Přenos dat, počítačové sítě a protokoly

Sítě peer-to-peer (P2P)

Ing. Petr Matoušek, PhD., M.A.

Fakulta informačních technologií VUT v Brně

matousp@fit.vutbr.cz

Problém malého světa

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Sociologický experiment *Problém malého světa* (The Small-World Problem)

- Studie amerického psychologa Stanleyho Milgrama z roku 1967 [1].
- Výzkumné otázky:
  - *Jaké je pravděpodobnost, že dva náhodně vybraní lidé na světě se znají?*
  - *Pokud se neznají osobně, před kolik známých je lze propojit?*

  Jinak řečeno, jak dlouhý by byl řetězec `X-a-b-c...-y-Z` pro projení X a Z?

Hypotéza: Existuje vůbec nějaká matematické struktura ve společnosti?

- Důsledek zajímavý pro různé obory: historii, sociologii, komunikaci, biologii a další.
- Např. historikové Henri Pirenne a George Duby se domnívají, že po pádu Říše římské byla na období cca 400 let přerušena komunikace mezi tehdejšími městy západní Evropy, čímž došlo k růstu izolace místních komunit a odlišnému vývoji.

$\Rightarrow$ Pokud určitá (skrytá) matematická struktura ve společnosti existuje, může hrát významnou roli ve vývoji a komunikaci.

Problém malého světa

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Jiná formulace problému malého světa

- Mějme universum, které obsahuje cca 200 miliónů bodů (obyvatel USA v r. 1960).
- Přes kolik mezilehlých uzlů můžeme vytvořit nejkratší cestu mezi dvěma libovolnými uzly?
- Cestu hledáme pouze na základě lokálních znalostí (není globální směrování).

Může to fungovat? (dva pohledy)

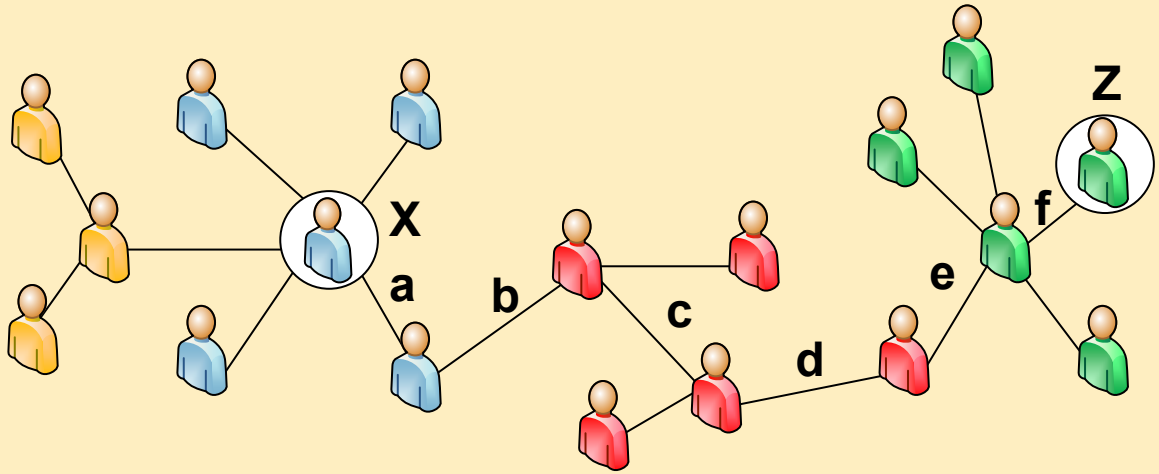
\+ Dva libovolní lidé X a Z kdekoliv na světě mohou být propojeni přes relativně malý počet známých.

– Existují nepřekonatelné propasti mezi různými skupinami ve světě. Nelze propojit dva lidi přes okruh známých, neboť jejich okruhy známých nemusí mít průnik.

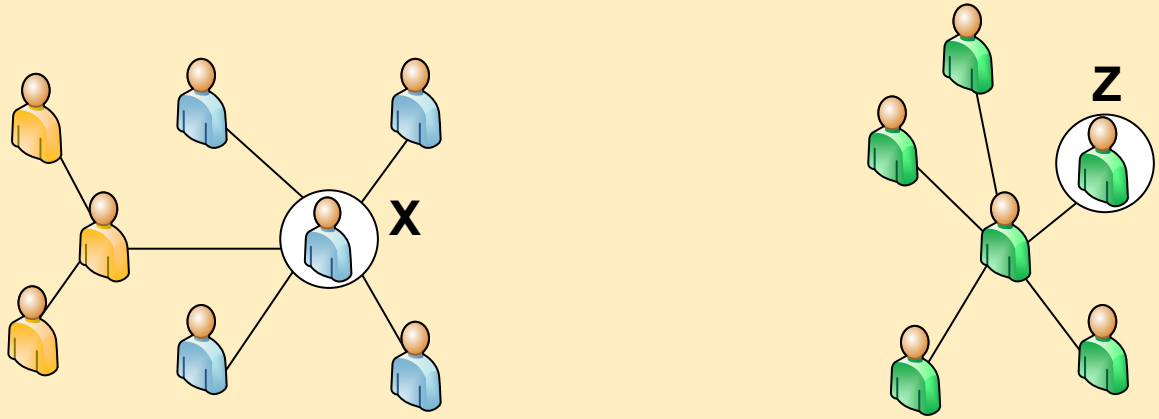
- Teoretické studie na MIT (Pool, Kochen)
- Obvykle má člověk cca 500 známých. Pravděpodobnost, že se znají dva Američané je 1:200,000. ⇒ Jak to ale ověřit v praxi?

Problém malého světa

Pohled 1: Propojení přes prostředníky existuje



Pohled 2: Propojení neexistuje kvůli chybějícím vazbám



Problém malého světa

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Milgramův experiment (výzkumný grand Harwardu, 680 USD)

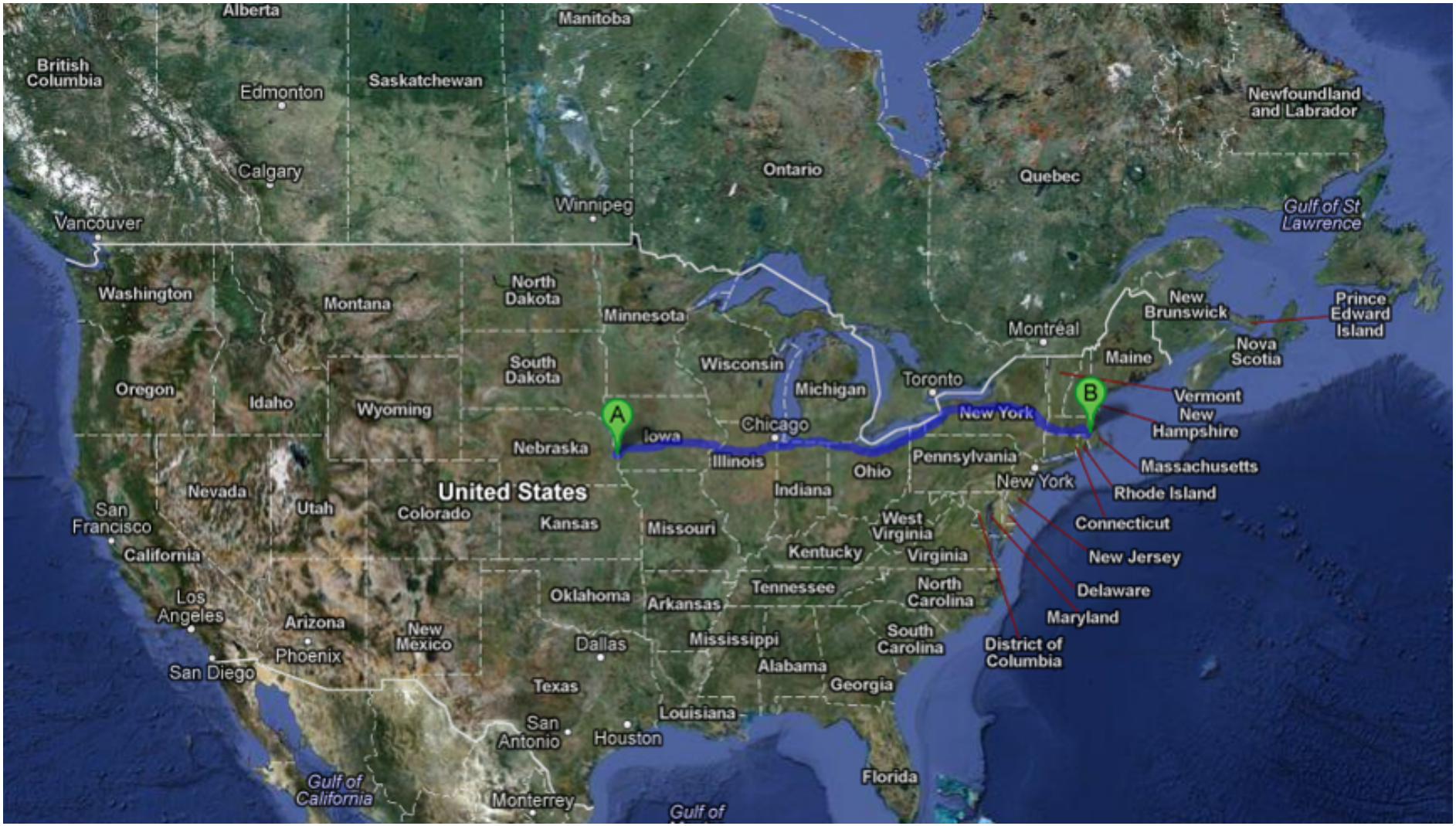
- Úkolem je doručit dopis příteli v městě Sharon, Massachusetts.
- Odesilatelem je 160 náhodně vybraných lidí z města Omaha, Nebraska.
- Dopis lze předávat pouze osobně skrze známé, které znám křestním jménem.
- Odesilatelé mají dostatek znalostí o adresátovi
  - Jméno, kde bydlí, povolání, kde studoval, apod.
  - Dopis předávají pouze na základě svých vlastních znalostí, vazeb a priorit.
- Předávající budou posílat na Harvard potvrzující dopisy pro sledování cesty.

Výzkumné otázky

- Bude to vůbec fungovat? Dojde nějaký dopis k cíli?
- Kolik lidí bude potřeba pro nalezení adresáta?
- Jaká bude vzdálenost (v km) mezi jednotlivými předáními?

Problém malého světa

- Vzdálenost měst Sharon a Omaha je cca 2.322 km (asi jako Brno a Madrid)



Problém malého světa

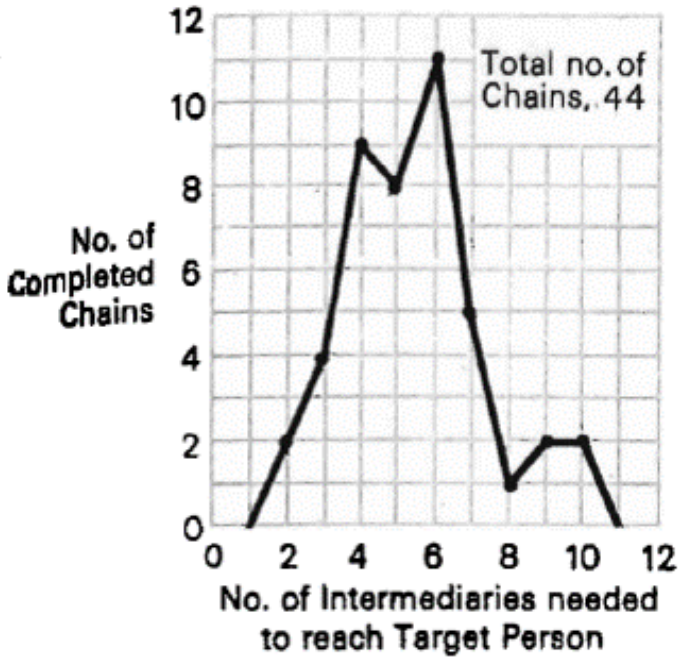
Výsledek Milgramova experimentu

Průběh prvního experimentu (Nebraska Study)

- Celkem odesláno 160 dopisů
- K cíli došlo 44
- Některé skončily blízko cíle

Výsledky:

- Nejkratší řetězec: dva prostředníci
- Nejdelší řetězec: jedenáct prostředníků
- Medián cesty: pět prostředníků



In the Nebraska Study the chains varied from two to 10 intermediate acquaintances with the median at five.

