

Network Data Repository for Researchers

*Saleem Bhatti, Tristan Henderson, Martin Bateman
School of Computer Science, University of St Andrews*



Computer Science
St. Andrews

Outline of talk

1. Outline of how to get started
2. What we have set up at St Andrews
3. CRAWDAD mirror
4. Questions
5. (Demo after talk)



Gathering network data

- Researchers are increasingly eager to test their ideas by taking measurements.
- UKERNA support for data capture becoming available.
- Sharing of data:
 - network traces, data used for papers.
 - reproducibility of results.
 - (applies to simulation studies too.)



Instrumentation and measurement

- Internet Measurement Masterclass in Cambridge, Dec 2006 - how to get started.
- However, it can take a long time to undertake:
 - sys/net admin interaction and agreement.
 - clarify legal aspects.
- Sys/net admins often reluctant to engage due to:
 - lack of previous experience of such activity.
 - insufficient knowledge of legal ramifications.



Local agreement

- Research Use of Network Traffic Data:
<http://www.ja.net/development/traffic-data/>
- Policy + SLA.
- MASTS project (UKLIGHT monitoring)
- Can be used locally:
 - local variant in use at St Andrews.
 - plus some additional guidelines agreed locally.
 - (will make available end of Aug 2007)



What you need to do

1. Define what data you need:
 - Unlikely to need full packet dumps.
2. Start talking to your *sys/net* admin staff:
 - Be patient ... these folk have a day job.
3. Ensure data users are briefed on DPA:
 - Ignorance is no defence.
4. Have a well-defined data distribution system:
 - Keep track of who uses data.



St Andrews agreement

- A modified version of the UKERNA document:
 - plus four Annexes
 1. List of users
 2. List of administrators
 3. Guidelines for producing publications
 4. Guidelines for using data

