Stochastic Analysis and File Availability Enhancement for BT-like File Sharing Systems

John.C.S.Lui

The Chinese University of Hong Kong

Outline

Outline

- **Introduction**

Outline

- Introduction
- Modeling the Performance

Outline

- Introduction
- Modeling the Performance
- Extension on heterogeneous Network

Outline

- Introduction
- Modeling the Performance
- Extension on heterogeneous Network
- Modeling the Availability

Outline

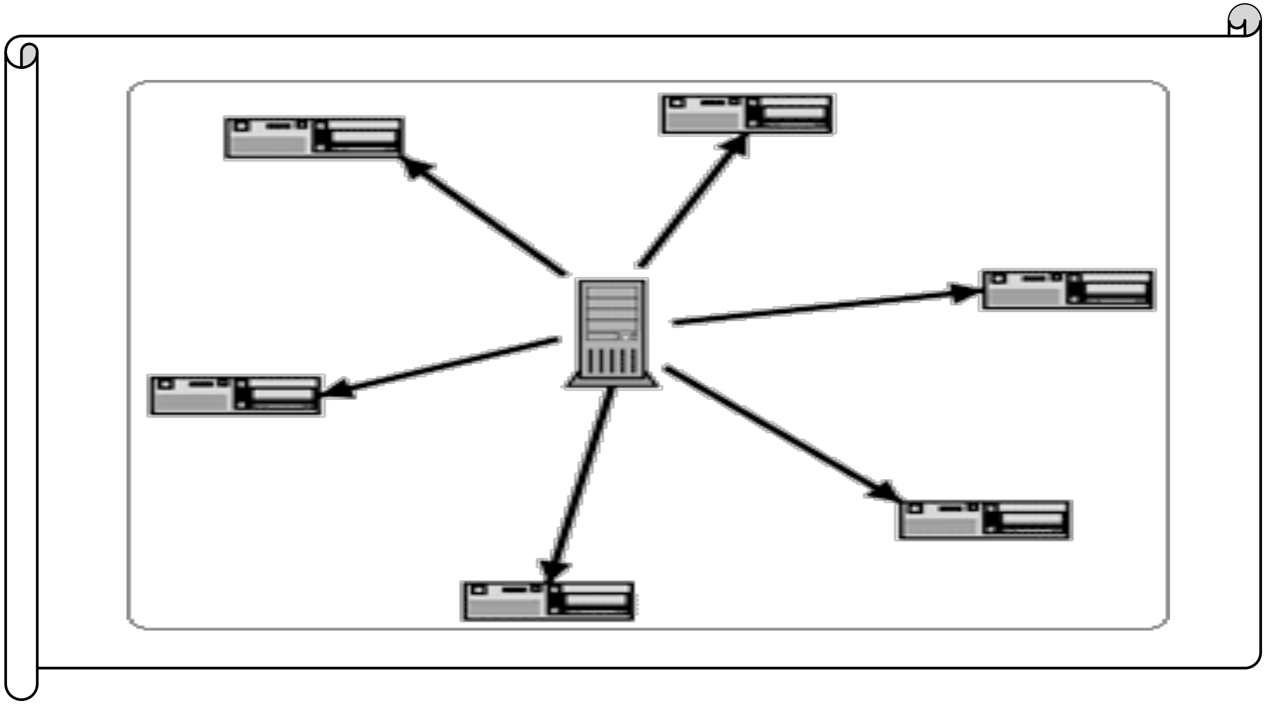
- Introduction
- Modeling the Performance
- Extension on heterogeneous Network
- Modeling the Availability
- Availability Improvement Algorithms

Outline

- Introduction
- Modeling the Performance
- Extension on heterogeneous Network
- Modeling the Availability
- Availability Improvement Algorithms
- Conclusion

