

# SOFTWARE ANALYTICS - AND HOW TO DRAW CONCLUSIONS OUT OF DATA

**Presenting the sw-analytics toolkit!**

# CONTEXT



- What are software metrics?
  - LOC, #defects, #contributors, cyclomatic complexity, ...
  
- Which metric is the best?
  - None?
  
- Can SW metrics be used to draw conclusions about software?

# SW-ANALYTICS TOOLKIT



- Scrapes VCS & issue tracker
  - Analyzes everything on function granularity
  - Stores in RDBMS
  
- Web-based “Exploratory visualization”
  - Subsystem – file - function

# TREEMAP - OVERVIEW

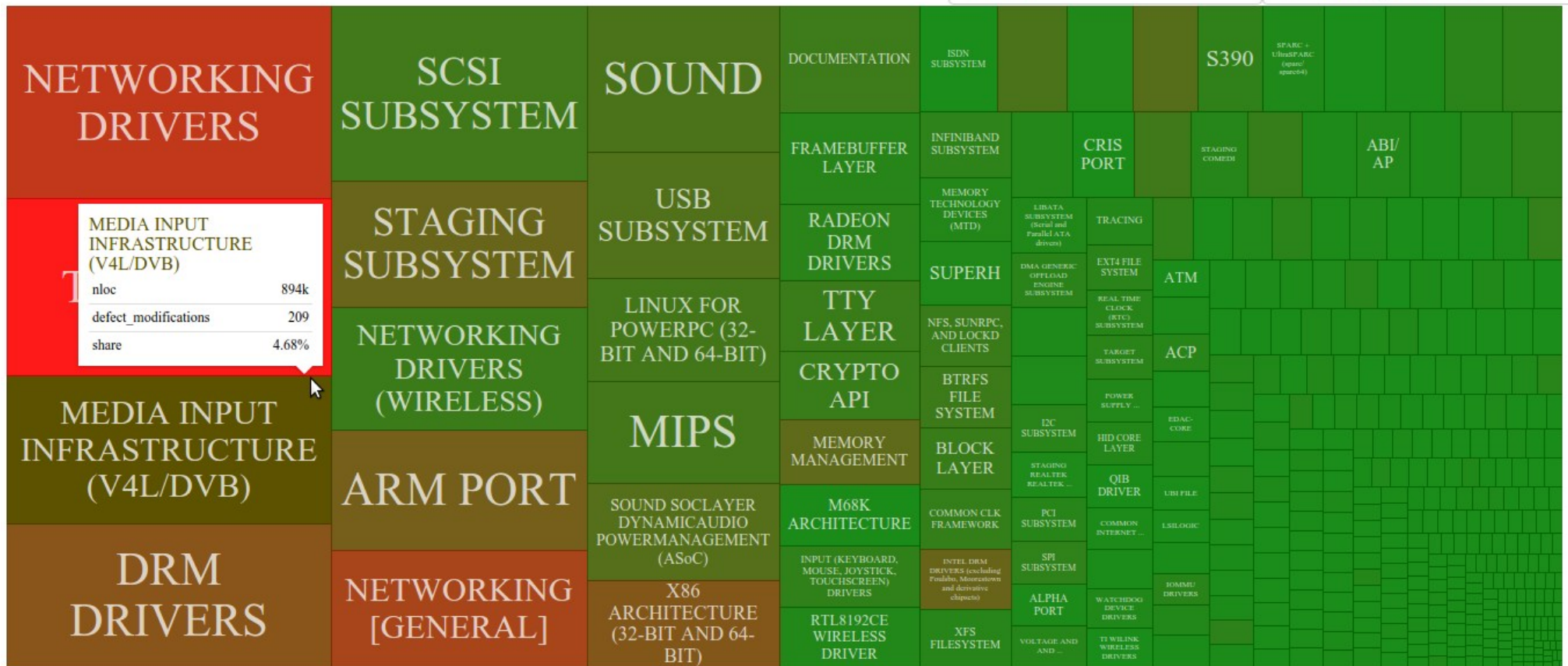


Last database update: (

Treemap Charts /docs

Root

Size: nloc Color: defect\_modifications



# TREEMAP - OVERVIEW

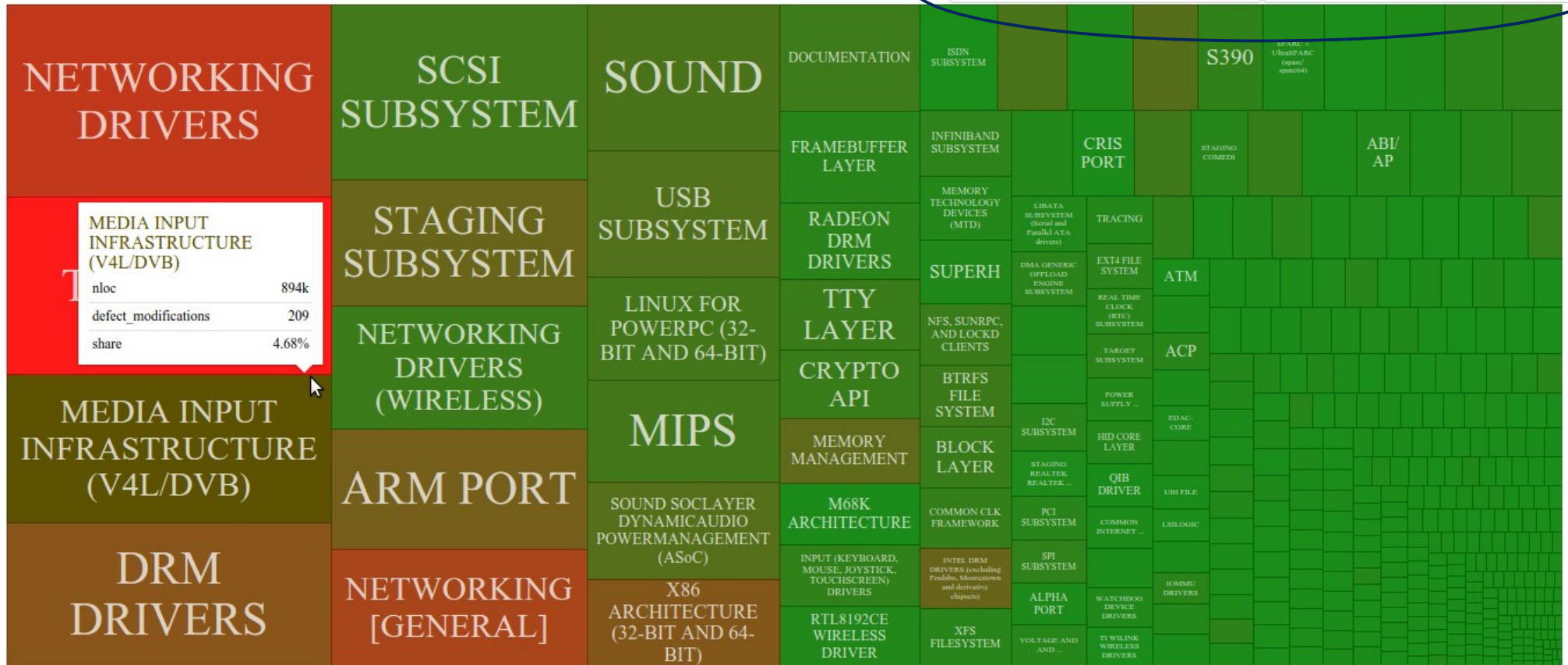


Last database update: (

Treemap [Charts](#) [/docs](#)

