

OpenVMS Caching

Neues von XFC

Thilo Lauer
Compaq Customer Support Center
München

COMPAQ

1C05

25. DECUS-Symposium 2002, Bonn

Agenda

- ◆ XFC und OpenVMS V7.3: wat denn nu???
- ◆ Caching unter OpenVMS
 - Formen von Data Caching
 - Historie von XFC
- ◆ XFC
 - Motivation
 - Funktionalität
 - Skalierung
 - Kontrolle
 - Performance

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 2

Status von XFC...

Die Kundensicht

- ◆ April 2001: OpenVMS V7.3 SSB beinhaltet XFC
- ◆ Aug. 2001: VMS73_XFC-V0100
- ◆ Oct. 2001: VMS73_XFC-V0100 put on hold
- ◆ Apr. 2002: immer noch kein neuer XFC-ECO

! Wir wollen XFC !

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 3

...Status von XFC...

Situation in OpenVMS Engineering:

- ◆ April 2001: nach Verzögerung wird XFC endlich mit OpenVMS V7.3 ausgeliefert
- ◆ Aug. 2001: einige Bug Fixes mit VMS73_XFC-V0100, "stabile XFC-Version erreicht"
- ◆ Oct. 2001: VMS73_XFC-V0100 put on hold, Data Corruptions entdeckt
- ◆ Jan. 2002: XFC-Test mit Kundenapplikation, wieder mehrere Probleme entdeckt
- ◆ Apr. 2002: V7.3-1 mit XFC enabled, V7.3-Backport

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 4

...Status von XFC

warum dauert das so lange?

- ◆ Krisen-Meetings, teilweise auf VP-Ebene, Aktionsplan
- ◆ zusätzliche Ressourcen für XFC Engineering
- ◆ Design Review
- ◆ Entwicklung neuer Test-Suites
- ◆ Tests mit duplizierten Kundenumgebungen
- ◆ Eigenes großes Cluster nur für XFC-Tests erstellt
- ◆ extensive Verbesserungen, V7.3-1 FT hat XFC enabled, Engineering-Prod.-Systeme arbeiten mit XFC

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 5

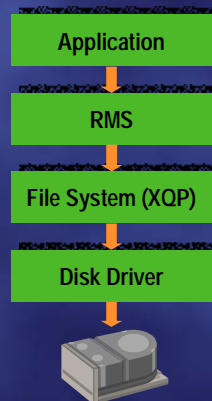
Was ist Caching?

- ◆ Wortstamm: cacher (frz.), kaschieren: etwas verstecken, verbergen
- ◆ Verborgen wird Leistungsengpass
 - CPU Caches: Durchsatz von Memory-Bus
 - I/O Caches: Durchsatz von Data Bus
- ◆ Caching verbirgt Leistungsengpässe auf Kosten anderer Ressourcen
 - sinnlos, wenn System generell überlastet ist

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 6

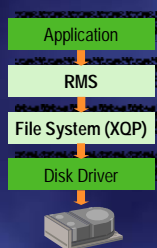
Caching unter OpenVMS I/O-Modell



COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 7

Caching unter OpenVMS... Virtual Block (VBN) Caches



◆ Produkte

- (RMS Local/Global Buffers)
- VIOC, XFC
- third party Caches

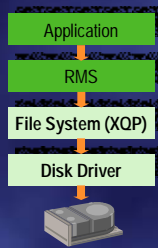
◆ Eigenschaften

- applikationsnah
- verstehen die File-Struktur: Readahead einfach
- verstehen *nicht* die On-Disk-Struktur

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 8

Caching unter OpenVMS... Logical Block (LBN) Caches



◆ Produkte

- RAM Disks
- (third party Caches)

◆ Eigenschaften

- verstehen keine File-Struktur
- manipuliert Device Driver I/O, applikationsfern
- gut für Storage Server
- "Interessante" Synchronisationsprobleme im Cluster

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 9

Caching unter OpenVMS Historie von XFC

- ◆ OpenVMS VAX V6.0: Virtual I/O Cache (VIOC)
- ◆ OpenVMS AXP V6.1: Portierung von VIOC auf Alpha: VaxCluster Cache (VCC)
- ◆ auch erwähnt als Virtual Data Cache (VDC)
- ◆ OpenVMS Alpha V7.3: XFC V1.0

