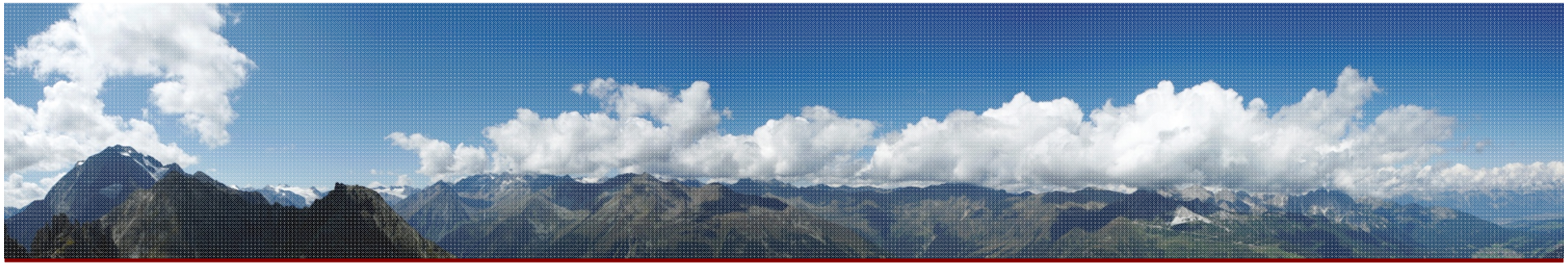
Web Services

Web Technologies



Where are we?

STI · INNSBRUCK

#	Title
1	Distributed Information Systems
2	Middleware
3	**Web Technologies**
4	Web Services
5	Basic Web Service Technologies
6	Web 2.0 Services
7	Web Service Security

Outline

STI · INNSBRUCK

- Motivation
- Technical solution
  - Exchanging Information over the Internet
  - Web Technologies for Supporting Remote Clients
  - Application Servers
  - Web Technologies for Application Integration
- Possible extensions
- Summary
- Resources

Motivation

Motivation

STI · INNSBRUCK

- Data and services often need to be shared across the boundaries of a single company or business unit:
  - Integration of different branches of the same company.
  - Automation of business processes that encompass several companies.
- The Web emerged as a technology for **sharing information** over the Internet.
  - It quickly became a medium for **connecting remote clients and servers** across the Internet.
  - More recently (with the advent of Web services) it became a medium **for integrating applications across the Internet**.
- This lesson aims at introducing basic Web technologies that are used to implement “Web” portion of Web services.

Technical Solution

Exchanging Information over the Internet

Exchanging Information over the Internet
Brief history

STI · INNSBRUCK

- In 1969 ARPA connected four universities in the US, building the network called ARPANET
  - The connected systems were autonomous and heterogeneous.
  - First standardization bodies were formed to govern the development of the network.
- One of the most prominent standards developed then is TCP/IP.
- Before the Web there were some other standards
  - Simple Mail Transfer Protocol (SMTP) – which is still the way to send e-mail
    - Later extended with Multi-purpose Internet Mail Extensions (MIME)
  - Telnet protocol
  - File Transfer Protocol (FTP)
    - Arrived soon after SMTP and Telnet.
    - Permitted anonymous file transfers.
  - Archie
    - Used FTP to create a distributed file system, index FTP archives and search through them.
  - Gopher
    - Simple client/server system and GUI for or distributing, searching, and retrieving text documents over the Internet.
- Core Web technologies (HTTP, HTML, Web servers and browsers) are evolution of those early technologies.

Exchanging Information over the Internet
Hyper Text Transfer Protocol (HTTP)

STI · INNSBRUCK

- Generic stateless protocol governing file transfer across a network.
- Originally developed by European Laboratory for Particle Physics (CERN)
  - The idea was to enable researchers to share their results and knowledge in a fast, easy and convenient manner.
- The same team came up with the name World Wide Web
  - The idea is today promoted and governed by WWW Consortium (W3C).
- Designed to support hypertext documents
  - HTTP supports Hyper Text Markup Language (HTML).
  - HTML defines standard set of markups used to render the information for human consumption.

