
Echtzeit–Videoverarbeitung

Hardwareplattformen für Echtzeit– Videoverarbeitung

Herbert Thoma

Gliederung

Echtzeitanforderungen

Prozessorarchitekturen

- Grundlagen
- Pipelining
- Parallele Befehlsausführung

Systemkomponenten

- Speicher
- Cache
- DMA

Beispiele

Echtzeitanforderungen 1

Garantierte maximale Reaktionszeit auf externe Ereignisse

- Steuerungs- und Regelungstechnik, z.B.: Endschalter bei Motorsteuerung

Datenverarbeitung mit der Geschwindigkeit, mit der die Daten erzeugt bzw. verbraucht werden

- Digitale Signalverarbeitung
- Digitale Modulation
- Kanalcodierung
- Bildverarbeitung
 - Effekte, Bildverbesserung
 - Mustererkennung, Qualitätssicherung
 - Codierung

Echtzeitanforderungen 2

Datenrate Fernsehsignal ITU–R BT601

- 720 x 576 pixel
- 25 fps
- 8 bit 4:2:2 Abtastung
- ➔ ca. 20 MByte pro Sekunde

Beispiel: Filter 5 x 5

- 25 Multiplikationen pro pixel
- 24 Additionen pro pixel
- ➔ 49 arithmetische Operationen pro pixel
- ➔ ca. 520 Mio. arithmetische Operationen pro Sekunde

Echtzeitanforderungen 3

Maßnahmen um Echtzeitanforderungen zu erfüllen

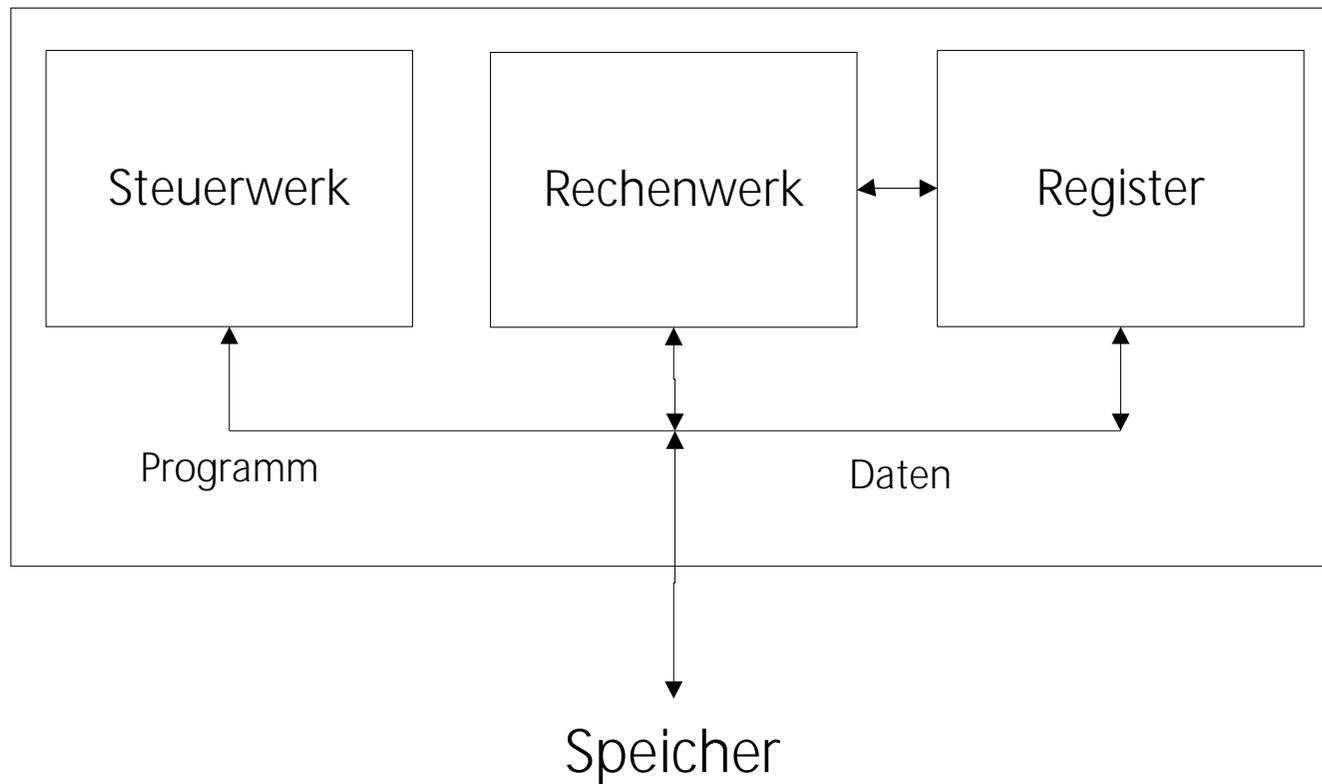
- Schnelle Prozessoren
 - Pipelining
 - Parallele Funktionseinheiten
 - Single Instruction Multiple Data (SIMD) Funktionseinheiten
- Schnelle Systemkomponenten
 - Speicher
 - Ein- / Ausgabe

- Dedizierte Hardware
 - MPEG-2 Decoder für DVD-Player

- Schnelle Algorithmen

Prozessoren 1

Prinzipschaltbild



Prozessoren 2

Architekturen

- Von Neumann
 - Gemeinsamer Speicher für Programm und Daten
 - Sequentieller Zugriff auf Programmbefehle und Daten
- Harvard
 - Getrennte Speicher für Programm und Daten, mehrere Datenspeicherbereiche
 - Gleichzeitiger Zugriff auf Programmbefehle und Daten
 - Typisch für Digitale Signalprozessoren (DSP)
- Kombinierte Architekturen
 - Extern (off chip) gemeinsamer Speicher für Programm und Daten
 - Intern (on chip) getrennte Speicher (Caches) für Programm und Daten
 - Übliches Konzept bei aktuellen Prozessoren

