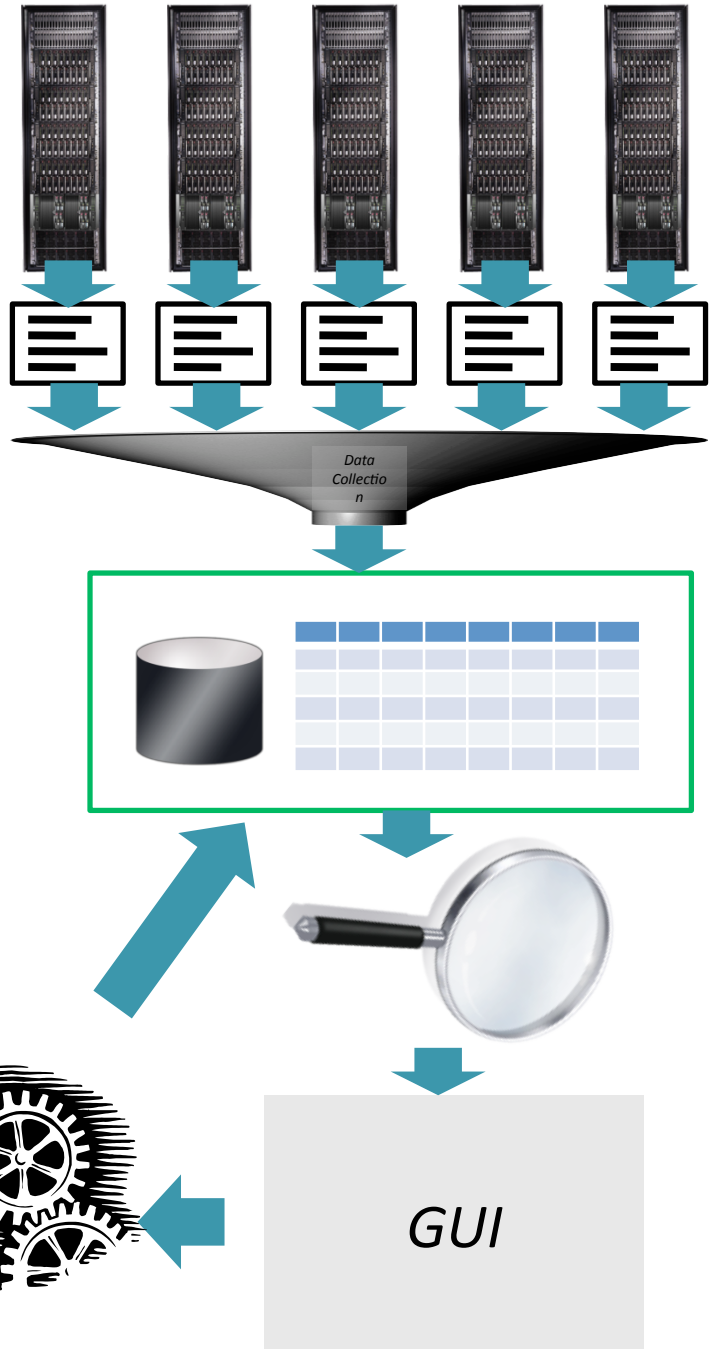
Artemis

Hunting for Bugs with Artemis

System Architecture



Distributed system

Dryad Overview

Logs

Data collection

Database

View

GUI

Conclusion s

Hunting for Bugs with Artemis

pptPlex Section Divider

The slides after this divider will be grouped into a section and given the label you type above. Feel free to move this slide to any position in the deck.

Hunting for Bugs with Artemis

Gabriela F. Crețu-Ciocârlie

Mihai Budiu

Moises Goldszmidt

Microsoft Research, Silicon Valley

WASL 2008

This presentation is built and should be viewed with pptPlex: http://www.officelabs.com/projects/pptPlex/Pages

Artemis Goal

One-stop shop for performance analysis of distributed systems

Principles

1) **Modular**:

