Network Coding in P2P-Systems

Christian Ortolf

Overview

- Introduction
- Galois fields
- Encoding/Decoding of Files
- Gain
  - Coupon Collector's problem
- Problems
- Security
- Other fields where it can be used
- Research in Freiburg

Introduction - History

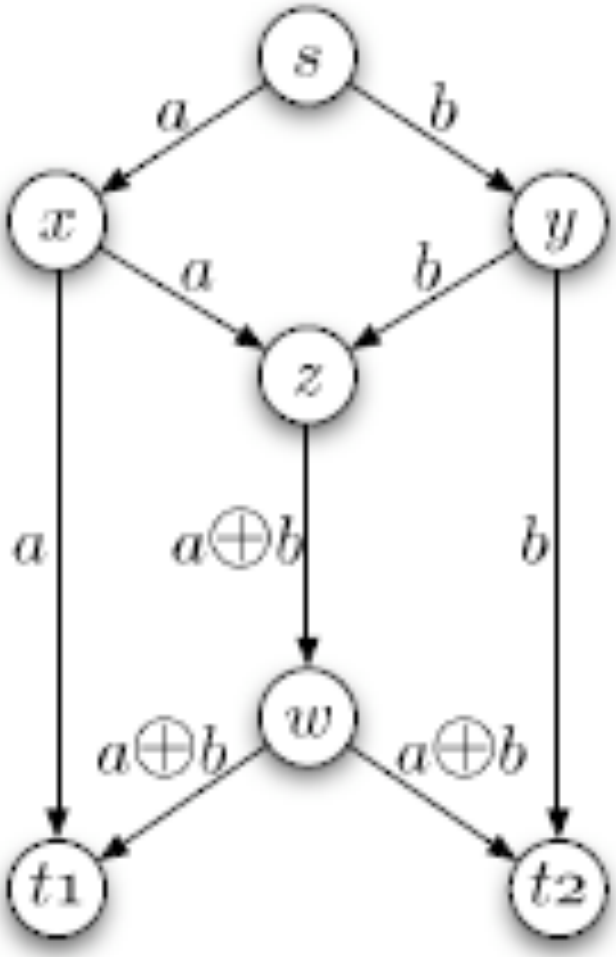
- 1999 first time used R.W.Yeung and Z. Zhang , "Distributed Source Coding for Satellite Communications"

- 2000 Definition of Network Coding, Max-Flow Min-Cut Theorem *in* ***Ahlswede, R., Cai, N., Li, S.Y.R., Yeung, R.W.: Network information flow.***

- ***2005 through Avalanche, Network Coding gets into the media, "Network Coding for Large Scale Content Distribution", C. Gkantsidis, P. Rodriguez***

- ***Since 2000 more than 200 Papers dealing with Network Coding***

Introduction – Canonical Example



- Max-Flow – Min-Cut Theorem
- Max Flow can't be reached here without coding

Galois field – Évariste Galois

- * 25. Oktober 1811 in Bourg-la-Reine
- † 31. Mai 1832 in Paris

Galois field

- Field
- Notation

$$GF\_p^n\_ \quad additional \quad F_q$$

- Example: M = {0,1} with + and * defined as:

