

クラウド上のIRIS - サイジングのベストプラクティス -

インターシステムズジャパン株式会社
セールスエンジニア
秦 信之

2023年10月18日

アジェンダ



1 **Cloud Resource**

2 **System Metrics**

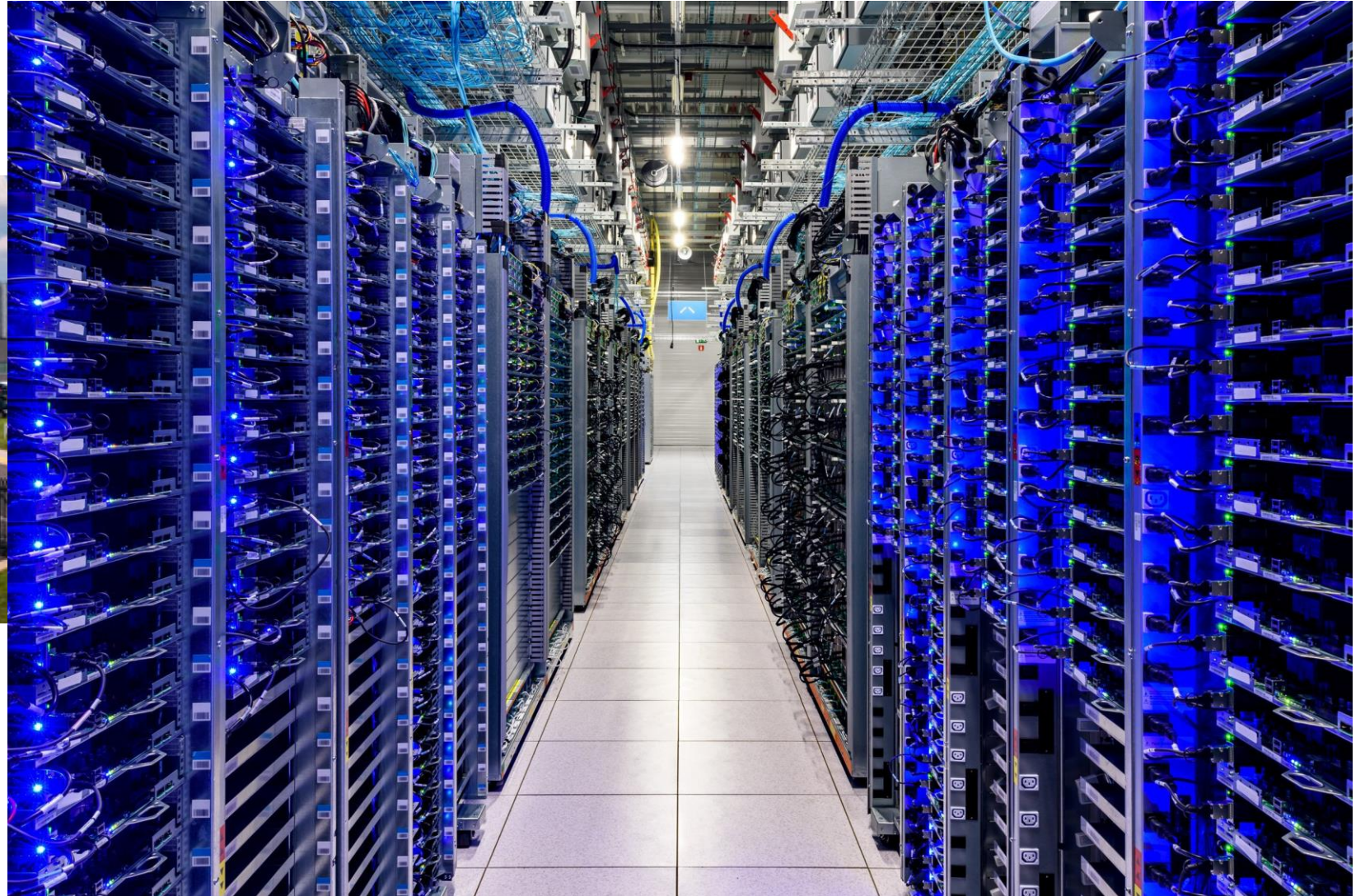
3 **Example**

4 **Summary**

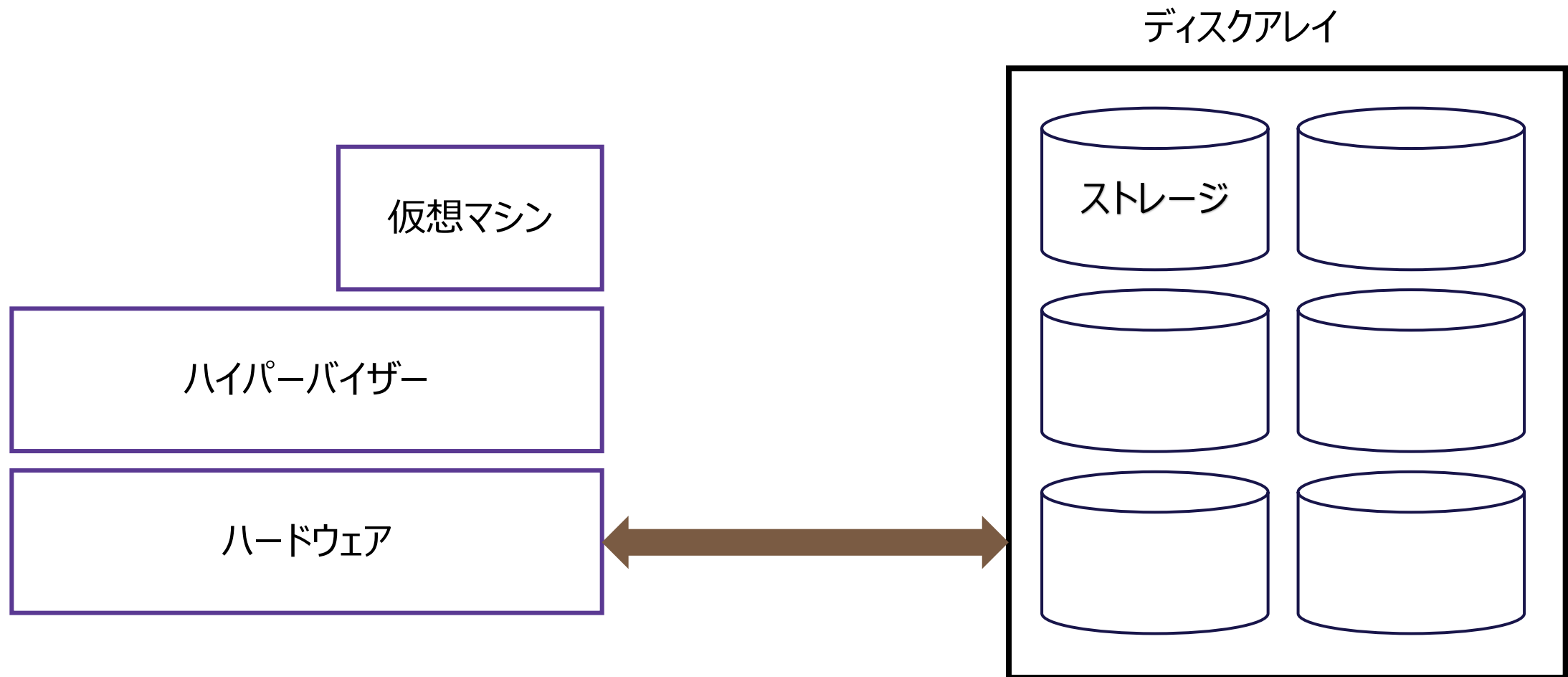
A vertical white line on the left side of the slide, with a small white dot at the top and a small white dot at the bottom.

