SMB Traffic Analyzer

Holger Hetterich
L3 Support Engineer
SUSE Linux Products GmbH

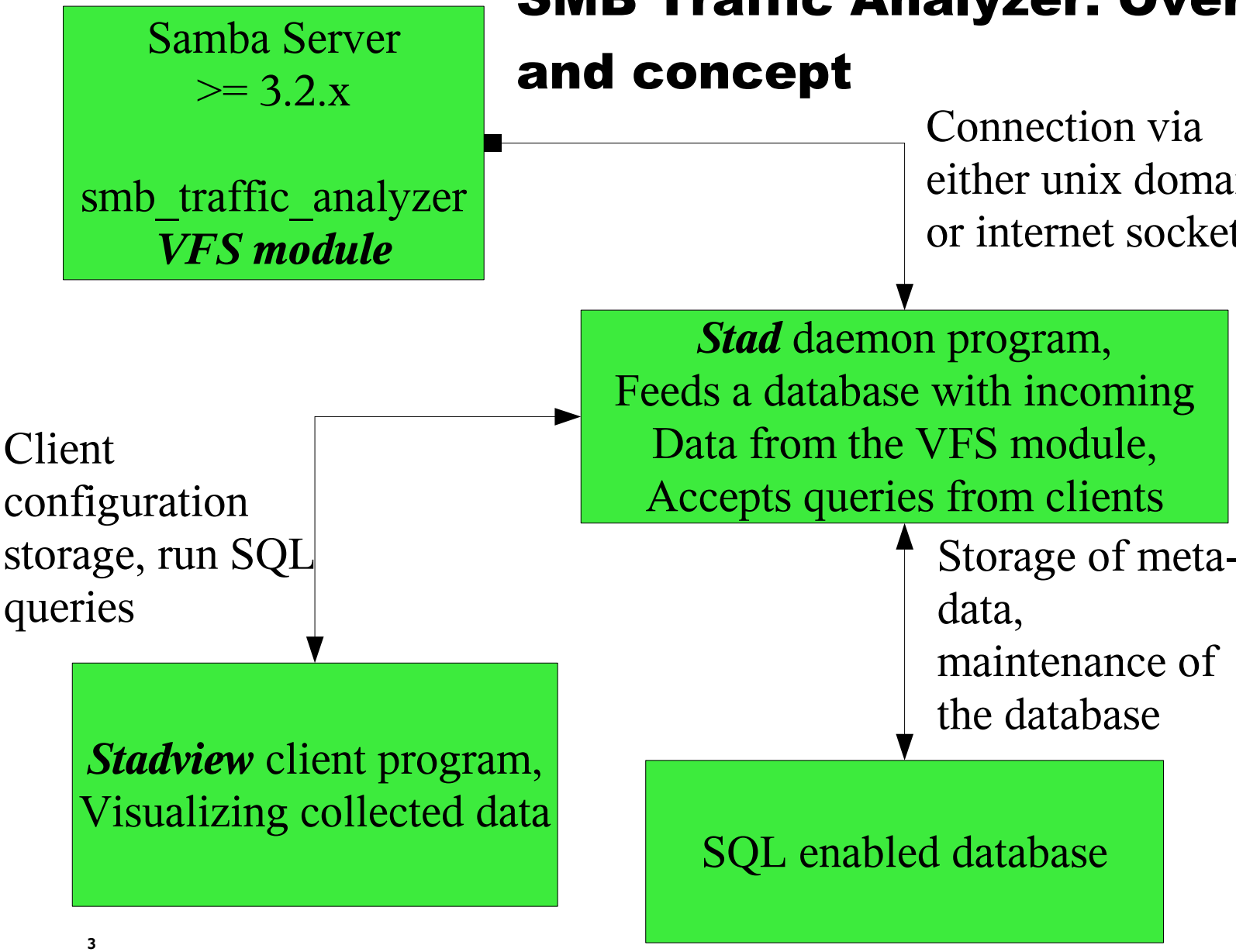
Novell

What is SMB Traffic Analyzer?

