

Einführung in z/OS Enterprise Computing

**Prof. Dr. Martin Bogdan
Dr. rer. nat. Paul Herrmann
Prof. Dr.-Ing. Wilhelm G. Spruth**

WS 2009/2010

Teil 12

Sysplex Coupling Facility

Literatur

Wilhelm G. Spruth, Erhard Rahm:

Sysplex-Cluster Technologien für Hochleistungs-Datenbanken.
Datenbank-Spektrum, Heft 3, 2002, S. 16-26.

Verfügbar (download):

<http://www-ti.informatik.uni-tuebingen.de/~spruth/publish.html>

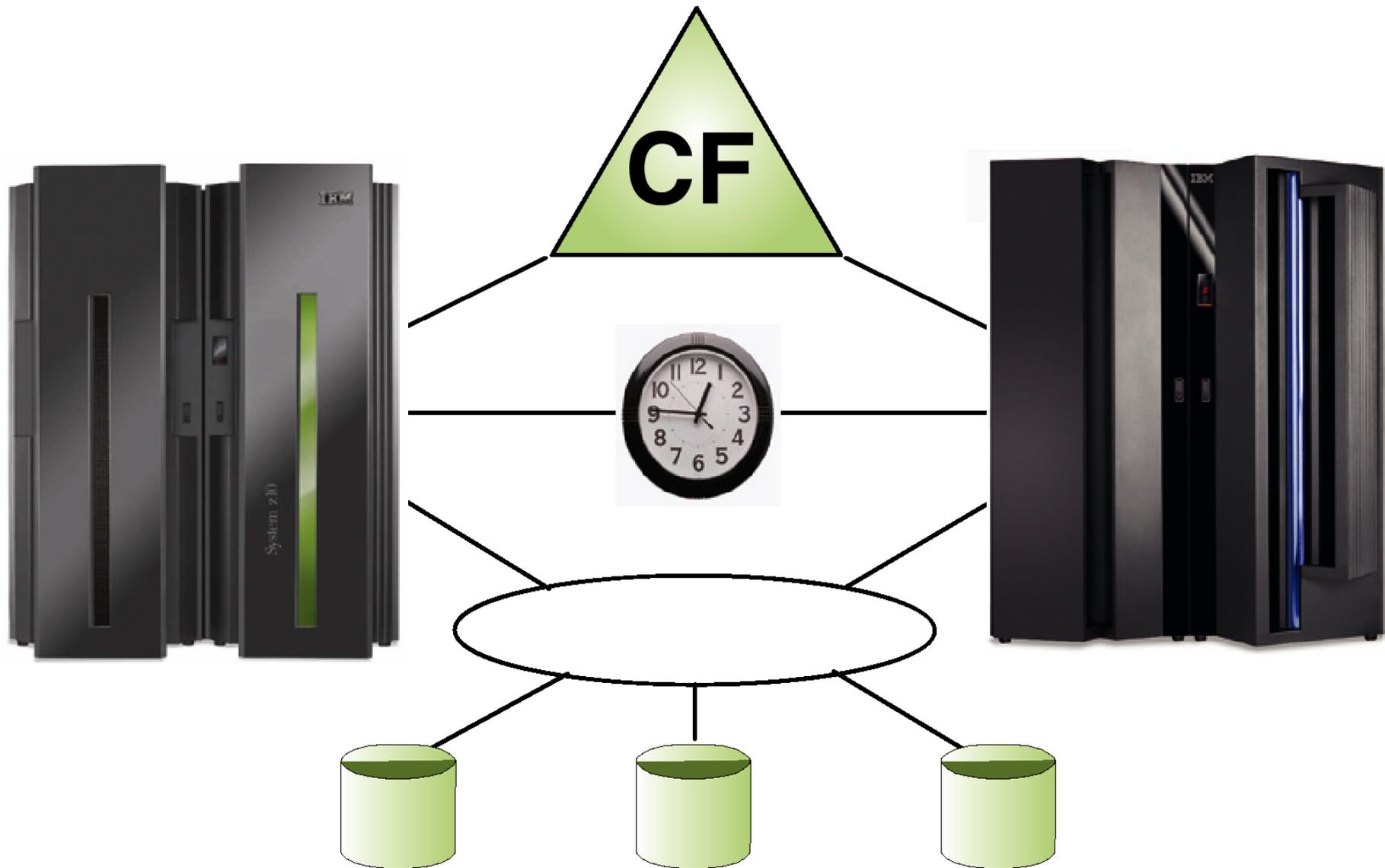
Sysplex Hardware:

Sonderheft des IBM Journal of Research and Development, Vol. 36, No.4, July 1992.