Cloud Resources

クラウドとは?



クラウドコンピューティング





仮想マシン – コンピューティングリソース

- 特徴
 - ELASTIC – 伸び縮みする, 融通性のある
- 複数のタイプ
 - 一般向け、メモリ最適化、ストレージ最適化等
 - タイプがvCPU, RAM, Network, Storageの組み合わせに反映
 - 同じタイプに複数の世代
- サイジングの目安
 - 目標CPU使用率 70-80%
 - 定期的なリソースの監視が必須
- オートスケーリング
 - 予測可能なアクセス急増減時にインスタンス数を調整



ストレージ

- 複数のタイプ
 - ブロック
 - ファイル
 - オブジェクト
- サイジングの目安
 - 必要十分なIOPS
 - 短期的に必要な容量
 - 定期的な空き容量監視
- 注意点
 - IOPSと帯域(MB/s)に上限あり

AWSの場合



- インフラストラクチャリソース

- インスタンス
- ストレージ
- ネットワーク

インスタンスサイズ vCPU メモリ (GiB)

r5n.large	2	16
r5n.xlarge	4	32
r5n.2xlarge	8	64
r5n.4xlarge	16	128
r5n.8xlarge	32	256
r5n.12xlarge	48	384
r5n.16xlarge	64	512
r5n.24xlarge	96	768
r5n.metal	96*	768

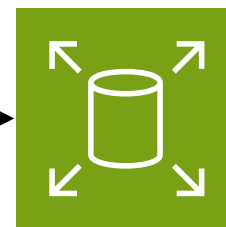


R5インスタンス

r5n.12xlarge
最大IOPS: 40,000 (16KiB I/O)
最大帯域: 9,500 (Mbps)
最大スループット: 1,187 (MB/s, 128KiB I/O)
ネットワーク: 50 (Gbps)

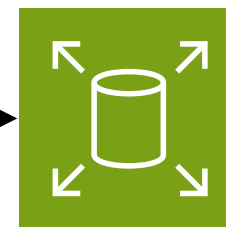
vCPU: 48
メモリー: 384 (GB)

EBSネットワーク



Amazon EBS

1個以上



Amazon EBS

EBS gp3
無料IOPS: 3,000
最大IOPS: 16,000
スループット: 125MB/s
最大スループット: 1,000MB/s

1桁ミリ秒のレイテンシーを、年間99%の時間でプロビジョ
ンされた数値の上下10%範囲内の性能を提供

2023/10/1現在 詳細はサービス提供会社にご確認下さい

リソースマッピング: AWSストレージ



ボリュームタイプ	最大 IOPS	最大スループット (MB/s)	最大IOPS – インスタンス	最大スループット (MB/s) – インスタンス
gp3	16,000	1,000	260,000	10,000
io2 Block Express	256,000	4,000	350,000	10,000

gp3

レイテンシー: 10ミリ秒以下

容量: 1 GB - 16 TB

価格: \$0.096USD/GB-1カ月

プロビジョンパフォーマンス価格:

- 3,000 IOPS
- \$0.006USD/プロビジョンドIOPS-1カ月(3,000 IOPS超過分)
- 125 MB/s スループット
- \$0.048USD/プロビジョンドMB/s-1カ月(125MB/s超過分)

io2 Block Express

レイテンシー: 1ミリ秒以下

容量: 4 GB – 64 TB

価格: \$0.142USD/GB-1カ月

プロビジョンパフォーマンス価格:

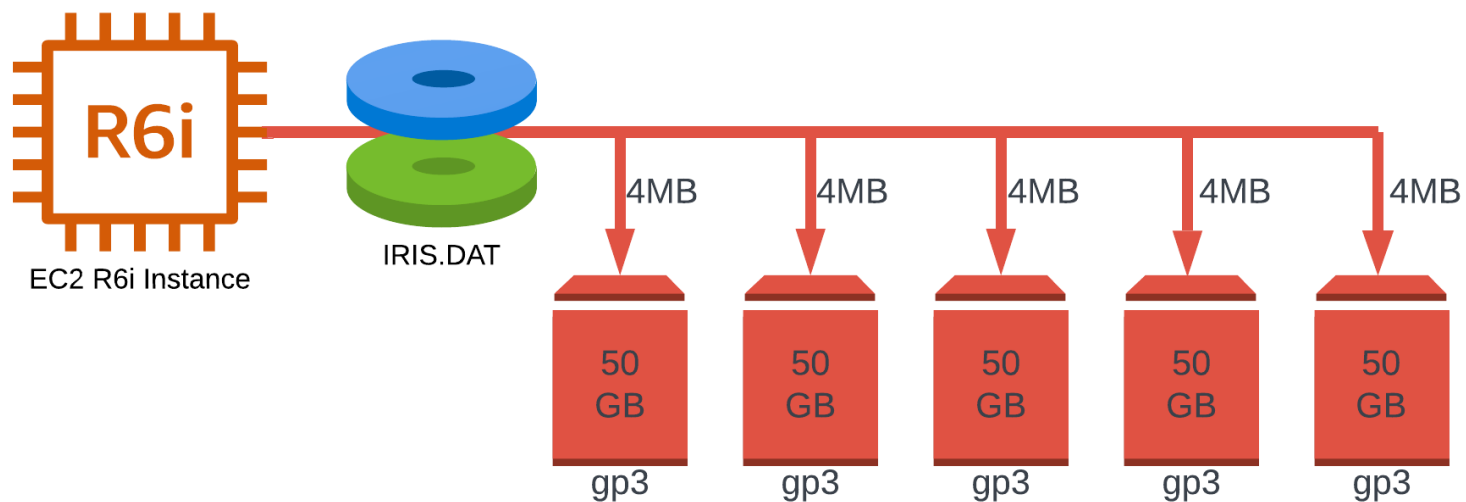
- \$0.074USD/プロビジョンドIOPS-1カ月(32,000 IOPSまで)
- \$0.052USD/プロビジョンドIOPS-1カ月(32,001から64,000 IOPSまで)
- \$0.036USD/プロビジョンドIOPS-1カ月(64,000 IOPS以上)

LVM PE ストライピング



- 複数のディスクデバイスで論理ボリュームを構成し、シングルファイルシステムとしてマウント
 - IOに関わるデバイスの数を増やすことで、データベースやファイルに小さい単位（例では4MB）のラウンドロビンでデバイスにアクセス

- $50 \text{ GB} \times 5 = 250 \text{ GB}$
- $3,000 \text{ IOPS} \times 5 = 15,000 \text{ IOPS}$ **もしくは**
- $125 \text{ MB} \times 5 = 625 \text{ MB/s}$

