Antfarm: Efficient Content Distribution with Managed Swarms

Ryan S. Peterson and Emin Gün Sirer

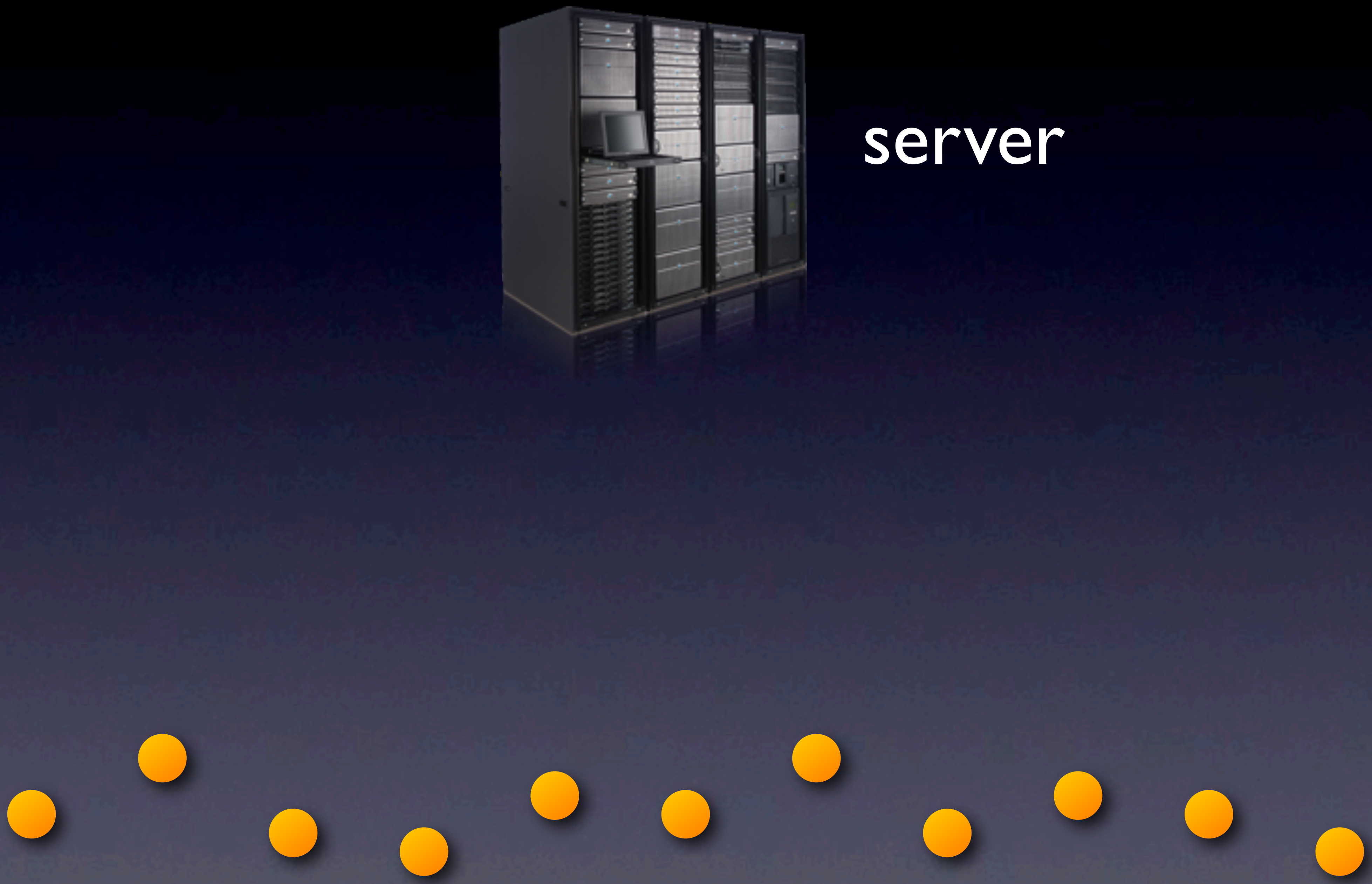
Department of Computer Science, Cornell University
United Networks, LLC

April 22, 2009

Problem Domain

What is the most efficient way to disseminate a large set of files to a large set of clients?

