Lezione n.2
SISTEMI P2P DI PRIMA GENERAZIONE: NAPSTER
Materiale didattico:lucidi

Laura Ricci

P2P:CARATTERISTICHE GENERALI

- Definizione di un Overlay: rete virtuale stabilita tra i peer mediante connessioni TCP per lo scambio di informazione di controllo (signalling network)

- Caratteristiche della overlay network
  - completamente indipendente dalla rete fisica, grazie al livello di astrazione TCP/IP
  - può essere strutturata in modo gerarchico (Gnutella Superpeers, JXTA rendez-vouz peers)
  - può includere un server centralizzato (look up server di Napster)

- I peer partecipano attivamente alla overlay network e possono
  - Fornire informazione
  - Ricercare informazione
  - Svolgere funzioni di routing

P2P: CARATTERISTICHE GENERALI

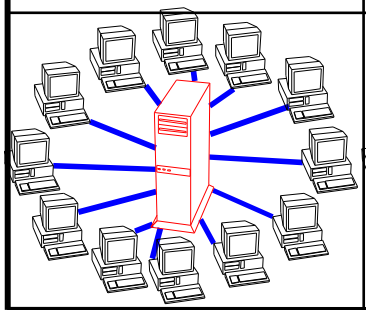
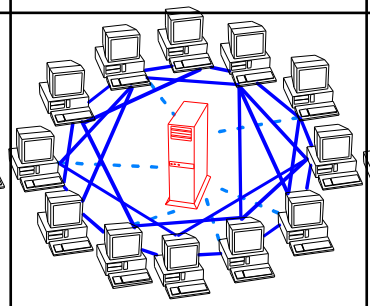
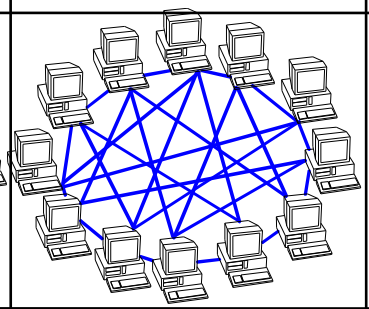
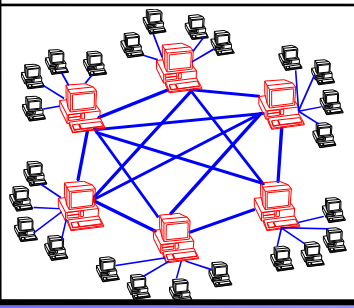
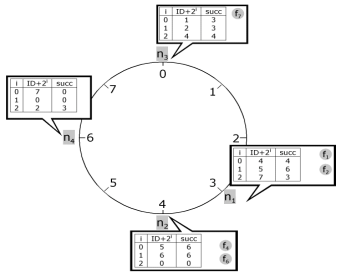
- L'informazione condivisa:
  - può essere memorizzata
    - sui nodi che la mettono a disposizione della rete
    - su altri nodi della rete

- può essere replicata su più nodi

- In generale la overlay network, (o signaling network), viene utilizzata per scambiare messaggi di controllo, in generale non per lo scambio dei dati
  - verifica periodica esistenza di altri peers (keep-alive)
  - routing di messaggi per ricerca di dati/ inserimento cancellazione di peer

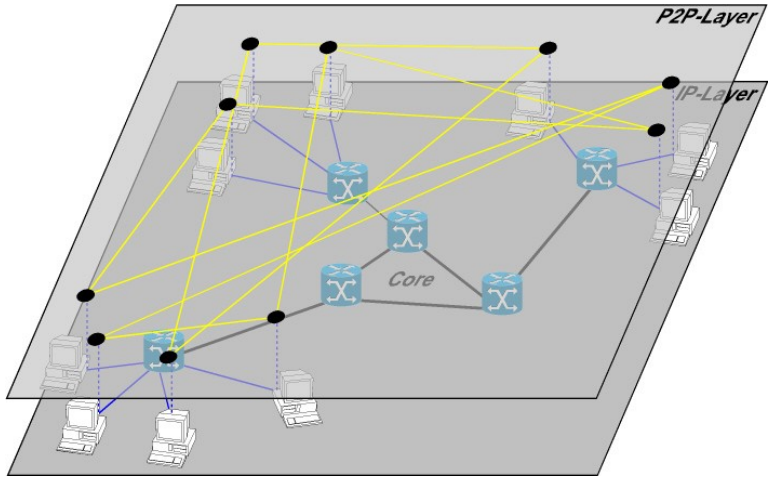
- Il trasferimento dei dati:
  - avviene su connessioni distinte da quelle definite dalla overlay network
  - in genere mediante il protocollo HTTP

P2P:CLASSIFICAZIONE GENERALE

Client-Server	***Peer-to-Peer: Caratteristiche Generali***			
• Il Server è l'unica entità in grado di memorizzare l'informazione condivisa → La rete è gestita dal Server in modo centralizzato 2. Il Server è il Sistema con la maggior capacità di calcolo 3. I client posseggono minor capacità di calcolo **Esempio:** WWW	1. Le risorse sono condivise tra i peers. 2. Le risorse offerte da un peer possono essere accedute direttamente da altri peer 3. Il peer fornisce e richiede risorse (Servent)			
	P2P Non strutturati			***P2P Strutturati***
	P2P Centralizzato	*P2P Puro*	*P2P Ibridp*	*P2P basato su DHT*
	• Soddisfano 1.,2.,3. • Esiste una entità centralizzata • Tale entità centralizzata svolge solo compiti di indicizzazione delle risorse **Esempio:** Napster	1. Soddisfano 1.,2.,3. 2. Non esiste alcuna entità centralizzata 3. La funzionalità del sistema non viene compromessa dall' eliminazione di un peer **Esempi:** Gnutella 0.4,	1. Soddisfano 1,.2.,3. • Esistono alcuni peer (superpeer) chehanno anche funzioni di indicizzazione 3. I superpeer sono determinati dinamicamente **Esempi:** Gnutella 0.6, JXTA, Kazaa	1. Soddisfano 1.,2.,3. 2. Non esistono entità centralizzate 3. La overlay network risulta strutturata 4. Overlay network strutturata **Esemnpi:** Chord, CAN,Pastry,Kademlia
		1st Gen.	**2nd Gen.**	
				

RIASSUNTO DELLA PRESENTAZIONE

1. Reti P2P centralizzate
   - Caratteristiche Base
   - Protocollo
   - Dicussione
2. Reti Peer to Peer Pure
   - Caratteristiche Base
   - Protocollo
   - Discussione
3. Reti Peer to Peer Ibride
   1. Caratteristiche Base
   2. Protocollo
   3. Discussione

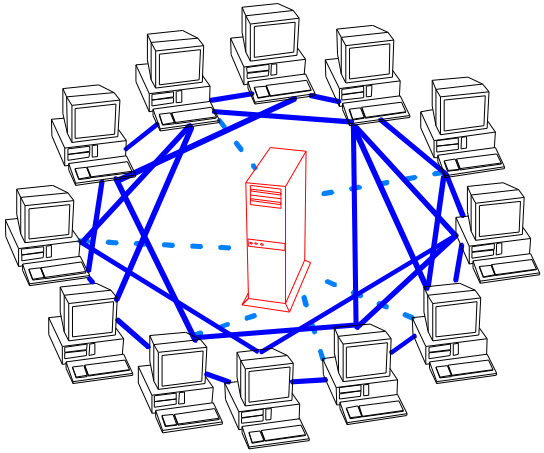


SISTEMI P2P CENTRALIZZATI: NAPSTER

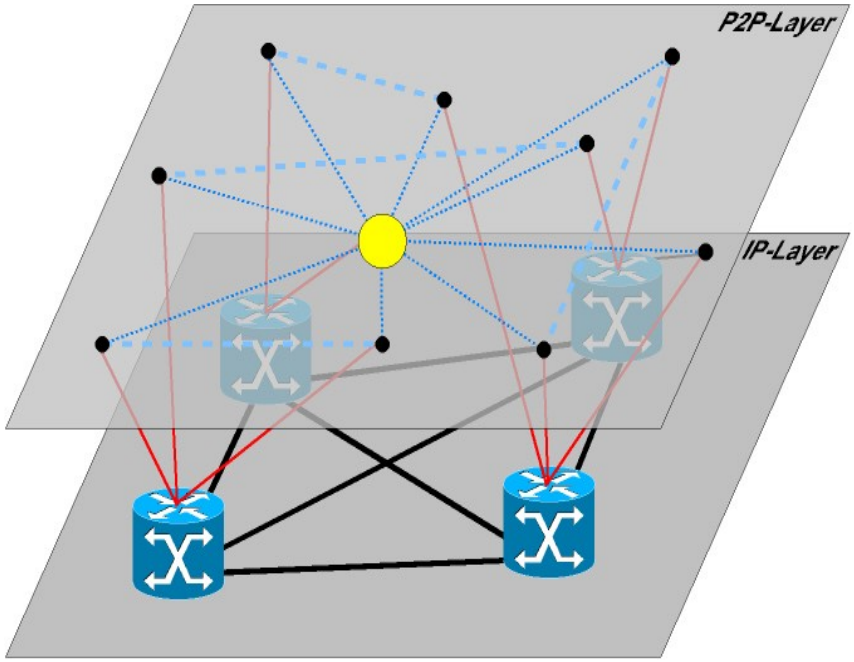
- Sul sito http: //opennap.sourceforge.net/napster.txt trovate la specifica completa del protocollo

- Tutti i peer si connettono ad un server centralizzato (broker) che indicizza I dati. La topologia dell'overlay network risultante è a stella: star overlay network.