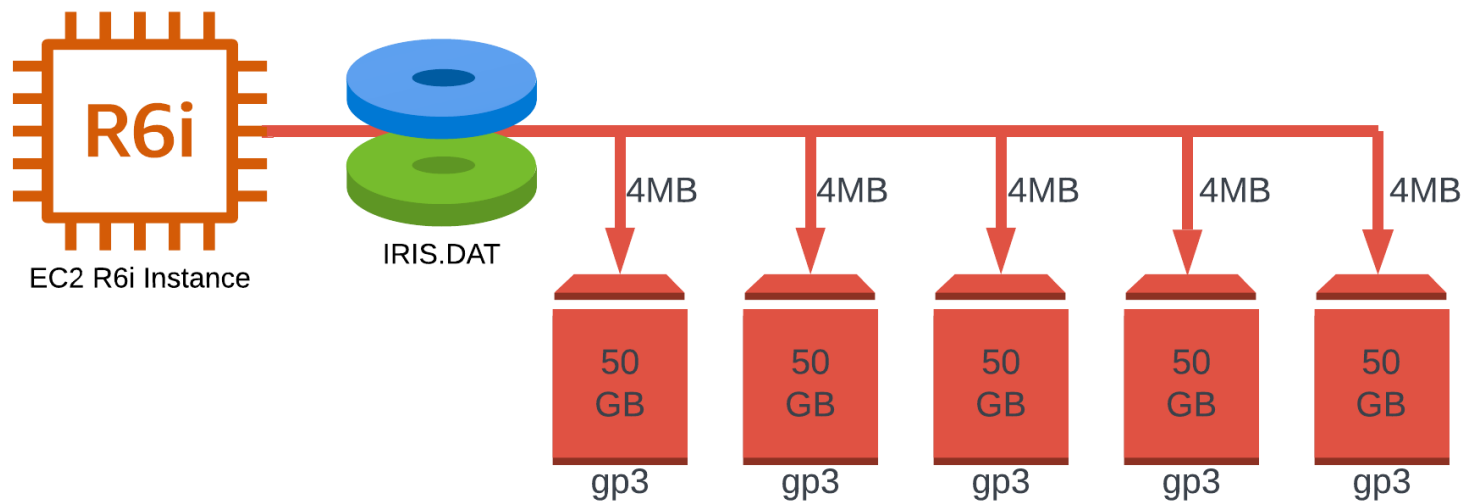
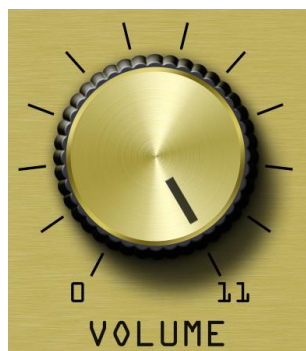


LVM PE ストライピング

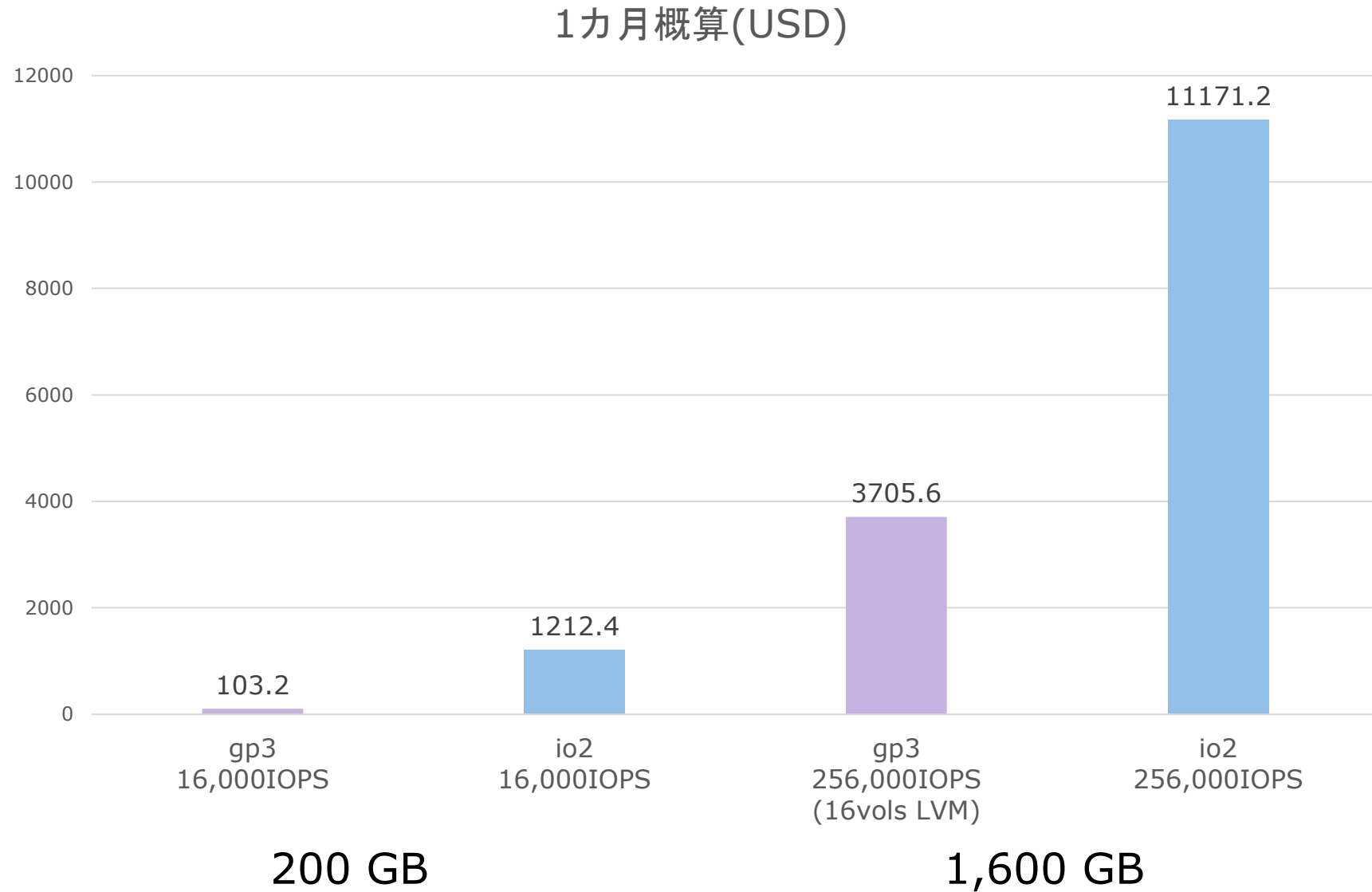


- 複数のディスクデバイスで論理ボリュームを構成し、シングルファイルシステムとしてマウント
 - IOに関わるデバイスの数を増やすことで、データベースやファイルに小さい単位（例では4MB）のラウンドロビンでデバイスにアクセス

- $50 \text{ GB} \times 5 = 250 \text{ GB}$
- $16,000 \text{ IOPS} \times 5 = 80,000 \text{ IOPS}$ **もしくは**
- $1,000 \text{ MB/s} \times 5 = 5,000 \text{ MB/s}$



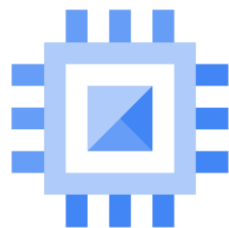
ストレージ費用月額 - 計算例



GCPの場合



- インスタンス
- ストレージ
- ネットワーク



ネットワーク



1個以上



n2-highmem-96
最大IOPS: 80,000
最大帯域: ネットワークの60%
最大スループット: 1,200MB/s
ネットワーク: 32Gbps -外向き

