



WebSphere. software

IBM WebSphere eXtreme Scale で実現する データ・アクセスのボトルネック解消

July 2009

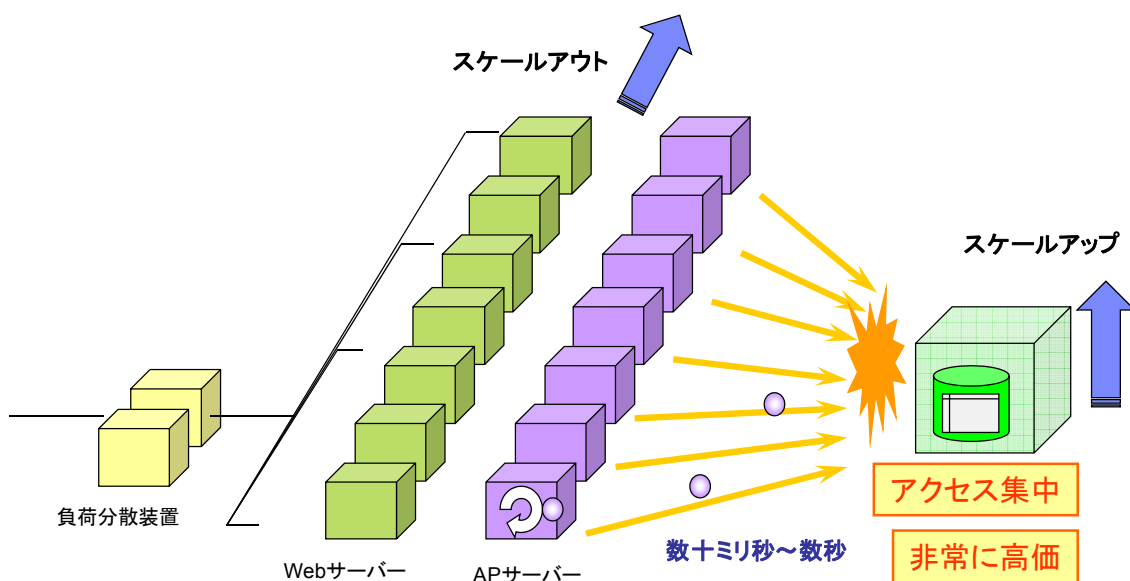
はじめに

IBM® WebSphere® eXtreme Scale は、インメモリー処理によるデータ層への高速アクセスと、リニアなスケーラビリティにより大量トランザクションの高速処理を実現する、データ・グリッド（分散キー・バリュー型データストア）製品です。大量のトランザクション処理（XTP: eXtreme Transaction Processing）や大規模 Web サイトでのレスポンス・タイムの高速化を実現します。企業内クラウドにおけるスケーラブルなデータ層の基盤としても注目を浴びています。

2009 年 6 月 26 日に WebSphere eXtreme Scale V7.0 が発表されました。本技術白書では、WebSphere eXtreme Scale の特長と V7.0 のハイライトをわかりやすく解説します。

スケールアップ方式によるデータベース増強の課題

Webアプリケーションのインフラ構成では、アプリケーション・サーバー層とデータベース・サーバー層を階層化することが一般的です。リクエストの増加に伴ってシステム全体のキャパシティを拡張する場合、アプリケーション・サーバー層では安価なサーバーの台数を追加していく、いわゆるスケールアウト方式が可能です。しかし、データベース・サーバー層では、データの整合性を維持するためなど、いくつかの理由によりスケールアウト方式による拡張が困難です。このため、従来はデータベース層のキャパシティ拡張手段としては、サーバーをより高性能なものに置き換えていく、いわゆるスケールアップ方式が一般的なアプローチです。

