Authentication

Mariano Ceccato

mariano.ceccato@univr.it

Passwords

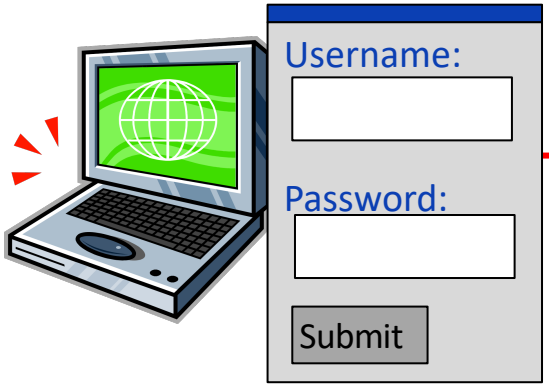
Lecture Outline

- Password Based Authentication
- Password Attacks
- Possible Countermeasures

User Authentication

- Three possible approaches
  - Something the user **knows**
    - **e.g PIN, password**
  - Something the user **has**
    - **E.g smart card, a key, a token**
  - Something the user **is**
    - **e.g fingerprint, face, voice**

Password authentication



1. End-user enter username and password

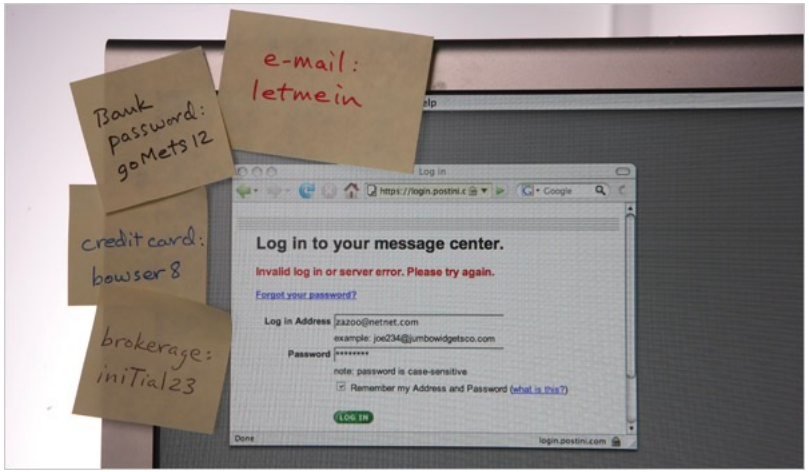
2. Client sends username/password across the network

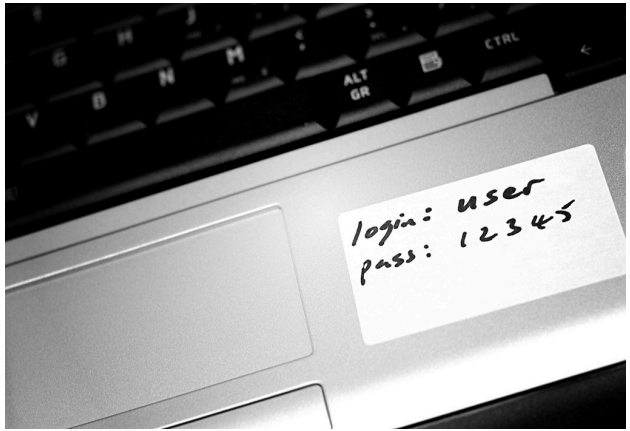
3. The server checks username/password

4. The authenticated identity is granted access to the system

Password overload problem

- Password is annotated (on the monitor, under the keyboard)
- The same password is reused on many different services
- Passwords are rarely changed
  - Online databases of stolen/guessed passwords





Predictable passwords

	2015	2014	2013	2012	2011
#1	123456	123456	123456	password	password
#2	password	password	password	123456	123456
#3	12345678	12345	12345678	12345678	12345678
#4	qwerty	12345678	qwerty	abc123	qwerty
#5	12345	qwerty	abc123	qwerty	abc123
#6	123456789	1234567890	123456789	monkey	monkey
#7	football	1234	111111	letmein	1234567
#8	1234	baseball	1234567	dragon	letmein
#9	1234567	dragon	iloveyou	111111	trustno1
#10	baseball	football	adobe123	baseball	dragon
#11	welcome	1234567	123123	iloveyou	baseball
#12	1234567890	monkey	admin	trustno1	111111
#13	abc123	letmein	1234567890	1234567	iloveyou

	2015	2014	2013	2012	2011
#14	111111	abc123	letmein	sunshine	master
#15	1qaz2wsx	111111	photoshop	master	sunshine
#16	dragon	mustang	1234	123123	ashley
#17	master	access	monkey	welcome	bailey
#18	monkey	shadow	shadow	shadow	passw0rd
#19	letmein	master	sunshine	ashley	shadow
#20	login	michael	12345	football	123123
#21	princess	superman	password1	jesus	654321
#22	qwertyuiop	696969	princess	michael	superman
#23	solo	123123	azerty	ninja	qazwsx
#24	passw0rd	batman	trustno1	mustang	michael
#25	starwars	trustno1	000000	password1	football

Password Reuse

- On average, just 6 unique passwords are used on 24 only accounts
  - Remembering a strong password is hard, so users just recuse the same password multiple times
- Once a password leaks, an adversary can attempt to reuse it on different services

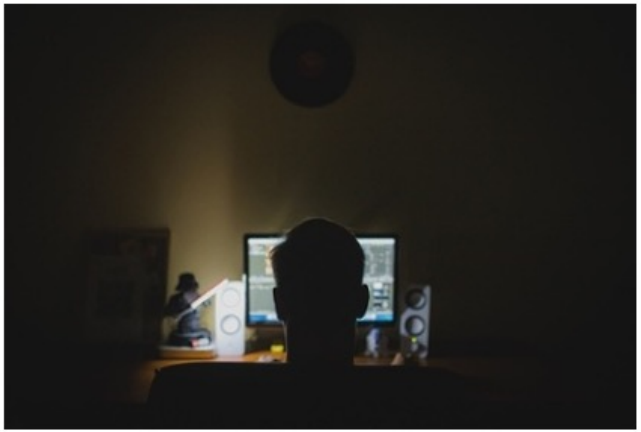
Some statistics

- At least 65% of people reuse passwords across multiple sites.
- A 13% of people use the same password for all passworded accounts and devices.
- Unfortunately, 48% of workers use the same passwords in both their personal and work accounts.
- The average person reuses each password 14 times
- An estimated 49% of employees only add a digit or change a character in their password when they're required to update it.
- Although 91% of participants in a recent survey understand the risk of password reuse, 59% admitted to doing it anyway.

63% of Data Breaches Result From Weak or Stolen Passwords

In its recent **2016 Data Breach Investigations Report**, Verizon Enterprise confirmed many industry trends that we see at ID Agent every day. The most glaring blind spot for organizations is how stolen credentials are the primary means by which hackers exploit their vital systems.

Credentials are the holy grail for hackers. In a study of 905 phishing attacks, the vast majority—91 percent—were after user credentials.



