Claims

1. A systematic distributed storage system (DSS) comprising

a plurality of storage nodes, wherein each storage node is configured to store one of a plurality of coded blocks, the coded blocks being linearly encoded from sub-blocks of a data file, each coded block being stored at a unique one of the storage nodes; the linear encoding consisting of XOR operations on the sub-blocks; and

a set of repair pairs of the storage nodes, for each of the storage nodes;

wherein the system is configured to use the respective repair pair of storage nodes to repair a lost or damaged coded block on a given storage node; and

wherein the repair pairs include one or more alternate pairs.

2. The system in claim 1 wherein the coded blocks are Non Maximum Distance Separable.

3. The system in claim 1 or claim 2 wherein the coding is binary Simplex coding.

4. The system in any one of the preceding claims wherein the linear encoding comprises:

$$z_j = \begin{pmatrix} \alpha_{j,1} & \alpha_{j,2} & \cdots & \alpha_{j,r} \end{pmatrix} \begin{pmatrix} o_{i,1} \\ o_{i,2} \\ \vdots \\ o_{i,r} \end{pmatrix} = \sum_{l=1}^{r} \alpha_{j,l} o_{i,l},$$

where $o_{i,1}, \ldots, o_{i,r}$ are the sub-blocks of the data file, $i = \left\lfloor \frac{j-1}{2^r-1} \right\rfloor + 1$, $\alpha_{j,l} \in F_2 (1 \leq l \leq \mathrm{r})$, and $\begin{pmatrix} \alpha_{j,1} & \alpha_{j,2} & \cdots & \alpha_{j,r} \end{pmatrix}$ is the binary representation of $j - \left\lfloor \frac{j-1}{2^r-1} \right\rfloor (2^r - 1)$ and $\lfloor \ \rfloor$ represents the integer floor.

5. The system in claim 3 wherein the simplex coding further comprises an added all-ones vector and then an overall parity check.

~~6. A distributed storage system DSS comprising~~

~~*h* non-empty nodes; and~~

~~data stored non-homogenously across the non-empty nodes according to the storing codes (*n*,*k*).~~

~~7. The system in claim 6 wherein the *h* non-empty nodes each having respective non-homogenous bandwidths.~~

~~8. The system in claim 6 or 7 wherein one of the non-empty nodes is a super-node with a significantly higher bandwidth, reliability and/or storage capacity than the remaining non-empty nodes, and a significantly higher proportion of the data is stored on the super-node.~~

~~9. The system in claim 8 wherein the super-node is a local host.~~

~~10. The system in claim 8 wherein the super-node is configured to store two or more systematic data sub-blocks using Maximum Distance Separable Coding.~~

~~11. The system in claim 8 wherein the super-node is configured to store two or more systematic data sub-blocks using Non Maximum Distance Separable Coding.~~

~~12. The system in claim 8 wherein the super-node is configured to store two or more parity data sub-blocks using Maximum Distance Separable Coding.~~

13. The system in claim 6 or 7 configured to optimise the distribution of data across the non-empty nodes to minimise the repair bandwidth and/or the download cost.

14. The system in claim 13 where h non-empty nodes store the same amount of information.

15. The system in claim 13 where h-1 non-empty nodes store the same amount of information.

16. The system in claim 6 or 7 further comprising a plurality of empty nodes and the method further comprising minimising h.

17. A method for determining linear erasure codes with local repairability comprising

selecting two or more coding parameters including r and δ;

determining if an optimal $[n, k, \delta]$ code having all-symbol (r, δ) locality ("$(r, \delta)_a$") exists for the selected r, δ; and

if the optimal $(r, \delta)_a$ code exists performing a local repairable code using the optimal $(r, \delta)_a$ code.

18. The method in claim 17 wherein the coding parameters further including n and k.

19. The method in claim 18 further comprising determining the lower bound of the required field size.

20. The method in any of claims 17 to 19 wherein when the coding parameters satisfy

a) $w \geq r + \delta - 1 - m$ and $r - v \geq u$

b) $w + 1 \geq 2(r + \delta - 1 - m)$ and $2(r - v) \geq u$

c) $r + \delta - 1) \mid n$, or

d) $m \geq (v + \delta - 1)$

an optimal $(r, \delta)_a$ code exists.

21. The method in claim 20 further comprising determining an optimal $(r, \delta)_a$ code using a first algorithm for (a) and (b) and a second algorithm for (c) and (d).

22. The method in claim 19 wherein the lower bound is determined using $\binom{n}{k-1}$.

23. The method in any of claims 17 to 22 wherein when the coding parameters satisfy

e) $(r + \delta - 1) \nmid n$ and $r \mid k$, or

f) $m < v + \delta - 1$ and $u \geq 2(r - v) + 1$

no optimal $(r, \delta)_a$ code exists.