



Present Can ~~Future~~ Web Technologies Profit to Accelerator Controls ?

An overview of present and future Web technologies (at CERN)

Alberto Pace - March 2001

alberto.pace@cern.ch

<http://cern.ch/alberto.pace>

CERN
Web Services



Using Web Technologies for File Storage and Application development

An overview of present and future Web technologies (at CERN)

Alberto Pace - March 2001

alberto.pace@cern.ch

<http://cern.ch/alberto.pace>

CERN
Web Services



Agenda

- ◆ CERN Web Architecture
- ◆ Services for End-Users
- ◆ Web Storage
- ◆ Web Technologies for Application Development

DESY Zeuthen Computing Seminar – March 2001



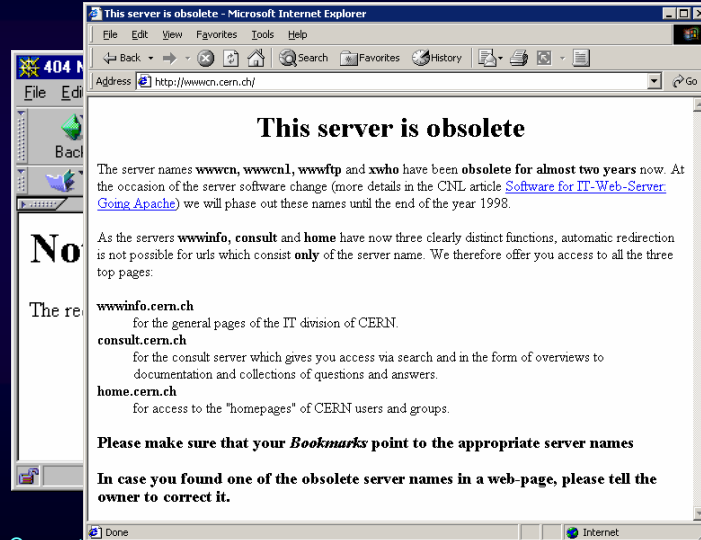
Yesterday's namespace

- ◆ [http://www.cern.ch/...](http://www.cern.ch/)
- ◆ [http://wwwinfo.cern.ch/...](http://wwwinfo.cern.ch/)
- ◆ [http://network.cern.ch/...](http://network.cern.ch/)
- ◆ [http://home.cern.ch/...](http://home.cern.ch/)
- ◆ [http://nicewww.cern.ch/...](http://nicewww.cern.ch/)
- ◆ [http://wwwas.cern.ch/...](http://wwwas.cern.ch/)
- ◆ [http://wwwlhc.cern.ch/...](http://wwwlhc.cern.ch/)
- ◆ ...

DESY Zeuthen Computing Seminar – March 2001



Was there a problem ?

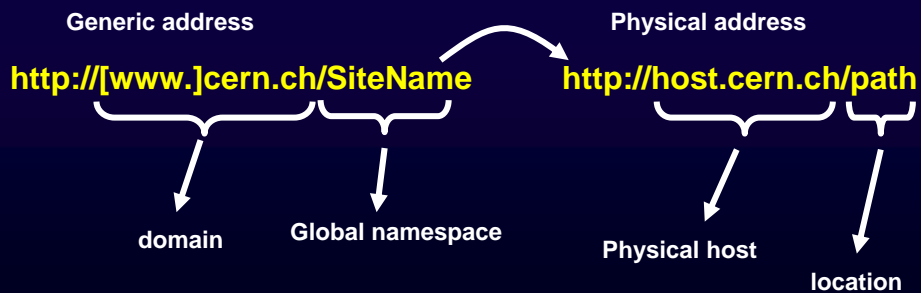


DESY Zeuthen Computing Seminar – March 2001



The solution ...

**A Database of Web sites,
A Unique Web namespace for CERN,
A Translation / Redirection Service ...**



DESY Zeuthen Computing Seminar – March 2001



Similar to the MAIL solution

Translation Service

Generic address

Physical address

First.last@cern.ch

user@host.cern.ch

http://cern.ch/SiteName

http://host.cern.ch/site

DESY Zeuthen Computing Seminar – March 2001



The redirection is recursive

- ◆ Once a 'SiteName' is registered, the redirection works for all subdocuments

Example:

`http://cern.ch/SiteName` -> `http://myhost.cern.ch/`

`http://cern.ch/SiteName/data/a.html` -> `http://myhost.cern.ch/data/a.html`

DESY Zeuthen Computing Seminar – March 2001



Advantages

- ◆ Sites can be reorganized and arranged
 - ◆ Possible migration from local servers to central servers
 - ◆ Possible migration from central servers to local servers

But not all problems solved !

DESY Zeuthen Computing Seminar – March 2001



More on the MAIL architecture

- ◆ The GENERIC address is useful for the Mail DELIVERY only

First.last@cern.ch

 - ◆ Unfortunately, “Mail Composing” tools are unable to resolve the generic address to the physical address when connecting to the user’s mailbox
 - ◆ Therefore, for the central mail services, given a user’s mailbox, the mail host can be found using the ‘mailbox’ DNS domain

user@user.mailbox.cern.ch

DESY Zeuthen Computing Seminar – March 2001



The Web has the same problem

- ◆ The **GENERIC** address is useful for HTML reading only

`http://cern.ch/SiteName`

- ◆ Unfortunately, “Web Authoring” tools are unable to gain read/write access from the generic address
- ◆ Therefore, we should also register (for the author’s use only) a web host alias that can be found using the ‘web’ DNS domain

`http://sitename.web.cern.ch/SiteName`

DESY Zeuthen Computing Seminar – March 2001



Not limited to the central servers

- ◆ In the mail architecture, the **Generic Address** (first.last@cern.ch) can point to the the **central web servers** (xxxx@mail.cern.ch) or to **locally managed servers** (xxxx@dxcoms.cern.ch) or even to servers outside CERN (xxxx@fnal.gov)
- ◆ Very similar for the web:
`http://cern.ch/xxxx`
`http://xxxx.web.cern.ch/xxxx`
can point to the the **central web servers** or to any **locally managed server**

DESY Zeuthen Computing Seminar – March 2001



Example

To READ only	READ and WRITE	Physical address (Never used)
http://cern.ch/pcdesktop	http://pcdesktop.web.cern.ch/pcdesktop	http://web2.cern.ch/pcdesktop
http://cern.ch/Physics	http://physics.web.cern.ch/Physics	http://web1.cern.ch/Physics
http://cern.ch/library	http://library.web.cern.ch/library	http://alice.cern.ch/library

DESY Zeuthen Computing Seminar – March 2001



Once registered ...

