

Harmonious Cloud
Hitachi vRAMcloud® Series

vRAMcloud®のご紹介

コンセプトペーパー

2012 年 4 月 発行
Ver. 1.0
株式会社 日立製作所

＜本書で使用する用語＞

ATM : Automatic Teller Machine
BA : Business Analytics
BCP : Business Continuity Plan
BI : Business Intelligence
CEP : Complex Event Processing
CMS : Cash Management System
CQL : Continuous Query Language
CRM : Customer Relationship Management
DBMS: DataBase Management System
DR : Disaster Recovery
DWH: Data WareHouse
GPS : Global Positioning System
POC : Proof Of Concept
POS : Point Of Sales
RAM : Random Access Memory
RDB : Relational DataBase
RFID: Radio Frequency Identification
SCM : Supply Chain Management
SFA : Sales Force Automation
SNS : Social Networking Service
SQL : Structured Query Language
TCO : Total Cost of Ownership

＜留意事項＞

本ホワイトペーパーで記載する数値は、日立環境での計測データであり、諸条件により結果は異なります。

＜他社商標注記＞

- GemFire は、VMware,Inc.の米国およびその他の国における商標または登録商標です。日立製作所がソフトウェアを提供し、サポートを実施している製品です。
- その他記載されている会社名、製品名は、それぞれの会社の商標または登録商標です。

本資料は、(株)日立製作所 情報・通信システム社 金融システム事業部が提供する vRAMcloud[®](*)のコンセプトをまとめたものです。さまざまな企業、組織において情報システムの企画・計画に携わる方々を読者として想定しています。

* クラウド環境でビッグデータの利活用に最適な先端テクノロジーを提供するソリューション群。

概要

本コンセプトペーパーは、ビッグデータを扱うために最適な、先端テクノロジーを集結させた vRAMcloud[®]のコンセプトをまとめたものです。

企業が扱う情報は、年々増加の一途をたどり、まさに「ビッグデータ時代」が到来しています。しかし、ビッグデータの内容は“玉石混交”のため、その中から本当に有効なデータを適切かつリアルタイムに把握できるかどうか、企業の事業オペレーション最適化や競争力強化に向けた重要な課題となっています。

日立が提供する vRAMcloud[®]は、膨大なビッグデータを高速に処理することで、今まで活用できず捨てていた情報や、分析できなかった情報を有効に利活用できるソリューション群です。

vRAMcloud[®]は、複数の業務アプリケーションで使用するデータをメモリ上で仮想的に共有する「分散キャッシュ」技術を実装した GemFire を中心に、「バッチジョブ超並列処理(グリッドバッチ)」、「ストリームデータ処理(CEP:複合イベント処理)」、「Hadoop」、分散キャッシュを最適にコントロールする「クラウド管理」、それらの製品群をつなぐ「フレームワーク」から構成されます。

(株)日立製作所 情報・通信システム社 金融システム事業部は、本ソリューションにより、お客様のビッグデータへの取り組みを強力に支援させていただきます。

なお分冊では、実証実験(POC:Proof Of Concept)の結果についてまとめておりますので、そちらもご参照ください。

Contents

「ビッグデータ時代」の到来.....	1
技術進化とクラウドの普及が追い風に.....	1
ビッグデータ活用は既に始まっている.....	2
ビッグデータ活用のメリット.....	3
金融業界でのビッグデータ活用.....	3
流通業界でのビッグデータ活用.....	3
産業・交通・社会インフラ分野でのビッグデータ活用.....	4
幅広い分野に広がる可能性.....	4
ビッグデータに対応する IT 基盤の条件.....	5
日 立のビッグデータ対応ソリューション「vRAMcloud®」.....	6
vRAMcloud®の目的.....	6
vRAMcloud®の構成.....	8
スケーラビリティ.....	11
BCP にも即対応.....	12
まとめ.....	13
今後の計画.....	13

「ビッグデータ時代」の到来

企業や組織で扱うデータ量は年々増加の一途をたどり、近年はまさに“情報爆発”と言われる時代となっています。企業で扱うデータは、業務データ、メール、画像・音声など、それぞれのファイル容量が増大しています。ここ数年は Facebook や Twitter に代表されるソーシャルメディアへの書き込み、RFID (Radio Frequency Identification) タグや各種センサーが時々刻々と生成するデータ、携帯電話・スマートフォンに組み込まれた GPS (Global Positioning System) から発生する位置情報なども氾濫しています。

これに伴い、企業が入手可能な情報も、文字列や数値などの「構造化データ」から、SNS (Social Networking Service) やメール、画像、センサーデータといった「非構造化データ」が主流となり、データ量もペタバイトからエクサバイトへと、数年前とは比べ物にならないほど膨大になっています。

このように、さまざまな拠点や端末から絶え間なく収集・蓄積されていく大量の情報、いわゆる「ビッグデータ」には、“ヒト”や“モノ”の行動やふるまいが多種多様なカタチで埋め込まれていると考えられます。特に SNS でリアルタイムに交わされるクチコミや、個人のブログから発信されるコメントなどは、消費者の商品購買や世論形成などに重要な影響を及ぼすことが広く認識されています。