Stolen Credentials Are a Big Problem That You May Not Know About

2014 eBay Data Breach

- In May 2014 eBay announced that they suffered a cyber attack
- The attack compromised a database containing encrypted passwords as well as other non-financial data
- The cyber attack was carried out after a small number of employee login credentials were stolen
- eBay forced the owners of compromised accounts to change their passwords.

2014 iCloud Attack

- August 2014 the Apple's iCloud account were hacked
- A collection of almost 500 private pictures of various celebrities, mostly women were posted on image boards and social media
- Victims' iCloud account information was obtained using phishing and brute force guessing
- Attacker indicated that one user created a fake email account called **appleprivacysecurity** to ask celebrities for security information.

How passwords are cracked

- **Brute force**: Automated guessing of billions of passwords until the correct string is found
- **Shoulder surfing**: observing someone typing their password
- **Searching**: IT infrastructure can be searched for electronically store password
- **Guessing**: personal information (name, birthdate, pet name) can be used to guess common passwords
- **Stealing**: insecurely stored passwords (handwritten near to the device) can be stolen
- **Key logging**: An installed keylogger intercepts passwords as they are typed
- **Interception**: password can be intercepted as they are transmitted over the network
- **Social engineering**: attackers use social engineering techniques to trick users into revealing their passwords

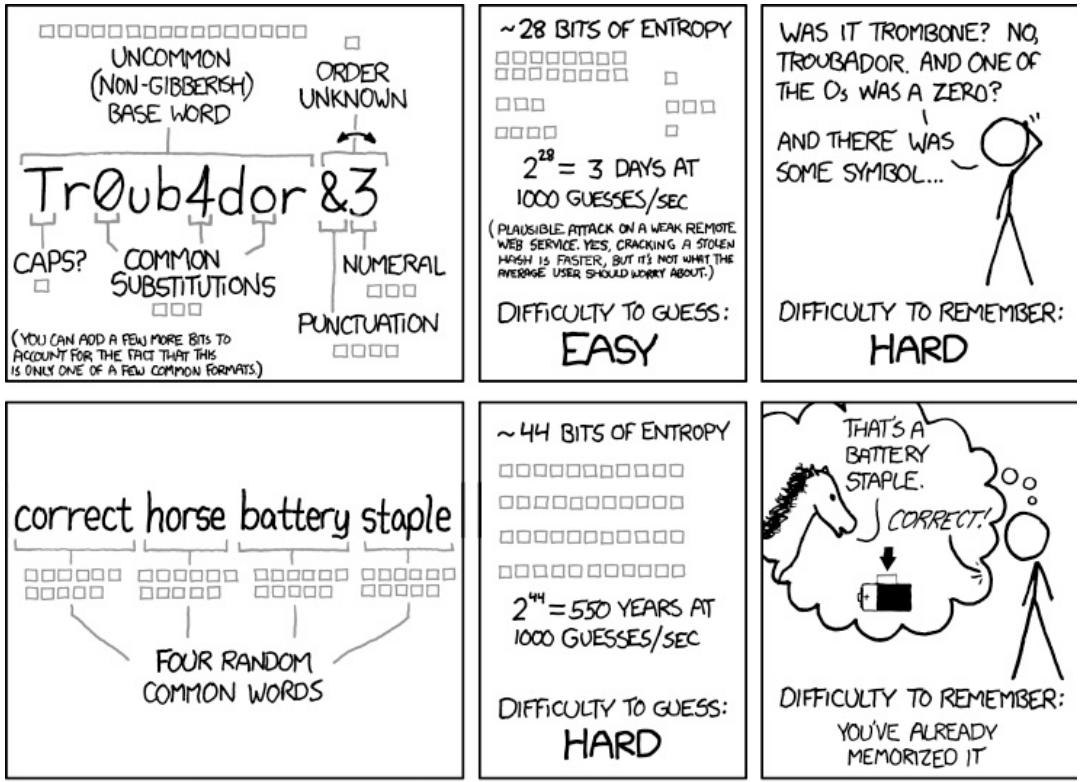
Brute Force Attacks

- Exhaustive search
  - Try all possible combinations of symbols up to a certain length
  - The size of the password space is $|A|^n$
- Assume a 8 characters password
  - Upper- and lowercase letters, digits, common symbols (96 possible characters)
  - 96^8 = 7.2 quadrillion password combinations

Brute Force Attacks – time required

Password	Length	Time to brute force
abcdefg	7 characters	.29 milllisec
abcdefgh	8 characters	5 hours
abcdefghi	9 characters	5 days
abcdefghij	10 characters	4 months
abcdefghijk	11 characters	10 years
abcdefghijkl	12 characters	2 centuries

Password metrics - xkcd



Password metrics

- Password **strength** measure the effectiveness of a password against **brute force** attack
- It estimates the number of trials an attacker has to make to guess the password correctly
- It is normally computed as $|A|^n$
  - A is the set of symbols composing the password
  - N is the length of the password
- Another measure is **entropy**
  - $\mathbf{Log_2}\ |A|^n = n\ \mathbf{Log_2}\ |A| = n\ \mathbf{Log}\ |A|/\text{Log}\ 2$
  - It is typically measured in bits
  - If the entropy of a password is **b** bits it means the attacker requires to 2^b attempts

Online dictionary attack

- Try passwords associated with the user
  - e.g name, name of friends, car brand
- Try words in a dictionary
- Try popular passwords

- Save attacker's time
- No guarantee the right password is found

Popular password guessing tools

- **Hydra – https://www.thc.org/thc-hydra**
  - Fast network logon password cracking tool
- **John The Ripper – http://www.openwall.com/john**
  - Password cracker to detect weak passwords
- **Cain and Abel - http://www.oxid.it/ca_um/**
  - Password cracking tool for Windows platform
- **Medusa – http://foofus.net/jmk/tools/medusa-1.4.tar.gz**
  - Password cracking tool that supports parallel attacks

Possible Countermeasures

- **Password policies**
  - Set password length: minimal password length should be prescribed
  - Set Password format: mix upper and lower case symbols, numerical, and non-alphabetical symbols
  - Avoid obvious passwords:12345, Forever1, John3:16, Monster1, Chicken1, …
- **Changing passwords**
  - Force users to change password regularly
- **Machine-generated passwords**
  - Pronounceable passwords are generated for the user

Limited effectiveness

Other countermeasures (more effective)

- **Lockout mechanics**
  - Lock user account after several unsuccessful login attempts
- **Throttling**
  - Time delays are introduced between consecutive failed login attempts
- **Protective monitoring**
  - Monitoring login to detect unusual use
  - Notify the user with details of attempted login
- **Password blacklisting**
  - Check if an input password is in a list of common words.

