静的にインターリーブされるかどうかと、公開されている仮想 NUMA トポロジがないかどうかを指定します。	True

遅延感度

仮想マシンの遅延感度は、遅延の影響を受けるアプリケーションのスケジュールリング遅延を最適化するように調整できます。

ESXi は優れたスループットを提供できるように最適化されています。仮想マシンは、遅延の影響を受けるアプリケーションの低遅延要件を満たすように最適化できます。遅延の影響を受けるアプリケーションの例には、VOIP やメディア プレーヤなどのほか、マウスとキーボードに頻繁にアクセスするアプリケーションがあります。

遅延感度の調整

仮想マシンの遅延感度を調整します。

前提条件

ESXi 6.7 では、遅延感度を [高] に設定すると、ハードウェア バージョン 14 を使用する仮想マシンをパワーオンするためには CPU の完全な予約が必要となります。

手順

- 1 vSphere Client で、仮想マシンを参照して移動します。
 - a データセンター、フォルダ、クラスタ、リソース プール、またはホストを選択して、仮想マシンを検索します。
 - b [仮想マシン] タブをクリックします。
- 2 仮想マシンを右クリックし、[設定の編集] をクリックします。
- 3 [仮想マシン オプション] をクリックし、[詳細] をクリックします。
- 4 [遅延感度] ドロップダウン メニューから設定を選択します。
- 5 [OK] をクリックします。

信頼性の高いメモリについて

ESXi は信頼性の高いメモリをサポートします。

一部のシステムには、信頼性の高いメモリ（システム内の他のメモリ部分よりも、ハードウェア メモリ エラーが発生しにくいメモリ部分）があります。さまざまな信頼性レベルに関する情報がハードウェアで公開されている場合、ESXi のシステム信頼性を高めることができる可能性があります。

信頼性の高いメモリの表示

信頼性の高いメモリがライセンスで許可されているかどうかを確認できます。

手順

- 1 vSphere Client で、ホストを参照して移動します。
- 2 [構成] タブをクリックし、[システム] をクリックします。
- 3 [ライセンス] を選択します。
- 4 [ライセンス機能] で、信頼性の高いメモリが表示されていることを確認します。

次のステップ

ESXCLI `hardware memory get` コマンドを使用して、信頼性の高いメモリと見なされているメモリ容量を確認できます。

1 GB ページでのゲスト vRAM のバッキング

vSphere 6.7 ESXi は、1 GB ページでのゲスト vRAM のバッキングに対して限定的なサポートを提供しています。

ゲスト メモリをバッキングするために 1 GB ページを使用するには、仮想マシンにオプション `sched.mem.lpage.enable1GPage = "TRUE"` を適用する必要があります。[設定の編集] を選択すると、詳細オプションでこれを設定できます。1 GB ページを有効にできるのは、パワーオフされている仮想マシンのみです。

1 GB ページが有効な仮想マシンには、十分なメモリ予約が必要です。そうしないと、仮想マシンはパワーオンできません。1 GB ページが有効な仮想マシン用のすべての vRAM は、パワーオン時に事前に割り当てられます。これらの仮想マシンには十分なメモリ予約があるため、メモリの解放の影響は受けず、メモリ消費は仮想マシンの有効期間全体に対して最長レベルを維持します。

1 GB ページでの vRAM のバッキングは便宜的で、1 GB ページはベスト エフォートで割り当てられます。これには、ホスト CPU が 1 GB の機能を持たない場合も含まれます。ゲスト vRAM を 1 GB ページでバッキングする機会を最大化するために、新規に起動したホストで 1 GB ページを必要とする仮想マシンを起動することを推奨します。これは、時間の経過とともにホスト RAM が断片化されるためです。

1 GB ページが有効な仮想マシンは、別のホストに移行できます。ただし、1 GB のページ サイズは、ソース ホストと同じ量でターゲット ホストに割り当てられないことがあります。また、ソース ホスト上の 1 GB ページでバッキングされている vRAM の一部が、ターゲット ホスト上の 1 GB ページでバッキングされていない場合もあります。

1 GB ページの便宜的な特徴は、1 GB ページでの vRAM のバッキングを保持しない可能性がある HA および DRS などの vSphere サービスにも影響します。これらのサービスは、ターゲット ホストの 1 GB の機能を認識せず、適正な判断を下す際に 1 GB でのメモリのバッキングを考慮しません。