つまりビッグデータとは、単純にデータ量の増加を意味しているのではなく、その多様性や発生頻度といった属性にも大きな特徴があり、分析方法や情報の組み合わせ方によって、従来なら専門家でさえ気づけなかった傾向や一定の法則性が浮き彫りとなり、革新的なサービスやビジネスモデルの実現につながるのではないかと期待されているのです。

そこで、この膨大な“情報の金脈”から、有益かつユニークな情報をスピーディーに抽出することで、サービス品質の向上や新ビジネスの創造、イノベーションの促進などにつなげていこうという動きが活発化してきました。これが現在、「クラウド」に続いてビジネスや産業、社会に大きなインパクトを与えるといわれる「ビッグデータ革命」です。

これまで IT の役割は、特定の目的で集められたデータを定型的に処理し、業務効率の向上を図ることに主眼が置かれていました。しかし今後、めまぐるしく変わる市場環境や顧客ニーズに追従しながら、グローバル規模で他社との差別化や競争優位性を確立していくためには、「ビッグデータ」と呼ばれるこれら膨大な情報の中から、いち早く新たな「知見」や「法則性」といった「価値」を見出し、効果的に活用していくことが何よりも強く求められているのです。

技術進化とクラウドの普及が追い風に

もちろんデータ量の巨大化とその戦略的な活用は、ここ数年で急に登場してきた話ではありません。金融業界やテレコム業界は以前から、膨大な取引情報や通信ログを事業戦略の立案や改善に活かしてきました。一般企業でも、日々の業務の中で多くのデータが生まれ続けており、製造業ならば部品の設計情報や生産現場のセンサーデータ、小売業なら店舗の POS データや、ポイントカードに関連付けられた購買履歴、E コマースの販売データなど、活用すれば必ず何か新しい発見や価値創造が期待できるデータは多岐にわたって存在しています。

このため「ビッグデータ」という概念が広く認識される以前から、IT を積極的に活用している企業では、データウェアハウス (DWH) やビジネスインテリジェンス (BI) といったシステムを活用し、新たな価値創出に挑戦してきました。

ただし、こうした取り組みには最新の IT アーキテクチャを駆使した環境整備や分析ツールの開発などに巨額のコストがかかるため、チャレンジへのハードルがかなり高かったのも事実です。

コスト面で折り合いがつかない場合は、大量に集まった貴重なデータを破棄するか死蔵せざるを得ず、活用するとしても一部のサマリデータに対して定期的にバッチ処理で集計・分析を行うという形が一般的だったはずです。

また、気象変動や災害予測、遺伝子情報などの解析では、スーパーコンピュータや特殊な分析ツールを適用しても、答えを導き出すまでに数週間から1か月という長い時間が必要でした。このため、新たな価値創出への欲求はありつつも、コストとデータの鮮度、期待効果との兼ね合いから、大量データの分析と活用へのチャレンジを見送る企業が少なくなかったのです。

いま改めてビッグデータの活用が脚光を浴びてきた背景には、CPU の高性能化やマルチコア化、ストレージやメモリ価格の低下など、ハードウェアの進化とコスト低減に加えて、膨大なデータを高速かつリーズナブルに処理できるソフトウェア技術が著しく進化してきたことが挙げられます。例えば、ビッグデータの分散処理を支えるオープンソースソフトウェアとして有名な「Hadoop」なら、これまで数週間かかっていたビッグデータの処理を 1 時間あるいは数十分といった短時間で行うことが可能となります。

こうした技術の適用が可能となれば、せっかく集めた宝の山(データ)をもう捨てる必要はありません。すべての生データを蓄積し、必要に応じて分析することで、他社への差別化や競争力の源泉に変えていくことが可能となるからです。

そして、これら膨大なデータの蓄積・処理・分析を効率的かつ低コスト、スケーラブルに実現できる「クラウド」の普及も追い風となり、幅広い企業や組織にビッグデータ活用が現実的なソリューションとして認知されるようになってきたのです。

ビッグデータ活用は既に始まっている

先進的な企業は、既にビッグデータをリアルなビジネスの成功に結びつけています。例えば Amazon.com が自社のクラウド基盤上に蓄積した膨大な Web ログを分析することで、ある商品を買った顧客に別の商品を薦めるレコメンデーションサービスを展開し、大きな成果を上げているのは有名な話です。

また、GPS 情報と連動した適切な場所でのクーポン券の発行や、膨大なクレジットカード利用パターンを分析した不正利用のリアルタイム検知、コミュニティサイトを構築しての新商品開発なども、すべてビッグデータ活用として現実に展開されているものです。

こうしてビッグデータと、高速な大量情報処理基盤を先行的に活用している企業は、それぞれ大きな先行者メリットを確保しています。

顧客からのニーズがますます多様化・高度化し、商品やサービスへの評価が急速なスピードで変化している今、他社との差別化を図り、確実に利益を確保していくためには、市場環境の変化を常に先取りし、いち早く次なるアクションへとつなげていくことが必要です。このため、既存の処理を大幅に高速化することで、顧客満足やサービス品質の向上をはかり、社内外に蓄積された“過去のデータ”だけでなく、SNS や各種センサー、GPS などから収集される“今のデータ”からも有益な知見を見出し、先手を打っていかうと、多くの企業がビッグデータ活用に本腰を入れ始めています。

