

# ***Elettra DAta aNalysis Tool: a data webhousing tool for heterogeneous log analysis***

***Roberto Pugliese***

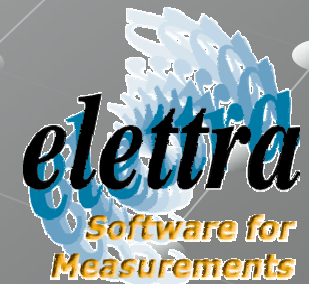
***Stefano Maraspin***

***Alessio Curri***

***Software for Measurements***

***Experiment Division***

***Sincrotrone Trieste S.C.p.A.***

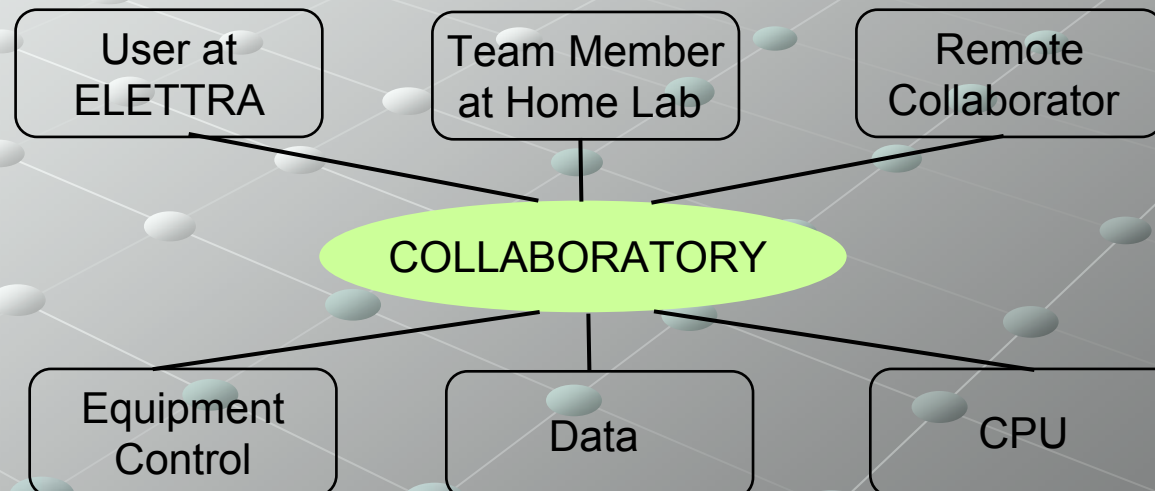


# Outline

- Elettra Virtual Collaboratory (EVC)
  - The portal application
  - The Collaboration Tools
  - Usage scenarios
- DANT an advanced Log Analyzer
  - Software Architecture
  - DANT at work
  - Status and future developments

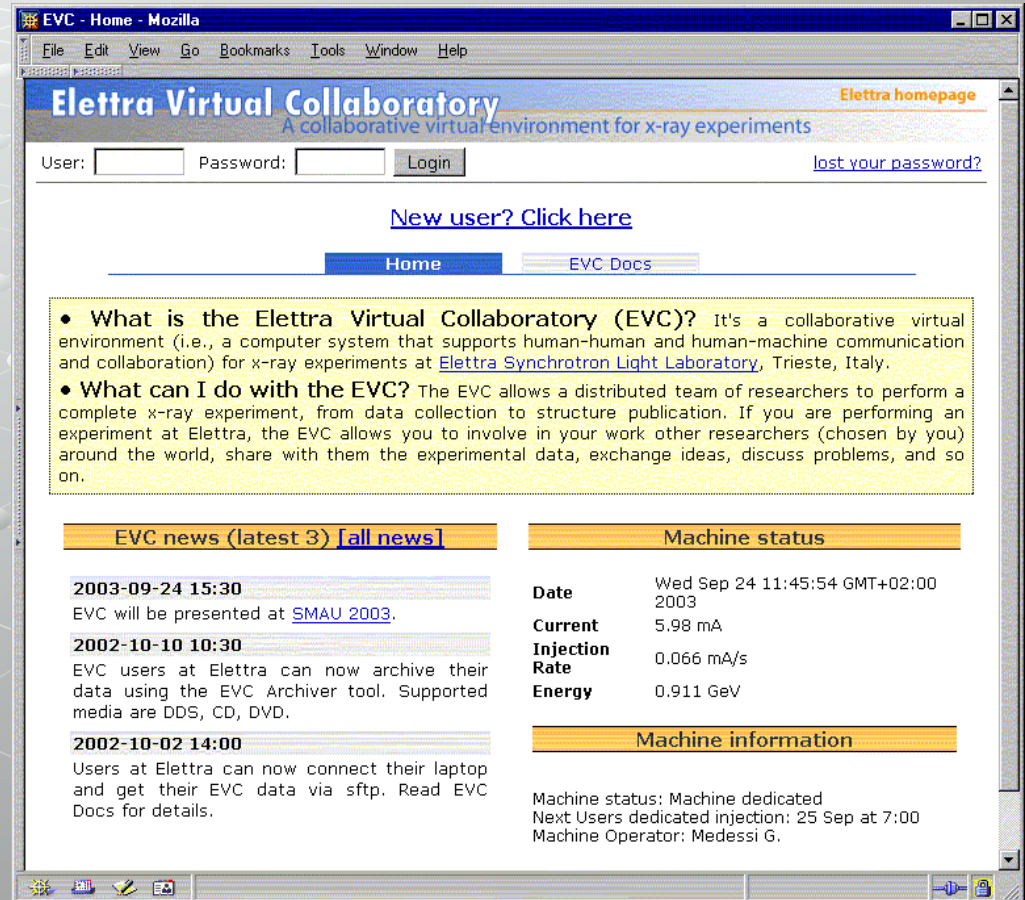
# What is the Elettra Virtual Collaboratory (EVC)?

- EVC is an example of virtual laboratory, a system which allows a team of researchers distributed anywhere in the world to perform a complete experiment on the equipped beamlines and experimental stations of Elettra.



# EVC in action: a web portal

- EVC is based on the “web portal” metaphor
- EVC supports four different user categories:
  - Visitors
  - Normal users
  - Project leaders
  - Staff



The screenshot shows a Mozilla browser window titled "EVC - Home - Mozilla". The page content includes a navigation bar with "Home" and "EVC Docs" buttons. A central text box explains the EVC and lists user actions. Below this are sections for "EVC news (latest 3)" and "Machine status".

