BÀI GIẢNG MÔN

AN TOÀN ỨNG DỤNG WEB & CSDL

CHƯƠNG 1 – TỔNG QUAN VỀ BẢO MẬT ỨNG DỤNG WEB

Giảng viên: PGS.TS. Hoàng Xuân Dậu

Khoa: An toàn thông tin

TÀI LIỆU THAM KHẢO

1. Hoàng Xuân Dậu, Bài giảng an toàn ứng dụng web và cơ sở dữ liệu, Học viện Công nghệ BCVT, 2017.
2. Bryan Sullivan, Vincent Liu, Web Application Security, A Beginner's Guide, McGraw-Hill, 2012.
3. Alfred Basta, Melissa Zgola, *Database Security*, Cengage Learning, 2012.
4. Dafydd Stuttard, Marcus Pinto, The Web Application Hacker's Handbook: Finding and Exploiting Security Flaws, John Wiley & Sons, 2011.
5. Ron Ben Natan, Implementing Database Security and Auditing, Elsevier Inc., 2005.
6. Mike Shema, Hacking Web Apps: Detecting and Preventing Web Application Security Problems, Elsevier Inc., 2012.
7. Roberta Bragg, Mark Rhodes-Ousley and Keith Strassberg, Network Security: The Complete Reference, McGraw-Hill Osborne Media, 2013.

TÀI LIỆU THAM KHẢO

7. Michael E. Whitman, Herbert J. Mattord, Principles of information security, 4th edition, Course Technology, Cengage Learning, 2012.
8. Denny Cherry, *Securing SQL Server: Protecting Your Database from Attackers*, Syngress, 2012.
9. Mark L. Gillenson, *Fundamentals of Database Management Systems*, 2nd edition, Wiley, 2011.
10. David Knox, Scott Gaetjen, Hamza Jahangir, Tyler Muth, Patrick Sack, Richard Wark, Bryan Wise, *Applied Oracle Security: Developing Secure Database and Middleware Environments*, McGraw-Hill Osborne Media, 2009.
11. Michael Gertz and Sushil Jajodia, *Handbook of Database Security Applications and Trends*, Springer, 2008.
12. Roberta Bragg, Mark Rhodes-Ousley and Keith Strassberg, *Network Security: The Complete Reference*, McGraw-Hill Osborne Media, 2013.

ĐÁNH GIÁ MÔN HỌC

- Các điểm thành phần:
  - Chuyên cần: 10%
  - Kiểm tra: 10%
  - Bài tập/thảo luận: 20%
  - Thi cuối kỳ: 60%

NỘI DUNG MÔN HỌC

Phần I – An toàn ứng dụng web

1. Tổng quan về bảo mật các ứng dụng Web
2. Các dạng tấn công lên các ứng dụng Web
3. Các biện pháp bảo mật máy chủ, ứng dụng và trình duyệt web
4. Bảo mật trong phát triển và triển khai ứng dụng web

Phần II – An toàn cơ sở dữ liệu

5. Tổng quan về an toàn cơ sở dữ liệu
6. Các cơ chế bảo mật cơ sở dữ liệu
7. Sao lưu, khôi phục dự phòng, kiểm toán và giám sát hoạt động CSDL

NỘI DUNG CHƯƠNG 1

1. Giới thiệu về dịch vụ web và kiến trúc các ứng dụng web
2. Các nguyên tắc bảo mật các ứng dụng Web
3. Các nguy cơ và lỗ hổng bảo mật trong các ứng dụng Web
4. Các phương pháp tiếp cận bảo mật các ứng dụng Web.

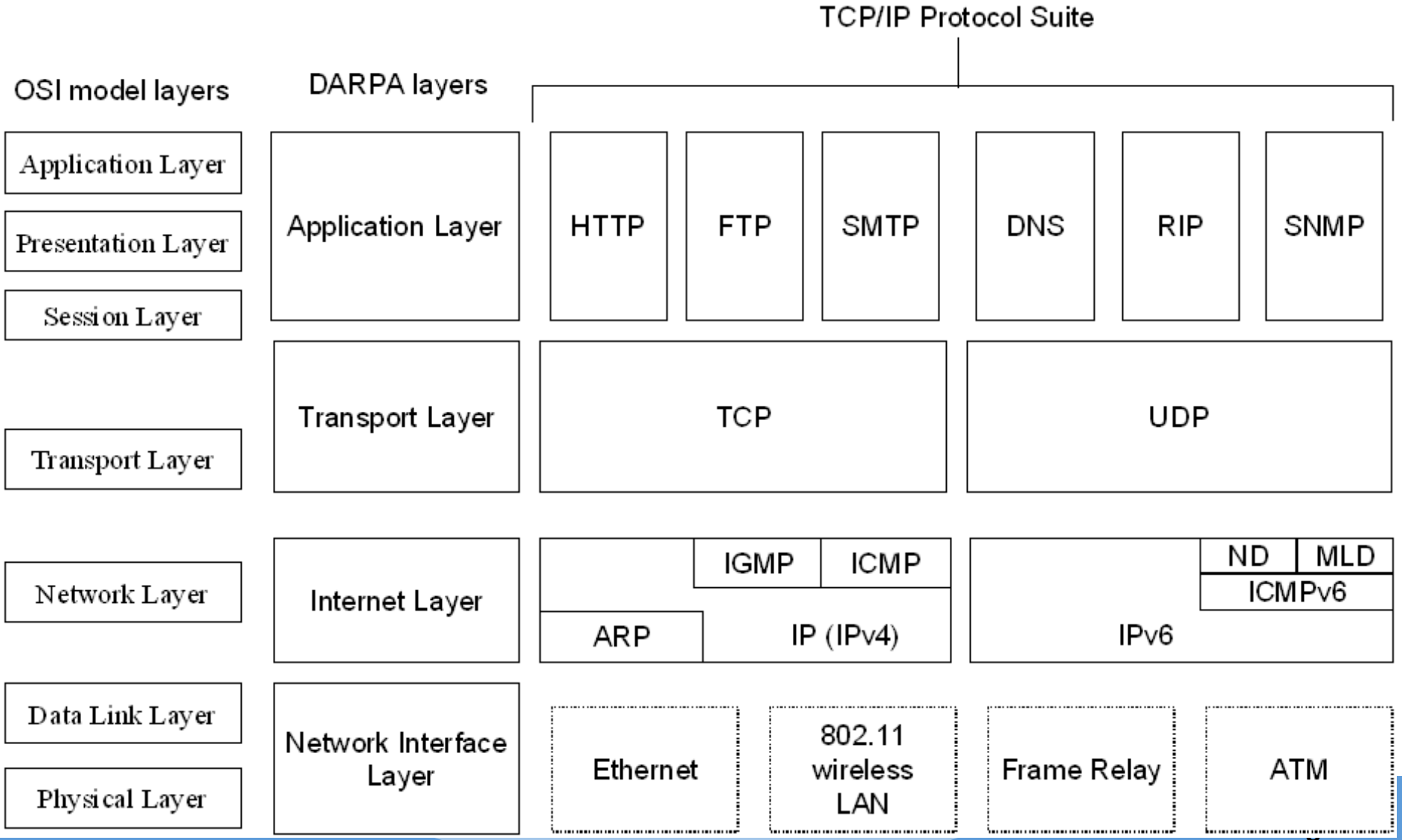
1.1 Giới thiệu về dịch vụ web và kiến trúc các UD web

