

Linux、WebLogic、Oracle9i RAC での大規模基幹システムの構築事例

Case Study: Constructing Large-Scale Backbone System Using Linux, WebLogic and Oracle9i RAC

山崎 貴弘 小池 洋介
Takahiro Yamazaki Yousuke Koike

概要

近年、オープンソースは低コストの魅力に加え、性能、信頼性、保守性、サポート性が向上し、業務システムでも十分に採用できるようになり、プロプラエタリソフトと同等になりつつある。そのため、企業利用も急速に高まりオープンソースの採用が本格化し始めてきていたが、大規模基幹システムへの採用事例は殆ど見られなかった。

本稿では、オープンソースと複数ベンダーの製品とを組み合わせるV社様(以下、V社)向け大規模基幹システムを再構築した事例について述べる。今回のシステム再構築は日本ヒューレット・パッカード社(以下、HP社)と協業し、パフォーマンスの事前検証、性能評価によるシステム構成の設計などを実施した。また、複数製品を組み合わせるときに保守・運用で課題となるパッチ適用についても言及する。

1. システム再構築範囲

V社は、ビジネスの発展、拡大と共に根本的なシステムの再構築を行ってきており、その当時の最新ITを駆使したビジネスを展開してきていた。今回もV社のビジネス拡大に伴い、今までの基幹システムで提供してきた、顧客管理、ビジネス情報管理、売上管理、単体決算、人事管理等では系統的に耐え切れなくなり、システムを刷新する必要があった。

また同時に、V社が提供していたインターネットサービスでもビジネスの成長にシステムが耐え切れなくなり、基幹システムと合わせて刷新することとなった。

基幹システム、インターネットシステム、それに開発環境を含めて、サーバ約50台規模のシステム再構築となった。

2. 適用技術の選定

今回のシステム再構築では、TCO(Total Cost of Ownership)の削減、ビジネスに連動したシステムの拡張、縮小を可能とし、

ビジネスモデルの進化に影響を受けない、安定性かつ可用性の高い基幹システムを最新のITを駆使して構築することが求められていた。

このような背景の中で、Linuxベースでのシステム再構築を検討していった。

2.1 事前検証

プロジェクト発足2ヵ月前に、UNIXベースのシステムからLinuxベースのシステムへマイグレーション可能かどうか、V社の要望に応える基幹システムとしての性能がLinuxベースのシステムで出せるかどうか、この2点の観点から事前検証を実施した。事前検証には、基幹システムで稼働中のクライアントアプリ、バッチプログラムをLinuxベースのシステム用にマイグレーションし、動作確認、パフォーマンス測定を実施した。

事前検証の結果では、Linuxサーバの動作、パフォーマンスは、商用UNIX専用に最適設計されたベンダーハードウェアと遜色ない性能評価であった。

この事前検証によりV社の要望にLinuxが基幹システムでも

応えられるという確証を得、Linuxベースでのシステム構成を前提とした。

2.2 最新技術動向調査

Linuxをとりまく環境は常に大きく変化し、より高性能なハードウェア、開発・運用をサポートするソフトウェア群、新たな開発手法など、様々なアーキテクチャが生まれている。それら様々なアーキテクチャを検討し、適用技術、ハードベンダーを選定した。

特に今回のシステム構築では、ハードベンダーに対して、コンサルティングベースでLinuxシステム構築の構想段階から一貫したサポート体制を組んでもらえるプロフェッショナル・サービスとしての役割を期待していた。そこで今回のシステム構築では、HP社と協業でシステム構築を行っていくこととなった。

2.2.1 BEA WebLogicの採用

APLサーバはシステム構築の根幹に関わるので、不具合発生頻度、製品シェア、当社が保有しているノウハウと合わせてしっかり選定する必要があった。

今回は、上記の判断に加え、操作性、柔軟性、システム構築速度も含め、BEA WebLogicが最適と判断した。

2.2.2 Oracle9i RAC (Real Application Clusters)の採用

V社のビジネスの発展、拡大に対応するビジネス・クリティカルなデータベースには、スケーラビリティ、可用性の高さだけでなく、サービス時間の延長を実現可能とする高速トランザクション処理能力が必要であった。

Oracle9i RACは、ビジネスの拡大に合わせて既存システムへのDBサーバ(ノード)追加でシステム拡張することが可能である。また、TAF (Transparent Application Failover)機能により、ハードウェア障害が発生した場合、自動的に残りの正常なノードに処理を分散することが可能であり、高スケーラビリティと高可用性の二つを兼ねそなえていた。