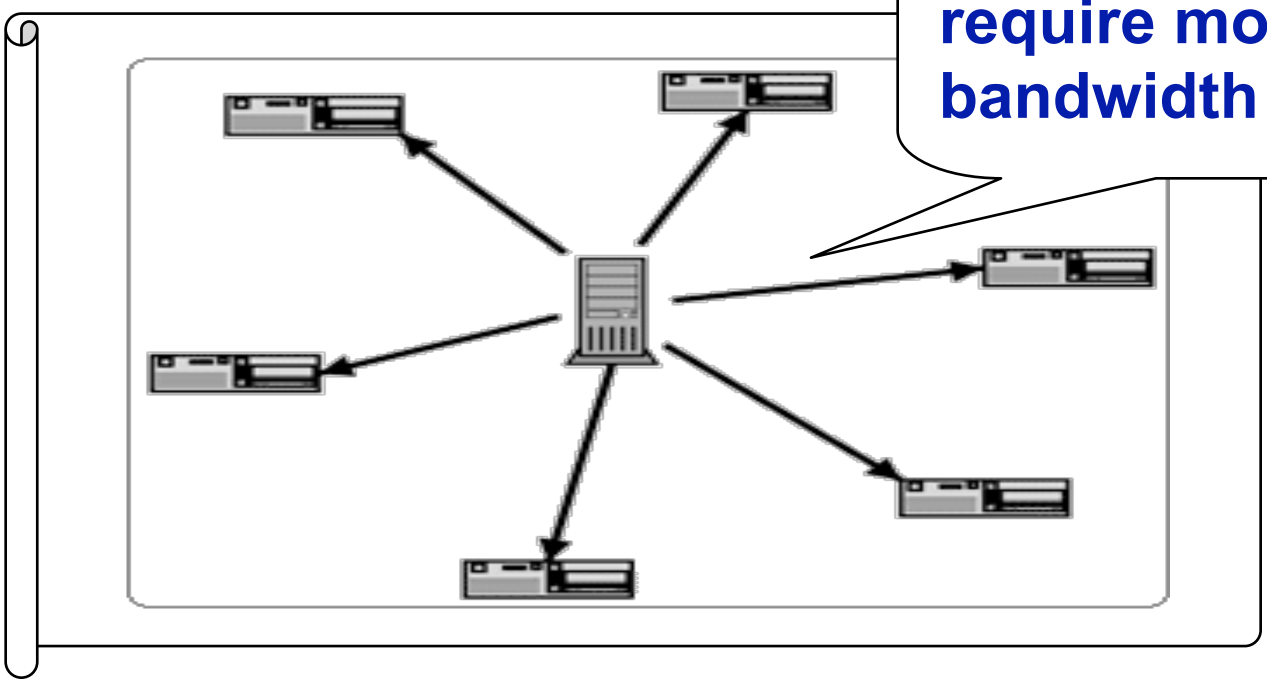
The Problem with C/S Publishing

The Problem with C/S Publishing



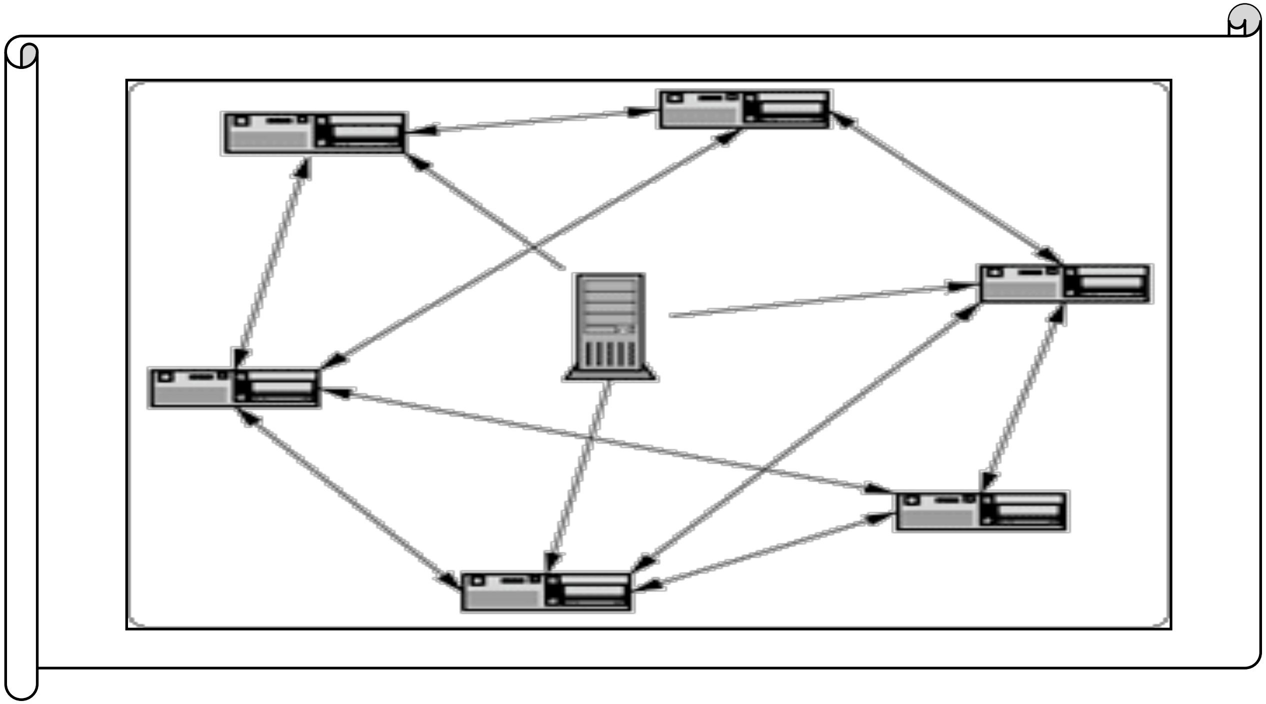
The Problem with C/S Publishing

More customers
require more
bandwidth



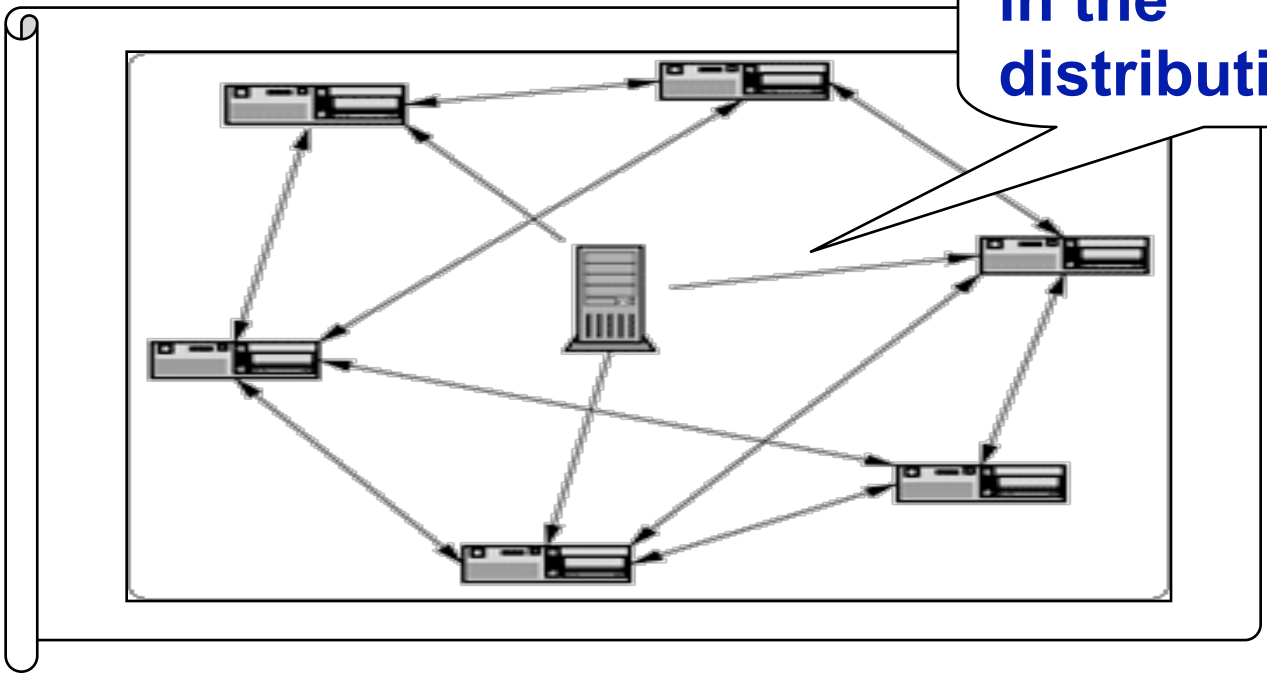
BitTorrent Solution

BitTorrent Solution



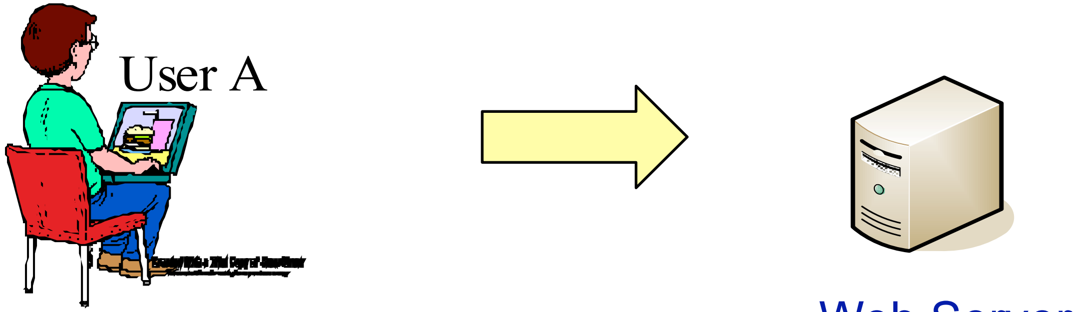
BitTorrent Solution

Users cooperate in the distribution



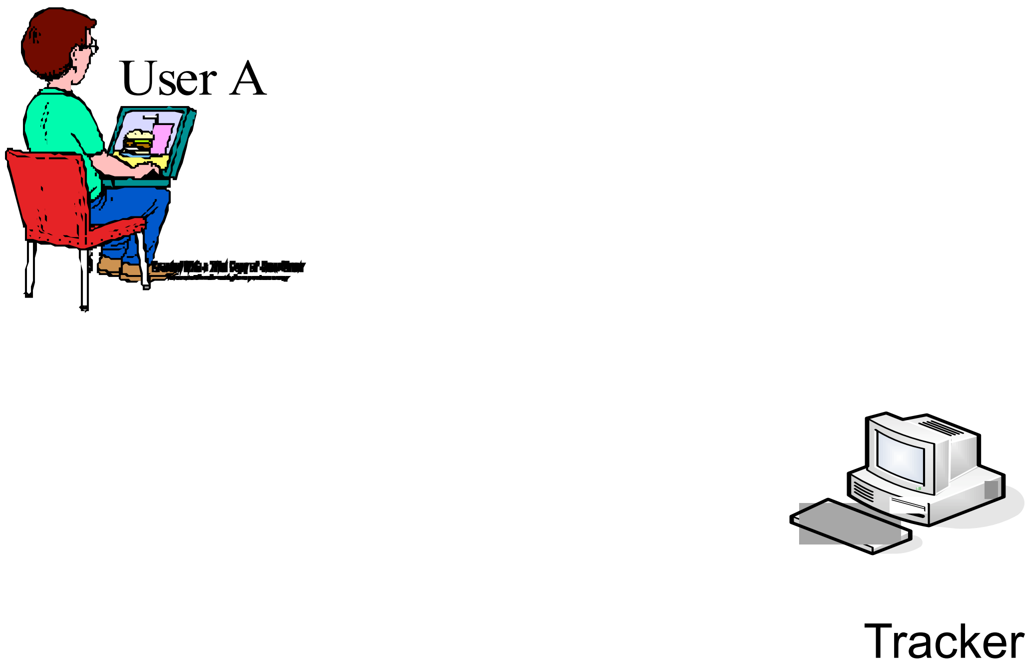
Publishing Content

Publishing Content

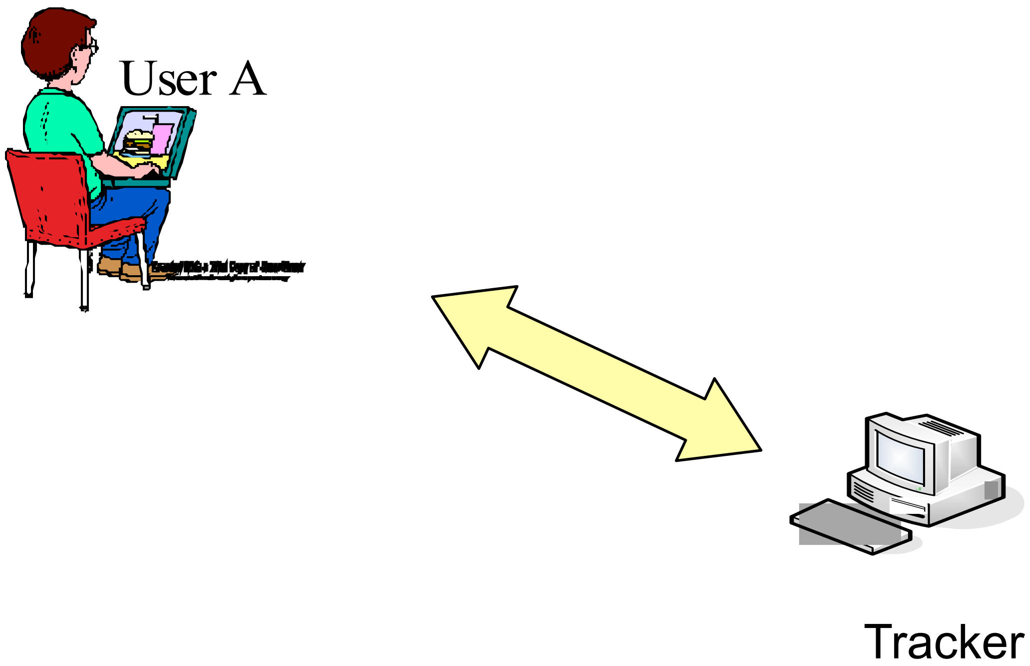


Publishing Content

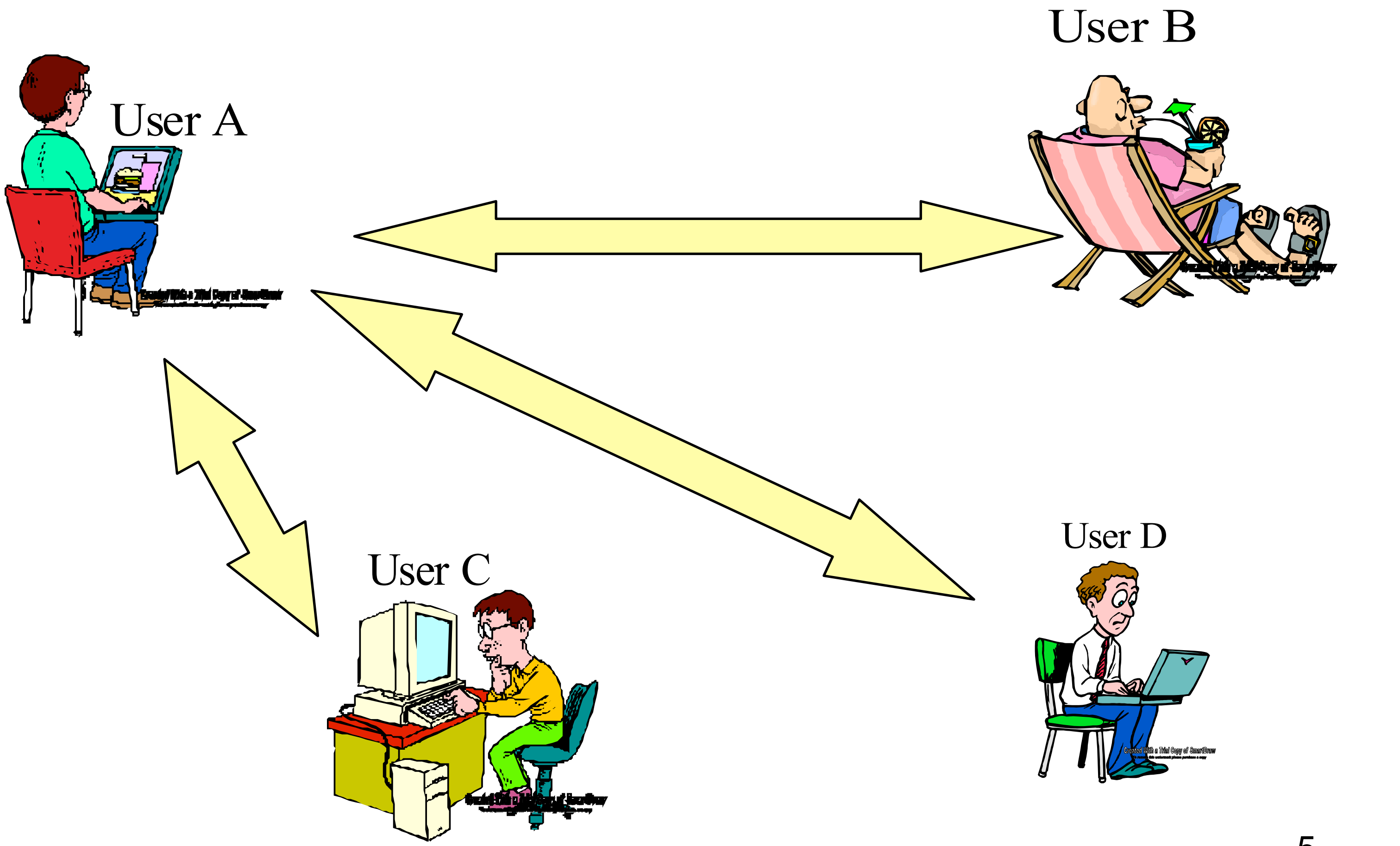
Publishing Content



Publishing Content



Publishing Content



Motivation

Motivation

- **To understand the BT protocol**

Motivation

- **To understand the BT protocol**
  - **Why is it good?**

Motivation

- **To understand the BT protocol**
  - **Why is it good?**
  - **Can we do better?**

Motivation

- **To understand the BT protocol**
  - **Why is it good?**
  - **Can we do better?**
- **Contribution:**

Motivation

- **To understand the BT protocol**
  - **Why is it good?**
  - **Can we do better?**
- **Contribution:**
  - **Analytical metrics**

Motivation

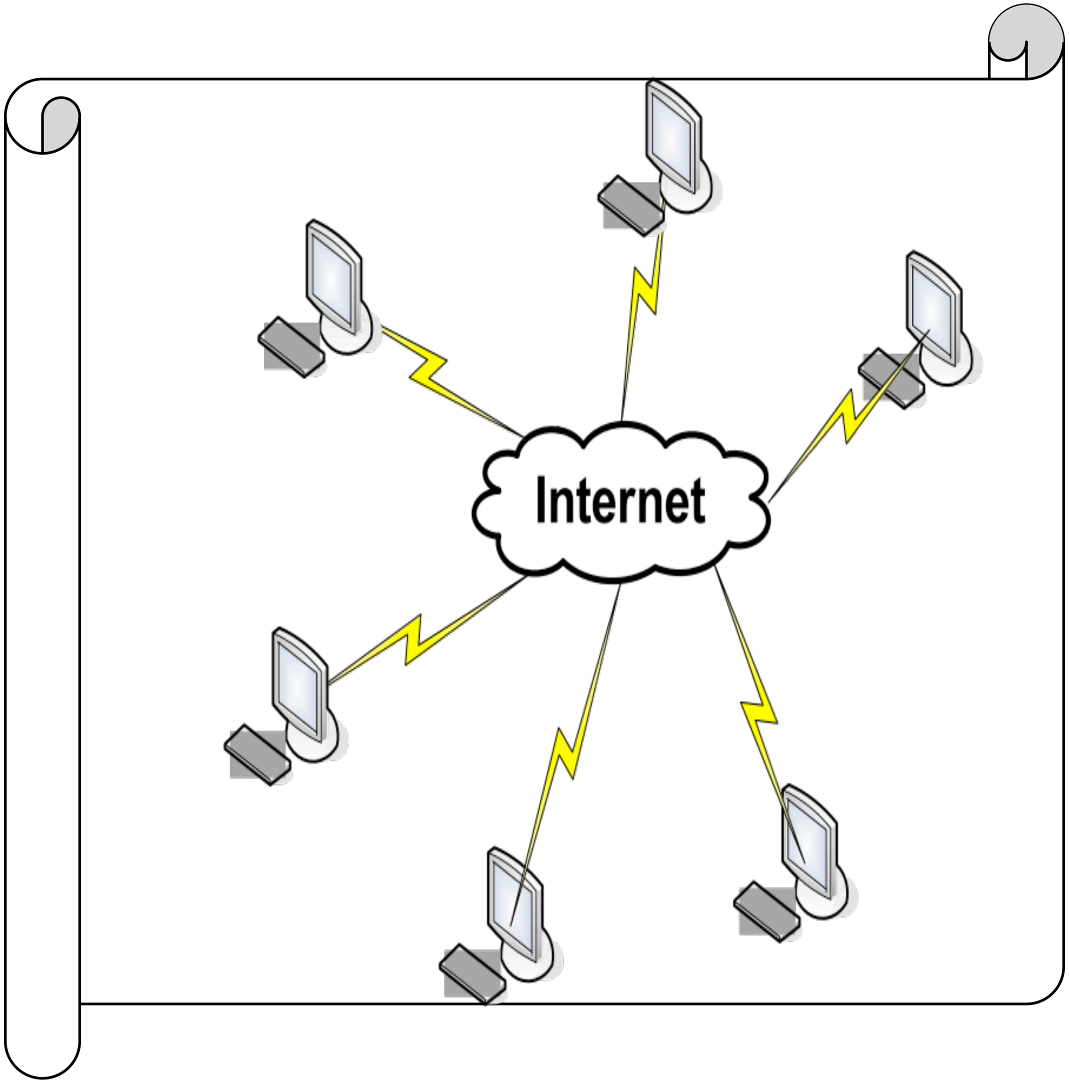
- **To understand the BT protocol**
  - **Why is it good?**
  - **Can we do better?**
- **Contribution:**
  - **Analytical metrics**
  - **Insights for Protocol-designers**

Related Work

- **X. Yang and et al, Infocom 2004:**
  - Simple Markov Model
  - Numerical calculation needed
- **L. Massoulie and et al, Sigmetrics 2004**
  - Detailed Markov Model
- **D. Qiu and et al, Sigcomm 2004**
  - Simple Fluid Model
  - Some analytical results are obtained