障害の定義

18

DRS の障害が発生すると、DRS アクションが生成できない理由（またはマニュアル モードによる推奨アクション）が表示されます。

このセクションで DRS の障害を定義します。

注： この章では、「メモリ」は物理 RAM または永続的なメモリを指します。

この章には、次のトピックが含まれています。

- 仮想マシンが固定されている
- 仮想マシンがいずれのホストとも互換性がない
- 異なるホストへの移動時に仮想マシン間の DRS ルール違反が発生する
- ホストに仮想マシンとの互換性がない
- 仮想マシン間の DRS ルールに違反する仮想マシンがホストにある
- 仮想マシン用の十分なキャパシティがホストにない
- ホストの状態が正しくない
- 仮想マシン用の物理 CPU の数がホストで不足している
- 各仮想マシン CPU 用の十分なキャパシティがホストにない
- 仮想マシンが vMotion 処理中である
- クラスタにアクティブなホストがない
- リソース不足
- HA 用として構成されたフェイルオーバー レベルを満たすための十分なリソースがない
- 互換性のあるハード アフィニティ ホストがない
- 互換性のあるソフト アフィニティ ホストがない
- ソフト ルール違反の修正が許可されていない
- ソフト ルール違反の修正による影響

仮想マシンが固定されている

DRS が仮想マシンで無効になっているためにその仮想マシンを移動できない場合、この障害が発生します。つまり、その仮想マシンは、登録先のホストに「固定」されています。

仮想マシンがいずれのホストとも互換性がない

仮想マシンを実行できるホストを DRS が見つけられない場合、この障害が発生します。

これはたとえば、仮想マシンの CPU またはメモリ リソースのニーズを満たせるホストがない場合、または仮想マシンで必要となるネットワーク アクセスまたはストレージ アクセスを提供できるホストがない場合に発生することがあります。

この問題に対処するには、仮想マシンの要件を満たすホストを用意します。

異なるホストへの移動時に仮想マシン間の DRS ルール違反が発生する

複数の仮想マシンが同じホスト上で実行され、別のホストに移動できないアフィニティ ルールを互いに持っている場合、この障害が発生します。

これは、すべての仮想マシンが現在のホストから vMotion で移動できるわけではないため発生します。たとえば、グループにある仮想マシンの 1 つで DRS が無効になっている場合です。

これを回避するには、このグループの一部の仮想マシンが vMotion で移動できない原因を確認します。

ホストに仮想マシンとの互換性がない

DRS が仮想マシンをホストに移行すると判断した場合に、そのホストに仮想マシンとの互換性がないことが検出されると、この障害が発生します。

これは、仮想マシンで必要となるネットワークまたはストレージへの接続がターゲットとなるホストにないため発生している可能性があります。この障害が発生する別の理由としては、ターゲットとなるホストの CPU が現在のホストと大きく異なる場合、ホスト間での vMotion がサポートされない点が挙げられます。

これを回避するには、すべてのホストの構成が統一されるようクラスタを作成し、vMotion がホスト間で互換性を持つようにします。

ホストが仮想マシンと互換性がない別の理由として、仮想マシンとホスト間で必要な DRS ルールが存在し、それによって DRS が特定の仮想マシンを特定のホスト上に配置できないようになっている点があります。

