エントリー高性能IAサーバ MAGNIA$_{TM}$ 3300

MAGNIA$_{TM}$ 3300 Entry Class High-Performance IA Server

梶原 茂弘 ■KAJIHARA Shigehiro　渡壁 健 ■WATAKABE Takeshi　内藤 賢一 ■NAITO Kenichi

Intel®(注1)の新しいサーバ用プロセッサXeon™(注2)を用いたエントリークラス2Way高性能IA(Intel Architecture)サーバ"MAGNIA$_{TM}$ 3300"を開発した。MAGNIA$_{TM}$ 3300では床置き，ラックマウントのどちらにも使えるコンパクトな筐体(きょうたい)の中に高性能，高信頼性，高拡張性を凝縮する必要がある。そのための熱対策を冷却シミュレーションと実験によりモデルの精度を向上させながら実施した。更に，ホットプラグ対応RAID(Redundant Arrays of Independent(Inexpensive) Disks)や独自のサーバ監視機能を搭載するなど高い信頼性を確保し，システムのセットアップや運用管理ソフトウェアの提供により，サーバに求められる機能も十分に備えている。MAGNIA$_{TM}$ 3300は，多様なニーズに応えるエントリークラスのサーバである。

Toshiba has developed a high-performance Xeon™ 2Way compact Intel® architecture (IA) server, the MAGNIA$_{TM}$ 3300. This server has the three major features of high performance, high reliability, and high expandability, and comes in chassis for either pedestal or rack mounting.

To achieve high performance and high expandability, cooling simulation was very useful for chassis design. The simulation model was improved by repeated experiments. The server also achieves high reliability by hot swappable RAID and an original server management method, as well as a low cost of ownership for setup and daily operation. These features make the MAGNIA$_{TM}$ 3300 an ideal solution as an entry class server.

1 まえがき

IAサーバ市場は成長を持続しており，特にボリュームの中心を占めるエントリーサーバは，国内でも2002年度後半で20％の伸びが予想されている。東芝は，このクラスとしてMAGNIA$_{TM}$3300を2002年9月に発表し，出荷を開始した(**図1**，**表1**)。この装置は，従来のエントリークラスに比べて次に挙げる特長を付加し，多様なニーズに対応した製品となっている。

図1．MAGNIA$_{TM}$ 3300 ― 机の下に設置可能なタワータイプと19型ラックに実装可能なラックタイプの2種類がある。

MAGNIA$_{TM}$ 3300 pedestal and rack mounting chassis

表1．MAGNIA$_{TM}$ 3300の主な仕様
Specifications of MAGNIA$_{TM}$ 3300

項　目	仕　様
プロセッサ	Xeon™1.8 GHz/ 2.2 GHz/ 2.4 GHz/ 2.8 GHz
プロセッサ数	1～2
メモリ(最大)	6 Gバイト(最小128 Mバイト) [DDR 200 MHz] ECC付
HDD(最大)	本体内蔵：876 Gバイト(10 k/15 krpm)
ディスクベイ	SCSI 18 Gバイト／36 Gバイト／ 73 Gバイト／146 Gバイト×6台
I/Oスロット	6スロット(PCI) (2×64ビット/100 MHz，2×64ビット/66 MHz， 2×32ビット/33 MHz PCI)
ネットワーク	10/100BASE-TX Ethernet(注3)×1 10/100/1000BASE-TX Ethernet×1
SCSI	2(Ultra160 SCSI)
添付ソフトウェア	サーバ設定支援ソフトウェア：SetupInstructor サーバ監視ソフトウェア：HarnessEye/web$_{TM}$
サポートOS	Windows® 2000(注4) Advanced Server，Linux

ECC：Error Correct Code

(1) 高性能　Intel® Xeon™プロセッサ(512 Kバイト L2キャッシュ)最大2個搭載，DDR-SDRAM(Double Data

(注1)，(注2)　Intel，Xeonは，米国Intel Corporationの商標。
(注3)　Ethernetは，日本における富士ゼロックス社の登録商標。
(注4)　Windowsは，米国Microsoft Corporationの米国及びその他の国における登録商標。

