

# **Trustworthy Internet: *Overlay Infrastructure for Trusted Computing and Communications***

Strategic Research Initiative of the GIGA  
technology program

Andrei Gurtov  
HIIT

GIGA day at PIMRC'06 11.9.2006

# Problem Statement

- Current Internet plagued by security issues
  - Top priority problem
- Phishing, viruses, worms, pharming, SPAM, Denial-of-Service (DoS) attacks
  - “The US Federal Trade Commission reports that identity theft now affects more than 10 million people every year representing an annual cost to the economy of \$50 billion. The Anti-Phishing Working Group reports that the frequency of these phishing attacks increases 24% every month”
- Current solutions are expensive (server certificates) or complex to use

# Our Approach

- Can't change deployed Internet infrastructure in near term
  - Consider deployment of IPv6 and DNSSEC
- Exploit recent advances in
  - public-key based security (Host Identity Protocol)
    - Allows authenticated and encrypted communication with mobility and multihoming
  - overlay networks (Internet Indirection Infrastructure, Distributed Hash Tables, Peer-to-peer)
    - Own routing protocols on top of existing IP networks between a set of hosts

# Example Scenario

- An Internet server receives an IP packet requesting secure association establishment
- The server would like to know if the source is trusted and thus worthwhile communicating with
- The server looks up the IP address in the OpenDHT overlay and receives the trust record
- The record contains information stored previously by this server and other servers on the reputation of the IP address
- The record is signed with certificates to prevent forgery
- The record tells that this address has been used for DoS attacks a long time ago, so a limited communication session is established

# Consortium



HIIT

- Prof. Martti Mäntylä
- Dr. Andrei Gurtov
- Dr. Pekka Nikander



TML/TKK

- Prof. Antti Ylä-Jääski

University of Helsinki

- Prof. Kimmo Raatikainen



# Project Information

- 100% Tekes funding
- HIIT (4py), TKK (2py), UH (1py)
- Duration: 3 + 2 years
- Budget: 600 000 €/year
- One person located at International Computer Science Institute (ICSI), Berkeley
- Related activities
  - Active at IETF/IRTF standardization
  - Teaching research seminar at TKK
  - Co-organizing Dagstuhl Seminar 06441 in Germany: Naming and Addressing for Next-Generation Internetworks, October 2006.