   Separate generic from
   application specific parts

2) **Extensible:**

   add new analyses via plug-ins

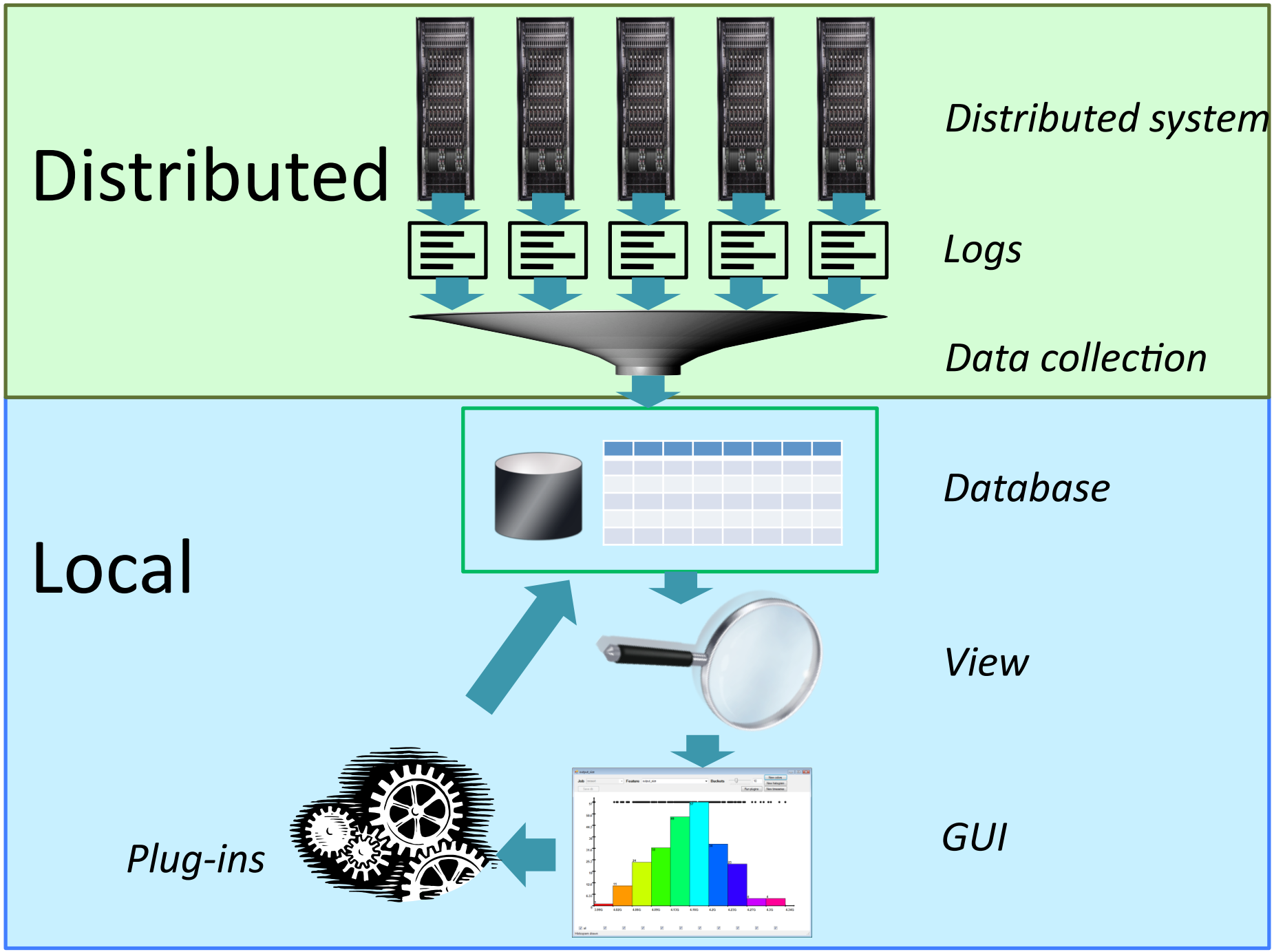
3) **Interactive:**

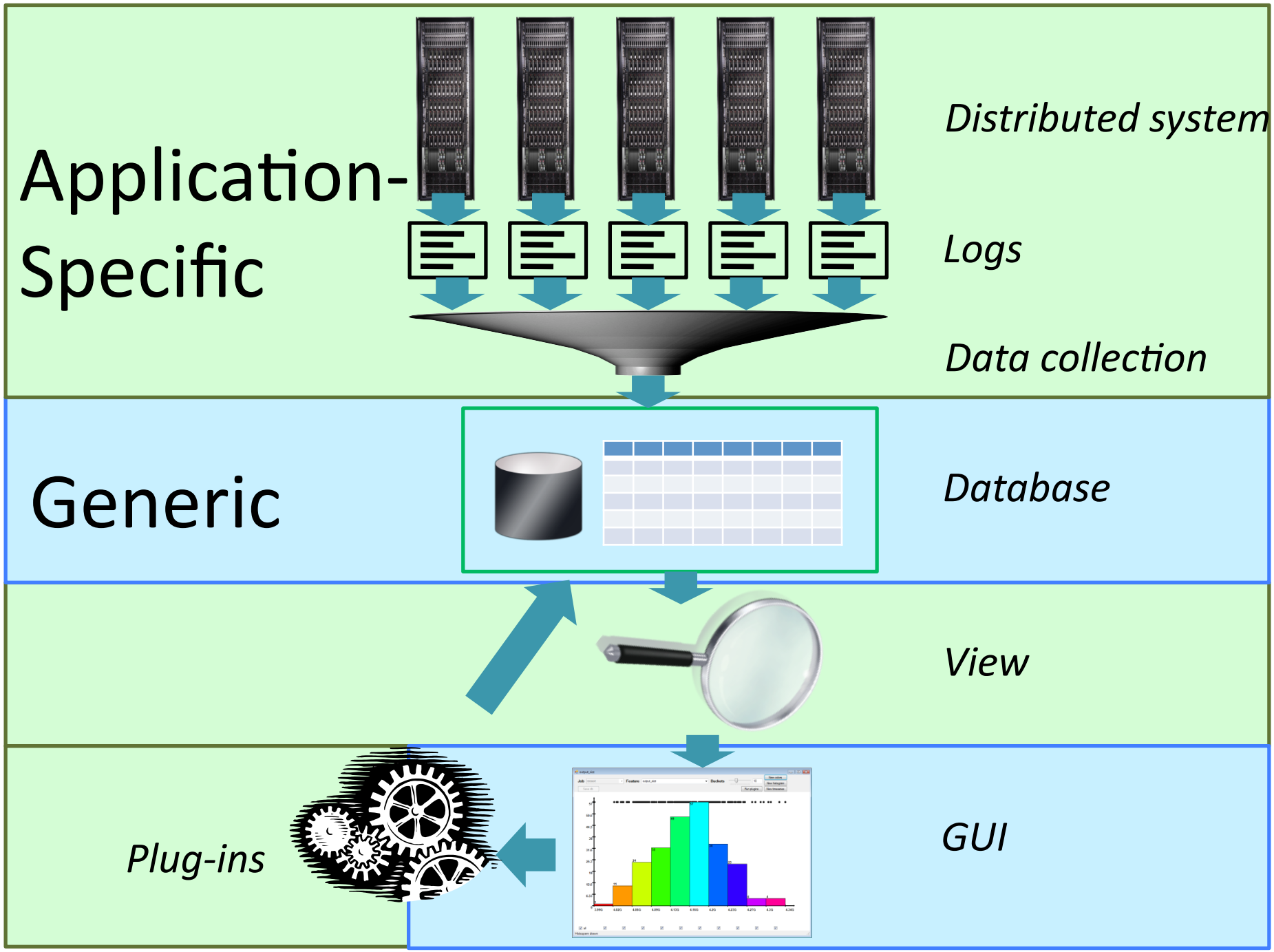
   human expert part of the analysis loop

System Architecture

pptPlex Section Divider

The slides after this divider will be grouped into a section and given the label you type above. Feel free to move this slide to any position in the deck.


Distributed
Distributed system
Logs
Data collection
Local
Database
View
GUI
Plug-ins

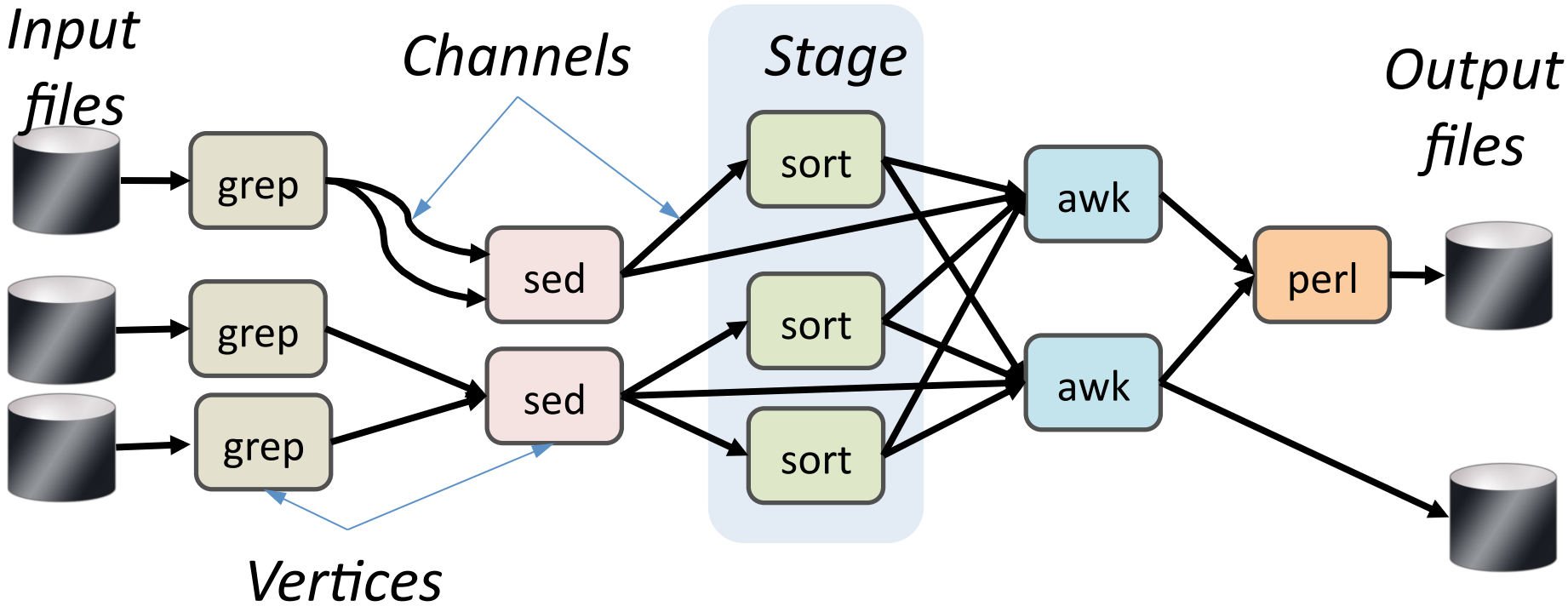

Application-Specific
Distributed system
Logs
Data collection
Generic
Database
View
Plug-ins
GUI