Sysplex Software:

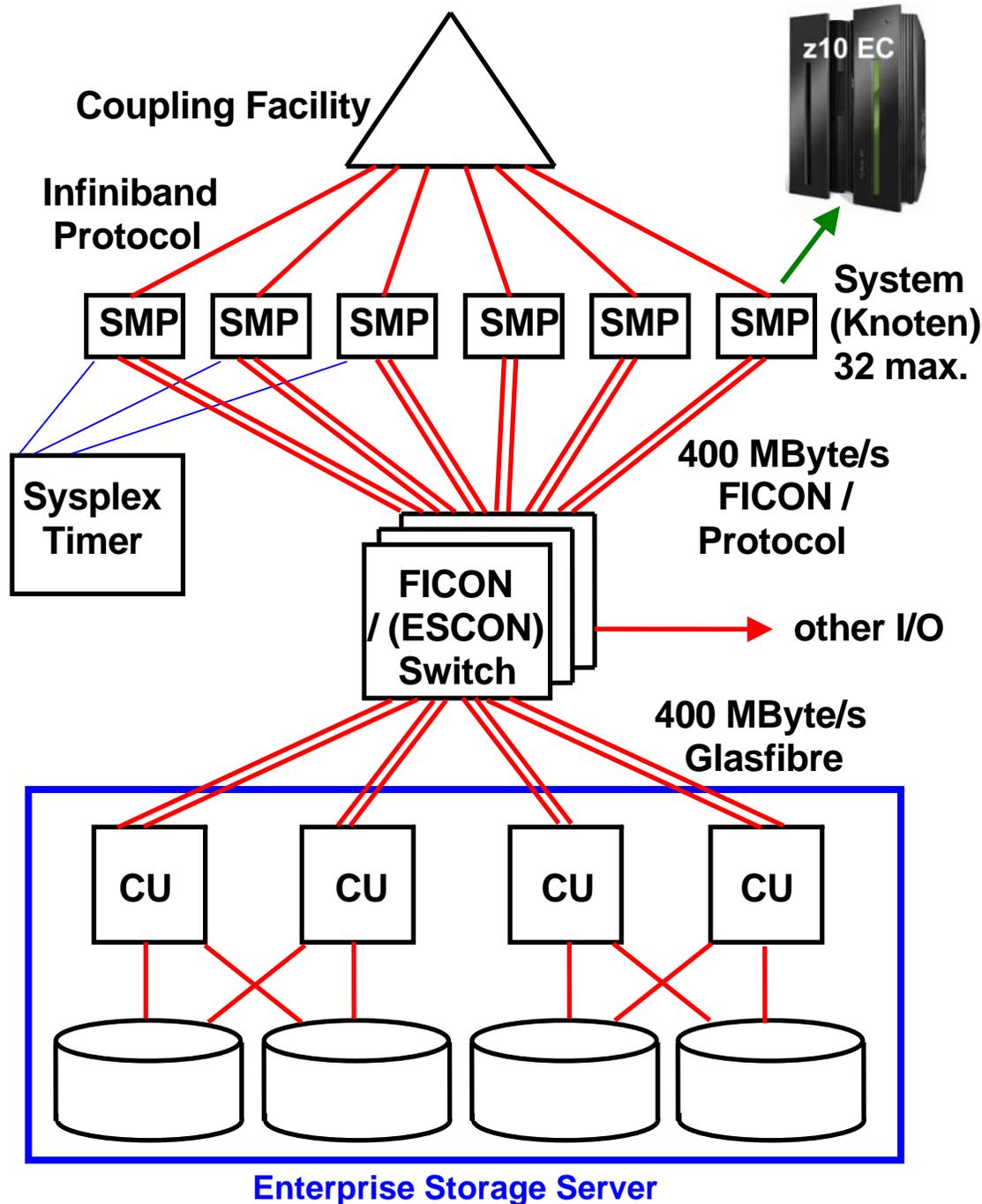
Sonderheft des IBM System Journal, Vol. 36, No.2, April 1997.

