Chair for Network Architectures and Services – Prof. Carle
Department for Computer Science
TU München

Master Course Computer Networks (IN2097)

Introduction to Network Resilience

Dr. Ali Fessi
Dr. Nils Kammenhuber
Prof. Dr.-Ing. Georg Carle

Technische Universität München

Overview

I. Terminology
II. Challenges in the current Internet
III. Resilience Mechanisms

Overview

I. **Terminology**
II. Challenges in the current Internet
III. Resilience Mechanisms

Terminolgy - Overview

1. The “*fault* ➔ *error* ➔ *failure*” chain
2. Fault tolerance
3. Resilience
4. Dependability
5. Security
6. Availability vs. Reliability

The "*fault* → *error* → *failure*" chain

- *Service*:
  - Sequence of the system's external state
- *Correct service* is delivered when the service implements the system function
- Definition
  - A *service failure*, or simply *failure*, is an event that occurs when the delivered service deviates from *correct service*
  - i.e., at least one external state of the system deviates from the correct service state
  - (de: Ausfall)

The "*fault* → *error* → *failure*" chain

- Definition
  - The deviation of an external state of the system from the correct service state is called an *error*
  - Thus, an error is the part of the total state of the system that may lead to its subsequent failure
  - (de: Defekt)
- Definition
  - The cause of an error (adjuged or hypothesized) is called a *fault*
  - (de: Fehler)

☞ "*fault* → *error* → *failure*"

Fault Tolerance

- Definition
  - A system is fault-tolerant if it can mask the presence of *faults* in the system by using *redundancy*
- Redundancy means
  1. *Replication* of the same object (software or hardware) or
  2. *Diversity*
     - Design or implementation
     - Hardware or software

Resilience

- Origin
  - Latin verb: "resilire" ~ jump back
- Resilience definition in different fields
  - Physics
    - A material's property of being able to recover to a normal state after a deformation resulting from external forces;
  - Ecology
    - Moving from a stability domain to another under the influence of disturbance;
  - Psychology and psychiatry
    - Living and developing successfully when facing adversity;
  - Business
    - the capacity to reinvent a business model before circumstances force to;

Resilience

- Definition:
  - "Resilience is the persistence of *dependability* when facing *changes*."

    J.-C. Laprie. "From Dependability to Resilience". In 38th International Conference On Dependable Systems and Networks. IEEE/IFIP, 2008.
- Changes can be particularly *attacks*

Dependability Attributes

- Availability
  - Readiness for correct service
- Reliability
  - Continuity of correct service
- Safety
  - Absence of catastrophic consequences on the user(s) and the environment
- Integrity
  - Absence of improper system alterations
- Maintainability
  - Ability to undergo repair and modification

Security Attributes

- "CIA" model
  - Confidentiality, Integrity, Availability
- Confidentiality
  - Absence of unauthorized disclosure of information
- Availability
  - Readiness for correct service
- Integrity
  - Absence of improper system alterations
- Notes:
  - CIA model actually not sufficient to describe "security"
  - "Security" addresses all kind of possible attacks which may lead to the deviation from correct service

Reliability vs. Availability

- The reliability of a unit at a point of time t is the probability that the unit is operational until t

  $R(t) = Pr$ [unit is operating until t]

- The availability of a unit at a point of time t is the probability that the unit is operational at t

  $A(t) = Pr$ [unit is operating at t]

MTTF & MTTR

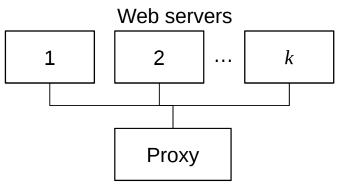
- Mean Time To Failure (MTTF)
  - Mean time between
    - Point of time when a unit is put into operation
    - Point of time when the unit fails for the next time
- Mean Time To Repair (MTTR)
  - Mean time between
    - Point of time when a unit fails
    - Point of time when the unit is put into operation again
- This results into an average availability

$$A_{avg} = \frac{MTTF}{MTTF + MTTR}$$

Examples

- DNS lookup (stateless service)
  - MTTF: 30 min
  - MTTR: 1 ms
  - A_{avg} = 0.998
- ☞ One can achieve
  - high availability
  - with low reliability (low MTTF)
  - if MTTR is sufficiently low
- Conference bridge (statefull service)
  - Each time, the bridge fails, participants need to re-dial
  - Even if MTTR is sufficiently low, it has to be guaranteed that the MTTF is sufficiently high to assure service quality