Root

Size: nloc Color: defect\_modifications





# TREEMAP - SUBSYSTEMS



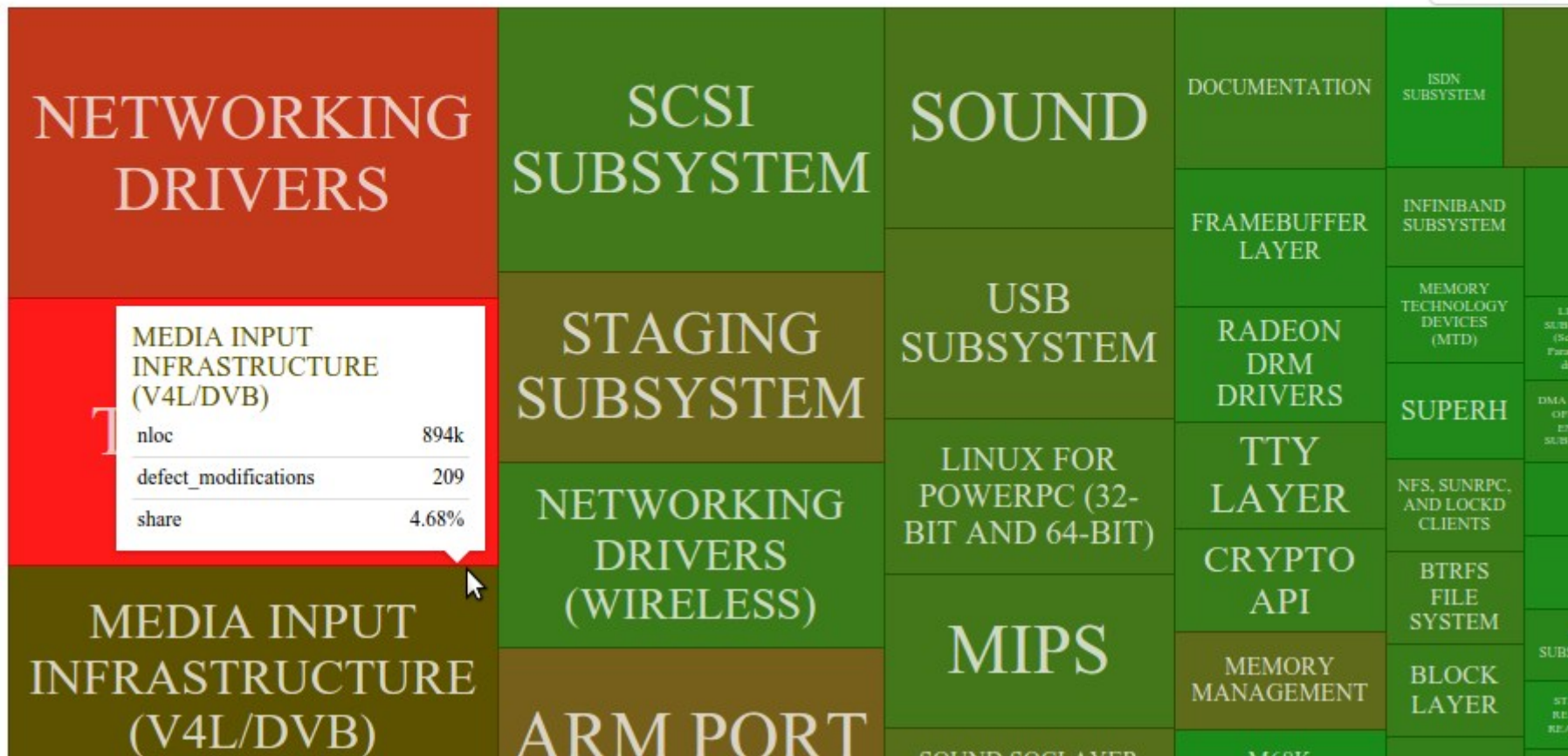
Treemap

Charts

/docs

Root

Size: nloc

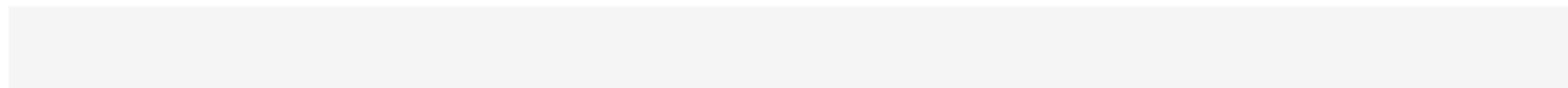


# TREEMAP - FILES

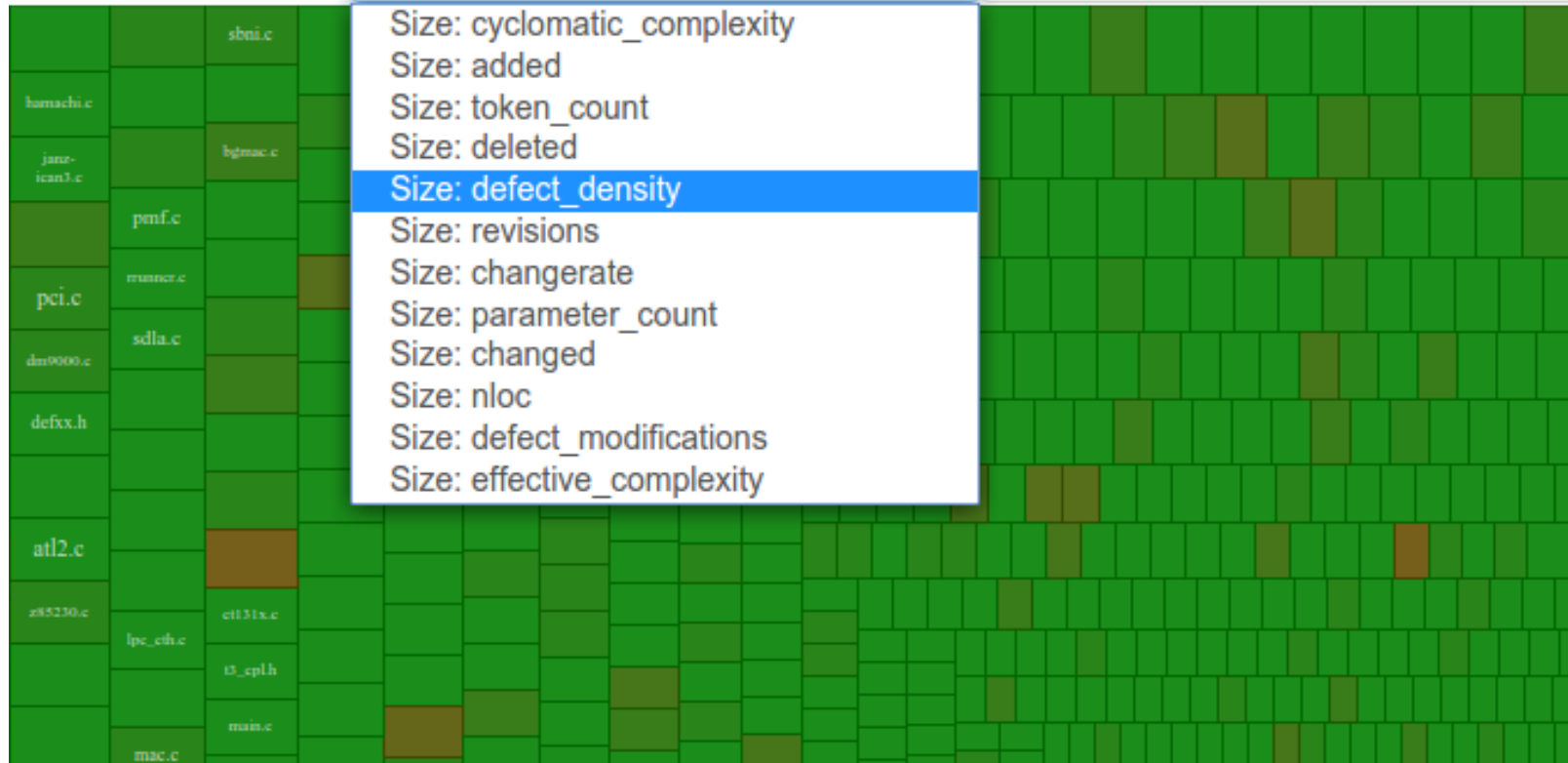


Charts /docs

Last database update: Oc



Size: nloc      Color: defect\_modifications



# TREEMAP - FUNCTIONS



Treemap