- ◆ The user can use any form:
 - ◆ <http://cern.ch/SiteName> - Read/Only
 - ◆ <http://SiteName.web.cern.ch/SiteName> - Read/Author/Write
 - ◆ <http://host.cern.ch/SiteName> - Read/Author/Write
- ◆ And in addition, all servers can be made aware of the namespace !
- ◆ Once registered, any of the following URLs will work, *whatever registered SiteName is used !*
 - ◆ <http://cern.ch/SiteName> - Read/only
 - ◆ <http://www.cern.ch/SiteName> - Read/only
 - ◆ <http://web.cern.ch/SiteName> - Read/only
 - ◆ <http://nicewww.cern.ch/SiteName> - Read/only
 - ◆ <http://home.cern.ch/SiteName> - Read/only

DESY Zeuthen Computing Seminar – March 2001



More on cern.ch subdomains ...

- ◆ mailbox.cern.ch → mail server load balancing
- ◆ print.cern.ch → print server load balancing
- ◆ web.cern.ch → official web sites load balancing
- ◆ home.cern.ch → personal web sites load balancing
- ◆ webtest.cern.ch → test web sites load balancing
 - ◆ **IMPORTANT: webtest domain visible *only* within CERN**

DESY Zeuthen Computing Seminar – March 2001



Site Aliases

- ◆ The user can register alternative Site Names (ALIAS)
 - ◆ **More descriptive names, old names, NickNames**
 - ◆ <http://cern.ch/CERN.Web.Services> (descriptive name)
 - ◆ <http://cern.ch/web> (nickname)
 - ◆ <http://cern.ch/WebOffice> (old name)
 - ◆ <http://cern.ch/Alberto.Pace> (descriptive name - Personal site)
 - ◆ Alias are mapped to existing sites
 - ◆ <http://cern.ch/Alberto.Pace> -> <http://cern.ch/pace>
 - ◆ A site can have an unlimited number of aliases
 - ◆ Aliases have less restrictions in the 'allowed characters' (tildas, dots, underlines, ...)

DESY Zeuthen Computing Seminar – March 2001



Renaming sites

- ◆ The Site Alias feature allows web authors to **rename web sites** when necessary without breaking existing hyperlinks that have been bookmarked or hardcoded in an unknown number of html documents worldwide

DESY Zeuthen Computing Seminar – March 2001



The new architecture ...

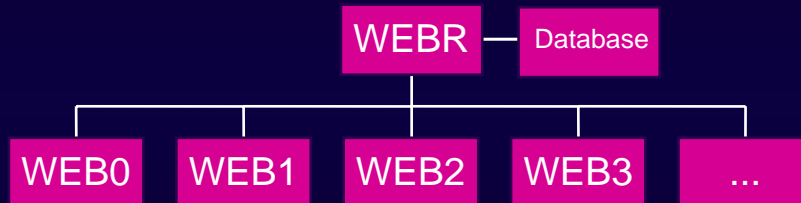
- ◆ ... is compatible with the existing infrastructure
- ◆ Existing servers and existing URLs integrates smoothly in the namespace
- ◆ No broken links
- ◆ Has allowed the evolution of the service
 - ◆ Migrations from local servers to central servers and vice-versa
 - ◆ Migrations between central servers, suppression of local servers
 - ◆ Split of large servers into smaller ones and vice-versa
 - ◆ Multiple central servers (differentiation possible, if necessary)
 - ◆ Server Load balancing
 - ◆ Stable HTML only service versus full CGI-BIN interfaces
 - ◆ Enhanced services, Multiple server platforms (!)
 - ◆ Test versus production sites
 - ◆ Personal versus official sites
 - ◆ ...

DESY Zeuthen Computing Seminar – March 2001



Physical Architecture

Web Redirector



- ◆ Heterogeneous pool of Web servers
 - ◆ Multiple OS (NT4, Win2000, Solaris, Linux)
- ◆ Web Sites can also be hosted on servers not managed by the Web Services team (locally managed servers)

DESY Zeuthen Computing Seminar – March 2001



Agenda

- ◆ CERN Web Architecture
- ◆ Services for End-Users
- ◆ Web Storage
- ◆ Web Technologies for Application Development

DESY Zeuthen Computing Seminar – March 2001



Today's Web Services *

- ◆ **Site Hosting**
 - ◆ a pool of PCs to host web sites of customer who do not want to maintain their own server
 - ◆ Load balanced using the web.cern.ch domain
 - ◆ Every web site has one (and only one) owner
 - ◆ Owners are *responsible* for the site content, can manage or delegate the site permissions and security
 - ◆ Flat Namespace (part of the CERN web namespace). Subsites possible but managed by owners of upper sites

*: Some restrictions to the availability to the Web services exists for non-Windows users and for non-FrontPage users

DESY Zeuthen Computing Seminar – March 2001



Today's Web Services * (2)

- ◆ **Authoring interfaces**
 - ◆ HTTP-PUT
 - ◆ FTP
 - ◆ OSE = Microsoft Office Server Extensions
See <http://www.microsoft.com/office/ork/2000/five/75t5.htm>
 - ◆ WebDAV = Distributed Authoring and versioning
See <http://www.w3.org/Protocols/>
 - ◆ **Authoring through the file system discouraged**
 - ◆ Platform Specific – important effort duplication

DESY Zeuthen Computing Seminar – March 2001



Today's Web Services * (3)