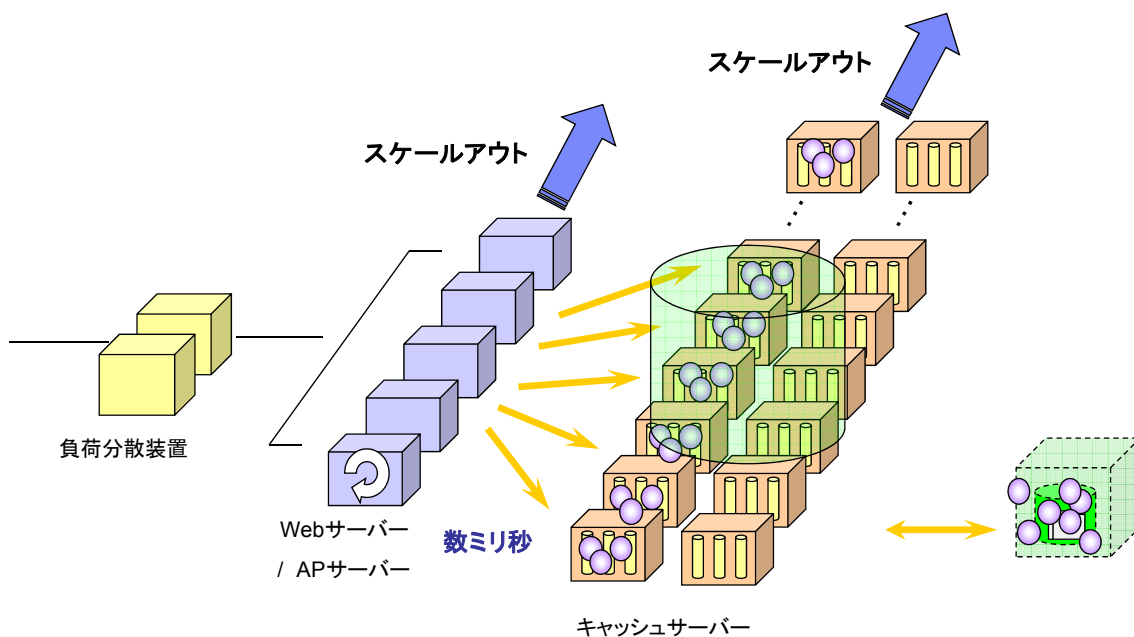


しかしながら、スケールアップ方式は、以下の点において考慮する必要があります。

- 性能向上のためにかかるコスト
- データ・アクセスの競合などによる性能向上の限界
- ハードウェアのスペックの上限
- スケールアップ時にサーバーのパーツ追加や筐体交換が発生するため、サービス停止が不可避である

スケールアウト方式でデータ・アクセスを高速化する WebSphere eXtreme Scale

データベース層に対してスケールアウト方式を適用し、そのメリットを享受することは不可能なのでしょうか。ひとつの解決策として、アプリケーション・サーバー層とデータベース層の間にキャッシュ層を設け、データベース・サーバーへの負荷を肩代わりさせるという方法があります。



キャッシュ層では、スケールアウト方式で多くのサーバーのメモリー上にデータをキャッシュすることで高速なデータ・アクセスと、データベース層の負荷軽減を同時に実現することができます。

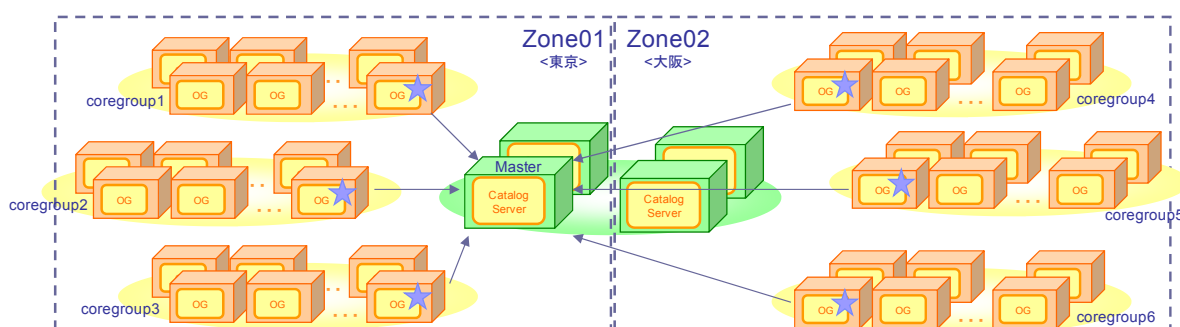
ただし、単純なキャッシュ機構では、データの整合性の維持や、複数のキャッシュ・サーバーへの効率的な負荷分散など、エンタープライズ・システムに求められる要件を満たすことは容易ではありません。このような高度なキャッシュ機構を実現するソリューションは一般的には「インメモリー・データ・グリッド」と呼ばれます。WebSphere eXtreme Scale(旧 WebSphere eXtended Deployment Data Grid)は、「インメモリー・データ・グリッド」を実現するための製品であり、証券取引システムなどの大量トランザクションを高速に処理する必要のあるシステムにおいて利用されています。