- Le connessioni tra i peers sono stabilite dinamicamente 'on demand' per lo scambio dei i dati



NAPSTER:TOPOLOGIA



Servent
Connection between 2 servents (TCP)
Connection between router & servent
Connection between routers (Core)

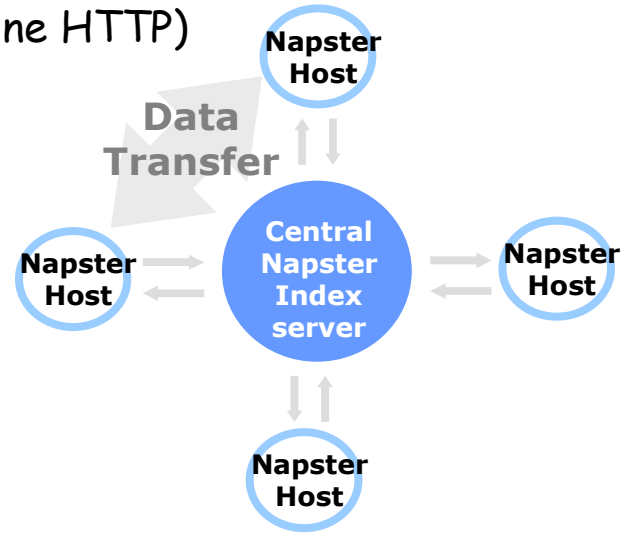
NAPSTER:PRINCIPI DI FUNZIONAMENTO

Comportamento del peer (ad alto livello)

- Si connette al Server Napster
- Invia al Server (push) le informazioni relative ai files musicali che intende condividere (metadata). Il server memorizza queste informazioni in un database centralizzato
- Ricerca un file, fornendo una lista di keywords inviate al server
- Riceve un insieme di coppie (file-peer) e sceglie una coppia in base a qulche criterio (bitrate, frequency, tipo di collegamento)
- Si connette al peer prescelto e scarica il file(connessione HTTP)

Ricerca : client server

Download: P2P



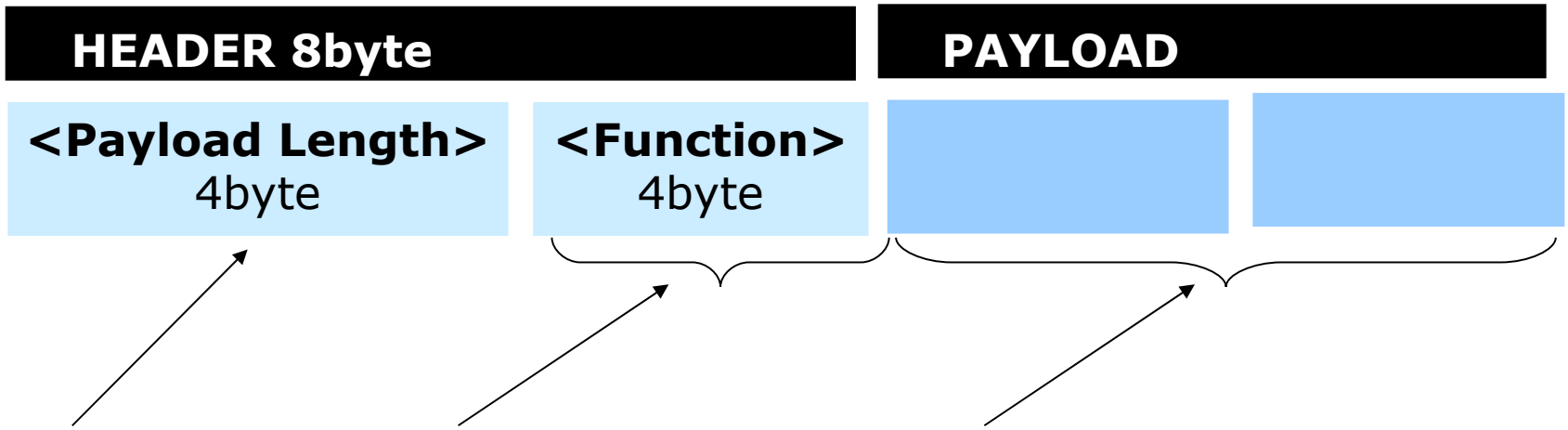
Caratteristiche

- Non esistono 'falsi negativi'Si garantisce l'individuazione di un file, se esso esiste
- Il servizio di indicizzazione costituisce "a Single Point of Failure".

NAPSTER: IL PROTOCOLLO

Protocollo plain text: il payload del messaggio contiene una stringa di caratteri ASCII

Definita una gran quantità di tipi di messaggi



Lunghezza in bytes del payload

Contiene il tipo del messaggio (e.g. login, search,…)

Contiene i parametri del messaggio(dipendente dal tipo di messaggio,e.g. keywords nel caso di search,…)

NAPSTER:REGISTRAZIONE

Client/Server Service

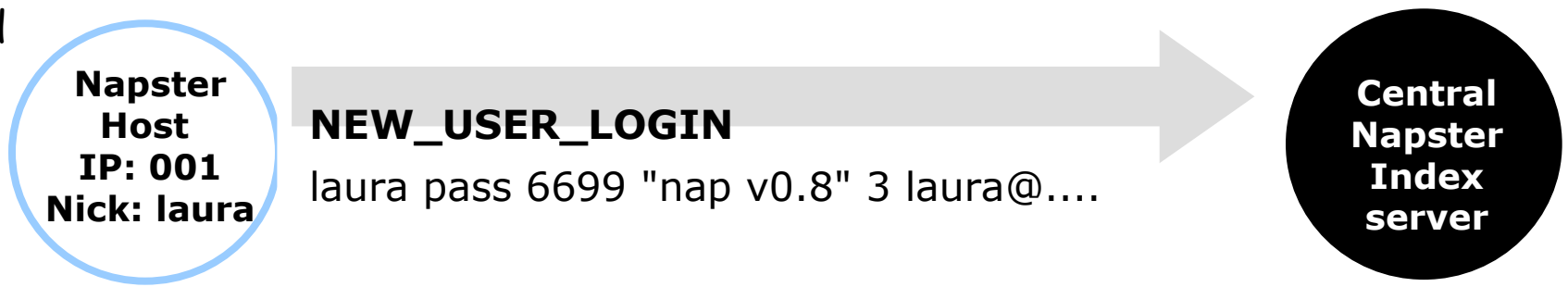
1: NEW USER LOGIN

<Nick>	<Password>	<Port>	<Link-type>	<e-mail>

- Utilizzato quando l'utente si collega per la prima volta per registrare un nickname ed una password con cui l'utente si sutenticherà

Messaggio di nick check per verificare l'unicità del nickname

- Il server risponde con un msg di nickname not registered/already registered/ invalid



NAPSTER LOGIN

Client/Server Service

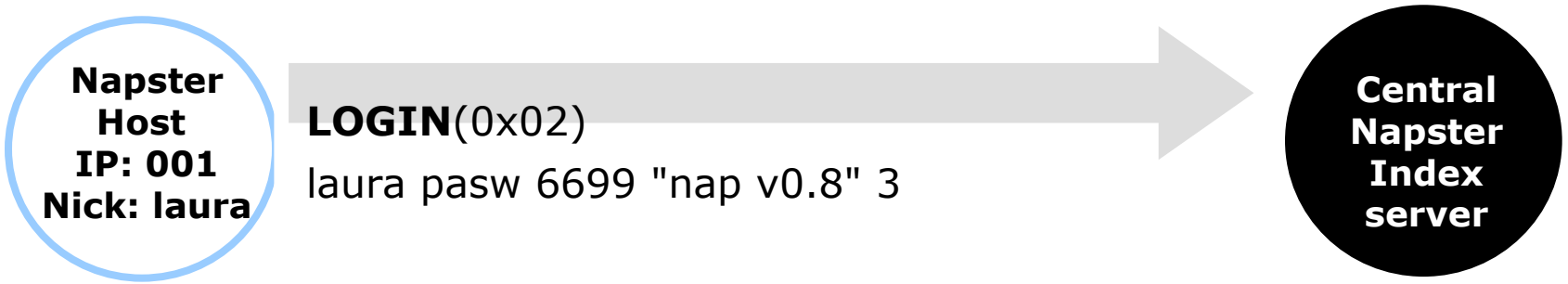
1: LOGIN (Function:0x02)

<Nick>	<Password>	<Port>	<Client-Info>	<Link-type>

2: LOGIN ACK (Function: 0x03)

3: NOTIFICATION OF SHARED FILE (0x64)

„<Filename>“	<MD5>	<Size>	<Bitrate>	<Freq>	<Time>



NAPSTER LOGIN

- Nick: nickname scelto dall'utente al momento della registrazione
- Password: Password di identificazione
- Porta: identifica la porta su cui un peer si mette in ascolto di richieste di connessione da parte degli altri peer per il trasferimento di files in modalità P2P
- Specifica la versione del client NAPSTER che il peer sta utilizzando per interagire con il server
- Link type: Valore intero che indica la banda del peer che sta effettuando il login