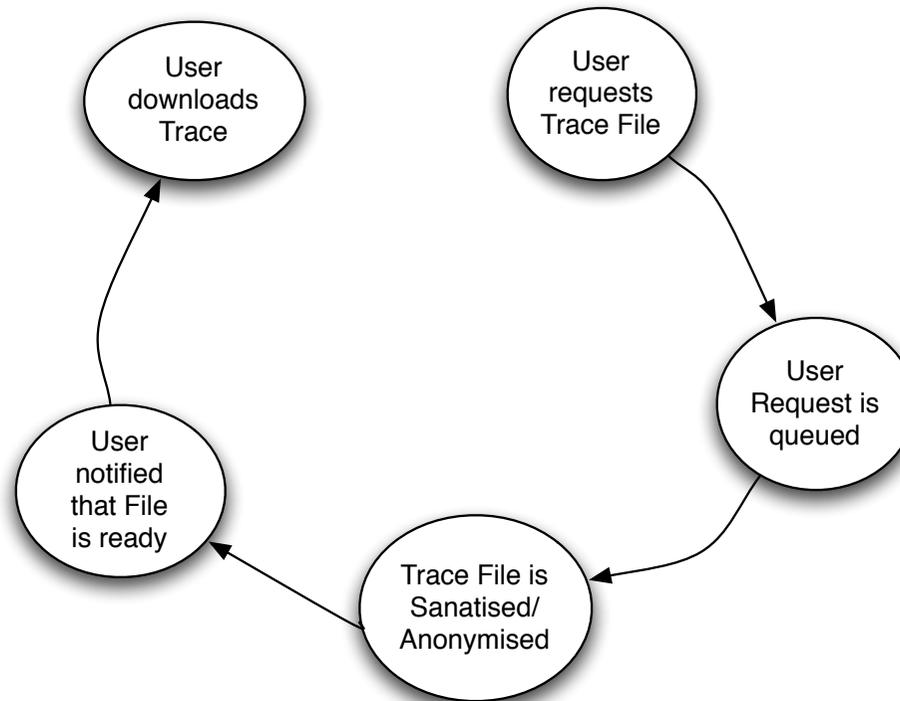


Outline of talk

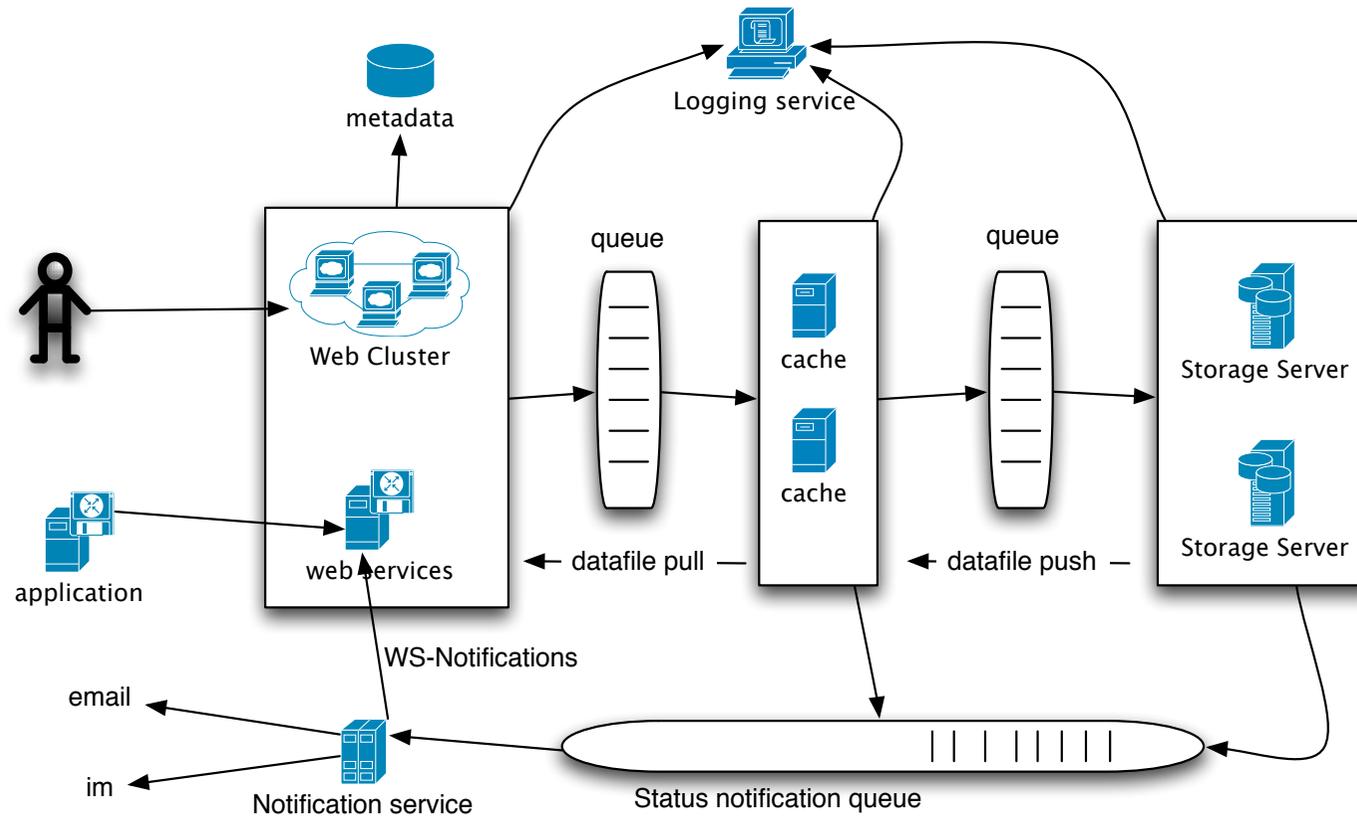
1. Outline of how to get started
2. Example of what we have set up
3. CRAWDAD mirror
4. Questions
5. (Demo after talk)



Process



Outline architecture



Outline of talk

1. Outline of how we got started
2. Outline of what we have set up
3. CRAWDAD mirror
4. Questions
5. (Demo - after talk)





RAW DAD Community Resource for Archiving Wireless Data At Dartmouth

<http://CRAWDAD.cs.dartmouth.edu>

Researchers in wireless networks need data to

- understand real network usage
- identify the real problems
- evaluate possible solutions

CRAWDAD is an NSF-funded project to build a wireless network data archive for the research community.

