# A study on the repair set of locally repairable codes


Mi-Young Nam, Jung-Hyun Kim, and Hong-Yeop Song

Yonsei Univ.



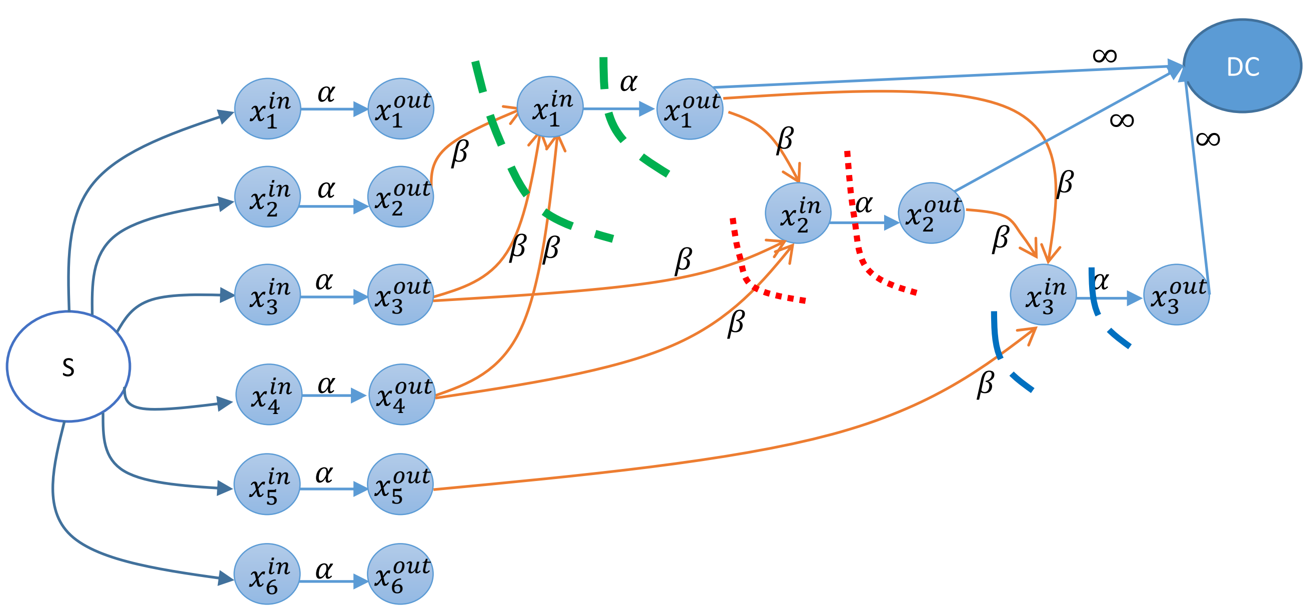
KICS Fall Conference 2015


## ▪ Regenerating Code

**ASSUMPTIONS:**

- ✓ The number of nodes ***r*** that should be contacted by a new comer for node repair
- ✓ The number of nodes ***K*** by a Data Collector for file retrieval
- ✓ $r \geq K$ is assumed
- ✓ **Any *r*** helper node should be able to repair a failed node

ANALYSIS using Information-Flow Graph




✓ The maximum file size that can be stored by using a Regenerating code is

$$B \leq \sum_{i=0}^{K-1} \min\{(r-i)\beta, \alpha\}$$

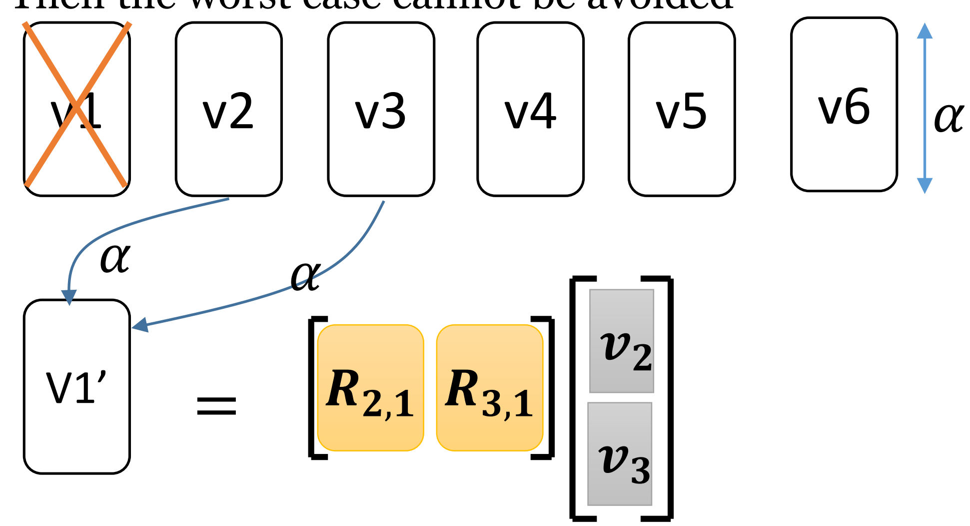
$$= \min\{3\beta, \alpha\} + \min\{2\beta, \alpha\} + \min\{\beta, \alpha\}$$

