

Login VSI による Parallels RAS スケーラビリティ テスト

1,000 ユーザー - Knowledge Worker

ホワイトペーパー | Parallels Remote Application Server | 2020

Parallels International GmbH
Vordergasse 59
8200 Schaffhausen
Switzerland
Tel: + 41 52 672 20 30
www.parallels.com

Copyright © 1999-2022 Parallels International GmbH. All rights reserved.

This product is protected by United States and international copyright laws. The product's underlying technology, patents, and trademarks are listed at http://www.parallels.com/about/legal/.

Microsoft, Windows, Windows Server are registered trademarks of Microsoft Corporation.
Apple, Mac, the Mac logo, macOS, iPad, iPadOS, iPhone, iPod touch are trademarks of Apple Inc., registered in the US and other countries.
Linux is a registered trademark of Linus Torvalds.
All other marks and names mentioned herein may be trademarks of their respective owners.

目 次

はじめに 4

スケーラビリティ 5

Parallels RAS のスケーラビリティ テスト 5

スケーラビリティ テストのための構成 5

テストプロセス 7

所 見 8

RD セッションホスト 9

RAS インフラストラクチャー コンポーネント 13

RAS Publishing Agent 13

RAS Secure Client Gateway 14

HALB 16

結 論 18

はじめに

Parallels Remote Application Server (RAS) は、包括的な仮想アプリケーションおよびデスクトップ配信ソリューションであり、従業員があらゆるデバイスからアプリケーションとデータを使用できるようにします。
シームレスで導入、設定、保守が簡単な Parallels RAS は、Microsoft RDS、Windows Virtual Desktop、主要なハイパーバイザーを介したアプリケーションとデスクトップの配信をサポートします。

本書は、Login VSI を使用して Parallels RAS 17.1 における、Knowledge Worker ワークロードを持つ RD セッションホスト上の約 1,000 ユーザー セッションで実施したスケーラビリティ テストの検証結果をまとめたものです。

スケーラビリティ

この章の内容

Parallels RAS のスケーラビリティ テスト 5
スケーラビリティ テストのための構成 5
テストプロセス 7
所 見.....8
RD セッション ホスト 9

Parallels RAS のスケーラビリティ テスト

Parallels RAS 構成を検証するために、Parallels のエンジニアは一連のパフォーマンス テストを実施しました。その目的は、VMware vSphere 仮想マシン上で実行される Parallels RAS セッションのスケーラビリティを分析することでした。このテストの一環として、Login VSI を使用して、一般的なユーザーのワークロードをシミュレートする RD セッションホスト サーバーへのユーザー接続を生成しました。

通常の Parallels RAS デプロイメントでは、ユーザーは Parallels クライアント アプリケーションを介して接続し、リモート アプリケーションやデスクトップにアクセスします。Login VSI クライアントがユーザー接続をシミュレートし、RAS Publishing Agent がそれを処理し、エンドユーザーと RD セッションホスト サーバー間のサービス接続をセットアップします。

スケーラビリティ テストのための構成

Parallels RAS のスケーラビリティ テストでは、以下のハードウェア コンポーネントで構成される Supermicro X9DRi-LN4+ を合計 10 台使用しました。

CPU	2x Intel Xeon E5-2695v2、2.6GHz、20 MB L3、115W TDP
RAM	128 GB, 1,600 MHz Micron DDR-4-2100 8GB x16
ストレージ	Host OS: 1 x 1TB SATA 7.2k RPM ワークロード VM とインフラ VM: 2 x 1TB SSD RAID なし

コンポーネント

コンポーネント	ソフトウェア バージョン / ビルド
Hypervisor	VMware ESXi 6.7
Hypervisor Management	VMware vSphere 6.7
Network Load Balancing	Parallels HALB Appliance 17.1.0 (Build 21783)
Domain Services (DNS, AD Functional level)	Microsoft Active Directory Domain Service
NTP server	Microsoft NTP Server
File Server for UPD	Microsoft SMB Server

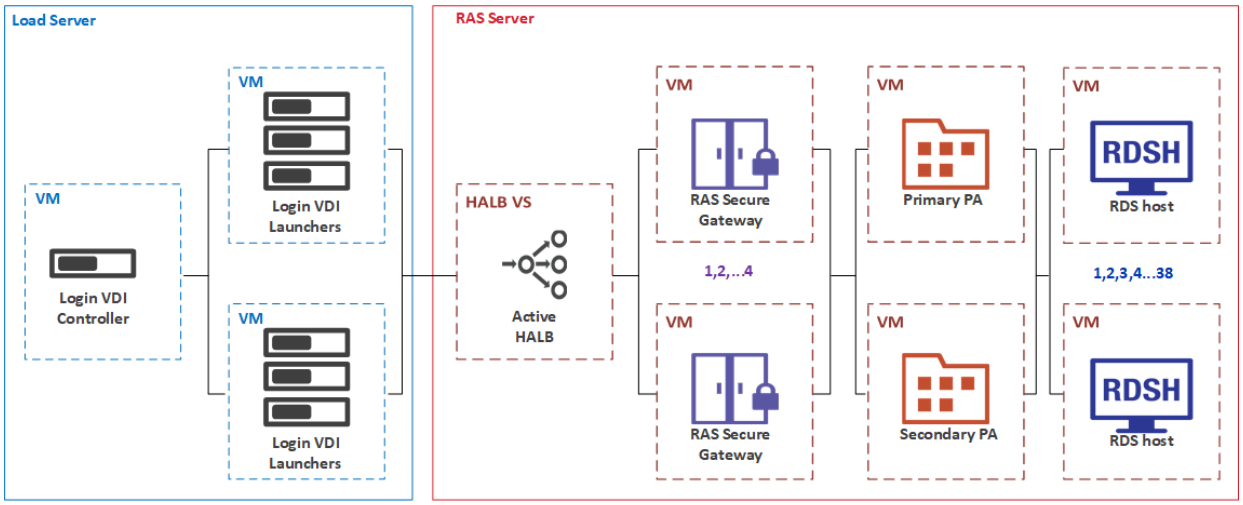
Parallels RAS VM の構成

Parallels RAS コンポーネント	VM の総数	各 VM の vCPU	各 VM の RAM
RAS Publishing Agent	2	2	8 GB
RAS Secure Client Gateway	3	2	8 GB
High Availability Load Balancing	1*	1	2 GB
RD Session Host	38	8	20 GB

