FH JOANNEUM
University of Applied Sciences

Presentation of the Masters Thesis

„Scanning the Internet for Security“

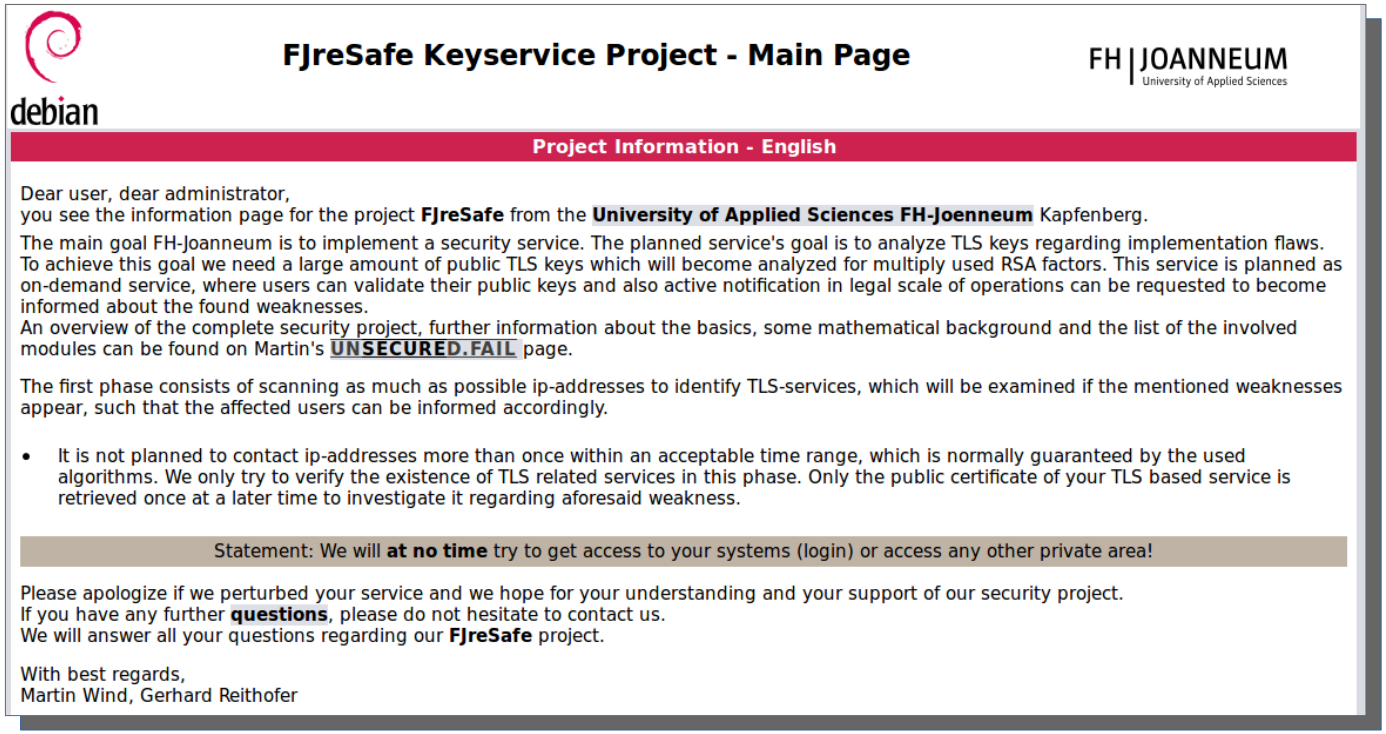
Gerhard Reithofer, FH-Joanneum Kapfenberg, 2017-01-16

Historical Background and Mathematical Basis

- Project work 2014 (FHJ Kapfenberg, ASE13)
  „Attacking RSA by Factoring Coprimes“
  - Algorithm from D. J. Bernstein,
    "Factoring into coprimes in essentially linear time",
    Journal of Algorithms 54 (2005)
  - Test and evaluation of the algorithm (Wind, Reithofer)
- Result: Idea for a security service at FH-Joanneum

Security Service: „FJreSafe“

FH
Joanneum
real
enhanced
Security
application
for
endusers



debian

FJreSafe Keyservice Project - Main Page

FH JOANNEUM University of Applied Sciences

Project Information - English

Dear user, dear administrator,
you see the information page for the project **FJreSafe** from the **University of Applied Sciences FH-Joenneum** Kapfenberg.

The main goal FH-Joanneum is to implement a security service. The planned service's goal is to analyze TLS keys regarding implementation flaws. To achieve this goal we need a large amount of public TLS keys which will become analyzed for multiply used RSA factors. This service is planned as on-demand service, where users can validate their public keys and also active notification in legal scale of operations can be requested to become informed about the found weaknesses.
An overview of the complete security project, further information about the basics, some mathematical background and the list of the involved modules can be found on Martin's **UNSECURED.FAIL** page.

The first phase consists of scanning as much as possible ip-addresses to identify TLS-services, which will be examined if the mentioned weaknesses appear, such that the affected users can be informed accordingly.

- It is not planned to contact ip-addresses more than once within an acceptable time range, which is normally guaranteed by the used algorithms. We only try to verify the existence of TLS related services in this phase. Only the public certificate of your TLS based service is retrieved once at a later time to investigate it regarding aforesaid weakness.

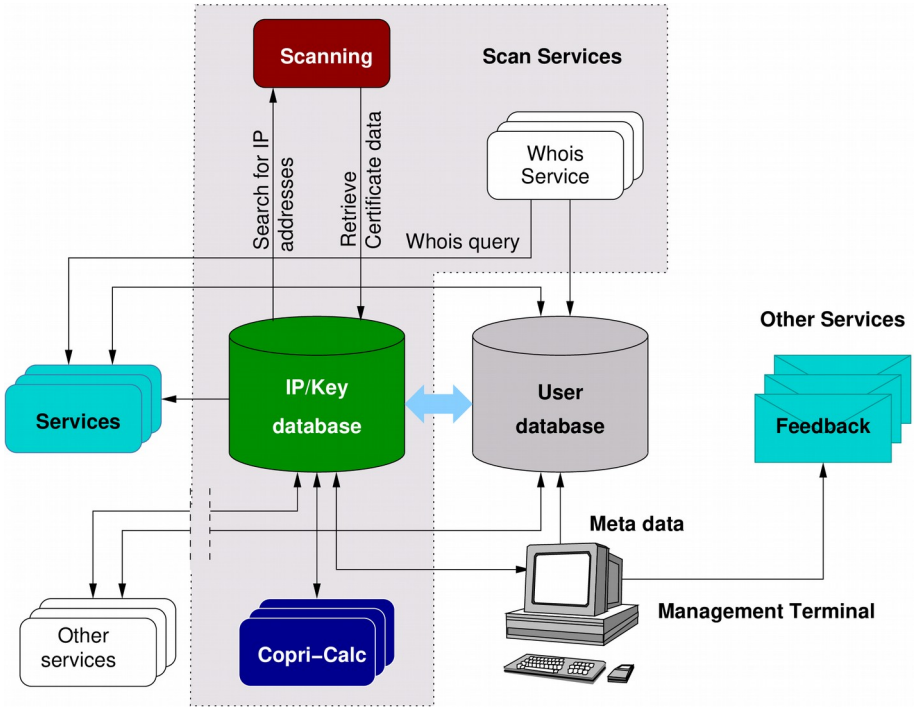
Statement: We will **at no time** try to get access to your systems (login) or access any other private area!