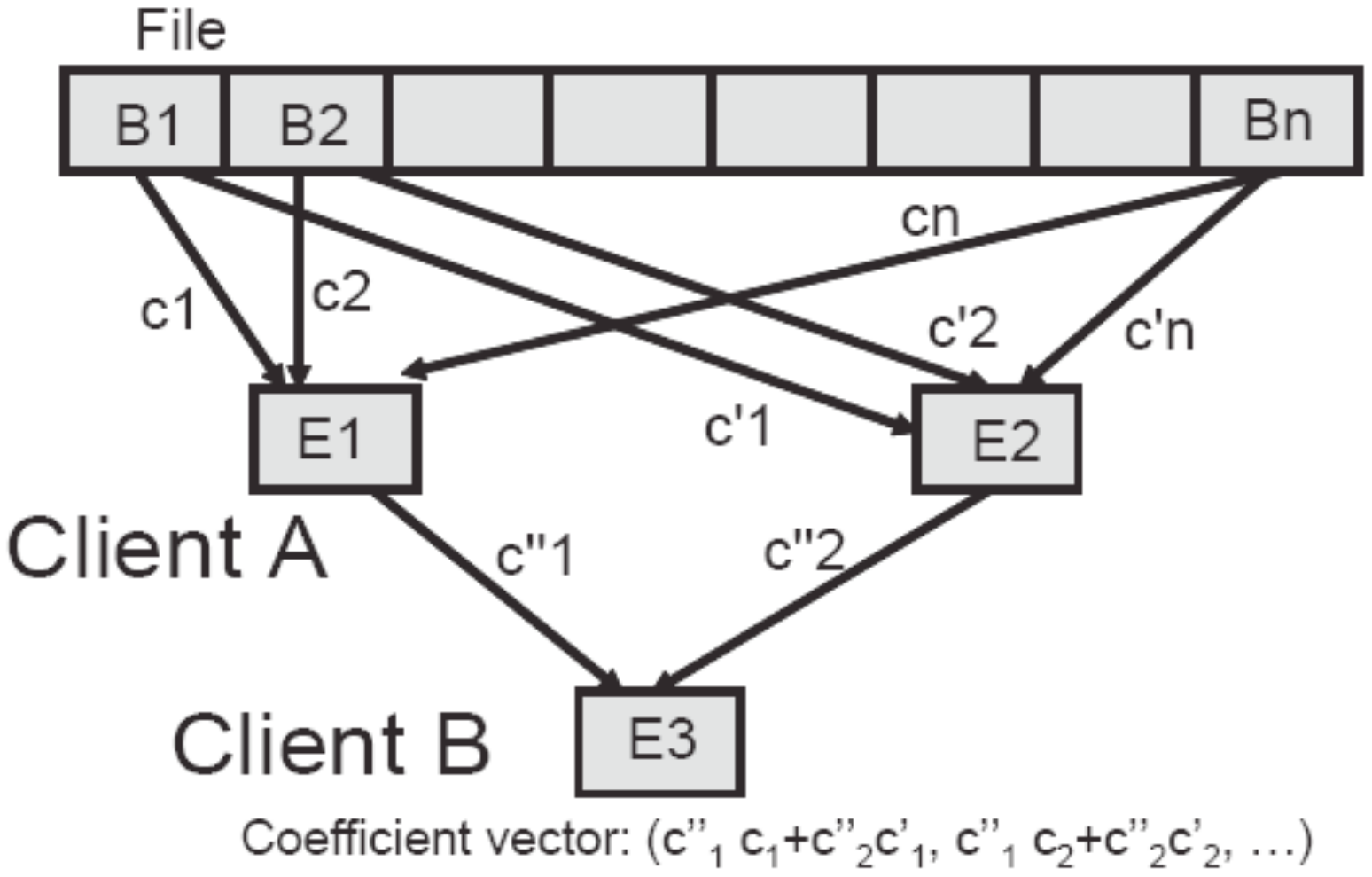
+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

- Construction: Irreducible Polynomes

Galois Field – Example

Encoding of files



- From "Network Coding for Large Scale Content Distribution", C. Gkantsidis, P. Rodriguez