仮想マシン間の DRS ルールに違反する仮想マシンがホストにある

仮想マシンがパワーオンされたり、vMotion の実行によって移動されたりしたときに、仮想マシン間の DRS ルールに違反すると、この障害が発生します。

仮想マシンのパワーオンまたは vMotion での移動は手動で実行できますが、vCenter Server による自動実行はできません。

仮想マシン用の十分なキャパシティがホストにない

仮想マシンを実行するための十分な CPU またはメモリ キャパシティがホストにない場合、この障害が発生します。

ホストの状態が正しくない

DRS によるアクションが必要な場合に、ホストがメンテナンス状態またはスタンバイ状態に切り替わると、この障害が発生します。

この障害に対処するには、ホストをスタンバイ モードまたはメンテナンス モードに切り替える要求をキャンセルします。

仮想マシン用の物理 CPU の数がホストで不足している

仮想マシン内の仮想 CPU 数に対応できる十分な CPU (ハイパースレッド) がホストのハードウェアにない場合、この障害が発生します。

各仮想マシン CPU 用の十分なキャパシティがホストにない

仮想マシンを実行するのに十分な CPU キャパシティがホストにない場合、この障害が発生します。

仮想マシンが vMotion 処理中である

仮想マシンが vMotion 処理中であるために DRS が仮想マシンを移動できない場合、この障害が発生します。

クラスタにアクティブなホストがない

仮想マシンの移動先となるクラスタに、メンテナンス モード以外で接続中のホストがない場合、この障害が発生します。

たとえば、すべてのホストが切断されている場合またはメンテナンス モードの場合、この問題が発生する可能性があります。

リソース不足

行なおうとした処理がリソースの構成ポリシーと競合している場合、この障害が発生します。

たとえば、リソース プールに割り当てられたメモリよりも多くのメモリがパワーオン処理で予約されている場合、この障害が発生することがあります。

リソースを調整してメモリを増やしたあと、再度この処理を試してください。

HA 用として構成されたフェイルオーバー レベルを満たすための十分なリソースがない

この障害は、フェイルオーバー用に予約された CPU またはメモリ リソースの HA 構成に違反がある場合、またはそれが DRS 操作には不十分である場合に発生します。

この障害は次のような場合に発生します。

- ホストがメンテナンス モードまたはスタンバイ モードに切り替わるよう要求された場合。
- 仮想マシンがパワーオンしようとしたときにフェイルオーバー違反が発生した場合。

互換性のあるハード アフィニティ ホストがない

必須となっている仮想マシンとホスト間の DRS アフィニティ ルールまたは非アフィニティ ルールを満たすホストを仮想マシンで利用できません。

互換性のあるソフト アフィニティ ホストがない

望ましい仮想マシンとホスト間の DRS アフィニティ ルールまたは非アフィニティ ルールを満たすホストを仮想マシンで利用できません。

ソフト ルール違反の修正が許可されていない

DRS 移行のしきい値が必須のみに設定されている。

これにより、必須ではない仮想マシンとホスト間の DRS アフィニティ ルールを修正する DRS アクションの生成が許可されません。

ソフト ルール違反の修正による影響

修正によってパフォーマンスに影響するため、必須ではない仮想マシンとホスト間の DRS アフィニティ ルールの修正が発生しません。

本オンライン ヘルプでは、vSphere[®] DRS (Distributed Resource Scheduler) の問題を次の特定のカテゴリごとに説明します： クラスタ、ホスト、および仮想マシンの問題。

注： この章では、「メモリ」は物理 RAM または永続的なメモリを指します。

この章には、次のトピックが含まれています。

- クラスタの問題
- ホストの問題
- 仮想マシンの問題

クラスタの問題

クラスタの問題により、DRS が適切に動作しなくなったり、不具合を報告しなくなる場合があります。

クラスタでの負荷不均衡

クラスタに、リソースの負荷不均衡があります。

問題

仮想マシンからの不均一なリソース要求およびホストの不均等なキャパシティが原因で、クラスタが不均衡な状態になる可能性があります。

原因

クラスタで負荷不均衡が発生する原因として考えられる内容は次のとおりです。

- 移行のしきい値が高すぎる。
しきい値が高いと、クラスタで負荷不均衡が発生する可能性も高くなります。
- 仮想マシン間の DRS ルールまたは仮想マシンとホスト間の DRS ルールによって、仮想マシンを移動できない。
- 少なくとも 1 台の仮想マシンで DRS が無効になっている。
- デバイスが 1 台以上の仮想マシンにマウントされているため、DRS が仮想マシンを移動して負荷を分散できない。

- DRS によって仮想マシンが移行されるホストとそれらの仮想マシンに互換性がない。つまり、移行しようとする仮想マシンに対して互換性のないホストが、クラスタ内に少なくとも 1 台あります。たとえば、ホスト A の CPU がホスト B の CPU に対して vMotion 互換でない場合、ホスト A はホスト B で実行されているパワーオン状態の仮想マシンと互換性がなくなります。
- 仮想マシンを移動させることは、現在置かれている場所で実行することよりも、仮想マシンのパフォーマンスにとって悪影響が大きくなる場合がある。これは負荷が安定しなかったり、仮想マシンの移動によって得られるメリットよりも移行コストのほうが高くなったりする場合に該当します。
- vMotion を有効にできない、またはクラスタ内のホストに対して設定されていない。

解決方法

負荷不均衡の原因となっている問題を解決します。

クラスタが黄色で表示される

リソース不足により、クラスタが黄色になっています。

問題

すべてのリソース プールおよび仮想マシンの予約を確保できるほど十分なリソースがクラスタにはないものの、実行中の仮想マシンの予約量を満たすには十分なリソースがある場合、DRS は動作し続け、クラスタは黄色に変わります。

