

# HP Moonshot の本番環境への導入



[www.hp.com](http://www.hp.com) 環境の主要ドメインを実行する新しいクラスのサーバー

## 目次

新しいクラスのサーバー .....	2
HPでのMoonshotの配備 .....	2
WWW/FTPアプリケーションアーキテクチャー .....	2
Tier 1 - キャッシュされるコンテンツ .....	3
Tier 2 - キャッシュされないコンテンツ .....	3
Tier 3 - データ.....	3
Moonshotシャーシ上のWWW/FTP配備トポロジ.....	3
実装 .....	4
ステップ 1 - 1 拠点のデータセンターにキャッシュレイヤーを配備 .....	4
ステップ 2 - すべてのデータセンターにキャッシュレイヤーを配備 .....	5
ステップ 3 - 1 拠点のデータセンターにApacheレイヤーを配備 .....	5
ステップ 4 - すべてのデータセンターにApacheレイヤーを配備.....	6
ステップ 5 - 1 拠点のデータセンターでhttp://ftp.hp.comおよびftp://ftp.hp.comをホスト .....	6
ステップ 6 - すべてのデータセンターでhttp://ftp.hp.comおよびftp://ftp.hp.comをホスト.....	6
メリットの実現 .....	7
今後の展望.....	7

新しい HP Moonshot System は、www.hp.com ホームページで発生する膨大なトラフィックを処理し、ソフトウェアドライバー/アップデートを配布する HP のダウンロードサイトへのアクセスを実現します。本書では、1 日あたりのヒット数が 3 億を超えるこの要求の厳しい環境に Moonshot System を確実に実装するために HP が行った段階的な方法について説明します。

## 新しいクラスのサーバー

HP Moonshot Systemは、ソーシャルメディア、クラウド、モバイル、およびビッグデータを背景とするITの課題に対処するために設計された、新しいクラスのサーバーです。HP Moonshot Webサーバーには、スマートフォンやタブレットで一般的なチップが組み込まれています。これにより、Moonshot Systemは、従来のサーバーと比べてエネルギー消費量を最大 89%、設置スペースを 80%、そして購入コストを 77%削減する、経済的に魅力的な新しいインフラストラクチャを実現しています<sup>1</sup>。また、管理、電力、冷却、およびネットワーク機能の共有による効率の向上もあります。このように大幅な削減が可能であるため、従来のITインフラストラクチャにおけるスペースおよびコスト面の制約を受けている今日の大規模データセンターにとって、Moonshot Systemは最適です。

HP Moonshot System は、Moonshot 1500 シャーシエンクロージャーと、カートリッジとも呼ばれる、アプリケーションに最適化された HP ProLiant Moonshot Server で構成されており、単一の 4.3U サーバーエンクロージャーで、最大 45 台のサーバーカートリッジ、2 台のネットワークスイッチ、および複数の補助コンポーネントを収容します。ラックあたり最大 1,800 台のサーバーをサポートできる HP Moonshot サーバーの専用スペースは、従来のサーバーで必要なスペースの 8 分の 1 に抑えられます。

## HP での Moonshot の配備

密度と経済性に優れた HP Moonshot System は、(100 近い Web サイトと何千もの DNS ドメイン名が含まれる) www.hp.com 環境で稼働中のアプリケーションをホストする、6 拠点到分散された HP のデータセンターのニーズを満たすのに最適です。www.hp.com 環境での Moonshot サーバーの初期配備では、www.hp.com ホームページへのトラフィックと、FTP プロトコルおよび HTTP プロトコルによってドライバー/ソフトウェアアップデートを配布する HP のダウンロードサイトへのアクセスをサポートしています。

Web ドメイン www.hp.com および ftp.hp.com は、外部向けの HP の Web サーバーによって処理される閲覧とダウンロードによるトラフィックの大部分に相当し、通常は、これらのドメインから 1 日あたり 400~500TB のデータが送出されます。

