



The LOBSTER project



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

Network Monitoring for Performance and Security

The LOBSTER project: Current State and Collaboration Opportunities

Evangelos Markatos

Institute of Computer Science (ICS)

Foundation for Research and Technology – Hellas (FORTH)

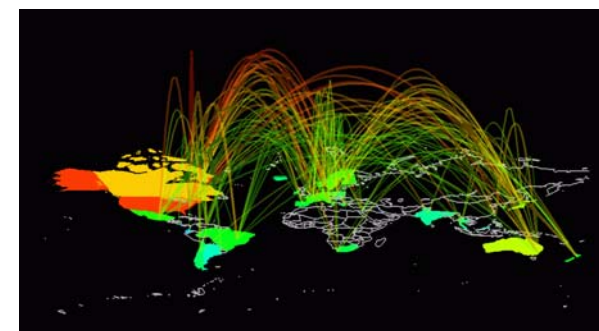
Crete, Greece



Roadmap



- Motivation
 - What is the problem?
 - Why is it important?
- Passive network traffic monitoring
 - The SCAMPI and LOBSTER approach
- Privacy protection
- How can you get involved?
 - Follow our activities
 - Install a sensor
 - Become part of the Infrastructure





Motivation

- Committee on **Research Horizons in Networking** (formed in 2001)
 - **David Patterson, Chair**, University of California at Berkeley
 - RISC processors, RAID storage, NOWs – clusters of workstations
 - **David Clark**, MIT Laboratory for Computer Science
 - The “father” of the “end-to-end” argument, on top of which the Internet design is based
 - **Anna Karlin, Jim Kurose, Edward D. Lazowska, David Liddle, Derek McAuley, Vern Paxson, Stefan Savage, Ellen Zegura**
- The committee was assigned to
 - “**formulate a fresh look at networking research**”
- They prepared a report
 - “**Looking over the fence at networks: a Neighbor’s View of Networking Research**”
 - They identified three “**Grand Research Challenges**”





lobster

<http://www.ist-lobster.eu>

Motivation: A GRAND Challenge in Networking



<http://www.ist-lobster.org>

The first GRAND Challenge in Computer Networking is to

“... develop and deploy the technology to make it possible to record a day in the life of the Internet...”



Committee on Research Horizons in Networking
Clark, Lazowska, Patterson, Paxson, Savage, Zegura,
2001



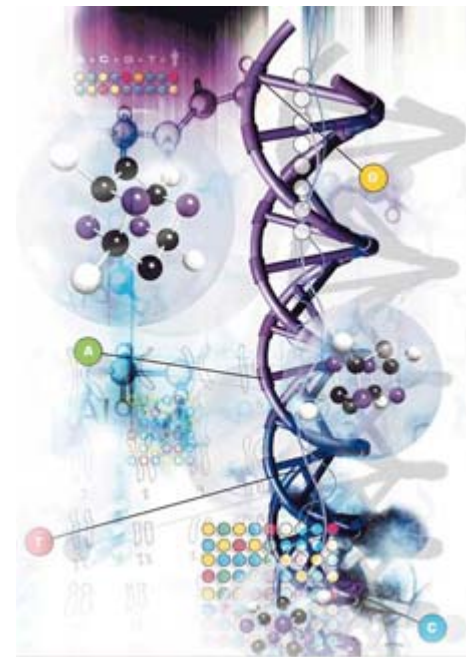
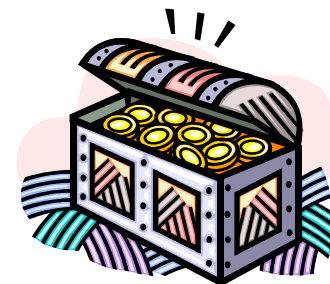
info@ist-lobster.org