1. Giao thức HTTP
2. Các thành phần của ứng dụng web
3. Kiến trúc ứng dụng web

1.1 DV web & KT UD web – Giao thức HTTP

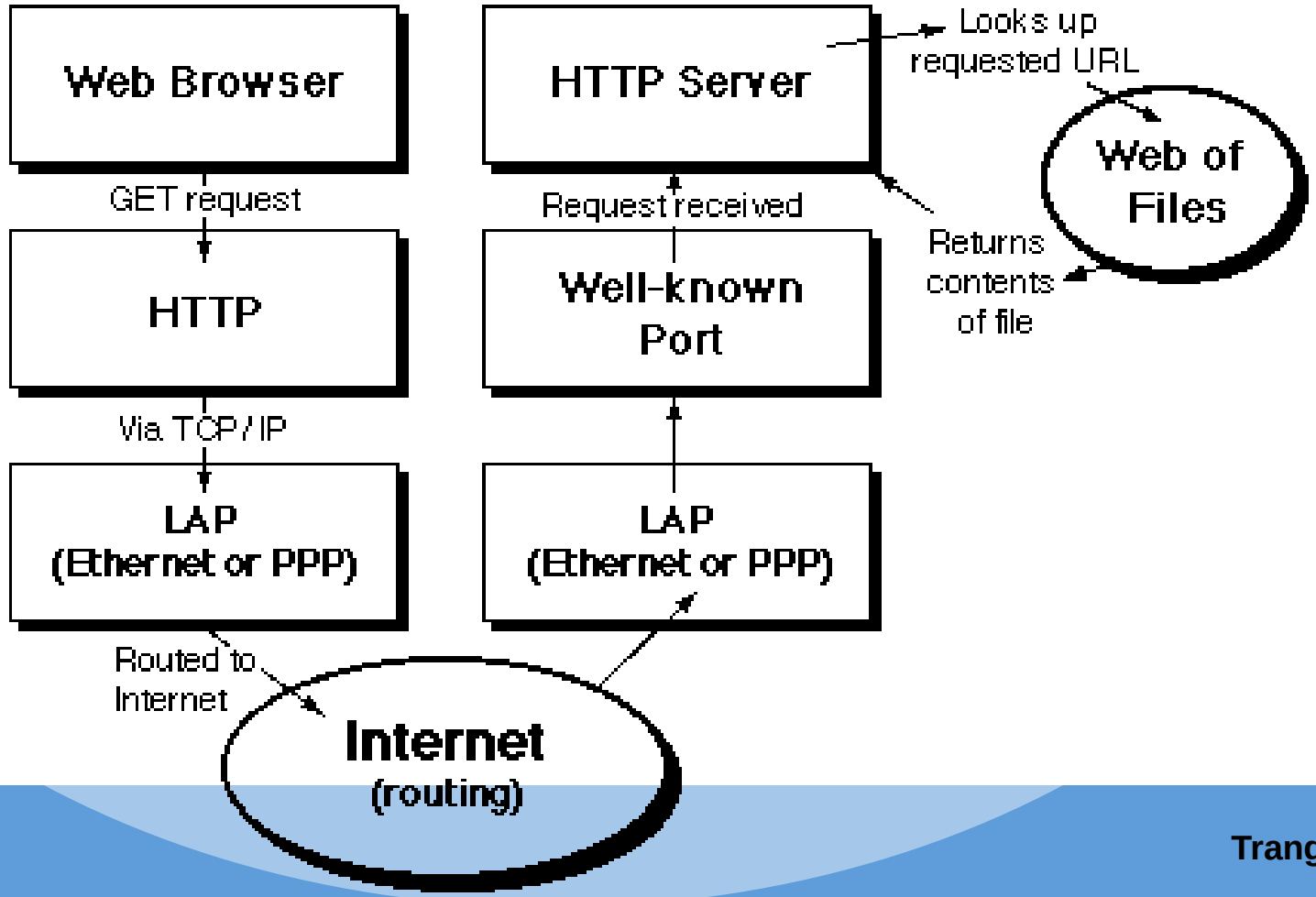
- Các ứng dụng web hoạt động dựa trên *giao thức truyền siêu văn bản* (HTTP - Hyper-Text Transfer Protocol):
  - HTTP là giao thức thuộc tầng ứng dụng của bộ giao thức TCP/IP chuyên dụng cho truyền *siêu văn bản*;
    - Cổng dịch vụ chuẩn của HTTP là 80;
  - Ngoài HTTP, HTTPS (Secure HTTP) còn được sử dụng cho các ứng dụng web có yêu cầu đảm bảo an toàn thông tin truyền giữa máy khách (Client) và máy chủ (Server);
    - Cổng dịch vụ chuẩn của HTTPS là 443.
  - HTTP hoạt động theo kiểu *yêu cầu – đáp ứng* (request - response) trong mô hình khách – chủ (client – server).

1.1 DV web & KT UD web – Giao thức HTTP



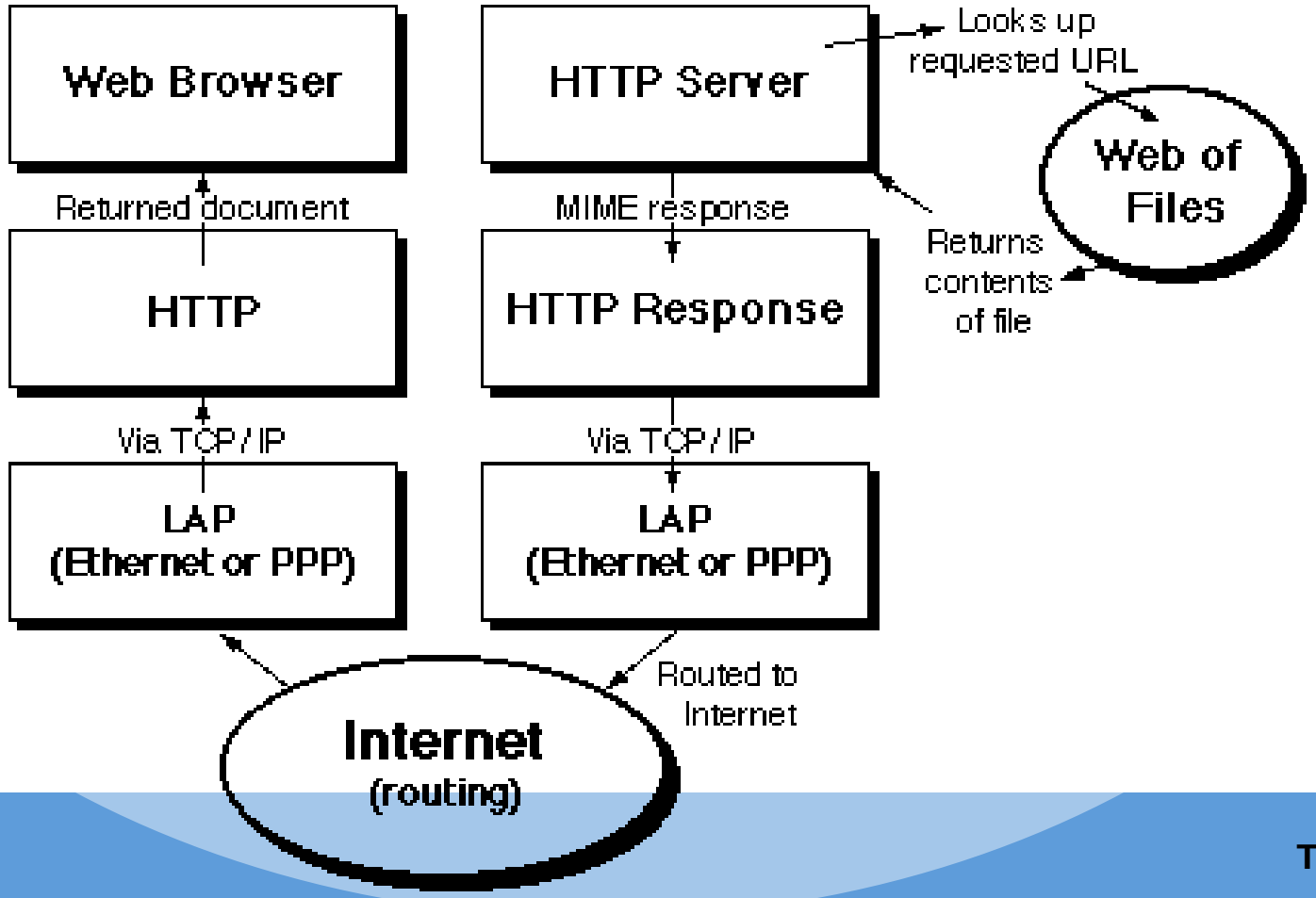
1.1 DV web & KT UD web – Giao thức HTTP

Giao tiếp giữa HTTP Client (Web Browser) và HTTP Server (Web Server): Client gửi yêu cầu (Request)