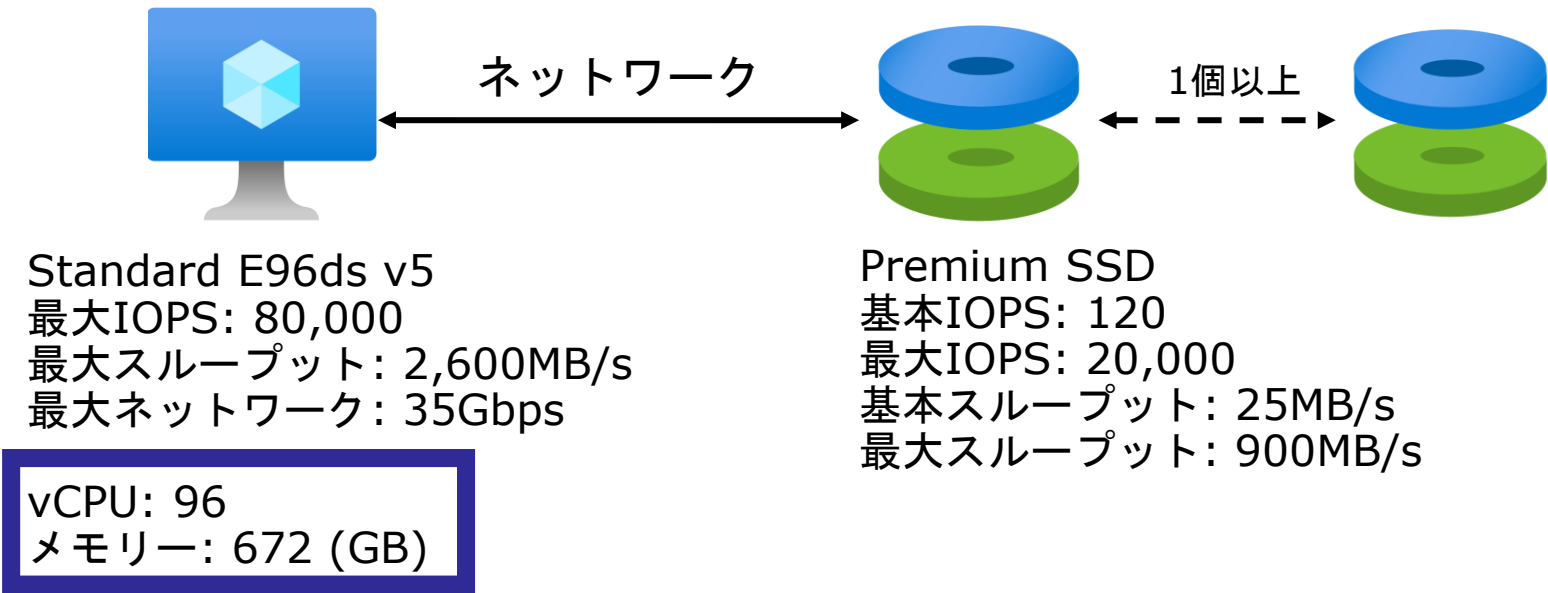
vCPU: 96
メモリー: 768 (GB)

ゾーンバランスPD
無料IOPS: 6 IOPS/GB
最大IOPS: 80,000
スループット: 0.28 MB/s /GB
最大スループット: 1,200MB/s

Azureの場合



- インスタンス
- ストレージ
- ネットワーク



Premium SSD	P1	P2	P3	P4	P6	P10	P15	P20	P30	P40	P50	P60
ディスクサイズ(GB)	4	8	16	32	64	128	256	512	1024	2048	4096	8192
IOPS(/s)	120	120	120	120	240	500	1100	2300	5000	7500	7500	16000



2023/10/1現在 詳細はサービス提供会社にご確認下さい



Collecting System Metrics

IRISアプリケーションをクラウドに移行

- まずはクラウドに移行して、後に最適化
 - Lift & Shift
- ベースライン
 - ユーザー体験を重視
 - 移行後でもユーザーは同じレスポンス時間で利用できるか
- キャパシティプランニング
 - インスタンスとストレージタイプを選択



システム指標の収集

- OS
 - mpstat, vmstat, iostat, Windowsパフォーマンスモニター等
- IRISドキュメント (“監視ガイド”で検索)
 - System Alerting and Monitoring (SAM)
 - 管理ポータル
 - **SystemPerformance (aka. pButtons)**
 - 履歴モニタ
- サードパーティアプリケーション
 - Nagios, Splunk,
 - Prometheus + Grafana

InterSystems IRIS Data Platform 2023.1

▼ 管理

About...

＞ システム管理ガイド

▼ 監視

About...

▼ 監視ガイド

管理ポータルを使用した
InterSystems IRIS の監視

InterSystems 診断レポートの使用
法

ログ・モニタの使用

システム・モニタの使用

^GLOSTAT を使用したグローバル
動作の統計収集

^PERFMON を使用したシステム・
パフォーマンスの監視

^PROFILE を使用したルーチン・パフ
ォーマンスの監視

^%SYS.MONLBL を使用したルー
チン・パフォーマンスの検証

^TRACE を使用したプロセス・パフ
ォーマンスのトレース

**^SystemPerformance を使用し
たパフォーマンスの監視**

^mgstat を使用したパフォーマンス
の監視

履歴モニタ

^BLKCOL を使用したブロック衝突
の監視

^PERFSAMPLE を使用したプロセ
スの監視



^SystemPerformance (a.k.a. pButtons)



- IRISに同梱
- IRISの監視ガイドにドキュメントあり
- 非常に小さいオーバーヘッド
- IRISの詳細と性能情報を出力
- ハードウェアとOSの詳細とOSの性能情報を出力





^SystemPerformance (a.k.a. pButtons)

- 24時間収集を毎日深夜0時に起動をスケジュール
- トrend把握には10秒毎、**キャパシティプランニングには5秒未満**
- 推奨: **出力ディレクトリ変更、フォルダのパーミッションチェック、圧縮**

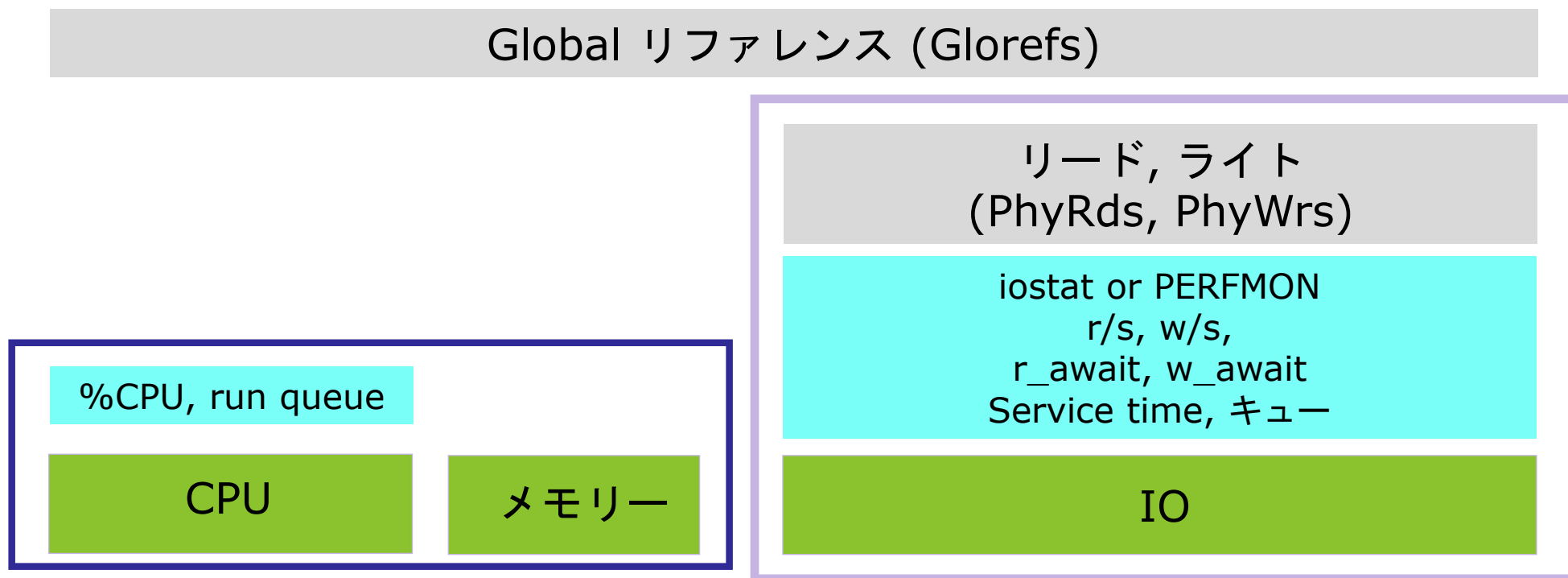
```
%SYS>d ^SystemPerformance
Re-creating command data for new ^SystemPerformance version.
Old command data saved in ^IRIS.SystemPerformance("oldcmds").
Current log directory: c:\intersystems\irishealth\mgr\
Windows Perfmon data will be left in raw format.
Available profiles:
   1 12hours      - 12 hour run sampling every 10 seconds
   2 24hours      - 24 hour run sampling every 10 seconds
   3 30mins       - 30 minute run sampling every 1 second
   4 4hours       - 4 hour run sampling every 5 seconds
   5 8hours       - 8 hour run sampling every 10 seconds
   6 test        - A 5 minute TEST run sampling every 30 seconds

Select profile number to run: █
```