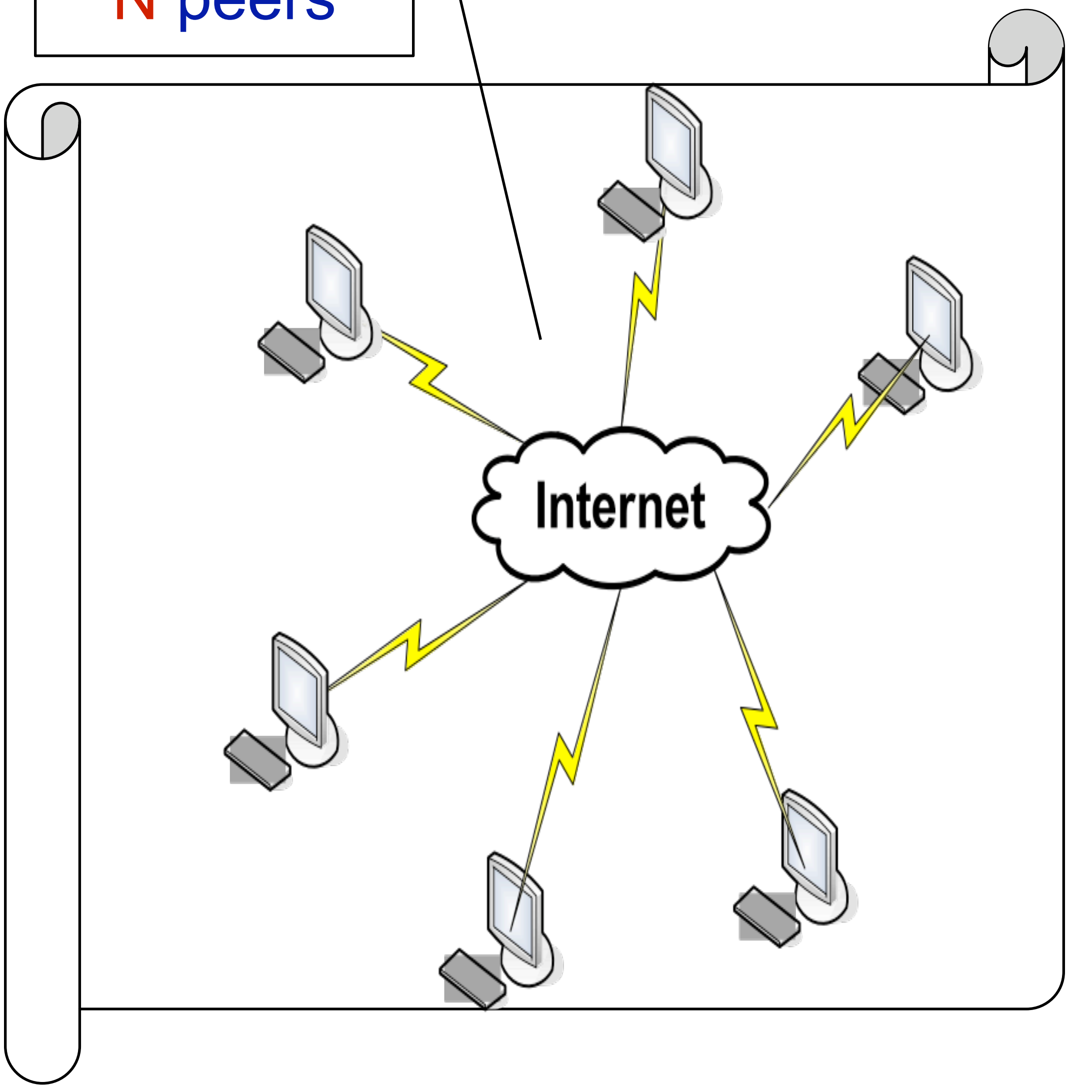
BitTorrent Model

BitTorrent Model



BitTorrent Model

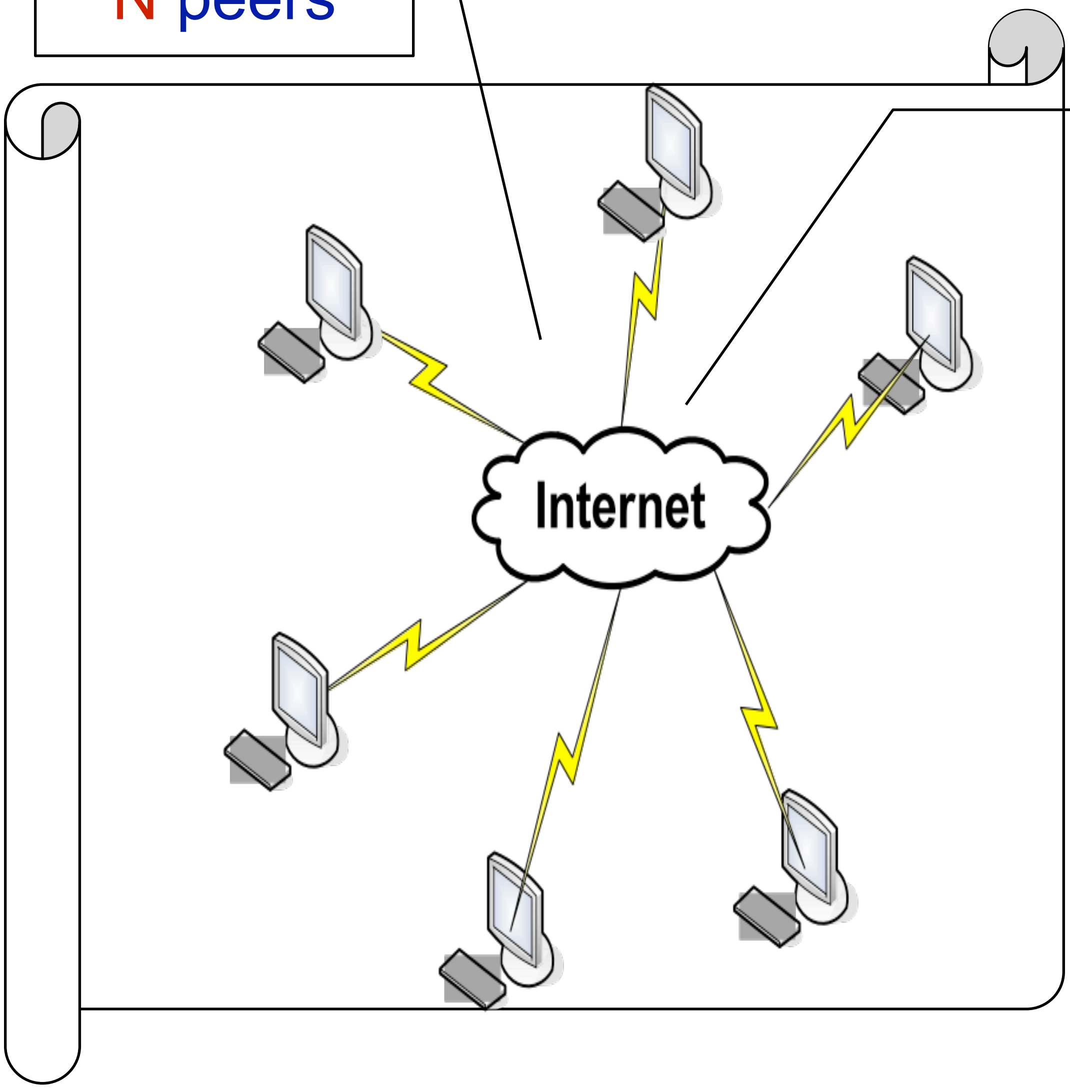
N peers



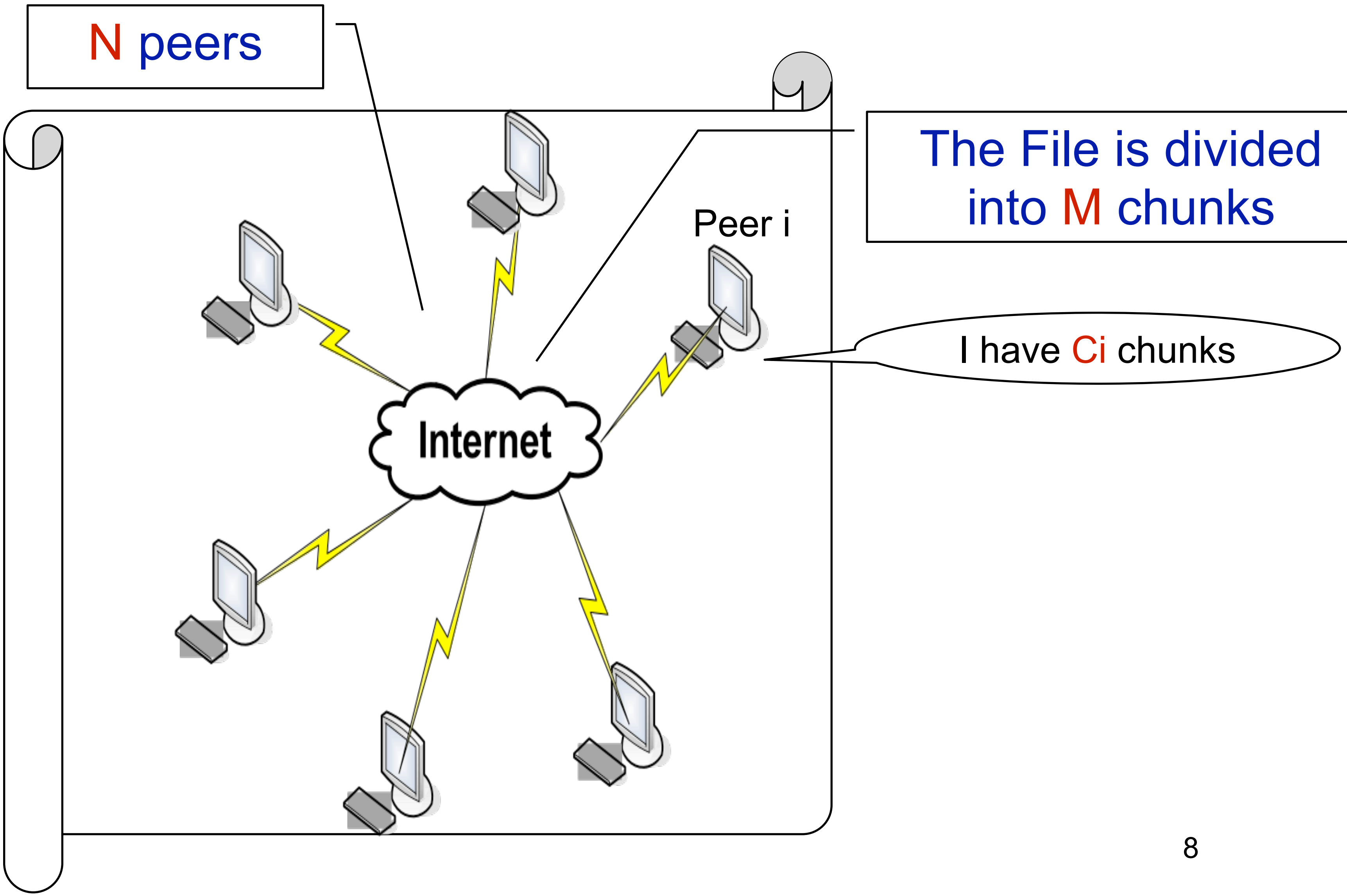
BitTorrent Model

N peers

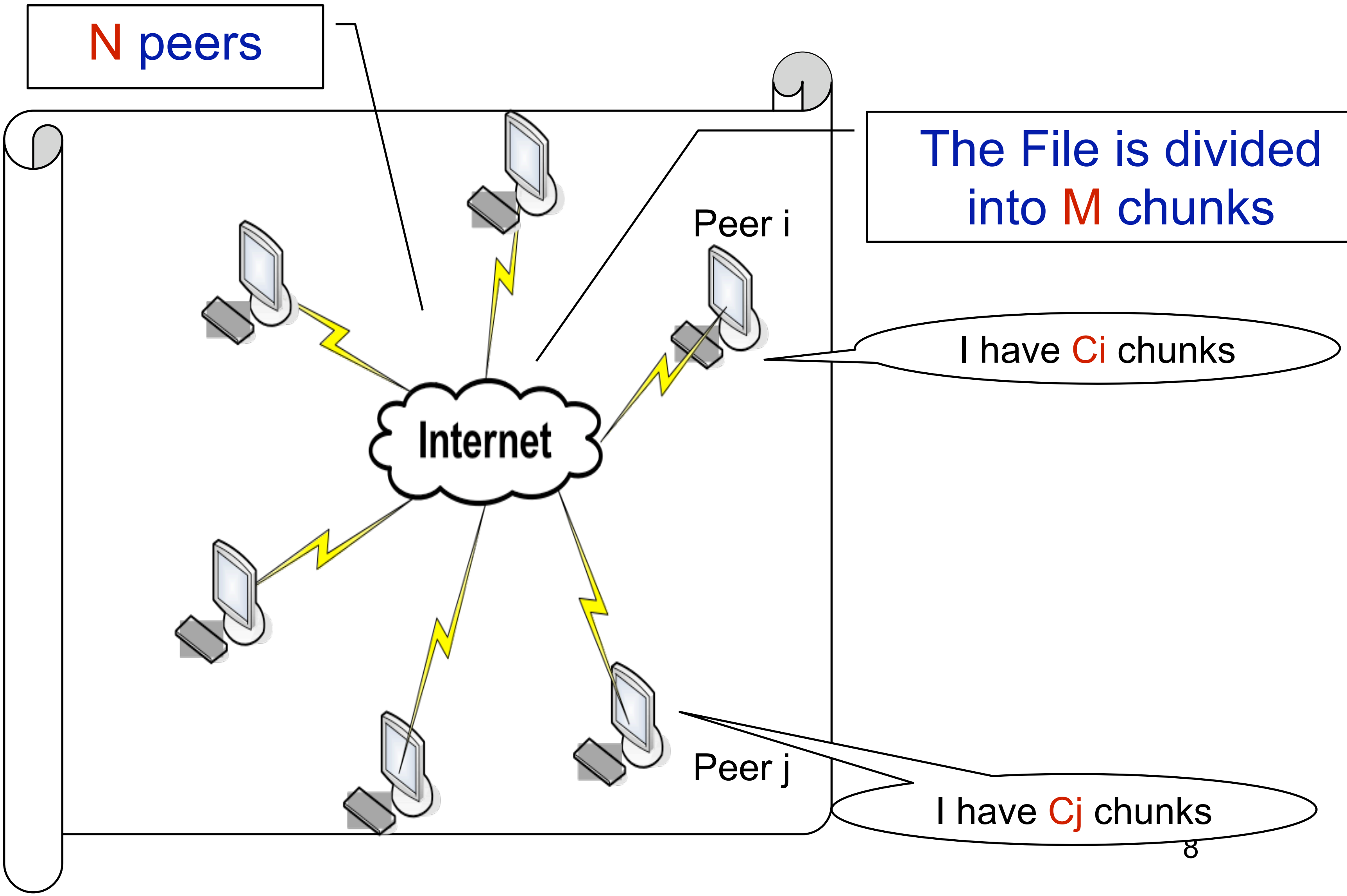
The File is divided into M chunks



BitTorrent Model

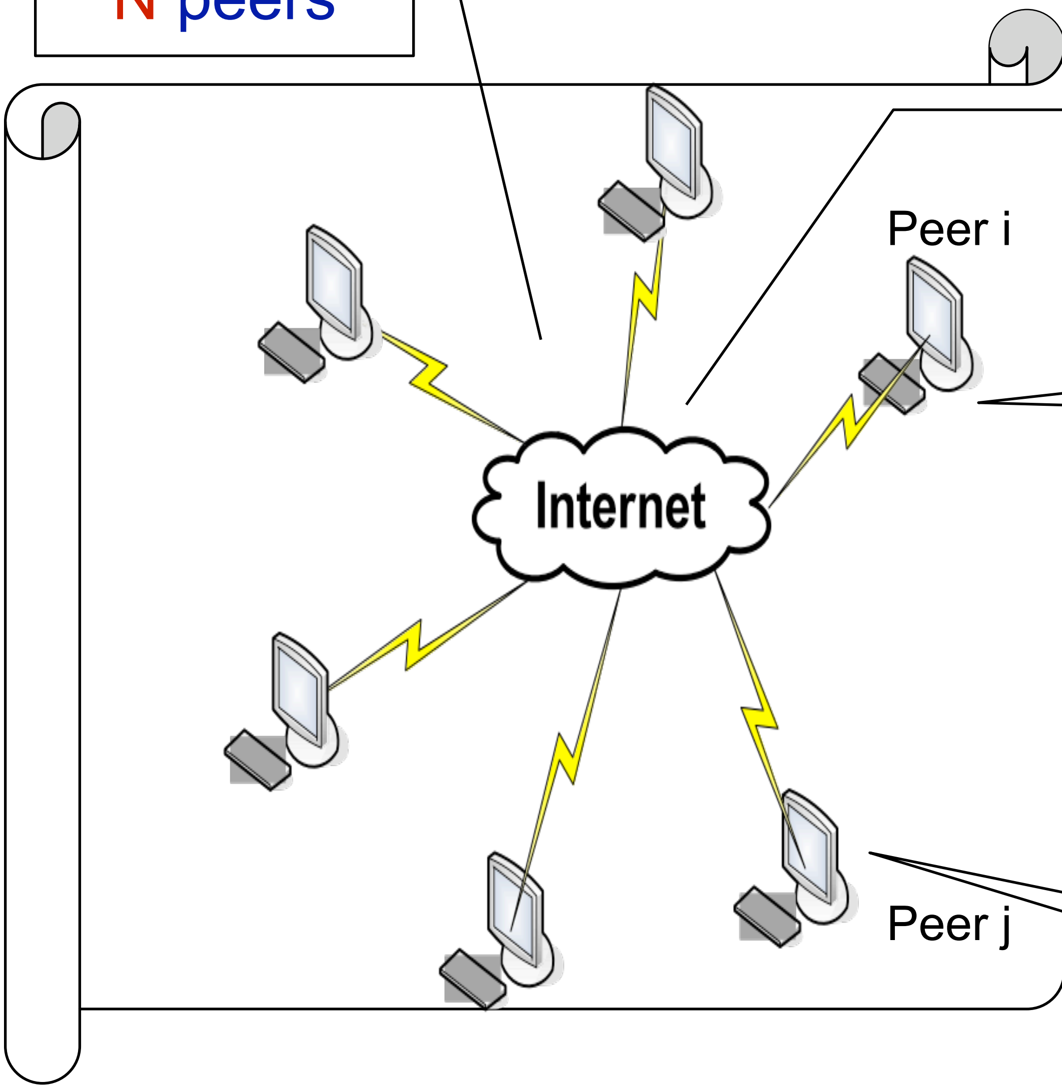


BitTorrent Model



BitTorrent Model

N peers



The File is divided into M chunks

I have Ci chunks

$$P_{i,j} = 1 \quad \text{When } c_i < c_j$$

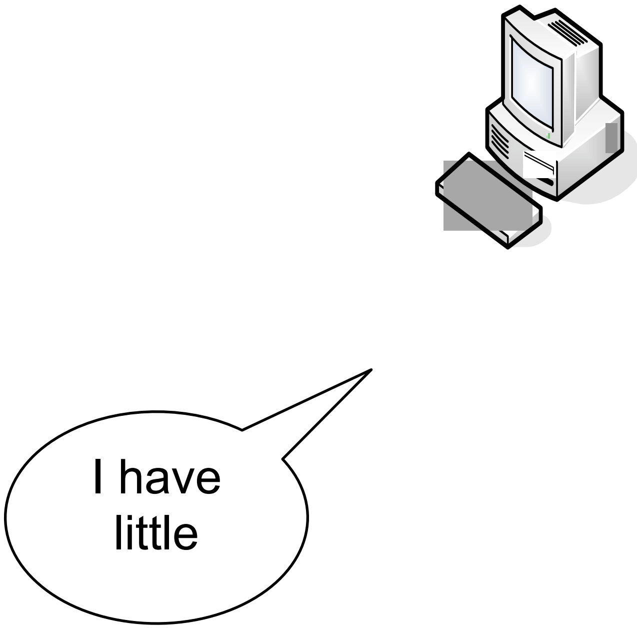
$$P_{i,j} = \quad 1 - \frac{\binom{c_i}{c_j}}{\binom{M}{c_j}}$$