Please apologize if we perturbed your service and we hope for your understanding and your support of our security project.
If you have any further **questions**, please do not hesitate to contact us.
We will answer all your questions regarding our **FJreSafe** project.

With best regards,
Martin Wind, Gerhard Reithofer

FH JOANNEUM University of Applied Sciences

Security-Service as Research Project at FHJ

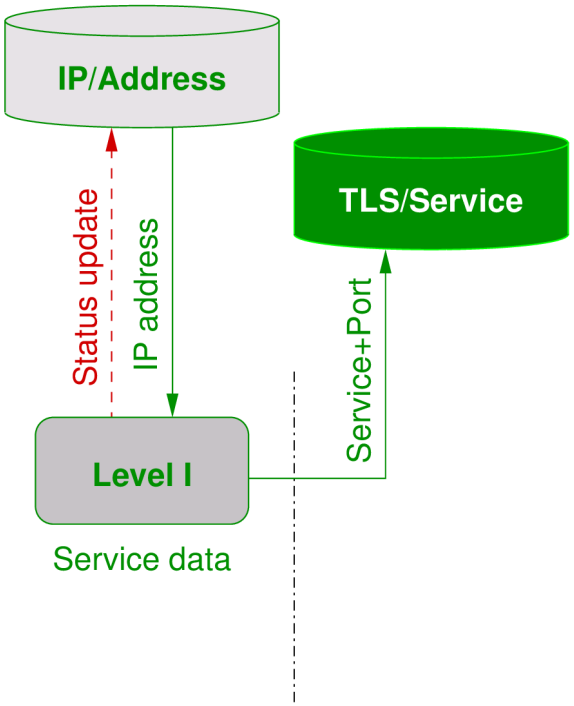


System Overview

- Scanning for RSA keys
- Managing (meta)data
- Calculating coprimes
- Informing affected users
- Providing a query service
- Providing additional support…

FH JOANNEUM University of Applied Sciences

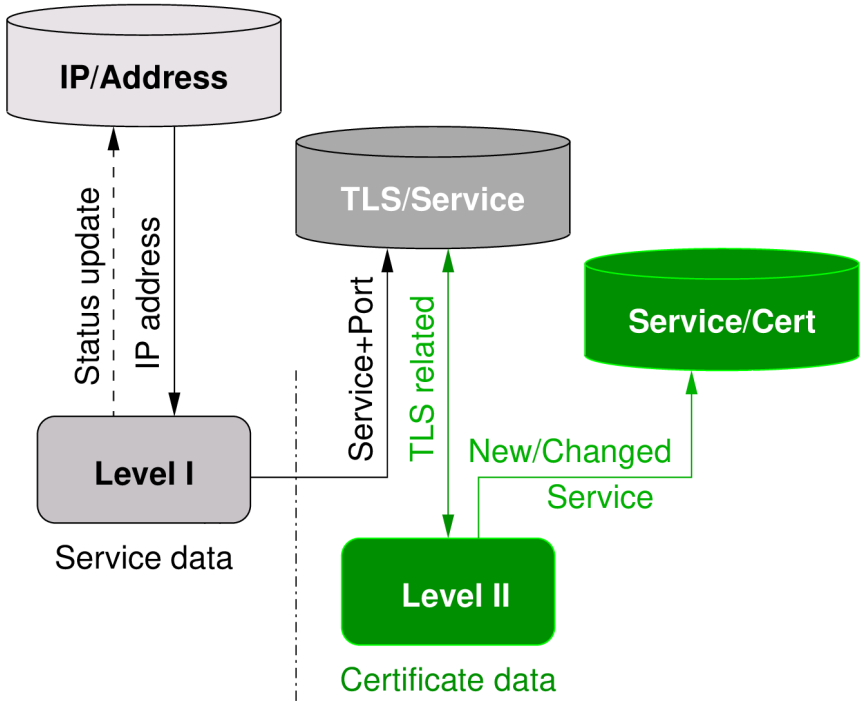
Scanning and collection of RSA public keys (MT's goal)



Scanning Level I Process

FH JOANNEUM University of Applied Sciences

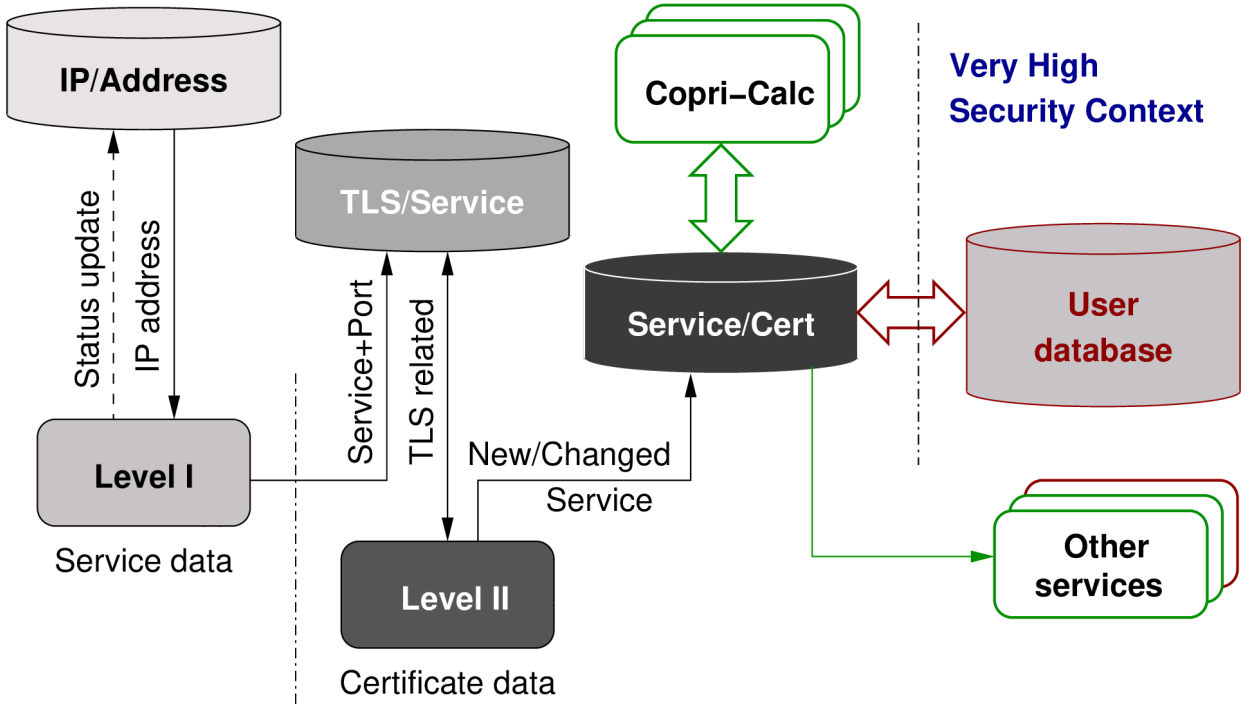
Scanning and collection of RSA public keys (MT's goal)