Examples



$$R_{system}(t) = R_{proxy}(t) \cdot R_{webserver\ pool}(t)$$

$$R_{webserver\ pool}(t) = 1 - (1 - R_{webserver}(t))^k$$

- Same holds for the availability

$$A_{system}(t) = A_{proxy}(t) \cdot A_{webserver\ pool}(t)$$

$$A_{webserver\ pool}(t) = 1 - (1 - A_{webserver}(t))^k$$

Overview

I. Terminology

II. Challenges in the current Internet

III. Resilience Mechanisms

Challenges in the current Internet

1. Topology Failures
2. Overload
3. Lack of Integrity
4. Software Faults
5. Domino Effects

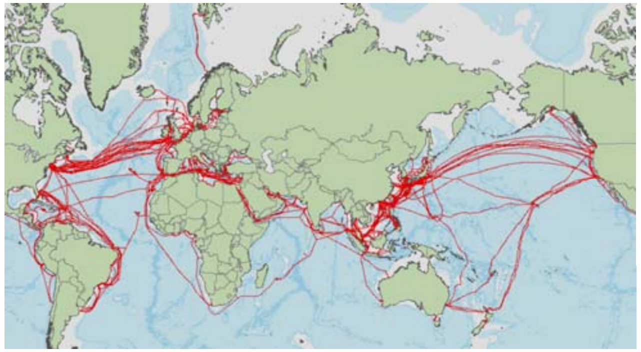
Challenges in the current Internet

1. **Topology Failures**
2. Overload
3. Lack of Integrity
4. Software Faults
5. Domino Effects

Topology Failures

- Failures in the “network graph”
- Network graph
  - Physical topology
  - Logical topology including service dependencies, e.g., DNS
    - ☞ *Dependency graphs*

Topology Failures; Sub-Marine Cables



- ~99% of inter-continental Internet traffic (less than 1% using satellites)
- High redundant
- But vulnerable to
  - Fishing and anchoring (70% of sub-marine cable failures)
  - Natural disasters (12%)
  - Cable theft

Submarine Cables; Natural Disasters

- Hengchun earthquake (December 2006)

Bloomberg.com

BloombergAnywhere

Updated: New York, Oct 27 1

QUOTE SEARCH NEWS SYMBOL LOOKUP Live TV

HOME NEWS MARKET DATA PERSONAL FINANCE TV and RADIO

news

Technology Currencies Forex Trading Videos ETFs CEO Commodi

Exclusive
Worldwide
Regions
Markets
Industries
Economy
Politics
Law
Environment
Science
Opinion
Spend

Asian Internet, Phone Services Hit by Taiwan Quakes (Update2)

Share | Email | Print | A A A

By Tim Culpan and Andrea Tan

Dec. 27 (Bloomberg) -- Internet and telephone services across Asia were disrupted, hampering financial transactions, after earthquakes near Taiwan damaged undersea cables.

``The repairs could take two to three weeks,'' said Leng Tai-feng, president of Chunghwa Telecom Co.'s international business. The Taipei-based company, Taiwan's largest phone operator, said two of its undersea cables were cut.

