A Couple of Web Servers, a Small Staff, Thousands of Users, and Millions of Web Pages...

How We Manage (Sort of)

MIT jag@mit.edu

Web Publishing at MIT

- **Six centrally supported web servers (web, map, tute, webx, search, counter)**
- **2,500 groups and over 20,000 users**
- **300,000 URLs from main web server alone**

- **600,000+ URLs on 1000 additional web servers that leverage central services**

Agenda

- **What does the infrastructure look like?**
- **What web tools are provided to web publishers?**
- **How do we maintain and support this?**

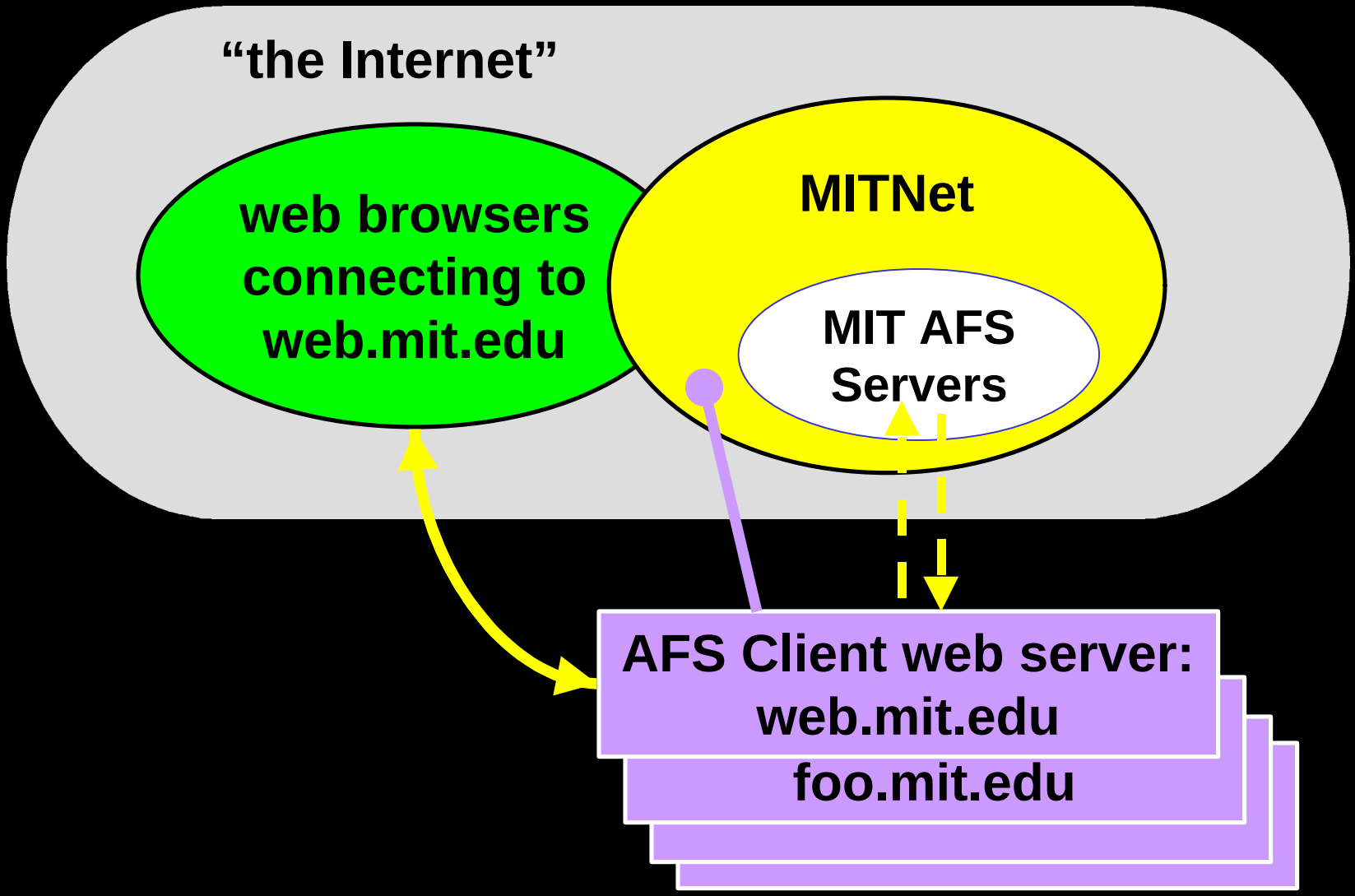
MIT Web Infrastructure

- **Leverage existing distributed computing environment (AFS)**
- **Allow web server to read AFS**
  - Store web pages in central AFS “lockers”
  - Main web servers read AFS and serve web pages