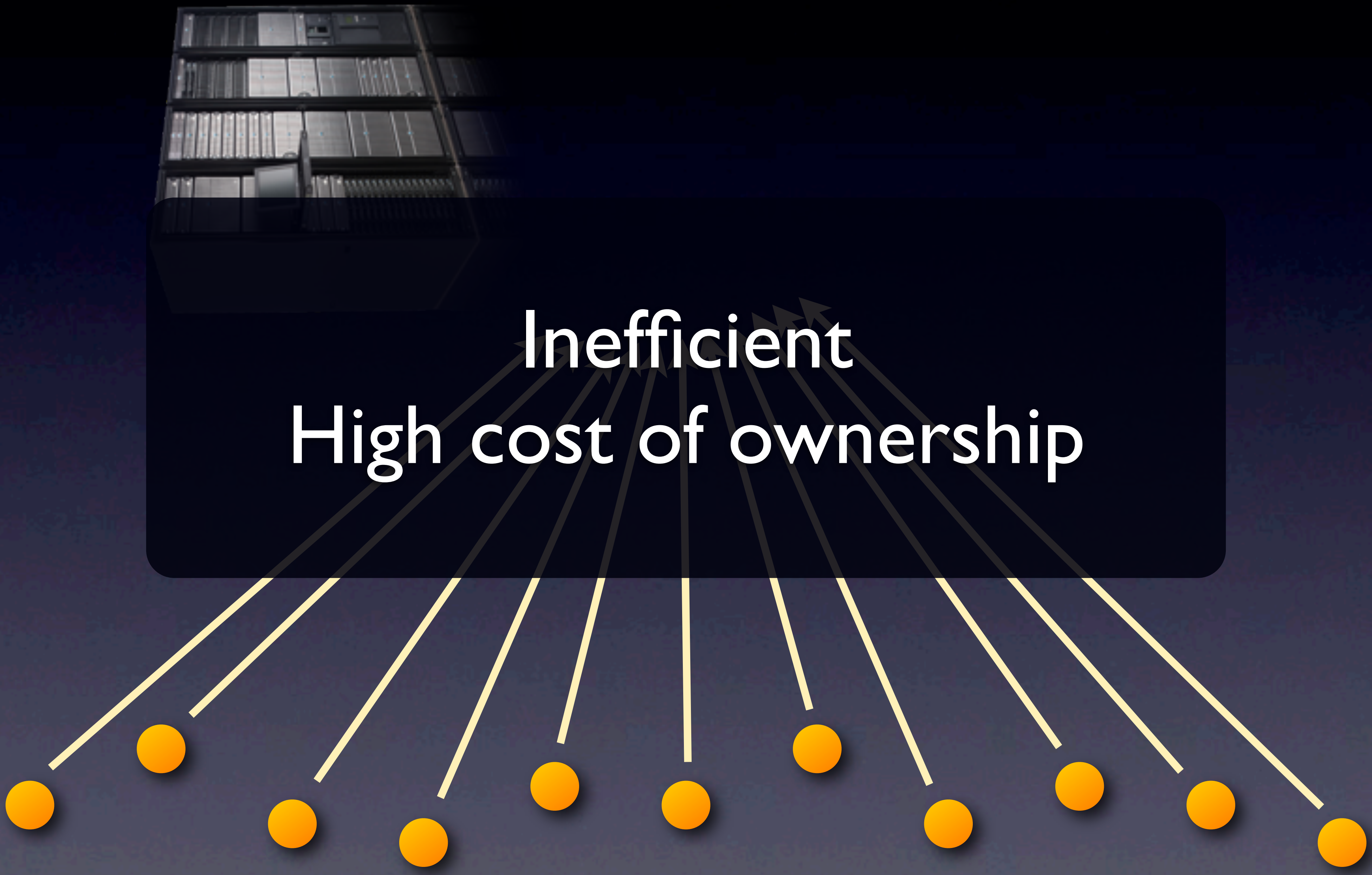
Client-Server



server

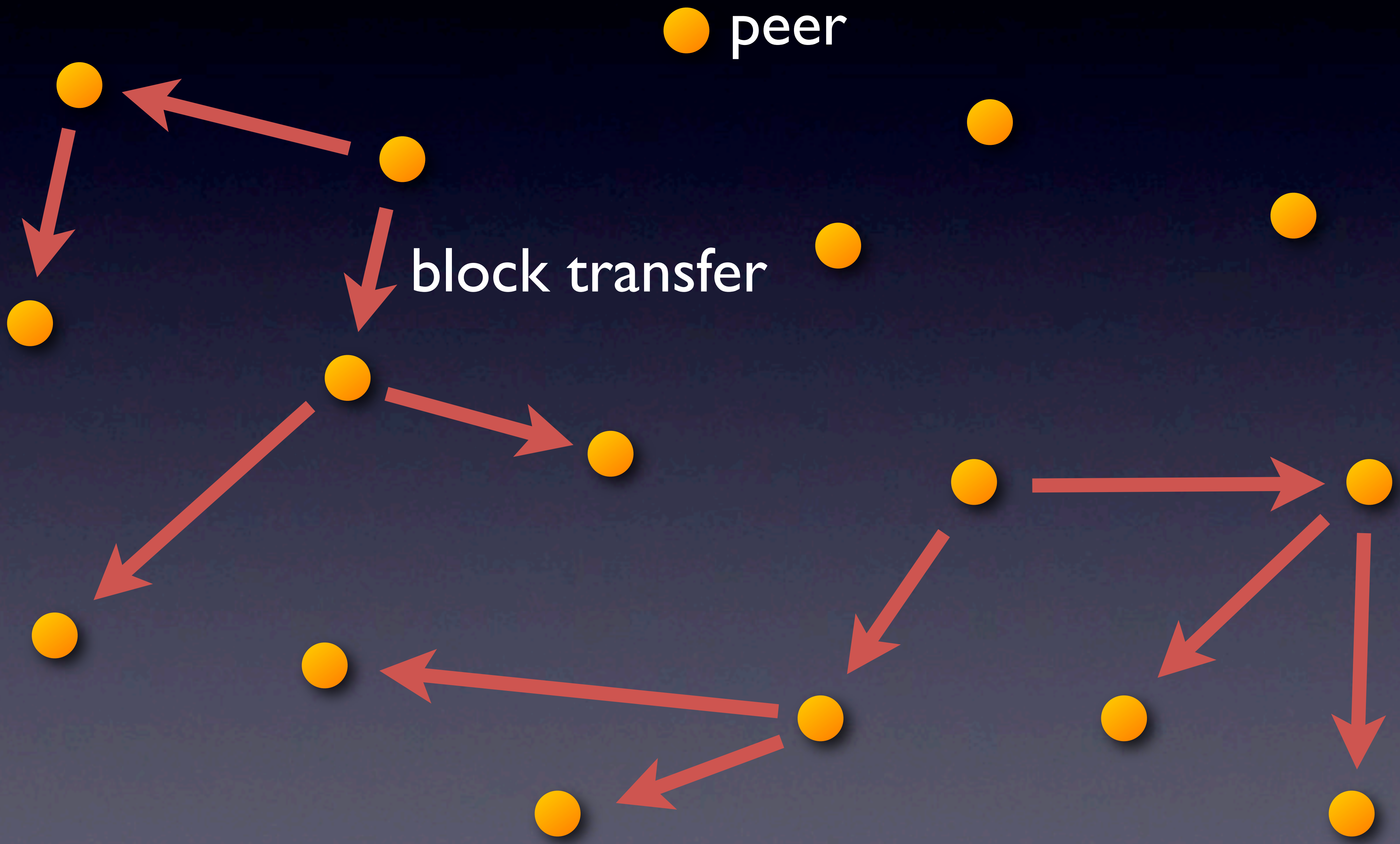
clients

Client-Server

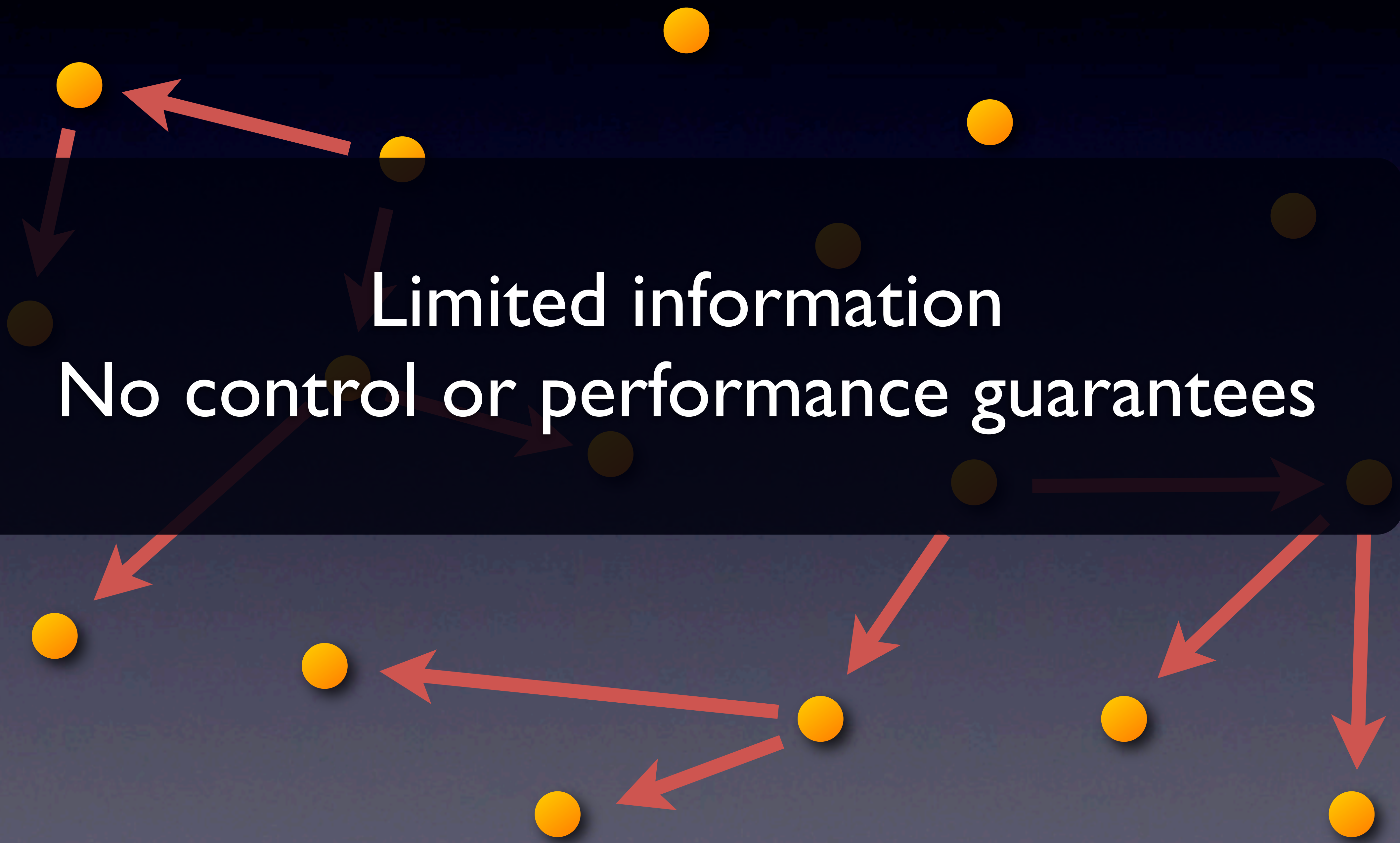


Inefficient
High cost of ownership

Peer-to-Peer



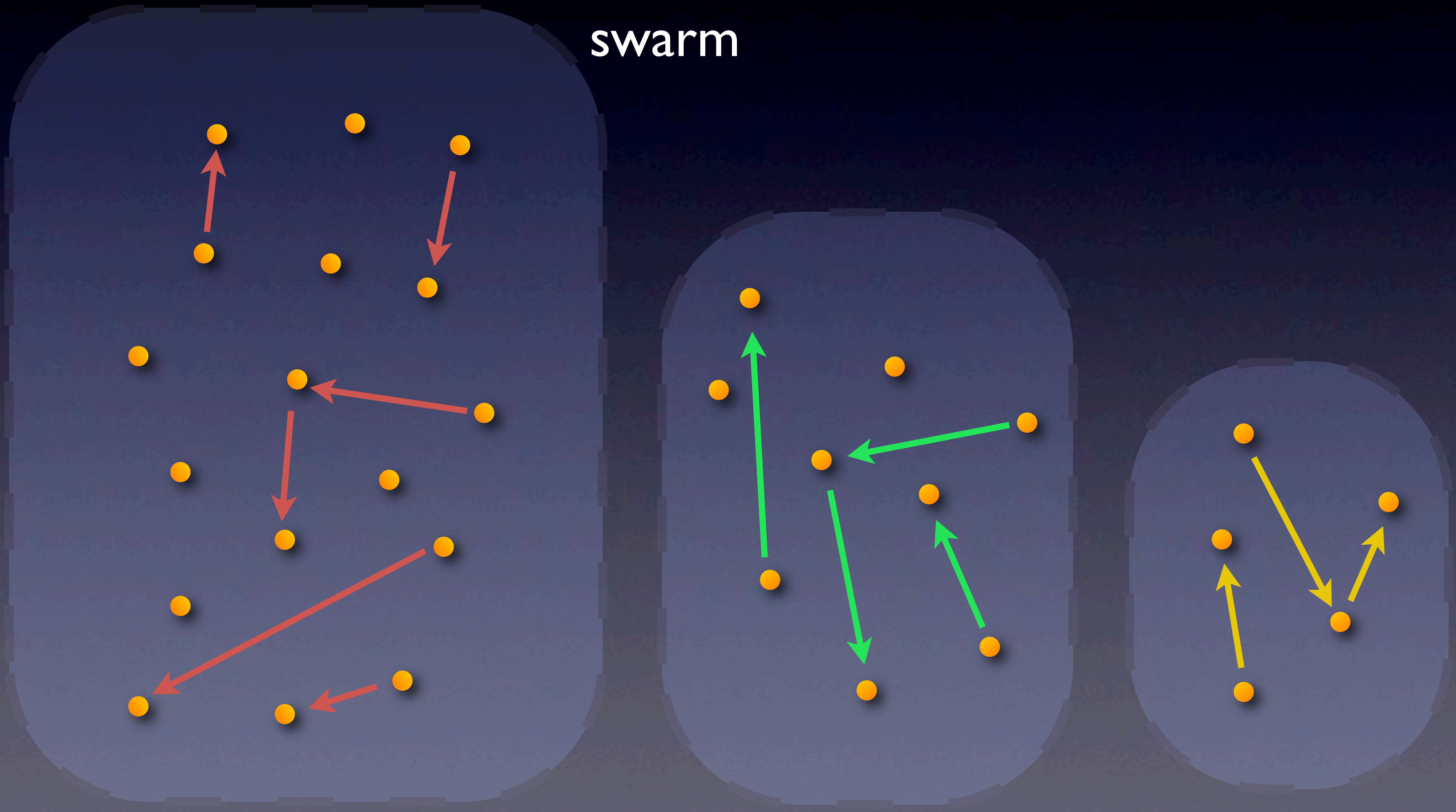
Peer-to-Peer



Limited information

No control or performance guarantees

Peer-to-Peer



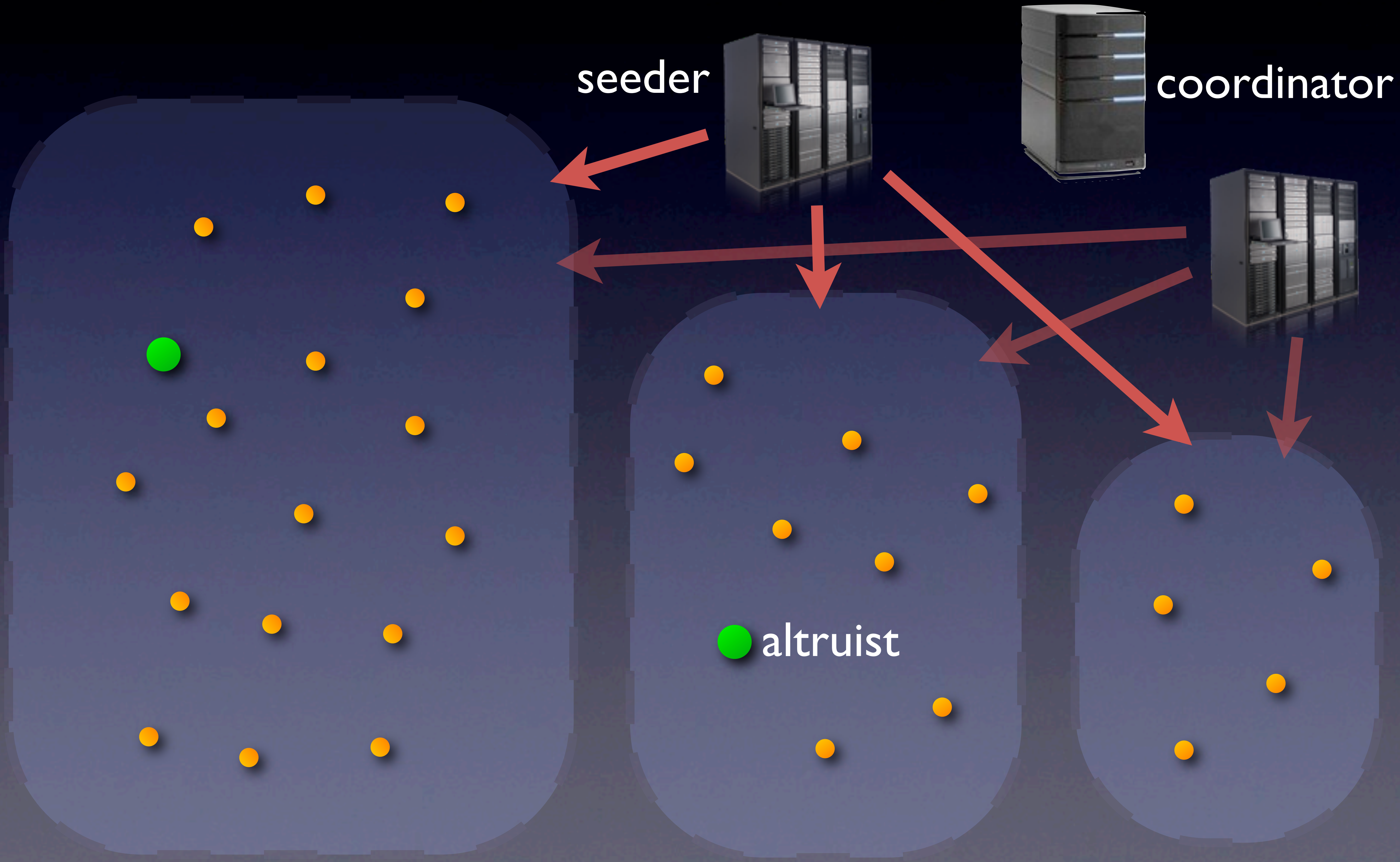
Antfarm Goals

- High performance
- Low cost of deployment
- Performance guarantees
  - Administrator control over swarm performance
- Accounting
  - Enables different resource contribution policies