Evangelos Markatos, FORTH



The Challenge

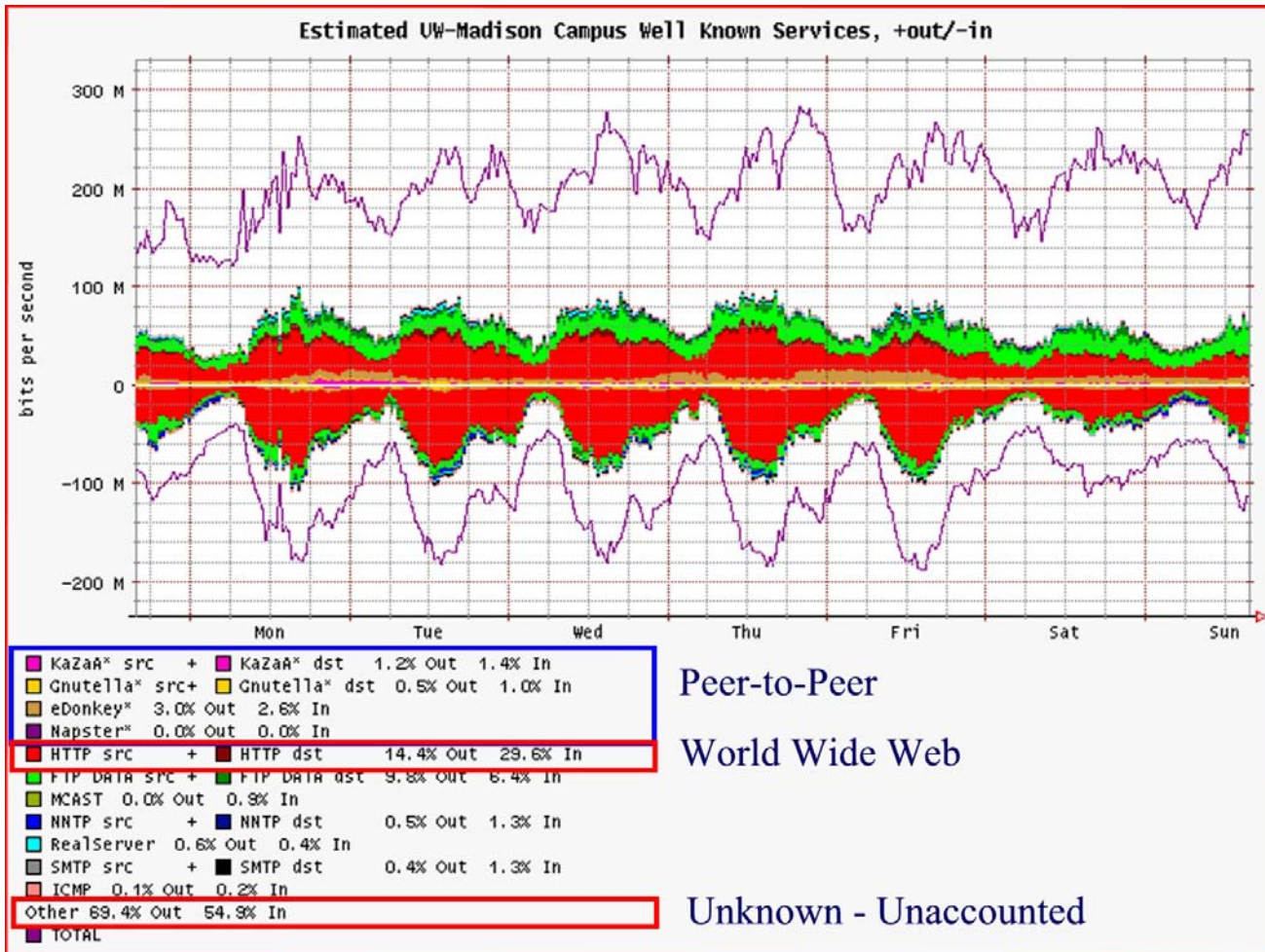
- Why is it important?
 - “a data set with **typical days for the next 10 years of the Internet** might be a **treasure chest for networking researchers**”
 - **Measurement-based GRAND challenges**, such as the **human genome**, have served to
 - Crystallize research issues, and
 - Mobilize research efforts
 - Good data are necessary for the **operational needs**
 - Why is my network slow?
 - Which route do my packets follow?
 - Why is a particular flow missing lots of packets?
 - How much peer-to-peer traffic is there?
- **Next GRAND Challenge in Networking Research:**
 - **Monitor a day in the life of the Internet**





Example Monitoring Question:

Which application generates all this traffic?



69% of the traffic is unaccounted-for

- Maybe belongs to p2p applications that use dynamic ports
- Maybe belongs to media applications
- The bottom line is:
 - We don't know



lobster

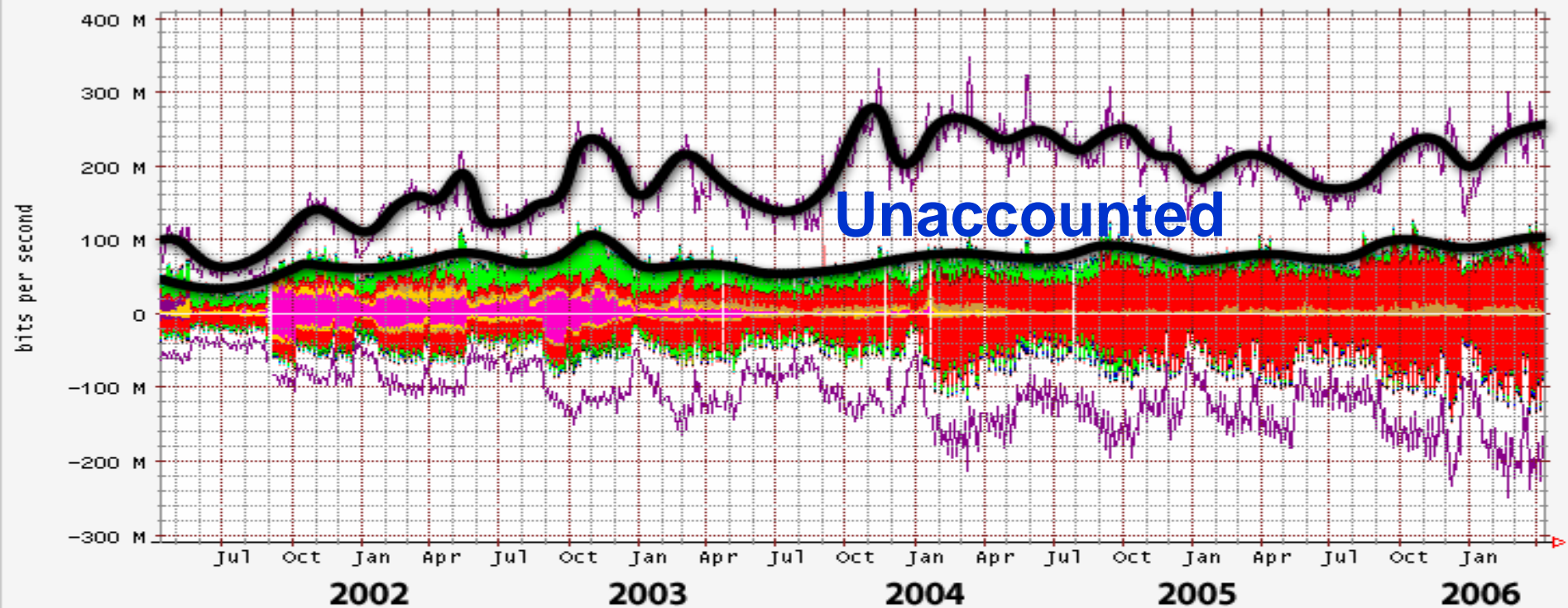
<http://www.ist-lobster.eu>

Unaccounted traffic is increasing



<http://www.ist-lobster.org>

Estimated UW-Madison Campus Well Known Services, +out/-in



KaZaA* src +	KaZaA* dst	5.1% Out	5.6% In
Gnutella* src+	Gnutella* dst	2.6% Out	2.6% In
eDonkey*		2.5% Out	1.8% In
Napster*		0.3% Out	0.2% In
HTTP src +	HTTP dst	18.2% Out	34.4% In
FTP DATA src +	FTP DATA dst	12.1% Out	6.2% In
MCAST		0.0% Out	1.0% In
NNTP src +	NNTP dst	0.5% Out	1.4% In
RealServer		1.1% Out	0.6% In
SMTP src +	SMTP dst	0.7% Out	1.5% In
ICMP		0.3% Out	0.2% In
Other		57.7% Out	45.0% In
TOTAL			



