PRISM and an Agenda for European Network Security Research

Another Turn of the Wheel: Mainframe, Desktop, Cloud, Peer

Christian Grothoff

Technische Universität München

02.07.2013

"Never doubt your ability to change the world." –Glenn Greenwald

Everybody Has Secrets

- Business & Trade Secrets
- Political opinions
- Illegal activities

Keeping Secrets

- Encryption: baseline
- Hide meta-data: state of the art
- Practice today?

Keeping Secrets

- Encryption: baseline
- Hide meta-data: state of the art
- Practice today?

Send everything to US in plaintext

- Guardian: “The PRISM program allows the intelligence services direct access to the companies servers.”
- Cooperating providers: Microsoft, Yahoo, Google, Facebook, PalTalk, YouTube, Skype, AOL, Apple

- Guardian: "The PRISM program allows the intelligence services direct access to the companies servers."
- Cooperating providers: Microsoft, Yahoo, Google, Facebook, PalTalk, YouTube, Skype, AOL, Apple
- PRISM enables real-time surveillance and access to stored content
- Data collected: E-mails, instant messages, videos, photos, stored data (likely files), voice chats, file transfers, video conferences, log-in times, and social network profiles
- Tiny part of NSA: \$20 M budget

US discussion focuses on spying on US citizens and legality under US law.

Frank Church (D-Idaho): "The NSA's capability at any time could be turned around on the American people, and **no American would have any privacy left**, such is the capability to monitor everything: telephone conversations, telegrams, it doesn't matter."

- NSA's tool to track global surveillance data
- 2,392,343,446 records from the US
- 97,111,199,358 records worldwide
- This is for March 2013 alone

- NSA's tool to track global surveillance data
- 2,392,343,446 records from the US
- 97,111,199,358 records worldwide
- This is for March 2013 alone
- Germany most surveilled country in Europe

- NSA's tool to track global surveillance data
- 2,392,343,446 records from the US
- 97,111,199,358 records worldwide
- This is for March 2013 alone
- Germany most surveilled country in Europe
- "leverages FOSS technology"

Other Programs

- "The SIGAD Used Most in NSA Reporting" $\Rightarrow$ there are more SIGINT tools
- Presentations list FARVIEW and BLARNEY
- Monitor fiber cables and infrastructure (IXPs?)
- "NSA collecting phone records of millions of Verizon customers daily" –Guardian

Other Programs

- "The SIGAD Used Most in NSA Reporting" $\Rightarrow$ there are more SIGINT tools
- Presentations list FARVIEW and BLARNEY
- Monitor fiber cables and infrastructure (IXPs?)
- "NSA collecting phone records of millions of Verizon customers daily" –Guardian

We do not know all about PRISM. Repr. Sanches (D-Calif.), after learning more during a briefing, said there is ... "significantly more than what is out in the media today (...) I believe **it's the tip of the iceberg**."

The Utah Data Center at Bluffdale

NSA's lastest expansion (2013):

- 1-1.5 million square feet
- \$2 billion building, \$2 billion hardware
- 65 MW power consumption
  SuperMuc: < 3 MW, 155,656 cores, ≈ 3 Peta FLOPS

$\Rightarrow$ Likely able to store and process all communication

Cyberwar

Presidential Policy Directive 20, issued October 2012 and released by Edward Snowden, outlines U.S. cyberwar policy:

"Offensive Cyber Effect Operations (OCEO) can offer unique and unconventional capabilities to **advance U.S. national objectives** around the world with little or no warning to the adversary or target and with potential effects ranging from **subtle** to severely damaging. (...)
The United States Government shall identify potential targets of national importance where OCEO can offer a favorable **balance of effectiveness and risk** as compared with other instruments of national power, establish and maintain OCEO capabilities integrated as appropriate with other U.S. offensive capabilities, and execute those capabilities in a manner consistent with the provisions of this directive."

Technical Cooperation

Bloomberg reports:

- US companies provide internal information to US secret services
- Companies from software, banking, communications hardware providers, network security firms
- Including technical specifications and **unpatched software vulnerabilities**
- In return, these **US companies** are given **access to intelligence information**
- Partners include: Microsoft, Intel, McAfee

History: ECHELON

- SIGINT collection network of AU, CA, NZ, UK and US
- Baltimore Sun reported in 1995 that Airbus lost a $6 billion contract in 1994 after NSA reported that Airbus officials had been bribing officials to secure the contract.
- Used to facilitate Kenetech Windpower's espionage against Enercon in 1994-1996.

Former US listening station at Teufelsberg, Berlin.

Does it matter?

MPI estimated losses due to industrial espionage damage in 1988 at DM 8 billion.

So how does the EU react to learning about PRISM?

Does it matter?

MPI estimated losses due to industrial espionage damage in 1988 at DM 8 billion.

So how does the EU react to learning about PRISM?

"Direct access of US law enforcement to the data of EU citizens on servers of US companies should be excluded unless in **clearly defined, exceptional and judicially reviewable** situations."
–Viviane Reding, EC vice-president in response to PRISM

History: Irak War

Katharine Gun leaked memo from NSA agent Frank Koza in 2003 about an American effort to monitor the communications of six delegations to the United Nations who were undecided on authorizing the Iraq War and who were being fiercely courted by both sides:

"As you've likely heard by now, the Agency is mounting a surge particularly **directed at the UN Security Council (UNSC)** members (minus US and GBR of course) for insights as to how to membership is reacting to the on-going debate RE: Iraq, plans to vote on any related resolutions, what related policies/negotiating positions they may be considering, alliances/dependencies, etc — the whole gamut of information that could give US policymakers an edge in **obtaining results favorable to US goals** or to head off surprises. In RT, that means a QRC surge effort to revive/create efforts **against** UNSC members Angola, Cameroon, Chile, Bulgaria and Guinea, as well as extra focus on Pakistan UN matters."

Not Just Monitoring

- US **controls** key Internet infrastructure:
  - Number resources (IANA)
  - Domain Name System (Root zone)
  - DNSSEC root certificate
  - X.509 CAs (HTTPS certificates)
  - Major browser vendors (CA root stores!)
- Encryption does not help if PKI is compromised!

Political Solutions?

Ron Wyden (US Senate intelligence committe) asked James Clapper, director of national intelligence in March 2013:

"**Does the NSA collect any type of data at all on millions or hundreds of millions of Americans?**"

Clapper replied:

"**No, sir.**".

The Enemy Within

"In February, the UK based research publication Statewatch reported that the **EU had secretely agreed** to set up an international telephone tapping network via a secret network of committees established under the "third pillar" of the Mastricht Treacty covering co-operation on law and order. (...) EU countries (...) should agree on **international interception standards (...) to co-operate closely with the FBI** (...). Network and service providers in the EU will be obliged to install **tappable** systems and to place under **surveillance** any person or group when served an interception order. These plans have never been referred to any European government for scrutiny (...) despite the **clear civil liberties issues** raised by such an **unaccountable** system. (...) The German government estimates that the mobile phone part of the package alone will cost 4 billion D-marks."

Scientific and Technological Options Assessment (STOA), "An Appraisal of Technologies of Political Control", European Parliament, PE 166499, 6 January 1998.

Technical Solutions

Can we develop technologies to solve problems created by technology?

Technical Solutions

Can we develop technologies to solve problems created by technology?

- Hack back?

Technical Solutions

Can we develop technologies to solve problems created by technology?

- Hack back?
- Monitor them?

Technical Solutions

Can we develop technologies to solve problems created by technology?

- Hack back?
- Monitor them?
- Move data to European cloud?

Technical Solutions

Can we develop technologies to solve problems created by technology?

- Hack back?
- Monitor them?
- Move data to European cloud?
- Decentralize data and trust!

Decentralize Everything

- Encrypt everything end-to-end
- Decentralized PKI
- Decentralized data storage
- No servers
- No authorities

Decentralize Everything

- Encrypt everything end-to-end
- Decentralized PKI
- Decentralized data storage
- No servers
- No authorities

$\Rightarrow$ No juicy targets for APTs

Decentralized vs. Centralized

Decentralized:	**Centralized:**
Slower	
No economics of scale	
More complex to use	
More complex to develop	
Hard to secure	
Hard to evolve	

Decentralized vs. Centralized

Decentralized:	**Centralized:**
Slower	Compromised
No economics of scale	
More complex to use	
More complex to develop	
Hard to secure	
Hard to evolve	

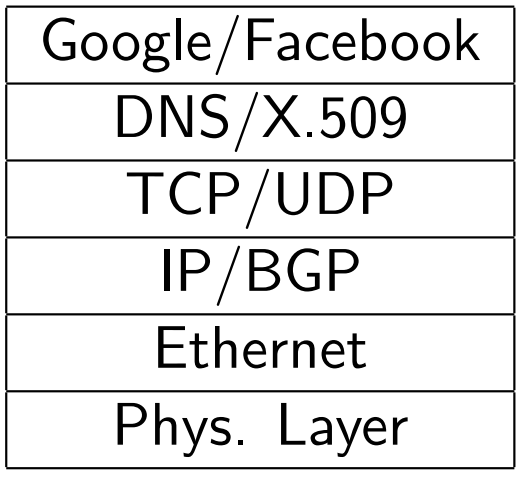
My Research and Development Agenda

Make decentralized systems:

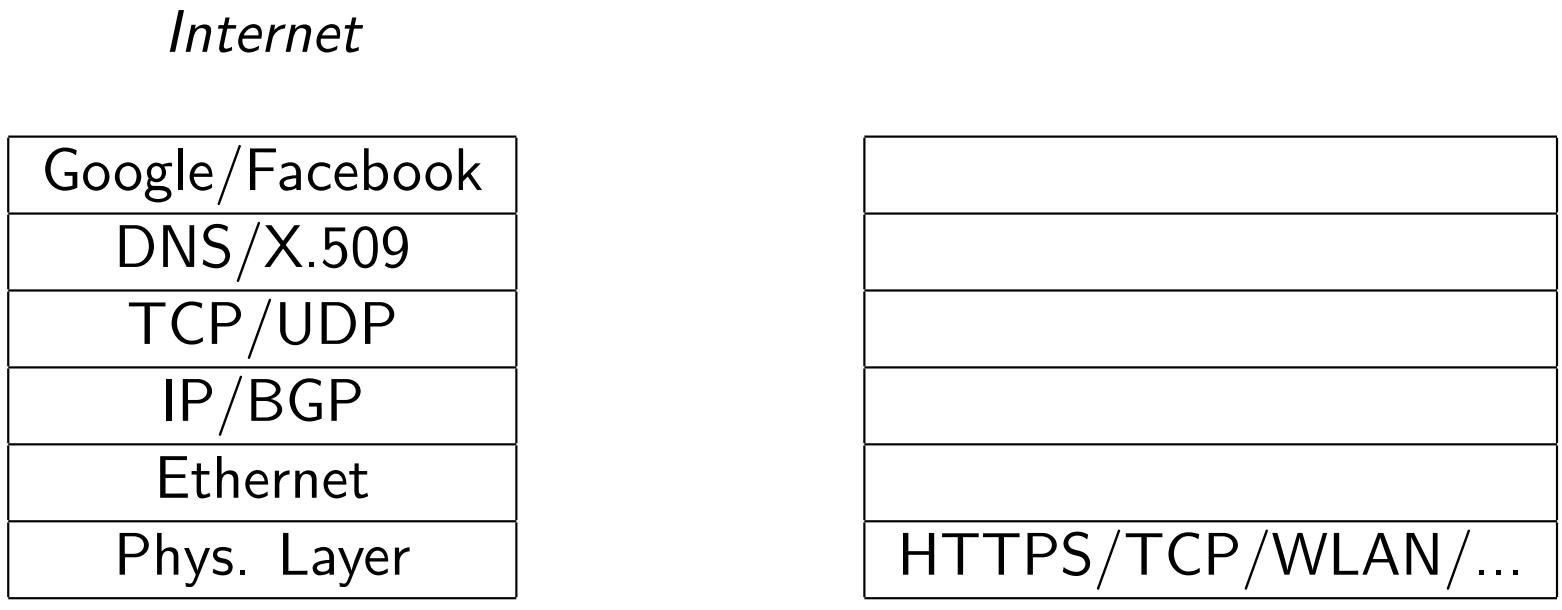
- Faster, more scalable
- Easier to develop, deploy and use
- Easier to evolve and extend
- Secure (privacy-preserving, censorship-resistant, available, ...)

Our Vision

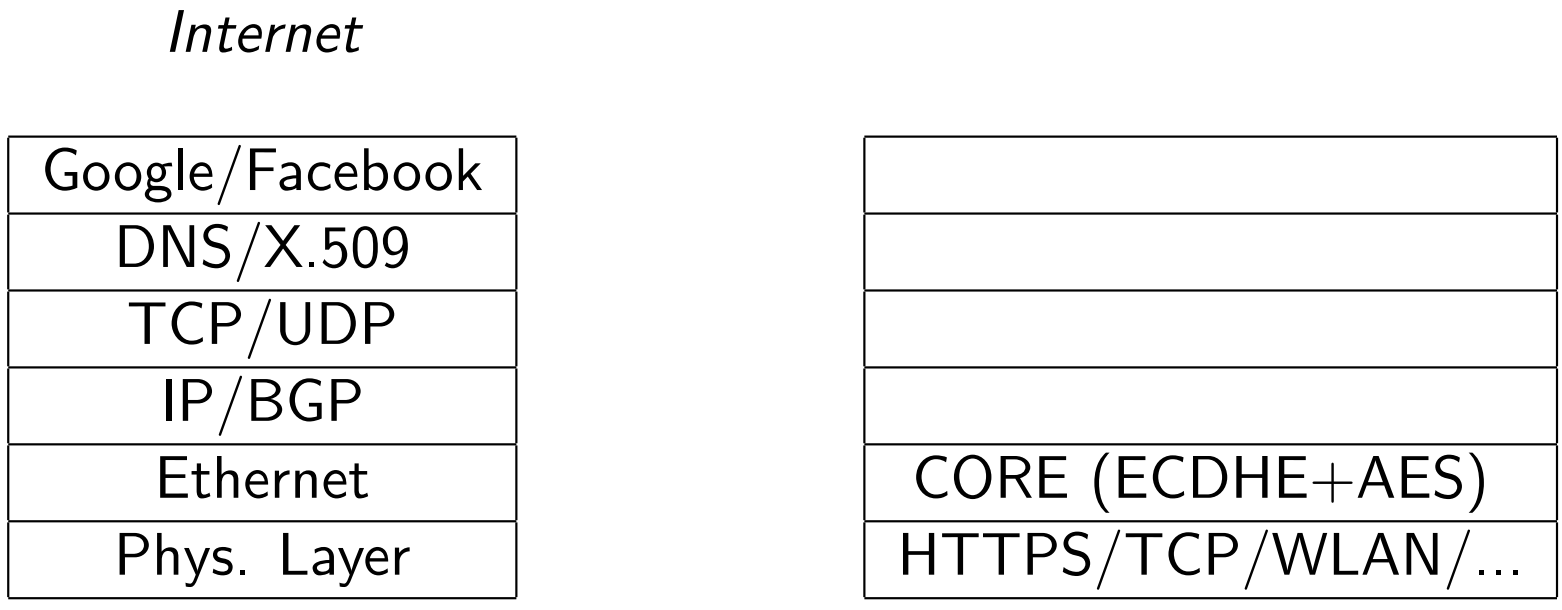
Internet



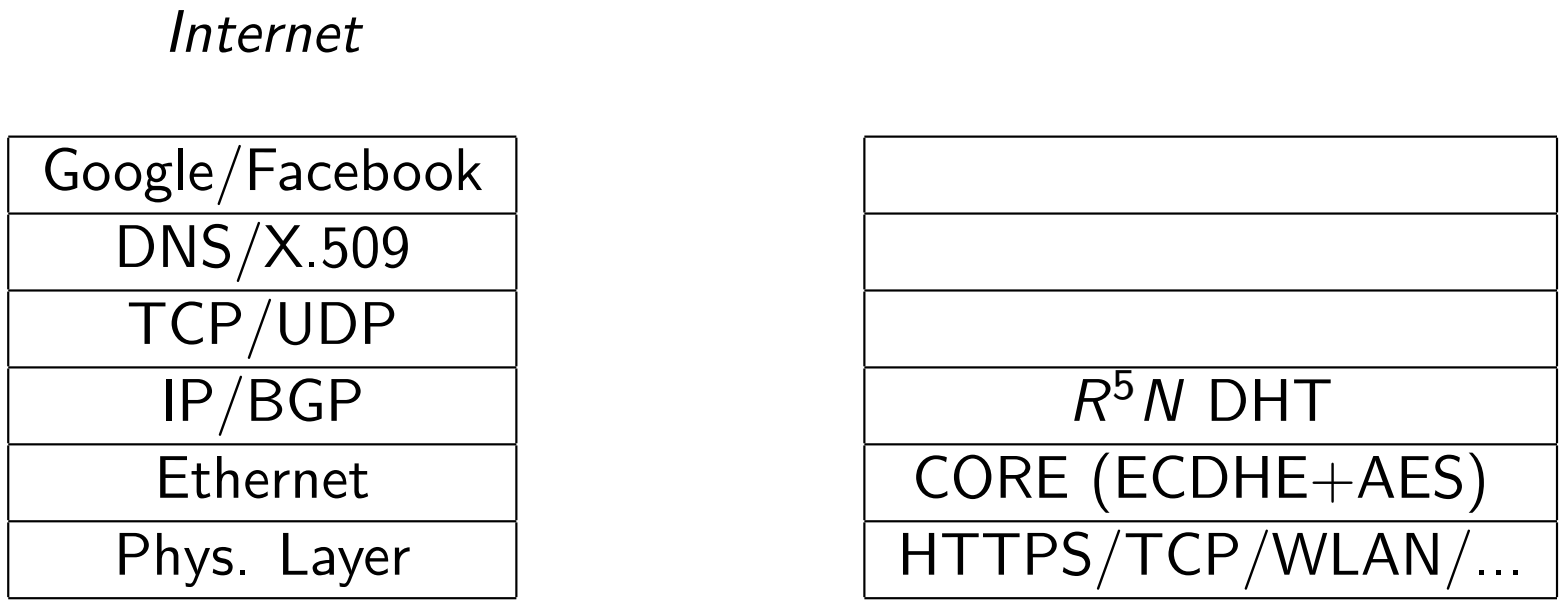
Our Vision


Internet
Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer
HTTPS/TCP/WLAN/...