Encoding - Example

Decoding of Files

- Check if Coefficient vector is Helpful

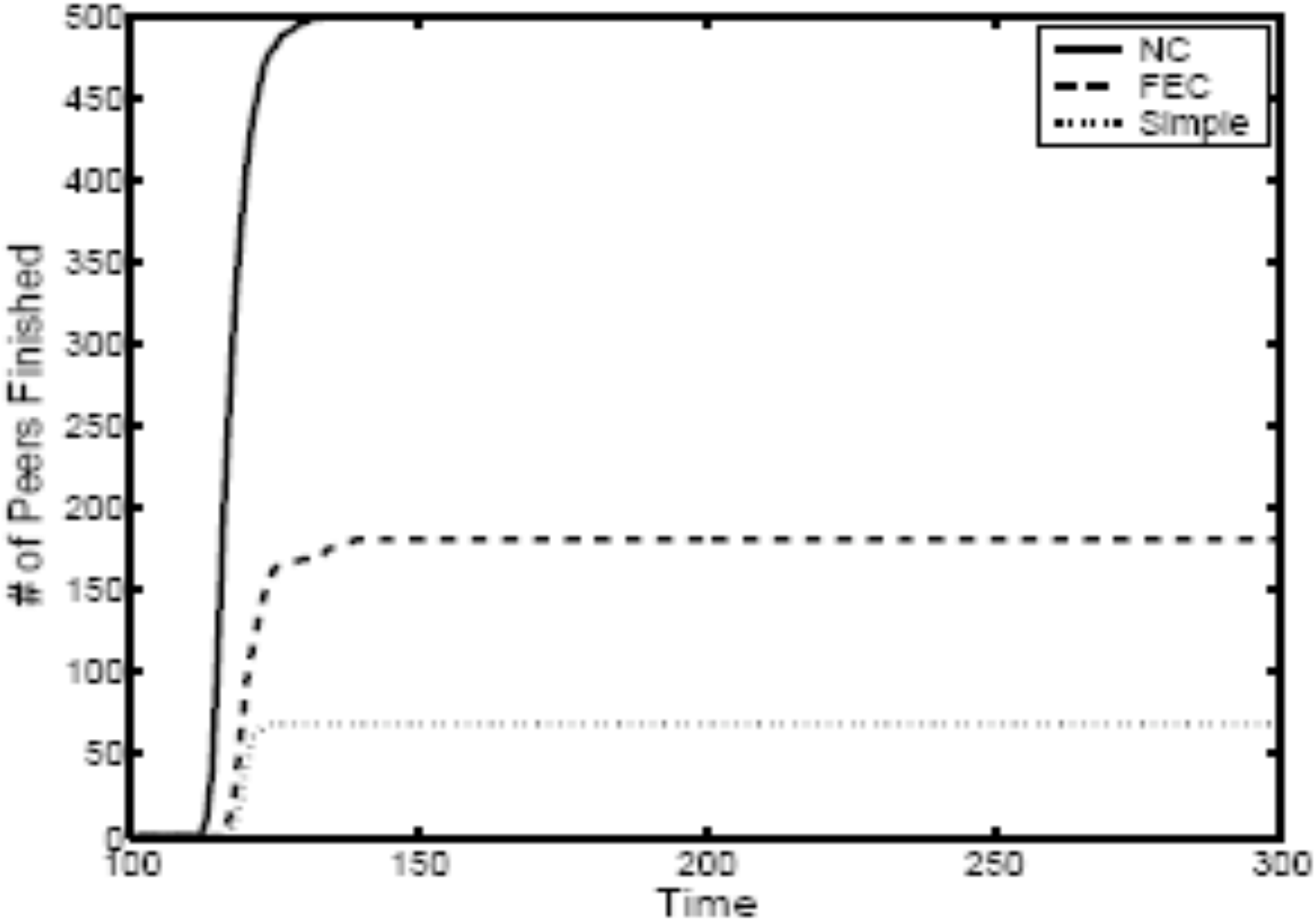
$$\begin{matrix} a_{11} & a_{12} & a_{13} & \_ & a_{1n} & z_{11} & z_{12} & 0 & \_ & 0 \\ 0 & a_{22} & a_{23} & \_ & a_{2n} & z_{21} & z_{22} & 0 & \_ & 0 \\ 0 & 0 & 0 & \_ & 0 & 0 & 0 & 1 & \_ & 0 \\ \_ & \_ & \_ & \_ & 0 & 0 & 0 & 0 & \_ & 0 \\ 0 & 0 & 0 & \_ & 0 & 0 & 0 & 0 & \_ & 1 \end{matrix}$$

- Download up to N Blocks
- Invert Matrix and Decode

Gain – Solves Coupon Collector's Problem

- Easier Protocol
  - Less Knowledge Exposed
  - Better Inscentive mechanisms possible
- High Chance to finish File if source leaves

