

B345 Internet Science and Technology

Week 11 lecture 1

Today's Lecture Learning Objective

- Understand the concept and importance of measurements for the web.

Why Measure?

- For content creators
- For web hosting companies
- For network operators
- For web/networking researchers

Content Creators

- User browsing patterns on web site.
- How users arrive.
- Where users go.
- Users' hardware/software resources.

Web Hosting Companies

- Billing.
- Allocate resources to sites.
- Selecting hardware/software.
- Performance refinement.

Network Operators

- Performance - caching.
- Allocate resources.
- Selecting hardware/software.
- Controlling Traffic Routing.

Web/Networking Researchers

- Evaluating performance of protocols.
- Determining what to include in protocols.
- Algorithms (eg. for path finding and traffic)

Measurement Techniques

- Server logs
- Client logs
- Proxy logs
- Packet monitoring
- Active measurements
- User-centric measurements

- Details next lecture...

Processing Measurement Data

- Parsing
- Filtering
- Transforming
- Analysing

- Details next lecture...

Characterizing Web Workloads

- From measurement data
- Workload models for characteristics:
 - HTTP messages
 - Resources
 - User behaviour
- More next lecture...

Using Workload Models

- Identifying performance problems
- Benchmarking web software and hardware
- Capacity planning

Standards in Measurement Jargon

- Australian Internet Industry Initiative (III) guidelines.
 - See required reading.
- The word "hit" very ambiguous - should NOT use.
- Some of the definitions not in line with common computing usage - so pay attention.

Measurement Jargon

- What is a "USER"?
 - Unique Audience
 - Unique Registration
 - Unique Browser
 - Unique IP
 - Unique Host

Measurement Jargon

- Page Request
- Page Download
- Site Session
- Visit
- Internet Session
- ISP Session

- Other activities to be considered later.