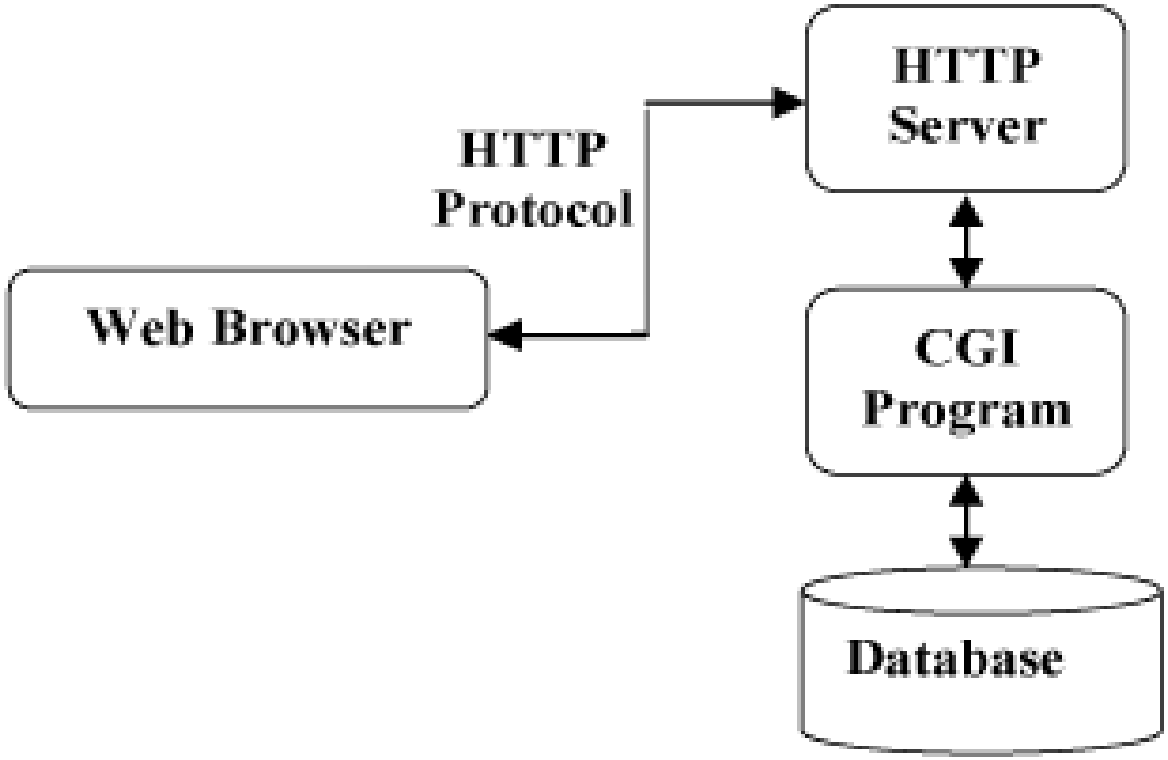
1.1 DV web & KT UD web – Giao thức HTTP

Giao tiếp giữa HTTP Client (Web Browser) và HTTP Server (Web Server): Server gửi trả đáp ứng (Response)



1.1 DV web & KT UD web – Giao thức HTTP

Giao tiếp giữa HTTP Client (Web Browser) và HTTP Server (Web Server) có sự tham gia của các chương trình chạy trên máy chủ (CGI) truy nhập cơ sở dữ liệu



1.1 DV web & KT UD web – Các thành phần của UD web

- Các thành phần của ứng dụng web:
  - Máy khách web/trình duyệt web (Web client/web browser)
  - Máy chủ web (web server)
  - URL/URI
  - Web session và cookies
  - Bộ diễn dịch và thực hiện các server scripts
  - Các server scripts (CGI – Common Gateway Interface)
  - Máy chủ CSDL
  - Hạ tầng mạng TCP/IP kết nối giữa máy khách và máy chủ web.

1.1 DV web & KT UD web – Các thành phần của UD web

- Trình duyệt web:
  - Là bộ phần mềm chạy trên máy khách có chức năng tạo yêu cầu, gửi yêu cầu và hiển thị kết quả trả về từ máy chủ web;
  - Các phương thức yêu cầu: GET, HEAD, POST
  - Có khả năng hiển thị nhiều loại dữ liệu của trang web: văn bản, hình ảnh, âm thanh, video,...
  - Hỗ trợ khả năng lập trình bằng các ngôn ngữ script (như javascript), xử lý các ngôn ngữ HTML, XML, CSS,...
  - Một số trình duyệt thông dụng: MS Internet Explorer, Google Chrome, Mozilla Firefox, Opera, Apple Safari,...

1.1 DV web & KT UD web – Các thành phần của UD web

- Máy chủ web:
  - Tiếp nhận yêu cầu từ trình duyệt web, xử lý yêu cầu và trả về đáp ứng (thường là trang web);
    - Nếu là yêu cầu truy nhập các file tĩnh, máy chủ web truy nhập hệ thống file cục bộ và gửi kết quả cho trình duyệt;
    - Nếu là yêu cầu truy nhập các file scripts, máy chủ web chuyển các scripts cho bộ xử lý scripts. Scripts có thể bao gồm các lệnh truy cập CSDL để xử lý dữ liệu. Kết quả thực hiện scripts được chuyển lại cho máy chủ web để gửi cho trình duyệt.

1.1 DV web & KT UD web – Các thành phần của UD web

- Máy chủ web:
  - Một số mã trạng thái đáp ứng:
    - 200: thành công
    - 404: lỗi không tìm thấy file/dữ liệu
    - 403: lỗi cấm truy nhập
    - 500: lỗi xử lý scripts trên máy chủ.

1.1 DV web & KT UD web – Các thành phần của UD web

- Máy chủ web:
  - Một số máy chủ web thông dụng:
    - Mozilla Apache web server
    - Microsoft Internet Information Services (IIS)
    - nginx (NGINX, Inc)
    - Google web services
    - IBM Websphere
    - Oracle web services
  - Các ngôn ngữ server scripts:
    - asp, asp.net
    - Java Servlet, JavaServer Pages
    - php, perl, python,…

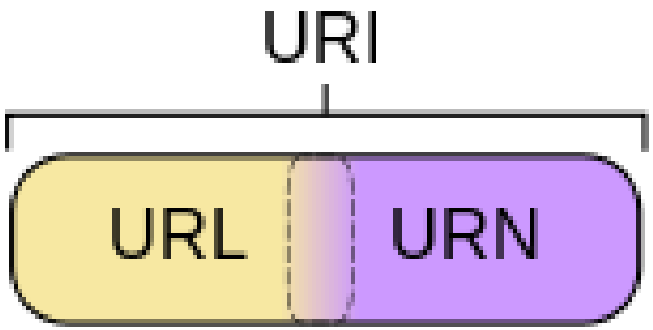
1.1 DV web & KT UD web – Các thành phần của UD web

❖ URL (Uniform Resource Locator):

- Còn gọi là địa chỉ web, là một chuỗi ký tự cho phép tham chiếu đến một tài nguyên;
- Dạng thông dụng:
  scheme://domain:port/path?query_string#fragment_id
  - scheme: chỉ giao thức truy cập (http, https, ftp,...)
  - domain: tên miền, ví dụ www.google.com
  - port: số hiệu cổng dịch vụ; với cổng chuẩn (http 80 hoặc https 443) thì không cần chỉ ra số hiệu cổng
  - path: đường dẫn đến tên file/trang
  - ?query_string: chuỗi truy vấn, gồm một hoặc một số cặp tên biến=giá trị. Ký tự và (&) được dùng để ngăn cách các cặp
  - fragment_id: một tên liên kết định vị đoạn trong trang.