処理するトラフィックの約 3 分の 1 は、HP の社外向け Web サイトの入口となる www.hp.com に対するもので、残り 3 分の 2 は、HTTP または FTP のいずれかによって ftp.hp.com からファイルをダウンロードする際に発生しています。

この実装では、最初のユースケースとして、キャッシュ、HTTP/HTTPS、および FTP ホスティングを HP Moonshot シャーシに配備しました。この最初の実装において、Moonshot カートリッジにより、大規模 Web アプリケーションで一般的な高度な要求/応答パターン、メモリ内キャッシュ配信、および Moonshot カートリッジ単体で使用可能な 1TB のストレージを超える大容量のファイルシステムを処理できることが実証されました。

これらの Web サイトは、シャーシに収容されているすべてのカートリッジにわたって配備されました。カートリッジにわたる各階層は、キャッシュできないオブジェクトを処理する Web 階層など、コンピューティング負荷の高い階層の割合を高くするように適切な重み付けを設定して分散させたため、ボトルネックは生じませんでした。

## WWW/FTP アプリケーションアーキテクチャー

WWW/FTP アプリケーションは、HTTP と HTTPS の要求用に Apache を実行し、FTP 用に Very Secure FTP Daemon (VSTPD) を実行する、単一階層のスケールアウト型 Web ホスティングスタックです。従来、WWW/FTP は、Intel® Itanium® プロセッサを搭載した 46 台の HP 製サーバーを使用して、(HP の 6 拠点のデータセンターにわたる) 分散型ファームでホストされていました。

<sup>1</sup> HP Moonshot サーバーと従来の x86 サーバーテクノロジーを比較した HP の社内調査結果による。

ソフトウェアスタックは、HP-UX 11.22 および 11.31、標準的な Apache および VSFTPD、そして 7TB のファイルシステムによってファイルを提供する直接接続ストレージからなる、従来の構成です。Moonshot への実装では、すべてのサービスを HP-UX から Red Hat® Enterprise Linux 6 (RHEL 6) オペレーティング環境に移行しました。

この新しいアプリケーションアーキテクチャーには、Web 階層、ミドルウェア階層、およびストレージ階層の 3 つの階層があります。

## Tier 1 - キャッシュされるコンテンツ

Apache レイヤーの負荷を削減するために、新たにキャッシュ階層を導入しました。サイトに対するこのエントリーポイントは、基盤となるコンポーネントにトラフィックを誘導し、複数のサーバーに要求を分散するオープンソースの TCP/HTTP ロードバランサー (リバースプロキシ) である HAProxy です。

キャッシュ機能とロードバランサー機能を実現するために、Web アプリケーションアクセラレータの Varnish Cache を同じ階層内に実装しました。イメージ、JavaScript、スタイルシートなど、すべての静的コンテンツは、最初のアクセスを経て Varnish によりキャッシュされます。Varnish Cache は、サイトで最も人気のある静的コンテンツをメモリに保持できる十分な大きさで構成されています。そのため、この階層は通常、ネットワークおよびメモリ I/O の制限を受けません。

FTP ホスティングについては、ftp://の要求を処理するために Varnish と合わせて VSFTPD を実装しました。

## Tier 2 - キャッシュされないコンテンツ

明示的に除外されたオブジェクトやキャッシュミスなど、キャッシュされない HTTP コンテンツは、Apache に渡されて処理されます。ファイルはディスクから取得され、ネットワーク経由で Varnish 階層に返されるため、この第 2 階層は通常、ディスク、CPU、およびネットワークの制限を受けません。

## Tier 3 - データ

HTTP と FTP のデータ階層はいずれも、第 3 階層として機能する、専用の複数台 1 組のカートリッジに実装します。アプリケーションによって処理されるファイルの合計サイズは約 7TB です。ホストに必要な単一ファイルシステムを使用できるように、オープンソースの分散ファイルシステムである Gluster を使用して、複数台のカートリッジにコンテンツをストライピングしました。Gluster ファイルシステムは、各ホスティングサーバーにマウントされています。この方式を使用すると、複数の Moonshot System カートリッジに含まれている 1TB ストレージが単一の共有ファイルシステムとして機能します。