原因

ホストのリソースがクラスタから削除された場合（ホストの故障など）にクラスタが黄色になることがあります。

解決方法

ホスト リソースをクラスタに追加するか、リソース プールの予約量を減らします。

リソース プールの不整合によりクラスタが赤色で表示される

DRS クラスタは無効なときに赤くなります。リソース プールのツリーが内部で不整合になっている場合に赤くなる場合があります。

問題

クラスタのリソース プールが内部で不整合になっている（子の予約量が親のプールの拡張不能な予約量よりも大きい場合など）と、クラスタは実行中の仮想マシンすべての予約を満たすために十分なリソースを確保できないため、クラスタが赤くなります。

原因

これは、vCenter Server が使用できない場合、または仮想マシンがフェイルオーバー状態にあるときにリソース プール設定が変更された場合に発生します。

解決方法

関係している変更を元に戻すか、リソース プールの設定を変更します。

フェイルオーバー キャパシティに違反しているためクラスタが赤色で表示される

DRS クラスタは無効なときに赤くなります。フェイルオーバー キャパシティに違反している場合に赤くなる可能性があります。

問題

ホストで障害が発生した場合、クラスタは仮想マシンをフェイルオーバーしようとしています。しかし、フェイルオーバー要件を満たしていても、すべての仮想マシンをフェイルオーバーできるほどのリソースを十分に確保できない可能性があります。

原因

HA 用として有効にされているクラスタが大量のリソースを損失しているためにフェイルオーバー要件を満たすことができない場合、メッセージが表示され、クラスタのステータスが赤に変更されます。

解決方法

クラスタのサマリ ページの最上部にある黄色のボックスで、構成に関する問題のリストを確認し、その問題の原因となっている問題を解決します。

クラスタの総負荷が低い場合、ホストがパワーオフされない

クラスタの総負荷が低い場合に、ホストがパワーオフされません。

問題

HA フェイルオーバー予約用として追加のキャパシティが必要なため、クラスタの総負荷が低いと、ホストがパワーオフされません。

原因

次の理由でホストをパワーオフできない場合があります。

- `MinPoweredOn{Cpu|Memory}Capacity` 詳細オプション設定の条件を満たす必要がある。
- 仮想マシンのリソース予約、仮想マシンとホスト間の DRS ルール、仮想マシン間の DRS ルール、DRS 非対応、使用可能なキャパシティを持つホストとの非互換が原因で、仮想マシンが少数のホスト上に統合できない。
- 負荷が安定しない。
- DRS で設定されている移行のしきい値が高すぎるため、必須の移行しか実行できない。
- vMotion が構成されていないために実行できない。
- パワーオフになっている可能性があるホスト上で DPM が無効になっている。
- 仮想マシンとホストとで互換性がないため、仮想マシンをほかのホストへ移行できない。
- ホストが Wake On LAN、IPMI、または iLO テクノロジーに対応していない。DPM がホストをスタンバイに切り替えるためには、このどちらかが必要です。

解決方法

クラスタの総負荷が低い場合に、ホストがパワーオフされない原因となっている問題を解決します。

クラスタの総負荷が高い場合、ホストがパワーオフされる

クラスタの総負荷が高い場合に、ホストがパワーオフされます。

問題

ホストや仮想マシンのパフォーマンスを低下させることなく、より少数のホスト上で仮想マシンを実行できると DRS が判断しました。また DRS には、使用率の高いホスト上からパワーオフがスケジュールされているホスト上に仮想マシンを移動するという制約があります。

原因

クラスタの総負荷が高くなりすぎています。

解決方法

クラスタの負荷を減らします。

DRS で vMotion の移行が実行されることがほとんど、またはまったくない

DRS で vMotion の移行がほとんど、またはまったく実行されません。

問題

DRS で vMotion の移行が実行されません。

原因

以下の問題のいずれか、または複数がクラスタ上に存在すると、DRS で vMotion の移行がまったく実行されません。

- DRS がクラスタ上で無効になっている。
- ホストに共有ストレージがない。
- クラスタのホストに vMotion ネットワークがない。
- DRS が手動で、移行を承認するユーザーがいない。

