Performance of Optimization Software -
an Update

INFORMS Annual 2011
Charlotte, NC
13-18 November 2011

H. D. Mittelmann
School of Math and Stat Sciences
Arizona State University
Services we provide

• Guide to Software: "Decision Tree"

• http://plato.asu.edu/guide.html

• Software Archive

• Software Evaluation: "Benchmarks"

• Archive of Testproblems

• Web-based Solvers (1/3 of NEOS)
We maintain the following NEOS solvers (8 categories)

Combinatorial Optimization  * CONCORDE [TSP Input]
Global Optimization  * ICOS [AMPL Input]
Linear Programming
  * bpmcpd [AMPL Input][LP Input][MPS Input][QPS Input]
Mixed Integer Linear Programming
  * FEASPUMP [AMPL Input][CPLEX Input][MPS Input]
  * SCIP [AMPL Input][CPLEX Input][MPS Input] [ZIMPL Input]
  * qsopt_ex [LP Input][MPS Input] [AMPL Input]
Nondifferentiable Optimization  * condor [AMPL Input]
Semi-infinite Optimization  * nsips [AMPL Input]
Stochastic Linear Programming  * bnbs [SMPS Input]
  * DDSIP [LP Input][MPS Input]
We maintain the following NEOS solvers (cont.)

Semidefinite (and SOCP) Programming

* csdp [MATLAB_BINARY Input][SPARSE_SDPA Input]
* penbmi [MATLAB Input][MATLAB_BINARY Input]
* pensdp [MATLAB_BINARY Input][SPARSE_SDPA Input]
* sdpa [MATLAB_BINARY Input][SPARSE_SDPA Input]
* sdplr [MATLAB_BINARY Input][SDPLR Input][SPARSE_SDPA Input]
* sdpt3 [MATLAB_BINARY Input][SPARSE_SDPA Input]
* sedumi [MATLAB_BINARY Input][SPARSE_SDPA Input]
Overview of Talk

• Current and Selected(*) Benchmarks
  – Parallel LP benchmarks
  – MILP benchmark (MIPLIB2010)
  – Feasibility/Infeasibility Detection benchmarks (MIPLIB2010)
  – slightly pathological MILP cases
  – MI(QC)QP benchmark

• Conclusions
COMBINATORIAL OPTIMIZATION

Concorde-TSP with different LP solvers (8-16-2011)

LINEAR PROGRAMMING

Benchmark of serial LP solvers (11-9-2011)
* Benchmark of parallel LP solvers (11-11-2011)
* Parallel Barrier Solvers on Large LP/QP problems (11-12-2011)
Large Network-LP Benchmark (commercial vs free) (11-10-2011)

MIXED INTEGER LINEAR PROGRAMMING

MILP Benchmark - serial codes (9-29-2011) (to be phased out)
* MILP Benchmark - MIPLIB2010 (11-11-2011)
* MILP cases that are slightly pathological (11-9-2011)
* Feasibility Benchmark (11-9-2011) (MIPLIB2010)
SEMIDEFINITE/SQL PROGRAMMING

Several SDP-codes on SDP problems with free variables (4-1-2008)
Several SDP codes on problems from SDPLIB (4-10-2008)
SQL problems from the 7th DIMACS Challenge (8-8-2002)
Newer SDP/SOCP-codes on the 7th DIMACS Challenge problems (4-7-2008)
Several SDP codes on sparse and other SDP problems (6-22-2011)
SOCP (second-order cone programming) Benchmark (10-29-2011)

NONLINEAR PROGRAMMING

Benchmark of commercial and other (QC)QP Solvers (11-11-2011)
AMPL-NLP Benchmark, IPOPT, KNITRO, LOQO, PENNLP, SNOPT & CONOPT (11-9-2011)

MIXED INTEGER NONLINEAR PROGRAMMING

* MI(QC)QP Benchmark (11-11-2011)

PROBLEMS WITH EQUILIBRIUM CONSTRAINTS

MPEC Benchmark (10-18-2010)
Important features of all our benchmarks

• Statistics of problems (dimensions etc)