Gain

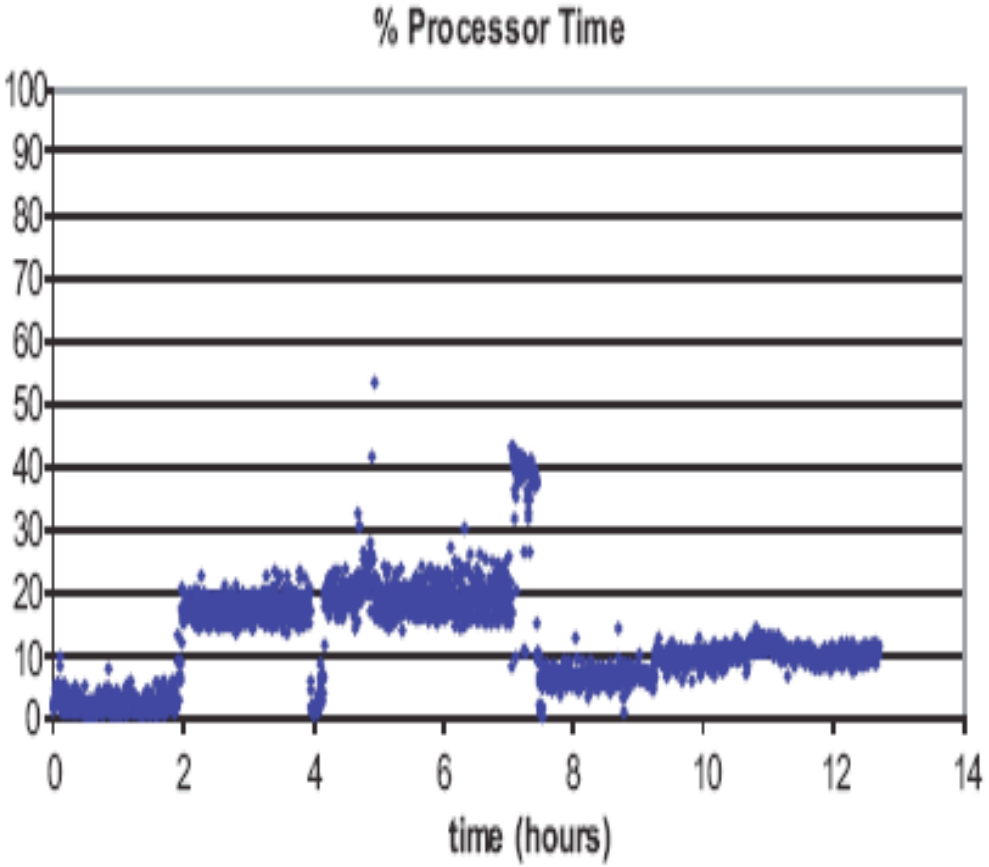


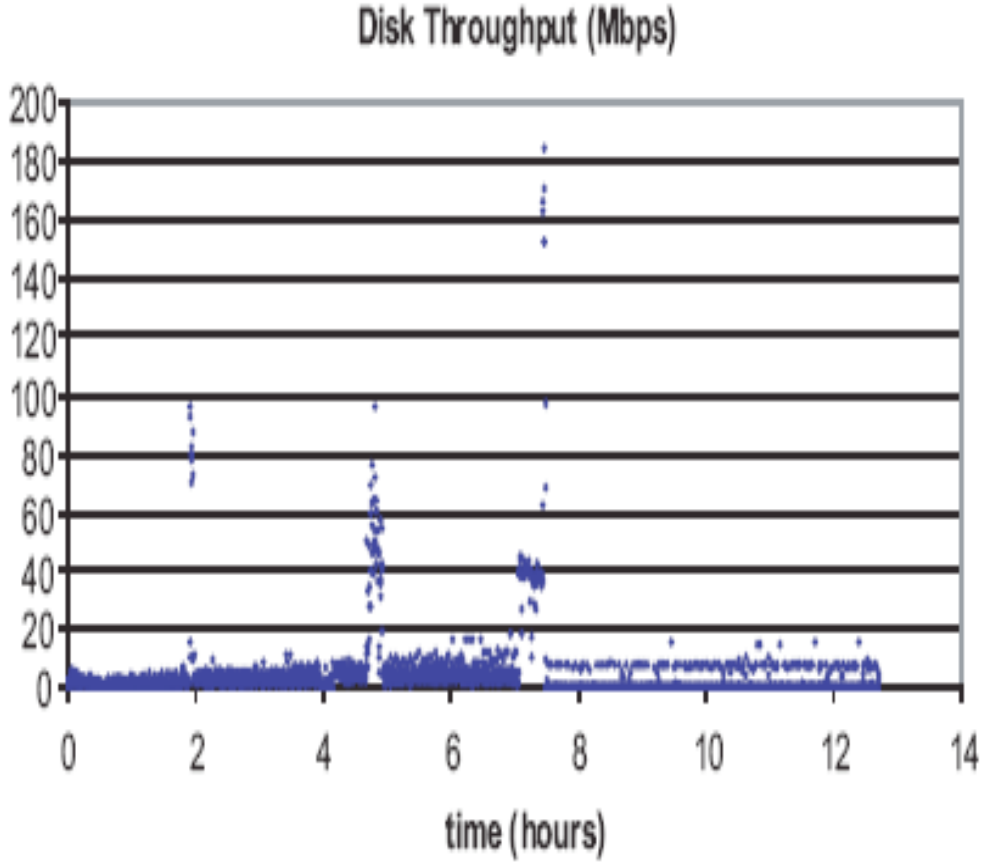
From "Network Coding for Large Scale Content Distribution", C. Gkantsidis, P. Rodriguez