Scanning Level II Process

FH JOANNEUM University of Applied Sciences

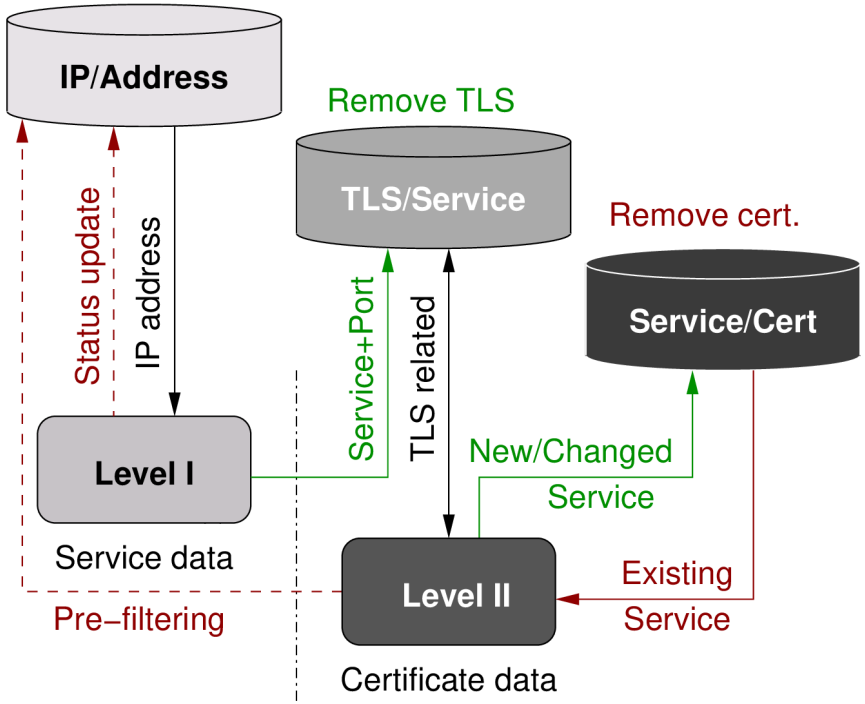
Scanning and collection of RSA public keys (MT's goal)



Service Interfaces

FH JOANNEUM University of Applied Sciences

Scanning and collection of RSA public keys (MT's goal)



Life-Cycle/Data Management/Data Transisions

Scan Process Classes/Life Cycle

Transition	Object Class	Input/Source
$A_1 \leftarrow A$	$A_1 \cdots filtered\ class\ A\ addr.$	$A \cdots all\ IPv\ 4\ addresses$
$S \leftarrow A_1 \times I_1$	$S \cdots probably\ TLS\ addr.$	$I_1 \cdots list\ of\ ports$
$S_1 \leftarrow S$	$S_1 \cdots filtered\ class\ S\ addr.$	
$C \leftarrow S_1 \pm I_2$	$C \cdots list\ of\ certificates$	$I_2 \cdots found/removed\ certs.$
$K \leftarrow C$	$K \cdots unique\ list\ of\ keys$	

Question in the context that similar databases exist (Scansio, Censys, EFF, ...):

Why rescanning the Internet?

- Service locality, trustworthy and data owning
- Timing constraints and data actuality
- Interest on specific data only (RSA keys)

Basic Assumptions and Estimations

Model Calculation where the processing of 1 IP needs 1 second:

$$\frac{2^{32}}{3600 \cdot 24} \approx 49710.2\, days \approx 136.1\, year$$

Estimation for processing the complete Internet (IPv4) in 2 weeks:

$$\frac{4294967296}{3600 \cdot 24 \cdot 14} \approx \frac{3500 \cdot IP}{sec}$$

Average data size for a TLS handshake is 2601.42 Bytes

$$2.6 \cdot \frac{Kb}{IP} \cdot 3500 \frac{IP}{sec} \cdot \frac{7622}{10000} \approx 7035 \frac{Kb}{sec} \approx 7.0 \frac{Mb}{sec}$$

Level I - Scanning Tools (services)

- nmap: universal command line scanner, many scan options, no list input
- zmap: development of the "Internet-Wide Scan Data Repository" group, no list input
- ipscan: GUI based Java application, multi threaded, much too slow
- masscan: speed optimized nmap compatible scanner, development state

Level II - Scanning Tools (certificates)

- Except for „openssl“ (slow) no single solution was found
- Solution: Self developed tool from a FH project 2014 extended to fulfill the new requirements:
  - Connect to a TLS service (if possible?)
  - Issue a TLS handshake (to verify TLS validity)
  - Retrieve certificate data (part of X.509 data)
  - Save retrieved meta data to the database

Used Technologies and Software (1)

- Multi platform scripting language Tcl/Tk
  - Process controlling framework (e. g. execution of binary tools)
  - Database interface to PostgreSQL for „managing“ the IP/Key data and key query tool
- Programming language ANSI C
  - Extending the existing TLS interface for Tcl for extracting additional X.509 data
  - IP-Shuffle tool for randomizing the IPV4 address table

Used Technologies and Software (2)

- HTML coding
  - Project information home page – central contact and communication point
- Programming language PHP for web tools
  - Web frontend and user interface for the certificate evaluation (query data for broken keys)
  - User interface for WHOIS service (integrated into the evaluation result information)