* 高可用性を実現するために、製品環境では 2 台以上の HALB デバイスを推奨します。

テスト環境を構成するすべての仮想マシンは、同一の仮想ネットワーク上にサイロ化されています。

テスト環境図



テストプロセス

スケーラビリティ テストでは、Parallels RAS 17.1 で Parallels Client for Windows 17.1 (x64) を使用し、Login VSI 4.1.40 でユーザー負荷を実行しました。Login VSI は、デスクトップ環境がサポートできる最大ユーザー数を測定するのに役立ちます。Login VSI は、ワークロードを Task Worker、Knowledge Worker、Power Worker、および Office Worker に分類します。

このテストでは、Knowledge Worker ワークロードを選択しました。Knowledge Worker は、2 (v) CPU 環境向けに設計されています。このワークロードは、システムにスムーズにストレスを与え、CPU、RAM、I / O の使用量を増加させる、バランスの取れた集中的なワークロードです。

Knowledge Worker ワークロードは、以下のアプリケーションを使用します。

- Internet Explorer 11 web browser
- Microsoft Outlook 2013
- Microsoft Word 2013
- Microsoft Excel 201
- Microsoft PowerPoint 2013
- Adobe PDF Reader DC
- Doro PDF Writer
- 7-Zip
- Windows Photo Viewer
- Freemind / Java

Login VSI Knowledge Worker ワークロードの詳細については、次のリンクを参照してください: https://support.loginvsi.com/hc/en-us/articles/360001046100-Login-VSI-Workloads-Default- workloads-information

Login VSI Knowledge Worker ワークロードを使用して、Parallels RAS 上の 1,000 ユーザーのワークロードをシミュレートしました。ログオン フェーズの期間は約 1 時間以下に設定し、ログオンレートは 4 秒に 1 セッションに設定しました。

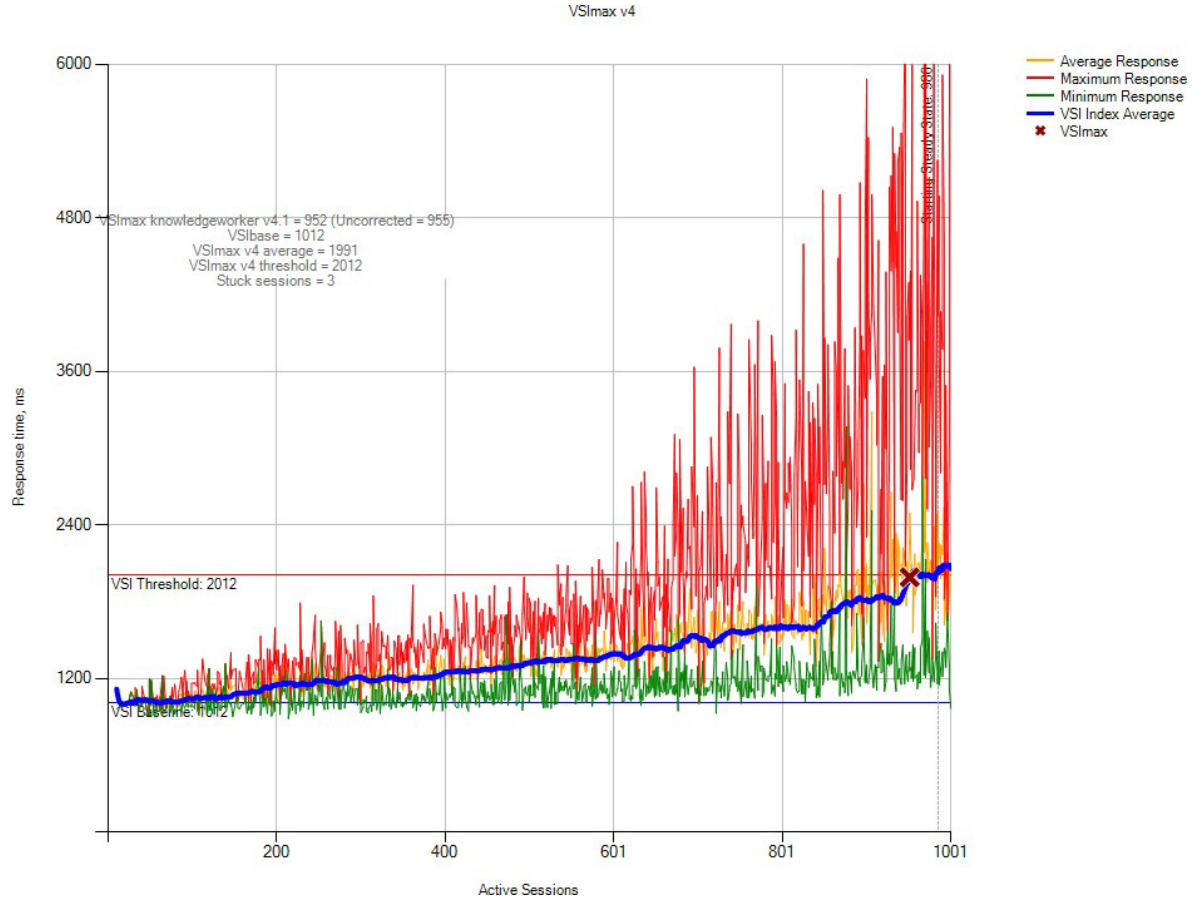
性能指標は、ユーザーログオン時、仮想デスクトップ取得時 (ランプアップ)、ユーザー ワークロード実行時 (定常状態)、ユーザー ログオフ時に取得しました。コンポーネントが適切にキャッシュされたタイミングを反映した一貫性のある測定値を得るため、各ワークロードは、Login VSI のパフォーマンス メトリクスを記録する前に 48 分間実行しました。VSI テストは、各 VM インスタンスで 3 回繰り返され、テストに成功したユーザーの平均人数を算出しました。