1.1 DV web & KT UD web – Các thành phần của UD web

- URI (Uniform Resource Identifier):
  - Là một chuỗi ký tự dùng để nhận dạng một địa chỉ web hoặc một tên;
  - URI có thể là URL hoặc URN (Uniform Resource Name)
    - URN được dùng để nhận dạng tên của tài nguyên
    - URL được dùng để tìm địa chỉ/vị trí của tài nguyên



1.1 DV web & KT UD web – Các thành phần của UD web

❖ Web session và cookies

- Web session (phiên làm việc) là một kỹ thuật cho phép tạo ra ứng dụng web có trạng thái (stateful) trên giao thức HTTP không trạng thái (stateless);
  - Máy chủ web tạo ra và lưu một ID cho mỗi Session theo yêu cầu của máy khách;
  - Thời gian mỗi phiên tùy thuộc vào cấu hình máy chủ web.
  - Ví dụ: Sau đăng nhập thành công, máy chủ web tạo một phiên làm việc cho người dùng và không yêu cầu thông tin đăng nhập với các yêu cầu truy nhập tiếp theo cho đến khi kết thúc phiên làm việc.

1.1 DV web & KT UD web – Các thành phần của UD web

- Web session và cookies
  - Cookie (còn gọi là HTTP cookie, hay Browser cookie):
    - Là một mẩu thông tin do website gửi và được lưu trên trình duyệt của người dùng, khi người dùng thăm website;
    - Khi người dùng thăm website trong tương lai, website có thể đọc lại thông tin trong cookie để biết các hoạt động trước đó của người dùng;
    - Cookie thường được sử dụng để lưu thông tin phiên làm việc và duy trì trạng thái phiên làm việc.

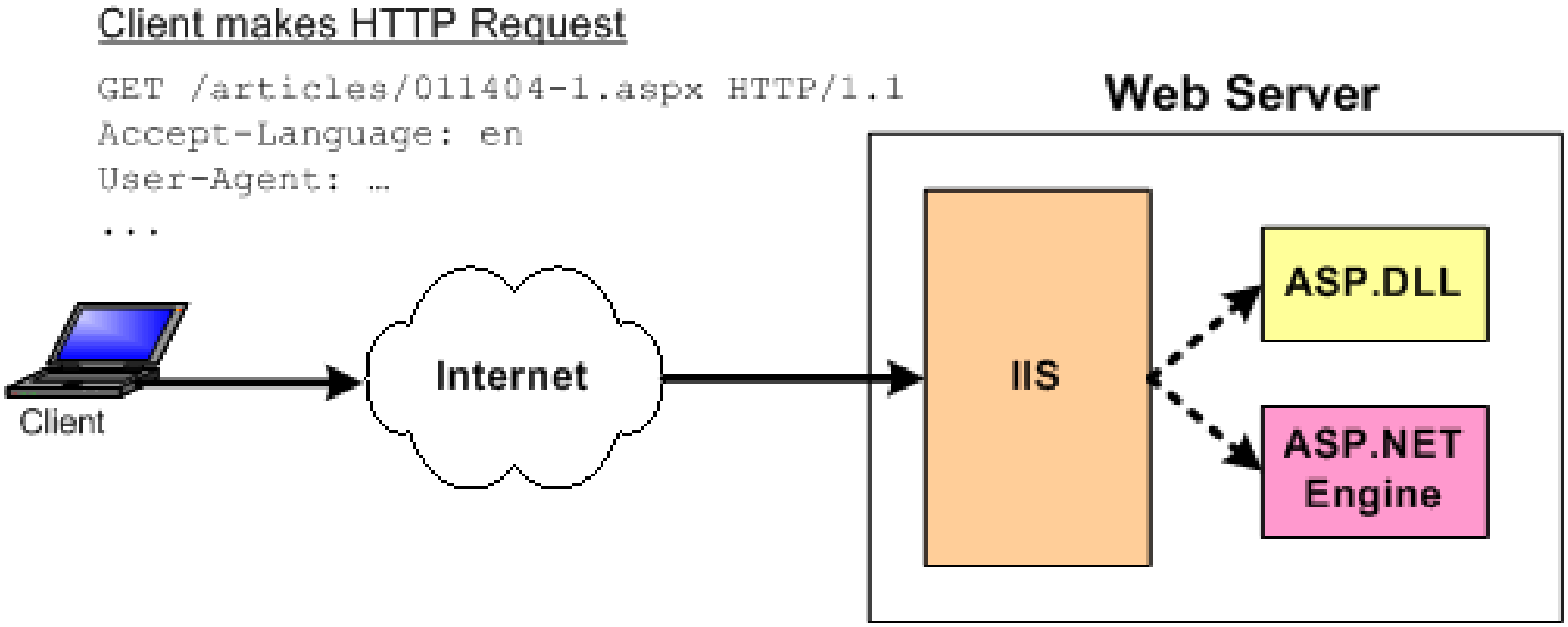
1.1 DV web & KT UD web – Các thành phần của UD web

❖ Bộ diễn dịch và thực hiện các server scripts

- Các bộ diễn dịch và thực hiện các server scripts là các engine có nhiệm vụ nạp, dịch và thực hiện từng dòng lệnh scripts trên máy chủ web;
- Do chúng làm việc theo chế độ thông dịch (interpretation) nên tốc độ thường chậm so với các ứng dụng đã được biên dịch ra mã thực hiện;
- Nhiều bộ diễn dịch và thực hiện các server scripts có thể được cài đặt và làm việc với một máy chủ web.
- Một số script engine thông dụng:
  - Microsoft ASP, ASP.NET
  - PHP engine
  - Perl, Python engine, JVM/JSP

1.1 DV web & KT UD web – Các thành phần của UD web

- Bộ diễn dịch và thực hiện các server scripts



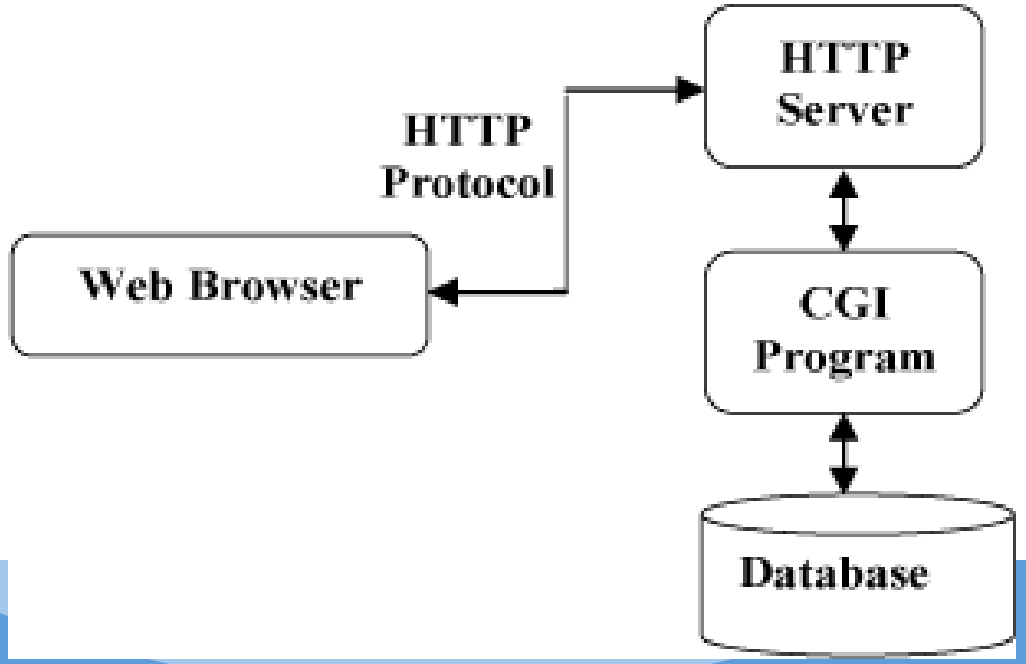
1.1 DV web & KT UD web – Các thành phần của UD web

- Các server scripts (CGI – Common Gateway Interface)
  - Các server scripts là các đoạn mã được nhúng vào các trang web HTML để thực hiện các công việc xử lý dữ liệu và trả về kết quả để tạo nội dung cho trang web;
  - Các server scripts được web server chuyển cho các script engine để dịch và thực hiện. Kết quả thực hiện scripts được chuyển lại cho web server;
  - Một số ngôn ngữ lập trình cho server scripts:
    - ASP (VBScript), ASP.NET (C#)
    - PHP
    - Perl
    - Python
    - JSP (Java),…

1.1 DV web & KT UD web – Các thành phần của UD web

❖ Máy chủ CSDL

- Máy chủ CSDL thường được sử dụng để chứa dữ liệu tạo các trang web động;
- Khi có yêu cầu truy vấn của người dùng, máy chủ web thực hiện các server scripts để truy cập và xử lý dữ liệu từ CSDL. Kết quả thực hiện scripts được chuyển lại cho web server để tạo nội dung trang web.