• Links to codes given

• Links to test problems given

• Links to full logfiles given

• Same selection for commercial/free codes
Reasons for updates

- New version of commercial software
  - CPLEX-12.4, GUROBI-4.6.0, KNITRO-7.2.1, MOSEK-6.0.0.124

- New versions of free software
  - CBC, CLP, SCIP, SYMPHONY
  - BONMIN, COUENNE, IPOPT, FEASPUMP2

- More multicore hardware
Benchmarks still in need of updates

• SDP Benchmark (parallel, iterative, nonlinear)

• MINLP Benchmark, not only MIQCQP

• Benchmarks in new areas
  – Compressive Sensing, other sparse optimization
  – Derivative free/nonsmooth optimization etc
  – More commercial codes
Overview of Talk

• Current and Selected(*) Benchmarks
  – Parallel LP benchmarks
  – MILP benchmark (MIPLIB2010)
  – Feasibility/Infeasibility Detection benchmarks (MIPLIB2010)
  – slightly pathological MILP cases
  – MI(QC)QP benchmark

• Observations and Conclusions
11 Nov 2011     Benchmark of parallel LP solvers

Logfiles of these runs at: plato.asu.edu/ftp/lp_logs/

This benchmark was run on a Linux-PC (2.67 GHz Intel Core 2 Quad).
The MPS-datafiles for all testcases are in one of (see column "s")

  mplib.zib.de/ [1]
  plato.asu.edu/ftp/lptestset/ [2]
  www.netlib.org/lp/data/ [3,7]
  www.sztaki.hu/~meszaros/public_ftp/lptestset/
  (MISC[4], PROBLEMATIC[5], STOCHLP[6], INFEAS[8])

The (dual) simplex and barrier methods of the following codes were tested:

  CPLEX-12.4.0.0   CPLEX
  GUROBI-4.6.0     www.gurobi.com/
  MOSEK-6.0.0.124  www.mosek.com
  XPRESS-7.2.1:    XPRESS

Scaled geometric mean of runtimes

\[
\begin{array}{cccccccccccc}
3.71 & 2.47 & 4.40 & 3.27 & 1 & 1.15 & 1.47 & 1.52 & 1.43 & 1.01 & 1.33 \\
\end{array}
\]

=================================================================================
  CPX-S  GRB-S  MSK-S  XPR-S  CPX-B  GRB-B  MSK-B  XPR-B  CPX-A  GRB-A  MSK-A 
=================================================================================

12
Parallel Barrier Solvers on Large LP/QP problems

H. Mittelmann (mittelmann@asu.edu)

Logiles at plato.asu.edu/ftp/barrier_logs/

CPLEX-12.4.0.0: CPLEX
GUROBI-4.6.0: GUROBI
MOSEK-6.0.0.124: MOSEK
XPRESS-7.2.1: XPRESS
IPOPT-3.10.1: IPOPT

The barrier methods (w/o crossover) of the above solvers were run on a 3.47 GHz Intel Xeon X5690 (6 cores, 48GB) on large LP problems from here. Times given are elapsed times in seconds.
Parallel Barrier Solvers on Large LP/QP problems