Prozessoren 3

Befehlssätze (Instruction Set Architecture, ISA)

- Complex Instruction Set Computer (CISC)
 - Komplexe Operationen pro Befehl
 - Variable Befehlsgröße
 - Direkte Speicherreferenzen in Befehlen
 - Komplexe Adressierungsmodi
 - Große Ausführungszeit pro Befehl
- Reduced Instruction Set Computer (RISC)
 - Einfache Basisoperationen pro Befehl
 - Feste Befehlsgröße
 - Load/Store Architektur
 - Einfache Adressierungsmodi
 - Kleine Ausführungszeit pro Befehl

CISC

ADD r1, *(r2 + 2 * r3 + 15), r4

RISC

MUL r3, 2, r5
ADD r2, r5, r5
ADD r5, 15, r5
LOAD *r5, r4
ADD r1, r4, r4

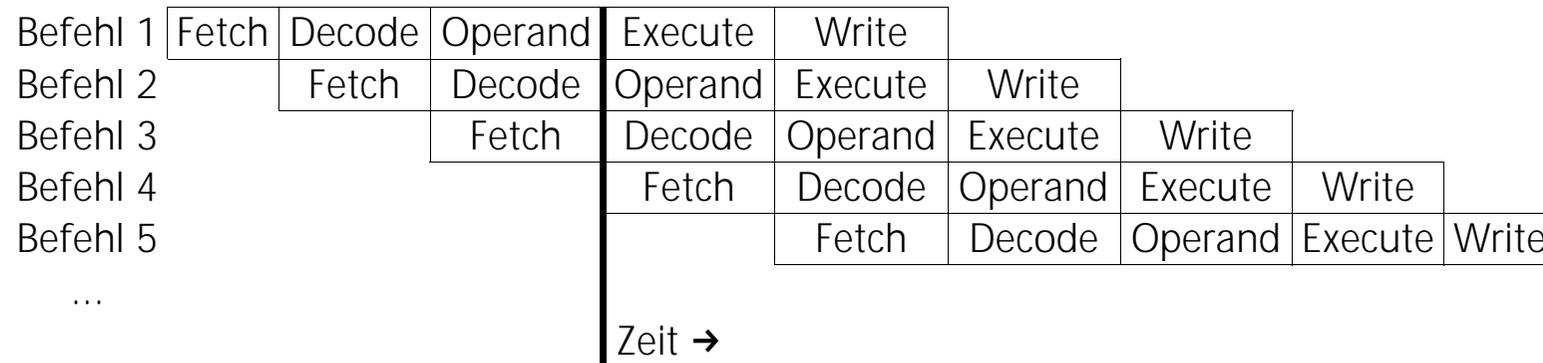
Pipelining 1

Schritte beim Ausführen eines Befehls

- Befehl holen (Fetch)
- Befehl decodieren (Decode)
- Operanden holen (Operand fetch)
- Befehl ausführen (Execute)
- Ergebnis speichern (Writeback)

Pipelining 2

Überlappende Ausführung der Befehle



Pipelining 3

Probleme

- Unterschiedliche Ausführungszeit verschiedener Befehle
 - Komplexe Befehle können mehrere Execute Stufen benötigen
 - Benötigt ein nachfolgender Befehl das Ergebnis, muß die Pipeline angehalten werden (Pipeline Stall)
- Variable Befehlslänge
 - Befehlslänge wird erst in der Decode Stufe festgestellt
 - Bei langem Befehl muß die Pipeline angehalten werden, um den Rest des Befehls zu holen
- Direkte Speicherreferenzen in Befehlen
 - Zusätzliche Adressberechnungen
 - Operand Fetch Stufe kann durch Speicherzugriff lange dauern
- Bedingte Verzweigungen
 - Sprungziel wird erst in der Execute Stufe ermittelt
 - Schon geholte und decodierte Befehle müssen eventuell verworfen werden
- Pipelinelänge
 - Lange Pipelines erlauben hohe Taktfrequenzen
 - Lange Pipelines verlieren viel Effektivität bei bedingten Sprüngen

Pipelining 4

Bedingte Verzweigungen

- Verzögerte Sprünge (Branch Delay Slots)
 - Schon in der Pipeline befindliche Befehle werden immer ausgeführt
- Bedingte Befehlsausführung (Conditional Execution, Predication)
 - Befehle werden abhängig von einer Bedingung ausgeführt
- Sprungvorhersage (Branch Prediction)
 - Angabe des wahrscheinlichen Sprungziels im Befehl (Static Branch Prediction)
 - Prozessor führt eine Liste der Sprungziele von Sprungbefehlen, die schon einmal ausgeführt wurden, besonders effektiv bei Schleifen
- Spekulative Befehlsausführung (Speculative Execution)
 - Benötigt parallele Funktionseinheiten

Pipelining 5

Beispiel Branch Delay Slot & Conditional Execution

- TMS320C6x Assembler
- Addition mit Clipping, vereinfacht ;-)

MV	64, A2	Schleifenzähler
MV	240, A3	Obere Clippinggrenze
MV	16, A4	Untere Clippinggrenze

Schleife:	LDBU	*A10++, A5	Lade Wert 1
	LDBU	*A11++, A6	Lade Wert 2
	ADD	A5, A6, A7	A7 = A5 + A6
	[A2] SUB	A2, 1, A2	Falls (A2 != 0) erniedrige Schleifenzähler
	[A2] B	Schleife	Falls (A2 != 0) springe zu Schleife, 5 delay slots
	GMPGT	A7, A3, A1	Falls (A7 > A3) A1 = TRUE
	[A1] MV	A3, A7	Falls (A1 == TRUE) A7 = 240
	GMPLT	A7, A4, A1	Falls (A7 < A4) A1 = TRUE
	[A1] MV	A4, A7	Falls (A1 == TRUE) A7 = 16
	STB	A7, *A12++	Speichere Ergebnis Sprung findet statt
	B	Adresse	Rücksprung
	NOP	5	Keine sinnvolle Aktion möglich