## Moonshot シャーシ上の WWW/FTP 配備トポロジ

WWW/FTP アプリケーションは、全体と個々の階層内の両方で、簡単にスケールアウトできます。そして潜在的なホットスポットとボトルネックには、適切なコンポーネント階層にインスタンスをさらに追加することによって対処できます。従来の本番環境から得た経験則と、各カートリッジのリソースの制約 (2 x 1GbE など) を鑑みて、コンポーネント階層を次のように分散しました。

### www.hp.com 用の 4 台の Varnish キャッシュロードバランサー

- これにより、アプリケーションに対する最低 4Gb の入出力が実現されます。

### http://ftp.hp.com 用の 8 台の HAProxy ロードバランサーと ftp://ftp.hp.com 用の VSFTPD

- これらのコンポーネントにより、ftp.hp.com に向かいバックエンドのレガシー FTP サーバーに誘導される HTTP トラフィックの負荷分散が実現されます。
- ftp.hp.com の FTP トラフィックは、Gluster によって提供される/pub ディレクトリにアクセスする VSFTPD を介して直接処理されます。

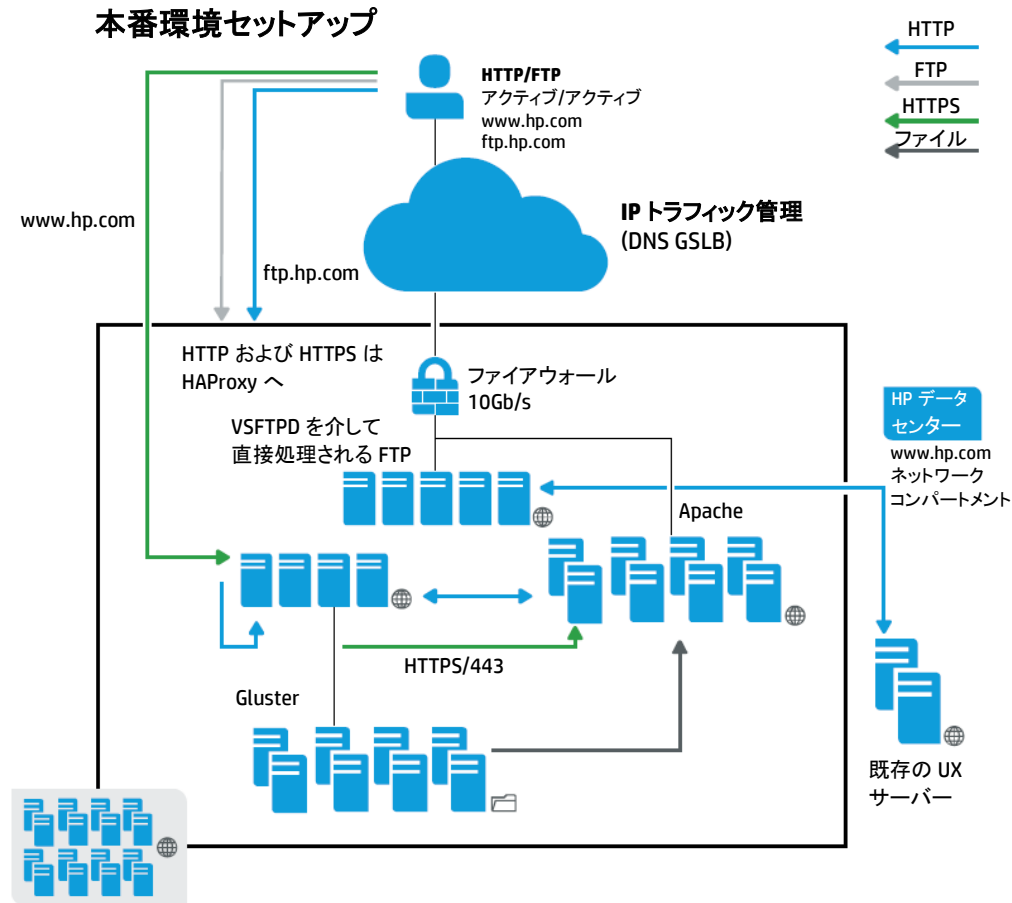
### 16 個の Apache インスタンス

- オブジェクト要求は Varnish を通過して Apache で処理されます。これまでと同様に、これは負荷がかかった状態におけるホットスポットであり、ボトルネックです。輻輳が発生しないように、Varnish に関して 2 倍の Apache インスタンスを配備しました。

