<table>
<thead>
<tr>
<th>Title</th>
<th>Case studies in internet publishing.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Josol, Regino Joel.</td>
</tr>
<tr>
<td>Citation</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>2000</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10220/3226">http://hdl.handle.net/10220/3226</a></td>
</tr>
<tr>
<td>Rights</td>
<td></td>
</tr>
</tbody>
</table>
Case Studies
In Internet Publishing
regino joel josol
business development manager
joel.josol@compaq.com

Is this your web site?
Web Site is not available.
We were overwhelmed by your interest.
What the Vatican Wanted To Do

- Store and serve a large number of documents
  - More than 300 MB of documents on the site today
  - Includes documents from the Roman Curia and 41 Dicasteri
- Support multiple languages:
  - Let users navigate the Web site in a choice of six languages
  - Offer all documents and speeches available in the original language (it can be any language used by the Pope in his travels) and in several translations (some documents have up to eight language versions)
- Manage news feeds from multiple sources, including:
  - Osservatore Romano (daily feed)
  - Sala Stampa (multiple daily feeds)
  - Holy Father office (on-spot feeds)
- Provide information on the Vatican Museums

Plans for Jubilee Year 2000

- Real-time worldwide broadcasting of:
  - Opening of the Holy Door
  - Urbi et Orbi blessing
  - Popes’ Speeches
- Online access to:
  - All papal publications since Trento Council (1545)
  - Osservatore Romano in six languages
  - Documents from “Secret Archive”
- Vatican Museum on line (an audio/video tour of the museum)
Other Vatican Plans

- Fill the 384MB up link offered by Telecom Italia
- Commercial services
  - Pilgrims' kiosk
  - Electronic wallet
  - Travel reservations
  - Room reservations
  - Happenings
  - Virtual library
  - Stamps
  - Coins
  - Books

Technical Issues

- To reduce the load on the Web servers, the majority of document preprocessing is done at the Virtual Press Room level.
- The Web site uses two separate Internet service providers to provide a link to the Internet, in case something goes wrong with one of the two providers.
Technical Requirements

- Allowed for growth of the Web site to support the increasing load
- Resistant to hardware and software failures
- All devices have redundant configurations, including power supplies, network cables, firewalls, and Web servers
  - For example, the security is guaranteed by a three-level firewall scheme, which includes a firewall cluster between each of the following:
    - The Internet and the Web servers
    - The Web servers and the Virtual Press Room/Internet Office
    - The Internet Office and the rest of the Vatican internal network
Initial Experience, Easter 1997

- Web site is opened after John Paul II's Easter Angelus speech

- In first three days:
  - 3 million hits
  - 14.5 GB are transferred over Internet
  - 2 Mbit per second line is busy at 1.5 Mbits per second

- Security Attacks
  - 20 minor attacks a day
  - 4 clever attacks a week
  - Recently, the site survives a 3-day sustained hacker attack
Vatican Survives 3-Day Hacking!

Washington Post’s Requirements:
High-End Web Publishing

- The Web site, as you’d imagine, had to handle frequent hourly updates
  - With this live Web site, the Post wanted to compete with the immediacy of local Washington, D.C. radio and TV news coverage
  - The Web site would include a version of each day’s printed newspaper
- The Web site would also offer Associated Press (AP) Online feeds, reference information, and hyperlinks to related information
- Readers should be able
  - to search for articles that matched the strings they input.
  - to submit letters, feedback on editorials, or movie reviews, or restaurant reviews, et cetera.
- To support revenue goals from advertisers, the server needed to track site and page access statistics and user demographic data.
Washington Post’s Main Goal

- To design and implement an editorial environment that provides
  - a robust, flexible, scalable, and manageable solution
  - which will allow them to become recognized as the premier newspaper industry’s service on the World Wide Web
- Content to draw traffic using information related to the Washington metropolitan area’s
  - Businesses
  - Arts
  - Entertainment
  - Sports
  - Real Estate
  - Job Market

Time Lines (Oct. 1995-June 1996)

- **Election Line project**
  - Project start, end of October, 1995
  - Design work during November, 1995
  - First implementation work December, 1995
  - First web site online, January, 1996
  - **Election Line**, ongoing tweaks through May, 1996

- **WashingtonPost.com project**
  - March, 1996, work starts on the Post
  - April, 1996, first prototypes on the Web
  - May, 1996, Web opened to Interchange users
  - June, 1996, doors fully open to the Post web site
  - July, 1996, project completes
  - October, 1996, web site runs out of CPU
Issues

- Experience issues in
  - Publishing
    - using NaviPress, HotDog Pro32 and WebEdit
  - Newsroom style of operation
    - Story Creation
    - Editing/Selection of Stories
    - Copy Editing
    - Release to Public Web Server
  - the Web
    - PLWeb search engine, AltaVista Forum for discussions
  - the Internet
    - scripting using PythonOO script

- Security Issues
  - Implemented 3-tier firewall scheme

Web Site Building Approach

- Agreed Upon Processes
  - Defining and creating a Web site hierarchy for the staging area, and mirrored on the public Web site
  - Defining where content resides
  - Determining how content is reviewed
  - Agreeing on who can release content

- Implementation Philosophy
  - Create all the content behind a firewall
  - Use a staging server
  - Release finalized content to the public Web site
  - Avoid the need to manage the public Web site content, live.
  - The public site is a reflection of the internal staging server
Washington Post Content Creation

Public Internet

Private Intranet

Public Web Server

Staging and Editorial Server

Firewall

Washington Post Environment

From Back End to the Web Site

This entire process can be bypassed to allow late-breaking news posting.
Summary

- Web publishing is a system of creating, managing, securing and distributing content.
- Requires robust, secure, scaleable servers that can handle the work load and volume of content.
  - Vatican web site took 3M hits in first days.
  - Vatican web site defended itself successfully in a 3-day hacker attack.
  - Washington Post stored 10 years worth of back issues.

- Compaq is your key technology partner.
  - Case studies were running on Tru64 UNIX on AlphaServers.
  - Content were prepared on either Linux or Windows 95 tools.