ビッグデータ活用のメリット

では、ビッグデータの活用によって、どのような業務革新やイノベーションが可能となるのか、さまざまな業務分野別に、その可能性を探ってみましょう。

金融業界でのビッグデータ活用

従来から大量データを蓄積し、効果的にハンドリングすることが求められていた金融業界は、ビッグデータ活用で最も大きな導入効果が期待できる業界の1つと言えます。

例えば、ATM やクレジットカードなどの利用によって収集される大量の操作・利用ログをリアルタイムに分析すれば、不正利用を早期に検知し、被害の拡大を未然に防ぐことが可能となります。

また、株価や出来高の分析によって売買注文を自動化するアルゴリズムトレード、株・為替などリアルタイムに変動する大量データを高速配信するサービス、あるいは過去の取引履歴やリアルタイムなサイト訪問履歴から顧客に最適な金融商品をベストなタイミングでレコメンデーションするサービスなども容易に実現することができるはずです。

さらに、従来なら日次単位のバッチ処理で行うことが多かったポートフォリオの時価評価算出などのポジション管理もリアルタイムに実施すれば、より鮮度の高い情報に基づいた迅速な意思決定や状況把握が可能となり、さまざまなリスク低減やビジネスチャンスの開拓につなげることが可能となります。

こうしたユーザーの行動予測と市場動向をリアルタイムにとらえた個客指向マーケティングの実現、不正検知の即時化、新たな知見から導き出される先行的なビジネスモデルの構築などは、まさにビッグデータ時代ならではの金融ソリューションと言えます。

流通業界でのビッグデータ活用

Google や Amazon.com、Facebook などの先進企業がビッグデータによって緻密な個客指向マーケティングに成功していることは既に明らかです。複数チャネルにわたる個々の顧客の購買履歴や行動特性を詳細に分析することで、従来に比べ、より適切な商品のレコメンデーションを迅速に行うことが可能となったからです。

これと同様、コンビニエンスストアなど多くの小売業が活用する POS データや、交通系の IC カードなどから日々収集される膨大な行動・利用情報も、これまでのように 1 日以上前のデータから次の行動を予測するのではなく、その日のうちに素早く分析してアクションへつなげることで、今までにない新しい価値を生み出すことができます。

例えば、店舗の POS データを数時間のうちに分析すれば、店舗間でより効率的に商品を融通し合え、生鮮品のセールをいち早く実施することなどが可能となります。

携帯電話やスマートフォンに組み込まれた GPS からの情報を解析しながら、優良顧客に最寄り店舗でのセール情報をタイムリーに発信することなども簡単です。

Facebook や Twitter などで発信された顧客の声をスピーディーに分析することで、クレーム対応を迅速化し、素早い改善策やサービス提供によって、顧客からの信頼を勝ち取ることも難しいことではなくなります。

企業活動に対する顧客の反応やクチコミがリアルタイムかつ広範囲に伝播している今、あらゆる顧客接点を確保しながら、一人ひとりの個客との絆を確保し、迅速なリスク対応を図っていく意味でも、ビッグデータの活用は欠かせない要件となっているのです。

産業・交通・社会インフラ分野でのビッグデータ活用

近年は電力、製造、輸送、交通、社会インフラなど、広範な領域で RFID や各種センサーから膨大な情報が発信されています。こうしたリアルタイムデータの収集と解析もビッグデータならではの高付加価値ソリューションと言えます。

例えば、膨大なセンサーデータをリアルタイムに処理することで、これまでなら熟練者の経験や勘に頼っていた機器・設備・インフラなどの異常検知を確実に自動的に行えるようになりますし、センサーデータを長期間にわたって解析すれば、機器・設備の故障分析や部品寿命などの予兆検知などにも役立ちます。

ここ数年、大きな注目を浴びているスマートグリッドも、ビッグデータを有効活用したソリューションの典型的な事例となります。各家庭に配備したスマートメーター（通信機能を備えた電力メーター）のデータを自動的に収集・分析することで、電力の需給バランスを細かくコントロールし、発電所や風力発電のパワーを貯めた蓄電池などからの送電を常に最適化していくことが可能となるからです。

交通システムの領域でも、走行中の自動車から GPS データや車両データを収集すれば、道路の走行状況や各車両の不具合情報などをリアルタイムに把握でき、より高精度な交通渋滞予測や迂回路の算出、事故を未然に防ぐ故障検知などが実現できるようになります。

幅広い分野に広がる可能性

もちろんこれ以外の業務分野でも、ビッグデータには大きな可能性が秘められています。例えば、顧客の個人情報保護のために社内のデータベースやファイルへのアクセスログ、重要施設への入退室ログ、監視カメラの映像ログなどを蓄積している企業は少なくないはずです。従来はこうした情報があったとしても、セキュリティ事案が明らかになった後の原因究明や被害状況の把握には多大な手間とコストがかかっていました。しかしビッグデータの分析技術を使えば、業務で発生するさまざまなイベントログを解析し、正常な状態とそうでない状態のパターンを容易に抽出できるため、将来的にはほぼリアルタイムに不正行為の予兆を発見し、コンプライアンス強化に向けた迅速・確実な対応が行えるようになります。

