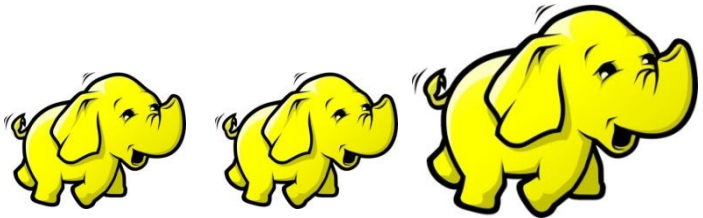
AITC
先端ＩＴ活用推進コンソーシアム

クラウド・テクノロジー研究部会

ビックデータを支えるクラウド技術
～今更聞けないHadoop入門(演習編)～

2013年01月26日

株式会社イーグル
菅井 康之



Agenda

- Hadoopの概要
  - お話したいことは多々ありますが、演習の時間を多く取りたいので手短に・・・

- 演習
  - Windows開発環境でMapReduceを実装
  - Windows上でスタンドアロンモードの動作確認
  - Linux上で擬似分散モード(1台)の動作確認
  - Linux上に構築したクラスタ環境での動作確認

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

ビッグデータ時代の到来

- 近年、情報通信の発達などにより、蓄積/収集するデータ量が爆発的に増大している
  - アクセスログ、RFIDデータ、モバイルセンサー、SNS投稿データ 等
  - 従来の常識を超える大量データ(テラ/ペタ/エクサバイト)

- 蓄積したデータを活用するためには高速処理するための基盤技術が必要となる
  - 高速処理の方向性には二軸ある
    - レイテンシ向上(低レイテンシ)
      - リアルタイム性の高いレスポンス(NoSQL、分散キャッシュなど)
    - スループット向上　←
      - 単位時間辺りの処理データ量の向上(Hadoopなど)

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

スループット向上

- 時間内に大量データを処理するニーズが増えている
  - アクセスログや購買履歴を解析し、レコメンドや広告表示
  - トラヒックデータを機械学習し、ネットワーク障害の予防に役立てる
  - WebをクロールしてHTMLをパース/スクレイピング
- スループットの向上には２つのアプローチがある
  - スケールアップ(性能向上-垂直)
    - Write時の整合性が保証しやすい。管理が容易
    - 高性能なサーバはある程度の性能から飛躍的に高価になる
      - 投資したコストと比例して性能が上がるわけではない
      - 性能向上には物理的な限界がある
  - スケールアウト(並列分散-水平)　←
    - 上手く並列分散出来ればコストをかけた分、性能を向上出来る
    - Readは早いがWriteが遅くなる傾向。管理が複雑

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

スケールアウト

- 例）1TBのデータを全て読み出す
  - 1台のドライブ(100MB/sの転送速度)で実行した場合
    - 1,000,000(MB) / 100(MB) = 10,000(s)
    - データを読み出すだけで2時間半以上かかる
  - 100台のドライブで並列に実行した場合
    - 1,000,000(MB) / 100(MB) / 100(台) = 100(s)
    - 並列分散処理することにより、スループットの向上を実現！・・・となれば何も苦労しない
    - 並列分散処理では越えなくてはいけない壁が多い
      - HDD障害、レプリケーション、ネットワーク負荷、データ結合、実行制御・・・

Hadoopの登場により、比較的手を出しやすい状態に
但し、Hadoopを使っても壁は残っている

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

Hadoopとは①

- 大規模データの分散処理を支えるJavaソフトウェアフレームワーク
  - Javaで実装されている
  - Streaming
    - Java以外の言語も利用可能(標準入出力を介してデータのやり取り)
- 分散に特化したアーキテクチャ
  - 分散する台数が少ないとむしろ遅くなる
    - オーバヘッドが大きい
- 大量データのバッチ処理に向いている
  - 少量のデータだとむしろ遅くなる
  - リアルタイム処理で使用するデータの準備などに利用
  - インデックスを使用しないフルスキャン向き
    - 一部のデータだけを利用する場合はRDBの方が速い可能性が高い