**Elettra Virtual Collaboratory**  
A collaborative virtual environment for x-ray experiments

User:  Password:   [lost your password?](#)

[New user? Click here](#)

[Home](#) [EVC Docs](#)

- **What is the Elettra Virtual Collaboratory (EVC)?** It's a collaborative virtual environment (i.e., a computer system that supports human-human and human-machine communication and collaboration) for x-ray experiments at [Elettra Synchrotron Light Laboratory](#), Trieste, Italy.
- **What can I do with the EVC?** The EVC allows a distributed team of researchers to perform a complete x-ray experiment, from data collection to structure publication. If you are performing an experiment at Elettra, the EVC allows you to involve in your work other researchers (chosen by you) around the world, share with them the experimental data, exchange ideas, discuss problems, and so on.

**EVC news (latest 3) [all news]**

**2003-09-24 15:30**  
EVC will be presented at [SMAU 2003](#).

**2002-10-10 10:30**  
EVC users at Elettra can now archive their data using the EVC Archiver tool. Supported media are DDS, CD, DVD.

**2002-10-02 14:00**  
Users at Elettra can now connect their laptop and get their EVC data via sftp. Read EVC Docs for details.

**Machine status**

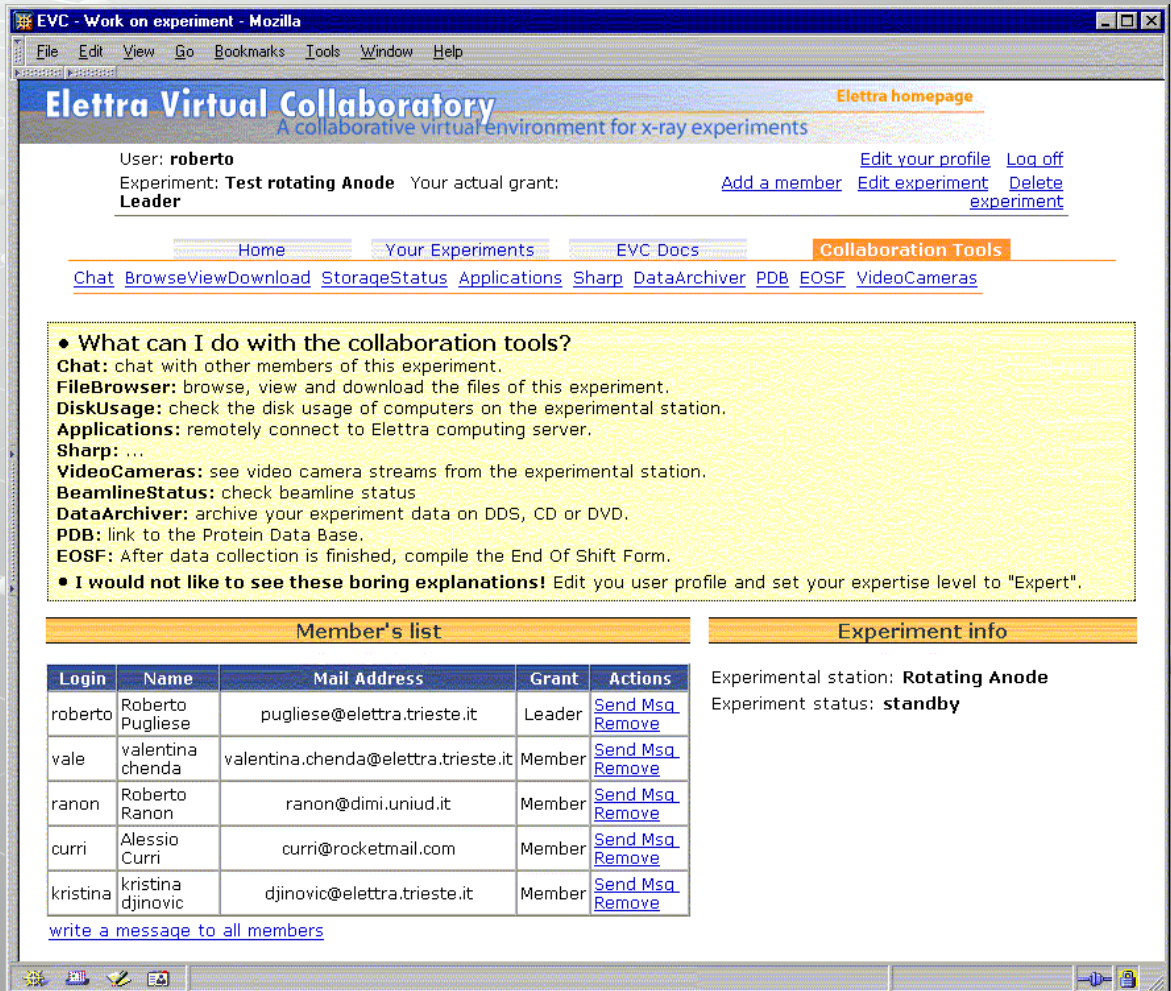
<b>Date</b>	Wed Sep 24 11:45:54 GMT+02:00 2003
<b>Current</b>	5.98 mA
<b>Injection Rate</b>	0.066 mA/s
<b>Energy</b>	0.911 GeV

**Machine information**

Machine status: Machine dedicated  
Next Users dedicated injection: 25 Sep at 7:00  
Machine Operator: Medessi G.

# Collaborating to an EVC project

- Scientists working to an EVC project can use many project related collaboration tools
- EVC presents an adaptive interface changing to suite the category and expertise level of the user



**Elettra Virtual Collaboratory**  
A collaborative virtual environment for x-ray experiments

User: **roberto** [Edit your profile](#) [Log off](#)  
 Experiment: **Test rotating Anode** Your actual grant: [Add a member](#) [Edit experiment](#) [Delete experiment](#)  
 Leader

[Home](#) [Your Experiments](#) [EVC Docs](#) [Collaboration Tools](#)

[Chat](#) [BrowseViewDownload](#) [StorageStatus](#) [Applications](#) [Sharp](#) [DataArchiver](#) [PDB](#) [EOSF](#) [VideoCameras](#)

