

# 次世代Dell PowerEdge サーバーで パフォーマンスを最適化するためのBIOS設定

## 著者:

Donald Russell  
Diego Esteves  
Waseem Raja

## 概要

Dell PowerEdgeサーバーには、お客様が最高のパフォーマンスを実現できるよう、さまざまなチューニング可能なパラメータが用意されています。このホワイトペーパーでは、最新世代のPowerEdgeサーバー（PowerEdge R660、R760、MX760c、C6620など）で使えるチューニング可能なパラメータの概要を説明し、さまざまなワークロード向けに推奨される設定を情報提供いたします。



図1. PowerEdge R660



図2. PowerEdge R760

次の表に、最新世代PowerEdgeサーバーのBIOS設定に関する推奨事項を示します。

表1. BIOS設定の推奨事項 - システム プロファイル設定

システム設定画面	設定	デフォルト	HPCおよびSPECcfp スピード重視の環境 向けパフォーマンスの ための推奨設定	低レイテンシー、 STREAM、および MLC環境で推奨される 設定	汎用ビジネス/ 科学 技術スループット向け (例: SPECcpu2017) 推奨構成
System profile settings	System Profile	Performance Per Watt [1]	Performance Optimized	First select Performance Optimized and then select Custom [1]	Custom
System profile settings	CPU Power Management	System DBPM	Maximum Performance	Maximum Performance	Maximum Performance
System profile settings	Memory Frequency	Maximum Performance	Maximum Performance	Maximum Performance	Maximum Performance
System profile settings	Turbo Boost [2]	Enabled	Enabled	Enabled	Enabled
System profile settings	C1E	Enabled	Disabled	Disabled	Disabled
System profile settings	C States	Enabled	Disabled	Disabled	Autonomous or Disabled [6]
System profile settings	Monitor/Mwait	Enabled	Enabled	Disabled [3]	Enabled
System profile settings	Memory Patrol Scrub	Standard	Standard [4]	Standard/Disabled [4]	Disabled
System profile settings	Memory Refresh Rate	1x	1x	1x	1x
System profile settings	Uncore Frequency	Dynamic	Maximum [5]	Maximum [5]	Dynamic
System profile settings	Energy Efficient Policy	Balanced Performance	Performance	Performance	Performance
System profile settings	CPU Interconnect Bus Link Power Management	Enabled	Disabled	Disabled	Disabled
System profile settings	PCI ASPM L1 Link Power Management	Enabled	Disabled	Disabled	Disabled

[1] システムのオーダー時の設定により異なる場合もあります。システムプロファイルのその他のデフォルト値は、この設定に依存します。まずパフォーマンスのプロファイルを選択し、そのうえで「カスタム」を選択し最適なプロファイルをロードして、さながら変更ができます。

[2] インテル Speed Select テクノロジー (インテル SST)とインテル ターボ・ブースト・テクノロジーは前世代から飛躍的に向上していますが、一部のターボ効果はすべてのワークロード条件で効果が保証されているわけではありません。お客様の環境でターボ・ブースト・テクノロジーを検証いただき、実際のワークロードで最も効果的な設定をお選びください。また、並行してDell Controlled Turboの活用もご検討ください。

[3] Monitor/Mwait命令の無効化は、論理プロセッサを無効化する際のみ使用ください。この設定は、Linuxのintel\_idleドライバーによるC-stateの強制を阻止します。