Parallele Befehlsausführung

Eine optimale Pipeline kann einen Befehl pro Takt ausführen

Voneinander unabhängige Befehle können parallel abgearbeitet werden

```
mul r1, r2, r3
mul r4, r5, r6
add r7, r8, r9
mul r10, r11, r12
add r13, r14, r15
```

```
mul r1, r2, r3
mul r4, r5, r6    add r7, r8, r9
mul r10, r11, r12 add r13, r14, r15
```

Bezeichnung: „multiple instruction issue“ Prozessor

Techniken

- Superskalare Prozessoren
- Very Long Instruction Word (VLIW) Prozessoren

Superskalare Prozessoren

Ermittlung der Parallelisierbarkeit von Befehlen durch den Prozessor zur Laufzeit

Keine direkte Kontrolle der parallelen Befehlsausführung durch den Programmierer oder Compiler

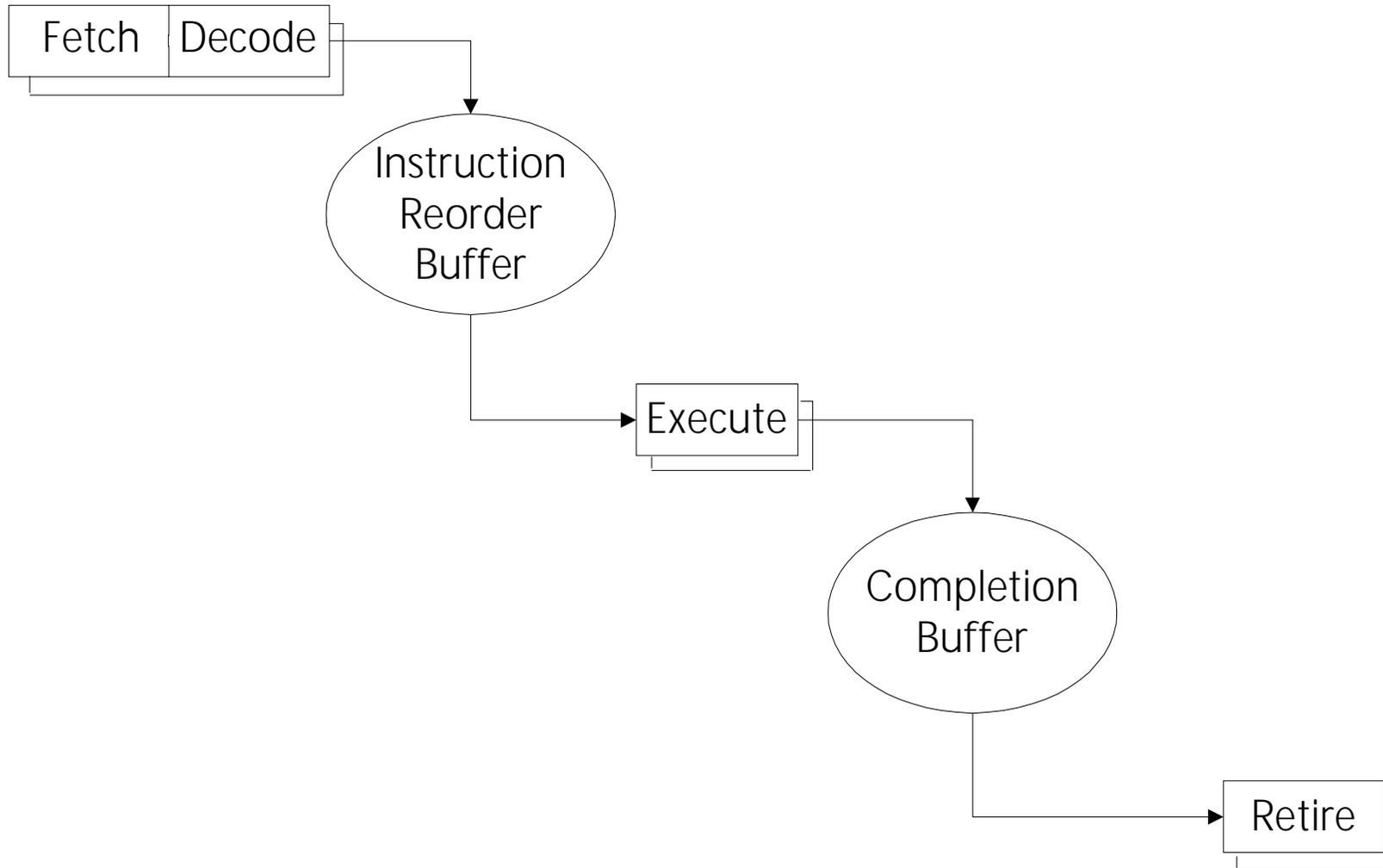
Techniken

- Speculative Execution
- Out-of-order Execution
- Register Renaming

Prozessoren

- Pentium, P Pro, P II, P III
- Alpha
- PowerPC
- SPARC II, SPARC III

Superskalare Pipeline



Very Long Instruction Word (VLIW)

Ermittlung der Parallelisierbarkeit von Befehlen bei der Programmentwicklung

Direkte Kontrolle der parallelen Befehlsausführung durch den Programmierer oder Compiler

Verbreitet bei Digitalen Signalprozessoren

- ADI SHARC: 48 bit Befehlslänge
- TI C8x: 64 bit Befehlslänge
- TI C6x: 32 – 256 bit Befehlslänge

Universalprozessoren

- Intel IA64: 128 bit Befehlslänge, Bezeichnung: EPIC Explicitly Parallel Instruction Computing

