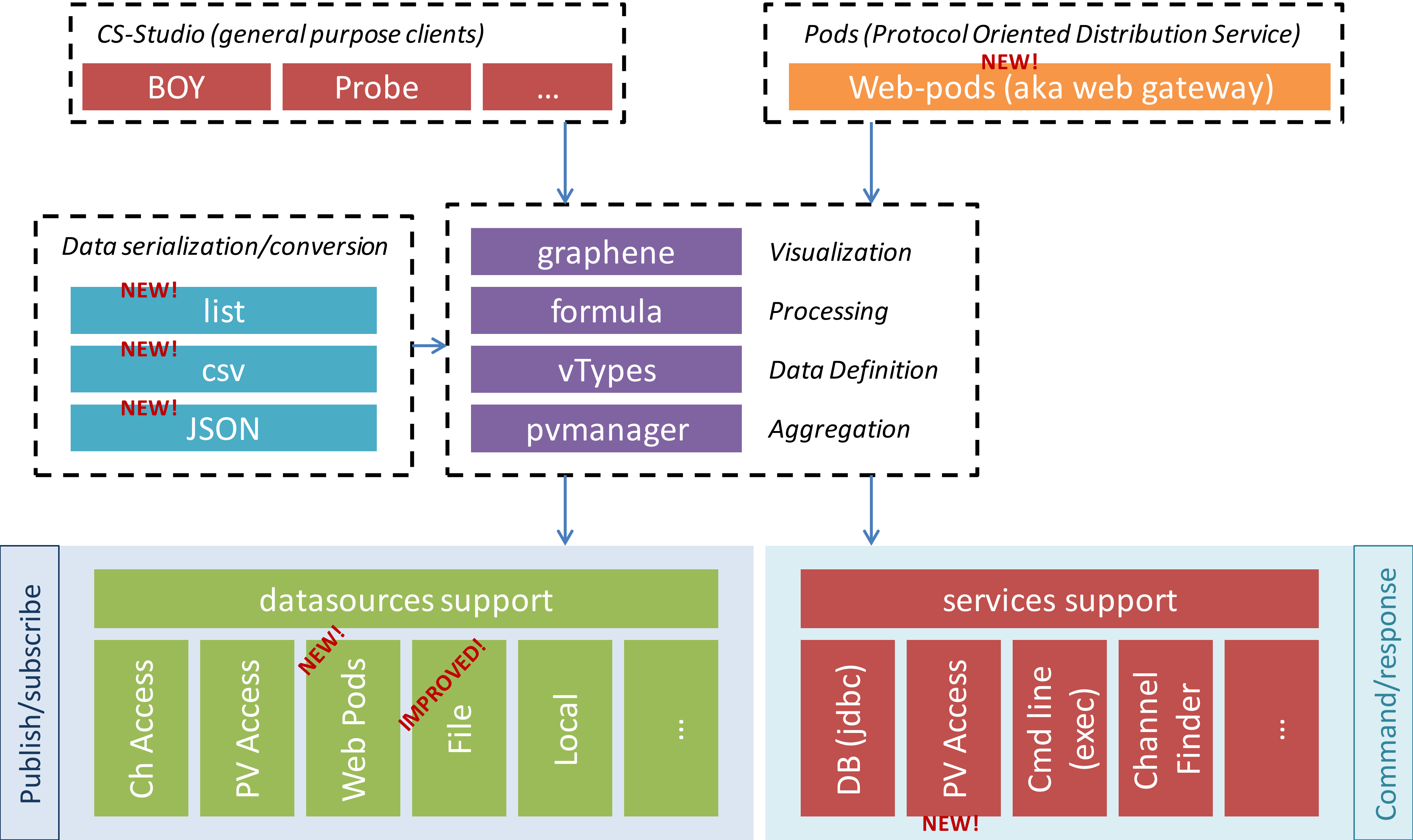
diirt
data integration in real time

Gabriele Carcassi

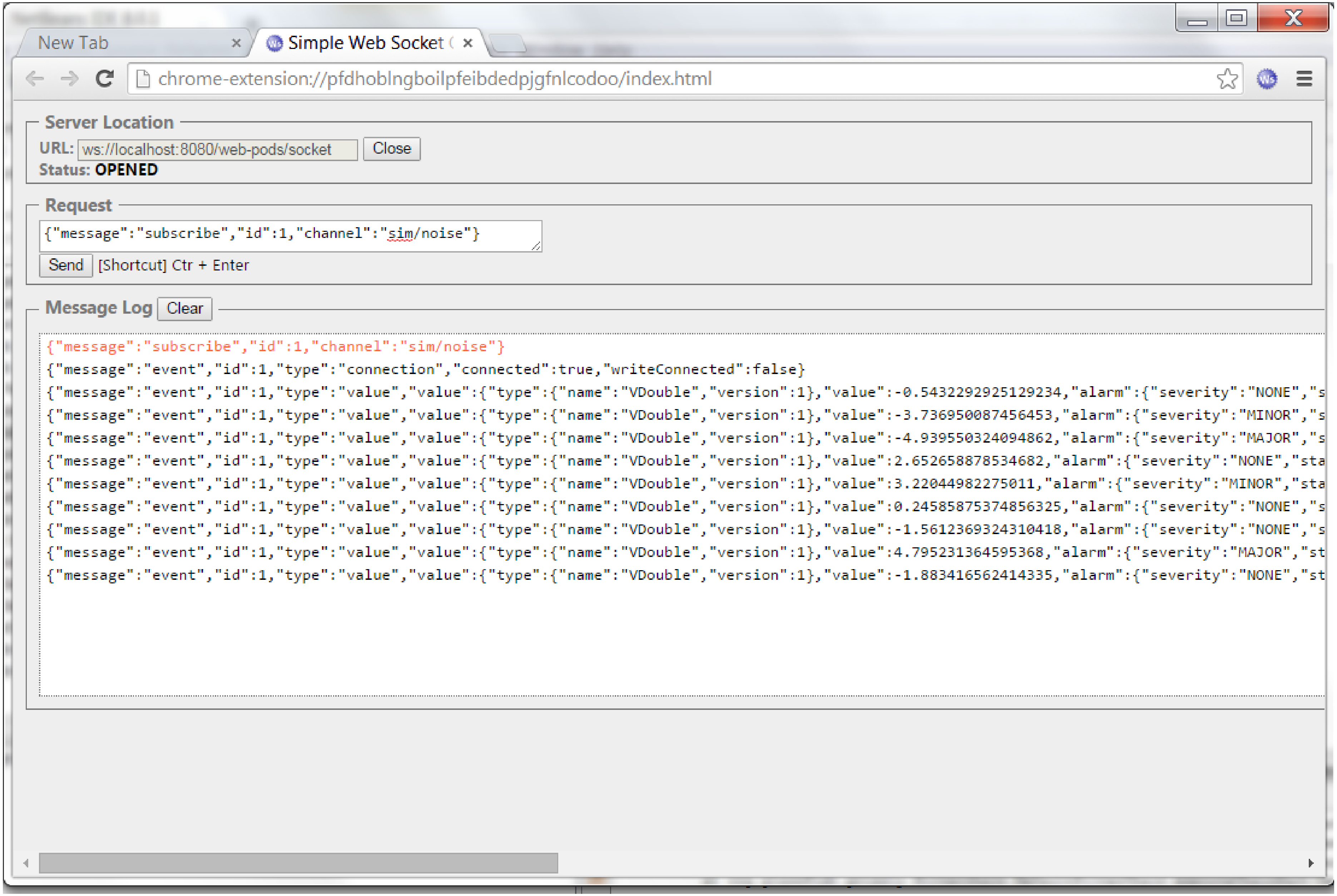
diirt: data integration in real-time

- A rebranding of pvmanager/graphene/vtypes as a separate layer
  - http://diirt.org
- Usage growing outside of CS-Studio, focus on it on its own
  - Improve coordination, modularity, configuration …
- Main two clients:
  - CS-Studio (Java based Eclipse RCP)
  - Web Pods NEW! (Web-socket + JSON “gateway”)

Data integration in real time (Diirt) architecture



WEB PODS



chrome-extension://pfdhoblngboilpfeibdedpjgfnlcodoo/index.html

Server Location

URL: ws://localhost:8080/web-pods/socket

Status: OPENED

Request

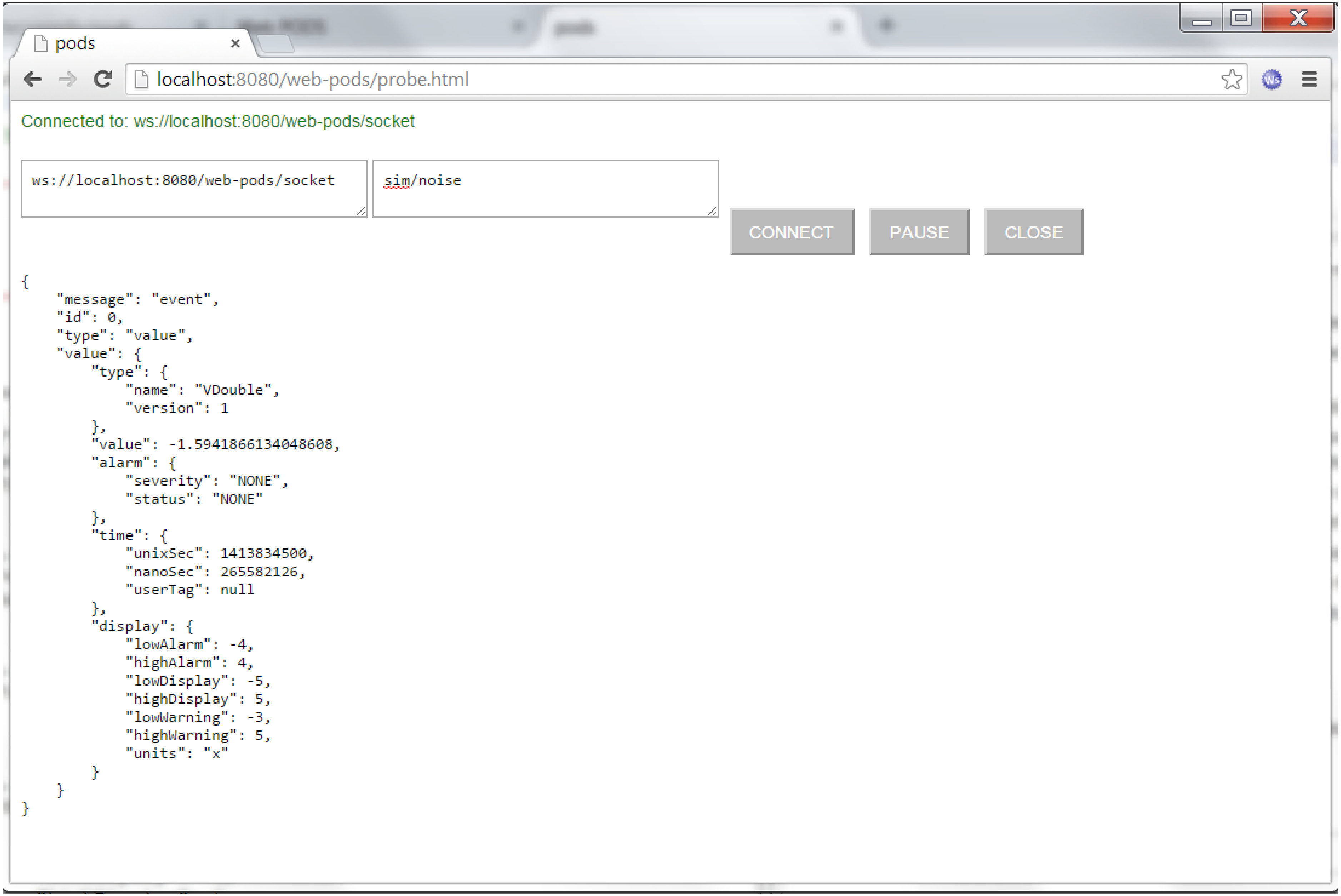
```
{"message":"subscribe","id":1,"channel":"sim/noise"}
```