WebSphere eXtreme Scale の特長 :

- (1) 極めて高いスケーラビリティ
千ノードのオーダーまで、リニアにスケールアウトが可能
- (2) 参照系だけでなく更新系にも適用可能
自己回復機能により、メモリー上のデータの信頼性を確保
- (3) Map/Reduce エージェントによる処理のスケールアウト
データだけでなく、ロジック処理もスケールアウトによる分散実行が可能
- (4) システム全体の可用性向上
サービスを停止することなく動的にノードを追加することが可能になり、可用性が向上

特長(1): 極めて高いスケーラビリティ

WebSphere eXtreme Scale の特長のひとつに、極めて高いスケーラビリティがあげられます。WebSphere eXtreme Scale は Java™ ベースのミドルウェアであり、複数の Java 仮想マシン上にオブジェクトを分散配置することにより、単一の Java 仮想マシンの限界を超えた広大なメモリー空間を利用することができます。WebSphere eXtreme Scale は Java 仮想マシン間の連携のために相互に通信を行っており、単純な peer-to-peer 通信ではノード数の増加に伴ってスケーラビリティが頭打ちになってしまいますが、WebSphere eXtreme Scale は通信ドメインを自動的にクラスター化することによってこの問題を回避しています。このような工夫により、1,000 ノード程度の規模までもリニアにスケールすることができます。



特長(2): 参照系だけでなく更新系にも適用可能

WebSphere eXtreme Scale のもうひとつの特長は、信頼性の高い自己回復機能を提供することです。

一般的に、キャッシュ・サーバー上では、データは Java 仮想マシンのヒープ・メモリー上にオブジェクトとして保持されているため、何らかの理由でサーバーがダウンした場合、データは失われてしまいます。しかし、WebSphere eXtreme Scale はレプリケーション・ポリシーを

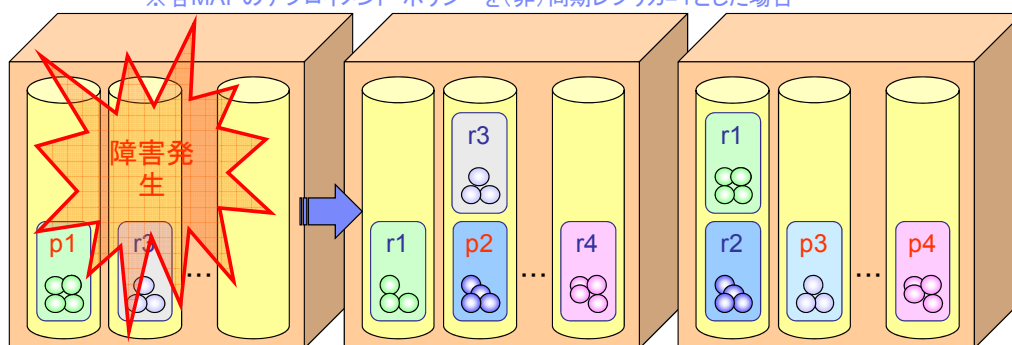
指定することで、データのレプリカ（複製）を自動的に作成するため、サーバーがダウンした場合でも、データが失われるリスクを回避することができます。

また、レプリカは同期・非同期の2タイプが作成可能です。同期レプリカは、レプリカの作成が厳密に保障されますが、データ更新時にノード間通信を伴うため若干オーバーヘッドがあります。逆に、非同期レプリカはデータ更新時のオーバーヘッドがない代わりに、障害発生のタイミングによってはレプリカの整合性が崩れる可能性があります。これらは要件によって使い分ける、ないしは併用することが可能であり、例えば「同一データセンター内は同期レプリカ」「災害対策センターへは非同期レプリカ」という構成をとることも可能です。

●自己回復

- ▶ 可用性確保のため、データのレプリカを同期/非同期で保持
 - データを保持する区画(シャード)には、プライマリとレプリカがある
 - レプリカの数や同期方法は構成ファイルで設定
- ▶ 障害発生時には、ポリシーが満たされるよう自動的に対応
 - レプリカがプライマリへと昇格
 - 必要数のレプリカがない場合は自動的に作成される

