CSC2231: Security in P2Ps

http://www.cs.toronto.edu/~stefan/courses/csc2231/05au

Stefan Saroiu

Department of Computer Science

University of Toronto

Outline

- **Power-law networks and flat overlays**
- **Sybil attack**
- **P2P routing attacks**
- **Discussion**

Outline

- **Power-law networks and flat overlays**
- **Sybil attack**
- **P2P routing attacks**
- **Discussion**

Power-Law Networks are here to Stay

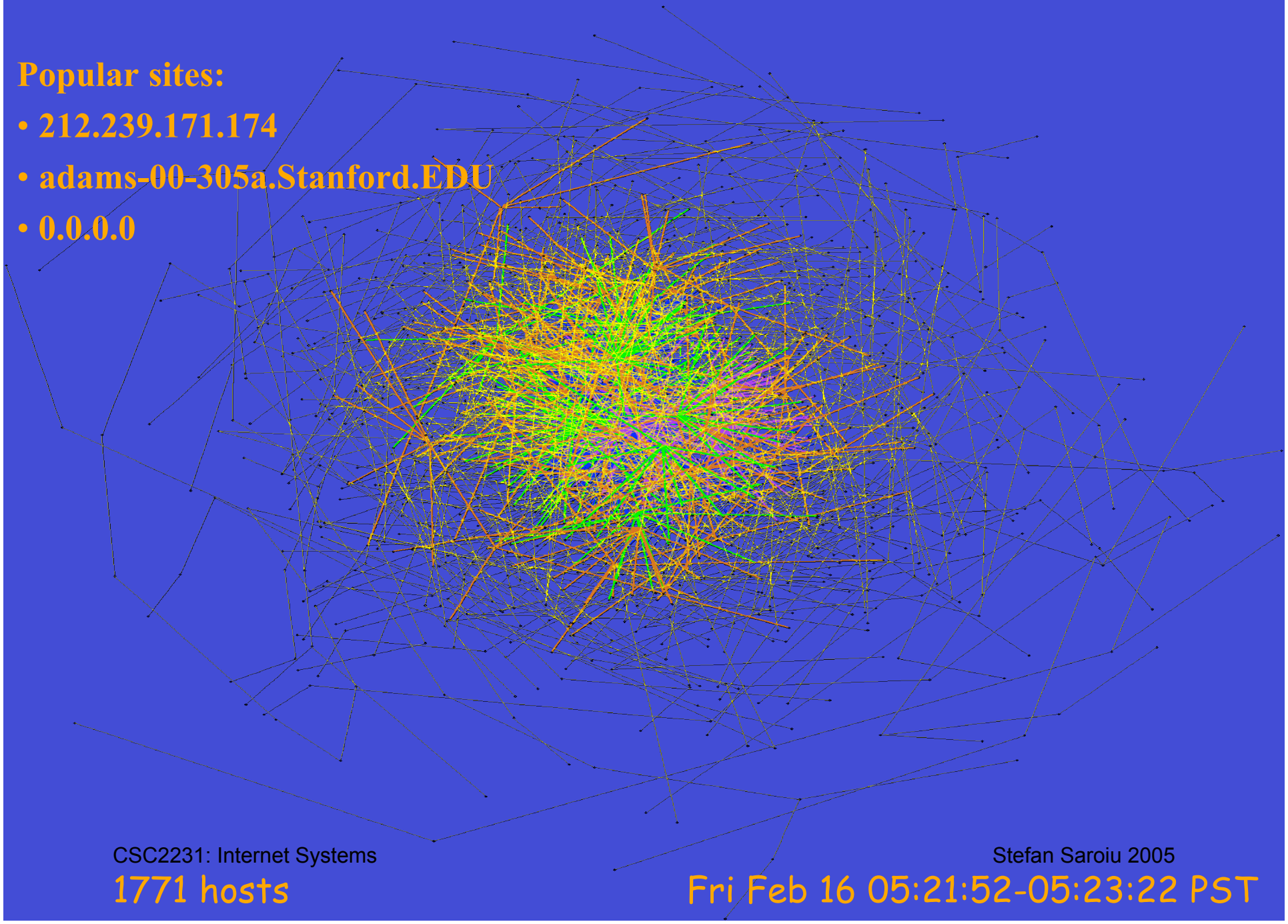
- **Barabasi and Albert showed that networks which…**
  - grow by continuous addition of new nodes
  - exhibit preferential attachment (likelihood of connecting to a node depends on the node's degree)
- **…power-law distribution of vertex degree**
- **Internet, WWW, Gnutella**