Our Vision


Internet
Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

Our Vision


Internet
Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer
R^5N DHT
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

Our Vision

Internet

Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer

Mesh (ECDHE+AES)
R^5N DHT
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

Our Vision

Internet

Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer

GADS
Mesh (ECDHE+AES)
R^5N DHT
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

Our Vision

Internet

Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer

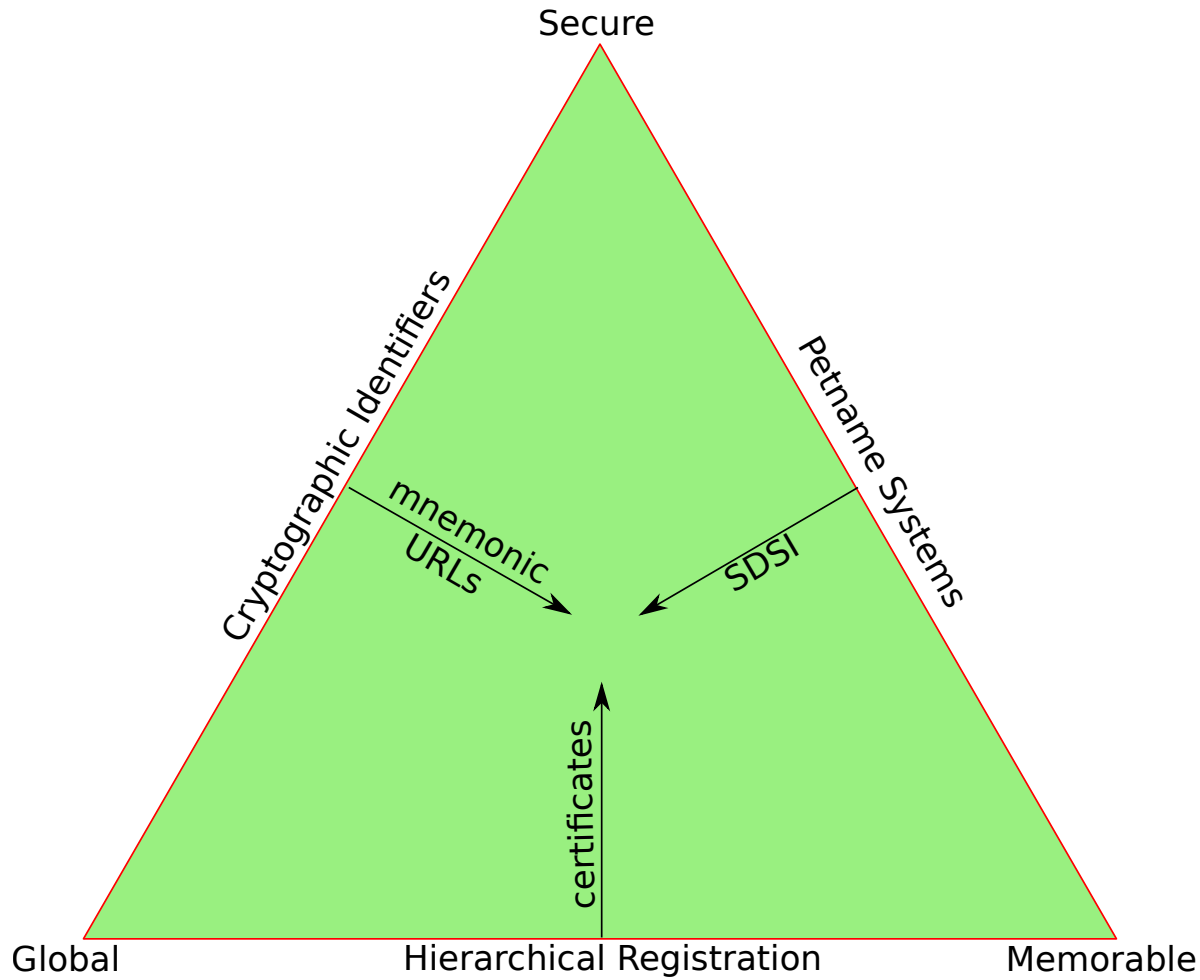
RegEx/PSYC
GADS
Mesh (ECDHE+AES)
R^5N DHT
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

Our Vision

Internet
Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer

GNUnet
RegEx/PSYC
GADS
Mesh (ECDHE+AES)
R^5N DHT
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

Decentralized Naming Systems[1]



Zooko's Triangle

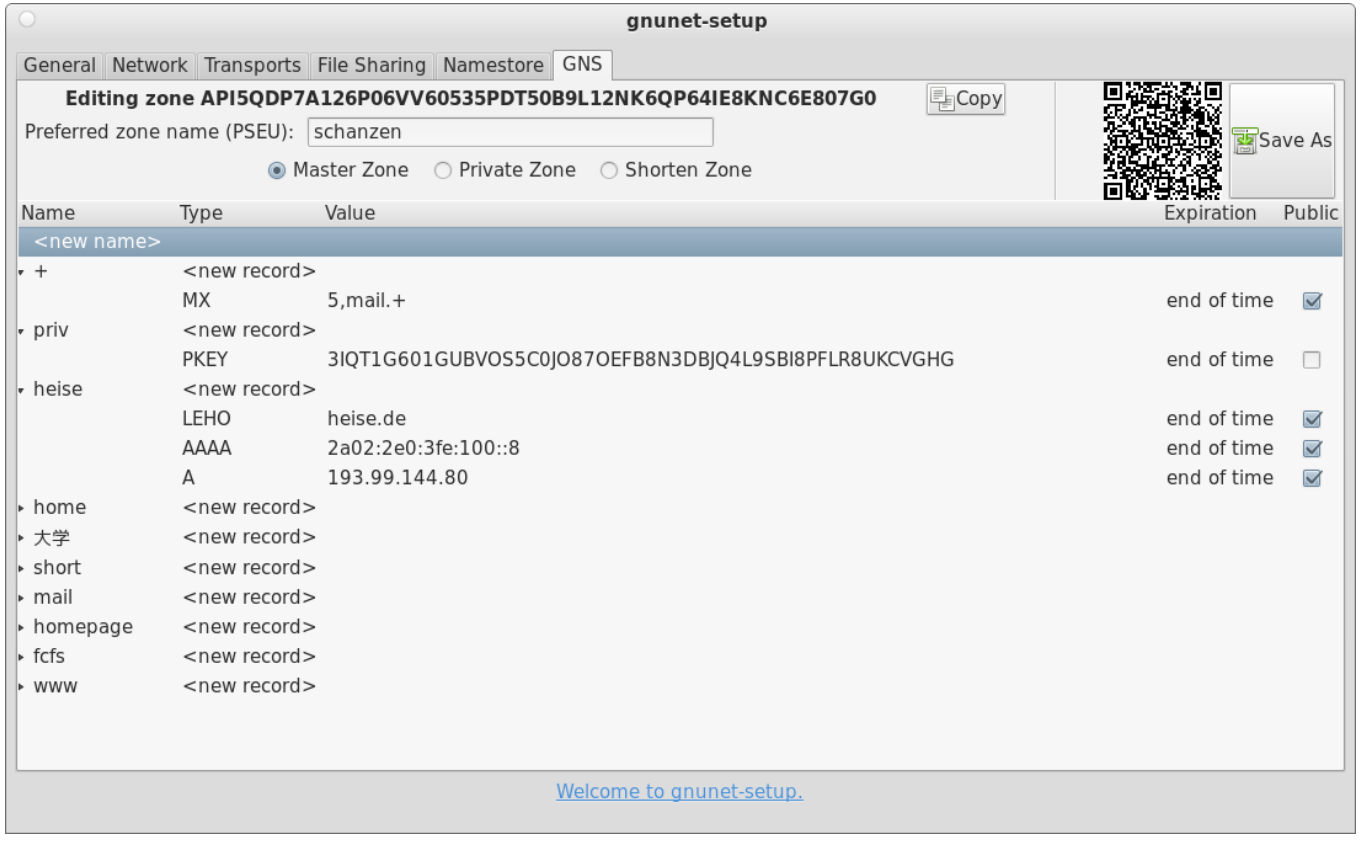
[1] Joint work with Martin Schanzenbach and Matthias Wachs

The GNU Alternative Domain System (GADS)

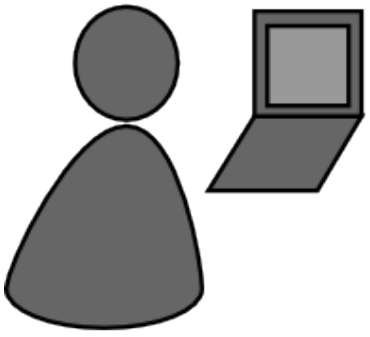
Decentralized PKI that can also replace DNS/DNSSEC:

- Signed Resource Records (RRs)
- Secure delegation provides **transitivity** (SDSI)
- Decentralized resolution (R^5N DHT)
- Every user manages his own zone