スケーラビリティ テストは、プラットフォームやソリューション全体のパフォーマンスを理解する上で重要な要素ですが、実際の本番でのワークロードの正確な測定値として推論されるべきではないことに留意することが重要です。アプリケーションがどのように動作するかをより適切に評価したいお客様は、カスタム ワークロード スクリプトを使用した独自の Login VSI スケール テストを実施する必要があります。また、Parallels RAS POC / Pilot をリクエストすることも可能です。このテストの目的は、可能な限りの密度を反映したベースラインを取得することであるため、Login VSI クライアント起動は、プロキシ SSL モードで Secure Gateway を経由するよう構成されました。

所 見

以下は、Knowledge Worker ワークロードのテスト結果です。VSImax v4 (特定のワークロードにおける最大ユーザー密度を示す) は、VSI Baseline と VSI Threshold の測定値から決定されます。VSI Baseline は、テスト前の Login VSI のベースライン応答時間測定値を表し、通常の Login VSI セッションがサンプリングされる前に決定されます。

VSImax v4 では、Task Worker のワークロードを実行する 1,000 ユーザーの密度が実証されました。私たちのテストでは、952 セッションで VSImax に到達しました。これは、現在のサーバーの仕様で UX の低下が観察される前に、すでに 952 セッションが同時処理していたことを意味します。



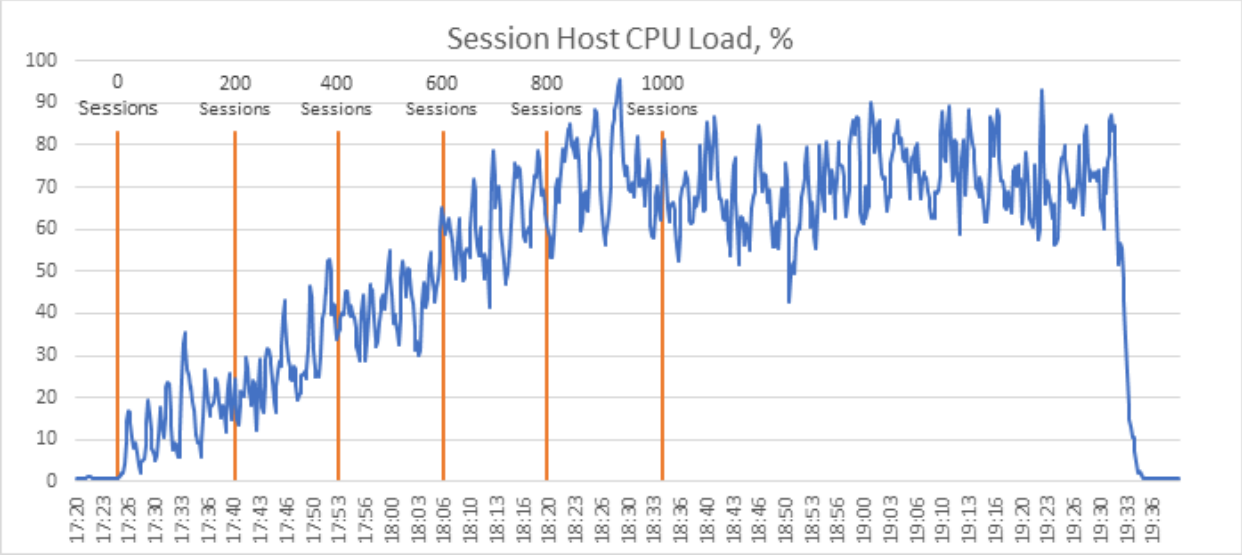
注：ログイン VSI テストでは、3 つのスタックセッションが検出されました。理想的にはスタックしたセッションがないことが望ましいのですが、この数は少なくそのようなセッションは VSImax スコアから差し引かれ、これは複数のテスト実行の平均数であることから、この結果を維持することに決定しました。

RD セッションホスト

CPU、メモリ消費量、ディスク I / O 応答時間、ネットワーク負荷に関する以下のテスト結果は、テストワークロード下での性能を評価するのに役立ちます。以下の各チャートは、1 台の平均的な RD セッションホスト サーバーから収集したデータです。38 台の RD セッションホスト サーバーと1,000 人の想定ユーザーが存在するため、1 台の RD セッションホスト サーバーは約 26 人のユーザーを収容しています。

次の 2 つのグラフでは、ユーザーの負荷が増加するにつれて、ユーザー数が VSImax v4 に近づいたところで CPU とメモリの使用量がピークに達しています。