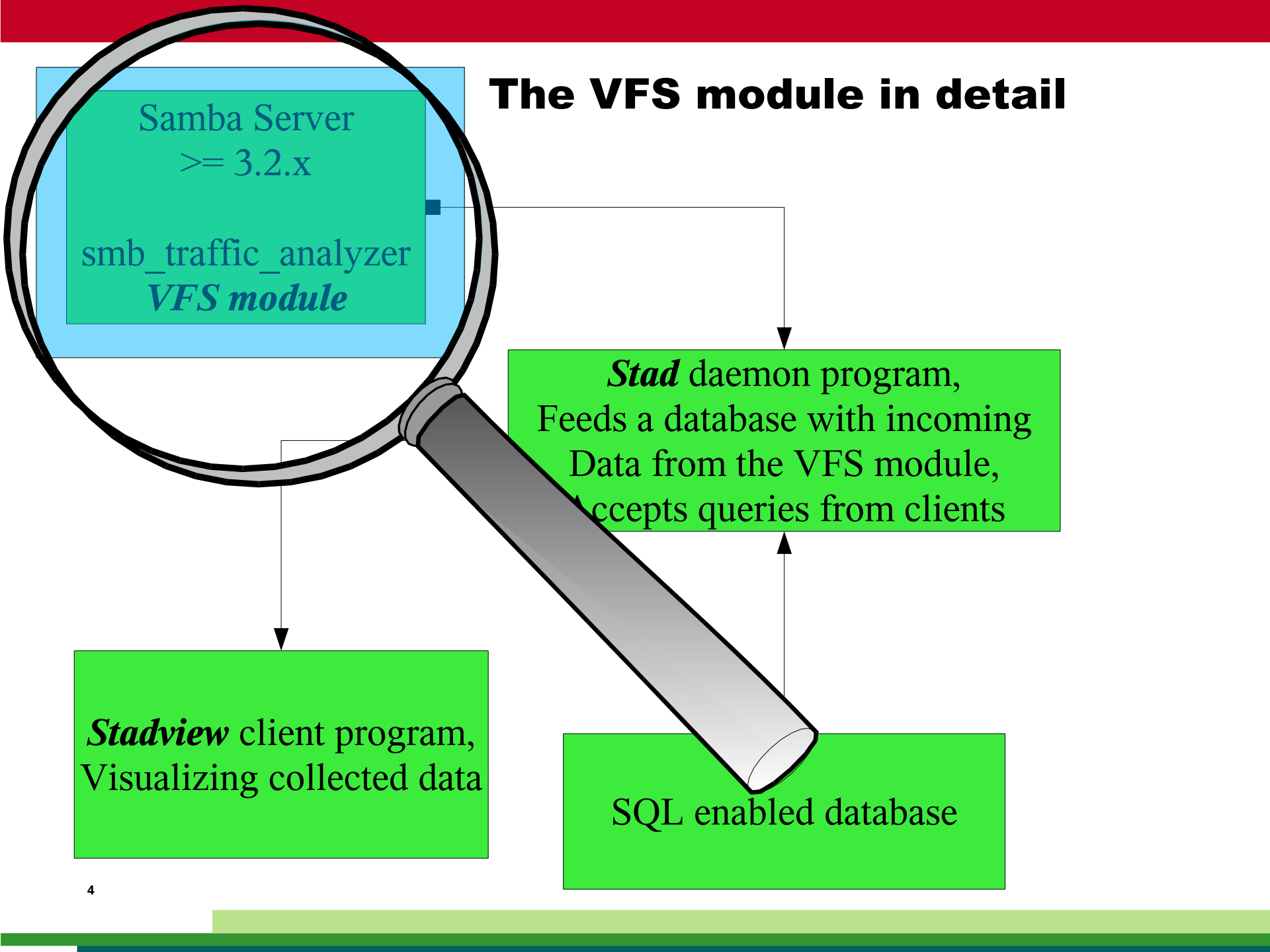
- Milestone 1 (current development)
  - Collect metadata of write and read processes on one or more samba servers.
  - Store data in a SQL queryable container.
  - Provide statistics based on this data, and visualize them

- Milestone 2 (long term goals) → later in this talk

SMB Traffic Analyzer: Overview and concept



The VFS module in detail



Looking at VFS
smb_traffic_analyzer(1)

"Sit in the VFS layer of samba, watch any Read or Write- file operation, and send data about this operations via network to a receiver."