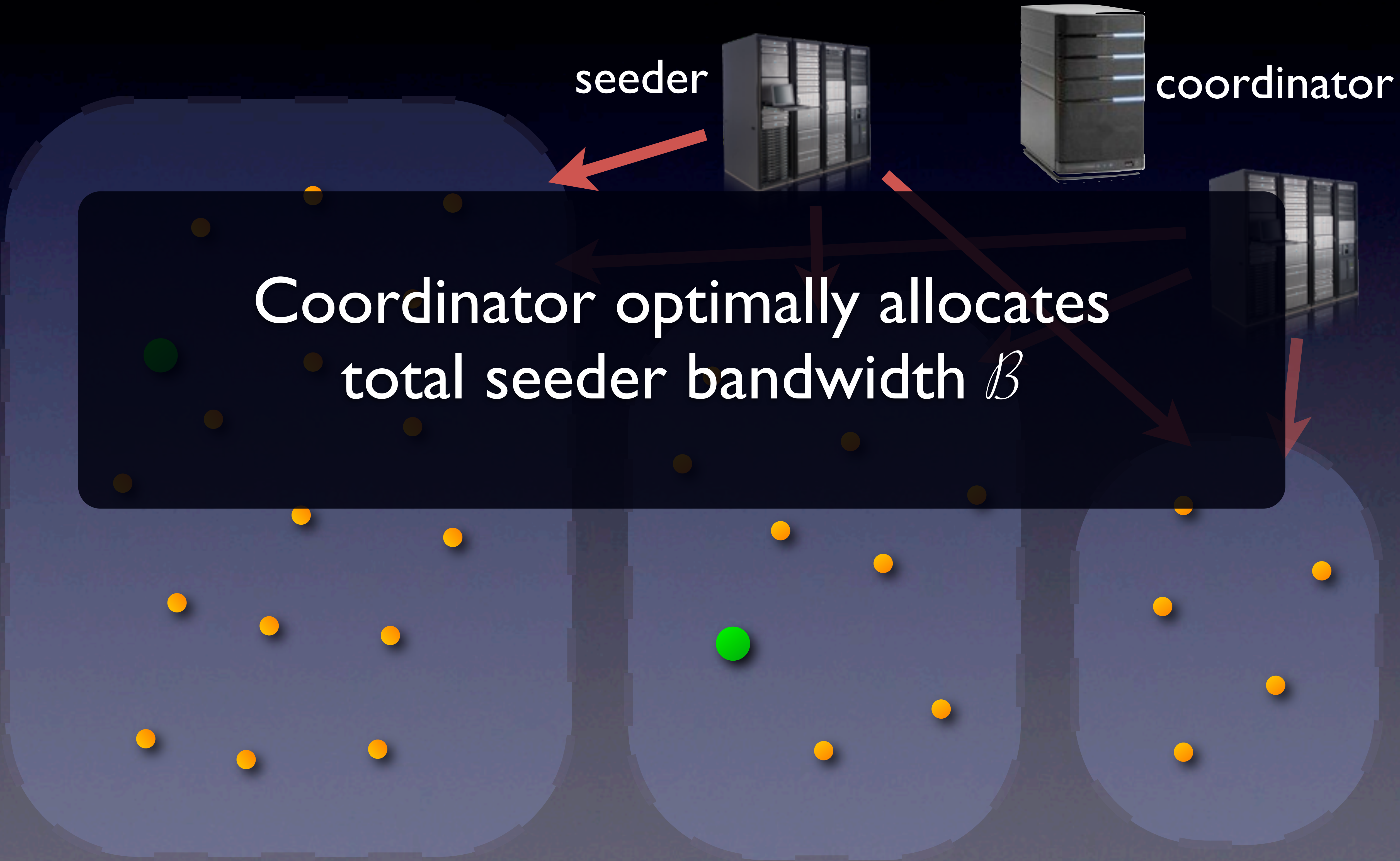
Antfarm Approach

- Key insight: view content distribution as an optimization problem
- Hybrid architecture
  - P2P swarming with a logically centralized coordinator
- Clean slate protocol

Antfarm System Model

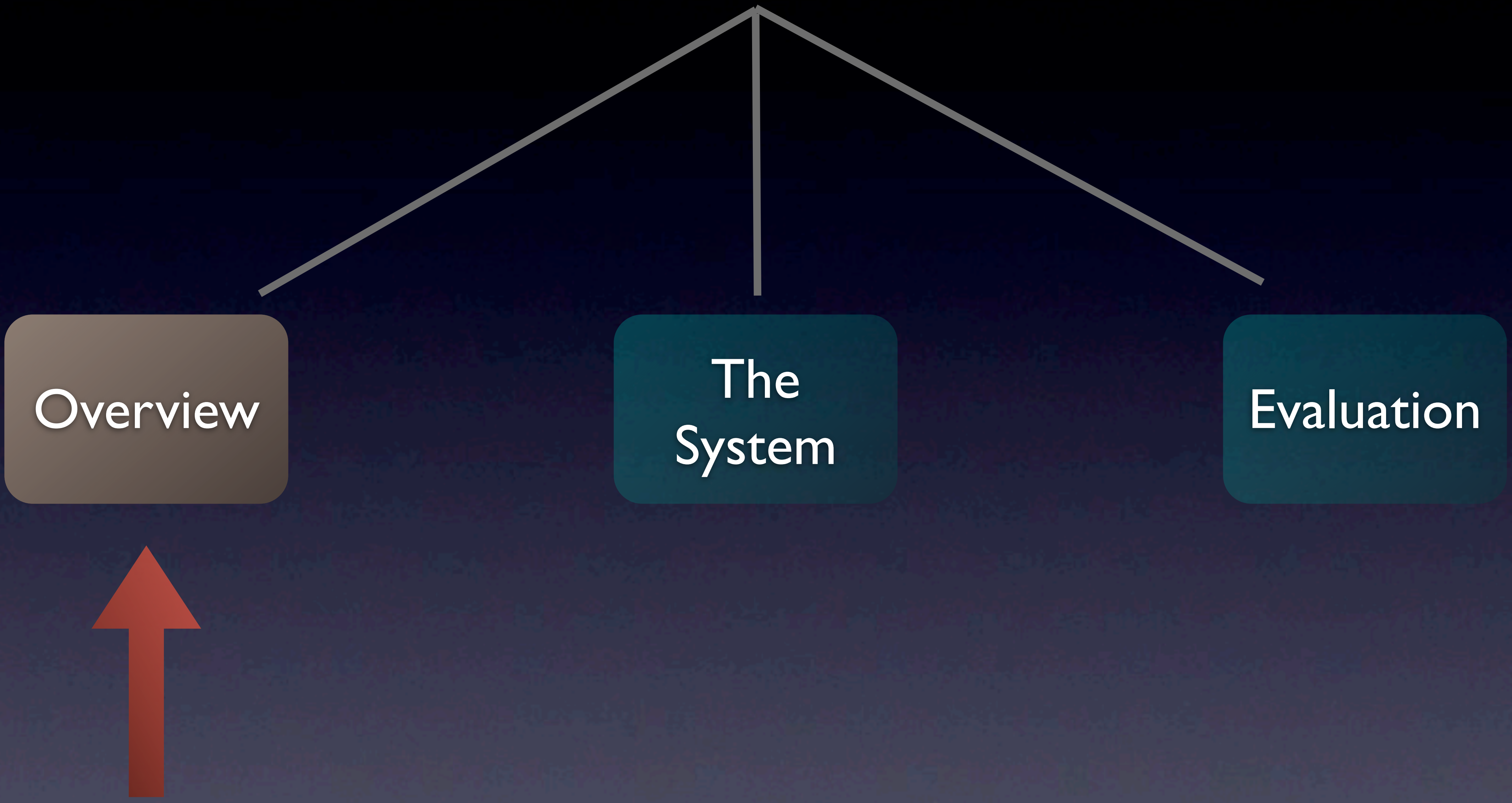


Antfarm System Model



Coordinator optimally allocates total seeder bandwidth $\mathcal{B}$

Antfarm

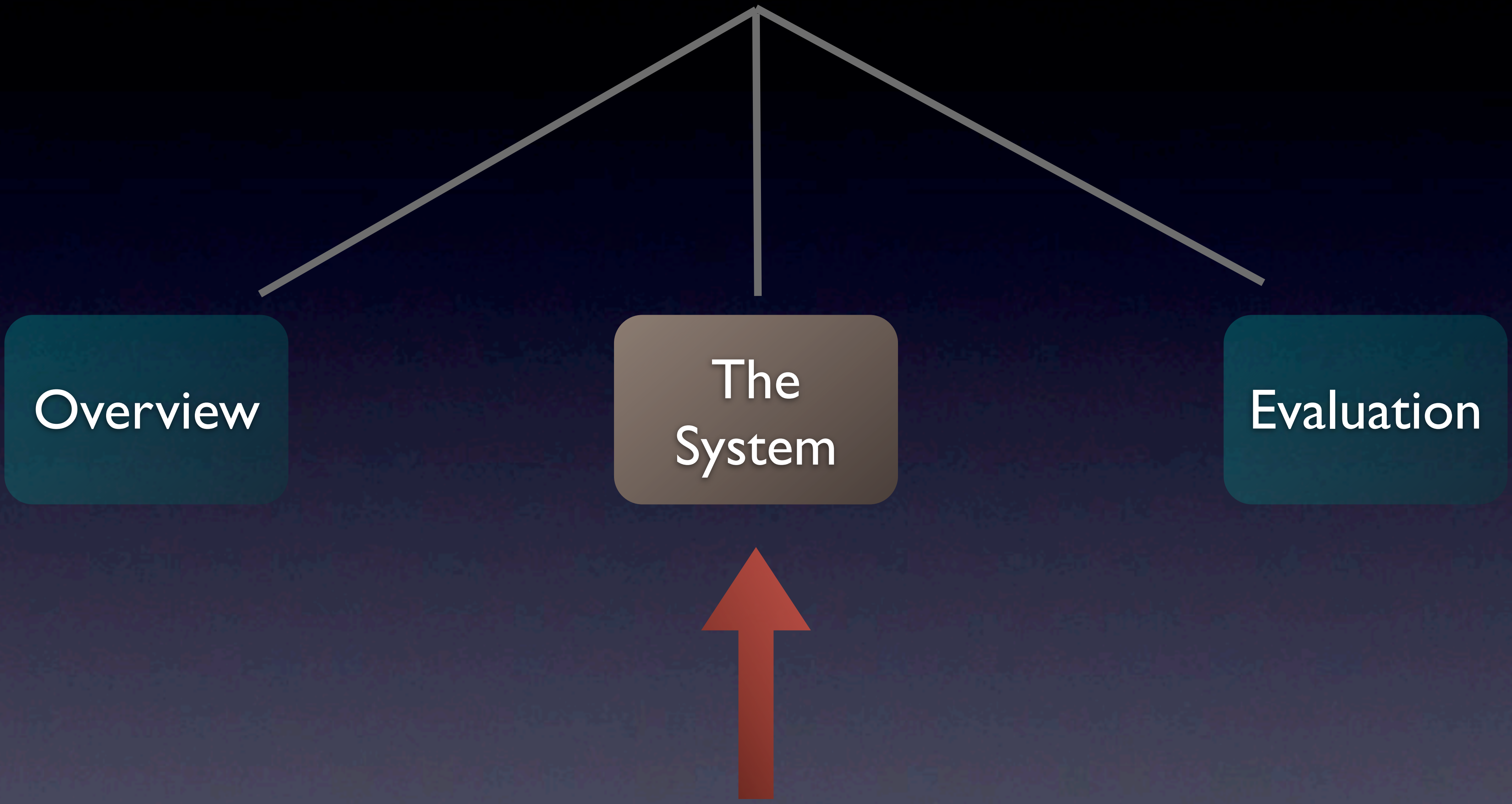


Overview

The System

Evaluation

Antfarm



Overview

The System

Evaluation

Strawman Coordinator

- One could schedule every data transfer in the system
    - All packets for all time
    - Unscalable, impractical!
- Antfarm coordinator makes critical decisions based on observed dynamics