Offline dictionary attack

- Attacker gains access to the password file
- Attacker obtains “encrypted password” or “hashed password”
- Attacker tries passwords from a “dictionary” of commonly used passwords and compares with encrypted or hashed password

- These attacks with current processor speeds take hours or days or even less

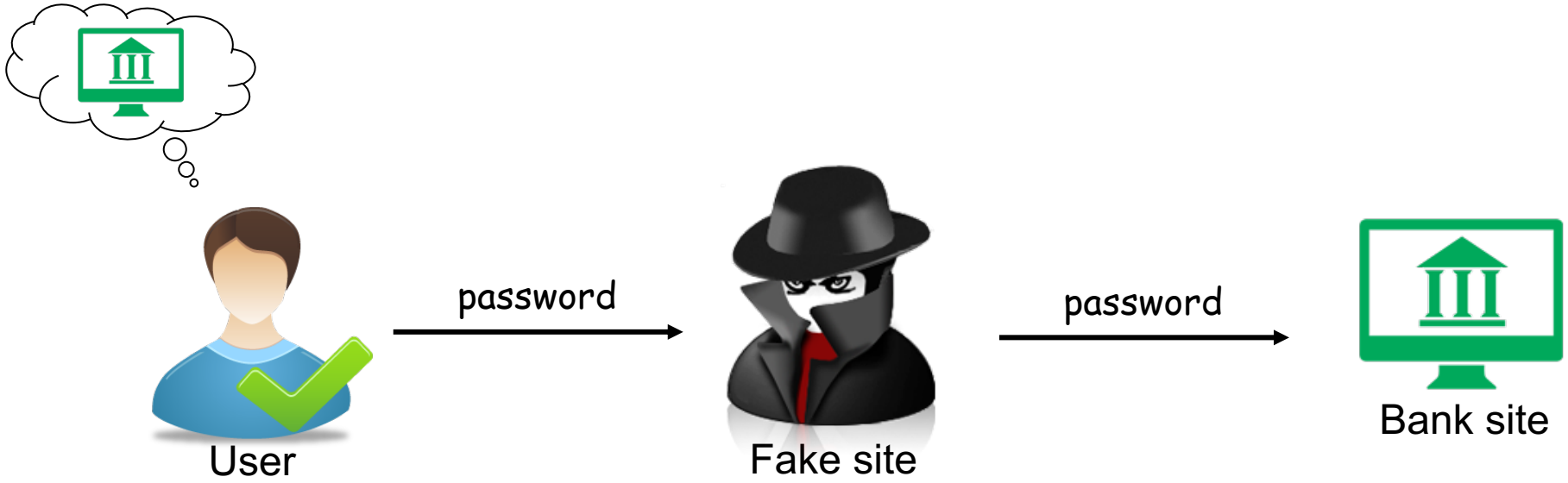
Countermeasures

- **Password salting**
  - Append a random number (salt) to the password
  - If the salt is *b* bits, the number of possible passwords is increased of factor 2^b
  - Dictionaries of hashed common passwords is useless

- **Password file access control**
  - Restrict access only to privileged users
  - Keep the hashed passwords separated from user IDs

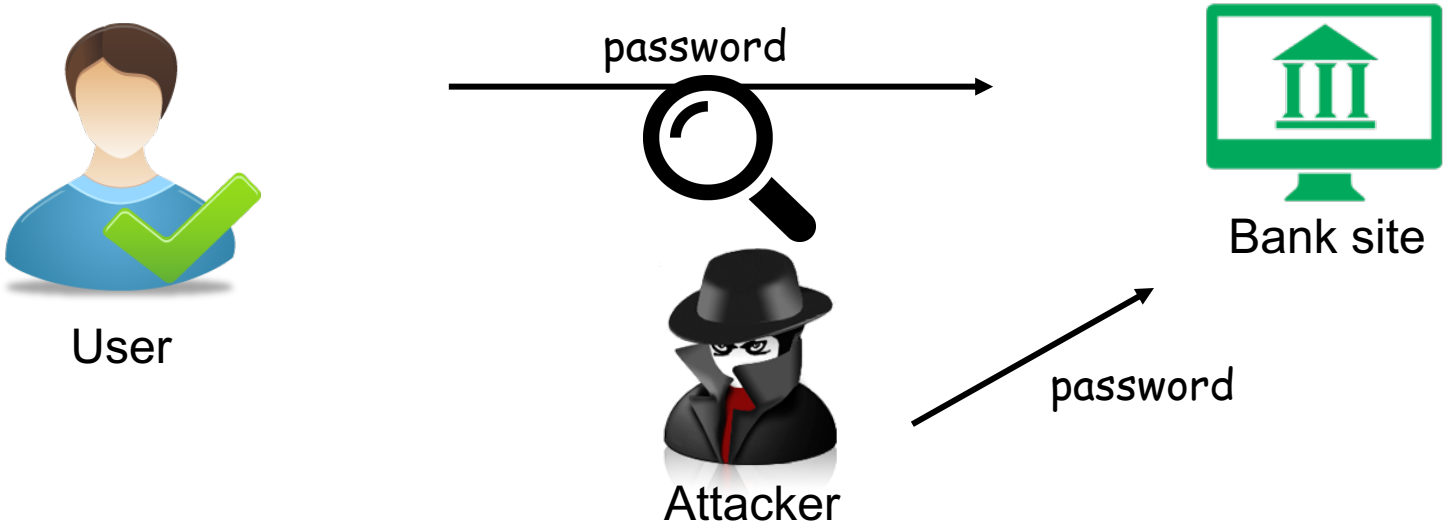
- **Fast reissuance of password**

Social engineering: Phishing



- Captured password can be used at target site
- **Countermeasure**: server-side authentication, e.g. SSL/TLS

Interception



- Clear text password is intercepted by the attacker
- **Countermeasure:** Encrypt the communication among users and web site, e.g. SSL/TLS protocols

Keylogger

- Small program that monitors each keystroke the user types on his keyboard
- Installed by attaching the program to an image or file and then send it via email
- Popular keyloggers
  - Refog
  - Revealer
  - KidLogger

Other attacks

- **Shoulder-surfing**
  - Attacker gathers passwords by watching over a person's shoulder while he/she is logging in
- **Dumpster-diving**
  - Attacker look into the trash for piece of papers or documents with written passwords
- **Countermeasure**: User awareness and training

Summary

- Password based authentication systems are not secure
  - Users use ease to guess passwords
  - Users reuse passwords across multiple web sites
- Password based authentication systems are vulnerable to various attacks
  - Social engineering and data breaches are on top of the list
- Effective countermeasures are
  - Account lockout and throttling
  - Predictive monitoring
  - Password blacklisting.