for $r = K = 3$ in the above example

## ▪ Traditional LRCs

**ASSUMPTIONS:**

- ✓ Every helper node sends its whole stored data to a newcomer node
  $: \beta = \alpha$
- ✓ Therefore, the repair set is disjoint
- ✓ Then the worst case cannot be avoided



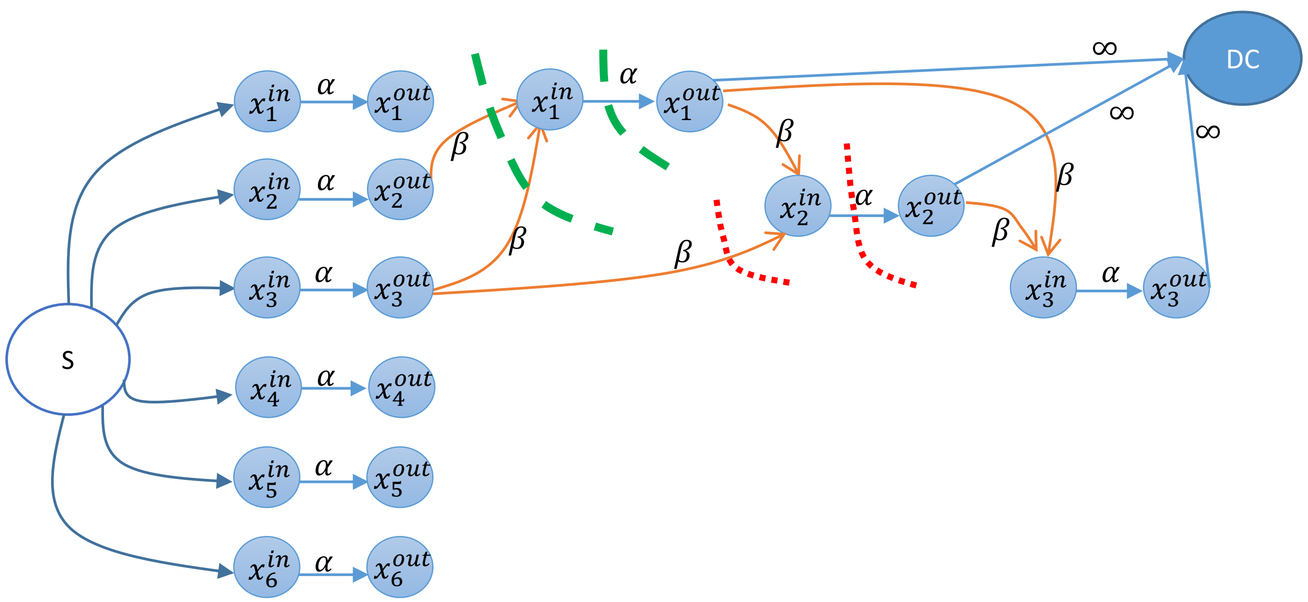

- $R_{2,1}$ is the repair matrix of size $(\alpha \times \alpha)$ that transforms a data $v_2$ into some useful data in a node V1
- $R_{3,1}$ is the repair matrix of size $(\alpha \times \alpha)$ that transforms a data $v_3$ into some useful data in a node V1
- Both $R_{2,1}$ and $R_{3,1}$ are full-rank otherwise $\beta < \alpha$ is possible
- Then node V2 can be repaired from node V1 and V3 and a node V3 also can be repaired from node V1 and V2.

$$v_k = R_{k,1}^{-1}\left(v_1 - \sum_{\substack{j=1,\\ j\neq k}}^{r=3} R_{j,1} v_j\right), for\ i = 2, 3$$