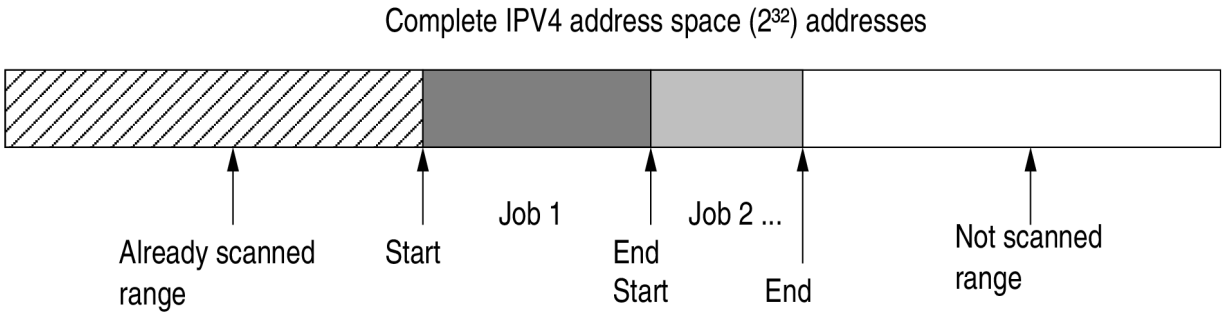
Final Solution - Requirements

1) Data collection:
   a) Level I scanning method
   b) Level II scanning method
2) Task improvement by using parallel processing
3) Process measurement and control methods
4) Prototype of the user interface for end users

Level I Scanning Method (1)

Randomized Linear IP-Table

A „pseudo database“ was created by using a flat file to define randomized order of IP addresses for scanning (prog: ipshuffle).



A „scan job“ is defined by its start offset and the job size. The „current job“ pointer is incremented by the blocksize.

Level I Scanning Method (2)

Sequential scan process - plugin

A control script executes a scan job:

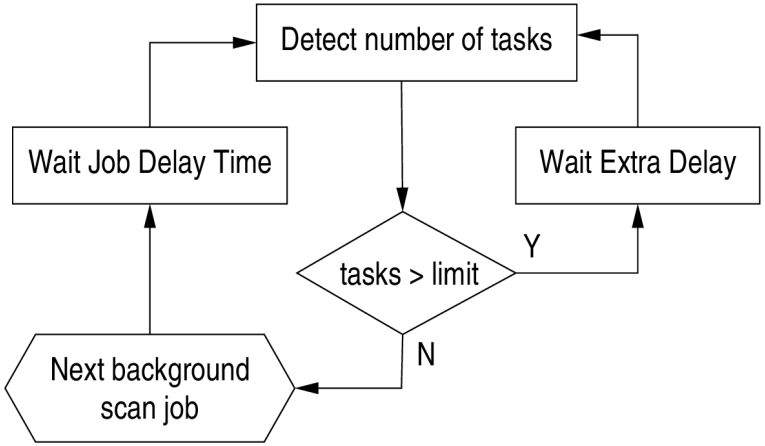
1. An IP range is read from the Randomized Linear IP-Table and a temporary file is created containing the addresses
2. The executable scan programm is started using the address table and a port list (defined by configuration)
3. The scan results are written to another temporary result file
4. The result file is parsed and the parsing results are written to the database (table „scan“)

Scanning Method (3)

Parallel scanning method

A process controller executes jobs parallel:

1) Detect the number of processes
2) If a maximum is reached?
   a) Wait an extra delay time
   b) Goto step 1)
3) Start a new „Sequential scan process“ in background
4) Wait a job delay time
5) Goto step 1)

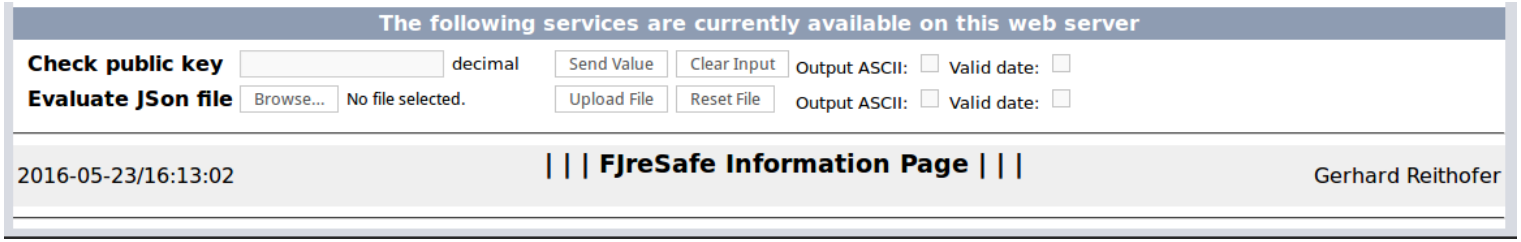


Data Post Processing

Creating the results

Data workflow (1)

- Public key database export to MPZ file is the Copri calculation input
- Calculation output is a JSON file containing the found broken keys
- This result file is used to request a Reverse Lookup to get the corresponding meta data (via web GUI or batch command)



Data Post Processing

Presenting the results

Data workflow (2)

The Reverse Lookup provides the corresponding meta data as web page or file download. Two hyperlinks allow a WHOIS service call or the download of the PEM certificate.

Found record: 232

Field	
id	8269145
ip_address	91.118.154.85
service	443
last_mod	2016-09-12 20:42:13.056744
status	ok
subject	CN=*.hybridserver.at,OU=Domain Control Validated - RapidSSL(R),OU=See www.rapidssl.com/resources/cps (c)14,OU=GT18085679
issuer	CN=RapidSSL SHA256 CA - G3,O=GeoTrust Inc.,C=US
chain	/OU=GT18085679/OU=See www.rapidssl.com/resources/cps (c)14/OU=Domain Control Validated - RapidSSL(R)/CN=*.hybridserver.at /C=US/O=GeoTrust Inc./CN=RapidSSL SHA256 CA - G3 /C=US/O=GeoTrust Inc./CN=RapidSSL SHA256 CA - G3 /C=US/O=GeoTrust Inc./CN=GeoTrust Global CA
rsa_n(hex)	a962be5a7c88d6dd7b7be622fbe41786e6bb2f0e3205ffa29a6ed5c9eb24e40be0c8a91d21d9eb392f0662e4d89905b0b57bfb6dfaefc9551624
rsa_n(dec)	21382965807535848790843472306668931154440197264519282881811976764362109695207958131966976237973740450737615229 77
rsa_e	65537
key_type	rsa
key_size	2048
not_valid_before	2015-07-07 02:14:47
not_valid_after	2018-06-04 09:26:21
certificate	Click for download

Back to previous page

... or goto to our Homepage.

