Perl Data Language (PDL)

Some examples

Perl Data Language (PDL): some examples

1

Contents

- what exists outside perl: matlab / octave, c/c++, numpy, R, mathematica / maxima

- benefits of perl as basis compared to other languages (regexp, I/O, data cleaning, flexio)

- great performance

- known uses and library stacks based on PDL (taken from http://pdl.perl.org/?page=users)

- a solution of running this in a multiple processor environment

- some examples of what Spoorgloren uses it for

Perl Data Language (PDL): some examples

What is PDL

- Perl Data Language extends perl (assumed known)

- In PDL arrays of numerical data are stored in binary form

- The usual perl operators are overloaded, and additional ones are added

- Calculations are at binary speed, with multiple-CPU support

- Numerical libraries are available (FFT, lapack)

Perl Data Language (PDL): some examples

short example PDL code

- perldl> $a=pdl ([0,2,3],[4,5,6])
- perldl> $a +=2
- perldl> print $a
  - [
  - [2 4 5]
  - [6 7 8]
  -]

Perl Data Language (PDL): some examples

PDL features

- Several matrix data types (int, float, double)

- Fast slicing, transposing, etc without data replication

- Easy re-dimensioning (3x2 matrix -> 6x1)

- Most numerical libraries are there (e.g. matrix operations)

- Use of perl (hashes !) to generate indices

- Good introduction at: http://perltv.org/v/introduction-to-the-perl-data-language

4/23/2014

Perl Data Language (PDL): some examples

5

Other solutions

- Scientists once used FORTRAN and Pascal

- c/c++ took over, but inconvenient I/O

- matlab / octave, R: very powerful, lame I/O

- mathematica / maxima: symbolic manipulation

- python+numpy: main alternative to PDL

Perl Data Language (PDL): some examples

Benefits of perl basis w.r.t. other solutions

- For unclean data perl is great, and fast: regexp, I/O
- data cleaning, hashing to coded numbers: great in perl
- PDL had own binary file handling flexio

Perl Data Language (PDL): some examples

Performance

- Text handling in perl is fast
- http://raid6.com.au/~onlyjob/posts/arena/
- PDL performance also good:
- http://www.freesoftwaremagazine.com/artic les/cool_fractals_with_perl_pdl_a_benchma rk

Perl Data Language (PDL): some examples

Spoor
Gloren

Known uses

2000 Emissions SNAP NATIONAL

- known uses and library stacks based on PDL (taken from http://pdl.perl.org/?page=users)
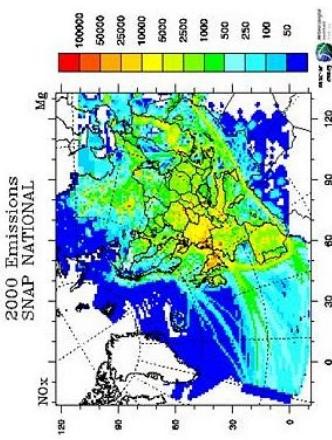
- UN-ECE's EMEP program monitors long-range air pollutants such as acid rain and ozone, over the whole of Europe

Perl Data Language (PDL): some examples

Known uses II

- ## Astronomy and Astrophysics

  - Artificial vision of magnetic fields on the Sun
  - High-speed galactic winds
  - Identifying galaxies from telescope images

- ## Meteorology and Terrestrial Monitoring

  - Measuring European air pollution for the U.N.
  - Real-time weather and forest fire detection from space

- ## Medical imaging

  - Brain scan visualization

Perl Data Language (PDL): some examples

Solution in multi-processor environment

- At Spoorgloren, data is usually delivered in daily of monthly updates

- Use file system and combination of ASCII and binary files to store update chunks

- Use "make -j" to start (re-) processing

- This is a simple solution for running in a multiple processor environment, not yet as multi computer (but not yet needed)

Perl Data Language (PDL): some examples

Spoor
Gloren

# Examples of Spoorgloren's usage

- passenger forecasts
- ECO driving
- passenger streams at stations
- Nightly energy usage reduction

Perl Data Language (PDL): some examples

Passenger forecasts

- Per-train level
- Including change in schedule

- Hoogenraad, J.H., De Vos, A.F., De Vries, J.J., 2013. NS-project AURORA voorspelt voor elke trein het aantal reizigers STAtOR 14 (3-4), 28-31