- Malý svět → stačí pět prostředníků.
- John Guare: "Šest stupňů odloučení" (Six Degrees of Separation)

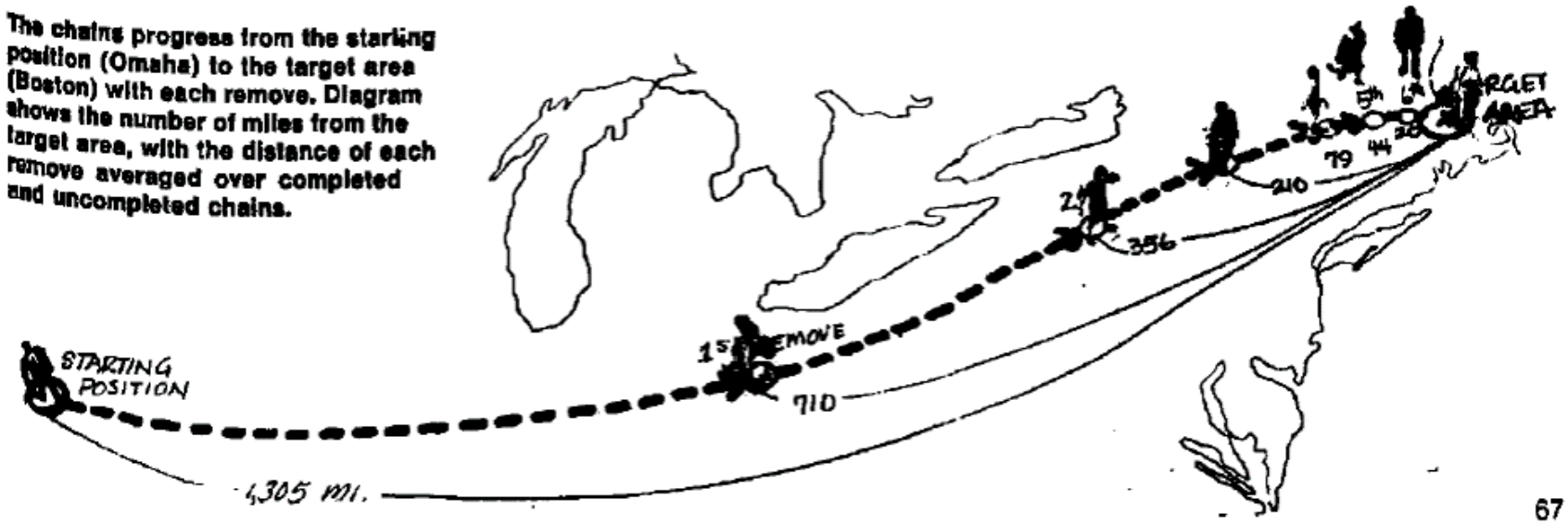
Problém malého světa

Výsledek Milgramova experimentu: další zajímavosti

Kansaský experiment (Kansas Study)

- Poměr předávání mezi muži a ženami
- Přátelé a příbuzní (123 : 22)
- 48% řetězců použili stejné poslední tři články.
- Vzdálenost se zkracuje směrem k cíli.

Female	→	Female	56
Male	→	Male	58
Female	→	Male	18
Male	→	Female	13



Problém malého světa

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Závěr pro nás

1. Jedinci, kteří používají pouze lokální informace, jsou velmi efektivní ve vytvoření nejkratší cesty mezi dvěma body v sociální síti.
   → Je možné najít efektivní směrování založené pouze na lokálních informacích.
2. Propojení mezi dvěma jedinci lze najít pomocí malé posloupnosti známých.

Aplikační otázky

- Existuje vždy takový řetězec spojující dva libovolné body?
- Existuje decentralizovaný algoritmus, který ho vždy najde?
- Jaká je časová náročnost výpočtu?

- Jon Kleinberg dokázal, že takový algoritmus existuje a stanovil podmínky pro jeho nalezení i hranice výpočtu, viz [2].

Obsah přednášky

Sítě peer-to-peer (P2P)

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Čím se liší P2P sítě od klasických sítí typu klient-server

- Jiná koncepce architektury ⇒ odlišná role uzlů
- Jiný způsob adresování ⇒ adresování obsahem
- Jiný způsob směrování ⇒ lokální rozhodování, specifická struktura sítí
- Další vlastnosti ⇒ decentralizovanost, samo-organizovatelnost

Co mají P2P sítě podobné s klasickými aplikačními službami?

- Vyžadují funkční IP infrastrukturu.
- Musí řešit adresování, směrování, zabezpečení a další.

Příklady sítí P2P

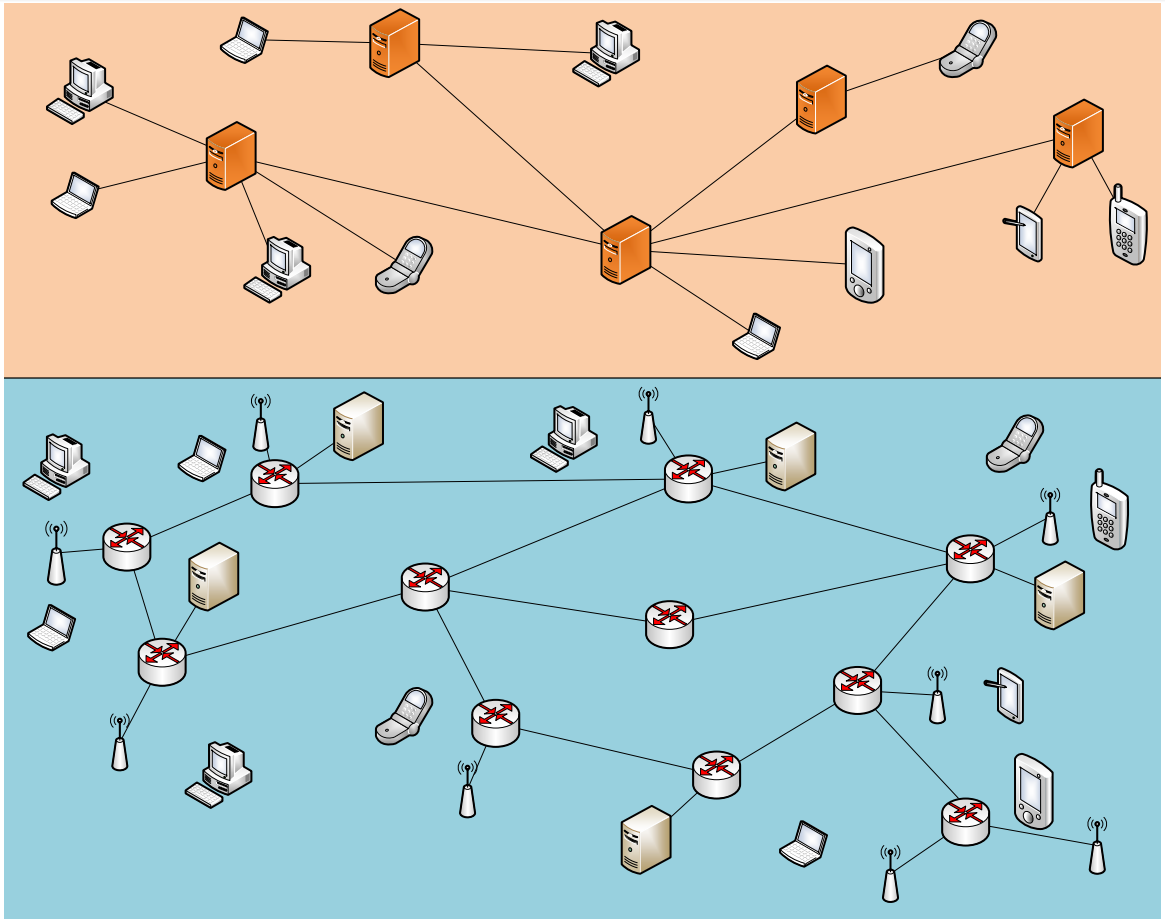
- Komunikace elektronických zařízení: Universal Plug-and-Play (UPnP), Bluetooth
- Sdílení objektů: Napster, Gnutella, KaZaA, eDonkey, BitTorrent
- Komunikace mezi uživateli: Skype, IM
- Sdílení výpočetního prostředí: seti@home, PlanetLab

Sítě peer-to-peer (P2P)

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Základem každé sítě P2P je tzv. logická síť (overlay)

- Logická síť je postavená nad existující síťovou strukturou.
- Logická síť definuje způsob propojení uzlů, směrování, vyhledávání informací, apod.



Sítě peer-to-peer (P2P)

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Definice sítě P2P

- *Dynamický soubor nezávislých uzlů (peers), které jsou propojeny a jejichž zdroje (objekty) jsou k dispozici ostatním uzlům v této síti.* [3, 4]

- Zdroje: výpočetní výkon, přenosové pásmo, disková kapacita, zařízení (tiskárny)
- Sdílené zdroje jsou přímo přístupné všem uzlům, ty je nabízejí a zároveň využívají.
- Síť obsahuje prostředky pro připojení uzlu k síti, hledávání a využití zdrojů, apod.

Typy sítí P2P

- *Pravé sítě (Pure):* odebrání libovolného uzlu ze sítě nemá vliv na ztrátu schopnosti sítě poskytovat služby.
- *Hybridní sítě:* pro svou činnost využívají centrální uzel pro poskytování části nabízených síťových služeb.
  - Centrální bod slouží k autentizaci, indexování, inicializaci uzlu, apod.

Vlastnosti sítí peer-to-peer

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Samo-organizovatelnost

- Decentralizovaná topologie, kde uzly spolupracují na vytvoření a udržování.
- Každý uzel zodpovědný za svůj lokální stav a část informací (zdrojů).
- Uzly mají částečný pohled na topologii sítě: směrování na nejbližší sousedy.
- Podobně se chovají sítě MANET (Mobile Ad hoc Networks).

Autonomní chování (samořiditelnost)

- Uzly se chovají dle svého nejlepšího rozhodování (sdílení zdrojů vs. free-riders).
- Rozhodování je lokální a nepredikovatelné $\Rightarrow$ má vliv na topologii sítě, směrování, rozmístění objektů.
- Uzly se mohou chovat zlomyslně.
- Problém s ověřování identity uzlů a důvěryhodností (decentralizované řízení).
- Možnost kolektivní zneužití zdrojů za špatným cílem.