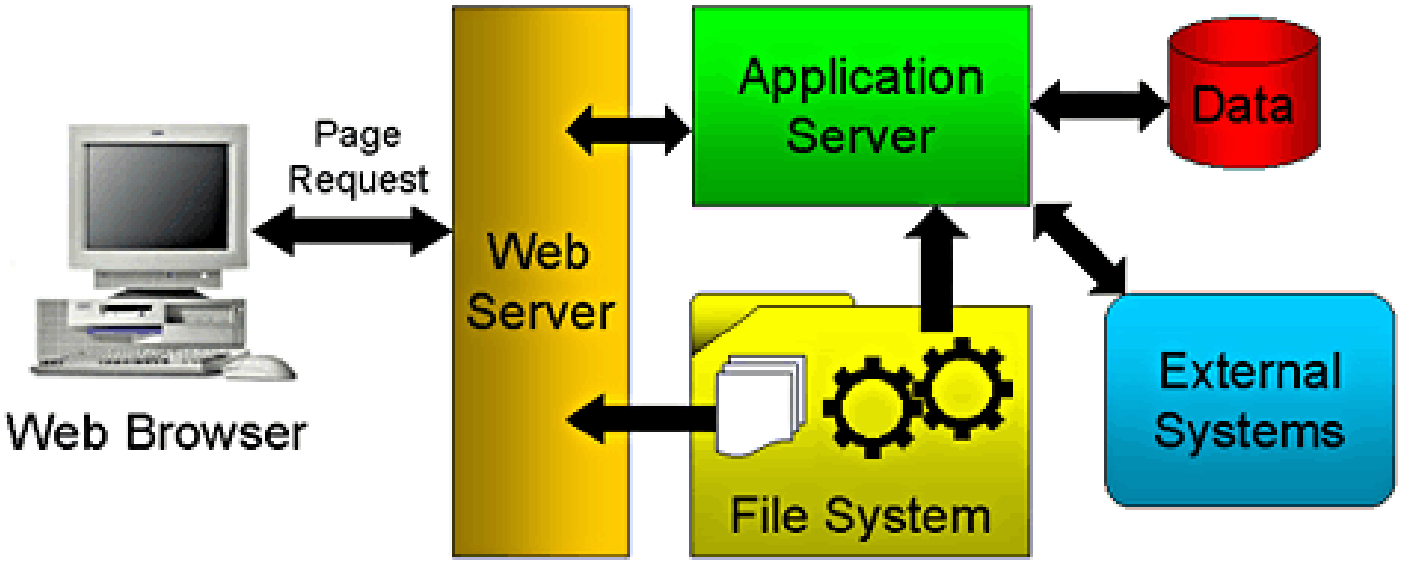


1.1 DV web & KT UD web – Các thành phần của UD web

- Hạ tầng mạng TCP/IP kết nối giữa máy khách và máy chủ web
  - Gồm tất cả các thiết bị tạo thành hệ thống truyền thông kết nối máy chủ web với máy khách web:
    - Switch
    - Router
    - Firewall
    - Cables,…

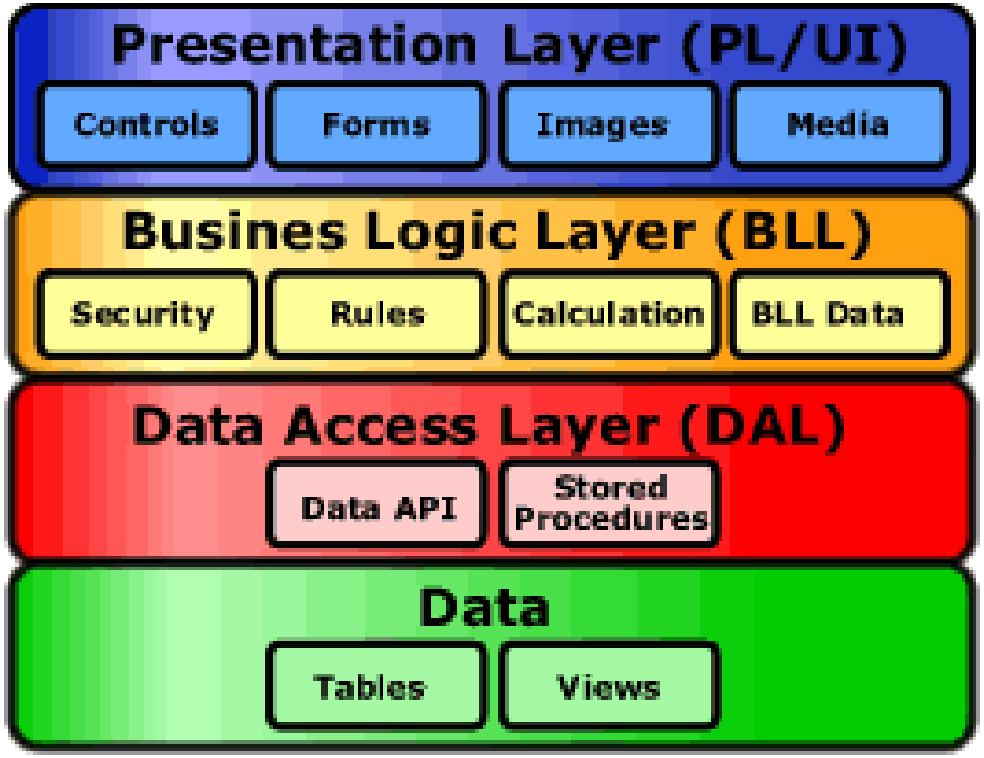
1.1 DV web & KT UD web – Kiến trúc của UD web

❖ Kiến trúc chuẩn của ứng dụng web



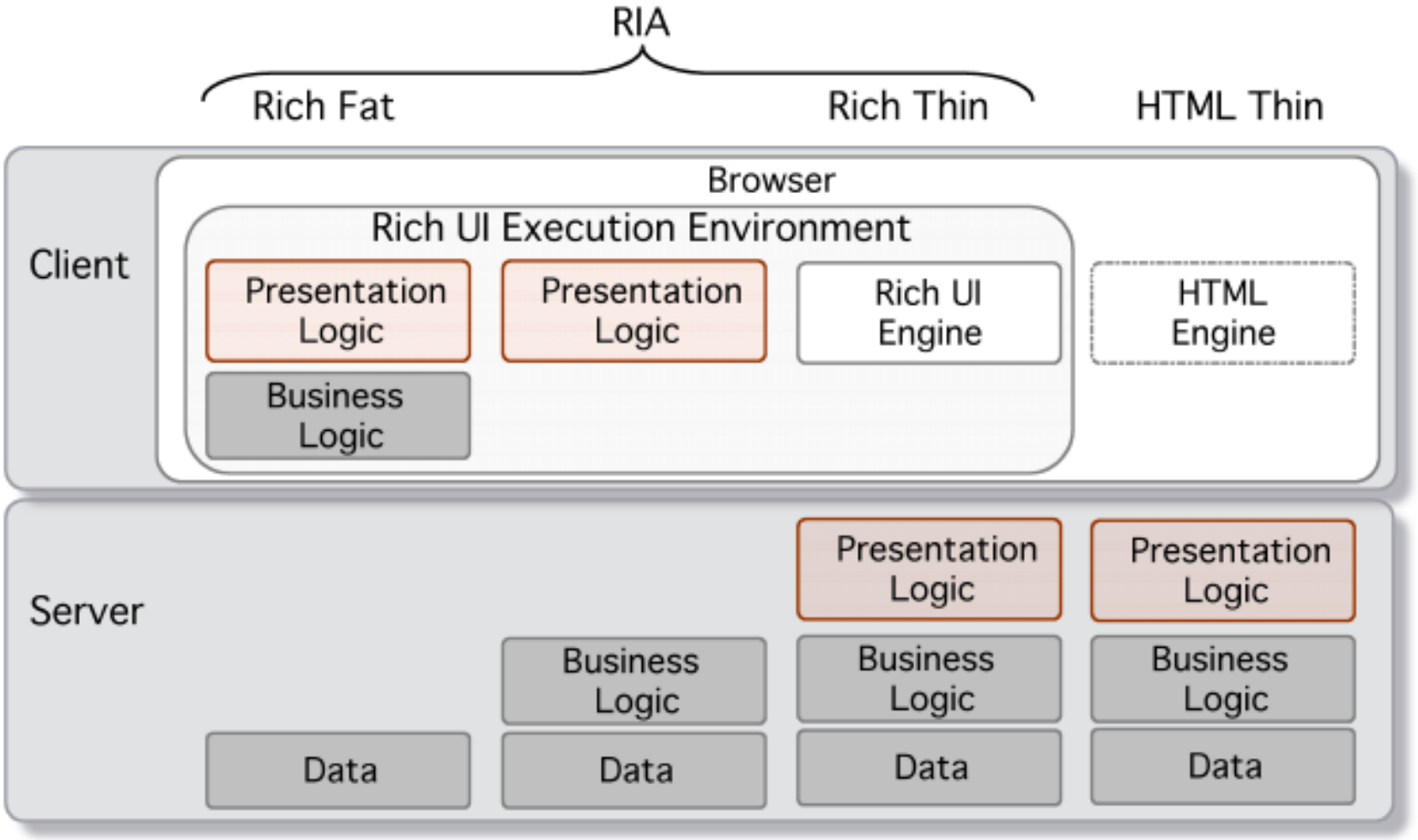
1.1 DV web & KT UD web – Kiến trúc của UD web