# Advisory Board

STONESOFT



Puolustusvoimat

Försvarsmakten | The Finnish Defence Forces



elisa

TeliaSonera

TietoEnator



HELSINKI

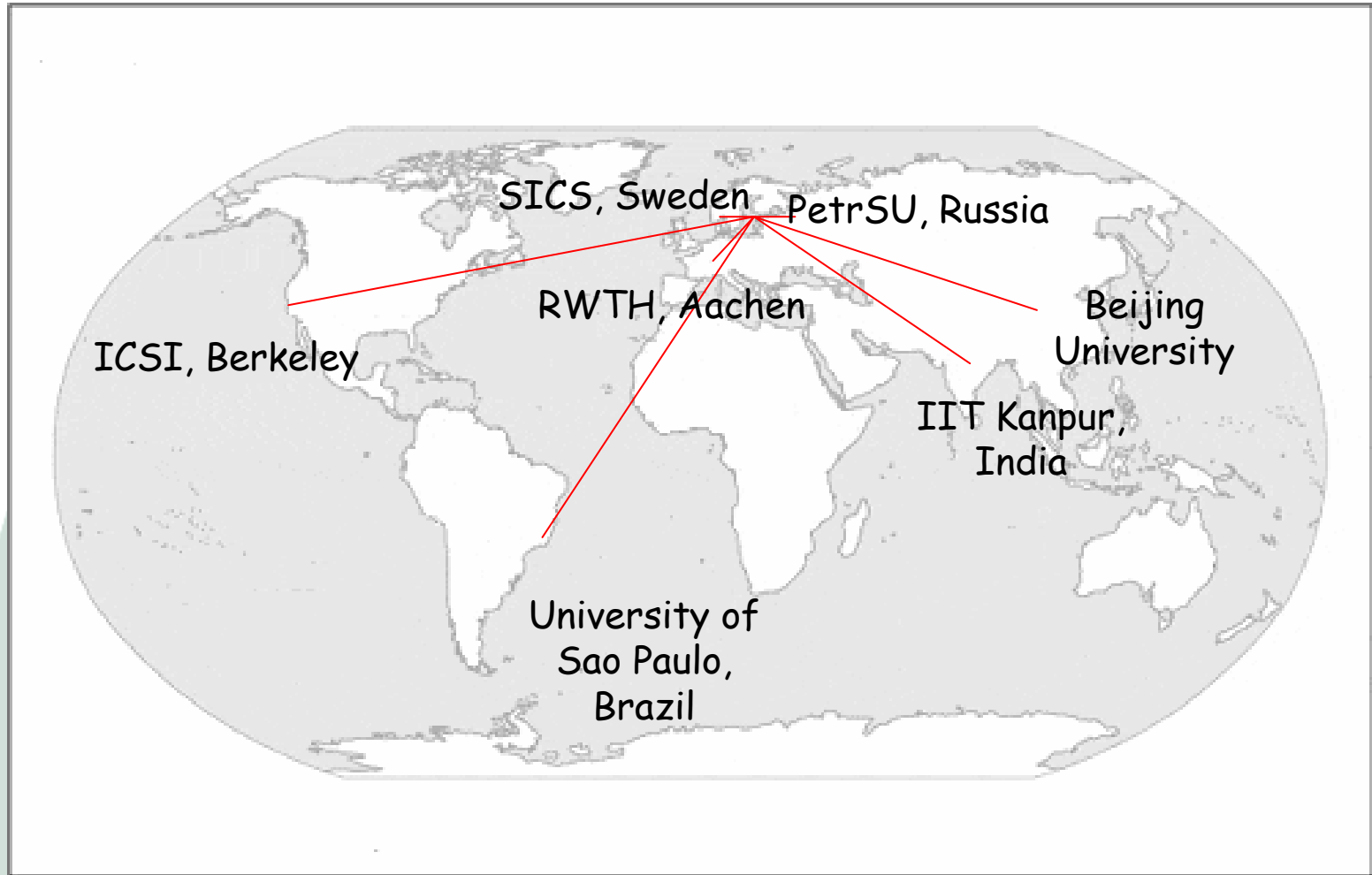
NOKIA  
Connecting People



TEKES

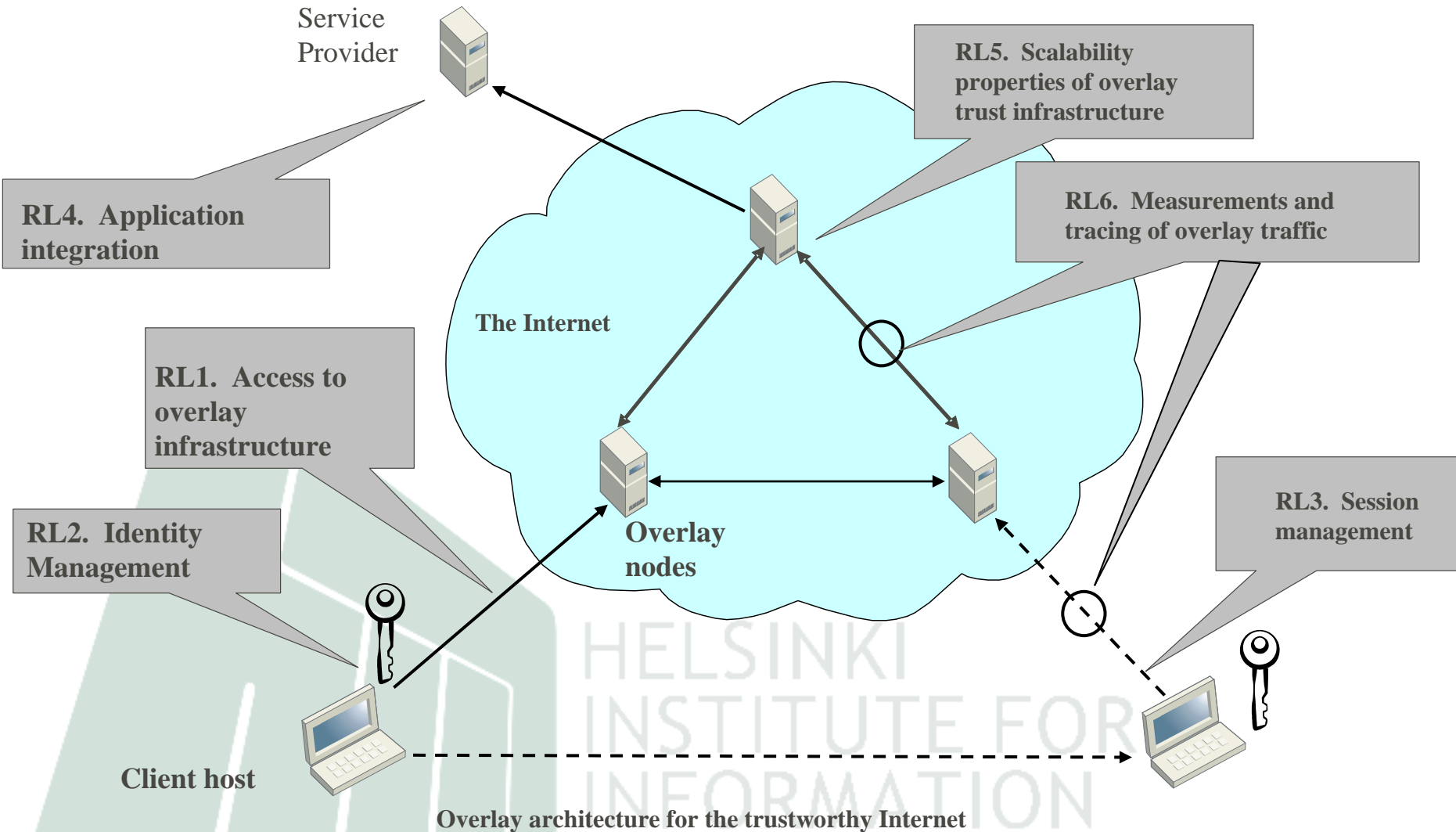
INFORMATION  
TECHNOLOGY

# International Connections



INFORMATION  
TECHNOLOGY

# Overlay Architecture



# RL1: Access to the Overlay Infrastructure

- Locating a nearby access node of the overlay infrastructure is an important component of the architecture
- Recently, IP anycast has been re-visited as a generic mechanism to find the nearest overlay node and enable gradual Internet evolution
- However, existing research has placed little consideration to the security aspects of such discovery and overlay connectivity
- Experimenting with Overlay-based **Anycast** Service Infrastructure (OASIS)
- Service discovery mechanisms to be extended to facilitate security and trust

# RL2: Identity Management

- Management of identities to enable trust relationships on different protocol layers
- In the current Internet, the host IP address is often the only base for authentication
  - Doesn't work in the presence of Network Address Translators
  - Doesn't allow mobility
- Previous experience with multiple identities support for Host Identity Protocol
  - Anonymous/changing identities to support user privacy
- The challenge of combining privacy and permanent identity for tracking reputation
- Robust exchange of identities (e.g. through device pairing)

# RL3: Session Management

- How to setup and manage the necessary session state over the overlay to forward packets based on trustworthiness of a session
- Too complex setup process could limit the scalability, while too lightweight setup process could result in low security
- Study the naming aspects of the sessions:
  - How to efficiently identify flows in a secure and privacy preserving manner?
  - In overall, what is the correct position of so called session identifiers in the Internet architecture?
  - Implementation of suspend/resume for TLS/SSL protocols