<table>
<thead>
<tr>
<th>problem</th>
<th>CPLEX</th>
<th>GUROBI</th>
<th>MOSEK</th>
<th>XPRESS</th>
<th>IPOPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>in</td>
<td>46454</td>
<td>186</td>
<td>898</td>
<td>309</td>
<td></td>
</tr>
<tr>
<td>pde_10</td>
<td>2271</td>
<td>747</td>
<td>3351</td>
<td>15906</td>
<td></td>
</tr>
<tr>
<td>pde_1</td>
<td>2516a</td>
<td>2023</td>
<td>6323</td>
<td>fail</td>
<td></td>
</tr>
<tr>
<td>pde_200</td>
<td>299a</td>
<td>526</td>
<td>1326</td>
<td>1528</td>
<td></td>
</tr>
<tr>
<td>pde_20</td>
<td>1896a</td>
<td>2349</td>
<td>16184</td>
<td>3186</td>
<td></td>
</tr>
<tr>
<td>pde_2</td>
<td>5269a</td>
<td>16164</td>
<td>60426</td>
<td>9889</td>
<td></td>
</tr>
<tr>
<td>qap_2</td>
<td>slow</td>
<td>fail</td>
<td>slow</td>
<td>1581</td>
<td></td>
</tr>
<tr>
<td>srd300</td>
<td>10091</td>
<td>fail</td>
<td>17369</td>
<td>&gt;60000</td>
<td></td>
</tr>
<tr>
<td>zib01</td>
<td>6010a</td>
<td>8254</td>
<td>9124</td>
<td>7113</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>problem</th>
<th>CPLEX</th>
<th>GUROBI</th>
<th>MOSEK</th>
<th>XPRESS</th>
<th>IPOPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2CTA3D</td>
<td>3271</td>
<td>6562</td>
<td>5681</td>
<td>2169</td>
<td></td>
</tr>
<tr>
<td>bdry2_0</td>
<td>1925a</td>
<td>998</td>
<td>4223a</td>
<td>2412</td>
<td>4609</td>
</tr>
<tr>
<td>bdry2_1</td>
<td>10977a</td>
<td>3919</td>
<td>fail</td>
<td>10111</td>
<td>30590</td>
</tr>
<tr>
<td>cont5_2_0</td>
<td>512</td>
<td>1123</td>
<td>31</td>
<td>564</td>
<td>1157</td>
</tr>
<tr>
<td>cont5_2_1</td>
<td>1746</td>
<td>4167a</td>
<td>65</td>
<td>1545</td>
<td>3242</td>
</tr>
<tr>
<td>cont5_2_2</td>
<td>7979</td>
<td>fail</td>
<td>155</td>
<td>4181</td>
<td>11414</td>
</tr>
<tr>
<td>twod_00</td>
<td>2499</td>
<td>8412</td>
<td>16526a</td>
<td>2848</td>
<td>9348</td>
</tr>
<tr>
<td>twod_0</td>
<td>12391</td>
<td>47549</td>
<td>fail</td>
<td>12841</td>
<td>41616</td>
</tr>
</tbody>
</table>

"a": reduced accuracy
<table>
<thead>
<tr>
<th>problem</th>
<th>constraints</th>
<th>variables</th>
<th>nonzeros</th>
<th>MPS-file</th>
</tr>
</thead>
<tbody>
<tr>
<td>in</td>
<td>1526203</td>
<td>1449074</td>
<td>6813128</td>
<td>561 MB</td>
</tr>
<tr>
<td>pde_10</td>
<td>13225799</td>
<td>4414198</td>
<td>39669001</td>
<td>1.96 GB</td>
</tr>
<tr>
<td>pde_1</td>
<td>26993999</td>
<td>9005998</td>
<td>80970001</td>
<td>4.1 GB</td>
</tr>
<tr>
<td>pde_200</td>
<td>5067399</td>
<td>3382596</td>
<td>15197001</td>
<td>824 MB</td>
</tr>
<tr>
<td>pde_20</td>
<td>13225799</td>
<td>8824196</td>
<td>39669001</td>
<td>2.2 GB</td>
</tr>
<tr>
<td>pde_2</td>
<td>26993999</td>
<td>18005996</td>
<td>80970001</td>
<td>4.5 GB</td>
</tr>
<tr>
<td>qap_2</td>
<td>57360</td>
<td>14400</td>
<td>203040</td>
<td>70.7 MB</td>
</tr>
<tr>
<td>srd300</td>
<td>1101467</td>
<td>2052911</td>
<td>59170079</td>
<td>1.6 GB</td>
</tr>
<tr>
<td>zib01</td>
<td>5887041</td>
<td>12471400</td>
<td>49877768</td>
<td>2.5 GB</td>
</tr>
<tr>
<td>L2CTA3D</td>
<td>210000</td>
<td>10000000</td>
<td>30000000</td>
<td>1.8 GB</td>
</tr>
<tr>
<td>bdry2_0</td>
<td>4001998</td>
<td>4003997</td>
<td>19991999</td>
<td></td>
</tr>
<tr>
<td>bdry2_1</td>
<td>12253498</td>
<td>12256997</td>
<td>61235999</td>
<td></td>
</tr>
<tr>
<td>cont5_2_0</td>
<td>1959681</td>
<td>1961081</td>
<td>11749904</td>
<td></td>
</tr>
<tr>
<td>cont5_2_1</td>
<td>3999656</td>
<td>4001656</td>
<td>23986126</td>
<td></td>
</tr>
<tr>
<td>cont5_2_2</td>
<td>8999641</td>
<td>9002641</td>
<td>53979959</td>
<td></td>
</tr>
<tr>
<td>twod_00</td>
<td>505284</td>
<td>511446</td>
<td>4856046</td>
<td></td>
</tr>
<tr>
<td>twod_0</td>
<td>989604</td>
<td>999306</td>
<td>9586066</td>
<td></td>
</tr>
</tbody>
</table>
Overview of Talk