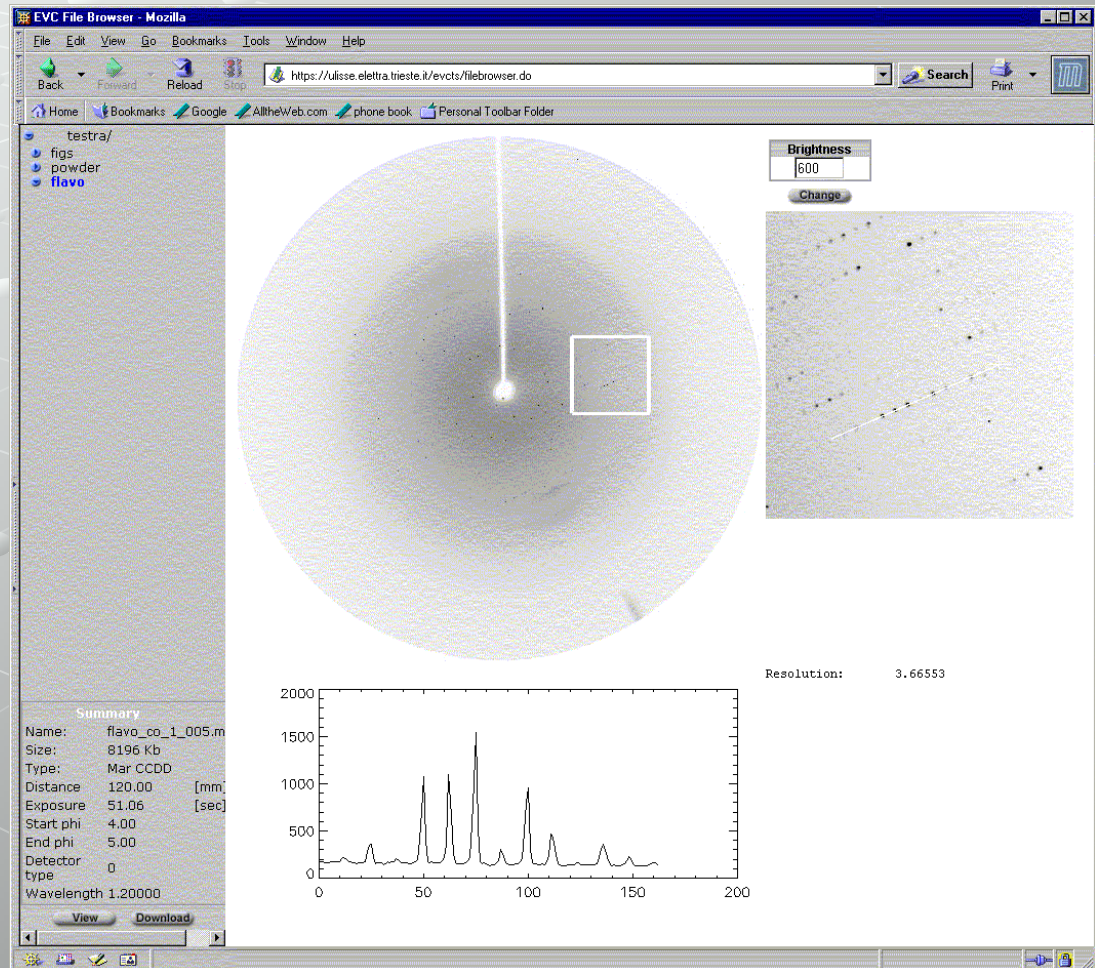
• What can I do with the collaboration tools?  
**Chat:** chat with other members of this experiment.  
**FileBrowser:** browse, view and download the files of this experiment.  
**DiskUsage:** check the disk usage of computers on the experimental station.  
**Applications:** remotely connect to Elettra computing server.  
**Sharp:** ...  
**VideoCameras:** see video camera streams from the experimental station.  
**BeamlineStatus:** check beamline status  
**DataArchiver:** archive your experiment data on DDS, CD or DVD.  
**PDB:** link to the Protein Data Base.  
**EOSF:** After data collection is finished, compile the End Of Shift Form.

• I would not like to see these boring explanations! Edit you user profile and set your expertise level to "Expert".

Member's list					Experiment info
Login	Name	Mail Address	Grant	Actions	
roberto	Roberto Pugliese	pugliese@elettra.trieste.it	Leader	<a href="#">Send Msg</a> <a href="#">Remove</a>	Experimental station: <b>Rotating Anode</b> Experiment status: <b>standby</b>
vale	valentina chenda	valentina.chenda@elettra.trieste.it	Member	<a href="#">Send Msg</a> <a href="#">Remove</a>	
ranon	Roberto Ranon	ranon@dimi.uniud.it	Member	<a href="#">Send Msg</a> <a href="#">Remove</a>	
curri	Alessio Curri	curri@rocketmail.com	Member	<a href="#">Send Msg</a> <a href="#">Remove</a>	
kristina	kristina djinovic	djinovic@elettra.trieste.it	Member	<a href="#">Send Msg</a> <a href="#">Remove</a>	
<a href="#">write a message to all members</a>					

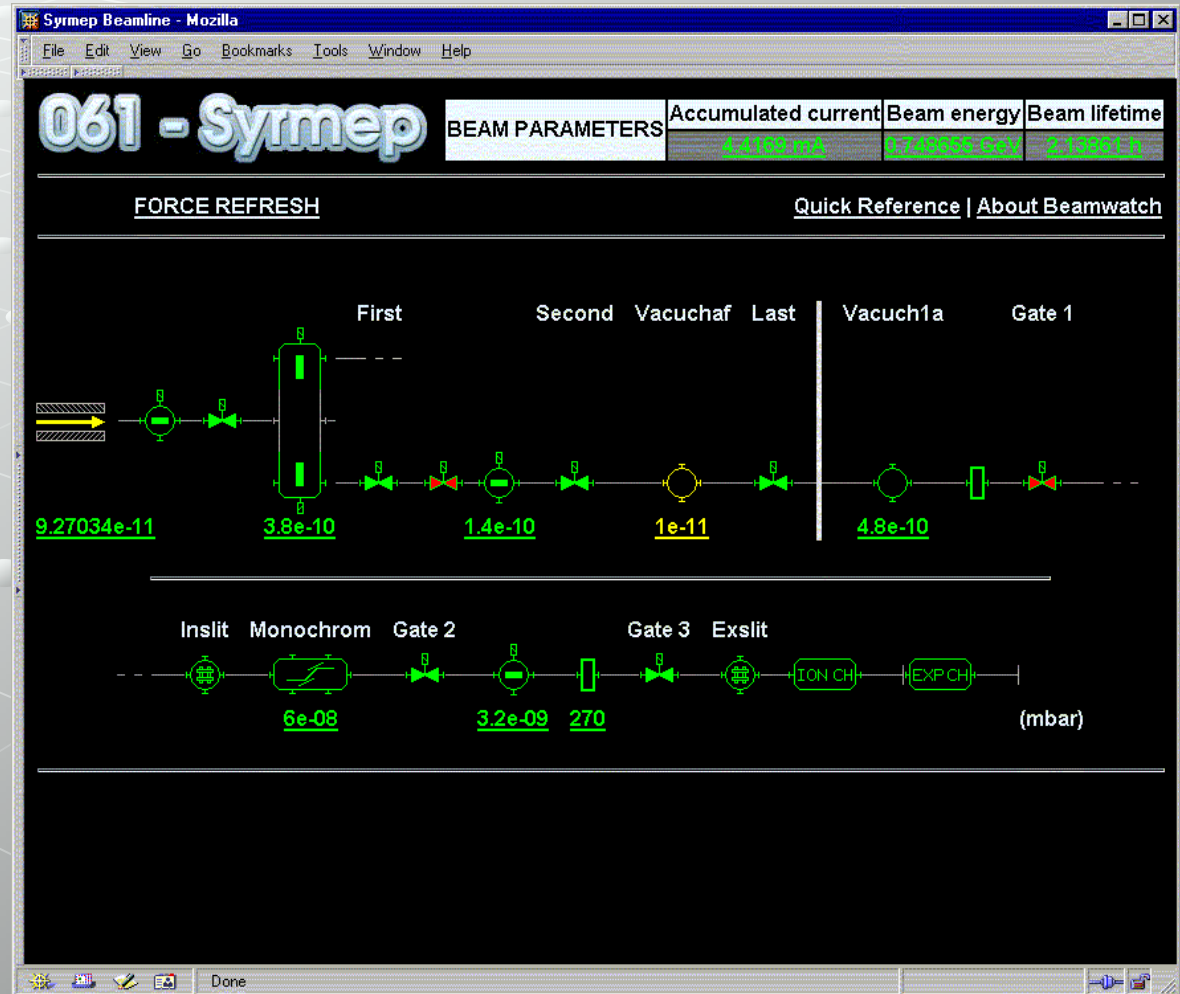
# Collaboration tools: scientific visualisation