- **Content management is left completely to the content owners**

Why use AFS?

- **Designed for larger networks**
- **User authentication and security through Kerberos**
- **Permissions set at directory level**
- **Supported 24x7 at MIT**
- **Reliable backup systems**
- **Scalable solution**

AFS Client Web Server



Web Server Configuration

- **Apache SSL**
  MIT build: http://web.mit.edu/apache-ssl
- **Suns running Solaris with AFS client**
- **Web server configuration files sit on local disk; content comes out of AFS**

- **Root on web server gives you no privileges on any other system**

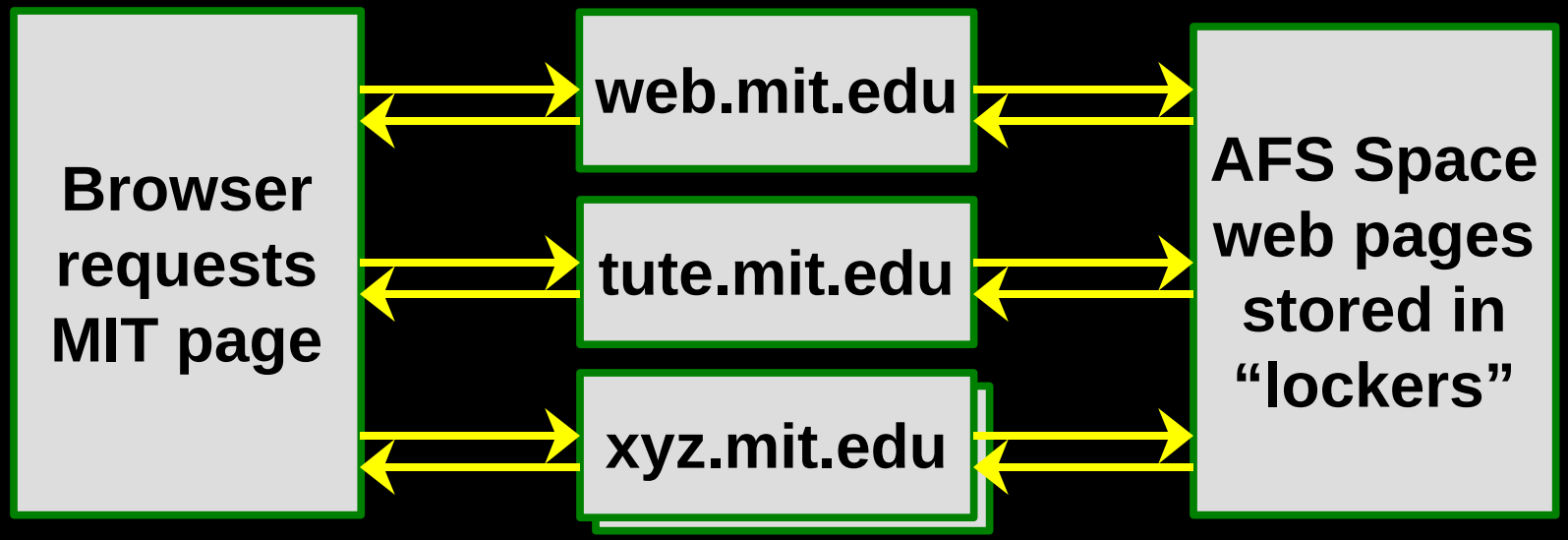
Stable platform for web pages

- **Users create html files with programs in AFS or on desktop**
- **Save or transfer files to location (“locker”) where they have write access**
- **See HTML page from http://web.mit.edu/locker/foo.html**