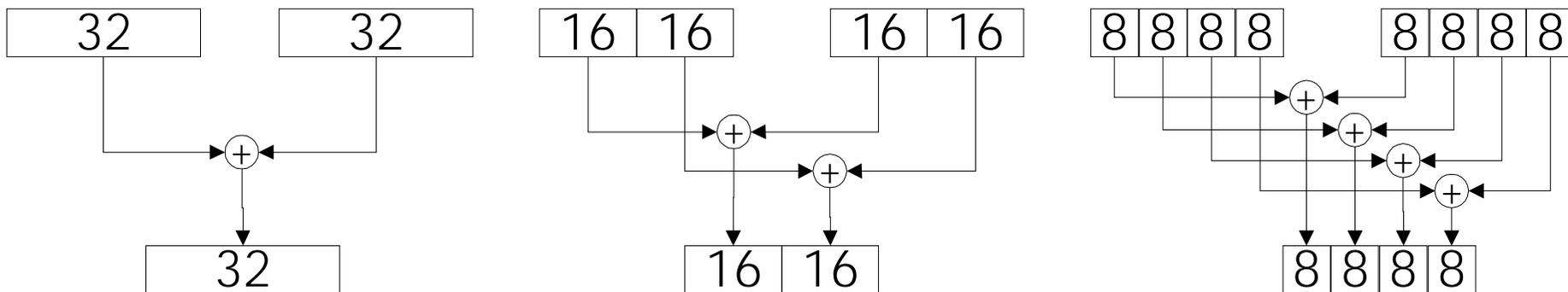
Single Instruction Multiple Data (SIMD) 1

Aufteilung eines Registers in mehrere Untereinheiten

Parallele Ausführung von Befehlen auf den Untereinheiten

Überlaufbehandlung

- Modulo: $240 + 17 = 1$
- Clipping: $240 + 17 = 255$



Single Instruction Multiple Data (SIMD) 2

x86

- MMX
 - Integer
 - 64 bit Register
- 3DNow!
 - Floating Point
 - 64 bit Register
- SSE „Internet SIMD Streaming Extensions“
 - Floating Point
 - 128 bit Register

PowerPC

- AltiVec
 - Integer & Floating Point
 - 128 bit Register

Sun SPARC

- VIS „Visual Instruction Set“
 - Integer
 - 64 bit Register

Speicher 1

Kenngroßen

- Größe (Mbyte)
- Bandbreite (Mbyte / s)
- Latenzzeit (ns)

Technologie

- Dynamisches RAM (DRAM)
 - Speicherzelle besteht aus 1 Transistor und 1 Kondensator
 - Höchste Integrationsdichte / Speichergröße
 - Mäßige Bandbreite
 - Hohe Latenzzeit
 - Benötigt Refresh
- Statisches RAM (SRAM)
 - Speicherzelle besteht aus 4 – 6 Transistoren
 - Mäßige Integrationsdichte
 - Hohe Bandbreite
 - Niedrige Latenzzeit

Speicher 2

DRAM

Typ	Taktfrequenz [MHz]	Busbreite [bit]	Bandbreite [MByte/s]	Latenzzeit [ns]
SDRAM 100	100	64	800	90
SDRAM 133	133	64	1064	75
DDR SDRAM 200	100 DDR	64	1600	70
DDR SDRAM 266	133 DDR	64	2128	60
RAMBUS 800	400 DDR	16	1600	70

Achtung: Die Werte für die Bandbreite gelten nur bei Zugriffen auf aufeinanderfolgende Speicherzellen!

Cache 1

Problem: Prozessortaktfrequenz \gg Speichertaktfrequenz

- Prozessortaktfrequenz 1 GHz
 - 128 bit Daten pro Takt (zwei 32 bit Befehle + ein 64 bit Datenwort)
- 16 GByte / s

Lösung: Schnelle Zwischenspeicher

Funktionsprinzip: Lokalität von Speicherzugriffen

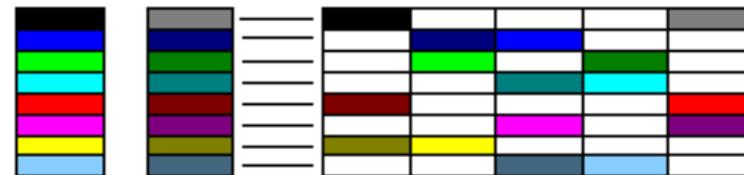
- Örtliche Lokalität
- Zeitliche Lokalität

Mehrstufige Cache Systeme

Cache 2

Organisation

- Direct Mapped
- Fully Associative
- Set Associative



Ersetzung von Cache-Inhalten

- Round Robin
- Least Recently Used

Cache 3

Schlechte Lokalitätseigenschaften von Algorithmen zur Videoverarbeitung

- Sequentieller Zugriff auf große Datenmengen
- Zufällige Zugriffsreihenfolge auf relative kleine Datenmengen
- ➔ Nur einmal benötigte Daten werden zwischengespeichert
- ➔ Oft benötigte Daten werden aus dem Cache verdrängt (Trashing)

Lösungsmöglichkeiten

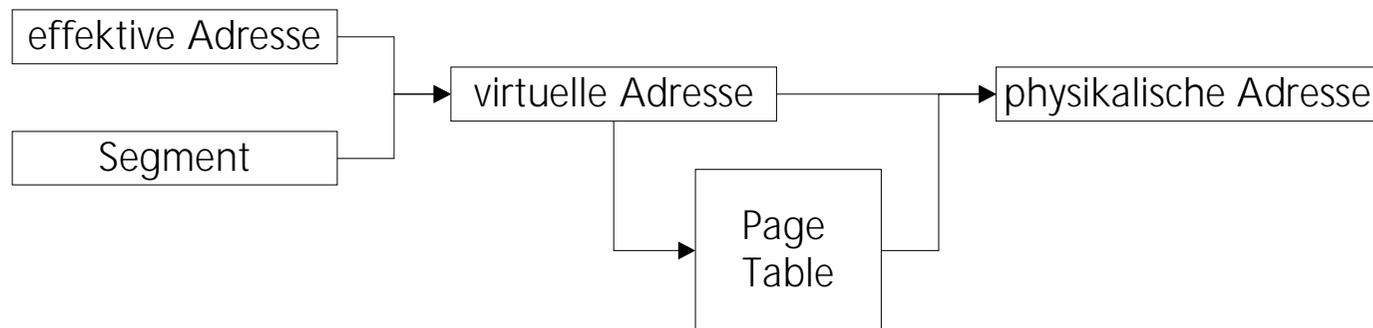
- Umgehung des Caches (Cache Bypass)
- Einfrieren des Cache-Inhalts (Cache Locking)
- Lokale Speicherbänke
 - Verwaltung unter Softwarekontrolle
 - Typisch für DSPs

