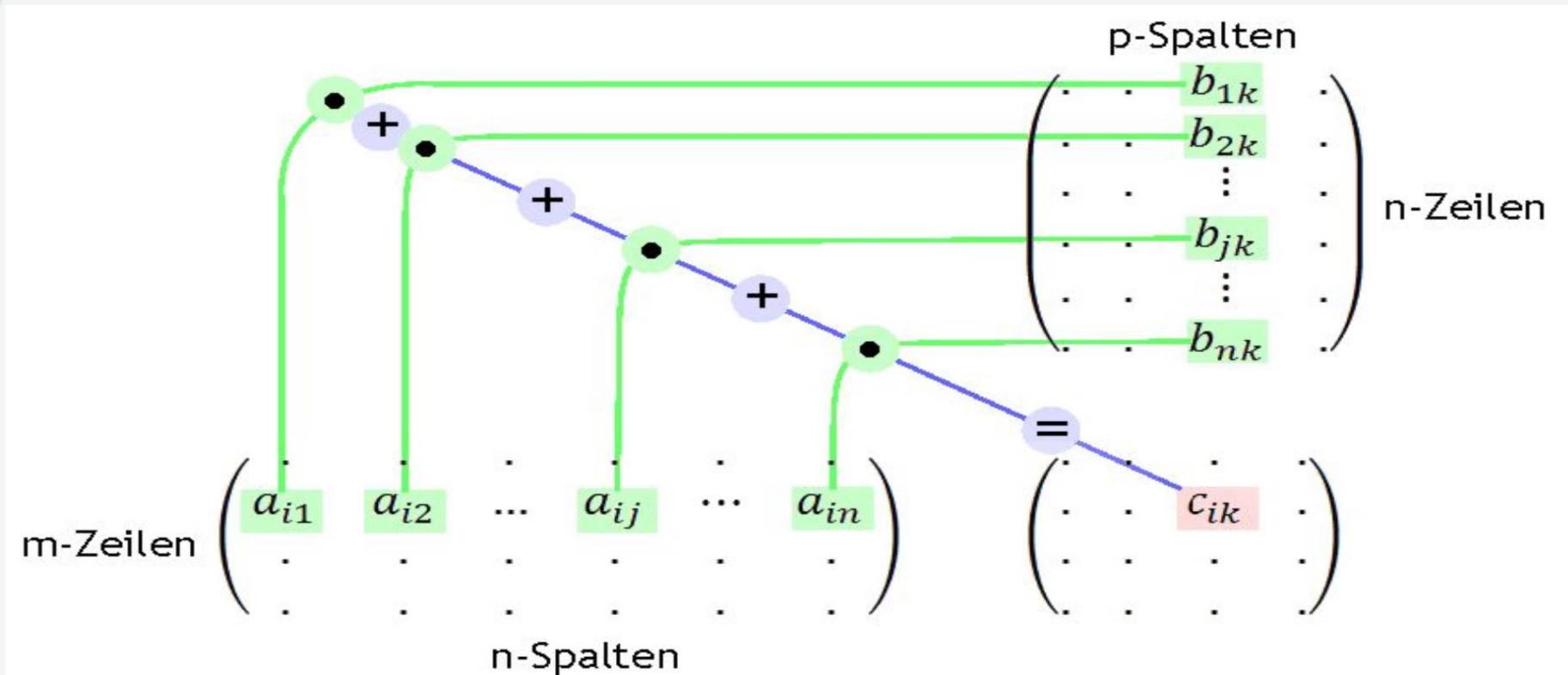


# Matrizenmultiplikation

- Charakteristik
- Speicherung: Stack oder Heap
- Datentypen: Int , Float, Double
- Verbesserung der Lokalität

# Matrizenmultiplikation



Vorname: Nouralrahman  
Nachname: Hussain

# Charakteristik

- Das Programm allokiert Platz in Heap oder Stack für 3 Matrizen
- Jede Matrix hat  $(1000*1000*4\text{byte})=4$  Megabyte Speichergröße. Bei manchen Versuchen ist die Zahl kleiner, sodass es mit dem Speichergröße anpasst.
- Installierung Erfolg durch random-numbers.
- Es werden 2 Matrizen miteinander multipliziert und das Ergebnis wird in einem 3ten Matrix gespeichert.

# Stack Heap

Heap

```
1  #include <stdint.h>
2  #include <stdio.h>
3  #include <stdlib.h>
4  #include <time.h>
5
6  #define N 1000 // Matrix size
7
8  uint64_t compute_timediff(clock_t start, clock_t end, uint8_t print) {
9      uint64_t ticks = end - start;
10     uint64_t nanosec = (ticks * 1000000000) / CLOCKS_PER_SEC;
11     if (print)
12         printf("NS elapsed: %llu ns\n", (unsigned long long)nanosec);
13     return nanosec;
14 }
15
16 int main() {
17     int i, j, k, sum;
18     clock_t start_alloc, end_alloc, start_init, end_init, start_mult, end_mult;
19
20     // Allokation der Matrizen
21     start_alloc = clock();
22     int **matrix_a = (int **)malloc(N * sizeof(int *));
23     int **matrix_b = (int **)malloc(N * sizeof(int *));
24     int **matrix_c = (int **)malloc(N * sizeof(int *));
25     for (i = 0; i < N; i++) {
26         matrix_a[i] = (int *)malloc(N * sizeof(int));
27         matrix_b[i] = (int *)malloc(N * sizeof(int));
28         matrix_c[i] = (int *)malloc(N * sizeof(int));
29     }
30     end_alloc = clock();
31
32     // Initialisierung der Matrizen
33     start_init = clock();
34     for (i = 0; i < N; i++) {
35         for (j = 0; j < N; j++) {
36             matrix_a[i][j] = rand() % 100;
37             matrix_b[i][j] = rand() % 100;
38             matrix_c[i][j] = 0;
39         }
40     }
41     end_init = clock();
42
43     // Matrixmultiplikation
44     start_mult = clock();
45     for (i = 0; i < N; i++) {
46         for (j = 0; j < N; j++) {
47             sum = 0;
48             for (k = 0; k < N; k++) {
49                 sum += matrix_a[i][k] * matrix_b[k][j];
50             }
51             matrix_c[i][j] = sum;
52         }
53     }
54     end_mult = clock();
55
56     // Berechnung der Zeiten
57     // printf("Allocieren\n");
58     compute_timediff(start_alloc, end_alloc, 1);
59     //printf("Initialisieren\n");
60     compute_timediff(start_init, end_init, 1);
61     //printf("Multiplizieren\n");
62     compute_timediff(start_mult, end_mult, 1);
63     //printf("wobei die Laenge der Matrizen = 1000\n");
64
65     // Freigabe des Speichers
66     for (i = 0; i < N; i++) {
67         free(matrix_a[i]);
68         free(matrix_b[i]);
69         free(matrix_c[i]);
70     }
71     free(matrix_a);
72     free(matrix_b);
73     free(matrix_c);
74
75     return 0;
76 }
77
```

Stack

```
1  #include <stdint.h>
2  #include <stdio.h>
3  #include <stdlib.h>
4  #include <time.h>
5
6  #define N 1000 // Matrix size
7
8  uint64_t compute_timediff(clock_t start, clock_t end, uint8_t print) {
9      uint64_t ticks = end - start;
10     uint64_t nanosec = (ticks * 1000000000) / CLOCKS_PER_SEC;
11     if (print)
12         printf("NS elapsed: %llu ns\n", (unsigned long long)nanosec);
13     return nanosec;
14 }
15
16 int main() {
17     int i, j, k, sum;
18     clock_t start_alloc, end_alloc, start_init, end_init, start_mult, end_mult;
19
20     // Allokation der Matrizen
21     start_alloc = clock();
22     int matrix_a[N][N];
23     int matrix_b[N][N];
24     int matrix_c[N][N];
25     end_alloc = clock();
26
27     // Initialisierung der Matrizen
28     start_init = clock();
29     for (i = 0; i < N; i++) {
30         for (j = 0; j < N; j++) {
31             matrix_a[i][j] = rand() % 100;
32             matrix_b[i][j] = rand() % 100;
33             matrix_c[i][j] = 0;
34         }
35     }
36     end_init = clock();
37
38     // Matrixmultiplikation
39     start_mult = clock();
40     sum = 0;
41     for (i = 0; i < N; i++) {
42         for (j = 0; j < N; j++) {
43             sum = 0;
44             for (k = 0; k < N; k++) {
45                 sum += matrix_a[i][k] * matrix_b[k][j];
46             }
47             matrix_c[i][j] = sum;
48         }
49     }
50     end_mult = clock();
51
52     // Berechnung der Zeiten
53     //printf("Allocieren\n");
54     compute_timediff(start_alloc, end_alloc, 1);
55     //printf("Initialisieren\n");
56     compute_timediff(start_init, end_init, 1);
57     //printf("Multiplizieren\n");
58     compute_timediff(start_mult, end_mult, 1);
59     //printf("wobei die Laenge der Matrizen = 1000\n");
60
61     return 0;
62 }
63
```

# Stack

## Vorteile

- Schnellere Allokation: "Push" und "Pop",
- Deterministische Speicherverwaltung: reproduzierbar. Minimierung von Fehlern wegen automatischer Bearbeitung

## Nachteile

- Begrenzte Größe: Stack Overflow
- Statische Speicherzuweisung: Zur Kompilierzeit bekannt

In den meisten Fällen soll nicht zu sehr auf die Geschwindigkeit der Speicherallokation konzentriert werden, da moderne Compiler und Laufzeitsysteme optimierte Mechanismen für die Speicherverwaltung verwenden, um die Leistung zu maximieren.

