

# Vectorwise と Hadoop を使用した ビッグデータ・ソリューション



## ～ ポーランド最大のSNSによる会員情報・Webアクセスログ分析 ～

「様々な分析データベースを比較した結果、パフォーマンスと費用対効果の観点から Vectorwise を選択しました。」

リサーチ・データウェアハウス部門 シニアプロジェクトコーディネータ Edward Mezyk

### 顧客：

NK (Nasza Klasa)

### 業種：

ソーシャルネットワーク  
サービス

### 課題：

3テラバイトのデータに対して、各種の検索を高速で実行し、その分析結果に基づいて、会員に様々なサービスを提供することにより、収入の拡大を図ること。

### ソリューション：

Vectorwise を使用し、単一のサーバー上に、高速・低価格な新しいデータウェアハウスシステムを構築し、それを巨大な Hadoop システムと統合した。

### 成果：

- ・大量データを対象に、毎日 50～90 種類の検索を実行
- ・1 秒未満の検索レスポンスタイムを達成
- ・ビッグデータに基づいた現実的かつ説得力のある意思決定を実現

### 概要

NK (NK.pl) は、ポーランドにおけるソーシャルネットワークサービスの一つで、学生や卒業生がオンライン上で集うことができるサービスです。2006年にポーランドWroclawのハイテクエリアに住む24才のコンピュータ専攻の学生、Maciej Popowiczによって始められました。NKは、“Nasza Klasa”の頭文字で、“我々のクラス”という意味です。このサイトは、イギリスのFriends ReunitedやフランスのCopains d'avant、アメリカのClassmatesやFacebookと同等のもので。しかし、ポーランドにおいて、NK.plはFacebookよりはるかに人気があり、Facebookの650万アクティブユーザ（ポーランド内）に対して、1300万以上のアクティブユーザを誇っています。Googleの検索ワード年間アクセスランキングによれば、NKは2008年、世界で7番目の急上昇検索ワードでした。NK.plの成功は、高品質のソーシャルネットワーク機能と多様なオンラインゲームの提供、そして的を絞った広告機能などによるものです。

NK.plは、ポーランドで最も広く利用され、かつ浸透しているソーシャルネットワーキングサイトです。大規模かつ多様なユーザを満足させるための各種の機能を提供すると共に、広告主が高い投資効果を実現するため、広告効果に関する様々な分析サービスも提供しています。常に変化する会員の好み、興味分野、要求を最大限把握するために、NKのプロダクトマネージャーは、様々なデータソースからクリックデータ、ページ遷移データ等を収集し、最も使用頻度の高い機能、最も人気のあるゲーム、あるいは広告のクリック経路・頻度といった分析を行っています。この分析により発見された様々な事実、傾向は、サイトの機能改善、広告収入の拡大、あるいは新サービスの提供などに活かされています。

### 課題

現在のデジタル時代におけるトレンドや好みは、極めて早いスピードで変化しています。これまで、NK.plは、会員の好みの変化やWebページの遷移の持つ意味について十分理解することができませんでした。NKが使用していたそれまでのトランザクションデータベースは、1300万の会員がアクセスする250TBのデータ量で、そのデータを検索しても現実的なパフォーマンスを得ることができなかったからです。データは1,500台のサーバーに渡っており、プロダクトマネージャがこのデータに対して検索を実行すると、レスポンスタイムが3週間から4週間もかかっていた。このようなレスポンスタイムは、NK.plにとって全く受け入れられるものではありません。検索結果を受け取った時には、もはや過去の分析結果になっており、意味がないからです。ソーシャルネットワークで提供するサービス内容、会員の得られるメリット、そしてNKのマーケティング活動を最高、最良のものにするためには、シンプルではあるものの大量データを対象にした検索を短時間に実行し、その日のうちに分析することが必要でした。しかも、極めて費用対効果の高い方法で実行することが必要でした。

## ソリューション

検索のレスポンスタイムを向上させるため、NK.plはデータウェアハウスを構築しました。トランザクションデータベースに格納されたデータに対して新しい検索システムを追加、構築しました。データウェアハウスは大きく二つに分かれます。一つは、大容量データの格納・分析システムで、Hadoopが採用されています。Hadoopでは、POSTやGETリクエスト、AdServerログなどといったWebアクセスログデータを格納、分析します。もう一つは、単一サーバーで構成される高速分析システムで、Vectorwiseが採用されています。

Vectorwiseがデータ分析基盤に統合される前の従来のデータウェアハウスは、検索のレスポンスタイムが非常に遅く使いものになりませんでした。しかし、現在では、Vectorwiseを使用することにより、最も重要な検索であっても“思考を妨げないスピードでの検索・分析”をすることができるようになりました。まず、トランザクションデータベースからデータを抽出して巨大なHadoopデータストレージを構築し、そこから更にデータを抽出し、Vectorwiseを基盤とした高速検索システムを構築しました。「様々な分析データベースを比較した結果、パフォーマンスと費用対効果の高さからVectorwiseを選択しました。」と、NKのリサーチ・データウェアハウス部門のシニアプロジェクトコーディネータ、Edward MezykはVectorwiseを選択した理由をこのように述べています。

## 成果

現在、NKはサーバー群から毎日1TBの追加データを収集しています。そのデータには、ユーザの属性、メッセージ、コメント、写真などが含まれます。プロダクトマネージャは、そこから更にVectorwiseに抽出、格納された大量データを対象に、毎日50～90の検索を実行していますが、レスポンスタイムは1秒未満です。それにより、素早くデータを分析し、現実的なアクションにつなげることができるようになりました。事実に基づいた説得力のある戦略的な意思決定が可能になったわけです。NKはVectorwiseの新しい分析システムによりそのパワーを得ることができました。「私たちのビジネスは、情報とそれを分析する能力で成り立っています。Vectorwiseによってそれを得ることができました。それこそが我々の現在と将来の収入を約束するものと言えます。」と、MezykはVectorwiseを評価しています。

## Actian社

Actian社(旧:Ingres社)は、データマネジメント製品を提供する先進的なソフトウェア会社です。ミッションクリティカルなトランザクションデータベース Ingres DB、大量データの分析を高速で行う革新的なデータベース Vectorwise、クラウドベースのアプリケーション環境 SkySafe の3製品を提供しています。Actian社の製品はいずれも圧倒的に高い価格性能比が評価されています。顧客は世界中で1万社を超えています。

## 株式会社ニューシステムテクノロジー

株式会社ニューシステムテクノロジーは、Actian社製品の日本総代理店です。パッケージ販売をビジネスの中核とし、セキュリティソフトウェアからユーティリティ、データベースへと品揃えの拡充を図っています。

国内総代理店



株式会社 ニューシステムテクノロジー  
〒105-0004 東京都港区新橋2-12-17  
新橋INビル7階

TEL: 03-3597-0031 FAX: 03-3597-0032  
<http://www.kknst.com> mail:info@kknst.com

パートナー