WBDB 2012

Paul Kent, SAS

@hornpolish

paulmkent
MPP Database – Simplified View

Key Points
- MPP hardware running database software
- A single logical table is stored in parts across multiple worker nodes
- “work” operates in parallel on the different parts of the table
- Aggregate memory large enough to fit the problem. YAY!
Idea Underlying SAS Alongside-the-DBMS

View the appliance as having a split personality

Yin
- An MPP database with an SQL interface

Yang
- A grid of nodes each of which can load TK extensions that mutually communicate with MPI (Message Passing Interface)

Alongside-the-database Yin and Yang
libname teradata joe;
proc hpreg data=joe.flights;
   class airline day(split);
   model delay=airline day duration ...;
   selection method=lasso;
run;
libname teradata joe;
proc hpreg data=joe.flights;
   class airline day(split);
   model delay=airline day
       duration ...;
   selection method=lasso;
run;

Alongside-the-Database
SAS Server

Appliance

General

Captains

Controller

Workers

tkgrid

Access Engine

MPI

TK

TK

TK

TK

TK

TK

SQL

SQL

SQL

SQL

SQL
libname hadoop joe;
proc hpreg data=joe.flights;
   class airline day(split);
   model delay=airline day duration ...;
   selection method=lasso;
run;

Alongside-Hadoop

SAS Server

libname hadoop joe;
proc hpreg data=joe.flights;
   class airline day(split);
   model delay=airline day duration ...;
   selection method=lasso;
run;

Hadoop Cluster

General

Captains

Controller

Workers

Access Engine

tkgrid

tkgrid

MPI

HDFS Block List

HDFS Blocks

HDFS Blocks

HDFS Blocks

HDFS Blocks

HDFS Blocks

HDFS Blocks
Math implemented so far...

Data Preparation / Exploration:
- HPSAMPLE (simple and stratified sampling)
- HPSUMMARY (base data reporting and analysis)
- HPDS2 (free-form code for data preparation and scoring)
- HPDMDB (class and variable summaries for modeling)
- HPIMPUTE (replace, encode, and flag missing values)
- HPBIN (discretize continuous inputs)

Regression Modeling:
- HPREG (linear regression with modern model selection methods)
- HPLOGISTIC (logistic regression with model selection)
- HPNLM (nonlinear regression)

Mixed Models:
- HPLMIXED (linear mixed models with both fixed and random effects)

Data Mining Methods:
- HPREDUCE (Supervised and unsupervised fast feature selection)
- HPNEURAL (feed forward neural network)
- HPFOREST (Random Forest – decision tree method)
- HP4SCORE (Scoring of a random forest model)
- HPDECIDE (profit and loss analysis)

Econometric Modeling:
- HPSEVERITY (estimates parameters of arbitrary continuous probability distributions)
- HPCOUNTREG (model count responses with Poisson type distributions)
So what do SAS want from the cluster?

1 – A good way to level blocks over datanodes on write
   - Today we use 0.21, write equal slices on node and concat()

2 – Fastest possible read into a C process
   - Today we use mmap()

3 – Network/Disk balance to compensate for #1

4 – Network to be available for MPI traffic as well

5 – plenty of excess CPU
Benchmarking Model of this

- SASload.write  -- create blocks on all nodes
- SASload.readstrategy – compute levelled map of which blocks to process on which hosts
  - Might be easy
  - Might require blocks be transhipped
- SASload.readblocks
- SASload.compute
  - do math that burns all cores
- SASload.output – write smaller amount
  - Possibly after memory stress on datanodes