Biometrics

Authentication

- The certification of user identity relies on (a combination of)
  - something the person knows (e.g., a password)

  - something the person has (e.g., a smart card, a radio key fob),

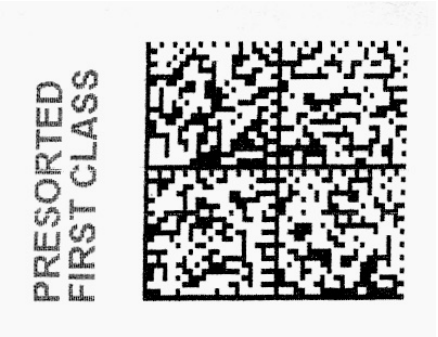
  - something the person is (e.g., a fingerprint, an eye retina)

Token-based authentication

- User relies on a token to be authenticated
  - Barcodes
  - One time password (OTP) devices
  - Magnetic stripe cards
  - Smart cards

Barcodes

- Developed in the 20th century to improve efficiency in grocery checkout.
- First-generation barcodes represent data as a series of **variable-width, vertical lines** of ink
  - One-dimensional encoding scheme.
- More recent barcodes are rendered as **two-dimensional patterns** using dots, squares, or other symbols that can be read by specialized optical scanners, which translate a specific type of barcode into its encoded information.



Authentication via Barcodes

- Since 2005, the **airline industry** incorporated two-dimensional barcodes into boarding passes, which are created at flight check-in and scanned before boarding
- This barcode encodes a **unique identifier** to let airport security look up the corresponding passenger's record with that airline
- Staff then verifies that the boarding pass was in fact purchased in that person's name (using the airline's database), and that the person can provide **photo id** for identification
- Barcodes are convenience but not secure
  - Barcodes are simply images, they are extremely easy to duplicate.

Authentication via OTP Devices

- Single factor OTPs
  - Generated a number every 30 or 60 seconds
  - Embed a secret that is used to generate the OTP

- Multi factor OTPs
  - They use a second factor authentication
  - They also contain a symmetric key and a nonce
  - An encryption algorithm is applied to obtain the OTP

Magnetic Stripe Cards

- Plastic card with a magnetic stripe containing personalized information about the card holder.
  - The first track of a magnetic stripe card contains
    - cardholder full name
    - an account number
    - format information
    - other data
  - The second track may contain
    - account number
    - expiration date
    - information about the issuing bank
    - data specifying the exact format of the track
    - other discretionary data.

Magnetic Stripe Card Security

- The magnetic stripe medium is easy to read and reproduce
- Magnetic stripe readers have low cost
  - attackers can read information off cards
- An attacker can easily clone existing cards with a magnetic stripe writer (only a little more expensive)
- Many scenarios require card holders to enter a PIN to use their cards (e.g., as in ATM and debit cards)

Authentication via Smart Cards

- Public and private key certificates of the user stored in the card's memory

- Challenge-response protocol with the reader
  - Users enters a PIN
  - Reader sends a **challenge** A
  - Smart card generates random value B,
  - The smart cards signs (A||B) with the **private key**
  - Reader verifies the signature with **public key.**

RFID tags

RFID Tag

Reader

Authentication Server

Types of Tags

- **Passive**
  - Operational power from reader radiated power

- **Active**
  - Operational power provided by battery

Authentication via RFID tags

Challenge: nonce x

Response: y=f(k, x)

y'=f(k, x)

- Function f is public
- Secret key K is known only to the tag and the reader
- The reader computes Y' = f(K,X) and verifies that Y=Y'

Electronic Passports

- All RFID communications are encrypted with a secret key, to protect sensitive information on the passport
- In many instances, however, this secret key is merely the passport number, the holder's date of birth, and the expiration date (in that order)
- All of this information is printed on the card, either in text or using a barcode or other optical storage method.
- The secret key is intended to be only accessible to those with physical access to the passport
  - An attacker with information on the owner (including when their passport was issued) may be able to easily reconstruct the key (passport numbers are typically issued sequentially)

Biometrics

- **Biometric** refers to any measure used to uniquely identify a person based on biological or physiological traits.

- Generally, biometric systems incorporate some sort of sensor or scanner to
  - read biometric information and
  - compare it to stored templates of accepted users

Examples of biometrics in real life

Requirements for Biometric Authentication

- **Universality:**
  - Almost every person should have this characteristic
- **Distinctiveness:**
  - Each person should have noticeable differences in the characteristic
- **Permanence:**
  - The characteristic should not change significantly over time
- **Collectability:**
  - The characteristic should have the ability to be effectively determined and quantified.

Biometric Identification

Biometric

Reader

Feature vector

=?

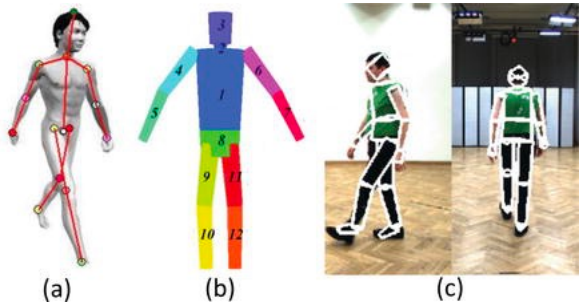
Reference vector

Match

No match

Candidates for biometric IDs

- Signature
- Fingerprints
- Retinal/iris scans
- DNA
- Voice recognition
- Face recognition
- Gait recognition



Other candidate for biometric IDs

Boffins take biometric logins to heart, literally: Cardiac radar IDs users to unlock their PCs

2026, when a change of heart will mean a pretty bad day

By Katyanna Quach 26 Sep 2017 at 05:01 20 SHARE ▼

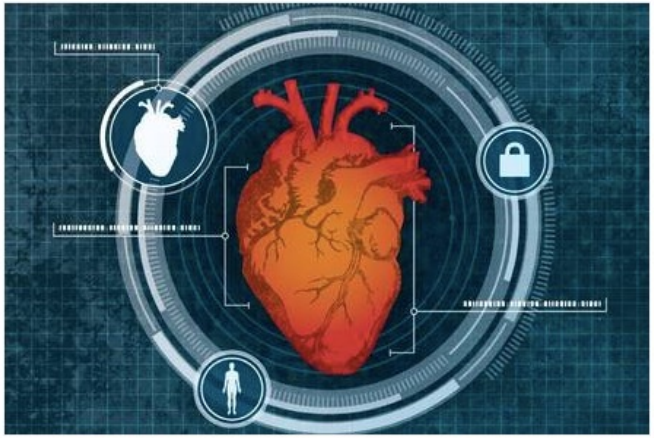


Image credit: Bob Wilder/University at Buffalo

Limitations of biometric authentication

- Accuracy of matching algorithm
  - False positives: access allowed to unauthorized user
  - False negatives: deny access to legitimate user
- Easy forging of biometric traits
  - Fingerprints left in many places
- Low user acceptance
  - User may not like to have their retina scanned.