- ◆ **Automated Site Creation/Registration Service**
 - ◆ **Users can create/delete web sites or registrations themselves, in real time**
 - ◆ **Users are authenticated on AFS or NICE**
 - ◆ **A moderation entity is notified in real time**

DESY Zeuthen Computing Seminar – March 2001



Today's Web Services * (4)

- ◆ **Support for Electronic Forms**
 - ◆ **With results saved to a file on the web server, sent to a user-written form handler, sent by E-mail, saved to a database**
 - ◆ **Used to generate dynamic queries to databases**
- ◆ **Support for Databases connection from web pages**
 - ◆ **Support for local (Access) and remote (Oracle) databases**
 - ◆ **Full Read and Write access**
 - ◆ **Reduced/zero programmatic effort**

DESY Zeuthen Computing Seminar – March 2001



Today's Web Services * (5)

- ◆ **Searching and Indexing**
 - ◆ **Both client-side (global) and server-side (fast, rich) indexing**
- ◆ **Access Control & Security**
 - ◆ **ACL on files, global authentication, login/password forms, IP address restrictions, document password protected ...**
- ◆ **Secure connections, HTTPS and SSL**
 - ◆ **Every Web page on the central Web server can be accessed using HTTP *or* HTTPS**
 - ◆ **Web Author can disable http access on a per page basis**

DESY Zeuthen Computing Seminar – March 2001



Today's Web Services * (6)

- ◆ **Cascading Style Sheets (CSS) and Themes**
- ◆ **Source Control**
 - ◆ **Document check-in and check-out**
 - ◆ **Multi-authored web sites**
- ◆ **Authoring tools**
 - ◆ **Support for FrontPage and Dreamweaver**
- ◆ **CGI, scripting**
 - ◆ **Standard CGI and ISAPI Interfaces**
 - ◆ **Support for Active Server pages (ASP)**
 - ◆ **Perlscript, Jscript, Vbscript**

DESY Zeuthen Computing Seminar – March 2001



Agenda

- ◆ CERN Web Architecture
- ◆ Services for End-Users
- ◆ Web Storage
- ◆ Application Development

DESY Zeuthen Computing Seminar – March 2001



File Storage ~~Ways to access a web site ...~~

Home Page - Microsoft Internet Explorer
Address: http://webservices.web.cern.ch/WebServices/
CERN Web Services Home Page
Search this site
Documentation for Web In case of problems: Use
ftp://pace@webservices.web.cern.ch/webservices/

HTTP (get)

\\WEB01WebServices - Microsoft Internet Explorer
Address: \\webservices.web.cern.ch\WebServices
private AuthoringDoc cgi-bin ClientDoc Communica... General
images File System default.asp

File System

ftp://pace@webservices.web.cern.ch/webservices/ - Microsoft Internet Explorer
Address: ftp://pace@webservices.web.cern.ch/webservices/
private AuthoringDoc cgi-bin ClientDoc Communica... General
images FTP session default.asp
wan150.doc wan150.doc wan170.doc wan180.doc wan190.doc wanie.doc

FTP session

http://webservices.web.cern.ch/webservices/ - Microsoft Internet Explorer
Address: http://webservices.web.cern.ch/webservices/
private AuthoringDoc cgi-bin ClientDoc Communica... General
images HTTP Read/Write WEB DAV WEB Folders
wan150.doc

**HTTP Read/Write
WEB DAV
WEB Folders**

2001



Web Servers vs File Systems

- ◆ **Web Servers are cross-platform alternatives to file servers**
 - ◆ **No need for platform specific solutions (AFS, Appleshare, DFS, IFS, NFS, Netbios, Novell, SMB, ...) for file sharing**
 - ◆ **This does not mean that there is no future for file systems, but only that file system can remain local to the platform and it is no longer a technology used to share information**
- ◆ **Web Sites as project space or home directories**
 - ◆ **There is no longer the need to distinguish between Web sites and home directories**
 - ◆ **End-users can create read-protected web sites with only one author**
 - ◆ **This enables the possibility of getting at files from any internet access point in the world**

DESY Zeuthen Computing Seminar – March 2001



WebDav Strong / Weak points

- ◆ **It is just a 'protocol' (based on http)**
 - ◆ **Multiple server implementation possible ...**
 - ◆ **Allows read and write access to files**
 - ◆ **It has source control (multiple authors, offline folders, versioning, workflow, ...)**
 - ◆ **See Thaoe examples**
 - ◆ **RFC exists to manage access control through WEBDAV**
 - ◆ **It inherits the HTTP security: simple and effective**

DESY Zeuthen Computing Seminar – March 2001



Suggestion # 1

- ◆ Have a look to the WEB-DAV specification on www.w3.org, www.webdav.org
- ◆ Have a look to the “Web folders” features available on all platforms where IE is available (HP-UX, Solaris, Macintosh, Windows)
- ◆ Try Gnome (with Nautilus file manager) on Linux
- ◆ If you are writing software...
 - ◆ **Avoid writing code that access platform specific file systems. Rather use HTTP to read or write files from your programs ...**

DESY Zeuthen Computing Seminar – March 2001



Comments

- ◆ We are already promoting the usage of Web storage for Windows (2000) users
 - ◆ **Personal web site as Home directory**
 - ◆ **Official Web sites as Project space**
 - ◆ **Native Built-in OS support**
 - ◆ **Standard technology -> Easy to outsource**
- ◆ We are looking in details on how to integrate DFS with Web storage services (this year)
- ◆ Should we promote the usage of Web storage also on other platforms ?
 - ◆ **Leave FOA for the discussion ...**

DESY Zeuthen Computing Seminar – March 2001



Agenda

- ◆ CERN Web Architecture
- ◆ Services for End-Users
- ◆ Web Storage
- ◆ Application Development

DESY Zeuthen Computing Seminar – March 2001



The Web as an application platform

- ◆ Since the web was invented
 - ◆ <http://www.yyy.com/xxx?param1=val1¶m2=val2>
 - ◆ The querystring or posted data could be used as parameters to remote procedure call invocation
- ◆ Unfortunately ...
 - ◆ No real programming model, no object model, no inheritance, no structures, very low level
- ◆ But ...
 - ◆ There is a huge commercial interest and technology is evolving very fast.
 - ◆ SOAP (Simple Object Access Protocol) is a messaging model based on HTTP/XML that standardizes existing practice of using HTTP (and XML) as an application protocol

DESY Zeuthen Computing Seminar – March 2001



XML example

```
<purchaseOrder orderDate="1999-10-20">
  <comment>Urgent !</comment>

  <item partNum="872-AA">
    <productName>Lawnmower</productName>
    <quantity>1</quantity>
    <Price>148.95</Price>
  </item>
</purchaseOrder>
```

DESY Zeuthen Computing Seminar – March 2001



What is XML ?

