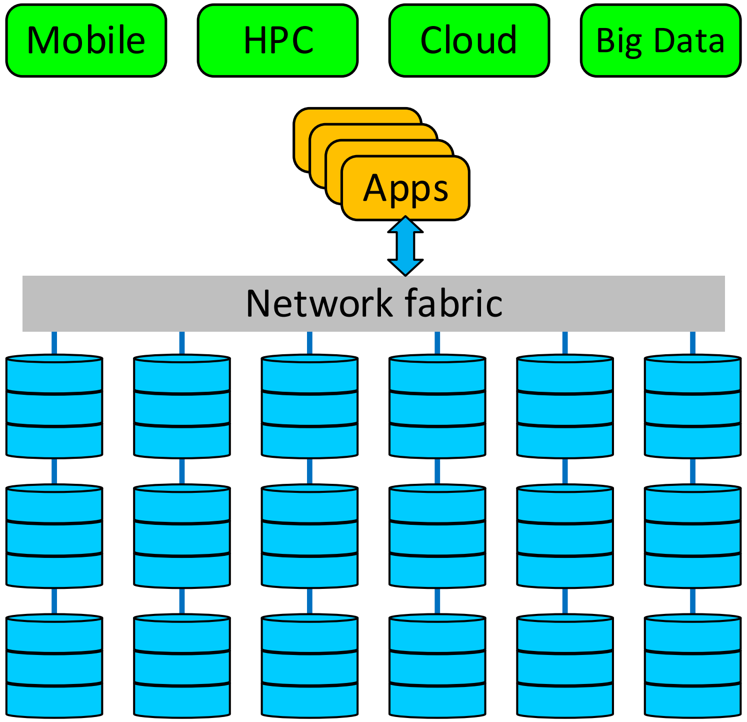
Modified Generalized Integrated Interleaved Codes for Local Erasure Recovery

Xinmiao Zhang
Dept. of Electrical and Computer Engineering
The Ohio State University

Outline

- Traditional failure recovery schemes for distributed storage system
- Locally recoverable erasures codes
- Generalized integrated interleaved (GII) codes for local erasure recovery
- Modified GII codes achieving locality improvement
- Comparisons and conclusions

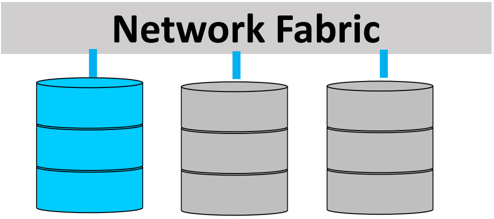
Failure Recovery for Distributed Storage



- A distributed storage system has many storage nodes, and redundancy is needed to tolerate and recover from failures
- The latency and network traffic overhead of recovering failed data packets or storage nodes largely affect the overall system performance