同様に、空港や駅、イベント会場などにおいて、多数の監視カメラから得られる画像をリアルタイムに分析すれば、通常の動きと違った行動を行う不審者を検知できるようになり、より高い安心・安全が確保できるようになります。

医療分野でも、多様なデータの活用による高度診療や医療事故の低減、オーダーメイド医療への適用などが期待されているほか、官公庁や自治体においても、天災を高精度で予測した災害対策や、大地震の際にも混乱を招かない避難誘導のシミュレーション、高度な世論分析による意志決定支援などに活かすことが可能となります。

このように、さまざまな発生源から集められる膨大なデータが潜在的に指し示す意味を迅速・的確に顕在化・知識化し、ビジネスや社会へフィードバックできるのがビッグデータ活用の本質的な価値となります。だからこそビッグデータは、企業や組織の付加価値向上だけでなく、社会問題全般の解決にも大きく寄与するものと期待されているのです。

そして、“現在”のリアルな情報と“過去”の蓄積情報を瞬時に分析・照合して、いち早く顧客ニーズの変化をとらえた業務オペレーションを実行し、新たな意思決定に役立てていく BA (Business Analytics) のアプローチも、これからのビッグデータ活用では不可欠な要件となってくるでしょう。

ビッグデータに対応する IT 基盤の条件

では今後、企業や組織がビッグデータ活用を推進していくためには、どのような IT 基盤が必要になってくるのでしょうか。日立は、次のように大きく3つの要件が求められると考えています。

- ①高速処理(オンライン、バッチ)の実現
- ②急激なトラフィック増減への柔軟な対応
- ③耐障害性と BCP(ビジネス継続性)への対応

①の「高速処理」は、爆発的に増え続けるビッグデータを高速かつリアルタイムに処理するためには最も欠かせない要件となります。ビッグデータの処理は、大量に蓄積されたデータを処理する「バッチ処理」と、時々刻々と収集されるオンライン系データを処理する「ストリームデータ処理」「分散キャッシュ」に大きく分けられ、それぞれに対応する基盤が求められます。また、それらの基盤へのデータ入出力やデータ解析を高速化するためのデータマネジメント層が必要となります。

②は、急激なデータ量の増加に対しても、処理や分析のレスポンスを低下させることなく柔軟に対応するための要件です。既存の RDB(Relational DataBase) では、ノード数を増加させても I/O がボトルネックとなることでスケーラビリティに限界があるため、ビッグデータの処理に適した新たなアーキテクチャの導入が求められています。

③は、継続的に膨大なデータがフィードされるビッグデータの処理において非常に重要な要件となるものです。次々と発生するデータを用いて事象の解析や製造製品のトレースなどを行うためには、ミッションクリティカルな IT 基盤が欠かせません。しかし、可用性の維持や BCP 対応に多大なコストがかかるようではビジネスの拡大に制限が出てくるのも確かです。ビッグデータ時代だからこそ、より低コストに可用性を維持し、万一の災害時でも事業継続を実現できるプラットフォームが求められているのです。

日立のビッグデータ対応ソリューション「vRAMcloud®」

そこで日立は、これらの要件を包括的に実現したビッグデータ対応のソリューション群「vRAMcloud®」を提供します。vRAMcloud®は、企業や組織で増加し続けるビックデータを扱うために最適な先端テクノロジーを集結させたもので、「仮想化技術」(Virtualized)を応用した「インメモリ処理技術」(RAM)および「クラウド管理技術」(Cloud)の頭文字をとって命名しました。

vRAMcloud®の目的

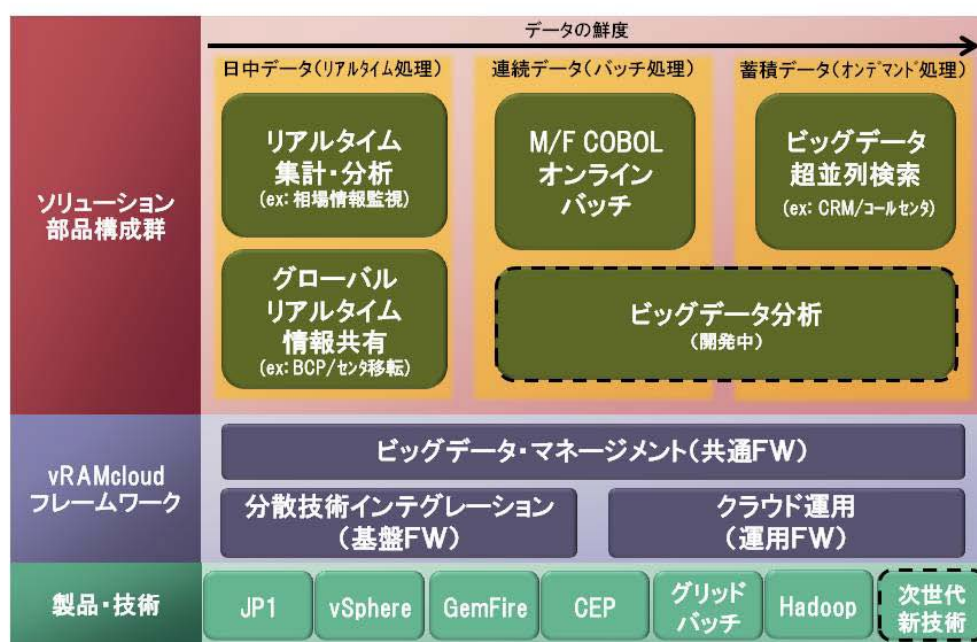
vRAMcloud®は、ストリームデータ処理や超並列処理などの製品群と、これら製品間をつなぐフレームワークから構成されており、お客さまのニーズに最適なビッグデータ利活用を実現するためのソリューション群です。