さらに、Oracle9i RACはオンライントランザクション処理の拡張性が高く、分散していたデータベースの統合が可能となり、バッチ連携処理時間の短縮によるサービス時間延長にも効果的であった。

表1 WEBアプリケーションサーバ比較

No.	製品	製品特性	考察
1	WebLogic Server (BEAシステムズ)	・ J2 EE・Web サービス等の最新技術への対応が速い	高いシェアを誇っており、実績・信頼性において評価は高い。
2	WebSphere Application Server (IBM)	・ J2 EE・Web サービス等の最新技術への対応が速い ・ IBM製ハード・ミドルウェアとの親和性が高い	WebLogicと比較しやや信頼性が劣るが、不採用を決めるほどでもない。IBM製のサーバを導入する場合は、親和性の高さから最適である。
3	SUN ONE Application Server (Sun Microsystems)	・ 開発言語C++の利用が可能	他ベンダー製品と比べるとやや費用が高い。
4	Oracle9i Application Server (Oracle)	・ Oracle DBとの親和性が高い ・ 開発言語PL/SQLの利用が可能	OracleのDBを採用する場合には統合環境のメリットがあるが、最新技術・機能の面で他ベンダー製品に比べると評価は低い。
5	Jakarta Tomcat	・ 商用・非商用を問わず無償で利用できる	コスト面のメリットが魅力的ではあるが信頼性やサポート面が不安。