Conventional Failure Recovery Schemes

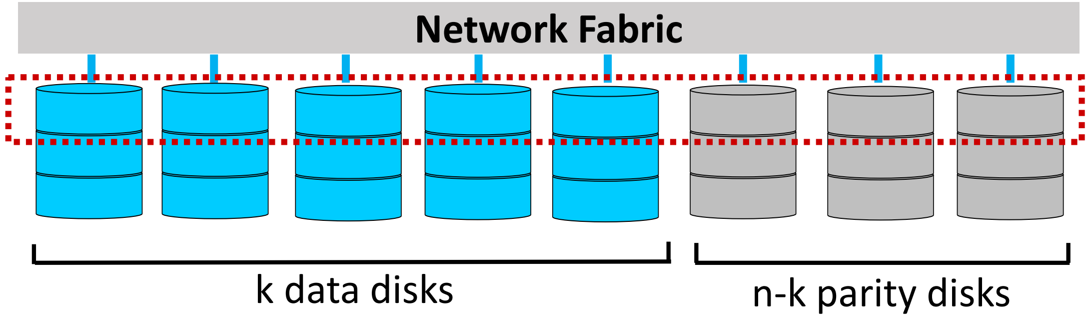
Triple replication



- Low repair overhead
- Very high redundancy

failures = erasures in coding terminology

Erasure coded



k symbols → Encoder → n symbols

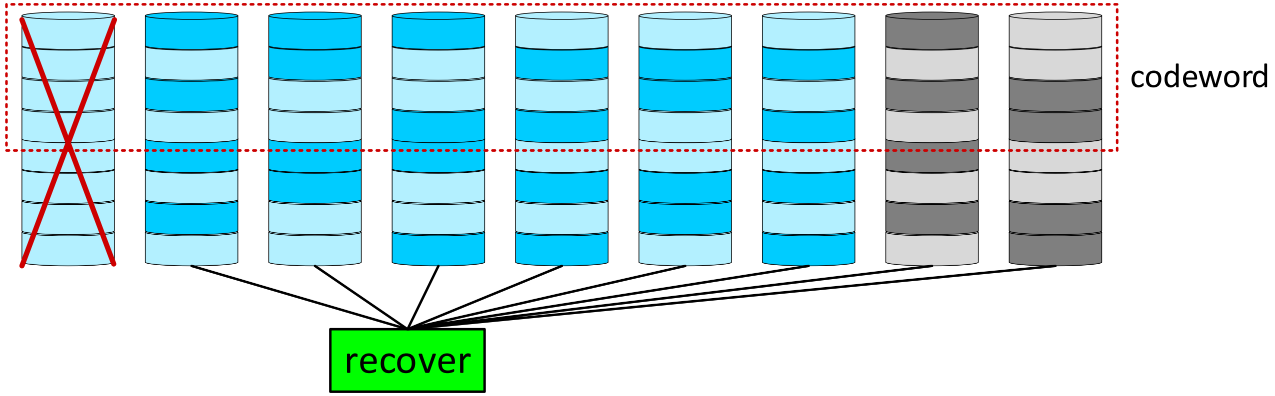
indices of $t' \leq t$ erased symbols → Decoder

k un-erased symbols → Decoder → t' recovered symbols

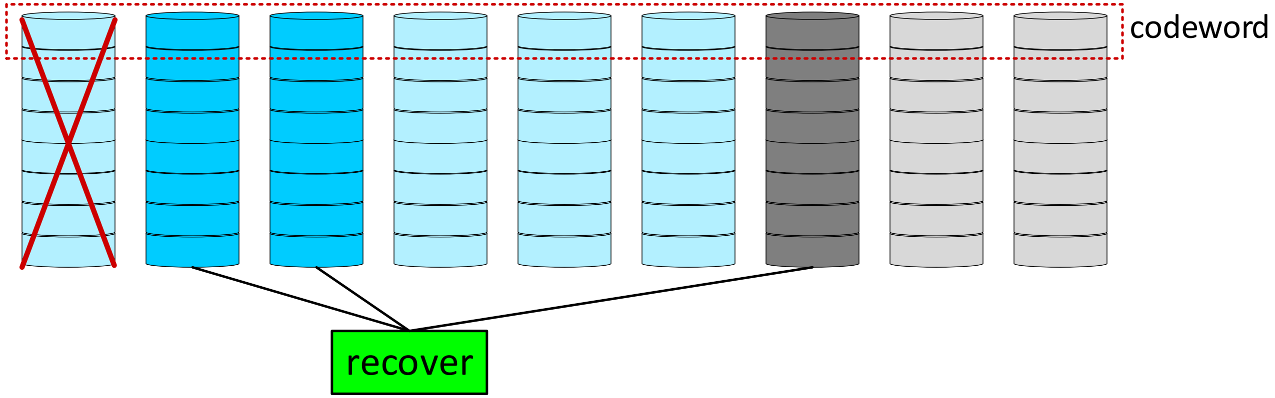
- An (n, k) code can recover at most $t = n - k$ erasures; codes meeting this bound are called maximum distance separable (MDS) codes
- Traditional erasure codes need k symbols to recover from any failures; have large repair overheads

Codes with Reduced Network Traffic for Failure Recovery

- Distributed storage needs coded failure recovery schemes accessing much less than k symbols (possible if actual erasure number $< t$: the erasure correction capability)
  - Lower network traffic
  - Reduced recovery latency
  - Better data availability