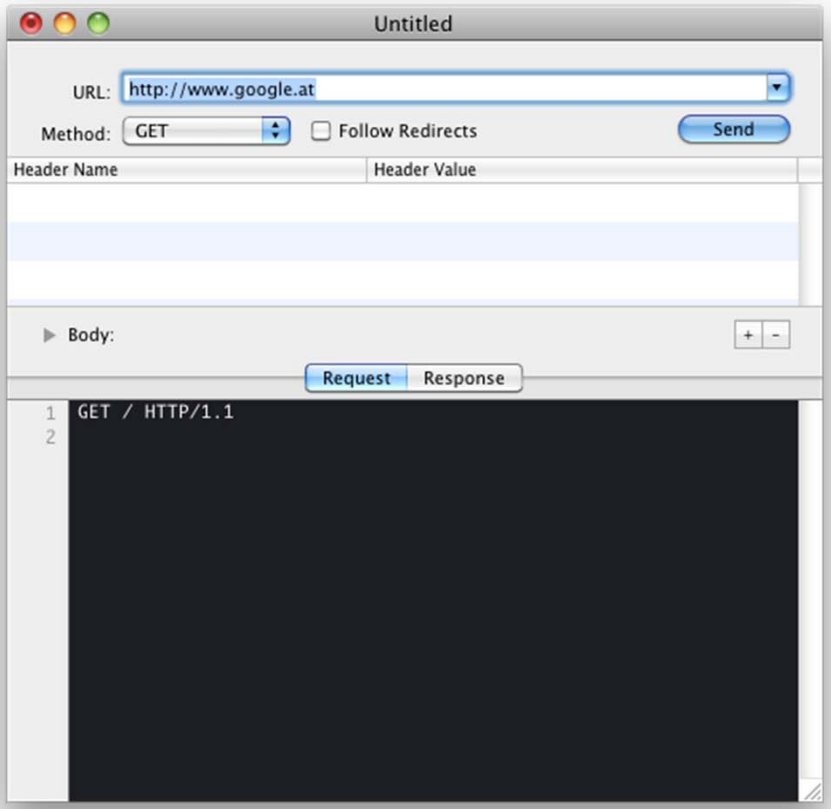
- HTTP documents are identified by Uniform Resource Identifiers (URIs).
  - URIs come in two flavors: Uniform Resource Locators (URLs) and Uniform Resource Names (URNs).
- URLs are the dominant way to identify documents over the Web.
  - In addition to identifying a resource, a URL provides a means to locate it
- A URL defines the *name of the protocol* (i.e., *scheme*) which should be used to access the document, the *address of the machine* where the resource is located, and *hierarchical description of the resource location* (and more like *query string*, and *anchor*).
- Documents can be static or dynamic
  - Dynamic documents are partially or in whole generated upon request.