IRIS と OS: ストレージ



- データベース – ランダムリード (連続的) と ライト (ライトデーモン)
- ライトイメージジャーナル (WIJ) – シーケンシャルライト (ライトデーモン)
- トランザクションジャーナル



IRIS と OS: ストレージ



入出力の種類	タイミング	方法	留意事項	入出力の種類	平均応答時間	最大応答時間	留意事項
データベースの読み取り (ほぼランダム)	インタフェースおよびユーザの処理ごとに継続的	インタフェース・クエリまたはユーザの処理によってディスク入出力が開始され、データを読み取り	データベースの読み取りは、デーモンが処理する Web ページ、SQL クエリ、または直接ユーザ処理によって実行されます。	データベースのランダム読み取り (キャッシュ処理なし)	<= 1 ms	<= 2 ms	データベース・ブロックは固定 8KB、16KB、32KB、または 64KB (ホストでのデータベース・キャッシュは大きいいため、ディスクへのほとんどの読み取りはキャッシュされません)。
データベースの書き込み (順序付けされるが非連続で実行)	約 80 秒ごとまたは保留中の更新がデータベース・キャッシュのしきい値のパーセントに達したとき (先に基準を満たした方)	データベース・ライト・デーモン (8 個のプロセス)	データベースの書き込みは、ライト・デーモンと呼ばれる一連のデータベース・システム・プロセスによって実行されます。ユーザ処理によってデータベース・キャッシュが更新され、トリガ (時間またはデータベース・キャッシュのパーセントが一杯になったとき) によってライト・デーモンを使用してディスクの更新が実行されます。通常、更新レートに応じて、書き込みサイクル中に数 MB から数 GB 書き込む必要があることを想定しています。	データベースのランダム書き込み (キャッシュ処理あり)	<= 1 ms	<= 2 ms	すべてのデータベース・ファイルの書き込みは、ストレージ・コントローラのキャッシュ・メモリによってキャッシュされることが想定されます。
WII 書き込み (連続)	約 80 秒ごとまたは保留中の更新がデータベース・キャッシュのしきい値のパーセントに達したとき (先に基準を満たした方)	データベース・マスタ・ライト・デーモン (1 個のプロセス)	ライト・イメージ・ジャーナル (WII) は、データベースの書き込みサイクル中のシステム障害から物理データベース・ファイルの整合性を保護するために使用されます。書き込みは、約 256KB のサイズごとに行われます。	ジャーナル書き込み	<= 1 ms	<= 2 ms	ジャーナル書き込みは連続で、4KB から 4MB までサイズが変化します。
ジャーナル書き込み (連続)	ジャーナル・データの 64KB ごとまたは 2 秒ごと、または ECP またはアプリケーションの同期要求	データベース・ジャーナル・デーモン (1 個のプロセス)	ジャーナル書き込みは連続で、4KB から 4MB までサイズが変化します。1 秒あたり数十回の書き込みから、ECP および個別のアプリケーション・サーバを使用した非常に大きい配置の場合には 1 秒あたり数千回の書き込みまでになる場合があります。				





アプリケーション

- ユーザーレスポンス時間
- トランザクション
 - 秒間の (処理|アクセス|データ追加) 件数
 - 同時アクセスユーザー
- アプリケーションの性能指標
 - コンポーネントレスポンス時間, Glorefs, Rourefs, 回数, ランダムクエリ時間

A vertical white line on the left side of the slide, with a small white dot at the top and a small white dot at the bottom.

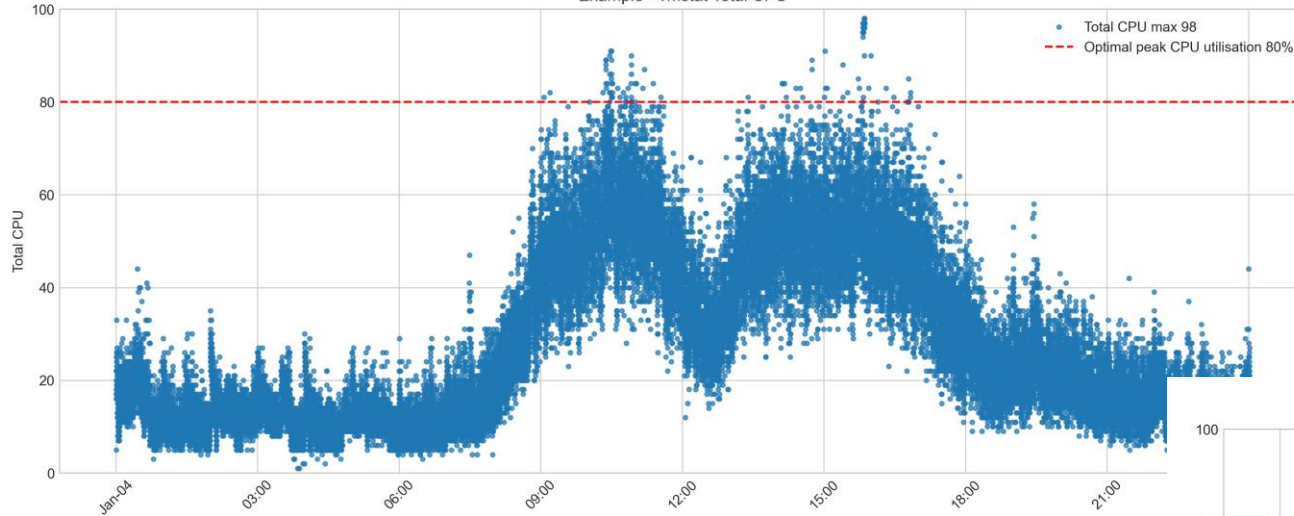
Example

オンプレミスからクラウドへの移行例

オンプレミス - 26 CPU, 204 GBメモリー

26 vCPU CPU E5-2680 v4 @ 2.40GHz - Wed, Jan 04, 2023 00:00:24 - Thu, Jan 05, 2023 00:01:12