Resilience to Failures

- **Power-law networks (Cohen et al.):**
  - very resilient in face of random node failures
    - a giant spanning cluster still exists
  - fairly resilient in face of cascading failures
  - very vulnerable in face of orchestrated attacks (towards high-degree nodes)

Popular sites:

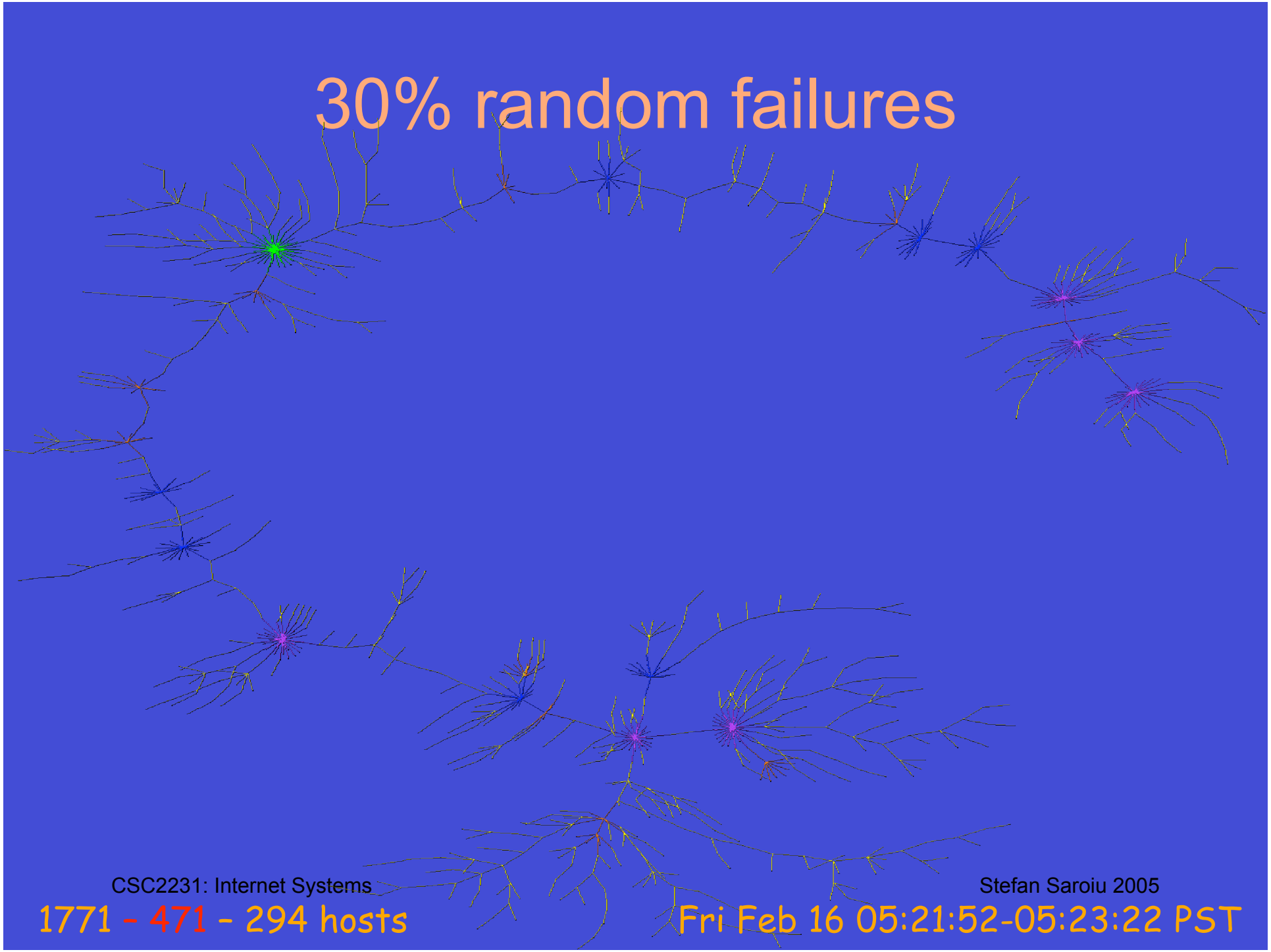
- **212.239.171.174**
- **adams-00-305a.Stanford.EDU**
- **0.0.0.0**