※2002年時点での評価

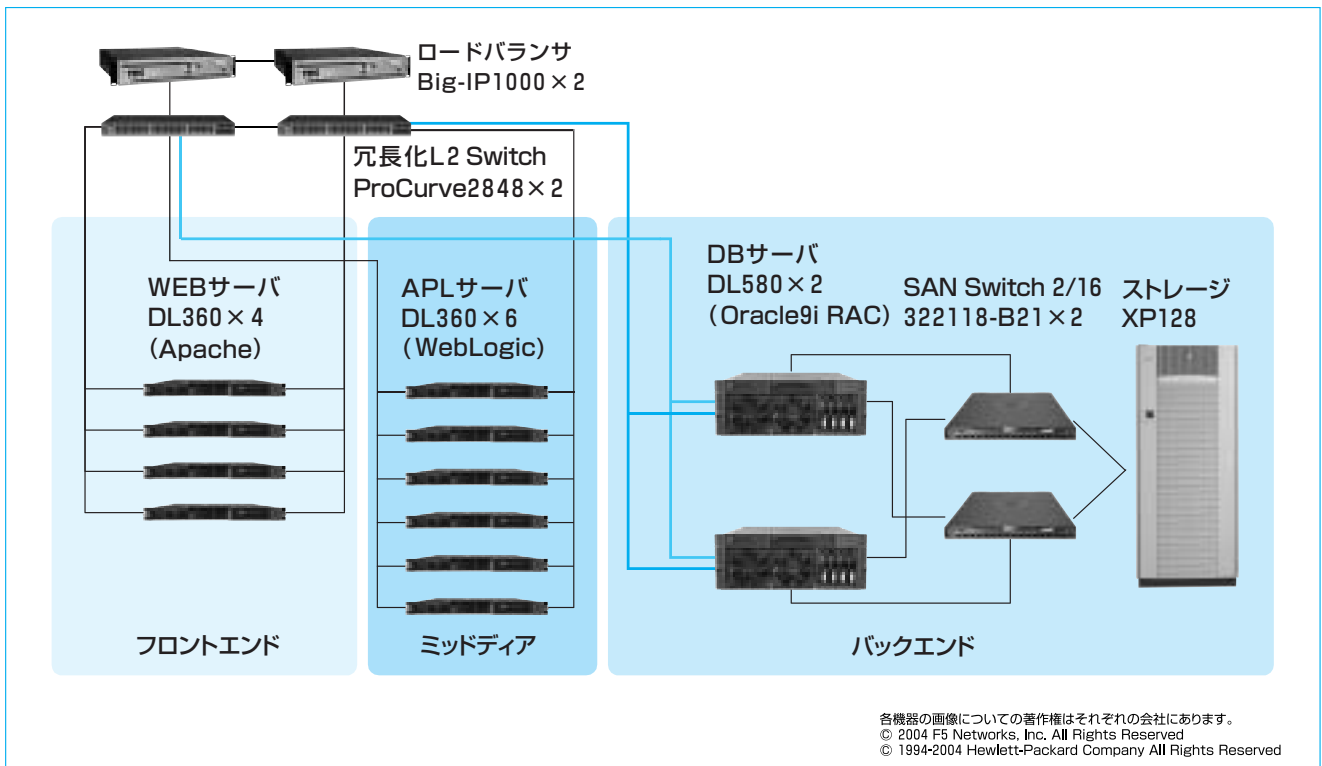


図1 3階層構成システム

2.2.3 3階層構成システム

V社のビジネス展開のスピードを考慮し、製品サイクルに依存しない拡張、縮小が可能で、ハイスペックな廉価製品を採用可能なシステム構成を採用した。

さらに、ロードバランサの導入により多重化を実現し、Linux、Apache、WebLogic、Oracle9i RACを組み合わせる事で、大規模基幹システムにも耐えうる、拡張性と管理性、負荷分散と多重化策、高スケーラビリティと高可用性を可能とした3階層構成システムを採用した。

3. システム性能評価

本番サーバ構成の確定、および可用性、最適なロードバランシング構成の確定を目的に、以下の観点からサイジング検証を実施した。

(1) 大量更新

Linux、Oracle RAC構成でバッチ処理による大量更新の動作検証を行い、その性能評価によって、主としてDBサーバ、ストレージ構成を確定させる。

(2) 大量検索

Linux、Oracle RAC構成で評価用アプリケーションに

よる大量検索の動作検証を行い、その性能評価によって、主としてDBサーバ、ストレージ構成を確定させる。

(3) 多重度

Linux、WebLogic構成で複数台のWEBサーバ、APLサーバ構成環境を構築し、ベンチマーク(ストレス)テストを実施し、その性能評価によって、主としてWEBサーバ、APLサーバの構成を確定させる。

3.1 DBサーバ、ストレージの選定

本番サーバ選定において、特に重点をおいたのが、DBサーバ、ストレージ機器の選定である。完全に冗長化されたWEBサーバ、APLサーバとは異なり、DBサーバ、ストレージに関しては、システム本稼動後に拡張することが極めて難しい。そのため、可能な限り高性能のDBサーバ、ストレージを導入する必要があった。

ストレージに関しては、当時、最高スペックであったストレージXP128を導入することとし、DBサーバに関しては、サーバ数、CPU数、CPUクロック数の異なる組み合わせで性能評価を行った(図2)。

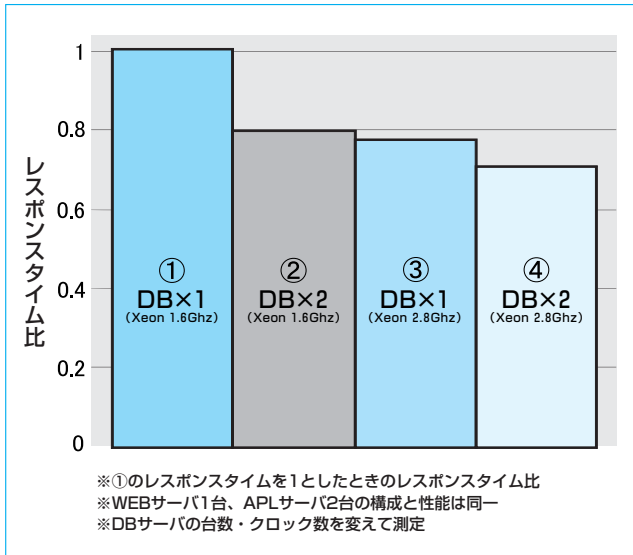


図2 DB性能評価

DBサーバの構成台数、CPUクロック数の性能比から、DBサーバの台数を増やすことにより大きく性能向上が見込まれ、また、CPUクロック数の向上により性能向上が見込まれた。ただし、CPU使用率が低く、CPU数の追加による性能向上はノード追加に比べ低いと評価した。

性能評価を実施するまでは、最大8wayまで搭載可能なDBサーバ(DL 760 G2)を採用する方向であったが、CPU数の追加による性能向上がノード追加に比べ低いという評価結果が出たため、最大4wayまで搭載可能なDBサーバ(DL 580 G2 4CPU)を採用し、システム開発当時(2003年9月)に動作