Multifactor authentication

- It requires two or more forms of authentication
  - Password + token (OTP device or smartcard)
  - Password + SMS on your mobile phone
  - Smart Card + fingerprint scanning

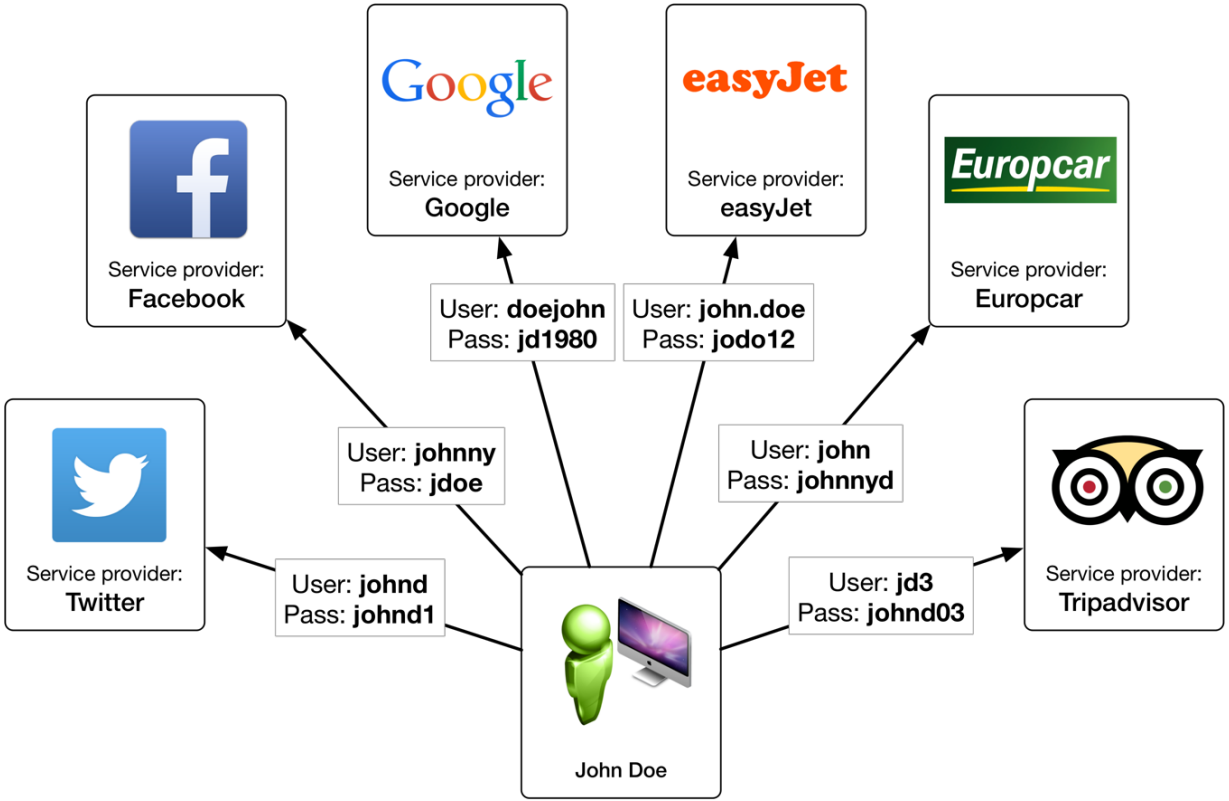
- Use for *remote access* prescribed by regulations
  - PCI DSS 3.0 for credit cards data
  - PDS2 for Internet banking transactions.

Summary

- Password based authentication is not secure
  - Can be guessed, Can be forgotten, Can be shared
- Alternative authentication approaches
  - Token-based
    - Can be shared, Can be lost or stolen, Can be forged
  - Biometric
    - Expensive, Not 100% accurate
- Possible solution: Multi-factor authentication

Single sign on

The problem


Google
Service provider:
Google
easyJet
Service provider:
easyJet
Europcar
Service provider:
Europcar
Service provider:
Facebook
User: doejohn
Pass: jd1980
User: john.doe
Pass: jodo12
User: johnny
Pass: jdoe
User: john
Pass: johnnyd
Service provider:
Twitter
User: johnd
Pass: johnd1
User: jd3
Pass: johnd03
Service provider:
Tripadvisor
John Doe

Digital identity management

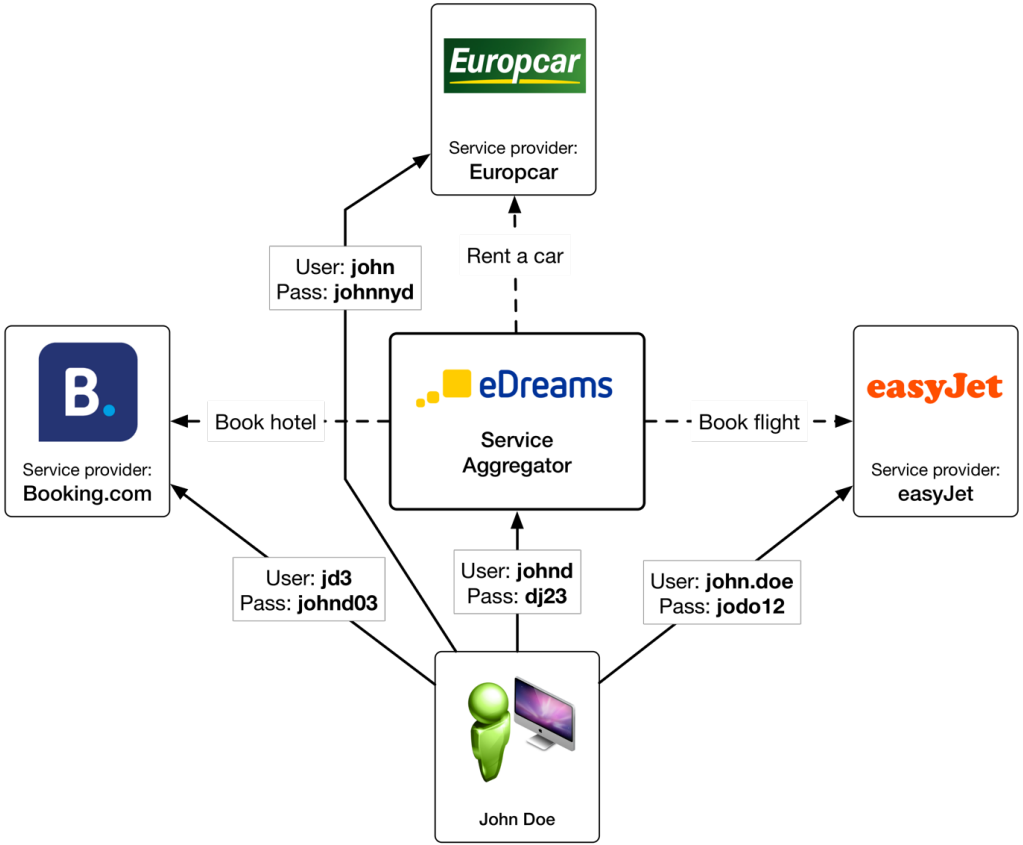
- A **digital identity** is the digital representation of the information know about a specific individual
  - Name and last name
  - National insurance number
  - Home address
  - Job title
  - User id and password

- A **digital identity management system** provides a centralized solution that manages users' digital identities and user access to resources/services
  - Maintain identity of the user and associates attributes to this identity
  - Verify identity of the user based on his/her identity attributes.