[Shortcut] Ctr + Enter

Message Log

```
{"message":"subscribe","id":1,"channel":"sim/noise"}
{"message":"event","id":1,"type":"connection","connected":true,"writeConnected":false}
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":-0.5432292925129234,"alarm":{"severity":"NONE","s
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":-3.736950087456453,"alarm":{"severity":"MINOR","s
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":-4.939550324094862,"alarm":{"severity":"MAJOR","s
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":2.652658878534682,"alarm":{"severity":"NONE","sta
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":3.22044982275011,"alarm":{"severity":"MINOR","sta
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":0.2458587537485632 5,"alarm":{"severity":"NONE","s
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":-1.5612369324310418,"alarm":{"severity":"NONE","s
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":4.795231364595368,"alarm":{"severity":"MAJOR","st
{"message":"event","id":1,"type":"value","value":{"type":{"name":"VDouble","version":1},"value":-1.883416562414335,"alarm":{"severity":"NONE","st
```

Chrome extension: Simple Web Socket Client



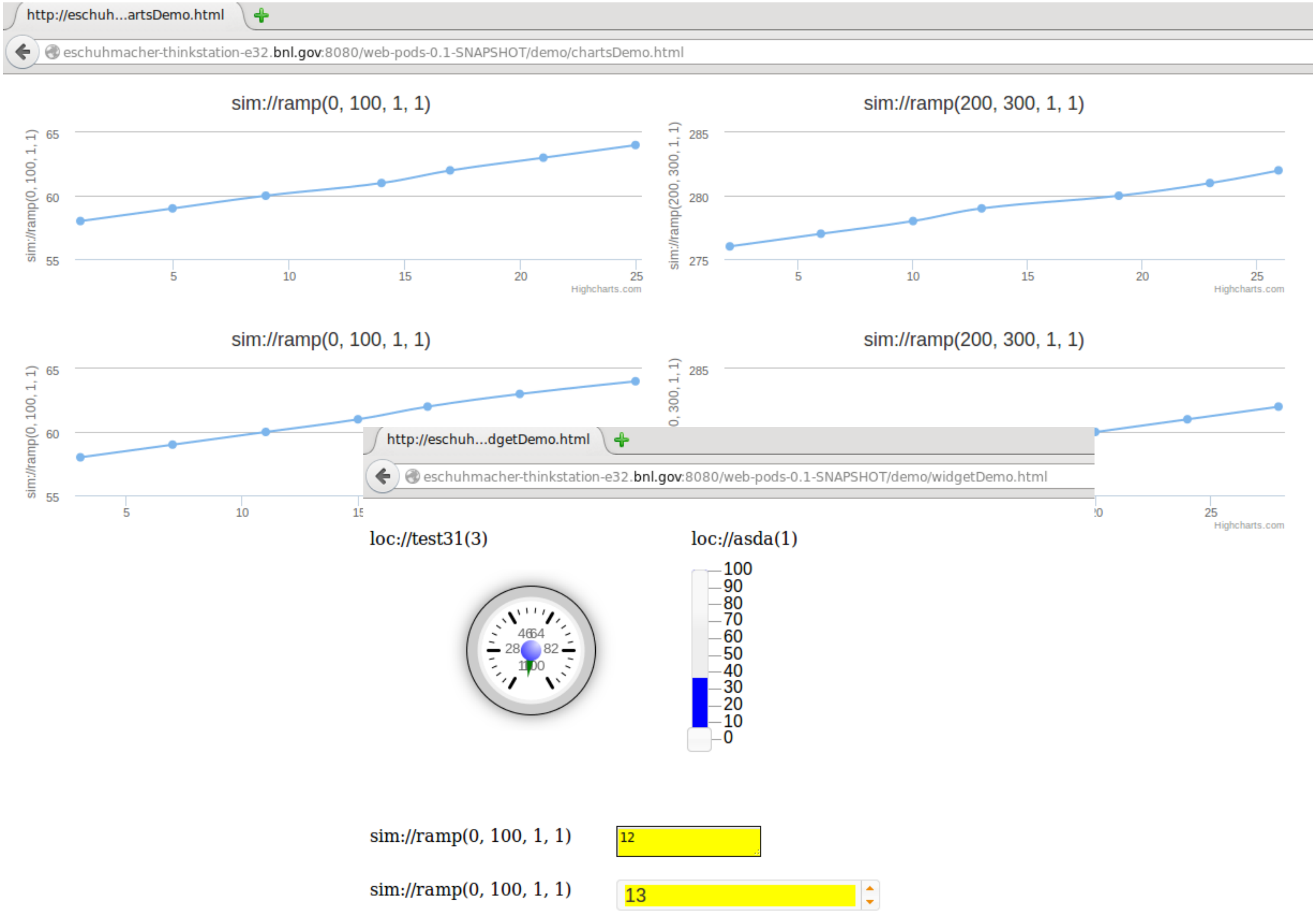
Connected to: ws://localhost:8080/web-pods/socket

ws://localhost:8080/web-pods/socket

sim/noise

```
{
    "message": "event",
    "id": 0,
    "type": "value",
    "value": {
        "type": {
            "name": "VDouble",
            "version": 1
        },
        "value": -1.5941866134048608,
        "alarm": {
            "severity": "NONE",
            "status": "NONE"
        },
        "time": {
            "unixSec": 1413834500,
            "nanoSec": 265582126,
            "userTag": null
        },
        "display": {
            "lowAlarm": -4,
            "highAlarm": 4,
            "lowDisplay": -5,
            "highDisplay": 5,
            "lowWarning": -3,
            "highWarning": 5,
            "units": "x"
        }
    }
}
```

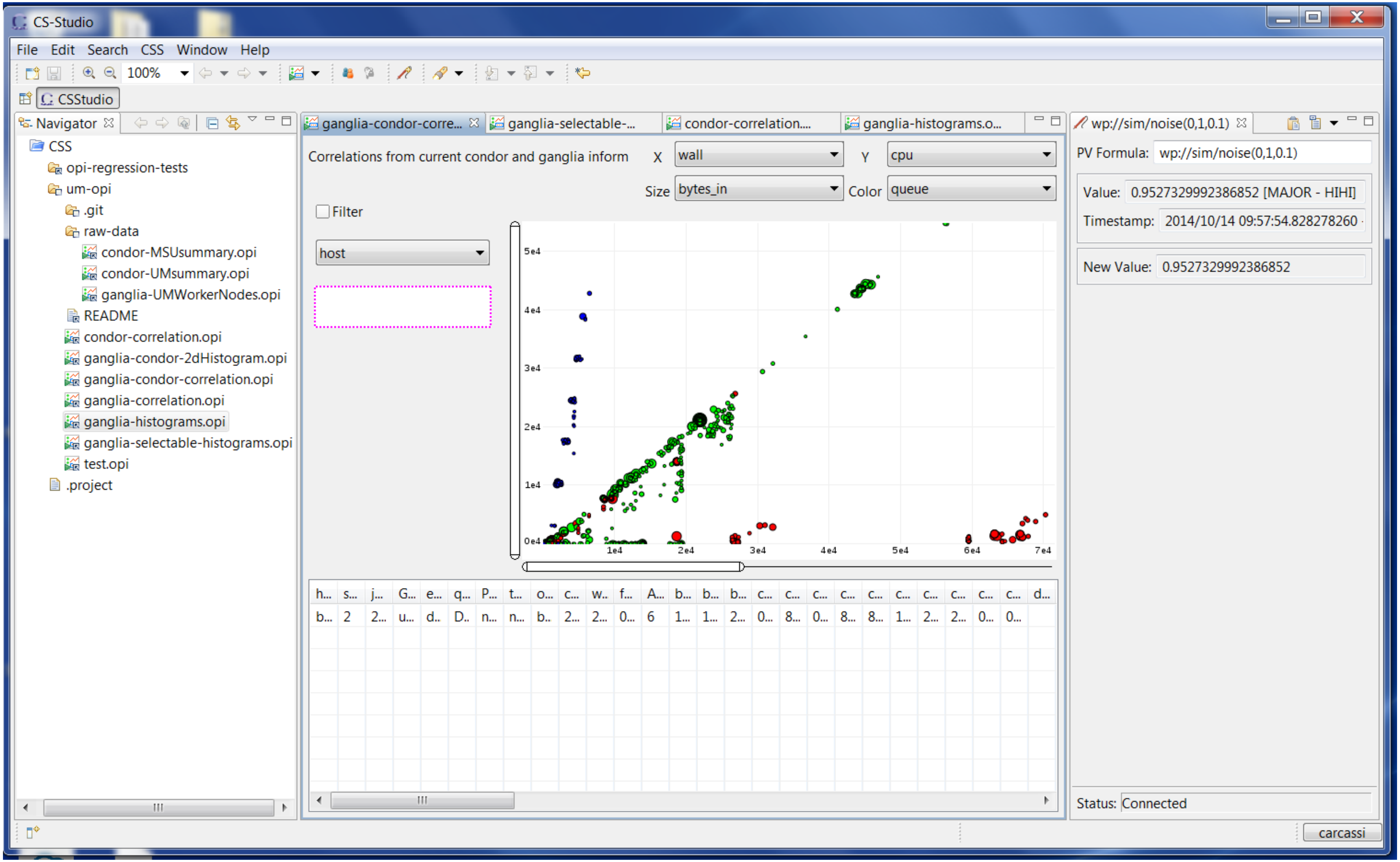
Test javascript client (Danielle Connolly, UMich)