Zone Management: like in DNS

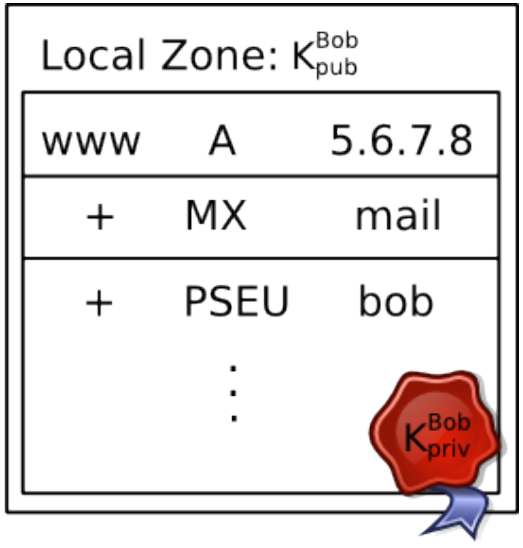


Name resolution in GADS



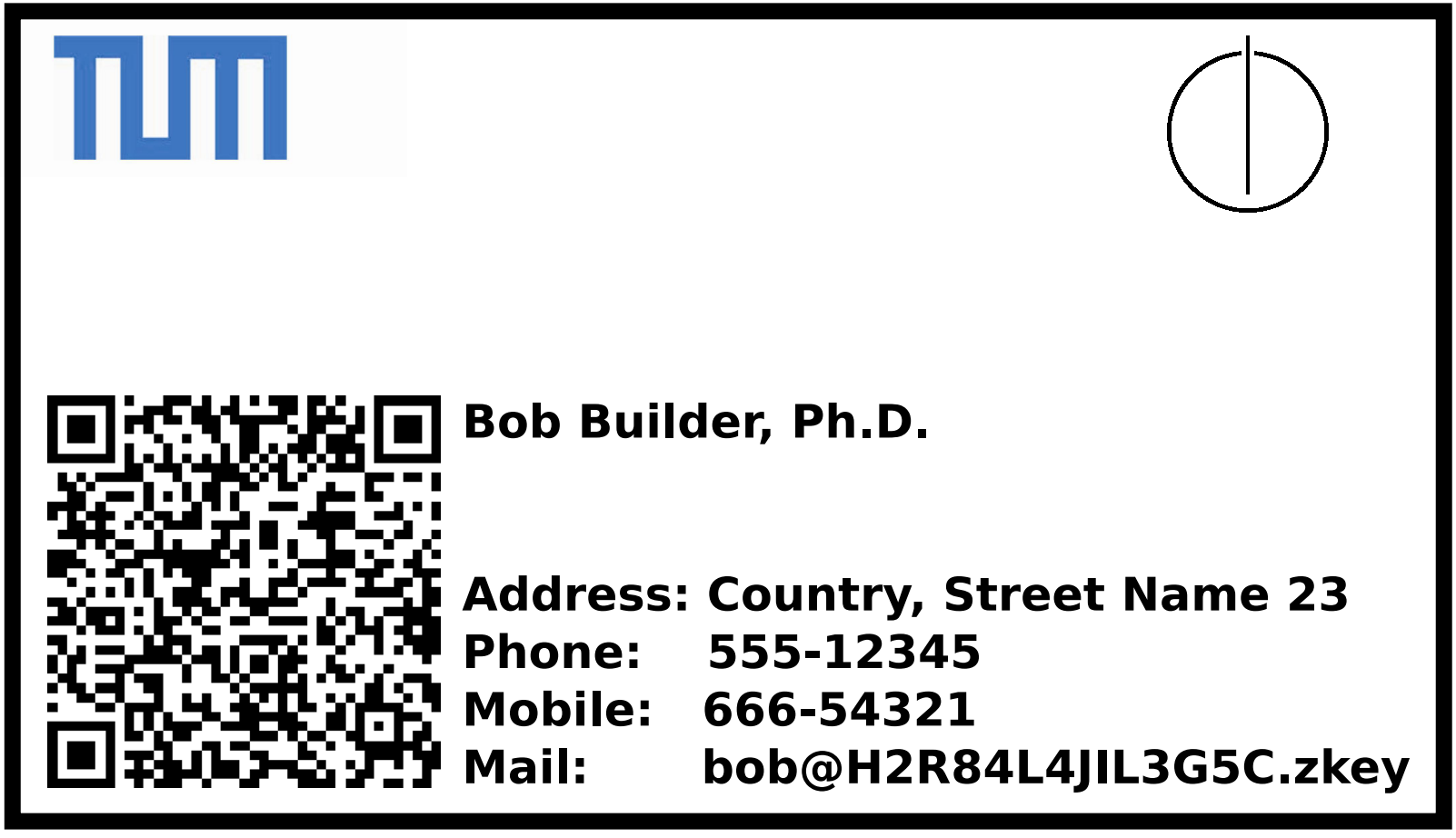
Bob

Bob's webserver



- Bob wants to be called **bob**
- Bob can reach his webserver via **www.gads**

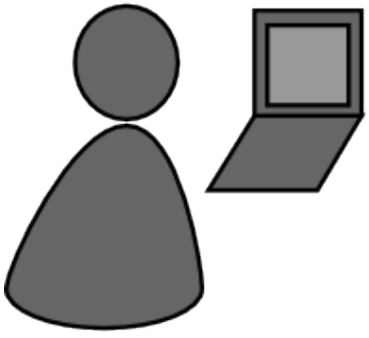
Secure introduction

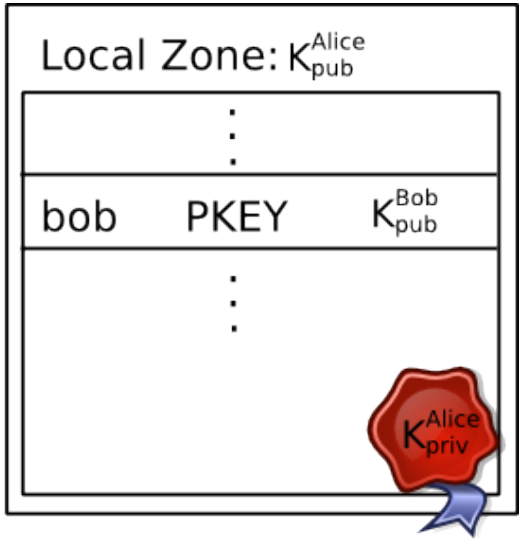


- Bob gives his public key to his **friends** via QR code

→ Bob's friends can resolve his records via *.*petname*.gads

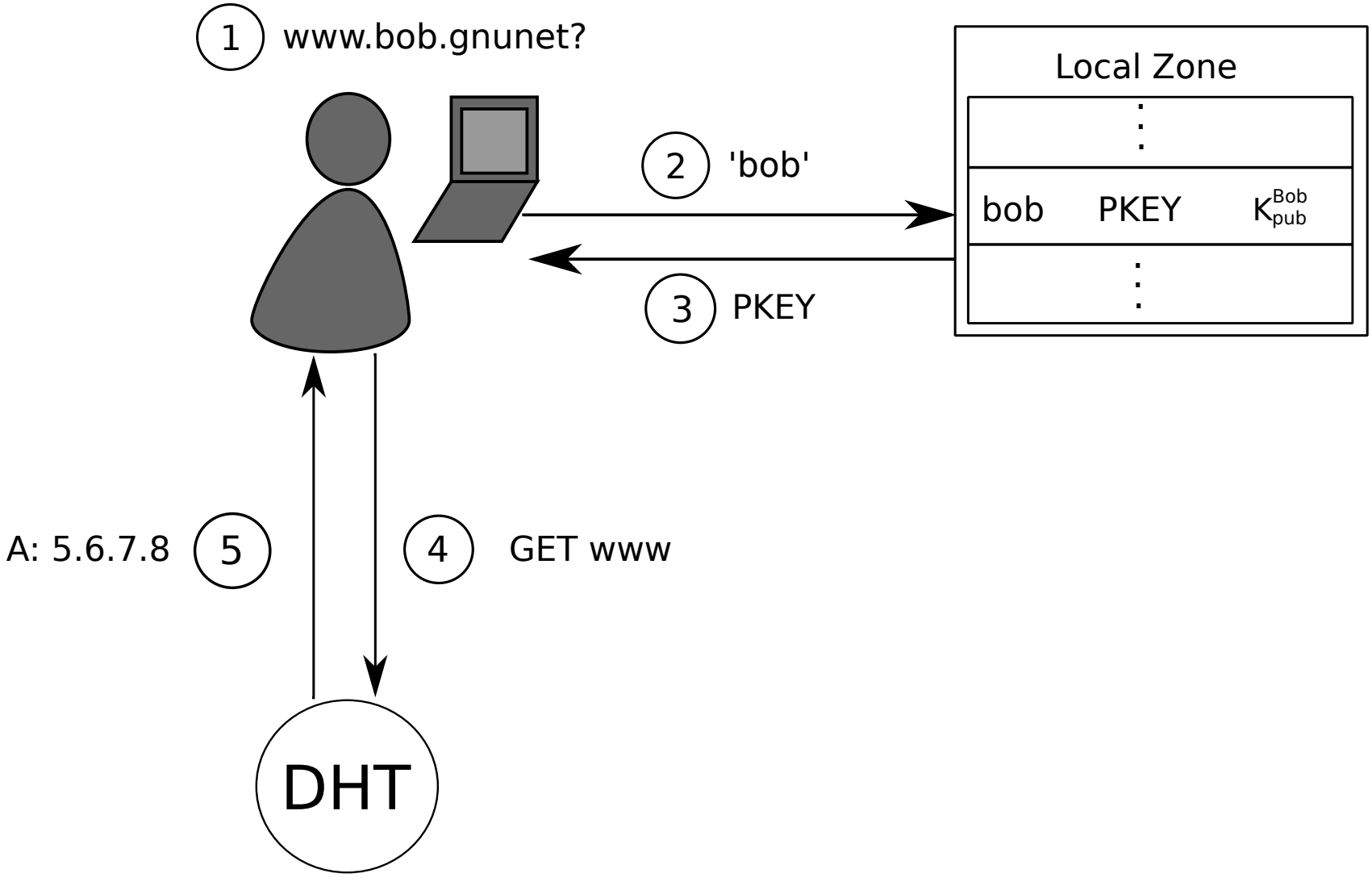
Delegation



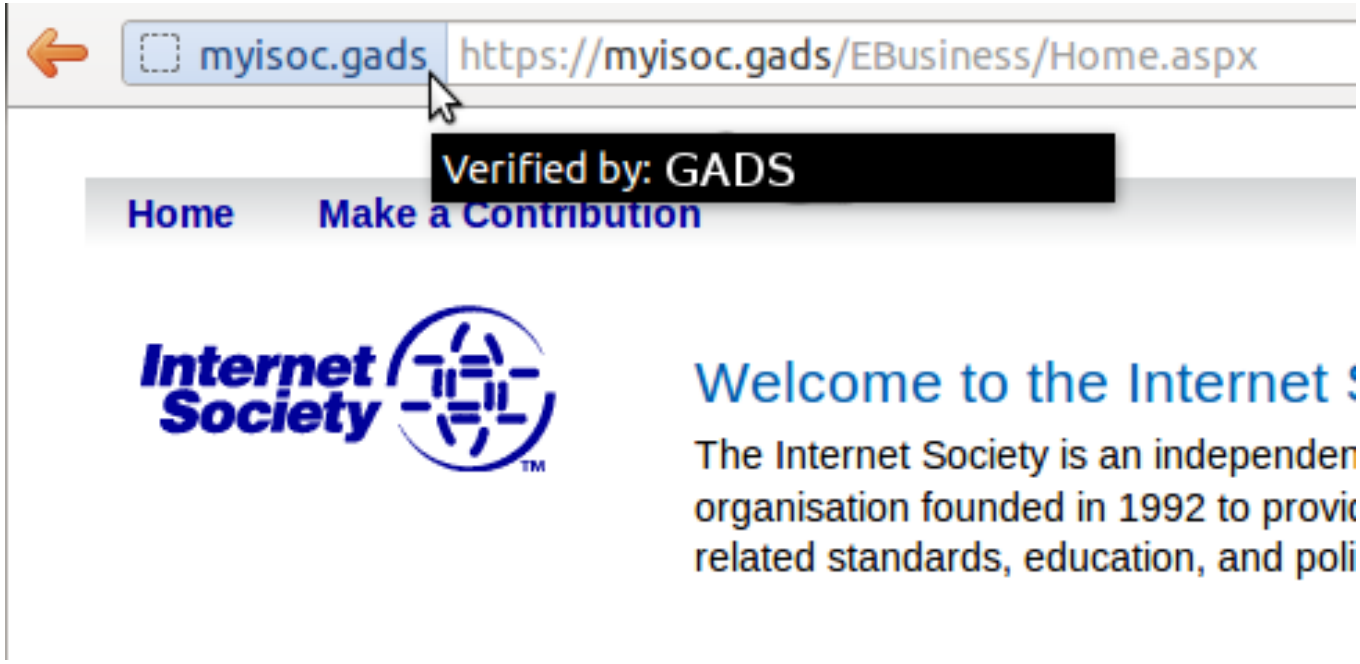


- Alice learns Bob's public key
- Alice creates delegation to zone **bob**
- Alice can reach Bob's webserver via **www.bob.gads**

Name Resolution



GADS as PKI (via DANE/TLSA)



GADS for GNUnet

Properties of GADS

- Decentralized name system with secure memorable names
- Decentralized name system with globally unique, secure identifiers
- QR codes for introduction, delegation used to achieve transitivity
- Supports standard DNS record types
- Can provide alternative PKI, validate TLS via TLSA records

Uses for GADS in GNUnet

- Pseudonymous file-sharing
- IP services in the P2P network (P2P-VPN) via "VPN" records
- Identities in social networking applications

Our Vision

Internet
Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer

GNUnet
RegEx/**PSYC**
GADS
Mesh (ECDHE+AES)
R^5N DHT
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

The Evolution Challenge[2]

- Features are frequently added to social applications
- Some require changes ("extensions") to data formats and messages
- Centralized, browser-based networks can easily update to new version
- Decentralized systems must transition *gracefully*

[2] Joint work with Carlo v. Loesch and Gabor Toth

Related Work: GNU libtool

Here are a set of rules to help you update your library version information:

1. Start with version information of 0:0:0 for each libtool library.
2. Update the version information only immediately before a public release of your software. More frequent updates are unnecessary, and only guarantee that the current interface number gets larger faster.
3. If the **library source code has changed** at all since the last update, then increment revision (c:r:a becomes c:r+1:a).
4. If any interfaces have been added, removed, or changed since the last update, increment current, and set revision to 0.
5. If any **interfaces have been added** since the last public release, then increment age.
6. If any **interfaces have been removed or changed** since the last public release, then set age to 0.

—taken from the GNU libtool manual.

Related Work: GNU libtool

There are three possible kinds of reactions from users of your library to changes in a shared library:

- Programs using the previous version may use the new version as **drop-in replacement**, and programs using the new version **can also work with the previous one**. In other words, no recompiling nor relinking is needed. In this case, bump revision only, don't touch current nor age.