以下の問題のいずれか、または複数がクラスタ上に存在すると、DRS で vMotion がほとんど実行されません。

- 負荷が安定しない、または vMotion に時間がかかる、あるいはその両方。移動が適切ではありません。
- DRS が仮想マシンをほとんどまたはまったく移動しない。
- DRS 移行のしきい値の設定が高すぎる。

DRS は次の理由によって仮想マシンを移動します。

- ユーザーが要求したホストの退避が、メンテナンス モードまたはスタンバイ モードに切り替わった。
- 仮想マシンとホスト間の DRS ルールまたは仮想マシン間の DRS ルール。

- 予約違反。
- 負荷の不均衡。
- 電源管理。

解決方法

DRS が vMotion の移行の実行を回避する原因となっている問題を解決します。

ホストの問題

ホストの問題によって、DRS が期待どおりに動作しない場合があります。

クラスタの総負荷が低い場合に、ホストをパワーオンしてキャパシティを増やすことを DRS が推奨する

ホストをパワーオンすることで、クラスタにキャパシティを追加するか、負荷の大きなホストを支援します。

問題

クラスタの総負荷が低い場合に、DRS がホストをパワーオンしてキャパシティを増やすことを推奨します。

原因

このように推奨されるのは、次のような理由が考えられます。

- クラスタが DRS-HA クラスタである。フェイルオーバーのキャパシティを増やすために、パワーオン状態のホストがさらに必要である。
- 一部のホストの負荷が高すぎるため、現在パワーオン状態のホスト上の仮想マシンをスタンバイ モードのホストに移動して負荷を均一化する。
- `MinPoweredOn{Cpu|Memory}Capacity` 詳細オプションの要件を満たすには、それに応じたキャパシティが必要である。

解決方法

ホストを起動します。

クラスタの総負荷が高い

クラスタの総負荷が高くなっています。

問題

クラスタの総負荷が高いのに、DRS がホストをパワーオンしないことがあります。

原因

DRS がホストをパワーオンしない理由として考えられるものは次の通りです。

- 仮想マシン間の DRS ルールまたは仮想マシンとホスト間の DRS ルールによって、仮想マシンをそのホストに移動できない。

- 仮想マシンが現在のホストに固定されているため、DRS でこの仮想マシンをスタンバイ モードのホストに移動し、負荷を均一化できない。
- DRS または DPM は手動モードに設定されており、さらに推奨が適用されていない。
- そのホストに、使用率の高いホスト上から移動される仮想マシンがない。
- ユーザー設定のため、または以前にホストがスタンバイを終了できなかったため、DPM がそのホスト上で無効にされている。

解決方法

DRS がホストをパワーオンするのを妨げている問題を解決します。

クラスタの総負荷が低い

クラスタの総負荷が低くなっています。

問題

クラスタの総負荷が低いのに、DRS がホストをパワーオフしないことがあります。

原因

DRS がホストをパワーオフしない理由として考えられるものは次の通りです。

- DPM (Distributed Power Management) がより適切なパワーオフ候補を検出した。
- vSphereHA でフェイルオーバー用のキャパシティがさらに必要である。
- 負荷がホストをパワーオフさせるほど十分に低くない。
- DPM は負荷が増加すると推測している。
- ホストで DPM が有効になっていない。
- 設定されている DPM のしきい値が高すぎる。
- ホストで DPM が有効になっているものの、適切なパワーオン方法がホストにない。
- DRS でホストを退避できない。
- DRS で設定されている移行のしきい値が高すぎるため、必須の移行しか実行できない。

解決方法

DRS がホストをパワーオフするのを妨げている問題を解決します。

要求されたホストを DRS が退避してメンテナンス モードまたはスタンバイ モードに切り替えることができない

要求されたホストを DRS が退避してメンテナンス モードまたはスタンバイ モードに切り替えることができません。

問題

ホストをメンテナンス モードまたはスタンバイ モードに切り替えようとする際、DRS が期待どおりにホストを退避しません。

原因

vSphere HA が有効な場合にホストを退避すると、HA フェイルオーバー キャパシティに違反することがあります。

解決方法

解決方法はありません。該当する場合は、ホストをメンテナンス モードまたはスタンバイ モードにする前に vSphere HA を無効にしてください。

DRS によって仮想マシンがホストに移動されない

DRS によってホストに仮想マシンが移動されません。

問題

DRS では、DRS が有効になっているクラスタに追加されたホストへの仮想マシンの移行は推奨されません。

原因

DRS が有効なクラスタにホストが追加されると、そのホストにデプロイされた仮想マシンはクラスタの一部になります。DRS は、クラスタに追加されたばかりのこのホストに、一部の仮想マシンを移行するように推奨できます。これが行われない場合、vMotion、ホストの互換性、またはアフィニティ ルールに問題があることがあります。考えられる原因は次のとおりです。