仮想化技術によって安価な PC サーバを並列化してコスト低減を図るとともに、データを複数のサーバのメモリ上で保持し、業務アプリケーションが高速にデータへアクセスできるインメモリ処理により、入出力や検索の超高速化を実現しています。

TCO (Total Cost of Ownership) を低減するアーキテクチャの提供に加え、急激な情報の増加やサーバの故障に対しても、ノードの追加・削除による動的なサーバ構成の変更が可能です。もちろんマシン構成変更の際にサービスを停止する必要はありません。

下図は、vRAMcloud®のソリューション体系を示しています。

企業のIT標準化を推進するソリューション群



※点線枠は将来提供予定

図 1 vRAMcloud ソリューション群

表 1 vRAMcloud ソリューション一覧

項番	ソリューション	概要	備考
1	リアルタイム集計・リアルタイム分析	大量データの収集、集計、分析を行う。 例：株注＋出来高集計、商品販売情報集計、 RFID・GPS による位置管理など	
2	グローバルリアルタイム情報共有	全世界でリアルタイムに情報共有を実現 例：グローバルブック	
3	M/F COBOL オンラインバッチ	大量データを対象としたバッチ処理の高速化。 例：CMS(キャッシュマネジメントシステム)、SCM 需給計算・シミュレーション	
4	ビッグデータ超並列検索	大量データを対象とした検索、更新、削除の高速化。 例：コールセンタ、窓口での顧客検索(CRM/SFA)	
5	ビッグデータ分析	開発中：大量データをさまざまな軸で高速に検索。 例：DWH(Data WareHouse)	

vRAMcloud®は、複数の業務アプリケーションで使用するデータをメモリ上で仮想的に共有する「分散キャッシュ」技術を実装した GemFire をコアとし、「バッチジョブ超並列処理(グリッドバッチ)」、「ストリームデータ処理(CEP:複合イベント処理)」、「Hadoop」、分散キャッシュを最適にコントロールする「クラウド管理」、それらの処理基盤を柔軟に連携する「フレームワーク」などから構成されます。