Charts

Root / NETWORKING DRIVERS / netdev.c

e1000\_flush\_desc\_rings

defect\_modifications 3

cyclomatic\_complexity 4

share 8.33%

Global Scope

e1000\_flush\_desc\_rings

e1000e\_cyclecou

e1000\_clean

e1000e\_down

e1000\_conf



# CHART – FILE & FUNCTION



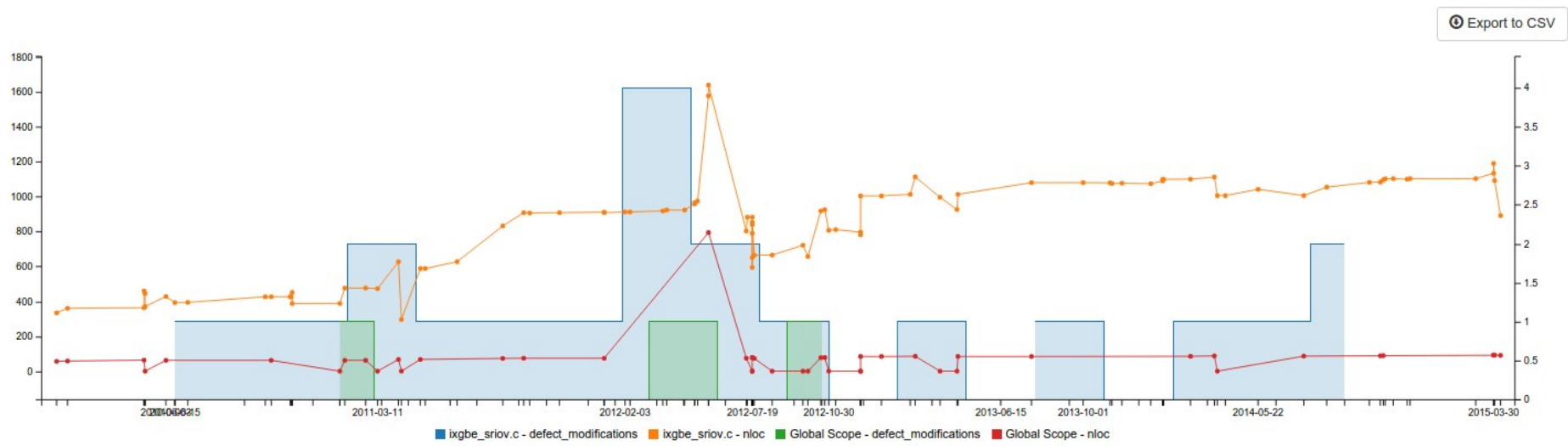
Last database update: Oct. 1

Treemap Charts /docs

Search...