Evaluation ECO driving (EZR)

EZR Besparing 2013 Landelijk

Pedestrian througput in stations

station Ut_UIT type st-per, selectie=CiCo, N>=100

1A (N=8746.6)	
11B (N=77265.2)	
12A (N=9463.8)	
12B (N=21934.8)	
14A (N=656.5)	
14B (N=116392.0)	
15 (N=110214.2)	
18A (N=31391.5)	
18B (N=12887.1)	
19A (N=32019.6)	
19B (N=48962.9)	
2 (N=52837.8)	
3 (N=56821.5)	
4A (N=10175.5)	
4B (N=981.9)	
5A (N=98483.6)	
5B (N=1508.0)	
7A (N=75174.3)	
7B (N=170.6)	
8A (N=324.7)	
8B (N=91418.6)	
9A (N=9828.5)	
9B (N=28872.4)	

fractie (%)

verblijfstijd
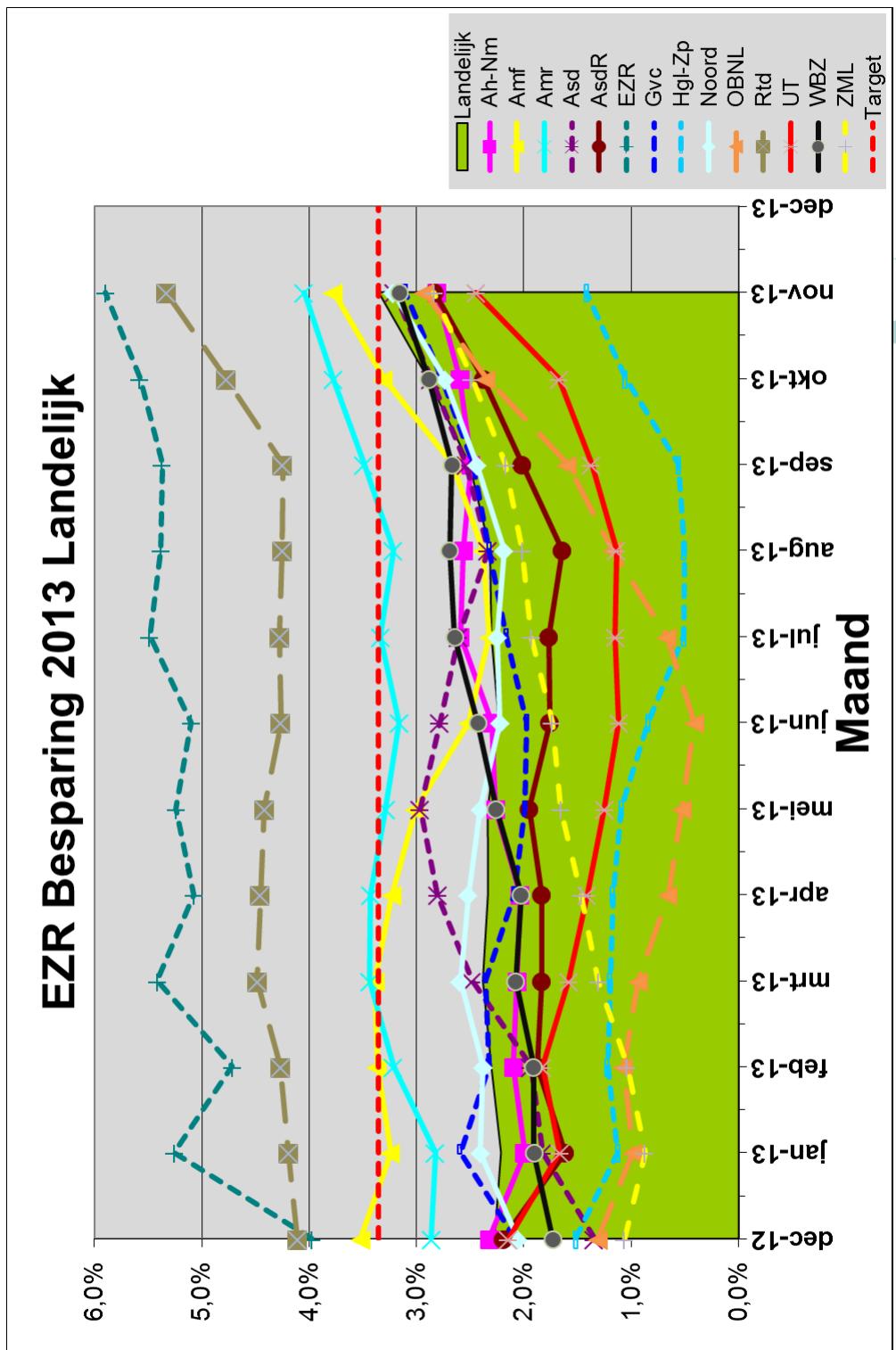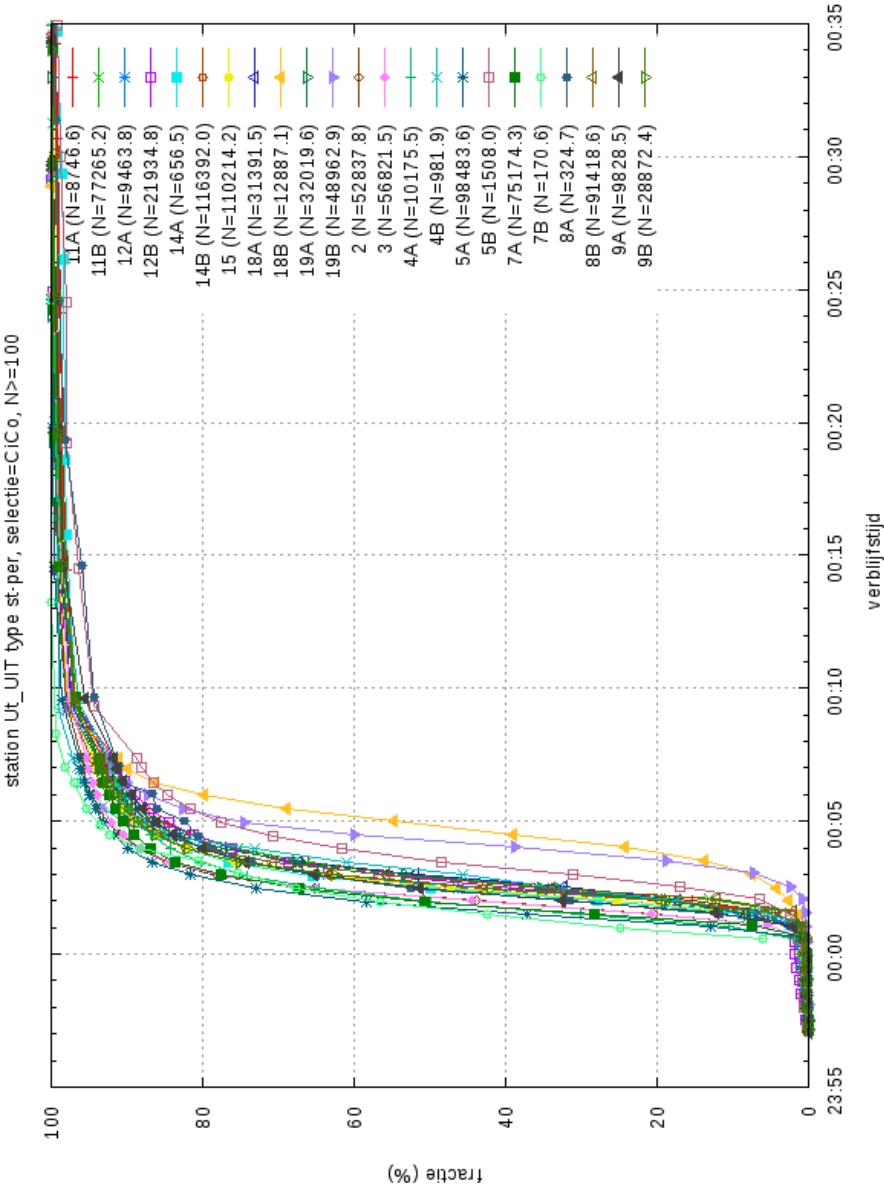
4/23/2014

Perl Data Language (PDL): some examples

Summary of our PDL usage

- Fast calculations in perl
- Cleaning and coupling of input data
- Usually batch oriented, data output (even though PDL has great graphics)
- For Dutch Rail

Perl Data Language (PDL): some examples