Example - vmstat Total CPU



インスタンスサイズ

vCPU

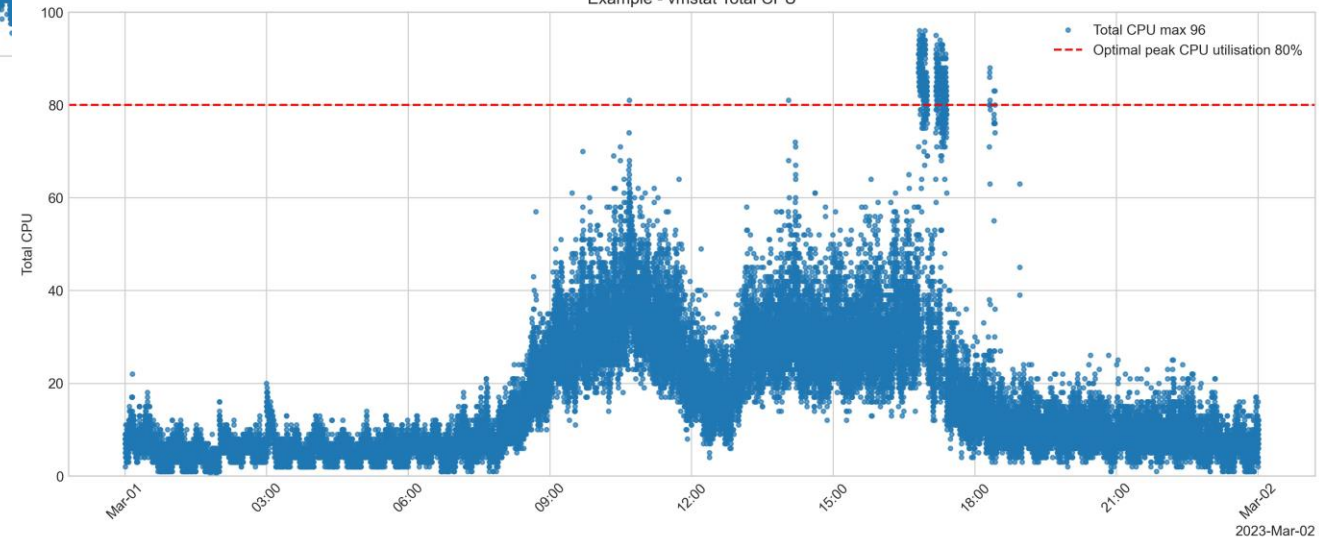
メモリー (GiB)

r5n.large	2	16
r5n.xlarge	4	32
r5n.2xlarge	8	64
r5n.4xlarge	16	128
r5n.8xlarge	32	256
r5n.12xlarge	48	384
r5n.16xlarge	64	512
r5n.24xlarge	96	768
r5n.metal	96	768

AWS - 48 vCPU, 384 GBメモリー

48 vCPU Platinum 8175M CPU @ 2.50GHz - Wed, Mar 01, 2023 00:00:25 - Thu, Mar 02, 2023 00:00:36

Example - vmstat Total CPU

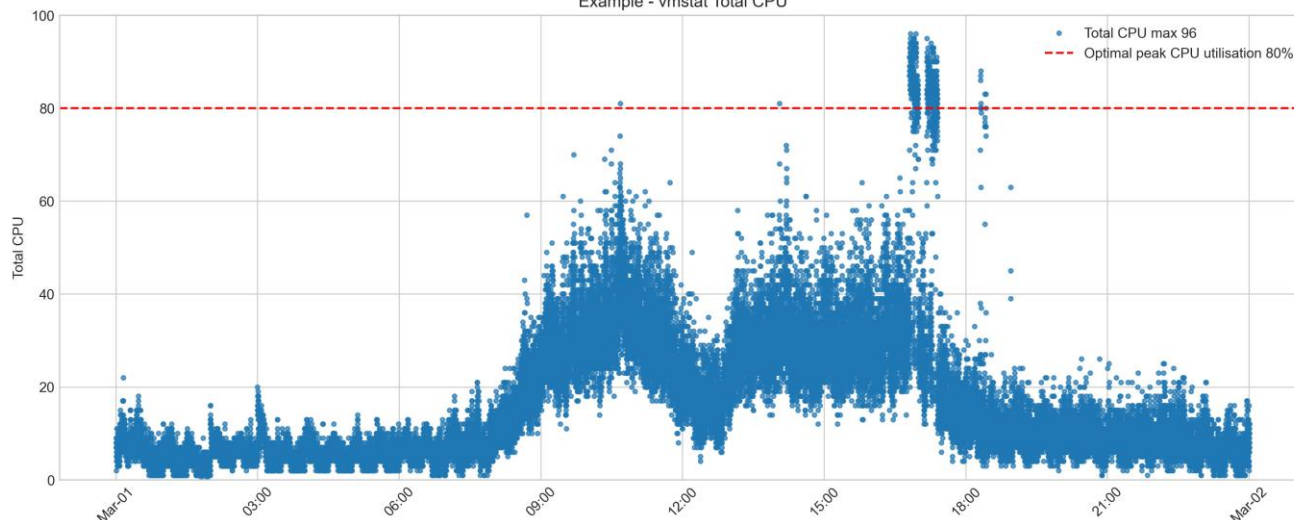


オンプレミスからクラウドへの移行例 - リサイズ

マイグレーション

48 vCPU Platinum 8175M CPU @ 2.50GHz - Wed, Mar 01, 2023 00:00:25 - Thu, Mar 02, 2023 00:00:36

Example - vmstat Total CPU

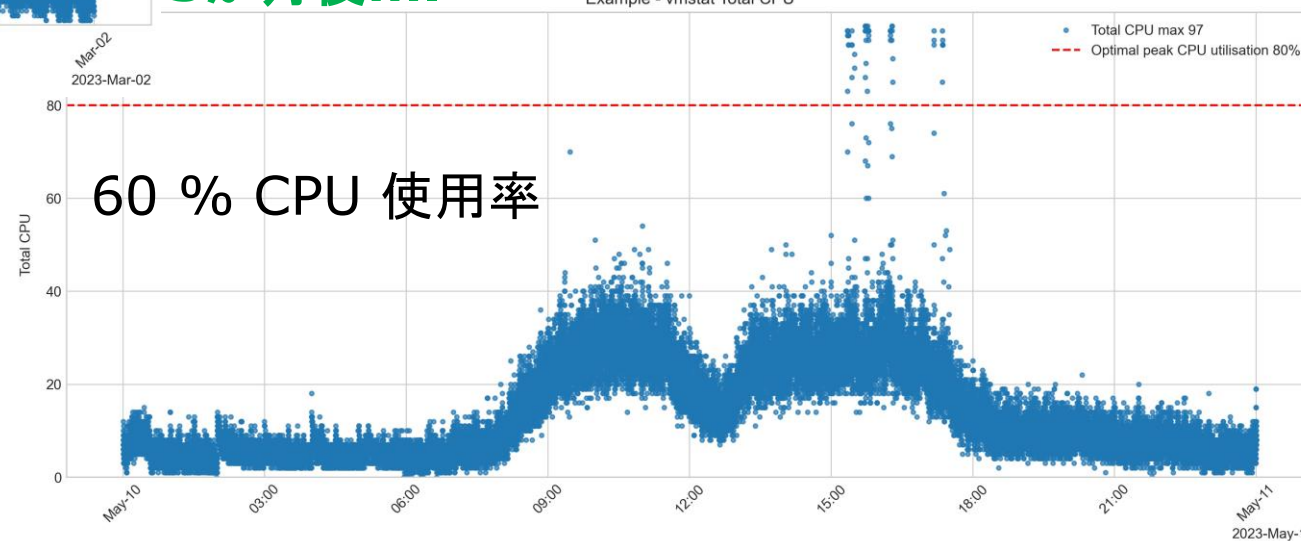


インスタンスサイズ	vCPU	メモリ (GiB)
r6i.large	2	16
r6i.xlarge	4	32
r6i.2xlarge	8	64
r6i.4xlarge	16	128
r6i.8xlarge	32	256
r6i.12xlarge	48	384
r6i.16xlarge	64	512
r6i.24xlarge	96	768
r6i.32xlarge	128	1,024
r6i.metal	128	1,024

3か月後....

