3.7 Techniken zur effizienten Cacheverwaltung

3.7.1 Vorbemerkung

- ❏ Wesentliche Faktoren, die die Effizienz beeinflussen
  - Design der Hardware
  - Zeitliche und örtliche Lokalität der Zugriffe in Programmen
  - Umgang des BS mit den Caches
- ❏ BS-Techniken
  - Layout von Adreßräumen
  - verzögerte Invalidierung von Caches
  - Ausrichtung der Anfangsadresse von Datenstrukturen

3.7.2 Layout des Adreßraums von Prozessen

© 1996, Universität Erlangen-Nürnberg, IMMD IV, F. Hofmann
Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig

- ❏ virtueller Cache: Die Anfangsadressen der Segmente Text, Data und Stack sollten so gewählt werden, daß sie nicht auf die gleiche Cachezeile abgebildet werden.
  Beispiel

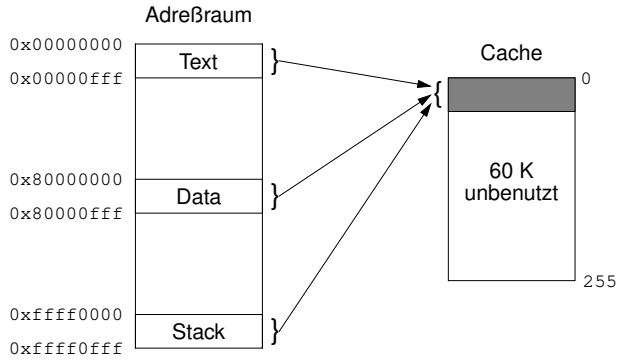
Region	Adresse
Text	0x00000000
Data/BSS	0x80000000
Stack	0xffff0000

Adresse in *Text*	Index
0x00000000	0
0x00000010	1
0x00000020	2
0x00000030	3
•	•
•	•
•	•
0x00000ff0	255

Adresse in *Data*	Index
0x80000000	0
0x80000010	1
0x80000020	2
0x80000030	3
•	•
•	•
•	•

Adresse in *Stack*	Index
0xffff0000	0
0xffff0010	1
0xffff0020	2
0xffff0030	3
•	•
•	•

© 1996, Universität Erlangen-Nürnberg, IMMD IV, F. Hofmann
Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig



© 1996, Universität Erlangen-Nürnberg, IMMD IV, F. Hofmann
Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig

Andere Wahl der Anfangsadressen

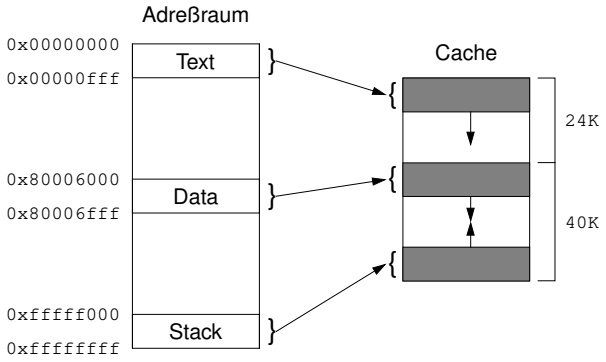
Region	Adresse
Text	0x00000000
Data/BSS	0x80006000
Stack	0xfffff000

Adresse in *Text*	Index
0x00000000	0
0x00000010	1
0x00000020	2
0x00000030	3
•	•
•	•
•	•
0x00000ff0	255

Adresse in *Data*	Index
0x80006000	1536
0x80006010	1537
0x80006020	1538
0x80006030	1539
•	•
•	•
•	•

Adresse in *Stack*	Index
0xffff0000	3840
0xffff0010	3841
0xffff0020	3842
0xffff0030	3843
•	•
•	•
•	•
0xffff0ff0	4095

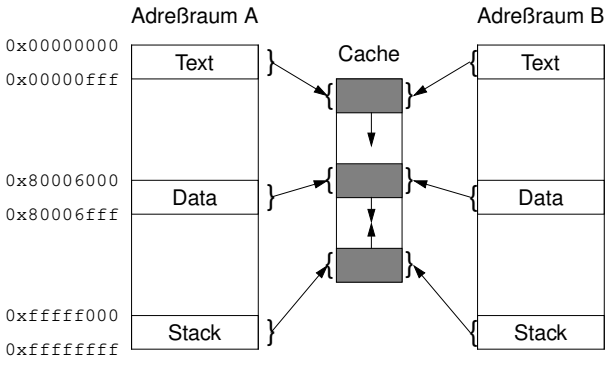
© 1996, Universität Erlangen-Nürnberg, IMMD IV, F. Hofmann
Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig



Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig

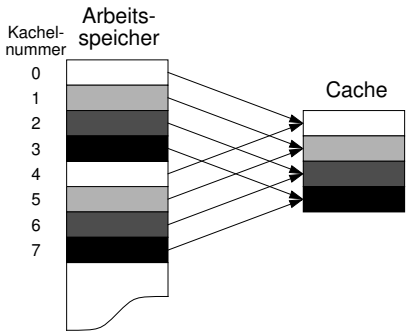
❏ Dynamische Vergabe der Anfangsadressen von Segmenten

Beispiel für die Problematik



Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig

❏ physikalischer Cache



Einteilung der Kacheln in Gruppen, die jeweils aus den Kacheln bestehen, die auf die gleiche Cache-Zeile abgebildet werden. Zuteilung gemäß Round-Robin.

Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig

3.7.3 Verzögertes Invalidieren

❏ virtueller Cache mit Schlüsseln

- Invalidierung nach exit kann verzögert werden bis zur Neuvergabe des Schlüssels
- Sammeln mehrerer exits bis Schlüssel wieder vergeben werden
- Kompliziert bei write-back, da Kacheln anderweitig vergeben worden sein könnten

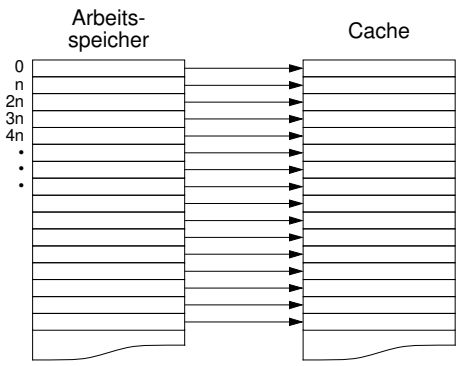
❏ virtueller Cache mit physikalischen Tags, physikalischer Cache ohne Busüberwachung

- zwei Listen für freie Kacheln
  - mit Zeilen im Cache (Kachel unsauber)
  - ohne Zeilen im Cache (Kachel sauber)
- bei exit
  - Kacheln in Liste der unsauberen einreihen
- bei Kachelzuordnung
  - aus Liste der sauberen Kacheln zuordnen
  - falls Liste der sauberen Kacheln leer, aber unsaubere Kacheln verfügbar, dann Cache invalidieren und alle unsauberen Kacheln in die Liste der sauberen umhängen
- analog bei Schrumpfen des Datensegmentes

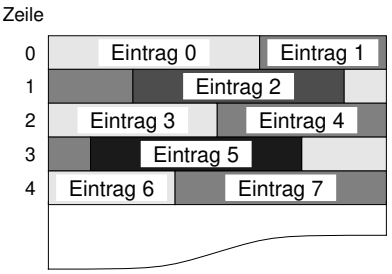
Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig

3.7.4 geschickte Wahl der Anfangsadressen von Datenstrukturen

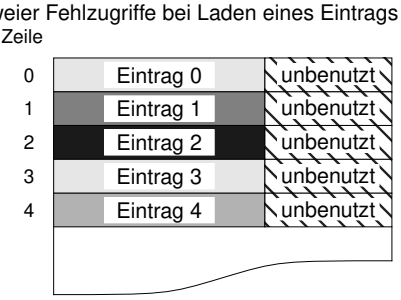
Abbildung bei Zeilengröße n



Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig



Vermeidung zweier Fehlzugriffe bei Laden eines Eintrags



Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig

3.7.5 Schlußfolgerung

- ❏ Sorgfältige Adreßraumgestaltung kann Effizienz deutlich erhöhen
- ❏ verzögerte Bereinigung kann benutzt werden, um die Zahl der Invalidierungen wegen nicht mehr zugeordneter Bereiche (infolge exit, brk oder sbrk) zu reduzieren
- ❏ Ausrichtung häufig benötigter Datenstrukturen auf Cachezeilen reduziert Fehlzugriffe

Reproduktion jeder Art oder Verwendung dieser Unterlage zu Lehrzwecken außerhalb der Universität Erlangen-Nürnberg ist ohne Genehmigung des Autors nicht zulässig