http://eschuh...artsDemo.html
eschuhmacher-thinkstation-e32.bnl.gov:8080/web-pods-0.1-SNAPSHOT/demo/chartsDemo.html
sim://ramp(0, 100, 1, 1)
sim://ramp(0, 100, 1, 1)
65
60
55
5
10
15
20
25
Highcharts.com
sim://ramp(200, 300, 1, 1)
sim://ramp(200, 300, 1, 1)
285
280
275
5
10
15
20
25
Highcharts.com
sim://ramp(0, 100, 1, 1)
sim://ramp(0, 100, 1, 1)
65
60
55
5
10
15
sim://ramp(200, 300, 1, 1)
0, 300, 1, 1)
285
20
25
Highcharts.com
http://eschuh...dgetDemo.html
eschuhmacher-thinkstation-e32.bnl.gov:8080/web-pods-0.1-SNAPSHOT/demo/widgetDemo.html
loc://test31(3)
46
64
28
82
10
100
loc://asda(1)
100
90
80
70
60
50
40
30
20
10
0

sim://ramp(0, 100, 1, 1) 12

sim://ramp(0, 100, 1, 1) 13

GWT client (Enrique Schuhmacher, BNL)



Correlations from current condor and ganglia inform

X: wall

Y: cpu

Size: bytes_in

Color: queue

☐ Filter

host

h...	s...	j...	G...	e...	q...	P...	t...	o...	c...	w..	f...	A...	b...	b...	b...	c...	c...	c...	c...	c...	c...	c...	c...	c...	c...	d...
b...	2	2...	u...	d..	D..	n...	n...	b..	2...	2...	0...	6	1...	1...	2...	0...	8...	0...	8...	8...	1...	2...	2...	0...	0...	

PV Formula: wp://sim/noise(0,1,0.1)

Value: 0.9527329992386852 [MAJOR - HIHI]

Timestamp: 2014/10/14 09:57:54.828278260

New Value: 0.9527329992386852

Status: Connected

Using CS-Studio off site through Web Pods

Web Pods server configuration

```
[xxx@diirt ~]# more .diirt/pods/web/mappings.xml
<?xml version='1.0' encoding='UTF-8'?>
<mappings version="1">
    <mapping channel="calibration/constants"
substitution="file:///path/to/file/table.csv"
permission="READ_ONLY"/>
    <mapping channel="public-.*" permission="READ_WRITE"/>
    <mapping channel="cf-tag-(.\w)" substitution="=cfQuery($1)"
permission="READ_ONLY"/>
    <mapping channel="sim/(.*)" substitution="sim://$1"
permission="READ_ONLY"/>
</mappings>
```

Access security planned, not yet implemented

- Use wss (like https) for authentication
- Use username/role/unix group/host for authorization

Be careful not to expose too much

Web Pods datasource configuration

```
[xxx@diirt ~]# more .diirt/datasources/wp/wp.xml
<?xml version='1.0' encoding='UTF-8'?>
<wp version="1">
    <!-- Change uri to the location of the web-pods socket -->
    <!--<connection socketLocation="ws://localhost:8080/web-
pods/socket"/>-->
    <connection socketLocation="ws://diirt.mysite.org/web-
pods/socket"/>
</wp>
```

Web Pods

- Rebroadcast data using Web sockets and JSON
  - Play nice with firewalls, get WAN notifications, data available to web tools, server is one class (no logic, a pass-through to pvmanager), should scale (different sockets on different servers, not tested)
  - Not a substitute to CA/PVA. JSON serialization problems (verbose, does not follow IEEE standards for double, no “delta” updates). Not intended for heavyweight applications.
- Possible uses
  - Active web pages
  - WAN gateway for CS-Studio and other tools
  - Easy way to make data available to CS-Studio and other tools
    - Setup a directory shared through web-pods
    - Write csv or jvtype files with scripts or text editors (updates are sent)
- Something similar could be done with PVA (Pva Pods)
  - Already designed pieces (e.g. mapping logic) to be reusable

PVMANAGER

pvmanager

- Performance improvements
  - Implemented “passive scan”
    - 65536 readers goes from ~25% to ~0% load (as reported by Windows Task Manager)
    - Still needs to be implemented for formulas
  - Faster channel creation
    - 65536 reader from ~28 to ~0.7 seconds
- Troubleshooting improvements
  - Creation stacktrace printed for unclosed readers/writers
  - Thorough debugging log for the CA datasource

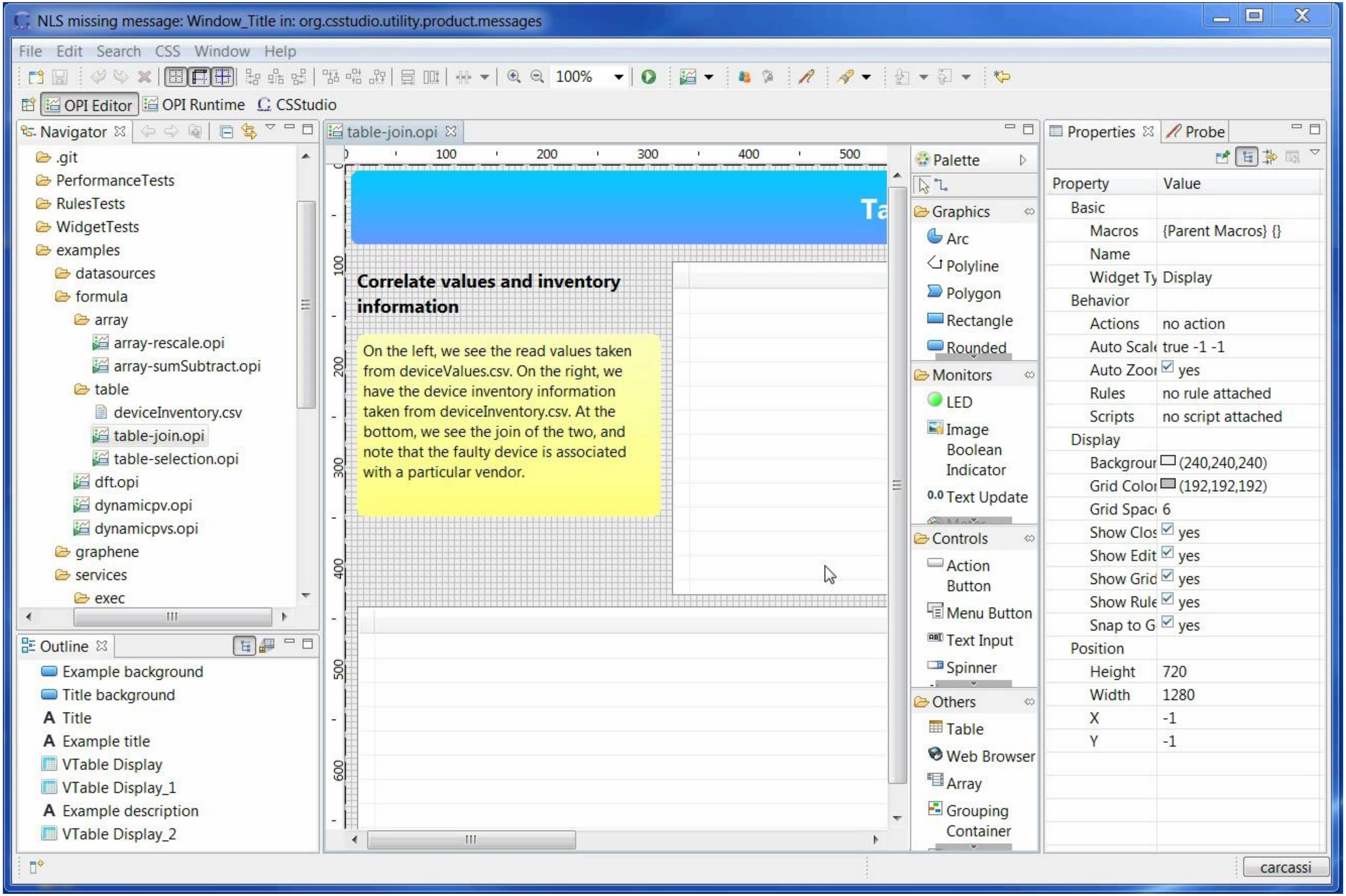
pvmanager

- Minor fixes for CA
  - Timestamp on disconnect, nanosec out of range, reconnect with type change, extra monitor after reconnect
- Increased support for PVA (Matej Sekoranja, Cosylab)
  - Update to 4.4 (as soon as released)
  - Support for services
    - Write an xml file with pva service location, arguments/result, drop it in a directory

pvmanager

- Upcoming changes in configuration
  - All pieces of diirt (datasources, services) will look at files in a common directory for configuration
    - Easier to distribute configuration
    - Same for WebPods, CS-Studio and all diirt clients
      - One set of rules, documentation, options, …
      - Easy to write a script that re-uses CS-Studio configuration (no extra setup required)
    - Default directory is $HOME/.diirt, but can be changed by environment variable or programmatically
      - CS-Studio could redirect it wherever
    - Should appear in CS-Studio 4.1