- ixgbe\_phy.c
- ixgbe\_ptp.c
- ixgbe\_sysfs.c
- ixgbe\_sriov.c
- Global Scope
- ixgbe\_check\_vf\_rate\_limit
- ixgbe\_link\_mbps
- ixgbe\_set\_vf\_rate\_limit
- ixgbe\_ndo\_set\_vf\_bw
- ixgbe\_ndo\_set\_vf\_spoofchk
- ixgbe\_check\_vf\_assignment
- ixgbe\_find\_enabled\_vfs
- ixgbe\_enable\_port\_vlan
- ixgbe\_disable\_port\_vlan

Metric	Histogram Interval	Currently Active
cyclomatic_complexity		Line Histogram None
added		Line Histogram None
token_count		Line Histogram None
deleted		Line Histogram None
changerate		Line Histogram None
parameter_count		Line Histogram None
changed		Line Histogram None
nloc		Line Histogram None
defect_modifications	<input type="text" value="100"/>	Line Histogram None



# CHART – FILE & FUNCTION



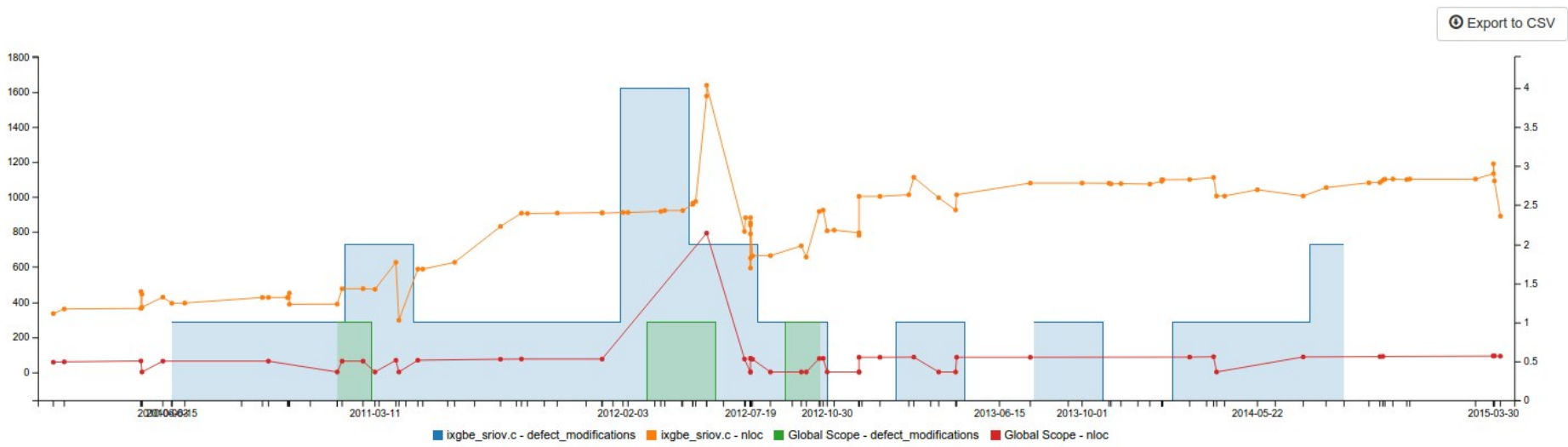
Last database update: Oct. 1

Treemap Charts /docs

Search...

- ixgbe\_phy.c
- ixgbe\_ptp.c
- ixgbe\_sysfs.c
- ixgbe\_sriov.c
  - Global Scope
  - ixgbe\_check\_vf\_rate\_limit
  - ixgbe\_link\_mbps
  - ixgbe\_set\_vf\_rate\_limit
  - ixgbe\_ndo\_set\_vf\_bw
  - ixgbe\_ndo\_set\_vf\_spoofchk
  - ixgbe\_check\_vf\_assignment
  - ixgbe\_find\_enabled\_vfs
  - ixgbe\_enable\_port\_vlan
  - ixgbe\_disable\_port\_vlan

Metric	Histogram Interval	Currently Active
cyclomatic_complexity		Line Histogram None
added		Line Histogram None
token_count		Line Histogram None
deleted		Line Histogram None
changerate		Line Histogram None
parameter_count		Line Histogram None
changed		Line Histogram None
nloc		Line Histogram None
defect_modifications		Line Histogram None