When $c_i \geq c_j$

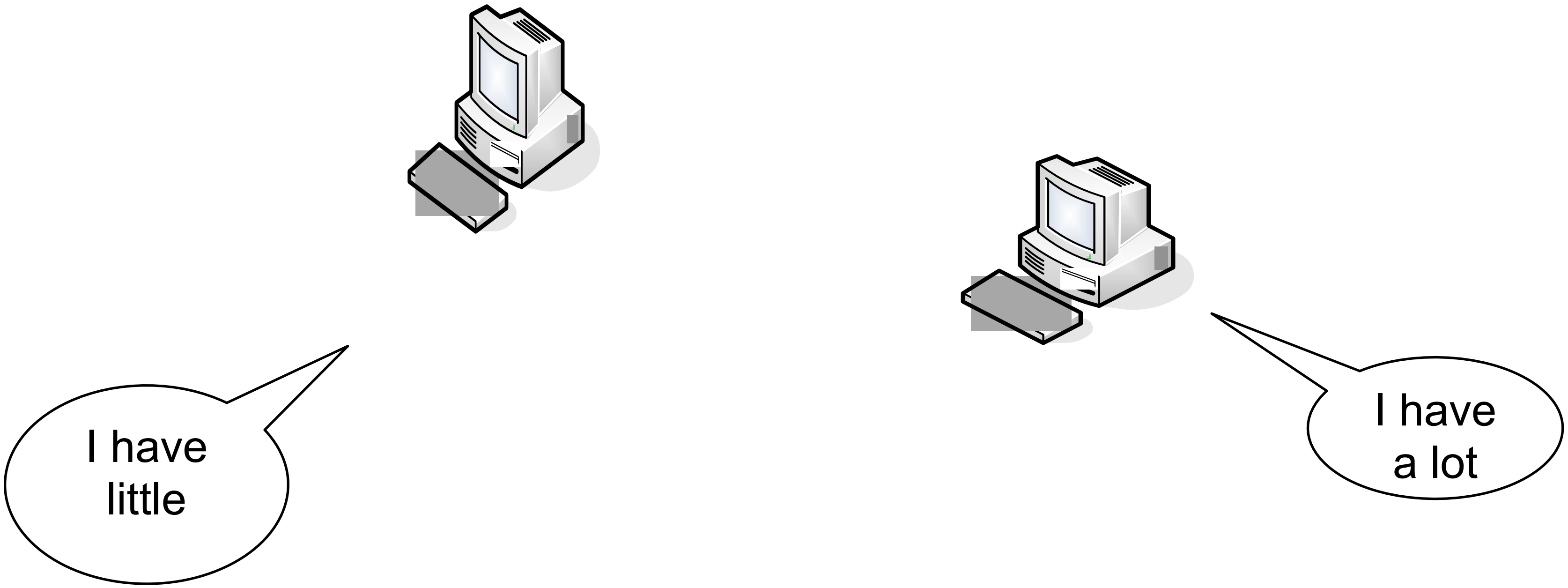
I have Cj chunks

Simplified Peer States

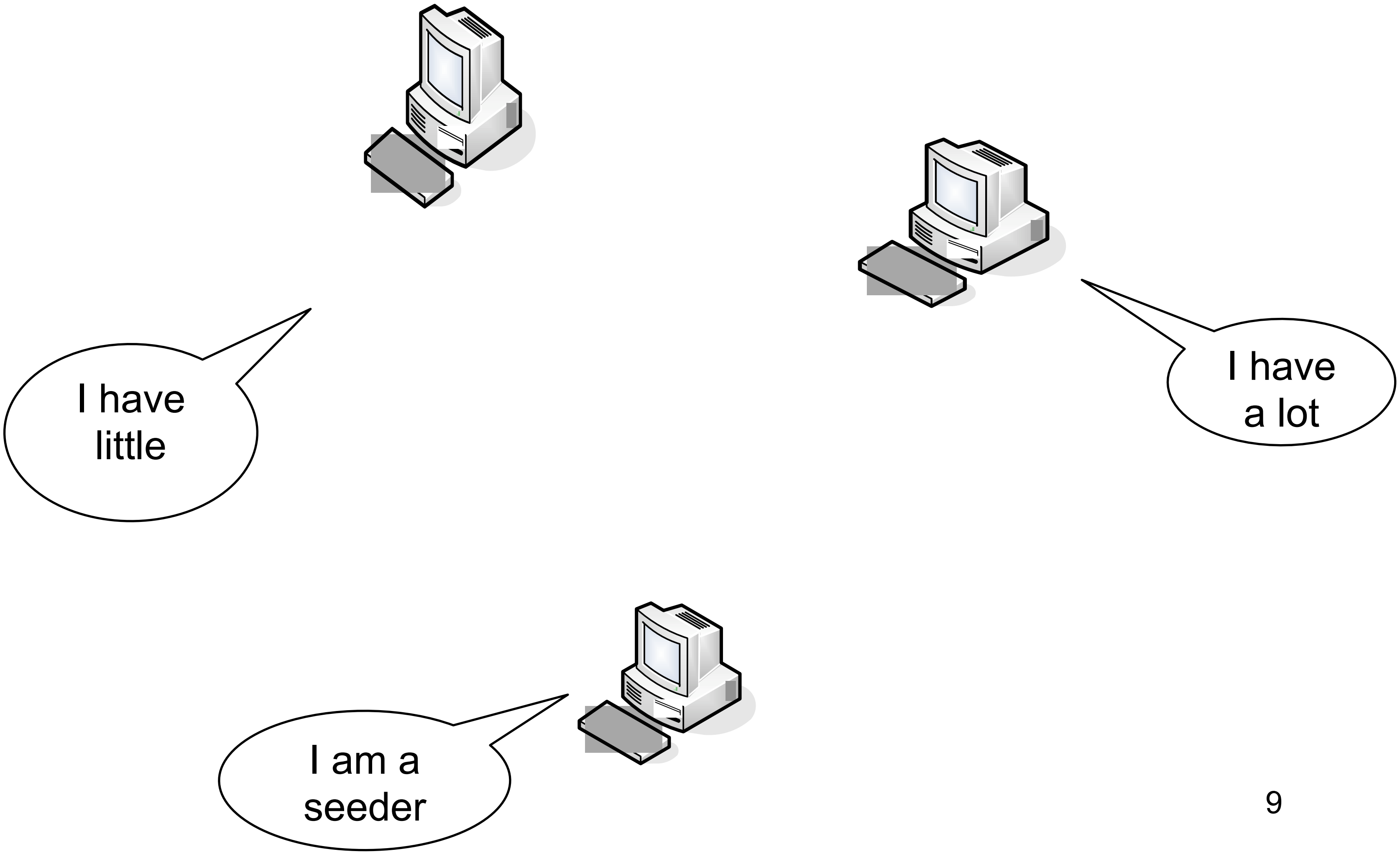
Simplified Peer States



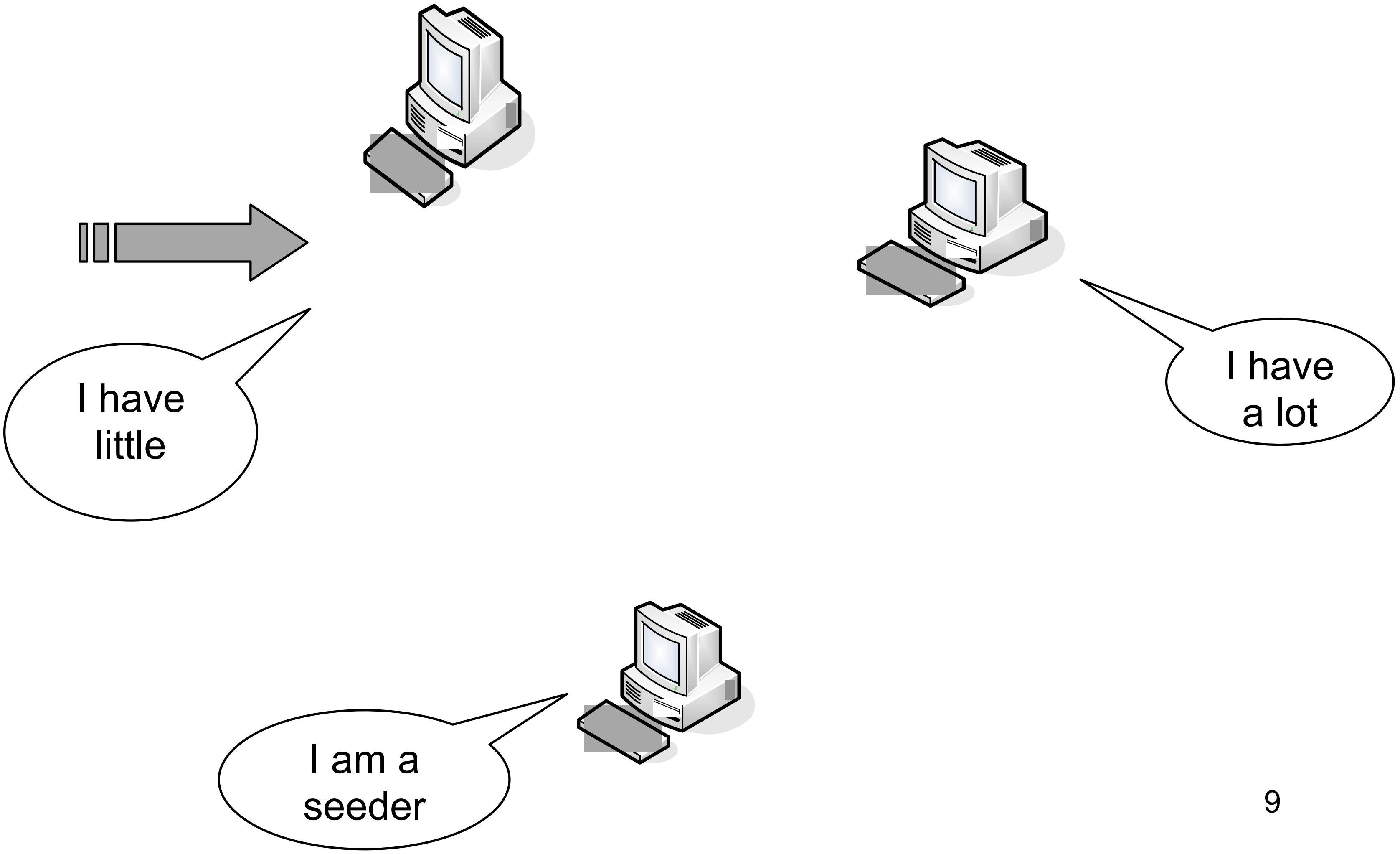
Simplified Peer States



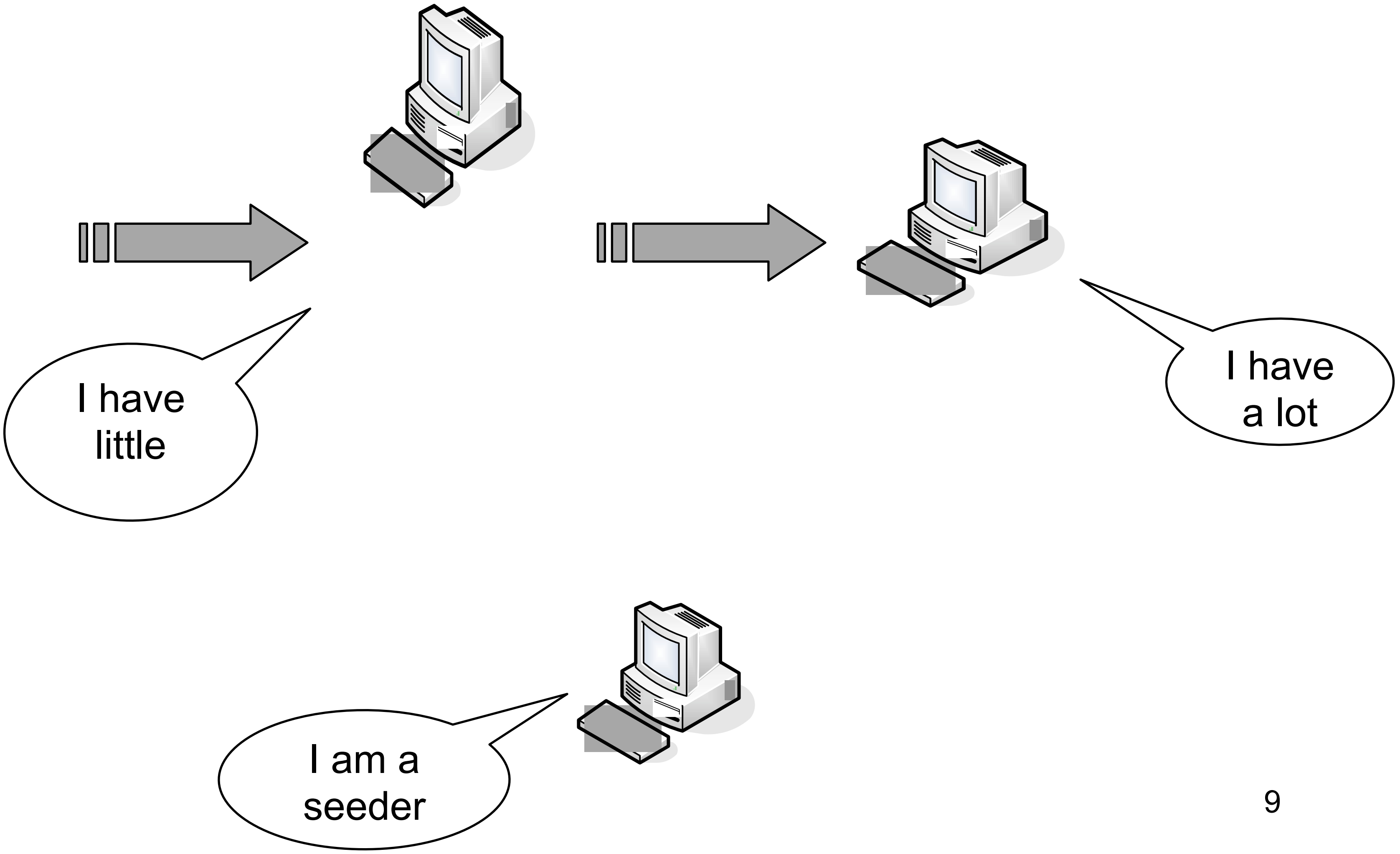
Simplified Peer States



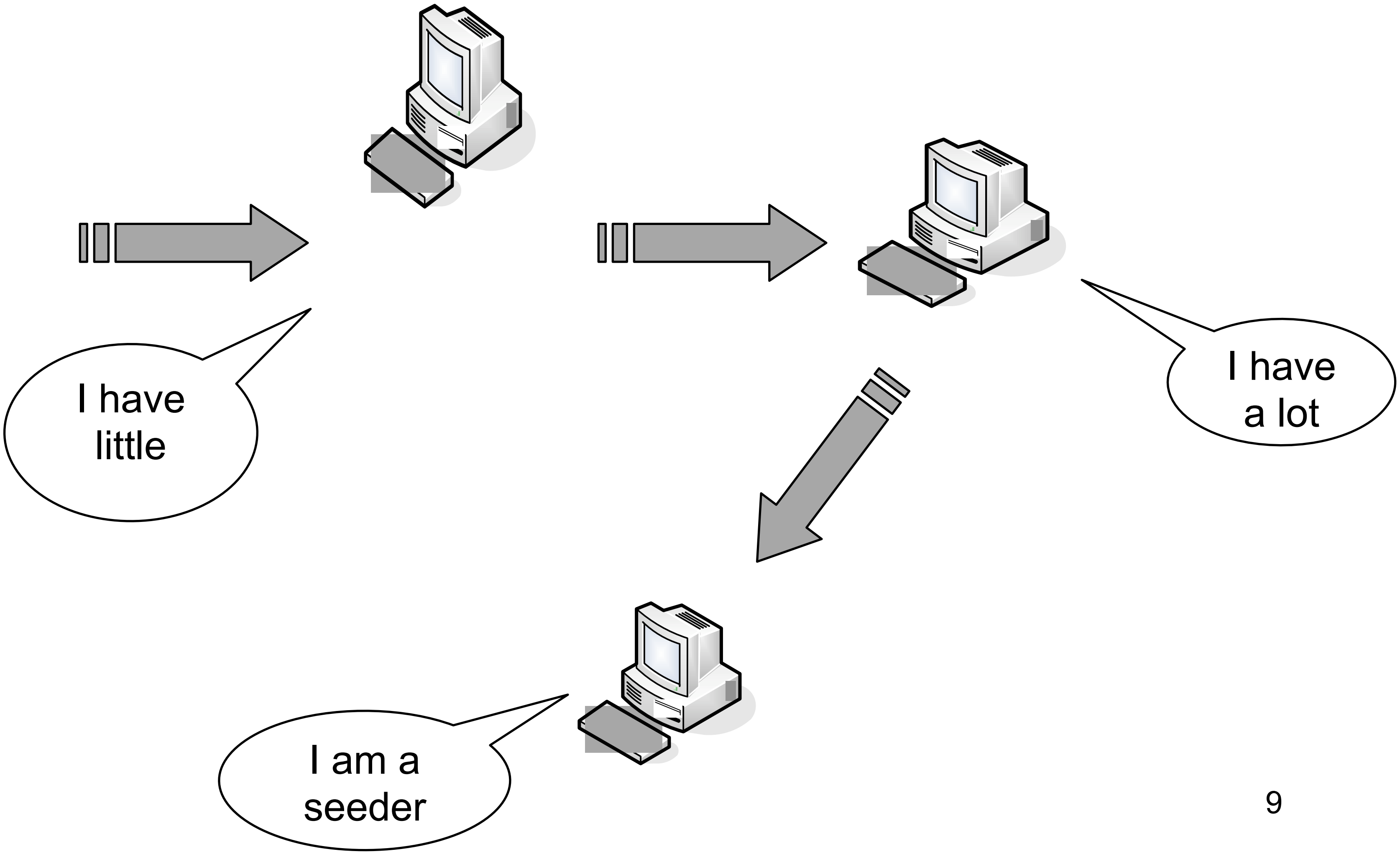
Simplified Peer States



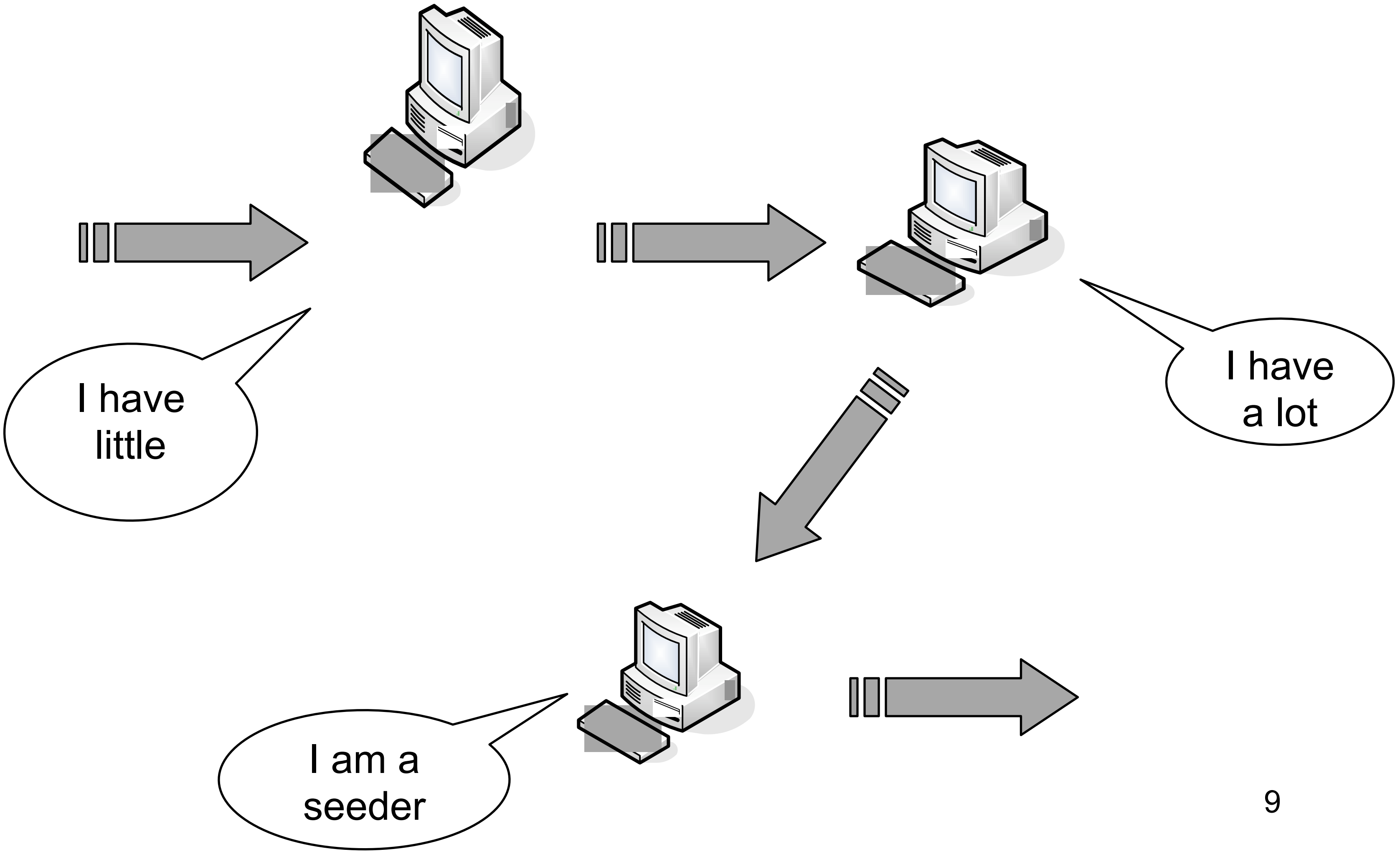
Simplified Peer States



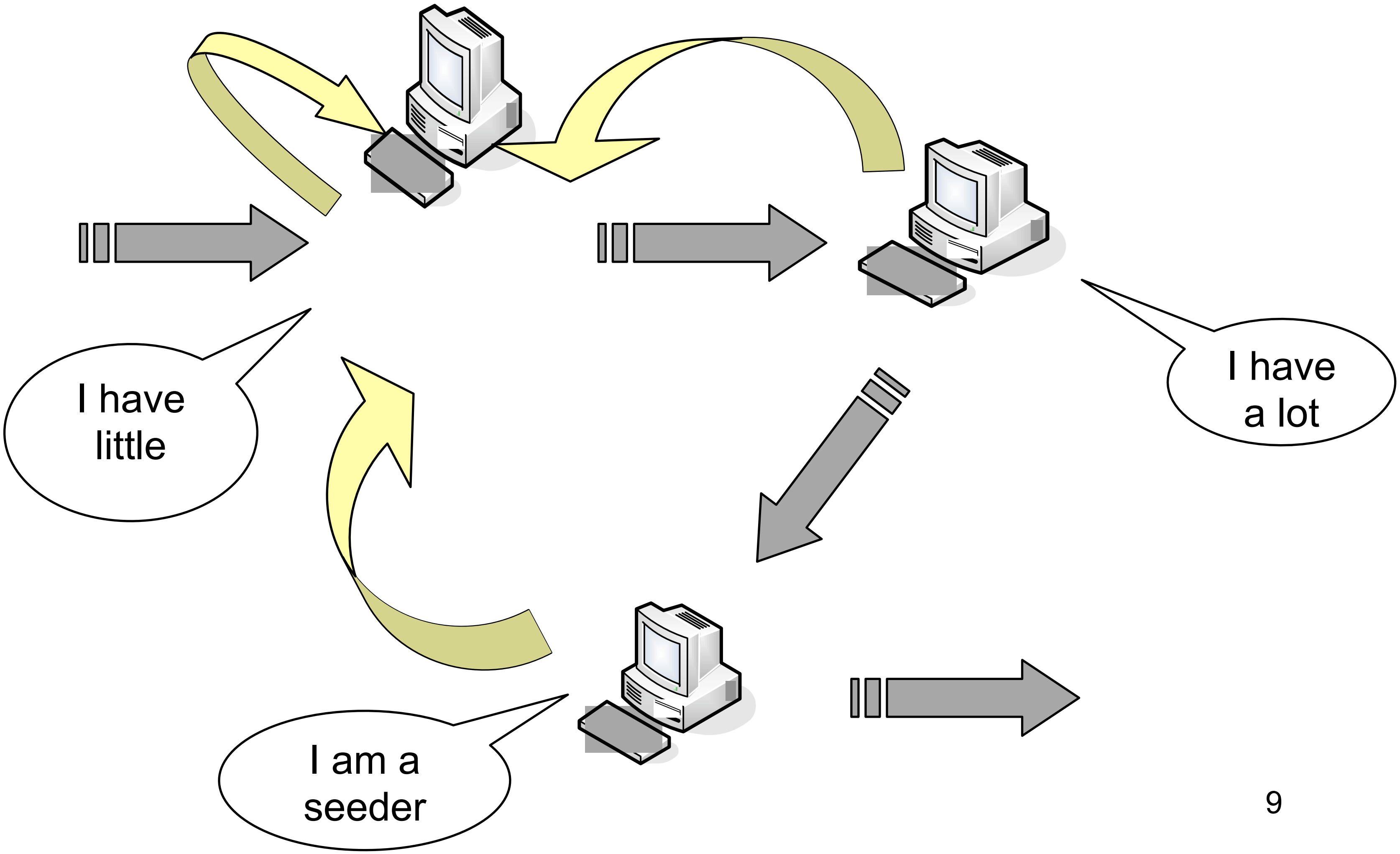
Simplified Peer States



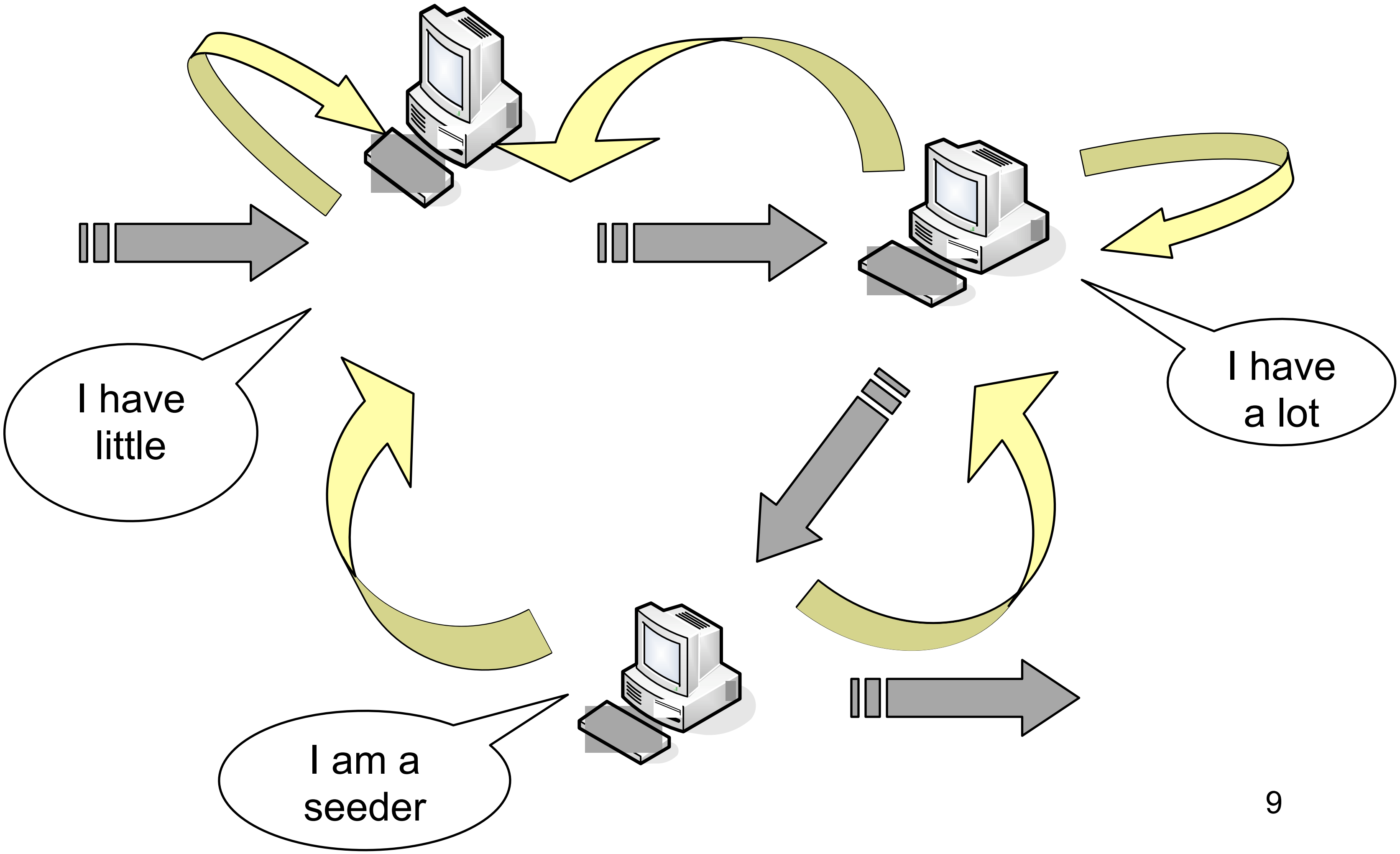
Simplified Peer States



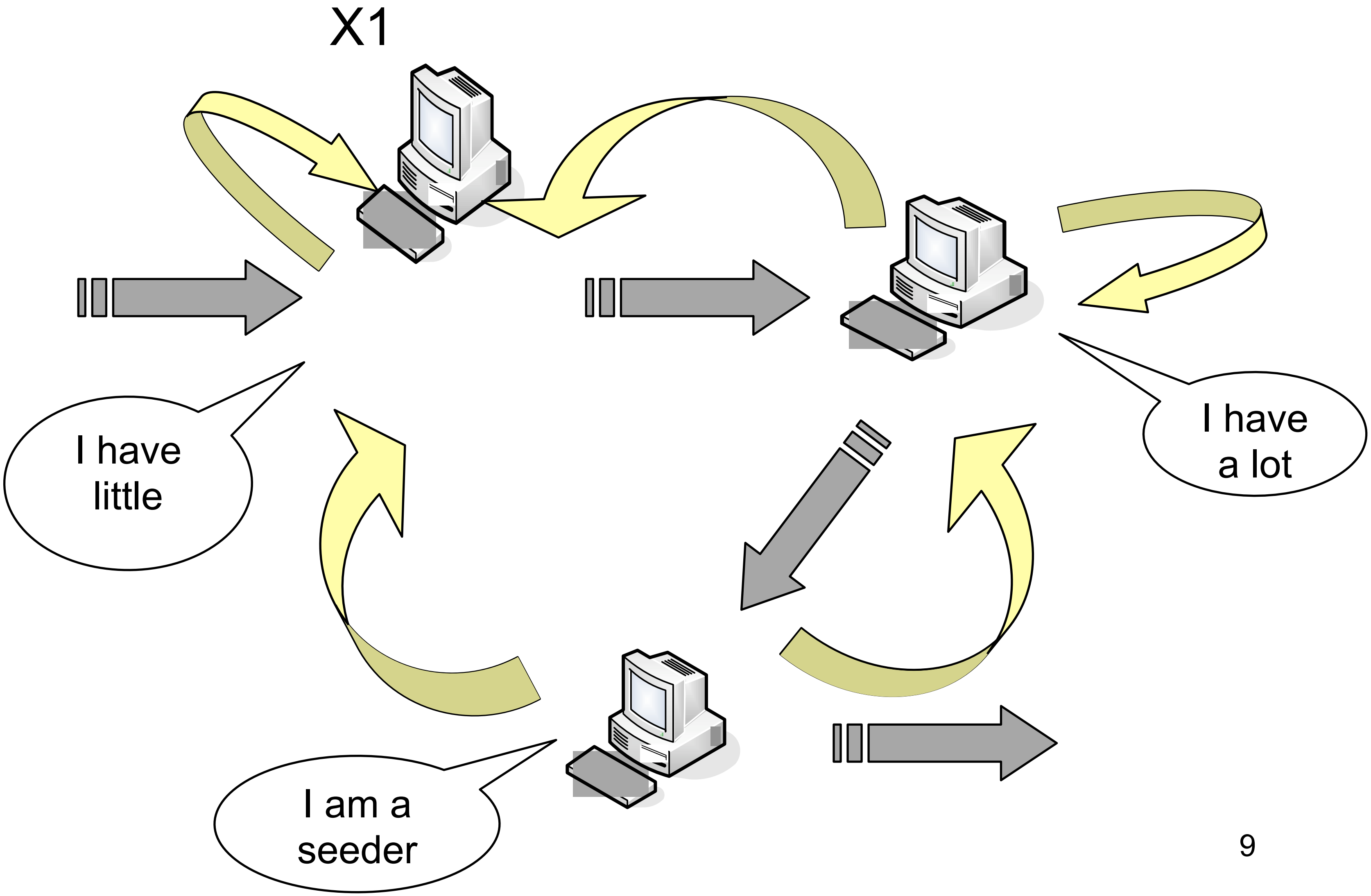
Simplified Peer States



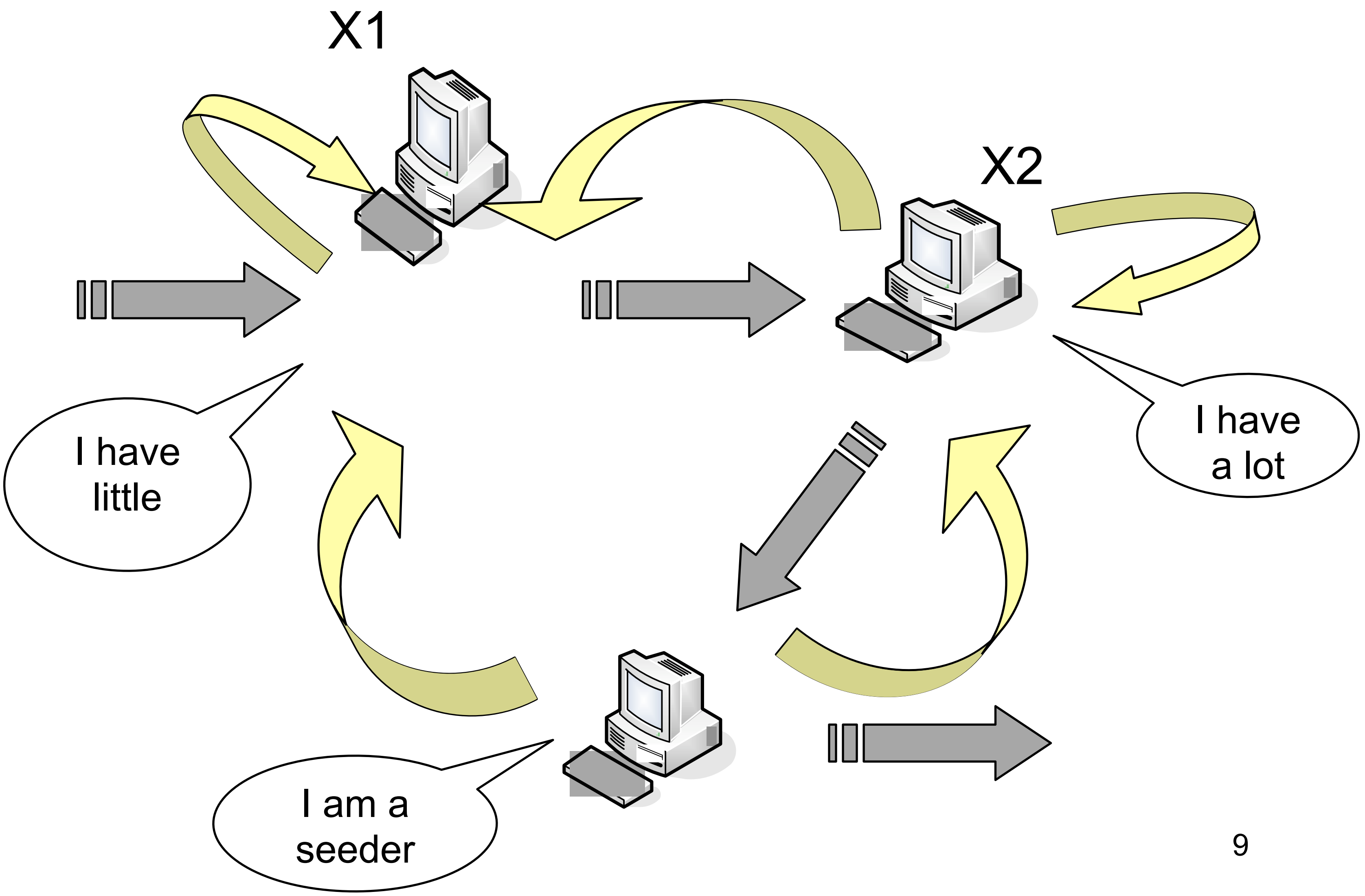
Simplified Peer States



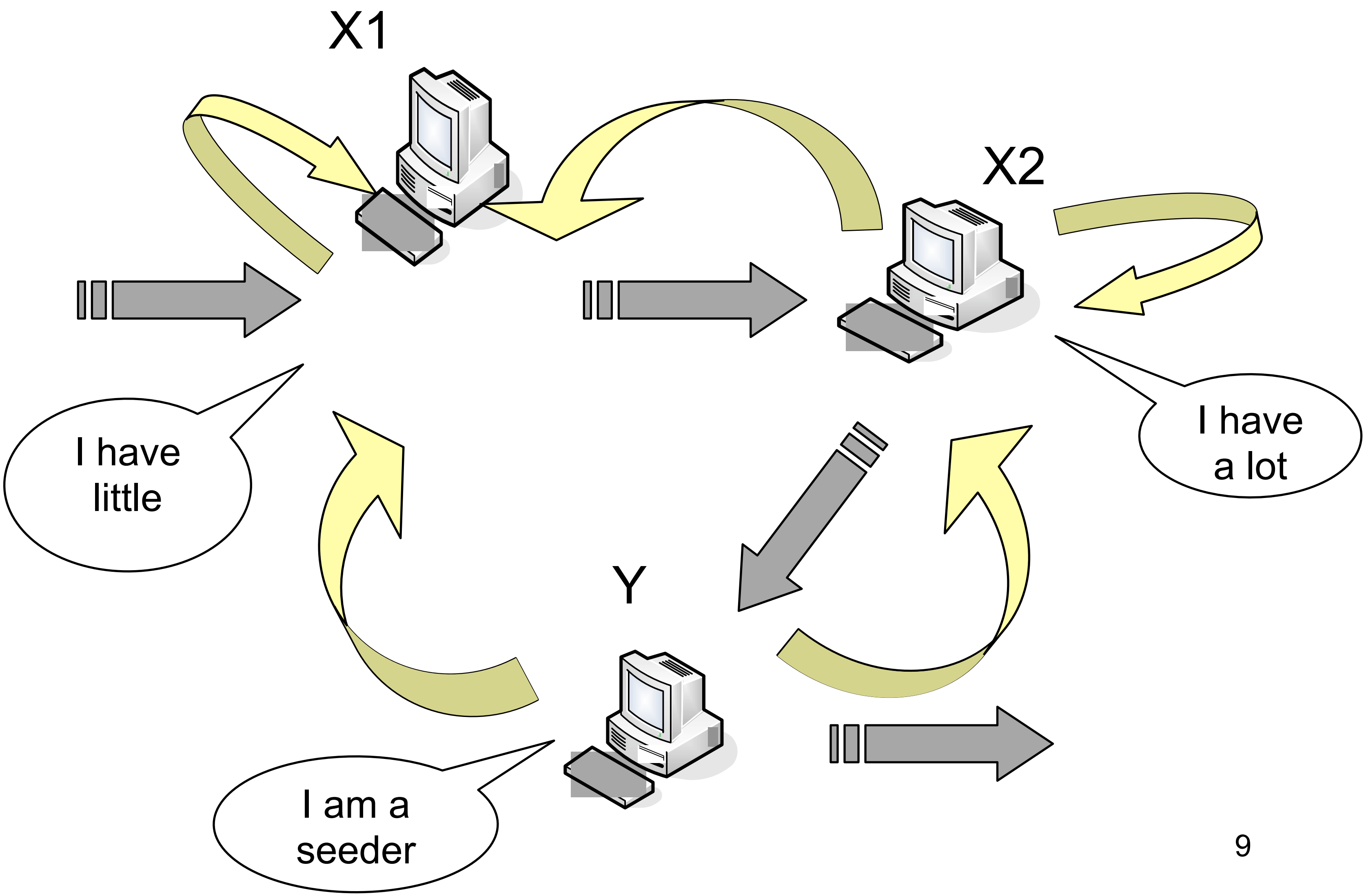
Simplified Peer States



Simplified Peer States



Simplified Peer States



Dynamics

Dynamics

$$\dot{X}1 = \text{ë} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \text{ã}Y$$