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

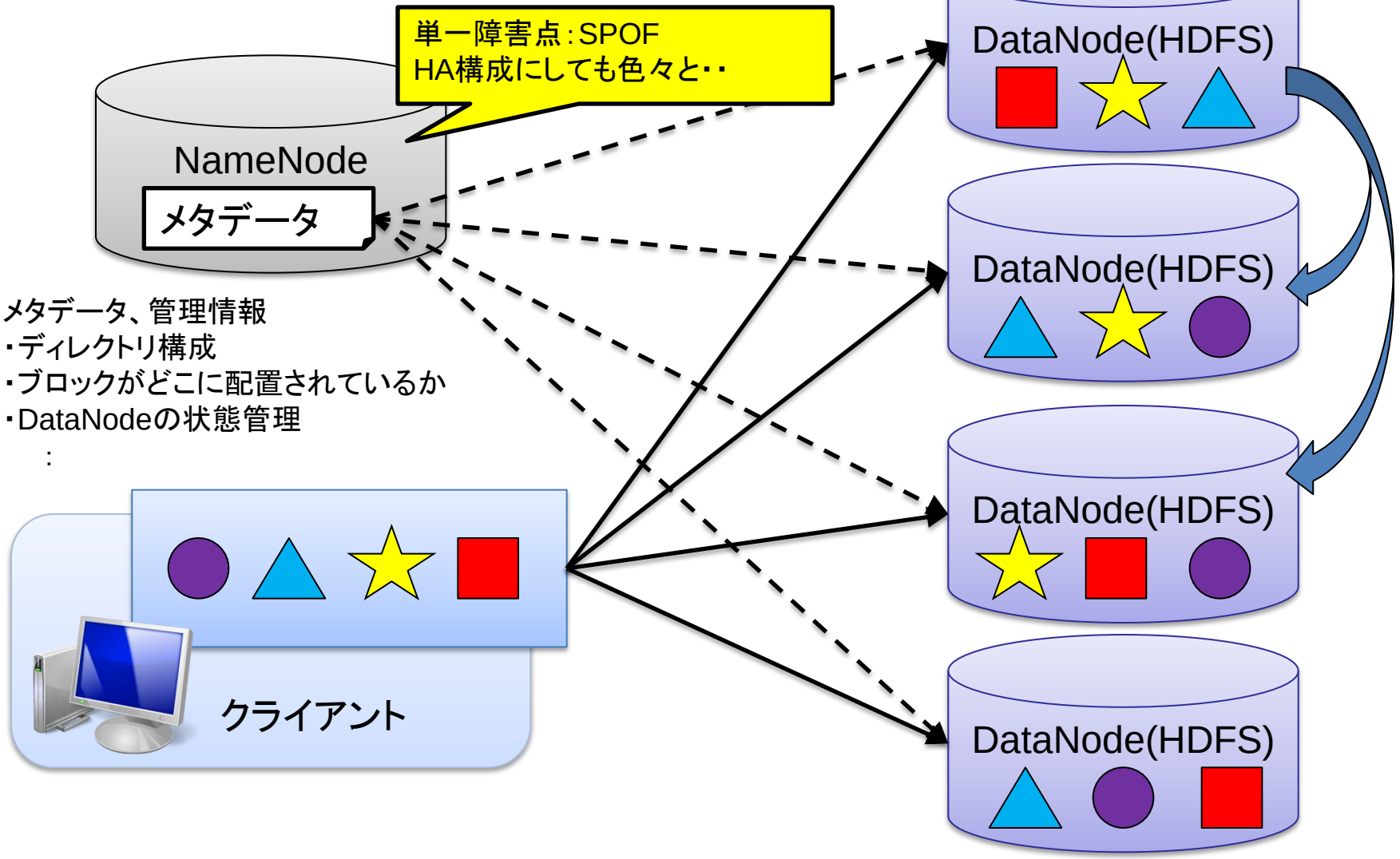
Hadoopとは②

- Hadoopの主要な構成要素
  - Hadoop分散ファイルシステム（HDFS：Hadoop Distributed File System）
    - 複数サーバにデータをコピーして配置
      - Read時に、HDDのI/Oを分散させるため
      - デフォルトは多重度が３→データノードは３台以上
      - HDFSはブロックサイズが64MByteなので、細かいファイルは苦手
        - harやHBaseを使って回避可能
  - Map/Reduceエンジン
    - 処理を分割して、複数サーバに分散
      - できるだけ、該当データを持っているノードに割り当てる
      - ローカリティを最大限に活かす
      - Mapだけで動作することも可能（手軽な分散処理）