CRAWDAD provides

- Archive of wireless-network traces
802.11, MANET, VANET, Sensor network, DTN, location, etc.
- Tools for trace collection
collect, process, sanitize, analyze, etc
- HOWTO documents in wiki
- Support for the research community
Event calendar
Bibliography
Specialist groups (MANET, Education)
Annual Workshop (free!)

Currently (March 2007),

- 598 users from 365 institutions around the world
- 23 data sets
- 12 tools
- 70 papers
- Staff:
full-time programmer plus 2 undergrads

Contact: crawdad@cs.dartmouth.edu

Hosted by Center for Mobile Computing at Dartmouth College

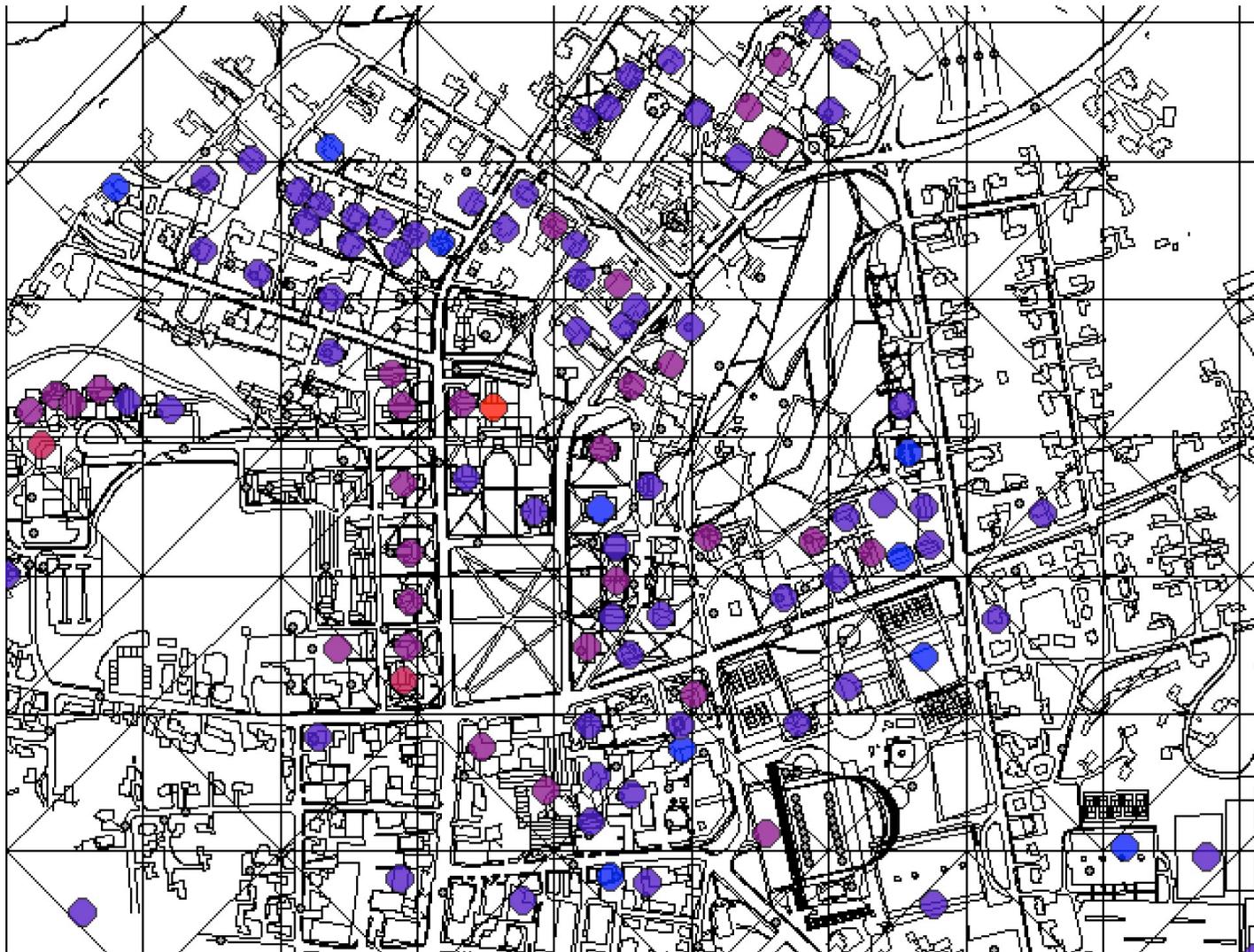


Computer Science
St. Andrews

Example: Dartmouth campus 802.11 WLAN

- Dartmouth campus, 500-1500 access points
- 6 years of mobility traces
 - location = access point