Main players

- **Subject**
  - System entity about which something can be asserted
- **Asserting Party or Identity Provider**
  - System entity that creates assertions about a subject
- **Relying Party or Service Provider**
  - System entity that consumes assertions about a subject

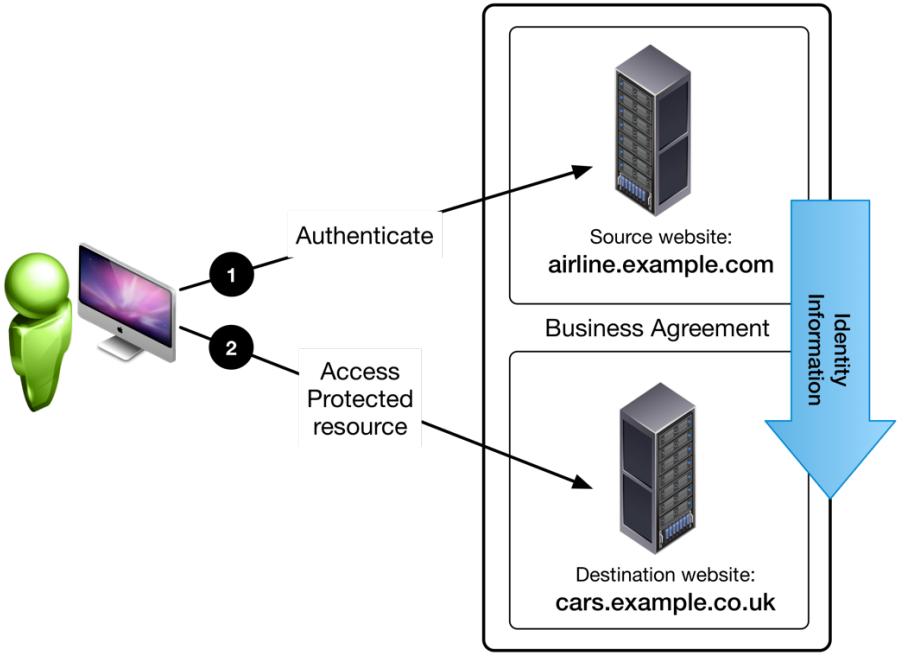
Single sign on



What is single sign on (SSO)

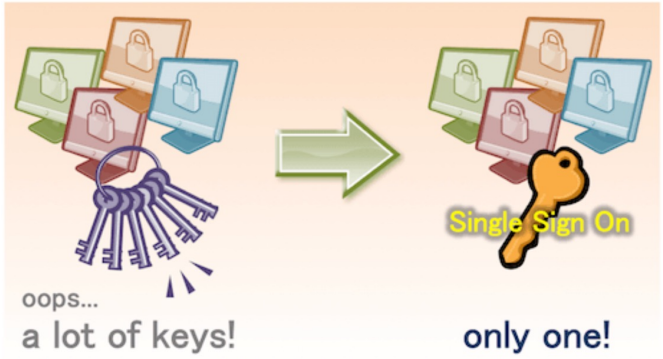
- User authenticates once and then access all the resources the user is authorized to use
- Authentication to the individual resources is handled by the SSO service in a manner that is transparent to the user
- SSO service maintains the identity information of the user
- When the user has to authenticate again for a resource, the SSO service does the job for the user

An Example of SSO

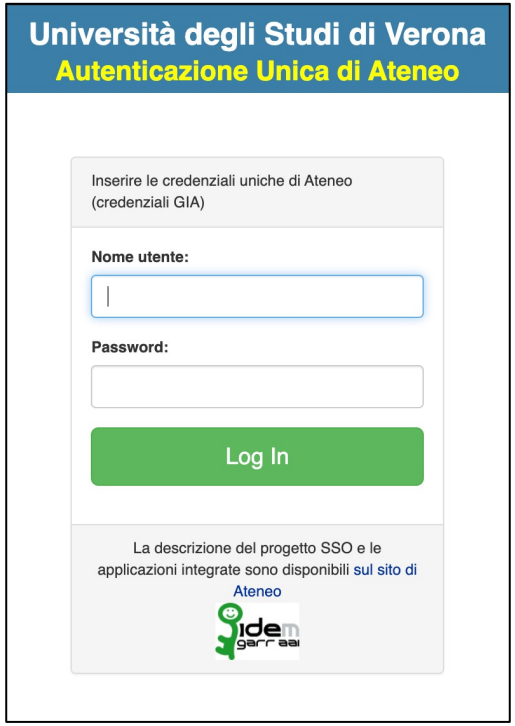


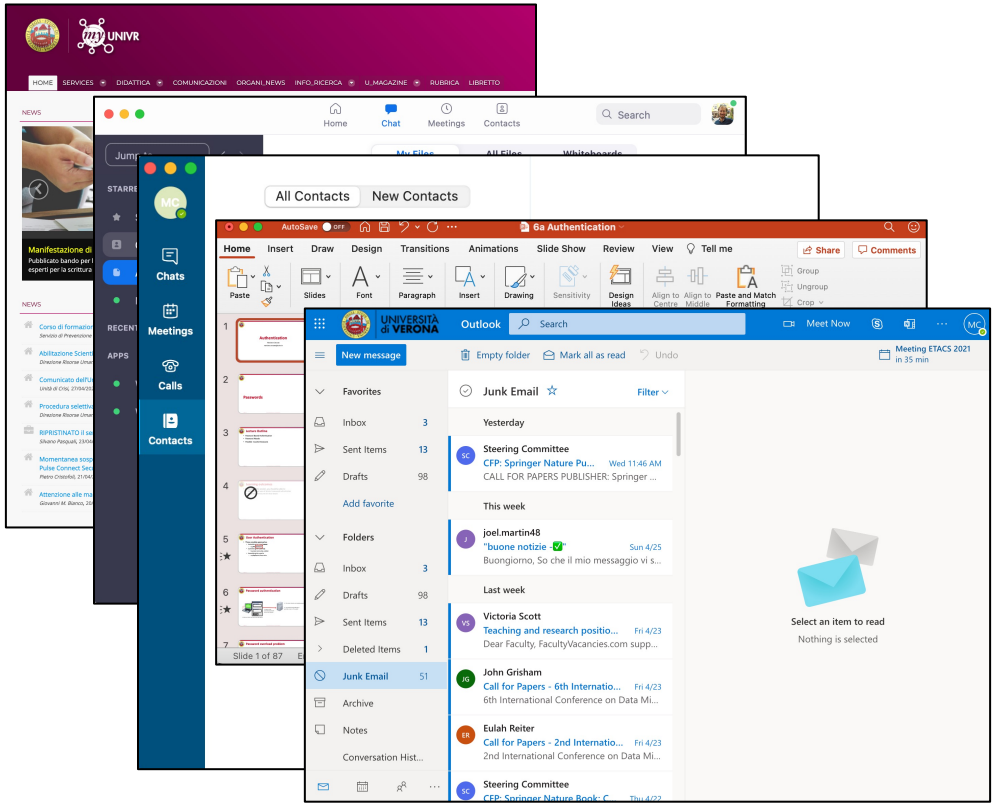
Advantages of SSO

- Reducing password fatigue from different username/password combinations
- Reducing time spent re-entering passwords for the same identity
- Reducing IT costs due to lower number of IT help desk calls about passwords