lobster

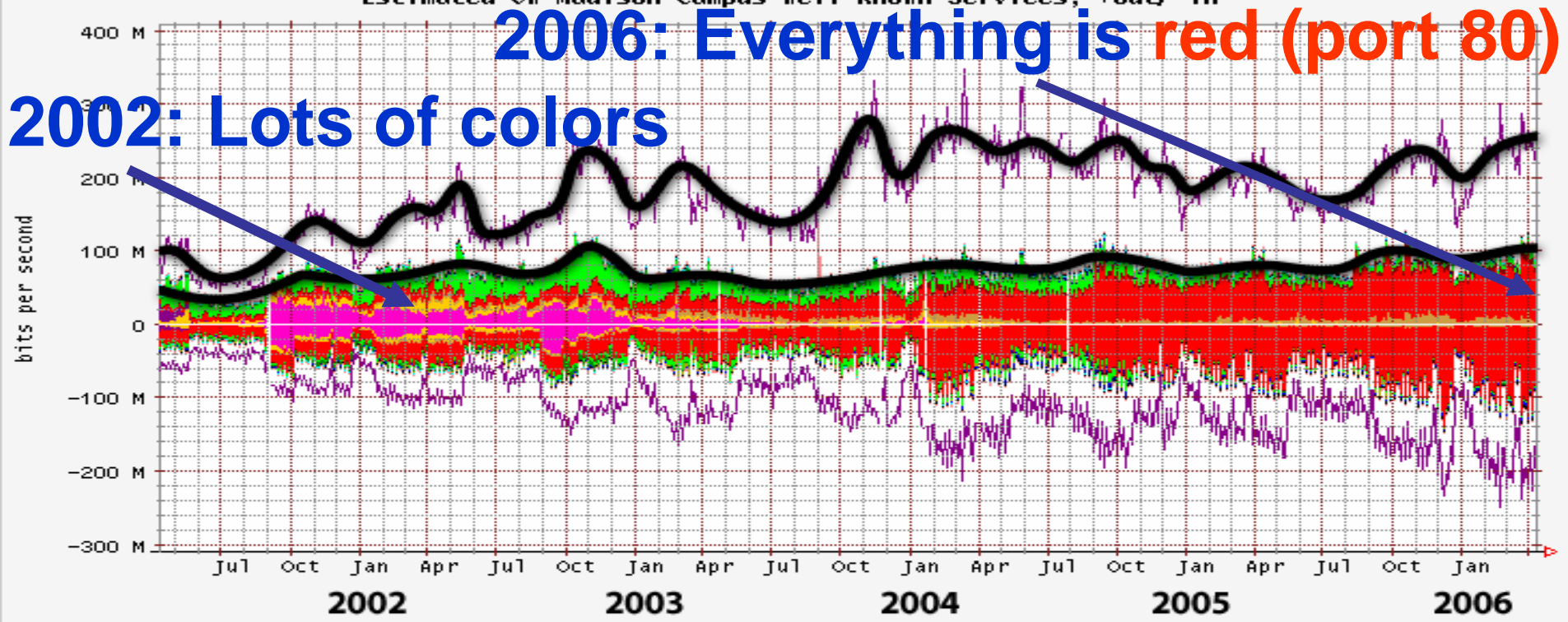
<http://www.ist-lobster.eu>

Even accounted-for traffic seems “suspicious”



<http://www.ist-lobster.org>

Estimated UW-Madison Campus Well Known Services, +out/-in



KaZaA* src +	KaZaA* dst	5.1% Out	5.6% In
Gnutella* src+	Gnutella* dst	2.6% Out	2.6% In
eDonkey*		2.5% Out	1.8% In
Napster*		0.3% Out	0.2% In
HTTP src +	HTTP dst	18.2% Out	34.4% In
FTP DATA src +	FTP DATA dst	12.1% Out	6.2% In
MCAST		0.0% Out	1.0% In
NNTP src +	NNTP dst	0.5% Out	1.4% In
RealServer		1.1% Out	0.6% In
SMTP src +	SMTP dst	0.7% Out	1.5% In
ICMP		0.3% Out	0.2% In
Other		57.7% Out	45.0% In
TOTAL			



Problem summary



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

- Our understanding of the Internet needs to be improved
- Grand Challenge: Monitor a day in the life of the Internet
- The **gap** between what **we can measure** and what **we need to measure** is large and getting larger



Solution?

- We need **better network monitoring**:
 - Faster: to cope with Gbit lines
 - More accurate: close the gap between what we know and what is really going on





SCAMPI-LOBSTER



Two projects towards meeting the challenge

Information Society
Technologies

<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

- Since 2001, we are trying to make contributions towards facing this GRAND challenge:
 - SCAMPI is an IST project: Funded by European Commission
 - Duration: 1/4/02-31/3/05
 - LOBSTER is a Specific Support Action Funded by European Commission
 - Two-year project, Duration 1/10/05-31/12/06