- Scientists can browse, visualise and process remotely scientific data in real-time as soon as the data is collected



# Collaboration Tools: Remote Beamline Control and Supervision

- Beamwatch presents a synoptic view of the beamlines
- Autorised people can thus operate remotely on the beamline instrumentation





# Collaboration Tools: Electronic Notebook

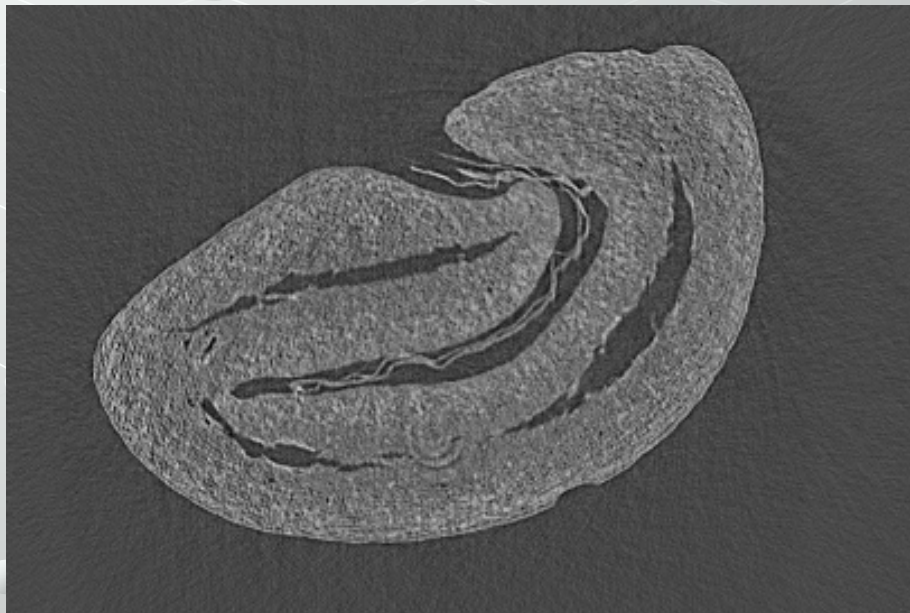
- Web application which substitutes the Beamline LogBook registering meaningful beamline events using a wiki-weblog methaphor
- Events can be entered manually or automatically by a program
- Texts and images are automatically indexed and hence easily searchable and browsable

The screenshot shows a Mozilla browser window titled "EVC - Notebook - Mozilla". The address bar contains the URL "https://ulisse.elettra.trieste.it/evcts/showNotes.do". The page header includes "Elettra Virtual Collaboratory" and "A collaborative virtual environment for x-ray experiments". The user is logged in as "roberto" with the experiment "Remote Operations" and the grant "Member". The interface features a navigation menu with "Home", "Your Experiments", "EVC Docs", and "Collaboration Tools". The main content area is divided into two columns. The left column, titled "Notes [List] [New]", displays a list of entries with timestamps and titles, each with "View" and "Edit" links. The right column contains a "Search" box and a "Calendar" for October 2003. The calendar shows the days of the week and dates, with some dates highlighted in blue. The browser's status bar at the bottom indicates "Document: Done (0.191 secs)".



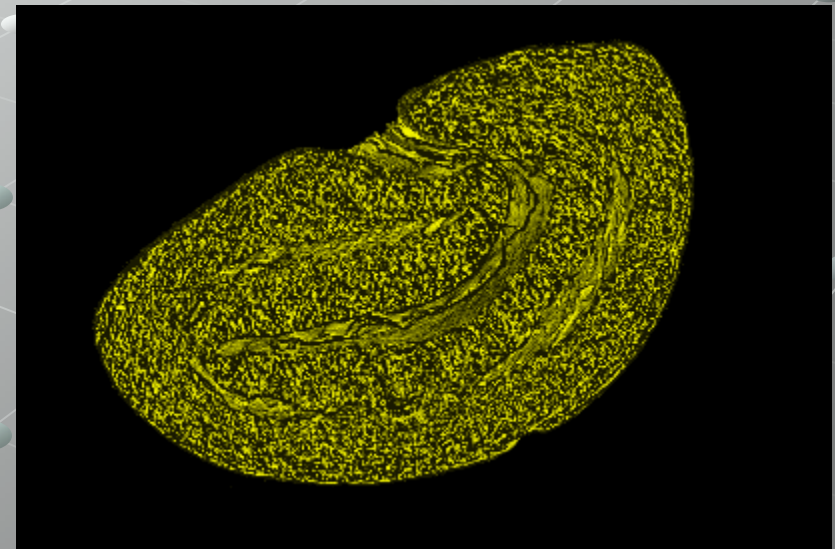
# EVC usage scenarios: real time micro-tomographic reconstructions

- EVC will be used to allow real-time micro-tomographic reconstructions on SYRMEP beamline, as soon as data is collected