- 2 detailed (3-4 month) packet traces
 - layer 2-4
- Used by over 200 researchers
 - mobility modelling, DTN, security, routing





Daily WLAN Users
(April 11, 2007)

- 1 - 10
- 10 - 100
- 100 - 400
- 400 - 800
- 800 - 1200



Computer Science
St. Andrews

Questions, queries, comments ...



Outline of talk

1. Outline of how to get started
2. Example of what we have set up
3. CRAWDAD mirror
4. Questions
5. (Demo after talk)





To view this page, you need to log in to area "WNDS" on localhost:8080.

Home - Network Monitoring Repository

https://localhost:8080/ Google

d-scent RSS Feed Hopback RSS Feed Apple (76) News (826)

Home - Network Monitoring Respoitory

- Home
- Traces
- Download
- Admin
- Infomation
- Help



Network Traces

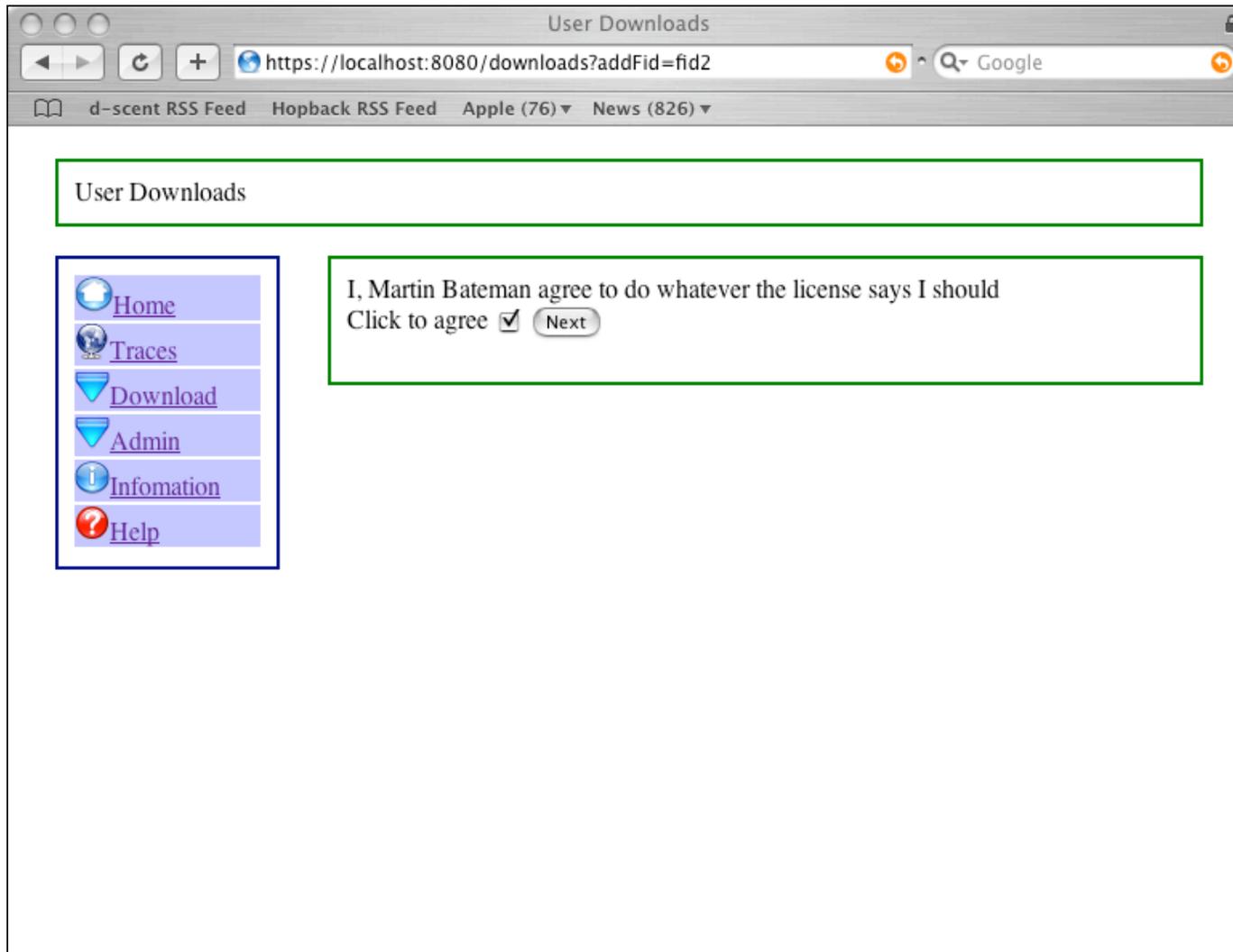
Home
Traces
Download
Admin
Information
Help