FORMULA



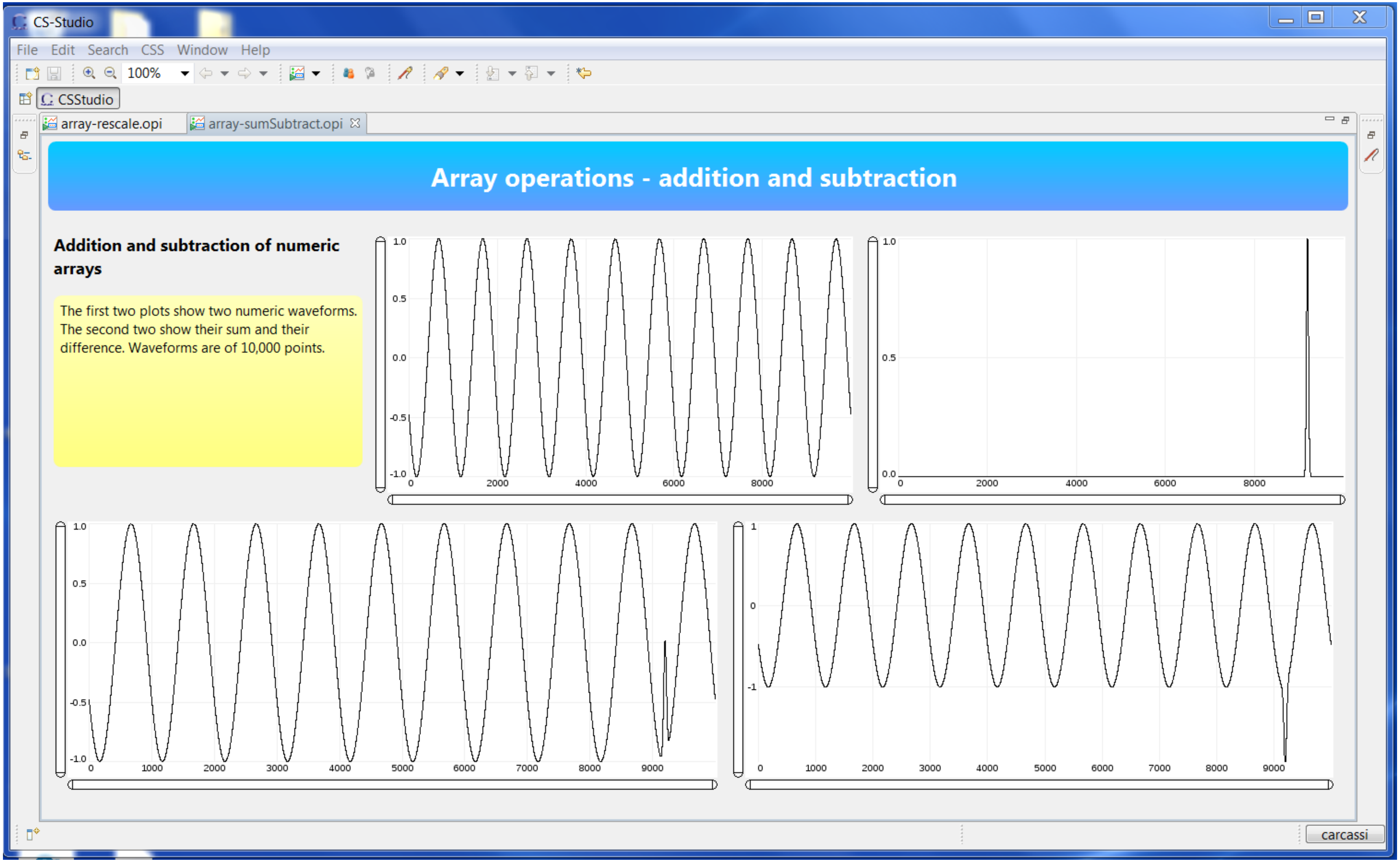
NLS missing message: Window_Title in: org.csstudio.utility.product.messages

Correlate values and inventory information

On the left, we see the read values taken from deviceValues.csv. On the right, we have the device inventory information taken from deviceInventory.csv. At the bottom, we see the join of the two, and note that the faulty device is associated with a particular vendor.

Property	Value
Basic	
Macros	{Parent Macros} {}
Name	
Widget Ty	Display
Behavior	
Actions	no action
Auto Scal	true -1 -1
Auto Zoo	☑ yes
Rules	no rule attached
Scripts	no script attached
Display	
Backgrour	□ (240,240,240)
Grid Colo	▭ (192,192,192)
Grid Spac	6
Show Clos	☑ yes
Show Edit	☑ yes
Show Grid	☑ yes
Show Rule	☑ yes
Snap to G	☑ yes
Position	
Height	720
Width	1280
X	-1
Y	-1

Table join: live pv with inventory information



Array operations - addition and subtraction

Addition and subtraction of numeric arrays

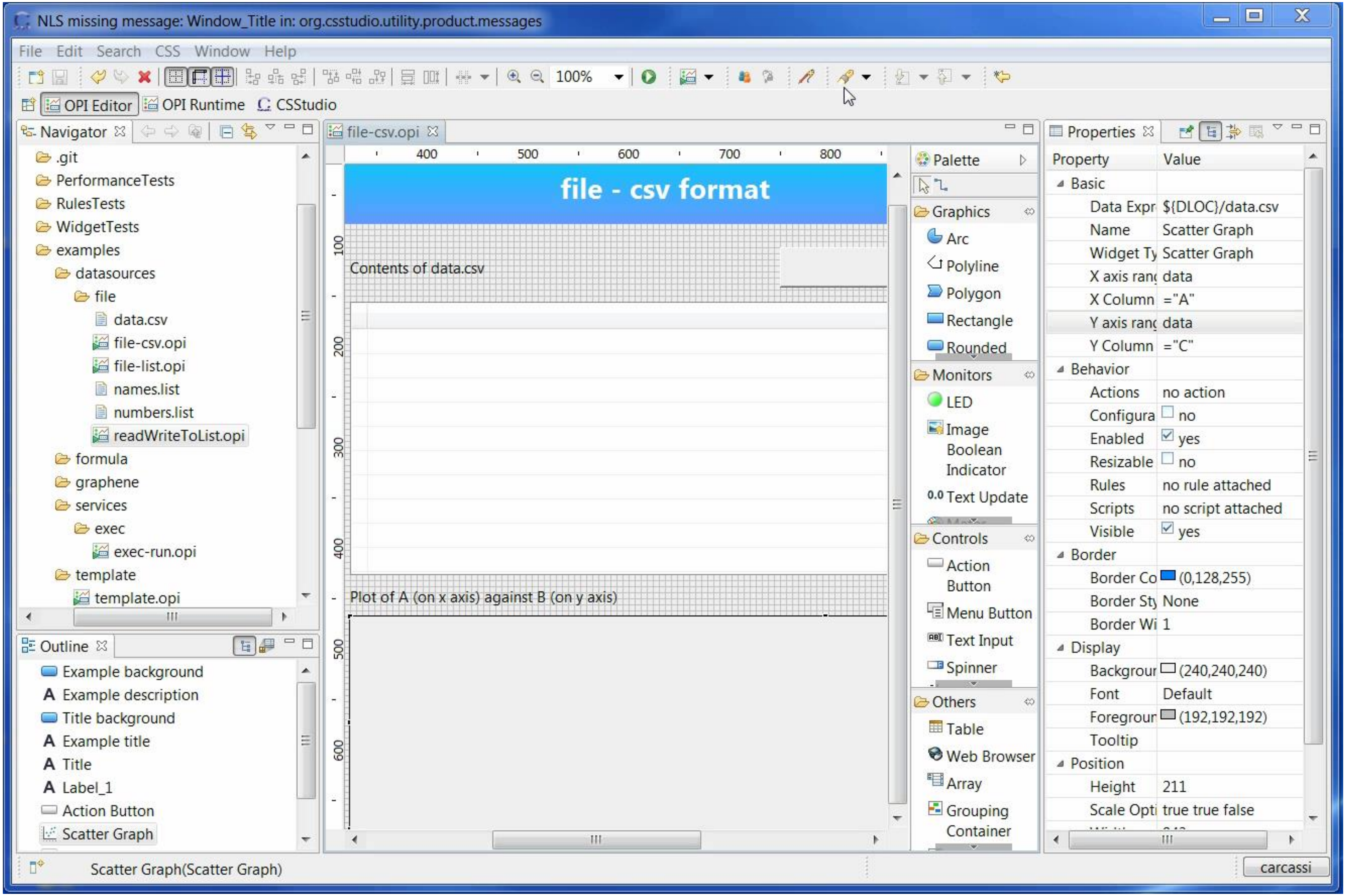
The first two plots show two numeric waveforms. The second two show their sum and their difference. Waveforms are of 10,000 points.

Improved array element-wise operations: add/subtract/multiply/divide/pow/…

Formula

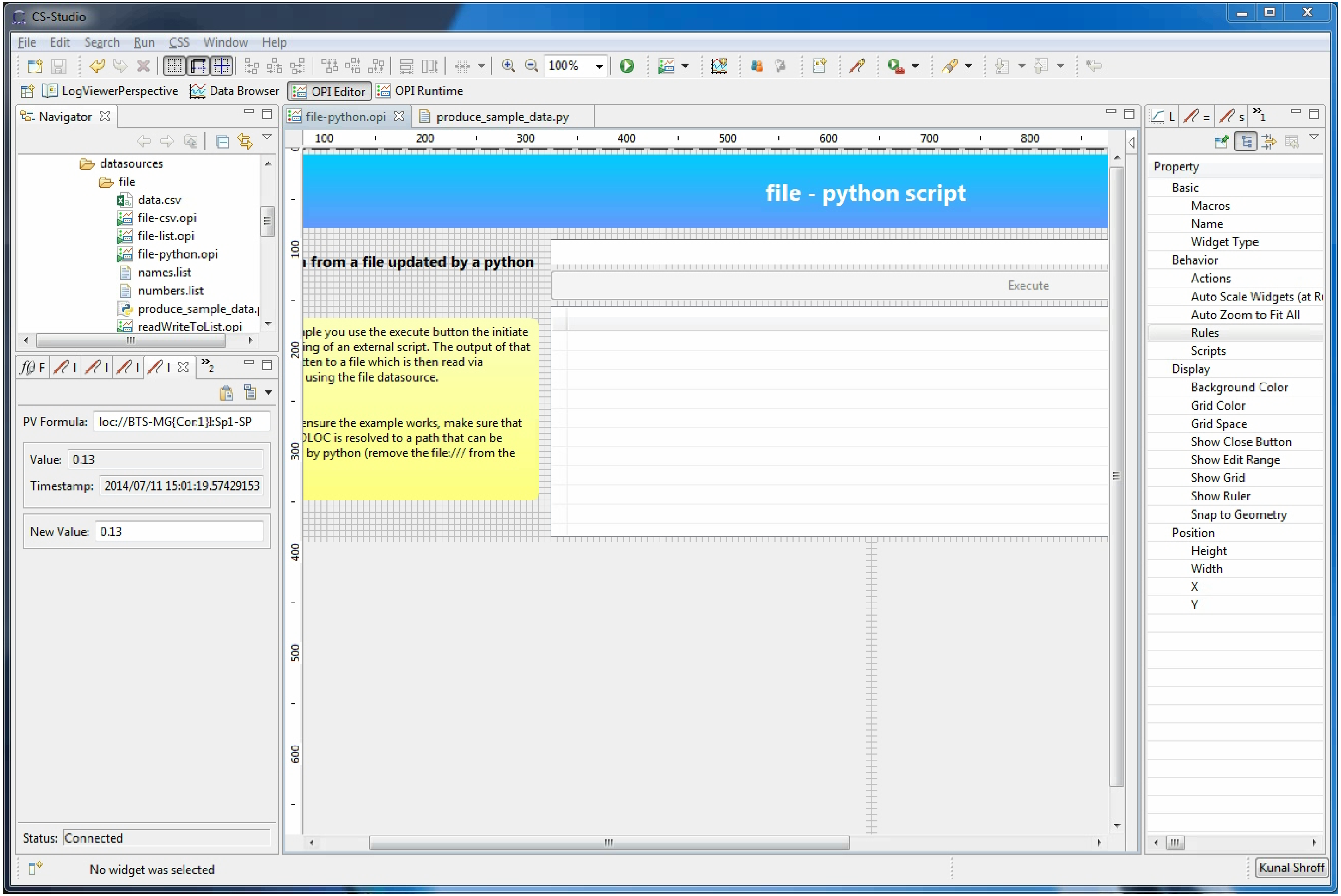
- Filters for tables
  - VTable tableValueFilter(VTable table, VString column, VType value)
  - VTable tableRangeFilter(VTable table, VString column, VNumber min, Vnumber max)
  - TRUE, FALSE, E, PI
- Other contributions
  - VInt indexOf(VEnum enum) (Tom Cobb, Diamond)
  - VImage image(PVANTNDArray array) (Matej Sekoranja, Cosylab)
    - Support for NTNDArray
  - VTable cf(VString query) (Kunal Shroff, BNL)
    - Query to channel finder