*Reconstructed stack of 279 slices*

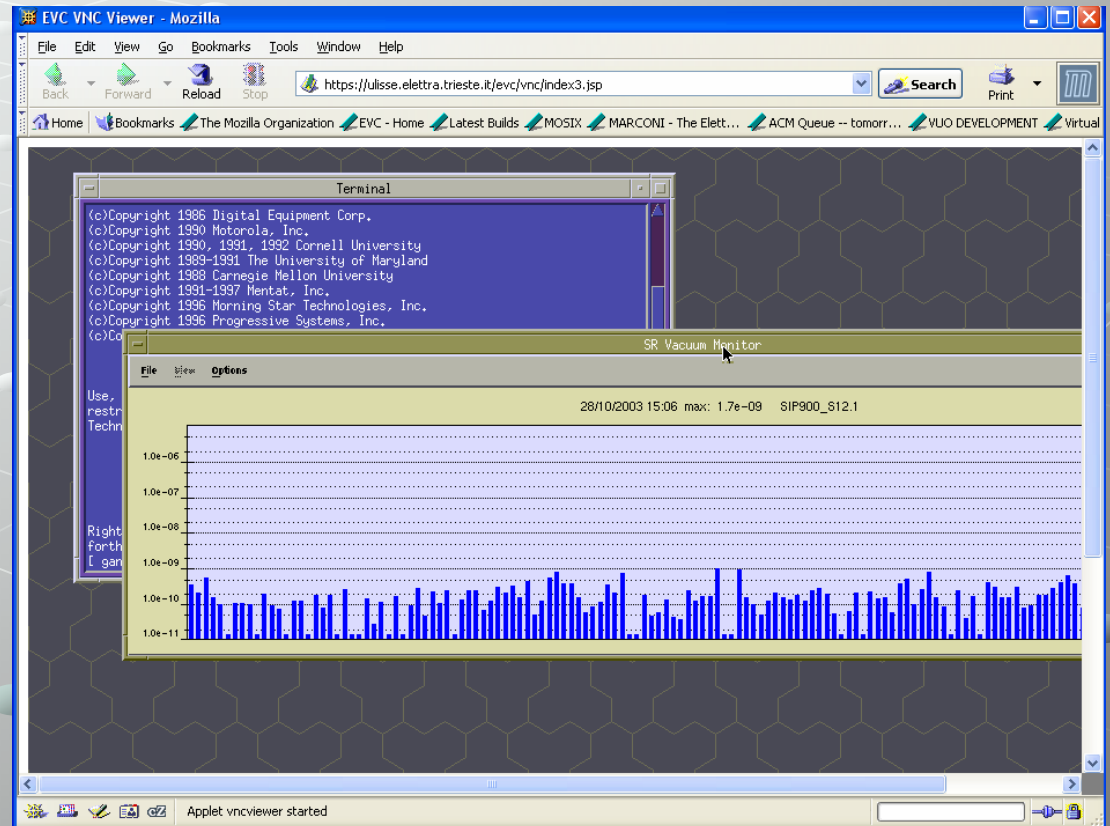
2 mm



*Volume rendering*

# EVC usage scenarios: remoting the ELETTRA control room

- A new experimental station was recently added to EVC: ELETTRA Control Room
- Available tools for this experimental station are:  
Chat, Scientific File Browser, OperatorConsole, ENotebook, LogAnalyzer

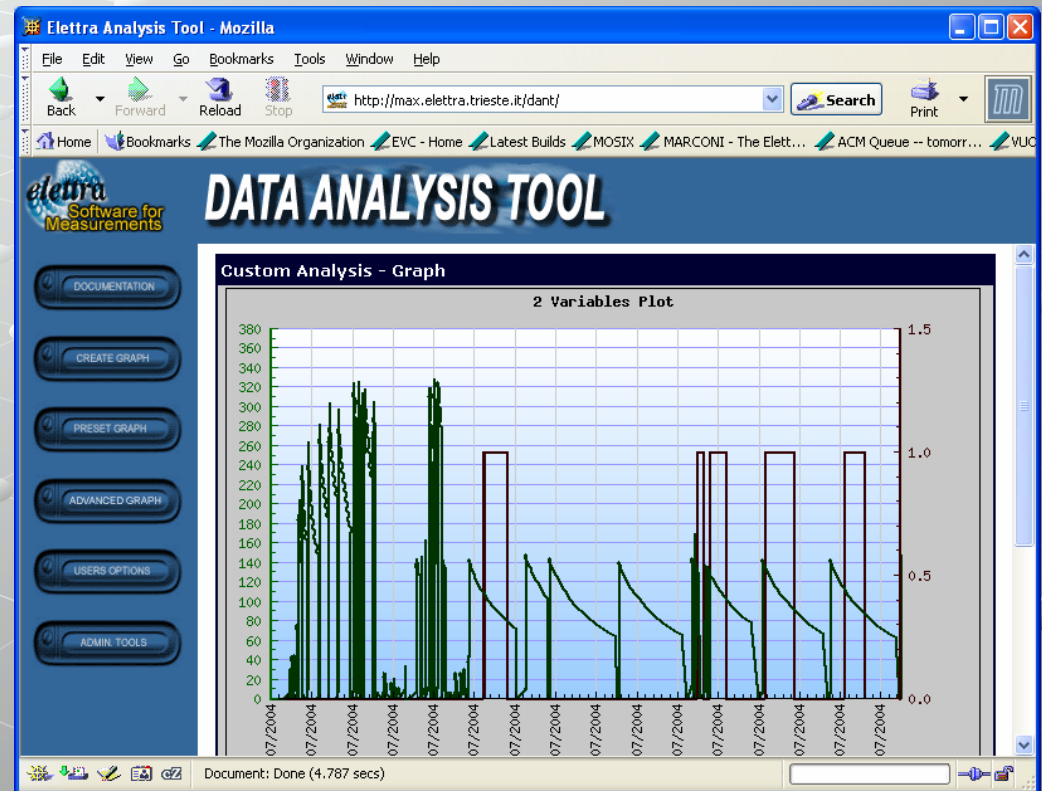




# Collaboration tools: Advanced LogAnalyzer

- Advanced LogAnalyzer is a web application which allows to select variables from the logfiles produced by different control and supervision systems, and to plot them in a user specified temporal interval

- Advanced LogAnalyzer is technically a data warehouse, modular both considering the data loading and the data visualisation (Visual Data-Mining)





# About DANT...

- Modular, web-based Log Analysis Tool
- Capable of storing a large amount of heterogeneous data into a unique, homogeneous data warehouse (thus facilitating information retrieval).
- Allows (remote) data analysis
- Doesn't require specific hardware or software to run. It can run properly on any PC with a moderate amount of RAM and CPU speed (IE 256MB 350Mhz).