Example: foo badpass 6699 "nap v0.8" 3

NAPSTER LOGIN

Client/Server Service

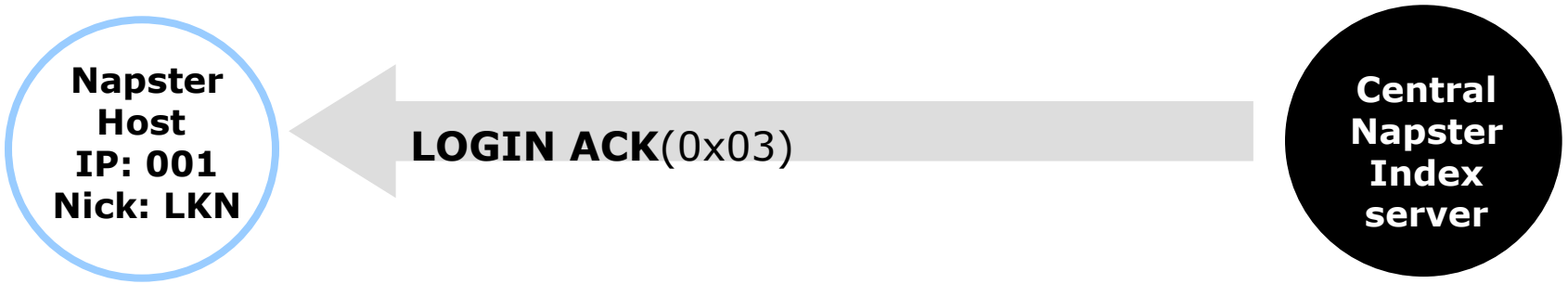
1: LOGIN (Function:0x02)

<Nick>	<Password>	<Port>	<Client-Info>	<Link-type>

2: LOGIN ACK (Function: 0x03)

3: NOTIFICATION OF SHARED FILE (0x64)

„<Filename>“	<MD5>	<Size>	<Bitrate>	<Freq>	<Time>



NOTIFICATION OF SHARED FILES

Client/Server Service

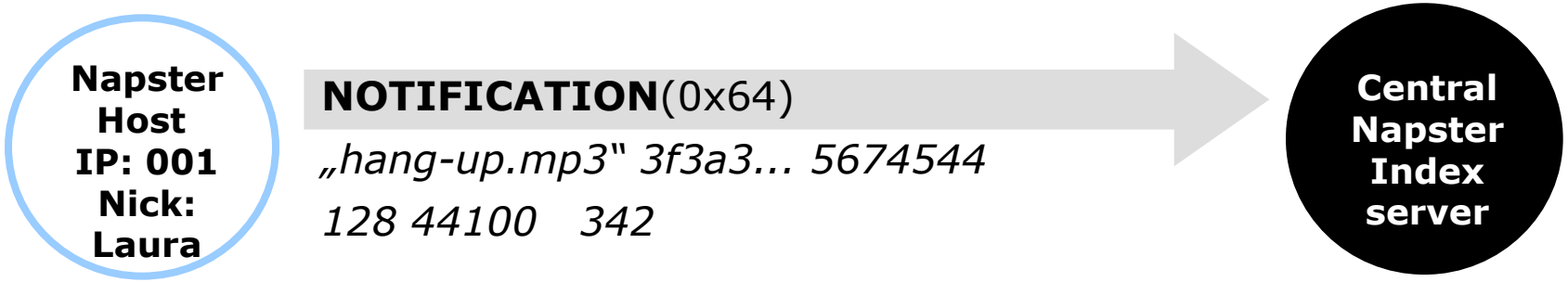
1: LOGIN (Function:0x02)

<Nick>	<Password>	<Port>	<Client-Info>	<Link-type>

2: LOGIN ACK (Function: 0x03)

3: NOTIFICATION OF SHARED FILE (0x64)

„<Filename>“	<MD5>	<Size>	<Bitrate>	<Freq>	<Time>



NOTIFICATION OF SHARED FILE

- MD5 (acronimo di Message Digest algorithm 5, Rivest 91) è un algoritmo di hashing utilizzato per la crittografia dei dati
- MD5 prende in input una stringa di lunghezza arbitraria e produce in output una stringa a 128 bit (32 valori esadecimali, indipendentemente dalla stringa di input).
- l'output (noto anche come "MD5 Checksum" o "MD5 Hash" o "fingerprint") restituito è con alta probabilità univoco e costituisce una firma digitale dell'input
- È praticamente impossibile ottenere date due diverse stringhe in input la medesima stringa in output
- È estremamente complesso risalire alla stringa di input partendo dalla stringa di output (la gamma di possibili valori in output è pari a 16 alla 32esima potenza).
- In Napster di solito si calcola l'hash dei primi 300K del file

NAPSTER LOGIN

- MD5 (fingerprint) – Introdotto originariamente per
  - identificare univocamente file MP3 identici offerti da utenti diversi ed identificati da parole chiave diverse
  - tenere traccia dei files duplicati nel sistema
  - facilitare il download del file da utenti diversi ed il resume dei download
- JAVA fornisce le classi per il calcolo dell'MD5 (o SHA1,2)

  MessageDigest md = MessageDigest.getInstance("MD5");
- Nel processo contro NAPSTER, gli amministratori di NAPSTER sostennero che era difficile individuare chi aveva registrato nel materiale illegale (sotto copyright)
  - Lo stesso materiale identificato da chiavi diverse
  - difficile il filtro su campi testuali

NAPSTER LOGIN

- Ma....nel processo contro NAPSTER fu messo in evidenza che le canzoni scaricate illegalmente sul sistema (ad esempio i CD appena lanciati dalle case discografiche) potevano essere individuati facilmente tramite il loro MD5

  The "fingerprints" of copyrighted sound recordings

  could, and should, be used by Napster to block access to plaintiffs

plaintiff(=parte lesa)

- Risultato: questo campo venne tolto in seguito alla causa legale sostenuta dalle case discografiche contro NAPSTER
- Altri campi:

  Size – Dimensione in byte dl file

  BitRate, Frequency – indicano la qualità dell'MP3

  Time - Durata della canzone

NAPSTER LOGIN

Bitrate

- File mp3 è un flusso di bit suddiviso in frames (simili ai fotogrammi di un film).
- Il bitrate è il valore che indica quanti bit vengono usati per codificare un secondo di musica .
- Si esprime in kilobit per secondo (kbps) e
- Maggiore sarà la quantità di bit utilizzati migliore sarà la resa
- Normalmente, la qualità in ascolto è proporzionale al bitrate, dunque bitrate sempre più alti garantiscono sicuramente qualità superiore.

Frequenza

- Frequenza di campionamento del suono

NAPSTER: RICERCA

1: SEARCH (Function: 0xC8)

[FILENAME CONTAINS „Search Criteria“] [MAX_RESULT <Max>]

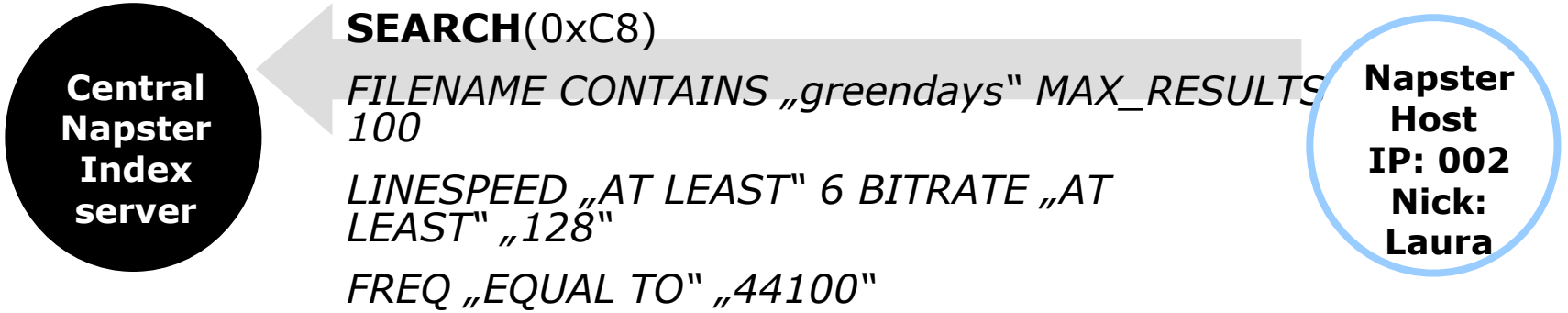
[LINESPEED <Compare> <Link-Type>]

[BITRATE <Compare> “<Bitrate>” [FREQ <Compare> “<Freq>”]

2: SEARCH RESPONSE (Function: 0xC9)