※各MAPのデプロイメント・ポリシーを(非)同期レプリカ=1とした場合



また、キャッシュ・サーバー上で更新されたデータは、ローダー・プラグインを利用してバックエンドのデータベースに永続化することができます。永続化の方法としては、「ライト・スルー (write through)」と「ライト・ビハインド (write behind)」のいずれかが選択可能です。

「ライト・スルー」の場合、メモリー上のデータ更新と同時にバックエンドのデータベースへの永続化が行われます。この場合、メモリー上のデータとバックエンドのデータベースのデータの整合性を保つことができますが、更新処理のパフォーマンスはデータベースへの直接更新と同等になります。

「ライト・ビハインド」の場合、メモリー上のデータ更新をキューイングしておき、一定時間間隔ないしは一定更新数ごとにまとめてデータベースへの永続化が行われます。この場合、ある一時点ではデータの整合性が崩れますが、以下のようなメリットが得られます。

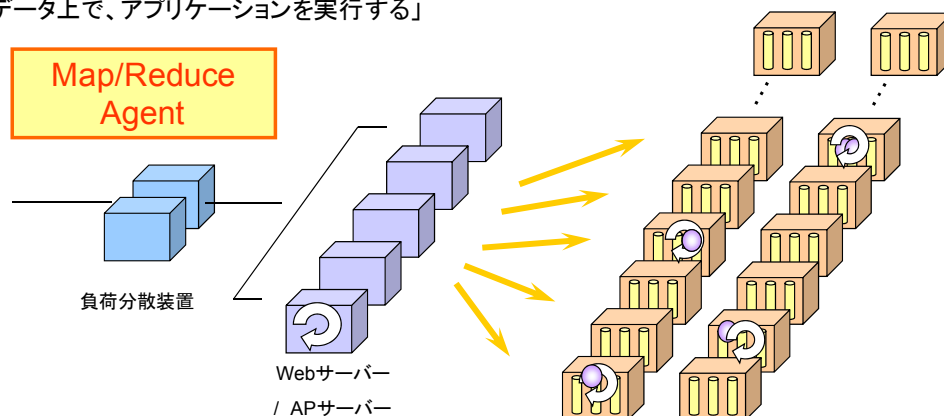
- 更新処理がメモリー上で完結するため、処理時間を大幅に短縮できる
- 永続化処理をまとめて行うことで、データベース・サーバーへの負荷を軽減できる
- データベース・サーバーが停止していても、メモリー上のデータでサービスが継続できるため、システム全体の可用性が向上する

特長(3): Map/Reduceエージェントによる処理のスケールアウト

スケールアウト方式のメリットをさらに活用する方法として、アプリケーション・ロジックの一部をキャッシュ・サーバー上で分散並列実行する機能も提供されています。これは、キャッシュ・サーバーのメモリーだけでなく、CPU も使ってしまおう、という発想です。この機能を Map/Reduce エージェントといいます。

● エージェント機能

- ▶ アプリケーション・ロジックを、データの近くで分散並列実行する
 - 真にリアル・タイムの応答性能を求められるシステムではネットワークにアクセスする時間も回避したい
 - データをアプリケーションに持ってくるのではなく、「アプリケーションをデータに持っていき」「データ上で、アプリケーションを実行する」



エージェントは Java で作成し、キャッシュ・サーバーに配布することで実行可能になります。エージェントは実行時に配布先キャッシュ・サーバーのメモリー上に格納しているデータしか参照しないため、ネットワーク経由のデータ参照が一切発生せず、非常に高速な処理が可能です。

また、多数のエージェントに分散並列実行させた結果のとりまとめは、WebSphere eXtreme Scale の API が自動的に行います。従って、アプリケーション側では並列度(エージェントの数)などを意識する必要はなく、容易に利用することができます。

特長(4): システム全体の可用性向上

WebSphere eXtreme Scale を導入するメリットは、データ・アクセスの高速化だけではありません。24 時間 365 日無停止で運用されているシステムで、キャパシティーを増強するという例を考えてみましょう。データベース層のスケールアップ方式では、サービスを停止せずにキャパシティーを増強することは非常に困難でした。

一方、WebSphere eXtreme Scale でキャッシュ層を構築した場合は、追加のサーバーを用意して、キャッシュ層に追加するだけで、キャッシュ層のキャパシティー増強を実現できます。そ

