

～2010年度 活動紹介セミナー～
クラウド・テクノロジー分野
(Hadoop、NoSQL等)

アドソル日進株式会社
荒本道隆

- 多種多様なクラウド関連技術がオープンソースや製品として出てきているが、「クラウド」という言葉の定義自体が幅広いいため、企業内システムでの適用できる領域・効果が分かりづらい。
- そこで、クラウド関連技術の幅広い情報収集をおこなって、調査結果を情報共有するとともに、Hadoop やNoSQL などの中で利用可能なものを実際に試用し、それらを使ってプロトタイプを試作してみることで、クラウドを企業内システムで活用するヒントや具体的なイメージを得るための活動を行う。

- IaaS (Infrastructure as a Service)
 - 仮想マシンやネットワークなどのインフラをサービスとして提供
 - 例: Amazon EC2, VMware
- PaaS (Platform as a Service)
 - OSやアプリを実行するためのミドルウェアをサービスとして提供
 - 例: Salesforce, Google App Engine, Windows Azure Platform
- SaaS (Software as a Service)
 - ソフトウェアの必要な機能だけをサービスとして提供
 - 例: 電子メール, 検索サービス, オンラインゲームなど
- MapReduce
 - 今までの規模を超えたスケールアウト
 - 例: Hadoop, BigQuery
- NoSQL
 - KVSに代表される、RDBとは異なるアーキテクチャのデータストア
 - 例: HBase, memcached, Cassandra, CouchDB, kumofs, kai, Flare, ROMA

社内調達とクラウド利用の比較

社内調達	IaaS	PaaS	SaaS
アプリケーション	アプリケーション	アプリケーション	アプリケーション
OS・ミドルウェア	OS・ミドルウェア	OS・ミドルウェア	OS・ミドルウェア
ハード・インフラ	ハード・インフラ	ハード・インフラ	ハード・インフラ



自分で面倒を見る範囲が
広がるので、管理コスト
が増える

視点の違いー1

- サービスを提供する側
 - 管理工数を削減するための仕組み作り
 - サービスレベルの維持
 - ビジネスモデルの確立
- サービスを利用する側
 - 自由にやりたいけど、細かい部分の面倒は見たくない
 - 低コスト・高品質・継続性
 - 使いたい時に、すぐに使いたい

- 想定する規模(利用者数、データ件数、容量)
 - 社内を対象
 - 国内の特定ユーザーを対象
 - 国内の不特定多数を対象
 - 世界の不特定多数を対象
 - 人間以外を対象(機器、センサー、動物、などなど)

活動内容

- 情報収集、事例研究（セミナー、勉強会）
 - クラウド関連の実装や技術に関する情報を収集する。
- クラウド関連の実装の試用（勉強会）
 - 利用可能な実装を試用してみる。
- プロトタイプ作成（部会）
 - 企業内システムを前提としたプロトタイプを作成する。

セミナーの開催

セミナーの開催－1

- クラウド関連の実装や技術、事例の紹介
 - 参加者が講師となる or 外部から講師を呼ぶ
 - 是非、講師としての参加をお願いします
 - 人に説明するための準備が、最高の勉強になります
- 目的
 - クラウド関連の動向調査と、勉強会で扱うテーマ探し

セミナーの開催－2

- 『第二回 先端IT活用シンポジウム』のスタイル？
 - － 2010/07/26(月) クラウドをテーマに開催済み
- 皆さんに質問
 - － 聞いてみたいこと
 - － 自ら説明したいこと
 - － 紹介できそうな事例

勉強会の開催

- 実際にクラウド関連の技術を使って、体験してみる
 - 既存の関連技術を知っていれば、クラウドの部分は少ないかも
 - 例：Servlet/JSP → Google App Engine
- 体験した時の気づきを共有する
 - 自分の業務領域外の気づきも得ることができる

- 皆さんに質問1：クラウド関連で、やってみたい領域は？
 - IaaS：Amazon EC2, VMware
 - PaaS：Salesforce, Google App Engine, Windows Azure Platform
 - SaaS：
 - MapReduce：Hadoop, BigQuery
 - NoSQL：HBase, memcached, Cassandra, CouchDB, kumofs, kai, Flare, ROMA
- 皆さんに質問2：PCを持ってくることは可能ですか？
 - USBメモリやInternetに接続しても良いか？
 - アプリの実行(できるだけ、インストーラーではなく、コピーで動くものを使用)
 - メモリ1GByte以上
 - 例：Hadoopはデフォルト状態で2.8GByte使用
 - OSの種類

例えばHadoopを使った開発

HadoopのJob開発

- Eclipseによるデバッグ(ローカルモード)

The screenshot displays the Eclipse IDE in a debug mode for a Java application named 'NewFileFilter'. The main window shows the source code of 'NewFileFilter.java' with a breakpoint set at line 159. The code snippet is as follows:

```
156 public void map(LongWritable key, Text value, Context context)
157     throws IOException, InterruptedException {
158
159     String line = value.toString();
160     log.info("Read line:" + line);
161
162     if (line == null || line.startsWith("#") || line.length() == 0){
163         return;
164     }
165     String srcfileFullpath = line;
166 }
```

The 'Variables' view on the right shows the state of the current thread:

Name	Value
this	NewFileFilter\$NewFileFilterMapper (id=37)
key	LongWritable (id=39)
value	Text (id=46)
context	Mappers\$Context (id=48)

The 'Console' window at the bottom shows the following log output:

```
NewFileFilter [Java Application] C:\Java\jdk1.6.0_16\bin\javaw.exe (2010/09/14 9:19:43)
1533 [main] INFO org.apache.hadoop.mapred.JobClient - Running job: job_local_0001
1544 [Thread-10] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 2
1714 [Thread-10] INFO org.apache.hadoop.mapred.MapTask - io.sort.mb = 100
1925 [Thread-10] INFO org.apache.hadoop.mapred.MapTask - data buffer = 79691776/99614720
1925 [Thread-10] INFO org.apache.hadoop.mapred.MapTask - record buffer = 262144/327680
2555 [main] INFO org.apache.hadoop.mapred.JobClient - map 0% reduce 0%
7656 [communication thread] INFO org.apache.hadoop.mapred.LocalJobRunner -
8560 [main] INFO org.apache.hadoop.mapred.JobClient - map 51% reduce 0%
```

プロトタイプ作成

- クラウド・テクノロジー研究部会（仮称）
- 企業内システムを前提としたプロトタイプを実成する
 - 公開を前提としたプロトタイプの実成
 - サービスを公開 or ソースコードを公開
 - 開発の途中経過も合わせて、できるだけオープンにしたい
 - SourceForge.JP などを使って共有？
 - 成果物予定
 - 設計書、ソースコード、環境構築手順書
- 別部会との協業
 - 例：サーバ：クラウド・テクノロジー研究部会
クライアント：インターネットデバイス アプリケーション部会
- 作るものは、まだ未定です

まとめ

- 要検討事項
 - 開催時刻は、昼間 or 夕方
 - 会場
 - 内容
 - 前提知識
 - 講師