- Introduced into the samba source tree at September 25, 2008
  - Included in 3.2.x and 3.3.x, SLE11 ships it.
- Fully transparent VFS module
- Configured easily via smb.conf
- Can operate either on a unix domain socket or on an internet socket

Looking at VFS smb_traffic_analyzer(2) – what data is stored?

- ***Length of a data transfer*** in bytes
- If the transfer was a ***Read or Write*** access
- The ***name of the file*** involved in the transfer
- The ***name of the user*** who initiated the transfer
- The ***name of the domain*** under which the transfer happened
- A ***timestamp*** including date and time to the millisecond

Looking at VFS smb_traffic_analyzer(3)

A sample share configured for smb_traffic_analyzer.

```
[pool1]
        path=/pool1
        read only = No
        vfs objects = smb_traffic_analyzer
           smb_traffic_analyzer:host = localhost
           smb_traffic_analyzer:port = 3490
```

Or activating the object in the global section activates all shares:

```
[global]
        vfs objects = smb_traffic_analyzer
           smb_traffic_analyzer:host = localhost
           smb_traffic_analyzer:port = 3490
```

Looking at VFS smb_traffic_analyzer(4) – Is this legal at all?

Exposing user related data is illegal in many countries.

Two ways of anonymization:

PREFIX + Hash-number:

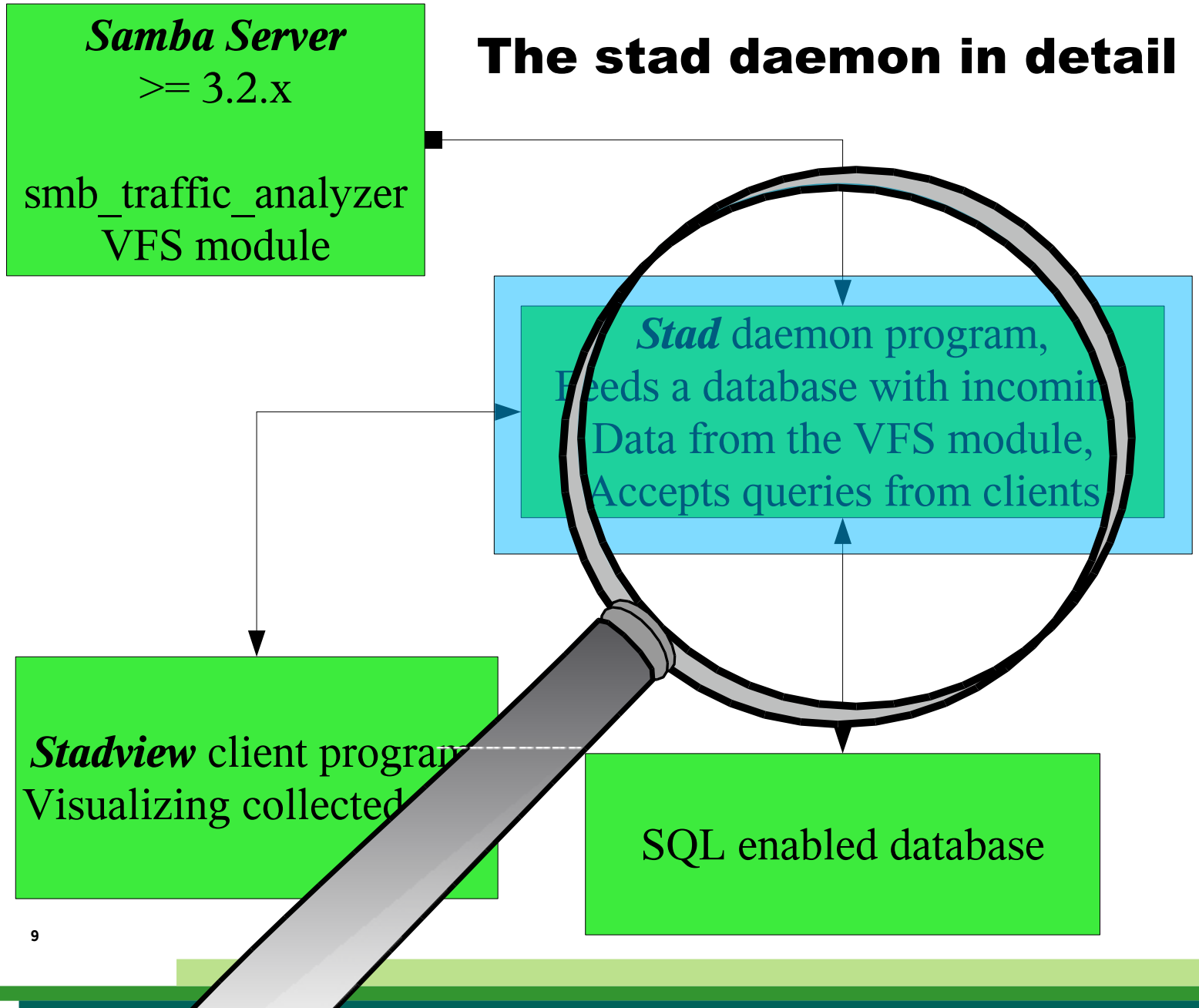
you can still recognize individual users:

```
smb_traffic_analyzer:anonymize_prefix = User
```

