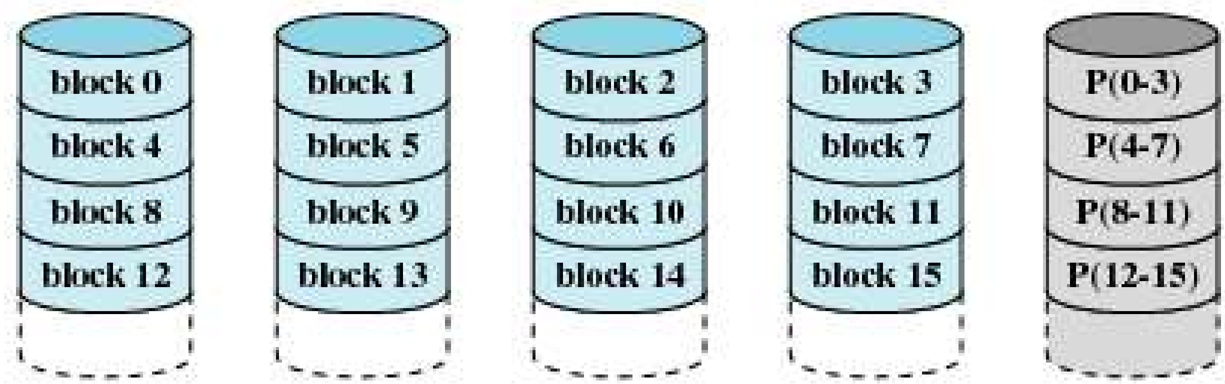
Data Storage Technologies & Networks
(CS C446, CS F446 & IS C446)

Lecture 20– Storage

RAID 4 (block-level parity)



(e) RAID 4 (block-level parity)

- Block inter leaved parity organization
  - Data transfer rate is slow, multiple read accesses can proceed in parallel (higher overall I/O rate)
  - Transfer rate of large reads are high (all disks can be read parallel)
  - Transfer rate of large writes are high (data & parity can be written parallel)

RAID – RAID 4

- RAID 3 enables high data transfer rates but
  - allows only one I/O at a time and
  - may suffer from worst case seek and rotational delays unless disks are synchronized.
- RAID 4
  - 1 bit of parity per D bits of data
    - 1 check disk per D data disks but with *block interleaving*
  - Operations:
    - Read (Normal):
      - Small reads:
        - Not all data disks are to be read
        - Independent reads can be on different (data) disks in parallel for reads smaller than stripe unit
      - Large reads:
        - Similar to RAID 3
    - Read (under 1 disk failure): read from all data disks and the check disk

RAID – RAID 4

- RAID 4
  - Operations:
    - Write :
      - Smaller than a stripe unit:
        - New parity = (Old data **XOR** New data) **XOR** old parity
          - So, 2 read operations and 2 write operations (1 data disk and 1 parity disk)
      - Parity disk is a bottleneck

- New parity (P_{ABCD}`) can be calculated from old parity ($P_{ABCD}$) and the difference between old data (D) and new data(D`)
- P_{ABCD}` = $P_{ABCD}$ XOR Δ where Δ = D XOR D`
- P_{ABCD} XOR D XOR D` = $P_{ABCD}$`

  P_{ABCD}` = A XOR B XOR C XOR D`
- Proof
  - P_{ABCD}` = A XOR B XOR C XOR D`

    D XOR D = 0.

    XOR with 0 will not change the original value
  - P_{ABCD}` = A XOR B XOR C XOR D` XOR [D XOR D]

    = A XOR B XOR C XOR D XOR [D XOR D`]