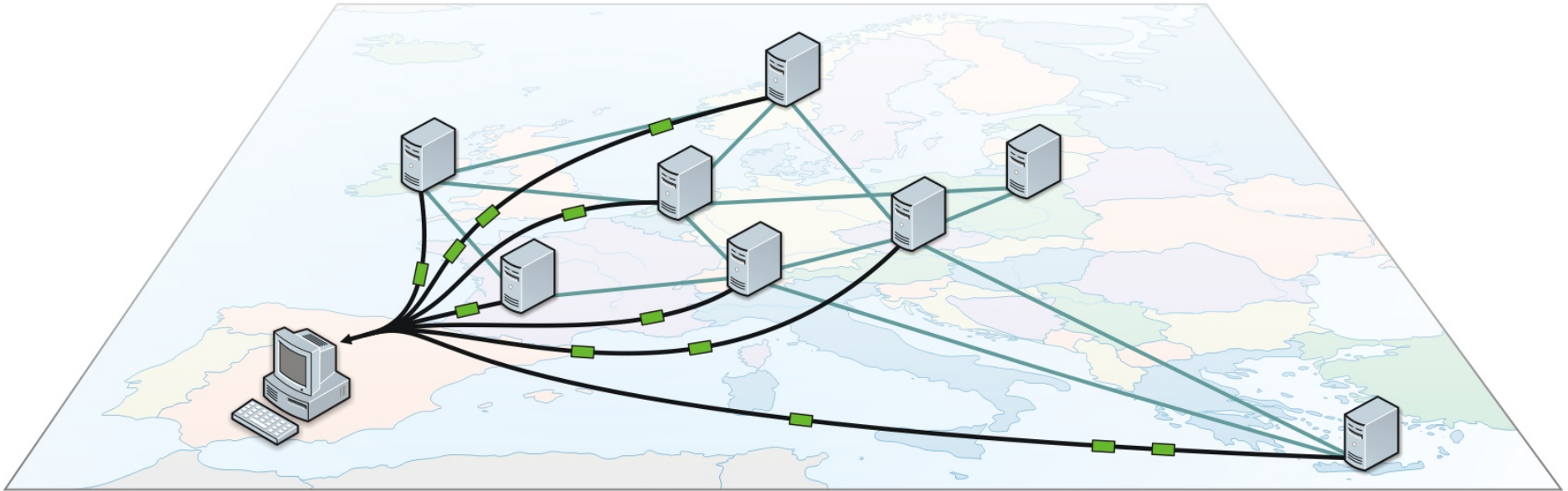
What has been developed so far?

- **A Network Monitoring Programming Environment**
 - **M**API (**M**onitoring **A**pplication **P**rogramming **I**nterface)
- An implementation of MAPI on top of several Monitoring cards
 - Regular network interfaces
 - Commodity Intel cards
 - Specialized network interfaces
 - Combo6 cards (by CESNET – Masaryk University)
 - DAG Cards (by Endace)





MAPI



- Programming Environment
 - MAPI (Monitoring API)
 - Up and Running
 - Use it to monitor several sensors at the same time



So, what is MAPI?

- Main Abstraction: The **Generalized NETWORK FLOW**
- A (generalized) Network Flow is a **subset** of the traffic:
 - All packets destined to **port 80**
 - All packets destined to www.cnn.com to **port 80** containing the string **“this is a test”**
- Network flows efficiently define a **subset** of the traffic
- What can you do with the packets of the network flow?
 - Capture them – dump them to disk
 - Post-process them later
 - To find time distributions, to observe suspicious patterns
 - Apply functions to them, e.g.
 - Count them, count the number of bytes they contain, etc.





Example: Search for a worm



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

- **SLAPPER worm**: from port 2002 to port 80, contains string "|00 00|E|00 00|E|00 00|@|00|"

```
int main() {
// Create a network flow consisting of SLAPPER worm packets
int fid;
struct mapipkt *pkt;
/* create a flow using the eth0 interface */
fd = mapi_create_flow ("eth0");
/* the bpf part of the signature */
mapi_apply_function (fd, "BPF_FILTER", "udp and src port 2002 and dst net 139.91.23 and dst port 80");
/* the content search part of the signature */
mapi_apply_function (fd, "STR_SEARCH", "|00 00|E|00 00|E|00 00|@|00|", 0, 100);
/* must use TO_BUFFER in order to read full packet records */
fid = mapi_apply_function (fd, "TO_BUFFER");
/* connect to the flow */
if( mapi_connect (fd) < 0) {
    printf("Could not connect to flow %d\n", fd);
    exit(EXIT_FAILURE);}
while(1) { /* forever, wait for worm packets */
    pkt = mapi_get_next_pkt (fd, fid);
    printf("\n Slapper worm packet! \n");
    print_IP_pkt(pkt);
}
}
```



What can I do with it?