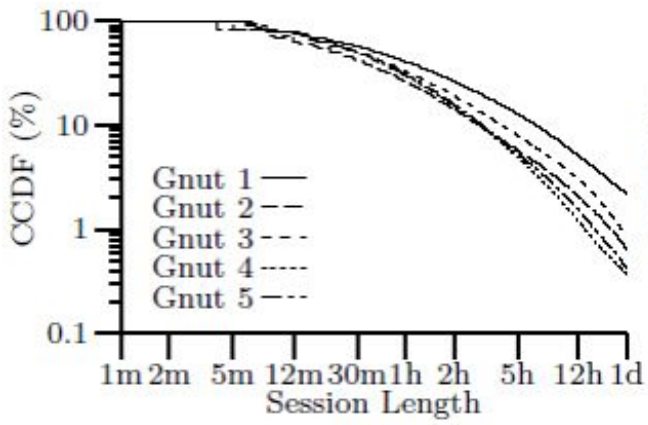
Vlastnosti sítí peer-to-peer

Spolehlivost

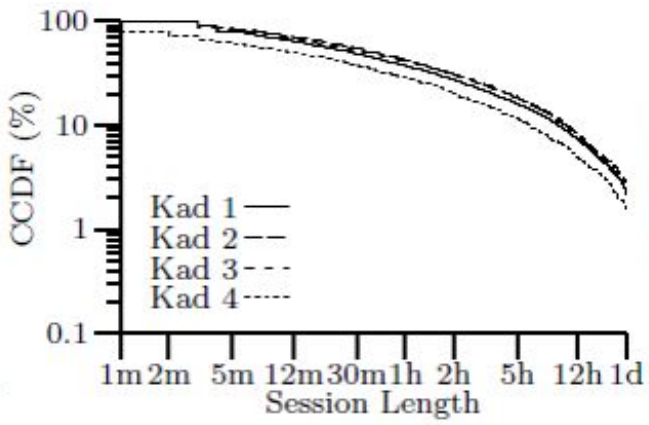
- Spolehlivost sítě roste s redundancí uzlů a informací.
- Kopie objektů jsou umístěny ve více uzlech.

Životnost uzlu (churn rate)

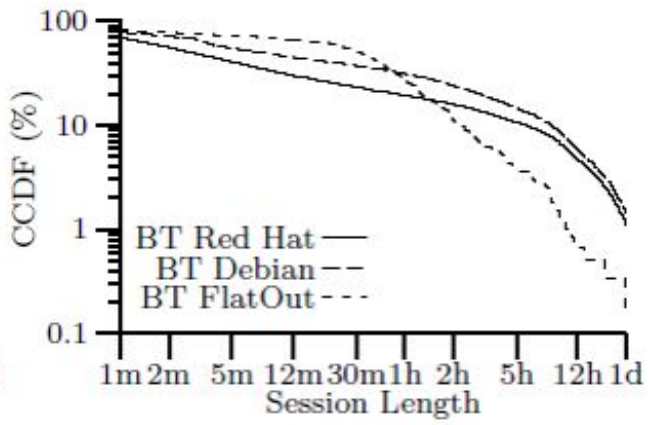
- Doba života uzlu je krátká a neodhadnutelná → problém s garancí služby.
- Závisí na subjektivním lokální rozhodnutí.
- Má vliv na směrování a vyhledávání: rychlost odlivu zákazníků (churn rate) [5].



(a) Gnutella



(b) Kad



(c) BitTorrent

Vlastnosti sítí peer-to-peer

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Srovnání s architekturou klient-server

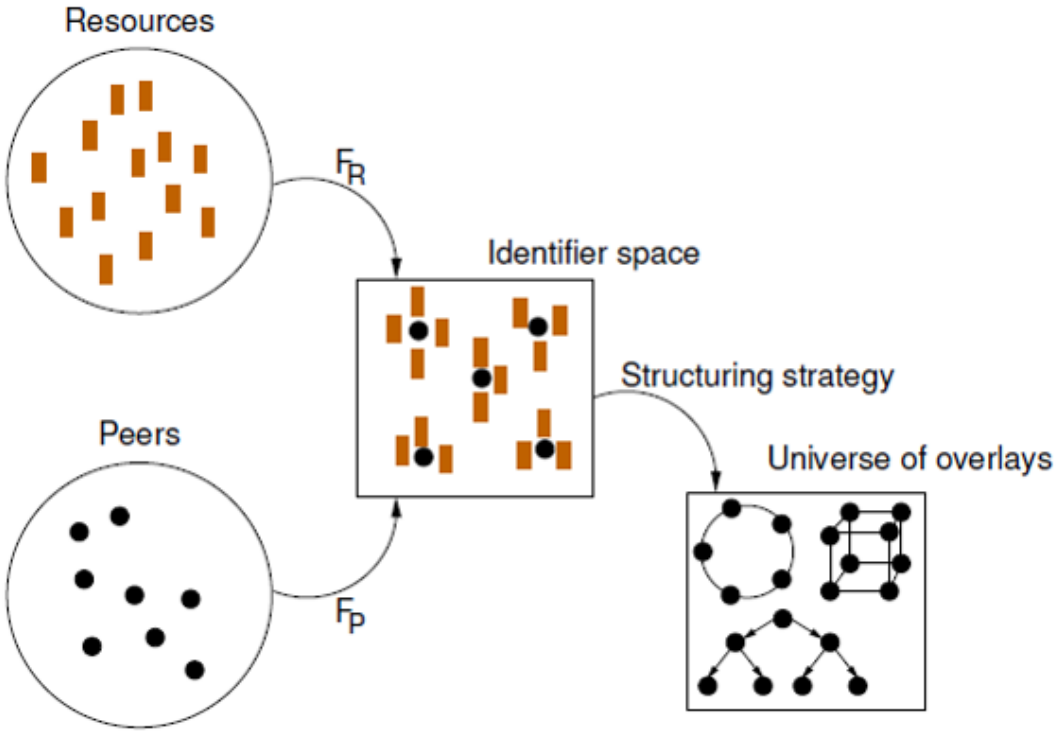
	Klient – server	Peer-to-Peer	*Výhody/nevýhody P2P*
Směr provozu	Asymetrický	Symetrický	*vs. xDSL, kabelový modem*
Topologie sítě	Stabilní	Dynamická	*Problém spolehlivosti*
Robustnost	Centrální bod	Distribuce zdrojů	*Kritický počet účastníků*
Rozšiřitelnost	Náročné	Součást návrhu	*Neomezený růst sítě*
Bezpečnost	Velký důraz	Problematické	*Chybí odpovědná autorita*
Správa a řízení	Centralizovaný model	Každý uživatel spravuje vlastní uzel	*Samo-organizovaná síť*
Poskytované zdroje	Omezené možnosti	Dynamicky rostoucí počet zdrojů	*Sdílení výpočetní prostoru, paměti, apod.*
Kvalita služeb	Garantovaná	Nelze zajistit	*Dynamicky se měnící*

Architektura sítí P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Referenční model sítě P2P [6]:

1. Jmenný prostor $\mathcal{I}$ (identifiers)
2. Množina uzlů P (peers)
3. Množina zdrojů R (resources)
4. Mapování zdrojů na identifikátory $F_R : R \mapsto \mathcal{I}$
5. Mapování uzlů na identifikátory $F_P : P \mapsto \mathcal{I}$
6. Struktura logické sítě



Množina uzlů P zpřístupňuje zdroje R v rámci jmenného prostoru $\mathcal{I}$ pomocí mapování F_R a F_P. Mapovací funkce vytváří vazbu mezi zdroji a uzly pomocí metriky blízkosti.

Referenční model sítě P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Jmenný prostor $\mathcal{I}$

- Prostor obsahuje metriku blízkosti (closeness) $d : \mathcal{I} \times \mathcal{I} \mapsto \mathcal{R}$
- Vlastnosti metriky:

$$\begin{aligned}
\forall x, y \in \mathcal{I} &: d(x, y) \geq 0 && (1)\\
\forall x \in \mathcal{I} &: d(x, x) = 0 && (2)\\
\forall x, y \in \mathcal{I} &: d(x, y) = 0 \Rightarrow x = y && (3)\\
\forall x, y \in \mathcal{I} &: d(x, y) = d(y, x) && (4)\\
\forall x, y, z \in \mathcal{I} &: d(x, z) \leq d(x, y) + d(y, z) && (5)
\end{aligned}$$

- Prostor $(\mathcal{I}, d)$ se nazývá *metrický* při splnění podmínek (1) – (5)
- Při splnění podmínek (1) – (3) jde o *pseudo-metrický prostor*.

- Jmenný prostor má vliv na adresování: každý uzel i zdroj obdrží identifikátor z $\mathcal{I}$.
- Metrika slouží pro lokalizaci a směrování požadavků.
- Metrika vede ke sdružování zdrojů a uzlů do klastrů podle blízkosti.

Referenční model sítě P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Mapování uzlů $F_P : P \Rightarrow \mathcal{I}$

- F_P přidělí uzlům jednoznačný identifikátor z $\mathcal{I}$.
- Každý uzel je zodpovědný za část jmenného prostoru (decentralizované řízení).
- Pro sousední uzly l_1, l_2, l_3 platí: $l_1 < l_2 < l_3$.
- Identifikátor může být odvozený z IP adresy či náhodný heš (např. SHA-1).

Mapování zdrojů $F_R : R \Rightarrow \mathcal{I}$

- Funkce F_R přidělí zdrojům identifikátor ze stejného jmenného prostoru $\mathcal{I}$.
- Způsob mapování je kritický pro úspěšné vyhledání zdrojů:
  - Sémantická blízkost zdrojů: podobné zdroje mají blízké identifikátory.
  - Rozložení identifikátorů má vliv na vytížení zdrojů.
- F_R se obvykle implementuje jako hešovací funkce $\rightarrow$ generuje uniformní rozložení.

Referenční model sítě P2P

Decentralizovaná správa jmenného prostoru (identifikátorů) funkcí $\mathcal{M}$ (Management)

- Funkce $\mathcal{M}$ definuje odpovědnost uzlů z P za konkrétní identifikátory, $\mathcal{M} : \mathcal{I} \mapsto 2^P$.
- Identifikátoru zdroje $i = F_R(r) \in \mathcal{I}$ je přidělena množina uzlů spravujících zdroj r:
  - Každý uzel p je zodpovědný za identifikátory $\mathcal{M}^{-1}(p)$.
  - Lokalizace zdroje r = vyhledání uzlu, který jej spravuje, tj. $\mathcal{M}(F_R(r))$.
  - Vyhledání zdroje implementuje $\mathcal{M}$ na základě výběru cesty (tj. směrování).

Vlastnosti funkce $\mathcal{M}$

- Funkce $\mathcal{M}$ může být úplná či parciální.
  - Parciální: identifikátory nemusí být vždy spojeny s nějakým uzlem.
  - Úplná (obvykle): každý identifikátor je zodpovědný za nějaký uzel.
- Stupeň replikace (kardinalita): $\mathcal{M}$ typicky obsahuje více než jeden prvek, tj. více uzlů je zodpovědných za jeden identifikátor.
- Identifikátory spojeny s nejbližšími uzly:
  $p \in \mathcal{M}(i) \Rightarrow d(F_p(p), i) = min_{q \in P} d(F_p(q), i)$.
- *Dynamické chování*: $\mathcal{M}$ se dynamicky mění tak, jak se mění uzly.

Geometrie sítě a směrování

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Struktura (geometrie) logické sítě P2P je popsána orientovaným grafem $G = (P, E)$.

- Dynamické chování reprezentováno jako posloupnost grafů: $G_i(P_i, E_i), G_{i+1}(P_{i+1}, E_{i+1}), G_{i+2}(P_{i+2}, E_{i+2}), \ldots$
- Připojení uzlu p' (operace join):
  - Přidání uzlu, tj. $P_{i+1} = P_i \cup \{p'\}$
  - Přidání hran, tj. $E_{i+1} = E_i \cup \{(p', m)\} \cup \{(n, p')\}$
- Odpojení uzlu p' (operace leave):
  - $P_{i+1} = P_i - \{p'\}$
  - $E_{i+1} = \forall m, n : E_i - \{(p', m)\} - \{(n, p')\}$

Směrovací tabulka

Každý uzel obsahuje lokální směrovací tabulku $R_p(V_p, E_P) \subseteq G$, která je součástí globální konfigurace sítě. Tabulka obsahuje množinu sousedních uzlů a hran k nim.