From: To:

Start	End	Description
2007-07-06 14:42:52.155	2007-07-06 14:42:52.155	This is trace data 1
2007-07-06 14:42:52.157	2007-07-06 14:42:52.157	This is trace data 2
2007-07-06 14:42:52.158	2007-07-06 14:42:52.158	This is trace data 3
2007-07-06 14:42:52.159	2007-07-06 14:42:52.159	This is trace data 4
2007-07-06 14:42:52.16	2007-07-06 14:42:52.16	This is trace data 5
2007-07-06 14:42:52.161	2007-07-06 14:42:52.161	This is trace data 6
2007-07-06 14:42:52.161	2007-07-06 14:42:52.161	This is trace data 7
2007-07-06 14:42:52.162	2007-07-06 14:42:52.162	This is trace data 8
2007-07-06 14:42:52.163	2007-07-06 14:42:52.163	This is trace data 9
2007-07-06 14:42:52.164	2007-07-06 14:42:52.164	This is trace data 10

[Next](#)





User Downloads

https://localhost:8080/downloads?agreed=on&addFid=fid2&requ Google

d-scent RSS Feed Hopback RSS Feed Apple (76) News (826)

User Downloads

- Home
- Traces
- Download
- Admin
- Infomation
- Help

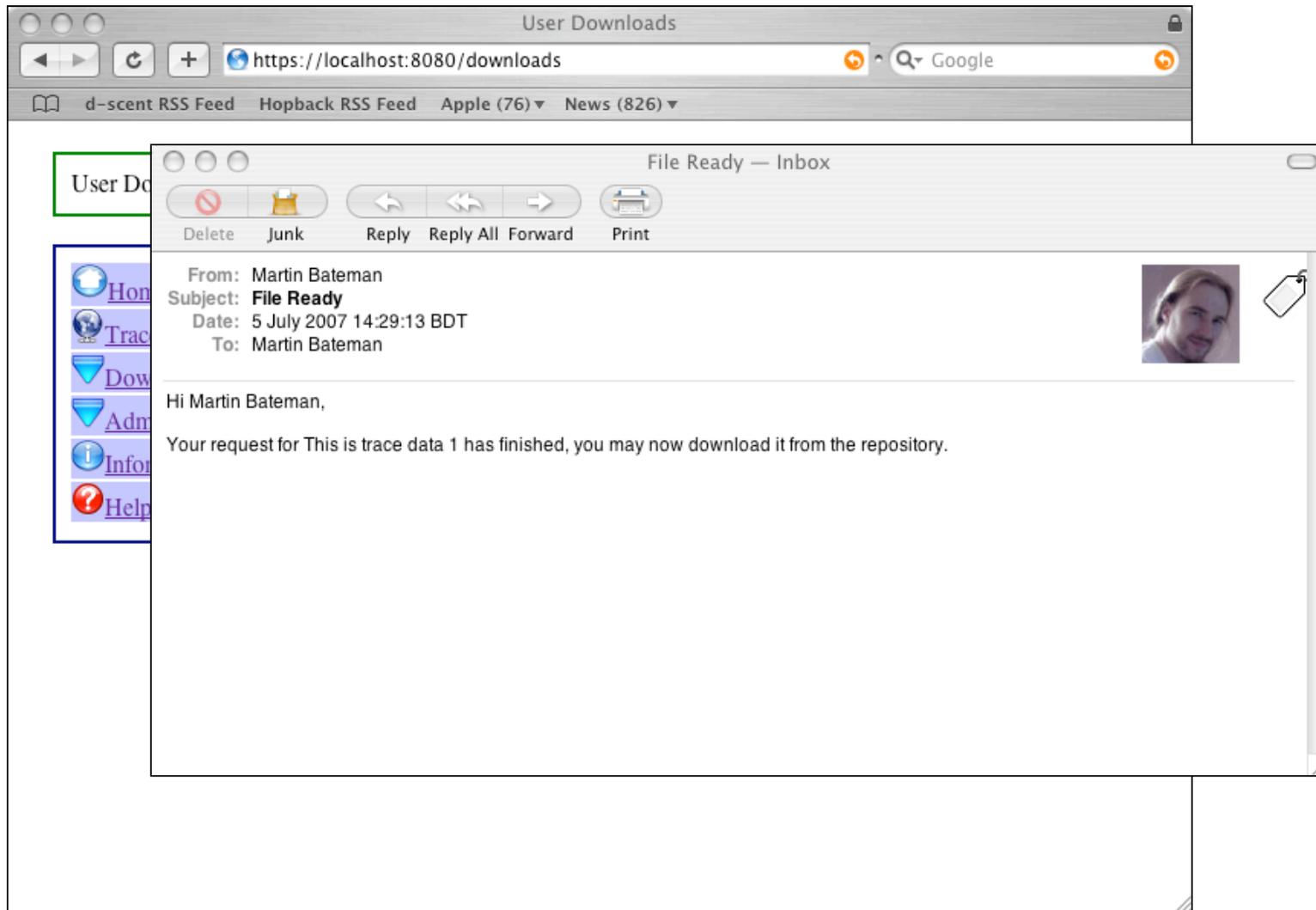
Trace id State

	fid2	Requested
--	------	-----------



The image shows a screenshot of a web browser window and an email client window. The browser window, titled "User Downloads", has the address bar set to "https://localhost:8080/downloads" and a search bar with "Google". Below the browser, the email client window "Anonymising Data - Inbox" is open. It features a toolbar with "Delete", "Junk", "Reply", "Reply