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

Dynamics

$$\dot{\mathrm{X}}1 = \text{ë} - R_1$$

$$\dot{\mathrm{X}}2 = R_1 - R_2$$

$$\dot{\mathrm{Y}} = R_2 - \text{ã}\mathrm{Y}$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

Dynamics

$$\dot{\mathrm{X}}1 = \text{ë} - R_1$$

$$\dot{\mathrm{X}}2 = R_1 - R_2$$

$$\dot{\mathrm{Y}} = R_2 - \text{ã}\mathrm{Y}$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\} / (sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\} / (sM/2)$$

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

Downloading rate per connection

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\} / (sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\} / (sM/2)$$

Dynamics

$$\dot{X}1 = \text{ë} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \text{ã}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

Dynamics

$$\dot{X}1 = \text{ë} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \text{ã}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

Connectivity

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

Dynamics

$$\dot{X}1 = \text{ë} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \text{ã}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

Bandwidth Constraint

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\} / (sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\} / (sM/2)$$

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

Chunk Size

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

Chunk Number

Dynamics

$$\dot{X}1 = \ddot{e} - R_1$$

$$\dot{X}2 = R_1 - R_2$$

$$\dot{Y} = R_2 - \tilde{a}Y$$

Transfer Rate from State 1 to State 2

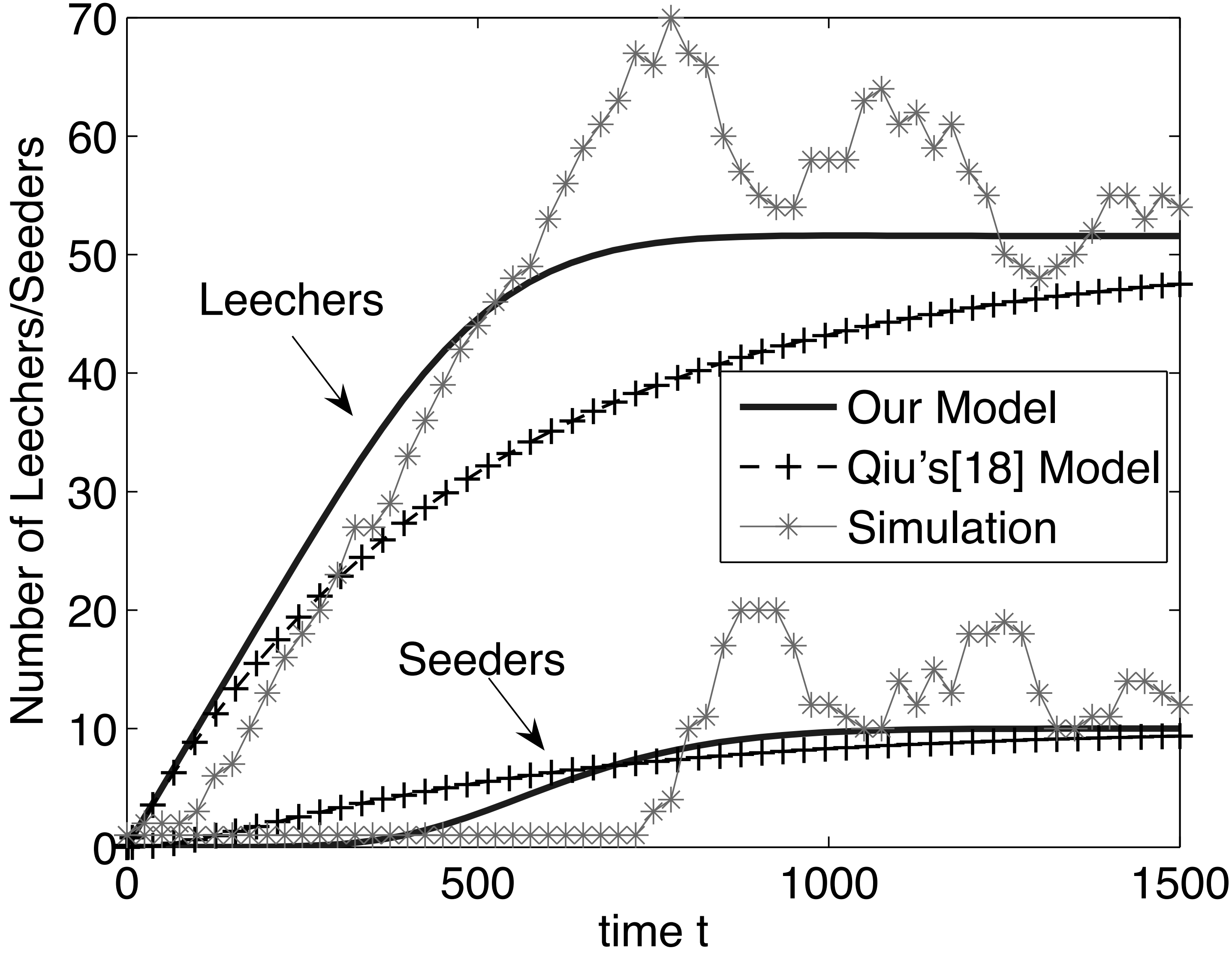
Transfer Rate from State 2 to State 3

$$R_1 = X1 \times \min\{\mu\rho(X1 + X2 + Y), B/2\}/(sM/2)$$

$$R_2 = X2 \times \min\{\mu\rho(X2 + Y), B/2\}/(sM/2)$$

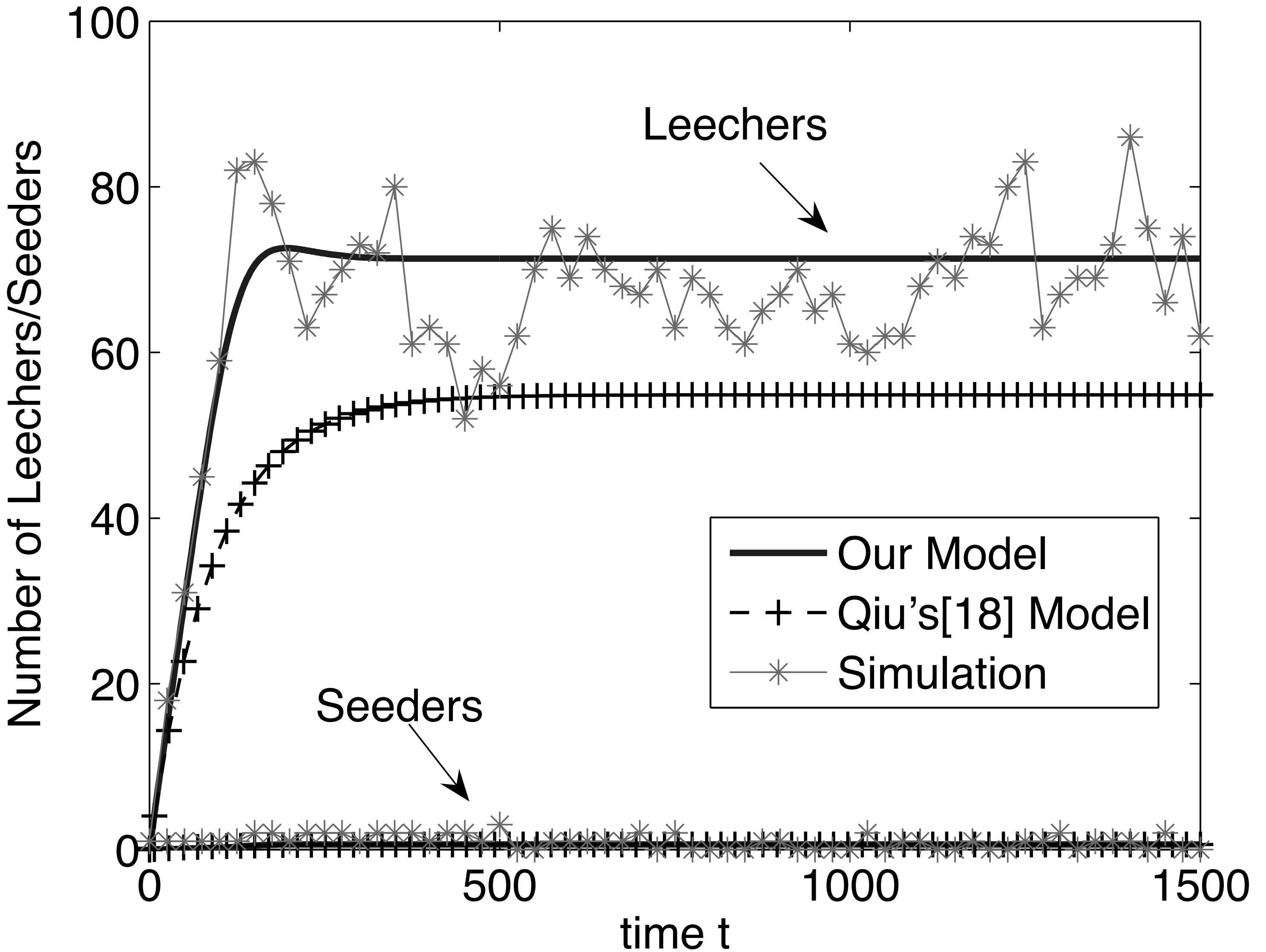
Compare the Dynamics with Simulation:

Compare the Dynamics with Simulation:



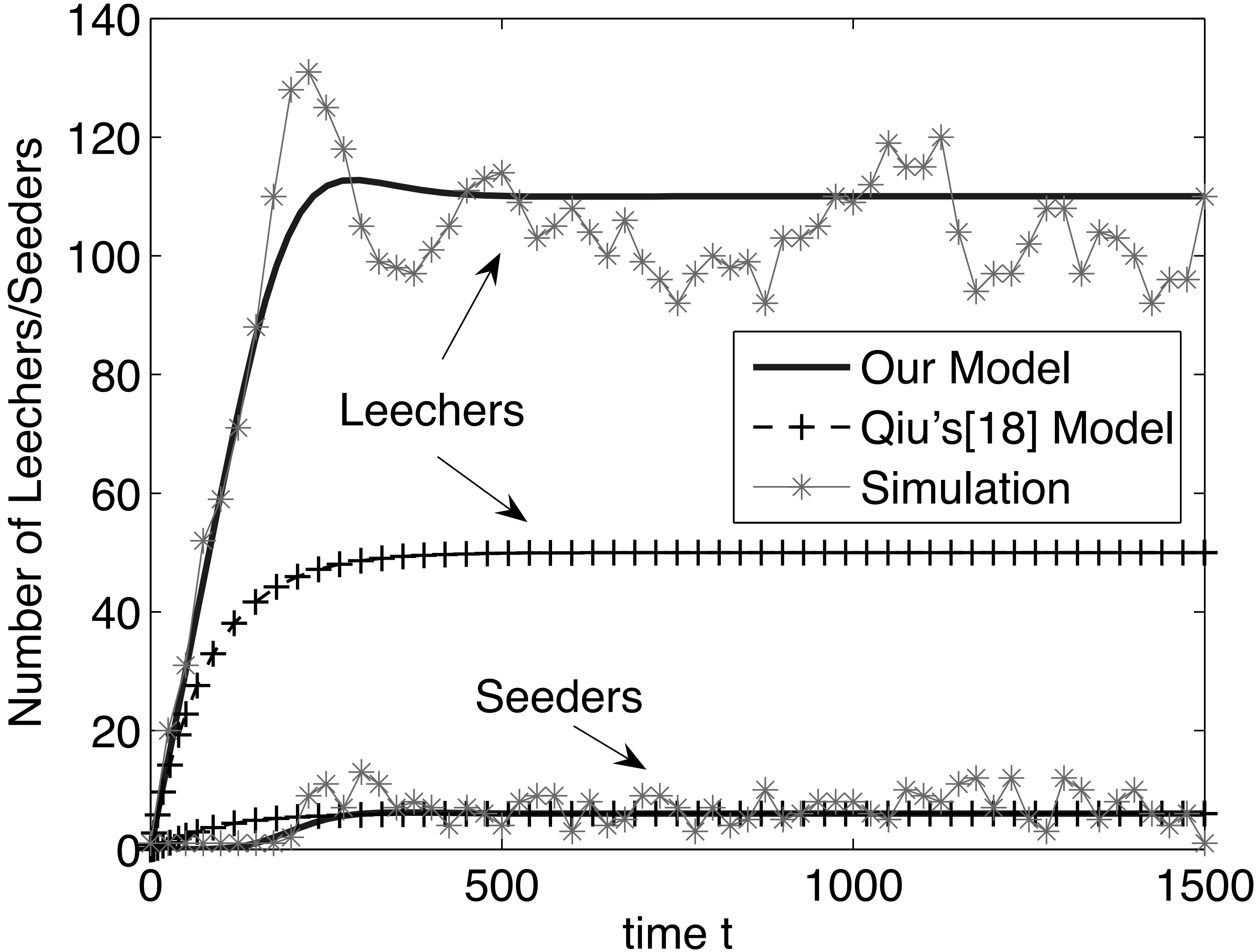
Compare the Dynamics with Simulation:

Compare the Dynamics with Simulation:



Compare the Dynamics with Simulation:

Compare the Dynamics with Simulation:



Compare the Dynamics with Simulation:

Steady State

$$\bar{T}_d = \frac{\bar{X}_1 + \bar{X}_2}{\lambda}$$

$$\bar{T}_p = \begin{cases} O(\bar{N}^2) & \textit{Case 1,} \\ O(\bar{N}) & \textit{Case 2 or 3.} \end{cases}$$

Steady State

- **By Little's Law, average downloading time in the steady state is derived by:**

$$\bar{T}_d = \frac{\bar{X}_1 + \bar{X}_2}{\lambda}$$

$$\bar{T}_p = \begin{cases} O(\bar{N}^2) & \textit{Case 1,} \\ O(\bar{N}) & \textit{Case 2 or 3.} \end{cases}$$

Steady State

- **By Little's Law, average downloading time in the steady state is derived by:**

$$\bar{T}_d = \frac{\bar{X}_1 + \bar{X}_2}{\lambda}$$

$$\bar{T}_p = \begin{cases} O(\bar{N}^2) & \textit{Case 1,} \\ O(\bar{N}) & \textit{Case 2 or 3.} \end{cases}$$

Steady State

- **By Little's Law, average downloading time in the steady state is derived by:**

$$\bar{T}_d = \frac{\bar{X}_1 + \bar{X}_2}{\lambda}$$

$$\bar{T}_p = \begin{cases} O(\bar{N}^2) & \textit{Case 1,} \\ O(\bar{N}) & \textit{Case 2 or 3.} \end{cases}$$

Steady State

- By Little's Law, average downloading time in the steady state is derived by:

$$\bar{T}_d = \frac{\bar{X}_1 + \bar{X}_2}{\lambda}$$

- The system throughput in steady state is derived by:

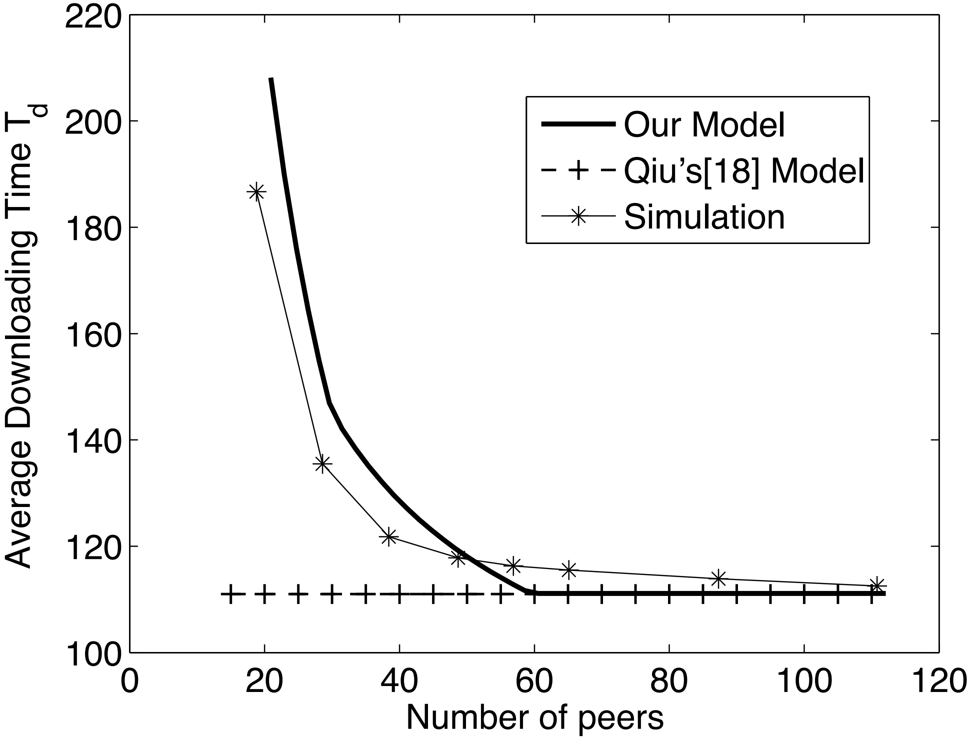
$$\bar{T}_p = \begin{cases} O(\bar{N}^2) & \textit{Case 1,} \\ O(\bar{N}) & \textit{Case 2 or 3.} \end{cases}$$

Insights: Scalability

- **The system shows good scalability:**
  - **Tp=O(N^2) in Case 1**
  - **Tp=O(N) in Case 2 & 3**

Insights: Scalability

- **The system shows good scalability:**
  - **Tp=O(N^2) in Case 1**
  - **Tp=O(N) in Case 2 & 3**



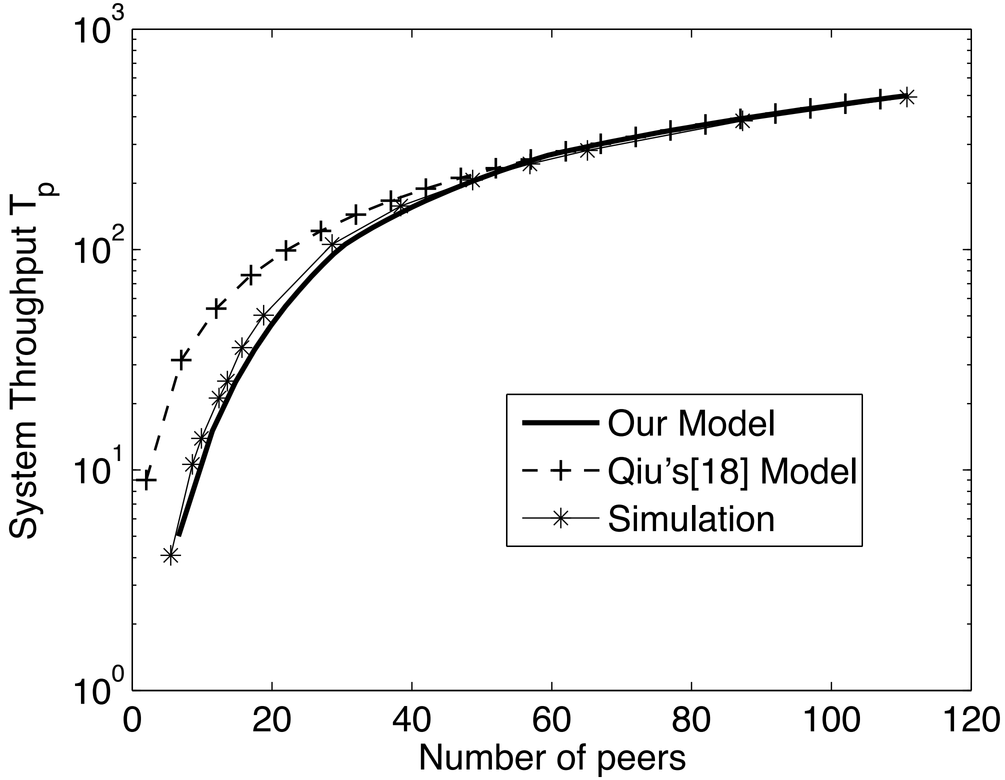
(a) $\bar{T}_d$ as the function of N

Insights: Scalability

- The system shows good scalability:
  - $Tp=O(N^2)$ in Case 1
  - $Tp=O(N)$ in Case 2 & 3

Insights: Scalability

- **The system shows good scalability:**
  - **Tp=O(N^2) in Case 1**
  - **Tp=O(N) in Case 2 & 3**



(b) $\bar{T}_p$ as the function of N

Insights: Scalability

- **The system shows good scalability:**
  - **Tp=O(N^2) in Case 1**
  - **Tp=O(N) in Case 2 & 3**

Insights: Popularity

- **The arrival rate λ represents the popularity of the served file.**

$$\frac{\partial \bar{T}_d}{\partial \lambda} = \begin{cases} -\frac{1+\sqrt{5}}{4\sqrt{\alpha}}\lambda^{-3/2} & \text{Case 1,} \\ -\frac{1}{2\sqrt{\alpha}}\lambda^{-3/2} & \text{Case 2,} \\ 0 & \text{Case 3.} \end{cases}$$

  - More popular the file is, less downloading time, in Case 1 and 2.
  - The downloading time keeps the same in Case 3.

Insights: Seeding

- Let $T_s = 1/\gamma$ be the average seeding time
- Increase seeding time T_s:
  - less downloading time T_d in case 1 and 2;
  - same downloading time T_d in case 3.
- Extreme situation: T_s=0:
  - Downloading time T_d won't be infinity

Insights: Topology

- The average degree of a peer in overlay:

$$\rho(N-1)$$

- This degree is affected by the list returned by tracker (30-60 by default)
- Larger ρ :
  - reduce T_d in case 1 and 2
  - won't help in case 3, only burden the network

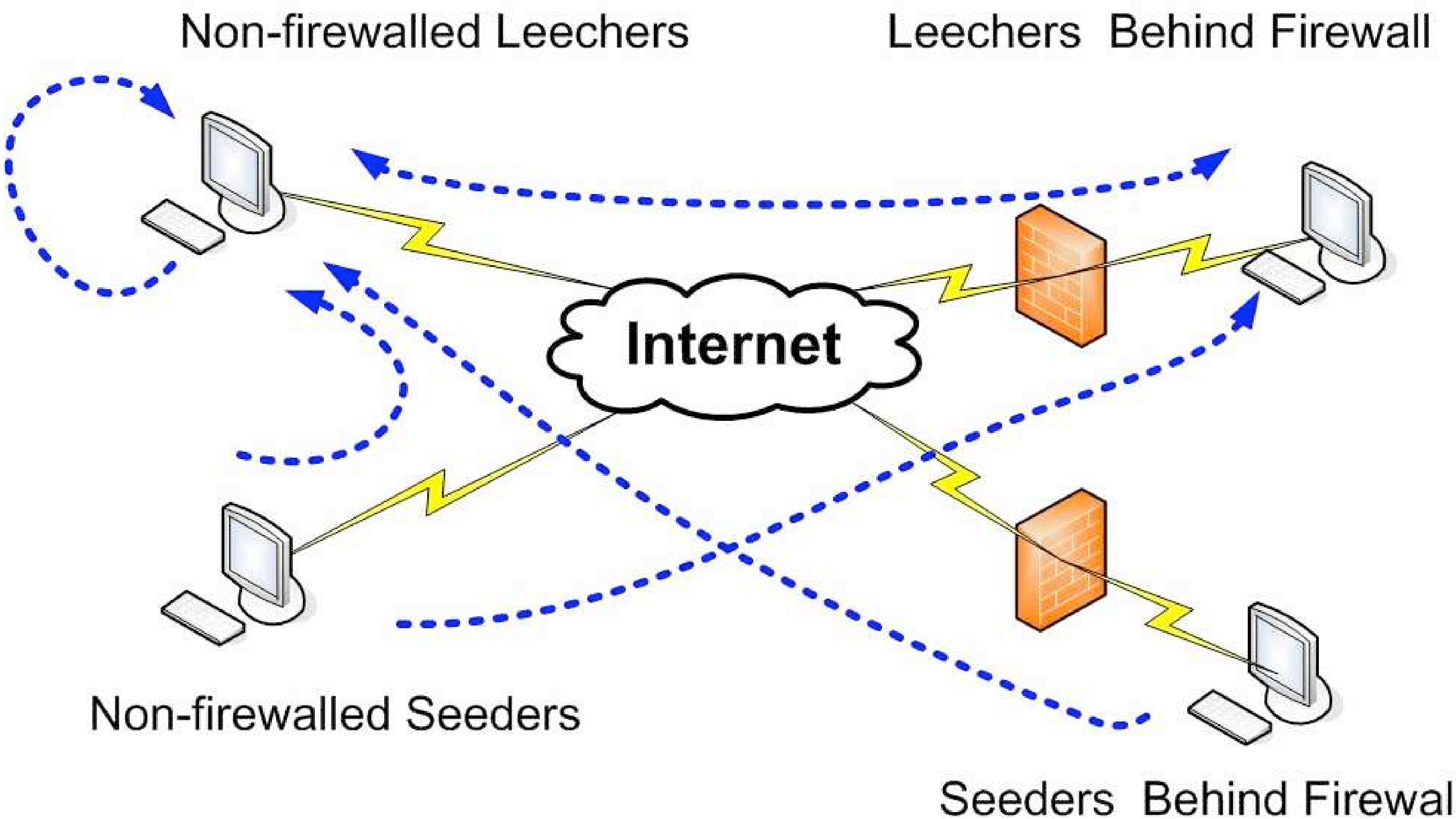
Insights: Bandwidth

- Larger B:
  - reduce T_d in case 2 and 3
  - won't help in case 1

Where you are

- Introduction
- Modeling the Performance
- Extension on heterogeneous Network
- Modeling the Availability
- Availability Improvement Algorithms
- Conclusion