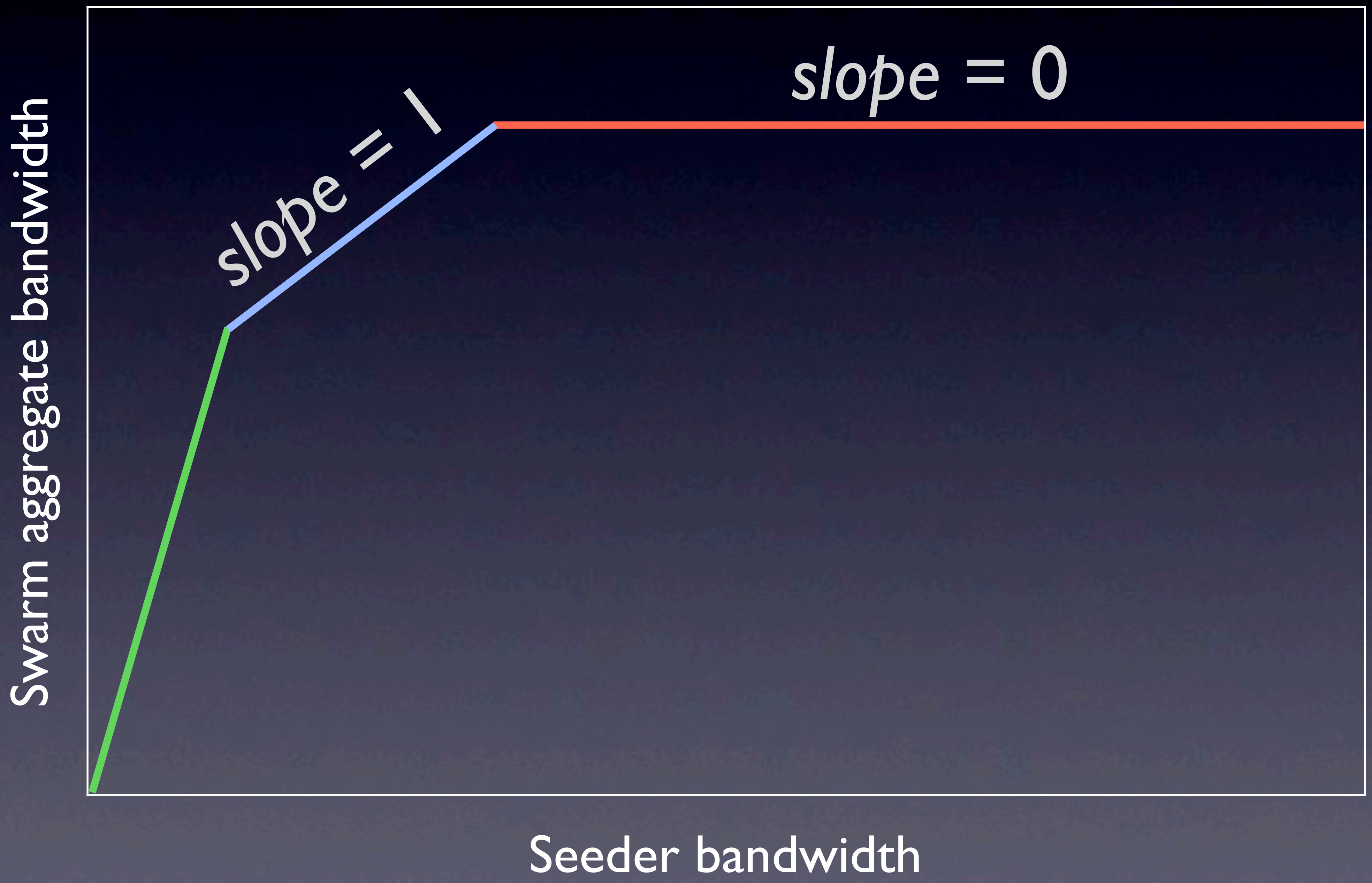
Antfarm Coordinator

- Models swarm dynamics
  - Measures and extracts key parameters
- Formulates optimization problem
  - Calculates optimal bandwidth allocation
- Enacts allocation decisions
  - Maximizes aggregate bandwidth
  - Minimizes average download time

