Lightweight Scheduling for Grid Applications

A. Lazarevic and L. Sacks
Outline

▪ Motivation
▪ Methodology
▪ Observed user behaviour
▪ Prediction Strategy
▪ Conclusions
Better Grid Scheduling

- Deadline based
- Transparent to the user
- (mostly) Self-managing
- More effective
  - Out-of-Order execution
  - Priority replaced with confidence level
If only...

- We knew...
  - Process’ execution time
  - Resource utilisation
  - Likelihood of a similar process arriving
  - Our confidence in these numbers

- We could...
  - Be flexible & robust
  - Treat anomalies in the same way as reg data
  - Adapt to change in circumstance & prediction quality
Methodology

Schedule performance monitoring & feedback
Confidence level and model improvements

Highly Granular Resource Monitoring

User-Orientated Deadline Based Scheduler
One-step-ahead predictions

Statistical Analysis:
- Job Arrival Rate
- Resource Utilisation
- Execution Time
- Social Factors
Methodology

- Raw Historical Data (Accounting, Resource Monitoring)
  - Other n-Dimensional Exploratory Techniques
  - Meta-Data Clustering (User / Group / Executable / Time of Day / Day of Week)
  - Temporal Analysis
  - Statistical Models: Linear/Polynomial Time-series ARIMA Wavelets ...
  - Exception Handling
  - 1...n Classes of Jobs
    - 1...n Classes of Models
      - Forecasts
      - Variable Time Horizon

- Quality Control: Prediction Errors
  - Class Temporal Properties
  - Class Stat. Properties
  - Deadline Based Scheduler (Out-of-order Execution, Statistical Multiplexing)
MATLAB Framework

Data Manipulation: Cleaning, Sorting, Meta-data extraction

Data Analysis & Preparation: Trends, Seasonality, Outliers, Smoothing, Differencing

Forecast Monitoring & QA: Residuals Analysis, Prediction Error Analysis, Prediction Variance & Confidence Level

Forecasting: Method selection, Model Building, Predicting

Anomaly Detection

Raw Logs, Accounting Files

Visualisation
Preliminary Data Analysis

▪ Data Source:
  ▫ 200 CPU Sun Grid Engine cluster @ UCL

▪ User community:
  ▫ 20 UK e-Science projects
  ▫ 5-10 different projects concurrently running

▪ Scope:
  ▫ 50,000 jobs (>5,500 CPU days)
  ▫ Jul 2004 – Dec 2004
Trivial Approach – All in

All Jobs (42,863)

Job_ID

WIClock

0 5000 10000 15000 20000 25000 30000 35000 40000 45000

0 5.0 50.0 500.0 5000.0 5E5 5E6

University College London
Temporal Job Selection

Wall Clock Execution Times - 1000 jobs

1000 Jobs, 15 hours of total execution time, used in all following analyses
Intra-Group Analysis - \textit{ocotir}

![Graph 1: Execution Time - ocotir Group](image1.png)

![Graph 2: Execution Time Histogram - ocotir Group](image2.png)

Total No: 19
Intra-Group Analysis - matsim

Execution Time - matsim Group

Execution Time Histogram - matsim Group

Total No: 933
Intra-Group Analysis - geogrs

Execution Time - geogrs Group

Execution Time Histogram - geogrs Group

Total No: 48
Temporal Locality?
Anomalies

- Detect sudden, significant and sustained changes in the quality of predictions
- Detecting them
  - Using a self-tuneable system
  - Flexible definition of an anomaly
- Dealing with them
  - Shorten the prediction horizon
  - Reduce the confidence interval
Anomaly Detection - Exp Smoothing

Line Plot (Spreadsheet11 11v*933c)

Line Plot (Spreadsheet11 11v*933c)
Prediction Methods

- Trend prediction (linear, polynomial...)
- Time-series
  - Exponential Smoothing (Holt, SES, etc)
  - Autoregressive & Moving Average
  - ARIMA (Box-Jenkins)
- Bayesian Statistics
- Wavelets
  - “All models are wrong, but some are useful”

George E.P. Box, Professor Emeritus
Forecasting method analysis done on a subset of *matsim* jobs (highlighted)
Trend Based Predictions – *matsim*

Actual v. Predicted WClock Values - *matsim* (0:250)
Predictor: 3rd order Poly Fit

- Actual Values
- Predicted Values
Trend Based Predictions - *matsim*

**WClock Percentage Prediction Error - matsim (0:250)**

- **Error Value**
- **Error Mean**
- **Error StD**
Conclusions

▪ Clear and pressing need for better Grid scheduling

▪ Patterns do develop in Grid utilisation

▪ These patterns vary between “groups”, and in time

▪ User-independent predictions with reasonable accuracy are possible

▪ Impact to the Grid utilisation, job workflow and administrative policies may be significant