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 10

eXtended File Cache (XFC): Motivation

VIOC

- ◆ Cache liegt im S0/S1 Space, dadurch Limit von ca. 1GByte Cache-Größe
- ◆ Hard Coded Limit der I/O-Size von max. 34 Blocks
- ◆ Statische Cache-Größe auf Alpha
 - (Default ist 6400 blocks!)
- ◆ Hard Coded "closed file"-Limit von 100
- ◆ kein Caching von Files, welche für Schreibzugriff geöffnet sind

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 11

XFC: Eigenschaften

- ◆ Alpha only
- ◆ Erstversion in V7.3
- ◆ Koexistenz mit VCC/VIOC
- ◆ benutzt S2 Space - skalierbarer, unlimitierter Cache
- ◆ Memory Allocation dynamisch und/oder statisch
- ◆ Readahead-Fähigkeit
- ◆ erweiterte Statistik-/Reporting-Funktionen
- ◆ keine "closed file"-Limitation
- ◆ schaltet sich erst bei Auftreten von Writes ab
- ◆ Basis-Plattform für zukünftige Weiterentwicklungen

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 12

XFC: Einschalten

- ◆ VCC_FLAGS
 - 0 = kein Caching
 - 1 = VIOC-Mode
 - 2 = XFC-Mode
- ◆ nicht dynamisch - Reboot erforderlich
- ◆ Defaults
 - in V7.3: XFC mode
 - aber: Compaq empfiehlt VIOC mode!!!
 - V7.3-1: XFC mode

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 13

XFC: weitere SYSGEN-Parameter

- ◆ VCC_MAXSIZE
 - obsolet: wird von XFC nicht benutzt
- ◆ VCC_MAX_CACHE (in MB)
 - Default ist -1 (= 50% des physikalischen Memory)
- ◆ VCC_MAX_IO_SIZE
 - Default ist 127 Blocks
- ◆ VCC_READAHEAD
 - Default ist 1 (= enabled)
- ◆ weitere VCC*-Parameter
 - werden gegenwärtig nicht benutzt

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 14

SHOW MEMORY/CACHE (bisher)

```
$ show mem/cache
System Memory Resources on 5-APR-2002 13:02:07.23

Virtual I/O Cache (Kbytes):  Total      Free      In Use
Cache Memory                3200      600      2600
$
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 15

SHOW MEMORY/CACHE (neu)

```
$ SHOW MEMORY/CACHE

System Memory Resources on 5-APR-2002 02:05:17.92

Extended File Cache (Time of last reset: 27-MAR-2002 18:34:05.39)
Allocated (Mbytes) 134.96 Maximum size (Mbytes) 1024.00
Free (Mbytes) 0.11 Minimum size (Mbytes) 0.23
In use (Mbytes) 134.85 Percentage Read I/Os 61%
Read hit rate 44% Write hit rate 0%
Read I/O count 612787 Write I/O count 386512
Read hit count 272095 Write hit count 0
Reads bypassing cache 1765 Writes bypassing cache 0
Files cached open 327 Files cached closed 263
Vols in Full XFC mode 0 Vols in VIOC Compatible mode 8
Vols in No Caching mode 0 Vols in Perm. No Caching mode 0
$
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 16

Neue SHOW MEM/CACHE-Kommandos...

SHOW MEMORY/CACHE, . . .