• Current and Selected(*) Benchmarks
  – Parallel LP benchmarks
  – MILP benchmark (MIPLIB2010)
  – Feasibility/Infeasibility Detection benchmarks (MIPLIB2010)
  – slightly pathological MILP cases
  – MI(QC)QP benchmark

• Conclusions
11 Nov 2011 Mixed Integer Linear Programming Benchmark (MIPLIB2010)

The following codes were run on the MIPLIB2010 benchmark set with the MIPLIB2010 scripts on an Intel Xeon X5680 (32GB, Linux, 64 bits, 2*6 cores), with one, four and twelve threads. (deterministically) and a time limit of 1 hour. These are updated and extended versions of the results produced for the MIPLIB2010 paper.

CPLEX-12.4.0.0: CPLEX
GUROBI-4.6.0: GUROBI
ug[SCIP/cpx/spx]: Parallel development version of SCIP (SCIP+CPLEX/SOPLEX/CLP on 1 thread)
CBC-2.7.4: CBC
XPRESS-7.2.1: XPRESS
GLPK-4.47: GLPK
LP_SOLVE-5.5.2: LP_SOLVE

Table for single thread, Result files per solver, Log files per solver
Table for 4 threads, Result files per solver, Log files per solver
Table for 12 threads, Result files per solver, Log files per solver

Statistics of the problems can be obtained from the iMIPLIB2010 webpage.
Geometric means of times

All non-successes are counted as max-time. The fastest solver is scaled to 1. The second line lists the number of problems (87 total) solved.

<table>
<thead>
<tr>
<th>threads</th>
<th>CBC</th>
<th>CPLEX</th>
<th>GLPK</th>
<th>GUROBI</th>
<th>LPSOLVE</th>
<th>SCIPC</th>
<th>SCIPL</th>
<th>SCIPS</th>
<th>XPRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>8.82</td>
<td>1.25</td>
<td>19.14</td>
<td>1</td>
<td>16.8</td>
<td>3.19</td>
<td>5.3</td>
<td>4.88</td>
</tr>
<tr>
<td>solved</td>
<td>41</td>
<td>73</td>
<td>3</td>
<td>78</td>
<td>5</td>
<td>61</td>
<td>52</td>
<td>56</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>threads</th>
<th>CBC</th>
<th>CPLEX</th>
<th>FSCIPC</th>
<th>FSCIPS</th>
<th>GUROBI</th>
<th>XPRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>10.27</td>
<td>1</td>
<td>5.78</td>
<td>9.41</td>
<td>1.06</td>
</tr>
<tr>
<td>solved</td>
<td>52</td>
<td>84</td>
<td>66</td>
<td>64</td>
<td>84</td>
<td>79</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>threads</th>
<th>CBC</th>
<th>CPLEX</th>
<th>FSCIPC</th>
<th>FSCIPS</th>
<th>GUROBI</th>
<th>XPRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12</td>
<td>11.1</td>
<td>1</td>
<td>8.08</td>
<td>11.9</td>
<td>1.07</td>
</tr>
<tr>
<td>solved</td>
<td>56</td>
<td>84</td>
<td>68</td>
<td>65</td>
<td>87</td>
<td>83</td>
</tr>
</tbody>
</table>