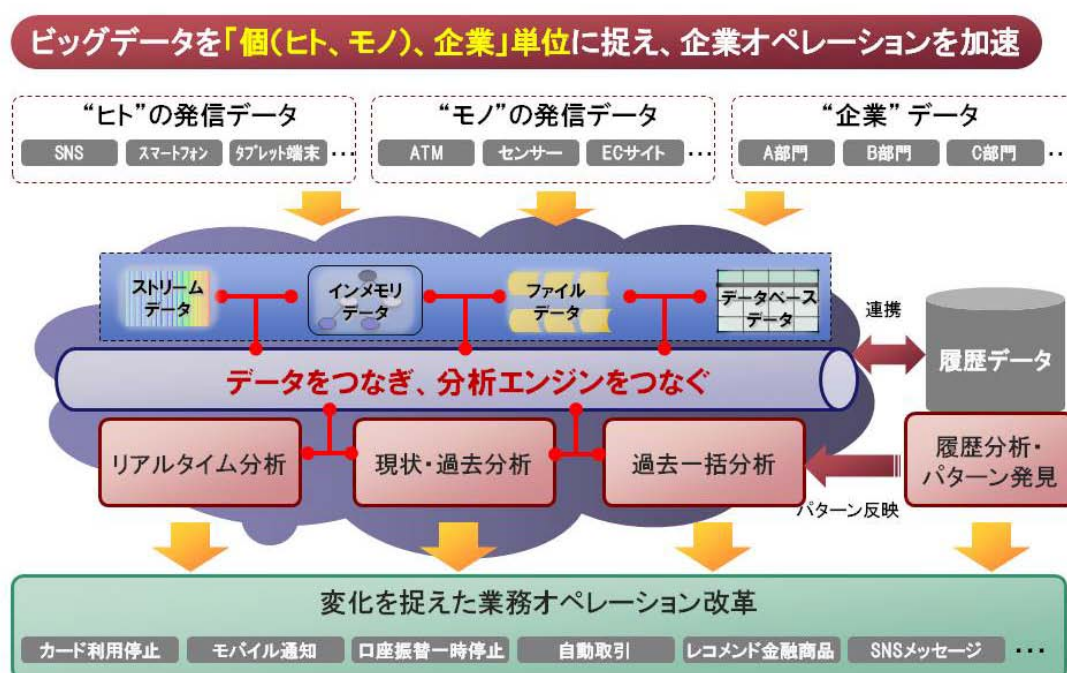


図 2 vRAMcloud のビジョン

vRAMcloudフレームワーク

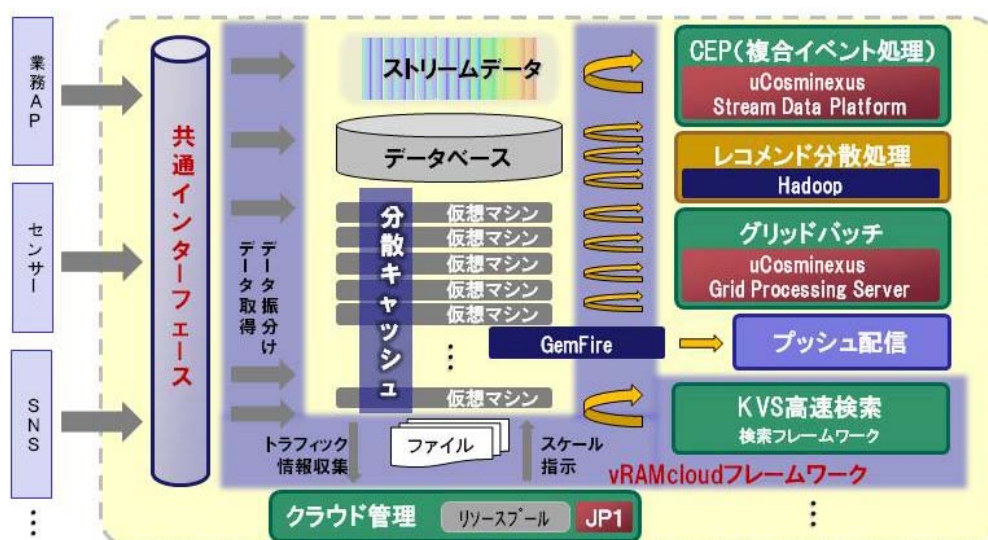


図 3 vRAMcloud ソリューションを構成する製品群とフレームワーク

日立は、ビッグデータを扱うお客さまの情報システムに対し、vRAMcloud®を構成する製品や技術、ソリューションの中から、必要となる構成要素を柔軟に組み合わせてご提供します。また、各構成要素をシームレスに連携させ、各機能を有効に利用できる過去事例や POC (実証実験) のノウハウの詰まったフレームワークも一緒にご提供するため、開発効率を大幅に向上させることが可能です。

表 2 vRAMcloud 構成要素一覧

項番	構成要素	説明	備考
1	分散キャッシュ・仮想化 GemFire+vSphere	高速なインメモリ処理を行う分散キャッシュエンジン。 差分連携を効率化するための、Push 配信機能も提供。 分散キャッシュは地理的条件に縛られないため、追加製品を購入することなく、BCP が可能です。 また、拠点を定めることで Follow-the-Sun が可能なシステムとなります。	
2	ストリームデータ処理	データ発生と同時にメモリ上でリアルタイム処理をする製品。 uCosminexus Stream Data Platform 時系列データ解析などに向け、“今”流れている情報を収集、分析できます。	
3	Hadoop	サーバを大量に並べた並列処理で、従来の DBMS では難しかった大量データのバッチ処理を高速化可能です。日立ならではの高信頼な運用技術も付加して提供できます。	オープンソースソフトウェアコミュニティ Apache Software Foundation にて開発・公開

項番	構成要素	説明	備考
4	バッチジョブ超並列処理 (グリッドバッチ)	バッチ処理を超並列化し高速に行う製品。 uCosminexus Grid Processing Server 実証実験では、超並列処理によりメインフレームのバッチ処理を最大81倍も効率化を実現。 現状プログラムは、ビジネスロジックをそのままに DB インタフェースを修正するだけで移行できます。	
5	クラウド管理	データ量の急激な増減に対し、分散キャッシュの動的かつ自動的な変更を実現するクラウド管理機能です。 日立的統合運用管理ツール「JP1」の機能を核に、プライベートクラウド、パブリッククラウド双方で最適なコストでの運用を可能とします。	

「ストリームデータ処理基盤」は、絶え間なく流れ込んでくる大量のデータを、データの発生と同時にメモリ上でリアルタイムに処理する製品です。集計・分析のシナリオ定義を、一般的なデータベース言語 SQL (Structured Query Language) を拡張したスクリプト言語 CQL (Continuous Query Language) で容易に記述できるため、SQL になじみのあるユーザーであれば、簡単にシナリオ定義を作成できます。

「バッチジョブ超並列処理基盤(グリッドバッチ)」は、安価なサーバを並列化することで IT コストの適正化を図りながら容易にスケールアウトを実現します。COBOL などで開発された既存バッチ処理からの移行容易性、データの分割配置方法の柔軟性、バッチ処理の終了時間厳守などの特徴があり、ミッションクリティカル領域におけるバッチ業務全体の高速化と高信頼化に大きなアドバンテージを持っています。

vRAMcloud[®]は、特定のサーバに負荷が集中したり、障害が発生した場合でも、動的にサーバを追加・削除できる仮想化アーキテクチャ(GemFire+ vSphere)を採用しています。これにより、小売業なら中元/歳暮/クリスマスシーズン、製造業なら新製品や限定商品の発表時期といった、平時よりも急激に高いデータ量やトラフィックが発生するような場合でも、ダイナミックにメモリ空間を拡張しながら柔軟に対応できるスケーラビリティを確保しています。