„<Filename>“ <MD5> <Size> <Bitrate> <Freq>

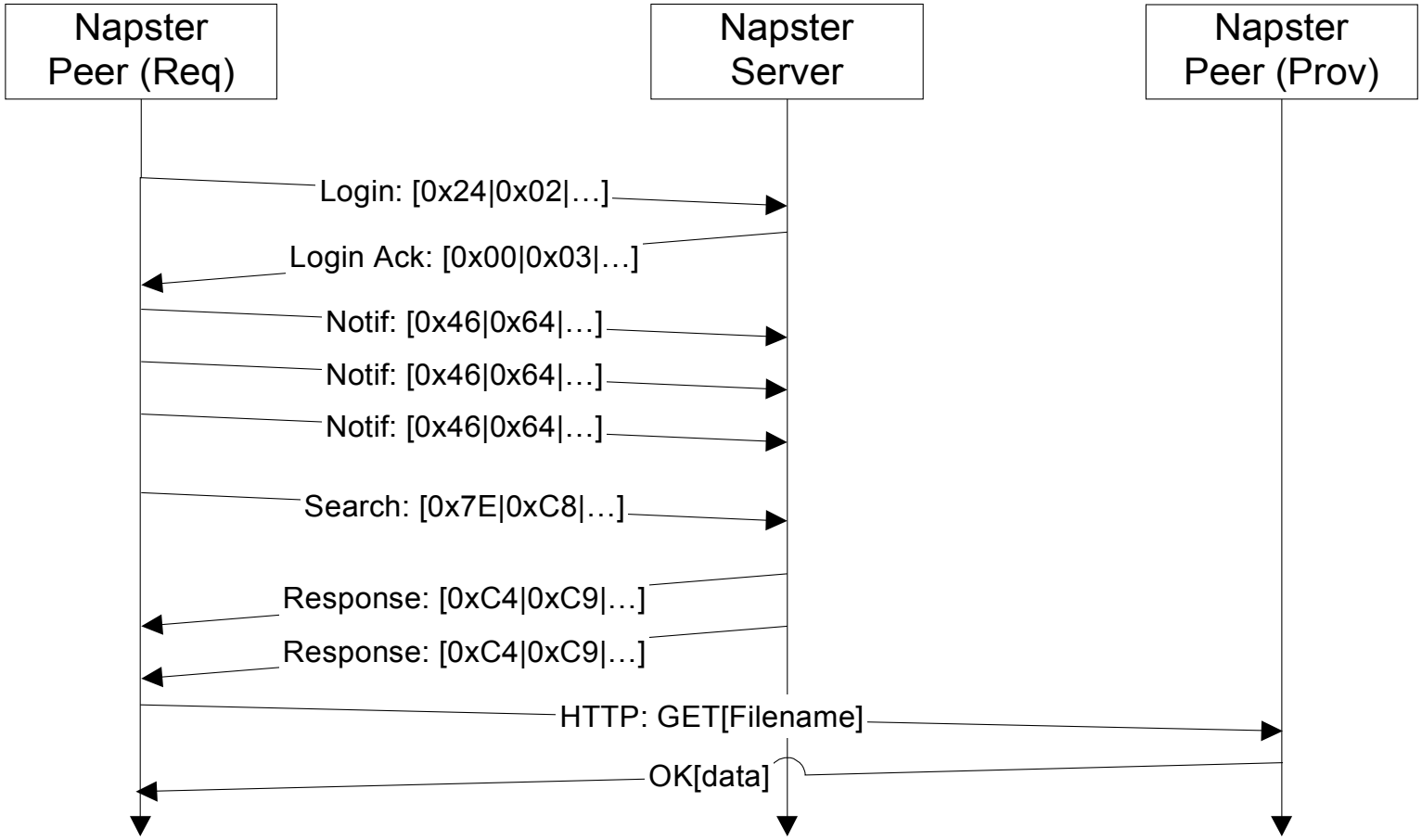
<Time> <Nick> <IP> <Link-Type>



NAPSTER: RICERCA

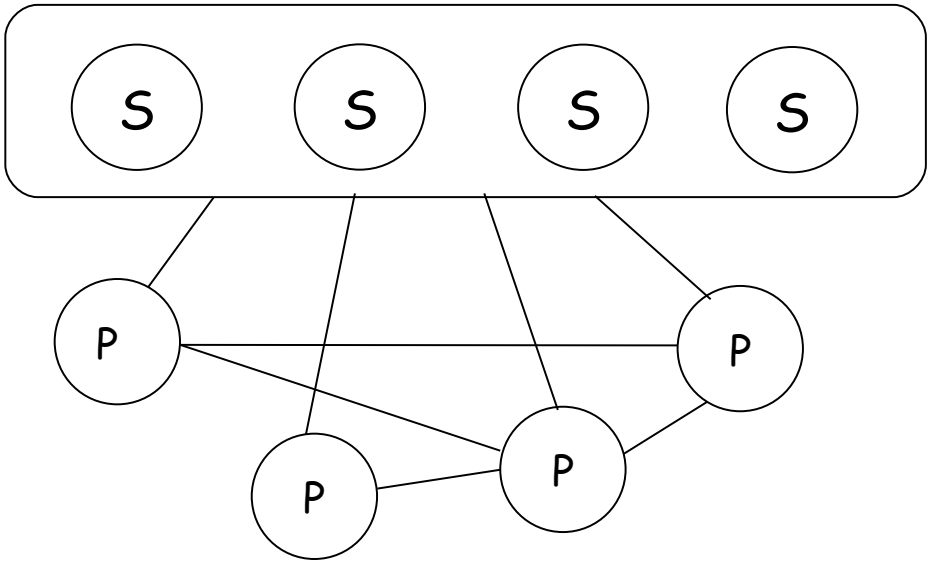
- Quando il server riceve un messaggio di SEARCH, ricerca nel proprio database un file che soddisfi i parametri della ricerca.
- Il database è creato con i dati ricevuti dai peer
- Se la query è soddisfatta, il server risponde con almeno un messaggio di serach response
- Il messaggio di query response contiene:
  - il nome completo del file
  - le sue caratteristiche. Viene spedito anche l'MD5 del file. Questo verrà utilizzato per controllare l'integrità del file scaricato
  - Le caratteristiche dell'host che possiede il file
- Il file viene scaricato utilizzando il protocollo HTTP

NAPSTER: IL PROTOCOLLO SEMPLIFICATO



Sequenza semplificata (rispetto al protocollo reale) di messaggi di una rete Napster con due peer. Il peer req richiede un file al Peer Prov

NAPSTER: IL PROBLEMA DELLA SCALABILITA'

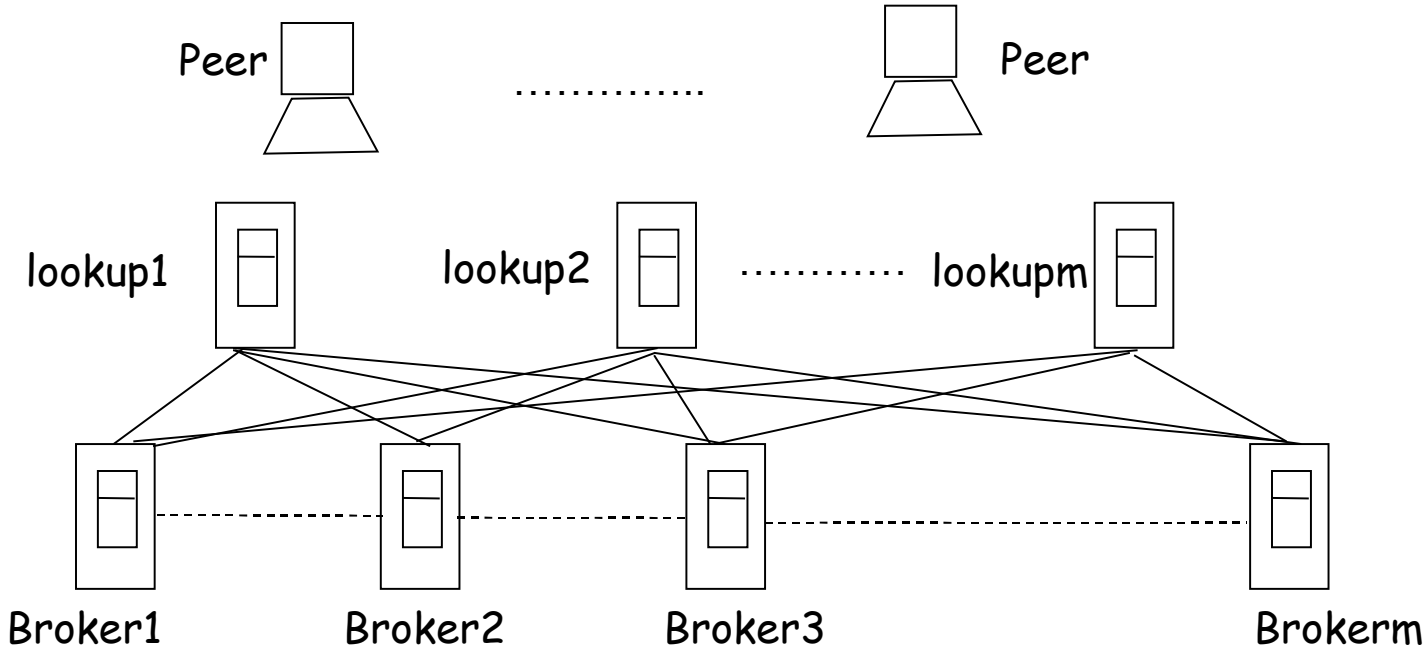


- Per aumentare la scalabilità del sistema si può definire una server farm
- I server sono localizzati in un'unica locazione geografica
- Single Point of Failure: Esiste comunque un unico punto di accesso al sistema (esempio il/i router che collega il cluster alla rete)
- i server condividono il database contenente le meta-informazioni

NAPSTER: IL PROBLEMA DELLA SCALABILITA'

Bilanciamento del carico: vengono definiti m lookup servers ed n brokers

- Brookers: gestiscono i metadata (informazioni sui files)
- Look-up servers: ricevono le richieste dagli utenti e smistano le richieste ai broker in modo da bilanciare il carico. Ogni look-up server è collegato ad ognuno dei broker per conoscere il suo carico
- I brokers sono interconnessi in modo da condividere i metadati