System Monitoring (1)

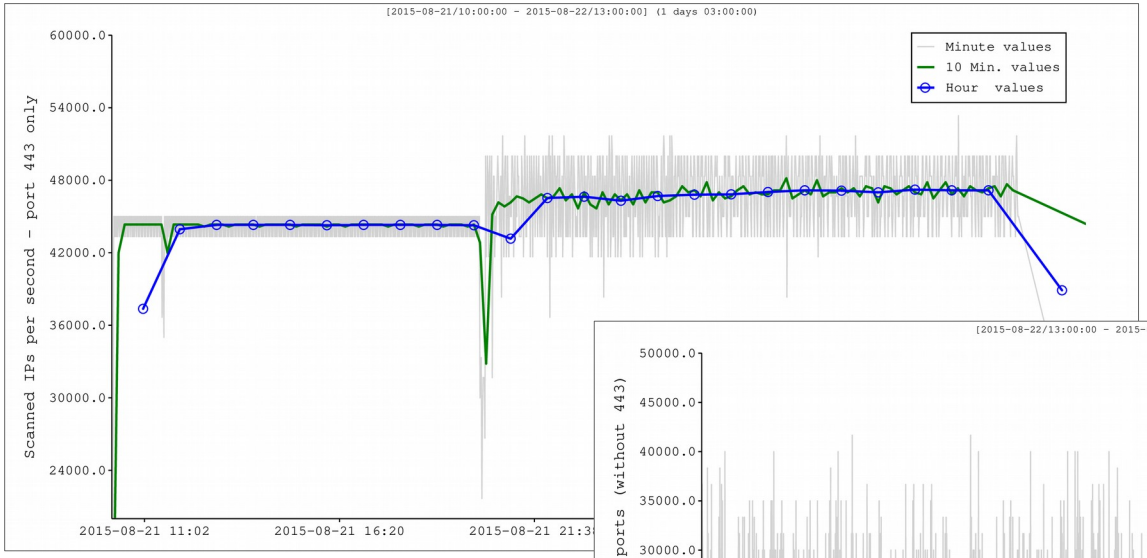
Used methods

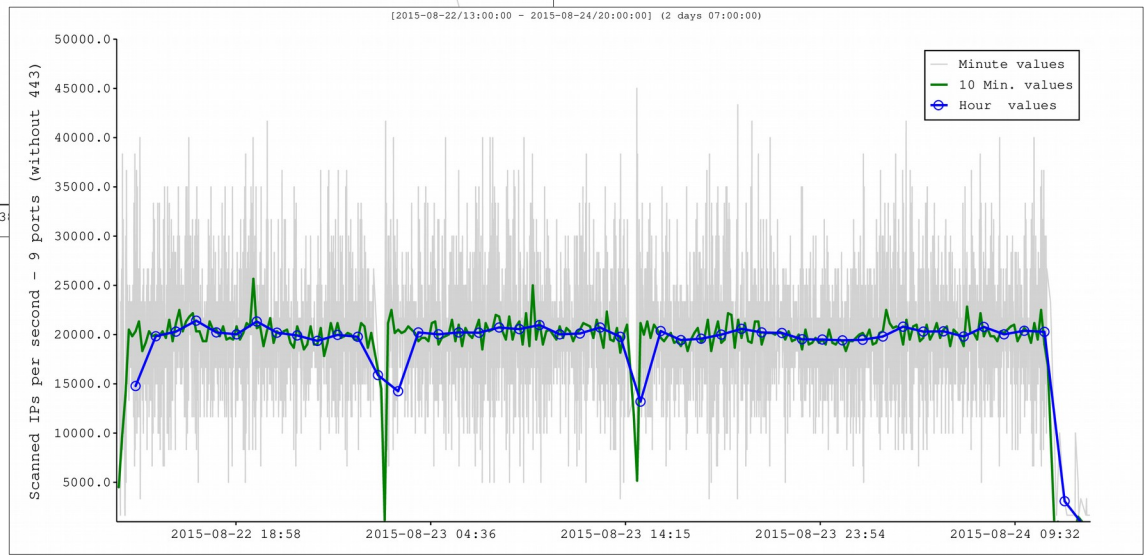
- Temporary result files
- Creation and evaluation of log files
- Existing and new monitoring tools
- Database queries

FH JOANNEUM University of Applied Sciences

System Monitoring (2)

Level I - single port scan vs. multi port scan





System Monitoring (3)

