Data at Scale

Erasure Correcting Codes

Remote Storage

- High Availability Measures
  - Replication
  - Erasure Correcting Coding

Erasure Correcting Codes

- Data broken into pages

$$P_1 = \{p_{1,0}, p_{1,1}, p_{1,2}, \ldots\}$$
$$P_2 = \{p_{2,0}, p_{2,1}, p_{2,2}, \ldots\}$$
$$P_3 = \{p_{3,0}, p_{3,1}, p_{3,2}, \ldots\}$$
$$\vdots$$

- Create parity pages

$$C_1 = \{c_{1,0}, c_{1,1}, c_{1,2}, \ldots\}$$
$$C_2 = \{c_{2,0}, c_{2,1}, c_{2,2}, \ldots\}$$
$$C_3 = \{c_{3,0}, c_{3,1}, c_{3,2}, \ldots\}$$
$$\vdots$$

Erasure Correcting Codes

- "Check" pages calculated using an erasure correcting code

$$(c_{1,i}, c_{2,i}, \ldots, c_{m,i}) = \Phi(p_{1,i}, p_{2,i}, \ldots, p_{n,i})$$

- *m*/*n*-code:
  - *m* data pages
  - *n* check pages
  - Any *m* of the *n*+*m* pages are sufficient to reconstruct all *m* data pages

Erasure Correcting Codes

- Related to error-correcting codes in telecommunications / networking
- Simplest code: Parity code
  - One check symbol
  - Is parity of the data symbols

$$c_{1,i} = p_{1,i} \oplus p_{2,i} \oplus \ldots \oplus p_{n,i}$$

  - Can calculate single lost data symbol as the parity of the survivors and the check symbol

Erasure Correcting Codes

- Group Activity:
  - Reconstruct the missing text

G	47	B	42	L	42
o	6f	u	75		6f
o	6f	e	65	O	7e
d	64	n	6e		6f
	20	a	61	S	61
e	65	s	73		58
v	76		20	T	37
e	65	n	6e		68
n	6e	o	6f		69
i	69	c	63		7e
n	6e	h	68		
g	67	e	65		
		s	73		

Erasure Correcting Codes

- Many codes are known
- Some involve algebraic objects such as finite fields
  - Example GF(2ˆ8)
    - Elements are bytes (string of 8 bits)
    - Addition is exclusive-or
      - 0010 0011 + 1011 1110 = 1001 1101
    - Multiplication can be defined as polynomial multiplication
      $00100011 \approx t^5 + t + 1$

Erasure Correcting Codes

- Multiplication in GF(2ˆ8)

$$00100011 \approx t^5 + t + 1$$

$$10111110 \approx t^7 + t^5 + t^4 + t^3 + t^2 + t$$

  - Multiply polynomials with coefficients in {0,1}
    - Caution: 1+1 = 0

$$t^{12} + t^{10} + t^9 + t^8 + t^7 + t^6$$

$$+t^8 + t^6 + t^5 + t^4 + t^3 + t^2$$

$$+t^7 + t^5 + t^4 + t^3 + t^2 + t$$

$$= t^{12} + t^{10} + t^9 + t$$

Erasure Correcting Codes

- Now divide this polynomial by a *generator polynomial*

$$t^8 + t^4 + t^3 + t + 1$$

- The remainder is the product
  - Group exercise: Calculate the remainder

Erasure Correcting Codes

- However, we do not have to do this multiplication every time
  - Can use tables and algebraic properties
  - Plank, Greenan, Miller: Can multiply 16 GF elements by a constant GF element with 6 assembly instructions
    - Using PSHUFB instruction

Erasure Correcting Codes

- Simple method to create erasure correcting codes
  - Start with Vandermonde matrix of size *n* by *n*+*m*
    - Defined by *n*+*m* <u>different</u> Galois field elements

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ \alpha_0 & \alpha_2 & \alpha_3 & \alpha_4 & \dots & \alpha_{n+m-1} \\ \alpha_0^2 & \alpha_2^2 & \alpha_3^2 & \alpha_4^2 & \dots & \alpha_{n+m-1}^2 \\ \alpha_0^3 & \alpha_2^3 & \alpha_3^3 & \alpha_4^3 & \dots & \alpha_{n+m-1}^3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_0^{n-1} & \alpha_2^{n-1} & \alpha_3^{n-1} & \alpha_4^{n-1} & \dots & \alpha_{n+m-1}^{n-1} \end{pmatrix}$$

Erasure Correcting Codes

- Use elementary row transformations to obtain

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & r_{0,0} & r_{0,1} & \dots & & r_{0,m} \\ 0 & 1 & 0 & \dots & 0 & r_{1,0} & r_{1,1} & \dots & & r_{1,m} \\ 0 & 0 & 1 & \dots & 0 & r_{2,0} & r_{2,1} & \dots & & r_{2,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & r_{n,0} & r_{n,1} & \dots & & r_{n,m} \end{pmatrix}$$

- Always possible

Erasure Correcting Codes

- This matrix has the remarkable property that any *n* by *n* sub-matrix is invertible
- Follows that multiplication with **G** defines an erasure correcting code
    - **d** — *n* data symbols written as a row vector
    - **c** — the same *n* symbols followed by *m* check symbols

$$\mathbf{c} = \mathbf{d} \cdot \mathbf{G}$$

Erasure Correcting Code

- These “linear" codes have remarkable properties
  - Can change a single page
  - Can calculate the new check pages from the old and the new values of the page
  -

Erasure Correcting Codes

- Are used in RAID Level 6 storage systems
- Can / could be used across system boundaries