Another example of SSO

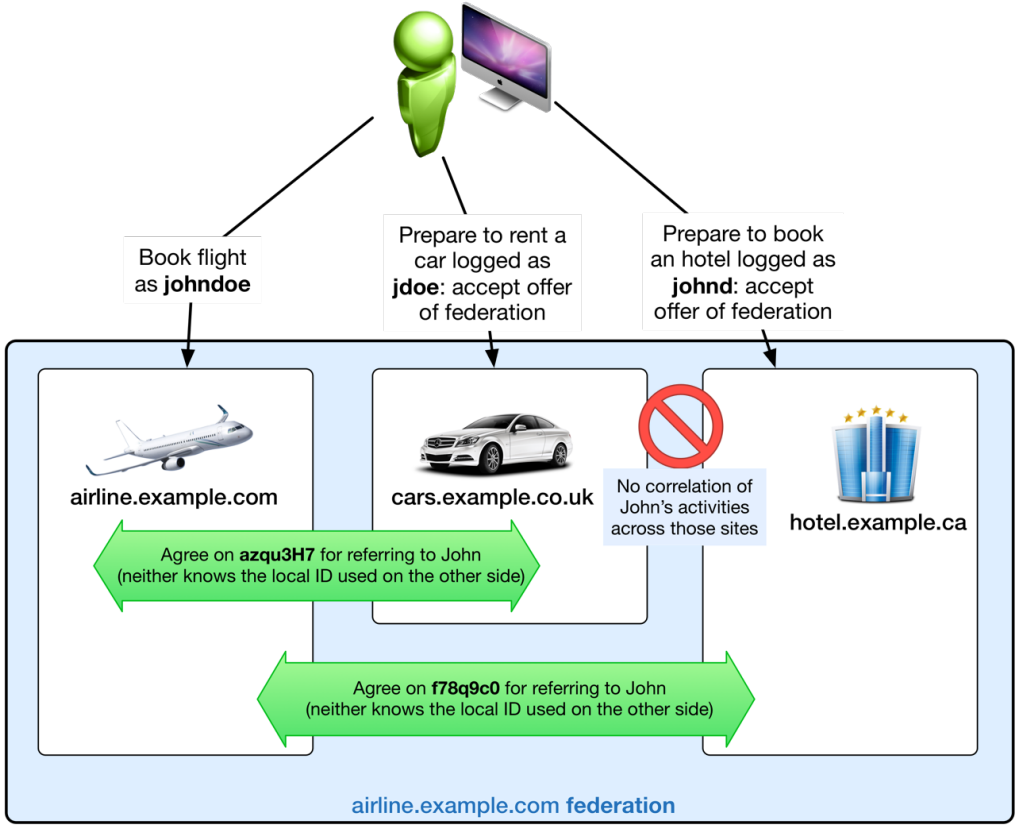




Federated Identity

- Organizations reach an agreement and establish a common, shared identifier to refer to a subject
- Facilitate sharing on identity information across different trust domains
- Facilitates SSO
  - An end-user that "logs into" any member of the federation has effectively logged into all of the members
- Reduces costs of maintaining and managing identities
  - No need independently and maintain collect identity information

An Example of Federated Identity



SAML

- XML-based framework for communicating
  - user authentication
  - attribute information

- SAML allows business entities to make assertions regarding the identity and attributes of a subject to other entities
  - E.g., identity provider to a service provider

- SAML is a flexible and extensible protocol designed to be used and customized if necessary by other standards

SAML Assertions

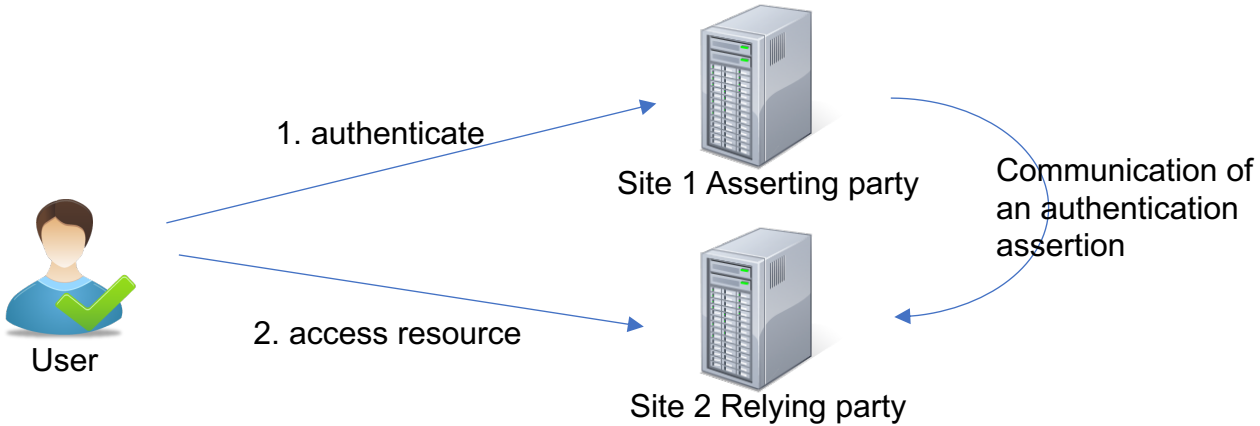
- Declaration of facts about a subject that an asserting party claims to be true

- Three types of assertions
  - **Authentication statements**
    - describe the means used to authenticate a subject
    - E.g., *Bill was authenticated at his company using a password*
  - **Attribute statements**
    - list attributes that a subject has
    - E.g., *Bill's email address is xxx@yyyyyyy.zzz*
  - **Authorization statements**
    - define the subject permissions
    - E.g., *Bill has permission to access resource X*

SAML does not support
- Performing authentication
- Granting Bill access to X

SAML

- SAML enables SSO and Identity Federation by providing a standard representation for attribute assertions and authentication assertion
- SAML Asserting Party/Identity Provider verifies identity of a user and issues an authentication assertion
- The user can present to a Service Provider the authentication assertion without authenticating again



If relying party is confident of the origin of the assertion (generated by the asserting party) can choose to log in users as if they had authenticated directly