- Kiến trúc logic 3 lớp (3-tier) của ứng dụng web: Lớp trình diễn (Presentation Layer), lớp Business Logic và lớp truy nhập dữ liệu (Data Access Layer)



1.1 DV web & KT UD web – Kiến trúc của UD web

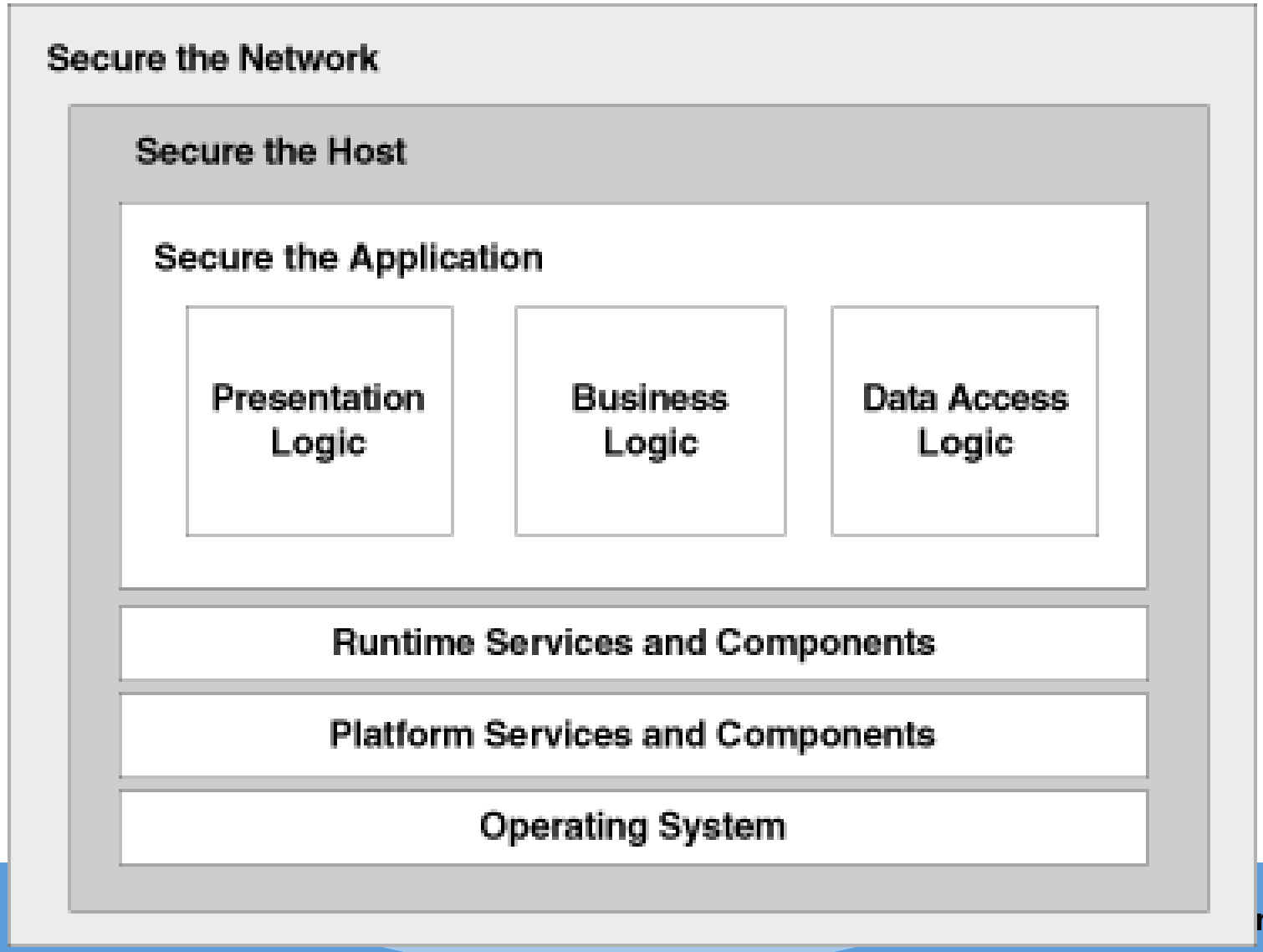
- Các dạng kiến trúc logic ứng dụng web



1.2 Các nguyên tắc bảo mật các ứng dụng Web

- Áp dụng nguyên tắc phòng vệ nhiều lớp, có chiều sâu (Defence in depth):
  - Lớp bảo mật mạng (Network)
  - Lớp bảo mật máy chủ (Host)
  - Lớp bảo mật ứng dụng (Application)