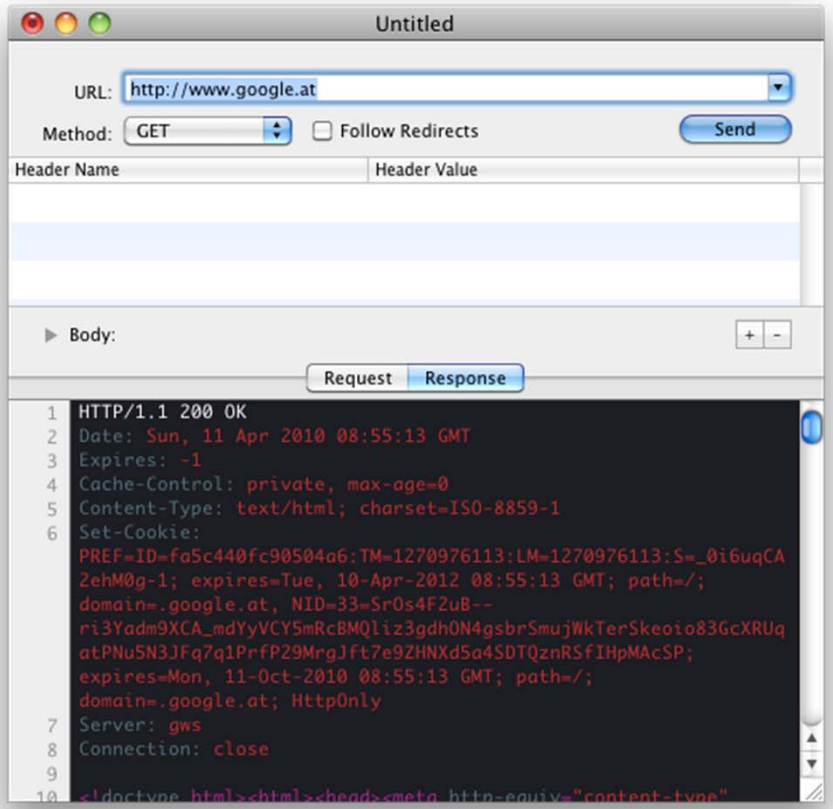
- HTTP underlying mechanism is Client/Server.
- HTTP typically relies on TCP/IP sockets.
- Starting from version 1.1 persistent connections are also supported.
- Most frequently used request methods are
  - OPTIONS – sends information about the communication options supported by a particular server,
  - GET – retrieves the specified document,
  - POST – appends or attaches the included data to the specified resource,
  - PUT – stores the included data at the location specified by the request, and
  - DELETE – deletes the resource indicated by the request.