- Programs using the previous version may use the new version as **drop-in replacement**, but programs using the new version may use **APIs not present in the previous one**. In other words, a program linking against the new version may fail with unresolved symbols if linking against the old version at runtime: set revision to 0, bump current and age.
- Programs may **need to be changed, recompiled, relinked** in order to use the new version. Bump current, set revision and age to 0.

Related Work: XML

- Extensible Markup Language
- Syntax is *extensible*
- Extensions have no **semantics**

PSYC

We are working on PSYC2, the successor to PSYC:

- More compact, mostly human-readable, faster-to-parse relative of XML/JSON/XMPP
- PSYC messages consist of a **state update** and a **method invocation**
- PSYC includes interesting ideas for social networking:
  - Stateful multicast
  - History
  - Difference-based updates
- PSYC addresses extensibility problem using **try-and-slice** pattern

PSYC State: Example

The PSYC state is a set of key-value pairs where the names of keys use underscores to create an **inheritance** relationship:

- _name
- _name_first
- _name_first_chinese
- _address
- _address_street
- _address_country

The data format for each state is fixed for each top-level label.

PSYC Methods: Example

A PSYC method has a name which follows the same structure as keys:

- _message
- _message_private
- _message_public
- _message_public_whisper
- _message_announcement
- _message_announcement_anonymous

Methods have access to the current state and a per-message byte-stream.

The Try-and-Slice Pattern

```
int msg (string method) {
  while (1) {
    switch (method) {
    case "_notice_update_news": // handle news update
      return 1;
    case "_notice":  // handle generic notice
      return 1;
    case "_message": // handle generic message
      return 1;
    // ...
    }
    int glyph = strrpos (method, '_');
    if (glyph <= 1) break;
    truncate (method, glyph);
  }
}
```

Advantages of Try-and-Slice

- Extensible, can support many applications
- Can be applied to state and methods
- Defines what backwards-compatible extensibility means:
  - Can incrementally expand implementations by deepening coverage
  - Incompatible updates = introduce new top-level methods

PSYC2 for GNUnet

Properties of PSYC

- Compact encoding (much smaller than XML/JSON/XMPP)
- Supports stateful multicast
- Supports message history (replay, see latest news, etc.)
- Extensible syntax and semantics

Uses for PSYC2 in GNUnet

- P2P social networking foundation (combine with GADS!)
- Pushes social profiles (state) to all recipients, no federation
- Replay from local database used as primary access method
- **My data is stored on my machine**
- Use secure multicast to support very large groups

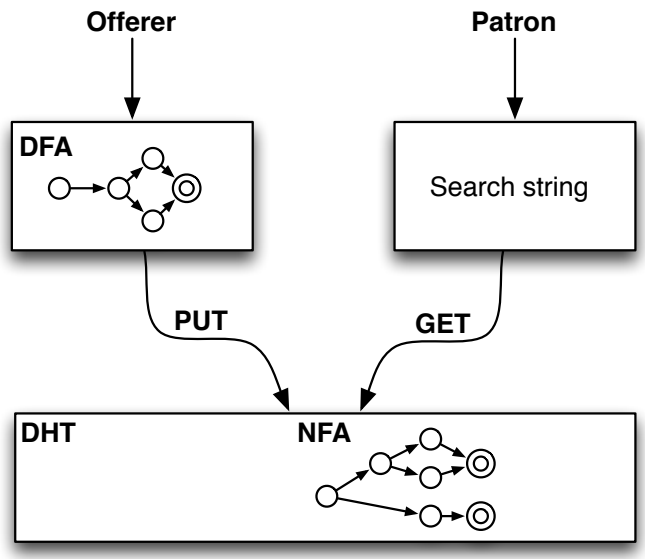
Our Vision

Internet
Google/Facebook
DNS/X.509
TCP/UDP
IP/BGP
Ethernet
Phys. Layer

GNUnet
RegEx/PSYC
GADS
Mesh (ECDHE+AES)
R^5N DHT
CORE (ECDHE+AES)
HTTPS/TCP/WLAN/...

Distributed Search via Regular Expressions: Idea[3]