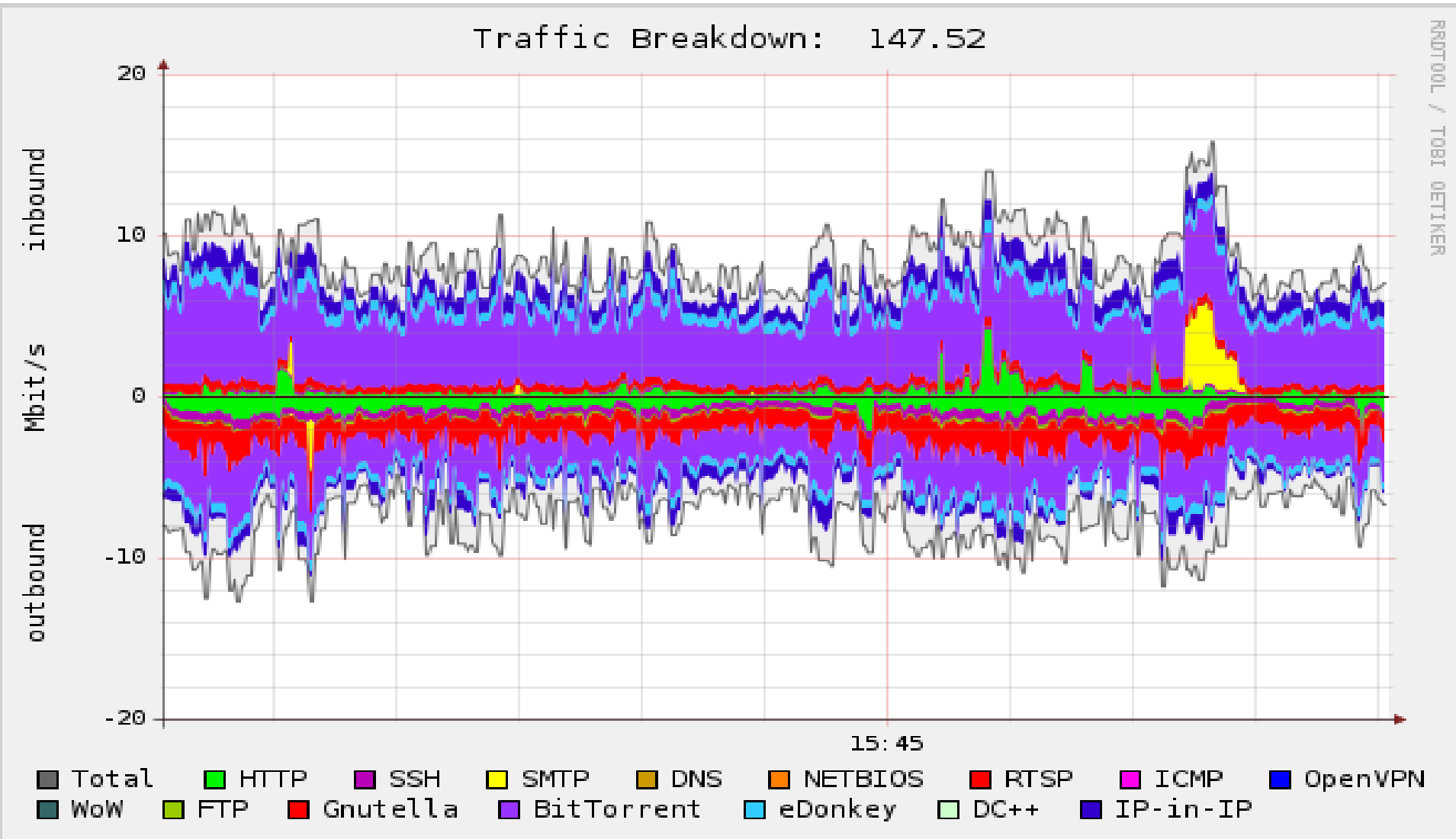
- Isolate and **study performance problems**
- **Troubleshooting**
 - I see weird port scans from port 1414, and 1515. Is anyone else seeing them?
- Gather **statistics**
 - How much Gnutella traffic do we have today?
- Observed **bandwidth for individual applications**
 - How many Mbps does my **GRID-enabled application** receives today?
- Identify **covert channels**
 - Is there any “**covert**” traffic masquerading as web traffic using port 80?
- Identify **BOTS**
 - Are there any compromised computers in my network engaged in suspicious activities?
- Monitor applications with **dynamic ports**
 - Teleconferencing, p2p, etc.





What have we done so far?

Applications: Traffic Categorization



RRDTOOL / TOBI OETIKER



What have we done so far?

Applications: Anomaly/Worm Detection



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

EAR Online Monitor

Parameters

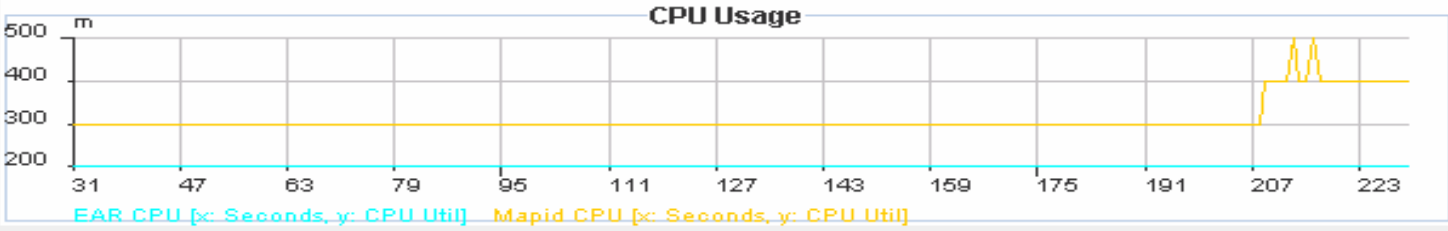
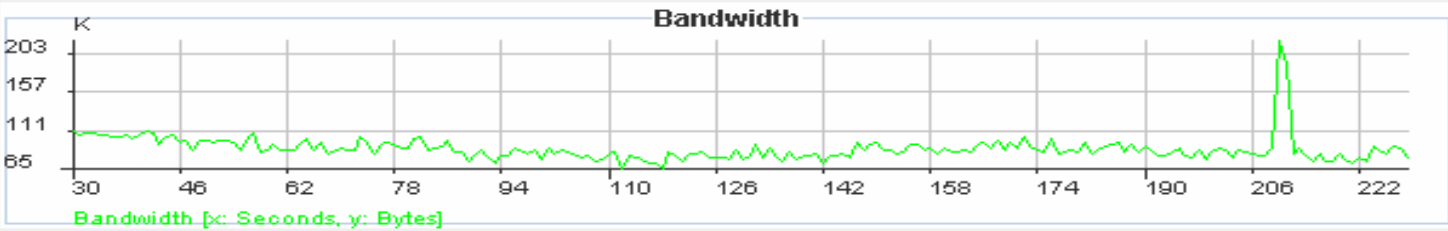
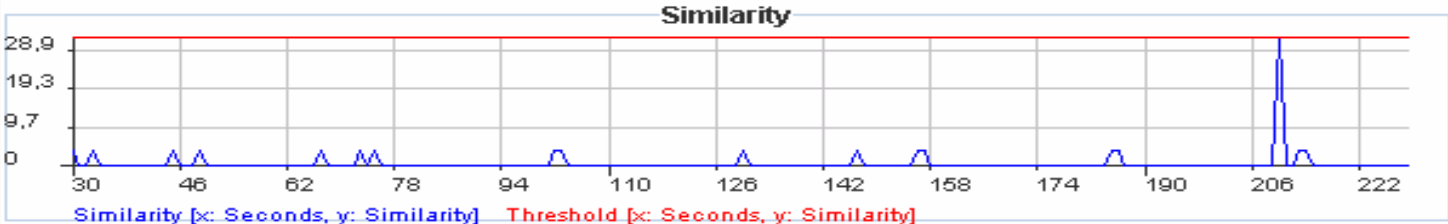
Sub-String Length (bytes)

