

DIE ÖSTERREICHISCHE BIBLIOTHEKENVERBUND  
UND SERVICE GMBH



obv sg

# MEASUREMENT OF INDIVIDUAL LICENSE USAGE IN CLUSTERED ALEPH INSTANCES

WOLFGANG HAMEDINGER  
THE AUSTRIAN LIBRARY NETWORK AND SERVICES LTD

8. SEPTEMBER 2008  
IGeLU 2008, MADRID

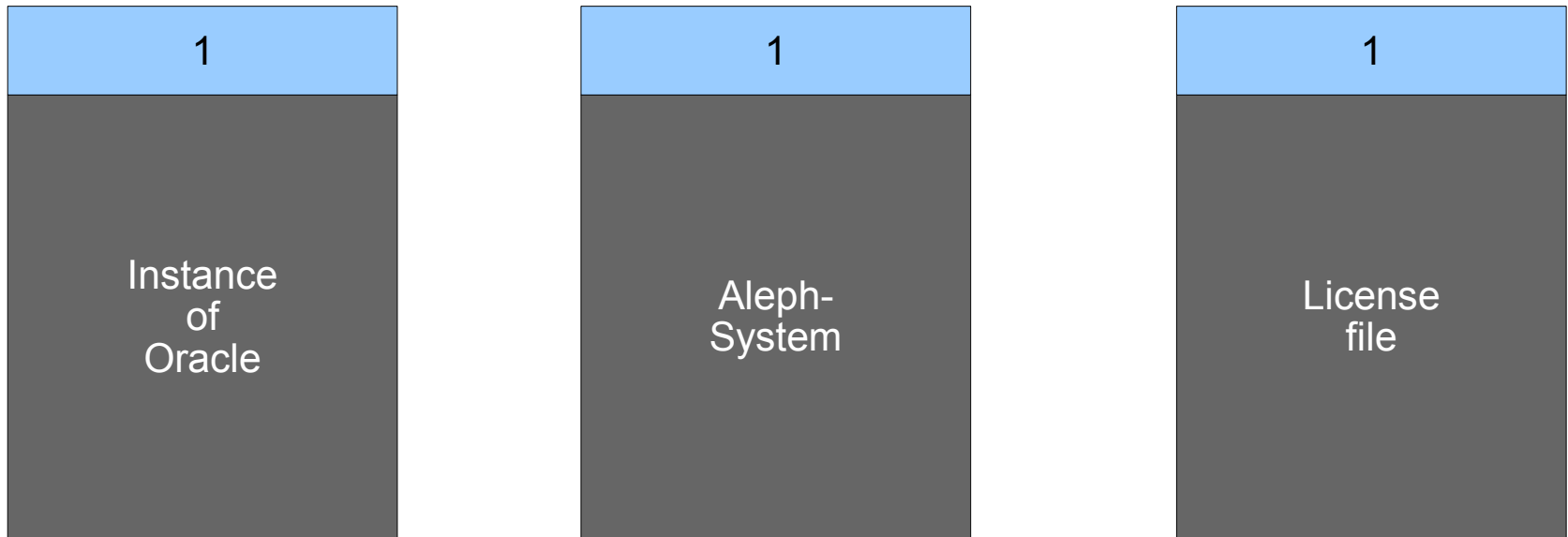
## Overview

- Problem description
- Starting position and ideas
- Used informations
- Description of solution – the big picture
- Example results
- Wishes
- Summary