Virtueller Speicher

Geschützter Adressraum für jeden Prozess

Paging zur Speicherverwaltung

Swapping zur Auslagerung von Speicherbereichen auf Massenspeicher



Cache für Zuordnung von virtuellen zu physikalischen Adressen (Translation Lookaside Buffer, TLB)

Direct Memory Access (DMA)

Datentransfer ohne Belastung des Prozessors

- Speicher ↔ Speicher
- Speicher ↔ Ein- / Ausgabegeräte

Anwendungen

- Massenspeicherzugriffe
- Kommunikationsschnittstellen
- Ein- / Ausgabe
 - Video
 - Audio
- Datenaustausch zwischen unterschiedlich schnellen Speicherbereichen

Komplexe DMA Operationen

- Mehrdimensionale Transfers
- Umsortieren
- Verkettung von DMA Transfers (DMA Chaining)

Beispiel Superskalarerer Prozessor PowerPC MPC7410 1

Maximal 2 Befehle können gleichzeitig starten bzw. enden

8 Execute Einheiten

- Branch Processing Unit
- 2 Integer Units
- Floating Point Unit
- 2 Vector Units (SIMD, AltiVec)
- Load Store Unit
- System Register Unit

3 Registerbänke

- 32 32bit Integer Register
- 32 64bit Floating Point Register
- 32 128bit Vector Register

Beispiel Superskalarerer Prozessor PowerPC MPC7410 2

On Chip Level 1 Cache

- 32 kByte Instruction Cache, 8 way set associative
- 32 kByte Data Cache, 8 way set associative

Controller für externen Level 2 Cache

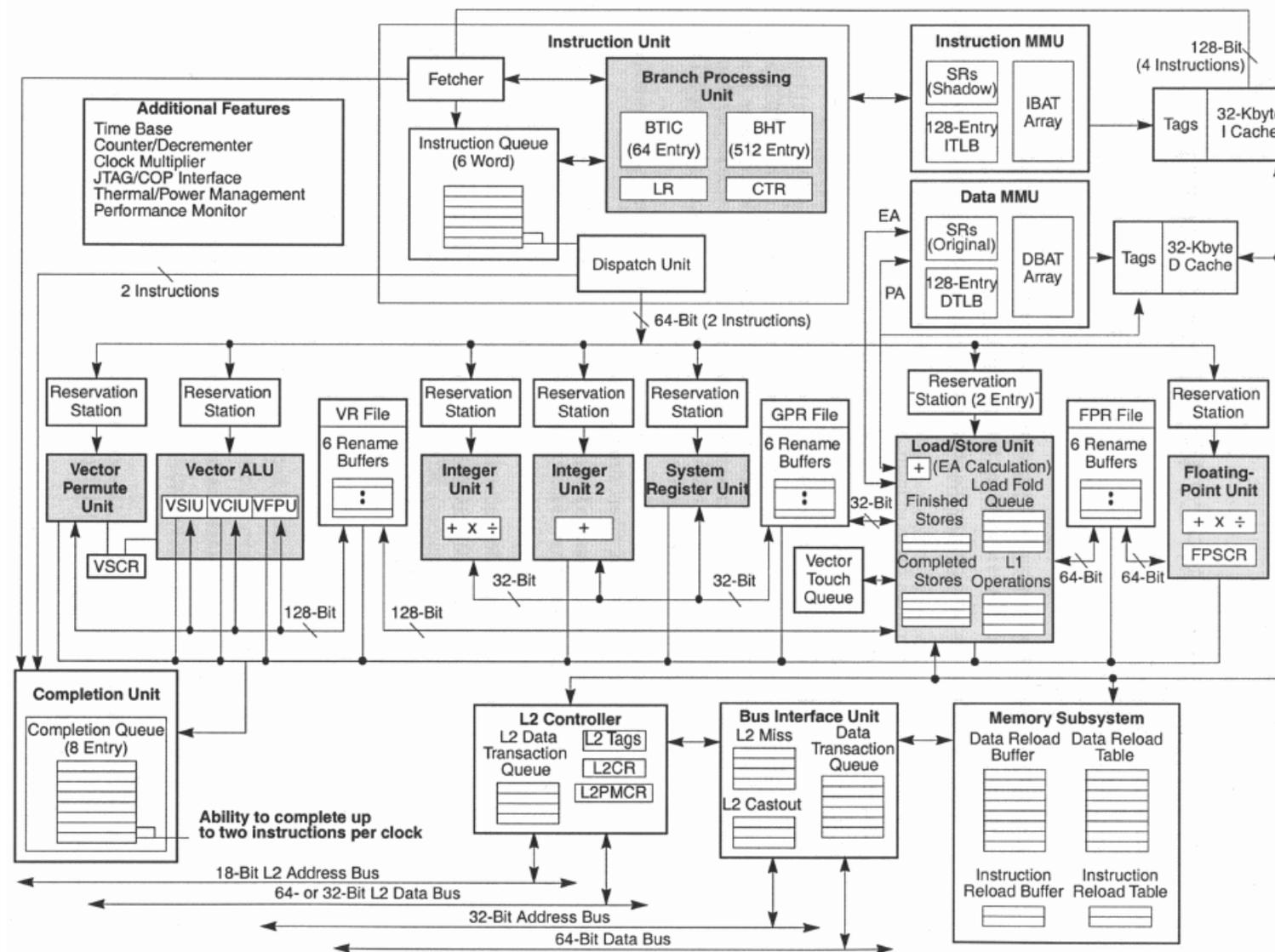
Memory Management Unit (MMU)