Dryad Overview

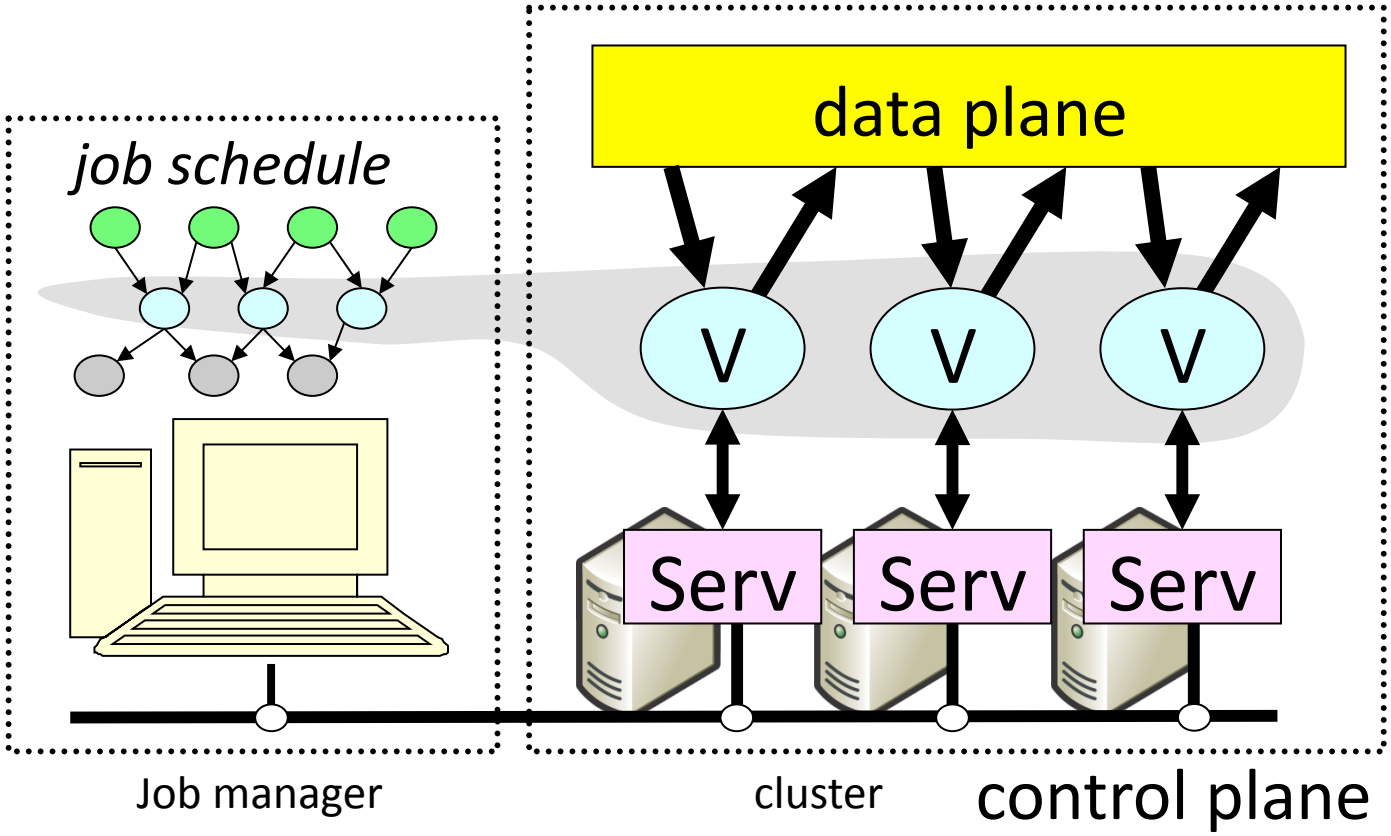
pptPlex Section Divider

The slides after this divider will be grouped into a section and given the label you type above. Feel free to move this slide to any position in the deck.

