



## 2009 年度下期未踏 IT 人材発掘・育成事業 採択案件評価書

### 1. 担当PM

藤井 彰人 PM(グーグル株式会社 エンタープライズ プロダクト マーケティング  
マネージャー、Mashup Awards 1-4 主宰)

### 2. 採択者氏名

チーフクリエイター:高木 潤一郎(東京大学大学院 工学系研究科電気系工学専攻  
融合情報学コース 修士1年)

コクリエイター:なし

### 3. プロジェクト管理組織

株式会社ゴーガ

### 4. 委託金支払額

5,000,000 円

### 5. テーマ名

時系列センサデータベースシステム TIVA の開発

### 6. 関連Webサイト

<http://tjun.jp/tiva/>

## 7. テーマ概要

本プロジェクトでは、複数のセンサからのセンサデータを長期間蓄積し、さまざまなセンサデータの組み合わせや過去のセンサデータの参照を可能とすることで、多様なアプリケーションから利用可能となるセンサデータベースシステム TIVA の開発を行う。

TIVA は、時間の経過とともに増加し続けるセンサデータを蓄積するため、システムがいっぱいになった場合にもノードの追加によって容易に拡張できる機構を持つ。また、アプリケーションからセンサデータを利用する際によく使われる時間範囲検索を高速に実現する。さらに、MySQL のストレージエンジンとして実装し、オープンソースで公開すること目標とする。

小規模・低コストで始めることができ後から拡張可能なシステムと、MySQL のインターフェースを提供することで、センサデータを扱う多くの研究者や開発者、また中小企業やスタートアップのベンチャーに利用して貰えると考えている。本ソフトウェアの実現によって、センサデータを用いたアプリケーションの開発を促進し、実空間の情報を使った新たなサービスやビジネスが創出されることを期待している。

## 8. 採択理由

Streaming Database や時系列 Database については、学術的な先進性はないものの、センサデータの効率処理は今後の IT における大きなテーマであり、適切なテクノロジーの普及が新たなサービスの可能性を広げると考える。提案者のニーズに即した提案と、MySQL ストレージエンジンへのマッピングの提案は、現実的かつ興味深く、未踏として採択したい。

## 9. 開発目標

本プロジェクトの目標は、全てのセンサデータ、つまり時刻と値を持つ時系列データを蓄積して、そのデータをアプリケーションから利用可能とするセンサデータベースシステム TIVA を実装することである。内部のデータ構造から、カラムの値に注目したデータ圧縮技術、Sparse Index と呼ばれる索引付け機能など、センサデータに特化した高速化技法を多数実装することを具体的な目標として設定した。

- (1) シンプルなデータ構造でデータを蓄積
- (2) 既存のデータベースより高速
- (3) SQLによる問い合わせ
- (4) MySQLのストレージエンジンとして実装

センサ情報の活用を促し、Webアプリケーション開発者等にも幅広く利用してもらうため、My SQL の代替としての実装も視野にいれることとした。

## 10. 進捗概要

### 開発初期

開発内容の確定と、開発ゴールの設定。

既存Key Value Storeとの性能比較。

### 開発中期

データ構造の見直しと、高速化手法の検討。

### 開発後期

SQL like な query interface の実装と、さらなる Performance Tuning

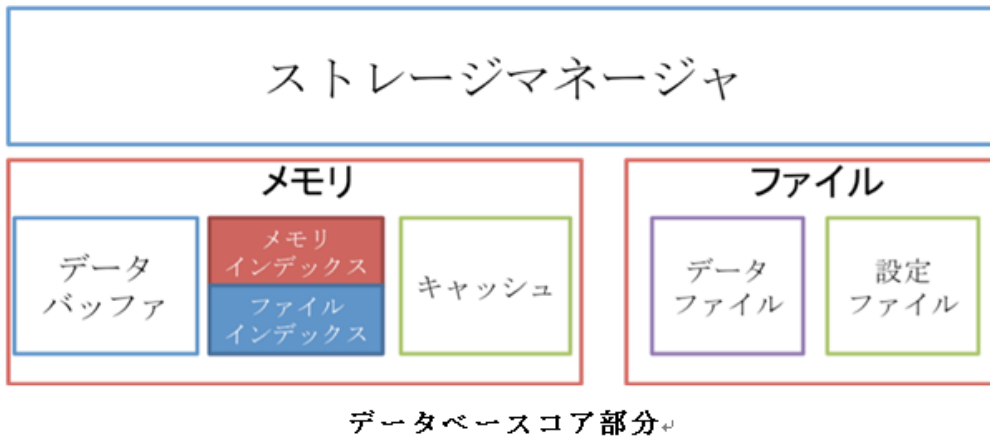
## 11. 成果

時系列データである、センサデータを効率よく、かつ高速に扱うためのデータベースを実装し、センサデータにおける、データ挿入、検索性能を他の代表的なデータベースと性能比較を行った。