- Pro množinu sousedních uzlů platí: $V_p \subseteq V, \forall (p, q) \in E \implies q \in V_P \wedge \nexists q \in V_p : (p, q) \in E$.
- Pro množinu hran platí: $E_p \subseteq E, \forall (p, q) \in E \implies (p, q) \in E_p$.

Směrování v sítích P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Různé typy topologií (geometrií) sítí P2P [7]

Topologie	Stupeň uzlu	Poloměr
de Bruijin	k	$log_k N$
Trie	k+1	$2log_k N$
Chord	$log_2 N$	$log_2 N$
CAN	2d	$\frac{1}{2} d.N.\frac{1}{d}$
Pastry	$(b-1)log_b N$	$log_b N$
Classic butterfly	k	$2log_k N(1-o(1))$

N: počet uzlů sítě, k: počet sousedů uzlu

Průměr topologií pro 10^6 uzlů [7]

Stupeň k	de Bruijn	Trie	Chord	CAN	Pastry	Classic butterfly
2	20	-	-	huge	-	31
3	13	40	-	-	-	20
4	10	26	-	1,000	-	16
10	6	13	-	40	-	10
20	5	10	20	20	20	8
50	4	8	-	-	7	7
100	3	6	-	-	5	5

k: počet sousedů uzlu

Směrování v sítích P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Relace sousedství $\mathcal{N}$ (neighborhood), tj. $\mathcal{N} : P \mapsto 2^P$

- Relace $\mathcal{N}$ popisuje strukturu sítě P2P, $\mathcal{N}(p)$ je množina sousedů.
- Důležitý je poloměr grafu (sítě uzlů):
  - Určujeme maximální vzdálenost mezi dvěma libovolným uzly grafu.

Směrování = předávání zprávy `route(p,m,i)` v síti P2P

- Hledáme cestu pro zprávu m směrovanou do uzlu p, který spravuje objekt i.
- Distribuovaný proces nalezení cesty v síti P2P na základě lokálních znalostí.
- Směrovací funkce $R : P \times I \mapsto 2^P$
  - Funkce vybere v každém uzlu p z množiny sousedů $\mathcal{N}(p)$ takový uzel q, který je nejblíže uzlu obsahujícímu objekt i, tj. $R(p, i) \in \mathcal{N}(p)$
  - Pro uzel q platí: $d(i, F_p(q)) < d(i, F_p(p))$ pro $q \in R(p, i)$

Chybí globální synchronizace směrování

- Mohou vzniknout nekonzistence v lokální směrovací tabulce.

Směrování v sítích P2P

- Vzdálenost v P2P nemusí odpovídat vzdálenost fyzické.
- Efektivnější síť P2P využívá informace o fyzické blízkosti [8].

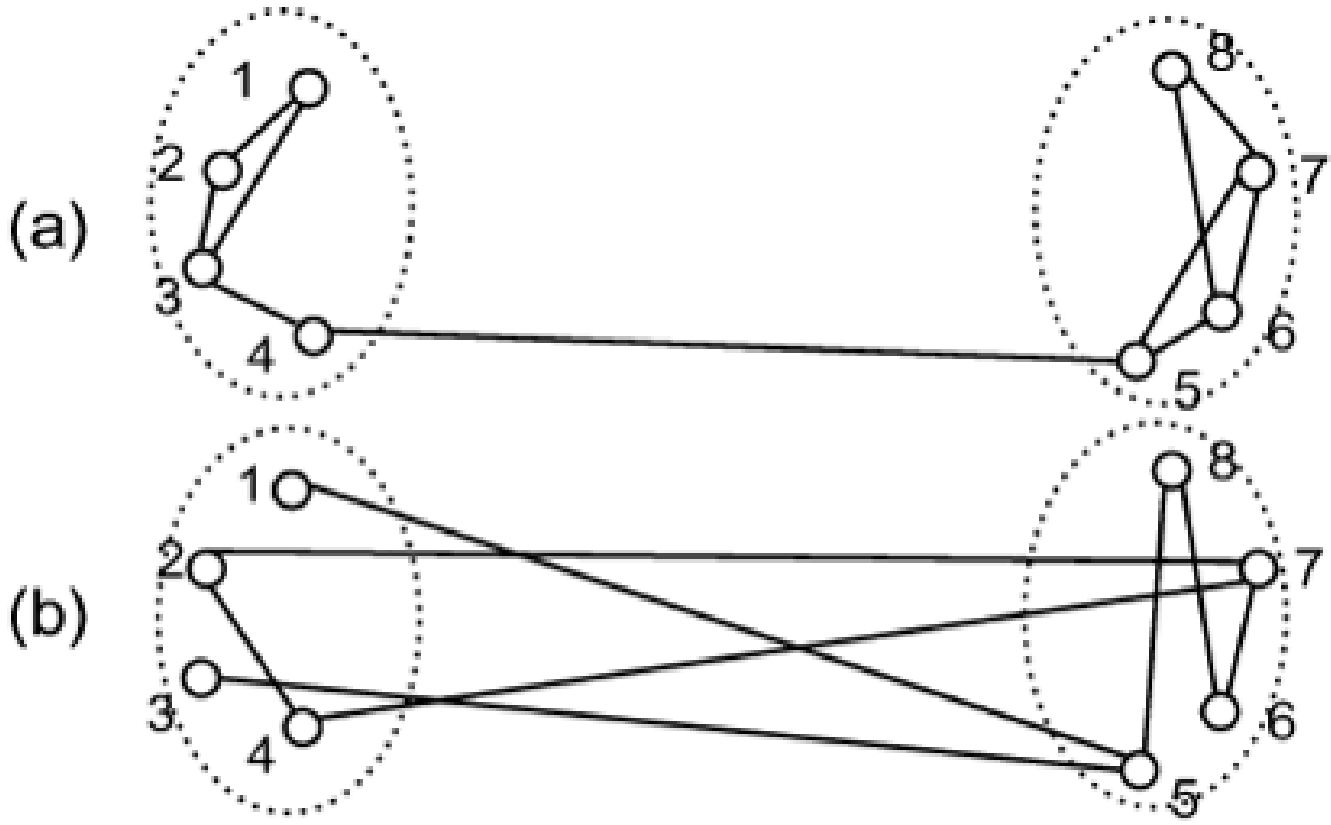


Fig. 1. (a) Illustration for locality-aware overlay and (b) randomly connected overlay.

Nestrukturované sítě P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Využívají poznatků ze sociálních sítí a malého světa

- Vyhledávání zdrojů probíhá tak, že kontaktujeme svou socální síť, která ví, kde informace je, nebo zná někoho, kdo tuto odpověď pozná.
- Watz a Strogatz zkoumali Milgramovu myšlenku "malého světa" na modelu náhodného grafu s malým průměrem [9]:
  - Využití sociálních vztahů při sdílení informací: bližší přátelé, zájmové skupiny
  - Každý uzel má své sousedy: využití tranzitivity

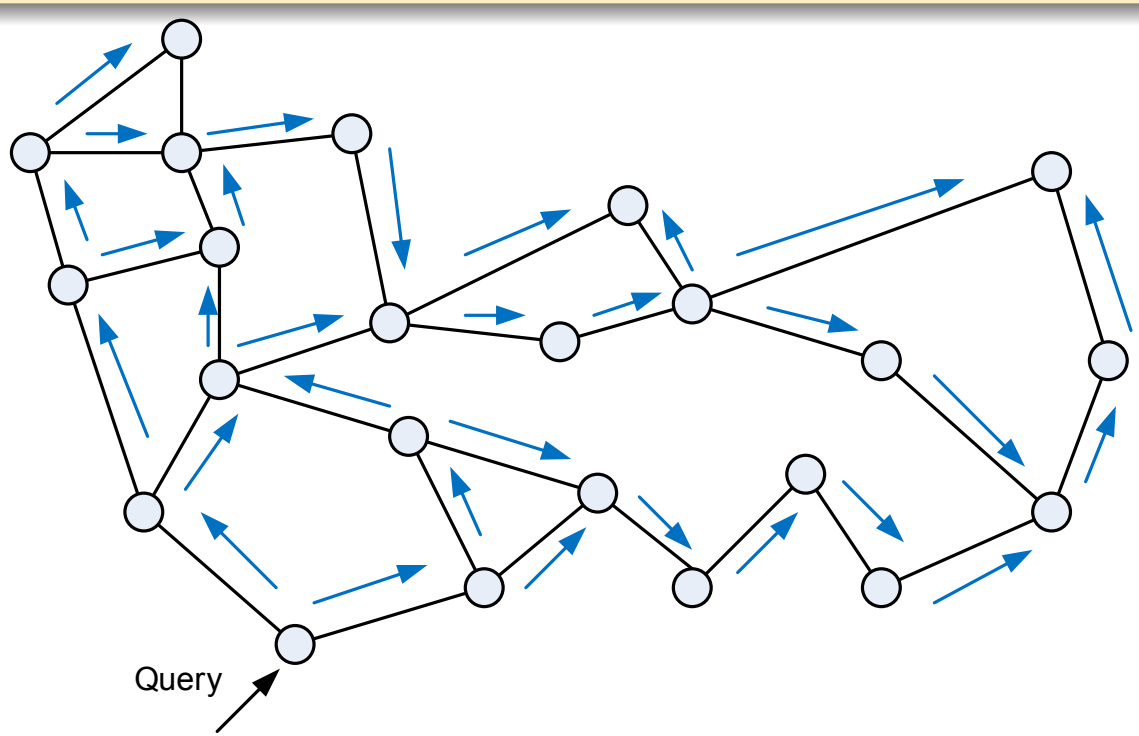
Vlastnosti

- Neexistuje struktura uložení informace: objekt je uložen v náhodně vybraném uzlu.
- Uzel si vyměňuje zprávy se svými sousedy:
  - Dotaz na vyhledání konkrétního objektu (zdroje)
  - Dotaz obsahuje například klíčové slovo, název souboru, apod.
- Nebezpečí zacyklení
  - Uzel obsahuje seznam identifikátorů zpráv, které zpracoval.
  - Využívá se hodnota TTL u zpráv.