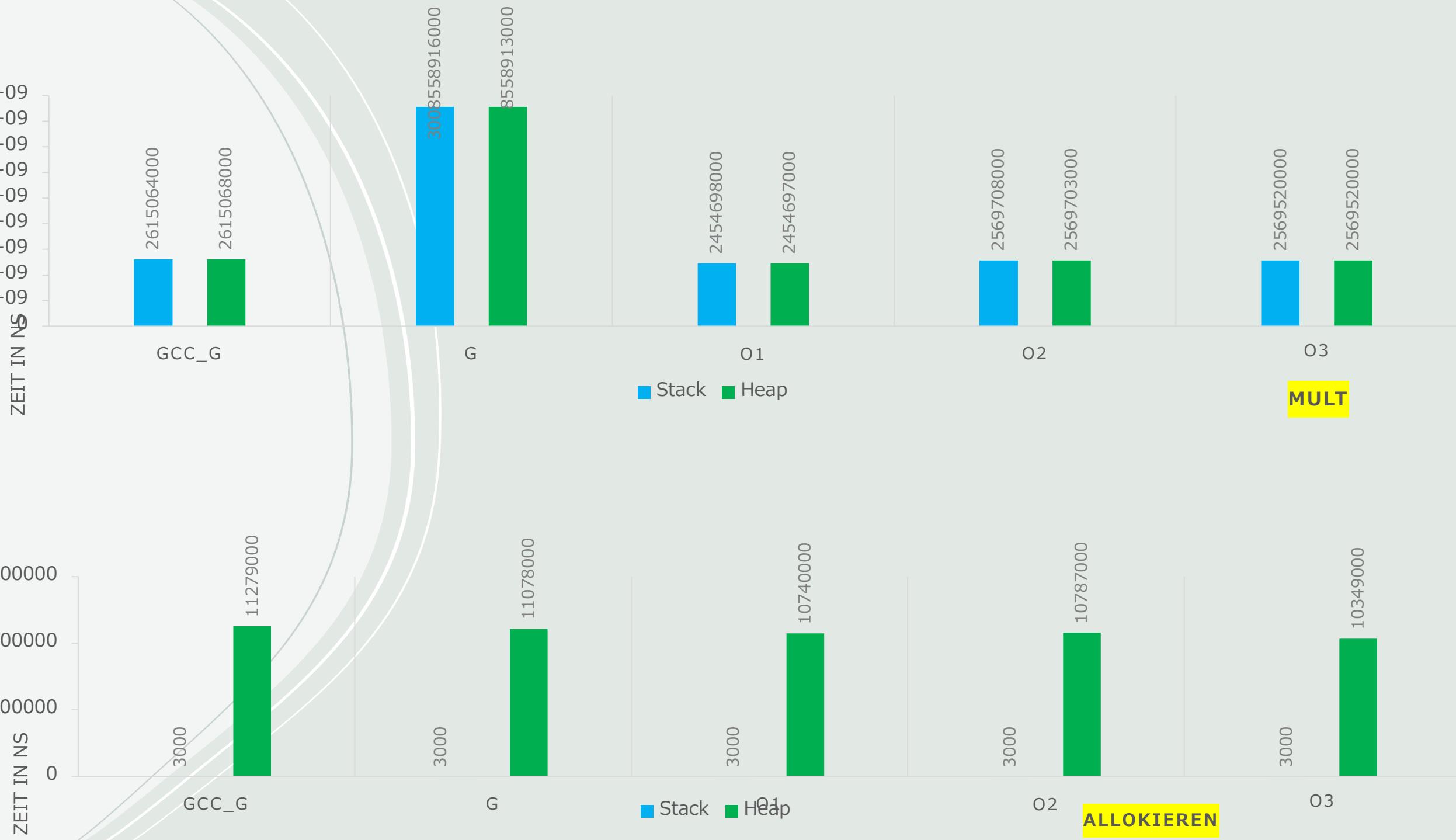
# Heap

## Vorteile

- Dynamische Speicherallokation: flexibler verwalten
- Größere Speicherkapazität: Im Vergleich zum Stack kann der Heap normalerweise eine größere Menge an Speicher zur Verfügung

## Nachteile

- Langsamere Allokation: Suche geeigneten Speicherblock
- Aktualisierung des Heap-Managers
- Manuelle Speicherverwaltung: "Memory Leak"



# Datentypen

int

```

1 #include <stdint.h>
2 #include <stdio.h>
3 #include <stdlib.h>
4 #include <time.h>
5
6 #define N 1000 // Matrix size
7
8 uint64_t compute_timediff(clock_t start, clock_t end, uint8_t print) {
9     uint64_t ticks = end - start;
10    uint64_t nanosec = (ticks * 1000000000) / CLOCKS_PER_SEC;
11    if (print)
12        printf("NS elapsed: %llu ns\n", (unsigned long long)nanosec);
13    return nanosec;
14 }
15
16 int main() {
17     int i, j, k, sum;
18     clock_t start_alloc, end_alloc, start_init, end_init, start_mult, end_mult;
19
20     // Allokation der Matrizen
21     start_alloc = clock();
22     int **matrix_a = (int **)malloc(N * sizeof(int *));
23     int **matrix_b = (int **)malloc(N * sizeof(int *));
24     int **matrix_c = (int **)malloc(N * sizeof(int *));
25     for (i = 0; i < N; i++) {
26         matrix_a[i] = (int *)malloc(N * sizeof(int));
27         matrix_b[i] = (int *)malloc(N * sizeof(int));
28         matrix_c[i] = (int *)malloc(N * sizeof(int));
29     }
30     end_alloc = clock();
31
32     // Initialisierung der Matrizen
33     start_init = clock();
34     for (i = 0; i < N; i++) {
35         for (j = 0; j < N; j++) {
36             matrix_a[i][j] = rand() % 100;
37             matrix_b[i][j] = rand() % 100;
38             matrix_c[i][j] = 0;
39         }
40     }
41     end_init = clock();
42
43     // Matrixmultiplikation
44     start_mult = clock();
45     for (i = 0; i < N; i++) {
46         for (j = 0; j < N; j++) {
47             sum = 0;
48             for (k = 0; k < N; k++) {
49                 sum += matrix_a[i][k] * matrix_b[k][j];
50             }
51             matrix_c[i][j] = sum;
52         }
53     }
54     end_mult = clock();
55
56     // Berechnung der Zeiten
57     // printf("Allocieren\n");
58     // compute_timediff(start_alloc, end_alloc, 1);
59     // printf("Initialisieren\n");
60     // compute_timediff(start_init, end_init, 1);
61     // printf("Multiplizieren\n");
62     // compute_timediff(start_mult, end_mult, 1);
63     // printf("wobei die Laenge der Matrizen = 1000\n");
64
65     // Freigabe des Speichers
66     for (i = 0; i < N; i++) {
67         free(matrix_a[i]);
68         free(matrix_b[i]);
69         free(matrix_c[i]);
70     }
71     free(matrix_a);
72     free(matrix_b);
73     free(matrix_c);
74
75     return 0;
76 }

```

float

```

1 #include <stdint.h>
2 #include <stdio.h>
3 #include <stdlib.h>
4 #include <time.h>
5
6 #define N 1000 // Matrix size
7
8 uint64_t compute_timediff(clock_t start, clock_t end, uint8_t print) {
9     uint64_t ticks = end - start;
10    uint64_t nanosec = (ticks * 1000000000) / CLOCKS_PER_SEC;
11    if (print)
12        printf("%llu ns\n", (unsigned long long)nanosec);
13    return nanosec;
14 }
15
16 int main() {
17     int i, j, k;
18     float sum;
19     clock_t start_alloc, end_alloc, start_init, end_init, start_mult, end_mult;
20
21     // Allocation of matrices
22     start_alloc = clock();
23     float **matrix_a = (float **)malloc(N * sizeof(float *));
24     float **matrix_b = (float **)malloc(N * sizeof(float *));
25     float **matrix_c = (float **)malloc(N * sizeof(float *));
26     for (i = 0; i < N; i++) {
27         matrix_a[i] = (float *)malloc(N * sizeof(float));
28         matrix_b[i] = (float *)malloc(N * sizeof(float));
29         matrix_c[i] = (float *)malloc(N * sizeof(float));
30     }
31     end_alloc = clock();
32
33     // Initialization of matrices
34     start_init = clock();
35     srand(time(NULL));
36     for (i = 0; i < N; i++) {
37         for (j = 0; j < N; j++) {
38             matrix_a[i][j] = ((float)rand() / RAND_MAX) * 100.0f;
39             matrix_b[i][j] = ((float)rand() / RAND_MAX) * 100.0f;
40             matrix_c[i][j] = 0.0f;
41         }
42     }
43     end_init = clock();
44
45     // Matrix multiplication
46     start_mult = clock();
47     for (i = 0; i < N; i++) {
48         for (j = 0; j < N; j++) {
49             sum = 0.0f;
50             for (k = 0; k < N; k++) {
51                 sum += matrix_a[i][k] * matrix_b[k][j];
52             }
53             matrix_c[i][j] = sum;
54         }
55     }
56     end_mult = clock();
57
58     // Calculation of times
59     // printf("Allocation\n");
60     compute_timediff(start_alloc, end_alloc, 1);
61     // printf("Initialization\n");
62     compute_timediff(start_init, end_init, 1);
63     // printf("Multiplication\n");
64     compute_timediff(start_mult, end_mult, 1);
65     // printf("where the length of matrices = 1000\n");
66
67     // Freeing memory
68     for (i = 0; i < N; i++) {
69         free(matrix_a[i]);
70         free(matrix_b[i]);
71         free(matrix_c[i]);
72     }
73     free(matrix_a);
74     free(matrix_b);
75     free(matrix_c);
76
77     return 0;
78 }