Verfügbar (download): [//www.research.ibm.com/journal](http://www.research.ibm.com/journal)



zSeries Coupling Facility

Großrechner bearbeiten mehrere 1000 Transaktionen/s unter ACID Bedingungen



Sysplex with Coupling Facility

Ein Sysplex besteht aus bis zu 32 System z Rechnern, die über FICON Glasfasern und FICON Switche mit Plattenspeichern in der Form von einem oder mehreren Enterprise Storage Servern verbunden sind. Ein Enterprise Storage Server emuliert mehrere Control Units (CU).

Weiterhin sind 1 oder in der Regel 2 Coupling Facilities vorhanden.

Eine Coupling Facility ist ein regulärer System z oder S/390 Rechner, auf dem „Coupling Facility Code“ an Stelle eines Betriebssystems läuft.

Weiterhin sind 1 oder 2 Zeitgeber (Sysplex Timer) vorhanden

Parallel Sysplex Cluster Technology

Mehrfache z/OS Systeme verhalten sich so, als wären sie ein einziges System (Single System Image).

Parallel Sysplex Cluster Technology Komponenten:

- **Prozessoren mit Parallel Sysplex Fähigkeiten**
- **Coupling Facility**
- **Coupling Facility Control Code (CFCC)**
- **Glasfaser Hochgeschwindigkeitsverbindungen**
- **ESCON oder FICON Switch**
- **Sysplex Timer**
- **Gemeinsam genutzte Platten (Shared DASD)**
- **System Software**
- **Subsystem Software**

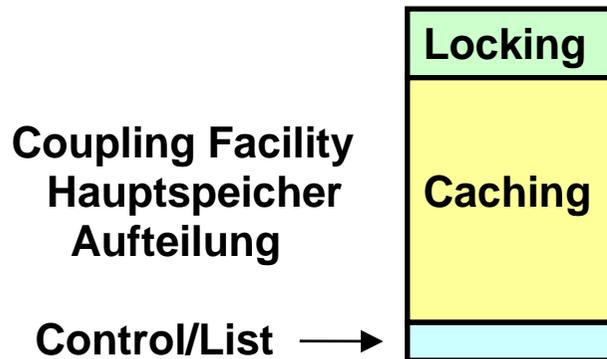
Die Coupling Facility ermöglicht Data Sharing einschließlich Datenintegrität zwischen mehrfachen z/OS Servern

Der Sysplex Zeitgeber (Timer) stellt allen z/OS und OS/390 Instanzen eine gemeinsame Zeitbasis zur Verfügung. Dies ermöglicht korrekte Zeitstempel und Ablaufsequenzen bei Datenbank Änderungen. Dies ist besonders bei Datenbank-Recovery Operationen wichtig.

Coupling Facility

Die Coupling (CF) Facility ist in Wirklichkeit ein weiterer zSeries Rechner mit spezieller Software (kein reguläres Betriebssystem). Die Aufgaben der CF sind:

- Locking
- Caching
- Control/List Structure



Die wichtigste Aufgabe der Coupling Facility ist ein zentrales Lock Management für die angeschlossenen Systeme. Der zentrale Lock Manager des SAP System R/3 hat in Ansätzen eine ähnliche Funktionalität.

Der größte Teil des Hauptspeichers der Coupling Facility wird als Plattenspeicher Cache genutzt. Der CF Cache dupliziert den Plattenspeicher Cache in den einzelnen Systemen. Cast out der CF Cache auf einen Plattenspeicher erfolgt über ein System.

Control und List Strukturen dienen der Sysplex Cluster weiten Verwaltung. Beispiel: RACF Sicherheits Subsystem.

