Compiled Matlab for LIGO Data Analysis

Lee Samuel Finn, for
Penn State University Relativity Group
https://gravity.psu.edu/~s2

LIGO-G030625-00-Z

Introduction & Motivation

- Goals:
  - » Development environment that facilitates rapid prototyping, testing, performance monitoring that maps directly to binary code suitable for distributed computing on cluster/enterprise/global grids.
  - » High-level language with library or useful abstractions for signal processing statistics
    - – Avoid low-level coding & reinventing the wheel
- Choice: Matlab
  - » Integrated development environment (interpreter, symbolic debugger, profiler, compiler)
  - » High-level language with linear algebra primitives
    - – Optimized for linear algebra (LAPACK on optimized BLAS, FFTW)
  - » Signal processing, statistics, filter design, system identification libraries
  - » Powerful integrated graphics

Example

```
function mloop(n)
% MLOOP - create vector, take
   log, dump to file
%
% mloop(n)
%
% n       expands vector 1:n

if (nargin ~= 1)
  error('usage: mloop n');
end

% cnvrt cmdlne to num
if isstr(n)
  n = str2num(n);
end
```

```
% Expand to vector and take
   log
x = 1:n;
y = log(x);

% Dump to text file
fid = fopen('gout.txt','w');
fprintf(fid,…
   '%d\t%10.3e\n',[x;y]);
fclose(fid);

return
```

I/O tools: reading channel data from frames

- Matlab interface to LDR database
  - » Hide all details of data location, organization from user
  - » Channel object: indexable view into specific channel

```
chan = chanstruct('H1:LSC-AS_Q');      % open channel for use
stim = 729331024;                      % start time
dur = 1234;                            % duration
[d,fs] = chanvector(chan,stim,dur);    % get data
…
[d,fs] = chanvector(chan,stim2,dur2);  % more data
…
[d,fs] = chanvector(chan2,stim3,dur3); % different channel
```

  - » Object-based implementation available; however, objects don't yet compile

Summary Conclusions

- BNETG & (almost all) data conditioning, event clustering and post-processing tools implemented
  - » All playground, other studies to date with compiled matlab on local cluster grid
- Performance?
  - » Fast on complex operations with long vectors; slow on simple operations, short vectors
  - » Hot-spots can be optimized with hand-coded c or c++
- License issues? None
  - » Runtime libraries, compiled binaries freely distributable
- Condor_compile for checkpointing
  - » Untested: waiting on RH9 operable condor_compile to investigate