以下のグラフは、平均的な RD セッションホストのテスト中の CPU 使用率を示しています。

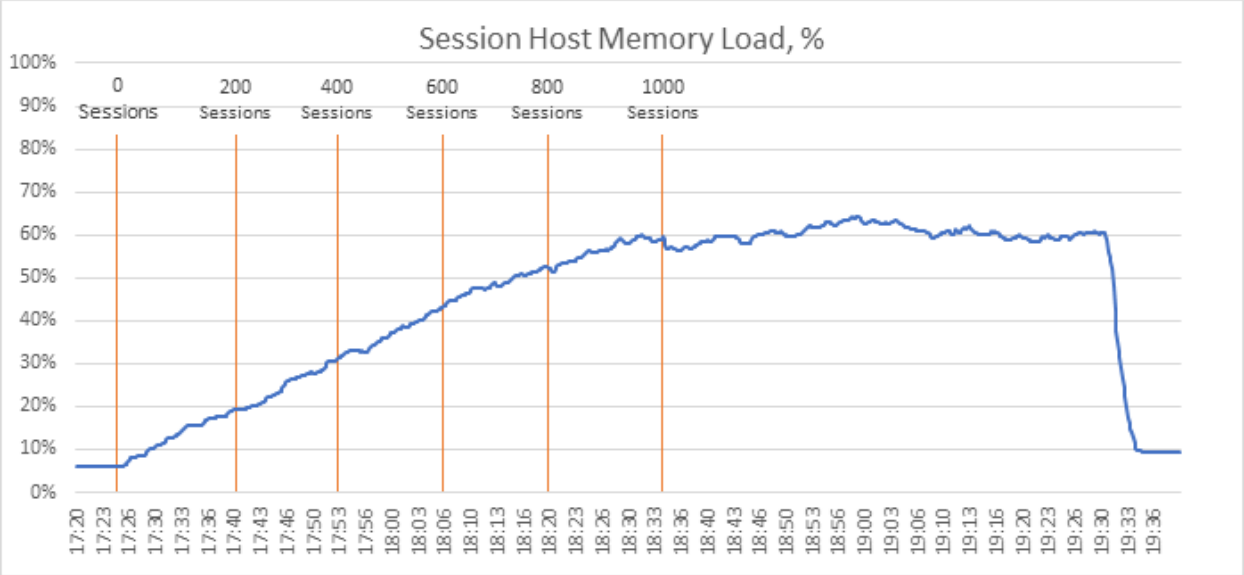


このグラフのログオン フェーズでは、多くのスパイクが見られますが、これはセッション ログオンによるもので、通常、新しいセッション ログオンごとにユーザー プロファイルの負荷がかかり、一般的に高負荷な操作となるためです。その後、定常テストフェーズでは、CPU 負荷は約 70.51 % にとどまり、最大値は 93.03 % と記録されました。これは、RD セッションホストの容量にまだ達していないことを示しており、VM 構成がテスト負荷に応じて適切に選択されていることを証明しています。

ワークロードフェーズのセッションホストの CPU 使用率は、セッション内のアクティビティに大きく

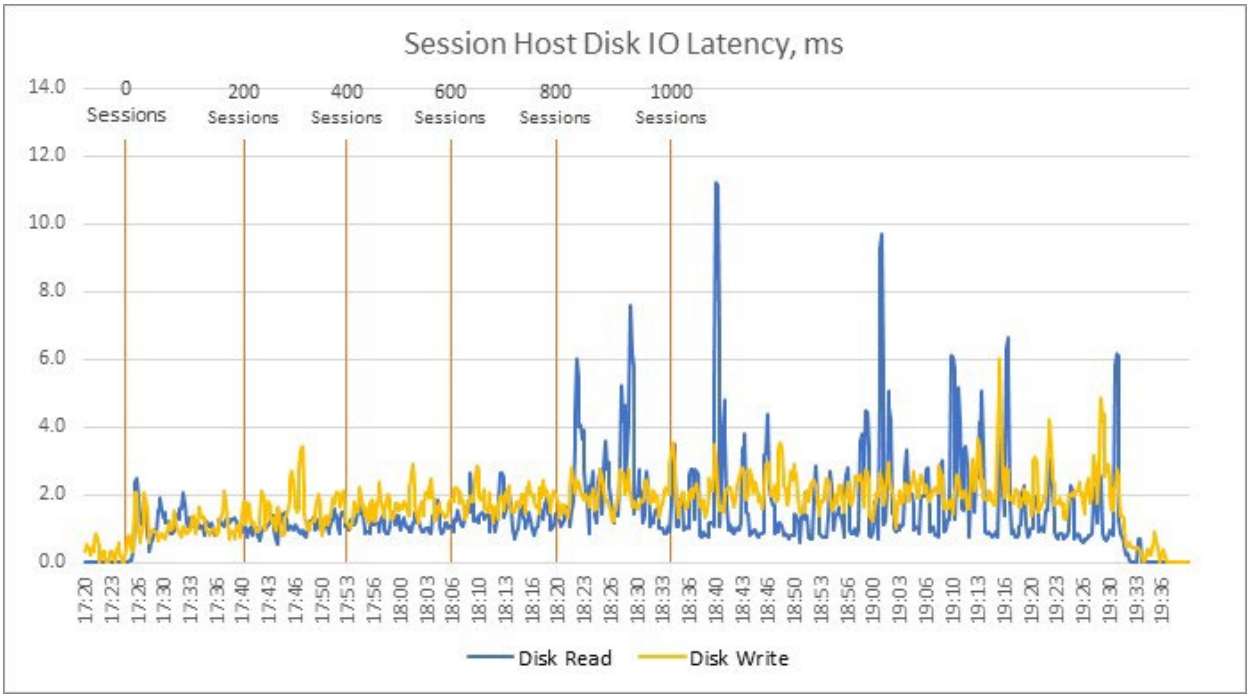
依存することを述べておきます。例えば、ビデオの再生はネットワークと CPU の使用量を増加させますが、より「静的な」アプリケーションで作業する場合は、画面の更新や I / O 操作が少なく、その結果、プロセッサの使用時間が短くなります。

以下のグラフは、平均的な RD セッションホストがテスト中に消費したメモリ (RAM) の平均値を示しています。負荷は、RAS Publishing Agent によって、テストに用いた 38 台の RD セッションホストすべてに均等に分散されています。RAM の消費量は、ログオン フェーズからワークロード フェーズまで着実に増加し、平均 59 % (11.8 GB) にとどまっています。

