

Distributed Data Management

Please make sure you can explain properly the following:

1. What is an overlay network?
2. What are the disadvantages of centralized P2P?
3. Why do we need TTL for flooding?
4. Should the TTL be rather high or rather low?
5. Does Gnutella 0.4 scales well? Why?
6. Can a pure P2P network function without ping pong? If no, why not? If yes, what are the disadvantages?
7. Why are low-bandwidth nodes a problem for pure P2P? Is the same true for hybrid P2P?
8. Why are routing tables useful?
9. How querying in a Hybrid P2P network is done when the system aims at preserving anonymity.