## License Model of Aleph

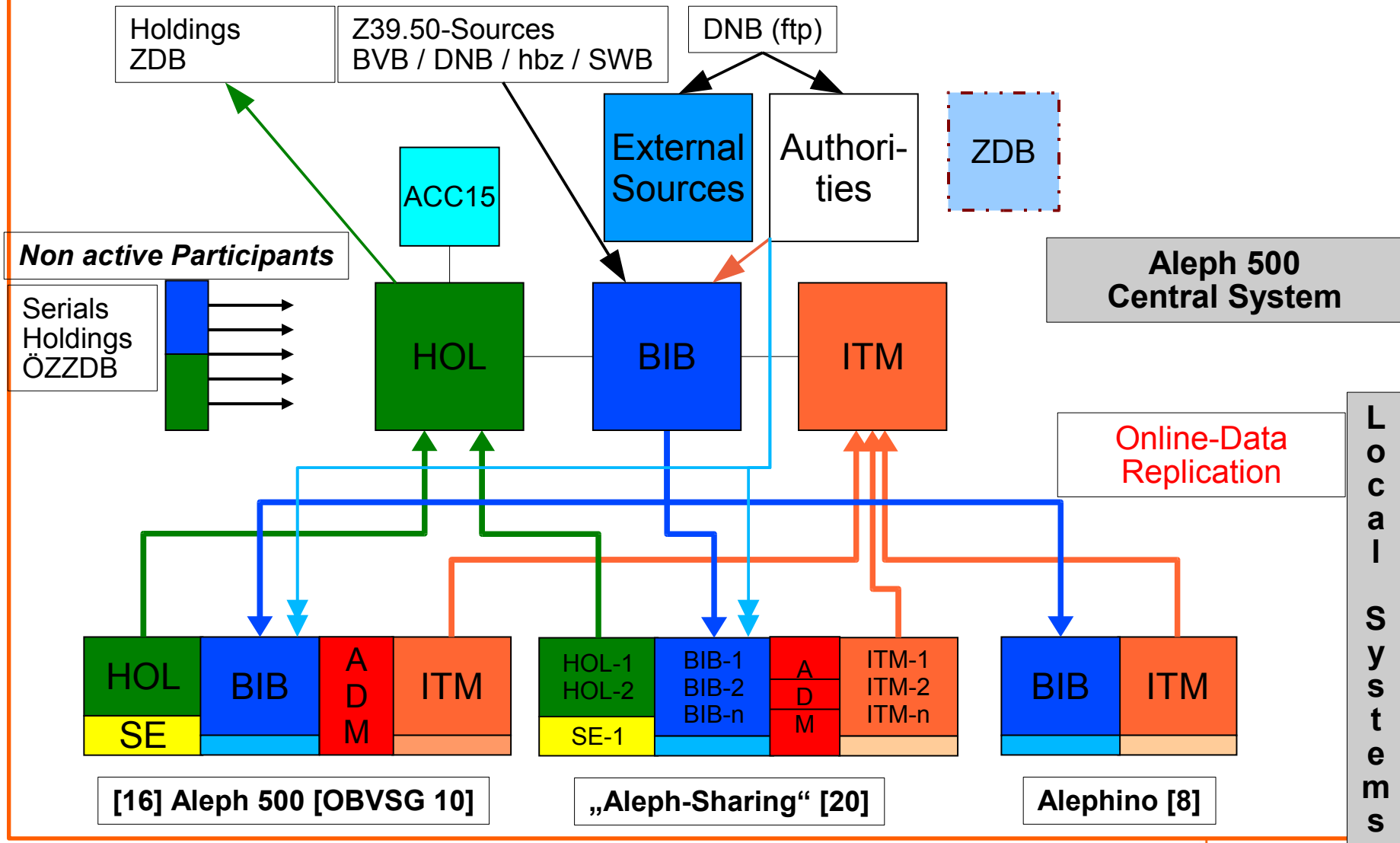
- Aleph uses the model of „Concurrent users“
- Activities of a certain user block a license for a defined period of time
- The standard value of this time out („Session Delta“) is 600 seconds
- There is an official tool to get the actual license usage – UTIL/Y/11