- ◆ ... is a method for putting structured data in a text file
- ◆ ... it looks a bit like HTML (but isn't HTML)
- ◆ ... it is just text (but isn't meant to be read)
- ◆ ... it corresponds to a family of technologies
- ◆ ... it is license-free, platform-independent and well-supported

- ◆ Why is XML important ?
 - ◆ **Because ... It has schemas and transformations ...**

DESY Zeuthen Computing Seminar – March 2001



XML Schema Example (simplified)

```
<schema>
  <complexType name="purchaseOrder">
    <element name="comment" type="string"/>
    <element name="item" minOccurs="0">
      <complexType>
        <element name="productName" type="string"/>
        <element name="quantity" />
        <element name="Price" type="decimal"/>
        <attribute name="partNum" type="SKU"/>
      </complexType>
    </element>
  </complexType>
</schema>
```

DESY Zeuthen Computing Seminar – March 2001



XML and XML Schemas

- ◆ A Schema defines a class of XML documents,
- ◆ Complex Type Definitions, Element & Attribute Declarations
- ◆ Allows “verification” of XML “instances” against a schema
- ◆ Allows the “interpretation” of the XML data (if verification succeeds)
- ◆ Allows the “transformation” of the XML data
 - ◆ **Why are transformation important ?**

DESY Zeuthen Computing Seminar – March 2001



The original web

Server



HTML 1.0 Document

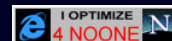
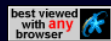
Client



DXHTML version 9 Super Browser

Mosaic 1.0

- ◆ Same Document sent to all clients, whatever capabilities they have
- ◆ To support all browser, must use lowest common denominator (no frames, no tables, no animation, no dynamic effects, no sound, no video, ...)
- ◆ Very difficult to move forward ...



DESY Zeuthen Computing Seminar – March 2001



The Web Today

Server



DXHTML 9.0 Document

HTML 1.0 Document

Client



DXHTML version 9 Super Browser

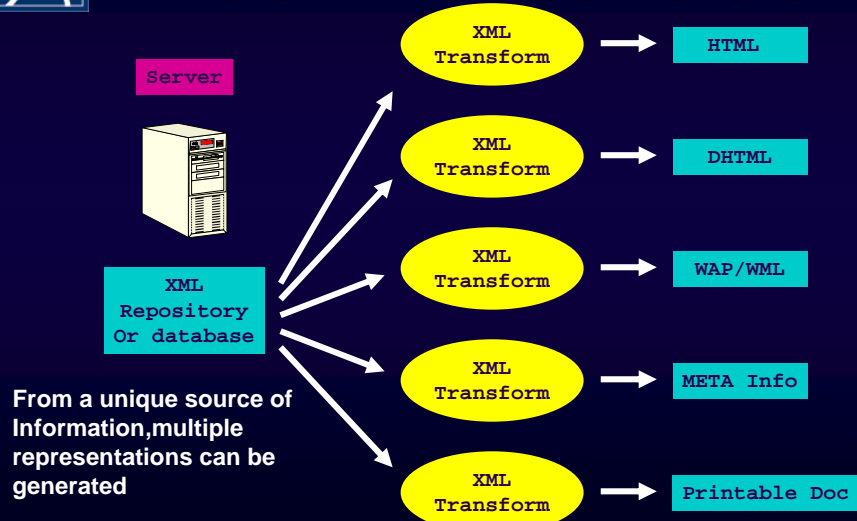
Mosaic 1.0

- ◆ Different documents sent to clients according to the capabilities they have
- ◆ The technology to select the proper document relies (often) on client side scripting
- ◆ Expensive and difficult to maintain: Multiple copies of every documents are necessary

DESY Zeuthen Computing Seminar – March 2001



The Web Tomorrow: XML and XML Transformations



DESY Zeuthen Computing Seminar – March 2001



XML Summary

- ◆ XML is the core technology that will be used in the web (and internet) arena in the next years.
- ◆ Now, Why HTTP ...

DESY Zeuthen Computing Seminar – March 2001



HTTP Example

(simplified)

HTTP Request

```
GET /bar/foo.txt HTTP/1.1
```

or

```
POST /bar/foo.cgi HTTP/1.1  
Content-Type: text/plain  
Content-Length: 11
```

```
Hallo Welt !
```

HTTP Response

```
200 OK  
Content-Type: text/plain  
Content-Length: 12
```

```
Hello World !
```

DESY Zeuthen Computing Seminar – March 2001



Why HTTP ?

- ◆ HTTP (hypertext transfer protocol)
 - ◆ ... has become the de facto protocol of the Internet, it is a real standard on top of which you can build
 - ◆ ... is available on all platforms !
 - ◆ ... is a simple protocol that requires little runtime support to work properly
 - ◆ ... is a session less protocol.
Few/no packets exchanged to set up/maintain sessions.
 - ◆ ... security is simple and effective
 - ◆ ... it is the only thing usable over firewalls

DESY Zeuthen Computing Seminar – March 2001



SOAP

- ◆ Simple Object Access Protocol
 - ◆ Fully based on HTTP (or Secure HTTP)
 - ◆ Remote Procedure call (RPC) model inspired by CORBA Internet Inter Object Protocol (IIOP)
 - ◆ Submitted as a W3C Note by Compaq, HP, IBM, Microsoft, SAP and other companies
 - ◆ <http://www.w3.org/Submission/2000/05/>