Nestrukturované sítě P2P

1) Záplava (flooding)

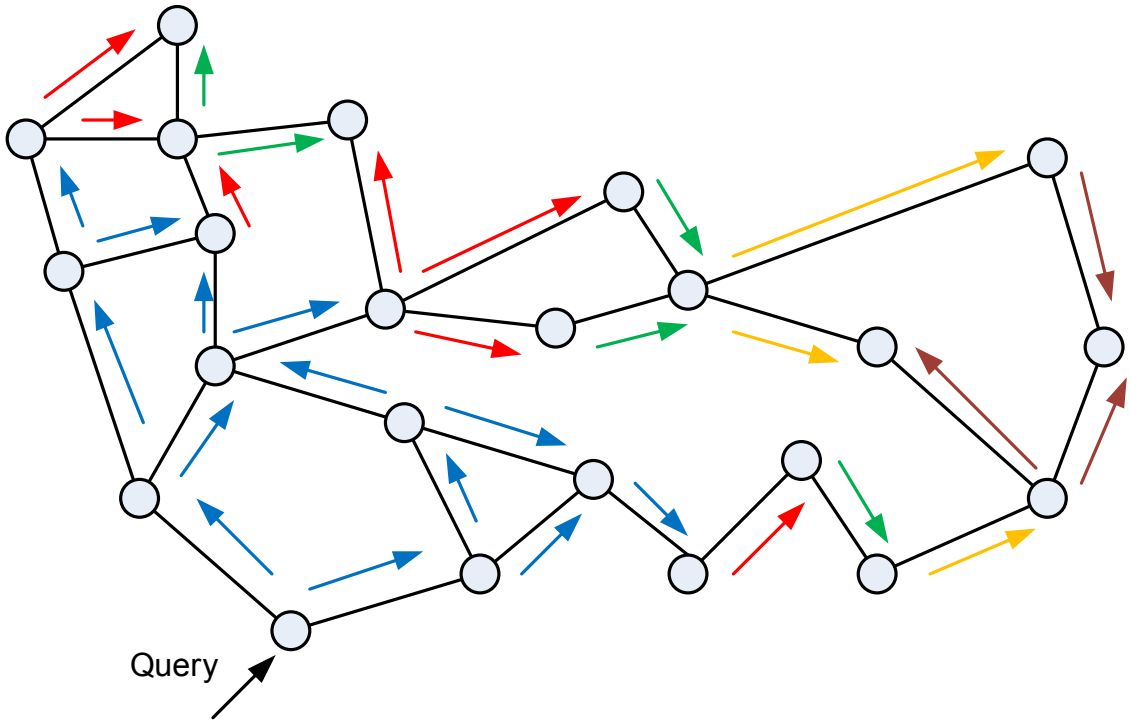
- Uzel pošle dotaz všem svým sousedům
  - Pokud soused obsahuje objekt, pošle zpět odpověď
  - Pokud nemá objekt, pošle zprávu svým sousedům (tranzitivita)
- Záplavu je možné omezit pomocí TTL ve zprávě
- Využívá např. Gnutella



Nestrukturované sítě P2P

2) Rozšiřující se kruh (expanding ring)

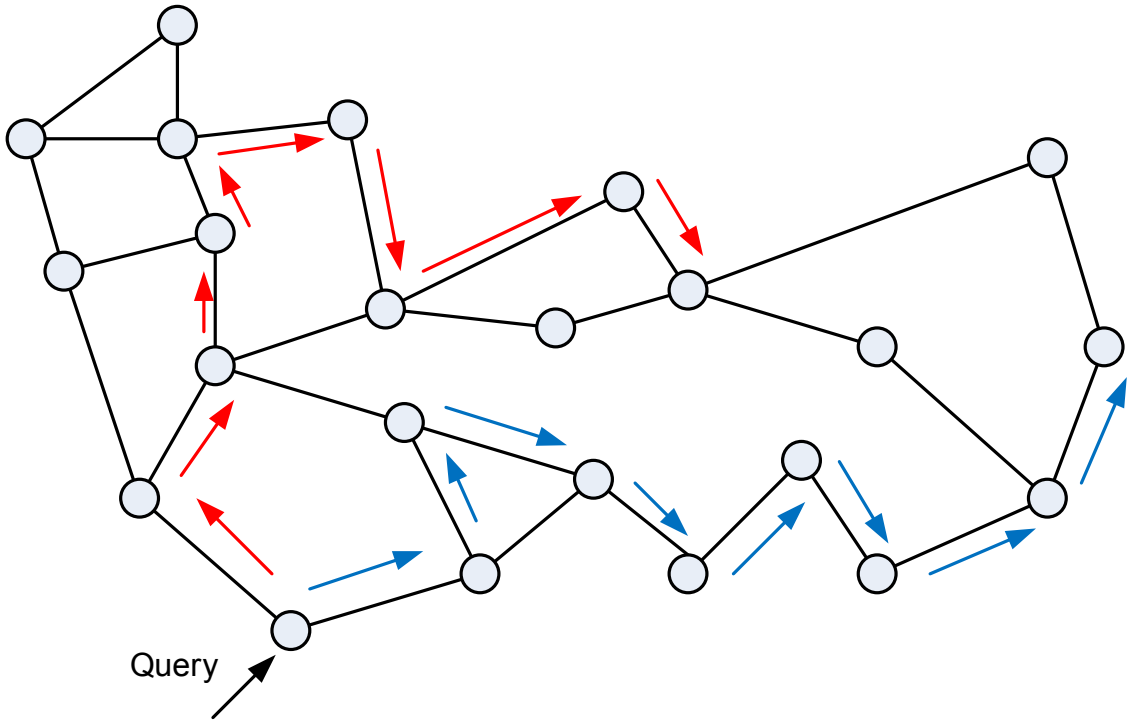
- Podobné jako záplava. Pošlu dotaz na objekt s malým TTL.
  - Pokud objekt najdu, hledání končí.
  - Pokud objekt nenajdu, zvýším TTL a pošlu dotaz znovu.
- Redukuje počet zpráv v síti



Nestrukturované sítě P2P

3) Náhodný průchod (random walk)

- Zpráva poslána náhodně vybraným sousedům
  - Možné poslat více sousedům současně
  - Potřeba kontrolovat vybraného souseda: neposílat dotaz zpět



Nestrukturované sítě P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

4) Hledání lokálního minima (Local Minimum Search, LMS) [10]

Definujme tento algoritmický problém:

- Máme množinu uzlů identifikovaných hodnotou x.
- Máme množinu objektů s identifikátorem w
  - Např. hash veřejného klíče, hash jména objektu
- Úkolem je umístit objekty do sítě uzlů tak, abychom je mohli najít rychle a spolehlivě, tj. jméno uzlu x by mělo být co nejblíže jménu ukládaného objektu w.

Princip hledání lokálního minima

- V nestrukturovaných sítích neexistuje mechanismus pro nalezení uzlu, který je nejbližší hledanému objektu.
- Místo globálního minima hledáme lokální minimum.
  - Uzly znají pouze své bezprostřední sousedy do vzdálenosti h kroků.
  - Neexistuje povědomí o globální topologii sítě.

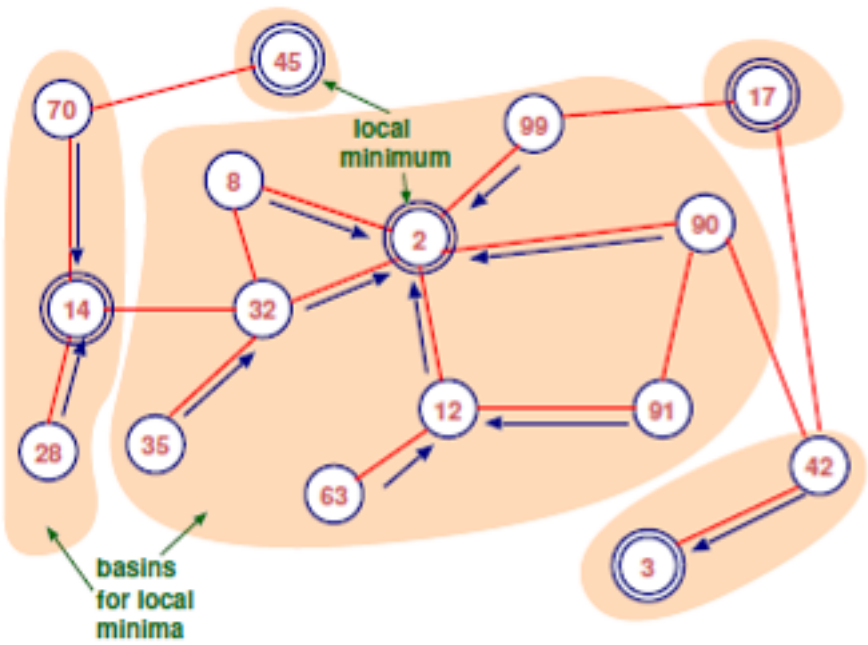
Při hledání používáme metriku vzdálenosti uzlu x od objektu w: $d(x, w)$

Nestrukturované sítě P2P

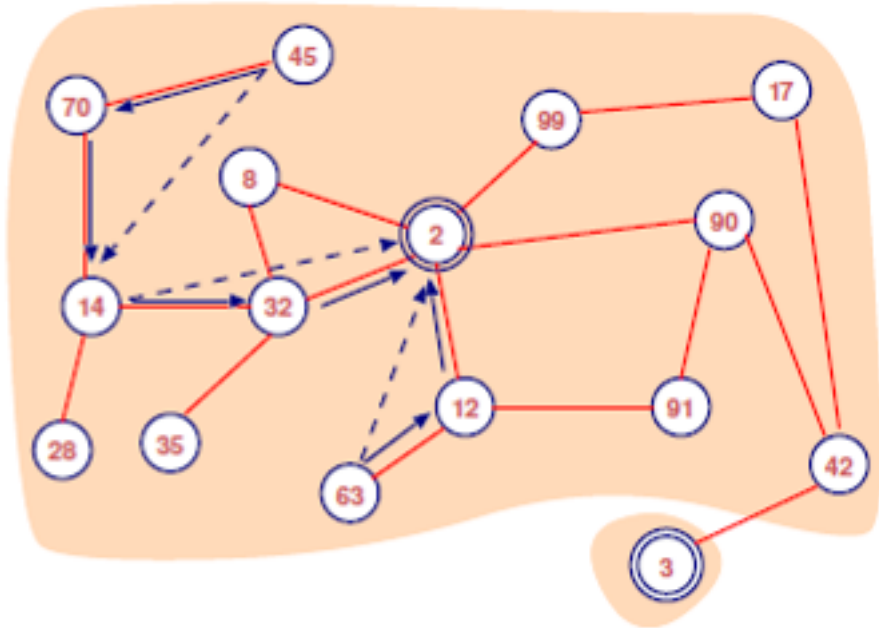
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Příklad

- Uzel u je lokálním minimem pro objekt, pokud je jeho ID nejbližší k ID objektu mezi jeho sousedy do vzdálenosti h kroků.
- Uzly jsou označeny vzdáleností $d(x, w)$.



(a) 1-hop Neighborhoods

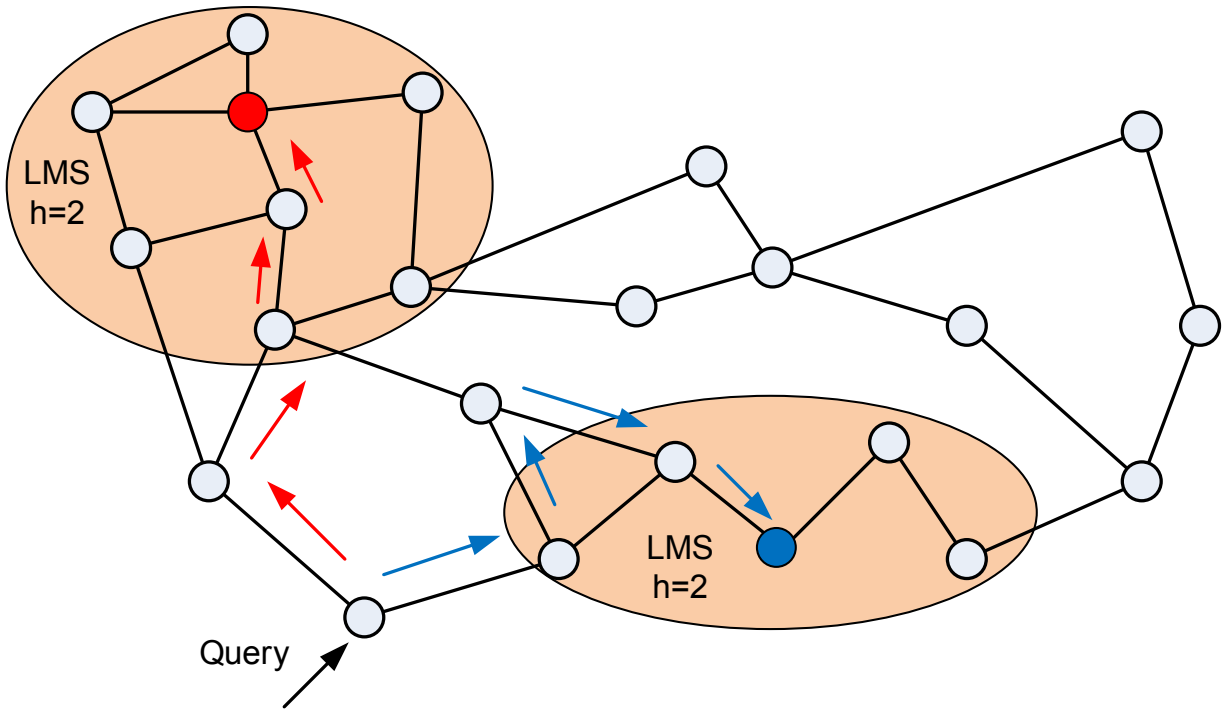


(b) 2-hop Neighborhoods

Nestrukturované sítě P2P

Algoritmus uložení objektu x

1. Uzel u vytvoří zprávu `probe(u,w,walk-length,path)` a pošle ji do sítě.
2. Síť procházíme náhodným průchodem, dokud `walk-length` > 0
3. Aktuální uzel v vypočítá $d(x, v')$ pro všechny své sousedy v'.
4. Předá zprávu uzlu s nejmenší metrikou, dokud nenajde lokální minimum.