[4] メモリー巡回スクラブの無効化に効果があるかどうか、お客様自身の環境でご確認いただけます。

[5] 動的モードでは、動的なUncore周波数の制限と引き換えにTDPに追加のヘッドルームをもたらします。最適な設定はワークロードに依存します。

[6] 空冷システムではAutonomous、Direct Liquid CoolingシステムではDisabledとなります。

図2. BIOS設定の推奨事項 - メモリ、プロセッサ、iDRAC設定

システム設定画面	設定	デフォルト	HPCおよびSPECcfp スピード重視の環境 向けパフォーマンスの ための推奨設定	低レイテンシー、 STREAM、および MLC環境で推奨される 設定	汎用ビジネス/ 科学 技術スループット向け (例: SPECcpu2017) 推奨構成
Memory settings	Memory Operating Mode	Optimizer	Optimizer [1]	Optimizer [1]	Optimizer [1]
Memory settings	Memory Node Interleave	Disabled	Disabled	Disabled	Disabled
Memory settings	DIMM Self Healing	Enabled	Disabled	Disabled	Disabled
Memory settings	ADDDC setting	Disabled [2]	Disabled [2]	Disabled [2]	Disabled [2]
Memory settings	Memory Training	Fast	Fast	Fast	Fast
Memory settings	Correctable Error Logging	Enabled	Disabled	Disabled	Disabled
Processor settings	Logical Processor	Enabled	Disabled [3]	Disabled [3]	Enabled
Processor settings	Virtualization Technology	Enabled	Disabled	Disabled	Disabled
Processor settings	CPU Interconnect Speed	Maximum Data Rate	Maximum Data Rate	Maximum Data Rate	Maximum Data Rate
Processor settings	Adjacent Cache Line Prefetch	Enabled	Enabled	Enabled	Enabled
Processor settings	Hardware Prefetcher	Enabled	Enabled	Enabled	Enabled
Processor settings	DCU Streamer Prefetcher	Enabled	Enabled	Disabled	Disabled
Processor settings	DCU IP Prefetcher	Enabled	Enabled	Enabled	Enabled
Processor settings	Sub NUMA Cluster	Disabled	SNC 2	SNC 4 on XCC SNC 2 on MCC	SNC 4 on XCC SNC 2 on MCC
Processor settings	UPI Prefetch	Enabled	Enabled	Enabled	Enabled
Processor settings	Dell Controlled Turbo	Disabled	Disabled	Enabled [4]	Disabled
Processor settings	Dell Controlled Turbo Optimizer mode	Disabled	Enabled [5]	Enabled [5]	Enabled [5]
Processor settings	XPT Prefetch	Enabled	Disabled	Disabled	Enabled
Processor settings	UPI Prefetch	Enabled	Disabled	Disabled	Enabled
Processor settings	LLC Prefetch	Disabled	Enabled	Disabled	Disabled
Processor settings	DeadLine LLC Alloc	Enabled	Enabled	Enabled	Disabled
Processor settings	Directory AtoS	Disabled	Disabled	Disabled	Disabled
Processor settings	Dynamic SST Perf Profile	Disabled	Disabled	Enabled	Disabled

システム設定画面	設定	デフォルト	HPCおよびSPECcfp スピード重視の環境 向けパフォーマンスの ための推奨設定	低レイテンシー、 STREAM、および MLC環境で推奨される 設定	汎用ビジネス/ 科学 技術スループット向け (例: SPECcpu2017) 推奨構成
Processor settings	SST-Perf- profile	Operating Point 1	Operating Point 1	Operating Point ? [6]	Operating Point 1
iDRAC settings	Thermal Profile	Default	Maximum Performance	Maximum Performance	Maximum Performance

[1] メモリ帯域幅重視の場合は最適化モードをご使用ください。Fault Resilient Mode（耐障害性モード）では帯域幅が最大で33%低下します。

[2] x4 DIMM搭載時のみ選択可能です。

[3] 論理プロセッサ（ハイパースレッディング）はSPEC CPU2017 INT\_RateやFP\_Rateなどのスループット指向のワークロードでより性能を発揮する傾向にあります。多くのHPCワークロードでこのオプションは無効化されます。

[4] Dell Controlled Turboは最大コア数（全コア数）でのターボの際に各コアのクロック周波数を維持する働きをし、ジッターを抑制します。ターボブーストを無効化する場合はこの機能も無効化してください。

[5] 液体冷却システムでのみ提供される選択肢です。

[6] ソフトウェアプログラムがベース周波数とターボ周波数の影響を受けるどうかに依存します。影響がある場合は利用するコア数を減らしその分より高いベース周波数およびターボ周波数で動作します。

## iDRACの設定における推奨事項

- 熱の観点で要求の厳しい環境では、iDRACのサーマル設定で冷却ファンの回転速度を上げてください。
- パフォーマンスが重視される環境ではいかなるパワーキャッピング機能も解除しておくべきです。

