Baseline w/o Middleware
And Gnuplot Tutorial

Advanced Systems Lab
Fall 2018
3 Clients connect to 1 Server

<table>
<thead>
<tr>
<th>Table: Control Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of servers</strong></td>
</tr>
<tr>
<td><strong>Number of client machines</strong></td>
</tr>
<tr>
<td><strong>Instances of memtier per machine</strong></td>
</tr>
<tr>
<td><strong>Threads per memtier instance</strong></td>
</tr>
<tr>
<td><strong>Virtual clients per thread</strong></td>
</tr>
<tr>
<td><strong>Workload</strong></td>
</tr>
<tr>
<td><strong>Multi-Get behavior</strong></td>
</tr>
<tr>
<td><strong>Multi-Get size</strong></td>
</tr>
<tr>
<td><strong>Number of middlewares</strong></td>
</tr>
<tr>
<td><strong>Worker threads per middleware</strong></td>
</tr>
<tr>
<td><strong>Repetitions</strong></td>
</tr>
</tbody>
</table>

- **Client VM 1**
- **Client VM 2**
- **Client VM 3**
- **Server VM 1**
3 Clients connect to 1 Server

You need to decide for yourself:
1) Granularity
2) Range

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of servers</td>
<td>1</td>
</tr>
<tr>
<td>Number of client machines</td>
<td>3</td>
</tr>
<tr>
<td>Instances of memtier per machine</td>
<td>1</td>
</tr>
<tr>
<td>Threads per memtier instance</td>
<td>2</td>
</tr>
<tr>
<td>Virtual clients per thread</td>
<td>[1..32]</td>
</tr>
<tr>
<td>Workload</td>
<td>Write-only and Read-only</td>
</tr>
<tr>
<td>Multi-Get behavior</td>
<td>N/A</td>
</tr>
<tr>
<td>Multi-Get size</td>
<td>N/A</td>
</tr>
<tr>
<td>Number of middlewares</td>
<td>N/A</td>
</tr>
<tr>
<td>Worker threads per middleware</td>
<td>N/A</td>
</tr>
<tr>
<td>Repetitions</td>
<td>3 or more</td>
</tr>
</tbody>
</table>

Decide on the error metric.

x-axis
1 Client connects to 2 Servers

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of servers</td>
<td>2</td>
</tr>
<tr>
<td>Number of client machines</td>
<td>1</td>
</tr>
<tr>
<td>Instances of memtier per machine</td>
<td>2</td>
</tr>
<tr>
<td>Threads per memtier instance</td>
<td>1</td>
</tr>
<tr>
<td>Virtual clients per thread</td>
<td>[1..32]</td>
</tr>
<tr>
<td>Workload</td>
<td>Write-only and Read-only</td>
</tr>
<tr>
<td>Multi-Get behavior</td>
<td>N/A</td>
</tr>
<tr>
<td>Multi-Get size</td>
<td>N/A</td>
</tr>
<tr>
<td>Number of middlewares</td>
<td>N/A</td>
</tr>
<tr>
<td>Worker threads per middleware</td>
<td>N/A</td>
</tr>
<tr>
<td>Repetitions</td>
<td>3 or more (at least 1 minute each)</td>
</tr>
</tbody>
</table>

Instances and threads have swapped!
1 Client connects to 2 Servers

<table>
<thead>
<tr>
<th>Number of servers</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of client machines</td>
<td>1</td>
</tr>
<tr>
<td>Instances of memtier per machine</td>
<td>2</td>
</tr>
<tr>
<td>Threads per memtier instance</td>
<td>1</td>
</tr>
<tr>
<td>Virtual clients per thread</td>
<td>[1..32]</td>
</tr>
<tr>
<td>Workload</td>
<td>Write-only and Read-only</td>
</tr>
<tr>
<td>Multi-Get behavior</td>
<td>N/A</td>
</tr>
<tr>
<td>Multi-Get size</td>
<td>N/A</td>
</tr>
<tr>
<td>Number of middlewares</td>
<td>N/A</td>
</tr>
<tr>
<td>Worker threads per middleware</td>
<td>N/A</td>
</tr>
<tr>
<td>Repetitions</td>
<td>3 or more (at least 1 minute each)</td>
</tr>
</tbody>
</table>

You need to decide for yourself:
1) Granularity
2) Range

x-axis

Decide on the error metric.
Description vs. Explanation

• Example sentence 1: We observe in Figure 1 that the throughput increases linearly with the number of clients, until X number of clients is reached.