### 16 個の Gluster インスタンス

- 7TB のファイルシステムに格納されるデータは 8 台のサーバーにわたってストライピングされ、冗長性を確保するためにもう 1 組の 8 台のサーバーに複製されます。

図 1. 本番環境セットアップ



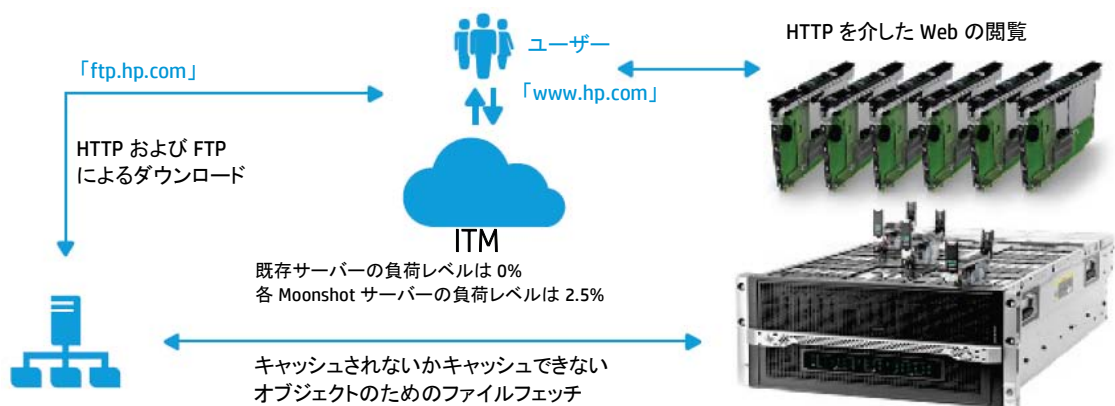
## 実装

WWW/FTP アプリケーションは、HP の 6 拠点のデータセンター（ジョージア州アトランタの 2 拠点、テキサス州ヒューストンおよびオースチンの各 2 拠点）でホストされています。Moonshot への移行は、サービスを中断しないために複数のステップを経て実施されました。

### ステップ1-1拠点のデータセンターにキャッシュレイヤーを配備

既存の WWW/FTP HP-UX サーバーが配備されているネットワークに最初の Moonshot シャーシを実装しました。PXE ブート機能を使用して RHEL 6 をインストールしてから、キャッシュされないオブジェクト用のホストとして既存の HP-UX Apache サーバーを使用して、HAProxy と Varnish を構成しました。

図 2. Moonshot の実装



Moonshot カートリッジと既存の Itanium®サーバーの機能を比較するために、同じ台数のカートリッジ (6 台) を実装しました。www.hp.com 用に DNS ベースの負荷分散レコードをセットアップしてから、6 台の Moonshot サーバーの IP アドレスを入力し、ロードバランサー構成に含まれている 6 台の Itanium の IP アドレスを削除して、このレコードを調整しました。その後、レガシーサーバーに送られた負荷レベルと同等になるまで、6 台のカートリッジに対する負荷レベルをゆっくりと上げました。