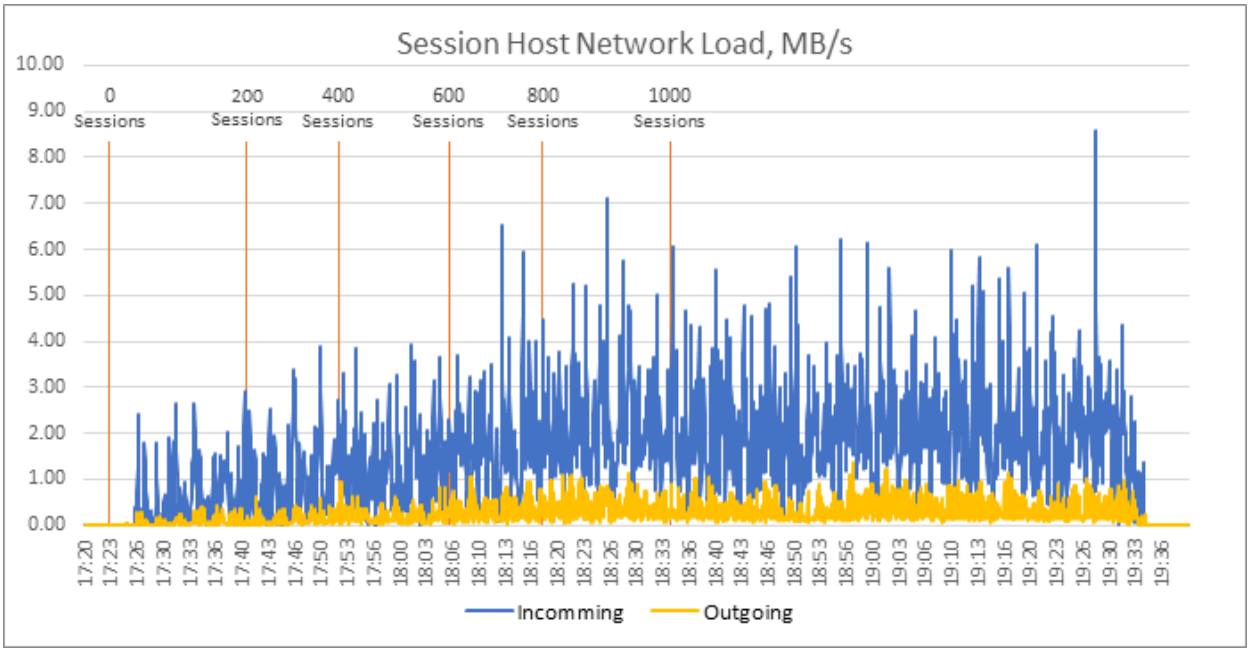


個々のユーザー セッションの作業負荷に依存する以外に、テストの定常フェーズで見られるように、RD セッションホストでのメモリ (RAM) 消費量も実行中のセッション数に大きく依存することに留意してください。

次のグラフは、ディスクの読み取りと書き込みの平均応答時間を示しています。ワークロード フェーズの平均書き込み I / O レスポンスタイムは約 0.646 ms、平均読み出し I / O レスポンスタイムは 2.01 ms です。



次の表は、RD セッションホストに出入りするデータのネットワーク転送速度を示しています。



Knowledge Worker ワークロードの場合、定常状態での平均送信帯域幅は、26 人のユーザーからなるテストグループ (1,000人の総ユーザー数を 38 台の RD セッションホストサーバーで割ったもの) で約 440 kB / s です。したがって、ユーザーあたりの送信転送レートは約 16.92 kB / s となります。受信トラフィックは、プロファイルのロードやローカル ネットワーク共有のファイルへのアクセスなど、主にローカル ネットワークの操作に関連するもので、その結果、受信トラフィックがより多く表示され、セッション ワークロードの性質によりグラフは大きく変動しています。

ワークロード フェーズのセッションホスト ディスクの I / O レイテンシーとネットワーク負荷の両方が、セッション内のアクティビティに大きく依存していることは、言及しておくべきことです。たとえば、ネットワーク共有からのビデオ再生は、ネットワークとディスク I / O カウンターの両方を増加させ、これらの操作は、対応するグラフに大きなスパイクを生成します。