```

double

```

1 #include <stdint.h>
2 #include <stdio.h>
3 #include <stdlib.h>
4 #include <time.h>
5
6 #define N 1000 // Matrix size
7
8 uint64_t compute_timediff(clock_t start, clock_t end, uint8_t print) {
9     uint64_t ticks = end - start;
10    uint64_t nanosec = (ticks * 1000000000) / CLOCKS_PER_SEC;
11    if (print)
12        printf("NS elapsed: %llu ns\n", (unsigned long long)nanosec);
13    return nanosec;
14 }
15
16 int main() {
17     int i, j, k;
18     double sum;
19     clock_t start_alloc, end_alloc, start_init, end_init, start_mult, end_mult;
20
21     // Allocation of matrices
22     start_alloc = clock();
23     double **matrix_a = (double **)malloc(N * sizeof(double *));
24     double **matrix_b = (double **)malloc(N * sizeof(double *));
25     double **matrix_c = (double **)malloc(N * sizeof(double *));
26     for (i = 0; i < N; i++) {
27         matrix_a[i] = (double *)malloc(N * sizeof(double));
28         matrix_b[i] = (double *)malloc(N * sizeof(double));
29         matrix_c[i] = (double *)malloc(N * sizeof(double));
30     }
31     end_alloc = clock();
32
33     // Initialization of matrices
34     start_init = clock();
35     srand(time(NULL));
36     for (i = 0; i < N; i++) {
37         for (j = 0; j < N; j++) {
38             matrix_a[i][j] = ((double)rand() / RAND_MAX) * 100.0;
39             matrix_b[i][j] = ((double)rand() / RAND_MAX) * 100.0;
40             matrix_c[i][j] = 0.0;
41         }
42     }
43     end_init = clock();
44
45     // Matrix multiplication
46     start_mult = clock();
47     for (i = 0; i < N; i++) {
48         for (j = 0; j < N; j++) {
49             sum = 0.0;
50             for (k = 0; k < N; k++) {
51                 sum += matrix_a[i][k] * matrix_b[k][j];
52             }
53             matrix_c[i][j] = sum;
54         }
55     }
56     end_mult = clock();
57
58     // Calculation of times
59     // printf("Allocation\n");
60     compute_timediff(start_alloc, end_alloc, 1);
61     // printf("Initialization\n");
62     compute_timediff(start_init, end_init, 1);
63     // printf("Multiplication\n");
64     compute_timediff(start_mult, end_mult, 1);
65     // printf("where the length of matrices = 1000\n");
66
67     // Freeing memory
68     for (i = 0; i < N; i++) {
69         free(matrix_a[i]);
70         free(matrix_b[i]);
71         free(matrix_c[i]);
72     }
73     free(matrix_a);
74     free(matrix_b);
75     free(matrix_c);
76
77     return 0;
78 }

```

# Datentyp

## Int

- 4byte 32bit
- ohne Genauigkeitsverlust
- weniger Speicherplatz
- schnellere Berechnungen
- Eingeschränkte Darstellung von Dezimalzahlen

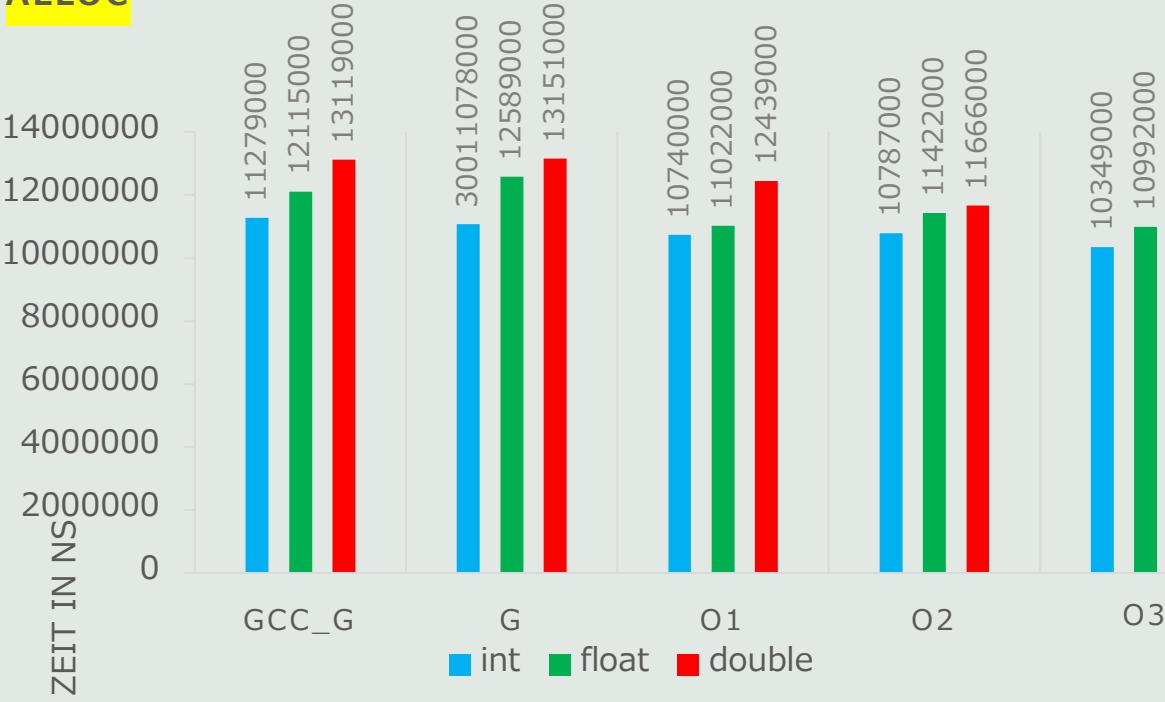
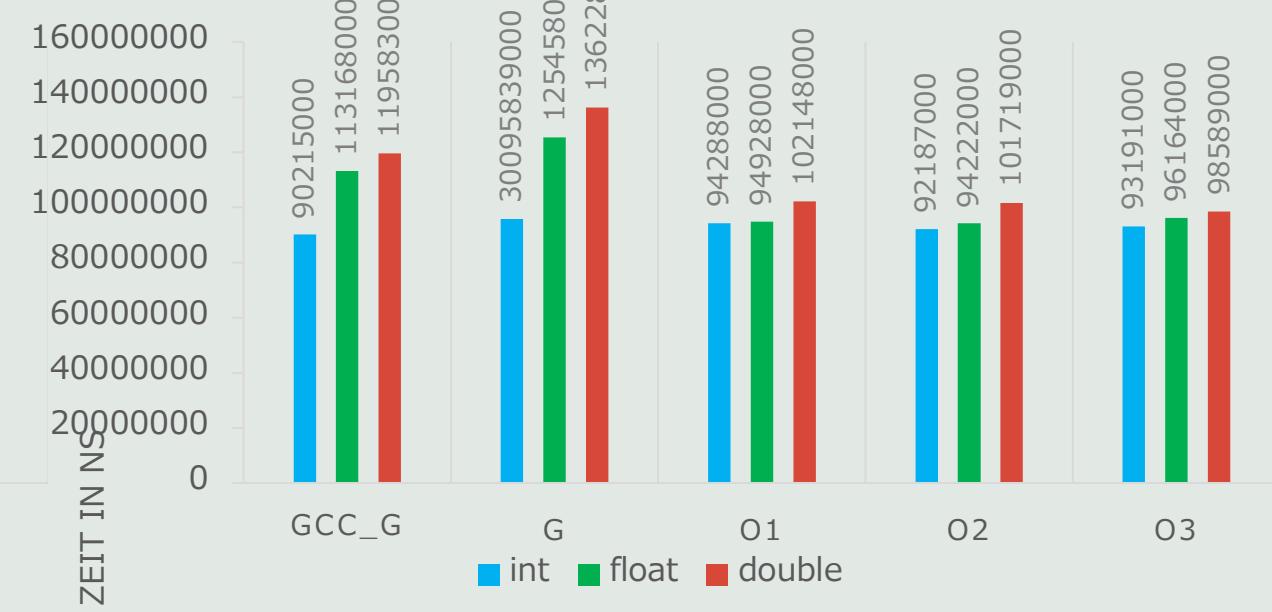
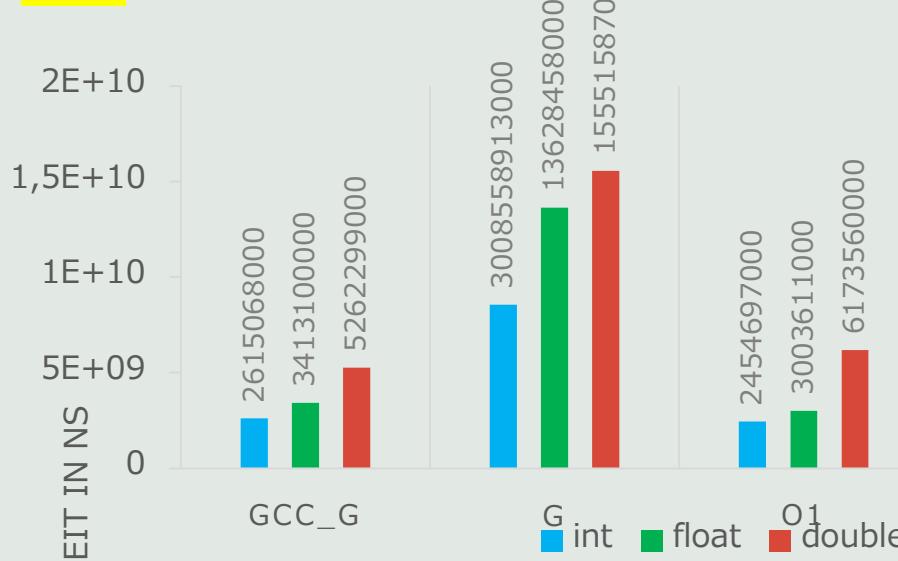
## Float

- 4byte 32bit
- Genauigkeitsverlust
- Speicherplatz mehr als int in der Regel
- Langsame Berechnungen
- Unterstützung für mathematische Funktionen
- Eingeschränkter Wertebereich

## Double

- 8byte 64bit
- Hohe Genauigkeit
- Größer Speicherbedarf
- Langsame Berechnungen
- Unterstützung für mathematische Funktionen

Die beste Entscheidung bezüglich der Geschwindigkeit wäre Int.