## BIOS設定に関する用語集

- System Profile: (デフォルト=Performance Per Watt)** : 特定の環境に合わせて個々の電力/性能機能を設定するのは難しい場合があります。そのため、お客様が最小電力使用量/音響レベル、最大効率、Energy Star最適化、最大パフォーマンスなどについてシステムを最適化するのに役立つメニューオプションを提供しています。
- Performance Per Watt DAPC (Dell Advanced Power Control)** : このモードはDellのプリセットを使って消費電力あたりパフォーマンスを最大化します。より省電力の方向に向け作用します。最大バス速度は重要でないアプリケーションにおいて、消費電力を削減しつつパフォーマンスを向上するための最良の機能を提供します。これはSPECpowerテストに適したモードになると予想されます。「効率 - 電力を優先」モードは、サーバー用のEnergy Starがリリースされる前に造られたプリセット動作モードを持つシステムとの下位互換性を維持します。
- Performance Per Watt OS** : このモードは、よりパフォーマンス重視の方向に作用しながら消費電力あたりパフォーマンスを最適化します。これはEnergy Starに適したモードになります。このモードは「Performance Per Watt (DAPC)」とは少し異なるという点にご注意ください。
- Performance** : このモードは、電力に関係なくシステムの絶対的なパフォーマンスを最大化します。このモードでは、消費電力は考慮されません。消費電力の増加に加え、冷却ファンの回転速度やシステムの熱出力などが増加する可能性があります。このモードではシステムの効率が低下する可能性があります。実行中のワークロードによっては絶対性能が向上する可能性があります。
- Custom** : カスタムモードでは、他の4つのプリセットモードでプリセットされていて変更することはできない、低いレイヤーの設定もユーザーが個別に変更できます。
- C-States** : C-Stateにより、CPUのアイドル電力が減少します。このモードには3つのオプションがあります。

- **Enabled** : 「Enabled」を選択するとオペレーティングシステムがC-stateへの移行を開始します。一部のオペレーティングシステムはACPIマッピング (intel\_idleドライバーなど) を無効にする可能性があります。
  - **Autonomous** : 「Autonomous」を選択すると、HALTおよびC1要求はハードウェア内でC6 要求に変換されます。
  - **Disabled** : 「Disabled」を選択すると、オペレーティングシステムでC0とC1のみが使用されます。C1は、OSが自動停止すると自動的に有効になります。
- **C1 Enhanced Mode** : C1E (C1 Enhanced) ステートを有効にすると、アイドル状態のCPUコアを停止させることで電力消費の抑制が可能になります。
  - **Turbo Mode** : ターボモードを有効にしておくと、CPUの全てのコアがフル活用されていない時に、全体的なCPUのパフォーマンスがブートされます。ターボモードに入ったCPUのコアは一時的にですがカタログ値よりも高い周波数で動作させることができます。
  - **Hyper-Threading** : ハイパースレッディングを有効にすると、オペレーティング システムは、物理コアあたり2つの仮想コアまたは論理コアを扱えるようになります。ワークロードは、可能な場合は仮想コアまたは論理コア間で共有できます。ハイパースレッディングの主な機能は、プロセッサリソースをより効率的に使用するために、パイプライン内の独立した命令の数を増やすことです。
  - **Execute Disable Bit** : Execute Disable Bit (無効化ビット実行) を使用すると、この機能をサポートするオペレーティングシステムで使用する場合、メモリを実行可能または非実行可能としてマークできます。これにより、コードが実行不可能なメモリで実行しようとしたときにオペレーティング システムにエラーを発生させるようにプロセッサを構成することで、システムのセキュリティを向上させることができます。
  - **DCA (Direct Cache Access)** : ネットワーク コントローラーなど、DCAに対応したI/Oデバイスはデータを直接CPUキャッシュに置けるため、応答時間を時間を改善します。
  - **Power/Performance Bias** : Power/Performance Biasは、CPUがどの程度積極的に電力管理され、ターボ状態に入るかを決定します。「プラットフォーム制御」では、システムが設定を制御します。「OS制御」を選択すると、オペレーティングシステムが制御できます。
  - **Per Core P-state** : コアごとのP-stateが有効になっている場合、各物理CPUコアは個別の周波数で動作できます。無効にすると、パッケージ内のすべてのコアは、すべてのアクティブなスレッドの中で最も解決された頻度で動作します。
  - **CPU Frequency Limits** : ターボとしての最大周波数と搭載CPUの定格周波数の間の任意の周波数に最大のターボ周波数を制限できます。
  - **Energy Energy Efficient Turbo** : 省エネルギー ターボが有効になっている場合、CPUの最適なターボ周波数はCPU使用率に基づいて動的に調整されます。
  - **Uncore Frequency Scaling** : 有効にすると、CPUアンコアはワークロードに基づいて速度を動的に変更します。
  - **MONITOR/MWAIT** : MONITOR/MWAITは、C-Stateを使用するために使用されます。
  - **Sub-NUMA Cluster (SNC)** : SNCは、アドレス範囲に基づいてLLC (Last Level Cache = ラストレベルキャッシュ) を分離クラスタに分割し、各クラスタをシステム内のメモリ コントローラのサブセットにバインドします。SNCは、LLCとメモリの平均遅延を改善します。SNCは、以前のプロセッサ製品ファミリーにあったCluster On Die (COD) に代わる機能です。マルチソケット システムの場合、すべてのSNCクラスタは一意的なNUMAドメインにマッピングされます。(IMCインターリーブも参照してください)。このBIOSオプションの値は以下です。
    - **Disabled**: このオプションが無効になっている場合、LLCは1つのクラスタとして扱われます。
    - **Enabled**: LLC容量をより効率的に使用し、コア/IMCの近接性によるレイテンシーを抑制します。これにより、NUMA対応オペレーティング システムのパフォーマンスが向上する可能性があります。