Impact of Firewall



Dynamics

Dynamics

$$\dot{X}p = \lambda p - Xp \min\{\mu\rho(Xp + Xf), B/2\}/sM$$

$$\dot{X}f = \lambda f - Xf \min\{\mu\rho Xp, B/2\}/sM$$

Dynamics

$$\dot{X}p = \lambda p - Xp \min\{\mu\rho(Xp + Xf), B/2\}/sM$$

$$\dot{X}f = \lambda f - Xf \min\{\mu\rho Xp, B/2\}/sM$$

Arrival Rate of non-firewalled peers

Dynamics

$$\dot{X}p = \lambda p - Xp \min\{\mu\rho(Xp + Xf), B/2\}/sM$$

$$\dot{X}f = \lambda f - Xf \min\{\mu\rho Xp, B/2\}/sM$$

Arrival Rate of non-firewalled peers

Arrival Rate of firewalled peers

Dynamics

$$\dot{X}p = \lambda p - Xp \min\{\mu\rho(Xp + Xf), B/2\}/sM$$

$$\dot{X}f = \lambda f - Xf \min\{\mu\rho Xp, B/2\}/sM$$

Arrival Rate of non-firewalled peers

Number of non-firewalled peers

Arrival Rate of firewalled peers

Dynamics

$$\dot{X}p = \lambda p - Xp \min\{\mu\rho(Xp + Xf), B/2\}/sM$$

$$\dot{X}f = \lambda f - Xf \min\{\mu\rho Xp, B/2\}/sM$$

Arrival Rate of non-firewalled peers

Number of non-firewalled peers

Arrival Rate of firewalled peers

Number of firewalled peers

Insights:

Insights:

- Without non-firewalled peers, the peers behind firewalls can not finish downloading

Insights:

- Without non-firewalled peers, the peers behind firewalls can not finish downloading
- Non-firewalled peers perform better

Insights:

- Without non-firewalled peers, the peers behind firewalls can not finish downloading
- Non-firewalled peers perform better
- The performance gap is related to the arrival rate

Insights:

- Without non-firewalled peers, the peers behind firewalls can not finish downloading
- Non-firewalled peers perform better
- The performance gap is related to the arrival rate

This gap can be very large even the two arrival rates are very close

Where you are

- Introduction
- Modeling the Performance
- Extension on heterogeneous Network
- Modeling the Availability
- Availability Improvement Algorithms
- Conclusion

File Availability

File Availability

- The probability that a peer can get ith chunk from its neighborhood:

File Availability

- The probability that a peer can get ith chunk from its neighborhood:

$$\gamma_i = 1 - \left(1 - \frac{h_i}{n}\right)^{\rho(n-1)} \approx 1 - e^{-\rho h_i}$$

File Availability

- The probability that a peer can get ith chunk from its neighbor

Number of copies of Chunk i

$$\gamma_i = 1 - \left(1 - \frac{h_i}{n}\right)^{\rho(n-1)} \approx 1 - e^{-\rho h_i}$$

File Availability

- The probability that a peer can get ith chunk from its neighbors

$$\gamma_i = 1 - \left(1 - \frac{h_i}{n}\right)^{\rho(n-1)} \approx 1 - e^{-\rho h_i}$$

Number of copies of Chunk i

Average number of neighbors

File Availability

- The probability that a peer can get ith chunk from its neighbors

$$\gamma_i = 1 - (1 - \frac{h_i}{n})^{\rho(n-1)} \approx 1 - e^{-\rho h_i}$$

Number of copies of Chunk i

Average number of neighbors

File Availability

- The probability that a peer can get ith chunk from its neighbors

$$\gamma_i = 1 - \left(1 - \frac{h_i}{n}\right)^{\rho(n-1)} \approx 1 - e^{-\rho h_i}$$

Number of copies of Chunk i

Average number of neighbors

File Availability

- The probability that a peer can get ith chunk from its neighbors:

$$\gamma_i = 1 - \left(1 - \frac{h_i}{n}\right)^{\rho(n-1)} \approx 1 - e^{-\rho h_i}$$

Number of copies of Chunk i

Average number of neighbors

- The probability that a peer can finish its download:

File Availability

- The probability that a peer can get ith chunk from its neighbor

Number of copies of Chunk i

$$\gamma_i = 1 - \left(1 - \frac{h_i}{n}\right)^{\rho(n-1)} \approx 1 - e^{-\rho h_i}$$

Average number of neighbors

- The probability that a peer can finish its download:

$$\Theta = \prod_{i=1}^{M} \gamma_i$$

Optimal Chunk Distribution

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Optimal Chunk Distribution

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Total Storage capacity

Optimal Chunk Distribution

- Maximize the probability of downloading all chunks:

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Total Storage capacity

Optimal Chunk Distribution

- Maximize the probability of downloading all chunks:

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Total Storage capacity

Optimal Chunk Distribution

- Maximize the probability of downloading all chunks:

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Total Storage capacity

Optimal Chunk Distribution

- Maximize the probability of downloading all chunks:

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Total Storage capacity

Optimal Chunk Distribution

- Maximize the probability of downloading all chunks:

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Total Storage capacity

Optimal Chunk Distribution

- Maximize the probability of downloading all chunks:

$$Max \quad \Theta$$

$$s.t. \quad \sum_{i=1}^{M} h_i \leq C$$

Total Storage capacity

- The optimal solution:
  - h1=h2=…hM=C/M

Optimal Chunk Distribution

- Maximize the probability of downloading all chunks:

$$Max \quad \Theta$$

Chunks should be distributed as equally as possible

$$\sum_{i=1}$$

- The optimal solution:
  - h1=h2=…hM=C/M

Define the Measurement

Define the Measurement

$$\min V(h_1, h_2, \ldots h_M) = \sum_{i=1}^{M} \frac{(h_i - \bar{h})^2}{M}$$

Define the Measurement

$$\min V(h_1, h_2, \ldots h_M) = \sum_{i=1}^{M} \frac{(h_i - \overline{h})^2}{M}$$

Optimization solution to this unconstrained l1-norm problem:

only increase hi when chunk i has the fewest copies

Define the Measurement

$$\min V(h_1, h_2, \ldots h_M) = \sum_{i=1}^{M} \frac{(h_i - \bar{h})^2}{M}$$

Optimization solution to this unconstrained l1-norm problem:

only increase hi when chunk i has the fewest copies

Use Rarest First Policy

Define the Measurement

$$\min V(h_1, h_2, \ldots h_M) = \sum_{i=1}^{M} \frac{(h_i - \bar{h})^2}{M}$$

Optimization solution to this unconstrained l1-norm problem:

only increase hi when chunk i has the fewest copies

Use Rarest First Policy

Synchronization problem occurs in high connectivity system

Where you are

- Introduction
- Modeling the Performance
- Extension on heterogeneous Network
- Modeling the Availability
- Availability Improvement Algorithms
- Conclusion

File Enhancement Algorithm

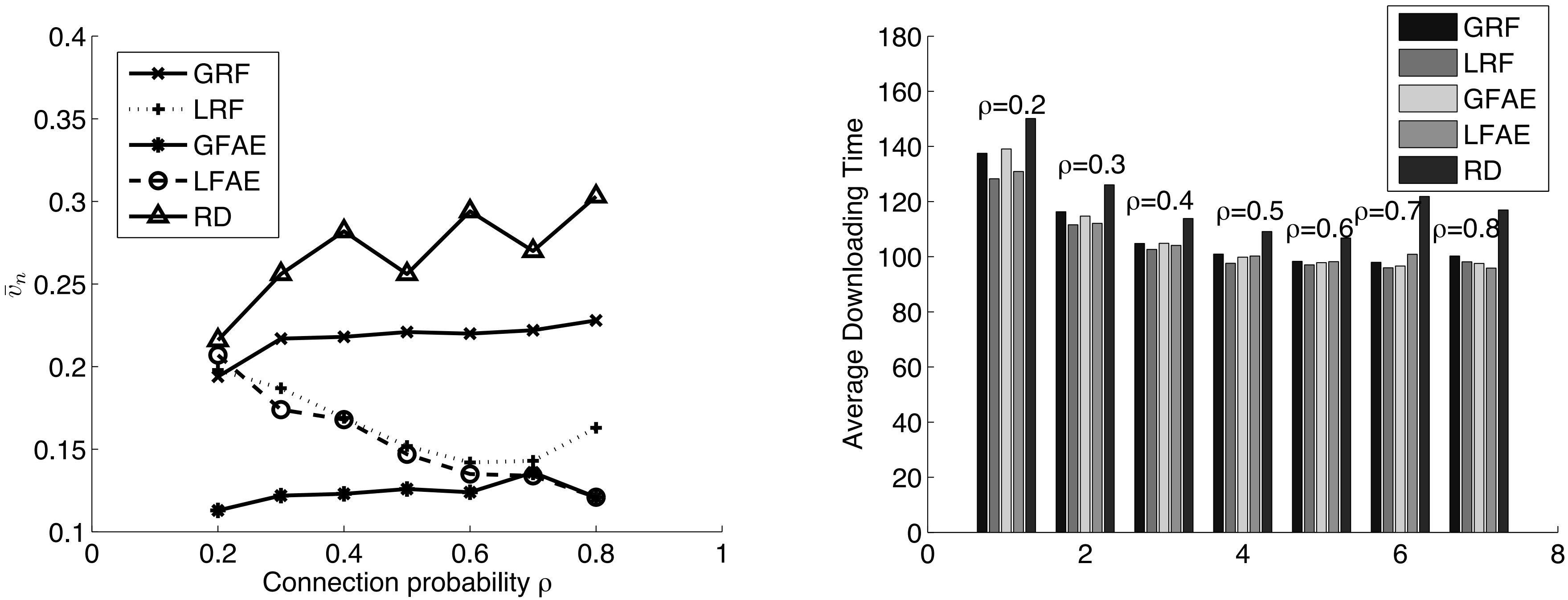
- **Choose chunk i probabilistically according to:**

$$\sigma_i = \frac{\Delta h_i}{\sum_{\forall \Delta h_j > 0} \Delta h_j}.$$

$$\Delta h_i = \begin{cases} \frac{\partial V}{\partial h_i} = \frac{2(\bar{h} - h_i)}{M} & \text{if } h_i \leq \bar{h} \\ 0 & \text{otherwise.} \end{cases}$$

Experiments

- Low bandwidth:

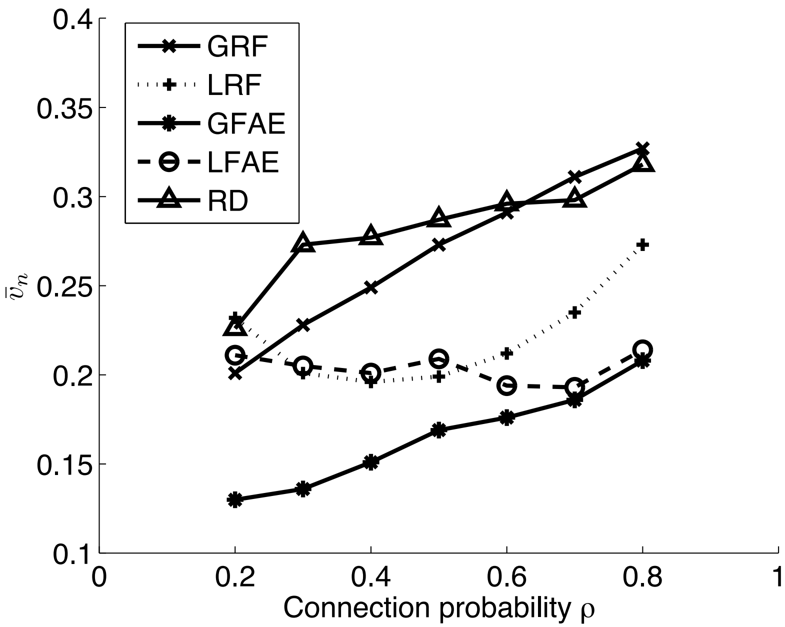


(a) Availability, $B = 4.5$

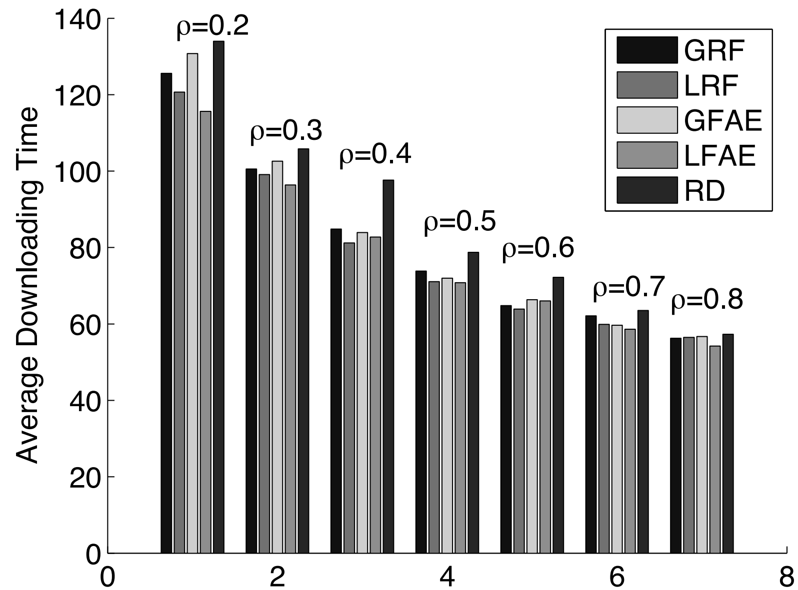
(b) Average downloading time, $B = 4.5$

Experiments

- High bandwidth:



(a) Availability, $B = 12$



(b) Average downloading time, $B = 12$

Conclusion

Conclusion

- **Propose the analytical model to understand BT**

Conclusion

- **Propose the analytical model to understand BT**
  - **On Throughput,**

Conclusion

- **Propose the analytical model to understand BT**
  - **On Throughput,**
  - **On Availability,**

Conclusion

- **Propose the analytical model to understand BT**
  - **On Throughput,**
  - **On Availability,**
  - **Sensitivity analysis on different system parameters.**

Conclusion

- **Propose the analytical model to understand BT**
  - **On Throughput,**
  - **On Availability,**
  - **Sensitivity analysis on different system parameters.**
- **Extend the model to consider peers**

Conclusion (Cont.)

- Validate the analytical result with extensive simulation (our model is more accurate than the Qiu's model)
- Propose new approach on chunk selection algorithm to enhance file availability