**ALLOC****INIT****MULT**

# Cache Performance durch Speicherung

```
1  #include <stdint.h>
2  #include <stdio.h>
3  #include <stdlib.h>
4  #include <time.h>
5
6  #define N 1000 // Matrix size
7
8  uint64_t compute_timediff(clock_t start, clock_t end, uint8_t print) {
9      uint64_t ticks = end - start;
10     uint64_t nanosec = (ticks * 1000000000) / CLOCKS_PER_SEC;
11     if (print)
12         printf("NS elapsed: %llu ns\n", (unsigned long long)nanosec);
13     return nanosec;
14 }
15
16 int main() {
17     int i, j, k, sum;
18     clock_t start_alloc, end_alloc, start_init, end_init, start_mult, end_mult;
19
20     // Allokation der Matrizen
21     start_alloc = clock();
22     int **matrix_a = (int **)malloc(N * sizeof(int *));
23     int **matrix_b = (int **)malloc(N * sizeof(int *));
24     int **matrix_c = (int **)malloc(N * sizeof(int *));
25     for (i = 0; i < N; i++) {
26         matrix_a[i] = (int *)malloc(N * sizeof(int));
27         matrix_b[i] = (int *)malloc(N * sizeof(int));
28         matrix_c[i] = (int *)malloc(N * sizeof(int));
29     }
30     end_alloc = clock();
31
32     // Initialisierung der Matrizen
33     start_init = clock();
34     for (i = 0; i < N; i++) {
35         for (j = 0; j < N; j++) {
36             matrix_a[i][j] = rand() % 100;
37             matrix_b[i][j] = rand() % 100;
38             matrix_c[i][j] = 0;
39         }
40     }
41     end_init = clock();
42
43     // Matrixmultiplikation
44     start_mult = clock();
45     for (i = 0; i < N; i++) {
46         for (j = 0; j < N; j++) {
47             sum = 0;
48             for (k = 0; k < N; k++) {
49                 sum += matrix_a[i][k] * matrix_b[k][j];
50             }
51             matrix_c[i][j] = sum;
52         }
53     }
54     end_mult = clock();
55
56     // Berechnung der Zeiten
57     //printf("Allocieren\n");
58     compute_timediff(start_alloc, end_alloc, 1);
59     //printf("Initialisieren\n");
60     compute_timediff(start_init, end_init, 1);
61     //printf("Multiplizieren\n");
62     compute_timediff(start_mult, end_mult, 1);
63     //printf("wobei die Laenge der Matrizen = 1000\n");
64
65     // Freigabe des Speichers
66     for (i = 0; i < N; i++) {
67         free(matrix_a[i]);
68         free(matrix_b[i]);
69         free(matrix_c[i]);
70     }
71     free(matrix_a);
72     free(matrix_b);
73     free(matrix_c);
74
75     return 0;
76
77
1  #include <stdint.h>
2  #include <stdio.h>
3  #include <stdlib.h>
4  #include <time.h>
5
6  #define N 1000 // Matrix size
7
8  uint64_t compute_timediff(clock_t start, clock_t end, uint8_t print) {
9      uint64_t ticks = end - start;
10     uint64_t nanosec = (ticks * 1000000000) / CLOCKS_PER_SEC;
11     if (print)
12         printf("%llu ns\n", (unsigned long long)nanosec);
13     return nanosec;
14 }
15
16 int main() {
17     int i, j, k, sum;
18     clock_t start_alloc, end_alloc, start_init, end_init, start_mult, end_mult;
19
20     // Allokation und Initialisierung der Matrizen
21     start_alloc = clock();
22     int *matrix_a = (int *)malloc(N * N * sizeof(int));
23     int *matrix_b = (int *)malloc(N * N * sizeof(int));
24     int *matrix_c = (int *)calloc(N * N, sizeof(int));
25     end_alloc = clock();
26
27     srand(time(NULL));
28
29     // Initialisierung der Matrizen
30     start_init = clock();
31     for (i = 0; i < N; i++) {
32         for (j = 0; j < N; j++) {
33             matrix_a[i * N + j] = rand() % 100;
34             matrix_b[i * N + j] = rand() % 100;
35         }
36     }
37     end_init = clock();
38
39     // Matrixmultiplikation
40     start_mult = clock();
41     for (i = 0; i < N; i++) {
42         sum = 0;
43         for (j = 0; j < N; j++) {
44             sum = 0;
45             for (k = 0; k < N; k++) {
46                 sum += matrix_a[i * N + k] * matrix_b[k * N + j];
47             }
48             matrix_c[i * N + j] = sum;
49         }
50     }
51     end_mult = clock();
52
53     // Berechnung der Zeiten
54     //printf("Allocieren\n");
55     compute_timediff(start_alloc, end_alloc, 1);
56     //printf("Initialisieren\n");
57     compute_timediff(start_init, end_init, 1);
58     //printf("Multiplizieren\n");
59     compute_timediff(start_mult, end_mult, 1);
60     //printf("wobei die Laenge der Matrizen = 1000\n");
61
62     // Freigabe des Speichers
63     free(matrix_a);
64     free(matrix_b);
65     free(matrix_c);
66
67     return 0;
68 }
```

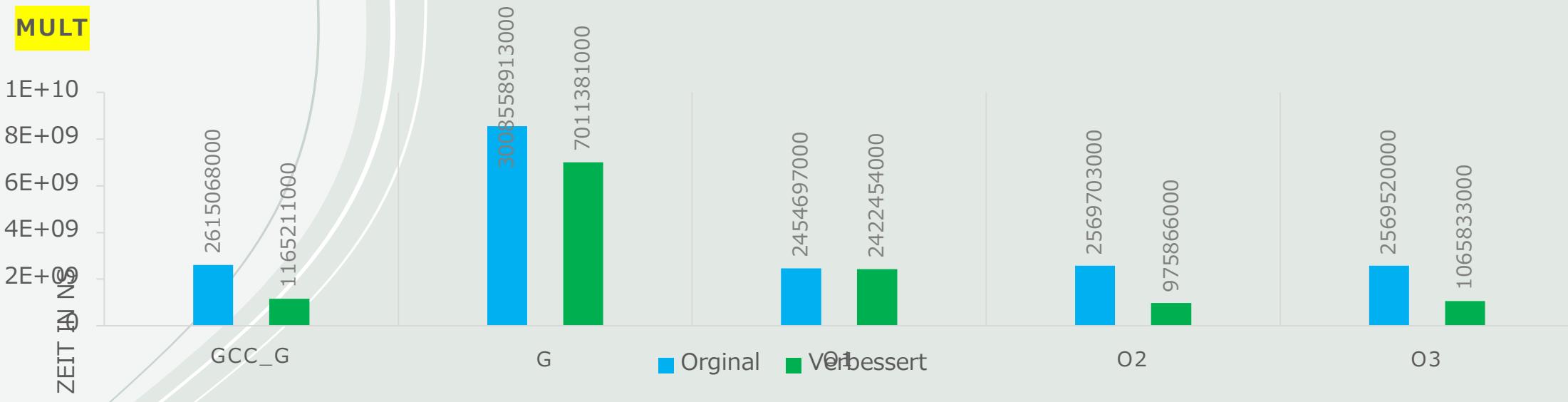
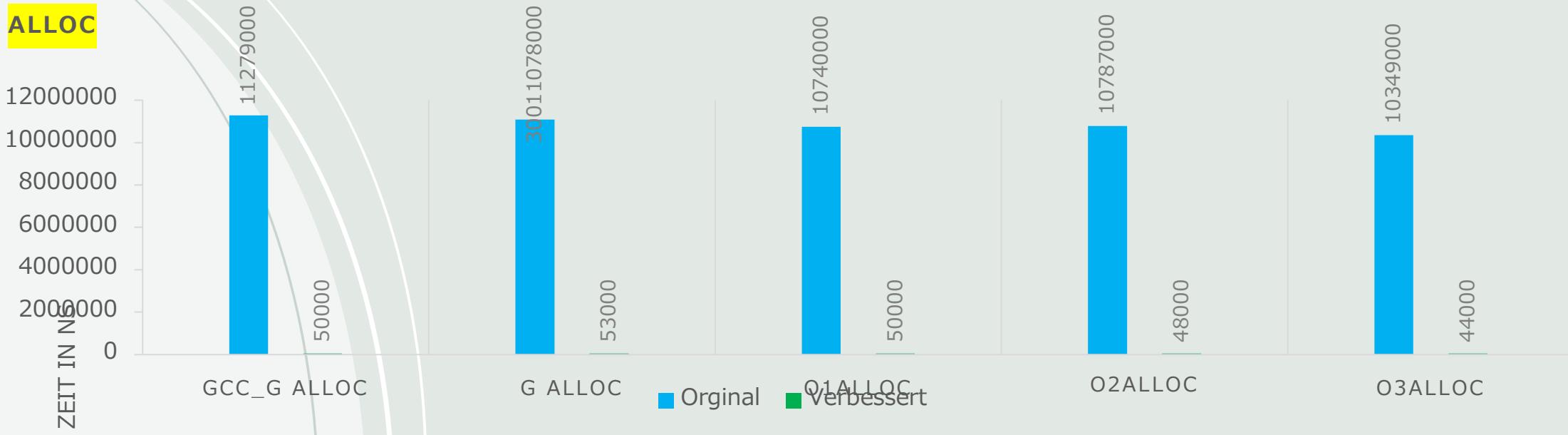
# Original Programm

Hier werden die Matrizen als Arrays von Zeigern auf Zeilen allokiert. Das führt zu einer schlechteren räumlichen Lokalität. Der Zugriff auf die Elemente in der innersten Schleife erfolgt über verschiedene Speicherbereiche, da die Zeiger auf unterschiedliche Zeilen verweisen. Dies kann zu einer höheren Wahrscheinlichkeit von Cache-Misses führen, da die benachbarten Elemente möglicherweise nicht im Cache gehalten werden.