CHAPTER 3

RAS インフラストラクチャー コンポーネント

この章の内容

RAS Publishing Agent.....13
RAS Secure Client Gateway 14
HALB 16

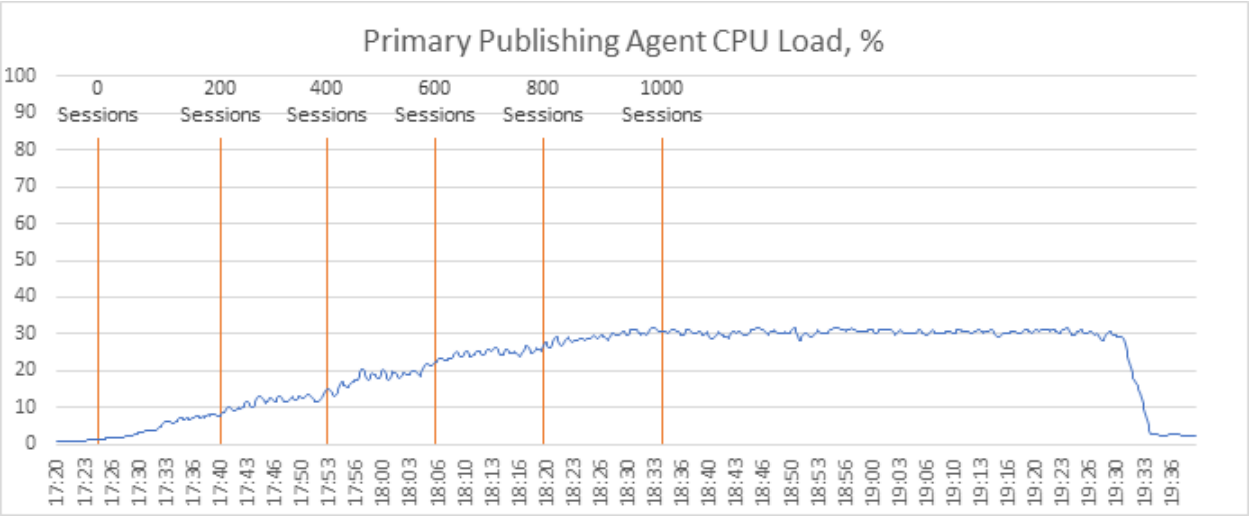
RAS Publishing Agent

RAS Publishing Agent (PA) は、公開されたアプリケーションとデスクトップのロード バランシングを提供します。

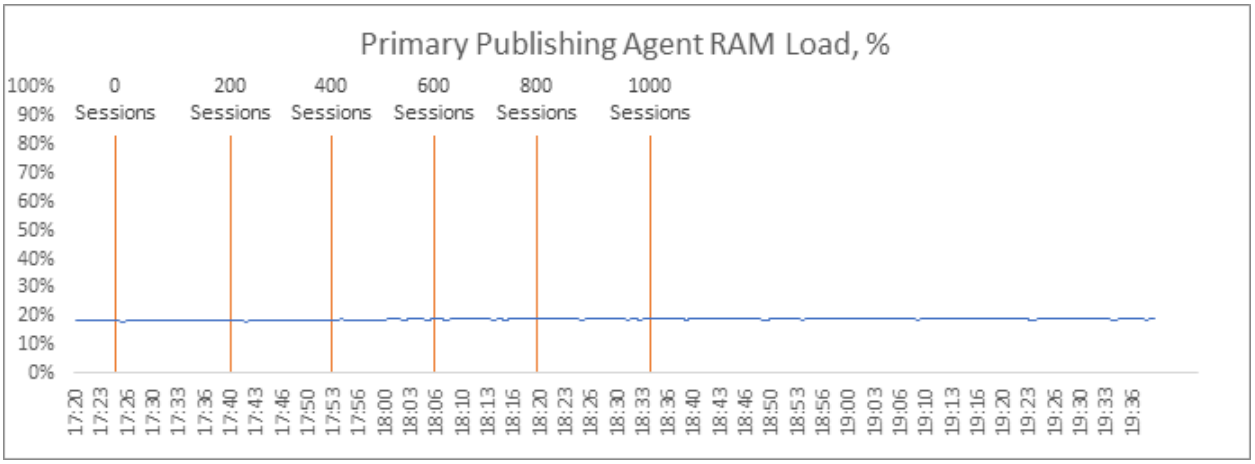
以下のグラフは、RAS 環境における 2 つの PA の平均 CPU 使用率を、ログオン時およびテスト中に起動したセッションの数で示したものです。
負荷はすべての PA、つまりプライマリとセカンダリの PA に均等に配分されています。

以下のグラフは、RAS 環境におけるプライマリ PA のテスト中の平均 CPU 使用率を示しています。テスト中のセカンダリ PA の CPU % 使用率は、プライマリ PA の CPU % 使用率とほぼ同じでした。

テスト全体を通して、PA の平均 CPU 使用率は平均 30 % で安定していることに注目できます。これは、1,000 セッションの同時実行というワークロードに対して、CPU のスペックが十分であったことを示しています。



以下のグラフは、RAS 環境における 2 つの PA の平均メモリ (RAM) 稼働率を、ログオン フェーズとその後のワークロード フェーズで起動したセッション数に基づいて示しています。負荷はすべての PA、つまりプライマリとセカンダリの PA に等しく分散されています。

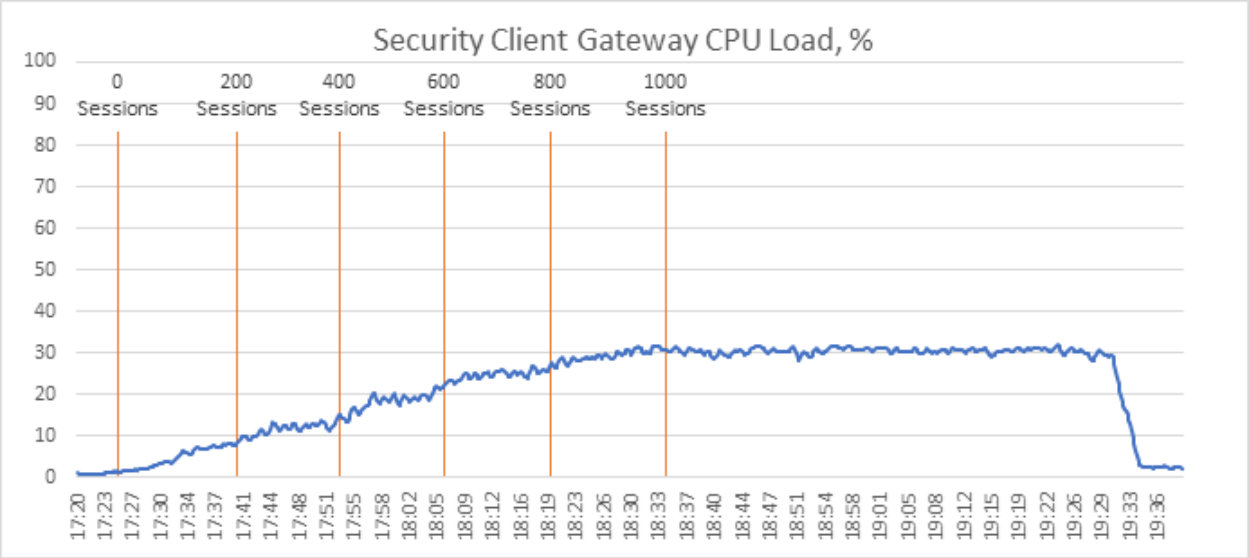


PA のメモリ (RAM) 利用率は、ほぼ全試験期間を通じて平均 19 % と安定しており、十分な利用が可能であることがわかります。これは、より少ない量のメモリが割り当てられたとしても、現在のユーザー数と作業負荷に対して十分であることを意味しています。

RAS Secure Client Gateway

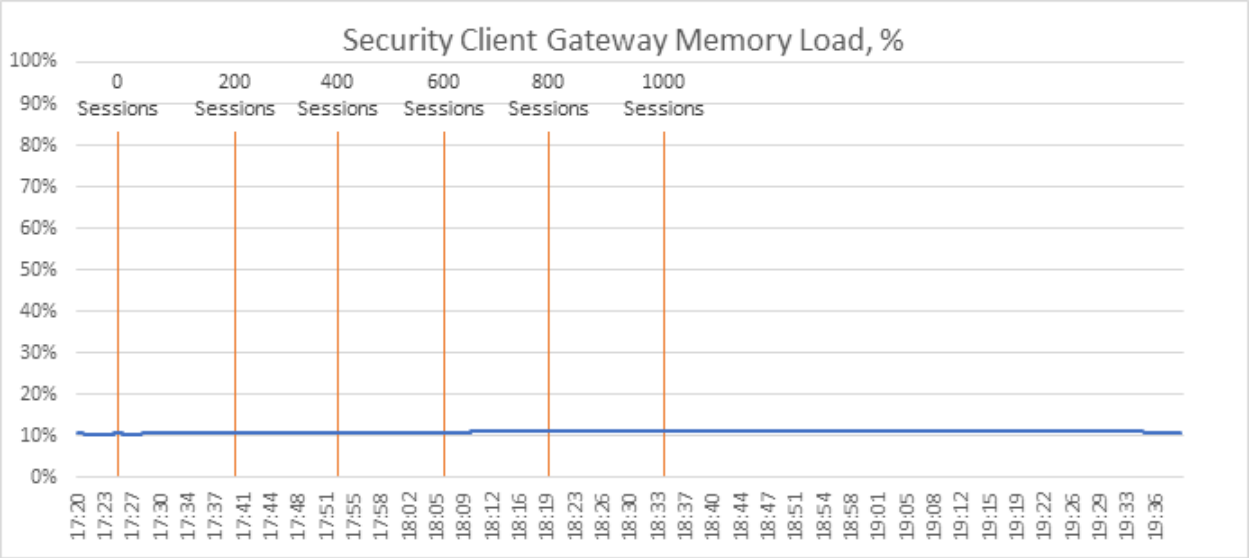
RAS Secure Client Gateway (SCG) は、すべての Parallels RAS データを 1 つのポートにトンネリングします。また、セキュアな接続を提供し、Parallels RAS へのユーザーの接続ポイントになります。