## Standard Situation

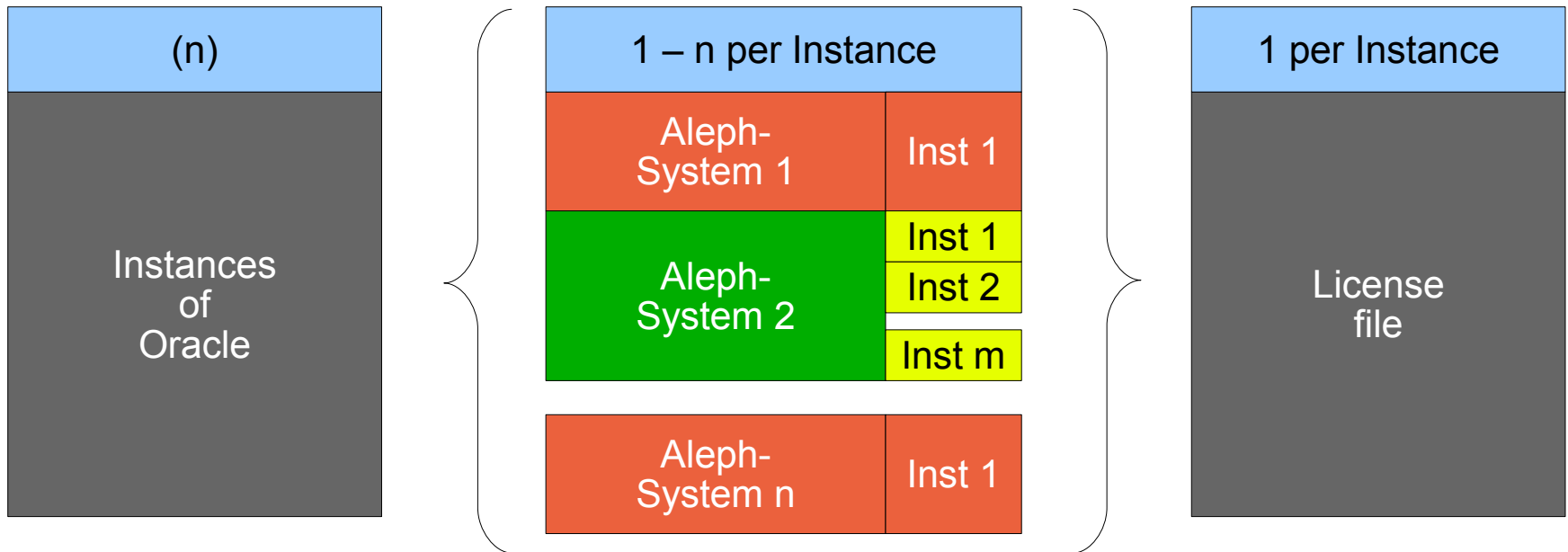


**So, what's the problem?**

# Data Sources and Data Flow within the Austrian Library Network



## Situation of a Host and Service Provider in General



- More than one Aleph-system within one Oracle instance
- Some Aleph-Systems shared by different institutions (ASP-model)
- One common license file for the whole instance
- “License limit exceeded”
  - Who has done it?
  - How often and why did it happen?

## Special Situation in the Austrian Library Network

- ALN was the first consortium with the architecture of a central catalogue, including a replication mechanism between systems – implemented in 1999
- A lot of tuning had to be done during the first time of operation
- GUI-Timeout was set to 300 seconds to optimize performance in version 11.5
- For every new release an adapted version of Aleph 500 had to be delivered to ALN
- In 2005 Ex Libris suggested to use the standard version for the future to avoid all the problems raising from this special case
- For adaptation of the license tables we had to find a way to measure the license usage with our timeout of 300 seconds and the standard value of 600 seconds simultaneously
- This new method should also solve the general problem of license measurement

It was the start of an interesting, somewhat tedious and unexpected time-consuming journey!

## Starting Position and Ideas

- Our first interests were only the staff licenses – no investigations for Web usage first
- Working version was 16.02
- Table z65 was known to us for a long time as source of information about license usage
  - obviously its undocumented
  - seems to give a snapshot of the current license situation
  - was used for long time to compare Ex Libris values with our own calculations
  - good fit, so we noticed changes of Session Delta immediately
  - unfortunately its not a “rich” table and contains only IP-addresses
- Assignment of IP-addresses or IP-ranges to single institutions did not look satisfying
  - “not natural”: actions are initiated by logged in persons with user-account and password, not by IP-addresses
  - especially for shared Aleph-systems affiliated, but different institutions may use the same IP-range
  - dynamically assigned IP-addresses can not be resolved
- Other sources for “information enrichment” with the login name had to be found
  - first candidate: Log file of pc-server



## The Main Source for License Usage: Table z65

### Structure

- **Z65\_REC\_KEY:** contains key <GUI|Z39>-<external IP>[-<internal IP>] or <IP :Port> (for SC-Server)
- **Z65\_TYPE:** <PC-SERVER|Z39-SERVER|SC-SERVER>
- **Z65\_TIME:** Update time (Unix format, seconds from begin of the “epoch”)
- **Z65\_CLIENT\_ADDRESS:** <external IP> (for Z39.50 use is external IP extern always 127.0.0.1 = localhost)
- **Z65\_START\_TIME:** Generation time of record (first update, Unix format)
- **Z65\_COUNT:** Update counter (every update increases by 1)