SAML Assertions Common Elements

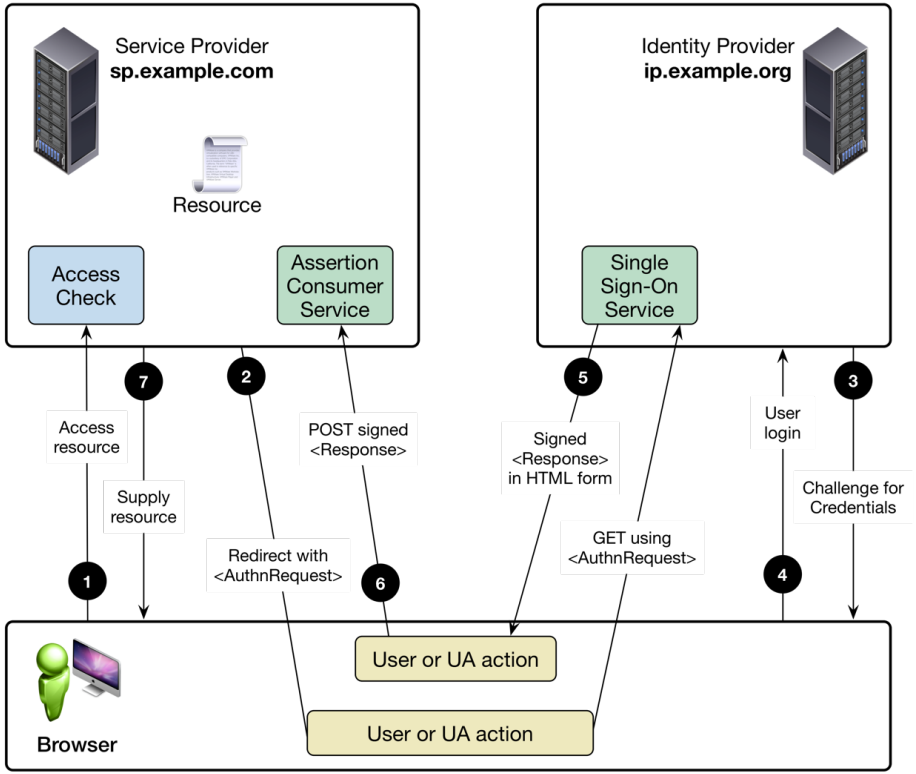
- **Assertion:** package of information supplying one or more statements made by a SAML authority

- Issuer and issuance timestamp
- Assertion ID
- Subject
  - Name and the security domain
- Conditions under which assertion is valid
  - SAML clients *must reject* assertions containing unsupported conditions
  - Special kind of condition: assertion validity period

SAML Concepts

- **Protocols**
  - Request/response protocols to obtain assertions
- **Bindings**
  - They detail how SAML request/response messages can be carried over underlying transport protocols
- **Profiles**
  - They define how the SAML assertions, protocols, and bindings are combined and constrained to support different use cases

SAML Use Cases – SP Initiated SSO



Advantages of SAML

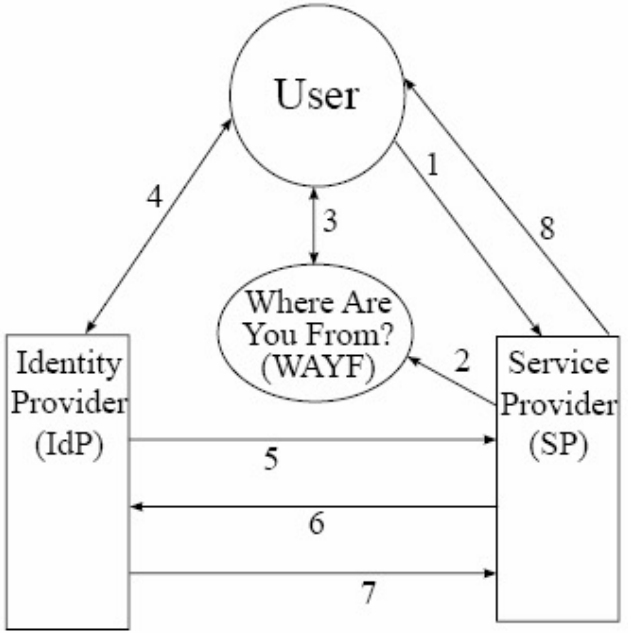
- **Platform neutrality**
  - SAML abstracts the security framework away from platform architectures and vendor implementations
  - Security is more independent of application logic
- **Loose coupling of directories**
  - SAML does not require user information to be maintained and synchronized between directories
- **Improved online experience for end users**
  - SAML enables Single Sign-On (SSO) by allowing users to authenticate at an identity provider and then access service providers without additional authentication
- **Reduced administrative costs and risk for service providers**
  - *Reuse* a single act of authentication (such as logging in with a username and password) multiple times across multiple services can reduce the cost of maintaining account information
    - Burden is transferred to the identity provider.

Shibboleth

- Consortium and opens source project that enables universities to control access to on-line resources and research across institutional boundaries
- Shibboleth enable students, faculty, and staff to access resources at partner institutions without needing to create separate, local IDs and passwords for each one
- Shibboleth system has profiled SAML for its particular requirements and has built privacy management into its architecture

Shibboleth Flow

1. The User attempts to access a Shibboleth-protected resource on the SP site
2. The User is redirected to the federation WAYF
3. The User select her home institution (IDP) from the list presented by the WAYF.
4. The Identity Provider authenticates the User
5. A one-time handle (session identifier) is generated for this User session and is sent to the SP
6. The SP uses the handle to request attribute information from the IDP for this User.
7. Based on the attribute information made available to it, the SP allows or refuses the User access to the resource.



OpenID Connect

- It provides a standard SSO protocol on top of OAuth 2.0 authorization framework
- Good for a range of domains
  - **consumer** applications
  - **social** applications
  - **enterprise** applications
- Wide-spread adoption is supported by major providers like Google, Microsoft and Facebook

OAuth2.0

- Authorization framework specified in RFC 6749, forms the basis of OpenID Connect
- An authorization server issues clients with **access tokens** to grant them access to protected HTTP resources
- OAuth 2.0 **access tokens** are employed in OpenID Connect to **allow** the **client application** to retrieve consented **user information** from the server

ID Token

- Asserts a user identity
- Specifies the identity provider
- May specify how and when the user was authenticated
- Has an issue and expiration date
- May contain additionally subject details e.g. name, email address
- Digitally signed and encrypted

ID Token Encoding

- Encoded as a **JSON Web Token** (JWT)
- The claims about the subject are packed in a simple JSON object
- Typically signed with the provider's private key or a shared secret issued to the client during registration

```
{
"iss" : "https://c2id.com"
"sub" : "alice",
"aud" : "s6BhdRkqt3",
"nonce" : "n-OS_WzA2Mj",
"exp" : "1311281970",
"iat" : "1311281970",
"auth_time" : "1311280970"
}
```

OpenID Providers

Google

Blogger

YAHOO!

flickr

myspace

WordPress.com

AOL

Summary

- SSO simplifies user authentication across different web sites
  - User login once at a web site and then can access resources to other web sites

- SSO implemented through two standardization efforts
  - SAML: XML-based framework
  - Shibboleth: SAML-based protocol for educational institutions
  - OpenID Connect: JSON-based framework

Recommended Readings

- Security Assertion Markup Language (SAML)
  - v2.0 Technical Overview. https://www.oasis-open.org/committees/download.php/27819/sstc-saml-tech-overview-2.0-cd-02.pdf
- Shibboleth
  - http://wiki.shibboleth.net
- OpenID Connect
  - http://openid.net/connect/