## Locally Repairable Code

**DIFFERENCE FROM REGENERATING CODE:**

- ✓ A code which satisfies $r < K$ is a Locally Repairable Code




- ✓ If we assume a random selection of helper nodes, then we cannot avoid the worst case above

$$B \le \sum_{i=0}^{r-1} \min\{(r-i)\beta, \alpha\}$$
$$= \min\{2\beta, \alpha\} + \min\{\beta, \alpha\} + 0$$

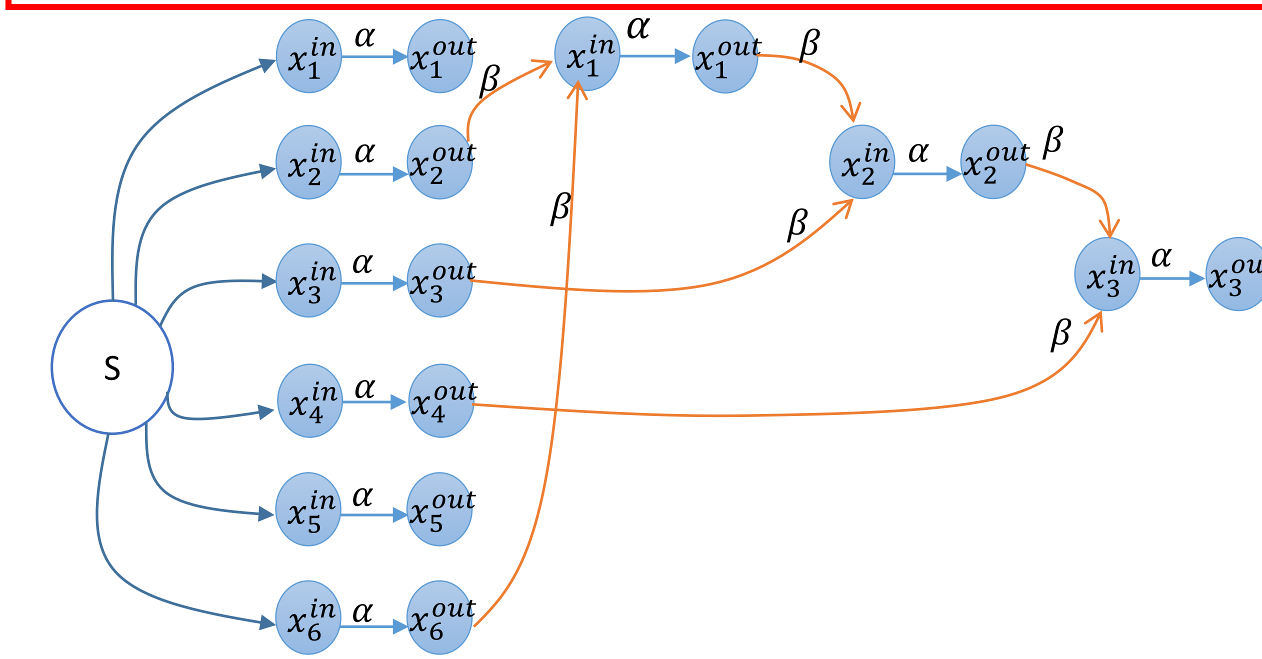
**DIFFERENCE FROM REGENERATING CODE:**

- ✓ For every node, a specific set of $r$ helper nodes, that participate in the repair process of that node, are pre-determined
  That is, fixed repair set is given
- ✓ Careful construction of repair set makes the code avoid the worst case performance above

## Repair set

- ✓ If we want to avoid the worst case in LRC the followings should be satisfied:
  i) $\beta < \alpha$,
  ii) $R(i) \nsubseteq \cup_{j \in S} R(j)$ ,
  $for\ all\ i \in [n]\ and\ for\ all\ S \subseteq [n] \backslash \{i\}, |S| \le K-1.$
- ✓ Repair set construction for codes with locality 2
  $R(i) = \{(i-1), (i+1)_{\ mod\ n}\}, for\ all\ i \in [n]$

(※ $R(i)$ is the repair set of node $i$)




## Concluding Remarks

- A structure of a repair set for LRC based on IFG
- Explicit repair-set for LRC with locality $r = 2$
- No explicit repair-set for LRC with locality $r > 2$