詳細は、TivaのWeb page <http://tjun.jp/tiva/> を参照されたい。現在、本ソフトウェアを説明するデモビデオが公開されている。

本プロジェクトでは、データベースのコア部分から、Terminal Interface、SQL interfaceまで、全ての機能を開発した。

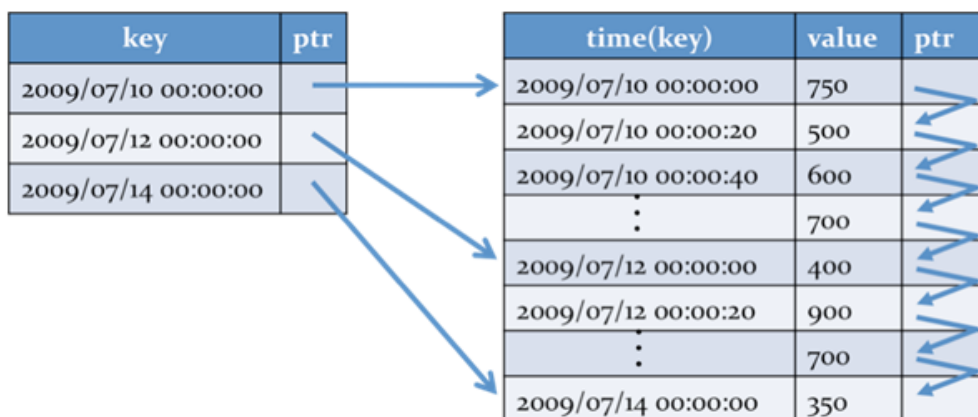
TIVA のシンプルなデータ構造と高速なデータ処理を実現するデータベースコア部分とDBMとしてのシンプルなインターフェース部分、そしてSQLライクなインターフェースを実現する部分と、そのSQLライクなインターフェースをターミナルから操作するターミナル対話型インターフェース、の4つの部分に分けられる。



TIVA のコア部分は、メモリ、ファイルシステムを含めたデータベースのコア部分である。TIVA の特徴として、センサデータの特徴である時刻と値のペアに着目したデータベースコア部分の設計実装、Sparse Index、非同期更新、SQL like なインターフェースの実装などを行っている。

●Sparse Index の実装

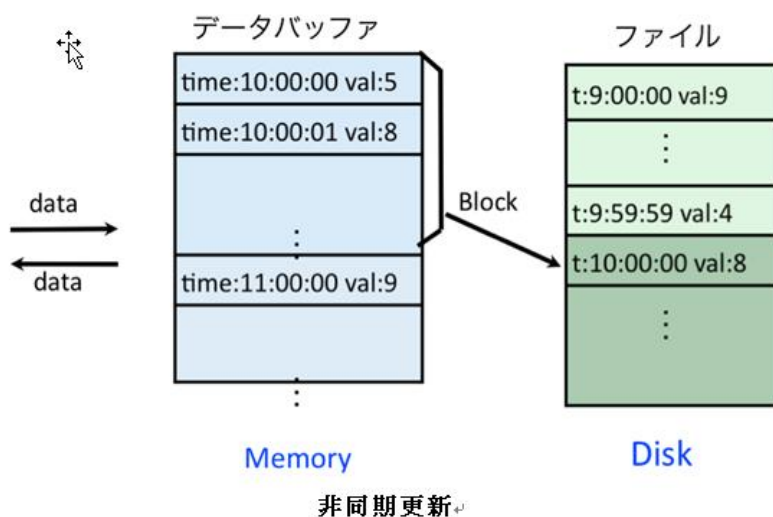
TIVADB における索引付けは、Sparse Index という手法を用いている。対象データが時系列データであるため、通常のデータベースのように全てのレコードに対して索引付けを行うのではなく、ブロック毎に索引付けを行う。TIVADB では全てのデータが時系列に並んでおり、キーである時刻が予め整列しているため、全てのレコードに対して索引付けを行う必要がない。



