

小口研究室 研究紹介 (2014年度)

(お茶の水女子大学理学部情報科学科)

ファットクライアントを利用した動画データ解析アプリケーションフレームワークの提案 (研究担当: 黒崎 裕子)

目的

- 一般家庭にも気軽にセンサが搭載可能に
- 一般家庭でのサーバやストレージの設置解析まで行うことは困難

クラウドの普及

クラウド上での問題点

- ネットワーク帯域
- クラウド資源の枯渇

クラウド(データセンター)

ファットクライアントを利用した動画データ解析アプリケーションフレームワークを提案、実装

センサ側端末で特徴抽出等の前処理を行う

- データ量の大きい動画データデータを小さくすることで、クラウド側端末の負担を軽減

Apache Stormを導入

- 負荷の高い処理を適宜平衡実行して高速化を図る

システム概要

リアルタイム分散処理システムApache Stormを採用

- 自動的にタスクの割り当てが行われる
- 実装が楽である

クラウド側の負担大

クラウド側計算機で前処理を行う場合

センサ側計算機で前処理を行う場合

アプリケーション概要

ファットクライアントモデルを採用し、動画データ解析アプリケーションを実装

リアルタイム分散処理システム: Apache Storm

リアルタイム解析フレームワーク: Jubatus

(1)WEBカメラからの画像取得

(2)Bag-of-Featuresを用いた特徴抽出

画像からSURF特徴量を抽出し、ただ1つの特徴ベクトルに変換

(3)Jubatusでの解析

Classifier APIを用いて受け取った特徴ベクトルを分類

Apache StormとJubatusを用いて実装したフレームワーク

実験結果

提案アプリケーションに関する2つの実験

①Apache Stormのアーバードの有無

②処理スレッドの増減による処理速度の変化

Apache Stormは自動負荷分散されるため、今後はパフォーマンスの向上が期待できる

ストリーム制御されたか、イベントが破棄された可能性が高く、今後詳しくデータ解析する必要あり

まとめ

- クラウドへの集中負荷を防ぐため、ファットクライアントモデルを採用した解析フレームワークを提案、実装
- Stormのオーバーヘッドは無視でき、複数センサを処理する場合、クラウド側サーバへの負荷の集中の回避の効果が期待
- 今後は複数センサを扱うフレームワークの実装、Apache Storm、Jubatusの分散解析処理環境構築

緊急地震速報に基づいたハイブリッドクラウドにおけるOpenFlowによるネットワークトラフィックの最適化 (研究担当: 原 瑠璃子)

研究背景

- センサ技術の普及
- 携帯型デバイスの発展
- クラウドやDCの増加

ビッグデータへの対応が重要となってきた

パースト的な負荷変動

大地震発生

情報処理基盤

安否確認

避難経路の確保

ビッグデータ処理は短時間で大きなシステムの負荷が起こる

研究目的

OpenFlowを用いた投機的なネットワーク制御を提案

従来のネットワーク制御

- 一般的に緩やかな負荷変動を念頭に検討されている
- 短時間で激しく変化するパースト的な負荷変動に追従することは難しい

通常システム

バックアップ退避

バックアップシステム

約30秒

大地震発生

実験概要

外部情報の取得

気象庁がTwitter上で発信している緊急地震速報をモニタリングし、地震情報を取得

本研究では、疑似的に緊急地震速報と同じ内容のツイートを流して実験を行う

まとめと今後の課題

重要なデータを迅速にバックアップする仕組みを提案した。実験環境としてOpenStackを用いてクラウド環境を構築し、OpenFlow制御における効率的なバックアップ処理を行う評価を行った。

今後の課題

- 今後はクラウド間でのバックアップ処理を検討し、評価を行いたい。

性能評価

緊急地震速報が送られてきたら、これをトリガとして、データファイルのバックアップ処理を下記のケースで開始する。

- バックグラウンドの負荷をそのままにした場合
- インスタンス間でiperfでトラフィックを発生させた場合
- インスタンス間でファイル転送を行った場合

通信制御時

背景あり (iperf)

背景あり (FTP)

異なるインスタンス間

通信制御時

背景あり (iperf)

背景あり (FTP)

データベース同期ミドルウェアによる遠隔バックアップ活用手法の検討 (研究担当: 細谷 柚子)

研究背景

- バックアップのパフォーマンスに課題

データ保護・管理対策としてバックアップの実施が約90%

1/3以上が現状に不満

遠隔バックアップの需要高まる

ローカルのリソースだけでは大規模災害時にバックアップごと失う

研究目的

DBリモートバックアップ確立のパフォーマンス向上

- サービスのためのトランザクション処理に影響を与えない
- バックアップを最新に更新できる

DBバックアップ確立ミドルウェアの検討

既存研究:Pangea

NTT 三島 健氏

LAN環境を前提としたDB同期ミドルウェア

- 更新処理: 全てのサーバ
- 照会処理: 1台のサーバ

提案手法:Pangea++

バックアップ処理を切り離し、更新処理のみ実行

更新処理を保管

commit契機で呼び出す

Synchronizer スレッド

Leader

Follower

Backup

評価実験

PangeaとPangea++の性能比較

- ベンチマーク:TPC-W(1200s実行)
- 仮想的なブラウザ(EB)がWebサーバに対してトランザクション処理を要求
- 遅延:RTT256ms(日米間、日欧間)

実験環境

	middleware	server
Local	Pangea	Leader1台,Follower1台
RTT256ms	Pangea	Leader1台,Follower1台(remote)
Pangea++	Pangea++	Leader1台,Follower1台,Backup1台(remote)

TPC-W workload

	read-only	update
browsing mix	95 %	5 %
ordering mix	50 %	50 %

実験結果

browsing mix

ordering mix

性能差

	throughput		response time	
	max	min	max	min
browsing mix	-20.79 %	+0.4 %	+1.9 秒	+0.04 秒
ordering mix	+31.77 %	+14.86 %	-2.9 秒	-1.21 秒

実験考察

- Pangea++ではwriterset保管がオーバーヘッド
- 更新処理が多い場合にはPangea++が有効
- 照会処理が多い場合にはPangeaが有効

今後の課題

- 高負荷時にwritersetが輻輳すると考えられる
- writersetのキューを溜めずに処理を行う
- バックアップにも処理を分散させる