Die Coupling Facility ist über Glasfaser Verbindungen mit einem optimierten Protokoll und spezieller Hardware Unterstützung mit den Knoten (Systemen) des Sysplex verbunden.

Coupling Facility Locking

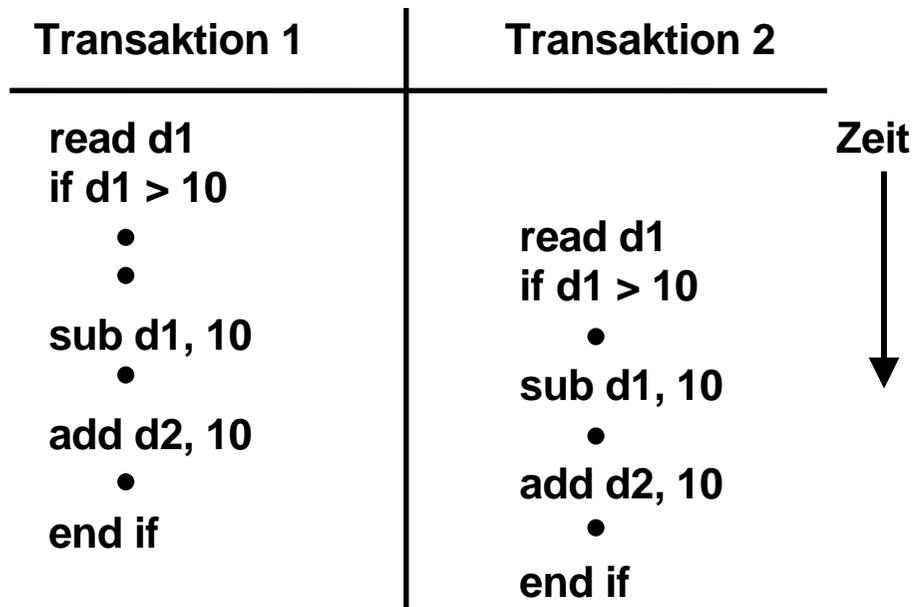
Coupling Facility Caching

Coupling Facility Control/List Strukturen

Locking Problem

Angenommen zwei Transaktionen, die **auf unterschiedlichen Systemen** des Sysplex laufen. Beide Transaktionen greifen auf die beiden Datenbank - Variablen d1 und d2 zu.

Anfangswerte: d1 = 15, d2 = 20 .



Die beiden Abhängigkeiten:

1. **Dirty Read**

Eine Transaktion erhält veraltete Information (hier der Fall)

2. **Lost update**

Eine Transaktion überschreibt die Änderung einer anderen Transaktion

müssen gesteuert werden.

Das Ergebnis ist: d1 = - 5, d2 = 40 , obwohl die if-Bedingung ein negatives Ergebnis verhindern sollte.

Benutzung von Locks (Sperren)

Transaktion 1	Transaktion 2
GetReadLock (d1) read d1 if d1 > 10	
GetWriteLock (d1) → <i>Nachricht an Transaktion 2</i> GetWriteLock (d2) sub d1, 10 add d2, 10 ReleaseLocks	GetReadLock (d1) read d1 if d1 > 10
end if	GetReadLock (d1) read d1 if d1 > 10 GetWriteLock (d1) GetWriteLock (d2) sub d1, 10 add d2, 10 end if

Das Senden einer Nachricht informiert Transaktion 2, dass ihr Wissen über den Zustand der beiden Variablen d1 und d2 nicht mehr gültig ist.

Transaktion 2 muss sich neu über den Zustand der Variablen d1 informieren, ehe sie ein Update von d1 vornimmt.

Frage: Woher weiß Transaktion 1, dass Transaktion 2 ein Interesse an der Variablen d1 hat (ein shared Lock besitzt) ?

Weitere Frage: Muss Transaktion 1 auch Transaktion 2 bezüglich des Write Locks für die Variable d2 benachrichtigen ?