## Possible Sources for Login Information [1]: Table z34 (IP Statistics)

### Advantages

- Hint from Ex Libris
- Documented table

### Disadvantages

- Does not contain any login information
- Contains only the external(?) IP, not the whole key from z65
- No further connection to z65

### Comment:

- Useless for our purpose

## Possible Sources for Login Information [2]: PC Server Log

### Advantages

- Standard log file for pc servers
- Clear text readable

### Disadvantages

- Not documented as far as I know
- Suffers from performance improvements in version 16.02
  - One pc-server handles many connections
  - No connection writes atomic to the log file
  - Consequence: under heavy load different information pieces vary randomly within the log file
- Information is not complete at all

### Comment:

- Promising, but disappointing
- Attempts to unravel the file may result in serious headache
- But it is useful to analyse relevant parts of transaction log

## Possible Sources for Login Information [2]: PC Server Log (Example A)

```
2006-05-09 07:54:16 88 [000] [log] Message length 164
2006-05-09 07:54:16 88 [000] [log] version 16.02 format ALEPH compression 1 encryption 0
2006-05-09 07:54:16 88 [000] [log] data block length 152
IP Address : 212.093.155.050:1079
           : 192.168.002.120:1079                - INTERNAL
Client Attr:
  LOGIN      - user
  PROFILE    - default
  UUID       - 0c31f8f3-776f-4e5c-af17-bd116d676450
  APP-TYPE   - CAT
  USER-NAME  - UBW__1
2006-05-09 07:54:16 88 [001] [log] Message length 168
2006-05-09 07:54:16 89 [001] [log] version 16.02 format ALEPH compression
1 encryption 0
2006-05-09 07:54:16 89 [001] [log] data block length 156
  APP-TYPE   - CAT
  USER-NAME  - UBW__1
IP Address : 141.201.085.157:1150
Client Attr:
  LOGIN      - xxxxxxxxxxxx
  PROFILE    - default
  UUID       - 448947e9-640f-4bcb-af89-c31a9ec83464
  APP-TYPE   - CAT
  USER-NAME  - UBS__1

SERVICE    : C1005
MODULE      : OPAC Services
DESCRIPTION: Retrieve Brief List
ACTION      : SET
PROGRAM     : pc_sear_c1005

SERVICE    : C1005
MODULE      : OPAC Services
DESCRIPTION: Retrieve Brief List
ACTION      : SET
PROGRAM     : pc_sear_c1005

2006-05-09 07:54:16 90 [002] [log] Message length 4953
2006-05-09 07:54:16 90 [002] [log] version 16.02 format ALEPH compression 1 encryption 0
2006-05-09 07:54:16 90 [002] [log] data block length 4941
IP Address : 212.093.155.050:4910
           : 192.168.002.119:4910                - INTERNAL
Client Attr:
  LOGIN      - user
  PROFILE    - default
  UUID       - 2d2a5a10-bf5e-437b-9b9d-d43d47dd7daa
  APP-TYPE   - CAT
  USER-NAME  - UBW__1
```

## Possible Sources for Login Information [2]: PC Server Log (Example B)

```
2006-05-09 08:18:12 03 [000] [log] Message length 164
2006-05-09 08:18:12 03 [000] [log] version 16.02 format ALEPH compression 1 encryption 0
2006-05-09 08:18:12 03 [000] [log] data block length 152
2006-05-09 08:18:12 03 [002] [log] Message length 160
2006-05-09 08:18:12 03 [002] [log] version 16.02 format ALEPH compression 1 encryption 0
2006-05-09 08:18:12 03 [002] [log] data block length 148
IP Address : 143.205.070.104:1533
Client Attr: IP Address :
212.093.155.050 LOGIN - :luser234 1050

PROFILE - : default 192.168.002.003
: UUID - 1050 ddba5d87-541a-4d1a-ba60-8b778c43a1c7 - INTERNAL

APP-TYPE - Client Attr: ACQ

LOGIN - USER-NAME - user UBK__1

PROFILE - default
UUID - 01835ee9-1dcb-49d5-bdfe-f5fe415d4b84
APP-TYPE - CAT
USER-NAME - UBW__1

SERVICE : C0205
MODULE : Catalog Services
DESCRIPTION: Get ACC Text
ACTION :
PROGRAM : pc_cat_c0205

SERVICE : C1005
MODULE : OPAC Services
DESCRIPTION: Retrieve Brief List
ACTION : SET
PROGRAM : pc_sear_c1005

2006-05-09 08:18:12 05 [002] [log] Wrote 148 bytes
2006-05-09 08:18:12 09 [000] [log] Wrote 1003 bytes
```