Problems

- Small Blocks make large Coefficient Vectors
  - 512KB blocks on 4GB file with GF(2^8) makes 64MB
- CPU usage
  - Encoding of aBlock O(n)
  - Decoding $O(n^2)$
  - Inverting $O(n^3)$

Problems - CPU





- ***From Anatomy of a P2P Content Distribution system Gkantsidis et al***
- Average: 1.4Mbit down 300Kbit/s up

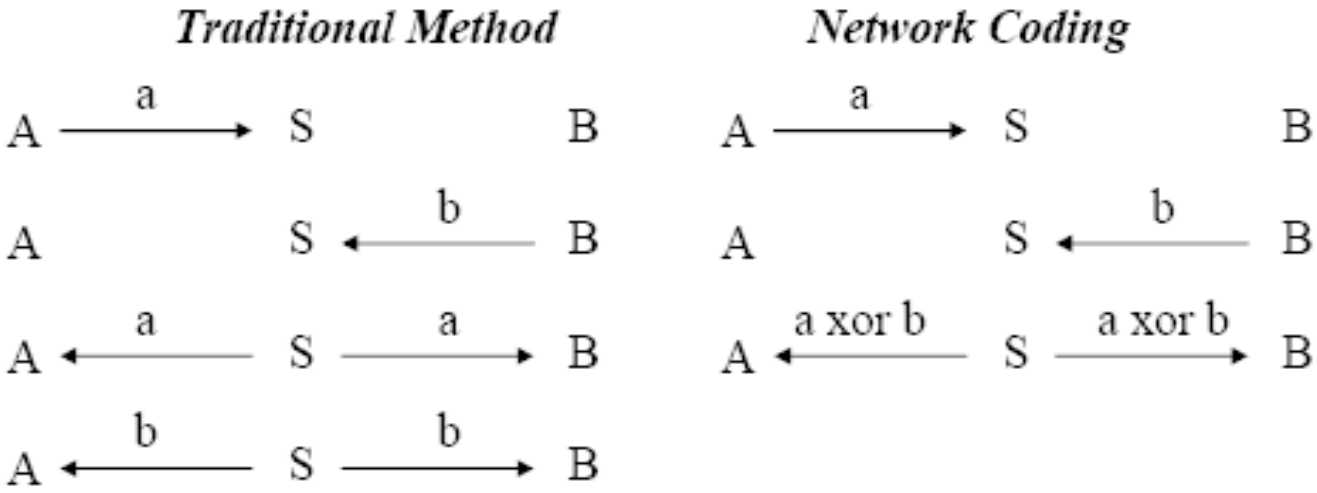
Security

- Usual Hashfunctions don't work
  - Homomorphic Hashfunctions as solution
    - Very slow
      - Cooperative hashing
        - Complex
        - Needs SRCs to protect against DoS
      - Bulk Hashing
  - Secure Randomn Checksums (SRCs)

- Harder to get info from a tapped wire
- Datainjection even without Hashing nearly impossible