Submarine Cables; Natural Disasters

- Hengchun earthquake (December 2006)
- Impact
  - Affected countries: China, Taiwan, Hong Kong, Philippines
  - China's Internet connectivity reduced by 70%
  - Hong Kong's Internet access completely disabled
- Recovery
  - BGP automatic re-routing helped to reduce disconnectivity
  - But resulted into congested links
  - Manual BGP policy changes + switch port re-configuration were necessary
  - Hong Kong's Internet users were still experiencing slow Internet connections 5 days after the earthquake

Submarine Cables; Failures in the Mediterranean Sea

- In Jan. + Feb. 2008, 3 successive events
- Impact
  - Affected countries: Egypt, Iran, India and a number of other middle east countries
  - Disruption of
    - 70% in Egypt
    - 60% in India

Submarine Cables; Cable Theft

- In March 2007, pirates stole an 11 kilometers section of the submarine cable connecting Thailand, Vietnam and Hong Kong,
- Impact: significant downgrade in Internet speed in Vietnam.
- Intention: The thieves wanted to sell 100 tons of cable as scrap.

Topology Failures; Routing

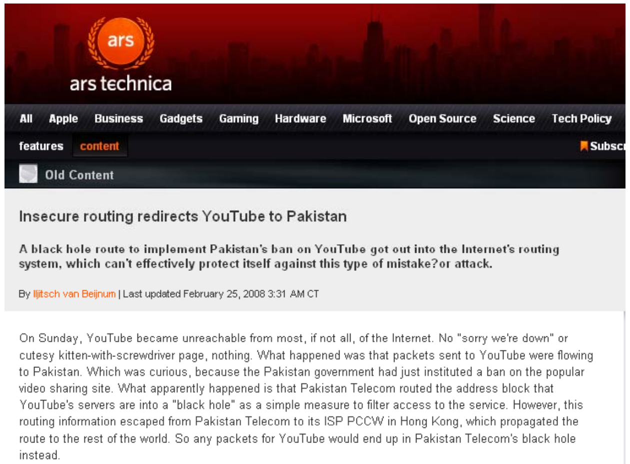
- Failures in the IP topology graph
  - Failures of routers (nodes)
  - Failure of links between routers
- Failure of links between routers generally caused by disconnection at lower layers
- Failure of routers
  - DoS attacks
  - Failures due to software bugs
  - Examples of reported bugs
    - Vulnerability to too long AS (BGP Autonomous Systems) paths
    - Long passwords to login to the router
    - Overflow of connection tables in some commercial firewalls

Topological Failures; Routing

- Time to Recovery
  - Intra-domain routing (OSPF, RIP, IS-IS, EIGRP): up to several 100ms
  - Inter-domain routing (BGP): up to several minutes

Topological Failures; Routing

- Other reasons
  - Misconfiguration which leads to false modification of the Internet topology



Insecure routing redirects YouTube to Pakistan

A black hole route to implement Pakistan's ban on YouTube got out into the Internet's routing system, which can't effectively protect itself against this type of mistake?or attack.

By Iljitsch van Beijnum | Last updated February 25, 2008 3:31 AM CT

On Sunday, YouTube became unreachable from most, if not all, of the Internet. No "sorry we're down" or cutesy kitten-with-screwdriver page, nothing. What happened was that packets sent to YouTube were flowing to Pakistan. Which was curious, because the Pakistan government had just instituted a ban on the popular video sharing site. What apparently happened is that Pakistan Telecom routed the address block that YouTube's servers are into a "black hole" as a simple measure to filter access to the service. However, this routing information escaped from Pakistan Telecom to its ISP PCCW in Hong Kong, which propagated the route to the rest of the world. So any packets for YouTube would end up in Pakistan Telecom's black hole instead.

Challenges in the current Internet

1. Topology Failures
2. **Overload**
3. Lack of Integrity
4. Software Faults
5. Domino Effects

Overload

- ❑ Topology failures are binary (link or node is up or down)
- ❑ But equipment in the network (routers, servers, etc.) have limited capacity
  - Queue length
  - CPU power
  - etc.