- vMotion がこのホストで構成されていない、または有効になっていない。
- ほかのホストの仮想マシンに、このホストとの互換性がない。
- 仮想マシン用の十分なリソースがこのホストにない。
- このホストへの仮想マシンの移動が、仮想マシン間の DRS ルールまたは仮想マシンとホスト間の DRS ルールに違反する。
- このホストが HA フェイルオーバー キャパシティ用として予約されている。
- デバイスが仮想マシンにマウントされている。
- vMotion のしきい値が高すぎる。
- DRS が仮想マシンで無効にされているため、仮想マシンをターゲット ホスト上に移動できない。

解決方法

DRS が仮想マシンをホストへ移動するのを妨げている問題を解決します。

DRS によってホストから仮想マシンが移動されない

DRS によってホストから仮想マシンが移動されません。

問題

仮想マシンがこのホストから移動されません。

原因

vMotion、DRS、またはホストの互換性の問題が原因となっている可能性があります。考えられる原因は次のとおりです。

- vMotion がこのホストで構成されていない、または有効になっていない。
- DRS がこのホストの仮想マシンに対して無効になっている。
- このホストの仮想マシンに、ほかのホストとの互換性がない。
- このホストの仮想マシン用の十分なリソースがほかのホストにない。
- このホストからの仮想マシンの移動が、仮想マシン間の DRS ルールまたは仮想マシンとホスト間の DRS ルールに違反する。
- このホスト上の 1 台以上の仮想マシンで DRS が無効になっている。
- デバイスが仮想マシンにマウントされている。

解決方法

DRS がホストから仮想マシンを移動するのを妨げている問題を解決します。

仮想マシンの問題

仮想マシンの問題によって、DRS が期待どおりに動作しない場合があります。

CPU またはメモリ リソース不足

仮想マシンが十分な CPU またはメモリ リソースを取得できません。

問題

仮想マシンの需要が、その仮想マシンのリソース確保量よりも大きい場合があります。この状態が発生すると、仮想マシンが十分な CPU またはメモリ リソースを取得できません。

原因

次のセクションでは、仮想マシンの確保量に影響を与える要因を説明します。

クラスタが黄色または赤く表示される

クラスタが黄色または赤く表示されている場合は、クラスタのすべての仮想マシンおよびリソース プールに対して構成されたリソース予約を確保するための十分なキャパシティがありません。特定の仮想マシンがその予約分を受け取っていない可能性があります。クラスタのステータス（赤または黄色）を確認し、問題を解決してください。

リソース制限が厳しすぎる

仮想マシン、仮想マシンの親リソース プール、またはそのリソース プールの先祖に構成されているリソース制限に、過度に厳しいものがある可能性があります。需要が、構成されているすべての制限値以上であるかを確認してください。

クラスタに負荷がかかりすぎる

仮想マシンが動作しているクラスタに十分なリソースがない可能性があります。また、仮想マシンのシェア値が大きいため、それに比例してほかの仮想マシンに多くのリソースが供与されています。需要がキャパシティよりも大きいかどうかを特定するには、クラスタの統計情報を確認してください。

ホストに負荷がかかりすぎる

ホストのリソースが超過しているかどうかを特定するには、ホストの統計情報を確認してください。超過している場合は、DRS が、そのホストで現在動作している仮想マシンをほかのホストに移動しない原因を検討してください。この状況は次の理由によって発生することがあります。

- 仮想マシン間の DRS ルールおよび仮想マシンとホスト間の DRS ルールに、現在の仮想マシンとホストとのマッピングが必要である。そのようなルールがクラスタで構成されている場合は、それらのルールの 1 つ以上を無効にすることを検討してください。次に DRS を実行し、状態が修正されているかどうかを確認します。
- この仮想マシンまたはその他の仮想マシンを必要なだけ DRS で移動させて、キャパシティを確保することができません。次のいずれかが原因で、DRS によって仮想マシンの移動が実行されません。
 - DRS がその仮想マシンで無効になっている。
 - ホスト デバイスが仮想マシンにマウントされている。
 - いずれかのリソース予約が過度に大きいため、クラスタのほかのホストで仮想マシンを実行できない。
 - 仮想マシンに、クラスタ内のほかのホストとの互換性がない。