48 vCPU Platinum 8259CL CPU @ 2.50GHz - Wed, May 10, 2023 00:00:28 - Thu, May 11, 2023 00:00:39

Example - vmstat Total CPU



60 % CPU 使用率

IOPSとスループットの上限を忘れずに



オンプレミスからクラウドへの移行例 - ストレージ



要件	データ (読み + 書き)	WIJ	ジャーナル	その他
IOPS	40,000	1.800	2,600**	e.g. EFS
スループット (MB/s)	600	450	50	

ボリュームタイプ	最大IOPS / ボリューム	最大スループット (MB/s) / ボリューム	最大IOPS / インスタンス	最大スループット (MB/s) / インスタンス
gp3	3,000 – 16,000	125 - 1,000	260,000	10,000

- **3,000 IOPS 付属** + 3,000IOPS以上は\$0.006/プロビジョンIOPS・月;
- **125 MB/s スループット 付属** + 124Mib/s以上は\$0.048/プロビジョンMB/s・月

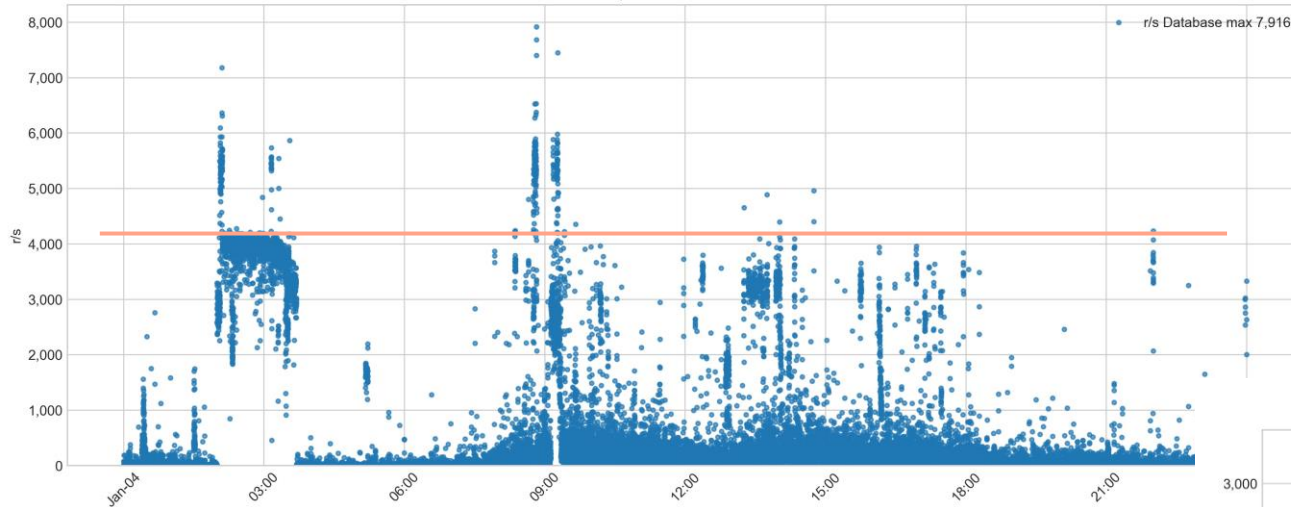
	データ gp3	WIJ gp3	ジャーナル gp3	その他
IOPS	4 x 10,000 = 40,000	1	1	
スループット (MB/s)	4 x 250 = 1 GB/s	450		

オンプレミスからクラウドへの移行例



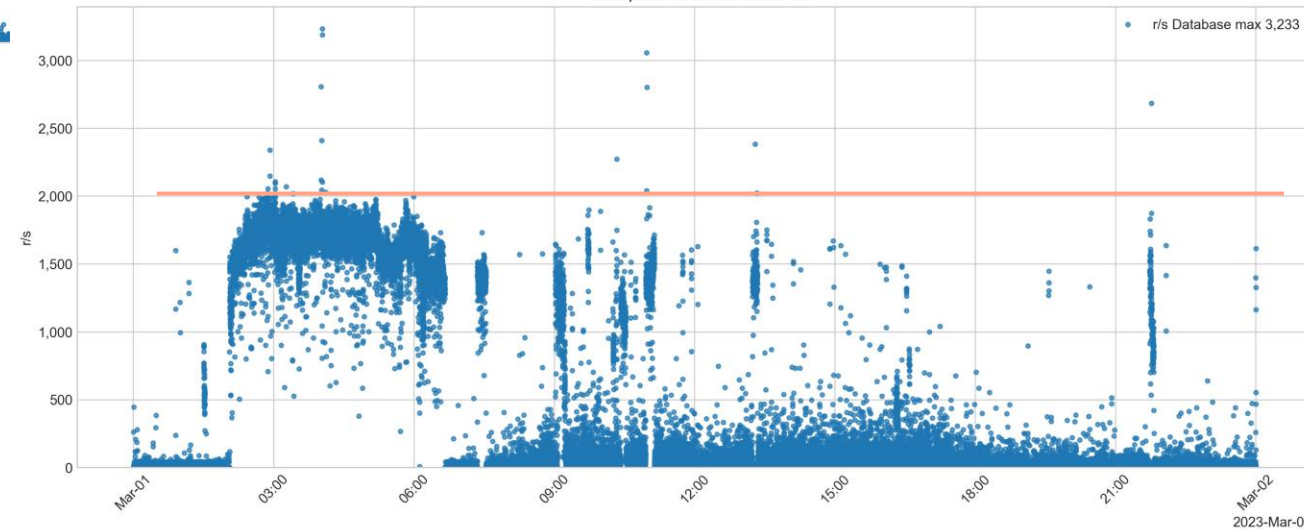
ビフォー(オンプレミス) – 4,000リードIOPS

Wed, Jan 04, 2023 00:00:24 - Thu, Jan 05, 2023 00:00:31
Example - iostat Database r/s



アフター(クラウド) – 2,000リードIOPS

Wed, Mar 01, 2023 00:00:26 - Thu, Mar 02, 2023 00:00:25
Example - iostat Database r/s