Dryad Application Structure



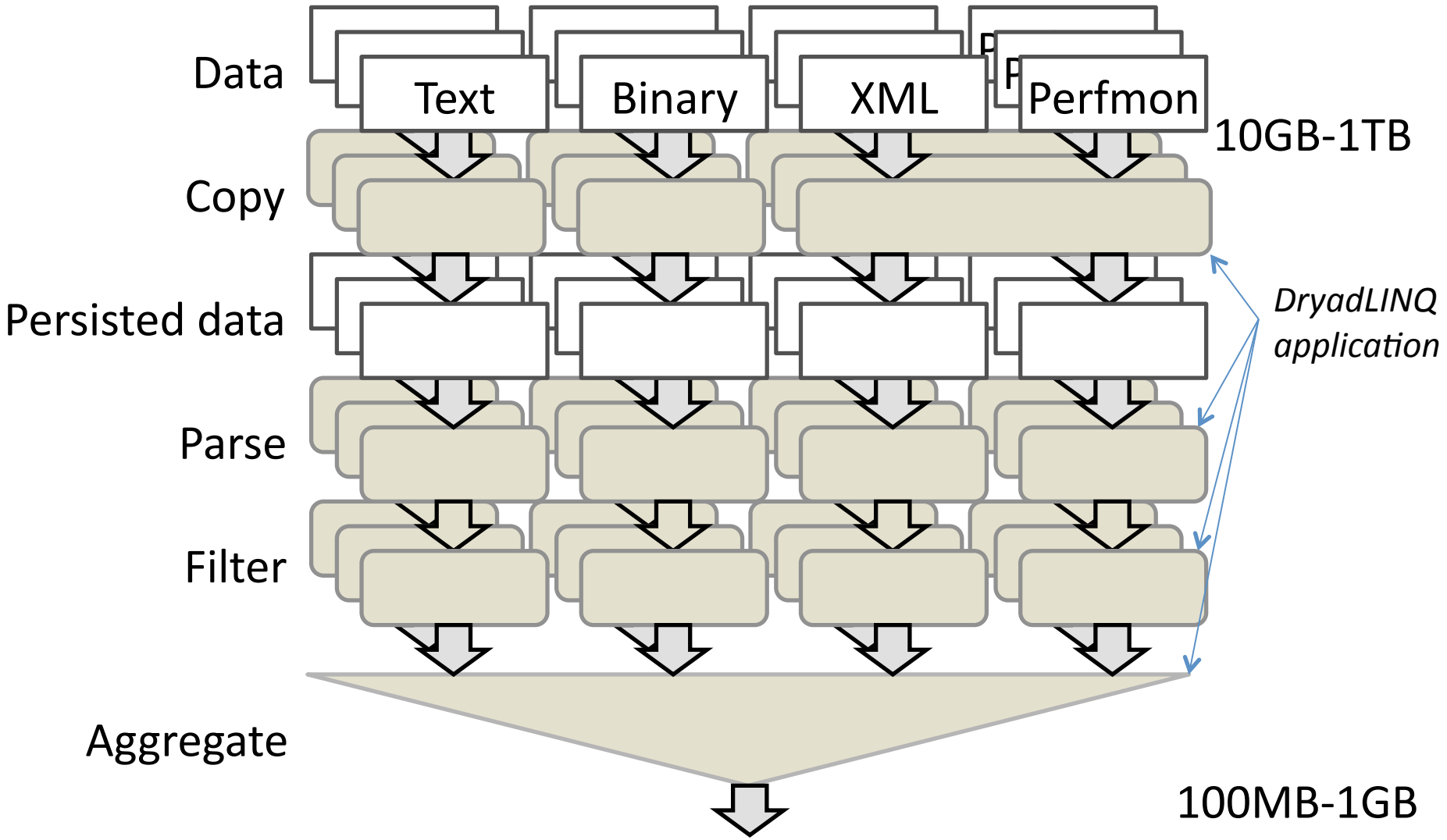
Dryad System Architecture



Data Collection

pptPlex Section Divider

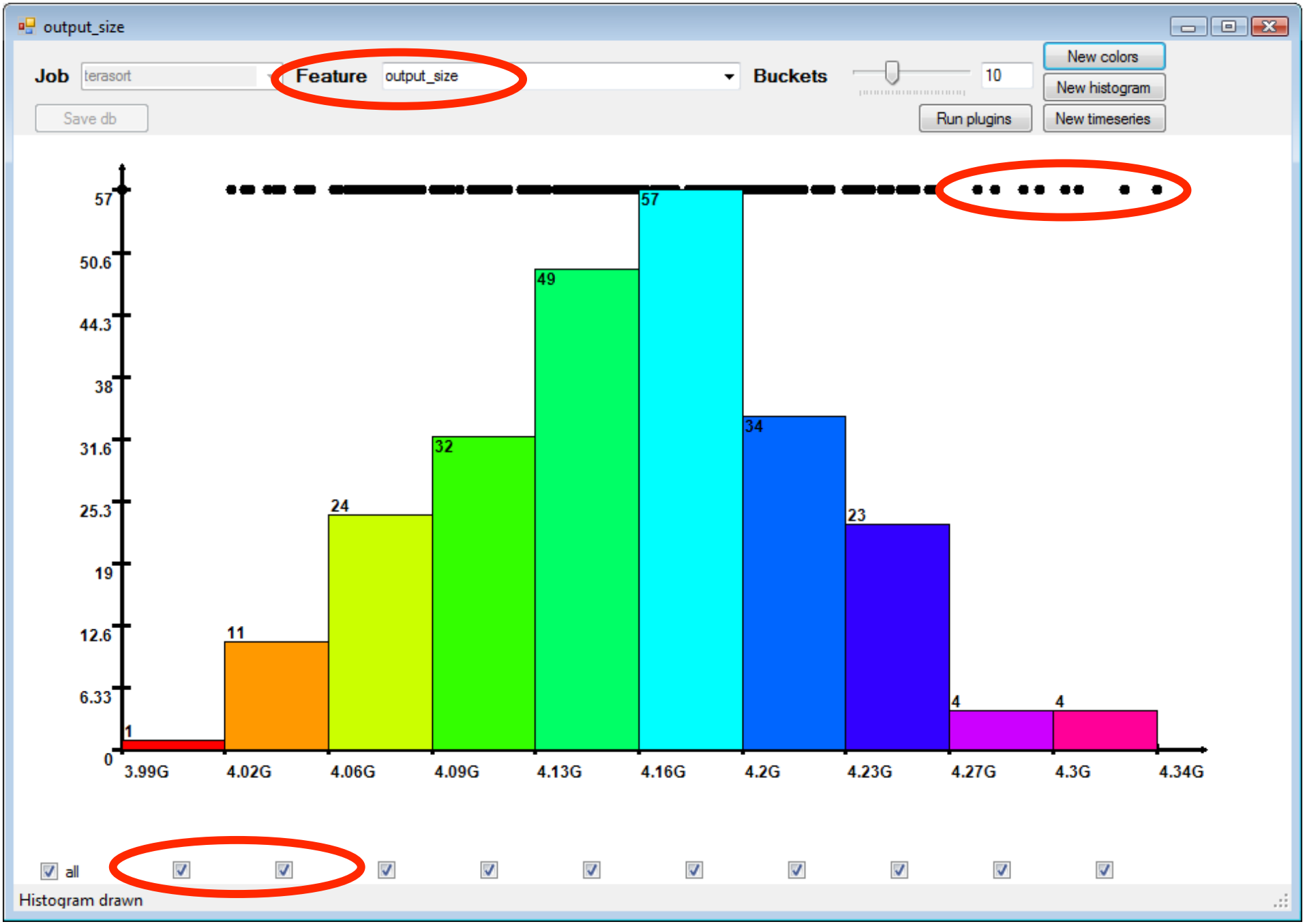
The slides after this divider will be grouped into a section and given the label you type above. Feel free to move this slide to any position in the deck.

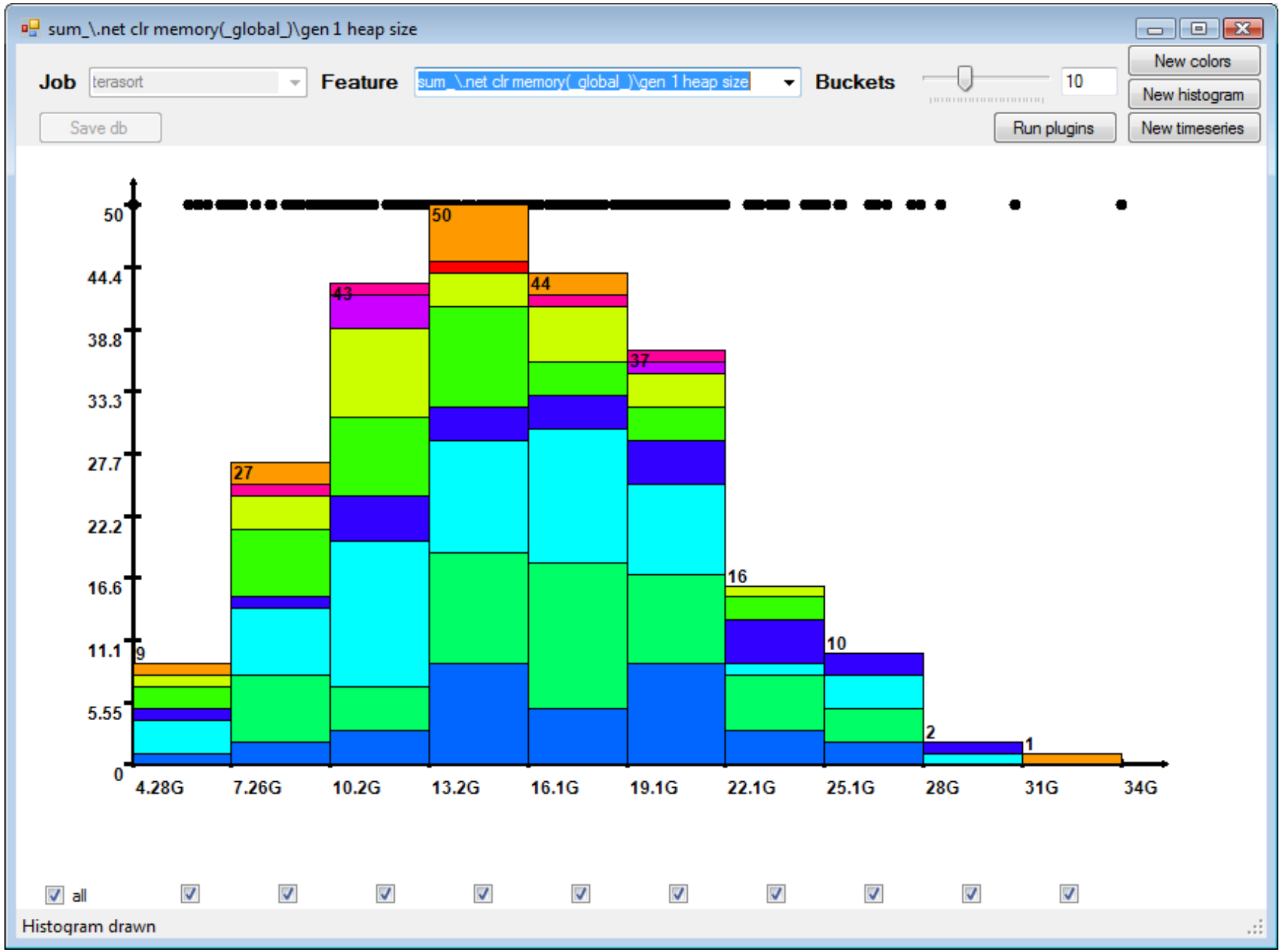

Data
Text
Binary
XML
Perfmon
10GB-1TB
Copy
Persisted data
DryadLINQ
application
Parse
Filter
Aggregate
100MB-1GB