Rate Synchronous DRAM)，Ultra320-SCSI（Small Computer System Interface）ハードディスクドライブ（HDD），Gビット Ethernetなど最先端のコンポーネントを搭載

(2) 高拡張性　コンパクトな床置筐体の中にHDD，バックアップ装置，拡張インタフェースを豊富に内蔵し，多様な運用形態に応じた拡張を筐体内で実現し，お客さまの各種ニーズに応えることが可能

(3) 高信頼性　独自のサーバ監視制御機構の実装，ホットプラグRAID，冗長電源，冗長ファンや2ポートLANなどによる高信頼性を実現

(4) 低運用コスト　セットアップからシステムの運用管理まで，システム構築が簡単に設定できる環境を提供

ここでは上記の特長化技術のほか，高性能化と高拡張性を実現するうえでハードルであった熱対策について述べる。

2 高信頼・高拡張性 2Way コンパクト高性能サーバ

MAGNIA™ 3300は，机の下に設置が可能な“タワータイプ”と，19型ラックに実装可能な高さ5U（1Uは1.75インチ＝44.45 mm）の“ラックタイプ”の2種類が用意されている。

筐体は213 mm（幅）×655 mm（奥行き）×465 mm（高さ）の大きさであり，この中にサーバとしての先端技術を凝縮してある。

(1) Intel® Xeon™ プロセッサを最大2個搭載可能

(a) ハイパー スレッディング テクノロジー

(b) Intel® NetBurst™(注5) マイクロアーキテクチャ

(2) 200 MHz DDRメモリ（最大6Gバイト）を採用　1Gバイト以上の大容量メモリユニットにおいては，メモリチップ故障に対する可用性を強化するChipkill(注6)技術を採用

(3) Ultra320対応SCSI HDDを採用（146 GバイトHDD最大6台搭載）

(4) 高速，高信頼性RAIDコントローラを採用　Ultra 320対応RAIDコントローラ，又は当社独自のRAID高速化技術を搭載した高信頼のRAIDコントローラ（U160対応）を選択可能

(5) Gビット Ethernetコントローラを標準で搭載（合計2ポート標準装備）

更にMAGNIA™ 3300は5インチベイが3個，ホットプラグ可能なHDDベイが6個，PCI（Peripheral Component Interconnect）スロットを6本備えている。この構成によりRAIDの自由な構築，オートローダ付バックアップ装置の装着など，単一の筐体で多様なニーズに応え，幅広いシステム構成を提供することができる。

（注5） NetBurstは，米国Intel Corporationの商標。
（注6） Chipkillは，米国IBM Corporationの商標。

3 高効率冷却設計

高性能CPUのXeon™を2個搭載，更に15 krpmの高速回転HDDが6台，その他I/O（Input/Output）ベイのストレージなど消費電力が大きいコンポーネントが内蔵され，これを縦置き，横置きのそれぞれで問題なく冷却されなければならない。

加えて，このシステムは冗長性による高可用性が特長である。通常のパソコン（PC）ではCPUの冷却にはアクティブヒートシンク（ファン付きヒートシンク）が一般的だが，これではファンが故障するとそのCPUを停止しなければならず高可用性を維持できない。

また，CPUそのものも従来の約2.5倍の発熱があり，単純に強力なファンにするだけでは解決しない。また，このクラスは小規模のオフィスに置かれる場合も多いことから低騒音であることも求められる。

そこで，**図2**に示すように，シミュレーションと実験機による評価を組み合わせながら効率の良い冷却を実現した。更に発熱に応じた最適なファン回転制御を盛り込み，オフィスにおける実使用環境での低騒音を実現した。