Exchanging Information over the Internet

Hyper Text Transfer Protocol (HTTP) - Example

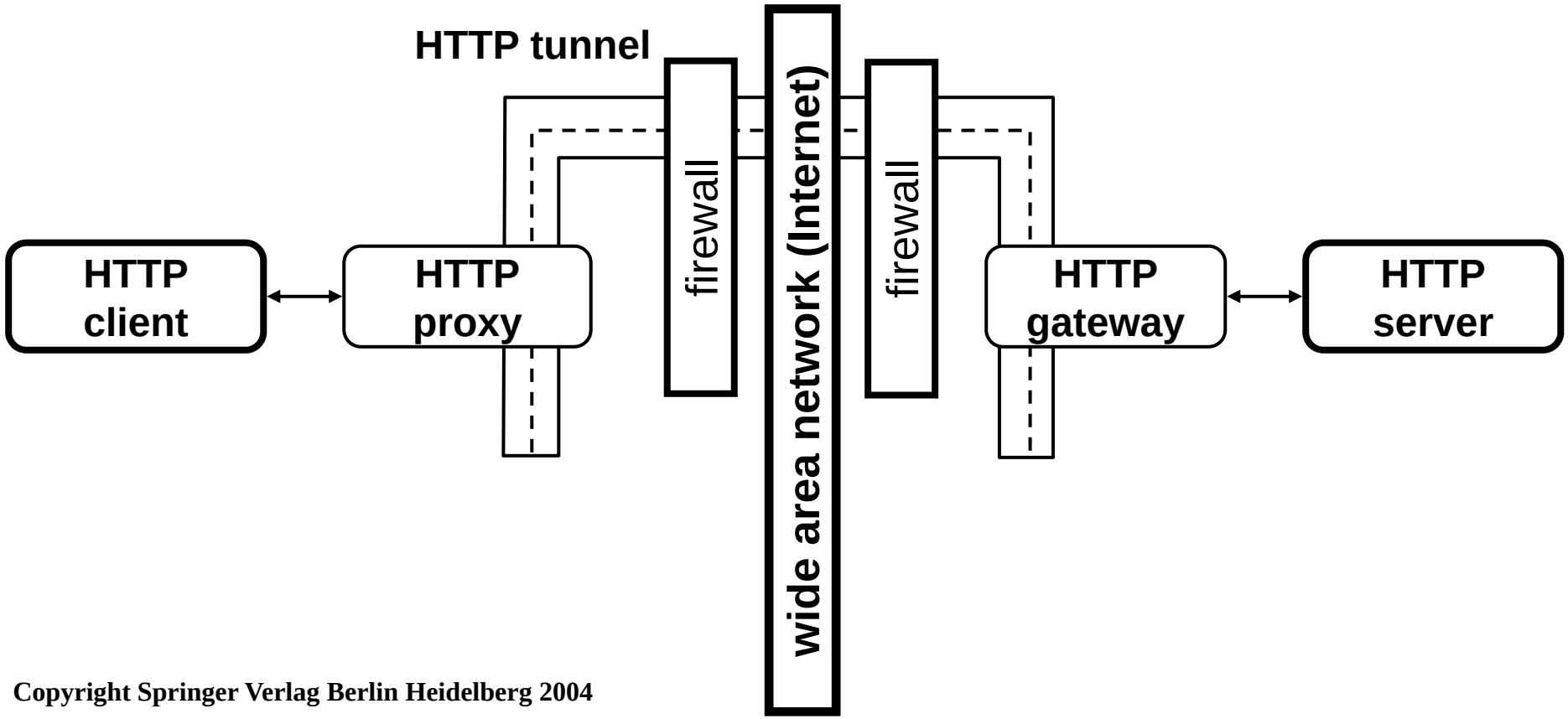
STI · INNSBRUCK





- Proxy (RFC 2616)
  - Intermediary program acting both as server and client for the purpose of making requests on behalf of other clients.
  - Potentially processes the URL and content.
- Gateway
  - Server acting as intermediary for some other server for the requested resource.
  - Acts on behalf of a server.
  - Potentially processes the URL and content.
- Tunnel
  - An intermediary program which is acting as a blind relay between two connections.