1.2 Các nguyên tắc bảo mật các ứng dụng Web

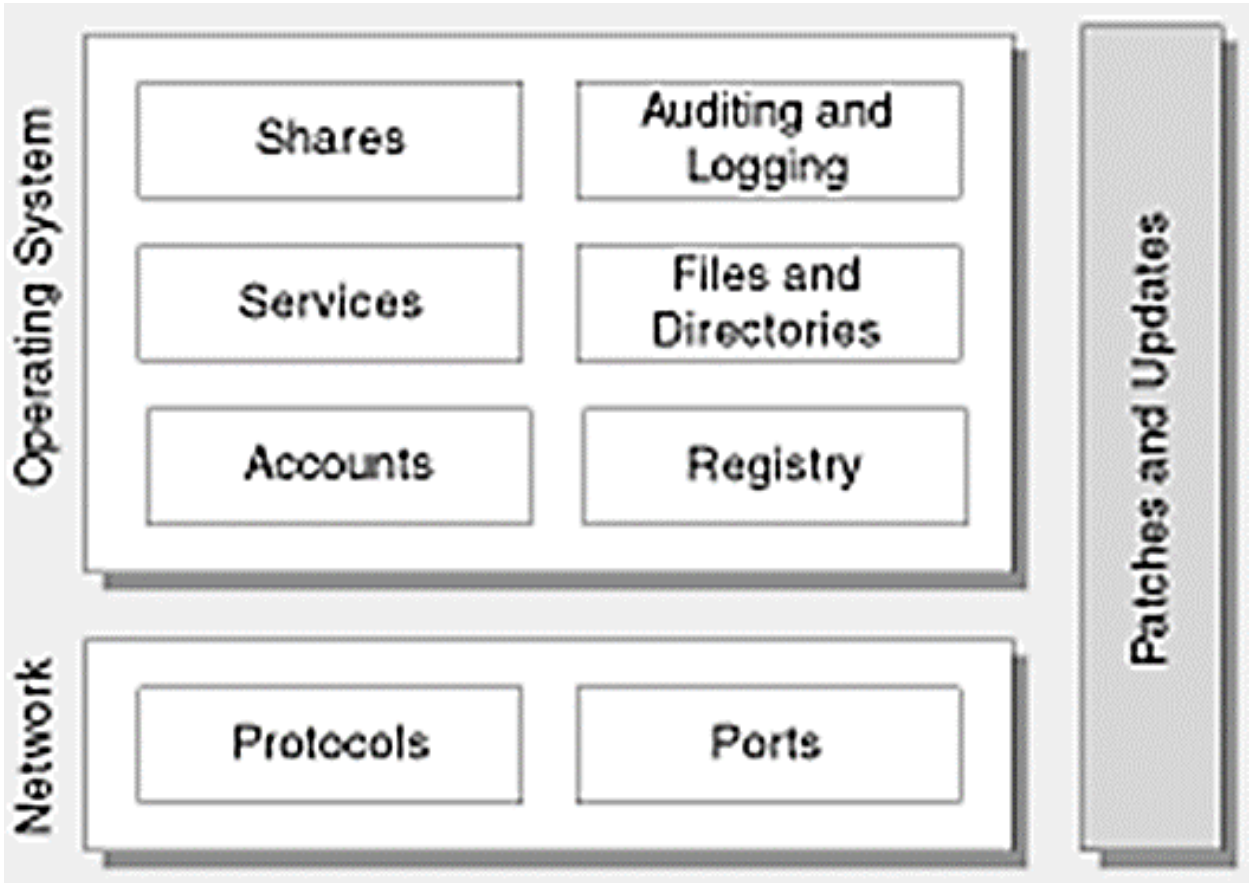


1.2 Các nguyên tắc bảo mật các ứng dụng Web

- Lớp bảo mật mạng (Network)
  - Các ứng dụng web cần hạ tầng mạng an toàn cho giao tiếp giữa máy chủ và máy khách;
  - Các thiết bị mạng cần được cài đặt và cấu hình theo chuẩn, đảm bảo an toàn:
    - Switch: bộ chuyển mạch
    - Router: bộ định tuyến
    - Firewall: tường lửa
    - IPS/IDS: hệ thống ngăn chặn/phát hiện đột nhập

1.2 Các nguyên tắc bảo mật các ứng dụng Web

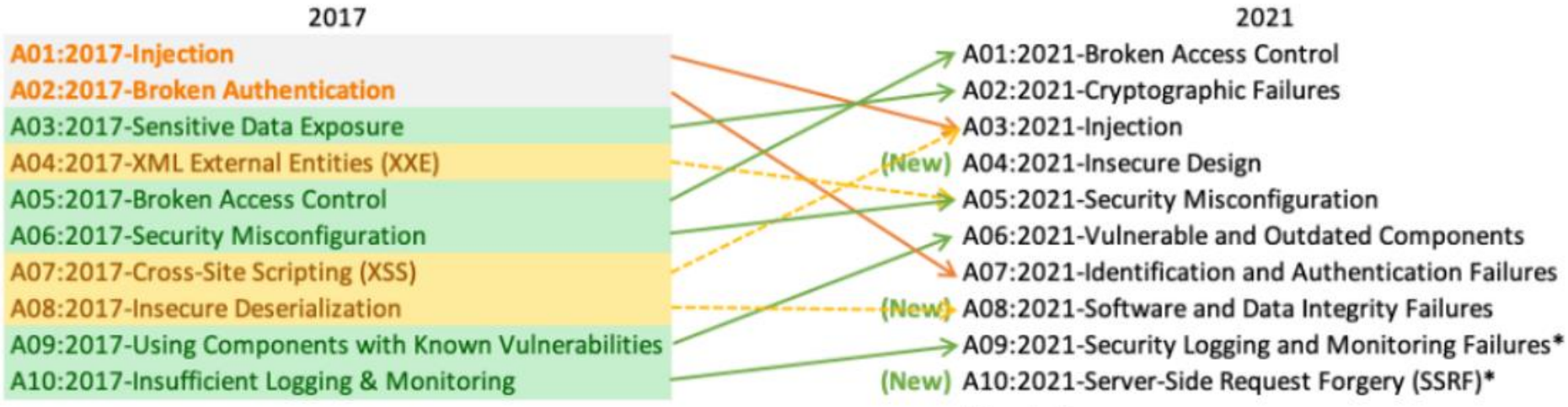
- Lớp bảo mật máy chủ (Host)
  - Bảo mật hệ điều hành
  - Bảo mật CSDL
  - Bảo mật các phần mềm/dịch vụ hệ thống



1.2 Các yêu cầu bảo mật các ứng dụng Web

- Lớp bảo mật ứng dụng (Application)
  - Xác thực/trao quyền
  - Cấu hình
  - Kiểm tra dữ liệu đầu vào
  - Quản lý phiên làm việc
  - Mã hóa dữ liệu
  - Quản lý các ngoại lệ
  - Ghi logs

1.3 Các nguy cơ và lỗ hổng bảo mật trong các UD Web



Top 10 lỗ hổng bảo mật ứng dụng web theo OWASP (https://owasp.org/Top10/)

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

- Injection (Chèn mã):
  - Buffer overflow
  - SQL injection
  - XPath/XQuery injection
  - LDAP lookups / injection
  - Shell command injection

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

- Xác thực và quản lý phiên yếu:
  - Khâu xác thực (authentication) và trao quyền (authorisation) được sử dụng khá phổ biến trong các ứng dụng web;
  - Nếu các khâu xác thực không đủ mạnh → là lỗ hổng để kẻ tấn công truy nhập đánh cắp thông tin.
  - Phiên làm việc (session) cũng cần được quản lý chặt chẽ;
  - Nếu không kẻ tấn công có thể lợi dụng để chiếm và điều khiển phiên làm việc của người dùng.
  - VD: đưa ID của phiên lên URL mà không có mã hóa, kiểm tra:

    http://www.error-site.com/test.aspx?session_id=12345

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

❖ Rò rỉ dữ liệu nhạy cảm:

- Nhiều ứng dụng web không có các cơ chế đủ mạnh để bảo vệ các dữ liệu nhạy cảm, như thông tin thẻ tin dụng, số an sinh xã hội và thông tin xác thực người dùng.
- Kẻ tấn công có thể đánh cắp, hoặc chỉnh sửa các thông tin nhạy cảm thể lạm dụng, hoặc trục lợi.
- Các dữ liệu nhạy cảm như mật khẩu cần được lưu dưới dạng mã hóa;
  - Mật khẩu Nên dùng các hàm băm 1 chiều (SHA)
- Hạn chế quyền truy nhập vào các files chứa thông tin nhạy cảm (lưu mật khẩu,…)

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

❖ Lỗi xử lý các thành phần XML bên ngoài (XXE)

- Tấn công khai thác lỗi XXE xảy ra khi một số bộ xử lý tài liệu XML cũ hoặc được cấu hình kém xử lý các tài liệu XML có chứa tham chiếu đến tham chiếu đến các thực thể bên ngoài.
- Các trường hợp khai thác lỗi XXE có thể gồm:
  - Các bộ xử lý XML có lỗi cho phép kẻ tấn công tải lên tài liệu XML hoặc nội dung độc hại trong tài liệu XML;
  - Lỗ hổng bảo mật trong mã chương trình;
  - Lỗ hổng bảo mật trong các thành phần phụ thuộc;
  - Lỗ hổng bảo mật trong tích hợp hệ thống.

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

- Kiểm soát truy cập yếu:
  - Việc hạn chế các tác vụ mà người dùng đã xác thực được phép thực hiện thường không được thực thi đúng cách;
  - Kẻ tấn công có thể khai thác những lỗ hổng này để truy cập vào chức năng và / hoặc dữ liệu trái phép, chẳng hạn như truy cập tài khoản của người dùng khác, xem các file nhạy cảm, sửa đổi dữ liệu của người dùng khác, thay đổi quyền truy cập, v.v.