Ergebnis: d1 = + 5, d2 = 30

Two-Phase Locking - Two-Phase Transaction

In Transaktionssystemen und Datenbanksystemen werden Locks (Sperrungen) benutzt, um Datenbereiche vor einem unautorisierten Zugriff zu schützen. Jedem zu schützenden Datenbereich ist ein Lock fest zugeordnet. Ein Lock ist ein Objekt welches (mindestens) über 4 Methoden und zwei Zustände S und E verfügt. Die Methode

- **GetReadLock** reserviert **S** Lock (shared),
- **GetWriteLock** reserviert **E** Lock (exclusive),
- **PromoteReadtoWrite** bewirkt Zustandswechsel S → E,
- **Unlock** gibt Lock frei.

Mehrere Transaktionen können ein S Lock für den gleichen Datenbereich (z.B. einen Datensatz) besitzen. Nur eine Transaktion kann ein E Lock für einen gegebenen Datenbereich besitzen. Wenn eine Transaktion ein S Lock in ein E Lock umwandelt, müssen alle anderen Besitzer des gleichen S Locks benachrichtigt werden.

Normalerweise besitzt eine Transaktion mehrere Locks.

In einer Two-Phase Transaktion finden alle Lock Aktionen zeitlich vor allen Unlock Aktionen statt. Eine Two-Phase Transaktion hat eine Wachstumsphase (growing), während der die Locks angefordert werden, und eine Schrumpf- (shrink) Phase, in der die Locks wieder freigegeben werden.

Two Phase Locking ist nicht zu verwechseln mit dem 2-Phase Commit Protokoll der Transaktionsverarbeitung

Locking Protokoll

Vorgehensweise:

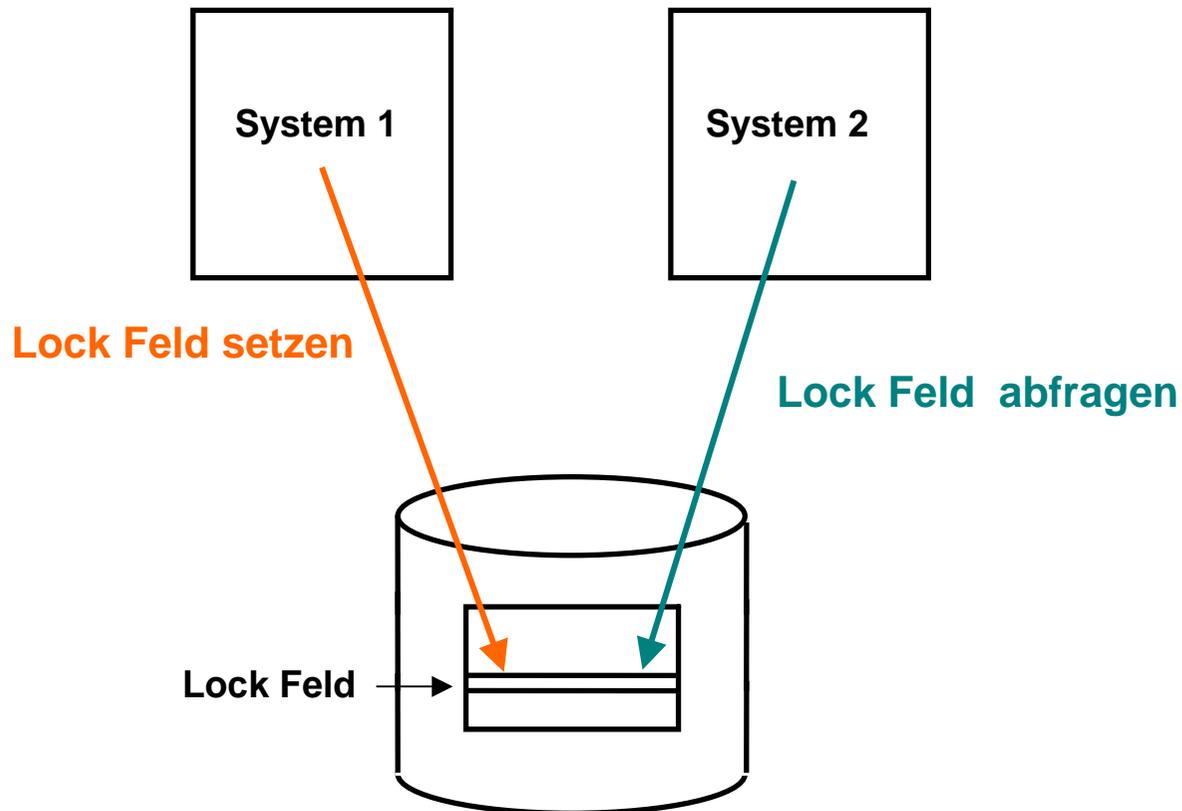
- Shared Lock (S) erwerben vor dem erstmaligen Lesen
- Exclusive Lock (E) erwerben vor dem erstmaligen Schreiben

