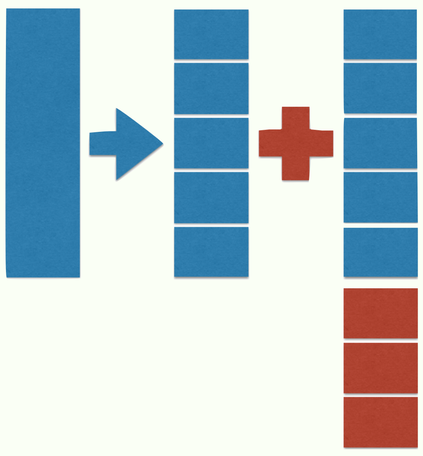
Applications of Erasure Coding to WLCG Data Distribution

Erasure Coding



File placement for Erasure-coded files:

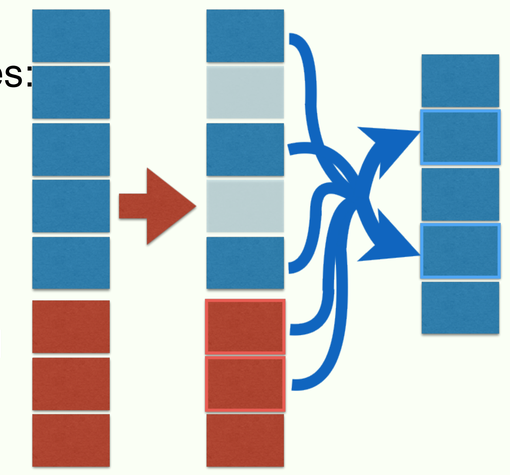
- ▶ Split file into k chunks (padding last chunk if needed).
- ▶ Generate m additional chunks via linear EC mapping.
- ▶ Distributed m+k chunks across filesystem (Grid)

Data distribution for EC coded chunks is fast, parallel and scales inversely with k

Bottlenecks on Restore

Chunk recovery for Erasure-coded files:

- ▶ Recover at least k chunks from distributed copies
- ▶ Reconstruct missing chunks from k coding copies.
- ▶ Rewrite regenerated chunks to Grid

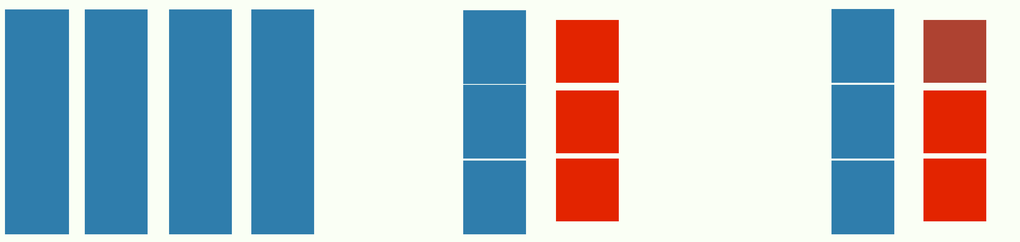


Chunk recovery for Erasure-coded files requires reading “whole file” to rebuild any one lost chunk.

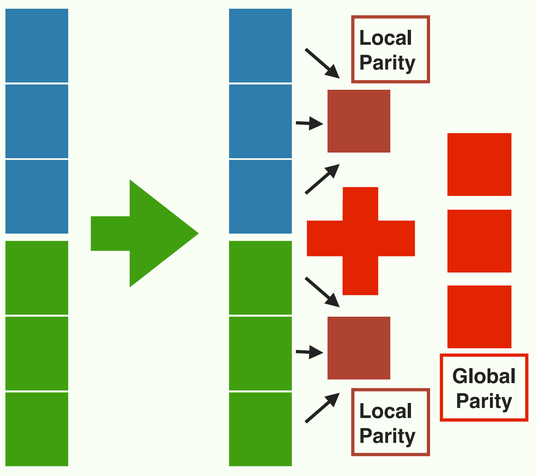
Unacceptable for World-distributed data.

An implementation, in Go, of Hitchhiker-style Local Reconstruction Codes for Grid Storage.

Conventional Replication Strategy	Naïve Erasure Coding	Local Reconstruction Coding
Integer multiples of file size needed.	*Rational* multiples of file size available.	*Rational* multiples of file size available.
Read parallelism = Number of replicas.	Read parallelism = Number of *chunks*.	Read parallelism = Number of *chunks*.
Repair = *copy* whole file again.	Repair = *read* whole file again, *copy single chunk*.	Repair = *read* local *subset* of file, *copy single chunk*.
Geographical distribution “coarse”	Geographical distribution “*fine*”	Geographical distribution “*fine*”, “*hierarchical*”



Local Erasure Codes



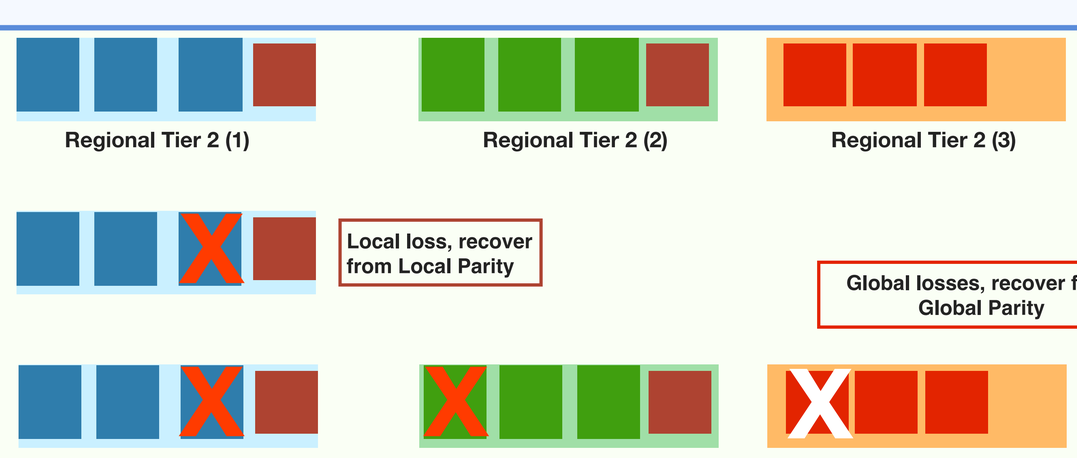
- ▶ **Modifications of standard Reed-Solomon style EC codes.**
- ▶ **Modify encoding to couple chunks in a local way**
- ▶ **Reconstruction of single lost chunk requires only locally coupled chunks.**
- ▶ **Reconstruction of significant losses breaks this assumption**

Work by Facebook, UC Berkeley

ref: ***XORing Elephants: Novel Erasure Codes for Big Data***
arXiv:1301.3791v1 [cs.IT] 16 Jan 2013

ref: ***A “Hitchhiker’s” Guide to Fast and Efficient Data Reconstruction in Erasure-coded Data Centers***
SIGCOMM14

Restoration



Chunk recovery for LEC files often requires reading only Local Cloud copies.

S.Skipsey, D Crooks*, G Roy, G Stewart, D Britton