NAPSTER: ARCHITETTURA

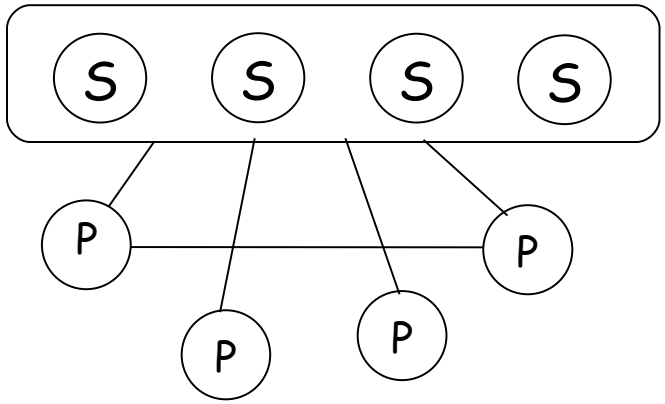
Bilanciamento del carico

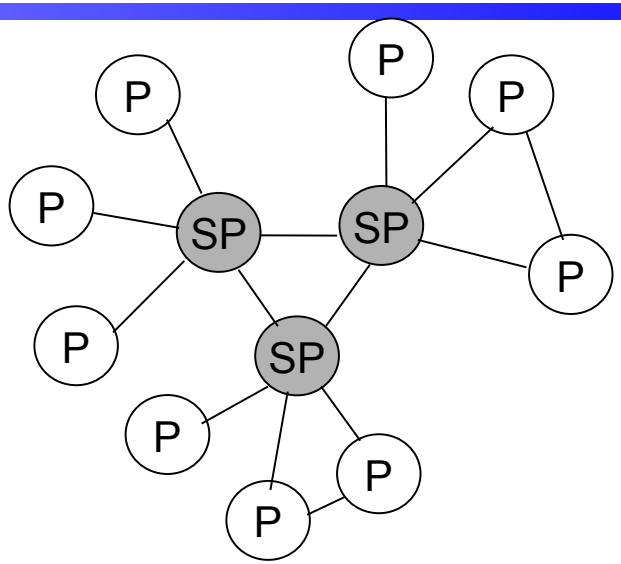
- L'intero sistema può essere acceduto mediante un unico punto di entrata, rappresentato mediante un unico nome simbolico (es: server.napster.com, porta 8875)

- Il DNS locale può effettuare il bilanciamento del carico, distribuendo le richieste ai diversi look-up servers o direttamente ai brokers

- Tecnica utilizzata per server che devono gestire un alto tasso di richieste (es: cnn.com, java.sun.com,…)
  - Si definisce un server farm
  - Il DNS associa al nome simbolico dell'host un insieme di indirizzi IP
  - Il DNS restituisce tutto l'insieme di indirizzi IP al richiedente, ma ruota l'ordine degli indiriirizzi. Poiché il client sceglie in genere il primo indirizzo IP, si ottiene un bilanciamento del carico sui diversi

SERVER FARM vs. P2P IBRIDO





- i servers sono collocati in un'unica locazione geografica

⇒

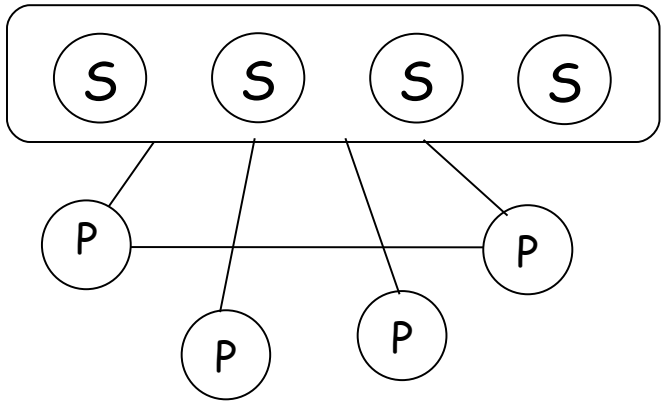
maggiore individuabilità

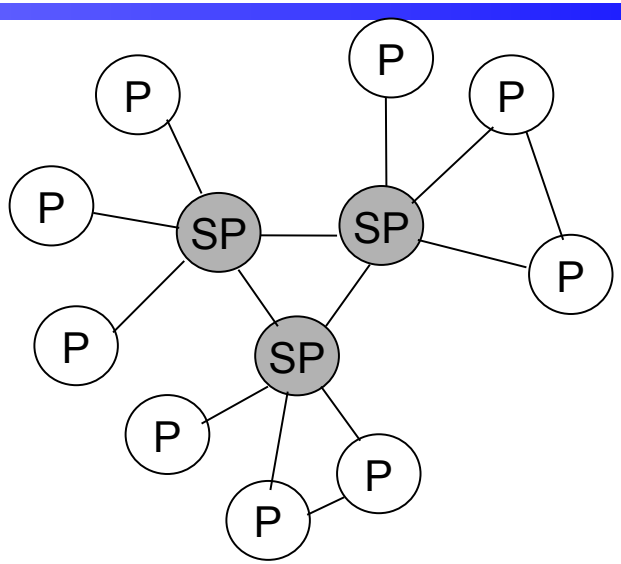
- un server ha solo funzionalità di coordinamento (full time activity)

- L'insieme di servers è geograficamente distribuito

- Il super peer è un peer che si assume anche funzionalità di coordinamento (part time activity)

P2P ibrido vs. Server Farm





- L' insieme dei servers è definito staticamente
- Unico punto di accesso al sistema
- scarsa tolleranza ai guasti

- Super peer sono eletti dinamicamente
- Nel caso di fault di un superpeer il sistema si riconfigura eleggendo altri superpeers

NAPSTER: IL PROTOCOLLO COMPLETO

- Un peer P riceva da un broker Napster una lista di peer che condividono una canzone ricercata. P sceglie un peer S dalla lista.

- La creazione della connessione tra P ed S viene coordinata dal broker B. Questo consente ai peer di accettare connessioni solo se "certificate" da B.
  - P notifica a B che intende scaricare un file da S
  - B notifica ad S la richiesta di P
  - S notifica a B la propria disponibilità ad accettare la connessione
  - B notifica a P la risposta di S
  - A questo punto P può aprire una connessione con S per il trasferimento dei dati

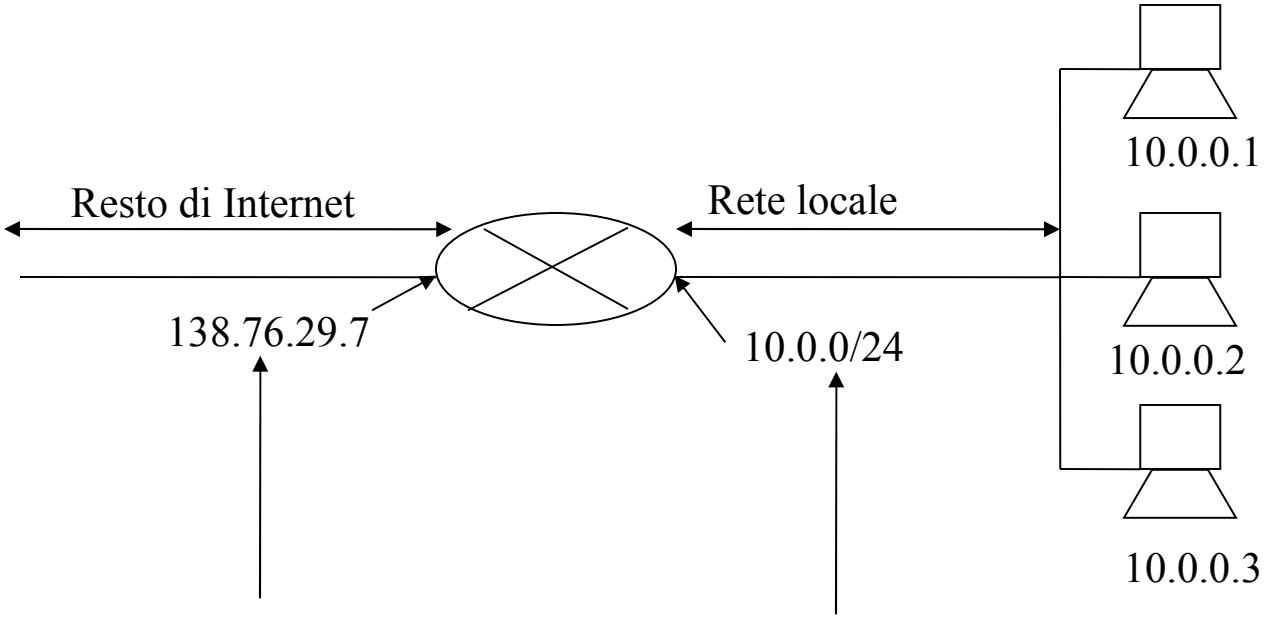
- Se S non è nattato/si trova a monte di un firewall
  - P apre una connessione verso S e il file viene scambiato su tale connessione

- Se S è nattato/si trova a monte di un firewall
  - Si deve eseguire una procedura di PUSH

NAT (NETWORK ADDRESS TRANSLATION)

- NAT (Network Address Translation) = Tecnica di filtraggio di pacchetti IP con sostituzione degli indirizzi (mascheramento)
- NAT = Consente la connessione di un insieme di hosts ad Internet utilizzando un unico indirizzo IP
- Vantaggi
  - risparmiare indirizzi IP (in attesa della piena diffusione di IPv6)
  - facilita l'amministazione della rete
  - aumenta la sicurezza

NETWORK ADDRESS TRANSLATION



Tutti i datagrammi che escono dalla rete hanno lo stesso indirizzo NAT 138.76.29.7, ma diversi numeri di porta

Tutti i datagrammi con sorgente/destinazione in questa sottorete hanno indirizzo 10.0.0/24 come sorgente/destinazione

NETWORK ADDRESS TRANSLATION

Motivazioni

- Risparmio di indirizzi IP: il provider attribuisce un solo indirizzo ad un insieme di hosts appartenenti alla stessa organizzazione. L'utente risparmia sul costo della connessione