# Verbesserte Programm

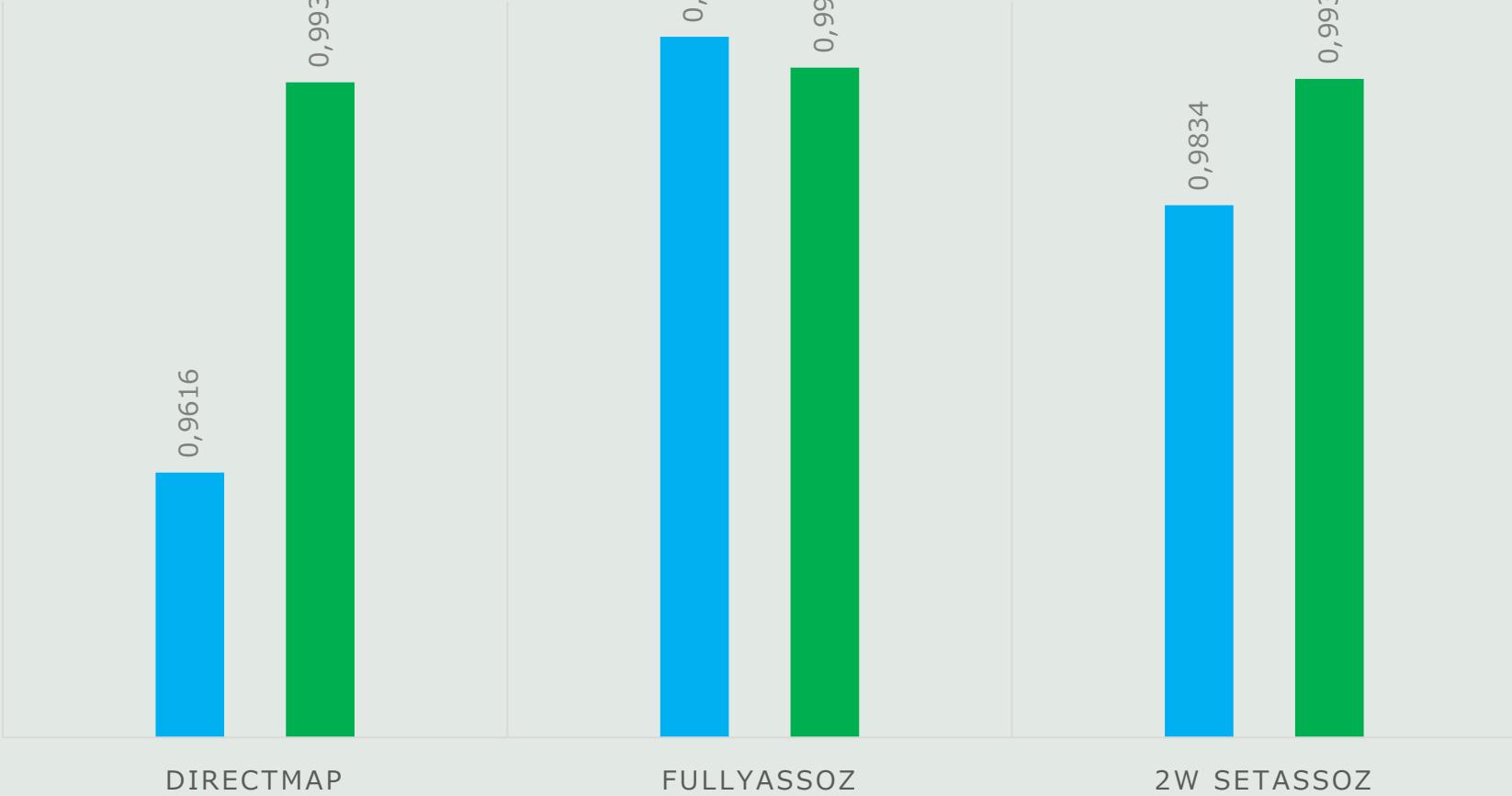
Hier erfolgt der Zugriff auf die Matrizenelemente in einem kontinuierlichen Speicherbereich. Dies kann dazu führen, dass die Elemente, die im Cache gehalten werden, zusammenhängend sind und die Wahrscheinlichkeit von Cache-Misses verringert wird. Dies liegt daran, dass beim Zugriff auf ein Element im eindimensionalen Array benachbarte Elemente ebenfalls im Cache vorhanden sein können.

Zusammenfassend lässt sich sagen, dass das verbesserte Programm in Bezug auf die Speichernutzung und das Cache-Verhalten effizienter ist.



# HIT RATE

■ Orginal ■ Verbessert



File Edit View Help

Ripes

OriginalDM

Cache configuration:

Preset: 32-entry 4-word direct-mapped

2<sup>N</sup> Lines: 5 Repl. policy: LRU

2<sup>N</sup> Ways: 0 Wr. hit: Write-back

2<sup>N</sup> Words/Line: 2 Wr. miss: Write allocate

Plot configuration: Statistics:

Numerator: Hits Denominator: Access count

Size (bits): 4896 Hit rate: 0.9616 Writebacks: 175

Ratio Moving avg. 50 cyc.

Hits: 11438 Misses: 457

Total Moving average

0 2499 4999 7498 9998

0 20 40 60 80 100 %

0 9998

Verbessert DM

Cache configuration:

Preset: 32-entry 4-word direct-mapped

2<sup>N</sup> Lines: 5 Repl. policy: LRU

2<sup>N</sup> Ways: 0 Wr. hit: Write-back

2<sup>N</sup> Words/Line: 2 Wr. miss: Write allocate

Plot configuration: Statistics:

Numerator: Hits Denominator: Access count

Size (bits): 4896 Hit rate: 0.9934 Writebacks: 10

Ratio Moving avg. 50 cyc.

Hits: 8977 Misses: 60

Total Moving average

0 2499 4999 7499 9999

0 20 40 60 80 100 %

0 9999

Access address: 31 98 4321 0  
00000000000000000000000000000000

Index V D Tag Word 0 Word 1 Word 2 Word 3

0	1	0	0x00000000	0x00004ecf	0x00090e80	0x01ac3615	0x8efb0c40
1	1	0	0x00000094	0x00000000	0x00000000	0x0001281c	0x00000000
2	1	1	0x00000094	0x00000000	0x00000000	0x00010424	0x00000000
3	1	1	0x00000000	0x727ea63f	0x9b9f8ac0	0x3e638ac5	0x856daea6
4	1	1	0x00000000	0x01a53334	0xf500b77	0x0ad12df	0x98455988
5	1	1	0x00000000	0xb22843f	0xeca83dc	0xd5facf27	0x6450251
6	1	1	0x00000000	0x2ccbe53	0xa43faaa6	0x954ecb80	0x15d45b6
7	1	1	0x00000000	0x0dff145e3	0x7dc947e	0x6376874	0x379c1bf
8	1	1	0x00000000	0xef1e3b81	0xbe1ce632	0x4cd41cb	0x1db0f4c4
9	1	1	0x00000000	0x1b5e8a4	0xa87a76dac	0x3523abb2	0x2a8f206b
10	1	1	0x00000000	0x65d03d61	0xabdcc09	0x5d3e7541	0x3b6573bb
11	1	1	0x00000000	0x48d7c31	0xbab02a8f	0x36cd844	0x7d62c6d0
12	1	1	0x00000093	0x00000000	0x00000000	0x00010120	0x000100d0
13	1	1	0x00000093	0x92371543	0xd9ef43ef	0x25af78fa	0xa5bf5da5
14	1	1	0x00000000	0x7f6b1e07	0x3134dee9b	0x4e4473d5	0x71f5f5f6
15	1	1	0x00000000	0x2e5a0531	0x8cf4837e	0x2d81f93a	0x00000000
16	1	0	0x00000093	0x00000000	0x00000000	0x00000000	0x00000000
17	1	1	0x00000097	0x00020000	0x00000000	0x00000000	0x00000000
18	1	1	0x00000097	0x00000000	0x00000000	0x00000000	0x00000000
19	1	1	0x00000097	0x00000000	0x00000000	0x00000000	0x00000000
20	1	1	0x00000097	0x00000000	0x00000000	0x00000000	0x00000000
21	1	1	0x00000097	0x00000000	0x00000000	0x00000000	0x00000000
22	1	0	0x00000097	0x00000000	0x00000000	0x00000000	0x00000000
23	1	1	0x00000093	0x00000000	0x00000000	0x95e5e781	0x75e44e66
24	1	1	0x0003ffff	0x00000000	0x00000000	0x00000000	0x00000000
25	1	1	0x0003ffff	0x00000000	0x00000000	0x00000000	0x00000000
26	1	1	0x0003ffff	0x00000000	0x00000000	0x00000000	0x0001101c
27	1	1	0x0003ffff	0x00000000	0x00000000	0x00000000	0x00000000
28	1	1	0x0003ffff	0x00000000	0x00000000	0x00000000	0x00000000
29	1	1	0x0003ffff	0x00000000	0x00000000	0x00000000	0x00010408
30	1	1	0x0003ffff	0x00000000	0x00000000	0x00000000	0x000108cc
31	1	0	0x003fffff	0x00000000	0x00000000	0x00000000	0x00000000

# Warum ist das verbesserte Programm besser bei Direct Mapping Strategie?

Beim Multiplizieren der Matrizen im verbesserten Programm erfolgt der Zugriff auf die Elemente in aufeinanderfolgenden Speicherbereichen. Da jeder Zeilenwert der Matrix einem bestimmten Line im Cache zugeordnet ist, werden auch die benachbarten Elemente auf demselben Line gespeichert.

Dadurch bleiben die benachbarten Elemente im Cache zusammen und können schneller abgerufen werden. Dies erhöht die Wahrscheinlichkeit, dass die benötigten Daten im Cache verfügbar sind, und somit steigt die Cache-Hit-Rate.

Darüber hinaus kann die lineare Allokation der Matrix im verbesserten Programm zu einer besseren räumlichen Lokalität beitragen, da benachbarte Elemente im Cache im selben Cache-Block gespeichert werden. Dies hilft, Cache-Miss-Situationen zu reduzieren.

**Original FA**

File Edit View Help

L1 Data Cache L1 Instr. Cache

Cache configuration:

Preset: 32-entry 4-word fully associative

2<sup>N</sup> Lines: 0 Repl. policy: LRU

2<sup>N</sup> Ways: 5 Wr. hit: Write-back

2<sup>N</sup> Words/Line: 2 Wr. miss: Write allocate

Plot configuration: Statistics:

Numerator: Hits Size (bits): 5216

Denominator: Access count

Ratio

Moving avg. 50 cyc.