# DANT Goals

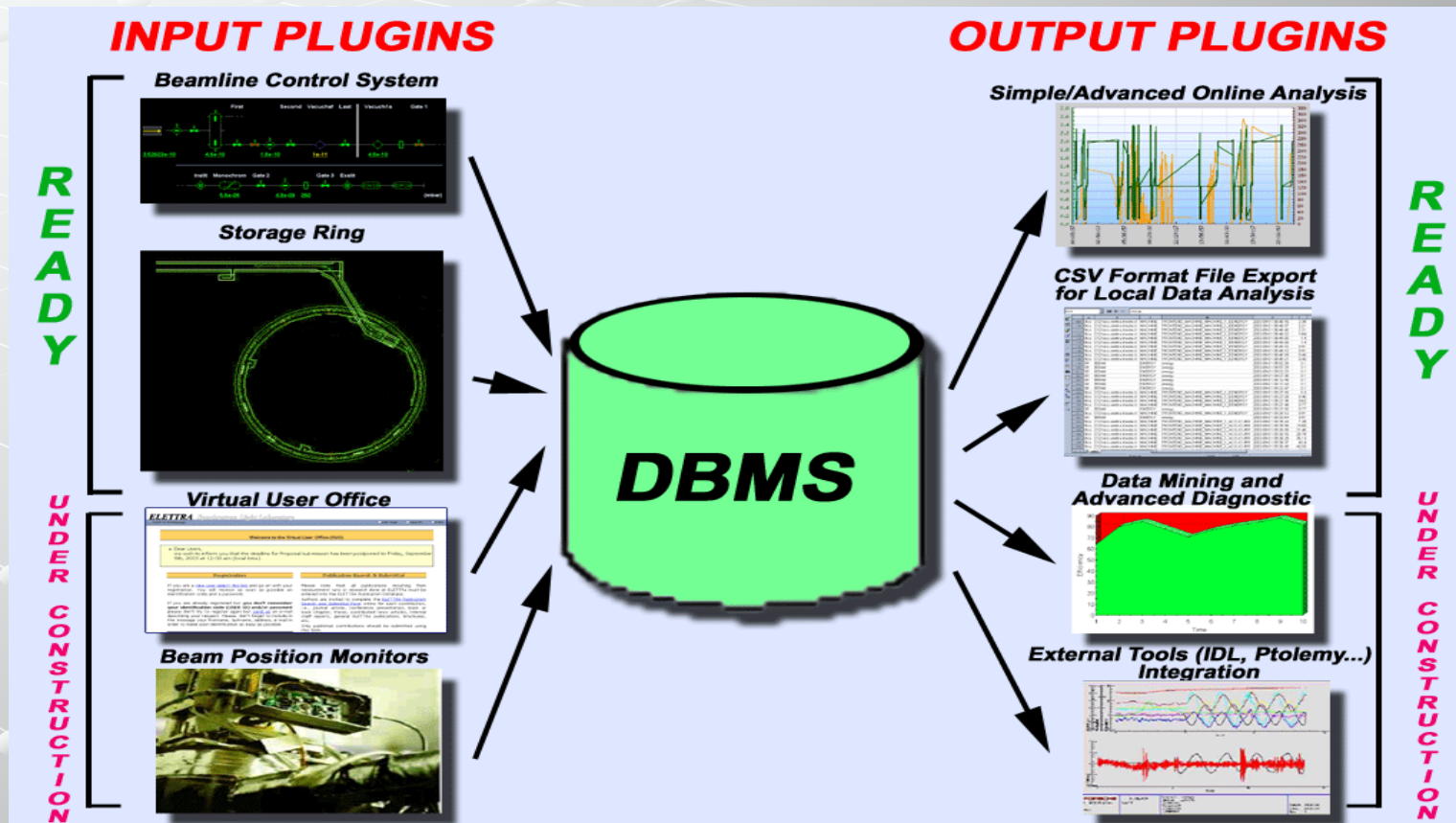
## ● Problem Context

- Elettra Scientific Instruments generate Log files with different structures and store them in different locations. This obviously doesn't make any analysis easy.
- Current log files often include significant redundancies and thus are easily prone to inconsistencies or unnecessary large storage requirements.

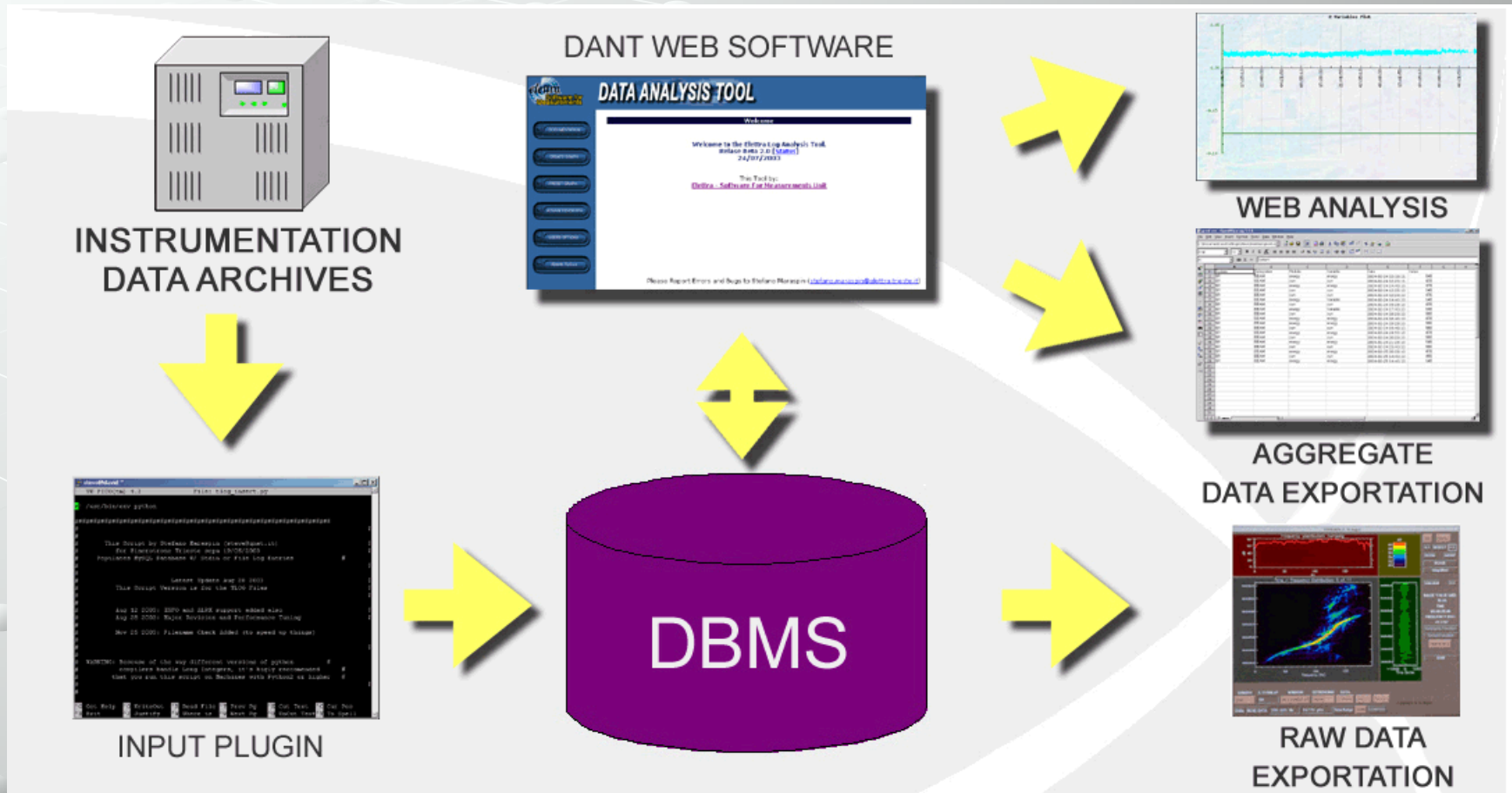
## ● Goals

- Collect all the instruments data logs into a unique, reliable and efficient data base system.
- Permit therefore easy data analysis or comparisons, possibly with other already available databases and tools (i.e. Elettra VUO).