# Data-Oriented Network Architecture -1

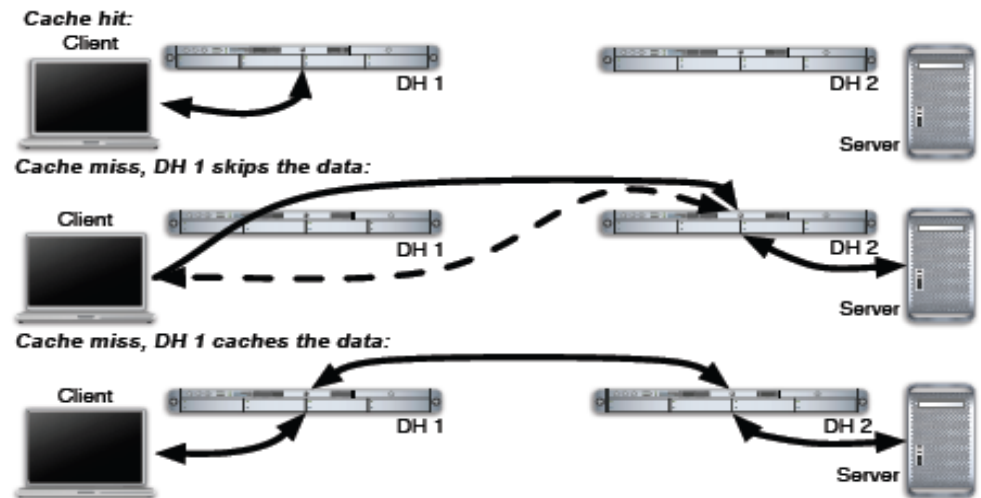
- Current Internet is built on a host-to-host model rather than host-to-data model
  - Data objects are identified using with a host-specific locator, e.g.  
<http://www.hiit.fi/images/logow.gif>
  - Data is authenticated using public-key infrastructure or securing a path to data
  - Fast majority of modern Internet use is data retrieval – just want a closest copy of data
  - Data replication and migration is complicated
    - Currently accomplished through DNS hacks e.g. by Akamai

# Data-Oriented Network Architecture -2

- Several researchers suggested moving from host-to-host paradigm to publish/subscribe
  - Data objects are named with flat (human unfriendly) 128-bit strings
  - Eases data caching, self-certifies object names, prevents DoS attacks
  - Routing on Flat Labels (ROFL) proposal from ICSI to route packets directly on data object ids instead of IP addresses
    - Interesting idea but faces scalability and deployment challenges

# Data-Oriented Network Architecture -3

- DONA is a continuation to ROFL from ICSI with contribution from TrustInet
  - DONA retains IP-routing while creating *fetch* and *register* operations for data objects on top
  - Data objects identified using public-private key pairs as in Host Identity Protocol



Three cases of fetching a data object through a Dissemination Handler (DH). Authoritative Resolvers (AR) take care of cache updates.

# RL4: Trust Infrastructure and Application Integration

- Storing reputation information for data objects or hosts through a distributed web of trust
- Overlay routing using *cycles* of node identifiers – without revealing IP addresses of overlay nodes
- Integration of applications to the trust overlay
  - Legacy applications
  - Overlay-aware applications
  - Human usability aspects of the interfaces
- Interface prototypes for HIP control GUI and possibly plugins to Firefox browser/Thunderbird email client

# Distributed Trust

- Storing reputation information for web servers, email addresses, hosts
- Generalizing the distributed web of trust model from Pretty Good Privacy (PGP)
  - PGP keys and reputation records are stored in key servers that are *loosely* synchronized
  - Adopting a single but distributed storage for reputation data using OpenDHT
- Challenges
  - Preventing Sybil attacks (bulk identity generation)
  - Who verifies certificates – could create a DoS attack by itself
  - Who is responsible at refreshing data stored in OpenDHT?

# RL5: Scalability Properties of a P2P Naming Architecture with Trust Requirements

- Scalability of large-scale overlays is an open research issue
- It is critical to understand, whether such overlays can support as large group of clients as DNS does today
- Analytical modeling is almost the only approach to study these issues, given immense difficulties for measuring or even simulating networks with millions of nodes
- When trust relationships are involved among communicating nodes, required properties from overlays are more challenging than for pure (static) naming purposes

HEL SINKI  
INSTITUTE FOR  
INFORMATION  
TECHNOLOGY

# RL6: Measurements, Traffic analysis, Experimentation

- The prototype architecture needs to be experimentally evaluated
- The use of PlanetLab/EMULAB/GENI as tools will allow real-world evaluation in various scenarios
  - Currently, own set of ~100 i3 overlay servers
- Using measurements to collect parameters in mid-scale infrastructure deployment to calibrate analytical models
  - Measuring hop latency between servers
  - Using public set of ~150 OpenDHT servers

# Thanks!

- Please also see TrustInet poster/OpenDHT demo today
- Further information: [gurtov@hiit.fi](mailto:gurtov@hiit.fi)
- Related InfraHIP project: <http://infrachip.hiit.fi>



HELSINKI  
INSTITUTE FOR  
INFORMATION  
TECHNOLOGY