- ◆ =(VOLUME=DKA200)
zeigt Volume-Level-Statistik
- ◆ =(TOPQI 0=3)
zeigt Files mit höchster I/O-Rate
- ◆ =(TOPHI TRATE)
zeigt Files mit höchster Hitrate im Cache
- ◆ =(FILE=DKA200: [VMS\$COMMON. SYSMGR] *. COM, OPEN, CLOSED)
zeigt Statistik für bestimmte Files
- ◆ alle Keywords können beliebig kombiniert werden

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 17

...Neue SHOW MEM/CACHE-Kommandos...

```
$ SHOW MEMORY/CACHE=(VOLUME=$14$DKA400, TOPHI TRATE=1, OPEN, CLOSED)
```

```
System Memory Resources on 5-APR-2002 04:32:06.93
```

```
Extended File Cache Top Hitrate File Statistics:
```

```
_$14$DKA400: (DISK$TEST), Caching mode is VI OC Compati ble
```

```
_$14$DKA400: [LAUER. GBLBUF. REGRES]GBL_STRESS. EXE; 40 (open)
```

```
Caching is enabled, active caching mode is Write Through
```

Allocated pages	6	Total QIOs	119382
Read hits	119322	Virtual reads	119382
Virtual writes	0	Hit rate	99 %
Read aheads	0	Read throughs	119358
Write throughs	0	Read arounds	24
		Write arounds	0

```
Total of 1 file for this volume
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 18

...Neue SHOW MEM/CACHE-Kommandos

```
$ SHOW MEMORY/CACHE=(FILE=DKA0: [VMS$COMMON.SYSMGR]LOGIN.COM)
```

System Memory Resources on 8-APR-2002 07:47:52.84

Extended File Cache File Statistics:

```
_$7$DKA0: [VMS$COMMON.SYSMGR]LOGIN.COM; 1 (closed)
```

Caching is enabled, active caching mode is Write Through

Allocated pages	1	Total I/Os	9
Read hits	8	Virtual reads	9
Virtual writes	0	Hit rate	88 %
Read aheads	0	Read throughs	9
Write throughs	0	Read arounds	0
		Write arounds	0

Total of 1 file for this volume

```
$
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 19

Caching-Einstellungen für XFC V1.0

- ◆ statische und/oder dynamische Memory-Allokierung
- ◆ erlaubter Memory-Verbrauch
- ◆ was soll gecached werden?
- ◆ maximale Größe von cached I/Os

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 20

Caching-Einstellungen -Memory Allocation...

- ◆ default: dynamisch
 - Memory wird angefordert/freigegeben abhängig von Systemauslastung
- ◆ statische Allokierung, um jeglichen Overhead der Cache-Verwaltung zu vermeiden
- ◆ Wie?
 - SYSMAN> RESERVED_MEMORY ADD VCC\$MIN_CACHE_SIZE /SIZE=500/ALLOCATE/NOPAGE_TABLES/NOZERO/NOGLOBAL
- ◆ VCC_MAX_CACHE setzen (nur Verkleinerung!):
 - =500: rein statisch
 - >500: statisch + dynamisch
 - <500: verschenkt Memory

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 21

Caching-Einstellungen -Memory Allocation...

```
$ SHOW MEMORY
System Memory Resources on 10-APR-2002 04:04:42.53

Physical Memory Usage (pages):      Total      Free      In Use    Modified
Main Memory (16.00Gb)              2097152    1637181   459561    410

Extended File Cache (Time of last reset: 9-APR-2002 10:40:14.76)
Allocated (Mbytes)                  1000.00    Maximum size (Mbytes)      8192.00
Free (Mbytes)                       952.14    Minimum size (Mbytes)      1000.00
In use (Mbytes)                      47.85     Percentage Read I/Os       97%
Read hit rate                        1%        Write hit rate              0%
Read I/O count                       3682287   Write I/O count             77407
Read hit count                       52919     Write hit count             0
Reads bypassing cache                171       Writes bypassing cache      0
Files cached open                    260       Files cached closed         174
Vols in Full XFC mode                 0         Vols in VIOC Compatible mode 13
Vols in No Caching mode               2         Vols in Perm. No Caching mode 0
:                                     :         :                             :
:                                     :         :                             :
Memory Reservations (pages):        Group      Reserved    In Use      Type
VCC$MIN_CACHE_SIZE                  ---        128000     128000     Allocated
Total (1000.00 Mbytes reserved)      128000    128000
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 22

Caching-Einstellungen - was soll gecached werden?

- ◆ erfordert Kenntnis der Applikationen
- ◆ für Volume:
\$MOUNT/[NO]CACHE
 - Vorsicht: schaltet auch die XQP-Caches für das Volume ab!
- ◆ für einzelne Files:
\$SET FILE/CACHE=[NO|WRITE THROUGH]
- ◆ für einzelnen \$QIO: Qualifier IO\$M_NOCACHE

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 23

Caching-Einstellungen - maximale Größe von cached I/Os...

- ◆ default: 127-block I/Os
- ◆ Beurteilungshilfe:
 - \$SHOW MEMORY/CACHE/FULL
 - SDA> XFC SHOW SUMMARY
- ◆ Einstellung mit VCC_MAX_IO_SIZE
- ◆ Reset der Counter mit:
 - \$SET CACHE/RESET

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 24

Caching-Einstellungen - ...maximale Größe von cached I/Os...