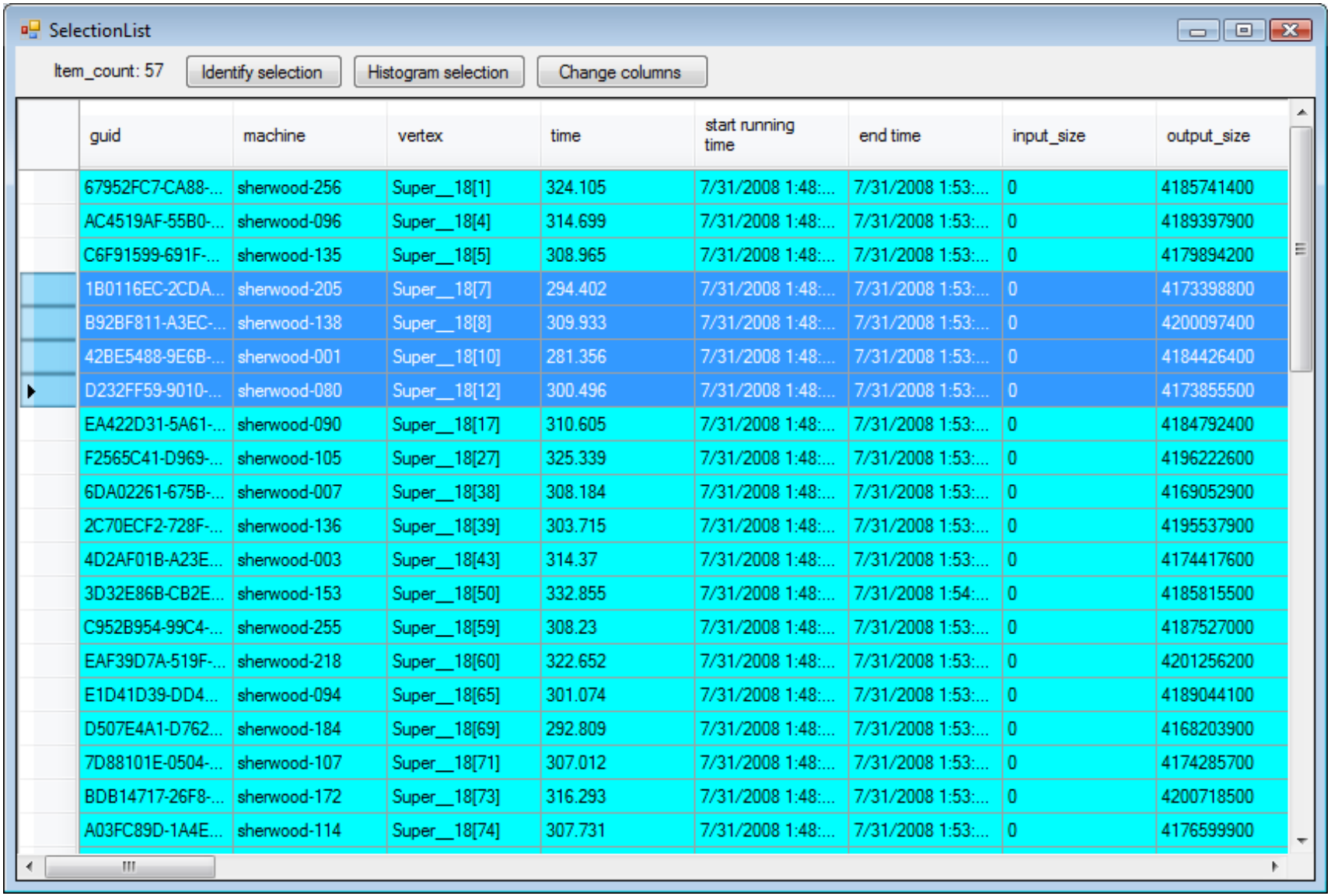
GUI

pptPlex Section Divider

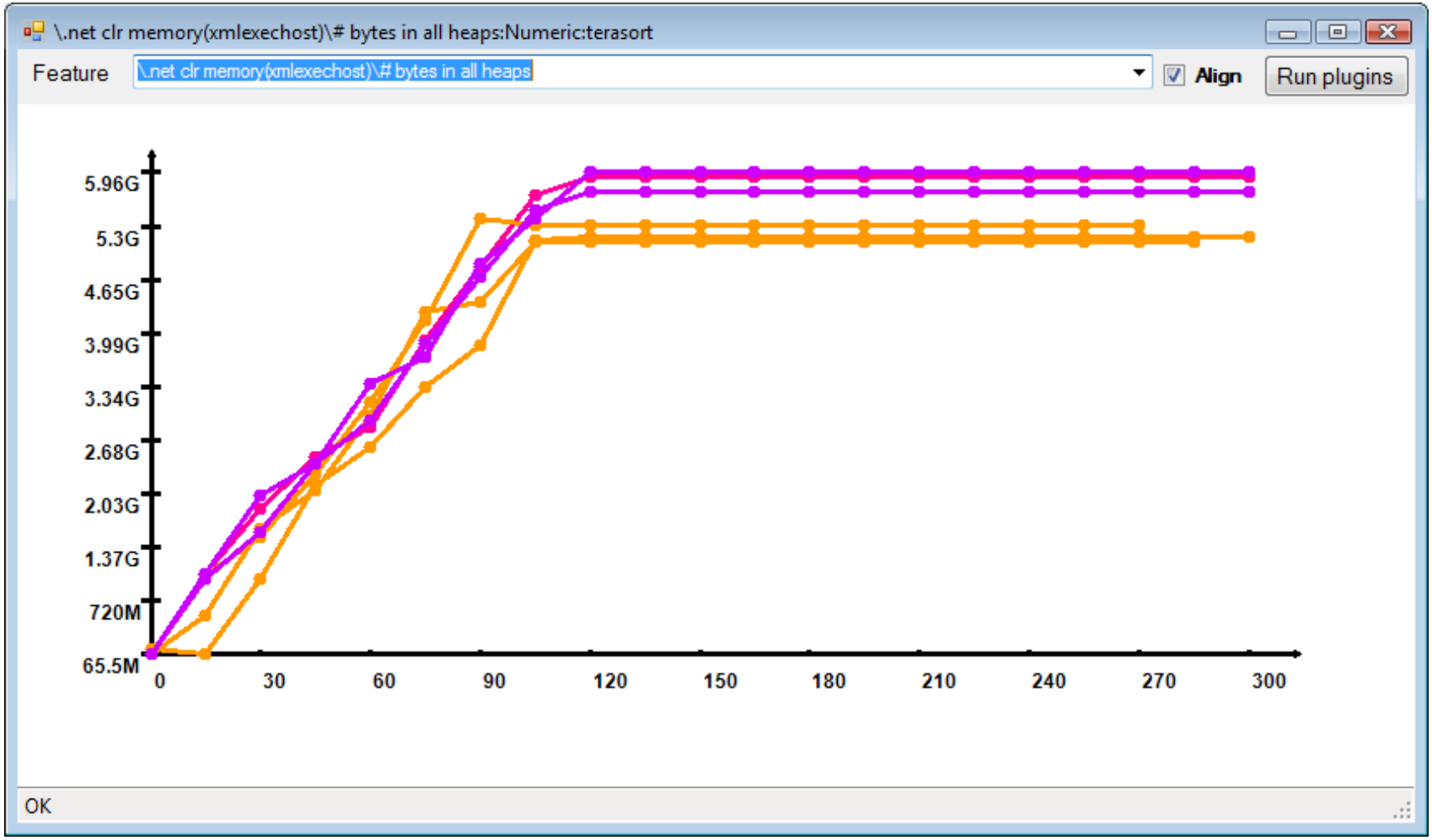
The slides after this divider will be grouped into a section and given the label you type above. Feel free to move this slide to any position in the deck.