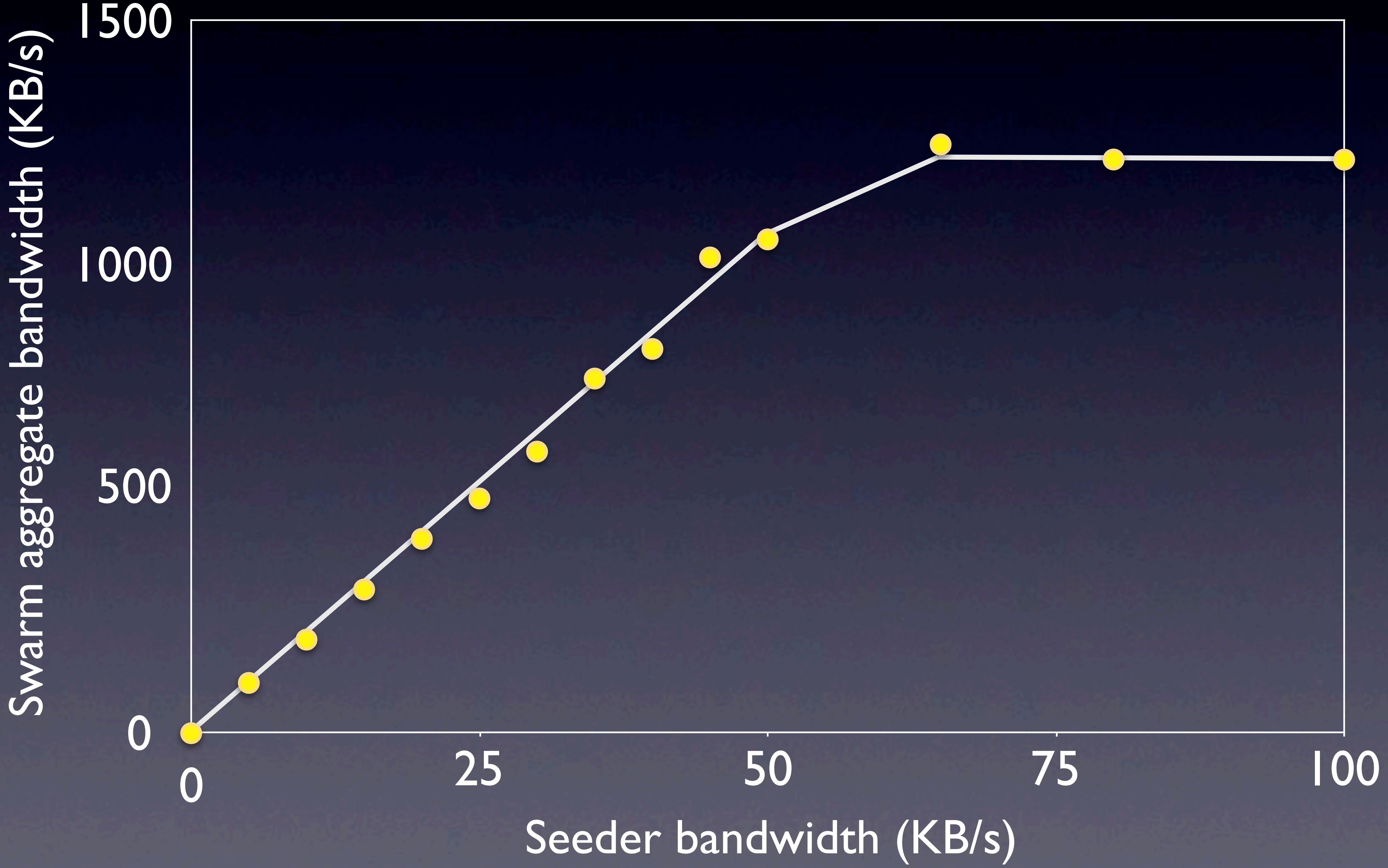
Antfarm Formalization

Maximize system-wide aggregate bandwidth subject to a bandwidth constraint

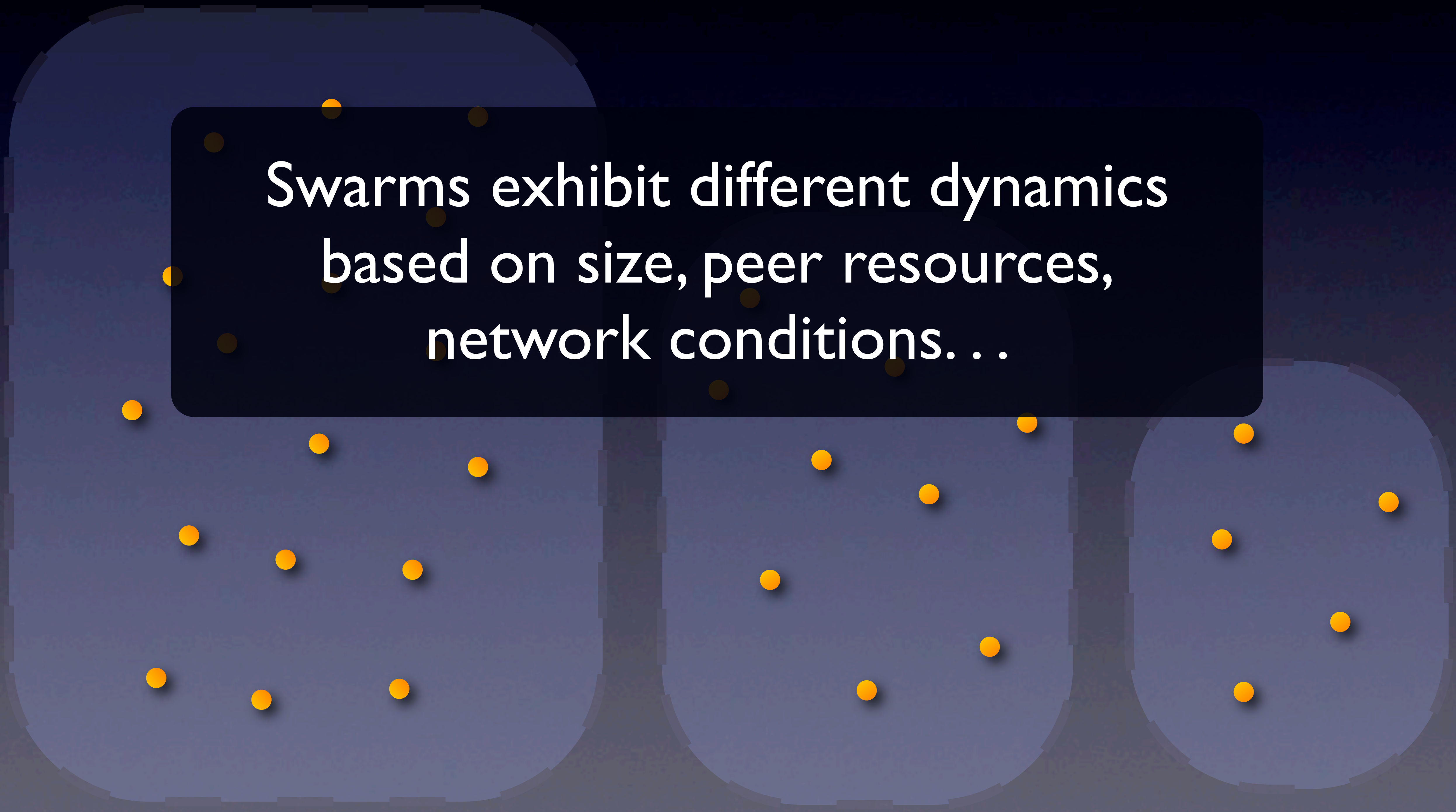
Response Curves



Response Curves

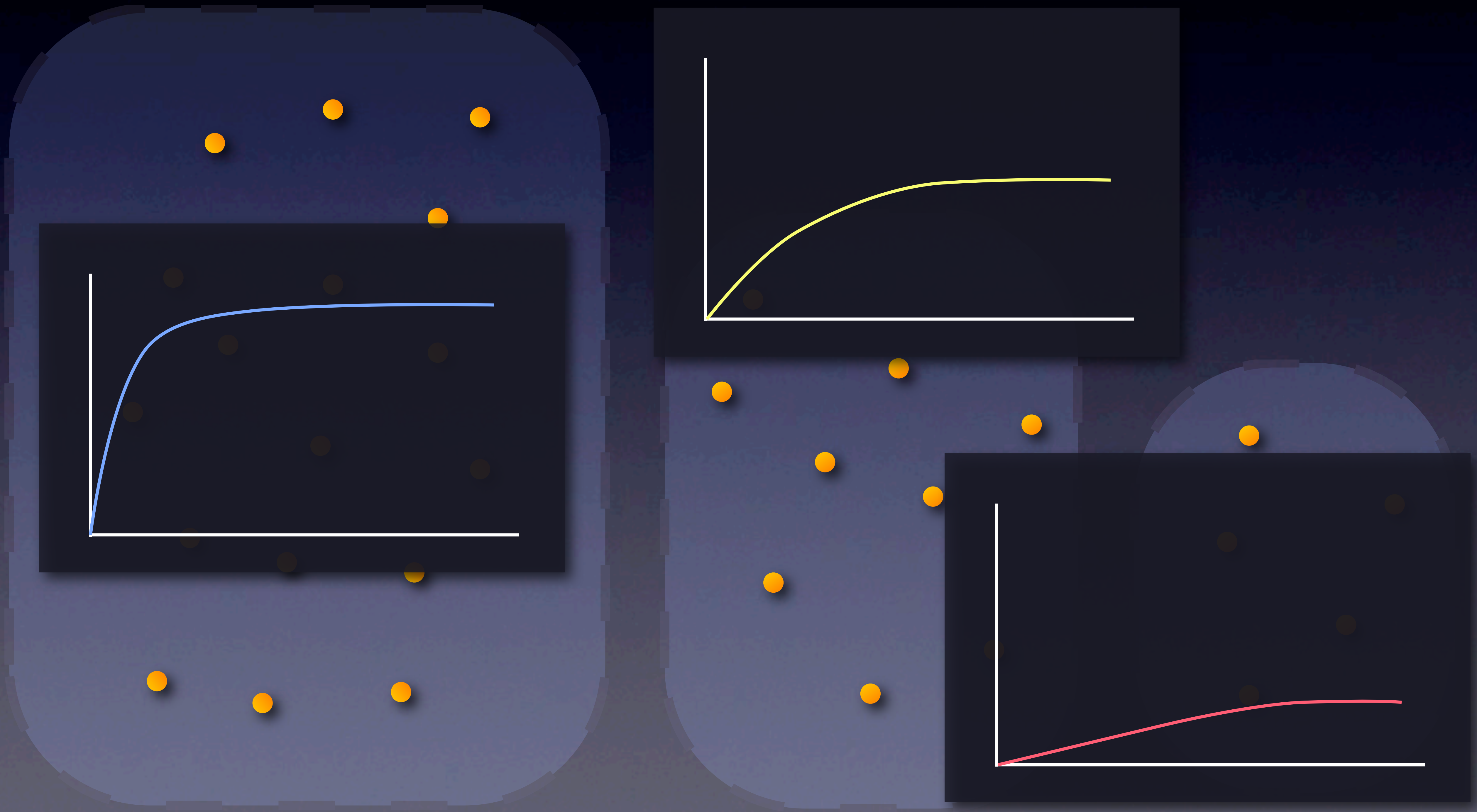


Swarm Dynamics

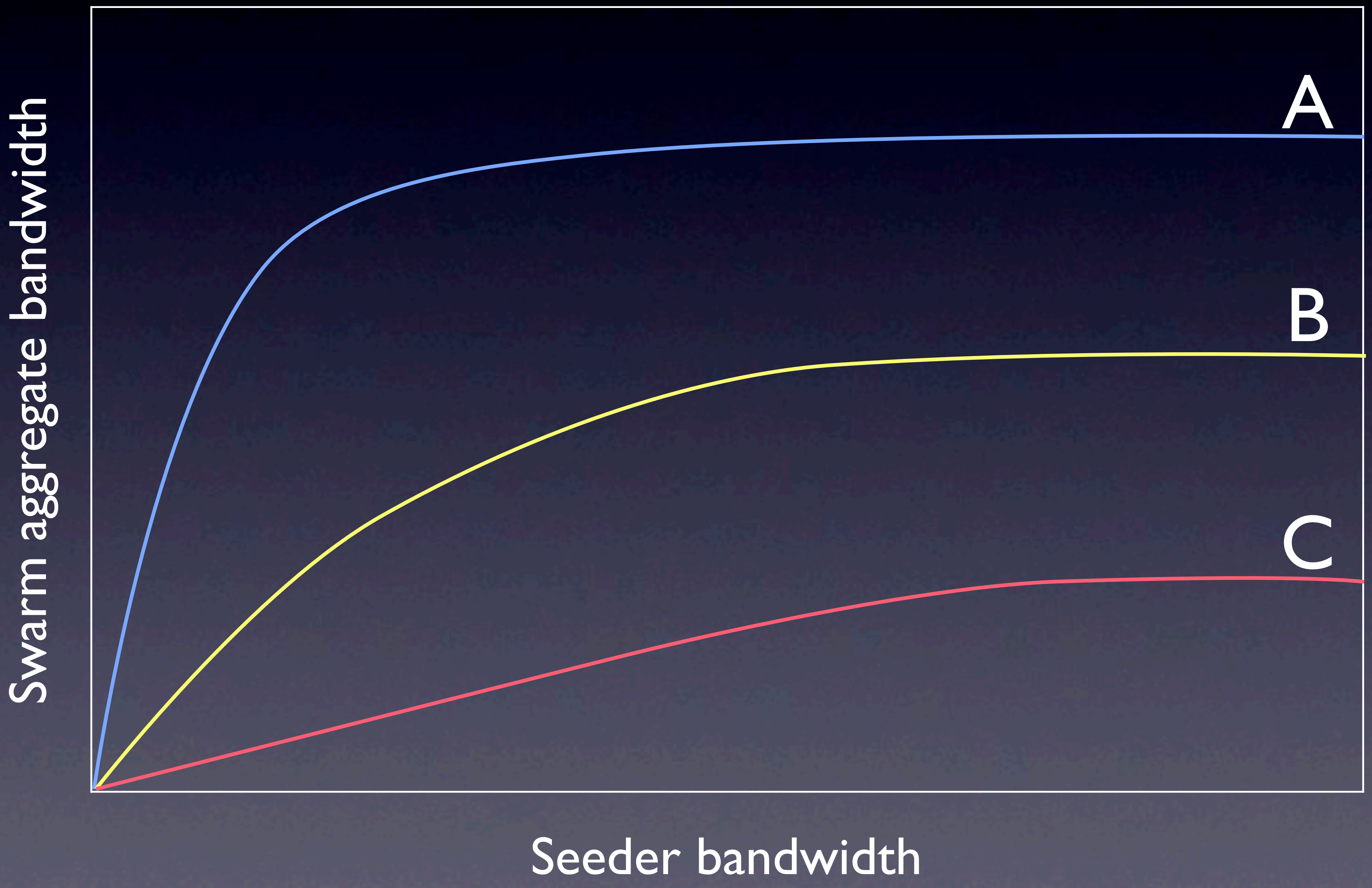


Swarms exhibit different dynamics based on size, peer resources, network conditions. . .