derzeitiger Status Anforderung	kein	Lesen shared	Schreiben exclusive
Lesen Share	bewilligt, share-mode	bewilligt, share-mode	abgelehnt, Mitteilung über Besitzer
Schreiben Exclusive	bewilligt, exclusive mode	bewilligt, Warnung über Besitzer	abgelehnt, Mitteilung über Besitzer

Mehrere Transaktionen können das gleiche Lock im Zustand S besitzen. Nur eine Transaktion kann ein Lock im Zustand E besitzen.

Eine Transaktion kann ein Lock vom Zustand S in den Zustand E überführen. Hierzu ist es erforderlich, dass eine Nachricht an alle anderen transaktionen geschickt wird, die das gleiche Lock im Zustand S besitzen. Mittels dieser Nachricht wird das Lock für ungültig erklärt.

Lock Verwaltung



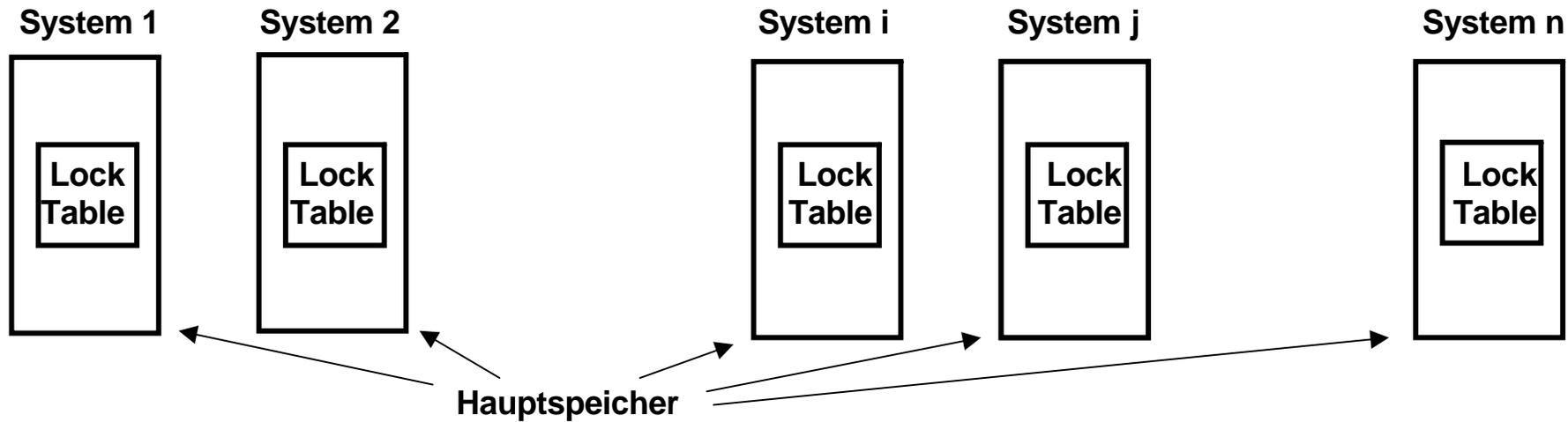
Einfachste Lösung:

Jeder zu schützende Datenbereich auf der Festplatte (z. B. eine Zeile in einer relationalen Datenbank-Tabelle) erhält ein zusätzliches Lock-Feld.