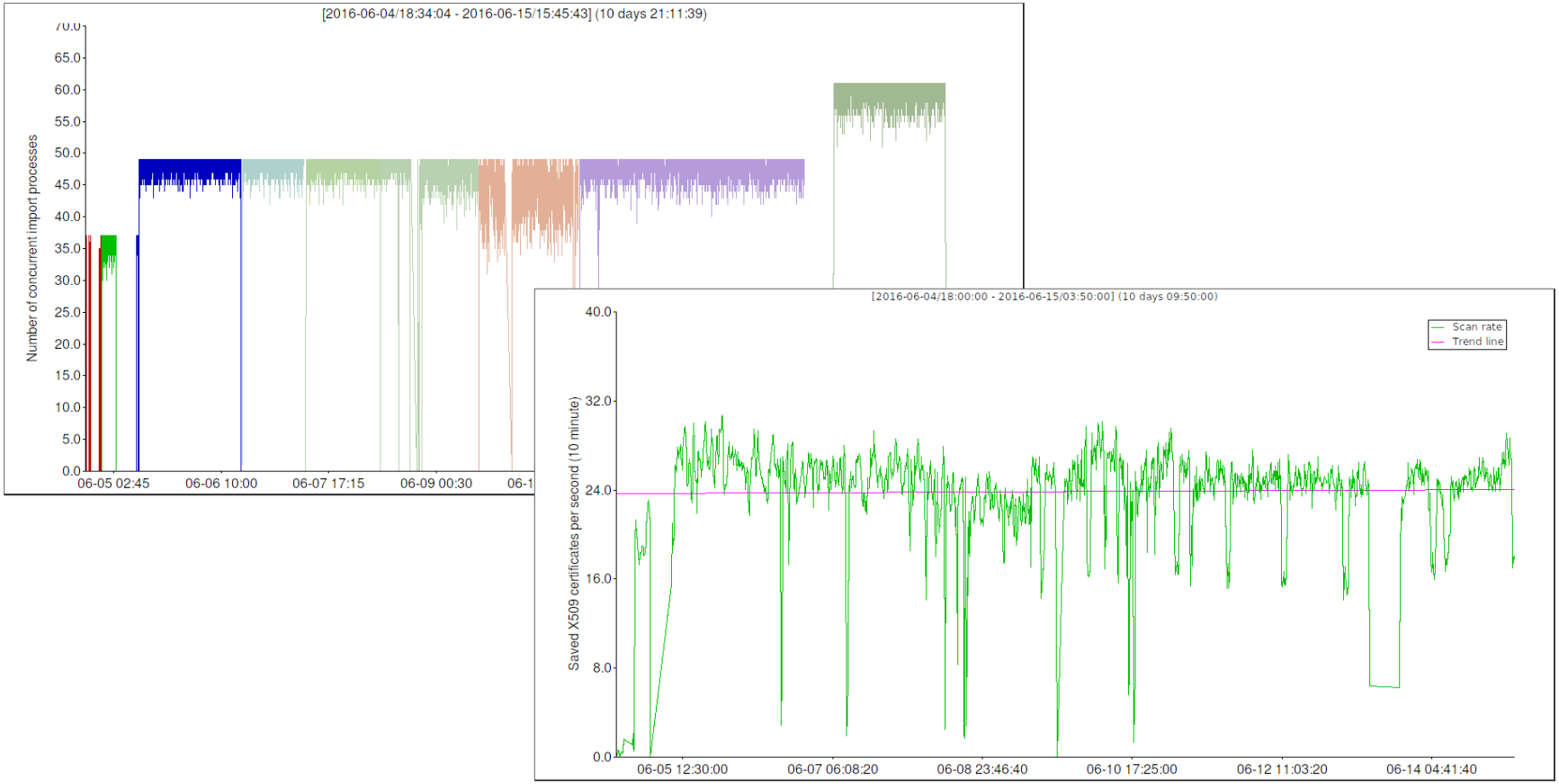
Scan Level I performance

Value	#1 p443 only	#2 9p (no 443)	#2/#1	Remark
IPs total	4,059,017,295	3,270,700,000	0.806	Scanned IPs
TLS rel.	3,575,192	22,936,589	6.415	Open TLS ports
Scan time	1d 02:32:56	2d 00:15:27	1.811	95576/173727 s
IPs/sec	42,469.0	18,826.7	0.443	Class A addr.
TLS/sec	37.4	132.0	3.529	Class S addr.
TLS/1M	880.0	7,010.0	7.966	1M=million IPs.
TLS/1M/port	880.0	778.8	0.885	Spec. rate

FH JOANNEUM University of Applied Sciences

System Monitoring (4)

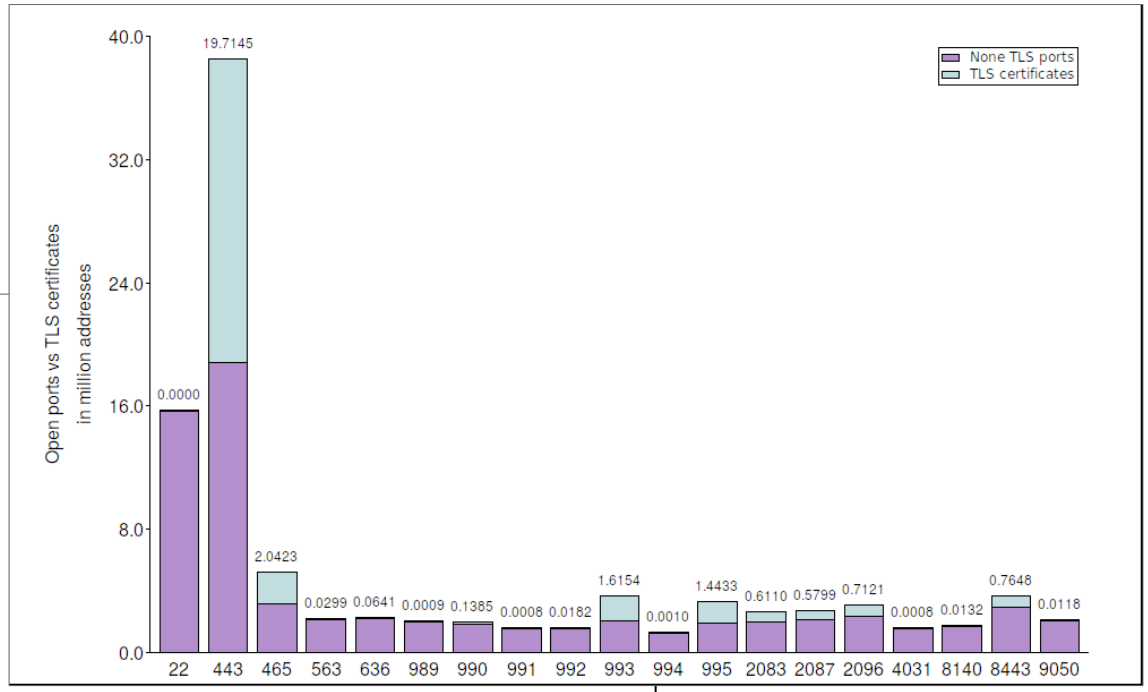
Level II – parallel tasks vs. scan rate

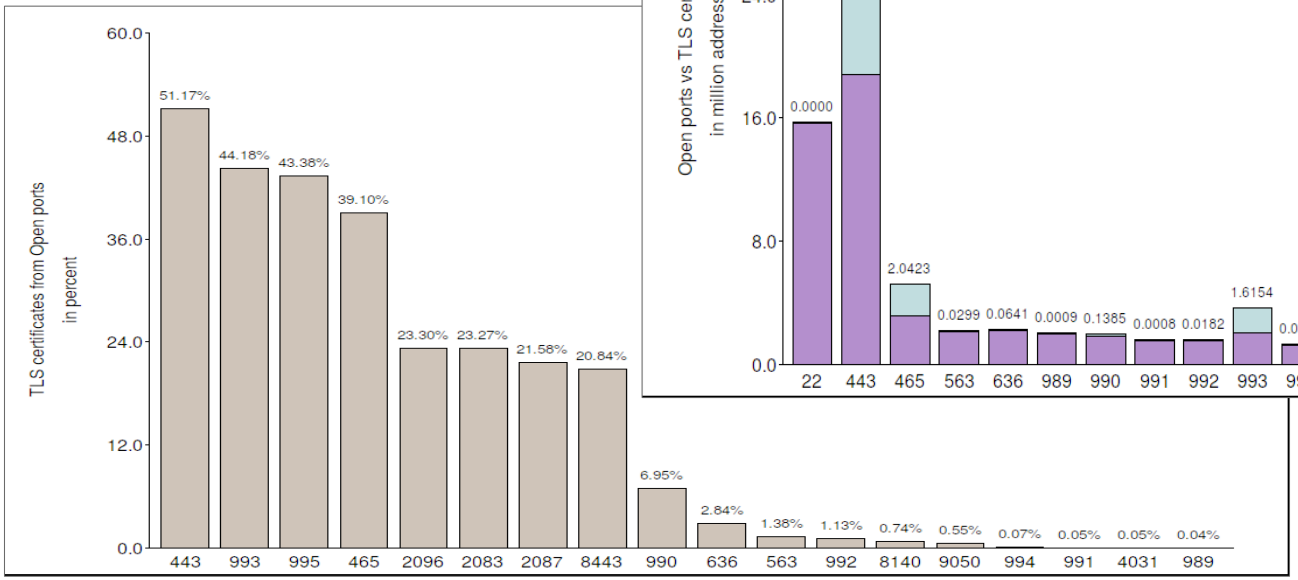


FH JOANNEUM University of Applied Sciences

Scan Results (1)