### 結果

6 台の Moonshot カートリッジで、www.hp.com に対するエンドユーザー要求のほぼ 12~14% に相当するデータセンターの全負荷を処理できました。1 日あたりの平均が 1 億ヒット程度であるため、各カートリッジで約 200 万ヒットが処理されたこととなります。キャッシュヒット率は、キャッシュ可能オブジェクトの 95% 程度となり、www.hp.com のヒット全体で見ると、およそ 60% でした。このように値に差が生じた原因は、複数のオブジェクトがキャッシュ処理から除外されたことにあります。

各カートリッジの平均 CPU 使用率は 2.5% でした。

## ステップ2 - すべてのデータセンターにキャッシュレイヤーを配備

追加の Moonshot シャーシ 5 台が正常にロールアウトされてから、6 拠点のデータセンターすべてにわたってキャッシュセットアップを複製しました。既存の WWW/FTP HP-UX サーバーが配備されているネットワークにシャーシを実装しました。PXE ブート機能を使用して RHEL 6 をインストールしてから、キャッシュされないオブジェクト用のホストとして既存の HP-UX Apache サーバーを使用して、HAProxy と Varnish を構成しました。

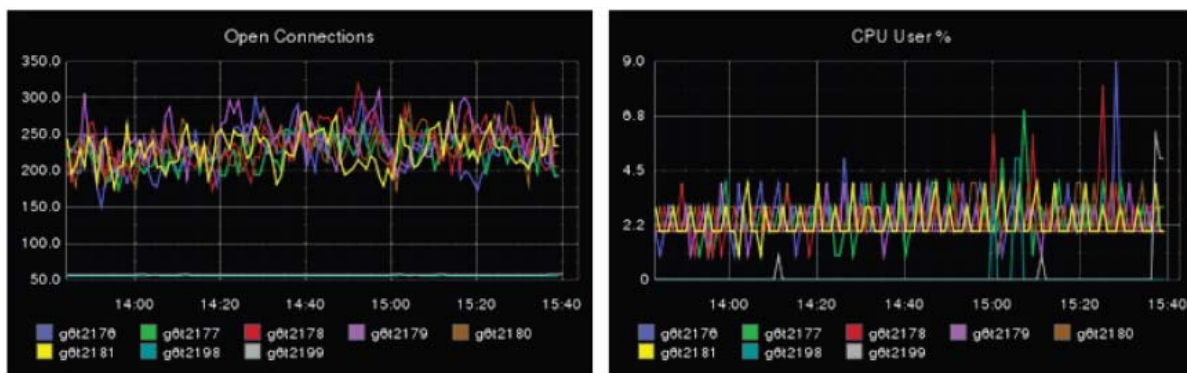
最後に、Moonshot カートリッジの IP アドレスのみを含むように www.hp.com に対する負荷分散を変更しました。

### 結果

これで、36 台の Moonshot カートリッジ (データセンターあたり 6 台) により、www.hp.com に対するエンドユーザー要求の負荷全体が処理されるようになりました。1 日平均の約 1 億ヒットについて、各データセンターでおよそ 1,600 万ヒットを処理します。キャッシュヒット率は、キャッシュ可能オブジェクトの 95% 程度であり、www.hp.com のヒット全体で見ると、およそ 60% です。ここでも、このように値に差が生じた原因は、複数のオブジェクトがキャッシュ処理から除外されたことにあります。

各カートリッジの平均 CPU 使用率はまだ 2.5% です。

図 3. キャッシュレイヤーにおける Moonshot カートリッジの CPU 使用率



## ステップ3 - 1拠点のデータセンターにApacheレイヤーを配備

www.hp.com 環境からレガシーサーバーを撤去するために、Apache を実装した 16 台 1 組のカートリッジと、Gluster を使用して 7TB のファイルシステムをホストする目的専用で使用する別の 16 台 1 組のカートリッジを構成しました。コンテンツは、スケジュール設定された複製ジョブを使用して既存の HP-UX サーバーから複製されました。この作業には、ファイルのコピーと同期を行うオープンソースのユーティリティである rsync を使用しました。すべてのドキュメントとオブジェクトにアクセスできるように、すべての Apache に Gluster ファイルシステムがマウントされています。