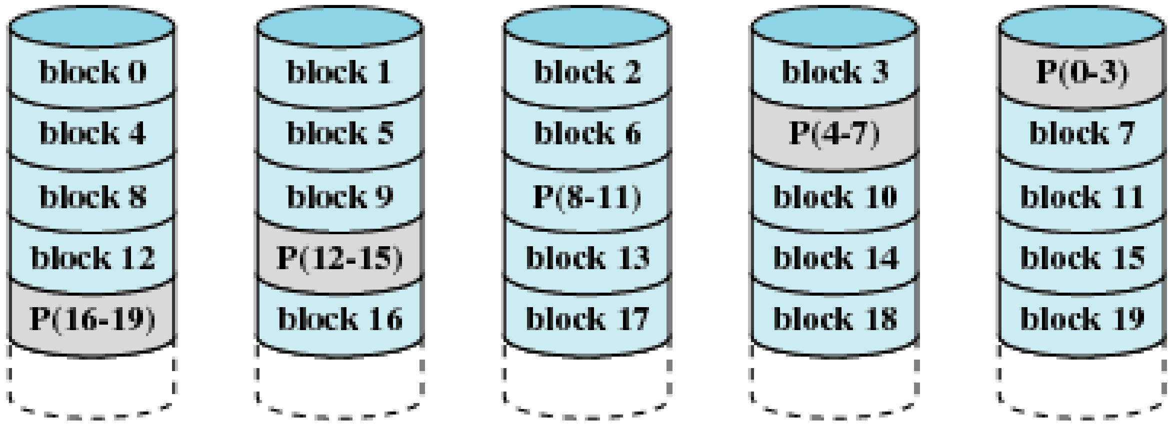
    = P_{ABCD} XOR [D XOR D`]

Example

A XOR B XOR C	D	D`	P_{ABCD}	D XOR D`	P_{ABCD} XOR D XOR D`	A XOR B XOR C XOR D`
0	0	0	0	0	0	0
0	0	1	0	1	1	1
0	1	0	1	1	0	0
0	1	1	1	0	1	1
1	0	0	1	0	1	1
1	0	1	1	1	0	0
1	1	0	0	1	1	1
1	1	1	0	0	0	0

$$P_{ABCD}` = \text{A XOR B XOR C XOR D`}$$
$$= P_{ABCD} \text{ XOR D XOR D`}$$

RAID 5 (block-level distributed parity)



(f) RAID 5 (block-level distributed parity)

- Block inter leaved distributed parity organization
  - Spread data & parity in among N+1 disks
  - Avoids potential over use of a single parity disk
  - Data transfer rate is slow, multiple read accesses can proceed in parallel (higher overall I/O rate)
  - Transfer rate of large reads are high (all disks can be read parallel), Transfer rate of large writes are high (data & parity can be written parallel)

RAID – RAID 5

- RAID 4 is rarely used commercially – because of the “parity disk bottleneck”
- RAID 5
  - Similar to RAID 4 in the sense that
    - 1 parity block is used per D data blocks (i.e. block interleaving)
  - But the parity blocks are not all placed in the same disk
    - They are distributed among all disks.
  - Different distributions are possible
    - Parity placement schemes
  - Preferred for messaging, data mining, RDBMS etc ...

RAID 5 - Parity Placement

- Right Asymmetric placement
  - Start with a RAID 0 (striping) arrangement
    - Insert parity blocks in (left-to-right) diagonal positions and right rotate the data blocks in a stripe

Column (disk)

P0	0	1	2	3
4	P1	5	6	7
8	9	P2	10	11
12	13	14	P3	15
16	17	18	19	P4

Row (stripe)

Right Asymmetric placement

- Derived from RAID 0 placement
- Push out data stripe units horizontally as parity stripe units are inserted
- For each successive parity stripe, the point at which the parity stripe unit is inserted is rotated one stripe unit towards right

RAID level 0

0	1	2	3	4
5	6	7	8	9
10	11	12	13	14
15	16	17	18	19
20	21	22	23	24

P0	0	1	2	3
4	P1	5	6	7
8	9	P2	10	11
12	13	14	P3	15
16	17	18	19	P4

(Right-Asymmetric)

RAID 5 - Parity Placement

- Left Asymmetric placement
  - Start with a RAID 0 (striping) arrangement
    - Insert parity blocks in (right-to-left) diagonal positions and left rotate the data blocks in a strip

0	1	2	3	P0
4	5	6	P1	7
8	9	P2	10	11
12	P3	13	14	15
P4	16	17	18	19

Left Asymmetric placement

- Derived from RAID 0 placement
- Push out data stripe units horizontally as parity stripe units are inserted
- For each successive parity stripe, the point at which the parity stripe unit is inserted is rotated one stripe unit towards left

RAID level 0

0	1	2	3	4
5	6	7	8	9
10	11	12	13	14
15	16	17	18	19
20	21	22	23	24

0	1	2	3	P0
4	5	6	P1	7
8	9	P2	10	11
12	P3	13	14	15
P4	16	17	18	19

(Left-Asymmetric)