# CHART - FILE & FUNCTION



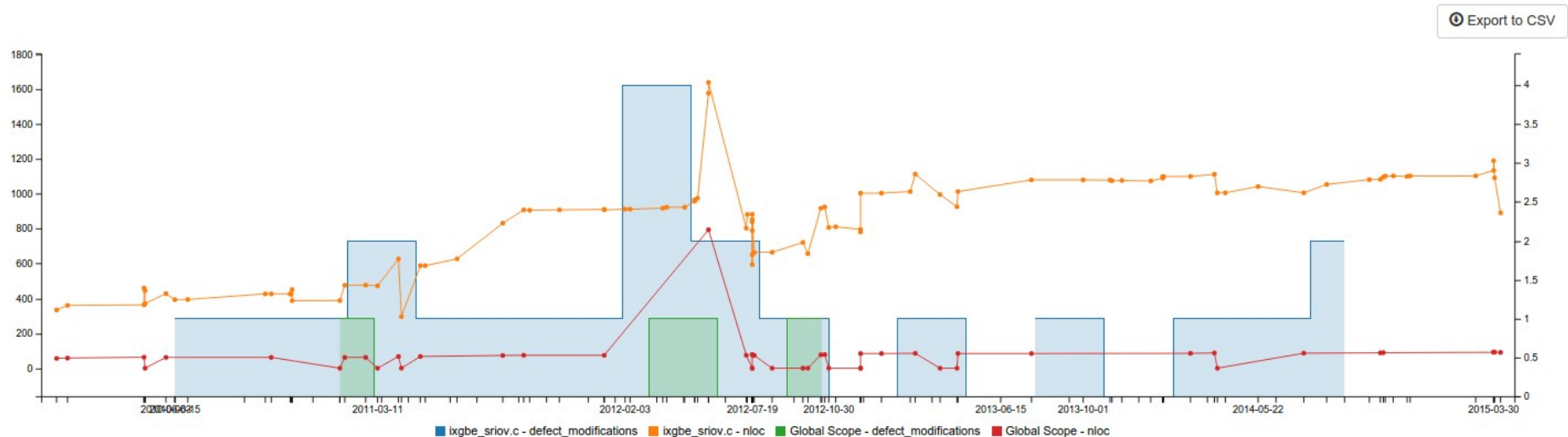
Last database update: Oct. 1

Treemap Charts /docs

Search...

- ixgbe\_phy.c
- ixgbe\_ptp.c
- ixgbe\_sysfs.c
- ixgbe\_sriov.c
  - Global Scope
  - ixgbe\_check\_vf\_rate\_limit
  - ixgbe\_link\_mbps
  - ixgbe\_set\_vf\_rate\_limit
  - ixgbe\_ndo\_set\_vf\_bw
  - ixgbe\_ndo\_set\_vf\_spoofchk
  - ixgbe\_check\_vf\_assignment
  - ixgbe\_find\_enabled\_vfs
  - ixgbe\_enable\_port\_vlan
  - ixgbe\_disable\_port\_vlan

Metric	Histogram Interval	Currently Active
cyclomatic_complexity		Line Histogram None
added		Line Histogram None
token_count		Line Histogram None
deleted		Line Histogram None
changerate		Line Histogram None
parameter_count		Line Histogram None
changed		Line Histogram None
nloc		Line Histogram None
defect_modifications	<input type="text" value="100"/>	Line Histogram None