18
Overview of Talk

• Current and Selected(*) Benchmarks
  – Parallel LP benchmarks
  – MILP benchmark (MIPLIB2010)
  – Feasibility/Infeasibility Detection benchmarks (MIPLIB2010)
    – slightly pathological MILP cases
  – MI(QC)QP benchmark

• Conclusions
Logfiles for these runs are at: plato.asu.edu/ftp/feas_bench_logs/

MILP problems from MIPLIB2010 were solved for a feasible point

The following codes were run on an Intel Xeon X5680 (32GB, Linux, 64 bits, 2*6 cores) with 4 threads:

- CPLEX-12.4.0.0: CPLEX
- FEASPUMP2: as implemented for interactive use at NEOS (utilizes CPLEX)
- GUROBI-4.6.0: GUROBI
- XPRESS-7.2.1: XPRESS
- CBC-2.7.4: CBC

Times given are elapsed times in seconds. A time limit of 1 hr was imposed. Geometric means of the times are listed. For objective values see logfiles.

<table>
<thead>
<tr>
<th>problem(30 tot)</th>
<th>CPLEX</th>
<th>FP2</th>
<th>GUROBI</th>
<th>XPRESS</th>
<th>CBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>geometric mean</td>
<td>4.23</td>
<td>12.94</td>
<td>1</td>
<td>13.5</td>
<td>35.6</td>
</tr>
<tr>
<td>problems solved</td>
<td>21</td>
<td>17</td>
<td>27</td>
<td>18</td>
<td>9</td>
</tr>
</tbody>
</table>
The following codes were run on the infeasible problems from MIPLIB2010 with the MIPLIB2010 scripts.

CPLEX-12.4.0.0: CPLEX
GUROBI-4.6.0: GUROBI
ug[SCIP/spx/cpx]: Parallel development version of SCIP
CBC-2.7.4: CBC
XPRESS-7.2.1: XPRESS

Table for 12 threads, Result files per solver, Log files per solver

Statistics of the problems can be obtained from the MIPLIB2010 webpage.

+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Geometric means of times (19 instances)

All non-successes are counted as max-time. The fastest solver is scaled to 1.

<table>
<thead>
<tr>
<th></th>
<th>CBC</th>
<th>CPLEX</th>
<th>FSCIPC</th>
<th>FSCIPS</th>
<th>GUROBI</th>
<th>XPRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>solved</td>
<td>6</td>
<td>16</td>
<td>12</td>
<td>14</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>33.4</td>
<td>1.06</td>
<td>15.2</td>
<td>19.6</td>
<td>1.00</td>
<td>3.22</td>
</tr>
</tbody>
</table>
Overview of Talk

• Current and Selected(*) Benchmarks
  – Parallel LP benchmarks
  – MILP benchmark (MIPLIB2010)
  – Feasibility/Infeasibility Detection benchmarks (MIPLIB2010)
  – slightly pathological MILP cases
  – MI(QC)QP benchmark

• Conclusions
MILP cases that are slightly pathological

H. Mittelmann (mittelmann@asu.edu)

9 Nov 2011

CPLEX-12.4.0.0: CPLEX to be added
GUROBI-4.6.0: GUROBI
ug[SCIP/spx]: FSCIP-Parallel development version of SCIP
CBC-2.7.4: CBC
XPRESS-7.2.1: XPRESS
SCIP-2.1.0: serial SCIP with CPLEX

These codes were run with the MIPLIB2010 scripts in default mode on an Intel Xeon X5680 (32GB, Linux, 64 bits, 2*6 cores) on problems from here. Times given are elapsed CPU seconds. Time limit 3 hrs. Available memory 24GB. This benchmark is not giving a representative impression of the relative performance of the codes.