• Example sentence 2: Since the rate of increase in response time changes suddenly at X number of clients, we deduce the system is saturated at…

• Example sentence 3: The throughput saturation is reached at X number of clients, because …

• Example sentence 4: The interactive law holds as expected, as shown in Figure 1.
Description vs. Explanation

• Example sentence 1: We observe in Figure 1 that the throughput increases linearly with the number of clients, until X number of clients is reached.

• Example sentence 2: Since the rate of increase in response time changes suddenly at X number of clients, we deduce the system is saturated at…

• Example sentence 3: The throughput saturation is reached at X number of clients, because …

• Example sentence 4: The interactive law holds as expected, as shown in Figure 1.
Summary

- It is OK to have *some* repetition from previous explanation sections, but DO NOT duplicate.
- Focuses more on differences.
- Focuses more on concluding remarks.
- Write a key take-away within ONE sentence.
Gnuplot 101

- Gnuplot is a command line tool for generating 2D and 3D plots of your data
- Works on Linux and Windows
- Can be used in interactive or scripted mode

<table>
<thead>
<tr>
<th>#</th>
<th>x</th>
<th>y</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>-1.1</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>1.0</td>
<td>0.97</td>
<td></td>
</tr>
</tbody>
</table>
How to organize your data?

• Output your experimental results as a comma/space/tab separated file

<table>
<thead>
<tr>
<th>#</th>
<th>x</th>
<th>value1</th>
<th>value2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>-1.1</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>1.0</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>-0.1</td>
<td>1.8</td>
<td></td>
</tr>
</tbody>
</table>

• You can also include text

<table>
<thead>
<tr>
<th>#</th>
<th>category</th>
<th>x</th>
<th>value1</th>
<th>value2</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>0.1</td>
<td>0.3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>0.4</td>
<td>-1.1</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>2.3</td>
<td>1.0</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>Fourth</td>
<td>0.1</td>
<td>-0.1</td>
<td>1.8</td>
<td></td>
</tr>
</tbody>
</table>
Hello world

- Start Gnuplot and type interactively, or create a text file with commands:

```
1 2
# age income
18 100
19 200
20 300
21 400
```

plot "results.txt" using 1:2 with lines title "hello world"
Gnuplot scripts

- Plots are set up declaratively – see online documentation for all commands

Start with `set term png/pdf/svg` to select the output type
- `set output “filename.png”`
- `set xlabel “x axis label”`
- `set ylabel “y axis label”`
- `plot “filename1” using 1:2 with lines title “first”,
  “filename2” using 1:3 with linespoints title “second”`
Column values can be combined

• It is possible to plot simple expressions:
  – Plot “file.txt” using 1:($2+$3+$4) with lines title “summed columns”

<table>
<thead>
<tr>
<th>#</th>
<th>t</th>
<th>req_type1</th>
<th>req_type2</th>
<th>req_type3</th>
<th>$2+$3+$4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>111</td>
<td>0</td>
<td></td>
<td>111</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>100</td>
<td>3</td>
<td></td>
<td>105</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>120</td>
<td>0</td>
<td></td>
<td>124</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>100</td>
<td>22</td>
<td></td>
<td>122</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>102</td>
<td>2</td>
<td></td>
<td>107</td>
</tr>
</tbody>
</table>

– You can use other arithmetic operations as well!
Example: Plotting response time

<table>
<thead>
<tr>
<th>#</th>
<th>CLIENTS</th>
<th>RESPONSE_TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>10000</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>20000</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>40000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>60000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>100000</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>150000</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>200000</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>300000</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>800000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1500000</td>
</tr>
</tbody>
</table>
Example: Plotting response time

```plaintext
set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5  # --- blue
set xlabel " "
set ylabel " "
set title " "
set xrange [-10:10]
set yrange [-1500000:1500000]
plot "data.txt" using 1:2 ls 1
```

What is wrong here?
Example: Plotting response time

set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5  # --- blue
set xlabel " "
set ylabel " "
set title " "
set xrange [-10:10]
set yrange [-1500000:1500000]
plot "data.txt" using 1:2 ls 1

- No labels or title.
- Legend is meaningless.
- Ranges are completely off.
- No lines.
Example: Plotting response time

set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5  # --- blue
set xlabel ""
set ylabel ""
set title ""
set xrange [-10:10]
set yrange [-1500000:1500000]
plot "data.txt" using 1:2 with lines ls 1

- No labels or title.
- Legend is meaningless.
- Ranges are completely off.
- No ticks.
Example: Plotting response time