```
$SHOW MEMORY/CACHE/FULL
I/O Statistics - Distributions
-----
Transfer Size:      Reads  Read Hits  Writes
1 Block IO:        3458   2545      1171
2 Block IO:         527    84         30
3 Block IO:         351   165         3
:                   :       :           :
7 Block IO:        236   170         6
8 Block IO:         621   549         1
:                   :       :           :
16 Block IO:       1802  1417         9
:                   :       :           :
64 Block IO:        850   386         0
:                   :       :           :
>127<256 Block IO:  10     :           0
>255   Block IO:   15     :           0
-----
Total s:           10905  6980      1230
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 25

Caching-Einstellungen - ...maximale Größe von cached I/Os

- ◆ CPU-Entlastung durch Limitierung der I/O-Größe
- ◆ sind keine freien CPU-Ressourcen vorhanden, kann Menge der XFC-behandelten I/Os begrenzt werden
- ◆ Cache-Effizienz wird gemindert - zugunsten anderweitig nutzbarer CPU-Leistung

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 26

XFC-Implementation

- ◆ 2 Images:
 - SYS\$XFCACHE.EXE
 - Normales "streamlined" Image
 - SYS\$XFCACHE_MON.EXE
 - "full tracing" Image mit erweiterten Checks
- ◆ Auswahl durch SYSTEM_CHECK = 0|1
- ◆ Oder Umbenennung in SYS\$XFCACHE.EXE

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 27

XFC-spezifische SDA-Erweiterungen

- ◆ Primär für Entwickler und Support gedacht
- ◆ Einige nützliche Kommandos:
 - SDA>XFC SHOW SUMMARY
 - SDA>XFC SHOW TRACE
 - SDA>XFC SHOW VOLUME/FULL
 - SDA>XFC SHOW FILE/ID=fi le_i d

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 28

...XFC-spezifische SDA-Erweiterungen...

SDA> SHOW SUMMARY

```

:
:
:
Virtual Reads:                    51968
  Sum of Read Around Count:        27
  and Read Through Count:         51941
Reads Completed:                  51968
Read Hits:                        36467
Read Cache Hit Percentage:        70.17 %
Total Synch Completion Count:     32783
Read Around due to Het. Cluster:  0
Read Around due to Modifiers:     0
Read Around due to Size:          27
Total reads past EOF:             12404
Total I/Os with read-ahead:       5824

Paging I/Os:                      455

```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 31

...XFC-spezifische SDA-Erweiterungen...

SDA> XFC SHOW TABLE

FHT: Contents of 1024 buckets

```

. 1 2 . . . 2 1 . . . 1 . . . 2 . . .
1 1 1 1 . . . . . 2 . . . 1 1 . .
. 1 1 . . 1 . . . 1 . . . 1 1 . .
. . . 1 1 . . . . 1 2 . . 2 . 1 1 .
. 1 . . . . . 1 . 1 . 1 . 2 . . .