# Software Architecture



# Software Architecture





# Software Features - 1

- **Abstract and Layered development Approach (Portability and Customization always kept in mind)**
  - **Abstract DB Connections:** DANT can run upon most common DB systems, including for example: Oracle, Mysql and PostgreSQL.
  - **Only software requirement** is to have a **php engine** installed on the (server) machine where the tool is intended to be run. Nowadays almost every web server supports php.
  - **Cross Platform Application:** being developed as a web tool, DANT offers full cross platform support. All of the processing and data handling occur on the server side, leaving to the connected clients just the task of displaying some HTML code in a web browser.



# Software Features - 2

## ● Low Cost

- Entirely based upon Open Source Software.

## ● Easy Deployment and Configuration

- There's a single configuration file keeping track of all system variables (IE Database Type, Maximum allocable memory..).
- This Configuration file can be additionally generated during installation, so that deployment and usage are straight forward.

## ● Modular (thus easily Scalable)

- Both Input (indexing) and output (presentation, visualization) extensions to the DB system are built as modules, thus allowing users to choose (or develop) the most adapt modules to the situation.

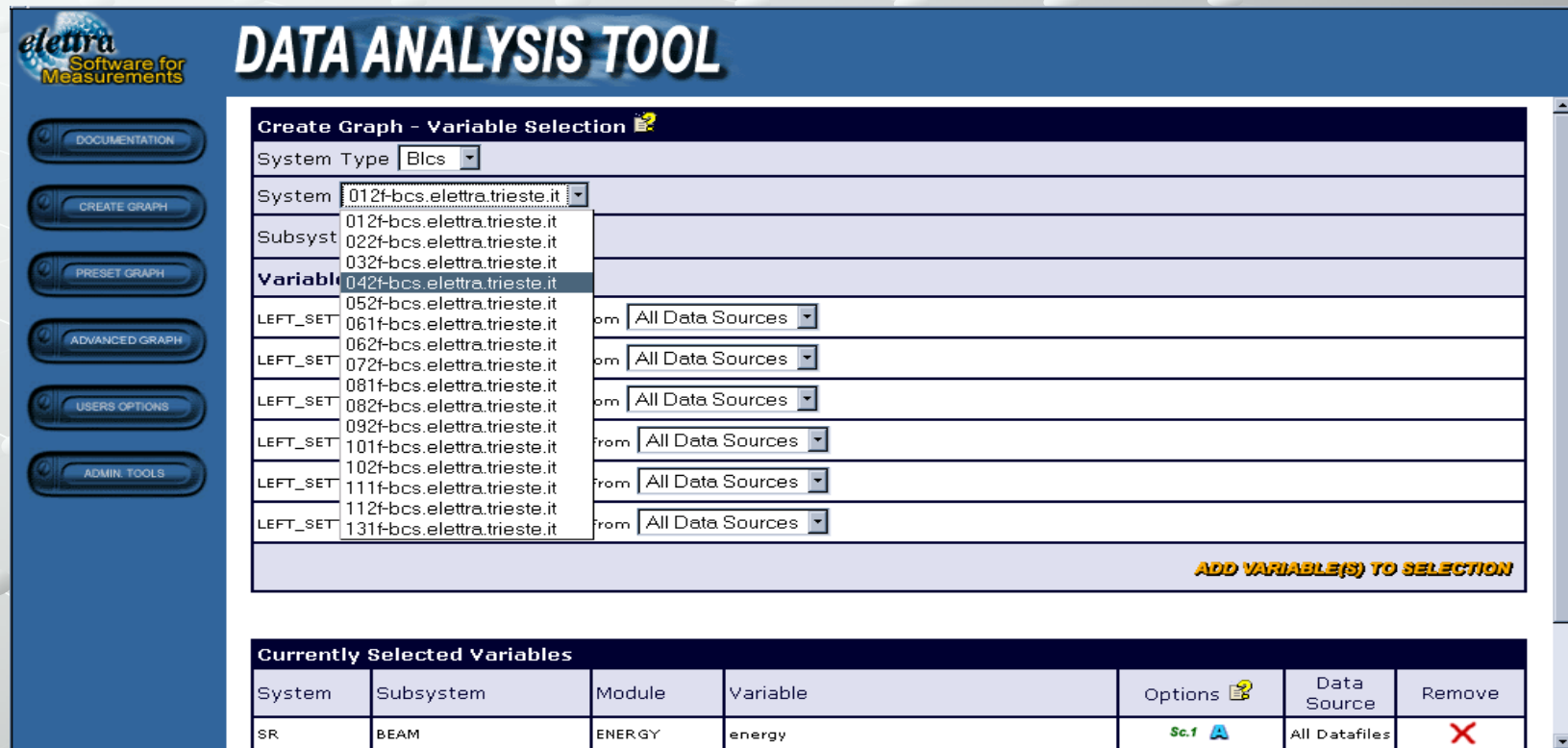


# Input Module Example: BCS Logs

- Implemented as a Python script
- Receives data either from STDIN, named pipes (remote Unix syslog) or Archived Log Files
- Capable of reading and Indexing compressed data (gzipped archives)
- Has SMART Hierarchical Indexing capability (no need to set up when new instruments are added).