(1) 冷却構造の検討　基本的な冷却構造を検討する。

(a) メインファン（冗長構成）で電源を除く筐体内すべて

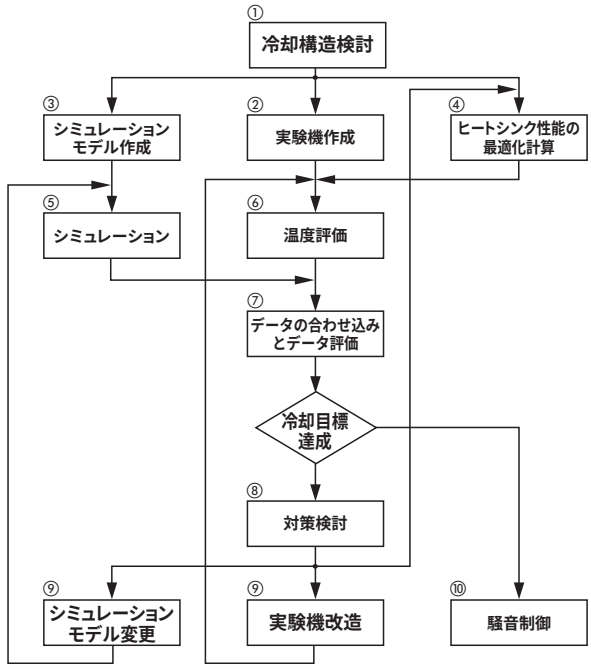

図2．冷却のための開発フロー ― シミュレーションと実験機の併用による冷却設計の流れを示す。
Flow of cooling development process

の冷却を行うものとし，CPUはエアダクトを用いて冷却する。

(b) ただし，PCIカード部は風が流れないため筐体内の風を吹きつける小型のファン(冗長構成)を付ける。

(2) 実験機作成　検討した冷却構造の実験機を作成する。

(3) シミュレーションモデル作成　実験機と同様のモデルを作成する。

(4) ヒートシンク性能最適化計算　冷却構造及び使用するファンの特性からヒートシンクの最適構造(フィンの厚さ，間隔，高さ，材質)を計算しヒートシンクを作成する。

(5) シミュレーション　モデルに発熱体のデータを与えシミュレーションを実施する。

(6) 温度評価　作成した実験機に発熱体を実装し温度評価を実施する(CPUはヒータチップを使用)。

(7) データの合わせ込みとデータを評価する。

(a) シミュレーションデータと実験機の測定データの合わせ込みを行い，モデル及び発熱体データの微調整を行う。

(b) シミュレーションデータと実験機の測定データを評価し，目標の達成度を判断する。

(8) 対策検討　冷却目標が達成できていなければ，冷却構造の見直しを検討する。

(9) 実験機改造　検討した構造を基に実験機の改造を行う。

以上の開発フローを数回まわすことにより，発熱が最大となる動作モードすなわち最高の能力を発揮する場合でも冷却が可能となった(**図3**，**図4**)。

(10) 騒音制御　一般的に通常の使用では負荷は動的に変動し，通常は負荷が低い時間が多い。通常状態でできる限り回転を抑えるために，メインファンはCPU以外の冷却が動作状態で保証できる回転数を最低とし，CPUの温度監視により4段階の回転制御を行った。

以上の対策を実行することで，従来機より大幅に性能が向上したにもかかわらず，従来機とほぼ同等の騒音レベルを実現した。

4 サーバ監視制御機構

MAGNIA™3300は，専用に新規開発したBMC(Baseboard Management Controller)と，後述するサーバ管理ソフトウェア HarnessEye/web™との連携でサーバ監視・制御機能を実現し，運用の利便性，可用性，保守性を向上させた(**図5**)。また，AMS(Availabillity Management Subsystem)カードと組み合わせることで，リモートサーバ管理を可能とする。