Add Web Tools (1994-97)

- **Email forms (cgiemail, cgiecho)**
- **Image map support**
- **Restricted access (MITNet)**
- **Search engine (Harvest)**
- **Recommended WYSIWYG web editor (Claris Homepage)**
- **Campus map**
- **Certificate-based authentication**
- **Web publishing training**

Web Topology & Tools (1997)



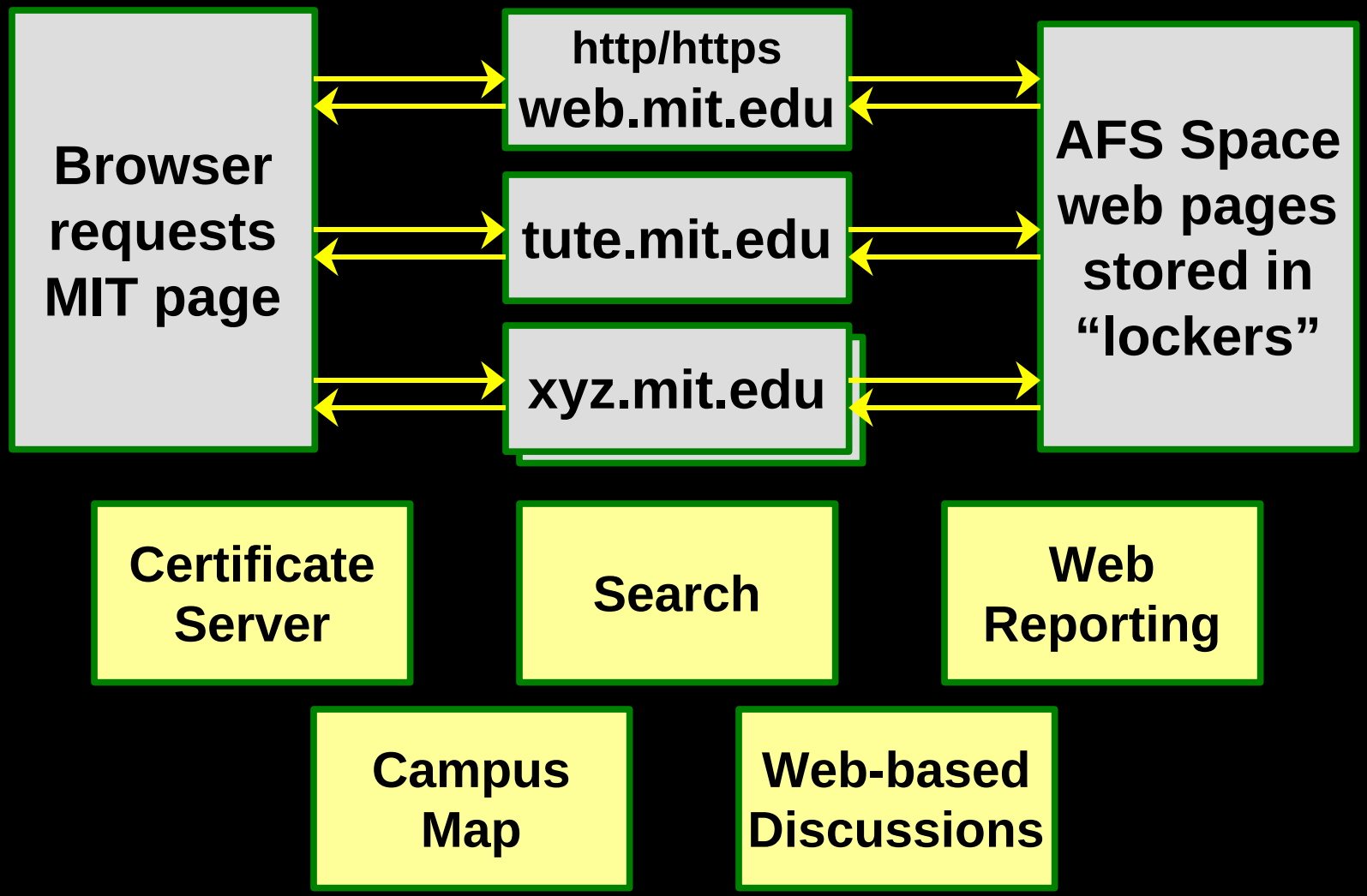
Certificate Server

Campus Map

Add More Web Tools (1998-99)

- **Discussion groups (WebCrossing)**
- **Search engine (Ultraseek)**
- **Web page usage statistics**
- **Server side includes**
- **Restricted access for individuals or lists**
- **Secure file transfer**
- **Recommended WYSIWYG web editor (Dreamweaver)**

Web Topology & Tools (1999)



Not all things...

- **Encourage groups to use as much of central web services as possible**
- **Recommend Apache-SSL for servers**
- **Recommend MIT-wide Integration guidelines**
- **Recommend vendors outside MIT as needed**
- **Can have server run by Administrative Server Services**

Use of Central Web Services

- **http://libraries.mit.edu**
  - Existing content in AFS
  - Additional content added, read by web server
  - Uses central search engine
- **http://mitsloan.mit.edu**
  - Web server reads files out of AFS
  - Uses main search engine with custom front end
  - Created custom Sloan “people finder”
- **http://student.mit.edu (WebSIS)**
  - Use of certificates to authenticate and encrypt

IS Organization

- **Information Systems (IS) focus is on "commons" services**
- **Partner with customers and vendors for solutions**
- **Processes: Discovery, Delivery, Integration, Service, Support**
- **Specific IS groups support specific web publishing functions**

Who supports infrastructure?

- **Three service groups to maintain web and AFS servers**
- **Three support groups for web publishing and training**

Server Service Groups

- **Network Operations maintains**
  - central MIT web servers, including web.mit.edu, search, counter…
  - Underlying services (DNS, NTP, Certificate server) for all web servers
- **Athena Operations maintains AFS**
  - Servers where content is stored, backups
  - Services for users of workstations
- **Administrative Server Services**
  - Maintain private servers for customers to create custom programs and content

Web Publishing Support

- **Faculty Liaisons support academic computing needs for courses and computing facilities**
- **Campus Wide Information Systems (CWIS) supports departments, labs, and organizations**
- **IS Training provides software and web publishing training**