- DB層スケールアップ型の限界をキャッシュ層で吸収、煩雑なメンテナンス作業不要
- システム全体を柔軟に構成変更が可能(日次、オンライン中など)

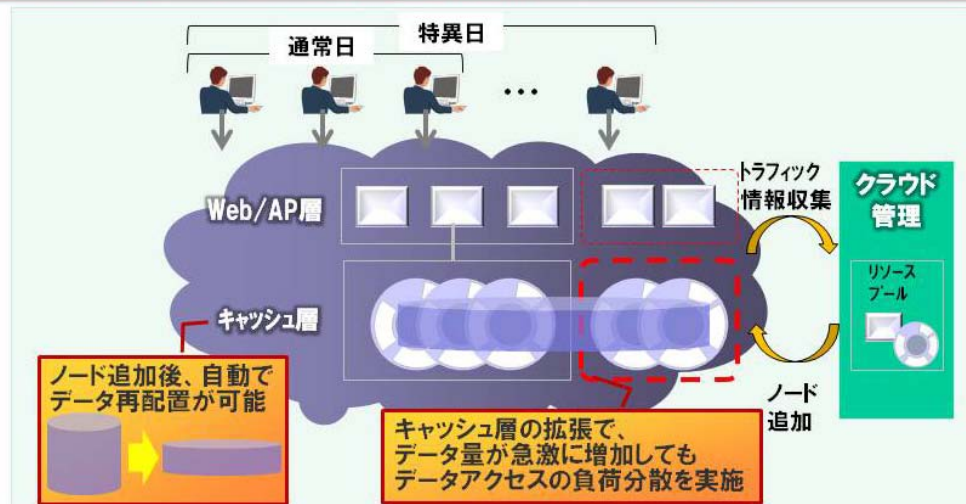


図 4 スケールアップとデータ量増加時のデータリバランス

サーバ(ノード)の追加を行った際も、vRAMcloud[®]ではノード間の処理の平準化を自動的に行うため、処理対象件数が増加しても一定したレスポンスを確保できます。従来のアプリケーションサーバを原因とするボトルネックを解消し、高速なデータアクセスを実現します。

また、ピーク時以降はサーバリソースを縮小し、最適なコストでシステムを安定的に運用できるため、初期投資の低減と TCO の削減にも寄与します。

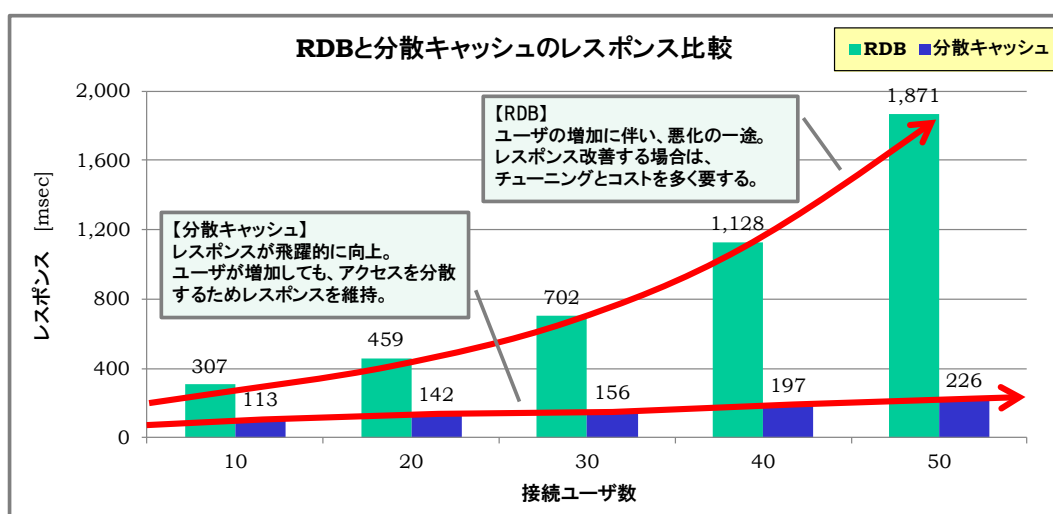


図 5 RDB と分散キャッシュのレスポンス比較

物理マシンのコア数が少なくても高いスループットが実現できるため、インフラコストの削減が可能です。また、今後のデータ量増加に伴う追加投資など、ランニングコストの削減も期待できます。下図は、RDB と分散キャッシュを用いた場合のコスト試算比較です。