Access address: 000000000010011101111100

Cache content:

V	D	LRU	Tag	Word 0	Word 1	Word 2	Word 3
1	1	28	0x000012f8	0x00020000	0x00000000	0x00000000	0x00000000
1	1	8	0x000012f9	0x00000000	0x00000000	0x00000000	0x00000000
1	1	19	0x00000007	0x00000000	0x00000000	0x31000000	0x00000000
1	0	10	0x000012a1	0x00000000	0x00000000	0x00000000	0x00000000
1	0	0	0x00001277	0x00000000	0x00000000	0x00000000	0x00000000
1	1	25	0x000012fd	0x00000000	0x00000000	0x00000000	0x00000000
1	0	1	0x000012f7	0x00012740	0x00000000	0x00012740	0xffffffff
1	1	5	0x00001288	0x00000000	0x00000000	0x0001288c	0x00000000
1	1	6	0x00001289	0x00000000	0x00010490	0x00000000	0x00000000
1	1	11	0x07fffffe	0x00000000	0x00000000	0x00000000	0x000100cc
1	0	9	0x00001273	0x00010074	0x00010120	0x000100d0	0x00000000
1	0	31	0x07ffffff	0x00000000	0x00000000	0x00000000	0x00000000
1	1	2	0x07fffffb	0x00000000	0x0000001c	0x00000000	0x00000000
1	1	7	0x07fffffa	0x00000000	0x00012890	0x00000000	0x00011088
1	1	21	0x07fffff9	0x00000000	0x00000000	0x00000000	0x00000000
1	0	30	0x000012b9	0x00012b88	0x00012b88	0x00012b90	0x00012b90
1	0	29	0x000012ba	0x00012b98	0x00012b98	0x00012ba0	0x00012ba0
1	0	23	0x000012b7	0x00012b68	0x00012b68	0x00012b70	0x00012b70
1	0	22	0x000012b6	0x00000000	0x00000000	0x00000000	0x00000000
1	1	24	0x07fffff8	0x00000000	0x00000000	0x00012b68	0x000108d8
1	1	27	0x07fffff7	0x00000000	0x00000000	0x00000000	0x00010edc
1	1	26	0x00001274	0x0000000c	0x00012a2c	0x00012a94	0x00012afc
1	1	3	0x07fffffc	0x00000000	0x00000000	0x00000000	0x00000000
1	1	4	0x07fffffd	0x00000000	0x00000000	0x00000000	0x00010474
1	1	13	0x00000000	0x00000000	0x00000000	0x00000000	0x00000000
1	1	12	0x00000001	0x00000000	0x00000000	0x00000000	0x00000000
1	1	16	0x0000127e	0x00000000	0x00000000	0x96e5e781	0x76e04ee6
1	1	20	0x00000002	0x00000000	0x00000000	0x00000000	0x00001a00
1	1	15	0x00000003	0x00000000	0x00000000	0x00000000	0x00000000
1	1	14	0x00000004	0x00000000	0x00000000	0x00000000	0x00000000
1	1	17	0x00000005	0x00000000	0x000e0000	0x00000000	0x00000000
1	1	18	0x00000006	0x00000000	0x00012f00	0x00000000	0x25000000

Plot configuration: Statistics:

Numerator: Hits Size (bits): 5216

Denominator: Access count

Ratio

Moving avg. 50 cyc.

Access address: 000000000010011101111100

Cache content:

V	D	LRU	Tag	Word 0	Word 1	Word 2	Word 3
1	1	11	0x7fffffe	0x00000000	0x00000000	0x00000000	0x000100cc
1	0	5	0x00001281	0x00000000	0x00000000	0x0001281c	0x00000000
1	1	24	0x00000006	0x2ccbef53	0xa43faaa6	0xe54ecb80	0x215d45b6
1	1	16	0x00000007	0x0df145e3	0x7dce947e	0x637687e4	0x6379c1bf
1	1	23	0x00000008	0xefe13b81	0xbe1ce632	0xdc4d1cbc	0x1db0f4c4
1	0	9	0x0000126c	0x00000000	0x00010074	0x00010120	0x000100d0
1	0	1	0x000012f0	0x000126d0	0x00000000	0x000126d0	0xffffffff
1	1	15	0x00000009	0x11b5e8a4	0x8a76dac0	0x3532abb2	0x2a8f206b
1	1	22	0x000000a0	0x68d03d61	0xabdcc09	0x5d3e7541	0x3b6573bb
1	1	13	0x000000d0	0x92371543	0xd9ef43ef	0x25af78fa	0xab5fda5d
1	1	14	0x000000b0	0x48d77c31	0xbb02af80	0x36cd844c	0x7d62c6d0
1	1	21	0x000000c0	0xdd2f22d3	0xc4e3db17	0x5cfda276	0xb8f6b4ce
1	1	2	0x07fffffb	0x00000000	0x00000000	0x00000000	0x00000000
1	1	7	0x07fffffa	0x00000000	0x00012890	0x00000000	0x00011088
1	1	21	0x07fffff9	0x00000000	0x00000000	0x00000000	0x00000000
1	0	30	0x000012b9	0x00012b88	0x00012b88	0x00012b90	0x00012b90
1	0	29	0x000012ba	0x00012b98	0x00012b98	0x00012ba0	0x00012ba0
1	0	23	0x000012b7	0x00012b68	0x00012b68	0x00012b70	0x00012b70
1	0	22	0x000012b6	0x00000000	0x00000000	0x00000000	0x00000000
1	1	24	0x07fffff8	0x00000000	0x00000000	0x00012b68	0x000108d8
1	1	27	0x07fffff7	0x00000000	0x00000000	0x00000000	0x00010edc
1	1	26	0x00001274	0x0000000c	0x00012a2c	0x00012a94	0x00012afc
1	1	3	0x07fffffc	0x00000000	0x00000000	0x00000000	0x00000000
1	1	4	0x07fffffd	0x00000000	0x00000000	0x00000000	0x00010474
1	1	13	0x00000000	0x00000000	0x00000000	0x00000000	0x00000000
1	1	12	0x00000001	0x00000000	0x00000000	0x00000000	0x00000000
1	0	31	0x000012af	0x00000000	0x00000000	0x00000000	0x00000000
1	1	29	0x07fffff8	0x00000000	0x00000000	0x00000000	0x00000000
1	1	8	0x000012f2	0x00000000	0x00000000	0x00000000	0x00000000
1	0	10	0x0000129a	0x00000000	0x00000000	0x00000000	0x00000000
1	1	3	0x07fffffc	0x00000000	0x00000000	0x00000000	0x00000000
1	1	6	0x00001282	0x00000000	0x00010424	0x00000000	0x00000000
1	1	4	0x07fffffd	0x00000000	0x00000000	0x00000000	0x00010408
1	1	28	0x00001277	0x00000000	0x00000000	0x96e5e781	0x76e04ee6
1	1	27	0x00000000	0x00004ecf	0x00090e8d	0x01ac3615	0x8efb0c40
1	1	20	0x00000001	0x8801cccd3	0x550282db	0x5c3129ea	0xb2540a96
1	1	26	0x00000002	0x0014ad1a	0x42c4ed9f	0x07d5695b	0x701c65ec
1	1	19	0x00000003	0x727ea63f	0x09bf8ac0	0x3e638ac5	0x86a0dae0
1	1	25	0x00000004	0x01a53334	0xf500bc77	0x0add12df	0x98465988
1	1	18	0x00000005	0xdb22843f	0xec8a3dcc	0xd5facf27	0xe6450351

Plot configuration: Statistics:

Numerator: Hits Size (bits): 5216

Denominator: Access count

Ratio

Moving avg. 50 cyc.

Access address: 000000000010011101111100

Cache content:

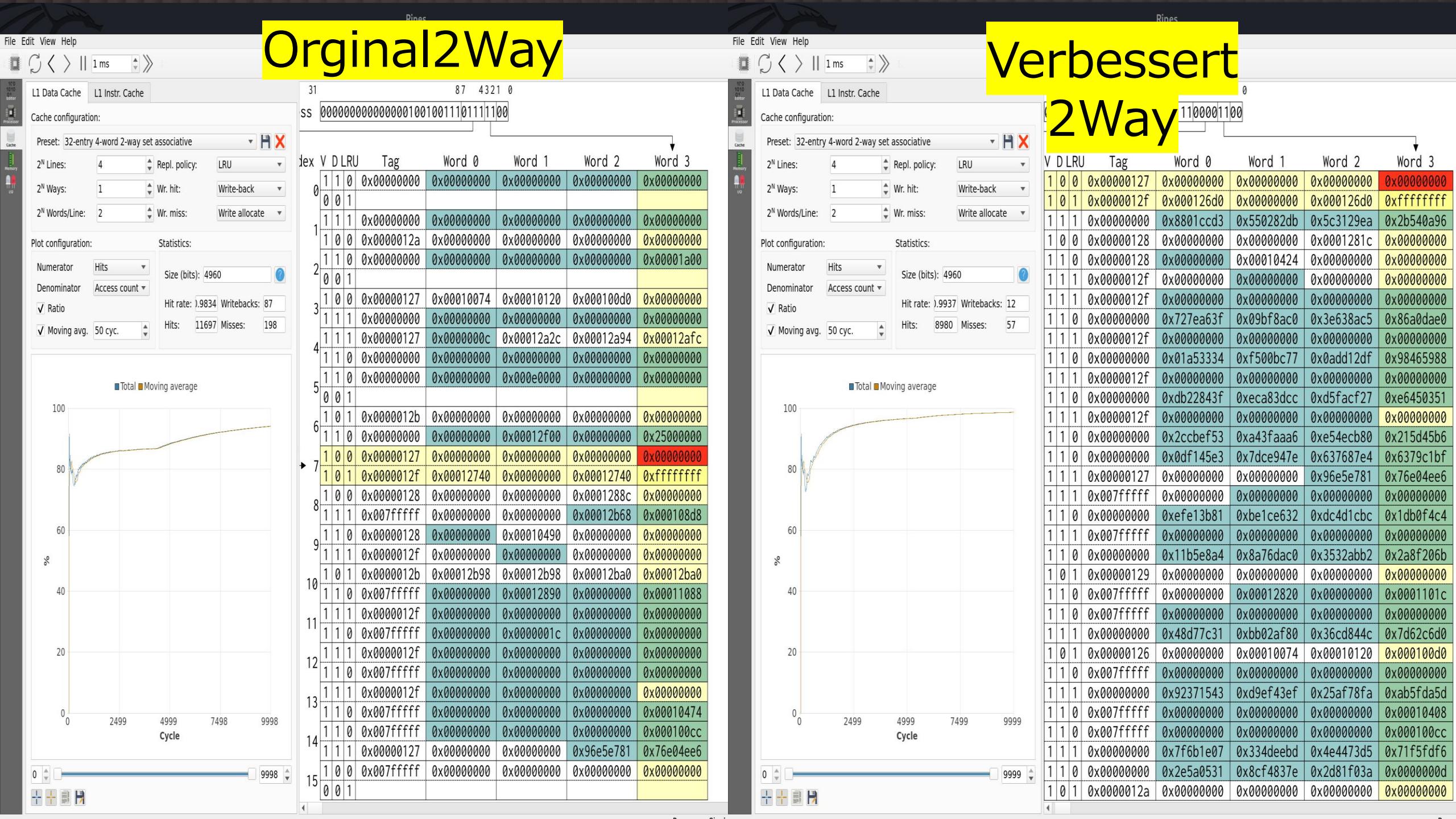
V	D	LRU	Tag	Word 0	Word 1	Word 2	Word 3
1	1	11	0x7fffffe	0x00000000	0x00000000	0x00000000	0x000100cc
1	0	5	0x00001281	0x00000000	0x00000000	0x0001281c	0x00000000
1	1	24	0x00000006	0x2ccbef53	0xa43faaa6	0xe54ecb80	0x215d45b6
1	1	16	0x00000007	0x0df145e3	0x7dce947e	0x637687e4	0x6379c1bf
1	1	23	0x00000008	0xefe13b81	0xbe1ce632	0xdc4d1cbc	0x1db0f4c4
1	0	9	0x0000126c	0x00000000	0x00010074	0x00010120	0x000100d0
1	0	1	0x000012f0	0x000126d0	0x00000000	0x000126d0	0xffffffff
1	1	15	0x00000009	0x11b5e8a4	0x8a76dac0	0x3532abb2	0x2a8f206b
1	1	22	0x000000a0	0x68d03d61	0xabdcc09	0x5d3e7541	0x3b6573bb
1	1	13	0x000000d0	0x92371543	0xd9ef43ef	0x25af78fa	0xab5fda5d
1	1	14	0x000000b0	0x48d77c31	0xbb02af80	0x36cd844c	0x7d62c6d0
1	1	21	0x000000c0	0xdd2f22d3	0xc4e3db17	0x5cfda276	0xb8f6b4ce
1	1	2	0x07fffffb	0x00000000	0x00000000	0x00000000	0x00000000
1	1	7	0x07fffffa	0x00000000	0x00012890	0x00000000	0x00011088
1	1	21	0x07fffff9	0x00000000	0x00000000	0x00000000	0x00000000
1	0	30	0x000012b9	0x00012b88	0x00012b88	0x00012b90	0x00012b90
1	0	29	0x000012ba	0x00012b98	0x00012b98	0x00012ba0	0x00012ba0
1	0	23	0x000012b7	0x00012b68	0x00012b68	0x00012b70	0x00012b70
1	0	22	0x000012b6	0x00000000	0x00000000	0x00000000	0x00000000
1	1	24	0x07fffff8	0x00000000	0x00000000	0x00012b68	0x000108d8
1	1	27	0x07fffff7	0x00000000	0x00000000	0x00000000	0x00010edc
1	1	26	0x00001274	0x0000000c	0x00012a2c	0x00012a94	0x00012afc
1	1	3	0x07fffffc	0x00000000	0x00000000	0x00000000	0x00000000
1	1	4	0x07fffffd	0x00000000	0x00000000	0x00000000	0x00010474
1	1	13	0x00000000	0x00000000	0x00000000	0x00000000	0x00000000
1	1	12	0x00000001	0x00000000	0x00000000	0x00000000	0x00000000
1	0	31	0x000012af	0x00000000	0x00000000	0x00000000	0x00000000
1	1	29	0x07fffff8	0x00000000	0x00000000	0x00000000	0x00000000
1	1	8	0x000012f2	0x00000000	0x00000000	0x00000000	0x00000000
1	0	10	0x0000129a	0x00000000	0x00000000	0x00000000	0x00000000
1	1	3	0x07fffffc	0x00000000	0x00000000	0x00000000	0x00000000
1	1	6	0x00001282	0x00000000	0x00010424	0x00000000	0x00000000
1	1	4	0x07fffffd	0x00000000	0x00000000	0x00000000	0x00010408
1	1	28	0x00001277	0x00000000	0x00000000	0x96e5e781	0x76e04ee6
1	1	27	0x00000000	0x00004ecf	0x00090e8d	0x01ac3615	0x8efb0c40
1	1	20	0x00000001	0x8801cccd3	0x550282db	0x5c3129ea	0xb2540a96
1	1	26	0x00000002	0x0014ad1a	0x42c4ed9f	0x07d5695b	0x701c65ec
1	1	19	0x00000003	0x727ea63f	0x09bf8ac0	0x3e638ac5	0x86a0dae0
1							

# Warum ist das verbesserte Programm schlechter bei fully assoziative- Strategie?

In einem voll assoziativen Cache kann jede Adresse im Speicher an beliebiger Stelle im Cache abgelegt werden. Dies bedeutet, dass keine feste Zuordnung der Speicheradressen im Cache erfolgt. Dadurch kann es zu Kollisionen kommen, wenn aufeinanderfolgende Adressen im Speicher an verschiedenen Positionen im Cache auftreten, was zu einer niedrigeren Hit-Rate führt.

Im verbesserten Programm können aufgrund des linearen Speicherlayouts und des Zugriffsmusters eindeutige Adressen benachbart im Speicher auftreten. Dies erhöht die Wahrscheinlichkeit von Kollisionen in einem voll assoziativen Cache und führt zu einer niedrigeren Hit-Rate.

Also das verbesserte Programm ist nicht für diese Cache-Variante geeignet.



# Warum ist das verbesserte Programm besser bei 2way set assoziative?

In Fall des verbesserten Programms, wo die Matrizen kontinuierlich abgelegt werden, gibt es eine höhere Wahrscheinlichkeit, dass kürzlich verwendete Adressen im Cache angesammelt sind. Daher besteht eine höhere Wahrscheinlichkeit für eine Verbesserung der Cache-Hit-Rate, da mehr Cache-Hits auftreten und weniger Cache-misses. Im Gegensatz dazu kann bei dem orginalen Programm, wo die Matrizen verstreut abgelegt werden, die Anzahl der Cache-misses zunehmen, da die nicht zusammenhängenden Daten im Hauptspeicher möglicherweise nicht gut zum Cache-Layout passen. Somit besteht eine höhere Wahrscheinlichkeit dafür, dass die benötigten Daten nicht im Cache vorhanden sind, was zu einer Verringerung der Cache-Hit-Rate führt.

Im Allgemeinen bietet das verbesserte Programm eine bessere Anordnung der Daten im Speicher, die besser mit dem 2-Wege-assoziativen Cache übereinstimmt. Daher wird erwartet, dass die Cache-Hit-Rate im verbesserten Programm im Vergleich zu dem orginalem besser ist.