Nestrukturované sítě P2P

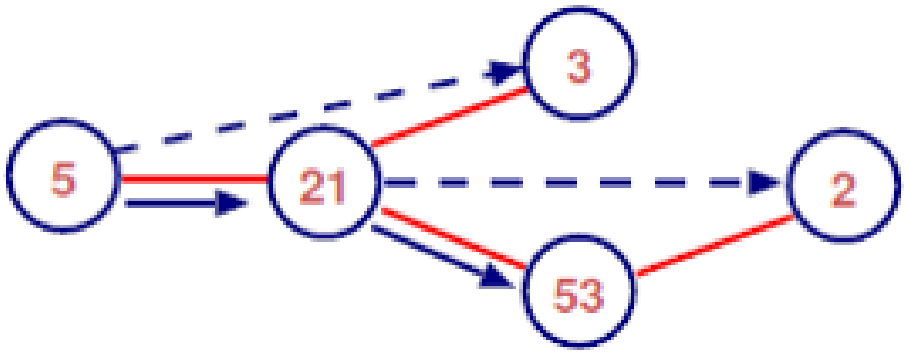
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Algoritmus vyhledání objektu x

- Probíhá jako uložení objektu: posíláme zpráva `search()`.
- Lokální minimum buď vrátí objekt nebo zprávu o chybě.
- Pokud není nalezen objekt, vysílající uzel pošle novou zprávu `search()`, která začíná náhodným průchodem.

Možnost nečekaného přesměrování dotazu

- Uzel 5 vidí uzel 3 jako svého souseda (h=2) a směruje mu zprávu.
- Uzel 21 zná lepší cíl a přepošle zprávu uzlu 53.



Nestrukturované sítě P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Porovnání směrovacích algoritmů u nestrukturovaných sítí P2P

- *Metoda záplavy a metoda rozšiřujícího se kruhu*
  - Jednoduchá implementace
  - Minimální paměťové i výpočetní nároky
  - Neefektivní, špatně rozšiřitelná
- *Metoda náhodného průchodu*
  - Hledání bez znalosti umístění objektu
  - Nalezení objektu může dlouho trvat
- *Metoda lokálního minima*
  - Přidává znalost: metriku vzdálenosti k objektu
  - Směrování podle jména objektu
  - Vyžaduje režii při hledání lokálních minim

Zhodnocení nestrukturovaných sítí:

- Neefektivní směrování, špatná lokalizace řídce se vyskytujících objektů.
- Směrování podle identifikátoru objektu (klíče).
- Příklady sítí: Napster, Gnutella, FastTrack, FreeNet, Gia, Tribler, INGA

Strukturované sítě P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Strukturované sítě

- Kombinují geometrické struktury a směrování.
- Využívají distribuované směrovací algoritmy:
  - Metriky: shoda prefixu, euklidovská či lineární vzdálenost, XOR, atd.
  - Velikost směrovací tabulku ovlivněna stupněm uzlů.

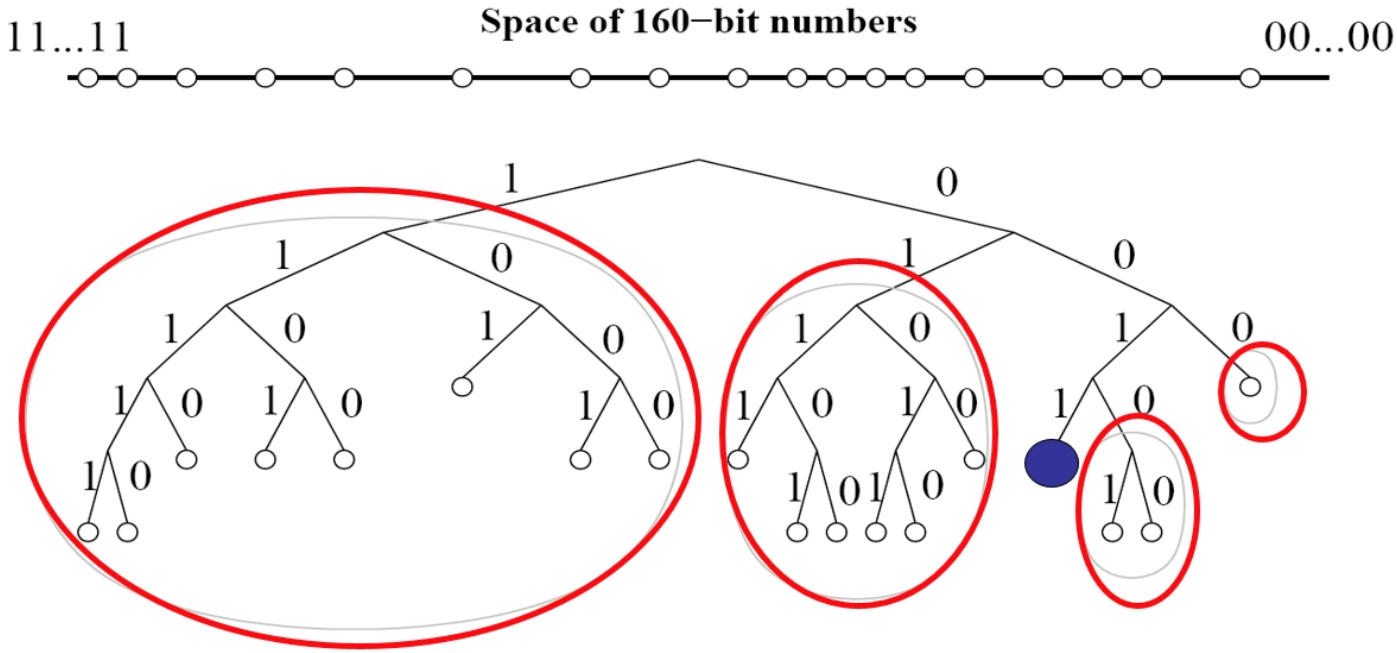
Příklad strukturované sítě Kademlia [11]

- Decentralizovaná, rozšiřitelná, samo-organizující se síť:
  - Každý uzel obsahuje informaci o dalších uzlech a souborech.
  - Identifikátory uzlů i souborů tvořeny jako 160bitový heš SHA-1.
  - Metriku blízkosti tvoří bitový XOR: $d(a, b) = a. \oplus b$
- Pro směrování používá distribuovanou hešovací tabulku (DHT).
- Složitost vyhledání je $\mathcal{O}(logN)$, kde N je počet uzlů sítě.
- Složitost připojení či odpojení uzlů je $\mathcal{O}(log^2 N)$.
- Používá se pro sdílení souborů v síti BitTorrent.

Struktura systému Kademlia

Jmenný prostor uzlů ⇒ binární strom

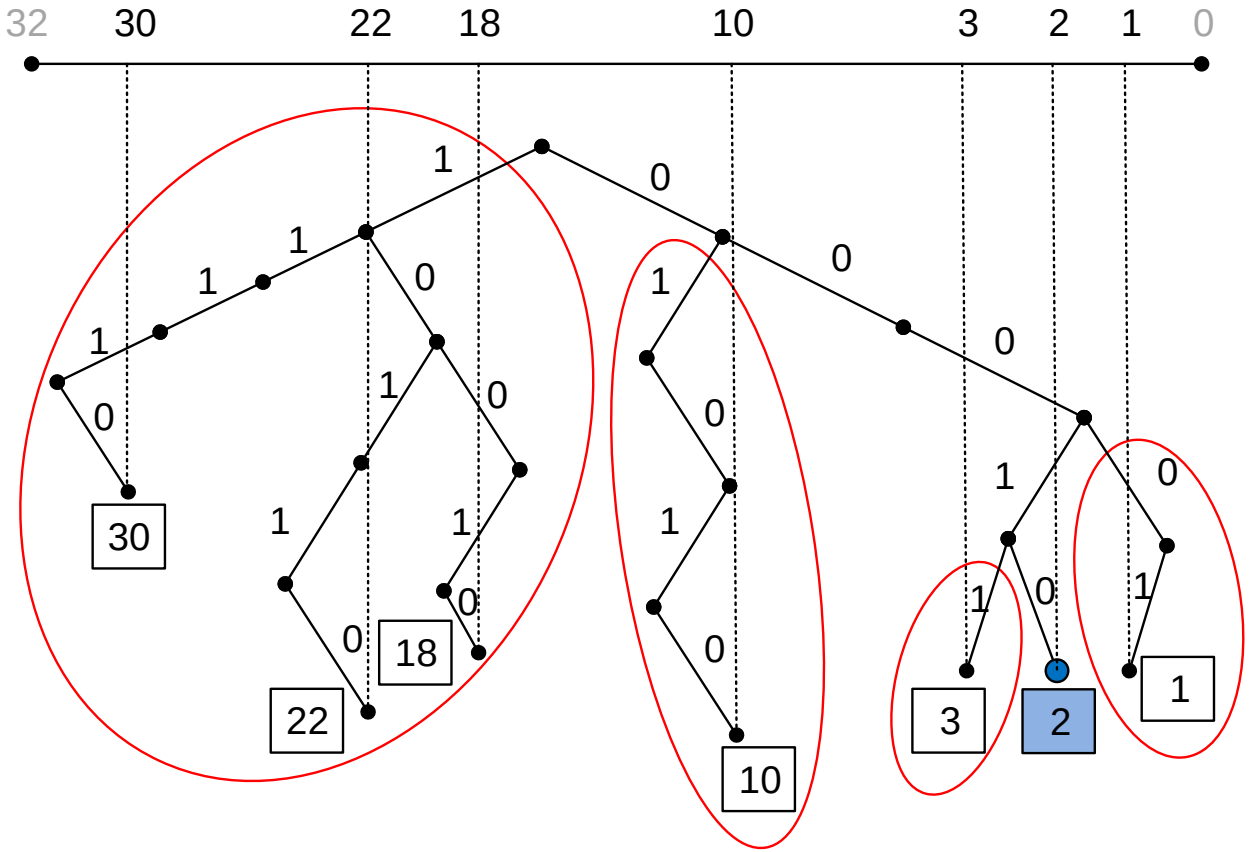
- Binární strom obsahuje známé sousedy (peers) daného uzlu.
- Strom je rozdělen na podstromy, které reprezentují celý stavový prostor.
- 1 až k uzlů z každého podstromu je uloženo do seznamu *k-bucket* v daném uzlu.
- Každý uzel v síti má směrovací tabulku.



Subtrees of interest for a node 0011.....