output_size
Job
terasort
Feature
output_size
Buckets
10
New colors
New histogram
Run plugins
New timeseries
Save db
57
50.6
44.3
38
31.6
25.3
19
12.6
6.33
0
1
11
24
32
49
57
34
23
4
4
3.99G
4.02G
4.06G
4.09G
4.13G
4.16G
4.2G
4.23G
4.27G
4.3G
4.34G
all
Histogram drawn


sum_\.net clr memory(_global_)\gen 1 heap size
Job
terasort
Feature
sum_\.net clr memory(_global_)\gen 1 heap size
Buckets
10
New colors
New histogram
New timeseries
Save db
Run plugins
50
44.4
38.8
33.3
27.7
22.2
16.6
11.1
5.55
0
9
27
43
50
44
37
16
10
2
1
4.28G
7.26G
10.2G
13.2G
16.1G
19.1G
22.1G
25.1G
28G
31G
34G
all
Histogram drawn



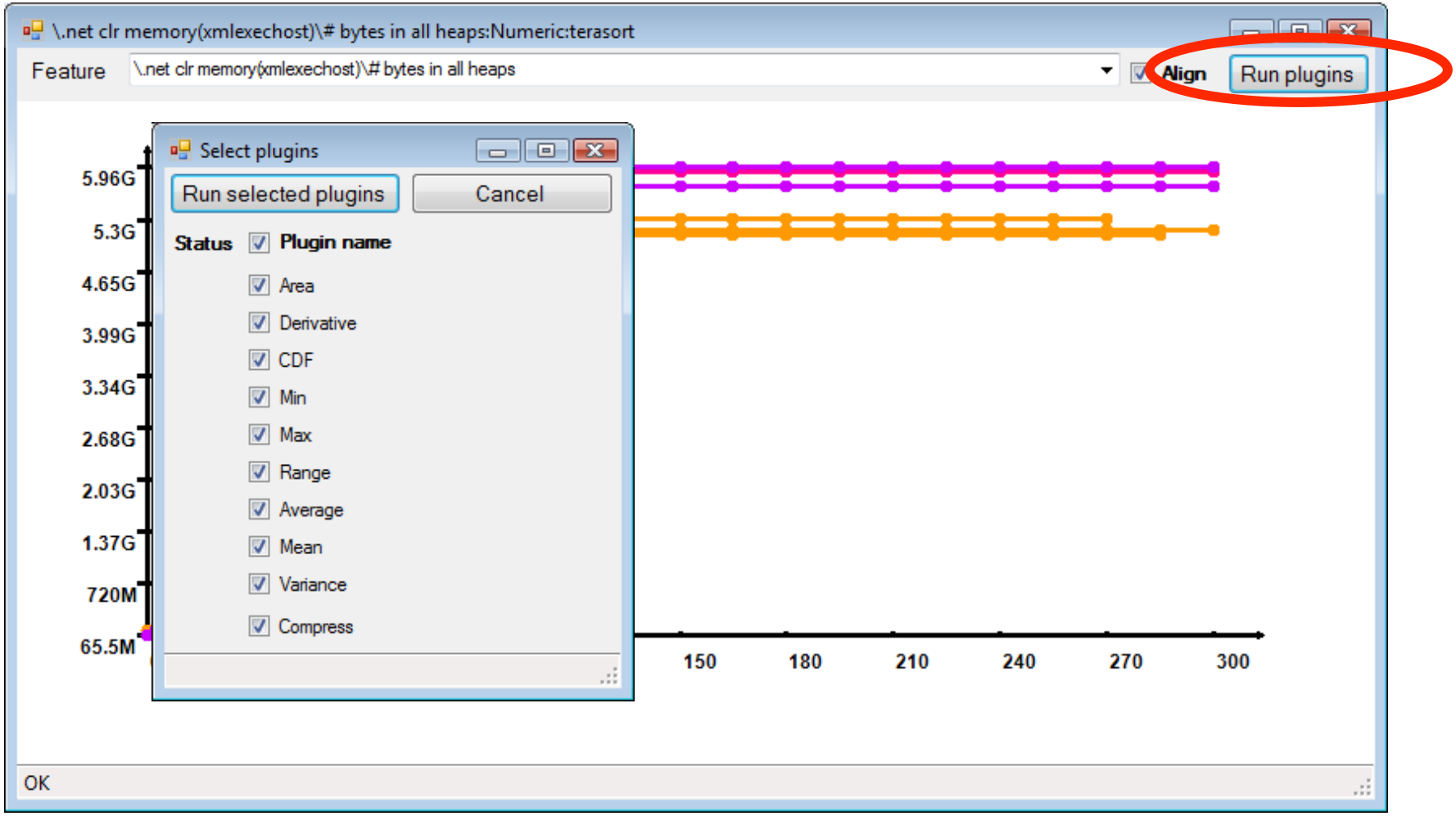
guid	machine	vertex	time	start running time	end time	input_size	output_size
67952FC7-CA88-...	sherwood-256	Super__18[1]	324.105	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4185741400
AC4519AF-55B0-...	sherwood-096	Super__18[4]	314.699	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4189397900
C6F91599-691F-...	sherwood-135	Super__18[5]	308.965	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4179894200
1B0116EC-2CDA...	sherwood-205	Super__18[7]	294.402	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4173398800
B92BF811-A3EC-...	sherwood-138	Super__18[8]	309.933	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4200097400
42BE5488-9E6B-...	sherwood-001	Super__18[10]	281.356	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4184426400
D232FF59-9010-...	sherwood-080	Super__18[12]	300.496	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4173855500
EA422D31-5A61-...	sherwood-090	Super__18[17]	310.605	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4184792400
F2565C41-D969-...	sherwood-105	Super__18[27]	325.339	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4196222600
6DA02261-675B-...	sherwood-007	Super__18[38]	308.184	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4169052900
2C70ECF2-728F-...	sherwood-136	Super__18[39]	303.715	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4195537900
4D2AF01B-A23E...	sherwood-003	Super__18[43]	314.37	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4174417600
3D32E86B-CB2E...	sherwood-153	Super__18[50]	332.855	7/31/2008 1:48:...	7/31/2008 1:54:...	0	4185815500
C952B954-99C4-...	sherwood-255	Super__18[59]	308.23	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4187527000
EAF39D7A-519F-...	sherwood-218	Super__18[60]	322.652	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4201256200
E1D41D39-DD4...	sherwood-094	Super__18[65]	301.074	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4189044100
D507E4A1-D762...	sherwood-184	Super__18[69]	292.809	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4168203900
7D88101E-0504-...	sherwood-107	Super__18[71]	307.012	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4174285700
BDB14717-26F8-...	sherwood-172	Super__18[73]	316.293	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4200718500
A03FC89D-1A4E...	sherwood-114	Super__18[74]	307.731	7/31/2008 1:48:...	7/31/2008 1:53:...	0	4176599900