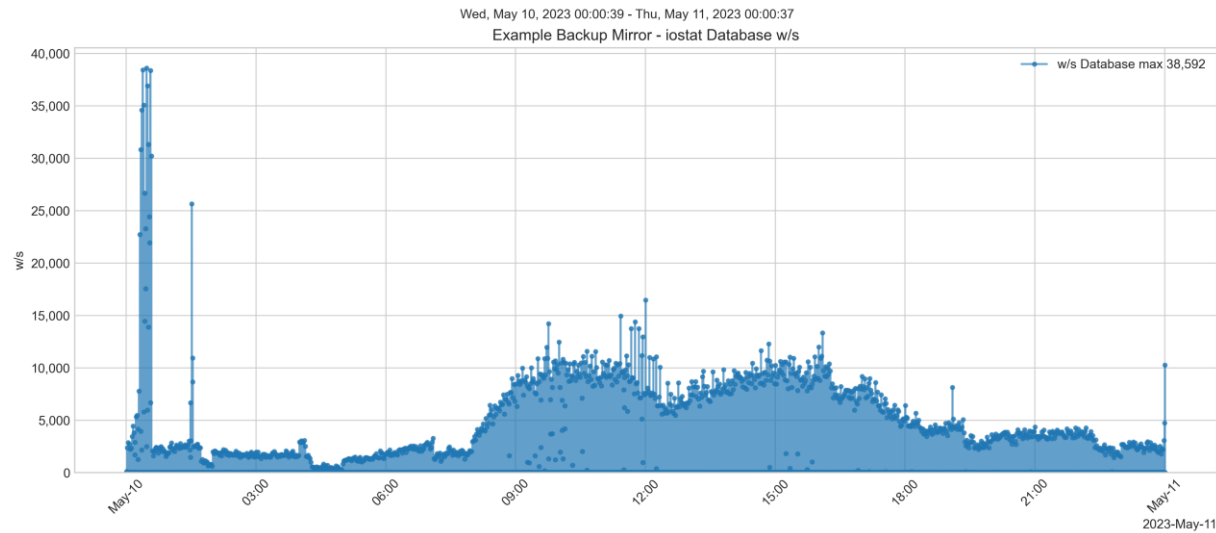


2023-Mar-02

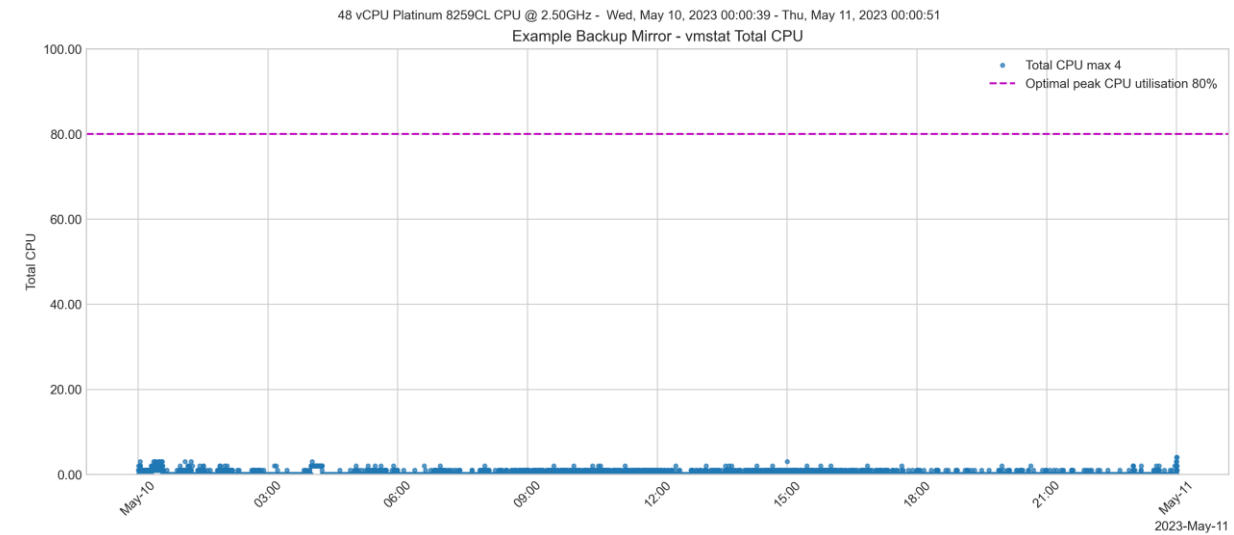
オンプレミスからクラウドへの移行例 - ミラー



ミラー - 同様のライトIOPS



低CPU使用率

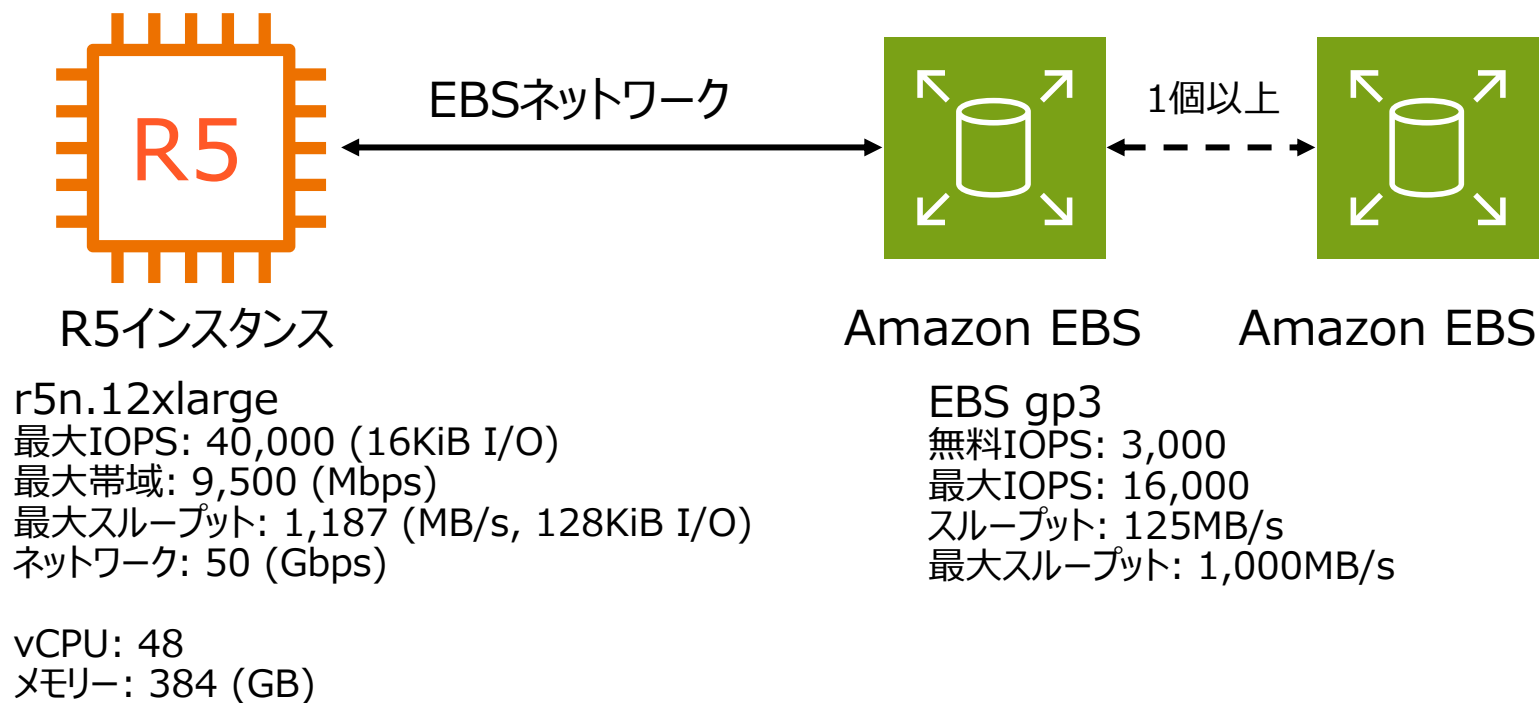


デプロイメント - AWS



- インフラストラクチャリソース

- インスタンス
- ストレージ
- ネットワーク



	データ gp3	WIJ gp3	ジャーナル gp3
IOPS	4 x 10,000 = 40,000	1	1
スループット (MB/s)	4 x 250 = 1 GB/s	450	

A vertical white line with small circular dots at both ends, positioned on the left side of the slide.

Summary



まとめ

- クラウドのサイジングで大きな余裕は不要
- パフォーマンスとディスクの空き容量は常時監視を
- サイズダウンも可能
- 新世代のタイプはより高い費用対効果
- 支払額もレビュー

ありがとう
ございました