DESY Zeuthen Computing Seminar – March 2001



SOAP Example

(simplified)

```
POST /path/foo.pl HTTP/1.1
Content-Type: text/xml
SOAPActor: interfaceURI#Add
Content-Length: nnnn
```

```
<soap:Envelope>
  <soap:Body>
    <Add>
      <arg1>24</arg1>
      <arg2>53.2</arg2>
    </Add>
  </soap:Body>
</soap:Envelope>
```

```
200 OK
Content-Type: text/xml
Content-Length: nnnn

<soap:Envelope>
  <soap:Body>
    <Response>
      <sum>77.2</sum>
    </Response>
  </soap:Body>
</soap:Envelope>
```

DESY Zeuthen Computing Seminar – M



Demo 1



DESY Zeuthen Computing Seminar – March 2001



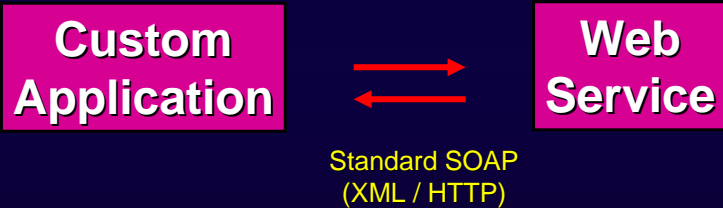
Demo 2



DESY Zeuthen Computing Seminar – March 2001



Demo 3



DESY Zeuthen Computing Seminar – March 2001



HTTP advantages ...

- ◆ **Over IIOP and DCOM (but also OPC, RMI, ...)**
 - ◆ **Easier protocol(s) (less programming skills)**
 - ◆ **No runtime support required to work properly**
 - ◆ **Straight cross platform porting**
 - ◆ **Usable over firewalls ! Global application possible.**
- ◆ **Over DCOM**
 - ◆ **Not connection oriented, no network overhead to create and maintain session state and information**
 - ◆ **Available on all platforms**
- ◆ **Over IIOP**
 - ◆ **ORB not required on both Client and Servers**
 - ◆ **No need for ORB licenses (\$\$\$). Only solution for global applications**

DESY Zeuthen Computing Seminar – March 2001



SOAP Strength

- ◆ **No new technology:** It simply standardizes existing practice of using HTTP (and XML) as an application protocol
- ◆ **No requirements on special runtime, no ORB, no traditional web server (e.g., Apache, IIS) required, no programming language implied, no programming model ...**
- ◆ **This is a *minimal* protocol for invoking methods on servers/services/components/objects**

DESY Zeuthen Computing Seminar – March 2001



SOAP Flexibility

- ◆ **SOAP can be viewed as another RPC protocol**
 - ◆ **Request contains in and inout parameters**
 - ◆ **Response contains inout and out parameters**
- ◆ **SOAP can be viewed as another ORPC protocol**
 - ◆ **Requests represent properties, method invocations, event triggering, callback and notifications**
- ◆ **SOAP can be viewed as a “messaging” protocol**
 - ◆ **Request contains a single serialized request object**
 - ◆ **Response contains a single serialized response object**
- ◆ **SOAP can be viewed as XML Transformations**
 - ◆ **Request contains an XML document**
 - ◆ **Server returns a transformed version**

DESY Zeuthen Computing Seminar – March 2001



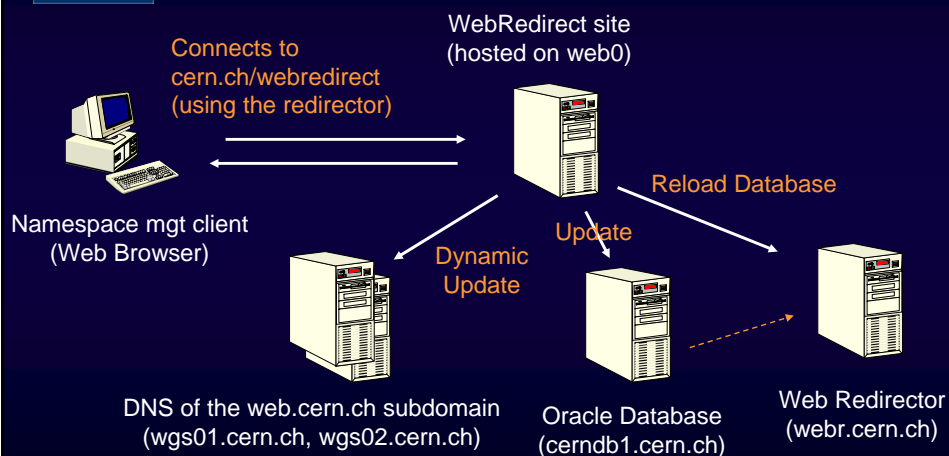
Creating SOAP applications

- ◆ You don't need to learn SOAP
- ◆ SOAP can be buried in your ORB product
 - ◆ A "CORBA version" using SOAP instead of IIOP
 - ◆ Your existing object can use soap without modification
- ◆ SOAP can be buried in your Web Server
 - ◆ Apache, ASP/ISAPI, JSP/Servlets
- ◆ SOAP can roll from your own components
 - ◆ Written in any language (Basic, Java, C, Perl, ...)
- ◆ Can be automatically generated from your development environment
 - ◆ Visual Studio .NET has built-in SOAP support. This includes Visual Basic, C++, C#.

DESY Zeuthen Computing Seminar – March 2001



Example: CERN Site Registration



DESY Zeuthen Computing Seminar – March 2001



Suggestion # 2

- ◆ Have a look to the SOAP submission on www.w3.org
- ◆ Have a look to the development tools available that generate HTTP/XML requests to manipulate your objects
- ◆ If you are writing global software (grid ?) ...
 - ◆ Use HTTP/XML for your RPCs, Object invocation, message passing ...

DESY Zeuthen Computing Seminar – March 2001



Comment

- ◆ The Web using HTTP/XML is more than cross OS development: It is a technology that integrates devices that are not computers
 - ◆ We can see more and more “black Boxes” that using http/xml offer “services”
 - ◆ Mobile phones, WebTV, Printers, Web Storage, PLCs, DAQ systems, ...