Service port distribution

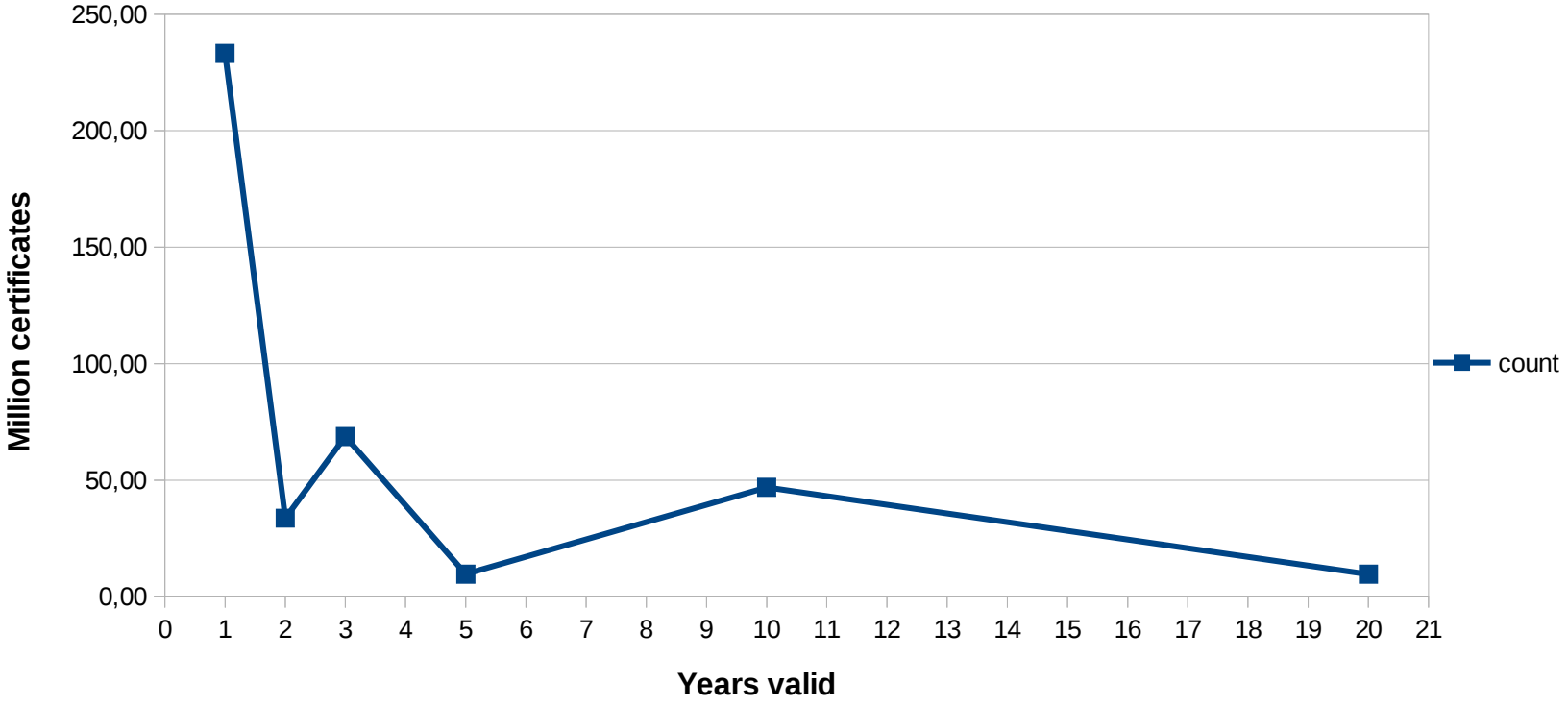




FH JOANNEUM University of Applied Sciences

Scan Results (2)