以下のグラフは、ログオン時に起動したセッション数に対する、4 つの SCG の平均 CPU 使用率を示しています。SCG あたり 250 セッションのピークに達するまで、RAS HALB アプライアンスを使用してすべての SCG に均等に負荷が分散されています。



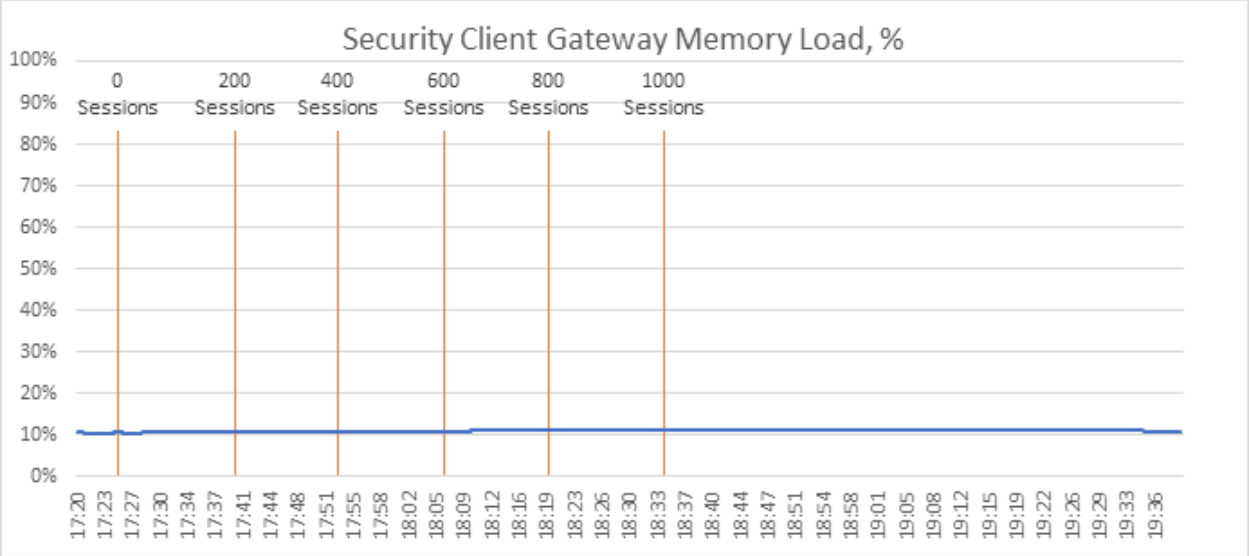
SCG の平均 CPU 使用率は、予想通りユーザーセッションの増加に伴い、着実に増加していることがわかります。また、テストフェーズの SCG の CPU 使用率は非常に安定しており、30 % 前後で推移しています。したがって、CPU の仕様は、4 つのゲートウェイで 1,000 の同時セッション、ゲートウェイあたり平均 250 人のユーザーというワークロードに対して十分に適切であったと結論づけることができます。

下のグラフは、テスト中の平均的な SCG のメモリ (RAM) 消費量を示しています。負荷はすべての SCG に均等に分散されています。



SCG の平均メモリ (RAM) 消費量は、1,000 ユーザー セッションに達した時点でも、常に平均 11 % (約 918 MB) の利用可能メモリで十分すぎるほどであることがわかります。これは、より少ない量のメモリが割り当てられたとしても、現在のユーザー数とワークロードに対して十分であることを意味します。

次のグラフは、SCG を介したネットワーク転送速度の平均値を示しています。予想通り、ネットワーク使用率はユーザー セッションのログオン回数に比例して増加します。作業負荷フェーズ (1,000 セッションに達した後) では、平均ネットワーク スループットは各 SCG で 9.90MB / s を記録し、すべての SCG を考慮すると合計 39.6 MB / s となり、これはユーザーあたり 39.6 KB / s に相当します。最大転送速度はワークロード フェーズで 12.77 MB / 秒、ユーザーあたり 51.08 KB / s と記録されています。