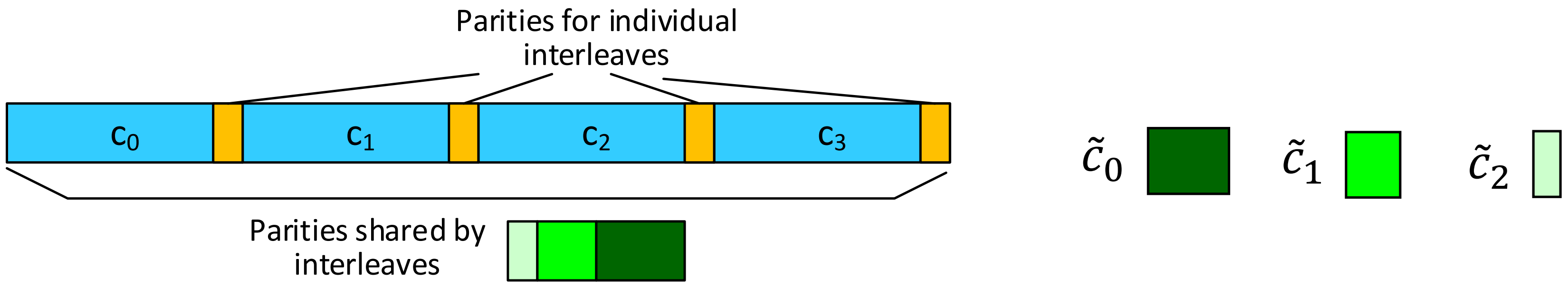


- Minimum storage regenerating codes
  - Array MDS codes
  - Read from a larger number of nodes but fewer symbols from each node



- Locally recoverable (LRC) erasure codes
  - Access fewer storage nodes
  - More redundancy is added to achieve locality

Generalized Integrated Interleaved (GII) Codes



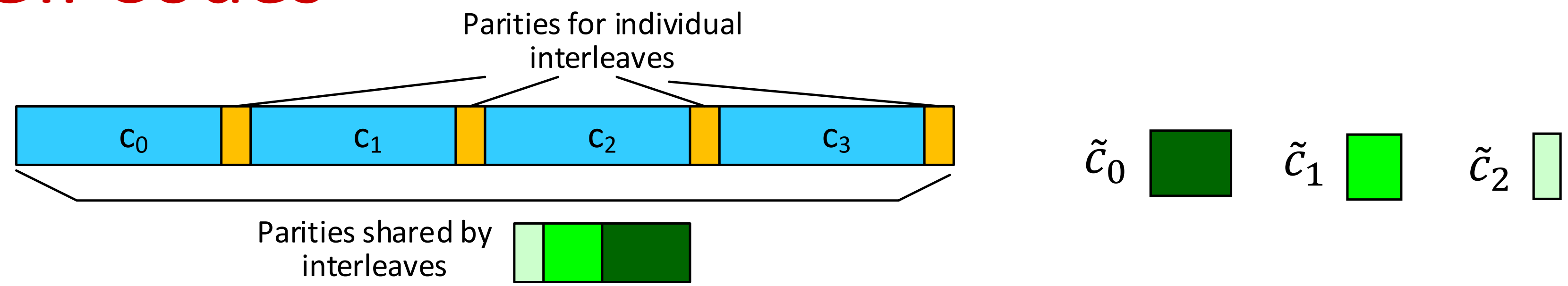
$$\begin{bmatrix} \tilde{c}_0 \\ \tilde{c}_1 \\ \vdots \\ \tilde{c}_{v-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{(v-1)} & \alpha^{2(v-1)} & \cdots & \alpha^{(v-1)(m-1)} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{m-1} \end{bmatrix}$$

Nesting matrix G

Correction capability $t_v \geq t_{v-1} \geq \cdots \geq t_1 \geq t_0$

$$\begin{array}{ccccccccc} \mathcal{C}_v & \subseteq & \mathcal{C}_{v-1} & \subseteq & \cdots & \subseteq & \mathcal{C}_1 & \subseteq & \mathcal{C}_0 \\ \Cup & & \Cup & & & & \Cup & & \Cup \\ \tilde{c}_0 & & \tilde{c}_1 & & & & \tilde{c}_{v-1} & & c_0, c_1, \cdots, c_{m-1} \end{array}$$

