Deployment Data Cleanup

D Goldsmith, R Wilkins

December 17, 2012

1 Introduction

Below is a description of the data summary process, how the information is processed an yields calculated.

It takes the form of a walk though of the yield calculation process, and includes code snippets from the R script used to generate the data.

The document was generated in R and Latex, using the Sweave plugin.

2 Scripts

This section has a description of the scripts used process the data, and combine all samples into one database.

These scripts are designed to work with the new format (location aware) database format.

They can found in the *dataclense* directory of the *cogent-house/djgoldsmith-devel* repository.

processCC.py Transfers current cost data from the old style sqlite database, into the new format database.

processAr.py Transfers data from an Archrock postgresql database into the new format database.

getStats.R R script that calculates yields for each deployment in a given database

calcKwh.R R script to caluclate KwH usage from current cost readings.

```
> library(RMySQL)
> library(ggplot2)
```

```
> drv <- dbDriver("MySQL")
> con <- dbConnect(drv,dbname="mainStore",user="root",password="Ex3lS4ga")
> allHouses <-  dbGetQuery(con,statement="SELECT * FROM House WHERE address !=
```