Kademlia: příklad směrovací tabulky

- Jmenný prostor o velikosti $N = 32$, k=3, délka ID je 5 bitů.
- Uzly ve stejném podstromu mají stejný prefix i stejnou vzdálenost od daného uzlu.
- Např. $d(2,3) = 2 \oplus 3 = 1, d(10 \oplus 2) = 8, d(22 \oplus 2) = 20, d(30 \oplus 2) = 28$, apod.



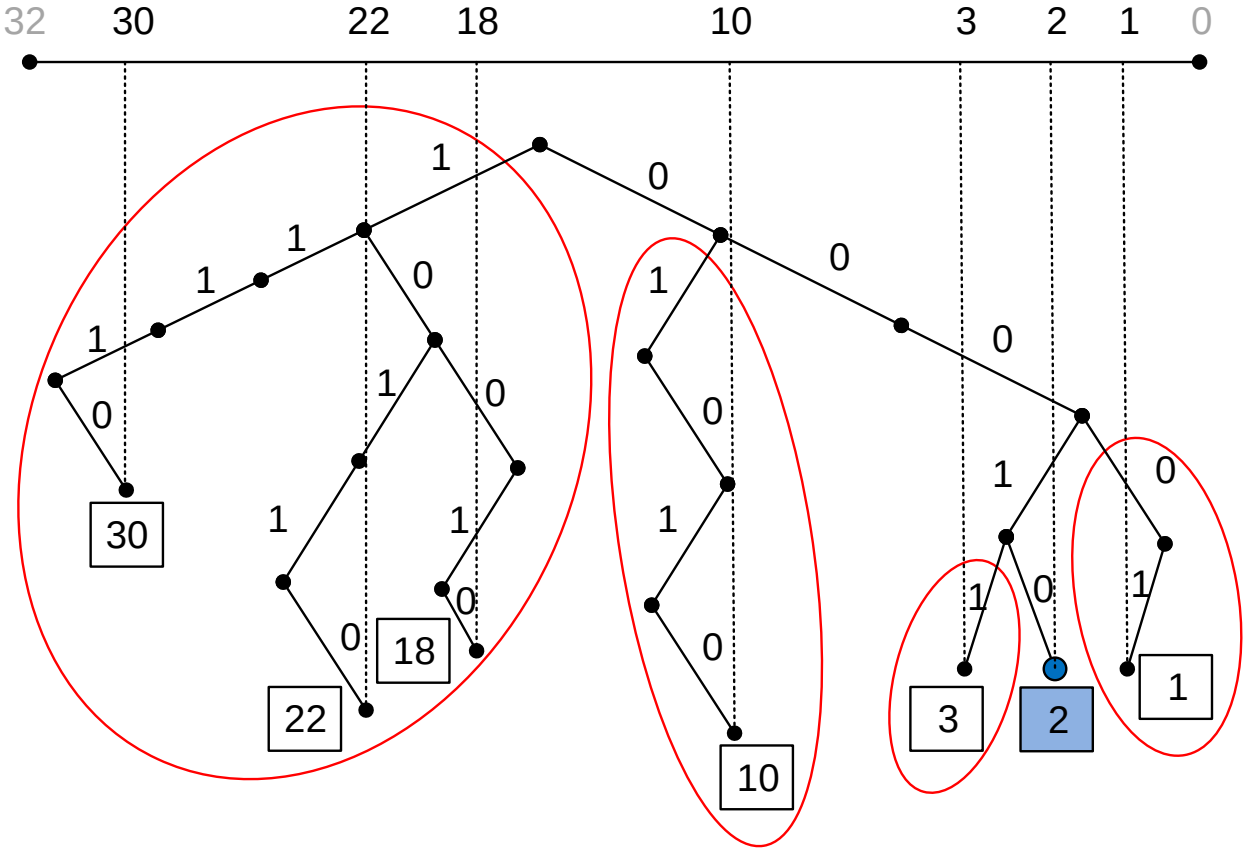
Node ID=2		
i	**Interval**	**Peers**
1	$[1,2)$	3
2	$[2,4)$	1
3	$[4,8)$	
4	$[8,16)$	10
5	$[16,32)$	18, 22, 30

$k = 3$ je stupeň replikace

- Řádek i obsahuje uzly podstromu ve dané vzdálenosti $[2^i, 2^{i+1})$ od daného uzlu.

Kademlia: vytváření směrovací tabulky

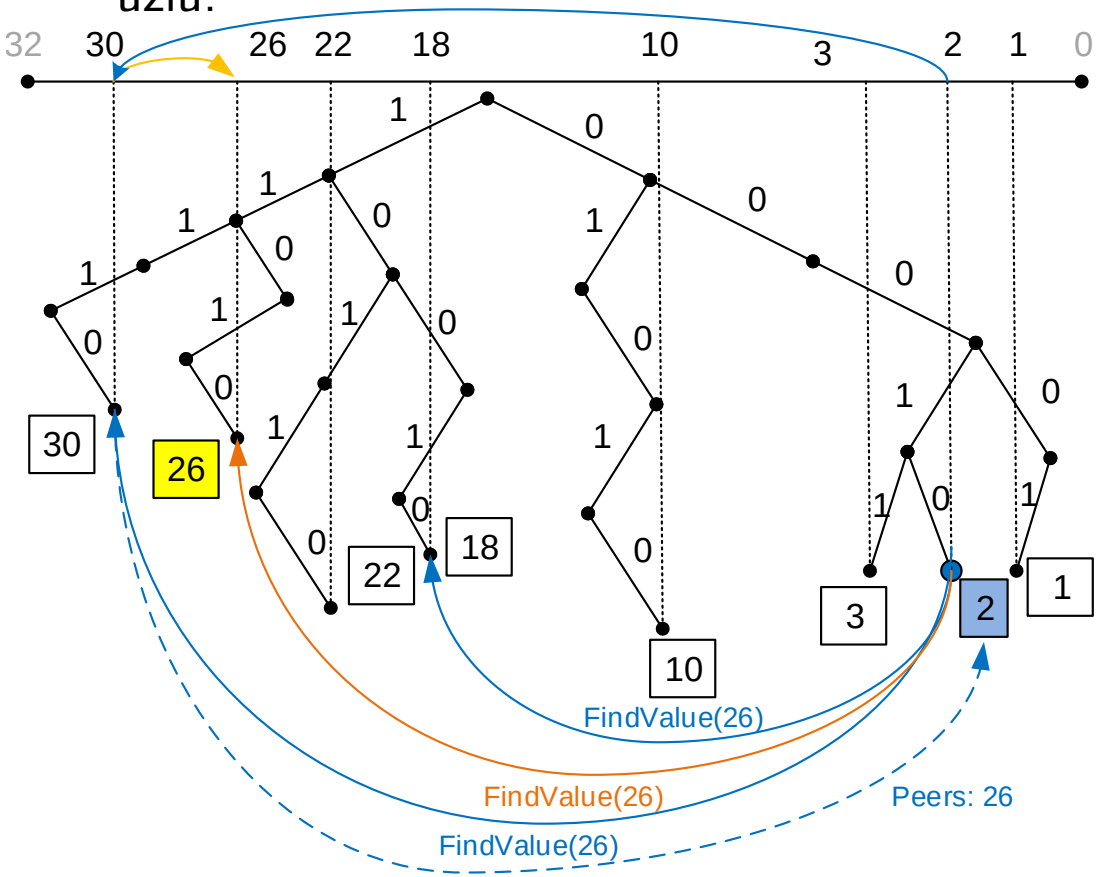
- Směrovací tabulka se aktualizuje dynamicky při příchodu zprávu od jiného uzlu.
- Pokud daný řádek obsahuje méně než k položek, přidá se nový uzel.
- Pokud je řádek plný, otestuje se dostupnost nejpozději přidaného uzlu. Pokud je nedostupný, nahradí se novým. Pokud je dostupný, nový uzel se do tabulky nepřidá. $\Rightarrow$ preferují se starší kontakty.



Node ID=2		
i	Interval	Peers
1	$[1, 2)$	3
2	$[2, 4)$	1
3	$[4, 8)$	
4	$[8, 16)$	10
5	$[16, 32)$	18, 22, 30

Komunikace v síti Kademlia

- `PING(nodeID)`: kontroluje, zda daný uzel je připojen.
- `STORE(fileID,nodeID)`: uloží do daného uzlu `fileID` a `nodeID`, který ho nabízí.
- `FIND_NODE(nodeID)`: daný uzel vrátí nejbližší uzly k uzlu `nodeID`.
- `FIND_VALUE(fileID)`: vrátí adresu uzlu obsahující soubor nebo seznam nejbližších uzlů.



Činnost uzlu ID=2 při hledání obsahu s fileID=26, $\alpha = 2$.

1. Uzel 2 vyhledá *k-bucket*: $d(2, 26) = 2 \oplus 26 = 24$.
2. Pošle dotaz na dva nejbližší uzly z $\{18,22,30\}$.
3. Každý oslovený uzel prohledá svou tabulku.
4. Uzel 30 vrátí odkaz na nejbližší uzel s nodeID=26.
5. Uzel 2 vyhledá soubor s fileID=26.

Síť BitTorrent

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Vlastnosti sítě BitTorrent

- P2P síť pro sdílení souborů.
- Síť využívá 160bitové identifikátory pro identifikaci uzlů (`peer ID`) i pro identifikaci sdílených souborů (`info_hash`).
- Více implementací: Mainline DHT (MLDHT), KAD, VUZE a další.
- Pro distribuci obsahu se využívá komunikace pomocí `trackeru` nebo distribuovaná hešovací tabulka DHT.

Síť BitTorrent

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Komunikace pomocí `trackerů`.

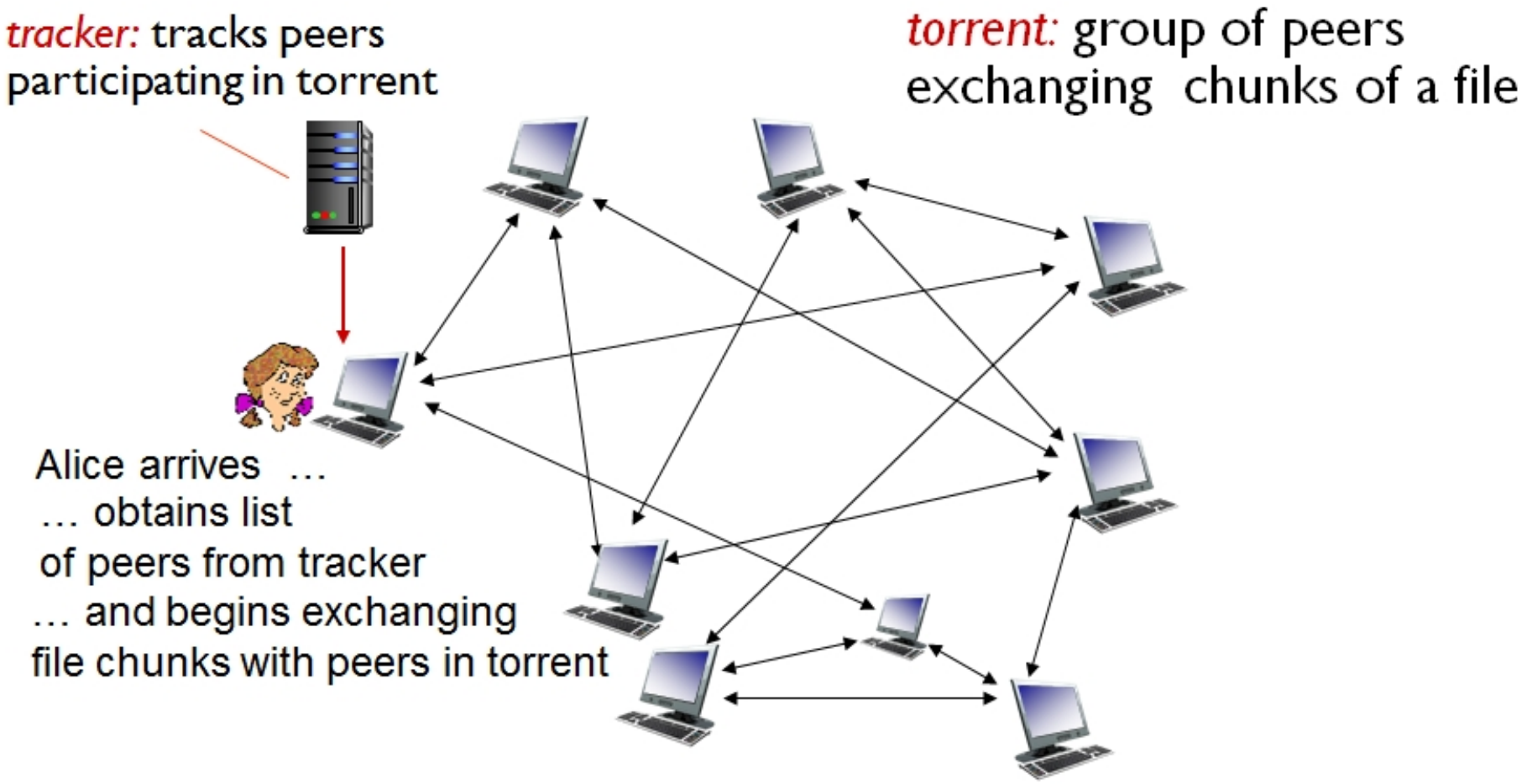
- `Swarm` [roj] = množina uzlů zapojených do sdílení souboru.
- `Tracker` [stopař] = uzel, který udržuje seznam uzlů zapojených do distribuce daného souboru.
- `Torrent` [příval] = soubor s informacemi o sdílených souborech (či adresářích) a s odkazem na `tracker`.
- Protokol BitTorrent = aplikační protokol pro distribuci souborů (nad TCP).

