Iowa Research Online
The University of Iowa's Institutional Repository

University of Iowa Libraries Staff Publications

12-4-1997

Managing a Cluster of Single-Purpose Web Servers: Administrative and Technical Issues

Paul A. Soderdahl
University of Iowa

Copyright 1997 Paul Soderdahl.

Hosted by Iowa Research Online. For more information please contact: lib-ir@uiowa.edu.

Managing a Cluster of Single-Purpose Web Servers: Administrative and Technical Issues

The University of Iowa Libraries-Wide Information System

While the preferred model for a web server at many institutions is to install a single server for the entire organization, the University of Iowa Libraries uses a cluster of single-purpose web servers across a number of different platforms in support of the Libraries' web project, known as the Libraries-Wide Information System (LWIS).

Employing this type of solution, the University of Iowa Libraries currently provides information and services on 13 distinct web servers, with 10 additional servers either in development or under consideration.

This hybrid solution offers a number of advantages over a single-server solution. First, such an implementation allows more content providers to deliver information in its existing format. By taking advantage of commercially available web server solutions for desktop applications, individual content providers can deliver their information over the web using a web-based interface without any modification to the original data or change in their workflow. This is part of the reason that the Libraries now has over 60 staff members delivering content over the web.

In addition, this solution allows the Libraries to offer new services quickly. As more desktop applications integrate a web server component, these applications can be brought online easily and inexpensively, obviating the need for custom programming or scripting. As a result, the web project can provide a large number of distinct individually inexpensive services, allowing for manageable growth. The system becomes very scalable, and can be sized up or down as needed, without any significant overhead in system infrastructure.

At the same time, this approach requires more coordination by the systems administration team. Because there is an underlying assumption that one size does not fit all, different servers, customized and optimized for individual applications, need to be administered collectively by a coordinated systems team.

Managing a Cluster of Single-Purpose Web Servers: Administrative and Technical Issues

The University of Iowa Libraries-Wide Information System

Information Servers

- LWIS-1, LWIS-2, LWIS-3 — Libraries' home page and general departmental information and services
- LWISTST — Test server where staff have accounts for posting departmental information
- ~LIB — Part of the University's central web server; used for Unix-based cgi's

Project Sites

- STAFFWEB — Web server for individual staff web pages
- TWIST — Web server for hosting sites related to the TWIST Project
- FIRSTCLASS — Conferencing server hosting conferences for the TWIST Project
- SDRCWEB — Web server for hosting the Scholarly Digital Resources Center
- BAILIWICK — Web server for hosting special topics sites created by Libraries staff

Forms Processing

- NETFORMS — Server dedicated to sophisticated handling of online forms

Database Interfacing

- TANGO — Web server interface to FileMaker Pro databases
- WEBZ — Implementation of OCLC's CIC Enhanced WebZ for Z39.50 & ILL
- WEBZTST — Development server for implementing OCLC's CIC Enhanced WebZ
- WEBSEARCH — Open Text's Livelink Search & Spider for web searching and SGML text
- INMAGIC — Web server interface to Inmagic databases
- TANGONT — Web server interface to Microsoft Access databases

Online Calendaring

- NUTD — Web server interface to Now Up-To-Date calendars
- ONTIME — Web server interface to OnTime calendars

System Administration

- MAIL — Mail server for virtual and real e-mail addresses for projects and departments
- PROXY — Proxy web server to allow remote access to IP-restricted databases
- BOLERO — Web server interface to real-time web stats using Butler SQL database
- SENTRY — Dedicated machine to monitor all other servers making sure they're online

CAUSE '97, Orlando
Paul A. Soderdahl, University of Iowa Libraries
http://staffweb.lib.uiowa.edu/psoderdahl/cause97.html

Libraries-Wide Information System

Information Servers

These servers are typical web servers, with thousands of HTML files, including a general home page for the University of Iowa Libraries and home pages for each of the individual departmental units within the Libraries.

The production servers are three identical machines, running simultaneously. This configuration, known as a RAIC (redundant array of inexpensive computers) provides several advantages over a single, more expensive server solution, including:

- minimized downtime
- robust maintenance
- smooth upscaling
- controlled testing

Libraries staff have accounts on a similarly configured, but separate, development server. Files are moved into production either ad hoc or on a timed basis.

Project Sites

These servers are standard web servers, but generally dedicated to a specific project.

Because any one individual project is generally not high traffic and may have a dozen or fewer authors and editors, an inexpensive desktop machine may be sufficient. With a web server infrastructure that can accommodate multiple single-purpose servers, individual projects of modest scope can warrant dedicated web servers at very little cost.

Among other advantages, this gives the project its own host name, resulting in a URL that can be easily remembered.

the university of iowa libraries

Libraries-Wide Information System

Forms Processing

Used widely in electronic commerce, online forms are often underutilized in educational institutions, with the exception of comments pages, which are, in effect, simple e-mail messages disguised as web forms.

By using commercially available forms processing software on a dedicated server, the University of Iowa Libraries supports a sophisticated but easy-to-use forms processing system.