## Possible Sources for Login Information [3]: Transaction Logs

### Advantages

- Lots of information
- Structured form

### Disadvantages

- Completely undocumented – normally used only for debugging needs
- Content may change at any time (seems to happen in version 18)
- Can not grow to more than 2 GB in size (32bit counter overflow?)
- Does also not have complete information (information which is in the pc server log is not in the transactions logs and the other side and vice versa)
- Many transactions do not contain the user, who caused the action
- Needs careful analysis using the pc server logs

### Comment:

- The best we found ...

## Information found in Transaction Logs

Selection of fields, which are important for measurement (extracted to a database table)

- PC server port
- Timestamp
- Key (same as in z65)
- Login name at the client pc (operating system level)
- Touched (used) library
- Login/role name in Aleph (often only system account "ALEPH")

## Further preparations

### Table z65\_history

- All changes to table z65 are collected into this table via a trigger

### Assignment table

- konk\_paraphe\_gruppe
  - Assigns login/role name to the belonging group or institution

### Additional assignments (guessed assignments)

- konk\_port\_lib\_gruppe
  - Assigns combination of port and touched library the belonging group or institution
- konk\_typ\_schluesel\_zuordnung
  - Assigns combination of record type and key to the belonging group or institution



## All Steps: The Big Picture [1]

z65\_history

- Creation of “blocked intervall” records (each entry in z65\_history raises from an action, which needs a license and blocks it for the timeout period; the process creates records of not intersecting time intervalls, where a license is used and can be assigned to a group or institution)

Transaction Logs

- Check of assignment table login names against login names in transaction table
- Assignment of institutions in transactions
  - insert assignment of resolvable login names into the transaction record
  - smoothing 1: completion of unassigned transactions in group (same port and key) using the login name of the user pc
- Check of assignment table port and lib against the combinations in transaction table

## All Steps: The Big Picture [2]

z65\_history

Transaction Logs

- First assignment try of blocked intervalls to groups or institutions
  - all transactions within the time intervall are looked up and grouped by port, login name, used library and assignment (within the transactions)
  - if there is one concrete assignment in the transaction group, this will be used
  - otherwise the combination of port and used library is tried
  - if this fails, the intervall is marked
    - ◊ SYSTEM
    - ◊ NOT\_ALLOWED
    - ◊ UNKNOWN (no transaction for this intervall)
- Smoothing 2: interpolation of missing (unknown) assignments from within the same group of key, type and timeout value