Table for 12 threads, Result files per solver, Log files per solver

Problems solved (total 22)

<table>
<thead>
<tr>
<th></th>
<th>CPLEX</th>
<th>GUROBI</th>
<th>FSCIP</th>
<th>CBC</th>
<th>XPRESS</th>
<th>SCIP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15</td>
<td>18</td>
<td>12</td>
<td>5</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>
Overview of Talk

- **Current and Selected(*) Benchmarks**
  - Parallel LP benchmarks
  - MILP benchmark (MIPLIB2010)
  - Feasibility/Infeasibility Detection benchmarks (MIPLIB2010)
  - slightly pathological MILP cases
  - MI(QC)QP benchmark

- Conclusions
11 Nov 2011

Mixed Integer (QC)QP Benchmark

H. Mittelmann (mittelmann@asu.edu)

Logiles at plato.asu.edu/ftp/miqp_logs/

The MPS-datafiles are in plato.asu.edu/ftp/miqp/ and the AMPL files in plato.asu.edu/ftp/ampl_files/miqp_ampl/ and egon.cheme.cmu.edu/ibm/files/

The following codes were run in default mode on a 2.66GHz Intel Core2 Quad, For accuracy reached, see logfiles. In the last columns are results for the QCQPs obtained by rewriting the QPs as: min t, subject to quadratic obj <= t plus constraints. SCIP does this transformation itself. SCIP uses relgap=0;

CPLEX-12.4.0.0: CPLEX
Bonmin-1.5.1: projects.coin-or.org/Bonmin (Bonmin: hybrid algorithm. with Cbc)
Couenne-0.3.1: projects.coin-or.org/Couenne
FilMINT: currently only at NEOS, (run locally)
GUROBI-4.6.0: gurobi.com
SCIP-2.1.0: scip.zib.de (with CPLEX and IPOPT)
XPRESS-7.2.1: XPRESS
Times given are user times in seconds. Time limit 10,800 seconds.

\[
\text{QP} \\
==
\text{Scaled geometric means of runtimes}
\]

\[
\begin{array}{ccccccc}
60 & 317 & 2.69 & 260 & 1 & 14.5 & 2.29 \\
\end{array}
\]

\[
\begin{array}{cccccccc}
\text{problem} & \text{Bonmin} & \text{Couenne} & \text{CPLEX} & \text{FilMINT} & \text{GUROBI} & \text{SCIP} & \text{XPRESS} \\
\end{array}
\]

\[
\text{QCQP} \\
====
\text{Scaled geometric means of runtimes}
\]

\[
\begin{array}{ccccccc}
11.8 & 35.98 & 2.46 & 30.0 & 1 \\
\end{array}
\]

\[
\begin{array}{cccccccc}
\text{problem} & \text{Bonmin} & \text{Couenne} & \text{CPLEX} & \text{FilMINT} & \text{XPRESS} \\
\end{array}
\]
Observations and Conclusions

- Observation: TTB (Tuning To the Benchmarks)
  - CPLEX-12.3 -> CPLEX-12.4 preview
    * better for MIPLIB benchmarks
    * worse for pathological ones
  - MOSEK-6.0.0.122 -> 6.0.0.124 Changelog
    * Improved the interior point optimizers for linear problems for certain hard problems.
Conclusions: Declare Winners?

- **MIPLIB-bench**: CPLEX/Gurobi-XPRESS
- **MIPLIB-feas**: Gurobi-CPLEX-XPRESS, FEASPUMP close
- **MIPLIB-infeas**: Gurobi, CPLEX close, XPRESS slow
- **Pathological**: Gurobi-CPLEX, SCIP,FSCIP respectable
- **MIQP**: Gurobi
- **MIQCQP**: XPRESS
- **LP**: CPLEX/Gurobi; **QP/QCQP**: XPRESS/MOSEK
Thank you!