これらの条件が仮想マシンに存在しているかどうか確認します。発生していない場合は、そのクラスタ内のほかの仮想マシンで発生している可能性があります。このような場合、DRS がクラスタを調整して仮想マシンの需要を解決することはできません。

- DRS の移行のしきい値を小さく設定して、問題が解決されたかどうかを確認してください。
- 仮想マシンの予約を増やします。

解決方法

仮想マシンが十分な CPU またはメモリ リソースを取得できない原因となっている問題を解決します。

仮想マシン間の DRS ルール違反または仮想マシンとホスト間の DRS ルール違反が発生する

DRS ルールは、仮想マシンが常駐すべき、あるいは常駐してはならないホスト、または同じホスト上に存在すべき、あるいは存在してはならない仮想マシンを指定します。

問題

仮想マシン間の DRS ルールまたは仮想マシンとホスト間の DRS ルールの違反があります。

原因

仮想マシン間の DRS ルールでは、選択した仮想マシンを同一のホスト上（アフィニティ）に配置するか、または仮想マシンを別々のホスト上（非アフィニティ）に配置することを指定します。仮想マシンとホスト間の DRS ルールでは、選択した仮想マシンを特定のホスト上（アフィニティ）に配置するか、特定のホスト上（非アフィニティ）に配置しないかを指定します。

仮想マシン間の DRS ルールまたは仮想マシンとホスト間の DRS ルールに違反する場合、DRS はルールに基づいて仮想マシンの一部またはすべてを移動できないことがあります。その仮想マシン、アフィニティ ルール内のほかの仮想マシン、またはそれらの仮想マシンの親リソース プールの予約が原因で、DRS がすべての仮想マシンを同じホストに配置できない可能性があります。

解決方法

- アフィニティ ルールに関連した障害については、DRS 障害パネルを確認してください。
- アフィニティ ルールにおけるすべての仮想マシンの予約量合計を計算します。この値がどのホストの使用可能容量よりも大きい場合は、ルールの条件を満たすことはできません。
- 親リソース プールの予約量合計を計算します。この値がどのホストの使用可能容量よりも大きい場合、リソースが 1 台のホストから取得されているのであれば、ルールの条件を満たすことはできません。

仮想マシンのパワーオン操作が失敗する

仮想マシンがパワーオンに失敗したというエラー メッセージが表示されます。

問題

仮想マシンがパワーオンに失敗します。

原因

リソースが不十分である場合、あるいは仮想マシンと互換性のあるホストがない場合に、仮想マシンがパワーオンに失敗する可能性があります。

解決方法

単一の仮想マシンのパワーオン、またはグループをパワーオンすることによる任意の仮想マシンのパワーオンを実行できる十分なリソースがクラスタにない場合は、クラスタまたはその親リソース プールで利用可能なリソースに照らし、仮想マシンに必要なリソースを確認してください。必要に応じて、パワーオンする仮想マシンの予約を減らすか、兄弟仮想マシンの予約を減らすか、クラスタまたはその親リソース プールで利用可能なリソースを増やします。

DRS によって仮想マシンが移動されない

ホスト上のリソースが不十分であるにも関わらず、最初のパワーオン時に DRS によって仮想マシンが移動されません。

問題

仮想マシンのパワーオン時に、その仮想マシンが登録されているホスト上に十分なリソースがない場合、DRS が期待どおりに仮想マシンを移行しません。

原因

DRS が仮想マシンを移動しない理由として考えられるものは次の通りです。

- DRS が仮想マシンで無効になっている。
- 仮想マシンにマウントされたデバイスがある。
- 仮想マシンに、ほかのすべてのホストとの互換性がない。
- その仮想マシンの各 CPU に対する十分な数の物理 CPU またはキャパシティがほかのホストにない。
- その仮想マシンの予約および必要なメモリを確保できる十分な CPU またはメモリ リソースがほかのホストにない。
- その仮想マシンを移動すると、アフィニティまたは非アフィニティ ルールに違反する。
- その仮想マシンの DRS の自動化レベルが手動に設定されており、さらに移行の推奨の承認がユーザーによって行われていない。
- DRS はフォールト トレランスが有効な仮想マシンを移動しない。

解決方法

DRS が仮想マシンを移動するのを妨げている問題を解決します。