- Facilità di amministrazione della rete:
  - si possono modificare gli indirizzi nella rete locale senza notificarlo al mondo esterno
  - si può cambiare l'indirizzo del provider senza modificare gli indirizzi della rete locale

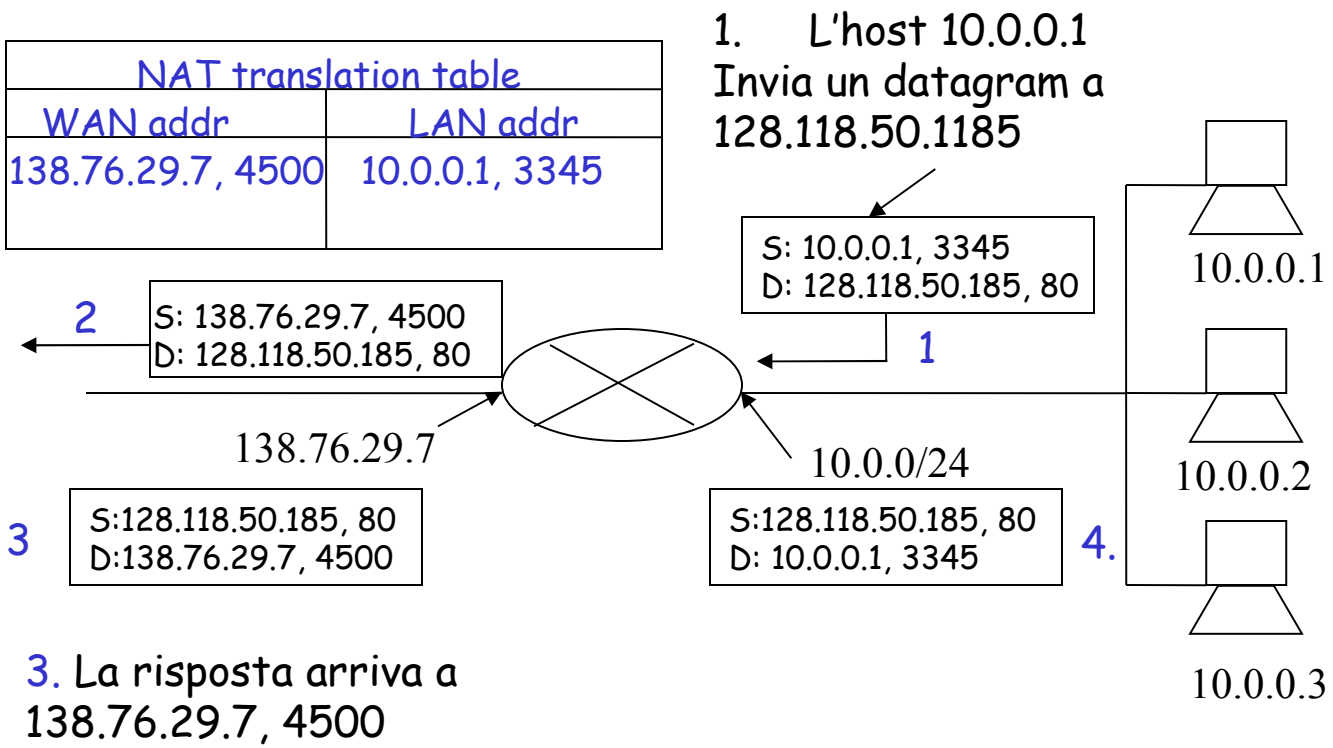
- Sicurezza gli hosts nella rete locale non sono visibili dall'esterno e quindi indirizzabili direttamente dall'esterno

NETWORK ADDRESS TRANSLATION

Funzioni di un router NAT

- sostituire in ogni datagram uscente la coppia (IP sorgente, #porta) con (NAT IP, #nuovaporta) dove #nuovaporta è un nuovo numero di porta generato dal NAT
- registrare in una tabella di traduzione la corrispondenza tra le due coppie
- sostituire in ogni datagramma entrante la coppia (NAT IP, #nuovaporta) con il corrispondente (IP sorgente, #porta) memorizzato nella tabella di traduzione

NAT:NETWORK ADDRESS TRANSLATION



In fase di ricezione, il router NAT cambia
l'indirizzo destinatario consultando l'indirizzo destinazione

NETWORK ADDRESS TRANSLATION

- NAT : funziona solo con datagram IP che trasportano pacchetti a livello trasporto spediti mediante il protocollo UDP o il protocollo TCP
- Gli indirizzi assegnati alle sottoreti interne appartengono ad una delle seguenti zone
  - 10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16
- Il router opera anche come dispositivo NAT
- Circa 60000 porte con 16 bit

  $\Rightarrow$

  circa 60000 connessioni aperte con un unico indirizzo IP

NETWORK ADDRESS TRANSLATION

- Vantaggi del NAT
  - Topologia della rete non visibile dall'esterno
  - Il NAT opera in modo trasparente per molte applicazioni
  - Spesso combinato con un firewall

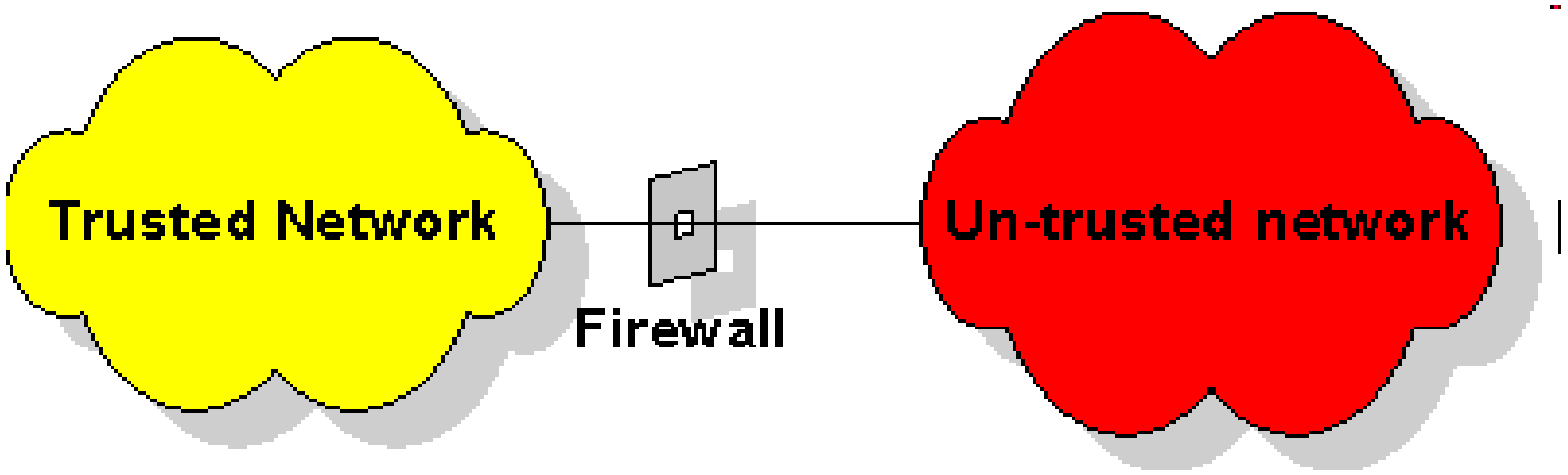
NAT: SVANTAGGI

- Rende più complessa la raggiungibilità degli host sulla rete. Un host che si trova dietro un NAT (host NATTATO) non può essere raggiunto da un host esterno alla rete locale. Questo rende più complesso lo sviluppo di applicazioni P2P

- Il router manipola i numeri di porta (livello trasporto), mentre dovrebbe operare solo fine al livello 3 (IP)

- Alcune applicazioni non sono trasparenti al NAT (esempio applicazioni che contengono indirizzi IP e numeri di porta nel payload)
  - Esempio:FTP utilizza due connessioni parallele, una per l'interazione con il server, l'altra per il trasferimento dati da e verso il server. I parametri della seconda (porta su cui spedire i dati) connessione sono inclusi nel payload della prima

FIREWALLS



• firewall: punto di controllo e di monitoraggio, collega reti con diversi livelli di affidabilità e delimita la rete da difendere (es:isola la rete interna di una organizzazione dalla rete pubblica)

•in generale, struttura hardware (o software) che separa una rete privata dal resto di Internet

• Consente la specifica di politiche di sicurezza per controllare e gestire il flusso di traffico tra il mondo esterno e le risorse interne

FIREWALLS: CARATTERISTICHE PRINCIPALI

Una semplice politica di sicurezza:

- Permettere agli utenti della trusted network di accedere a servers esterni (eventualmente scegliendo quali), ma non permettere accessi dall'esterno verso la trusted network (ad esempio ad un server locale)

- Per implementare politiche simili, il firewall deve conoscere
  - L'applicazione a cui ci si intende connettere
  - La direzione della connessione