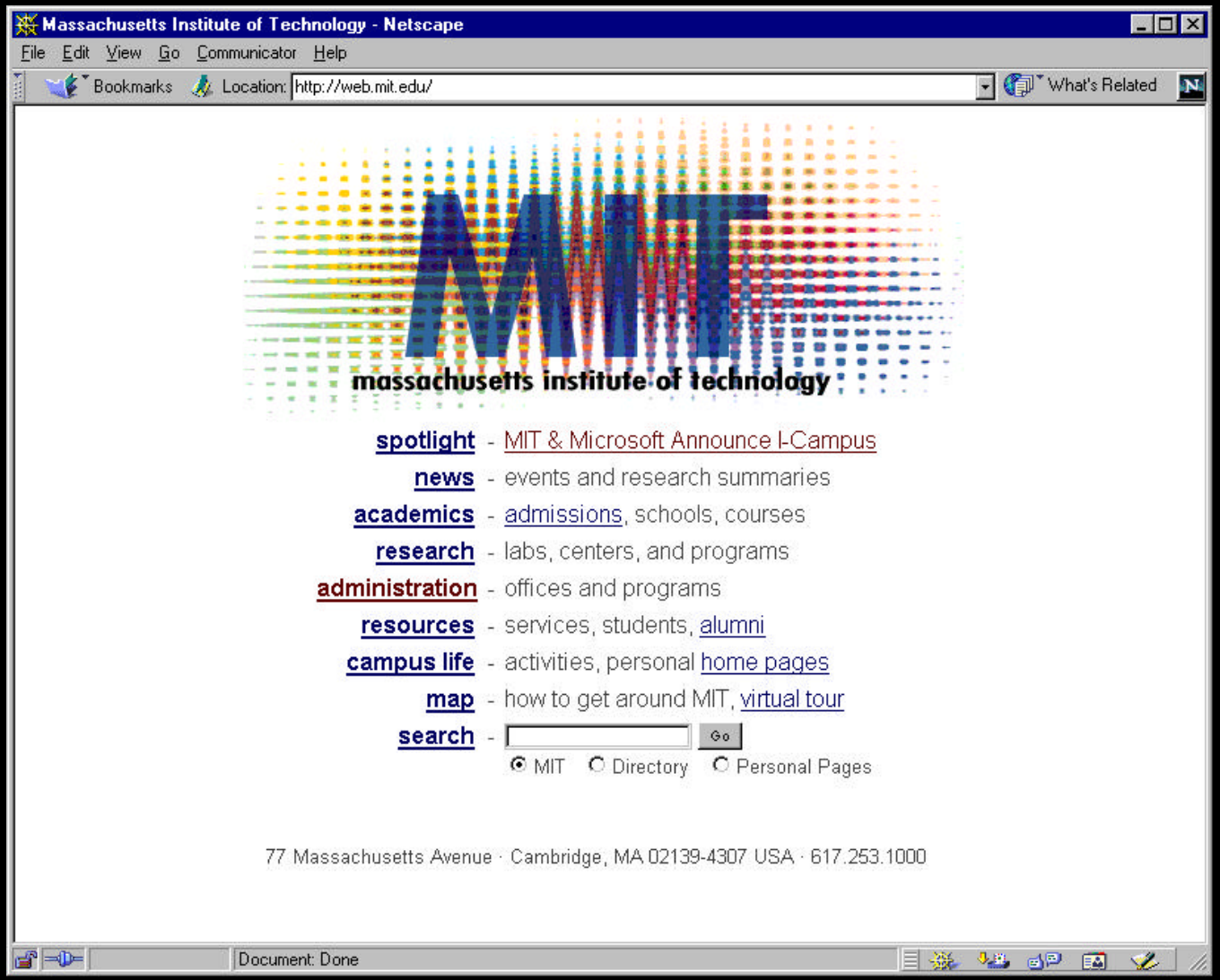
Relationships

- **Web Publishing at MIT works because of relationships between the service and support groups**
- **Client needs are heard through ongoing outreach**
- **Promote guidelines without being “police”**
- **http://web.mit.edu/guidelines.html**

What We're Working On

- **Institute-wide Events Calendar**
- **Secure credit card transactions**
- **Module-based publishing (XML?)**
- **Database-generated content**
- **Better "indexing" of web content**
- **Portals, the next generation**

MIT “Enterprise” Portal



Lessons Learned

- **Ability to scale for growth**
- **AFS + 1 web server not enough**
- **Special purpose web servers required**
- **What users see is what they want**
- **Influence through guidelines and relationships**

Q & A

- **This presentation is available from http://web.mit.edu/cwis/lisa/**
- **Campus Wide Information Systems http://web.mit.edu/cwis/**
- **Network Operations http://web.mit.edu/network/**
- **Transarc's AFS http://www.transarc.com/Product/EFS/**

- **jag@mit.edu, salemme@mit.edu**