PREFIX: map any username to one string.

```
smb_traffic_analyzer:anonymize_mode = Total
```

The stad daemon in detail



Looking at the stad daemon(1)

"Be as fast as possible at accepting data from the VFS module, put the data into a SQL enabled storage. At the same time, maintain the database at the user's wish, and accept SQL queries about the data from clients."

- Configured via command line or ini-style config file
- Caching data into the sytem RAM
- Accepts multiple clients and VFS modules

Looking at the stad daemon(2)

Simplest possible way to run stad:

```
$ stad -i 3490
```

- use database /var/lib/staddb with sqlite3
- accept VFS connection on internet socket port 3490
- accept client connections on internet socket port 3491

Advanced configuration is done in a config file, see the manpage.

Looking at the stad daemon(3) – data processing plugins

Why not a simple text file?

Stad supports a plugin architecture for it's data processing. If there is no interest in a networked client or SQL queryable information, an other plugin can do the output.

Currently shipping:
- ***sqlite3*** Plugin
- ***CSV text file*** plugin
  (loadable with OpenOffice, Excel and friends)

In Development:
- ***MySQL*** Plugin
- ***syslog*** plugin

Looking at the stad daemon(4) – performance impact

Thinkpad X61 configured as “all in one” system, copying ***2.8 GB*** in ***3.381 file*** objects to a share.

→ running stad, the database and the samba server

Performance reduction : 39 %

By changing the conservative default parameters of stads system RAM usage, enlarging it's memory area to hold data packages for the db, it was possible to reach a

Performance reduction: 12%

Dell QuadCore Desktop system as Samba server, and ***Blade*** server running stad and the db, copying ***4.8 GB*** to a share.

→ stad and db separated from the samba server

Performance reduction : 16%

Looking at the stad daemon(5) – storage is limited

Without a regulating process, the database that is maintained by stad would grow and grow.

This parameters in the config file will run the maintenance process ***every hour***, and delete any data that is older than ***5 days*** from the database.