  - Used to connect two networks.
  - Doesn't process anything.

- Intermediary systems are enabling integration in the Web environment.



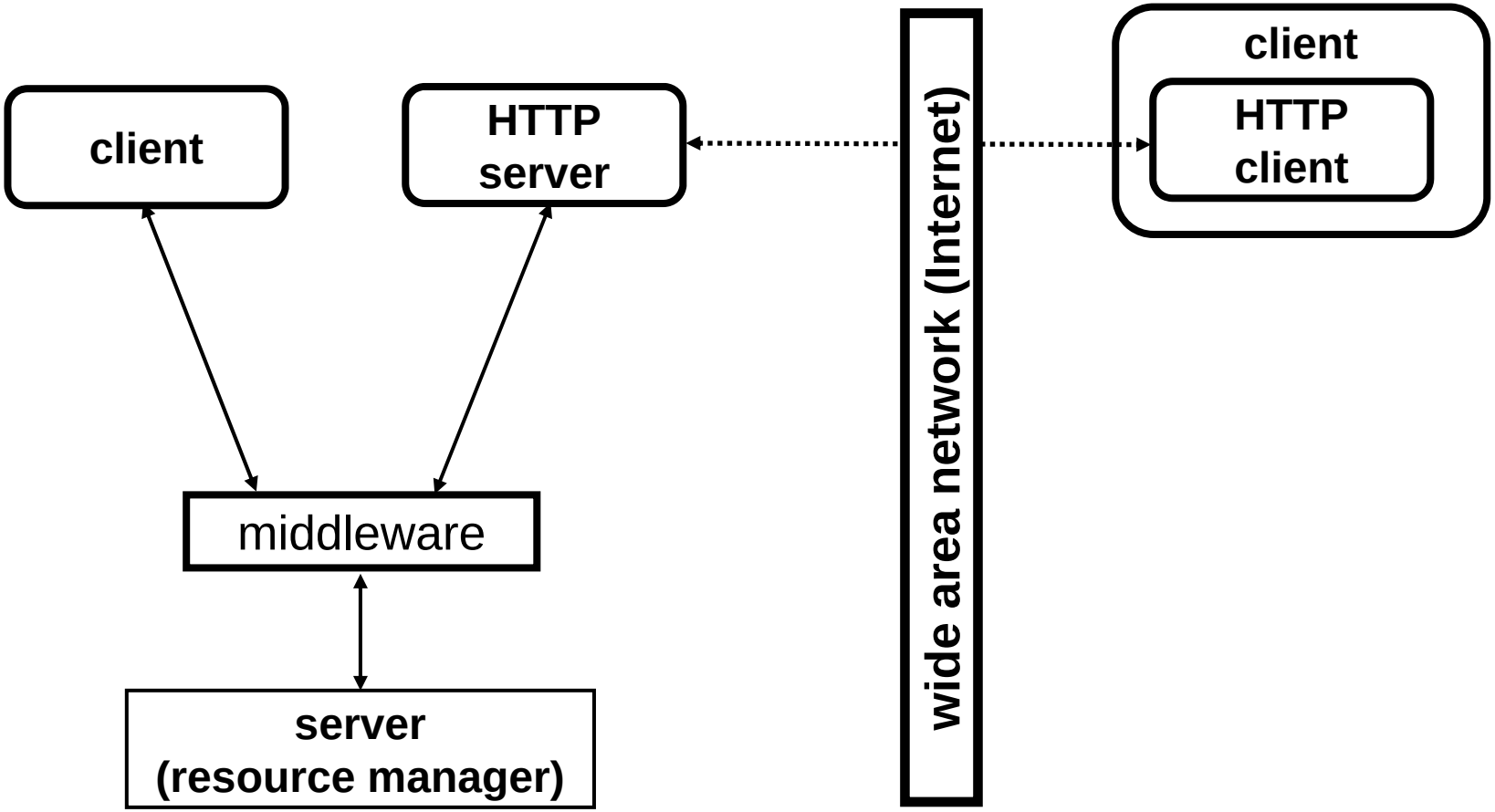
- No data encryption
  - Secure Socket Layer (SSL) developed by Netscape (1996), and its successor Transport Layer Security (TLS).
    - Relies on public key encryption to protect data transferred over TCP/IP.
  - Hyper Text Transfer Protocol over TLS/SSL (HTTPS)
    - Allows Web server and client to use TLS/SSL to authenticate to each other and establish an encrypted connection between themselves.