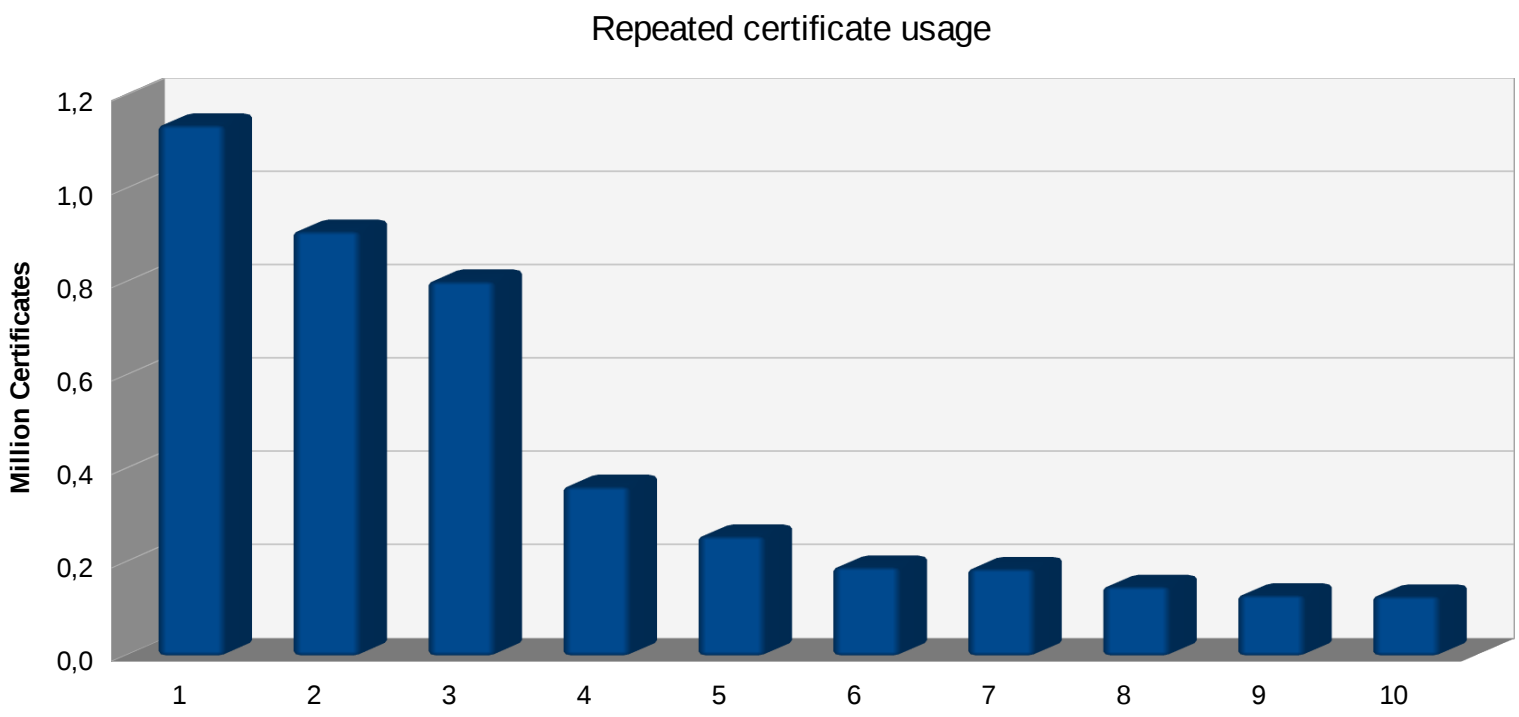
Certificate lifetime statistcs



FH JOANNEUM University of Applied Sciences

Scan Results (3)

20% of all certificates were used multiple times



Scan Results (4)

Invalid or ridiculous validation periods

- Very long running certificate (7985 years).

  Valid from Jan 4 23:44:33 2015 until Dec 31 23:59:59 9999.

- One second time validation period. Not Before date is identical to Not After date.

  Valid from Jan 27 18:48:24 2011 and valid until is also Jan 27 18:48:24 2011.

- Minus one day validation period. Not Before date 1 day later than Not After.

  Valid from May 25 10:09:43 2016 until May 24 10:09:43 2016.

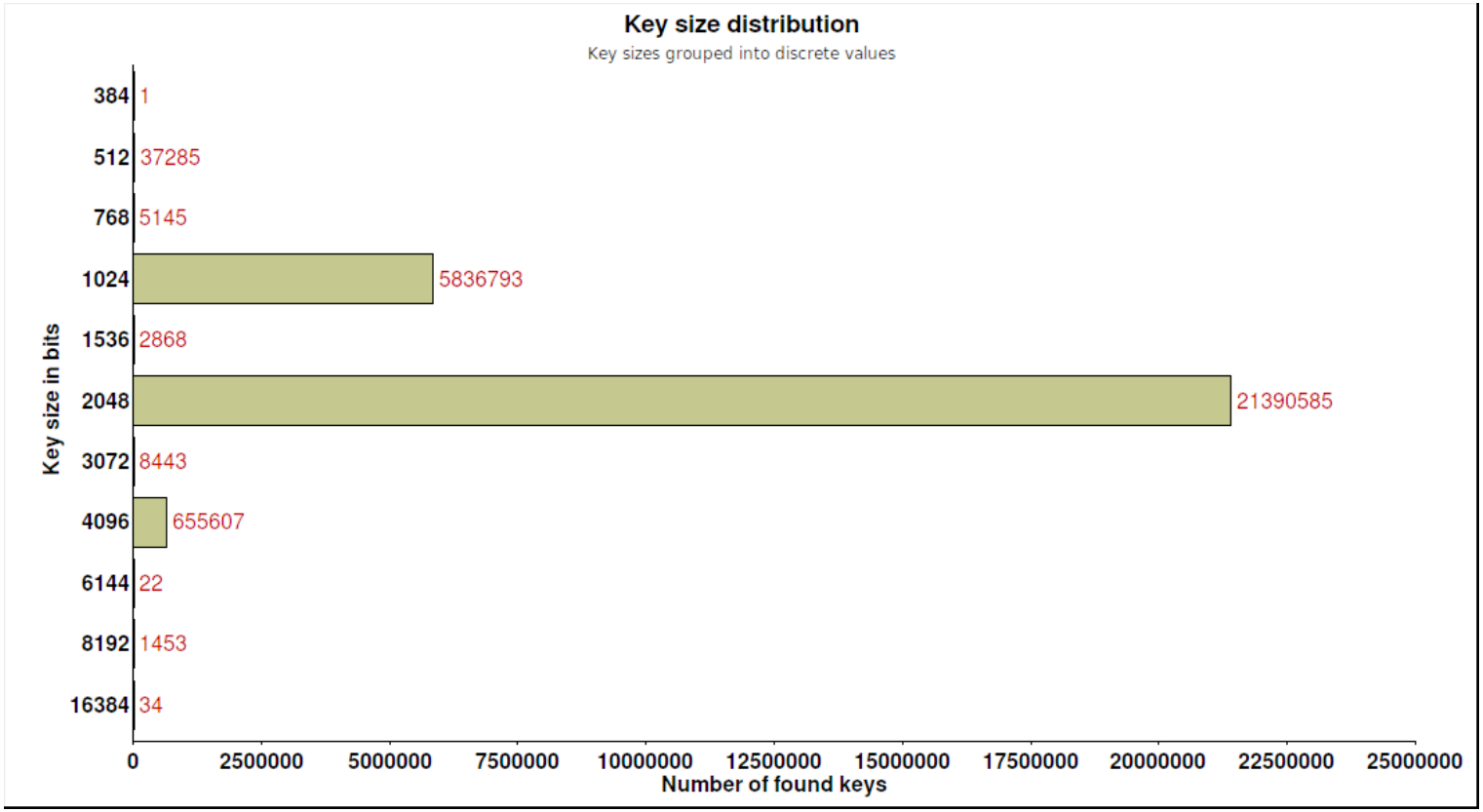
- Invalid validation period of -26,6 years.

  Valid from Sep 23 16:49:29 2007 until Feb 21 10:21:13 1981.

FH JOANNEUM University of Applied Sciences

Scan Results (5)

Key size statistics



Final Discussion

„The optimization step will never end!“

- Port scanner masscan is in some parts in „development state“ (e. g. certificate extract)
- Parallel processing optimization (bottlenecks „disc IO“ and concurrent database access)
- Spare use or partitioning of resources (multi server)
- Optimize software, the database use and data model
- Challenge the „completeness“ pretension and prepare a draft for IPV6

FH JOANNEUM
University of Applied Sciences

„Thank you for your attention!“

Are there any additional questions?