```

```

0(589)
1(359)
2(70)
3(6)

```

```

Total number of CFBs: 517
Longest chain length: 3
Shortest chain length: 0
Average chain length: 0.50

```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 32

...XFC-spezifische SDA-Erweiterungen...

```
XFC> sh file/id=18/stat
```

```
CFB Address:                FFFFFFFF7BDE2570
Flink:                      FFFFFFFF7BDE2300
Blink:                      FFFFFFFF7BDE2820
Access Count:                1
Write Access Count          0
Volume (CVB):               FFFFFFFF7C000050
Quiescing:                  False
File (FCB):                 FFFFFFFF8138A000
Volume Id:                  FFFFFFFF7C000050
File Id:                    0000000000000018
External FID:                (24, 19, 0)
Predicted Next VBN:         0000002D ( 45)
Active Caching Mode:        Write Through
Active I/O count:           0
File Bad:                   False
Flush Fail Status:          00000000 ( 0)
No Readahead Reasons:       0
Active Readaheds:           0
Caching disabled:          False
:                            :
:                            :
:                            :
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 33

...XFC-spezifische SDA-Erweiterungen

```
XFC> sh file/id=18/stat
```

```
:                            :
:                            :
:                            :
Total I/Os to this file:    6
Read I/Os to this file:    6
Write I/Os to this file:    0
Read Hits:                  3
Hit Rate:                   50.00 %
Average Overall I/O response time to this file in microseconds: 4757
Average Cache Hit I/O response time to this file in microseconds: 443
Average Disk I/O response time to this file in microseconds: 9070
Accuracy of I/O resp time: 100 %
Read Ahead Count:          2
Read Through Count:        6
Write Through Count:       0
Read Around Count:         0
Write Around Count:        0
```

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 34

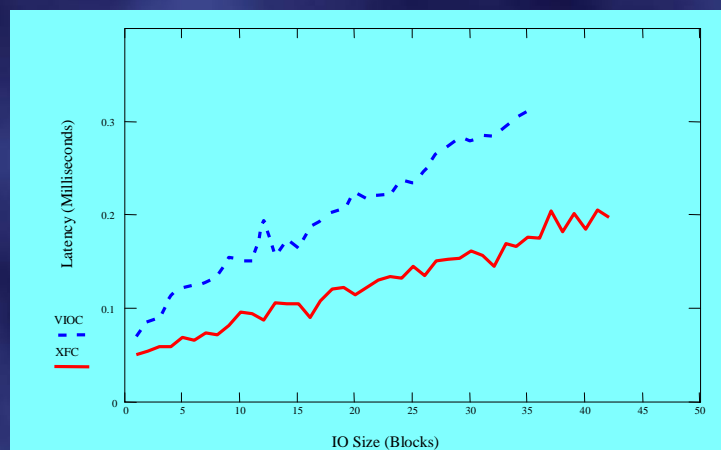
XFC-Performance...

- ◆ spezifisch für jedes System und jede Applikation
- ◆ meist künstliche Test-Suites, die nicht unbedingt reale Anforderungen wiedergeben
- ◆ Realer Test: Antwortzeit-Verhalten Ihrer Applikation auf Ihrem System
- ◆ einige Beispiele...

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 35

...Performance - Read I/O Latency



Test conditions: Single Size, Random VBN, 10 minute samples, File size 450,000 blocks

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 36

...Performance - Elapsed Time: Lesen kleiner Dateien

- ◆ 1.000 block File, 100.000 I/Os
 - VIOC: 5 min 54 sec
 - XFC: 4 min 14 sec

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 37

...Performance - Elapsed Time: Lesen großer Dateien

- ◆ 1.500.000 block File, 300.000 I/Os
 - VIOC: 50 min 24 sec
 - XFC: 17 min 36 sec
- ◆ 2.000.000 block File, 300.000 I/Os
 - VIOC: 54 min 22 sec (+8%)
 - XFC: 18 min 17 sec (+4%)

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 38

XFC Futures/Phantasies

- ◆ XFC V2.0 im Planungs-Stadium
- ◆ Features
 - erweiterte Kontrolle durch System Manager
 - File Priorities
 - Memory Allocation
 - erweitertes File I/O Monitoring
 - bessere Performance
 - Spinlock bottleneck
 - /NOCACHE-Handling
 - Memory Reclamation
 - Cache Hits
 - AMDS Interface

Keine Garantie!!!

Keine Garantie!!!

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 39

Fragen?

Thilo.Lauer@compaq.com

COMPAQ

Thilo Lauer, Neues von XFC (1C05), Seite 40