All", "Forward", and "Print" buttons. The email header shows: "From: Martin Bateman", "Subject: Anonymising Data", "Date: 5 July 2007 14:29:12 BDT", and "To: Martin Bateman". A small profile picture of a man is visible next to the "To" field. The email body contains the text: "Hi Martin Bateman, Your request for This is trace data 1 is currently being anonymised." To the left of the email client is a sidebar with navigation links: "Home", "Traces", "Down", "Admin", "Info", and "Help".





License Agreement

License Agreement

- Home
- Traces
- Download
- Admin
- Information
- Help

I, Martin Bateman agree to do whatever the license says I should
Click to agree



The screenshot shows a web browser window with the title "User Downloads". The address bar contains "https://localhost:8080/downloads". The browser's search engine is set to Google. Below the address bar, there are several RSS feeds: "d-scent RSS Feed", "Hopback RSS Feed", "Apple (76)", and "News (826)".

The main content area is divided into two sections:

- User Downloads**: A header section.
- Navigation Menu**: A vertical list of links with icons:
 - Home (Globe icon)
 - Traces (Globe icon)
 - Download (Download icon)
 - Admin (Download icon)
 - Infomation (Information icon)
 - Help (Question mark icon)
- Trace id State**: A table showing the status of a download:

Trace id	State
fid2	Downloaded