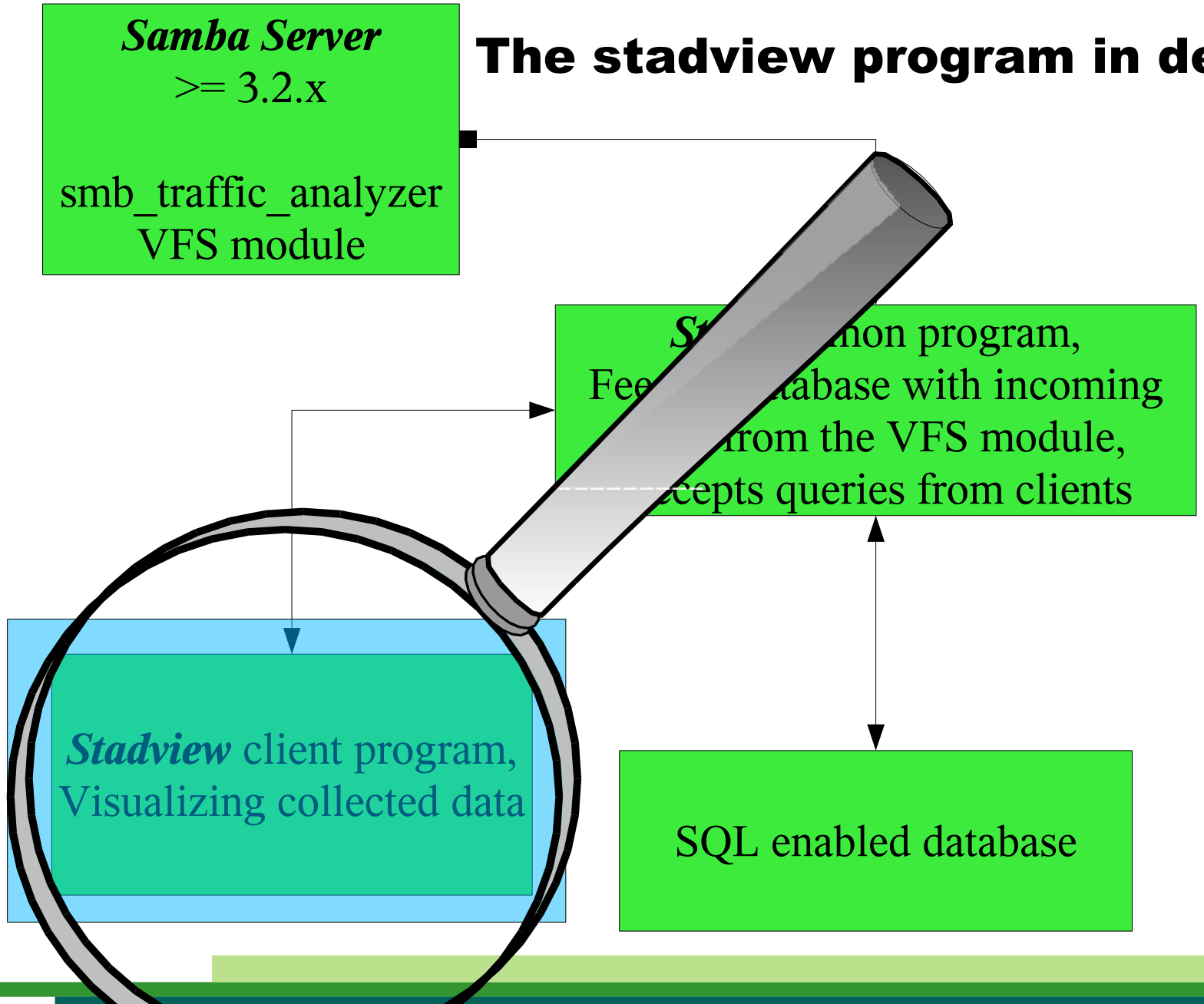
```
[maintenance]
maintenance_timer = 01:00:00
maintenance_timer_config = 5,00:00:00
```

Looking at the stad daemon(6) – stadtorture - A test utility for stad

Stadtorture is a tool utilizing libsmbclient to produce traffic on a server.

- → creates a file set on two samba shares
- → copies files around with pauses
- → can record it's own run and playback
  - → turn into a benchmarking tool