FILE DATASOURCE



Property	Value
Basic	
Data Expr	${DLOC}/data.csv
Name	Scatter Graph
Widget Ty	Scatter Graph
X axis ranı	data
X Column	="A"
Y axis ranı	data
Y Column	="C"
Behavior	
Actions	no action
Configura	□ no
Enabled	☑ yes
Resizable	□ no
Rules	no rule attached
Scripts	no script attached
Visible	☑ yes
Border	
Border Co	(0,128,255)
Border Sty	None
Border Wi	1
Display	
Backgrour	(240,240,240)
Font	Default
Foregrour	(192,192,192)
Tooltip	
Position	
Height	211
Scale Opti	true true false

CSV file update: change a file, update the OPI



file - python script

n from a file updated by a python

ple you use the execute button the initiate
ing of an external script. The output of that
tten to a file which is then read via
using the file datasource.

ensure the example works, make sure that
OLOC is resolved to a path that can be
by python (remove the file:/// from the

PV Formula: loc://BTS-MG{Cor:1}I:Sp1-SP

Value: 0.13

Timestamp: 2014/07/11 15:01:19.57429153

New Value: 0.13

Status: Connected

No widget was selected

BOY screen for a python script (Kunal Shroff, BNL)

File datasource

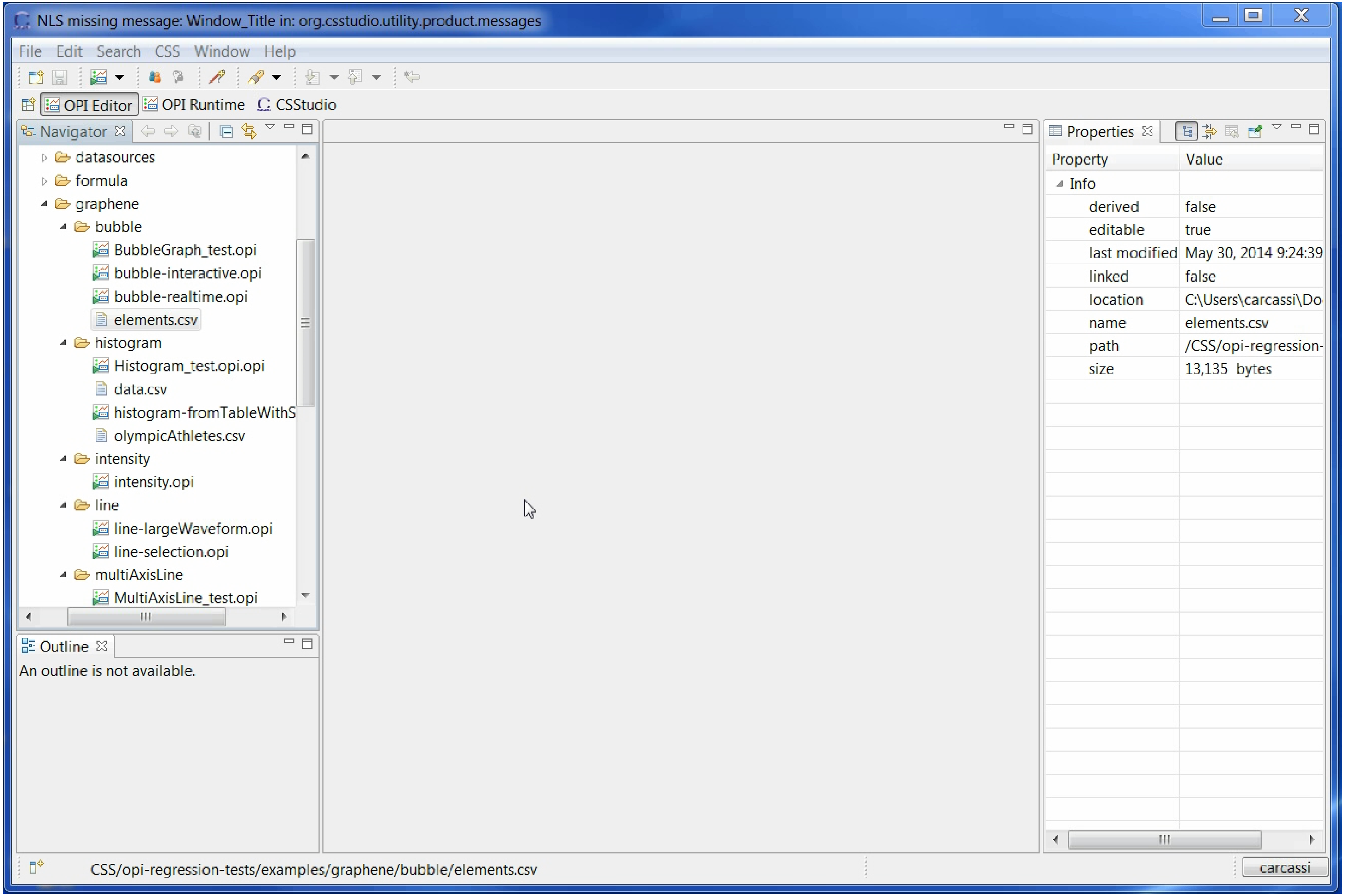
- Data formats (can be extended)
  - Comma separated files (.csv)
    - Separator can be , ; TAB SPACE
    - Strings can be delimited by “ ”, can be multiline
    - Loaded as a table (VTable)
  - Supports lists (.list)
    - Either numbers or strings
    - Treated as a numeric or string array (VNumberArray or VStringArray)
  - JSON vTypes (.jvtype)
    - Verbose but flexible
    - https://github.com/diirt/diirt/wiki/JSON-serialization-for-vTypes-v1
- Change notifications
  - If the file changes, values are automatically updated
  - Does not work on network mounts (e.g. NFS)

GRAPHENE

Graphene

- Status
  - Many graphs at the beta stage now
    - Good understanding on how to bind them to types
    - Suitable performance
    - Suitable real-time handling for auto axes
    - Still need to finish the last 20% to actually make them useful “drop-ins”
      - Really need help here to guide development through user feedback

BUBBLE GRAPH



NLS missing message: Window_Title in: org.csstudio.utility.product.messages
File Edit Search CSS Window Help
OPI Editor
OPI Runtime
CSStudio
Navigator
Outline
Properties

- datasources
- formula
- graphene
  - bubble
    - BubbleGraph_test.opi
    - bubble-interactive.opi
    - bubble-realtime.opi
    - elements.csv
  - histogram
    - Histogram_test.opi.opi
    - data.csv
    - histogram-fromTableWithS
    - olympicAthletes.csv
  - intensity
    - intensity.opi
  - line
    - line-largeWaveform.opi
    - line-selection.opi
  - multiAxisLine
    - MultiAxisLine_test.opi

An outline is not available.

Property	Value
Info	
derived	false
editable	true
last modified	May 30, 2014 9:24:39
linked	false
location	C:\Users\carcassi\Do
name	elements.csv
path	/CSS/opi-regression-
size	13,135 bytes

CSS/opi-regression-tests/examples/graphene/bubble/elements.csv

carcassi

Dynamic bubble graph: periodic elements

Bubble graph

- 3 numbers (x,y,size) + 1 string (color)
- Allows to find correlations
- Typical use: sift through data to find problems
  - Each bubble represents an item (e.g. a router, a bpm, an analysis job in a cluster, …)
  - You expect some relationship between variables (e.g. setpoint=readback, current $\propto$ setpoint, …)
  - Find anomalies and look for other variable that correlate