1771 hosts

Fri Feb 16 05:21:52-05:23:22 PST

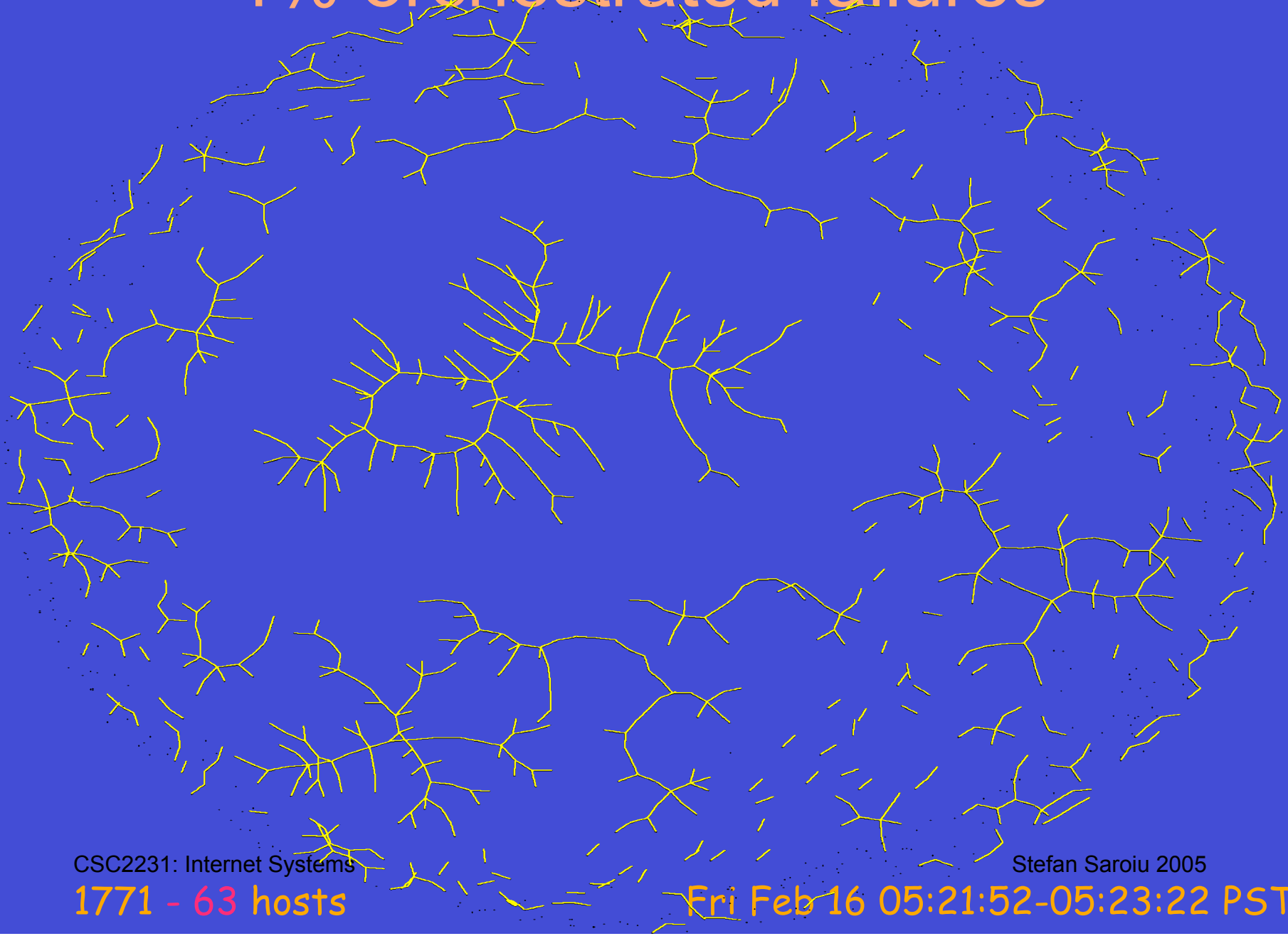
30% random failures



1771 - 471 - 294 hosts

Fri Feb 16 05:21:52-05:23:22 PST

4% orchestrated failures



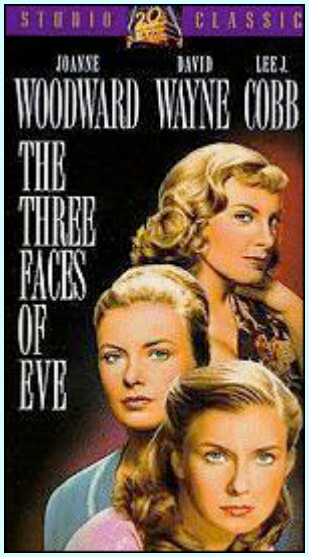
1771 - 63 hosts Fri Feb 16 05:21:52-05:23:22 PST