RAID 5 - Parity Placement

- Right Symmetric placement
  - Start with a RAID 4 arrangement
    - Align parity blocks in (left-to-right) diagonal positions by right rotating the stripe

P0	0	1	2	3
7	P1	4	5	6
10	11	P2	8	9
13	14	15	P3	12
16	17	18	19	P4

Right Symmetric placement

- Derived from RAID 4 placement
- Derived by right rotations of entire parity stripes from the RAID 4 placement

RAID level 4

0	1	2	3	P0
4	5	6	7	P1
8	9	10	11	P2
12	13	14	15	P3
16	17	18	19	P4

Symmetric

P0	0	1	2	3
7	P1	4	5	6
10	11	P2	8	9
13	14	15	P3	12
16	17	18	19	P4

(Right-Symmetric)

RAID 5 - Parity Placement

- Left Symmetric placement
  - Start with a RAID 4 arrangement
    - Align parity blocks in (right-to-left) diagonal positions by left rotating the stripe

0	1	2	3	P0
5	6	7	P1	4
10	11	P2	8	9
15	P3	12	13	14
P4	16	17	18	19

Left Symmetric placement

- Derived from RAID 4 placement
- Derived by left rotations of entire parity stripes from the RAID 4 placement

RAID level 4

0	1	2	3	P0
4	5	6	7	P1
8	9	10	11	P2
12	13	14	15	P3
16	17	18	19	P4

0	1	2	3	P0
5	6	7	P1	4
10	11	P2	8	9
15	P3	12	13	14
P4	16	17	18	19

(Left-Symmetric)

RAID 5 - Parity Placement

- Extended Left Symmetric
  - Derived from RAID 0 placement
  - Derived by pushing out data stripe units vertically as parity stripe units are inserted
  - For each successive parity stripe, the point at which the parity stripe unit is inserted is rotated one stripe unit towards the left
  - In arrays with only one row of disks
    - Extended left symmetric is identical to left symmetric

Extended-Left-Symmetric

0	1	2	3	P0
10	11	P2	13	4
P4	21	12	23	14
20	31	22	P6	24
30	P8	32	33	34

5	6	7	P1	9
15	P3	17	8	19
25	16	27	18	P5
35	26	P7	28	29
P9	36	37	38	39

- In arrays with multiple rows of disks
  - If each five element column represents a distinct disk, then P1 is the parity corresponding to stripe units 4, 5, 6 and 7 NOT 5, 6, 7 and 9.
  - P2 corresponds to stripe units 8, 9, 10 and 11

Row 0

0	1	2	3	4
10	11	12	13	14
20	21	22	23	24
30	31	32	33	34

Row1

5	6	7	8	9
15	16	17	18	19
25	26	27	28	29
35	36	37	38	39

Step 1

0	1	2	3	P0
10	11	12	13	4
20	21	22	23	14
30	31	32	33	24
				34

5	6	7	8	9
15	16	17	18	19
25	26	27	28	29
35	36	37	38	39

Step 2

0	1	2	3	P0
10	11	12	13	4
20	21	22	23	14
30	31	32	33	24
				34

5	6	7	P1	9
15	16	17	8	19
25	26	27	18	29
35	36	37	28	39
			38	

Step 3

0	1	2	3	P0
10	11	P2	13	4
20	21	12	23	14
30	31	22	33	24
		32		34

5	6	7	P1	9
15	16	17	8	19
25	26	27	18	29
35	36	37	28	39
			38	

Step 4

RAID 5 - Parity Placement

- Flat Left Symmetric
  - Derived from the extended left symmetric placement by grouping all the parity together and placing them at identical offsets within each disk

Flat-Left-Symmetric

0	1	2	3	4
5	6	7	8	9
10	11	12	13	14
15	16	17	18	19
P4	P3	P2	P1	P0

Flat Left Symmetric

- Reading
  - All the disk heads with in the same row skip over parity at the same time
    - Reduces synchronization time
- Writing
  - Performance is likely to be worse than the extended left symmetric placement
    - Parity stripe unit is located at a different offset with in the disk relative to its corresponding data stripe units

RAID 5 – Parity Placement

- Desirable Placement Properties
  - Orthogonal RAID requirement:
    - Stripe units belonging to the same (parity) stripe should not be placed in the same column
  - Parity and data should be distributed over all disks