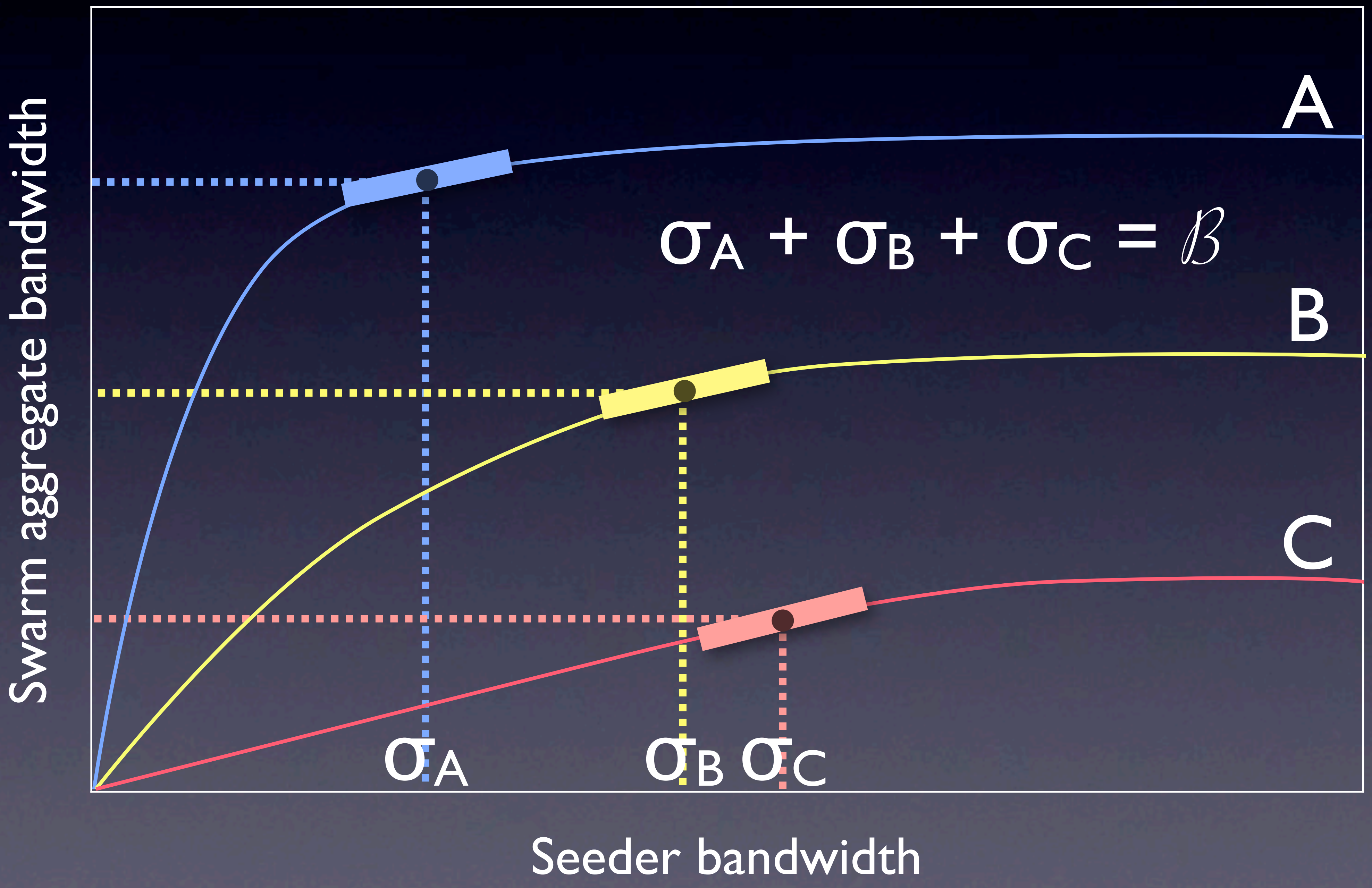
Swarm Dynamics



Antfarm Optimization



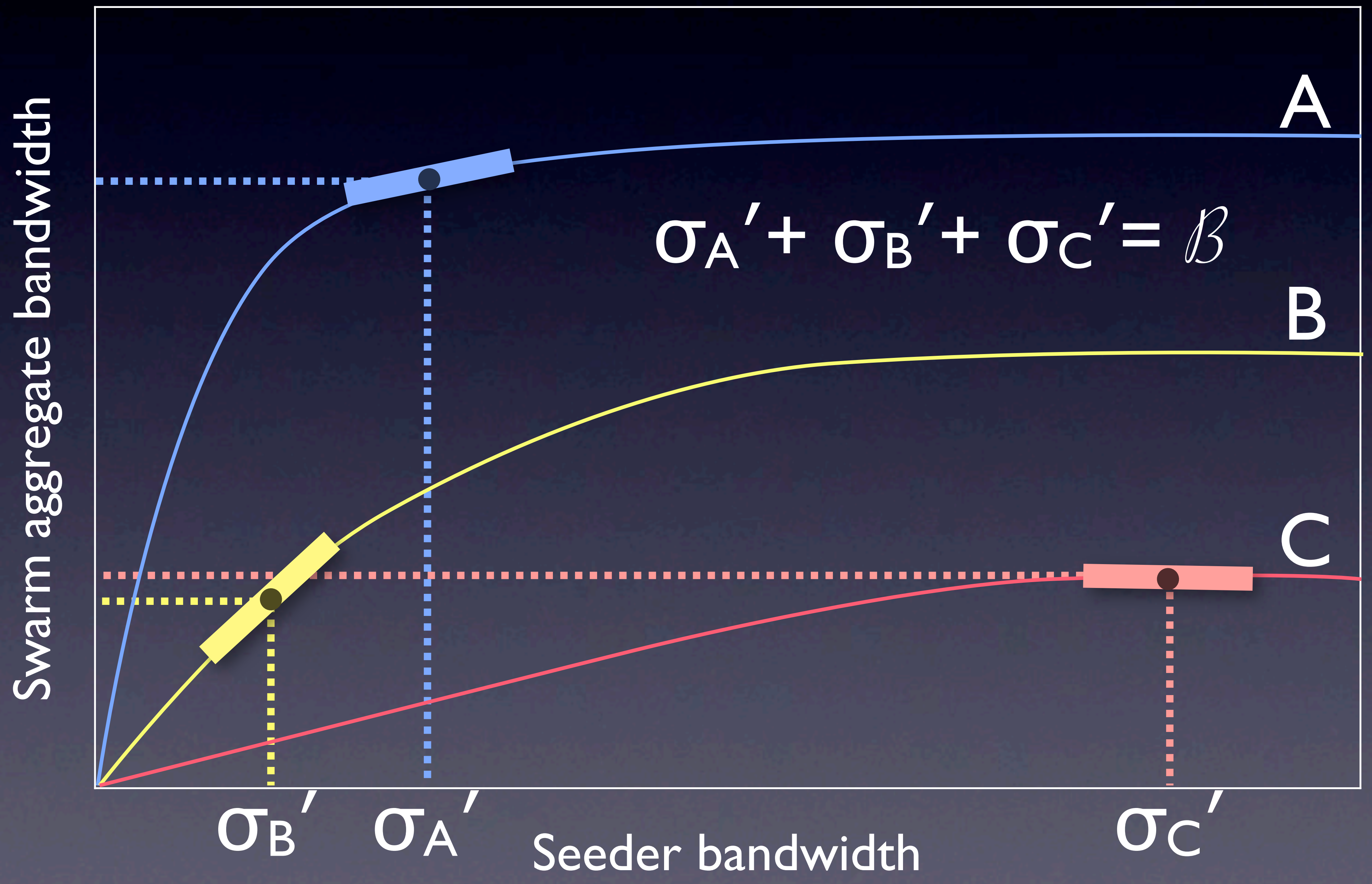
Antfarm Optimization



Performance Control

- Can provide swarm performance guarantees
  - Guarantee minimum level of service
  - Prioritize swarms

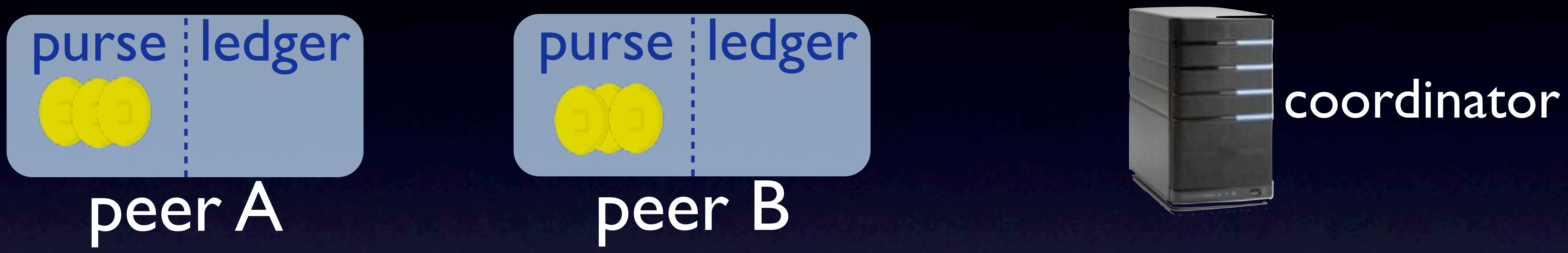
Antfarm Allocation



Adapting to Change

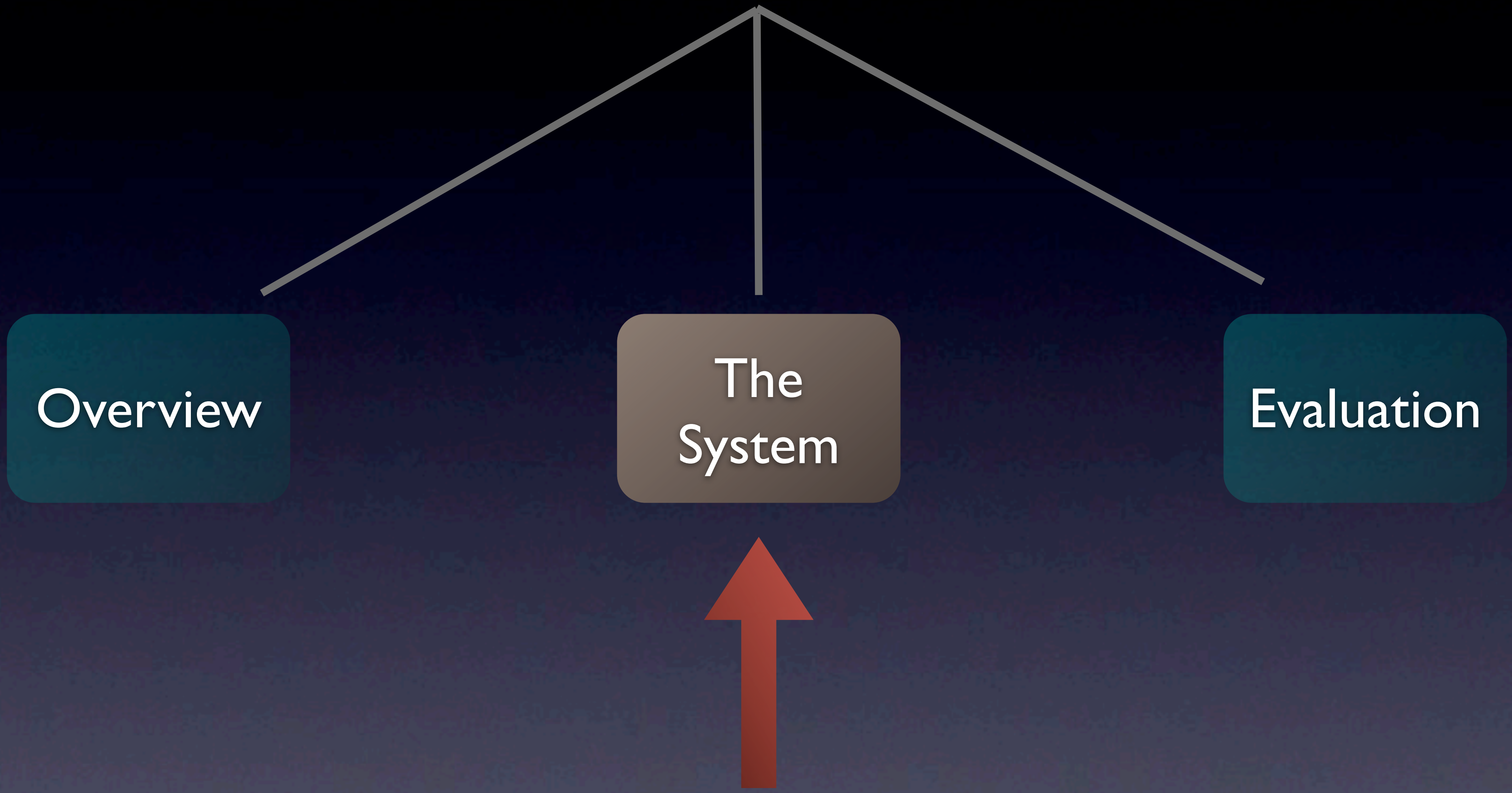
- Swarm dynamics change
  - Churn
  - Network conditions
- Antfarm updates response curves
  - Coordinator explores around point of operation

Wire Protocol



- Coordinator mints small, unforgeable tokens
- Peers trade each other tokens for blocks
- Peers return spent tokens to the coordinator as proof of contribution