\.net clr memory(xmlexechost)\# bytes in all heaps:Numeric:terasort
Feature
\.net clr memory(xmlexechost)\# bytes in all heaps
Align
Run plugins
5.96G
5.3G
4.65G
3.99G
3.34G
2.68G
2.03G
1.37G
720M
65.5M
0
30
60
90
120
150
180
210
240
270
300
OK

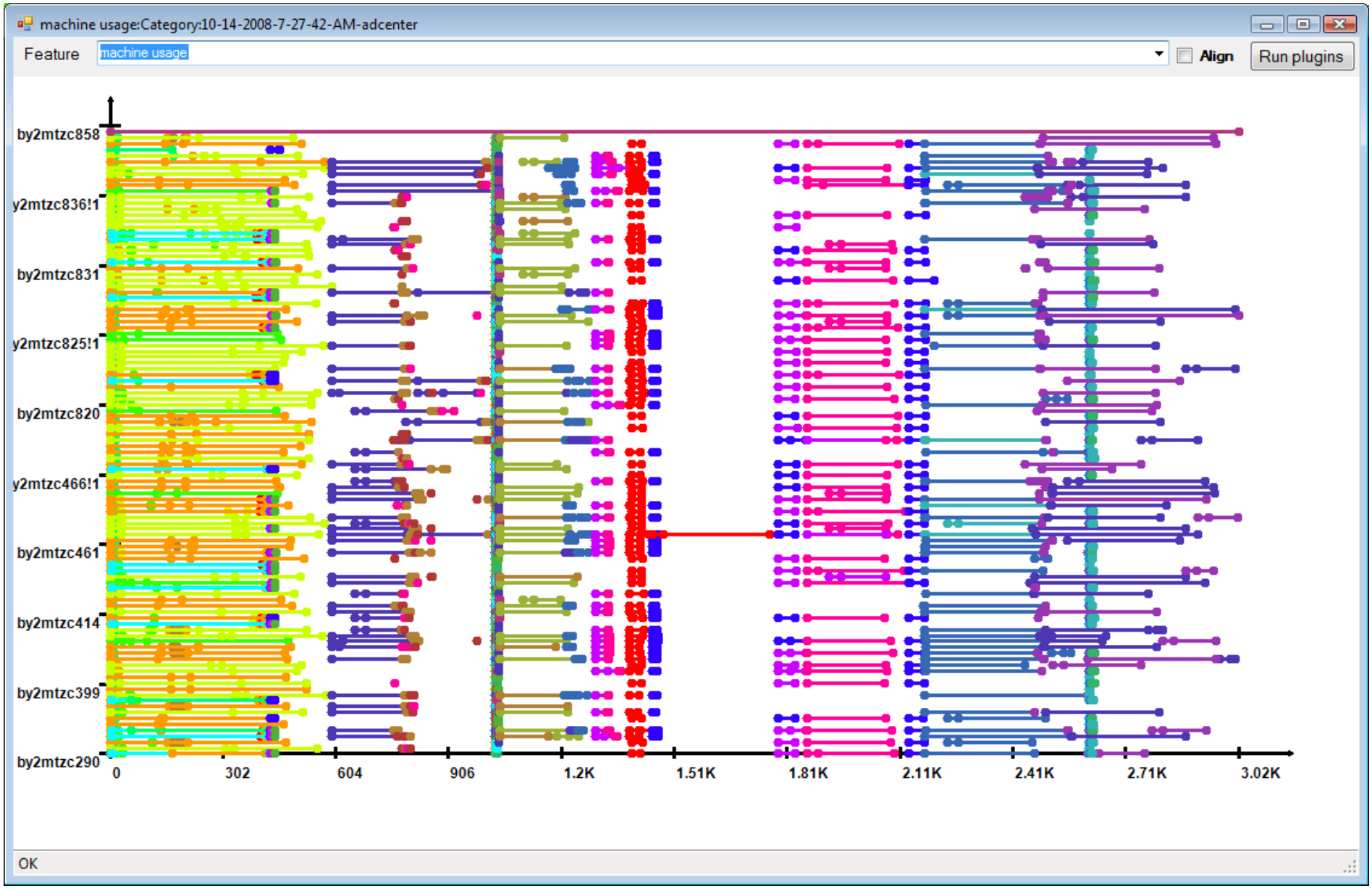
Plug-ins

pptPlex Section Divider

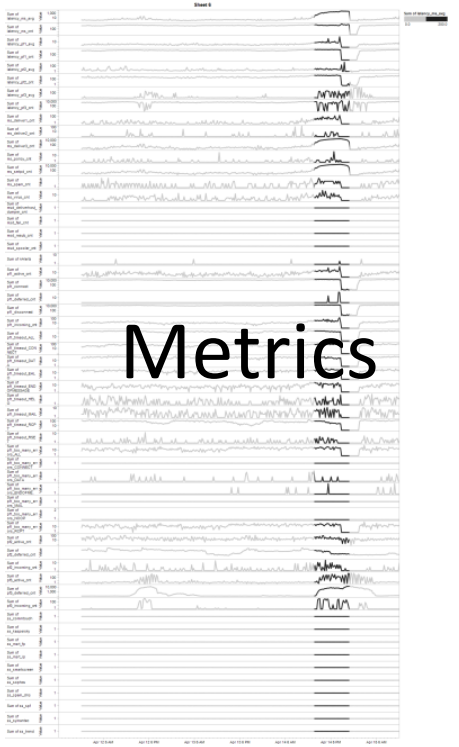
The slides after this divider will be grouped into a section and given the label you type above. Feel free to move this slide to any position in the deck.


\.net clr memory(xmlexechost)\# bytes in all heaps:Numeric:terasort
Feature
\.net clr memory(xmlexechost)\# bytes in all heaps
Align
Run plugins
Select plugins
Run selected plugins
Cancel
Status
Plugin name
Area
Derivative
CDF
Min
Max
Range
Average
Mean
Variance
Compress
5.96G
5.3G
4.65G
3.99G
3.34G
2.68G
2.03G
1.37G
720M
65.5M
150
180
210
240
270
300
OK

