



INDEXBASIERTES SHORT-READ-MAPPING AUF FPGAS

Zwischenvortrag Großer Beleg

Frank Hoffmann

Dresden, 28.6.2012



Index

Alignment Algorithmus

FPGA Design

01 Indexerstellung

DB:	C	G	A	T	A	C
Index:	0	1	2	3	4	5

Index	DB[Index:END]
0	CGATAC
1	GATAC
2	ATAC
3	TAC
4	AC
5	C

01 Indexerstellung

DB:		C	G	A	T	A	C
Index:		0	1	2	3	4	5

Index	DB[Index:END]
4	AC
2	ATAC
5	C
0	CGATAC
1	GATAC
3	TAC

01 Indexerstellung

DB:	C	G	A	T	A	C
Index:	4	2	5	0	1	3

Index	DB[Index:END]
4	AC
2	ATAC
5	C
0	CGATAC
1	GATAC
3	TAC

01 Suche im Index

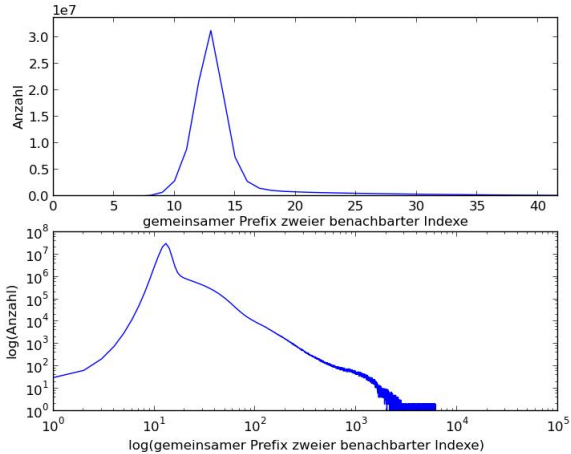
- Read: ACTC
- jeden Offset im Index suchen

Beispiel für Offset 0:

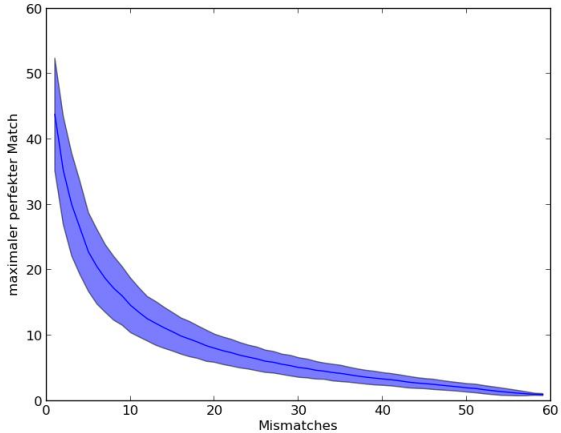
Index	DB[Index:END]
10	A A C A
42	A A G C
32	A C C A
78	A C T A
37	A C T G
91	A T A G
61	C A G G
⋮	⋮ ⋮ ⋮ ⋮
⋮	⋮ ⋮ ⋮ ⋮

← ACTC

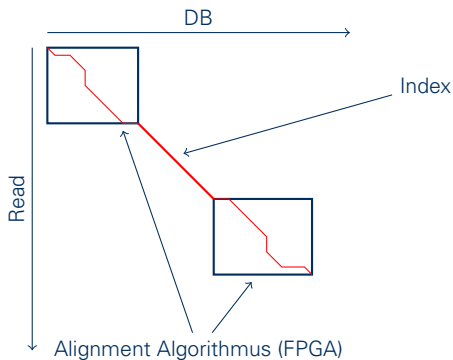
Charakteristik eines Indexes der Größe 112640310.