Antfarm

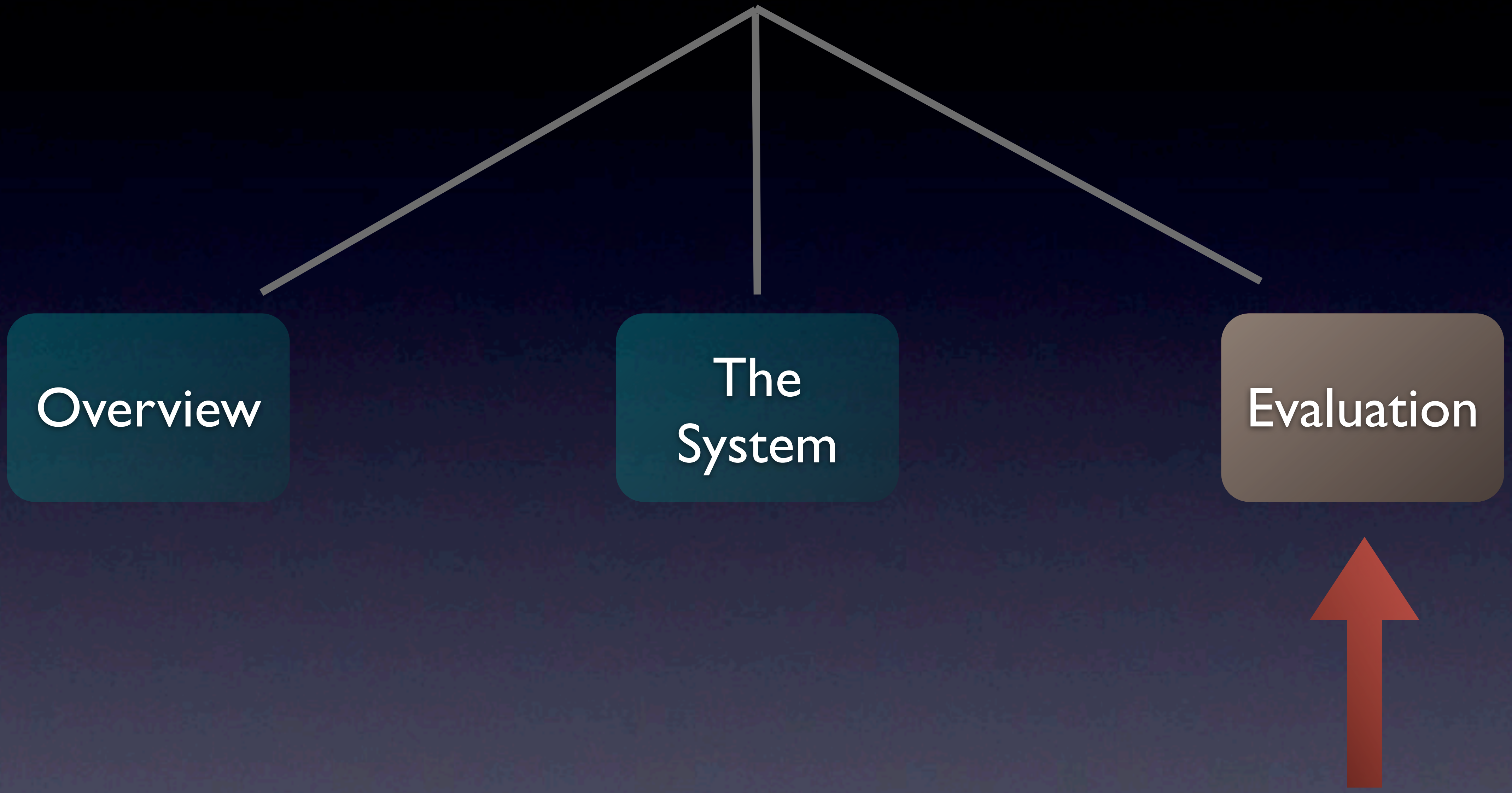


Overview

The System

Evaluation

Antfarm

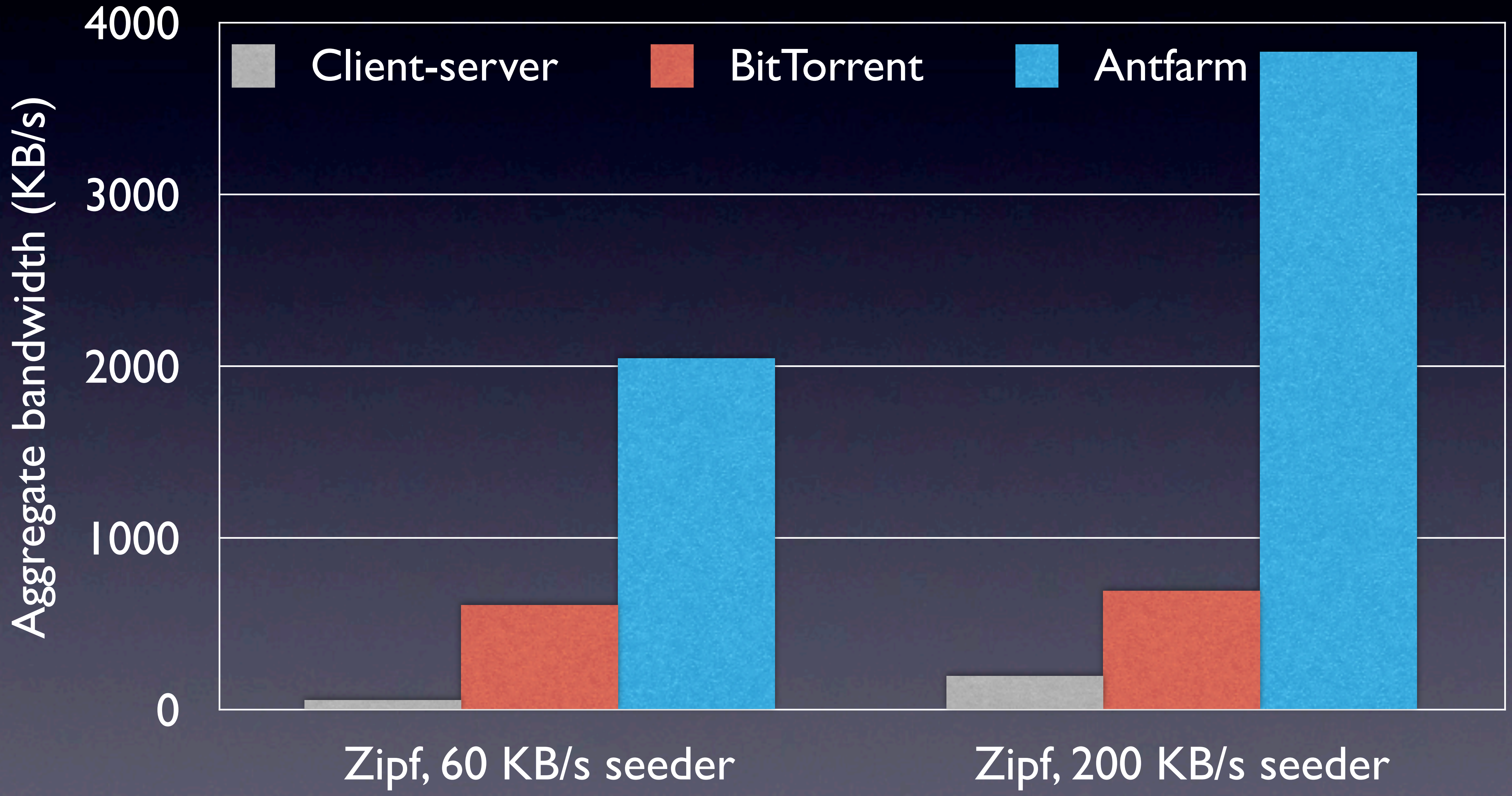


Overview

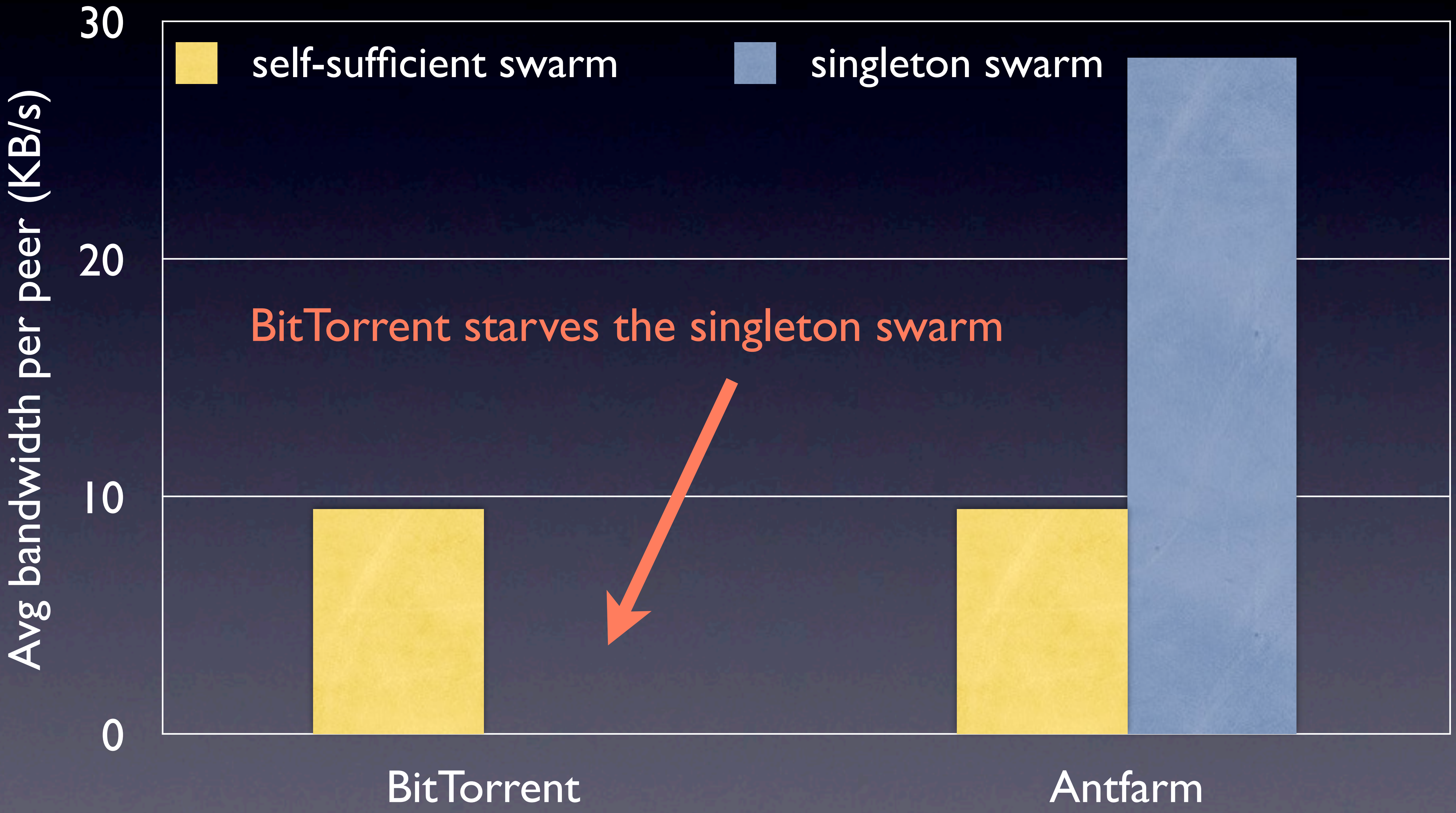
The System

Evaluation

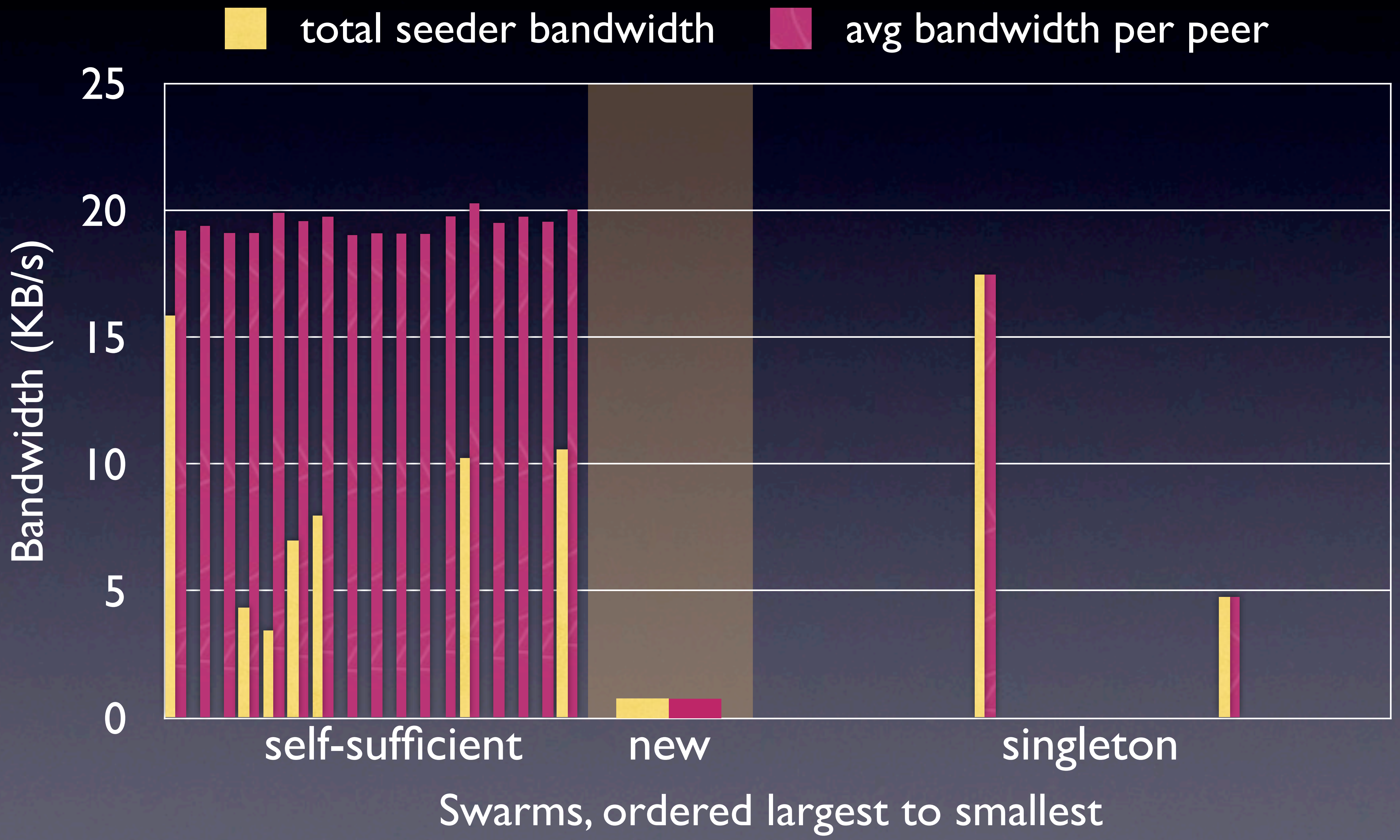
Antfarm Performance



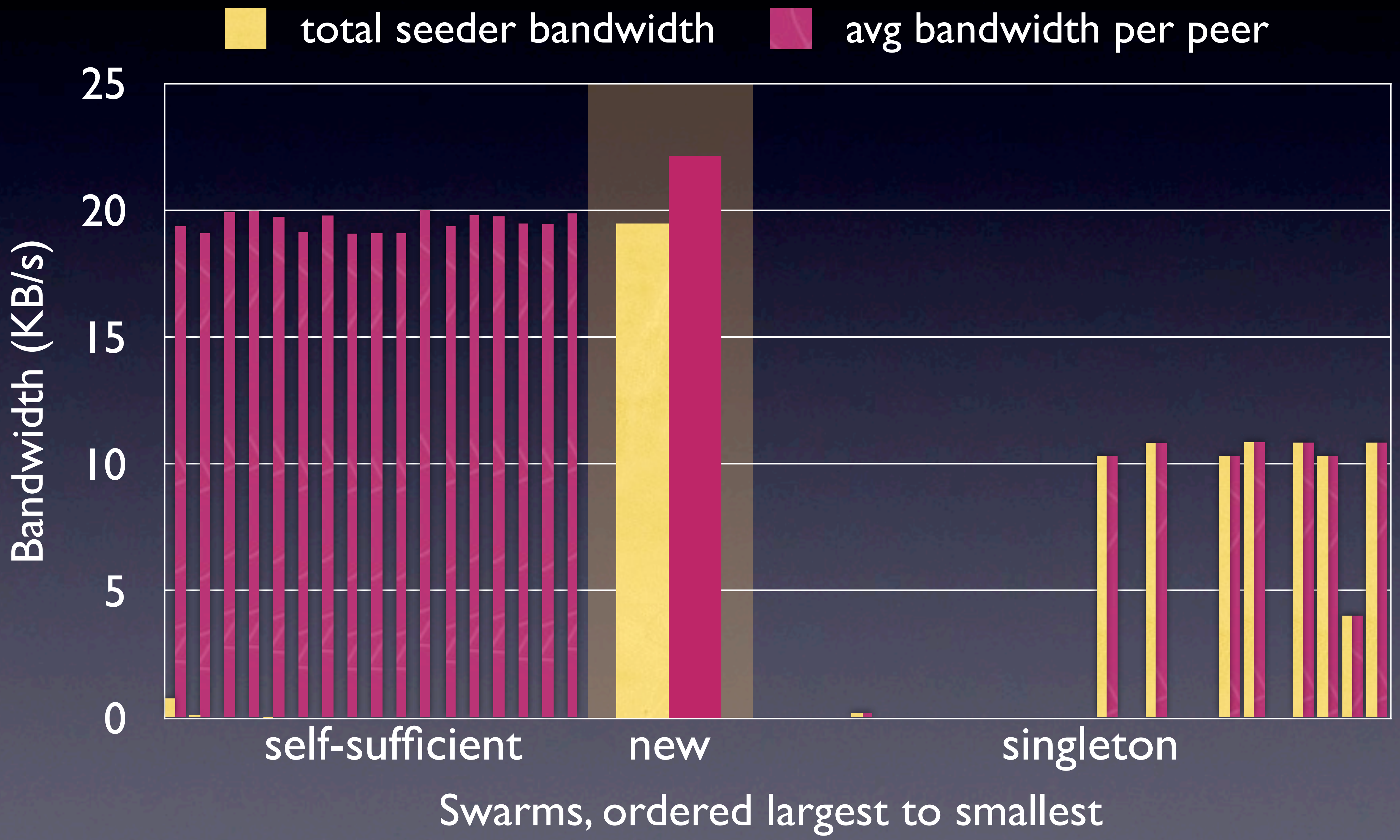
Swarm Starvation



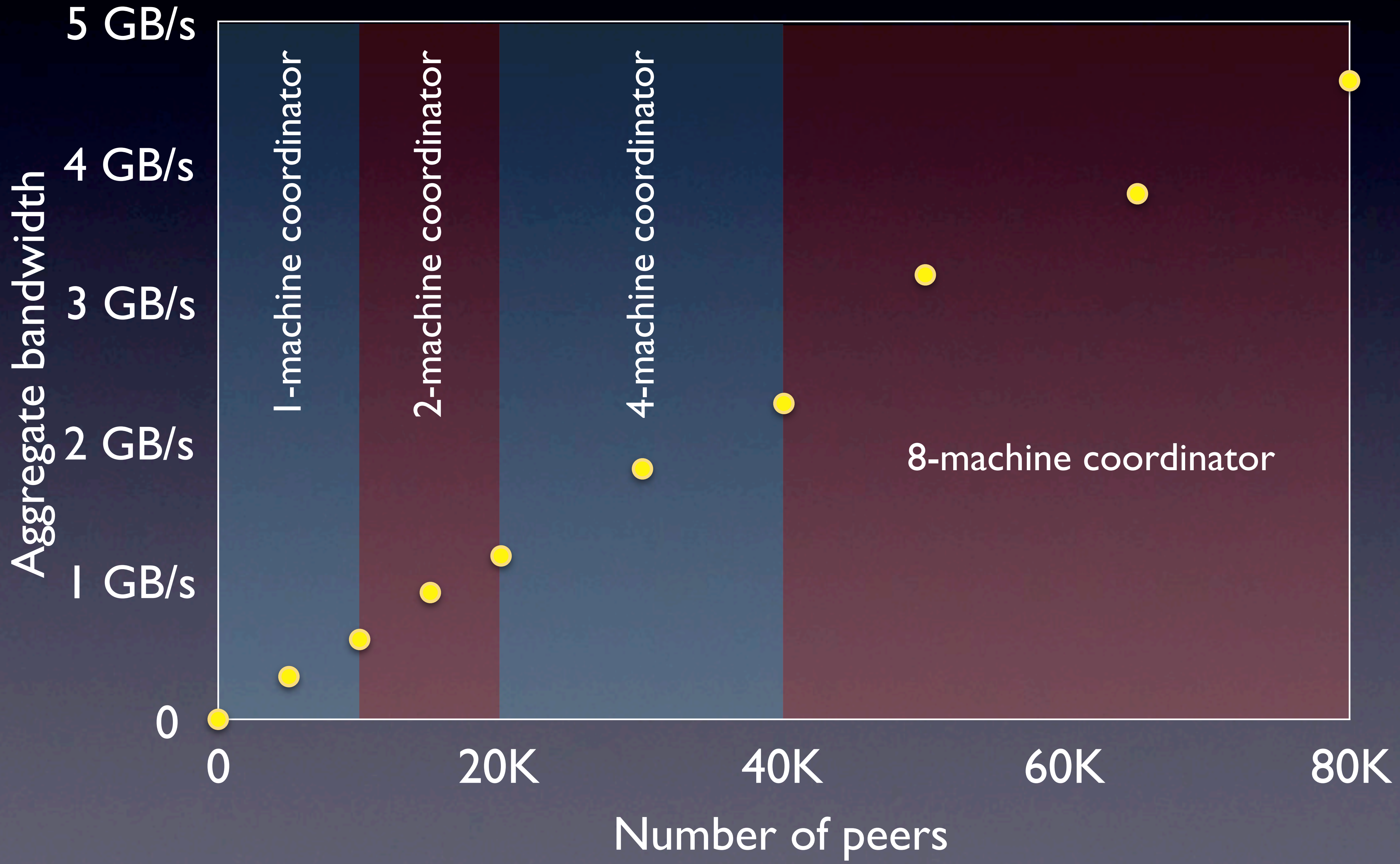
BitTorrent: Starves New Swarm



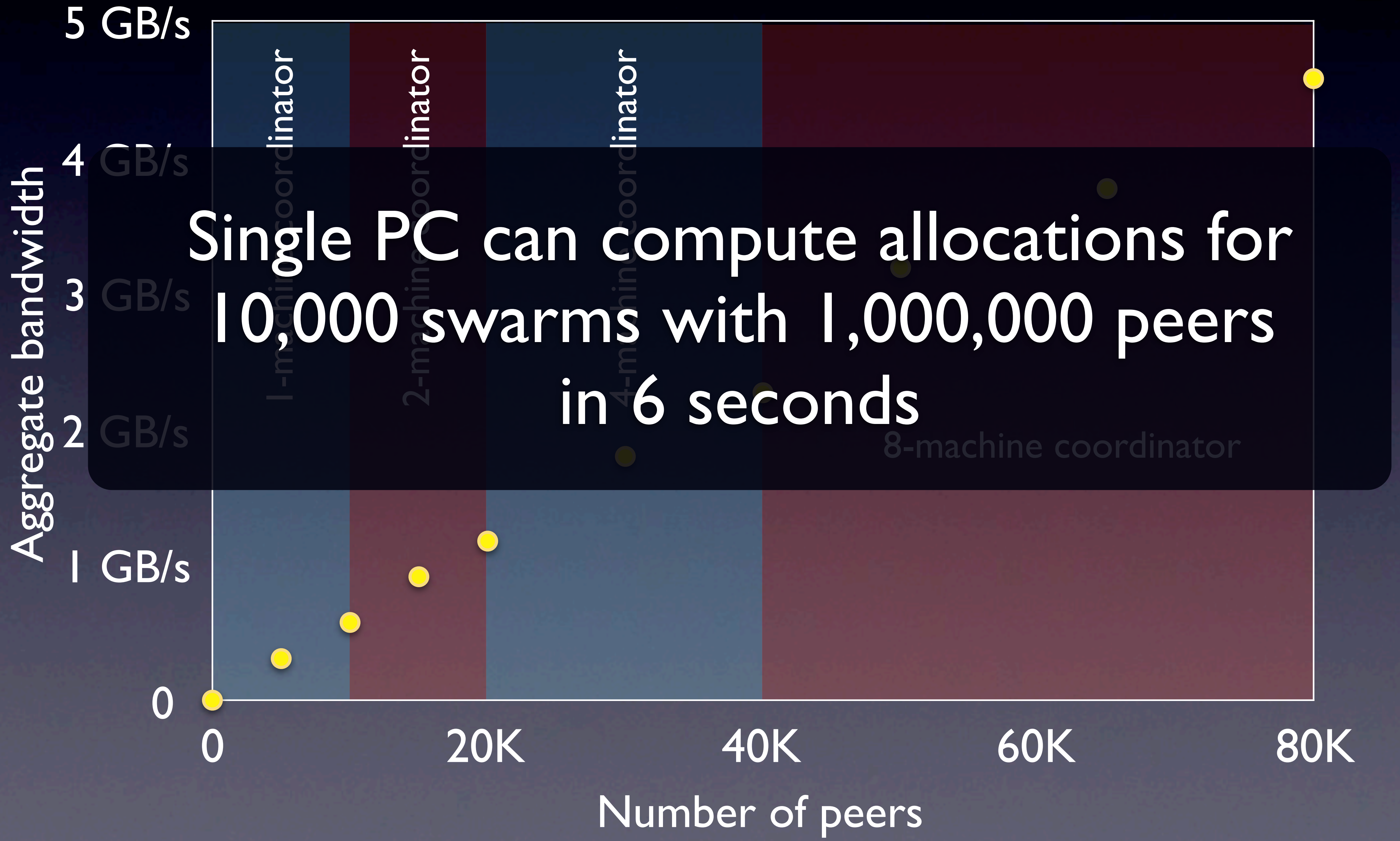
Antfarm: Seeds New Swarm



Scalability



Scalability



Single PC can compute allocations for 10,000 swarms with 1,000,000 peers in 6 seconds

Antfarm Implications

- No fine-tuning
- Subsumes hacks devised for BitTorrent
  - Share ratio
  - Manual pruning

Related Work

- Content Distribution Networks
    - Akamai, CoBlitz, CoDeeN, ECHOS, Coral, Slurpie, YouTube, Hulu, GridCast, Tribler, Joost, Huang et al. 2008, ...
- P2P Swarming
    - BitTorrent, BitTyrant, PropShare, BitTornado, BASS, Annapureddy et al. 2007, Guo et al. 2005, ...
- Incentives and microcurrencies
    - Dandelion, BAR Gossip, Samsara, Karma, SHARP, PPay, Kash et al. 2007, ...

Conclusions

- Model swarm dynamics and allocate bandwidth optimally
- Novel hybrid architecture
- PlanetLab deployment shows that Antfarm outperforms client-server and P2P

Questions?