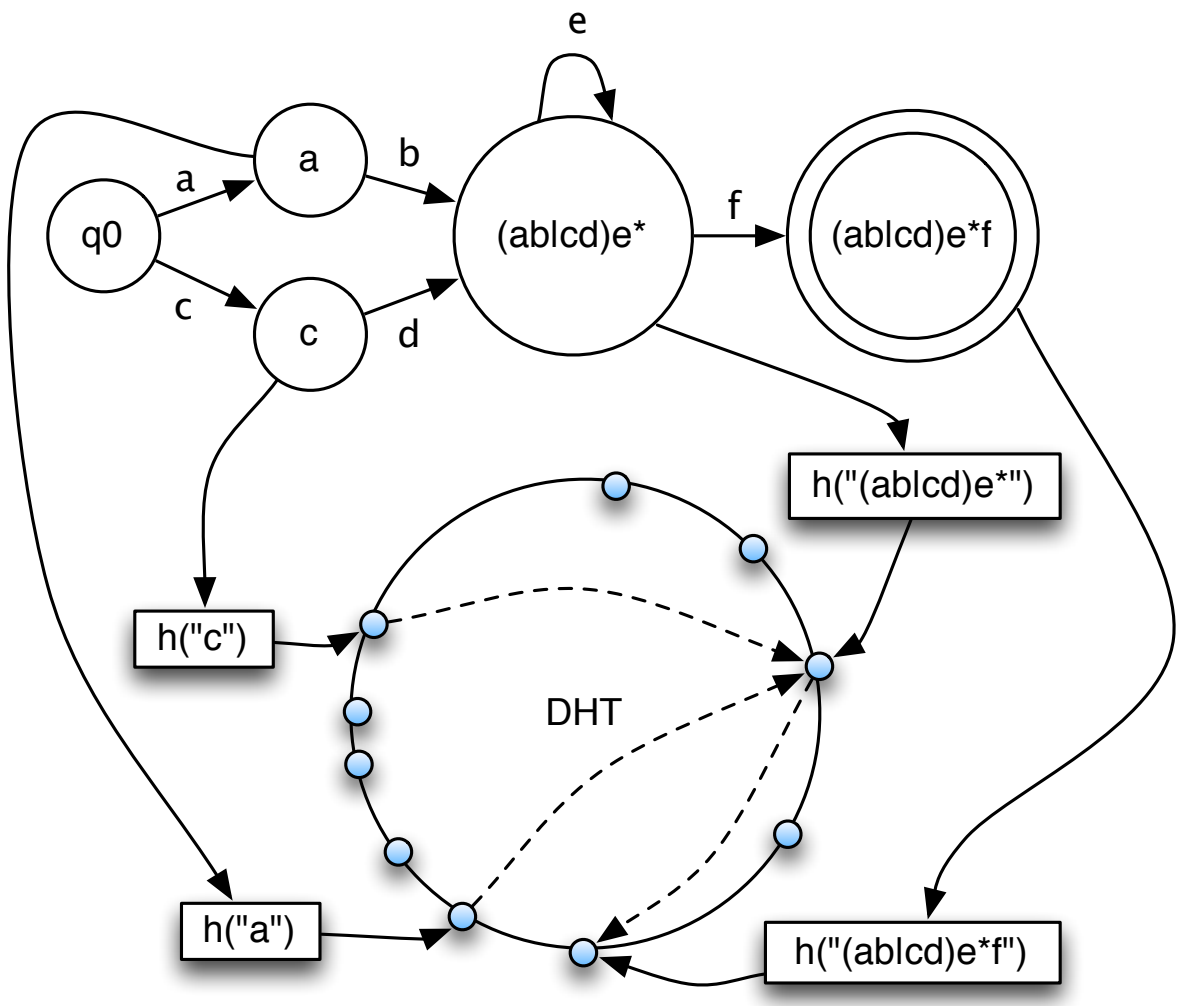
1. Offerer creates regular expression describing service
2. Regular expression is compiled to a DFA
3. DFA is stored in the DHT
4. Patron matches using a string



[3] Joint work with Max Szengel, Ralph Holz, Bart Polot and Heiko Niedermayer

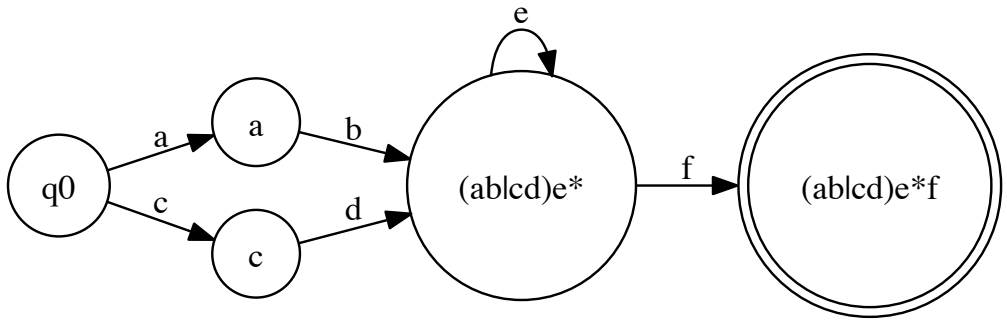
Problem: Mapping of States to Keys

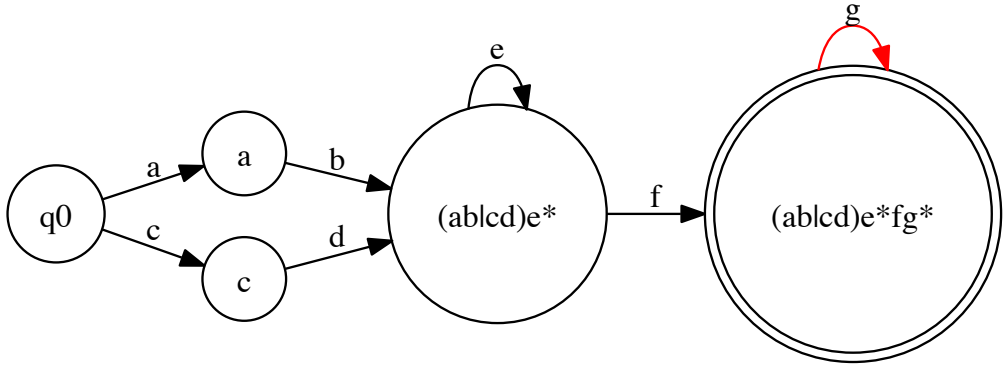
Regular expression $(ab|cd)e^*f$ and corresponding **DFA**



Problem: Merging of DFAs

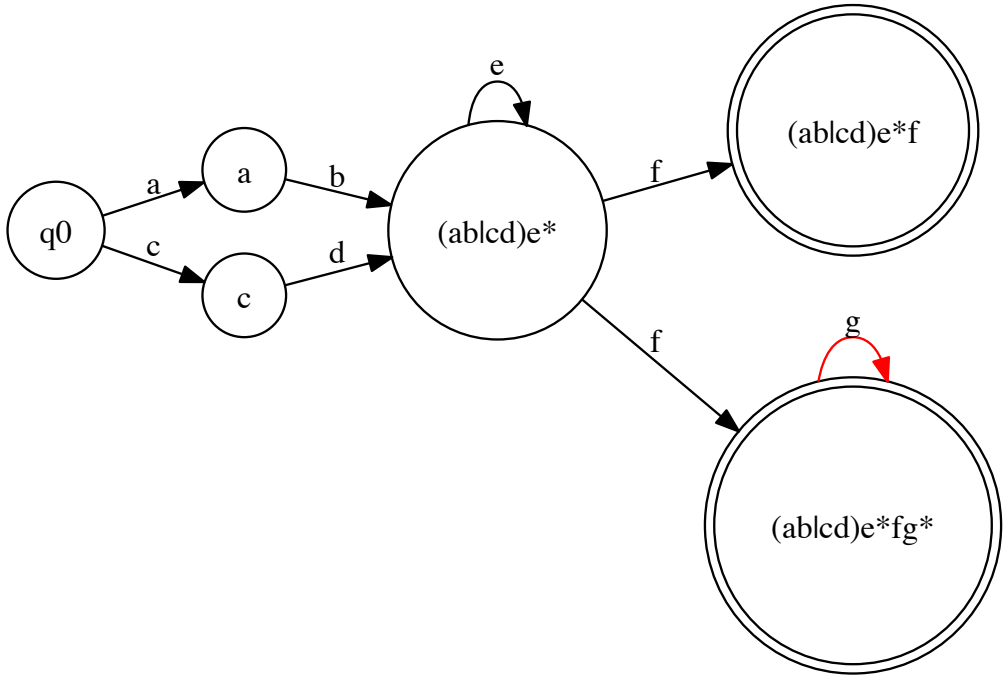
Regular expressions $(ab|cd)e^*f$ and $(ab|cd)e^*fg^*$ with corresponding **DFAs**





Problem: Merging of DFAs

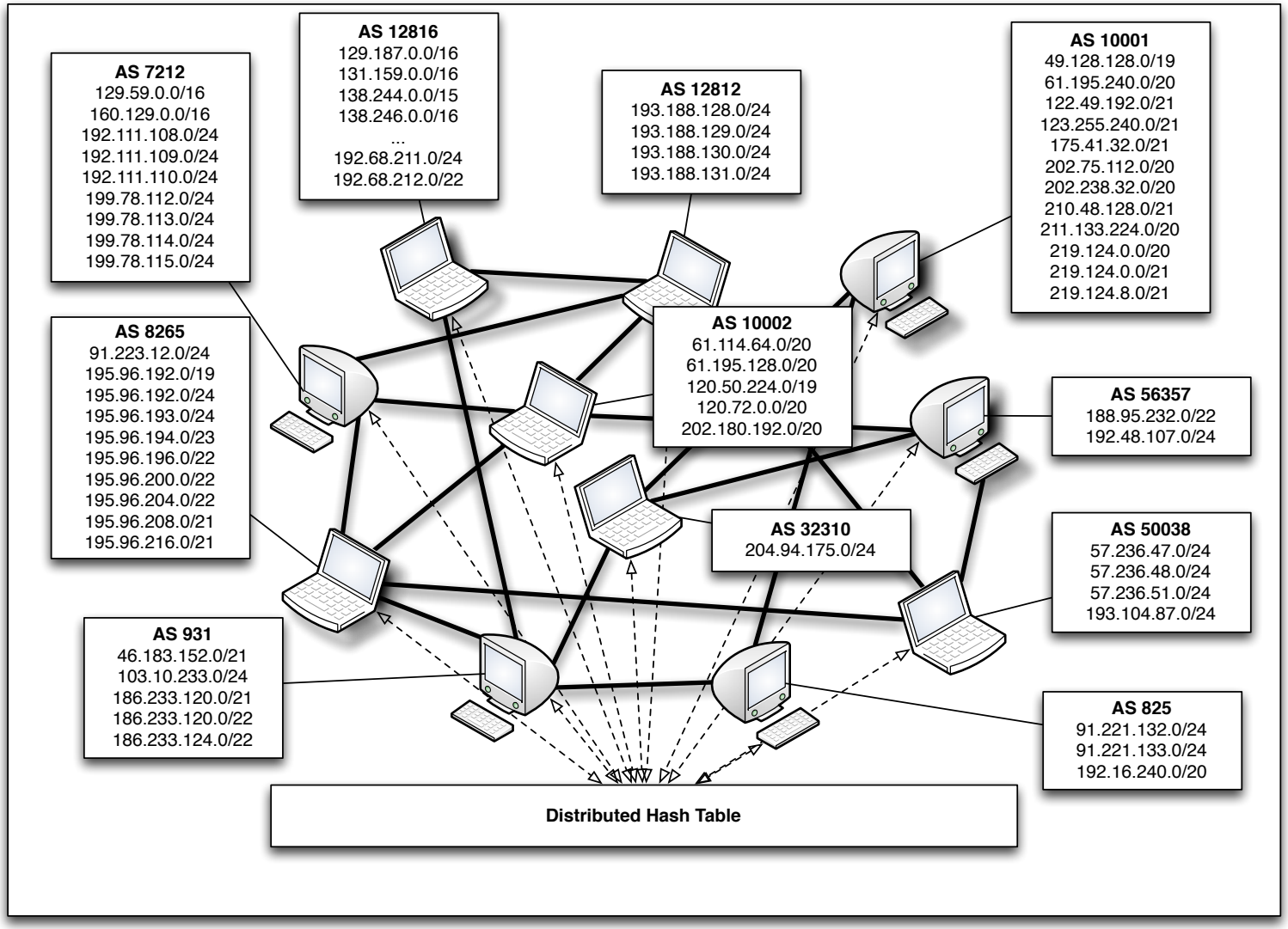
Merged **NFA** for regular expressions $(ab|cd)e^*fg^*$ and $(ab|cd)e^*f$