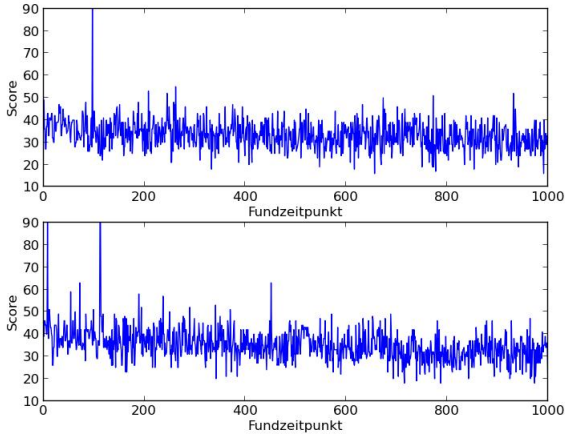
Einfluss zufälliger Mismatches auf maximale Übereinstimmungen.



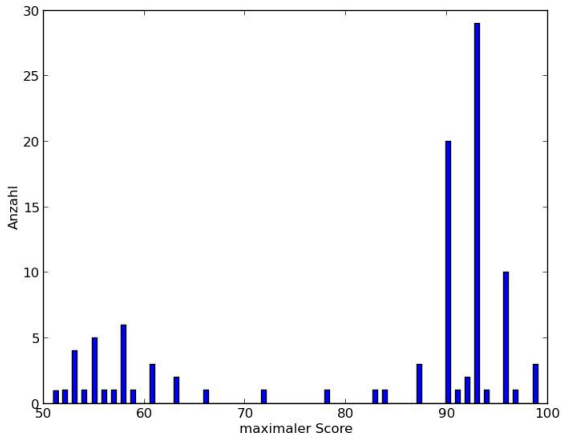
02 Alignment Algorithmus



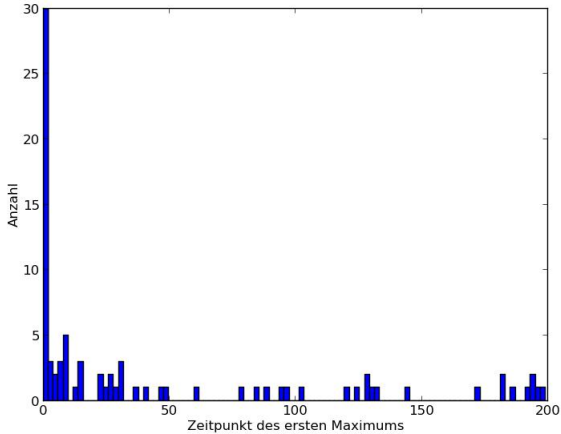
Suchbeispiele.



100 Sequenzen von je 60 Basenpaaren mit 10 Mismatches.
Überprüfung der ersten 1000 Indexe.



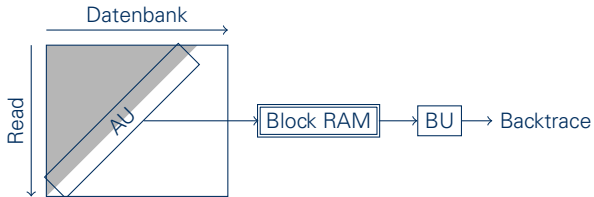
100 Sequenzen von je 60 Basenpaaren mit 10 Mismatches.



Optimierungen:

- SIMD für die Stringvergleiche.
- Lookup-Table um die Suche teilweise zu beschleunigen.
- SSD für Zugriffsbeschleunigung beim Suchen.
- Alignment der Enden auf den FPGA auslagern.

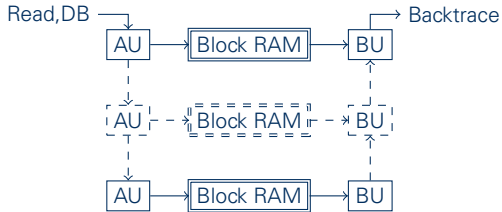
03 FPGA Design



AU: Alignment Unit

BU: Backtrace Unit

03 FPGA Design



AU: Alignment Unit
BU: Backtrace Unit

03 FPGA Design