`Příklad souboru torrent pro OS Debian`

```
d8:announce41:http://bttracker.debian.org:6969/announce7:comment35:"Debian
CD from cdimage.debian.org"13:creationdatei1520682854e9:httpseedsl142:
https://cdimage.debian.org/cdimage/release/9.4.0//srv/cdbuilder.debian.org
/dst/deb-cd/weekly-builds/amd64/iso-dvd/debian-9.4.0-amd64-DVD-1.iso142:
https://cdimage.debian.org/cdimage/archive/9.4.0//srv/cdbuilder.debian.org
/dst/deb-cd/weekly-builds/amd64/iso-dvd/debian-9.4.0-amd64-DVD-1.isoe4:
infod6:lengthi3977379840e4:name28:debian-9.4.0-amd64-DVD-1.iso12:
piecelengthi1048576e6:pieces75880
```

Síť BitTorrent

Příklad komunikace BitTorrent [12]



Síť BitTorrent

Komunikace pomocí DHT [BEP 5]

- Decentralizovaný systém komunikace postavený na principu Kadmelia.
- Využívá aplikační protokol DHT nad UDP.
- Každý uzel obsahuje směrovací tabulku jmenného prostoru $(0..2^{160})$.
  - Tabulka rozdělena na řádky (`k-buckets`), kde $k = 8$.
  - Dostupnost uzlů se testuje každých 15 minut.
- Síť DHT tvoří uzly (nodes), které obsahují adresu peerů.
- Příkazy DHT: `ping`, `find_node`, `get_peers`, `announce_peer`

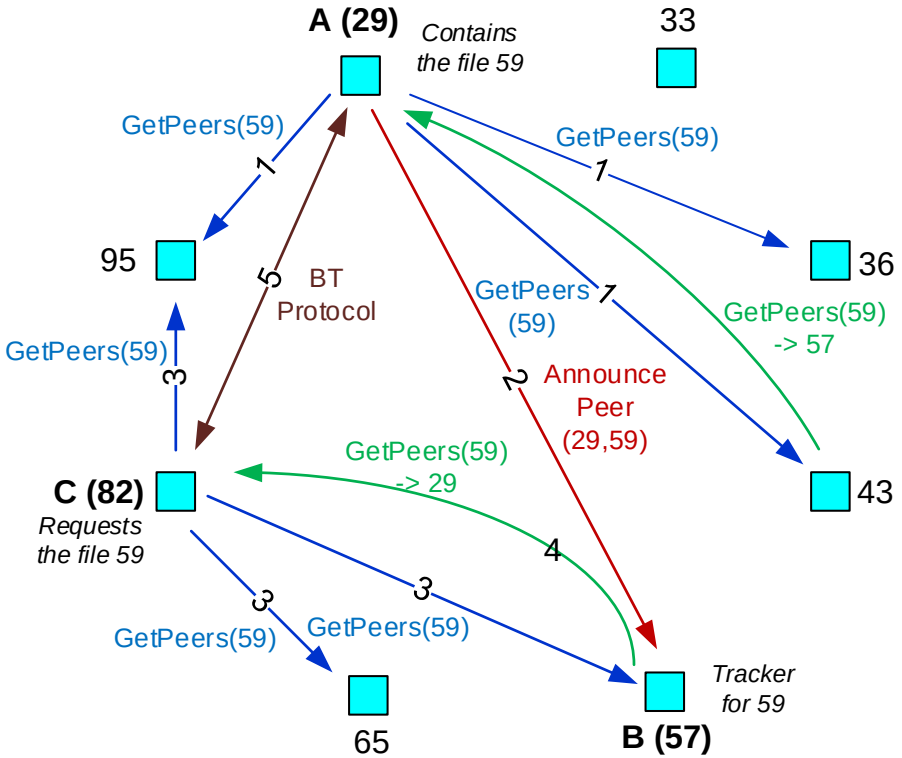
Příklad komunikace

```
get_peers Query = {"t":"aa", "y":"q", "q":"get_peers", "a":
{"id":"abcdefghij0123456789", "info_hash":"mnopqrstuvwxyz123456"}}
bencoded=d1:ad2:id20:abcdefghij01234567899:info_hash20:mnopqrstuvwxyz
123456e1:q9:get_peers1:t2:aa1:y1:qe
announce_peers Query = {"t":"aa", "y":"q","q":"announce_peer","a":{"id":"abcde34
56789", "implied_port": 1, "info_hash":"mnopqrstuvwxyz123456", "port": 6881}}
bencoded = d1:ad2:id20:abcdefghij012345678912:implied_porti1e9:info_hash20:mnop
uvwxyz1234564:porti6881e5:token8:aoeusnthe1:q13:announce_peer1:t2:aa1:y1:qe
```

Síť BitTorrent

Komunikace pomocí modifikované DHT

- `PING`: testování dostupnosti uzlu. Nedostupný uzel je vyřazen ze směrovací tabulky.
- `FIND_NODE`: hledání nejbližších k-uzlů na základě metriky blízkosti.
- `GET_PEERS`: Pro zadaný `info_hash` hledá množinu nejbližších uzlů.
- `ANNOUNCE_PEER`: uzel oznamuje, že patří do množiny `swarm` pro distribuci souboru.



Příklad [13]: Uzel **A** obsahuje soubor x=59. Uzel **C** hledá tento soubor.

1. **A** vyhledá pomocí `GET_PEERS` nejbližší uzel k x. Je to uzel **B**.
2. **A** oznámí **B** pomocí `ANOUNCE_PEER`, že vlastní soubor x. **B** si to uloží.
3. **C** hledá uzly nejbližší k x.
4. **B** pošle **C** seznam uzlů (`swarm`) pro x.
5. **C** se připojí do `swarm` a stáhne soubor x z **B** pomocí protokolu BitTorrent.

Strukturované sítě P2P

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Zhodnocení strukturovaných sítí P2P

- Klíčovým parametrem je geometrie (struktura) sítě.
- Potřeba implementovat operace pro připojení, odpojení, vyhledání.
- Směrovací algoritmus musí konvergovat k cíli.
- Nutné sledovat stav sousedů a pravidelně aktualizovat směrovací tabulku, případně tabulky sousedů.

Příklady strukturovaných sítí

- Kadmelia
- BitTorrent
- Skype

Otázky k opakování

- Vysvětlete experiment Stanleyho Milgrama a jeho vliv pro návrh sítí P2P?
- Definujte sítě P2P. Popište jejich vlastnosti, výhody a nevýhody při použití.
- Popište rozdíly architektur sítí P2P a klient-server.
- Popište referenční model sítí P2P.
- Čím se liší strukturované a nestrukturované sítě P2P.
- Jaké znáte směrovací algoritmy u nestrukturovaných sítí P2P?
- Popište metodu hledání lokálního minima.
- Popište směrování v síti Kadmelia.
- Popište chování sítě BitTorrent za pomoci trackerů a pomocí DHT.

Použitá literatura I

[1] Stanley Milgram.
The Small World Problem.
Psychology Today, 67(1):61–67, 1967.

[2] Jon Kleinberg.
The Small-world Phenomenon: An Algorithmic Perspective.
In *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*, STOC '00, pages 163–170, New York, NY, USA, 2000. ACM.

[3] John Buford, Heather Yu, and Eng Keong Lua.
P2P Networking and Applications.
Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008.

[4] R. Schollmeier.
A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications.
In *Proceedings of the First International Conference on Peer-to-Peer Computing*, P2P '01, pages 101–, Washington, DC, USA, 2001. IEEE Computer Society.

[5] Daniel Stutzbach and Reza Rejaie.
Understanding Churn in Peer-to-peer Networks.
In *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*, IMC '06, pages 189–202, New York, NY, USA, 2006. ACM.

[6] Karl Aberer, Luc Onana Alima, Ali Ghodsi, Sarunas Girdzijauskas, Seif Haridi, and Manfred Hauswirth.
The essence of P2P: A reference architecture for overlay networks.
In *Proc. of the Fifth IEEE International Conference on Peer-to-Peer Computing*, 2005.
Konstanz, Germany.

[7] Dmitri Loguinov, Anuj Kumar, Vivek Rai, and Sai Ganesh.
Graph-theoretic analysis of structured peer-to-peer systems: routing distances and fault resilience.
In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 395–406. ACM, 2003.

Použitá literatura II

[8] Xin Yan Zhang, Qian Zhang, Zhensheng Zhang, Gang Song, and Wenwu Zhu.
A construction of locality-aware overlay network: mOverlay and its performance.
IEEE Journal on Selected Areas in Communications, 22(1):18–28, Jan 2004.

[9] Duncan J. Watts and Steven H. Strogatz.
Collective dynamics of "small-world" networks.
Nature, 393(6684):440–442, June 1998.

[10] Ruggero Morselli, Bobby Bhattacharjee, Aravind Srinivasan, and Michael A. Marsh.
Efficient lookup on unstructured topologies.
In *PODC '05: Proceedings of the twenty-fourth annual ACM symposium on Principles of distributed computing*, pages 77–86, New York, NY, USA, 2005. ACM Press.

[11] Petar Maymounkov and David Mazières.
Kademlia: A Peer-to-Peer Information System Based on the XOR Metric.
In *Revised Papers from the First International Workshop on Peer-to-Peer Systems*, IPTPS '01, pages 53–65, London, UK, UK, 2002. Springer-Verlag.

[12] James F. Kurose and Keith W. Ross.
Computer Networking: A Top-Down Approach Featuring the Internet.
Addison-Wesley, 6th edition, 2012.

[13] L. Wang and J. Kangasharju.
Measuring large-scale distributed systems: case of BitTorrent Mainline DHT.
In *IEEE P2P 2013 Proceedings*, pages 1–10, Sept 2013.