Bubble graph

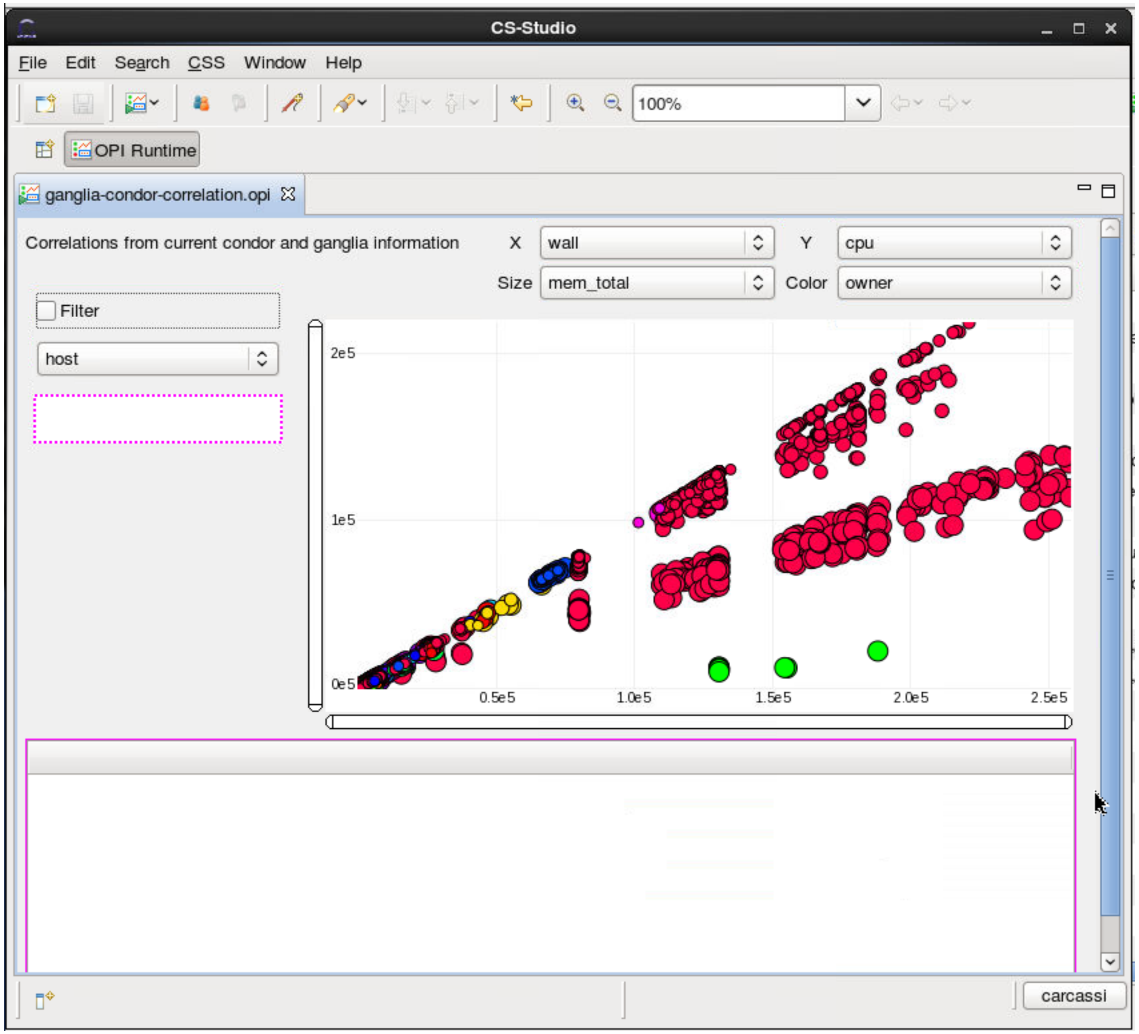
Case study: analysis jobs efficiency

- X axis – wall clock time
- Y axis – CPU time

Expect a straight line, some job are below the straight line (they are stuck waiting more)

Turns out those jobs were on all new hardware: disk was the bottleneck

The data used is actually the join from two tables: one from the job scheduler (condor) and one from the system monitoring (ganglia)



Correlations from current condor and ganglia information

X: wall

Y: cpu

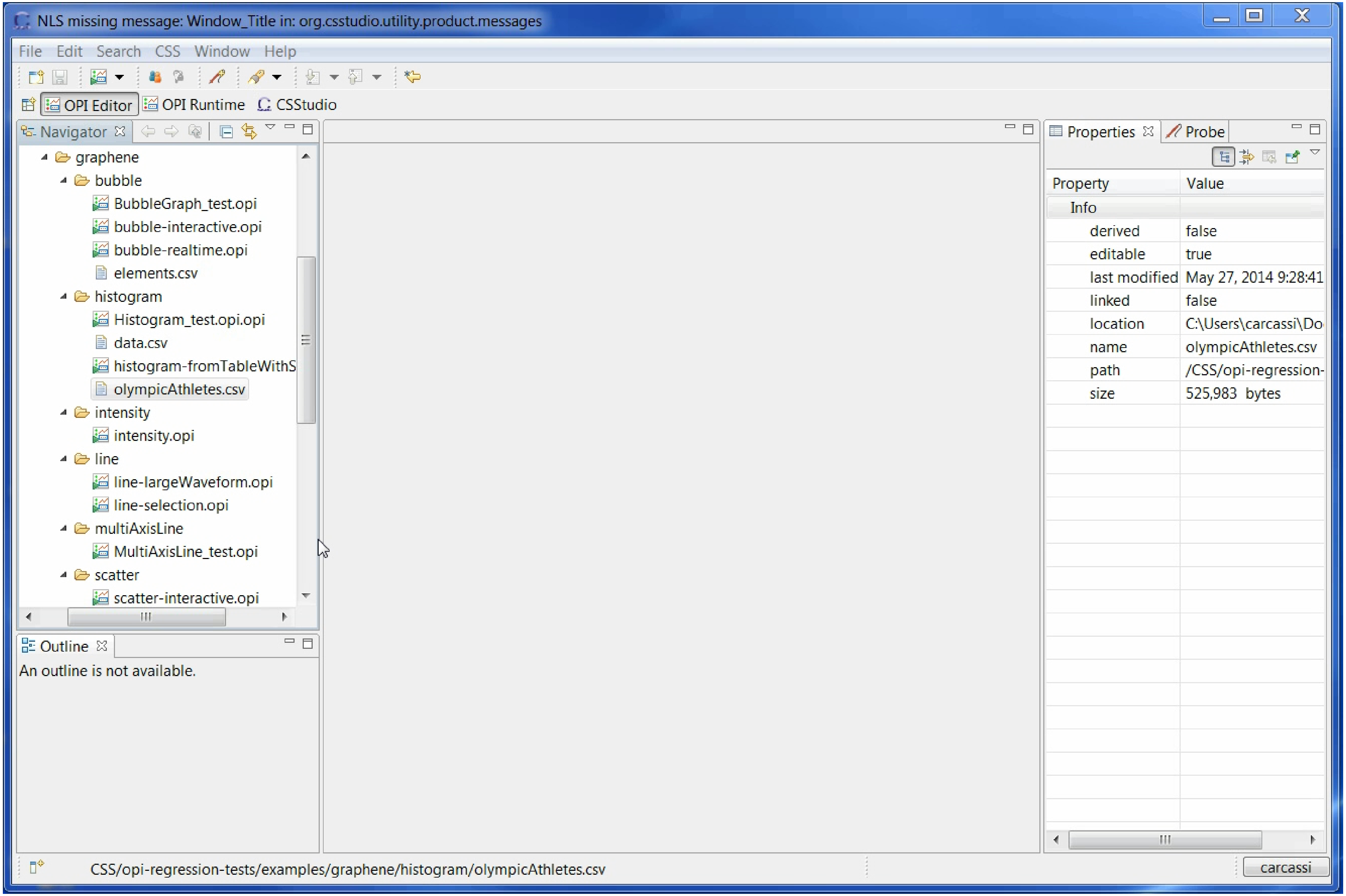
Size: mem_total

Color: owner

□ Filter

host

HISTROGRAM



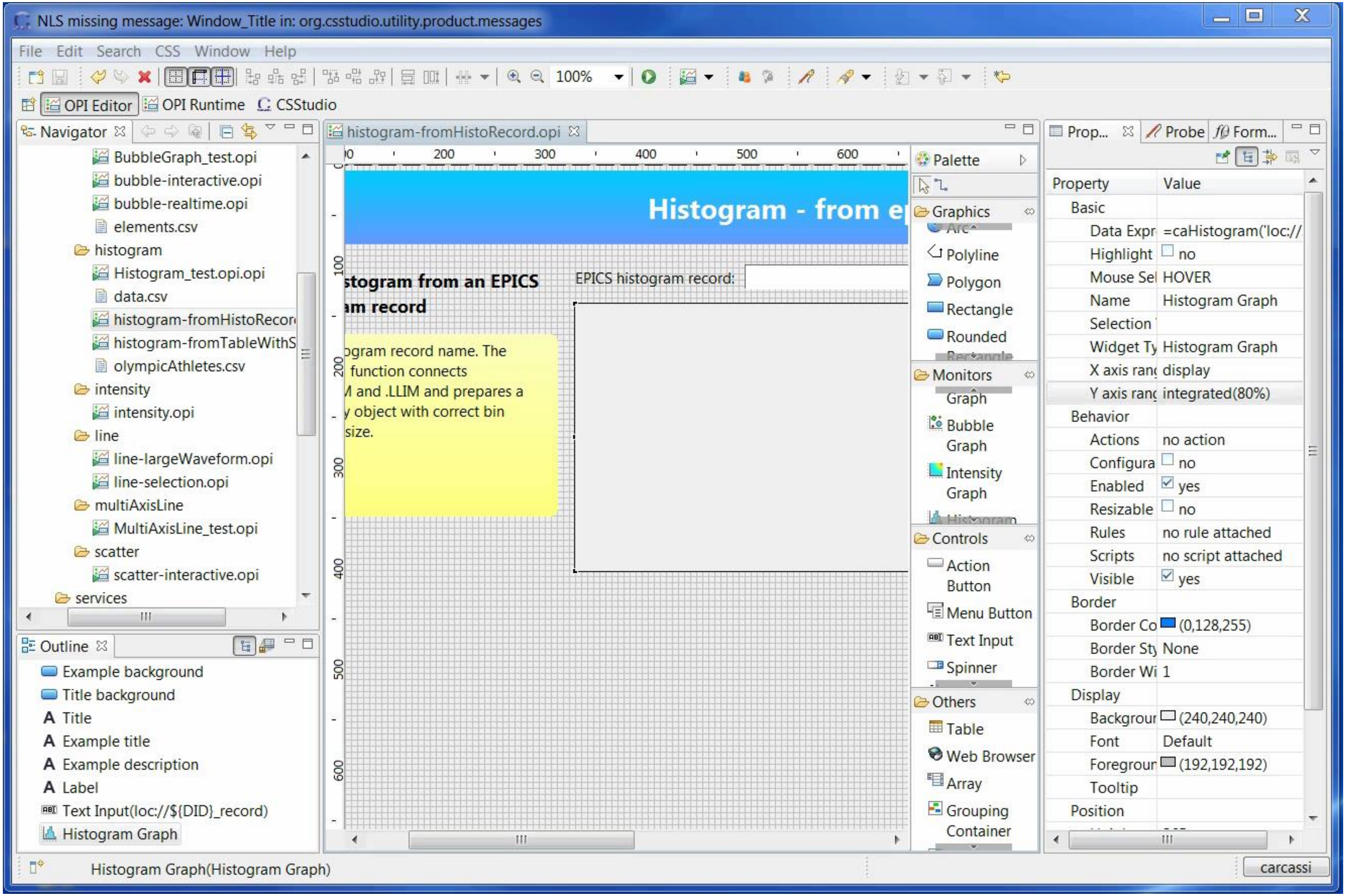
NLS missing message: Window_Title in: org.csstudio.utility.product.messages

Property	Value
Info	
derived	false
editable	true
last modified	May 27, 2014 9:28:41
linked	false
location	C:\Users\carcassi\Do
name	olympicAthletes.csv
path	/CSS/opi-regression-
size	525,983 bytes

An outline is not available.