Evaluation

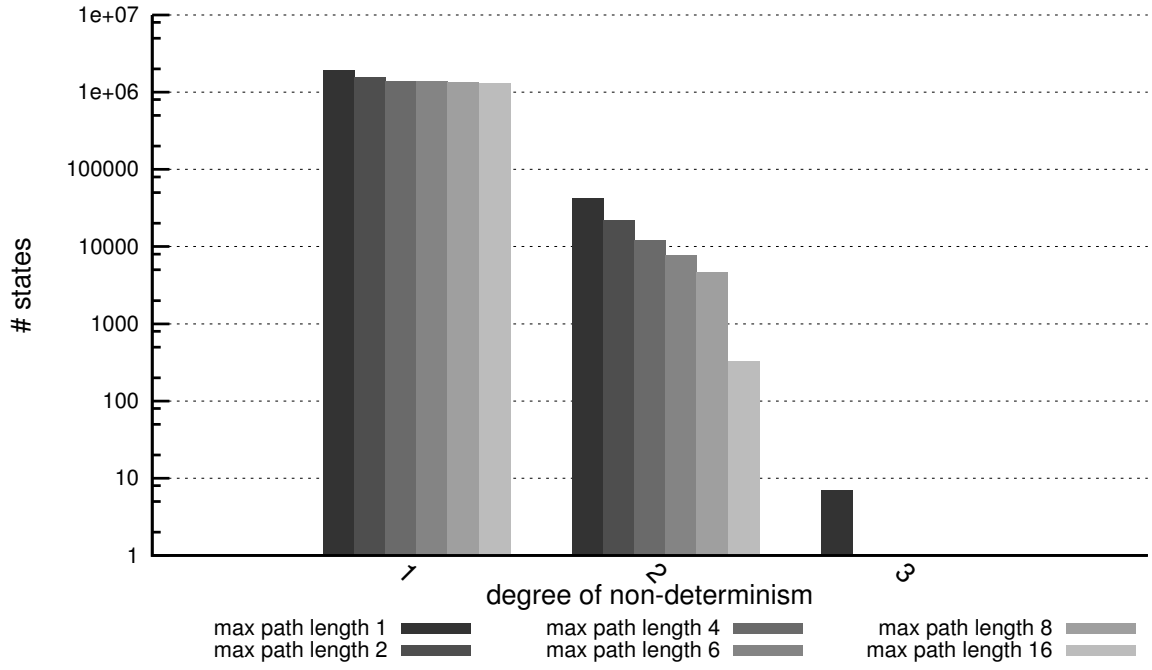
- Implementation in GNUnet
- Profiling of Internet-scale routing using regular expressions to describe AS address ranges
- CAIDA AS data set: Real AS data

Evaluation


AS 7212
129.59.0.0/16
160.129.0.0/16
192.111.108.0/24
192.111.109.0/24
192.111.110.0/24
199.78.112.0/24
199.78.113.0/24
199.78.114.0/24
199.78.115.0/24
AS 12816
129.187.0.0/16
131.159.0.0/16
138.244.0.0/15
138.246.0.0/16
...
192.68.211.0/24
192.68.212.0/22
AS 12812
193.188.128.0/24
193.188.129.0/24
193.188.130.0/24
193.188.131.0/24
AS 10001
49.128.128.0/19
61.195.240.0/20
122.49.192.0/21
123.255.240.0/21
175.41.32.0/21
202.75.112.0/20
202.238.32.0/20
210.48.128.0/21
211.133.224.0/20
219.124.0.0/20
219.124.0.0/21
219.124.8.0/21
AS 8265
91.223.12.0/24
195.96.192.0/19
195.96.192.0/24
195.96.193.0/24
195.96.194.0/23
195.96.196.0/22
195.96.200.0/22
195.96.204.0/22
195.96.208.0/21
195.96.216.0/21
AS 10002
61.114.64.0/20
61.195.128.0/20
120.50.224.0/19
120.72.0.0/20
202.180.192.0/20
AS 56357
188.95.232.0/22
192.48.107.0/24
AS 32310
204.94.175.0/24
AS 50038
57.236.47.0/24
57.236.48.0/24
57.236.51.0/24
193.104.87.0/24
AS 931
46.183.152.0/21
103.10.233.0/24
186.233.120.0/21
186.233.120.0/22
186.233.124.0/22
AS 825
91.221.132.0/24
91.221.133.0/24
192.16.240.0/20
Distributed Hash Table

Evaluation: Results of Simulation

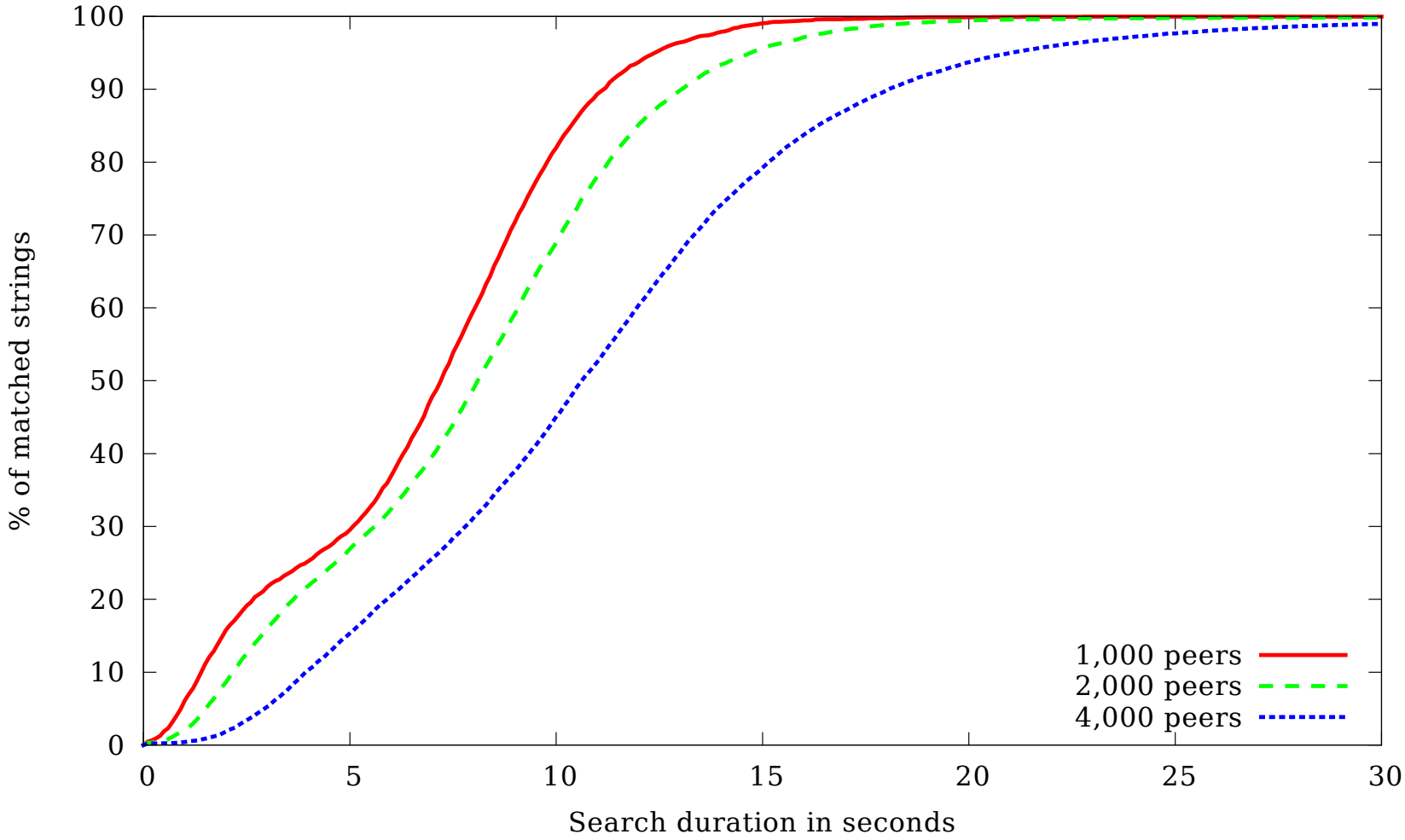
Degree of non-determinism at states in the merged NFA



Dataset: All $40,696$ ASes

Evaluation: Results of Emulation

Search duration averaged over five runs with randomly connected peers.