- **Snoop Preference** : ワークロードに基づいて適切なスヌープモードを選択します。スヌープモードには2種類あります。
  - **HS w. Directory + OSB + HitME cache**: ほとんどのワークロードにとって最適な設定です。(デフォルト設定はこちらです。)
  - **Home Snoop**: 帯域幅 (BW) 重視のワークロードで最適な設定です。
- **XPT Prefetcher** : XPT プリフェッチャーは、LLC (Last Level Cache = ラストレベルキャッシュ) に送信される読み取り要求が、その読み取りのコピーをメモリー コントローラー プリフェッチャーに投機的に発行できるようにするメカニズムです。
- **UPI Prefetcher** : UPIプリフェッチャーはDDRバス上でメモリの読み取りを早期に開始するメカニズムです。UPI受信パスは、メモリコントローラプリフェッチャーへのメモリ読み取りを生成します。
- **Patrol Scrub** : Patrol scrub は、すべての DIMMに対してバックグラウンド メモリ スクラブを実行する、メモリRAS機能です。この機能は、パフォーマンスに悪影響を及ぼす可能性があります。
- **DCU Streamer Prefetcher** : データ キャッシュ ユニット (DCU) ストリーマ プリフェッチャーを有効または無効にします。このオプションは、デフォルトで有効に設定されています。
- **LLC Dead Line Allocation** : 一部の インテルCPUキャッシュ スキームでは、Last Level Cache (LLC = ラストレベルキャッシュ) にMid-Level Cache (MLC = ミッドレベルキャッシュ) の削除が埋められます。特定の行がMLCから削除されLLCに移された場合、CPUコアは削除されたMLCの行に "dead" のフラグを立てることができます。これは、同じ行が再度読み取られる可能性が低いことを意味します。このオプションを使用すれば、dead となった行はキャッシュから落とされるため、このオプションを無効にしても、これがLLCを満たすことはありません。このBIOSオプションの値は以下です。
  - **Disabled**: このオプションを無効にすると、MLCのデッドラインがLLCに埋められないため、LLCのスペースを節約できます。
  - **Enabled**: LLCのスペースが利用可能な場合は、MLCのデッドラインを日和見的にLLCに埋めます。
- **Adjacent Cache Prefetch** : スレッドの少ないアプリケーションと一部のベンチマークでは、隣接するキャッシュ ラインのプリフェッチを有効にすることでメリットが得られます。デフォルトは [Enabled] です。
- **Intel Virtualization Technology** : インテル バーチャライゼーション・テクノロジーを使用すると、プラットフォームは複数のオペレーティング システムとアプリケーションを独立したパーティションで実行できるため、1 つのコンピューター・システムを複数の仮想システムとして機能させることができます。デフォルトは [Enabled] です。
- **Hardware Prefetcher** : スレッドの軽いアプリケーションと一部のベンチマークでは、Hardware Prefetcherを有効にすることでメリットが得られます。デフォルトは [Enabled] です。
- **Trusted Execution Technology** : インテルTrusted Execution Technology (インテルTXT) を有効にします。デフォルトは [Disabled] です。



For more info,  
visit the [Servers  
Info Hub](#)



[Contact us](#) for  
feedback and  
requests



Follow us for  
PowerEdge  
news