CSS/opi-regression-tests/examples/graphene/histogram/olympicAthletes.csv

Histogram from table values: Olympic medal winners



Histogram - from e

stogram from an EPICS
am record

EPICS histogram record:

ogram record name. The
function connects
/l and .LLIM and prepares a
y object with correct bin
size.

Property	Value
Basic	
Data Expr	=caHistogram('loc://
Highlight	□ no
Mouse Sel	HOVER
Name	Histogram Graph
Selection	
Widget Ty	Histogram Graph
X axis rang	display
Y axis rang	integrated(80%)
Behavior	
Actions	no action
Configura	□ no
Enabled	☑ yes
Resizable	□ no
Rules	no rule attached
Scripts	no script attached
Visible	☑ yes
Border	
Border Co	(0,128,255)
Border Sty	None
Border Wi	1
Display	
Backgrour	(240,240,240)
Font	Default
Foregrour	(192,192,192)
Tooltip	
Position	

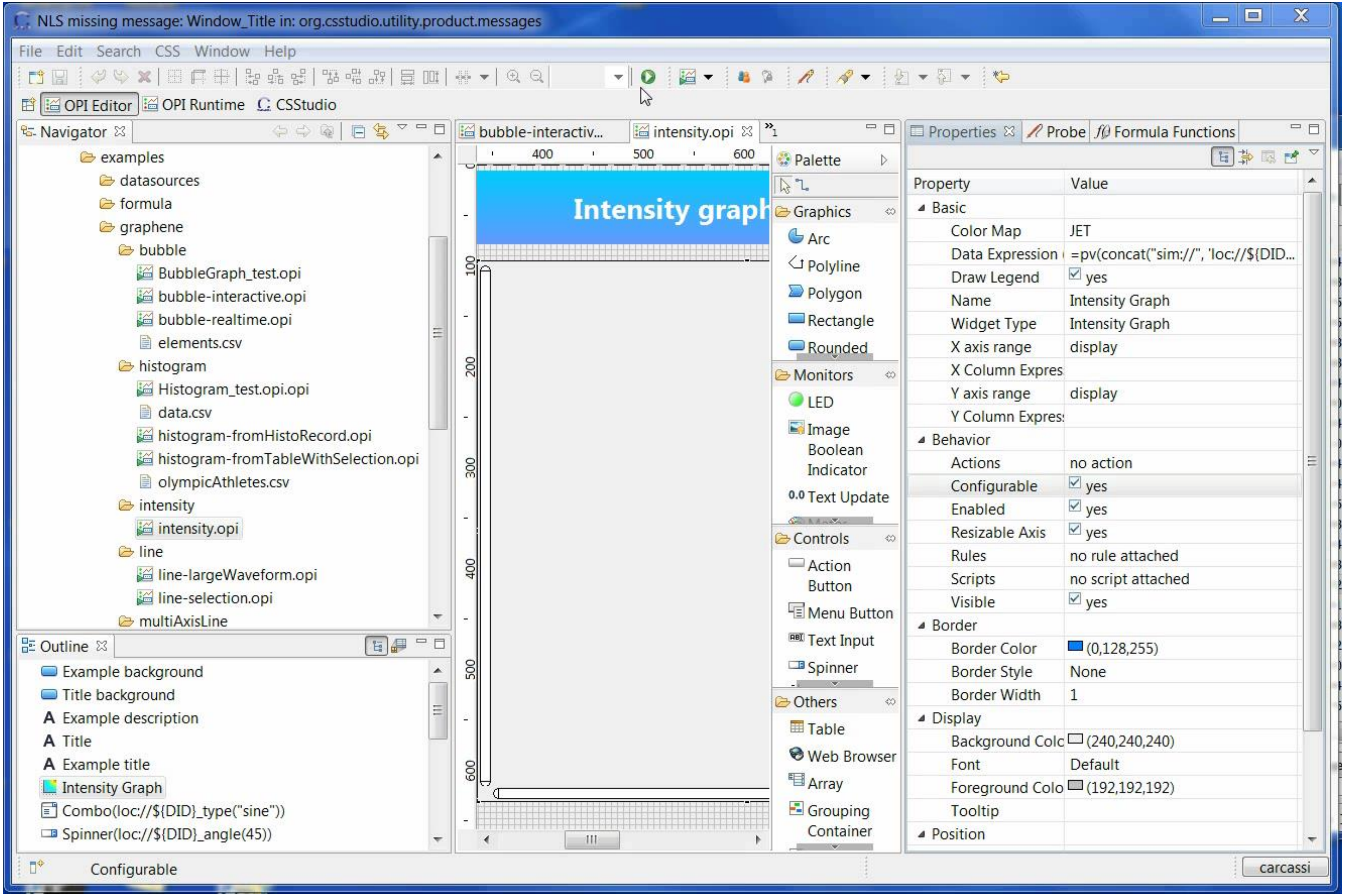
Histogram Graph(Histogram Graph)

Histogram from IOC: just put the record name

Histogram

- The data
  - caHistogram(<VString>): expects a histogram ca record and gets all the information
  - histogramOf(<VNumberArray>): creates a histogram out of the array
    - Auto-ranges, integrates so that bins are not redefined at every update
- The graph
  - Support selection on click or mouse over
    - Broadcasts the sub-array of the selected bin

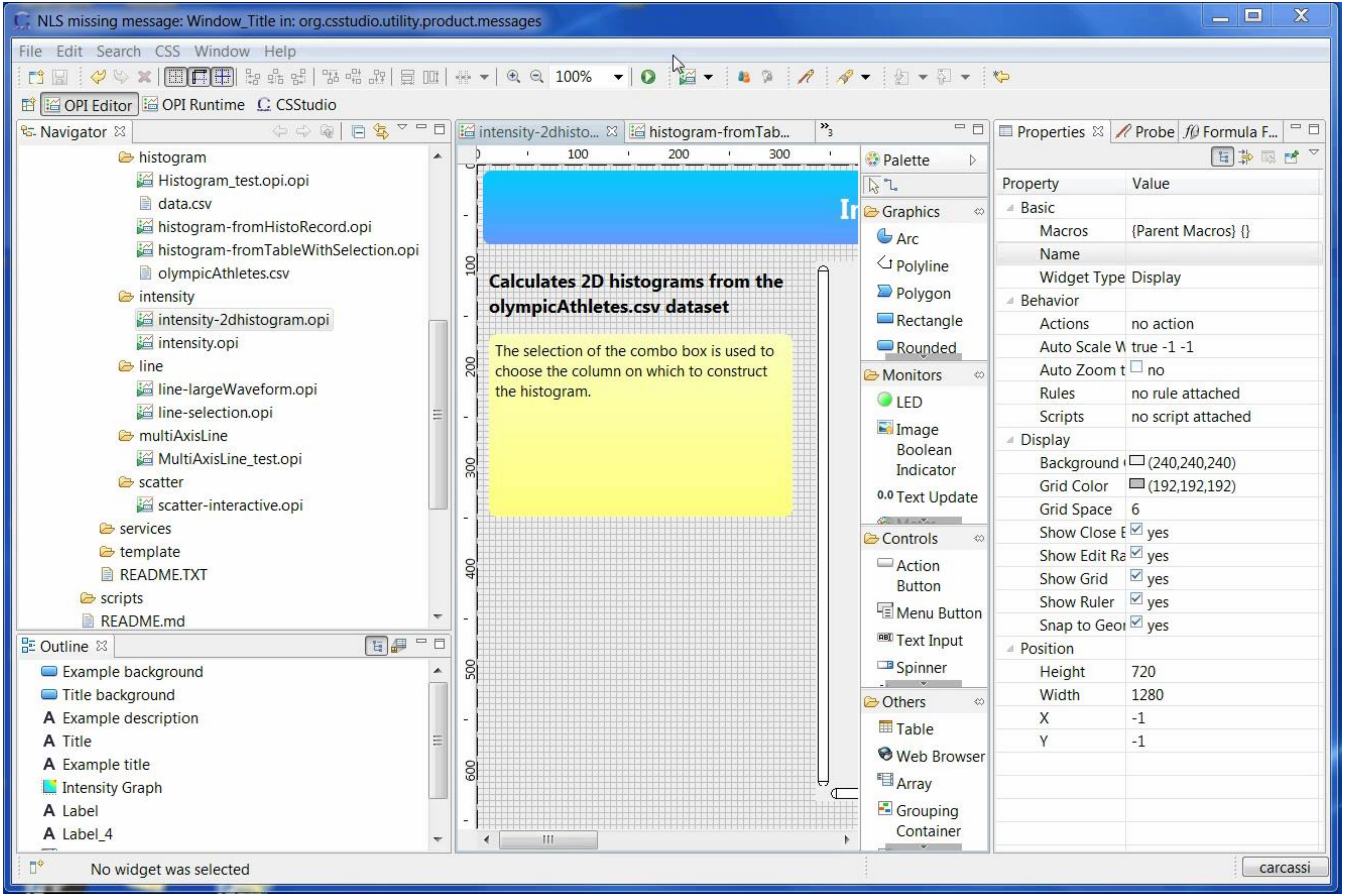
INTENSITY GRAPH



NLS missing message: Window_Title in: org.csstudio.utility.product.messages

Property	Value
Basic	
Color Map	JET
Data Expression	=pv(concat("sim://", 'loc://${DID...
Draw Legend	☑ yes
Name	Intensity Graph
Widget Type	Intensity Graph
X axis range	display
X Column Expres	
Y axis range	display
Y Column Expres	
Behavior	
Actions	no action
Configurable	☑ yes
Enabled	☑ yes
Resizable Axis	☑ yes
Rules	no rule attached
Scripts	no script attached
Visible	☑ yes
Border	
Border Color	(0,128,255)
Border Style	None
Border Width	1
Display	
Background Colo	□ (240,240,240)
Font	Default
Foreground Colo	(192,192,192)
Tooltip	
Position	