確認の取れていた最高クロック数(Xeon 2.8GHz)のCPUを採用することとした。さらに、冗長化構成と性能比も考慮に入れ、DBサーバ2台(DL 580 G2 4CPU Xeon2.8GHz メモリ8GB)の構成が最適であると判断した。

3.2 WEBサーバ、APLサーバ構成の選定

単純に高可用性を重視し、横並びに各サーバを均等分散した構成では、各サーバへのオーバーヘッドの増加などにより、必ずしも良いパフォーマンスを得ることができなかった。

サイジング検証において、APLサーバの台数を増やすことで性能向上が見込まれたが、WEBサーバの台数増加時には、ApacheProxyPluginによるロードバランス処理にバラつきがあったため、特定のAPLサーバへの負荷集中状況が発生しパフォーマンス劣化の可能性があったことが分かった。

そのため、冗長化と性能を考慮し、WEBサーバ2台(DL 360 G3 1CPU Xeon 2.8GHz メモリ1GB)、APLサーバ3台(DL 360 G3 1CPU Xeon 2.8GHz メモリ1GB)を2セット組み合わせると最適であると判断した。

均等負荷分散構成(図3)と最適化構成(図4)では、可用性の面からも性能の面からも最適構成の方が上であった。

Linuxベースに限らず3階層構成で構築する際は、必ずしも均等負荷分散構成が最適だとは限らないので、評価・分析を行い、最適な構成を判断すべきである。

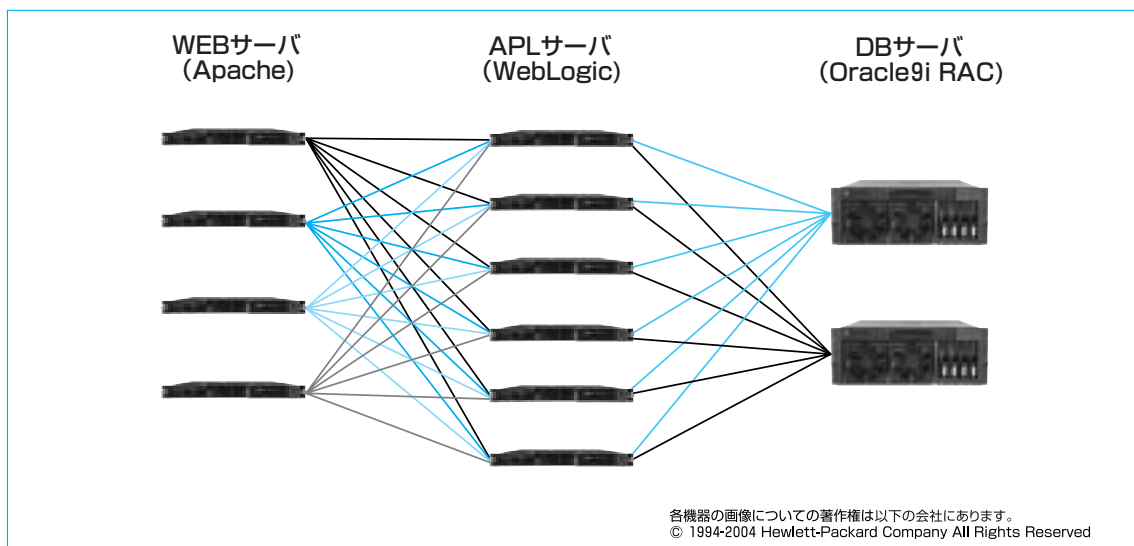


図3 均等負荷分散構成

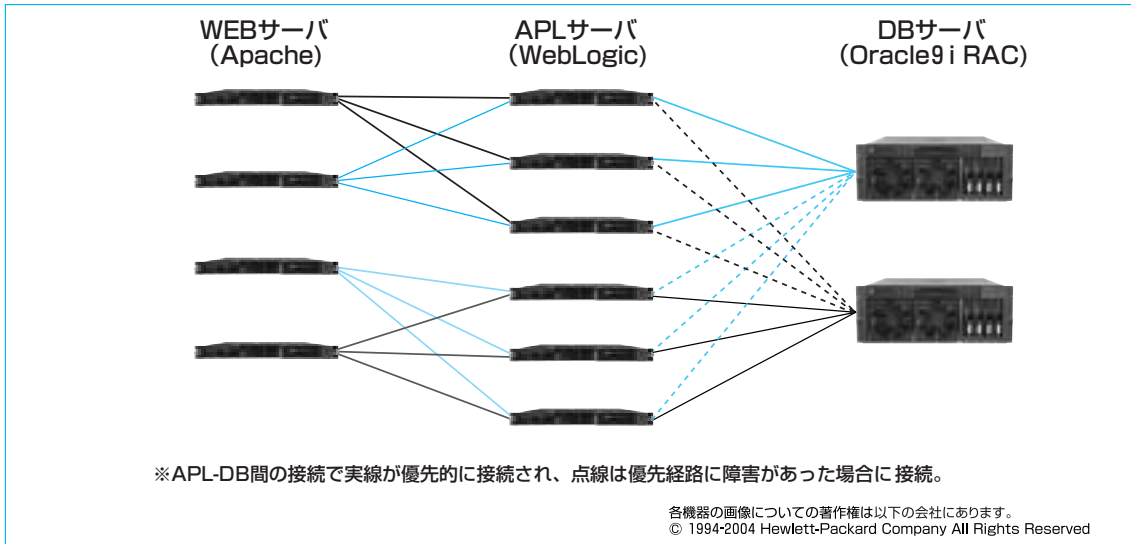


図4 冗長化と性能を考慮した最適化構成

4. システム構築

4.1 HP社との役割分担

当社としては、前例のないLinuxベースでの大規模基幹システム構築実績を作り上げること、それと合わせて、HP社が保有しているプロフェッショナル・サービスとしてのITコンサルティングとソリューションを組み合わせたトータルなインフラ構築サービス技術を習得することを今回のシステム構築の目的としていた。