- No labels or title.
- Legend is meaningless.
- Ranges are completely off.
Example: Plotting response time

```bash
set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5   # --- blue
set xlabel " "
set ylabel " "
set title " "
set xrange [0:11]
set yrange [10000:1500000]
plot "data.txt" using 1:2 with linespoints ls 1
```

What is wrong here?
Example: Plotting response time

set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5  # --- blue
set xlabel " "
set ylabel " "
set title " "
set xrange [0:11]
set yrange [10000:1500000]
set yrange [10000:1500000]
plot "data.txt" using 1:2 with linespoints ls 1

- No labels or title.
- Legend is meaningless.
- y-axis range does not start from 0.
Example: Plotting response time

set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5  # --- blue
set xlabel " "
set ylabel " in thousands "
set title " "
set xrange [0:11]
set yrange [0:1600]
plot "data.txt" using 1:($2/1000) with linespoints ls 1

- No labels or title.
- Legend is meaningless.
- x-axis ticks are not correct
Example: Plotting response time

set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5   # --- blue
set xlabel "Number of Clients"
set ylabel "Response time (s)"
set title "Response time over Number of Clients"
set xrange [0:11]
set yrange [0:1600]
plot "data.txt" using 1:($2/1000):xticlabel(1) with linespoints ls 1 notitle
Example: Plotting Error Bars

set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5  # --- blue
set xlabel "Number of Clients"
set ylabel "Throughput (transactions/second)"
set title "Throughput vs. Number of clients"
set xrange [0:11]
set yrange [0:80]
plot "data.txt" using 1:2:3:xticlabel(1) with errorbars title "System X, config ABC" ls 1

<table>
<thead>
<tr>
<th># CLIENTS</th>
<th>THROUGHPUT</th>
<th>STDDEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.2</td>
<td>1.1</td>
</tr>
<tr>
<td>2</td>
<td>20.1</td>
<td>1.2</td>
</tr>
<tr>
<td>3</td>
<td>28.9</td>
<td>1.4</td>
</tr>
<tr>
<td>4</td>
<td>37.0</td>
<td>4.5</td>
</tr>
<tr>
<td>5</td>
<td>40.1</td>
<td>5.5</td>
</tr>
<tr>
<td>6</td>
<td>42.2</td>
<td>7.0</td>
</tr>
<tr>
<td>7</td>
<td>43.0</td>
<td>14.4</td>
</tr>
<tr>
<td>8</td>
<td>43.3</td>
<td>22.2</td>
</tr>
<tr>
<td>9</td>
<td>44.4</td>
<td>17.0</td>
</tr>
<tr>
<td>10</td>
<td>42.1</td>
<td>2.2</td>
</tr>
</tbody>
</table>
Example: Plotting Multi Lines

```
set style line 1 lc rgb '#0060ad' lt 1 lw 2 pt 7 ps 1.5
set style line 2 lc rgb '#ad6000' lt 1 lw 2 pt 7 ps 1.5
set style line 3 lc rgb '#60ad00' lt 1 lw 2 pt 7 ps 1.5
set xlabel "Number of Clients"
set ylabel "Throughput (transactions/second)"
set title "Throughput vs. Number of clients"
set xrange [0:11]
set yrange [0:80]
plot "data.txt" index 00:00 using 1:2:3:xticlabel(1) with errorlines title "System X, config ABC" ls 1,
"data.txt" index 01:01 using 1:2:3:xticlabel(1) with errorlines title "System X, config DEF" ls 2,
"data.txt" index 02:02 using 1:2:3:xticlabel(1) with errorlines title "System X, config GHI" ls 3
```

Legends are important in multi-line graphs.
Example: Plotting Stacked Bar Charts

```
set style data histogram
set style histogram rowstacked
set style fill solid 1.0 border -1
set boxwidth 0.8 relative
set xlabel "Number of Clients"
set ylabel "Throughput (transactions/second)"
set title "Throughput vs. Number of clients"
set xrange [-1:10]
set yrange [0:100]
plot "DATA_FILE" using 2 title "Throughput 1", " using 3 title "Throughput 2"
```
Linear vs. Log Scale

Response time over Number of Clients

Response time (s)

Number of Clients
Recommendations

- As you progress and rerun experiments save results in different files – you can regenerate graphs for different versions of the result.
- Keep style consistent over graphs.
- Use bash scripts for exporting graphs!
Alternatives

• Bash (GNUplot), Python (matplotlib), Windows (Excel), Latex (pgfplots) ...

• No hand-drawn graphs!