の際、サービスを停止する必要はありません。WebSphere eXtreme Scale は新たに追加されたノードを自動的に検知し、キャッシュの再配置を行います。これはキャパシティーを縮小する場合も同様であり、サービスを停止せずにキャパシティーを動的に変更できるため、負荷の増減が大きいシステムなどでは特に効果が期待できます。

WebSphere eXtreme Scale V7.0 の新機能と強化点

2009年6月26日に WebSphere eXtreme Scale の最新バージョンである V7.0 がリリースされました。ここでは、V7.0 での追加機能のハイライトをご紹介します。

1) メモリー利用効率の向上

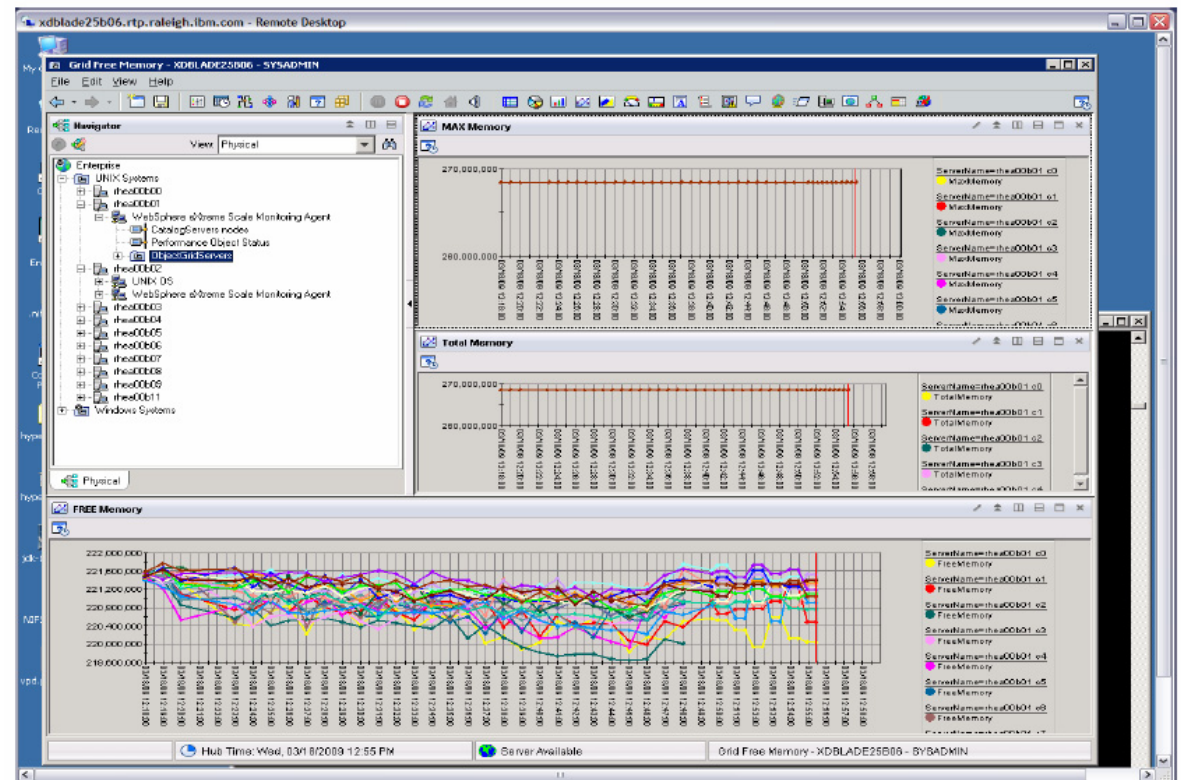
キャッシュ・サーバー上のメモリー管理方式を改良し、データ本体やインデックスに関するメモリー利用効率を向上しています。同時に、GC(ガベージ・コレクション)によるパフォーマンス・オーバーヘッドの削減も期待できます。

2) リアルタイム Java 対応

WebSphere Real Time(別製品)上での WebSphere eXtreme Scale の利用が可能になりました。WebSphere Real Time が提供するメトロノーム GC は、ガベージ・コレクションを一定時間内に終了させることを可能にします。従来はガベージ・コレクションが発生したタイミングでレスポンス・タイムが一時的に悪化する可能性がありましたが、WebSphere Real Time を利用することで安定したレスポンス・タイムを維持することが可能になります。