Intensity graph performance



NLS missing message: Window_Title in: org.csstudio.utility.product.messages

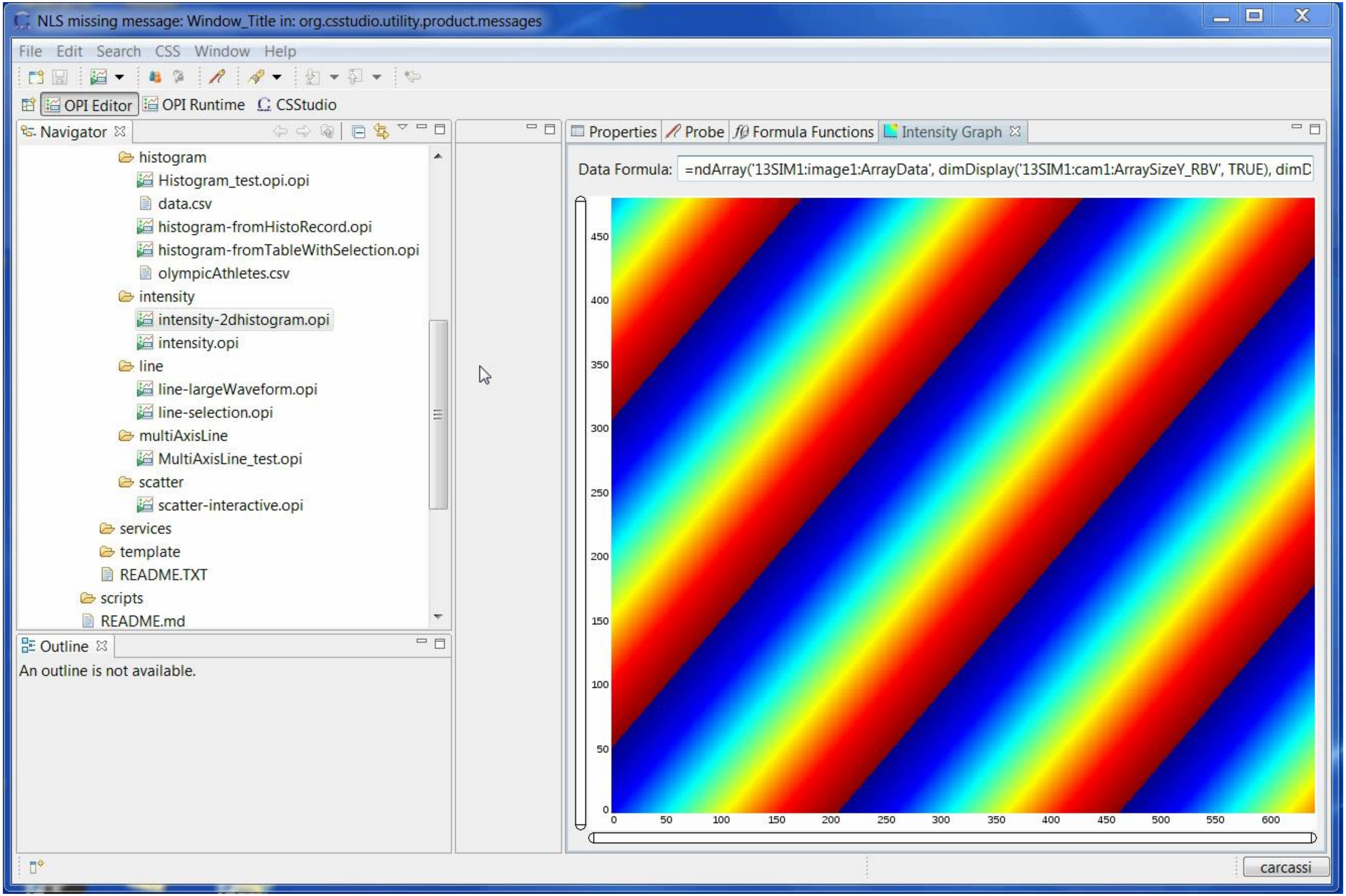
Calculates 2D histograms from the olympicAthletes.csv dataset

The selection of the combo box is used to choose the column on which to construct the histogram.

Property	Value
Basic	
Macros	{Parent Macros} {}
Name	
Widget Type	Display
Behavior	
Actions	no action
Auto Scale W	true -1 -1
Auto Zoom t	□ no
Rules	no rule attached
Scripts	no script attached
Display	
Background	□ (240,240,240)
Grid Color	■ (192,192,192)
Grid Space	6
Show Close E	☑ yes
Show Edit Ra	☑ yes
Show Grid	☑ yes
Show Ruler	☑ yes
Snap to Geo	☑ yes
Position	
Height	720
Width	1280
X	-1
Y	-1

No widget was selected

Intensity graph for a 2D histogram



NLS missing message: Window_Title in: org.csstudio.utility.product.messages

An outline is not available.

Data Formula: =ndArray('13SIM1:image1:ArrayData', dimDisplay('13SIM1:cam1:ArraySizeY_RBV', TRUE), dimD

Intensity graph from area detector

Intensity graph

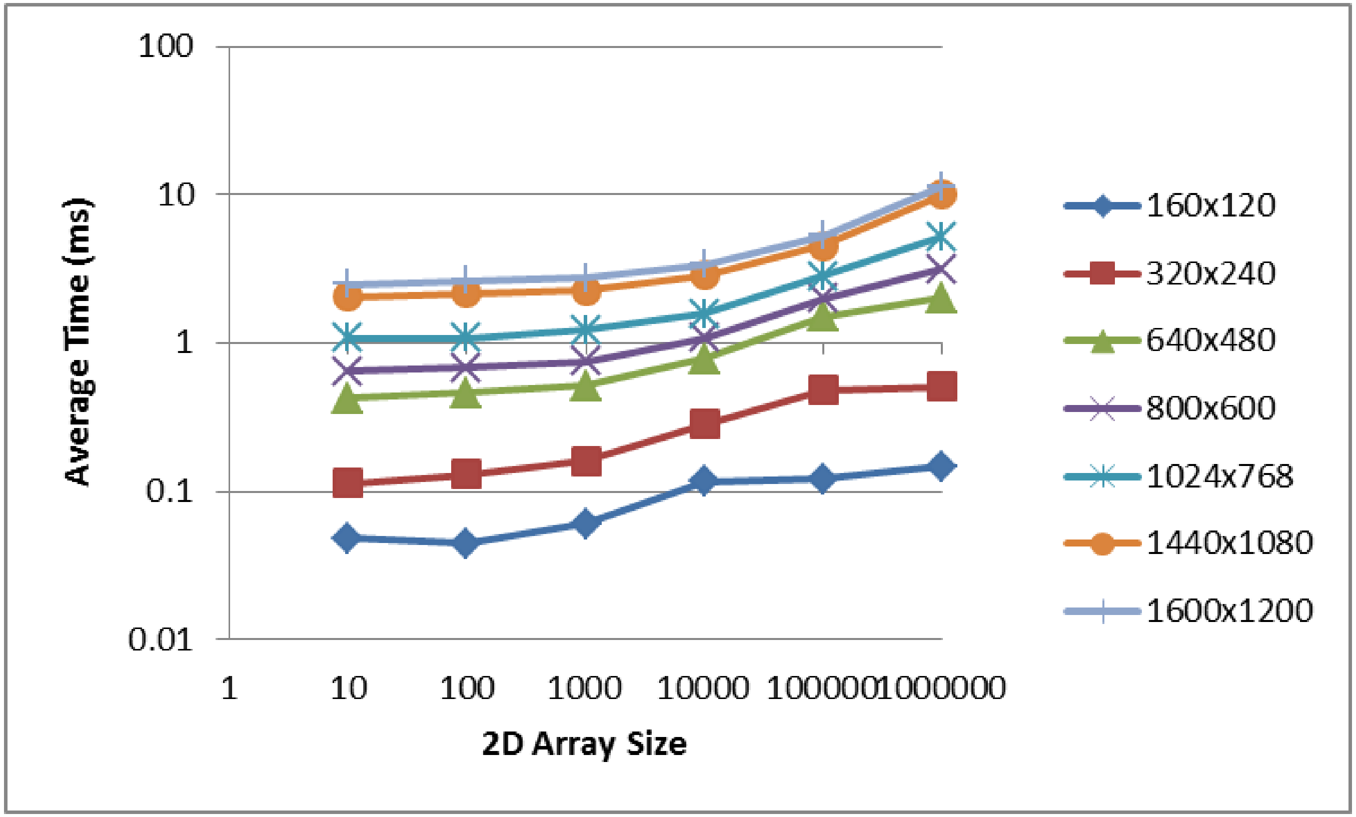
- The data
  - Cameras
    - ndArray('13SIM1:image1:ArrayData', dimDisplay('13SIM1:cam1:ArraySizeY_RBV', TRUE), dimDisplay('13SIM1:cam1:ArraySizeX_RBV', FALSE))
      v3 needs to compose the ndArray from different pvs
      Investigating how to have caAreaDetector(<VString>)
      - Maybe caAreaDetector(“13SIM1”, “1”) ?
    - v4 will (hopefully) give the combined object
    - Need to figure out the color format and other details
  - 2D histogram
    - =histogram2DOf(<VTable>, <VString>, <VString>): creates a 2D histogram from two columns of a table
- The graph
  - Handles non square cells (e.g. different bin sizes)
  - Region of interest not finished

Intensity graph

- Win7 – i7-3820QM – 2.70GHz – Java 7

<15 ms per plot even at high resolution

No optimization for special cases



Conclusion

- diirt: Data Integration In Real-Time
  - Natural evolution of past 5 years of work on pvmanager
  - May have some answers to “How do we integrate the data from different domains?”
- Modular design keeps paying off
  - JSON serialization developed for Web Pods easily reused for file datasource
  - Formulas and datasources developed for CS-Studio reused for Web Pods
- The value is given by the product of the features, not their sum
- But contributors can work independently on their own piece
  - Come and tell me how hard or how easy it was!