DESY Zeuthen Computing Seminar – March 2001



Before Ending

- ◆ UDDI - Universal Discovery Description and Integration
 - ◆ White pages
 - ◆ Yellow pages
 - ◆ Green pages
- ◆ Allow businesses to register information about the Web Services they offer so that other businesses can find them

DESY Zeuthen Computing Seminar – March 2001



Before Ending

- ◆ Links shown on WebDAV
 - ◆ Webdav Specifications (RFC 2518)
 - ◆ <http://andrew2.andrew.cmu.edu/rfc/rfc2518.html>
 - ◆ Webdav.org
 - ◆ <http://www.webdav.org/>
 - ◆ IETF Webdav Working group
 - ◆ <http://www.ics.uci.edu/~ejw/authoring/>
 - ◆ Nautilus (and Gnome) for Linux
 - ◆ <http://nautilus.eazel.com/>
 - ◆ Sun Open Link Virtuoso
 - ◆ http://solutions.sun.com/catalogs/all/Software_Development_Tools/ORBs_Middleware/49501.html
 - ◆ Microsoft
 - ◆ <http://www.microsoft.com/Tahoe>
 - ◆ Other
 - ◆ <http://www.lyra.org/pipermail/dav-announce/2000q1/000051.html>
- ◆ Links shown on SOAP
 - ◆ W3c.org
 - ◆ <http://www.w3.org/TR/2000/NOTE-SOAP-20000508/>
 - ◆ Apache
 - ◆ <http://xml.apache.org/soap/>
 - ◆ Developer Mentor
 - ◆ <http://www.develop.com/soap/>
 - ◆ Sun Open Link Virtuoso
 - ◆ http://solutions.sun.com/catalogs/all/Software_Development_Tools/ORBs_Middleware/49501.html
 - ◆ Microsoft
 - ◆ <http://msdn.microsoft.com/xml/general/soapspec.asp>
 - ◆ <http://msdn.microsoft.com/net>

DESY Zeuthen Computing Seminar – March 2001



Conclusion

- ◆ The web is more than a technology to share documents ...
 - ◆ The web as a file system (WebDav)
 - ◆ The web as an application platform (HTTP/XML)

<http://cern.ch/web>

<http://cern.ch/win>

DESY Zeuthen Computing Seminar – March 2001



Agenda

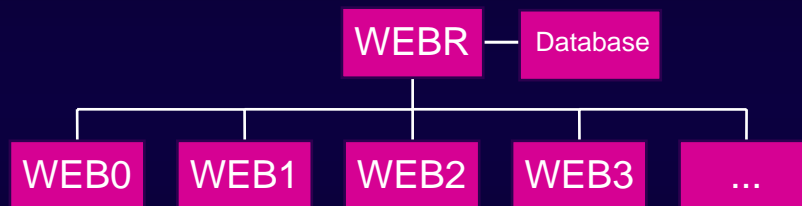
- ◆ CERN Web Architecture
- ◆ Services for End-Users
- ◆ Web Storage
- ◆ Application Development
- ◆ Bonus material, extra features and goodies

DESY Zeuthen Computing Seminar – March 2001



Physical Architecture

Web Redirector



- ◆ Webr is the redirector, which does not host web sites
- ◆ Webx x=0,1,..n are the web servers
- ◆ Web Sites can also be hosted on servers not managed by the Web Services team (locally managed servers)

DESY Zeuthen Computing Seminar – March 2001



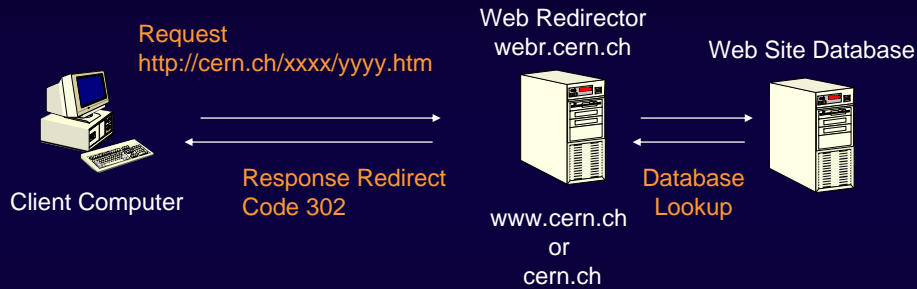
Generic and Physical Site Names

- ◆ What happens when the user registers
 - ◆ <http://cern.ch/alberto.pace> -> <http://myhost.cern.ch/pace>
- ◆ The following happens
 - ◆ pace.web.cern.ch becomes a DNS alias of myhost.cern.ch
 - ◆ <http://pace.web.cern.ch/pace> works in R/W
 - ◆ The SiteName 'pace' is registered in the cern.ch domain
 - ◆ <http://cern.ch/pace> works in Read mode
 - ◆ The SiteName 'alberto.pace' is registered in the cern.ch domain as an alias of 'pace'
 - ◆ <http://cern.ch/alberto.pace> works in Read mode

DESY Zeuthen Computing Seminar – March 2001



Namespace implementation



- ◆ www.cern.ch and cern.ch have the address of webr.cern.ch
- ◆ All unqualified requests (www.cern.ch and cern.ch) are handled by the redirector
- ◆ Web site database stored on Oracle and cached by the redirector
- ◆ The redirector does not have files locally stored

