

Hadoopによる利用状況の見える化

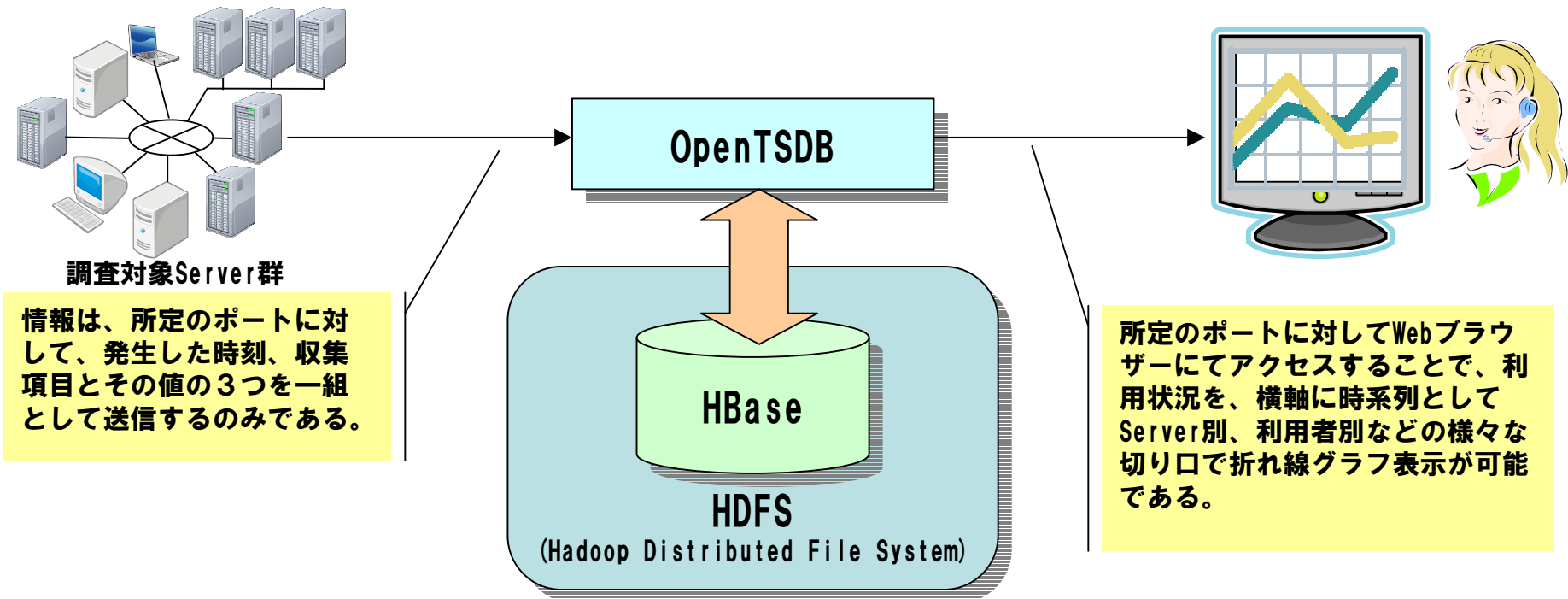
1. 目的

利用者単位での課金を目的として、複数の組織に管理されているServerの利用状況を一元的かつ、視覚的に把握する。

2. 課題と対策

- (1) 調査対象のServerが増加することが見込まれているため、スケールアウトが可能な構成とする。ファイルシステムはHDFS (Hadoop Distributed File System)、データベースはKVS系のHBaseを導入する。
- (2) 調査対象のServerが異なる組織で管理されているため、情報収集に複雑な仕組みを構築できない。また、利用状況の収集項目も追加変更が予測されるため、収集情報の簡素化、収集結果となる利用状況のグラフ表示の自動化を考慮しOpenTSDBを導入する。