さらに今後、当社とHP社で協業し、Linuxなどのオープン環境での大規模基幹システムの構築と運用サービスの拡販を目指すという目的もあり、HP社からの技術移管も含めながら、以下のサーバ構築作業をHP社と協業で実施した。(その他周辺システムの構築作業も実施)。

- (1) OS (Red Hat Linux Advanced Server 2.1 Update 2)
- (2) ロードバランサ (BIG-IP1000 R)
- (3) WEBサーバ (Apache 1.3.27_3)
- (4) APLサーバ (WebLogic Server 8.1 J Advantage Edition)
- (5) DBサーバ (Oracle9i RAC)
- (6) 共有ストレージ (XP128)
- (7) ファイバチャンネルスイッチ (Fibre Channel SAN Switch 2/16 322118-B 21)
- (8) バッチサーバ (MC ServiceGuard)

開発環境への導入・構築は当社が主体で実施し、本番環境の導入・構築はHP社が主体で実施した。

4.2 ディストリビューションとバージョン

Linuxでサーバ構成設計を実施していく上で、ディストリビューションを決定する必要がある。Linuxに関しては、当時、HP社で動作確認が取れていたRed Hat Linux Advanced Server 2.1 Update 2 (kernel:2.4.9-e.27)を採用し、WEBサーバにはApache 1.3.27を採用することとした。システム構築当時、もっと新しいバージョンも出ていたが、動作確認が取れる時期が未定だったのと、安定性の面から上記の組み合わせで実装することとした。

4.3 サーバ構成設計

HP社から導入実績のある推奨設定値を提示してもらい、当社でも設定内容を再度検討し、サーバ構成を決定していった。特にインストールパッケージに関しては、セキュリティなどの不具合影響を最小限にするために、保守・運用も考慮した必要最低限のものに絞り込んだ。