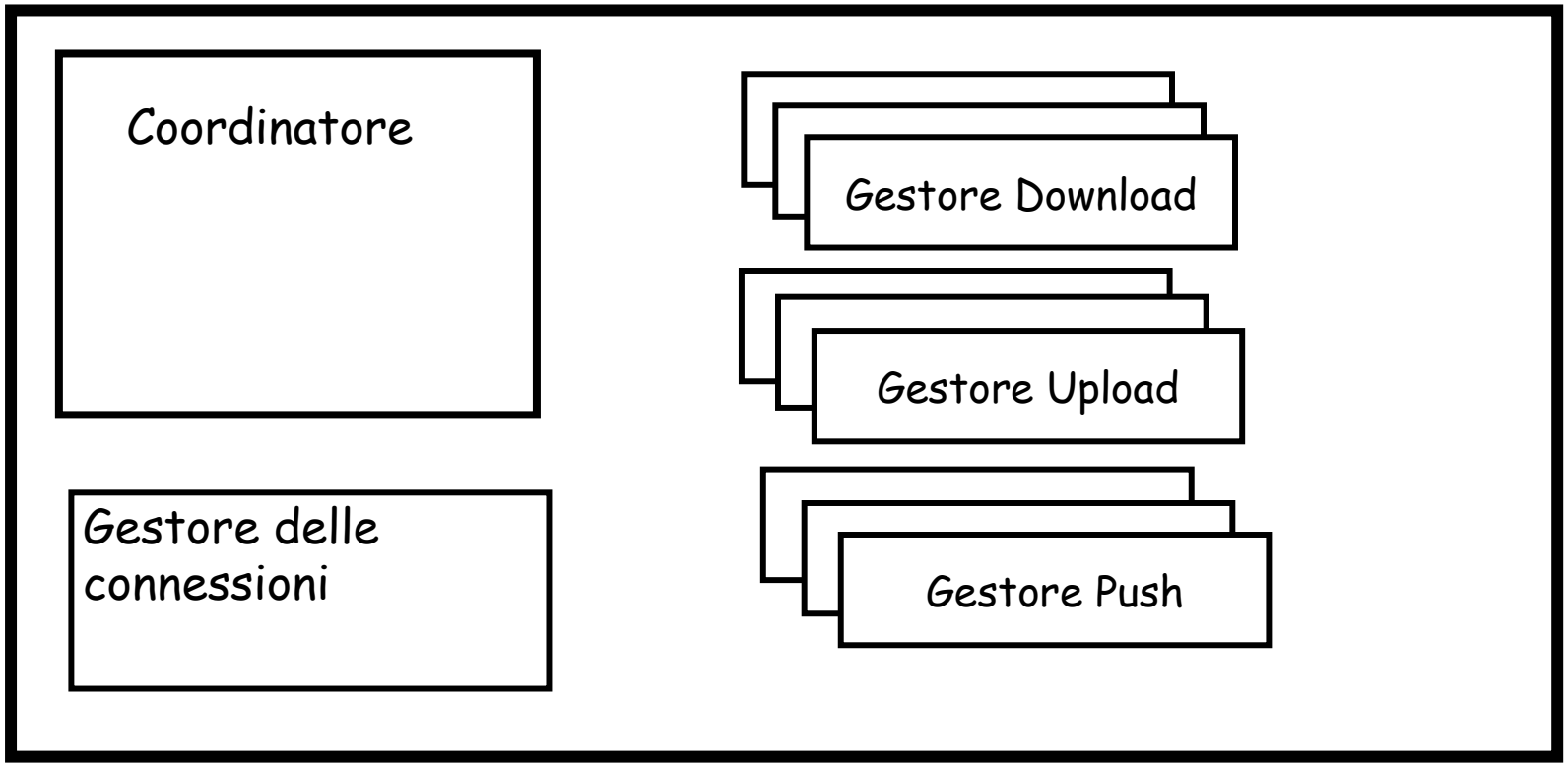
- Applicazione individuabile mediante il numero di porta (porta di ricezione, es porta 80 per www, porta 25 per e-mail,…)

- Stabilire la direzione della comunicazione può risultare più complesso

PEER NATTATI

- Alcuni peer possono trovarsi a monte di un NAT/firewall
  - In questo caso, in generale, il peer non può accettare connessioni dall'esterno
  - Il peer può però aprire connessioni verso l'esterno
  - Almeno uno dei due peer deve accettare connessioni dall'esterno, altrimenti la comunicazione non risulat possibile, a meno che entrambi I peer facciano riferimento a rely peer accessibili dall'esterno
- Nel caso in cui il peer P voglia scaricare un file da un peer S nattato, si effettua una operazione di PUSH
  - S apre una connessione verso P
  - S deve essere informato della porta su cui P sta accettando connessioni

NAPSTER: STRUTTURA DI UN PEER



NAPSTER: STRUTTURA DI UN PEER

- Il coordinatore gestisce:
  - le connessioni e le comunicazioni con il look-up server e quindi con il Broker
  - l'interazione con l'utente: ricerca di files, download, rimozione di files, etc.

- Il gestore delle connessioni si occupa di gestire tutte le richieste di connessione provenienti dagli altri peers (richieste di connessione per upload o per push)

- I gestori di download, upload,push, gestiscono lo scambio diretto dei dati tra i peers
  - Possono esistere più istanze di questi gestori. Ogni istanza si occupa di un singolo trasferimento
  - Push: utilizzato per peer che si trovano a monte del firewall.

NAPSTER: STRUTTURA DI UN PEER

- Tutte le comunicazioni avvengono tramite TCP/IP
- Per unirsi alla rete Napster il peer
  - Apre una connessione con un look-up server (porta 8875)
  - Il look-up server restituisce l'indirizzo di uno dei brokers (porta 6699)
  - Si connette al broker ed invia l'informazione relativa ai dati che vuole condividere (metadata)
  - Attende messaggi dal broker, oppure connessioni da altri peer (porta 6699) o eventi generati dall'utente locale

NAPSTER:STRUTTURA DI UN PEER

- Il download peer-to-peer di un file viene coordinato dal broker

- Download di un file: P1 vuole scaricare un file dal peer P2
  - P1 invia al proprio broker la richiesta di download (download request)
  - Il broker invia a P2 una richiesta di upload (upload request)
  - P2 invia al broker l'ack (upload accept+porta su cui effettuare la connessione)
  - Il broker invia a P1 l'ack (download accept)
  - P1 apre una connessione verso P2
  - Il file viene spedito da P2 sulla connessione aperta (upload)

- Il coordinamento centralizzato consente di introdurre un livello di sicurezza un peer non accetta una connessione se prima non ha ricevuto una richiesta di upload

NAPSTER: STRUTTURA DI UN PEER

- Se il peer P1 che possiede il file F non si trova a monte di un firewall/NAT, il peer P2 che vuole scaricare F apre una connessione sulla porta specificata dal server
- P2 accetta la connessione
- P1 invia una GET, poi invia un messaggio

  <mynick> "<filename>" <offset>

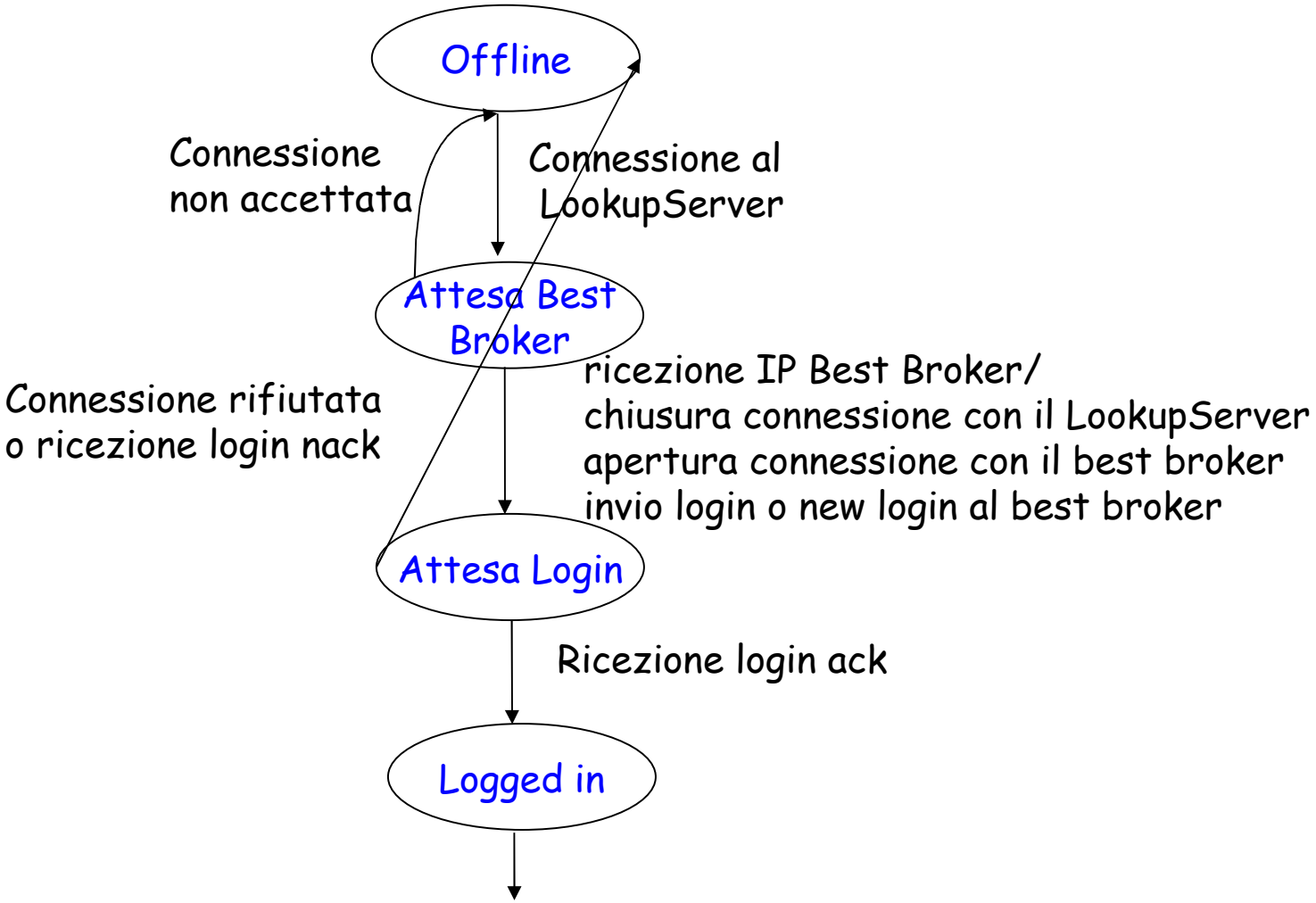
  <offset> è l'offset in F che indica il primo byte che si intende scaricare (0 la prima volta)
- P2 restituisce la lunghezza del file, oppure un messaggio di errore
- Se non ci sono errori P2 inizia ad inviare il file