Outline

- Power-law networks and flat overlays
- **Sybil attack**
- P2P routing attacks
- Discussion

Fundamental problems for P2P

- **One can have, some claim, as many electronic personas as one has the time and energy to create. – *Judith S. Donath.***
- **"Sybil attack…."**
  - Correct assumption: attacker has access to infinite resources
  - It registers *many times* using different identities
  - Cannot distinguish whether these are real or not. Real clients get poor service
  - Impossible to get around in a fully decentralized manner



Identities and Entities

- **Ideally want one to one mapping**
  - Virtual servers idea was a many-to-one
- **How to convince that two different identities correspond to two different entities**
  - Perform a task that no single entity can perform
  - Have others vouch that these are two different entities

Observations on performing tasks

- **The # of identities one entity can create is proportional to the ratio of an entity's resources to the resources of the weakest peer**
  - Can use computational resources (i.e., solve a puzzle)
  - Can use communication resources (i.e., use large packets)
  - Can use storage resources (i.e., challenge large data)
- **If the check on entities is not done simultaneously, one entity can create an infinite # of identities**
  - Can't be used for computational, communication resources
  - Could be used for storage, however it's extremely wasteful

Observations on Vouching

- **If one accepts an identity vouched by q accepted identities, a malicious set F of nodes can create an arbitrary number of identities if |F| > q**
- **If the accepting identities is not coordinated in time across the system, one can create an infinite # of identities**

Outline

- Power-law networks and flat overlays
- Sybil attack
- **P2P routing attacks**
- **Discussion**

P2P Routing

- **A key identifier space**
- **A node identifier space**
- **Rules for associating keys to nodes**
- **Per-node routing tables**
- **Rules for forwarding packets to neighbors**
- **Rules for updating routing entries when nodes join/leave**

- **Routing attacks:**
  - Attacks on the ID space (manipulating Ids)
  - Attacks on the routing tables (manipulating latencies/proximity)
  - Attacks on forwarding protocol (ignore messages)

Node ID Assignment

- **Example of possible attacks:**
  - Surround a target node
  - Partition the overlay network
  - Become the root of a certain document
- **How are IDs generated?**
  - Randomly in FreeNet
  - Hashing something in Chord/Pastry
  - Both are easy to manipulate
- **Solution: centralized certification authority for Ids**
  - Single point-of-failure for the system (DoS attacks)
  - Small overlay networks are problematic
  - Dynamic IDs DHTs problematic

Attacks on Routing Tables

- **Fake being the closest node**
- **If f faulty nodes, probability of receiving bad update**
  - f + (1-f)f …. quite a bit
- **Solution: use a backup routing table**
  - trade complexity for security
    - Sounds like a really bad idea
- **Potential issues:**
  - Bootstrapping nodes
  - Bugs
  - Probability of receiving bad updates is still f + (1-f)f

Attacks on forwarding

- **Attacks:**
  - Ignore forwarding requests
  - Be the root of a key and ignore lookups
    - (1-f)^(log N)
    - f = 10%, routing fails 35% of the time
- **Solution: failure test**
  - Have several roots for a document
  - The attack only works when faulty nodes are becoming the roots for the document
    - This means that the root set ID density is higher than the node ID's density in the overlay

Additional Attacks

- **Content pollution**
  - Fair-exchange is hard to enforce in a fully-decentralized manner
    - A central authority is a great solution (i.e., escrow)
- **Rapid joins and leaves**
- **Lots of key inserts**
- **Lots of negative lookups**
- **Inconsistent behavior**
  - “good citizen” to half of the network, “bad citizen” to the other half

High-level Lesson from P2P Security Paper

- **Simple idea:**
  - Build a P2P around a verifiable invariant
  - Check for invariant to enforce security
- **Observation:**
  - Very hard to do in practice
  - A centralized check for the invariant would simplify the problem tremendously

Discussion

- **How decentralized should a P2P be?**

Discussion

- **How decentralized should a P2P be?**
  - CDNs are semi-decentralized:
    - Central authority, management, etc..
    - Can be used to bootstrap security/trust

Discussion

- **How decentralized should a P2P be?**
  - CDNs are semi-decentralized:
    - Central authority, management, etc..
    - Can be used to bootstrap security/trust
- **Is the Internet decentralized?**
  - If yes, how can it avoid these problems?
  - How come BGP works?