Sparse Index

- 非同期更新の実装

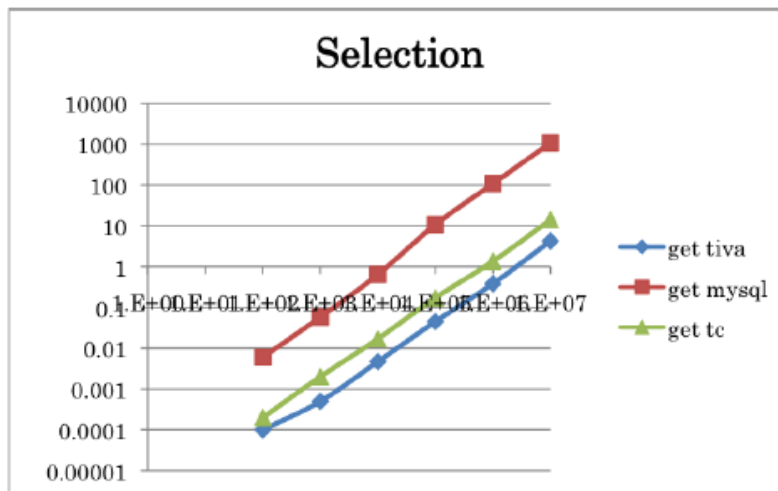
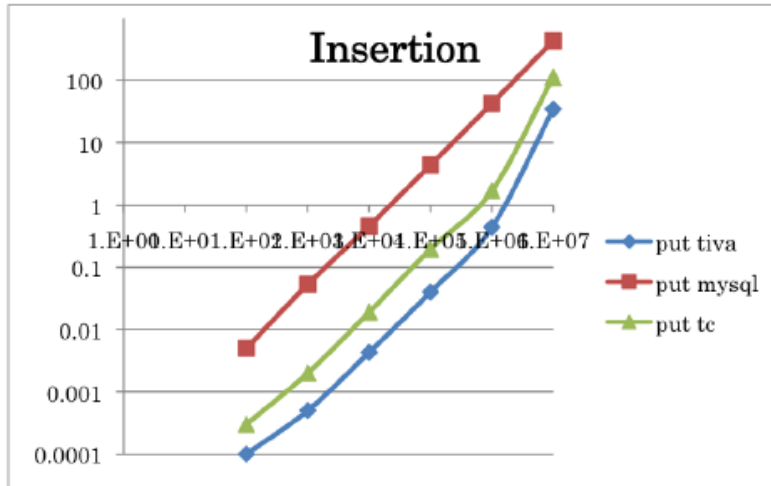
TIVA では、センサデータの特徴を勘案し、プログラムから挿入された新規データをそのままファイルに書き込まず一時的にメモリに保持し、一括して書き込むという非同期更新を行っている。



- SQL likeなインターフェースの実装

TIVA では、ユーザは GET や PUT といった関数だけでなく、SELECT や INSERT といった SQL ライクなクエリでデータベースの利用を行うことが可能である。

性能評価に用いたクエリは 100 件から 1000 万件までのデータの挿入と、1000 万件レコードを登録した状態での 100 件から 1000 万件のレコードの検索である。MySQL 及び、Tokyo Cabinet との比較でも高い性能を記録している。



## 12. プロジェクト評価

センサデータを効率よく管理また簡単に利用可能なソフトウェア基盤は、各種センサの携帯電話やスマートフォンへの実装が進む中、新たなサービスを開発する意味においても、その重要性を増している。過去にもStreaming Databaseをはじめとする研究が行われているが、MySQLのようにだれもが気軽に利用できるセンサデータベースの開発は、現在各種センサがインターネットに接続される現在こそ取り組むべき課題である。

センサデータを効率よく管理できるデータベースの必要性を自身の研究経験から痛感し、センサデータに特化したデータベースを開発するという取り組みは、一見地

味ではあるが、未踏ソフトウェア事業らしいITのコア技術の研究開発であり、デバイス  
大国日本らしいプロジェクトと言える。

高木氏は、開発目標としたセンサデータベースを、一から、しかも一人で開発する  
というプロジェクトを遂行し、当初の目標であるセンサデータベースを開発した。セン  
サデータの特徴に着目し、基本的なデータ構造から組み上げ、sparse indexや、各  
種高速化のための実装も行い、SQL likeなインタフェースも実装したことは、高く評価  
したい。用途特化型とは言え、特に他製品 MySQL、Tokyo Cabinetよりも高速な  
実装を本短期間の開発中に実現できたことは大きな成果である。

MySQLと同様な使い勝手を当初の目標としていたが、まずは同様製品よりも高速  
なデータベースコア実装を目指したため、本内容の実装は本未踏開発期間中には行  
わないこととした。

### 13. 今後の課題

データベースとして、今後、センサデータに特化した更なる機能実装、性能アップを  
期待したい。

センサデータを活用した革新的なサービスの開発には、多くの開発者のアイデアや  
試行錯誤が必要である。このためにも開発者がセンサデータを気軽に使える環境が  
重要である。TIVAをその基盤技術として普及させることも大きな課題として認識し、  
今後の啓蒙活動を行って欲しい。

#### 1)センサデータに特化した更なる機能実装

アイデアとして提示されていたイベント毎のインデックス実装や、データの圧縮技術  
の実装等、センサデバイスデータの特性に応じたデータベース機能のさらなる拡充  
が望まれる。

#### 2)さらなるコンパクト化と高速化

未踏開発期間中には、調査が十分に行えなかった、データベースの高速化技法や  
ファイルシステム等の研究を行い、更なる高速化、コンパクト化を期待したい。

#### 3)MySQLストレージエンジンへのマッピング

当初目標としていた、MySQL ストレージエンジンへのマッピングは、今後も引き続  
き開発目標として欲しい項目である。基盤ソフトウェアの発展には、利用者の増加  
が発展のkey pointとなるため、コア技術同様にユーザーニーズを調査し「使っても  
らえる」技術の開発が今後の課題となる。