## All Steps: The Big Picture [3]

z65\_history

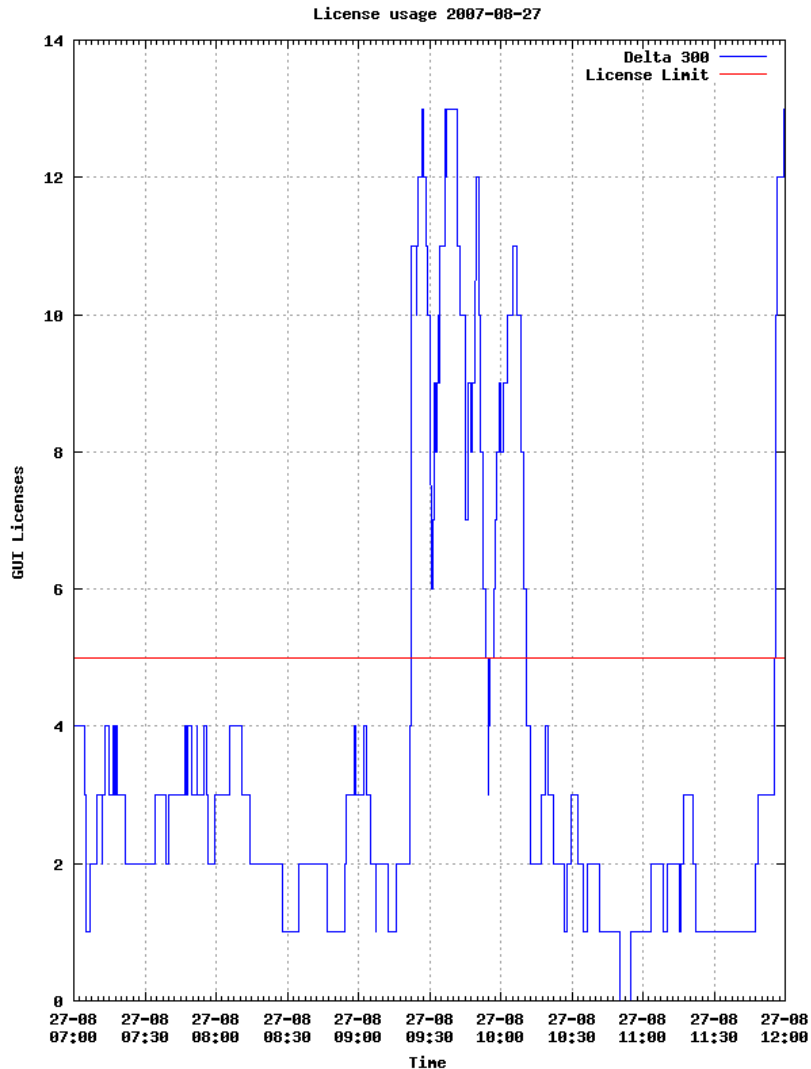
Transaction Logs

- Check of assignment table type and ip-key against the combinations in table of blocked intervalls without any assignment
  - Typically used for the replication process, which does not generate entries in the transaction logs
- Smoothing 3: lookup of missing assignment via the type and ip-key table
- Consolidation of blocked intervalls
  - Combines adjacent intervalls with same key and assignment to one record
  - Assigns marked intervalls more concise information
- Calculation of license usage

## Some Results [1]

- **Every** access to the pc server uses a license – even if the login is not valid
  - Denial of service attack possible
- The replication process uses a license
- Program was used with the finest possible resolution of 1 second
  - Evaluation period: almost a year
  - 166.316.566 records generated (two timeout periods – 300 and 600 seconds)
- Accuracy (Instance with 5 systems)
  - Comparison with the values of UTIL/Y/11
  - Differences only for under 3 thousandth of all measuring points
  - Differences with absolut value bigger than 1 only for under 2 tenthousandth of all measuring points
  - Bigger differences mostly during maintenance (short) times
- Assignment
  - Assignment are not always possible due to insufficient information in transaction logs
  - Noise rate is low at all
  - IP-range should be used within the smoothing procedures too
  - 100% assignment rate seems to be impossible in the current environment

## Some Results [2]



Extracted data for one institution (of five):

- A lot of license usage early in the morning
  - Reason?
- Somebody seems to switch on all computers within a short time between 9:00 and 9:30 am
- The institution consumes rather often much more than their license limit would allow
  - They are using the licenses of others in this times ...
- Some investigations about the permanent license usage in off business hours may lead to optimizations

## Development Status: Still Work in Progress

- Management tools for the consortium structure
  - Automated generation of assignment tables
- Analyze the changed structure of transaction logs in version 18
  - Adaptation of the loader program
- Extraction and Visualisation of individual data
- Use of round robin methods to limit the amount of stored data
- Integration of another smoothing step via use of IP-range assignments

## Wishes to Ex Libris

- Bear in mind that there are complex installations somewhere out there
- Better log files would help enormously
- But before you make it worse, let it as it is ...
- I hope for the new URM-systems and the new “Software as a service” paradigm

## Summary

- The developed method works for whole instances with different timeouts
- Its no realtime procedure
- Needs time and resources
- At the moment it is possible to separate the license usage of different institutions within one Aleph instance with only small error rate
- But there are limitations due to the restriction of available informations
- The procedure is fragile – all information its based on is undocumented and may change any time
- Work is not finished yet, but I hope that all of it will become obsolete in the new design



**The End**

**There are lots of details left  
But it is enough now**

**Thanks for your attention!**