Running MATLAB @NSC

ANDERS SJÖSTRÖM

Parameter sweep

Two-dimensional cantilevered truss subjected to forcing

Serial version

```
for ii = 1:numel(nGrid)

    [isModelAssumptionValid,Y,bars,groundDofs,actualToReduced]=trussCantilever(nGrid(ii),hGrid(ii),aGrid(ii),L);

    if isModelAssumptionValid

peakVals(ii) = max(max(abs(Y(:,2:2:end/2))));

    else

        fprintf('Linear model assumption not valid for N = %d and A = %e\n',nGrid(ii),hGrid(ii),aGrid(ii));

        peakVals(ii) = nan;

    end

if ~isempty(hTopAxes)

    hTopAxes.Children.ZData(ii)=log10(abs(peakVals(ii)));

    drawnow

    end

if showTruss

    plotTruss(Y,bars,L,nGrid(ii),hGrid(ii),groundDofs,actualToReduced)

    end

end
```

Parfor version

```
D = parallel.pool.DataQueue;
D.afterEach(@(x) updateSurface(hTopAxes,x));
parfor ii = 1:numel(nGrid)
[isModelAssumptionValid,Y,bars,groundDofs,actualToReduced]=…
trussCantilever(nGrid(ii),hGrid(ii),aGrid(ii),L);
    if isModelAssumptionValid
peakVals(ii) = max(max(abs(Y(:,2:2:end/2))));
        send(D, [peakVals(ii), ii]);
Else
     fprintf('Linear model assumption not valid for N =…
 %d and A = %e\n',nGrid(ii),hGrid(ii),aGrid(ii));
        peakVals(ii) = nan;
    end
if ~isempty(hTopAxes)
            drawnow
    end
if showTruss
     plotTruss(Y,bars,L,nGrid(ii),hGrid(ii),groundDofs,actualToReduced)
    end
end
```

LUND
UNIVERSITY

Parfeval version

```
% Note: with many iterations, overhead will be large, so we submit groups of iterates

p = gcp;

N = min([3*p.NumWorkers, numel(nGrid)]);   % similar to what parfor does (recommended for non-trivial cases)

%N = numel(nGrid);     % Basically turns off the grouper (an over-ride)

% Asynchronous submission
for ii = 1:N

    idx=ii:N:numel(nGrid);  % take every Nth sample, so we can see results fill in

    % Note that parfeval will submit, but not waitfor results

    f(idx)=parfeval(p, @Grouper, 6, nGrid(idx), hGrid(idx), aGrid(idx), L, idx); %#ok<AGROW>

end


% Can do other work while I wait ( Asynchronous)
```

Parfor version cont.

```
for ii=1:N

    [~, thisResult, Y, bars, groundDofs, actualToReduced, thisIdx] = fetchNext(f);

    peakVals(thisIdx) = thisResult;

     if ~isempty(hTopAxes)

        hTopAxes.Children.ZData(thisIdx)=log10(abs(peakVals(thisIdx)));

        drawnow

    end

% Plot trusses

    if showTruss

        plotTruss(Y,bars,L,nGrid(thisIdx),hGrid(thisIdx),groundDofs,actualToReduced);

    end

end

mainComputationTime = toc(t0);

if ~isempty(hTopAxes)

    visualizeParamSweep(nVals, hVals, aVals, peakVals, hTopAxes);

end
```

Large dataset

Demo

Airline dataset demo by Ken Atwell
- Data from 1987 through 2008
- (we use the first 9 years)
- Each year have 0.5 GB data

Using hadoop databases

```
setenv('HADOOP_HOME', '/path/to/hadoop/install')
setenv('SPARK_HOME', '/path/to/spark/install');
cluster = parallel.cluster.Hadoop;
% Optionally, if you want to control the exact number of workers:
cluster.SparkProperties('spark.executor.instances') = '16';
mapreducer(cluster);

ds = datastore('airlinesmall.csv');
varnames = {'ArrDelay', 'DepDelay'};
ds.SelectedVariableNames = varnames;
ds.TreatAsMissing = 'NA';

tt = tall(ds)
```

Where to learn more?

https://blogs.mathworks.com/loren/

https://se.mathworks.com/help/distcomp/use-tall-arrays-on-a-spark-enabled-hadoop-cluster.html

https://se.mathworks.com/help/distcomp/

https://se.mathworks.com/campaigns/portals/mdcs-resources-for-hpc.html