3) レポートニング/モニタリング機能の強化

IBM Tivoli® Monitoring (ITM)、および Hyperic HQ との統合が可能になりました。これらの製品により、WebSphere eXtreme Scale の稼働状況を視覚的に確認できるだけでなく、統合的な運用監視に組み込むことが可能になります。



4) 「ドロップイン」キャッシュ機能の強化

WebSphere eXtreme Scale は、アプリケーション・サーバーに対するプラグイン機能も提供しています。このプラグイン機能を利用することで、アプリケーションに修正を加えることなく、WebSphere eXtreme Scale を活用することが可能です。

プラグイン機能として、以下が提供されています。

- HTTP セッション拡張プラグイン
- Hibernate/JPA セカンドレベル・キャッシュ・プラグイン
これらに加え、V7.0 では以下の機能が追加されました。
- WebSphere Application Server DynaCache プラグイン

まとめ

本書では、WebSphere eXtreme Scale が提供する「インメモリー・データ・グリッド」のテクノロジーによって、従来は適用が困難であったデータ・アクセス層に対しても、スケールアウト方式によるキャパシティー増強が可能になることをご紹介します。

WebSphere eXtreme Scale 適用のメリットは、単にデータ・アクセス層のスケールアウトが可能になることだけではなく、自律的な高可用機能によってシステムの可用性を向上し、メンテナンス停止を伴わないキャパシティー管理が実現可能になるなど、多岐にわたります。

キャッシュ層の導入は、データベースに対してだけでなく、メインフレーム上で稼働するサービスへの呼び出しなどをオフロードするためにも有効です。データ・アクセスに関する負荷が集中している箇所があれば、ぜひ WebSphere eXtreme Scale の適用をご検討ください。

関連リンク集

[発表レター・製品ページ]

WebSphere eXtreme Scale V7.0 発表レター

<http://www.ibm.com/jp/domino02/NewAIS/aisextr.nsf/ByLetterNo/SWA09048>

WebSphere eXtreme Scale 製品ページ

<http://www.ibm.com/software/jp/websphere/apprtransaction/extremescale/>

WebSphere eXtreme Scale 評価版ダウンロード

<http://www.ibm.com/developerworks/jp/downloads/ws/wsdg/>

[developerWorks]

WebSphere eXtreme Scale 技術情報

<http://www.ibm.com/developerworks/jp/websphere/category/wxd/wxs.html>

「クラウドの実現技術としての WebSphere eXtreme Scale」

http://www.ibm.com/developerworks/jp/websphere/library/wxd/wxs_cloud/index.html

「インメモリー技術を利用した超高速トランザクションへの対応」

http://www.ibm.com/developerworks/jp/websphere/library/wxd/wxs_soliddb/index.html

WebSphere eXtreme Scale を使ったコンプレックス・ビジネス・イベント処理

http://www.ibm.com/developerworks/jp/websphere/library/bpm/wbe61_extremescale/1.html



© Copyright IBM Corporation 2009

日本 アイ・ビー・エム株式会社
ソフトウェア事業
〒106-8711
東京都港区六本木三丁目2番12号

当資料に記載の肩書きや数値、固有名詞等は英語版掲載時のものであり、変更されている可能性があります。

IBM、IBM ロゴ、ibm.com、Tivoli、WebSphere は、世界の多くの国で登録された International Business Machines Corp.の商標です。

他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。

Java およびすべての Java 関連の商標およびロゴは、Sun Microsystems, Inc.の米国およびその他の国における商標です。

他の製品名およびサービス名等はそれぞれ IBM または各社の商標である場合があります。

現時点での IBM の商標については、
<http://www.ibm.com/legal/copytrade.shtml> をご覧ください。

本文書は、お知らせのみを目的にして提供されま
す。できる限り正確を期すよう努めてはいますが、
内容は現時点のものであり、いかなる明示または
暗黙の保証も負いかねます。また、本情報は IBM
の現在の製品計画および方針に基づいており、予
告なく変更される場合があります。IBM は本文
書およびその他関連文書の使用に起因するいか
なる損害についても責任を負いません。本文書は、
IBM(または IBM のサプライヤーまたはライセ
ンサー) にいかなる保証責任を負わせるものでは
なく、また、IBM ソフトウェアの使用に際し適
用される、プログラムのご使用条件の内容も変更
するものではありません。