DESY Zeuthen Computing Seminar – March 2001



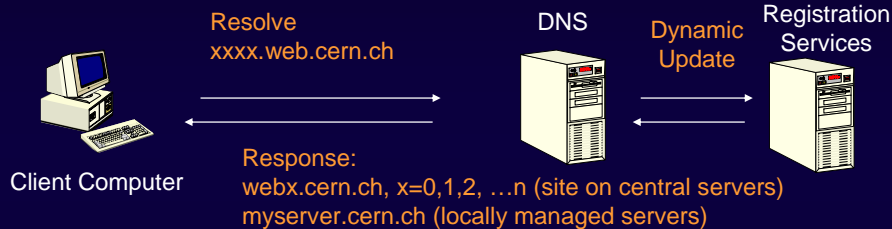
Redirector response

- ◆ **Depends on the web site type stored in the database**
 - ◆ **Normal sites**
 - ◆ http://www.cern.ch/xxx -> http://xxx.web.cern.ch/xxx
 - ◆ The web server hosting the site is resolved using the DNS
 - ◆ Supports locally managed servers
 - ◆ **Aliases**
 - ◆ http://www.cern.ch/Per.Hagen -> http://www.cern.ch/hagen
 - ◆ Are resolved recursively inside the redirector (no double redirection seen by the web client)
 - ◆ Allows illegal characters in the site name (%, -, _)
 - ◆ **Hard-coded exceptions**
 - ◆ http://www.cern.ch/zzzz -> http://myserver.cern.ch/aaaa/b/c
 - ◆ Fits all other possibilities:
 - ◆ Deep, nested sites
 - ◆ Sites stored in the root of a locally managed server

DESY Zeuthen Computing Seminar – March 2001



Server Resolution using DNS



- ◆ The DNS resolves the physical address of the form xxx.web.cern.ch (which represents “the web server hosting site xxx”)
- ◆ Can be the address of a “central” web server (webx x=0,1,...n)
- ◆ Can be the address of a “locally managed” server

DESY Zeuthen Computing Seminar – March 2001



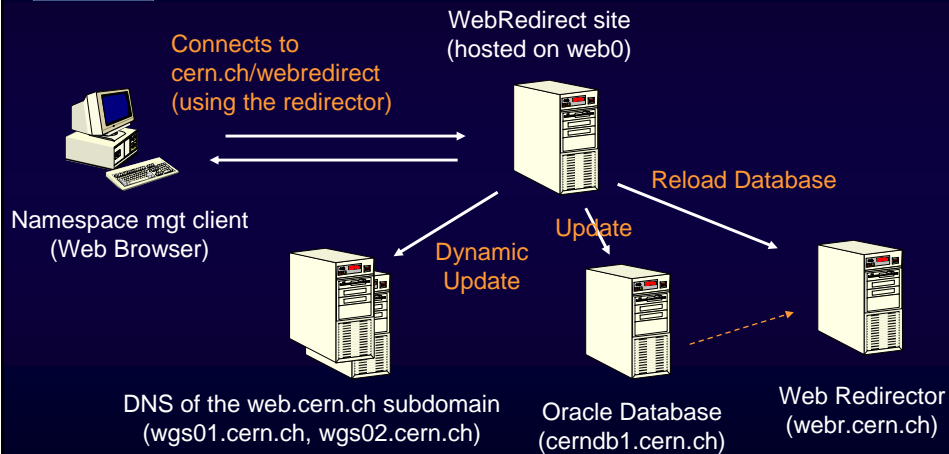
Site Registration Services

- ◆ Web0.cern.ch has a special role: it hosts the registration site
- ◆ It is through the site <http://cern.ch/webredirect> (hosted on Web0) that the database and the redirector is managed
- ◆ The redirector itself is completely data driven from the site database stored on oracle (cerndb1/webreg).
- ◆ The only message that can be sent to the redirector is a “database modified” signal that invalidates the redirector cached database and forces a database reload.
 - ◆ This signal is sent to the redirector by reading the URL <http://cern.ch/ReloadDatabaseData>

DESY Zeuthen Computing Seminar – March 2001



Site Registration Services



DESY Zeuthen Computing Seminar – March 2001



Compatibility with the past

- ◆ To allow migration of existing sites from / to the central servers we have to be able to also handle requests of the type **http://xxx.cern.ch**
- ◆ To have this compatibility Site Owners can register the hostname xxx.cern.ch as an alias of webr.cern.ch
- ◆ The redirector will intercept all requests sent to xxx.cern.ch and translate them according to the Web Site database:
 - ◆ **http://xxx.cern.ch/... -> http://cern.ch/xxx/... -> ...**
- ◆ A flag in the database attached to the site name, allows to ignore the interpretation of the hostname as the sitename
 - ◆ **http://www.cern.ch/... -> http://cern.ch/... -> ...**
 - ◆ **http://web.cern.ch/... -> http://cern.ch/... -> ...**
 - ◆ **http://webr.cern.ch/... -> http://cern.ch/... -> ...**

DESY Zeuthen Computing Seminar – March 2001



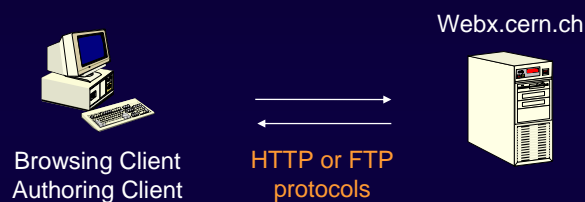
The Webx Servers (x=0,1,...n)

- ◆ Can be uniform (all equal) or specialised (all different)
 - ◆ Current strategy: Try to be uniform (ease of management) and differentiate only where necessary
- ◆ Current and foreseen situation
 - ◆ Two Web server OS: NT or SunOs (a migration to Linux can be envisaged if the Sun HW has to be replaced)
 - ◆ Two Web server daemons: IIS and Apache
 - ◆ Web sites can be local on the Web server or stored on remote file systems (AFS, Novell)
 - ◆ Servers can be specialised: Gateways, Scripts, static html only, ...
- ◆ Many, many combinations possible. This is what is (currently) exposed to end users:
 - ◆ Web sites on Central servers authored using http/ftp
 - ◆ Web sites stored on AFS, authored from AFS clients
 - ◆ Web sites stored on Novell, authored from a NICE client