# Write-Back und Write-Through

Write-  
Through

Original Verbessert

2397

1884

WRITE BACK

Write-Back

Original Verbessert

175

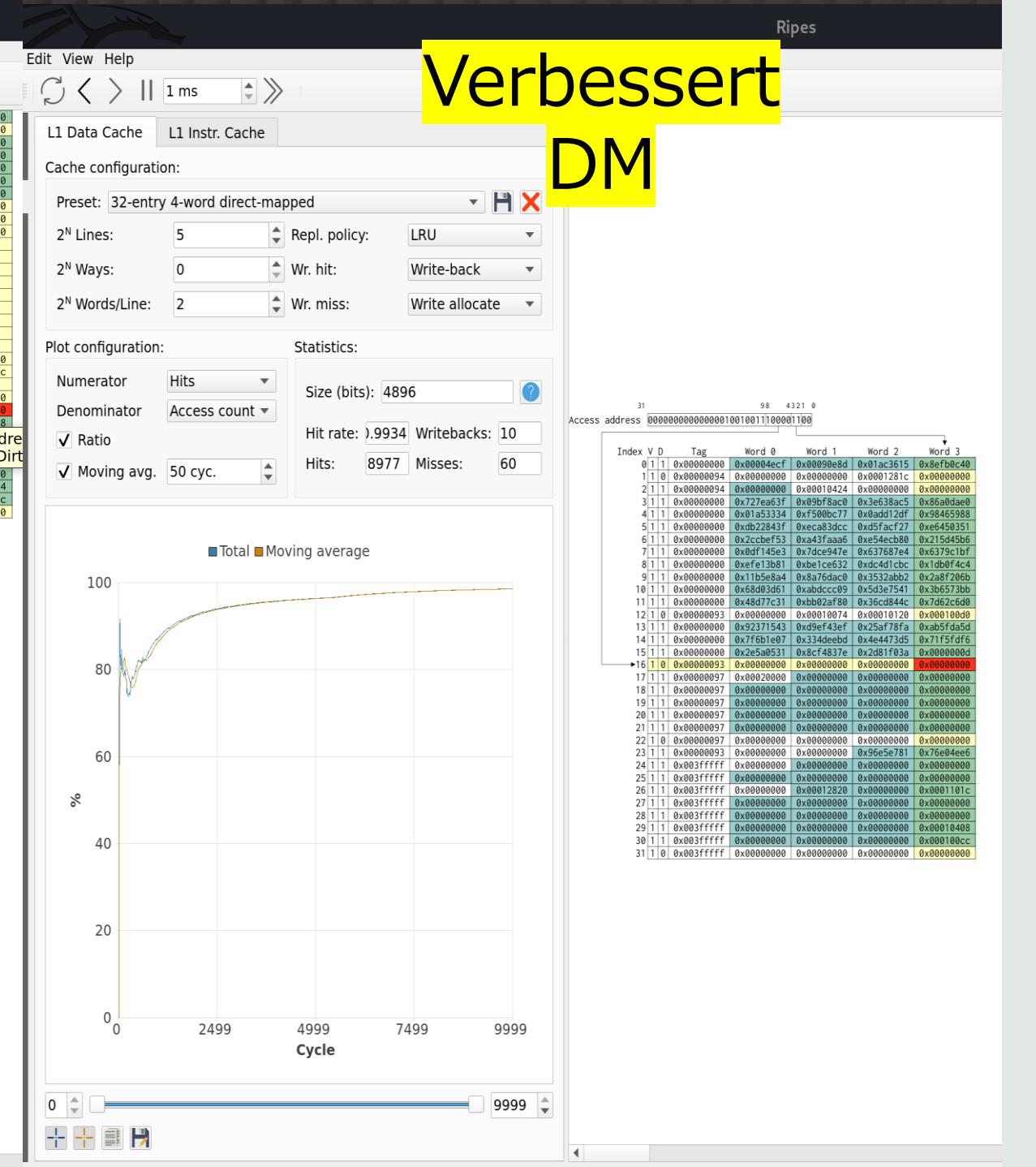
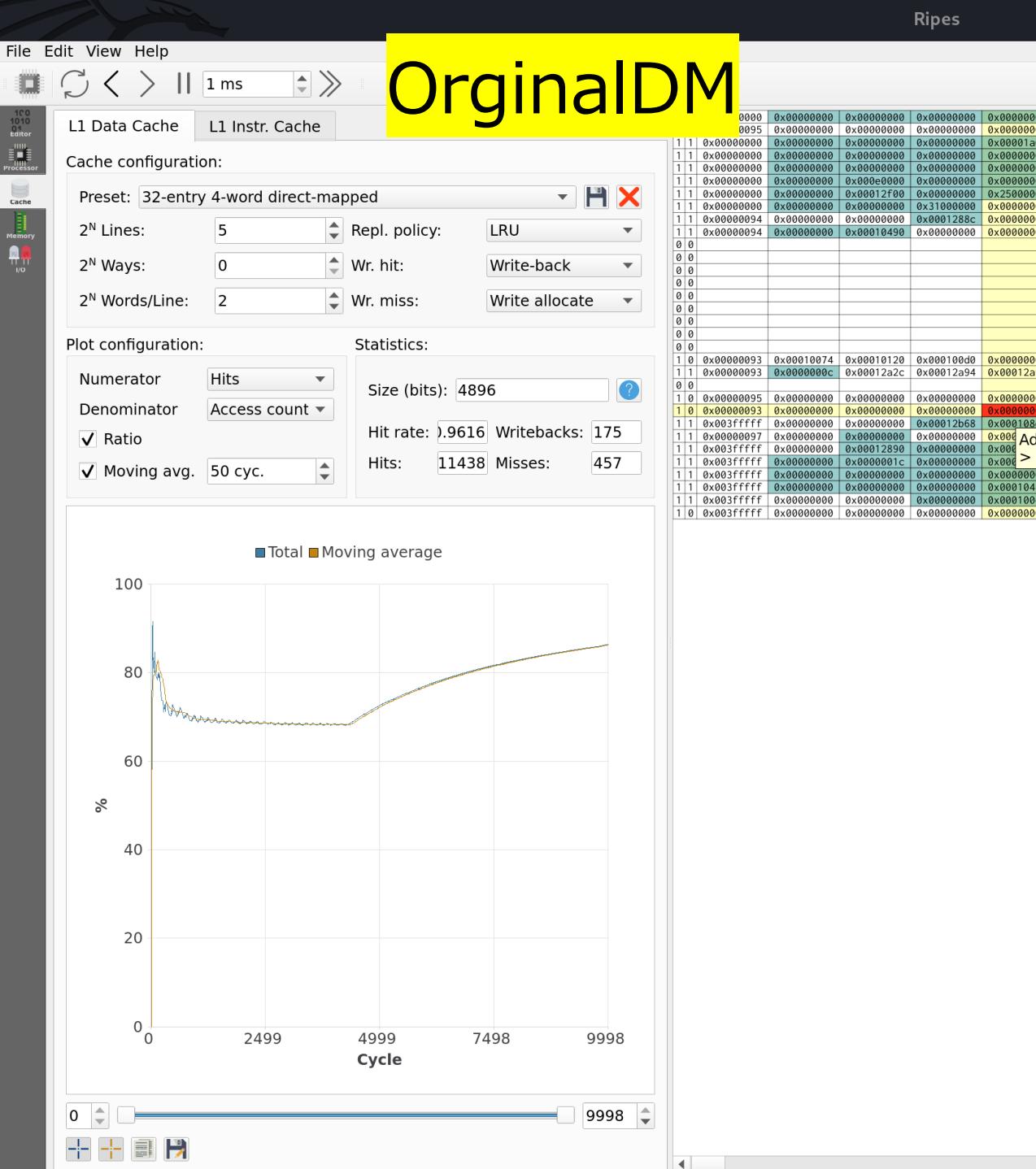
10

WRITE BACK

# Die Vorteile der Write-Back-Strategie

---

- Reduzierte Hauptspeicherzugriffe
- Bessere Ausnutzung des Cache-Speichers
- Geringere Latenzzeit für Schreiboperationen



**Original WT**

**Verbessert WT**

**Ripes**

**Cache configuration:**

**Preset:**

**2<sup>N</sup> Lines:** 5 **Repl. policy:** LRU

**2<sup>N</sup> Ways:** 0 **Wr. hit:** Write-through

**2<sup>N</sup> Words/Line:** 2 **Wr. miss:** Write allocate

**Plot configuration:**

**Statistics:**

**Numerator:** Hits **Denominator:** Access count

**Hit rate:** 0.9616 **Writebacks:** 2397

**Ratio**  **Moving avg.** 50 cyc.

**Size (bits):** 4864

**Access address:** 00000000000000001001001110000000

**Plot configuration:**

**Statistics:**

**Numerator:** Hits **Denominator:** Access count

**Hit rate:** 0.9934 **Writebacks:** 1884

**Ratio**  **Moving avg.** 50 cyc.

**Size (bits):** 4864

**Access address:** 00000000000000001001001110000000

**Index V Tag Word 0 Word 1 Word 2 Word 3**

0 1 0x00000000 0x00004ecf 0x00090e8d 0x01ac3615 0x8efb0c40

1 1 0x00000094 0x00000000 0x00000000 0x0001281c 0x00000000

2 1 0x00000094 0x00000000 0x00000000 0x00000000 0x00000000

3 1 0x00000000 0x727ea63f 0x09bf8ac0 0x3e638ac5 0x86a0da0e

4 1 0x00000000 0x01a53334 0xf500bc77 0x0add12df 0x98465988

5 1 0x00000000 0xdb22843f 0xecaa83dc 0xd5facf27 0x6e450351

6 1 0x00000000 0x2ccbf53 0xa443faa6 0x54ecbb80 0x215d45b6

7 1 0x00000000 0x0df145e3 0x7dce947e 0x637687e4 0x6379c1bf

8 1 0x00000000 0xefe13b81 0x0be1ce632 0x3dc4d1cbc 0x1db0f4c4

9 1 0x00000000 0x11b5e844 0x8a76dac0 0x353dab2 0x2a8f206b

10 1 0x00000000 0x68d03d61 0xabddcc09 0x5d3e7541 0x3b6573bb

11 1 0x00000000 0x48d77c31 0xbb02a180 0x36cd844c 0x7d62c6d0

12 1 0x00000093 0x00000000 0x00000000 0x00001074 0x0000100d

13 1 0x00000000 0x92371543 0x9d9ef43ef 0xab5fd5d 0x25af78fa

14 1 0x00000000 0x7f6b1e07 0x334deebd 0x4e4473d5 0x71f5fd6

15 1 0x00000000 0x2e5a0531 0x8cf4837e 0x2d81f03a 0x0000000d

16 1 0x00000093 0x00000000 0x00000000 0x00000000 0x00000000

17 1 0x00000097 0x00020000 0x00000000 0x00000000 0x00000000

18 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

19 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

20 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

21 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

22 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

23 1 0x00000093 0x00000000 0x00000000 0x965e781 0x76e04e6

24 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

25 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

26 1 0x003fffff 0x00000000 0x00000000 0x00011280 0x00000000

27 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

28 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

29 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00010408

30 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x0000100cc

31 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

**Plot configuration:**

**Statistics:**

**Numerator:** Total **Denominator:** Moving average

**Access address:** 00000000000000001001001110000000

**Plot configuration:**

**Statistics:**

**Numerator:** Total **Denominator:** Moving average

**Access address:** 00000000000000001001001110000000

**Index V Tag Word 0 Word 1 Word 2 Word 3**

0 1 0x00000000 0x00004ecf 0x00090e8d 0x01ac3615 0x8efb0c40

1 1 0x00000094 0x00000000 0x00000000 0x0001281c 0x00000000

2 1 0x00000094 0x00000000 0x00000000 0x00000000 0x00000000

3 1 0x00000000 0x727ea63f 0x09bf8ac0 0x3e638ac5 0x86a0da0e

4 1 0x00000000 0x01a53334 0xf500bc77 0x0add12df 0x98465988

5 1 0x00000000 0xdb22843f 0xecaa83dc 0xd5facf27 0x6e450351

6 1 0x00000000 0x2ccbf53 0xa443faa6 0x54ecbb80 0x215d45b6

7 1 0x00000000 0x0df145e3 0x7dce947e 0x637687e4 0x6379c1bf

8 1 0x00000000 0xefe13b81 0x0be1ce632 0x3dc4d1cbc 0x1db0f4c4

9 1 0x00000000 0x11b5e844 0x8a76dac0 0x353dab2 0x2a8f206b

10 1 0x00000000 0x68d03d61 0xabddcc09 0x5d3e7541 0x3b6573bb

11 1 0x00000000 0x48d77c31 0xbb02a180 0x36cd844c 0x7d62c6d0

12 1 0x00000093 0x00000000 0x00000000 0x00001074 0x0000100d

13 1 0x00000000 0x92371543 0x9d9ef43ef 0xab5fd5d 0x25af78fa

14 1 0x00000000 0x7f6b1e07 0x334deebd 0x4e4473d5 0x71f5fd6

15 1 0x00000000 0x2e5a0531 0x8cf4837e 0x2d81f03a 0x0000000d

16 1 0x00000093 0x00000000 0x00000000 0x00000000 0x00000000

17 1 0x00000097 0x00020000 0x00000000 0x00000000 0x00000000

18 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

19 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

20 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

21 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

22 1 0x00000097 0x00000000 0x00000000 0x00000000 0x00000000

23 1 0x00000093 0x00000000 0x00000000 0x965e781 0x76e04e6

24 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

25 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

26 1 0x003fffff 0x00000000 0x00000000 0x00011280 0x00000000

27 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

28 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000

29 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00010408

30 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x0000100cc

31 1 0x003fffff 0x00000000 0x00000000 0x00000000 0x00000000