- Separate MMUs für Code und Daten
- TLBs mit je 128 Einträgen

Weitere periphere Einheiten

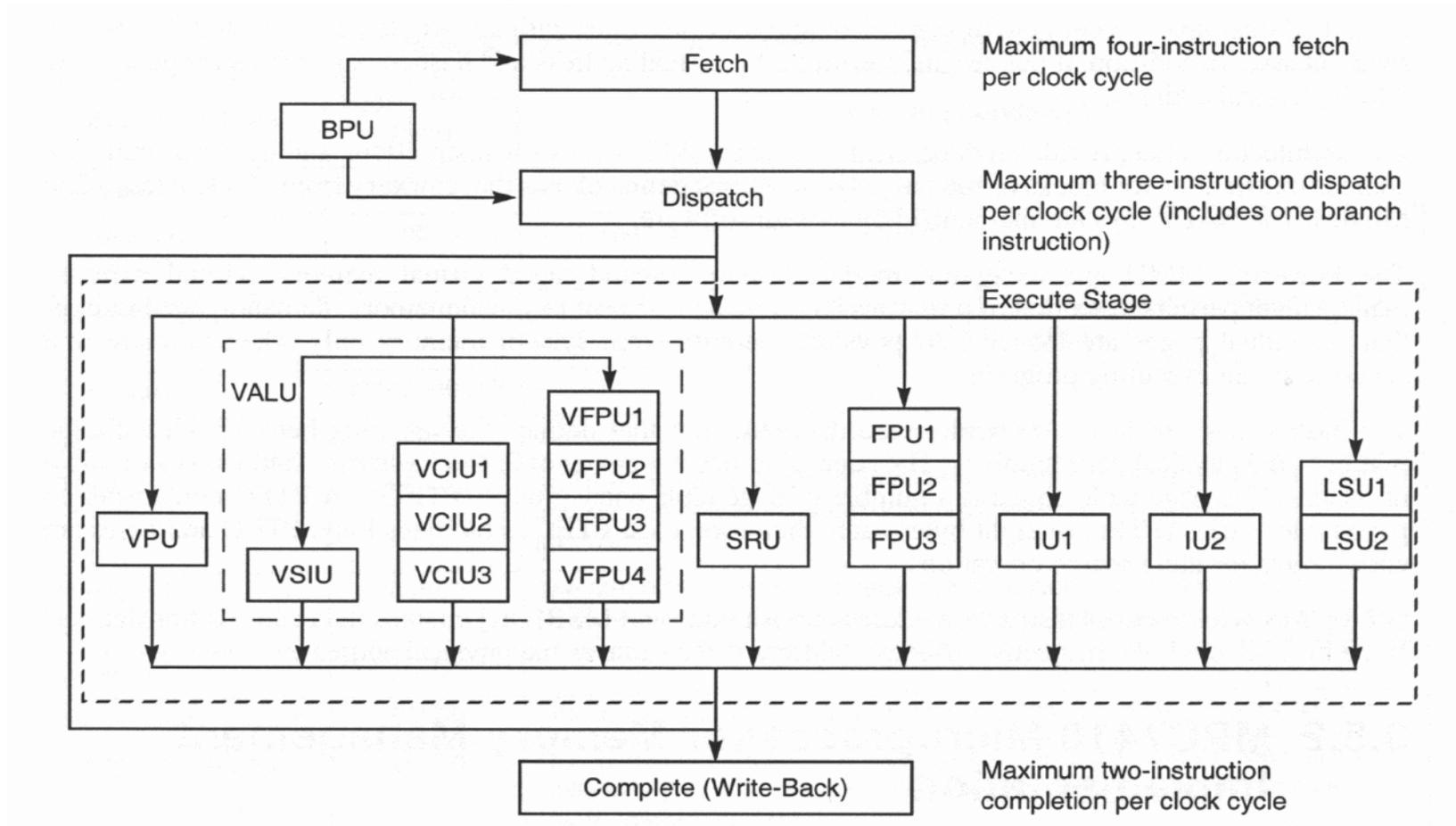
- Timer
- Takt PLL
- Power Management

Beispiel Superskalärer Prozessor PowerPC MPC7410 3



Beispiel Superskalarerer Prozessor PowerPC MPC7410 4

Pipeline



Beispiel VLIW Prozessor TMS320C6201 1

8 parallele Funktionseinheiten

- 2 .L Units
 - Addition, Subtraktion
 - Vergleich
 - Logische Operationen
- 2 .S Units
 - Addition, Subtraktion
 - Logische Operationen
 - Schiebeoperationen
 - Verzweigungen
- 2 .M Units
 - Multiplikation
- 2 .D Units
 - Addition, Subtraktion
 - Laden, Speichern

2 Registerbänke

- 16 Register pro Bank
- 32 bit Datenbreite

Beispiel VLIW Prozessor TMS320C6201 2

On Chip Speicher

- 64 kByte Programmspeicher
 - Lokaler Speicher oder Cache
- 64 kByte Datenspeicher
 - 2 Zugriffe gleichzeitig möglich (dual ported)

External Memory Interface (EMIF)