RegEx Search for GNUnet

Properties of RegEx Search

- Capability discovery in DHT-based P2P networks using regular expressions
- Linear latency in the length of the search string
- Suitable for applications that can tolerate moderate latency

Uses for GADS in GNUnet

- Network search
- Discovery of matching services, such as VPN exit nodes
- Topic-based subscriptions in messaging (decentralized MQTT)

Conclusion

- Everybody has something to hide
- Decentralization creates challenges for research

Conclusion

- Everybody has something to hide
- Decentralization creates challenges for research

We must decentralize or risk to loose control over our lives.

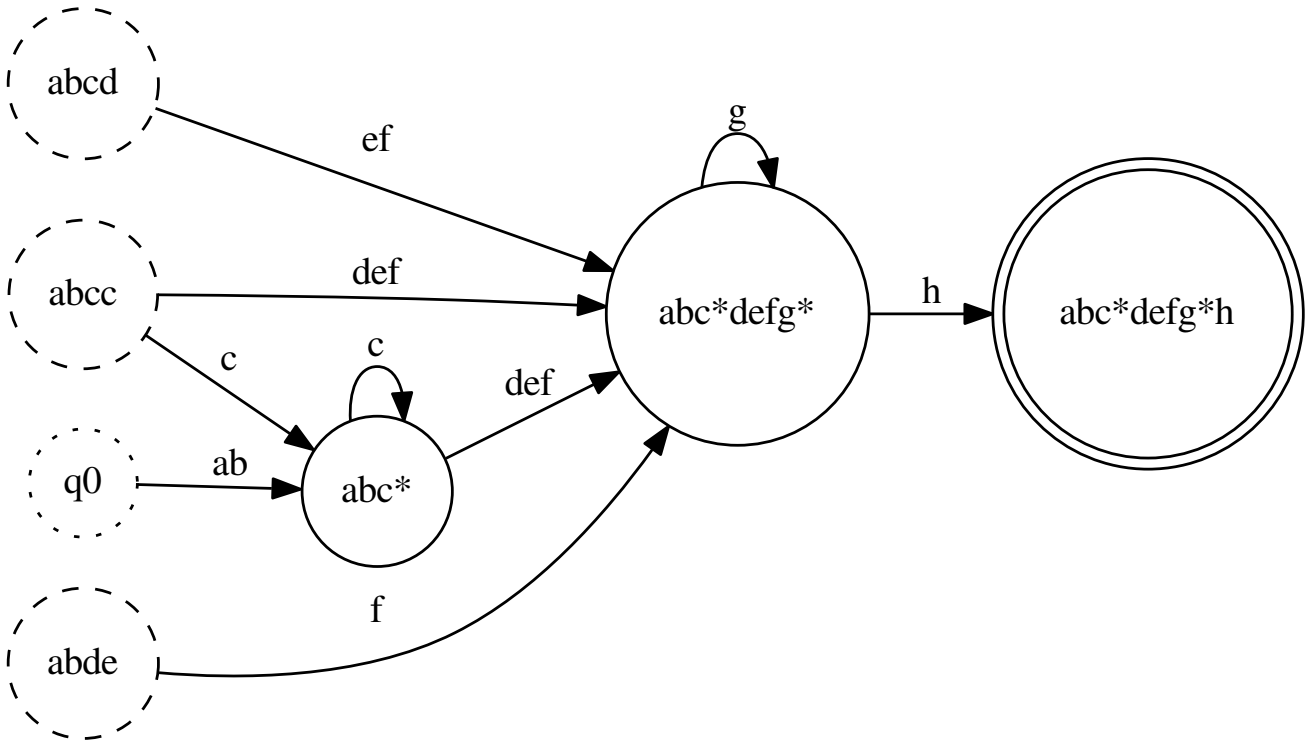
Do you have any questions?

References:

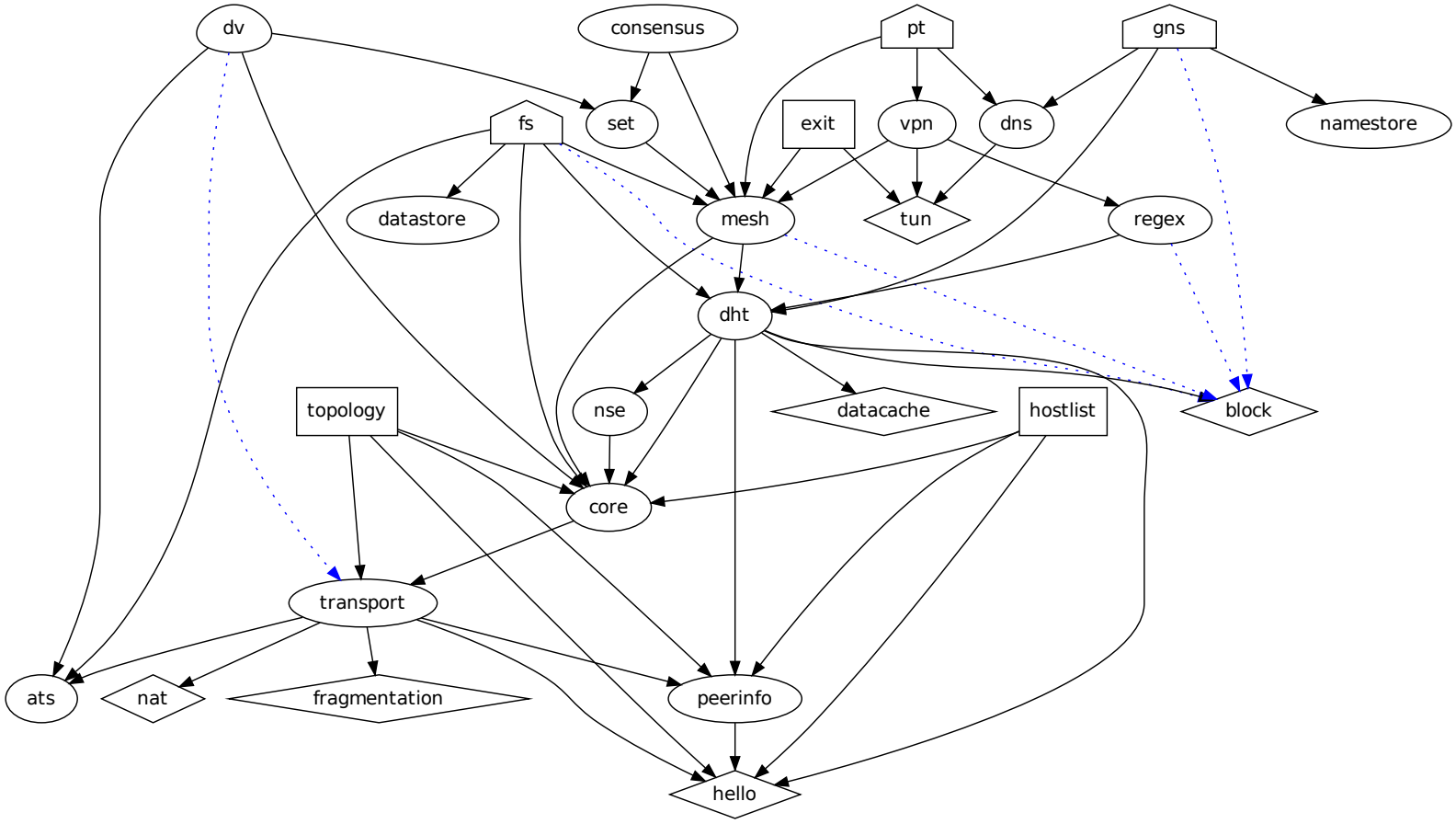
- Glenn Greenwald and Ewen MacAskill. *NSA Prism program taps in to user data of Apple, Google and others*. In **The Guardian**, June 7 2013.
- George Zornick. *Remember When NSA Surveillance Was Used to Help Launch the Iraq War?*. In **The Nation**, June 11, 2013.
- Michael Riley. *U.S. Agencies Said to Swap Data With Thousands of Firms*. In **Bloomberg**, Jun 14, 2013.
- Rudolf Wagner. *US-Spionage: Lauschangriff auf die Konkurrenz in Europa*. In **Der Spiegel**, Jan 7, 2001.
- Gerhard Schmid. *Report on the existence of a global system for the interception of private and commercial communications (ECHELON interception system) (2001/2098(INI))*. In **European Parliament Session Document**, July 11, 2001.
- Martin Asser. *Echelon: Big brother without a cause?* In **BBC News Online**, July 6, 2000.
- Nathan Evans and Christian Grothoff. *R5N. Randomized Recursive Routing for Restricted-Route Networks*. **5th International Conference on Network and System Security**, 2011.
- M. Schanzenbach *Design and Implementation of a Censorship Resistant and Fully Decentralized Name System*. **Master's Thesis (TUM)**, 2012.
- M. Szengel. *Decentralized Evaluation of Regular Expressions for Capability Discovery in Peer-to-Peer Networks*. **Master's Thesis (TUM)**, 2012.

Problem: Decentralizing the Start State

Regular expression: abc^*defg^*h and $k = 4$.



GNUnet: Framework Architecture



GNUnet: Envisioned Applications