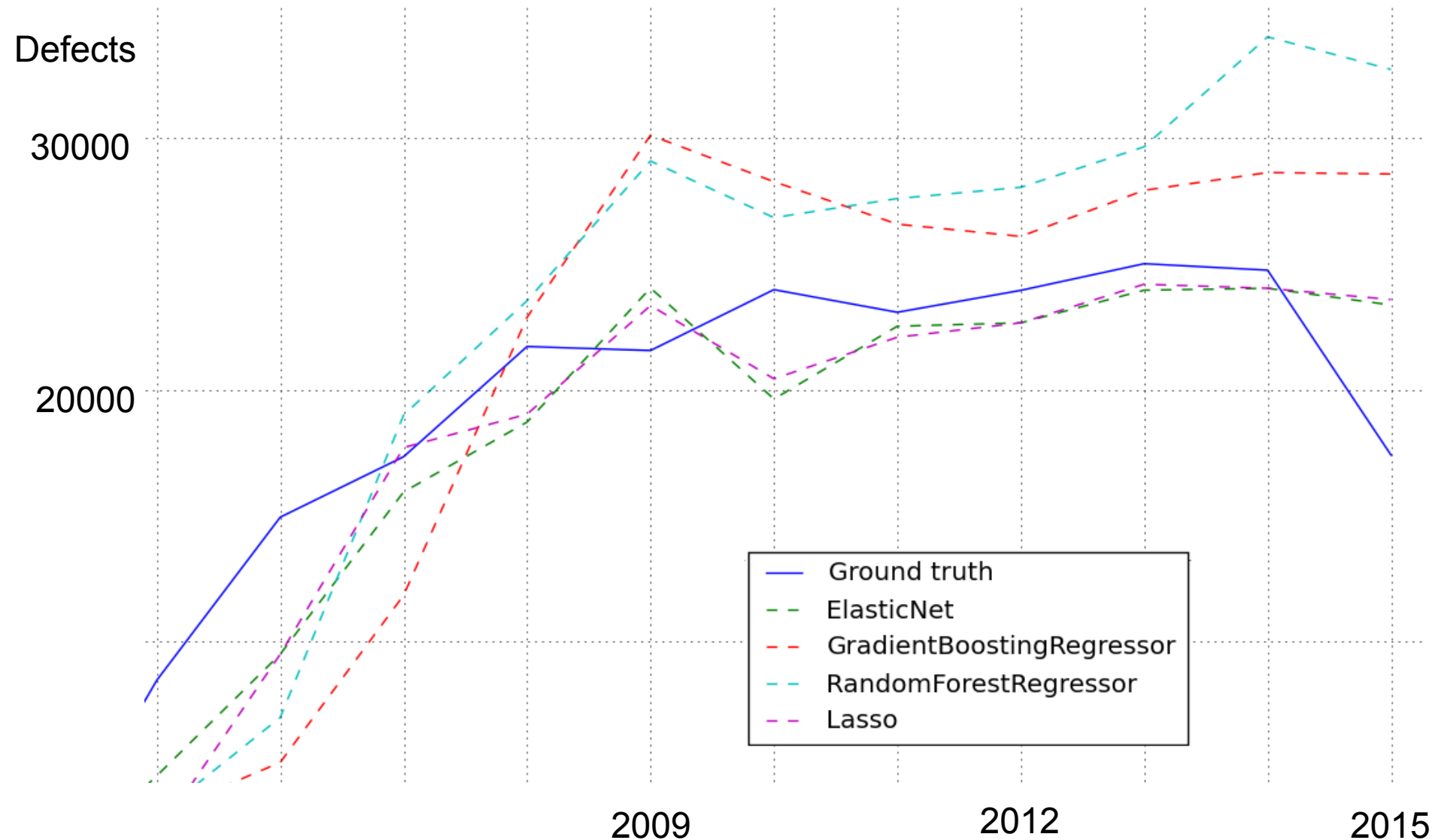
# COMING ATTRACTIONS



- Defect predictions
  - Subsystem – file – function
  
- Automatic redesign/refactoring identification
  - Effort distribution
  
- .....