Destinations Threshold

Time Threshold (milli-seconds)

Status

Uptime : 0:8:40
 Bytes Processed : 48221910
 New Alerts : 19
 Total Alerts : 19
 Current Parameters : 300/5/2000





Witty worm Snort signature



```
alert udp any 4000 -> any any (msg:"ISS PAM/Witty Worm  
Shellcode"; content: "|65 74 51 68 73 6f 63 6b 54 53|"; depth:246;)
```



DiMAPI code

```
fd = mapi_create_flow("<scope>");  
mapi_apply_function(fd, BPF_FILTER,  
    "udp and src port 4000");  
mapi_apply_function(fd, STR_SEARCH,  
    "|65 74 51 68 73 6f 63 6b 54 53|", 0, 246);
```





DiMAPI Applications: Intrusion Detection (2/2)



lobster

<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

```

Shell - Konsole <2>
mikepo@castro:~/review> ./worm 139.91.70.59:eth1 139.91.70.65:eth2
scope: 139.91.70.59:eth1 139.91.70.65:eth2

*** Witty worm packet ***
0:4:23:AB:A4:A1 > 0:7:E9:17:B4:8F type 0x800 IPv4
10.1.0.3:1487 > 10.1.0.5:0 UDP TTL:64 TOS:0x0 ID:41822 IpLen:20 DgmLen:1053
Len: 1025
45 00 04 01 d3 b4 00 00 71 11 dd a9 db 9a 9c a1      E.....q.....
41 ad da a4 0f a0 c4 24 03 ed dd 38 05 00 00 00      A.....$.8....
00 00 00 12 02 00 00 00 00 00 00 00 00 00 00      .....
00 02 2c 00 05 00 00 00 00 00 00 00 6e 00 00 00      ..,.....n....
00 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 .....
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 .....
41 02 05 00 00 00 00 00 00 00 de 03 00 00 00 00 00 .....
00 00 00 00 00 00 00 00 00 00 01 00 00 01 00 00 .....
01 00 00 1e 02 20 20 20 20 20 20 28 5e 2e 5e      ..... (^.^
29 20 20 20 20 20 20 69 6e 73 65 72 74 20 77 69      )      insert wi
74 74 79 20 6d 65 73 73 61 67 65 20 68 65 72 65      tty message here
2e 20 20 20 20 20 20 28 5e 2e 5e 29 20 20 20 20      .      (^.^)
20 20 20 89 e7 8b 7f 14 83 c7 08 81 c4 e8 fd ff      .....
ff 31 c9 66 b9 33 32 51 68 77 73 32 5f 54 3e ff      .1.f.32Qhws2_T>.
15 9c 40 0d 5e 89 c3 31 c9 66 b9 65 74 51 68 73      ..@.^..1.f.etQhs

```



lobster

<http://www.ist-lobster.eu>

Potential LOBSTER applications: Early-warning systems

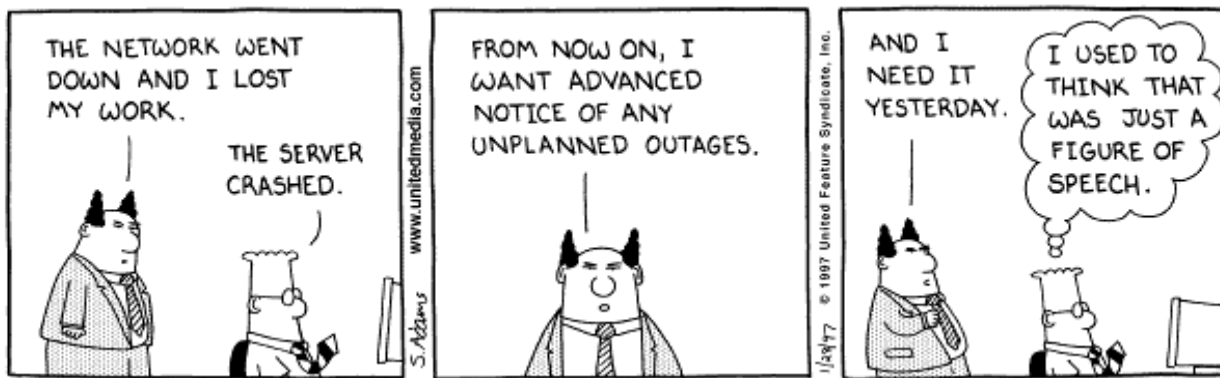


Information Society
Technologies

<http://www.ist-lobster.org>



- Automatic detection and fingerprinting of zero-day worms
 - How?
 1. Find lots of “similar” incoming packets [ICC 05]
 - Which have not been seen before
 2. “Execute” the incoming request
 - If the incoming request executes for a long time it is suspicious [SEC 05] [DIMVA 06]
- Use Passive Monitoring to Complement Honeypots
 - With Shadow honeypots: “stand-by servers” [USENIX SEC 05]
 - Network-level inspection: it finds “suspicious” packets
 - Shadow honeypots: receive “suspicious” packets and serve them





What do I need to start monitoring?

- A regular PC
 - With a regular network Interface
 - Put it in promiscuous mode
 - Mirror your traffic from a router to the PC
 - Start collecting packets
- What is the cost of the hardware?
 - Minimal – use an old PC
 - It works well for 10-100 Mbps lines
 - It even works for 1 Gbps lines for lightweight monitoring
 - not heavy string searching



- For higher speeds (10Gbps) you may need
 - A better PC (better I/O and memory bandwidth)
 - A better network Interface (Endace, Combo)



- But you can always start small, inexpensive, and grow into it



What software do I need?