The stadview program in detail

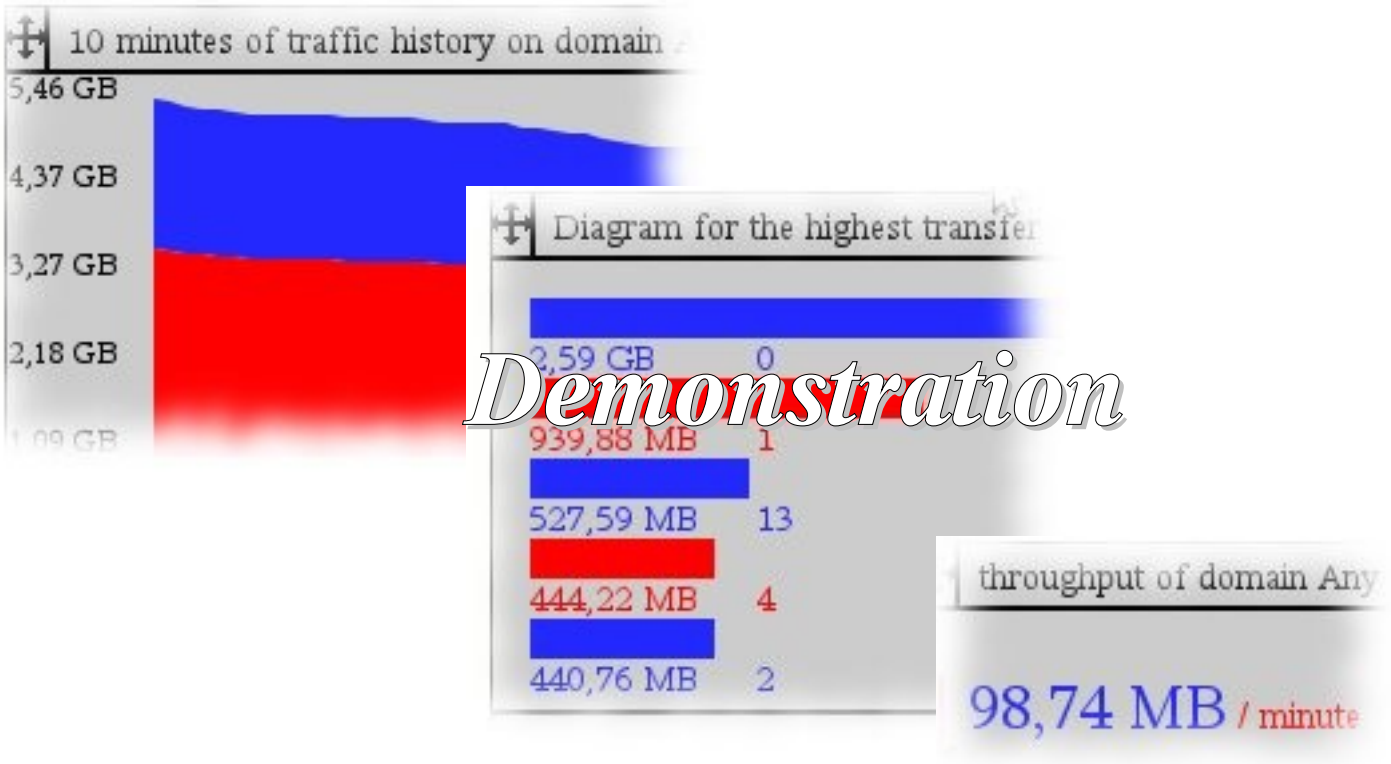


Looking at the stadview program(1)

Show statistics about the data stored in the database. Update this information in real-time. Also try not to crash while being demonstrated at sambaXP :)
At some day, run on the Windows platform.

- utilitzing GTK and Cairo graphics
  - easy output to different devices
- no local configuration
- manage multiple sessions on a stad server

Looking at the stadview program(2)



Where do we go to?

- Milestone 2 (long term goals)
  - Create appliances for stad:
    - Out of the box samba server/stad/stadview
    - Out of the box stad/stadview ready for network integration
  - Secure connections between VFS module/stad/stadview.
  - Collect **any** VFS operations (like full-audit)
  - Make table based applets for stadview
  - Make samba / stad a native Microsoft Event Viewer Source
  - Have stadview running on Windows

SMB Traffic analyzer – project data

Homepage:
http://holger123.wordpress.com/smb-traffic-analyzer/
Any component is ***GPLv3***.

Documentation:
Detailed manpages with examples for all components.

Around five people are working on SMB Traffic analyzer project since January 2008.

Main contact in case of questions, patches or suggestions:
ozzy@metal-district.de (Holger Hetterich)
hhetter@novell.com

SMB Traffic analyzer

http://holger123.wordpress.com/smb-traffic-analyzer/

QUESTIONS AND ANSWERS