# Sample Use Case – Step 1

## Step 1 – Variable Selection



**DATA ANALYSIS TOOL**

**Create Graph - Variable Selection**

System Type: Blcs

System: 012f-bcs.elettra.trieste.it

Subsystem: 022f-bcs.elettra.trieste.it

Variable: 042f-bcs.elettra.trieste.it

LEFT\_SET: 052f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 061f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 062f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 072f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 081f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 082f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 092f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 101f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 102f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 111f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 112f-bcs.elettra.trieste.it from All Data Sources

LEFT\_SET: 131f-bcs.elettra.trieste.it from All Data Sources

**ADD VARIABLE(S) TO SELECTION**

Currently Selected Variables						
System	Subsystem	Module	Variable	Options	Data Source	Remove
SR	BEAM	ENERGY	energy	Sc.1 A	All Datafiles	X

# Sample Use Case – Step 2

## Step 2 – Time Window/Options Definition


**Create Graph - Step 2**

**Period Selection**

Day A  /  /  Time A  :  :

Day B  /  /  Time B  :  :

(Date Format is: DD/MM/YYYY HH:mm:ss)

**Output Plugin Selection** 

First Scale is Logarithmic

Second Scale is Logarithmic

Show Horizontal Grid

Show Vertical Grid

Gradient Background

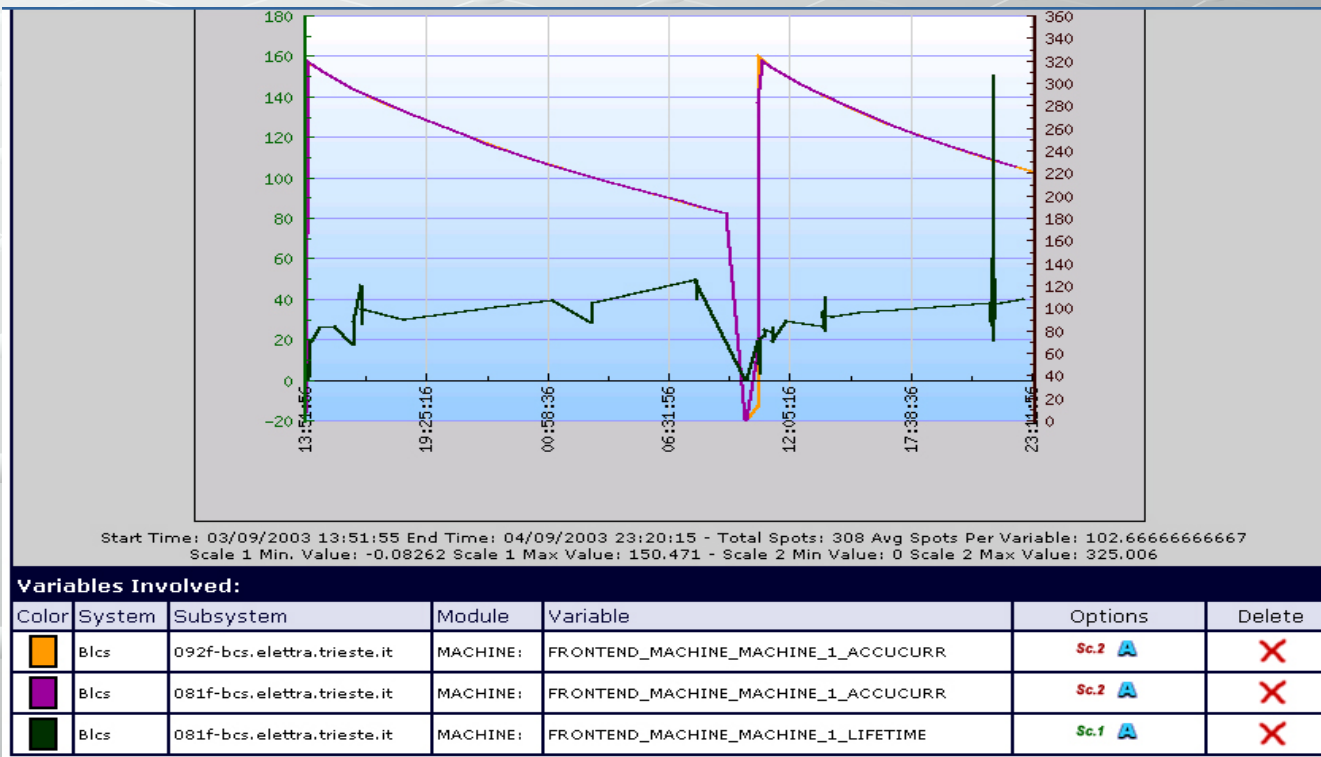
Show Variables Value on Plot

Show Row Data Also

**PROCEED WITH CURRENT SETTINGS**

# Sample Use Case – Step 3

## Step 3 – Graphic Render/Data Export



# Advanced Analysis

## ● Combined Plots Drawing Capability

- Used not to plot variable values in a timeline but rather relate different variables together in a single plot, displaying the result of a certain expression.
- Sample Use Case: Measure Machine Efficiency and find out actual beamtime usage during a specific User's Shift.

Currently Selected Variables						
ALIAS	System	Subsystem	Module	Variable	Data Source	Remove From List
V1	Blcs	092f-bcs.elettra.trieste.it	MACHINE:	FRONTEND_MACHINE_MACHINE_1_ACCUCURR	All Datafiles	X
V3	Blcs	081f-bcs.elettra.trieste.it	MACHINE:	FRONTEND_MACHINE_MACHINE_1_ACCUCURR	All Datafiles	X
V4	Blcs	081f-bcs.elettra.trieste.it	MACHINE:	FRONTEND_MACHINE_MACHINE_1_LIFETIME	All Datafiles	X
V5	Blcs	012f-bcs.elettra.trieste.it	DIGI71:	FRONTEND_SETTINGS_DIGI_1_B25	All Datafiles	X
V6	Blcs	012f-bcs.elettra.trieste.it	DIGI71:	FRONTEND_SHUTTER_DIGI_1_B12	All Datafiles	X

**BACK TO VARIABLE SELECTION**

Custom Formulas			
You can use variables alias to combine different values in a "function style" representation plots. Y scale will always be time. <a href="#">Click here for a full list of available operators and functions</a>			
z	v4	on	Scale 1
v	v1/v3	on	Scale 1
u	v5 AND v6) AND (v4 > 220)	on	Scale 2

**PROCEED WITH CURRENT SETTINGS**

# DANT Status ...

- Currently Supported Logs: **Beamline Control System, Beamline Access Control System Storage Ring**
- Average archived records Indexing (hourly): **35.000**
- Maximum Acceptable Fetched Values for a Single Plot [upon current hardware]: **35.000**
- Average Monthly Stored Records: **around 25.000.000**
  - Probably swap-in/swap-out of older data will be required
- Record Size: **94bytes**
- Average Compression Factor **1:8**

# ... and future developments

## ● Integration with Other Analysis Tools/Instruments

- Add Java Support (applet analysis)
- Permit Realtime Analysis
- New Presentation Modules development (IDL)

## ● Integration with Other Data Sources

- VUO (Virtual User Office)
- BPM (Beamline Position Monitor)
- [...]



# Acknowledgements

- Fulvio Bille', Michele Turcinovich, Roberto Borghes, Claudio Scafuri, Massimiliano Chiardone, Kristina Djinovic from ELETTRA
- Dorian Lamba, Alberto Cassetta from CNR