以下に，主要な機能を挙げる。

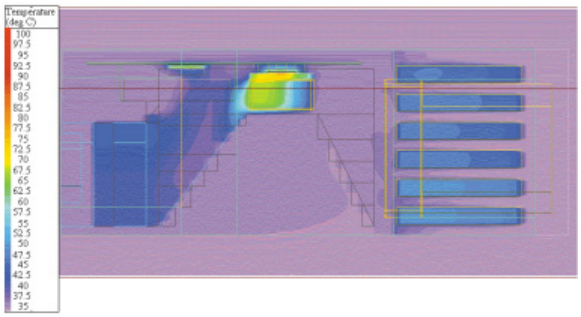

図3．当初の温度分布 ― 筐体を上から見た図で，中央上がCPU，右側にHDDが6台ある。CPUから左下に熱がダクト内部を拡散しているが，流れが悪くCPUの冷却が不十分で過熱した状態の温度分布を示す。

Thermal analysis of original status

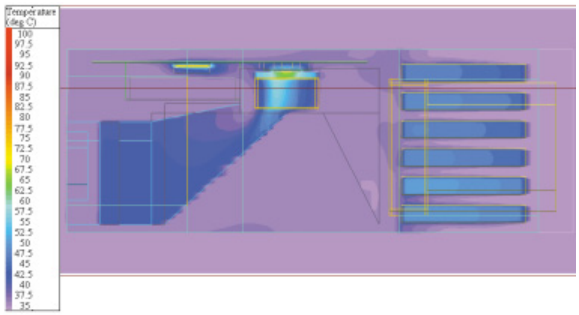

図4．冷却設計後の温度分布 ― ヒートシンクの改善，ダクト形状の改善，筐体内部の空気の流れの最適化により改善された温度分布を示す。ダクト内部の熱の拡散が改善されCPUの温度が下がっている。

Thermal analysis after cooling design

(1) 基本サーバ監視機能

(2) 従来より改良したファン制御

(3) Chipkill機能による最大4ビットまでのメモリエラー補正

(4) マザーボード上の電圧変換回路監視

(5) ハードウェア制御のCPUスロットリング機能

4.1 基本サーバ監視機能

MAGNIA™3300は，サーバ監視機能として，最大で9種類の電圧監視，4個のファン回転数監視，2個のCPU温度監視，2個の電源モジュール監視，後述する電圧変換回路監視をサポートする。

これら監視対象の故障を含め，システムダウンの要因となるハードウェア異常や，電源on/off などシステムにとって重要な状態の変化は，BMCが管理するSEL(System Event Log)に記録する。BMCは，かりに基本ソフトウェア(OS)がダウンしても独立して動作できるため，ログデータは消滅しない。SELによりシステム管理者はハードウェア クラッシュの原因をすばやく突き止められ，短時間でシステムを復旧できる。SELにはBMC単体で最大1,000件記録可能とした。SEL情報は随時HarnessEye/web™で収集・保存し，長期