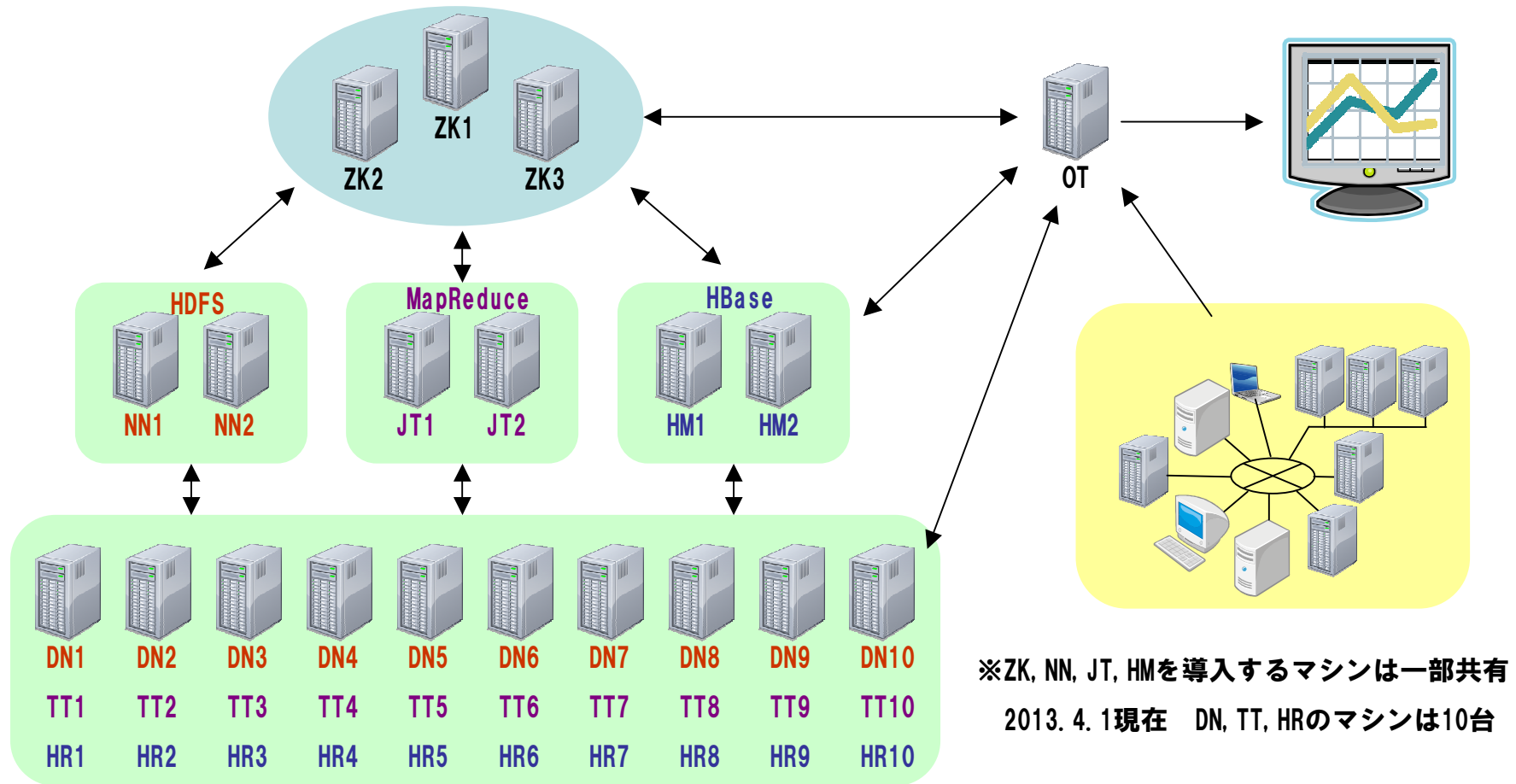
3. 全体イメージ



Hadoopによる利用状況の見える化

4. プロダクト構成

HadoopのDistributionは、Cloudera社製のCDH4. 2. 0にて構築した。



※ZK, NN, JT, HMを導入するマシンは一部共有
2013. 4. 1現在 DN, TT, HRのマシンは10台

凡例	ZK : Zookeeper Server	NN : Name Node	DN : Data Node	● : Quorum 構成
		JT : Job Tracker	TT : Task Tracker	■ : HA 構成
	OT : OpenTSDB	HM : HMaster	HR : HRegion Server	■ : 調査対象Server