Any staff who can write in HTML can quickly learn the HTML-like tags for advanced forms processing that can perform any combination of the following:

- send a formatted e-mail message anywhere
- create a customized response
- create an HTML page on-the-fly and store it to the web server
- append to a tab-delimited file to create spreadsheets on-the-fly

CAUSE '97, Minneapolis
Paul A. Soderdahl, University of Iowa Libraries
http://staffweb.lib.uiowa.edu/psoderdahl/cause97.html

Libraries-Wide Information System

Database Interfacing

While static HTML pages are useful for delivering certain types of information, a web server can become much more powerful when it serves data from a back-end database and builds the user-friendly HTML page on-the-fly.

This gives content providers who maintain the data the ability to use a database management system best suited for given content. The information can then be delivered directly to the user from its original data source.

By taking advantage of commercially available web server solutions for desktop database applications such as FileMaker Pro, Inmagic, Microsoft Access, or for any SQL databases, individual content providers can deliver their databases over the web using a web-based search interface without any modification to the original data or change in their workflow.

Libraries-Wide Information System

Online Calendaring

This is just one example of the way that more and more desktop and network applications are beginning to include a web server built right into the application. Thus, with minimal setup required, the application itself is configured to run as a web server, and any user can interact with the application through a web interface.

The University of Iowa Libraries is using this feature built into Now Up-To-Date calendar software and has a dedicated web server for Now Up-To-Date calendars. We are beginning to experiment with other calendaring software applications.

With a web server infrastructure that can accommodate these application-specific web servers, more applications can be brought online easily and inexpensively, obviating the need for custom programming or scripting.

Libraries-Wide Information System

System Administration

The University of Iowa Libraries currently supports four servers that enhance the Libraries' web project, but do not provide content directly to end users:

- a mail server that allows Libraries' staff and other content providers to create dummy e-mail addresses for any project, large or small
- a proxy server that allows authorized users off-campus to access IP-restricted databases
- a stats processing server that records all transactions from the other web servers into a relational database, providing content providers with the ability to perform summary statistics on-the-fly
- a sentry server that monitors all of the other servers in the system and can automatically reboot machines that go down

CAUSE '97, Minneapolis
Paul A. Soderdahl, University of Iowa Libraries
http://staffweb.lib.uiowa.edu/psoderdahl/cause97.html

ABSTRACT

While the preferred model for a web server at many institutions is to install, maintain, and support a single server for the entire organization, the University of Iowa Libraries uses an alternate model providing a number of different web servers across a number of different platforms in support of the Libraries' web project. This hybrid solution offers several advantages over a single-server solution. Such an implementation allows more content providers to deliver information in its existing format, and also allows the Libraries to offer new services quickly as more desktop applications integrate a web server component. Employing this type of solution, the University of Iowa Libraries currently provides information and services on 10 distinct web servers, with five additional servers either in development or under consideration. The solution enables the Libraries to embrace the web for delivery of important information and services to its clientele.

OUTLINE

I. Infrastructure
  A. Traditional Infrastructure – single server solution
    1. WindowsNT – Microsoft, O'Reilly, Netscape
    2. UNIX – Netscape, Apache, NCSA httpd
  B. University of Iowa Libraries – hybrid solution
    http://www.lib.uiowa.edu/lwis/infrastructure.html
    1. Regular HTML files – production server
      a. RAIC (redundant array of inexpensive computers)
        http://www1.arcade.uiowa.edu/
        http://www2.arcade.uiowa.edu/
        http://www3.arcade.uiowa.edu/
      b. IBM RS/6000
        http://www.uiowa.edu/~lib/
    2. Development server
      a. LWISTST
        http://www4.lib.uiowa.edu:8080/
    3. Forms processing server
      a. NetForms
        http://netforms.lib.uiowa.edu:8001/
    4. Calendaring software
      a. Now Up-to-Date
        http://now.lib.uiowa.edu:8002/
      b. On Time
        under development
    5. Database backends
      a. FileMaker Pro / Tango
        http://tango.lib.uiowa.edu:8003/
      b. Microsoft Access
        under development

c. OCLC CIC-Enhanced WebZ
http://webz.lib.uiowa.edu:4000/
d. Inmagic
under consideration

6. Full-text searching
a. OpenText Livelink Search and Spider
under development

7. Special projects
a. Teaching With Innovative Style and Technology
http://twist.lib.uiowa.edu/
b. StaffWeb
under development

II. Advantages
A. Use data in its existing form
1. Bibliographies in FileMaker Pro and Access
2. Online Public Access Catalog (NOTIS)
3. Calendaring software
4. Dynamic conversion to HTML
B. Empower content providers
1. Interfaces to desktop applications with no need to learn scripting
2. Ability to create sophisticated forms without scripting
3. Ability to create sophisticated search interfaces without scripting
C. Cost-effectiveness
1. Make use of commercially available software
2. Minimize special-purpose programming
3. Minimize development time for programming/scripting
4. Controlled growth for upsizing/downsizing

III. Issues
A. Change in server management
1. Increase in management required
2. Decrease in system-level programming, perl scripting needed
B. Expertise widely distributed
1. Content providers more empowered
2. Create local pools of expertise in various applications
3. Administration centralized, expertise decentralized

IV. Likely Outcomes
A. Balancing centralization and autonomy
B. Most appropriate solution for large, diversified institutions
C. Enable content providers to adopt web publishing across more situations

---

A/V REQUIREMENTS

1. Overhead projector and screen
2. Small table
3. Easel or bulletin board