☞ Overload (congestion) is not rare

Lack of Congestion at the Network Layer

- ❑ Routing protocols react to the failure of a link or a router.
- ❑ But not to network congestions
- ❑ ARPANET had some mechanisms to react to congestions
- ❑ But they resulted into oscillations
- ❑ Congestion control was introduced in the Internet as enhancement of TCP
- ❑ But TCP has
  - no knowledge about the network topology
  - no way of re-wiring the traffic path in case of congestion

DoS Attack vs. Flash Crowds

- ❑ Big challenge
  - Ambiguous differentiation between DoS attacks and *__flash crowds__*
  - *__Flash crowds__*: unusual but legitimate traffic
  - Even if attacks are identified as such, it remains difficult to separate between malicious and legitimate traffic and to eliminate the malicious traffic

DoS Attacks

- ❑ Some DoS attacks have a political or ethnical reasons


Low graphics | Accessibility help
BBC NEWS
Watch One-Minute World News
News services Your news when you want it
News Front Page
Africa
Americas
Asia-Pacific
Europe
Middle East
South Asia
UK
Business
Health
Science & Environment
Technology
Entertainment

Last Updated: Thursday, 17 May 2007, 14:52 GMT 15:52 UK

E-mail this to a friend | Printable version

The cyber raiders hitting Estonia

As Estonia appeals to its Nato and EU partners for help against cyber-attacks it links to Russia, the BBC News website's Patrick Jackson investigates who may be responsible.

Estonia, one of the most internet-savvy states in the European Union, has been under sustained attack from hackers since the ethnic Russian riots sparked in late April by its removal of a Soviet war memorial from Tallinn city centre.

днем победы!

SEE ALSO
- Estonia hit by 'Moscow cyber war' 17 May 07 | Europe
- Hackers attack heart of the net 07 Feb 07 | Technology
- Playing Estonia's political cards 12 May 07 | Europe
- News fuels Russian internet boom 10 Apr 06 | Europe
- Country profile: Estonia 30 Apr 07 | Country profiles

RELATED INTERNET LINKS
- Estonian foreign ministry
- Russian government
- Kaspersky Lab

Challenges in the current Internet

1. Topology Failures
2. Overload
3. **Lack of Integrity**
4. Software Faults
5. Domino Effects

Lack of Integrity

- Majority of Internet traffic (signaling and data) is not integrity-protected
- This leads to several security vulnerabilities
  - ARP poisoning
  - Forged BGP announcements
  - Forged DNS responses
  - SPAM SPAM SPAM SPAM SPAM SPAM SPAM SPAM SPAM SPAM
  - etc.

Challenges in the current Internet

1. Topology Failures
2. Overload
3. Lack of Integrity
4. **Software Faults**
5. Domino Effects

Software Faults

- Developments faults
  - Introduced during the development phase
- Configuration faults
  - Introduced during the deployment phase

Software Faults

- Examples
  - Buffer overflows in server or router implementation
  - BGP Youtube misconfiguration
  - On Jan. 31st 2009, Google search engine marked every search result with "This site may harm your computer";
    Root cause: Database of suspected sites was mistakenly extended by ‚/'
  - Software update of the Authentication Server (Home Location Register HLR) of T-Mobile on April 21st 2009
    - Impact: phone calls and text messaging were not possible for 4 hours

Challenges in the current Internet

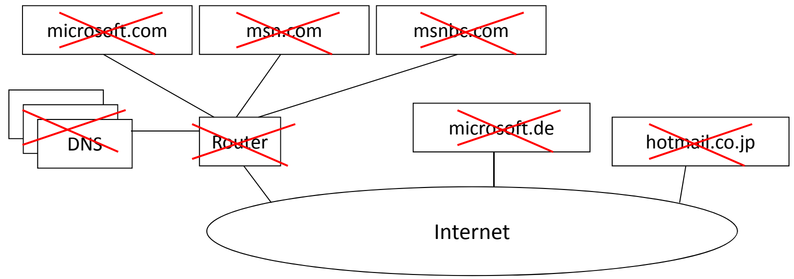
1. Topology Failures
2. Overload
3. Lack of Integrity
4. Software Faults
5. **Domino Effects**