Decoding of GII Codes



Interleave syndromes

nested syndromes

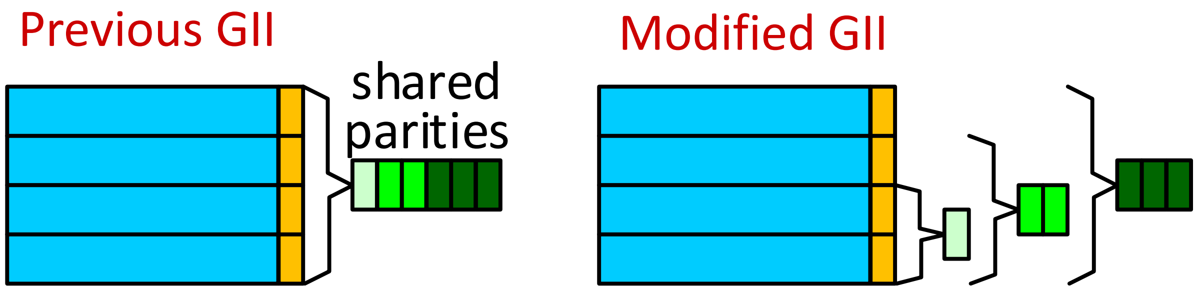
$$\begin{bmatrix} S_j^{(l_1)} \\ S_j^{(l_2)} \\ \vdots \\ S_j^{(l_b)} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{l_1} & \alpha^{l_2} & \cdots & \alpha^{l_b} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(b-1)l_1} & \alpha^{(b-1)l_2} & \cdots & \alpha^{(b-1)l_b} \end{bmatrix}^{-1}}_{\text{Syndrome conversion matrix}} \begin{bmatrix} \tilde{S}_j^{(0)} \\ \tilde{S}_j^{(1)} \\ \vdots \\ \tilde{S}_j^{(b-1)} \end{bmatrix}$$

$l_1, l_2, \cdots, l_b$: indices of interleaves with more than t_0 erasures (exceptional interleaves)

- t syndromes are needed to correct t erasures
- Higher-order syndromes for the interleaves are generated from the nested syndromes
- Syndrome conversion matrix is always invertible
- Each nested syndrome is generated by utilizing all interleaves

Need all the interleaves if any of them has more than t_0 erasures!

Modified GII Codes



$$
G' = \begin{array}{c}
\left[\begin{array}{ccccccc}
1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
0 & \alpha & \cdots & \alpha^{v-2} & \alpha^{v-1} & \cdots & \alpha^{m-1} \\
0 & 0 & \cdots & \alpha^{2(v-2)} & \alpha^{2(v-1)} & \cdots & \alpha^{2(m-1)} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & \alpha^{(v-1)^2} & \cdots & \alpha^{(v-1)(m-1)}
\end{array}\right]
\end{array}
\begin{array}{l}
t_v \\ t_{v-1} \\ t_{v-2} \\ \\ t_1
\end{array}
$$

