



**HAL**  
open science

# The Mojette Erasure Code for Distributed File Systems

Dimitri Pertin, Benoît Parrein, Nicolas Normand

► **To cite this version:**

Dimitri Pertin, Benoît Parrein, Nicolas Normand. The Mojette Erasure Code for Distributed File Systems. ACM EuroSys, Apr 2014, Amsterdam, Netherlands. 2013. hal-01149426

**HAL Id: hal-01149426**

**<https://hal.science/hal-01149426>**

Submitted on 27 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# The Mojette Erasure Code for Distributed File Systems

Dimitri Pertin, Benoît Parrein, Nicolas Normand

► **To cite this version:**

Dimitri Pertin, Benoît Parrein, Nicolas Normand. The Mojette Erasure Code for Distributed File Systems. ACM EuroSys, Apr 2014, Amsterdam, Netherlands. 2013. <hal-01149426>

**HAL Id: hal-01149426**

**<https://hal.archives-ouvertes.fr/hal-01149426>**

Submitted on 27 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Mojette Erasure Code for Distributed File Systems

Dimitri Pertin<sup>1,2</sup>, Benoît Parrein<sup>1</sup>, Nicolas Normand<sup>1</sup>

<sup>1</sup>LUNAM Université, Université de Nantes, IRCCyN UMR 6597, Nantes, France  
<sup>2</sup>Fizians SAS, Nantes, France

## 1 Data protection is expensive

In distributed storage systems, failures are the norm

### Replication

- $n$  copies
- $n - 1$  storage overhead
- ok for hot data



Figure 1: Data triplication (200% overhead)

### Erasure Coding (EC)

- same protection
- only  $\frac{n}{k} - 1$  overhead
- limited to cold data

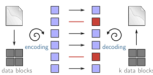


Figure 2: (6,4) Erasure coding (50% overhead)

**Problem:** implementations of Reed Solomon codes are computational bottlenecks for I/O-intensive applications  
 → Galois field multiplication  $GF(2^m)$  is expensive

**Our goal:** design fast ECs for Distributed File Systems  
 → intense random access over small blocks (e.g. 4 kB)

## 2 The fast Mojette erasure code

### The $(n, k)$ non-systematic Mojette encoding

- $k$  lines of the grid (i.e. data blocks)
- $n$  projections  $(p, q)$  (i.e. parity blocks)
- based only on additions (XORs)
- reconstruction if  $\sum |q_i| \geq k$
- $(1 + \epsilon)$  optimal

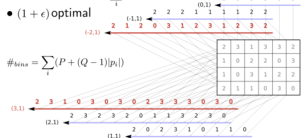


Figure 3: (6,4) Mojette encoding (mod 4) over a 6x4 discrete grid of data

### The Mojette decoding

- decoding dependency graph
- deterministic path
- complexity is  $\mathcal{O}(kPQ)$

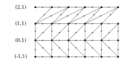


Figure 4: Mojette dependency graph

## 3 Performance comparison with Reed Solomon codes

Intel Xeon @3.10 GHz

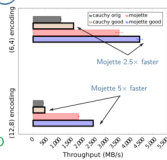
### Criteria

- blocks of 4 kB
- input: 704 MB video
- output: desired protection

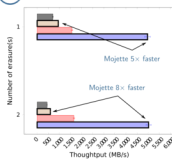
### In competition

- cauchy RS (Jerasure 1.2)
- cauchy good: sparse matrix
- mojette (RozoFS)
- mojette good: optimized I/O

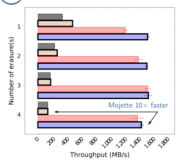
### 1 Encoding performance



### 2 (6,4) Decoding perf



### 3 (12,8) Decoding perf



## 4 Conclusion and future works

**Conclusion:** Mojette provides better performance, independently of the number of erasures

### Future works:

- systematic Mojette for better performance
- repair problem

## 5 Acknowledgments

Supported by **FEC4CLOUD** (ANR-12-EMMA-0031-01)

Mojette code available at <http://github.com/rozofs/rozofs>

contact me: dimitri.pertin@univ-nantes.fr