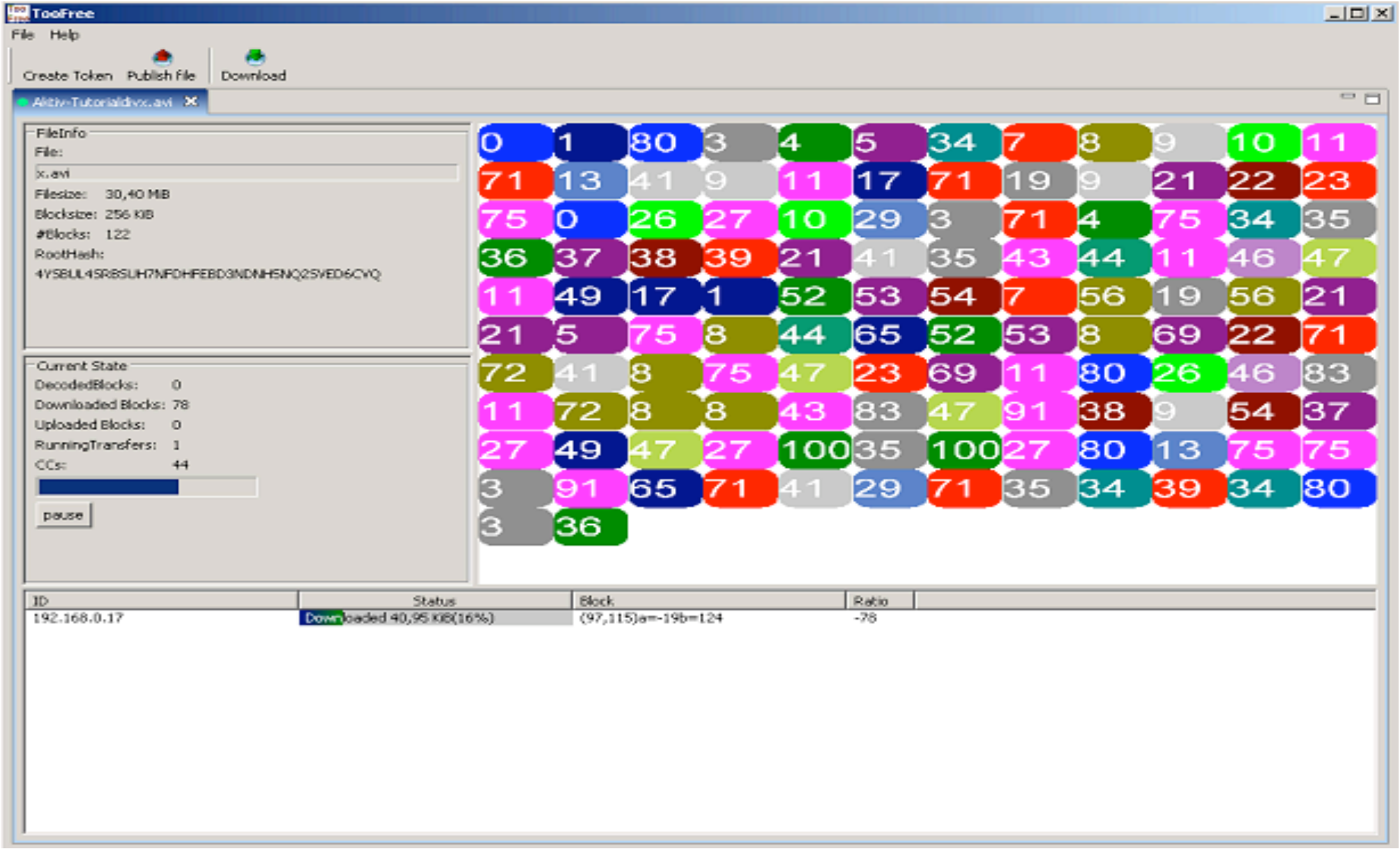
Other Fields where it can be used

- In Multicast Networks (example from NC Primer)



- Reduce Cost in WSN / Storage redundancy
- Network Tomography
- P2P - SAN

Research in Freiburg - TooFree



Finding solutions for the too large CPU cost

Thank you for
your Attention

Questions?