Domino Effects

- Any kind of challenges mentioned above may lead to other challenges
  - E.g., failure of a server in a server pool may lead to overload of neighboring servers
  - Router failures may lead to congestion of neighboring links and routers
  - DNS failure may lead to unavailability of other services,

Domino Effects

- E.g., DoS attack on Microsoft router on 24th + 25th Jan. 2001 lead to unavailability of DNS and thus of services located in other MS sites



Overview

Resilience Mechanisms

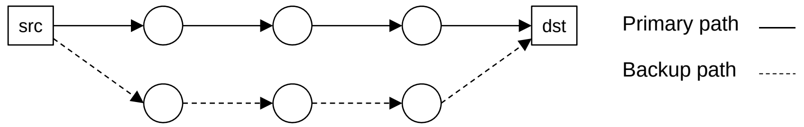
Resilience Mechanisms

Topology-based Resilience Metrics

- Several metrics exist
- But not all are useful

- Definitions
  - *k*-link (edge) connectivity is the minimal number of links whose removal would disconnect the graph
  - *k*-node (vertex) connectivity is the minimal number of nodes whose removal (including removal of adjacent links) would disconnect the graph
  - A *k-regular graph* is k-node-connected if there are k node-disjoint paths between any pair of nodes.

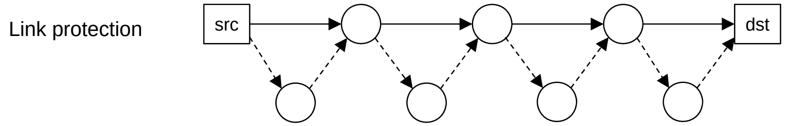
Path Protection

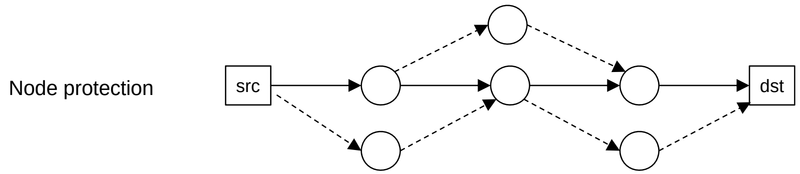


- Traffic is forwarded using backup path in case of failure
- Source needs to monitor the operation of primary path
  - ☞ Info about node or link failure needs to be propagated back to src

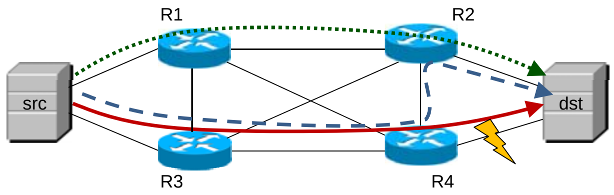
Local Protection

- Node or link failures are detected locally and backup paths are used until routing re-converges
  - ☞ This can reduces the MTTR by the order of a magnitude compared to *path protection*
  - ☞ Contra: higher signaling and equipment overhead





Example

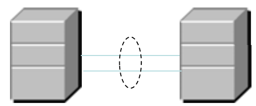


- Location protection at IP layer
- Routing protocol: OSPF
- Local protection according to IP Fast Reroute (IPFRR) (RFC 5714)

1. Normal operation: Routing from src to dst via R3 and R4
2. After failure of link between R4 and dst: Rerouting from R4 to dst via R2
3. Then, info is propagated in the network, OSPF routing converges and a new path is used from src to dst via R1 and R2.

IEEE 802.3ad: Link Aggregation

- IEEE Link Aggregation allows for bundling
  - several physical Ethernet connections
  - into a logical one
- Connection between
  - Two hosts
  - Two Ethernet switches
  - Host and switch



- IEEE Link Aggregation allows for increasing bandwidth
- But is also a fault tolerance mechanism
  - If a cable is plugged out,
    - e.g., for maintenance reasons,
  - the two layer-2 devices remain connected.