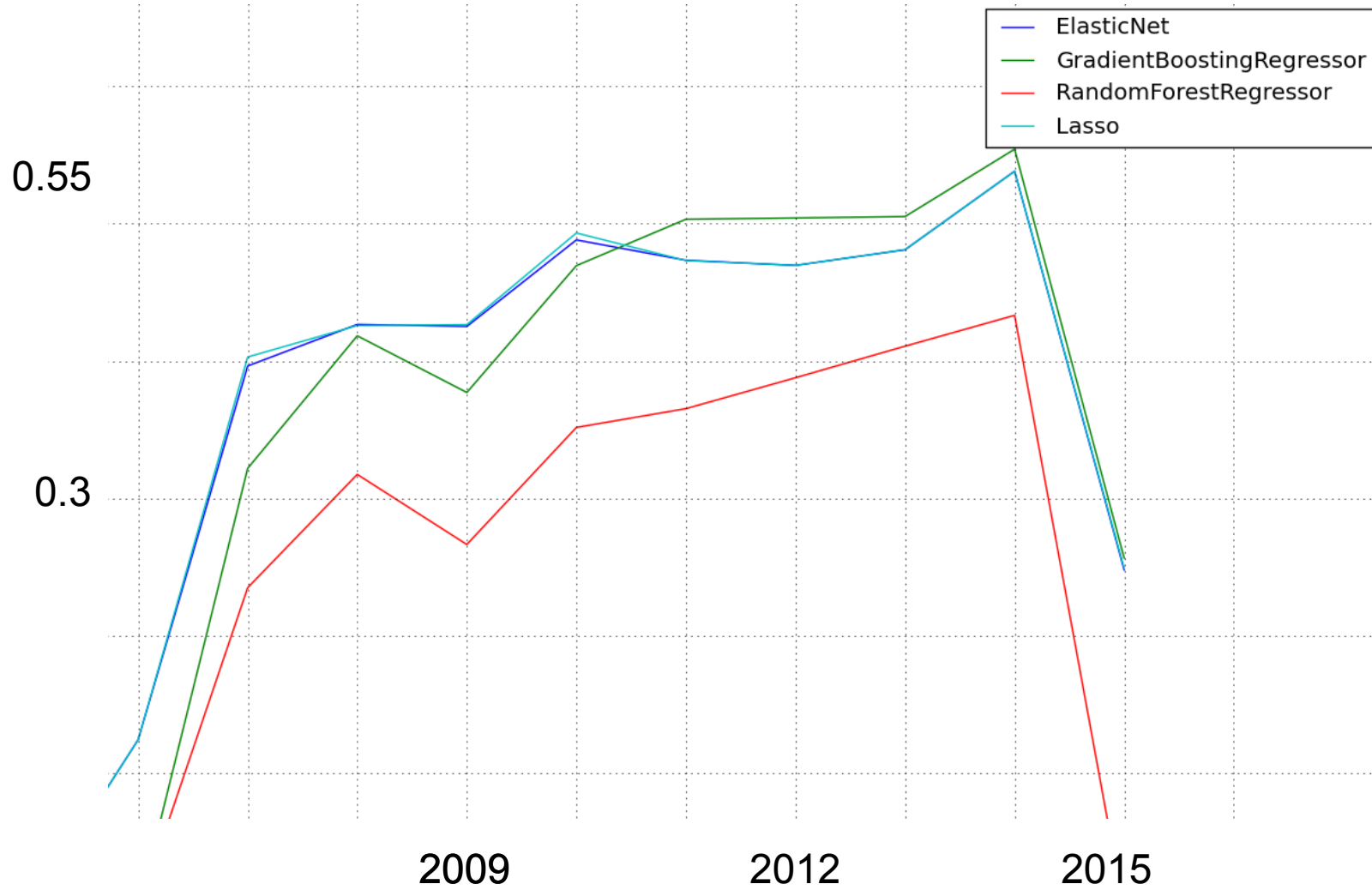
# DEFECT PREDICTION

## SUM OF ALL DEFECTS



# DEFECT PREDICTION

## R<sup>2</sup> OVER TIME



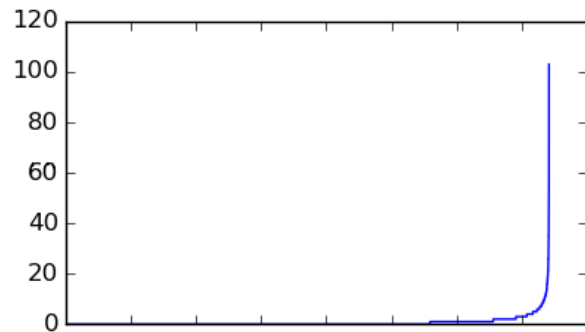


# DEFECT PREDICTION

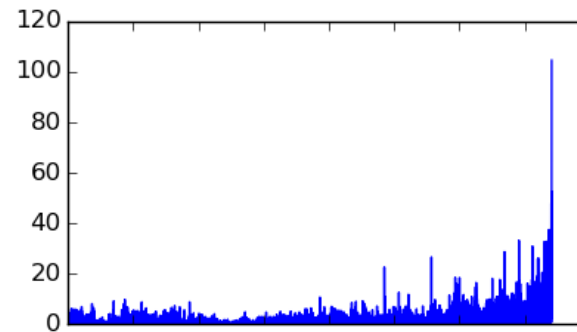
## PER FILE FOR 2014



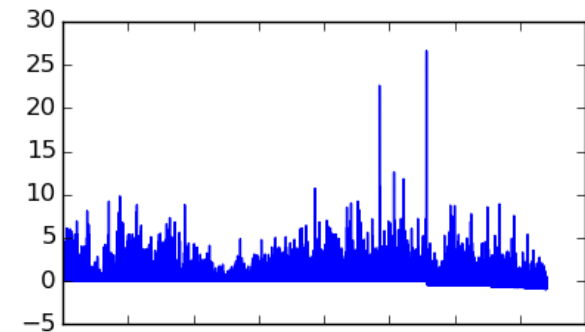
### RandomForestRegressor



Ground truth



Predicted



predicted/truth

# CONCLUSIONS



- No single metric exists that will tell you what your problem is!
- SW-analytics toolkit helps you analyze your code in a robust way
- SW-analytics toolkit have some really cool features coming up! =)



# SW-ANALYTICS TOOLKIT:

<http://www.github.com/jderehag/swat>



**ERICSSON**