NAPSTER: STRUTTURA DI UN PEER

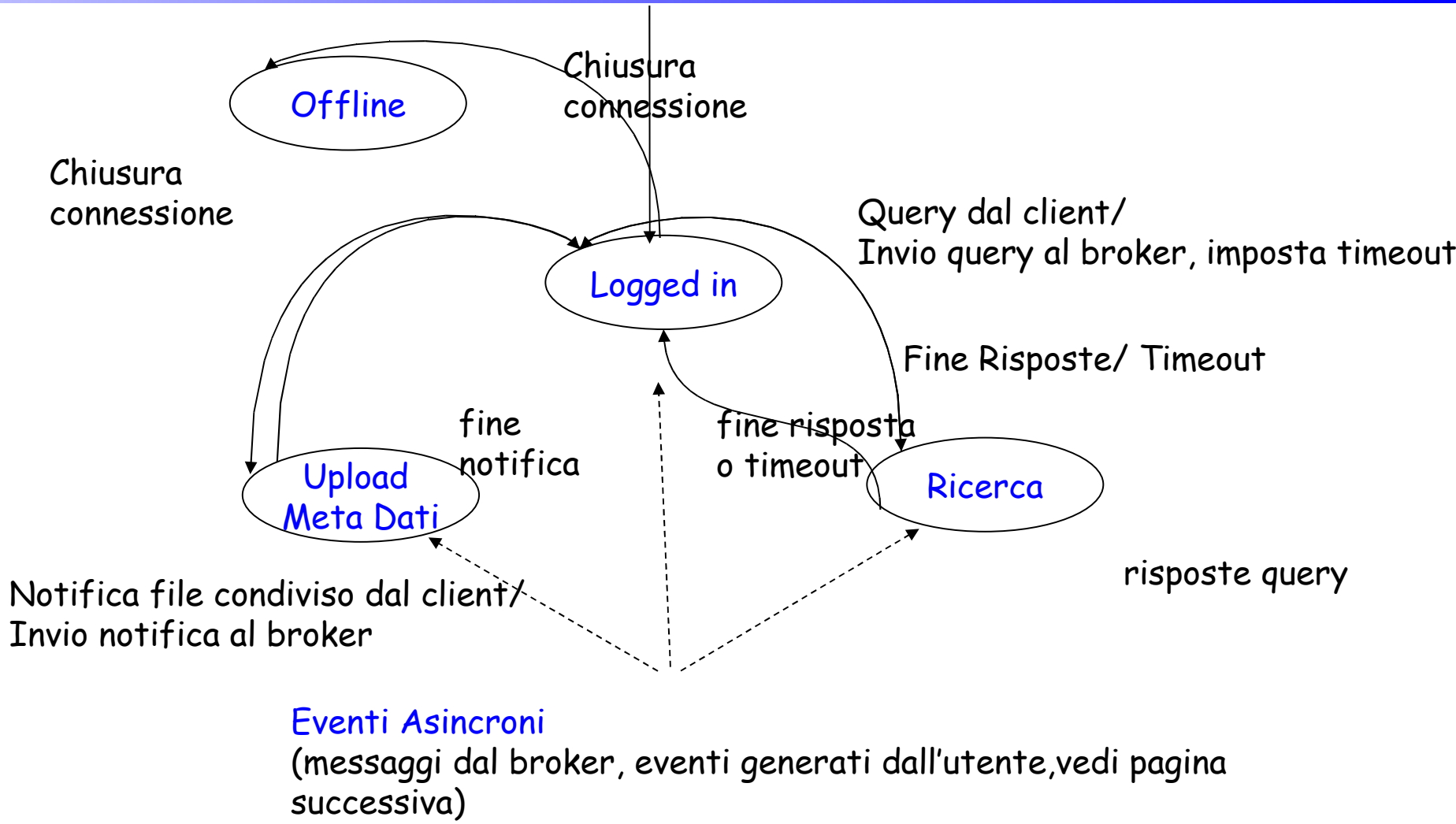
P1 vuole scaricare un file da un peer P2 che si trova a monte di un firewall

- P1 invia al proprio broker la richiesta di download (download request)
- Il broker invia a P2 una richiesta di upload (upload request)
- P2 invia al broker l'ack indicando la disponibilità di effettuare l'upload, ma l'impossibilità di accettare connessioni dall'esterno (upload accept, porta = 0)
- Il broker invia a P1 l'ack (download accept, porta =0)
- P1 dichiara di accettare la connessione da P2 (alternate download request) e specifica la propria porta
- Il broker invia a P2 l'ack (alternate download ack)
- P2 apre una connessione verso P1
- Il file viene spedito da P2 sulla connessione aperta (push)

STRUTTURA DI UN PEER:IL COORDINATORE



STRUTTURA DI UN PEER: IL COORDINATORE



STRUTTURA DI UN PEER: IL COORDINATORE

Eventi Asincroni

- Ricezione di statistiche da parte del server
- Richiesta di cancellazione di un file (dall'utente) / invio meta dati al broker
- Richiesta di download da parte dell'utente / invio download request al server
- Richiesta di upload da parte del broker / invio (upload accept + porta, porta ≠ 0 se posso accettare connessione, porta = 0 se non posso accettare connessioni, perché a monte di un firewall) oppure accept failed
- Ricezione di dowload ack con porta > 0/ apertura connessione con il peer remoto, attivazione di un'istanza del Gestore di Download
- Ricezione di download ack con porta = 0/ invio di alternate-download request al server o rifiuto ad accettare connessione (se il peer si trova a monte di un un NAT o di un firewall)
- Ricezione di alternate download ack/ attivazione di un'istanza del Gestore di Push

NAPSTER: GESTIONE DELLE CONNESSIONI

Gestore delle connessioni: schema generale

Richiesta PUSH/
attivazione di
una istanza del Gestore
di Download

Richiesta GET/
attivazione di una istanza del
Gestore di Upload

- Gestore delle connessioni: si mette in attesa di ricevere una richiesta di connessione (esempio accept su un ServerSocket in JAVA)
- Se la richiesta di connessione arriva da un peer che intende scaricare il file (richiesta GET) si attiva un'istanza del gestore di Upload per inviare i dati
- Se la richiesta di connessione arriva da un peer che possiede il file, ma si trova a monte di un NAT (richiesta di PUSH), il peer stesso poi mi invierà i dati su quella connessione e quindi si deve attivare una istanza del gestore d Download per ricevere i dati

NAPSTER: STRUTTURA GENERALE DEL PEER

- Gestore Download = riceve il file dal peer remoto e lo memorizza sul disco (esegue una sequenza di receive)
- Gestore Upload = spedisce il file al peer remoto che lo ha richiesto (esegue una sequenza di send)
- Gestore Push = apre una connessione verso il peer remoto e spedisce il file richiesto (come sopra, ma apre anche la connessione)

NAPSTER: VALUTAZIONE DEL PROTOCOLLO

- Consideriamo un utente che condivide 10 files e richiede un file che è condiviso da altri 20 utenti
- Traffico sulla rete generato dal protocollo (solo messaggi per l'implementazione del protocollo, non download

$$(\text{login} + \text{login-ack}) + 10^* \text{notif} + 1 * \text{serch} + 10 * \text{response}$$

login = 40 bytes, login_ack= 4 bytes, notif = 74 bytes,

search = 130 bytes, response = 200 bytes

$$40 + 4 + 10^* 74 + 130 + 10^* 200 = 2914 \text{ bytes}$$

- 2914 bytes per la gestione del protocollo

NAPSTER: VANTAGGI E SVANTAGGI

- Svantaggi: Server centralizzato
  - Collo di bottiglia
  - Limitata scalabilità
  - Poco affidabile (single point of failure)
- Vantaggi
  - Ricerca veloce (one hop lookup)
  - Ricerca completa = assicura che, se una informazione esiste nel sistema, essa viene individuata
  - L'entità centralizzata garantisce funzionalitàm di controllo/sicurezza
  - Minimizzazione dei messaggi spediti sulla rete. Non richiesti messaggi per la riconfigurazione dinamica della rete (vedi keep-alive in Gnutella)
- Idee di base ripresa da
  - BitTorrent
  - E-mule- E-donkey
  - .....

EVOLUZIONE DI NAPSTER

- Definizione di un insieme di servers geograficamente distribuiti
  - Il server possiede solo funzionalità di indicizzazione (non può essere un peer)
  - Alto numero di servers più o meno interconnessi
  - Si stabilisce staticamente quali nodi eseguono i servers
  - Full time activity
  - Esempio: e-mule, protocollo e-donkey

- Alcuni peer assumono anche funzionalità di indicizzazione.
  - Super Peers, Rendez-vouz Peer
  - I super Peer vengono eletti dinamicamente e cooperano
  - Part Time Activity:assumono funzione di servers, ma continuano a funzionare come peers
  - Es: JXTA