Multihoming

- *Multihoming* refers to a network setup where a host or a network is connected to the Internet via more than 1 connection
- It can be applied in various contexts
  - Host Multihoming
    - An IP host connected via multiple network interfaces
    - Each network interface might be connected to a different access network
  - Multihoming at the transition point between networks
    - An enterprise network connected to the Internet via multiple ISPs
    - BGP peering with multiple providers

Resilience Mechanisms

1. Topology Protection
2. **Congestion Control**
3. Signaling Integrity
4. Server Redundancy
5. Virtualization
6. Overlay and P2P Networks

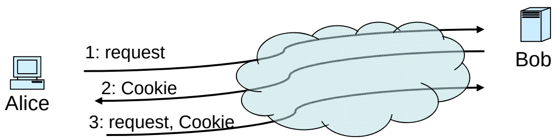
Congestion Control

- TCP congestion control
- Traffic Engineering
- Protection again DoS attacks
  - Rate limiting: vulnerable to
    - “false positives”, i.e., legitimate traffic is classified as malicious
    - “false negatives”, i.e., malicious traffic is classified as legitimate
  - Cookies

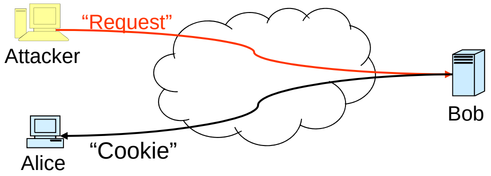
Traffic Engineering

- Addresses network congestion at the network layer
- Goals
  - Optimize network throughput, packet loss, delay
- Input
  - Network topology
  - Traffic matrix (may change over time, e.g., daily patterns)
- Output
  - (Eventually modified) link weights used to compute routing tables

Denial-of-Service Protection with Cookies (1)



- Upon receiving a request from Alice, Bob calculates a Cookie and sends it to Bob.
- Alice will receive the Cookie and resend the request with the Cookie together.
- Bob verifies that the Cookie is correct and then starts to process Alice's request.
- An attacker that is sending requests with a spoofed (i.e. forged) source address will not be able to send the Cookie.



Denial-of-Service Protection with Cookies (2)

- Cookies discussion:
  - Advantage: allows to counter simple address spoofing attacks
  - Drawbacks
    - Requires CPU resources
    - In some applications, e.g., DNS, it might be easier to respond to the request than generating the cookie
    - Requires one additional message roundtrip.
    - Network may remain congested

Resilience Mechanisms

1. Topology Protection
2. Congestion Control
3. **Signaling Integrity**
4. Server Redundancy
5. Virtualization
6. Overlay and P2P Networks

Signaling Integrity; "ARP" protection

- Manual configuration, e.g., ARP messages with wrong matching (IP to MAC) are discarded
  - ☞ Too costly
- IPv6 SEcure Neighbor Discovery (SEND) (RFC 2461 and 2462)
  - Uses a Cryptographically Generated Address (CGA)

Routing prefix	Hash62(Host public key)

Signaling Integrity; DNSSEC

- ❑ Protects DNS responses with cryptographic signatures
- ❑ In a dedicated DNS record: the RRSIG record (RFC4034)
- ❑ DNS Records can be verified with a “chain of trust”
  - Public key of the DNS root zone must be known by clients
- ❑ Authority delegation is restricted to sub-domains
  - e.g., system administrator of “net.in.tum.de” can not sign records for “lrz.de”
  - Note: this is not the case for PKIs currently used in the web

Signaling Integrity; BGP Security

- ❑ Not trivial
- ❑ Can not be solved by simply adding message integration protection of BGP announcements
  - E.g., what is if “Pakistan Telecom” signs BGP announcements for a Youtube prefix?
- ☞ Integrity of BGP announcements needs to be validated by a combination of
  - ☞ topology authentication,
  - ☞ BGP path authentication and
  - ☞ announcement's origin authentication

Signaling Integrity

- ❑ Domain Keys Identified Mail (DKIM)
  - Allows for validation of a domain name associated with an email address
  - An organization takes responsibility for a message in a way that can be validated by a recipient
  - Prominent email service providers implementing DKIM
    - Yahoo, Gmail, and FastMail.
    - Any mail from these organizations should carry a DKIM signature

Signaling Integrity

- ❑ Spammers can still sign their outgoing messages
  - ☞ DKIM should be used with reputation:
    - Email messages sent by a domain that is known for signing good messages can be accepted
    - while others may require further examination.