- Protocol is stateless
  - Information is not shared across HTTP request/response roundtrips.
  - Application developer is responsible for maintaining the relationships (i.e., state).
  - HTTP Cookies developed by Netscape (1994).
    - Enabling deployment of small data structures on the client machine on behalf of Web server.
    - They can maintain state information, can be used for personalization, tracking, session management.

Technical Solution

Web Technologies for Supporting Remote Clients

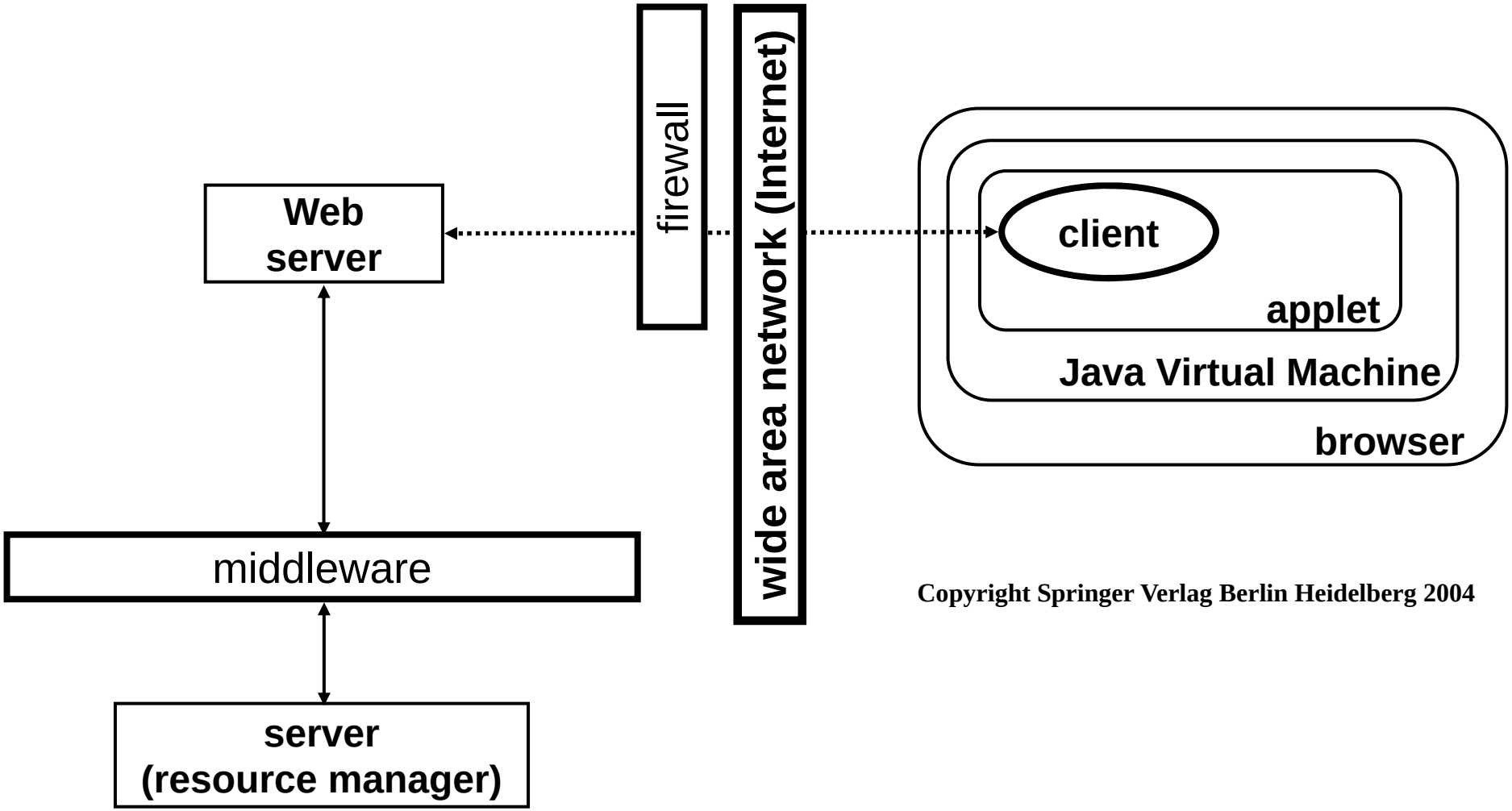
- Conventional middleware is assuming operation inside of the safe company boundaries.
- Information systems today are opening for some other users (e.g. customers)
  - Usage of Automatic Teller Machines (ATMs) by banks gives customers easier access to their accounts.
  - Manual work when dealing with customers disappears which reduces costs for banks.
- ATMs are not in a personal possession and they still incur some costs for customers (they need to travel to use the provided services)
  - Once the customer owns its personal ATMs (i.e. client) possibilities are endless – advanced applications, no usage constrains, etc.
- These are Business-To-Customer (B2C) operations
  - Customer is directly accessing company services.
  - Without Web technologies it would be quite complex to achieve efficient B2C.



- Building sophisticated applications at the client side (i.e. on Web browser side) is difficult.

- An applet is a Java program which can be embedded in an HTML document
  - Introduced by Sun Microsystems with the first version of Java language (1995).
  - The program is executed inside a Java Virtual Machine (JVM) in a controlled manner.
  - Client code (Java classes and associated artifacts) is sent to the client.
  - Applets turn the Web browser into an application-specific client without complex (re)configurations and installation procedures.
  - Applets are transient
    - Their lifetime is associated to the running browser instance.
    - Inadequate to support complex client code or frequent interactions.

STI · INNSBRUCK



- CGI is usually used to enable server to serve content from dynamic sources (e.g. publishing information from databases).

- It basically enables HTTP server to interface with external applications (they can serve as “gateways” to the local information system).
  - CGI establishes binding between a program and a requested URL.
  - Program arguments are sent as part of requested URL.
  - CGI programs can be written in various languages.
  - A program is started as a separate process and it interacts with the underlying middleware.