- Off-the-shelf libraries
 - Pcap (<http://sourceforge.net/projects/libpcap/>)
 - MAPI (<http://www.ist-lobster.org/downloads/>)
- Applications
 - Some are provided with MAPI
 - Counting
 - Traffic Categorization
 - DoS attack detection, etc.
 - Write your own
 - And **share it!**
 - (<http://mapi.uninett.no/index.php?page=download>)

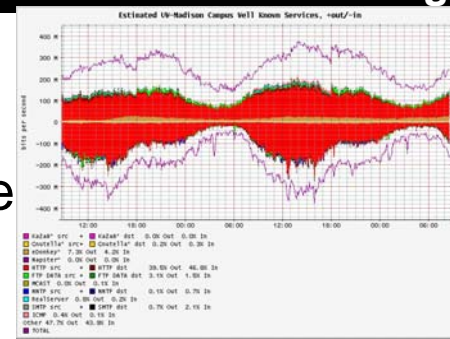


How about my privacy?

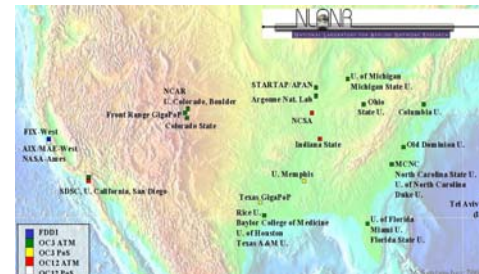


- Concerned about the privacy of
 - My users, my network, my topology
- Before I give you LOBSTER's solution, let's contemplate on this question:
- Am I you willing to share information?

- **No, absolutely not. I will not reveal any information.**
 - **I will not reveal even the load of my network**



- Maybe I can share some general statistics
 - Much like U of Wisconsin does
- I am willing to share the headers of the packets (IP addresses anonymized, payload stripped)
 - Just like NLANR does today
- I would like to share
 - all information with my local administrators
 - Some information with associated researchers
 - Anonymized information with the rest of the world



- **Yes – share everything**

LOBSTER can support all of the above



Sharing - Anonymization

- Anonymization Framework
- Flexible of anonymization
 - Per-field anonymization
 - Examples
 - Anonymize src/dst IP addresses and strip the payload (NLNR traces)
 - **Prefix preserving** src/dst IP address anonymization and strip the payload (Dartmouth traces)
 - For FTP traces: strip ftp passwords, strip files names transferred (LBL ftp traces)



Evangelos Markatos, FORTH



lobster

<http://www.ist-lobster.eu>

Anonymization Tool



<http://www.ist-lobster.org>

Anonymization is also available as a standalone tool from <http://www.ist-lobster.org/downloads/>

Example:

Original Packet:

```

07/20-15:35:01.849758 0:D0:D3:6:90:0 -> 0:10:54:45:F0:A8 type:0x800
len:0x162
147.52.67.200:4155 -> 195.167.100.117:80 TCP TTL:125 TOS:0x0 ID:5941
IpLen:20 DgmLen:340 DF
***A*** Seq: 0x6051E9BF Ack: 0x8BF955BB Win: 0xFB9E TcpLen: 20
47 4 4 20 2F 62 61 6E 6E 72 73 2F 79 65 6E GET /banners/yel
6C 61 77 6E 65 74 2F 73 70 6C 72 74 2F 73 70 6F lownet/sport/spo
72 74 59 6E 67 62 65 74 32 30 78 33 30 30 2E rtingbet230x300.
73 71 56 3F 67 6F 3D 68 74 70 25 33 41 2F 2F swf?go=http%3A//
79 61 54 61 2E 61 64 6D 61 2E 67 72 2F 63 6C yoda.adman.gr/cl
69 61 5B 2F 32 35 37 35 2F 39 37 36 2F 2F 25 ick/2575/6976//%

```

Anonymized Packet:

```

07/20-15:35:01.849758 0:D0:D3:6:90:0 -> 0:10:54:45:F0:A8 type:0x800
len:0x162
129.0.0.5:4155 -> 193.0.0.11:80 TCP TTL:125 TOS:0x0 ID:5941 IpLen:20
DgmLen:340 DF
***AP*** Seq: 0x6051E9BF Ack: 0x8BF955BB Win: 0xFB9E TcpLen: 20

```



Who can benefit from LOBSTER?



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

- **NRNs/ISPs**
 - Better Internet traffic monitoring of their networks
 - Better understanding of their interactions with other NRNs/ISPs
- **Security analysis/researchers**
 - Access to anonymized data
 - Access to anonymized testbed
 - Study trends and validate research results
- **Network and security administrators**
 - Access to a traffic monitoring infrastructure
 - Access to early-warning systems
 - Access to software and tools



lobster

How can you get involved? Visit us from time to time



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

<http://www.ist-lobster.org/>

<http://www.ist-lobster.eu/>

LOBSTER > LOBSTER - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://www.ist-lobster.org/> Go Google

lobster Large-scale Monitoring of Broadband Internet Infrastructures

Information Society Technologies

[Home](#) [Contact](#)

About LOBSTER

Partners

Publications

Downloads

Announcements

Calendar & Events

Related Sites

Private Area

Welcome to LOBSTER!

LOBSTER is a pilot European Infrastructure for accurate Internet traffic monitoring. Based on passive monitoring, and capitalizing on our experience, the LOBSTER infrastructure will be unique in Europe and among the only two similar infrastructures that exist in the world today.

Thus, LOBSTER will make a significant contribution to "IST 2.3.5 Research Networking test-beds" whose objective is "to integrate and validate ... state-of-the-art technology that is essential for preparing the future upgrades in the infrastructure deployed across Europe".

In addition, LOBSTER will contribute to the overall IST Workprogramme, whose objective is to ensure European leadership in the generic and applied technologies", by enhancing the European leadership in network monitoring technologies.

LOBSTER is a successor of the [SCAMPI project](#).