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

HDFSの構成①

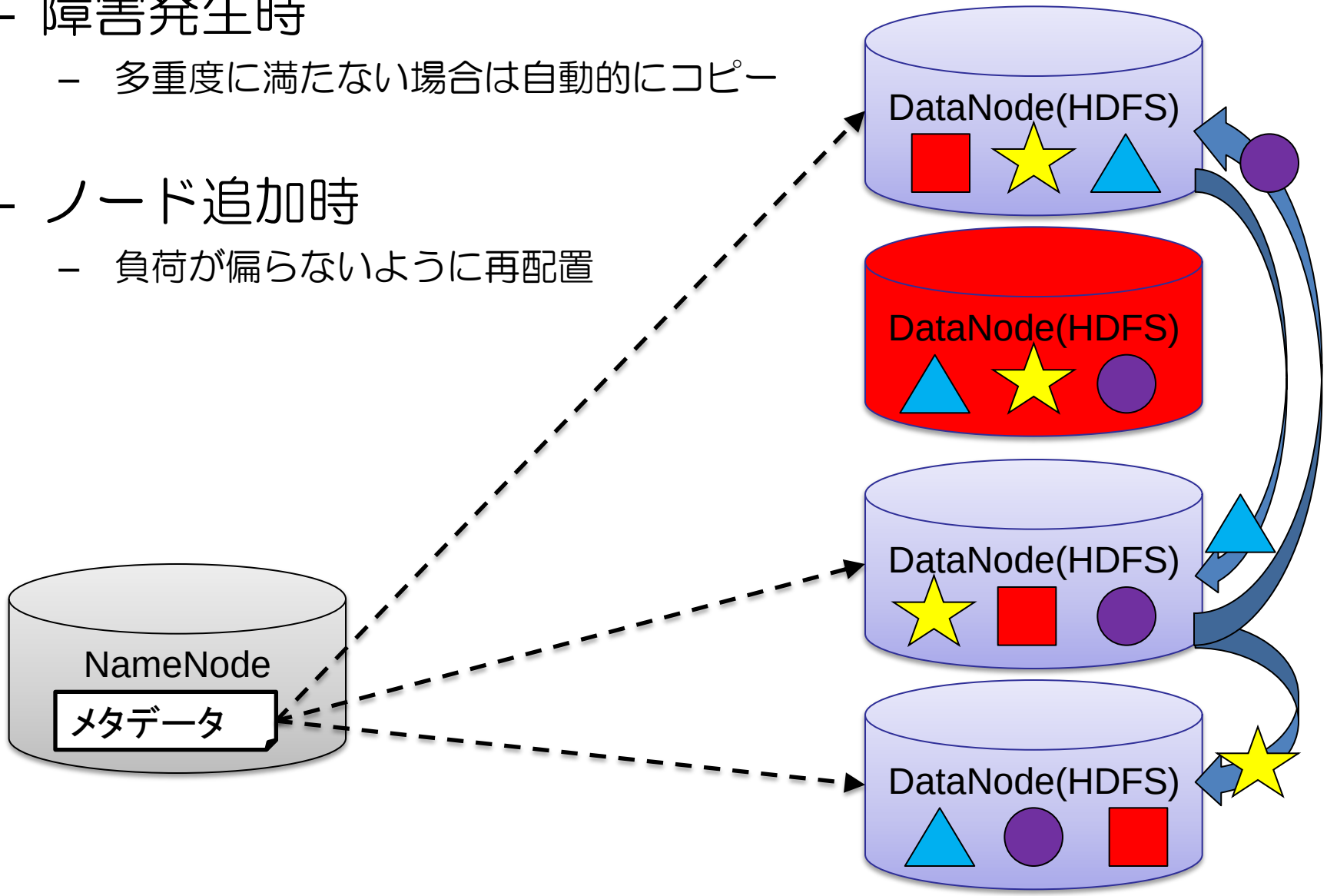
- 一つのブロックを複数のDataNodeで保持することで
  DataNodeの障害が発生してもデータが失われない



Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

HDFSの構成②

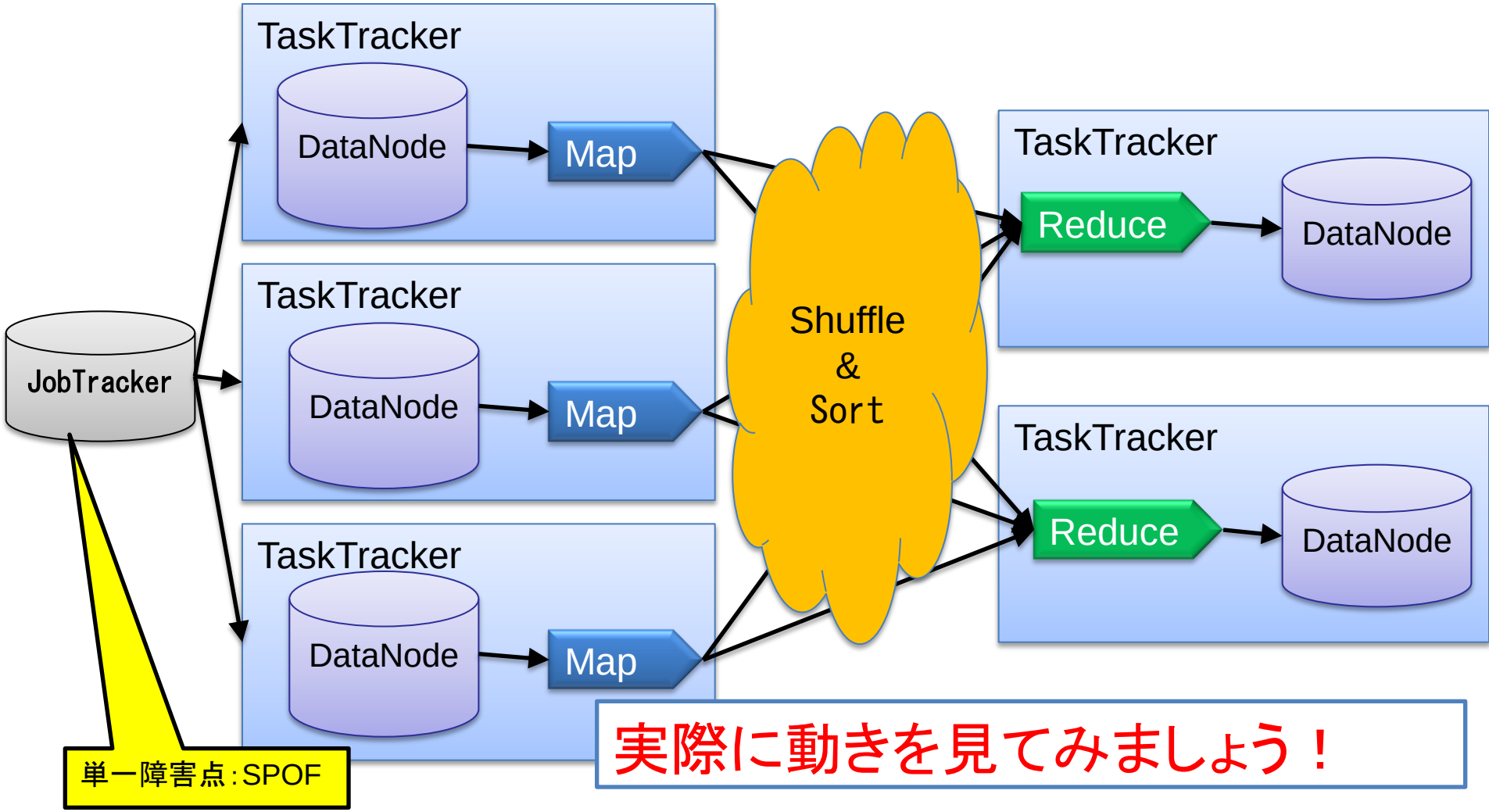
- 障害発生時
  - 多重度に満たない場合は自動的にコピー
- ノード追加時
  - 負荷が偏らないように再配置



Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

MapReduceの構成

- できるだけ、該当データを持っているノードにMapを割り当てる
- JobTracker = ジョブの実行制御, TaskTracker = ジョブの実行係



実際に動きを見てみましょう！

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習 環境について

- ハンズオン終了後は使用前に戻します
  - C:¥aitc で全て作業を行います
    - C:¥aitc
      - Eclipse
      - hadoop-0.20.2
      - Cygwin
  - ※hadoopのtmp dirとしてC:¥tmpが使用される(default設定)

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習α 環境構築①

- Windowsにて、Hadoopが動作する環境を構築する(デバッグ環境)
  - Java (http://java.com)
    - 1.6以降
  - Cygwin (http://cygwin.com)
    - HadoopはLinuxを正式サポート
    - Hadoop内部でchmodやbashコマンドを実行
    - Cygwinにパスを通すことで実行可能に
  - Hadoop (http://hadoop.apache.org)
    - バージョン0.20.2が一番無難
      - 0.20.2X=>1.0系, 0.23.X=>2.0系
    - 1.0以降はpermission処理の問題によりWindowsで動作しない
    - Hadoop自体をデバッグしたい場合は ant.jar(http://ant.apache.org/)が必要
  - Eclipse (http://www.eclipse.org)
    - HadoopにEclipseのpluginが含まれているが、使用しない
    - Deprecatedなソースが生成される。メンテナンスされていない模様

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習α 環境構築②

- 今回のハンズオンで使用する構成
  - Eclipse+Java
    - Javaがインストールされていないpcのため、JRE同封のEclipseを使用
    - http://mergedoc.sourceforge.jp
    - 余計なもの(Tomcat, JDK)が多いので、eclipseのみZIP圧縮
  - Cygwin
    - インストール時にレジストリを更新してしまうため、他のPCでインストールしたものをZIP圧縮
  - Hadoop 0.20.2
    - http://hadoop.apache.org/releases.html#Download
    - Download a release now!のリンクからミラーサイトを選択
    - 0.20.2が無いのでreadme.txtに記載のリンク
      - http://archive.apache.org/dist/hadoop/common/

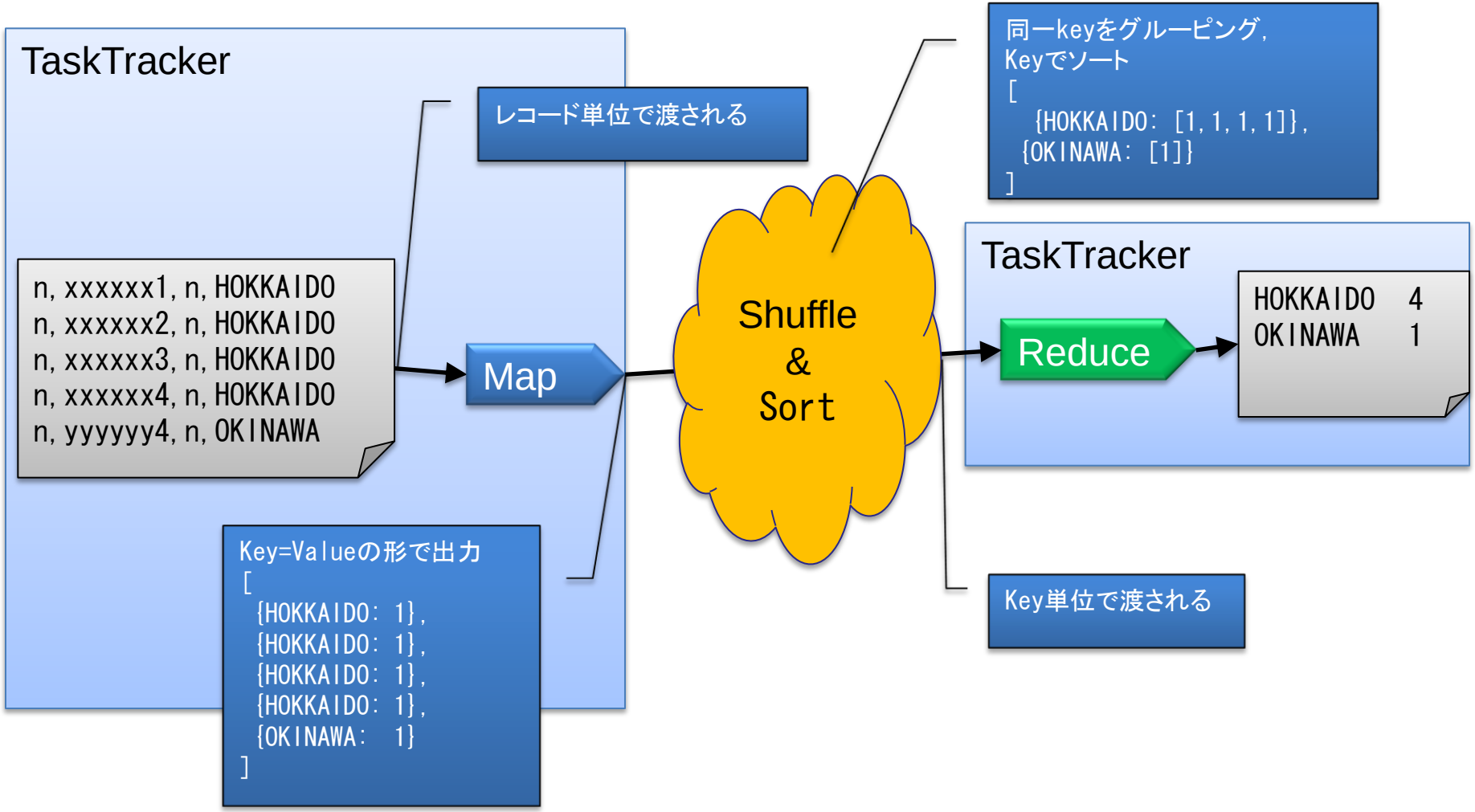
Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習α サンプルの実行

