Random Linear Coding in Distributed Storage Systems

Adam Visegradi, MTA SZTAKI

2013, BOINC Workshop

The presented work is supported by the FP7 Capacities Programme under contract nr RI-312297 (IDGF-SP)

Volunteer Storage

- Unreliable node → need **redundancy**
- Possibilities
  - Replication
    - *k* blocks of data
    - *r* factor of replication
    - *kr* copies of each block are stored
  - Erasure coding
    - *k* blocks of data
    - *n*>*k* distinct coded blocks are generated
    - Any *k*'>=*k* of the coded blocks will be sufficient to reconstruct data

- http://boinc.berkeley.edu/trac/wiki/VolunteerDataArchival
  - 40 blocks, 20 node failures: 2000% vs. 50% redundancy

Volunteer Storage

- Need **redundancy**
- Possibilities
  - Replication
  - Erasure coding
  - **Random Linear Coding**

Coeff. GF(2)	Packet
1	A
0	B
1	C
1	D
Coded block:	A **xor** C **xor** D

Volunteer Storage

- Need **redundancy**
- Possibilities
  - Replication
  - Erasure coding
  - **Random Linear Coding**

Coeff. GF(2^w)	Packet
a	A
b	B
c	C
d	D
Coded block:	aA **xor** bB **xor** cC **xor** dD

Why RLC?

- Data reconstruction is **not** required
  - Storage overhead is still needed centrally
  - Using *sparse* coding, this overhead is lower
- Same redundancy with lower storage overhead; better **reliability**
- *Assuming* authenticity, complete **decentralization** is possible

Why not RLC?

- Compute intensive

k	64kB	128kB	256kB	512kB	1MB	2MB	4MB	8MB	16MB	32MB	64MB	128MB	256MB	512MB
8	0.0009	0.0016	0.0028	0.0076	0.0108	0.0214	0.049	0.0986	0.1946	0.378	0.7273	1.455	2.9411	5.8764
16	0.0016	0.0035	0.0066	0.0107	0.0277	0.0365	0.0811	0.1566	0.3072	0.648	1.2729	2.5983	5.4194	10.960
32	0.0023	0.0055	0.0099	0.0256	0.0346	0.0663	0.1304	0.2976	0.5828	1.139	2.2639	4.5539	9.3095	19.017
64	0.0043	0.0086	0.0203	0.0439	0.0655	0.1459	0.278	0.5567	1.1023	2.2089	4.4876	8.8185	18.069	36.353
128	0.0134	0.0182	0.0362	0.081	0.1414	0.31	0.5904	1.1449	2.2859	4.5113	9.0622	18.172	36.547	74.225

Thank you!

https://github.com/avisegradi/rnc-lib

Adam Visegradi <adam.visegradi@sztaki.mta.hu>