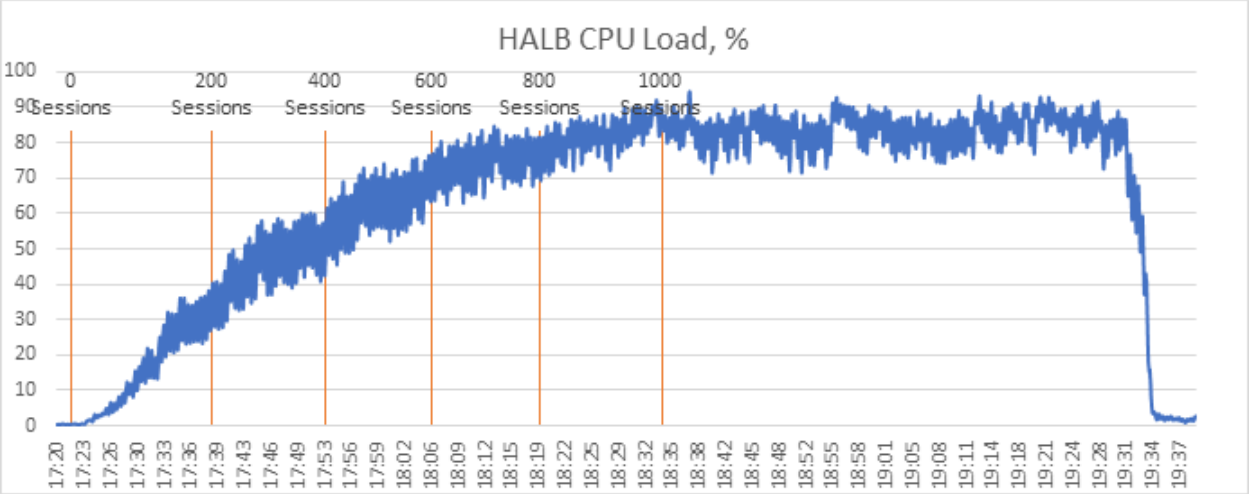


しかし、ワークロード フェーズにおける SCG の CPU とネットワークの使用量は、セッション中のアクティビティに大きく依存することに留意する必要があります。たとえば、ビデオの再生はネットワークと CPU の使用量を増加させますが、Microsoft Word のような画面更新の少ない「静的」なアプリケーションでの作業では SCG リソースはあまり使用されません。

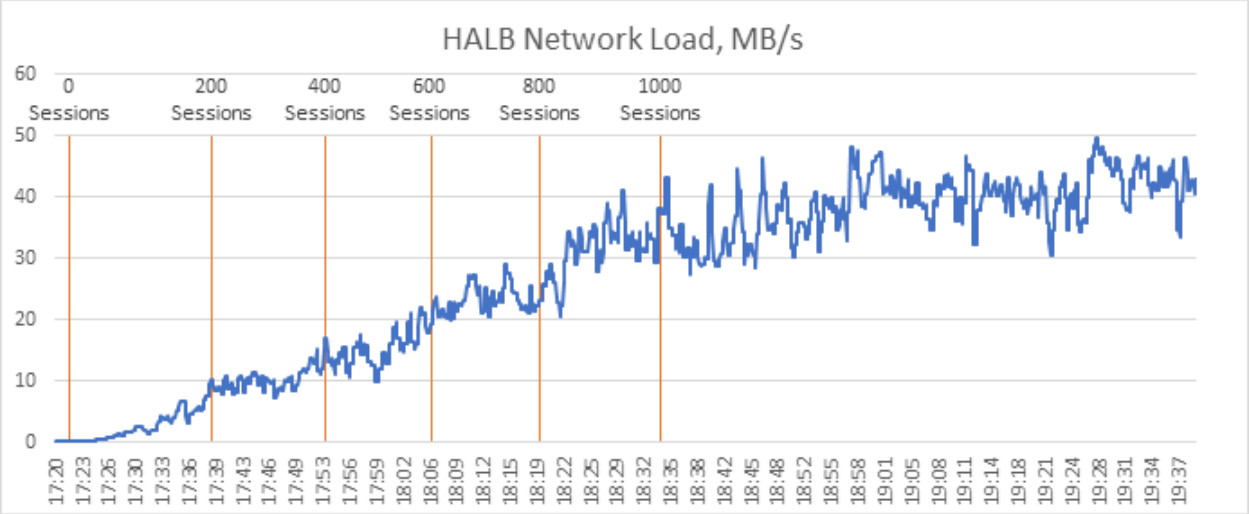
HALB

HALB (High Availability Load Balancing) は、RAS セキュア クライアント ゲートウェイに負荷分散を行うアプライアンスです。Parallels HALB アプライアンスは、オペレーティング システムがインストールされ、関連するすべての設定が構成された、事前構成済みの仮想マシンで構成されています。すべてのリモートセッション トラフィックは HALB を通過します。

下のグラフは、HALB アプライアンスの平均 CPU 使用率を示しています。テストフェーズの平均 CPU 負荷は 76.44 % で、HALB のハードウェアの限界にまだ達していないことを意味します。



次のグラフでは、ログオン時の HALB 経由のネットワーク転送速度がテスト時の平均値 39.16 MB / s に達するまで、順調に伸びていることが分かります。最大転送速度は、49.74 MB / s (ユーザーあたり 50.9 KB / s) で、ワークロード フェーズで確認されました。



HALB スケーラビリティの追加情報については、KB https://kb.parallels.com/125229 を参照してください。

CHAPTER 4

結 論

本書で紹介する Parallels RAS のスケーラビリティの結果、Knowledge Worker ワークロードを使用する 952 の Login VSI セッションが、所定のハードウェア構成で正常に起動することが確認されました。特に、8 個の vCPU と 20 GB の RAM を搭載した合計 38 台の RD セッション ホスト サーバーでユーザー体験を損なうことなく、これらのセッションに対応することができました。

VMware ESXi 6.7 上に Parallels RAS を以下のように導入しました。

Parallels RAS コンポーネント	VM の総数	各 VM の vCPU	各 VM の RAM
RAS Publishing Agent	2	2	8 GB
RAS Secure Client Gateway	4	2	8 GB
High Availability Load Balancing	1	1	2 GB
RD Session Host	38	8	20 GB

負荷テストとスケーラビリティ テストは、プラットフォームとソリューション全体のパフォーマンスを理解する上で重要な要素ですが、この文書で得られた結果と提示された結果は、実際の本番ワークロードの正確な測定値として推論されるべきものではないことに留意することが重要です。アプリケーションがどのように動作するかをより適切に評価したいお客様には、独自のワークロード サンプルで独自の負荷およびスケーラビリティ テストを実施することをお勧めします。さらに、Parallels RAS の概念実証 (POC) またはパイロット版をリクエストして、本番環境に移行する前に設計、導入、サイジングを支援することができます。

Parallels RAS の詳細、機能、および利点については、https://parallels.com/ras をご覧ください。

■ 本内容のお問い合わせ先:
コーレル株式会社 ジャパン マーケティング
〒108-0075 東京都港区港南 2-16-5 NBF 品川タワー 6 階
お問い合わせフォーム: https://jp.learn.corel.com/volume-licensing/inquiry/
E-mail: ras-jp@parallels.com

■ 各種ビジネス向け情報を掲載:
コーレル ディスカバリー センター ジャパン (https://jp.learn.corel.com/category/business/)
ディスカバリー センターには、Parallels RAS や他製品に関する数々のお役立ち情報・資料を掲載しております。是非、ご活用ください。