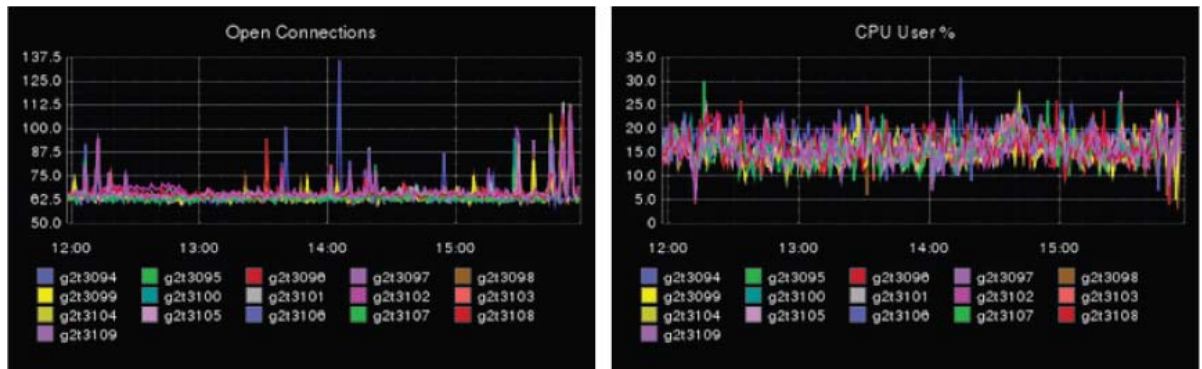
最後のステップでは、Itanium Apache サーバーを撤去して、16 台の Moonshot Apache サーバーで置き換えるように、Tier 1 サーバーの Varnish 構成を変更しました。これで、Varnish は、キャッシュできないオブジェクトに対する要求を、Moonshot シャーシに收容された Linux Apache サーバーに送るようになりました。

## 結果

16 台の Moonshot カートリッジでデータセンターに対する要求全体のおよそ 30~40%が処理されます。平均キャッシュヒット率は 60~70%です。

各カートリッジの平均 CPU 使用率はほぼ 20%です。

図 4. Apache レイヤーにおける Moonshot カートリッジの CPU 使用率



## ステップ4 - すべてのデータセンターにApacheレイヤーを配備

残りのデータセンター構成からレガシーサーバーを撤去するために、ステップ 3 と同様のセットアップを残り 5 拠点のデータセンターで実施しました。

## ステップ5 - 1拠点のデータセンターでhttp://ftp.hp.comおよびftp://ftp.hp.comをホスト

ftp.hp.com のダウンロードトラフィックが、この WWW/FTP アプリケーションにおける最大のトラフィック要素です。ftp.hp.com に対するおよそ 2 億ヒットのうち、99%は HTTP からの要求で、残りの 1%が FTP からのヒットです。ftp.hp.com の DNS ロードバランサーを変更するには、Moonshot ファームで HTTP 要求だけでなく FTP 要求も処理できる必要があります。

数時間のテストを経て、FTP キャッシュ用に割り当てた台数の Moonshot カートリッジでは、前述した量のダウンロードファイルを処理するには不十分であったと判断しました。本書の執筆時点では、ユーザー接続を処理するために HAProxy とともに FTP フロントエンドサーバーが構成されています。このとき、4 台のレガシー Apache サーバーで構成されるファームにわたって、ftp.hp.com に対する HTTP プロトコル要求の負荷分散が行われます。

ftp.hp.com に対する FTP プロトコル要求は、ステップ 3 で説明されている Apache サーバー同様、フロントエンドサーバーにインストールされ、Gluster ファイルシステムからファイルを読み取る VSFTPD によって直接処理されます。

ftp.hp.com に対する要求は、ステップ 3 で説明されている Apache サーバー同様、Tier 1 サーバーで稼働し、16 台の Gluster カートリッジによって実現されている Gluster ファイルシステムからファイルを読み取る、VSFTPD サーバーによって処理されます。

ftp.hp.com 用に DNS ベースの負荷分散レコードをセットアップしてから、6 台の Moonshot サーバーの IP アドレスを入力して Itanium の IP アドレスを削除し、このレコードを調整しました。その後、レガシーサーバーに送られた負荷レベルと同等になるまで、6 台のカートリッジに対する負荷レベルをゆっくりと上げました。