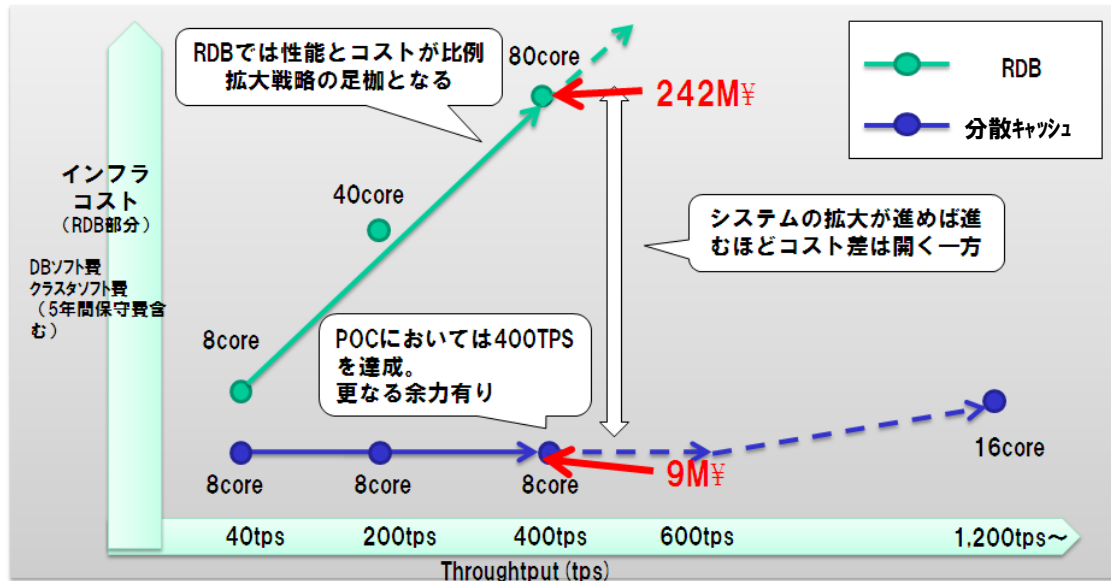


図 6 RDB と分散キャッシュのインフラコスト比較

BCP にも即対応

継続的に膨大なデータがフィードされるビッグデータの処理と活用においては、データの確実な保全性と、可用性の高いミッションクリティカルな IT 基盤が欠かせません。

vRAMcloud®で適用している分散キャッシュ技術はインメモリ処理を採用しているため、ディスク上にデータを保持するRDBなどと比べて耐障害性が懸念されがちですが、vRAMcloud®では二重化、三重化によるシステム構築が可能のため、こうしたリスクを極小化することが可能です。高可用構成の場合、たとえ障害が発生しても、操作しているユーザーはサーバが切り替わったことにすら気づかないまま、実行していた操作を継続できます。

さらに vRAMcloud®の分散キャッシュ技術は、複数のノード間でデータレプリケーションを行える仕組みを持っているため、同一拠点だけでなく遠隔地の他拠点ともDR(Disaster Recovery)による分散化対応を容易に実現することができます。バックアップサイトへのデータ転送は、パラメータ設定だけで行うことができ、DR サイトへは差分情報のみを転送することで、ユーザーに接続先を意識させることなくBCP 対応が完了します。

つまりvRAMcloud®を適用すれば、特別なBCPソフトを用意することなくミッションクリティカルなビッグデータ対応基盤が構築でき、事業継続性の向上にも寄与するのです。

まとめ

日立が提供する vRAMcloud[®]は、分散キャッシュソフトウェア「GemFire」、オープンソースソフトウェア「Hadoop」、日立のクラウドサービスプラットフォーム「Cosminexus」のストリームデータ処理や超並列処理などの製品群と、これらの製品間をつなぐフレームワークで構成しています。お客様のビッグデータ利活用ニーズに対して、適材適所で最適な製品とフレームワークを提供します。

安価な PC サーバによる最小構成からスタートし、スケーラブルに拡張できるアーキテクチャのため、「まだビッグデータを活用する予定がない」とお考えのお客様も、まずは適用評価から始めて、その効果を実感していただくことが容易に行えます。

日立では、導入に関する POC (Proof Of Concept:実証実験) 支援サービスなども用意しておりますので、ぜひご活用ください。

ビッグデータを高速に処理できる環境を持つことの付加価値には、計り知れないものがあります。実世界から間断なく流れ込んでくる膨大な情報をリアルタイムに分析すれば、より迅速な意思決定やリスク管理、現場状況の把握などに生かすことができるほか、バッチ処理の短時間化によるオンラインサービス時間の増加などで、顧客満足やビジネスチャンスの向上にも寄与します。

また、プライベートクラウド、パブリッククラウドのご選択をいただければ日立のデータセンター内で最適なクラウド環境をご用意するため、繁忙期のピーク処理量に合わせたシステムリソースの事前準備は不要となるサービスを今後ご提供していきます。

さらに、処理件数の増加に対応した自動的なスケールアウトで処理速度を一定に保つことができるため、保守費用の削減や戦略投資分野への資金割り当ての推進にも寄与します。

現在、あらゆる業務分野において、より鮮度の高いデータに基づいた迅速な意思決定や状況分析、スピードを付加価値とした新ビジネスの創出が強く求められています。日立の vRAMcloud[®]を、お客様の企業価値向上と競争力の強化に、ぜひお役立てください。

今後の計画

ビッグデータ活用では、膨大なデータをリアルタイムに分析し、時々刻々と変化する状況を見極めながら、一歩先をゆく行動をとるための仕組み作りが重要なポイントとなります。そこで日立は現在、vRAMcloud[®]の構成要素の1つとして、従来型の BI (Business Intelligence) の発展形として、多様なデータから将来予測に基づく意思決定やビジネスの最適化までを支援する BA (Business Analytics) 機能の開発に取り組んでいます。

本資料のほか、vRAMcloud[®]を適用した実証実験の結果や具体的な技術的特徴をまとめたホワイトペーパーも発行しています。併せてご参照ください。

以上