- 都道府県毎に郵便番号の保有数をカウント
  - 元データ
    - 郵政省　郵便番号
    - http://www.post.japanpost.jp/zipcode/dl/roman.html
  - 以下のディレクトリに「KEN_ALL_ROME.CSV」を配置
    - C:¥aitc¥work¥input
  - サンプルコード、環境はworkspaceの中
    - クラスファイル右クリック＞実行
  - 実行結果は以下のディレクトリに出力
    - C:¥aitc¥work¥output
    - 再度実行する際には、ディレクトリごと削除する必要あり
  - サンプルは２つ作成。Hadoop APIの書き方を変えている
    - 新しいAPIだとLinux上でエラー(NoClassDefFound)となるため、以後古いAPIを使用する(サンプル２)
      - 新API pkg：org.apache.hadoop.mapreduce
      - 旧API pkg：org.apache.hadoop.mapred
        - Deprecatedが出ますが、動作には支障ありません

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

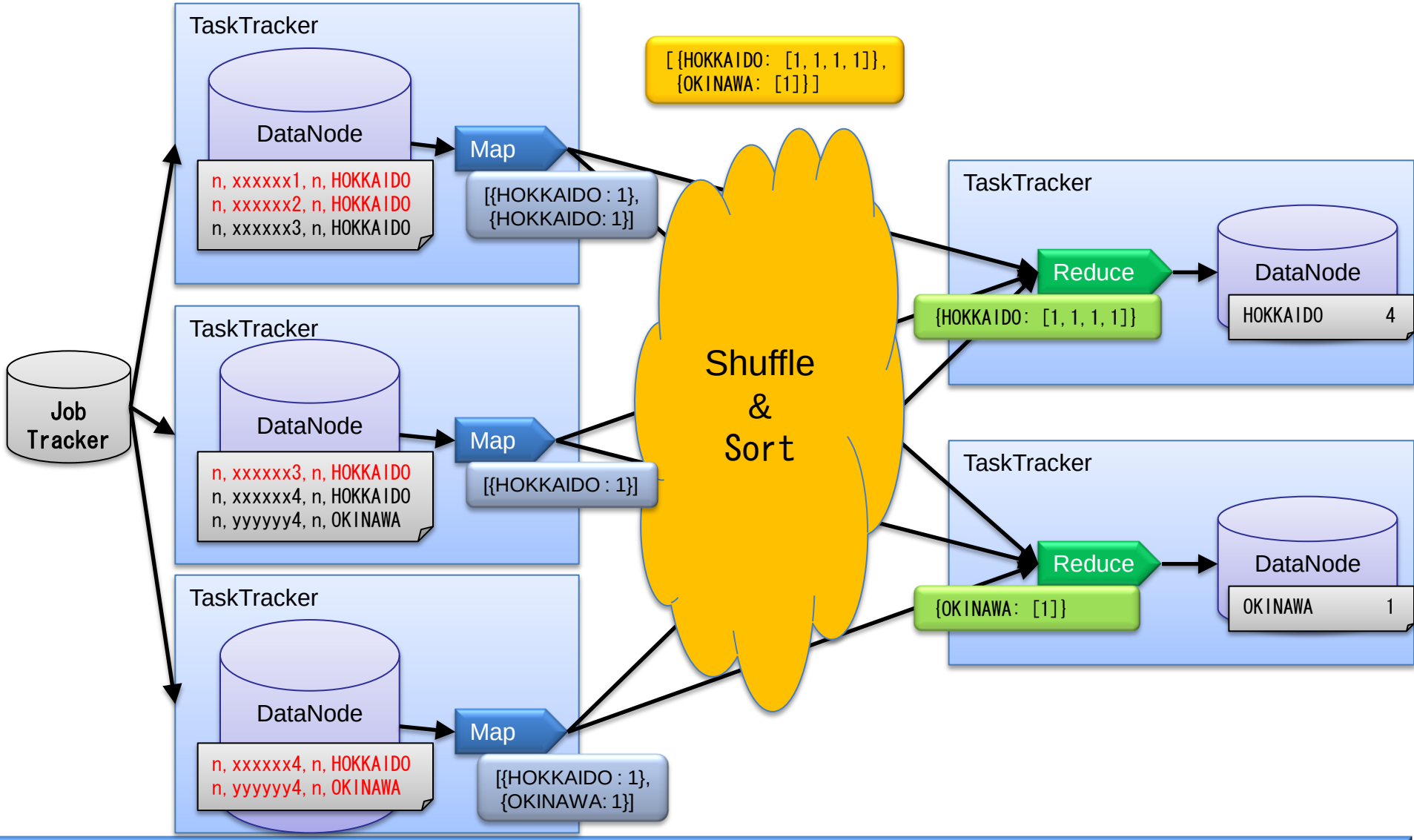
演習α サンプル１の実行



Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習α サンプル１の実行

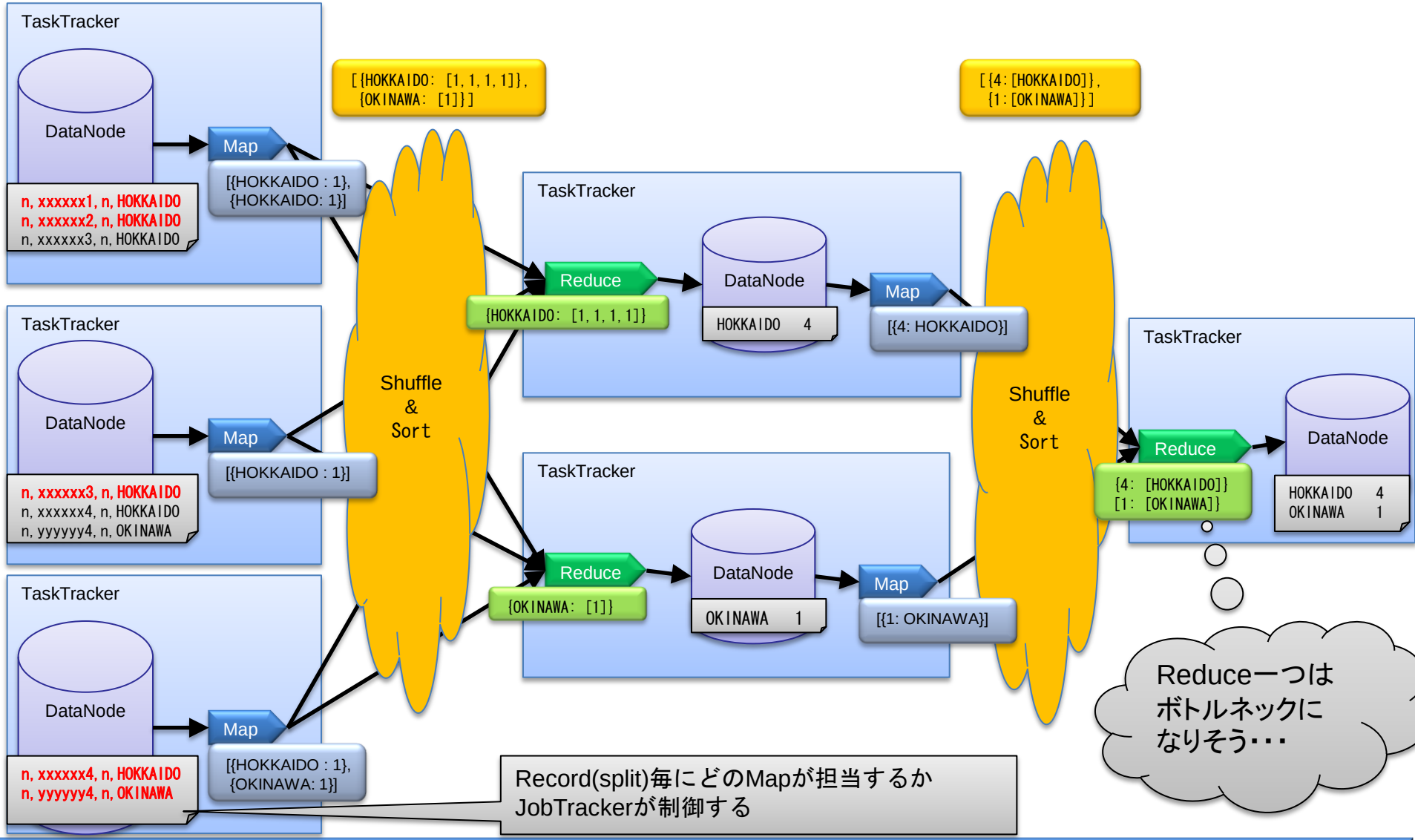
– 複数ノードで動作した際のイメージ



Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習α サンプル２の実行

– Map/Reduceを繋ぐことも可能(JobTracker省略)



Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習① Windowsで実装

- 使用するデータ
  - 防災情報XMLのフィード情報
    - 防災情報XMLとは
      - http://xml.kishou.go.jp/
    - 20日間で・・・
      - プッシュフィード： 8,263本　12,239,733Byte
      - 防災情報XML本体 ：17,826本　320,194,258Byte
    - データが沢山あって嬉しい
      - 大量データを用意するのが大変
        - ダミーデータ
      - 宝の山
      - 誰でも登録可能

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習① Windowsで実装

- サンプル
  - XMLファイルを読み込むサンプルのMapを用意
  - 日付毎に「雪」を含んだ情報の件数をカウント

- 演習内容
  - 日付,情報種別,場所(発表官署)毎に件数を集計
    - Base:/feed/entry (一つのfeedに複数のentry)
    - 日付
      - updated
    - 情報種別
      - title
    - 場所(発表官署)
      - author/name
  - input/outputのディレクトリパスが環境によって異なるため、標準入力でパスを受け付ける
    - スタンドアロン: c:¥xxx¥xxx / 分散環境: hdfs:///xxx/xxx

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習① Windowsで実装

- Eclipseのプロジェクト作成～実装
  - ファイル＞新規＞Javaプロジェクト
  - ビルドパスに設定
    - プロジェクト右クリック＞プロパティ＞Javaのビルドパス＞ライブラリ＞外部Jarの追加
      - C:¥aitc¥hadoop-0.20.2¥hadoop-0.20.2-core.jar
      - C:¥aitc¥hadoop-0.20.2¥lib 直下の*.jarを全て
      - ※ hadoopとは関係ありませんが、Xpathを使用するため lib 配下の*.jarもビルドパスに含めてください
  - コードを実装

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習① Windowsで実装

- 動作確認するための設定
  - Mainクラス 右クリック > 実行 > Javaアプリケーション
    - 初回はエラーとなる
  - Mainクラス右クリック > 実行 > 実行の構成
    - Javaアプリケーションに先ほど実行したMainクラス用の構成が作成されている
    - VM引数の設定（引数タブ）
      - -Xmx1000m -Dhadoop.log.dir=C:¥aitc¥hadoop-0.20.2¥logs -Dhadoop.log.file=hadoop.log -DHadoop.home.dir=C:¥aitc¥hadoop-0.20.2 -Dhadoop.id.str=host -Dhadoop.root.logger=INFO,console -Dhadoop.policy.file=hadoop-policy.xml
    - ちなみに、この状態で実行するとcygwinが必要な理由が分かります
      - “chmod”コマンド実行エラー
    - 環境変数の設定（環境タブ）
      - 変数: Path, 値: C:¥aitc¥cygwin¥bin

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Linuxの環境設定
  - 現状の環境
    - Javaはインストール済み
      - JAVA_HOME=/usr/local/java
    - FTPログイン
      - aitc/aitc01
  - Hadoop環境設定
    - hadoop-0.20.2.tar.gzをFTPでアップ
    - サーバ上でファイルの展開

```
> tar xvfz hadoop-0.20.2.tar.gz
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Linuxの環境設定
  - Pathの設定
    - ~aitc/.bash_profile

```
JAVA_HOME=/usr/local/java
HADOOP_HOME=/home/aitc/hadoop-0.20.2

PATH=$JAVA_HOME/bin:$HADOOP_HOME/bin:$PATH:$HOME/bin

export PATH JAVA_HOME HADOOP_HOME
```

    - ※PATH反映のため編集後、再度コンソールに入り直す

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Linuxの環境設定
  - Hadoop分散環境の設定（初期状態はスタンドアロン）
    - $HADOOP_HOME/conf/core-site.xml

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
  <property>
     <name>fs.default.name</name>
     <value>hdfs://localhost:9000</value>
  </property>

  <property>
     <name>hadoop.tmp.dir</name>
     <value>/home/aitc/hdfs</value>
  </property>
</configuration>
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Linuxの環境設定
  - Hadoop分散環境の設定（初期状態はスタンドアロン）
    - $HADOOP_HOME/conf/hdfs-site.xml

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
  <property>
     <name>dfs.replication</name>
     <value>1</value>
  </property>
</configuration>
```

    - １台で動作するため、レプリケーション数は１

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Linuxの環境設定
  - Hadoop分散環境の設定（初期状態はスタンドアロン）
    - $HADOOP_HOME/conf/mapred-site.xml

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
  <property>
     <name>mapred.job.tracker</name>
     <value>localhost:9001</value>
  </property>

  <property>
     <name>mapred.system.dir</name>
     <value>/mapred/system</value>
  </property>
</configuration>
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Linuxの環境設定
  - Hadoop分散環境の設定（初期状態はスタンドアロン）
    - $HADOOP_HOME/conf/hadoop-env.sh

```
# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME.  All others ar
e
# optional.  When running a distributed configuration it is best to

# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

# The java implementation to use.  Required.
# export JAVA_HOME=/usr/lib/j2sdk1.5-sun
export JAVA_HOME=/usr/local/java
:
:
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Linuxの環境設定
  - SSHのパスフレーズ無し設定
    - 以下のコマンド実行