DESY Zeuthen Computing Seminar – March 2001



Web sites on Central Servers

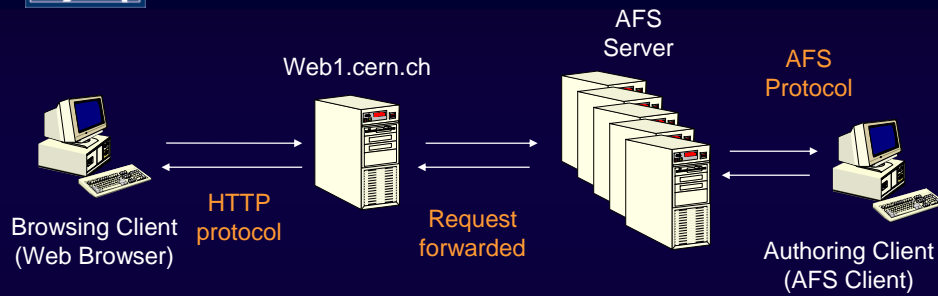


- ◆ All services supported
 - ◆ FrontPage and Dreamweaver support
 - ◆ E-Forms, Database connections, Access Control, Source Control, SSL, Indexing, Scripting (CGI-Interface, Perl, VB, Java).

DESY Zeuthen Computing Seminar – March 2001



Sites Stored on AFS



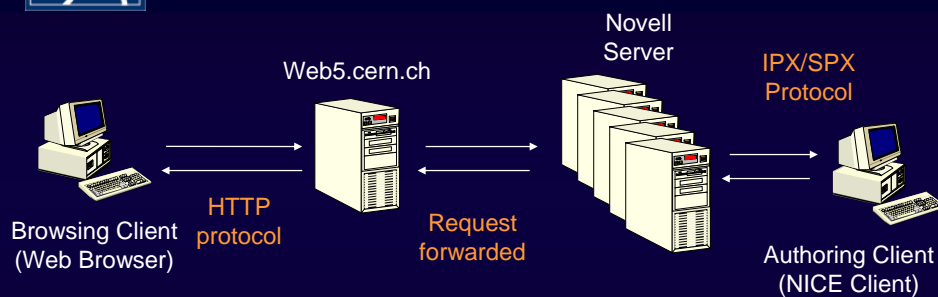
◆ Supports

- ◆ **Static HTML, Access Control, Scripting (Perl).**
- ◆ **Scripting allows Forms, Database access,**
- ◆ **FrontPage and Dreamweaver work but not officially supported (authoring done via the file system)**

DESY Zeuthen Computing Seminar – March 2001



Sites stored on Novell



◆ Supports

- ◆ **Static HTML only**
- ◆ **FrontPage and Dreamweaver work but not officially supported (authoring done via the file system)**

DESY Zeuthen Computing Seminar – March 2001



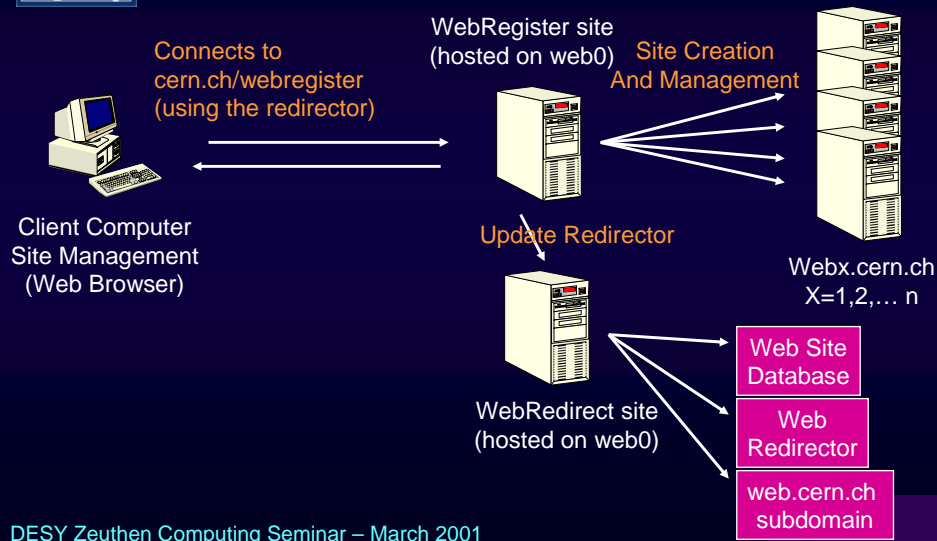
Sites Creation & Mgt Services

- ◆ Web0.cern.ch has a special role: it hosts the site creation and management interface
- ◆ It is through the site <http://cern.ch/webregister> (hosted on Web0) that all Web sites are created/deleted/maintained
- ◆ The Site management interface uses the site registration interface (WebRegister) to manage the redirector (webr) and the site database stored on oracle (cerndb1/webreg).

DESY Zeuthen Computing Seminar – March 2001



Sites Creation & Mgt Services



DESY Zeuthen Computing Seminar – March 2001



Site Creation example

- ◆ When a user creates a web site he will
 - ◆ Have a virtual web site created on a machine in the computer centre `webx.cern.ch` ($x = 1, 2, \dots n$)
 - ◆ The machine hosting his web site will have the DNS alias `sitename.web.cern.ch`
 - ◆ Sitename will be register in the redirection database. The site will be available using `http://cern.ch/sitename`
 - ◆ The site and its owner will be registered in the central Oracle database

DESY Zeuthen Computing Seminar – March 2001



Migrating existing servers

- ◆ Simple migration paths exist to move/suppress web sites to/from the computer centre
- ◆ Example:
 - ◆ Existing physical server `http://xxx.cern.ch`
 - ◆ User creates the site `http://cern.ch/xxx` on the central services
 - ◆ Data is moved from `xxx.cern.ch` to `xxx.web.cern.ch`
 - ◆ The computer `xxx.cern.ch` is removed from the network. Its host name is registered as an alias of the central redirection machine (`webr.cern.ch`)
- ◆ From now on:
 - ◆ Request `http://xxx.cern.ch/path/...` will be received by the redirector. These will be equivalent to requests `http://cern.ch/xxx/path/...` and therefore translated to `http://xxx.web.cern.ch/xxx/path/...`
 - ◆ All works. Data migrated, no broken links.

DESY Zeuthen Computing Seminar – March 2001