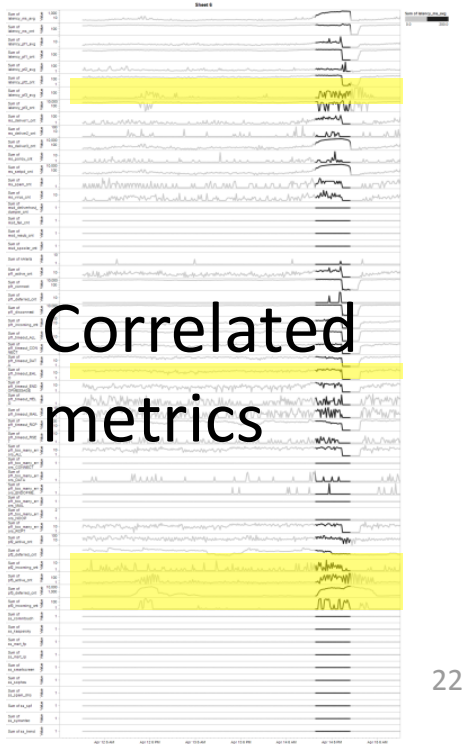
Machine Utilization Plug-in



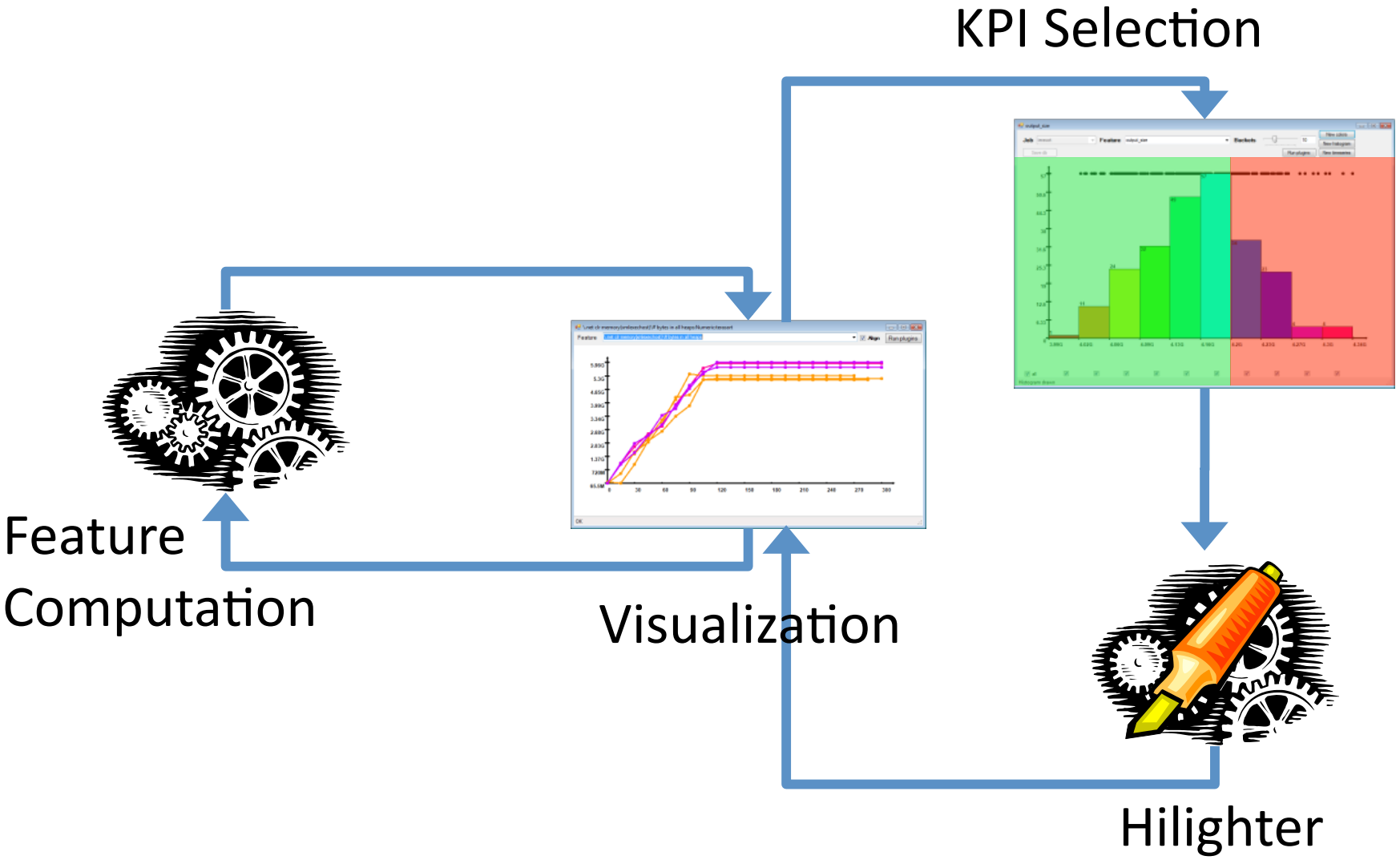
Complex statistics: HiLighter plug-in



Binary search over
logistic regression with
L1 regularization

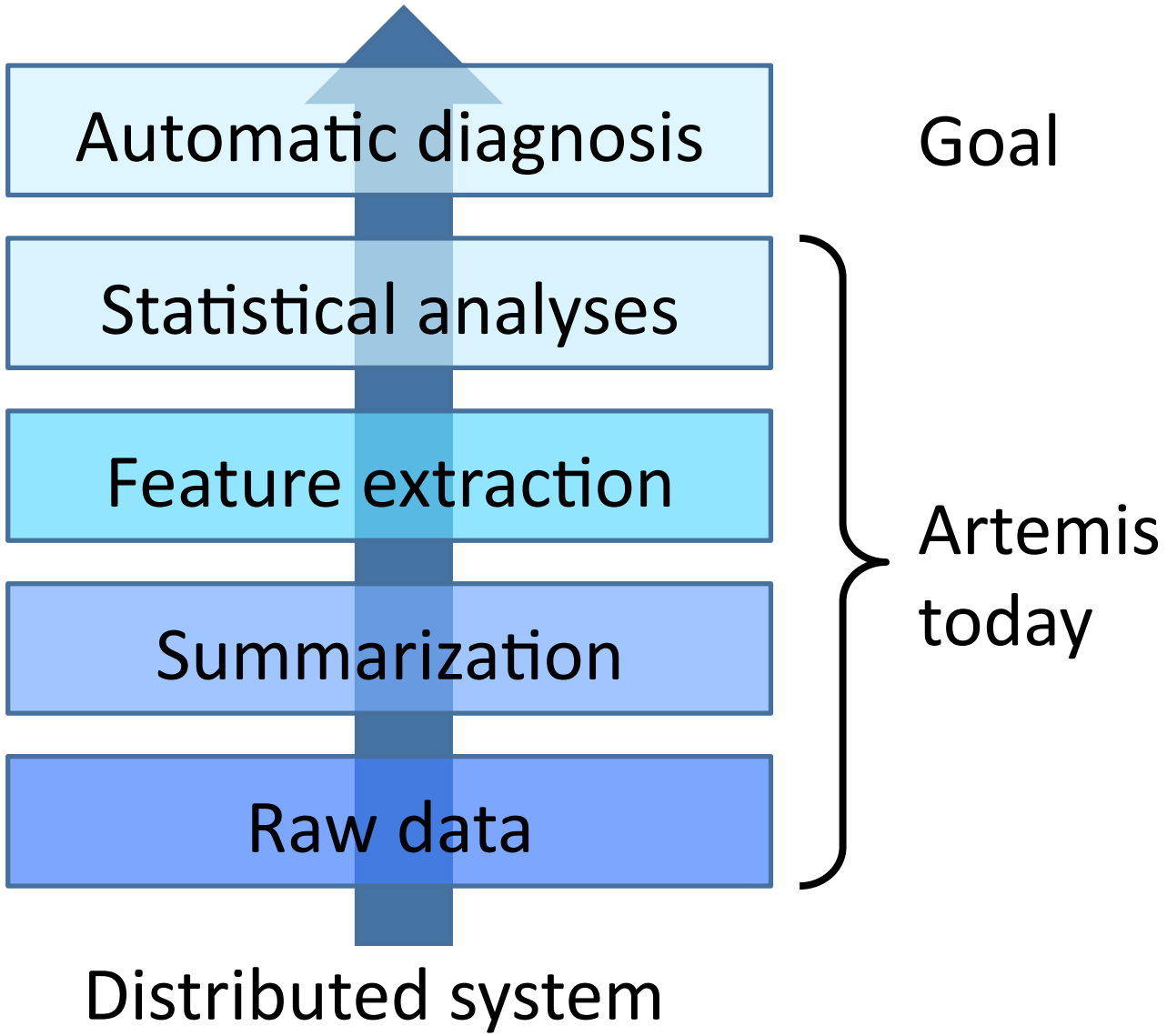


Interactive Analysis



Conclusions

pptPlex Section Divider

The slides after this divider will be grouped into a section and given the label you type above. Feel free to move this slide to any position in the deck.

Conclusions