Scripting, Plotting, Latex

Fan Zhang
Outline

• Learn to do your work in scripts with one example
  – Major commands that are covered (Examples borrowed from Bill Garrison’s materials)
    • grep, sed, awk
    • plot
    • latex

• Some useful links
Suppose we have this task

• Given a file with different lines of information
• Generate a report about it
  – See example: bad_name.bones
  – Generate a report about the connection time
Steps we need to do

• Extract the data lines
  – 64 bytes from pishon.cs.pitt.edu (130.49.222.82): icmp_req = 21
    Ql = 255 & me = 0.239 ms

• Extract the detailed numbers
  – pishon.cs.pitt.edu 0.239ms

• Plot about each individual

• Generate a PDF report
Extracting the data lines

• Extract the lines
  – Use grep
    • `grep icmp_req bad_name.bones > lines.dat`

• Grep cheatsheet
Get the figures out

• Use sed
  – sed 's/64 bytes from ([A-Za-z0-9\.-]*) ([0-9.]*):
icmp_req=([0-9]*) ttl=[0-9]* time=([0-9.]*):
ms/\2,\1,\3/' < lines.dat > parsed.dat

• sed tutorial:

• Essential:
  – Sed ‘s//’ < old > new
Split the data

• Use awk
  
  for i in {1..24} 
    do 
      awk "/^$i,/] < parsed.dat | sed "s/^$i,\([A-Za-z0-9\.-]*\),\([0-9\.-]*\)/\1 \2/" >> data_files/d$i.dat 
    done 

• Awk tutorial
More about awk

• Counting the size of files
  – `ls -l *.* | awk '{sum+=$5} END {print sum}'`

• Counting the ram used by each user
  – `ps aux | awk ‘NR!=1{a[$1]+=$6;} END {for (i in a) print i “, “ a[i] “KB”;}’`
More about AWK

• BEGIN { }
  – Before processing

• {}
  – Processing

• END {}  
  – After processing
Plot the data

• Plot one file
  set title "Test number **i**"
  set xlabel "site"
  set ylabel "time (ms)"
  set term png size 600, 800 enhanced font "Vera,12"
  set output "pngs/plot_*i*.png"
  set xtics rotate
  plot "data_files/d*i*.dat" using 2:xtic(1)

• Gnuplot tutorial
  – http://people.duke.edu/~hpgavin/gnuplot.html
Basics of gnu plot

• Customizing scales
  – xtic, ytic

• Specifying columns
  – Using 2:1

• Set style
  – set style data histogram
Generate plot files for different data files

- Using what we have learned
  
  for i in {1..24}
  do
    sed "s/\*\*i\*\*/$i/" < plot_i.gp > gp_scripts/plot_$i.gp
  done
Calculate the average and plot

• sed "s/\*\*i\*\*/AVG/" < plot_i.gp > gp_scripts/plot_AVG.gp

• gnuplot gp_scripts/plot_AVG.gp
Generate the report

• Using latex
  – # create report.tex, a latex file with all plots
  – echo '\documentclass{article}' > report.tex
  – echo '\usepackage{graphicx}' >> report.tex
  – echo '' >> report.tex
  – echo '\begin{document}' >> report.tex
  – echo '' >> report.tex
  – for i in {1..24}
    – do
    –   echo '\includegraphics[width=\textwidth]{pngs/plot_$i.png}' >> report.tex
    – done
  – echo '\includegraphics[width=\textwidth]{pngs/plot_AVG.png}' >> report.tex
  – echo '\end{document}' >> report.tex
  – # compile latex file into pdf
  – pdflatex report
Basic knowledge of latex

• \begin{} \end{} 
  – Document, section, paragraph

• Introducing packages 
  – \usepackage{}

• Insert graph, table 
  – \includegraphics{}, \begin{tabular}...

• Math formulas 
  – $, $$
Latex

• Latex tutorial:

• Writing papers with latex

• Tools that make latex easier:
  – texmaker
  – http://www.xm1math.net/texmaker/
Putting things in the cloud

www.writelatex.com
• Yay!