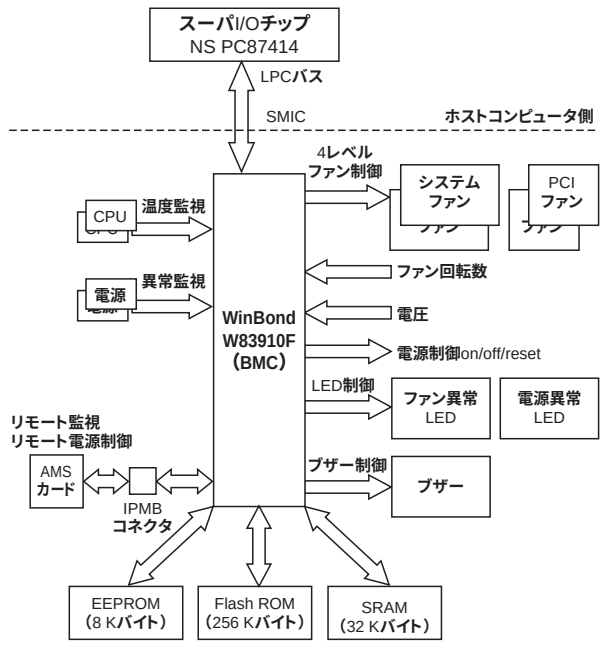

図5．BMCを中心とするサーバ監視・制御機構 ― MAGNIA™ 3300の主なシステム監視・制御の概要を示す。

Server management of MAGNIA™ 3300

間のデータ収集や，リモート端末から参照が可能である。

4.2　従来より改良したファン制御

MAGNIA™3300のファン構成は，主冷却用の大型ファンと補助用の小型ファンから成る。それぞれ1個ずつの基本構成か，2個ずつの冗長構成を取りうる。BMCは，基本構成側のファン速度をCPU温度によって4段階制御する。従来機種よりもいっそう制御方法を改善し，冷却効率向上，騒音抑制を図っている。

従来機種ではすべてのファンを同様に制御していたが，今回は，大型ファンと小型ファンを別々のしきい値により異なる速さに制御できるよう設計し，最終的に小型ファンは回転数を一定に保つこととした。

また，従来機種では待機中の冗長側ファンは停止させていたが，今回，待機中は低速で回転させておく仕組みを採用した。騒音抑制，冷却効率向上，ファン故障防止に効果がある。

更に，従来機種では1個のファンが故障した際にシステム内の全ファンを冗長側に切り替えるのに対し，今回は，故障した基本ファンだけを冗長ファンに切り替えることでシステム可用性を向上させた。

4.3　Chipkill機能によるメモリエラー補正

MAGNIA™3300は，メモリエラー補正機能が強化されている。従来の1ビット補正に比べ，Chipkill機能の採用により最大4ビットまでのエラーを補正し稼働し続けることを可能とした。エラー発生の際は，補正可能だったか，どのメモリで発生したかの情報をSELに残す。これにより，交換すべきメモリを特定できる。もし固定的なメモリ故障に至った場合は，ブート時にメモリ縮退し，残りのメモリで可能な限りシステムを稼働させようと試みる。

4.4　マザーボード上の電圧変換回路監視

MAGNIA™3300は，システムとして従来よりも多種の電圧を必要としており，それらはマザーボード上の電圧変換回路により作成される。そのため従来どおりの電圧監視に加え，電源系の監視を強化する目的で，マザーボード上の電圧変換回路の故障監視も行う。この監視により，BMCが故障を検出すると電源を自動的にオフにして，システムを保護する仕組みを採用した。

4.5　ハードウェア制御のCPUスロットリング機能

CPUは高速，高消費電力化の一途をたどっており，単位時間当たりの温度上昇率が増大している。このため今後の高速CPUでは，ファン制御による冷却がCPUの急激な温度上昇に追いつけない可能性が出てきた。そこで今回は，ハードウェアによるCPUスロットリング機能を採用した。CPUがあまりに急激に高温になった場合に，自動的にCPUのクロックを間引き，発熱を抑える技術である。ハードウェアによる自動制御のため，迅速に最小限の期間のみ作用させることができる。

5　あとがき

MAGNIA™3300は高性能，高信頼性，高拡張性を備え，将来も安心してお客さまにご使用いただけるエントリークラスサーバである。今後とも，時々刻々と変化する市場ニーズに応える製品開発に注力していく。

梶原　茂弘　KAJIHARA Shigehiro

デジタルメディアネットワーク社 デジタルメディアデベロップメントセンター PCサーバ設計部グループ長。コンピュータの高信頼化，高性能化の研究・開発に従事。電子情報通信学会会員。

Digital Media Development Center

渡壁　健　WATAKABE Takeshi

デジタルメディアネットワーク社 デジタルメディアデベロップメントセンター PCサーバ設計部主務。サーバコンピュータの高信頼化ソフトウェア開発に従事。情報処理学会会員。

Digital Media Development Center

内藤　賢一　NAITO Kenichi

デジタルメディアネットワーク社 デジタルメディアデベロップメントセンター ソフトウェア第二部主務。サーバソフトウェアの開発に従事。情報処理学会会員。

Digital Media Development Center