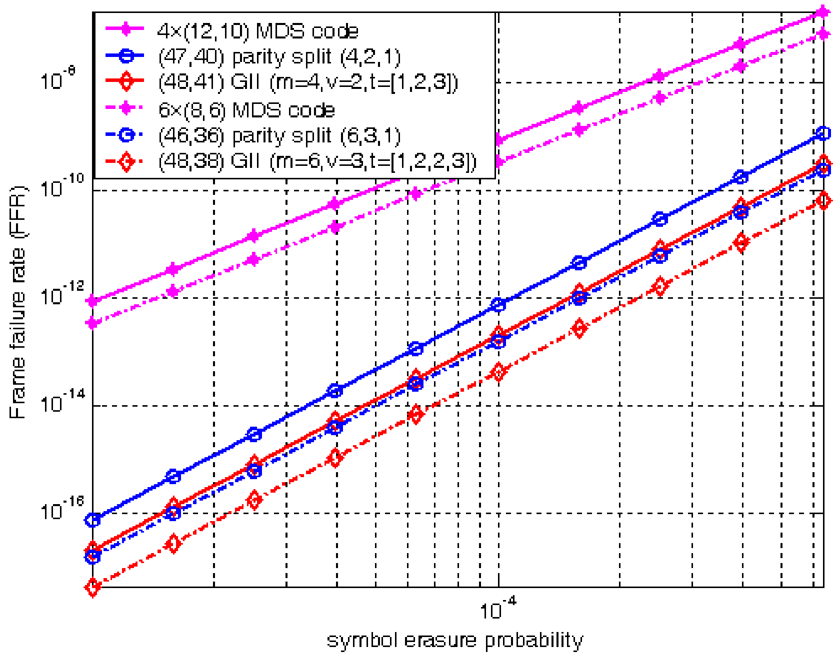
- ➢Less powerful nestings involve fewer interleaves
- ➢Form the syndrome conversion matrix using the bottom rows of G' as much as possible
  - The selected nestings should have sufficient correction capability
  - The selected nestings should cover every exceptional interleave
  - Consecutive nestings are used to simplify the selection

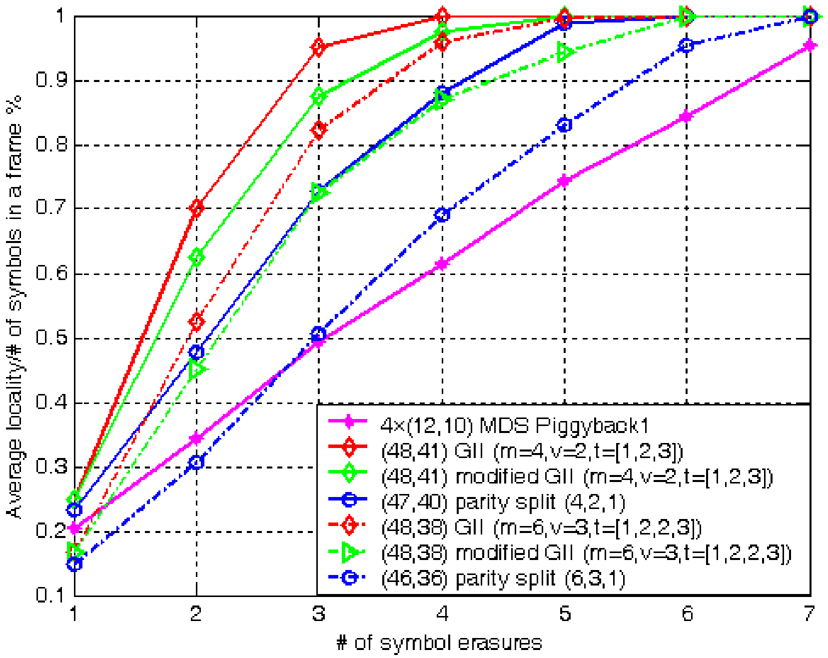
Invertibility of Syndrome Conversion Matrix

- The columns of the syndrome conversion matrix correspond to the exceptional interleaves
- All-zero columns can be avoided, but the syndrome conversion matrix may still have zero entries
- The syndrome conversion matrix is invertible if the number of interleaves does not exceed the values in the following table for given number of nested codewords (v) and finite field order

v	$GF(2^4)$	$GF(2^5)$	$GF(2^6)$	$GF(2^7)$	$GF(2^8)$	$GF(2^9)$
2	16	32	64	128	256	512
3	5	6	12	22	28	62
4	5	6	8	12	15	20

Correction Capability and Locality Comparisons





- Modified GII codes preserve the same correction capability as the original GII codes for most practical settings
- Modified GII codes require fewer interleaves to utilize the shared parities when there are fewer extra erasures to correct
- Have very small implementation overhead compared to the GII codes
- Achieve good tradeoff on locality and correction capability

Conclusions

- Modified GII codes substantially improve the locality for erasure correction over prior GII codes
- Modified GII codes do not bring any correction capability degradation for most practical settings
- Modified GII codes achieve good tradeoff on the locality and correction capability compared to other LRC codes
- Further locality improvement can be achieved by multi-layer integrated interleaved codes