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

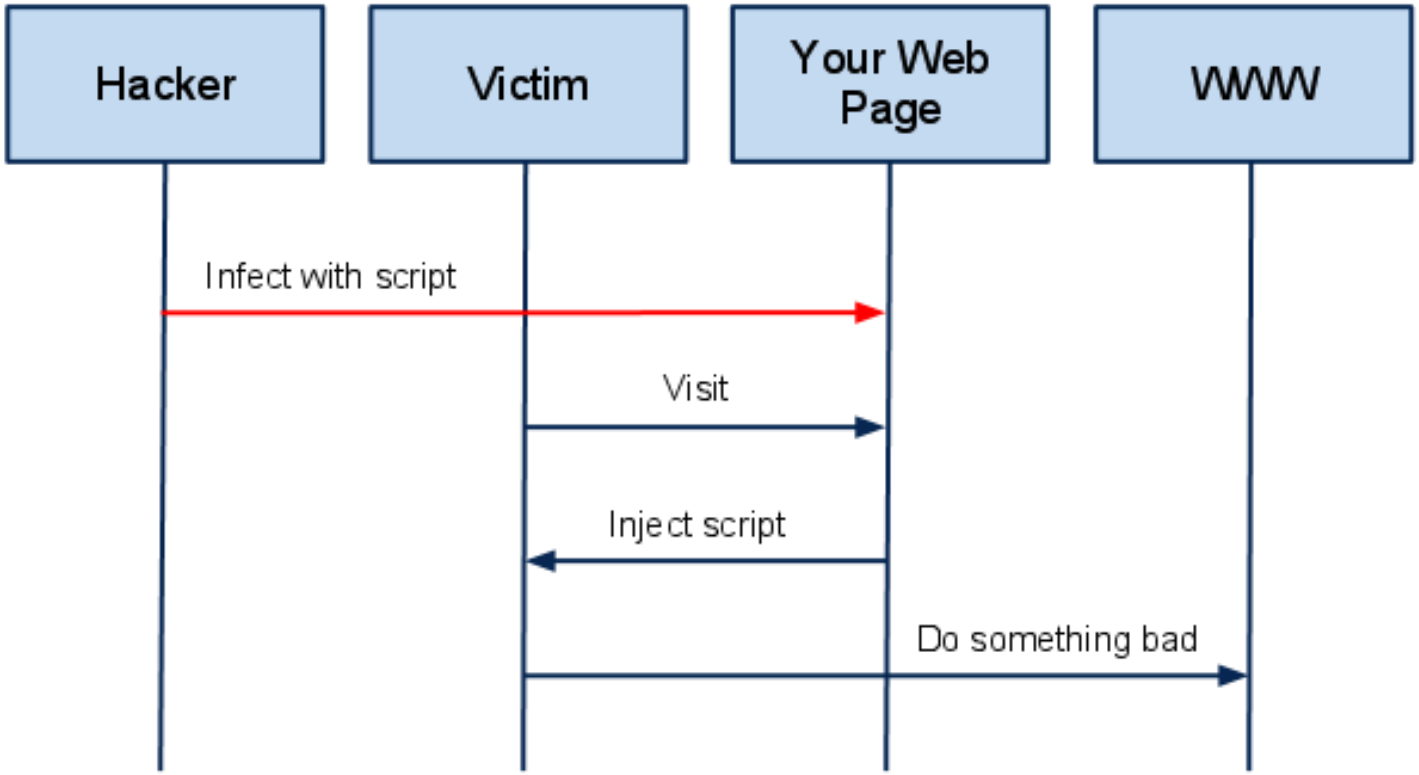
- Lỗi cấu hình an ninh: Đây là dạng lỗi khá thường gặp, bao gồm:
  - Cấu hình ngầm định không an toàn hoặc không đầy đủ
  - Lưu trữ trên đám mây mở
  - HTTP header cấu hình sai
  - Thông báo lỗi chứa thông tin nhạy cảm
  - Hệ điều hành, các framework, thư viện, ứng dụng không được cấu hình đúng, hoặc không được cập nhật.

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

- XSS – Cross-Site Scripting:
  - Là dạng tấn công trình duyệt người dùng
  - Kẻ tấn công chèn mã script (thường là javascript) vào các trang web có lỗi XSS
  - Khi người dùng mở các trang này thì mã script của kẻ tấn công được thực hiện giúp đánh cắp thông tin lưu trong trình duyệt người dùng.

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

❖ XSS – Cross-Site Scripting:



A High Level View of a typical XSS Attack

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

❖ Giải tuần tự hoá không an toàn :

- Tuần tự hoá và giải tuần sự hoá (serialization/ deserialization):
  - Tuần tự hoá là quá trình chuyển đối tượng thành chuỗi byte;
  - Giải tuần tự hoá là quá trình chuyển chuỗi byte thành đối tượng.
- Nguy cơ:
  - Giải tuần tự hoá không an toàn có thể dẫn đến k. năng thực hiện mã từ xa;
  - Giải tuần tự hoá không an toàn cũng có thể bị sử dụng để thực hiện tấn công, như tấn công phát lại, chèn mã, hoặc leo thang đặc quyền.

1.3 Các nguy cơ và lỗ hổng bảo mật–Top 10 OWASP 2017

❖ Giải tuần tự hoá không an toàn :

- Ví dụ:
  - Một trang web PHP sử dụng kỹ thuật tuần tự hoá đối tượng để chuyển thông tin phiên người dùng (gồm ID, tên, mật khẩu băm,…) thành 1 chuỗi và lưu dưới dạng 1 cookie như sau:

    a:4:{i:0;i:132;i:1;s:7:”Mallory”;i:2;s:4:”user”; i:3;
    s:32:”b6a8b3bea87fe0e05022f8f3c88bc960″;}

  - Một tin tặc có thể chỉnh sửa chuỗi trên để biến hắn thành người quản trị:

    a:4:{i:0;i:1;i:1;s:5:”Alice”;i:2;s:5:”admin”;
    i:3;s:32:”b6a8b3bea87fe0e05022f8f3c88bc960″;}

  - Nếu tin tặc có thể chỉnh sửa thông tin và giải tuần từ (chuyển chuỗi byte thành đối tượng phiên) thì có thể chiếm quyền người quản trị và chiếm quyền điều khiển ứng dụng web.

1.3 Các nguy cơ và lỗ hổng bảo mật trong các UD Web

- Sử dụng các thành phần chứa lỗ hổng đã biết:
  - Các thành phần, bao gồm các thư viện, các framework và các mô đun phần mềm hầu như được chạy với quyền truy cập đầy đủ như người dùng kích hoạt ứng dụng;
  - Nếu một thành phần có chứa lỗ hổng bị khai thác có thể gây ra việc mất mát nhiều dữ liệu, hoặc máy chủ có thể bị chiếm quyền điều khiển.
  - Các ứng dụng sử dụng các thành phần chứa lỗ hổng đã biết có thể làm suy giảm khả năng phòng vệ của ứng dụng và cho phép thực hiện nhiều loại tấn công lên hệ thống.

1.3 Các nguy cơ và lỗ hổng bảo mật trong các UD Web

- Thiếu cơ chế giám sát và ghi log:
  - Thiếu cơ chế giám sát và ghi log cùng với việc thiếu hoặc có cơ chế phản ứng kém hiệu quả cho phép kẻ tấn công thực hiện tấn công tiếp tục vào các hệ thống, hoặc duy trì kiểm soát hệ thống;
  - Nhiều nghiên cứu về tấn công, khai thác cho thấy thời gian trung bình cho phát hiện các vi phạm là khoảng 200 ngày, và chủ yếu được phát hiện bởi bên ngoài, không phải do bên trong hệ thống nhờ việc giám sát và ghi log.

1.4 Các phương pháp tiếp cận bảo mật các ứng dụng Web

- Luôn thực hiện kiểm tra dữ liệu đầu vào
  - Không bao giờ tin người dùng
  - Kiểm tra kích thước, định dạng và nội dung dữ liệu
  - Sử dụng các bộ lọc
- Giảm thiểu các giao diện có thể bị tấn công
  - Hạn chế người dùng truy nhập trực tiếp vào các hệ thống CSDL
  - Phân quyền truy nhập ở mức "vừa đủ" cho công việc.
- Phòng vệ có chiều sâu.