Resilience Mechanisms

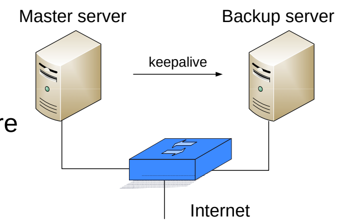
1. Topology Protection
2. Congestion Control
3. Signaling Integrity
4. **Server Redundancy**
5. Virtualization
6. Overlay and P2P Networks

Server Redundancy

- Server redundancy as a *fault tolerance* mechanism
- Servers instances may be
  - in the same LAN or
  - different sub-networks ☞ *Geographic diversity*
- Supporting mechanisms
  - IP Takeover
  - NAT Takeover
  - DNS

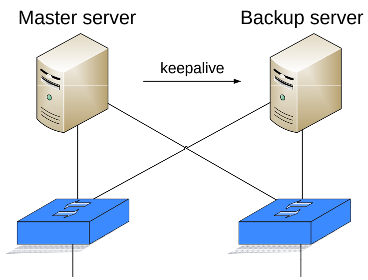
Server Redundancy; IP Takeover

- Simple redundancy mechanism
- Backup server receives periodic "keep alive" messages from master server, e.g., every 10ms
- In case of no response
  - Backup server broadcasts an ARP message in the LAN
  - From now on, all IP traffic is forwarded to the backup server
- Drawbacks
  - Existing session state gets lost
  - Ethernet switch is a single point of failure



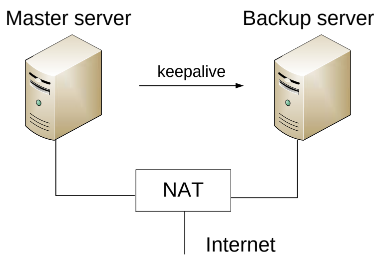
Server Redundancy; IP Takeover with 2 Switches

- Both master and backup servers are connected to 2 switches
- Same procedure with ARP
  ☞ Incoming requests from both switches is forwarded to the backup server
- Any component (server or switch or cable) can be removed, e.g., for maintenance reasons, while the service keeps on being available



Server Redundancy; NAT Takeover

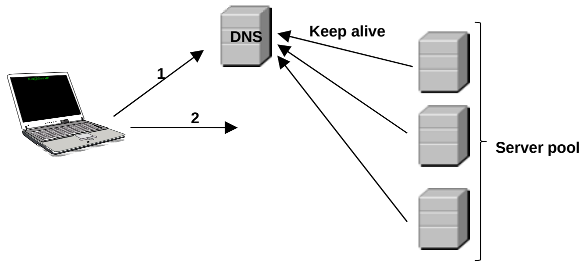
- ❑ Similar to IP Takeover
- ❑ "Keep alive" messages from backup to master server
- ❑ Change NAT binding upon lack of response from master server
  - ☞ Incoming requests are forwarded to the backup server



- ❑ Note: Master and backup server do not have to be in the same LAN

Server Redundancy; DNS

- ❑ DNS can provide several IP addresses for the same name
- ❑ By monitoring the availability of servers from a server pool, unavailable servers can be removed from DNS responses



- ❑ Moreover, DNS responses can be adjusted according to the current load
  - ☞ See, e.g., Content Distribution Networks (CDN)

Resilience Mechanisms

1. Topology Protection
2. Congestion Control
3. Signaling Integrity
4. Server Redundancy
5. **Virtualization**
6. Overlay and P2P Networks