## 結果

これで、6 台の Moonshot カートリッジによって 1 時間あたり約 170 万件の要求、つまり、カートリッジあたり毎秒約 60 件の要求が処理されるようになりました。

## ステップ6 - すべてのデータセンターでhttp://ftp.hp.comおよびftp://ftp.hp.comをホスト

ステップ 5 で説明したセットアップを正常に複製してから、Tier 1 FTP サーバーの Moonshot IP アドレスを 36 個だけを含むように、ftp.hp.com の DNS ロードバランサー構成をもう 1 度変更しました。FTP ダウンロード要求用にファイルサーバーとしてデータセンターあたり 2 台の Itanium システムを維持する一方で、画像をホストしている Itanium システムを www.hp.com から撤去できました。

## メリットの実現

WWW/FTP アプリケーションは、HP の各データセンターにある 1 台 (合計 6 台) の Moonshot シャーシでホストされるようになりました。サーバーの台数は 46 台から 270 台に増加しましたが、全体的なスペースと電力消費量は大幅に削減されました。各 HP-UX サーバーは 2.5kW~3.3kW を消費し、データセンターのラックのほぼ半分を占有しますが、Moonshot 1500 シャーシの電力消費量は 1kW 程度で、必要なスペースもラックのわずか 10 分の 1 で済みます。

以下に、6 拠点のデータセンターにわたってロールアウトされた Moonshot ファームによる削減効果の平均想定値を示します。

表 1.6 拠点のデータセンターでの Moonshot による削減効果の平均想定値

	HP-UX	Moonshot	増減
サーバー	46	270 台+12 台のレガシーサーバー	500%増加
電源	約 138kW	約 6kW+レガシーサーバーの 30kW	74%減少
スペース	23 台のラック	1 台のラックの 10 分の 6+レガシーサーバーのラック 6 台	70%減少

## 今後の展望

WWW/FTP アプリケーションプロジェクトへの実装が成功すれば、今後数か月から数年のうちに、より大規模な www.hp.com 環境で HP Moonshot System をさらに活用できるようになると期待しています。次期 Moonshot リリースでは、ftp.hp.com に対するファイルサーバー機能を置き換えることに重点を絞ります。その後、Moonshot プラットフォームに移行できるアプリケーションの識別、およびこの方法の拡張と実運用化に継続的に取り組み、広範な www.hp.com 環境に Moonshot サーバーを実装して行きます。

### リソース

[ケーススタディを参照](#)

[ケーススタディの動画を視聴](#)

### 詳細情報

[hp.com/go/moonshot](http://hp.com/go/moonshot) (英語) [hp.com/jp/moonshot](http://hp.com/jp/moonshot) (日本語)

## メールニュース配信登録

<http://hp.com/go/getconnectedjp>

© Copyright 2013 Hewlett-Packard Development Company, L.P. 本書の内容は、将来予告なく変更されることがあります。HP 製品およびサービスに対する保証については、当該製品およびサービスの保証規定書に記載されています。本書のいかなる内容も、新たな保証を追加するものではありません。本書の内容につきましては万全を期しておりますが、本書中の技術的あるいは校正上の誤り、省略に対しては責任を負いかねますのでご了承ください。

インテル Itanium および Itanium は、米国および米国以外の国における Intel Corporation の商標です。Red Hat は、米国およびその他の国における Red Hat, Inc. の商標です。Java は、Oracle Corporation およびその関連会社の登録商標です。