Hot Topics

- NERD software available from the [Downloads section](#)

Latest News

- Published deliverable [D1.4 - Integrated Architecture Definition](#) - 29 September 2005
- Published deliverable [D1.2 - Common Access Platform Definition](#) - 25 July 2005
- Published deliverable [D1.1a - Anonymization Framework Definition](#) - 25 July 2005
- Published deliverable [D1.3 - First-tier Encryption Definition](#) - 21 July 2005

Internet



How can you get involved? Follow the LOBSTER activities



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

- Subscribe to LOBSTER news
 - <http://www.ist-lobster.org/announcements/>



The screenshot shows the 'Announcements' page of the LOBSTER website. At the top left is the 'lobster' logo and the text 'Large-scale Monitoring of Broadband Internet Infrastructures'. At the top right is the 'Information Society Technologies' logo. Below the header is a navigation bar with 'Home > Announcements' and a 'Contact' link. A left sidebar contains links for 'About LOBSTER', 'Partners', 'Publications', 'Downloads', 'Announcements' (highlighted), 'Calendar & Events', 'Related Sites', and 'Private Area'. The main content area is titled 'LOBSTER Announcements' and features a subscription link: 'Subscribe to lobster-news@ist-lobster.org'. Below this is a link to the 'LOBSTER Announcement archive' and a paragraph explaining the mailing list. A subscription form includes an '*E-mail:' field, radio buttons for 'Subscribe' and 'Unsubscribe', and a 'Submit' button. A note states '* denotes required field'. At the bottom of the page is a link to the 'LOBSTER Announcement mailing list policy'.



Hosted by [TERENA](#)

This page was last updated on 19 January 2005.

Evangelos Markatos, FORTH

info@ist-lobster.org



lobster

How can you get involved? Would you like to experiment with it? Install a passive monitor



<http://www.ist-lobster.eu>

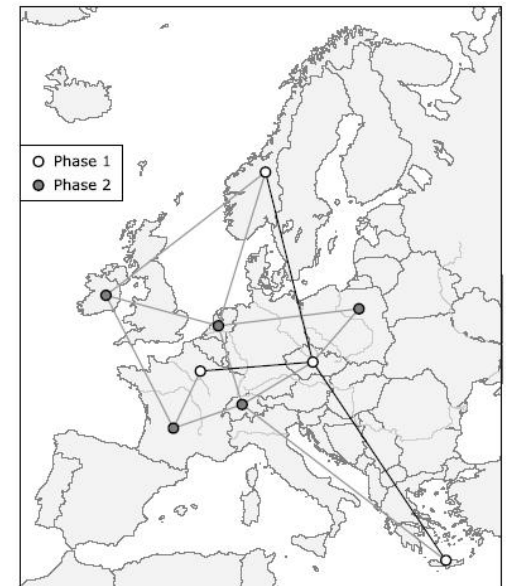
<http://www.ist-lobster.org>

- How?
 1. Get a regular PC
 2. Read the FAQ
 - <http://www.ist-lobster.org/about/faq.html>
 3. Start capturing packets
- Receive help and feedback from LOBSTER
- Join the MAPI mailing list
 - <http://mapi.uninett.no/>
- Use it for your own monitoring purposes



How can you get involved? Become a part of the LOBSTER team

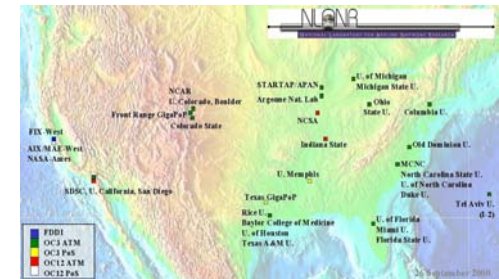
- Would you like to share data?
 - Even highly anonymized data?
- Stay tuned
 - LOBSTER phase two is coming up:
 - In July 2006 the LOBSTER infrastructure will start accepting members





Collaborations

- NRNs
 - Node installation
- GigaCampus
 - >100 PM sensors in Norway
- Far East
 - Shanghai Jiao Tong University (Mao Weihua)
 - Interested in network monitoring collaboration
 - and common proposals
 - Institute for Infocomm Research (Singapore)





Summary

- Networking GRAND Challenge:
 - “Monitor a day in the life of the Internet”
- LOBSTER/SCAMPI are about (Passive) Traffic Monitoring
- MAPI (Monitoring Application Programming Interface)
 - Write your own applications
- Concerned about privacy?
 - LOBSTER provides an anonymization infrastructure
 - All you need to do is answer the following question:
 - “How much data am I willing to share and with whom?”
- How can you get involved
 - Just browse www.ist-lobster.org from time to time
 - Join our email list
 - Install a passive monitor for your network
 - Link your monitor to the LOBSTER infrastructure



The LOBSTER project



<http://www.ist-lobster.eu>

<http://www.ist-lobster.org>

Network Monitoring for Performance and Security

The LOBSTER project: Current State and Collaboration Opportunities

Evangelos Markatos

Institute of Computer Science (ICS)

Foundation for Research and Technology – Hellas (FORTH)

Crete, Greece



lobster

<http://www.ist-lobster.eu>



Information Society
Technologies

<http://www.ist-lobster.org>

- Back up slides



LOBSTER partners

- **Research Organizations**
 - ICS-FORTH, Greece
 - Vrije University, The Netherlands
 - TNO Telecom, The Netherlands
- **NRNs/ISPs, Associations**
 - CESNET, Czech Republic
 - UNINETT, Norway
 - FORTHNET, Greece
 - TERENA, The Netherlands
- **Industrial Partners**
 - ALCATEL, France
 - Endace, UK





I have more than one Internet connection How can I monitor it?

- Install several passive monitors
- Would you like to see all of them as one?
 - Use DiMAPI (Distributed Version of MAPI)
- DiMAPI enables you to monitor several lines as if they were one:
- The SCOPE abstraction:
 - SCOPE is a set of lines (interfaces) to monitor
- `mapi_create_flow("host1:eth2, host2:/dev/dag0, host3:eth1");`
- Fully compatible with MAPI
 - All previous functions work as usual