```
> ssh-keygen -t dsa -P '' -f ~/.ssh/id_dsa

> cat ~/.ssh/id_dsa.pub >> ~/.ssh/authorized_keys
> chmod 600 ~/.ssh/authorized_keys
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- Hadoopの動作確認
  - NameNodeのフォーマット

```
> hadoop namenode -format
```

  - NameNode, DataNodeの起動(JobTracker, TaskTrackerも含む)

```
> start-all.sh
```

  - NameNode, DataNodeの停止(JobTracker, TaskTrackerも含む)

```
> stop-all.sh
```

  - ※困ったときは、stop-all.shで停止後、/tmp配下と/home/aitc/hdfs (hadoop.tmp.dir)を消して、再度フォーマットからやり直してみてください

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- 作成したジョブの配置、実行
  - JARファイルの作成
    - プロジェクト右クリック＞エクスポート＞実行可能JARの作成
  - FTPでサーバに配備
    - どこでもOK
  - 入力ファイルをHDFSに配備

```
> hadoop fs -put input hdfs:///input
> hadoop fs -put data hdfs:///data
```

  - 実行

```
> hadoop jar xxxxx.jar hdfs:///input hdfs:///output
```

  - 実行結果確認

```
> hadoop fs -cat /output/part-r-00000
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- HDFS Tips
  - HDFS上のディレクトリ削除

```
> hadoop fs -rmr /xxxxx/xxxxx
```

  - HDFS上のファイル一覧

```
> hadoop fs -ls /xxxxx/xxxxx
```

  - HDFS上にファイル配備

```
> hadoop fs -put <LOCAL> <HDFS>
```

  - HDFS上のファイル内容確認

```
> hadoop fs -cat /xxxxx/xxxxx
```

  - HDFSからファイル取得

```
> hadoop fs -get <HDFS> <LOCAL>
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習② Linuxで動作確認

- 動作状況を見るには・・・
  - HDFS
    - http://<IP>:50070
  - ジョブ
    - http://<IP>:50030
  - 細かい実行情報が出力されているので見ているだけでも面白い

- サンプルで動作確認

```
#InputファイルをHDFSに配置
> hadoop fs -put LICENSE.txt hdfs:///LICENSE.txt
#Wordカウントのサンプル実行
> hadoop jar hadoop-0.20.2-examples.jar ¥
                  wordcount hdfs:///LICENSE.txt hdfs:///output
#結果確認
> Hadoop fs -cat hdfs:///output/part-r-00000
```

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

演習③ 分散環境で動作確認

- 擬似分散環境で動作確認出来た方は声を掛けてください

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

時間が余ったら・・・

- Combiner/Partitionerの話
  - ネットワークの帯域節約
  - ネットワークトポロジー
- エコシステムの話
- Hadoopカンファレンスの話
- Hadoopバージョンの不思議

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

Hadoopエコシステム

- Hadoop単独で使用するのではなく、エコシステムを組み合わせて使用する
  - エコシステムがHadoop人気の一つの理由

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

Hadoopカンファレンス

- ビックサイトで開催
  - 1000人近くの人が参加
  - Hadoop人気はまだ続いている模様

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

最後に・・・

- 次回のクラウド部会では、RDF/SPARQLのハンズオンを行う予定です
  - RDFは、LinkedOpenDataなどで広く活用が進んできているデータ構造です
    - ProjectLAでも取り組んでいます
  - SPARQLはRDFのクエリ言語です
    - RDFの三つ組構造(トリプル)を扱うことが出来ます
  - 是非、この機会にRDF/SPARQLに触れてみてください
  - クラウド部会に遊びに来て下さい！

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

付録：AmazonEMRでの実行①

- 事前作業
  - credentials.jsonに認証関係の設定を書き込む
  - Job(JAR), InputファイルをS3に配置

- EMR上での実行
  - ワンタイムでの実行 (毎回EMRの起動/停止)

```
./elastic-mapreduce --create ¥
    --name aaaa ¥
    --master-instance-type m1.large ¥
    --slave-instance-type m1.small ¥
    --num-instances 10 ¥
    --jar s3://xxxxx/xxxxx.jar ¥
    --args s3://input/,s3://output
```

Master: 1台
Slave: 9台

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

付録：AmazonEMRでの実行②

- EMR上での実行
  - 繰り返し実行
    - 起動　※JOBフローのID(j-xxxxxxx)が払い出されるので、以後IDを指定

```
./elastic-mapreduce –create --alive ¥
    --name aaaa ¥
    --master-instance-type m1.large ¥
    --slave-instance-type m1.small ¥
    --num-instances 10 ¥
```

    - 実行

```
./elastic-mapreduce --jobflow j-xxxxxxx ¥
    --jar s3://xxxxx/xxxxx.jar ¥
    --args s3://input/,s3://output
```

    - 停止

```
./elastic-mapreduce --terminate --jobflow j-xxxxxxx
```

- Linux上に構築した際との違い
  - Hadoopクラスタを自分で組む必要が無い
    - パラメータでいくつ起動するか指定可能
  - データファイル、JobはHDFSではなくS3に配置する
  - 基本的な考え方に違いは無い

Copyright © 2012 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

AITC
先端ＩＴ活用推進コンソーシアム

END