Virtualization

- ❑ Different virtualization techniques, e.g., KVM, Xen, etc.
- ❑ Can be used to enhance resilience of network services
  - Start new servers from existing images *on demand*, e.g.,
    - To address overload situations
    - In case servers in other locations crash

Resilience Mechanisms

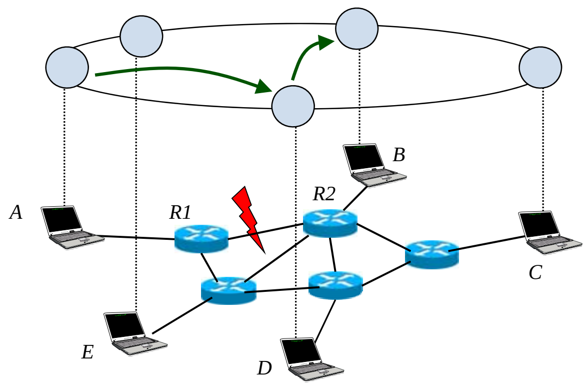
1. Topology Protection
2. Congestion Control
3. Signaling Integrity
4. Server Redundancy
5. Virtualization
6. **Overlay and P2P Networks**

Overlay Routing

- Overlay networks
  - Are networks built on top of existing networks
  - They typically provide additional functionality not provided at the „underlay" network
- Overlay routing
  - End hosts can organize themselves in a P2P network
  - and provide routing using the overlay in case the underlay routing fails

Overlay Routing

- Example
  - Upon link failure between R1 and R2
  - *A* can reach *B* via *D* or *C*



Overlay Routing

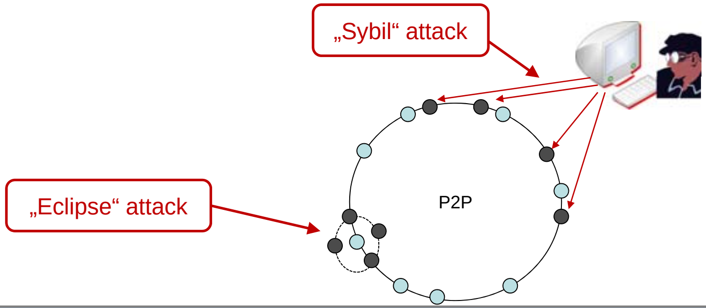
- Typical reasons for lack of connectivity in the underlay
  - Misconfigured middleboxes (firewalls, NATs)
  - Slow BGP convergence
- Systems supporting overlay routing
  - Tor
    - while it is actually designed with anonymization in mind, it provides overlay routing and can be useful in case of network partial failures
  - Skype
    - Skype supernodes typically provide connectivity for Skype clients behind firewalls or NATs

P2P Networks

- ❑ Resilience properties
  - Decentralization
  - Geographic diversity
  - Ability to cope with “churn”
    - “Churn” means that peers join and leave at any time
    - ☞ Replication of each data item on several peers
    - ☞ Autonomic recovery from stale P2P routing tables

P2P Networks

- ❑ Drawback: several attacks are possible
  - Sybil attacks:
    - Attacker participate with several fake identities
    - In order to control a portion of the network
  - Eclipse attacks,
    - Attacker control the neighborhood of a peer or content
    - In order to make unavailable for other participants in the P2P networks
  - etc.



P2P Networks

- ❑ Common approaches
  - ☞ Managed P2P networks (or supervised P2P networks)
  - ☞ E.g., Google File System (GFS), Skype

Summary

I. **Terminology**
   - ❑ The “*fault* ➔ *error* ➔ *failure*” chain
   - ❑ Fault tolerance, Resilience, Dependability, Security
   - ❑ Availability vs. Reliability

II. **Challenges in the current Internet**
   - ❑ Topological Failures, Overload, Lack of Integrity
   - ❑ Software Faults, Domino Effects

III. **Resilience Mechanisms**
   - ❑ Topology Protection, Congestion Control, Signaling Integrity
   - ❑ Server Redundancy**,** Virtualization, Overlay and P2P Networks