BIG-IP、Apache、WebLogic、Oracle9i RAC、ストレージXP128に関しても、当社で設定内容を再検討し、HP社と

パッケージグループ (APLサーバ)		
<input checked="" type="checkbox"/> プリンタサポート	<input checked="" type="checkbox"/> 旧X Window システム	<input checked="" type="checkbox"/> GNOME
<input type="checkbox"/> KDE	<input type="checkbox"/> マルチメディアサポート	<input checked="" type="checkbox"/> ネットワークサポート
<input type="checkbox"/> ダイアルアップサポート	<input checked="" type="checkbox"/> メッセージ、ウェブツール	<input type="checkbox"/> 画像処理
<input type="checkbox"/> ニュースサーバ	<input checked="" type="checkbox"/> NFSファイルサーバ	<input type="checkbox"/> Windowsファイルサーバ
<input checked="" type="checkbox"/> Anonymous (匿名)FTPサーバ	<input type="checkbox"/> SQLデータベースサーバ	<input type="checkbox"/> Webサーバ
<input type="checkbox"/> ルータ/ファイアウォール	<input type="checkbox"/> DNSネームサーバ	<input checked="" type="checkbox"/> ネットワーク管理ワークステーション
<input type="checkbox"/> 著作/出版	<input checked="" type="checkbox"/> Emacs	<input checked="" type="checkbox"/> ユーティリティ
<input checked="" type="checkbox"/> 旧式アプリケーションサポート	<input checked="" type="checkbox"/> ソフトウェア開発	<input checked="" type="checkbox"/> カーネル開発
<input type="checkbox"/> Windows 互換/相互運用	<input type="checkbox"/> ゲーム	<input checked="" type="checkbox"/> Advanced Server

図5 Linuxパッケージグループ

協議の上、サーバ構成の決定を実施した。

ただし、本番用ストレージXP 128の設定を実施できるのはHP社だけであり、しかもXP 128のディスク構成は容易には変更できない。そのため、Oracle9i RAC+ストレージXP128 + Fibre Channel Switchの設定に関しては、サーバ構成設計の時点で最終確定させておく必要があった。

アプリケーション開発者に本稼動5年後でのデータ容量を割り出してもらい、1TBのディスク構成をサーバ構成設計段階で最終確定させた。

5. システム保守

V社向け基幹システムは2004年4月のリリース以降、特に大きな問題は発生していない。V社からもパフォーマンス的にもまったく問題ないとの報告を受けており、Linuxにおける性能、信頼性、保守性が一段と向上してきているのは事実のようである。

ただし、安定稼動しているディストリビューション、バージョンはその時点だけのものと考えるべきである。V社向けシステム開発当時(2003年9月)から現在(2004年8月)で、既に150以上ものパッチモジュールが発表されている。

本稼動後のパッチモジュール適用に対応していかなければ、安定性、信頼性を確保することはできない。しかし、パッチとファームウェア、ドライバとの相性問題もある上、オープンソースの進化のスピードについていけないパッケージソフトが存在するため、パッチを適用しても問題が起こらないか必ず事前に確認しておく必要がある。

構築時はTCOの削減を実現できたが、保守・運用費だけに関して言えば、オープンソースを採用する前と比べてコストメリットはあまり見られていない。如何に最適なパッチモジュールを効率良く、コストをかけずに適用し、さらなるTCOの削減を実現させていけるかが今後の課題である。

現在では、Linux用のパッチ自動適用パッケージソフトなども出始めているので、それらを採用するののも一つの手段であるが、安全性を確保してパッチを適用するならば、事前動作確認、最適なパッチモジュールの選定作業は外すことはできない。

V社向けシステムでは、以下のような段階を踏んで動作検証を実施し、確実かつ安全性を確保してパッチを適用していく予定である。

(1) 開発環境の各レイヤのサーバごとにパッチを適用し動作

確認を実施。

- (2) 冗長化したサーバを本番稼動とは切り離し、順次パッチを適用し動作確認を実施。
- (3) 本稼動中に適用できないものは、計画停止時にパッチを適用し動作確認を実施。

6. おわりに

Linuxシステムの採用、および複数ベンダー製品を含めてサービスを提供するためには、十分な事前検証、性能評価を行うことが大切である。この工程なしでは、今回のLinux、WebLogic、Oracle9i RACでの大規模基幹システム構築の実現も難しかった。それだけLinuxの進化は早く、最新のハードウェア、ソフトウェア、ミドルウェア、ファームウェア、ドライバなどの相性をその時々で確認していく必要がある。

それでも、TCOの削減、システムの拡張、縮小を容易にし、ハイスペック機種が採用できるLinuxシステムは魅力的であり、今後も積極的にLinuxの採用を進めていきたい。

当社は今回の大規模基幹システムの構築実績をもとに、今後、HP社と協業で大規模基幹システムの構築と運用サービスの拡販を目指していく。



山崎 貴弘

Takahiro Yamazaki

- ・システム開発事業本部
第二システム開発部
- ・Linuxシステム構築をはじめとするWEBシステム業務全般に従事



小池 洋介

Yousuke Koike

- ・システム開発事業本部
第二システム開発部
- ・Linuxシステム保守業務に従事