- Direkte Anschlußmöglichkeit (glueless) für verschiedene Speichertypen

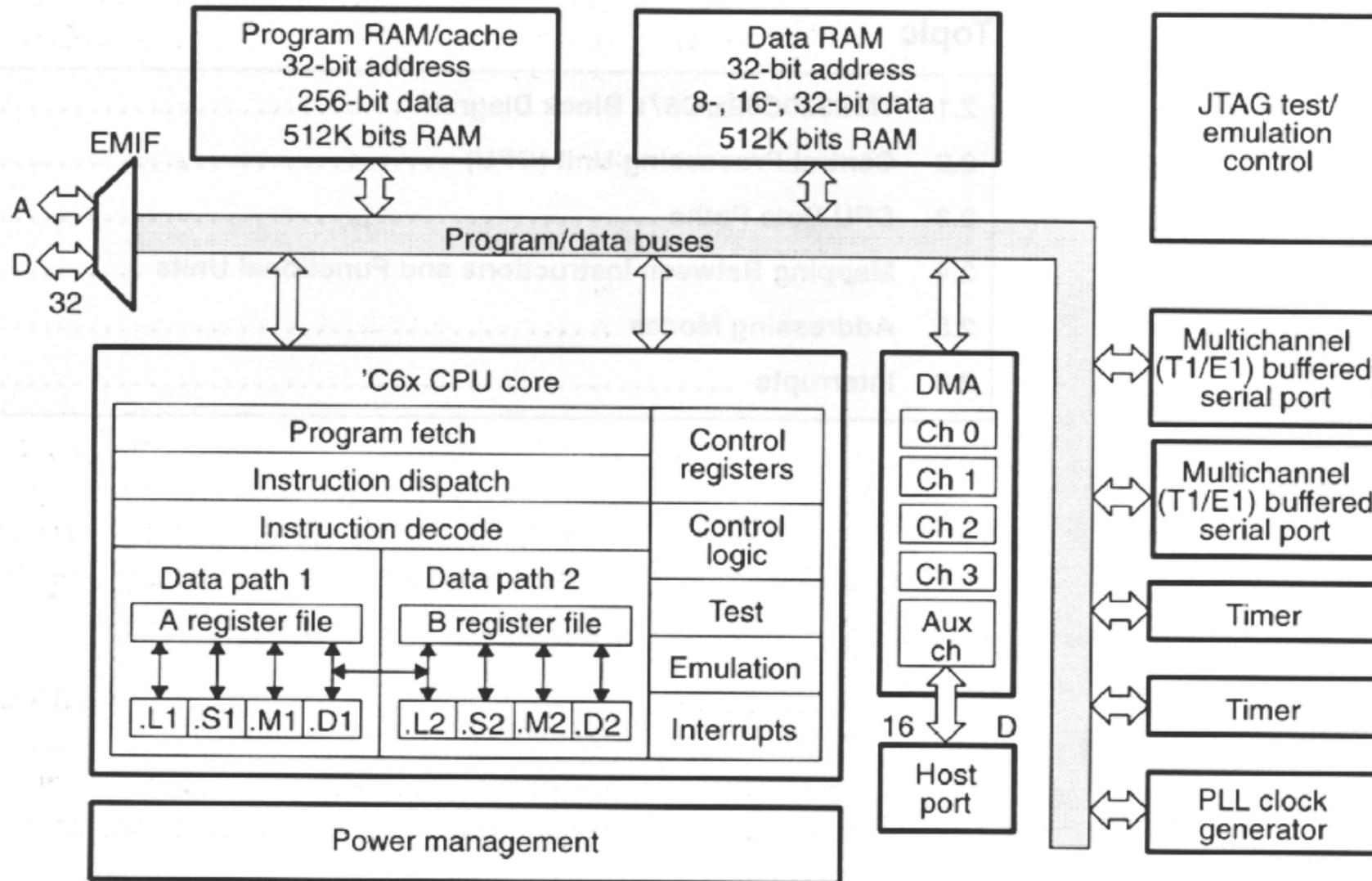
DMA Controller

Serielle Schnittstellen

Timer

Hostport

Beispiel VLIW Prozessor TMS320C6201 3



Beispiel Video Encoder System 1

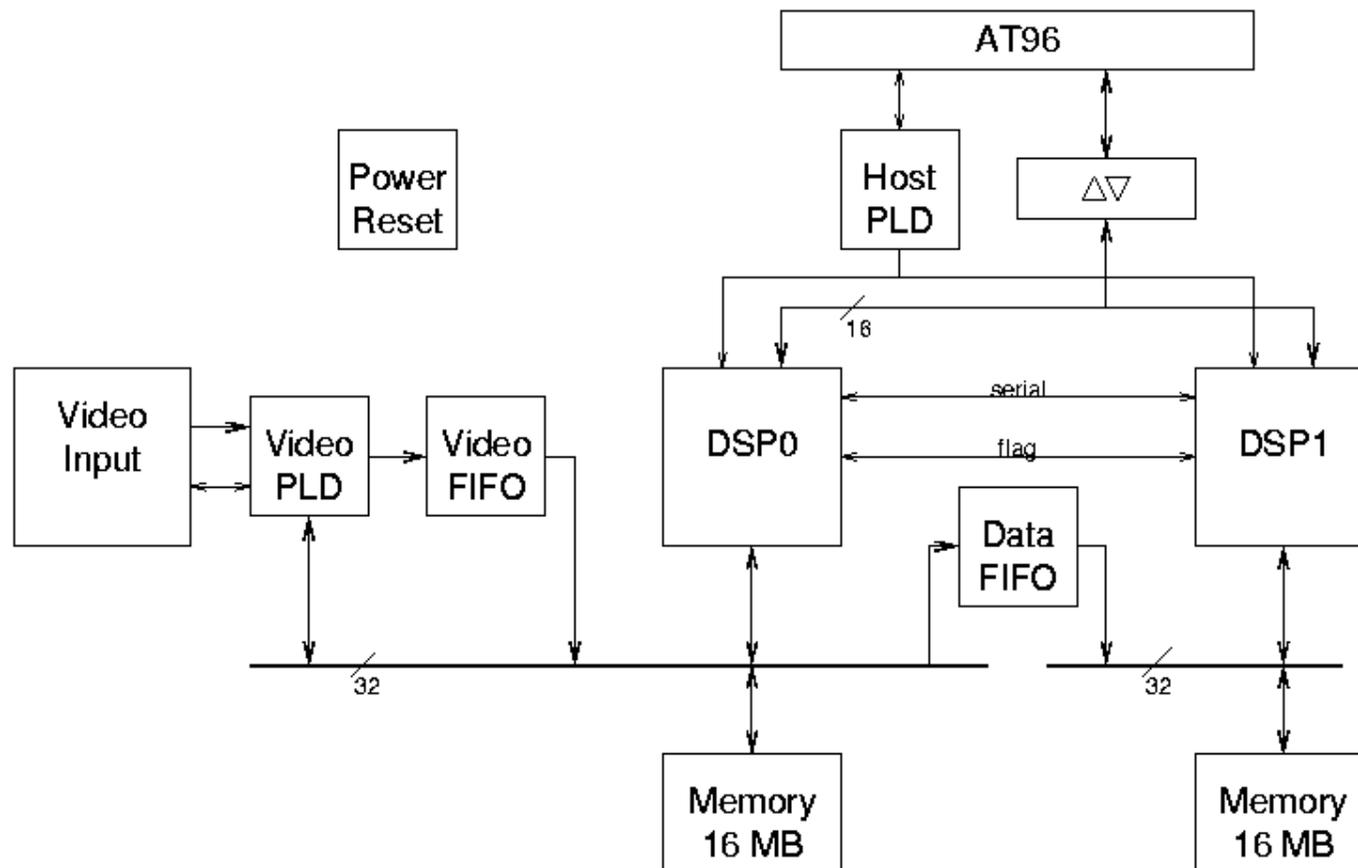
H.263 Echtzeit Encoder

- Maximale Bildgröße: CIF (352 x 288)
- 12,5 fps bei CIF (mehr bei kleineren Bildgrößen)
- Bitrate: 10 kbps bis 1 Mbps

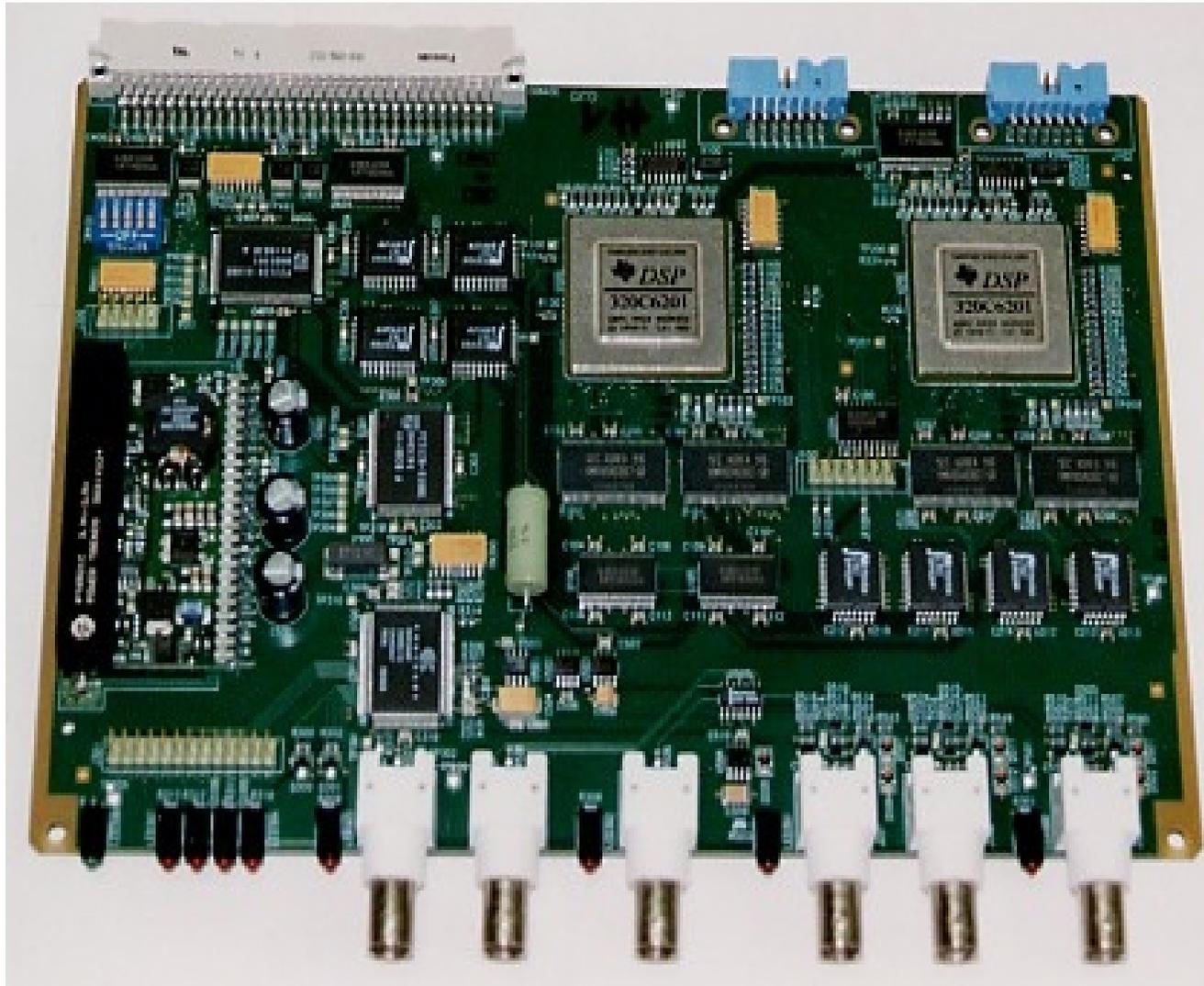
Funktionsblöcke

- Video Eingang
 - FBAS, Y/C
 - PAL, NTSC, SECAM
 - Größenskalierung
- 2 DSPs TMS320C6201 200MHz
 - DSP 0
 - Vorverarbeitung
 - Bewegungsschätzung
 - DSP 1
 - Prädiktion
 - Transformation, Quantisierung
 - Huffman Codierung

Beispiel Video Encoder System 2



Beispiel Video Encoder System 3



Zusammenfassung

Hardwareplattformen für Echtzeitvideoverarbeitung

- Prozessoren
- Systemkomponenten
- Systeme

Hardware als Grundlage zur Videoverarbeitung

Software: Nächste Vorlesungen