Bei einem Zugriff wird das Lock zunächst geprüft und dann gesetzt, ehe ein Zugriff erfolgt.

Nachteil: Die erforderlichen zusätzlichen Zugriffe auf den Plattenspeicher sind in Hochleistungssystemen nicht akzeptabel.

Lösung: die Lock Tabellen werden in den Hauptspeichern der angeschlossenen Systeme gehalten.



Verteilte Lock Tabelle

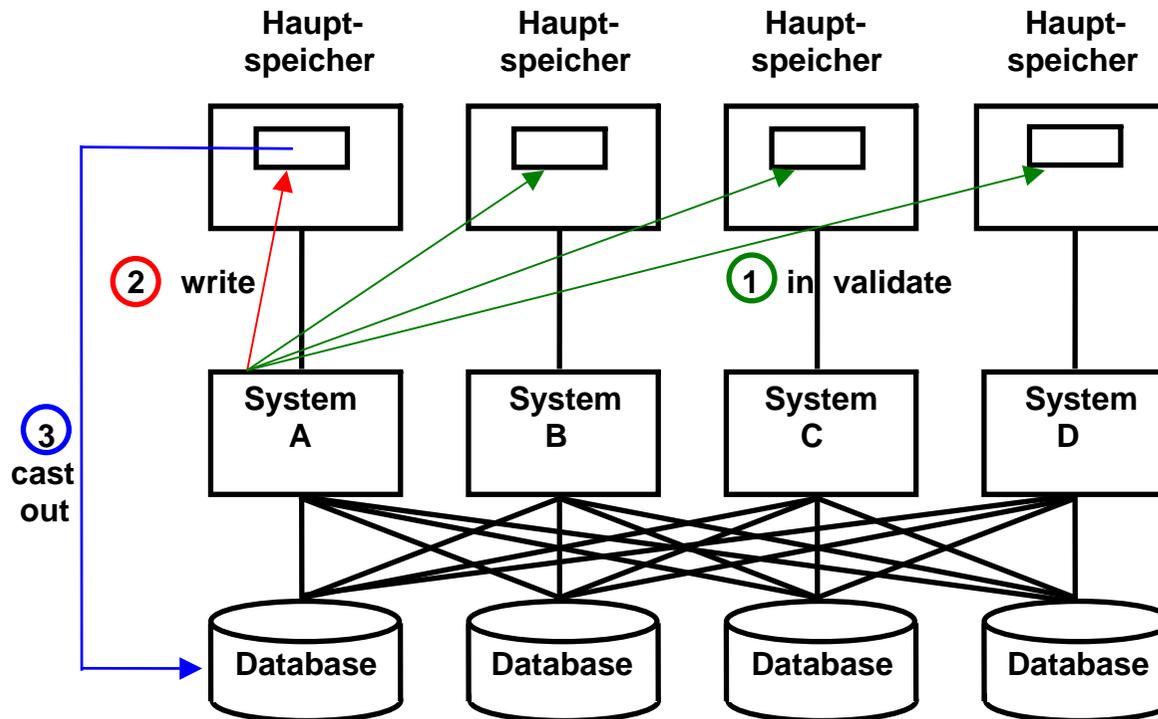
Decentralized Locks

Zugriffe auf Tabellen mit Locks erfolgen recht häufig. Deswegen werden Lock Tabellen in der Regel im Hauptspeicher gehalten. Bei einem SMP mit mehreren Prozessen aber nur einem Hauptspeicher ist dies unkritisch.

Bei einem Cluster mit mehreren Hauptspeichern sind auch mehrere Lock Tabellen in den Hauptspeichern der beteiligten Systeme vorhanden, die alle auf dem gleichen Stand gehalten werden müssen (verteilte Lock Tabelle) .

Zur Auflösung von Lock Konflikten erfolgt bei jeder Änderung in einer Lock Tabelle entweder ein Broadcast (Invalidate-Broadcast Kohärenzsteuerung) oder eine gezielte Nachricht von System i an System j.

Ersteres erfordert die Verarbeitung der laufenden Transaktion auszusetzen. Der Overhead kann 20 ms betragen,

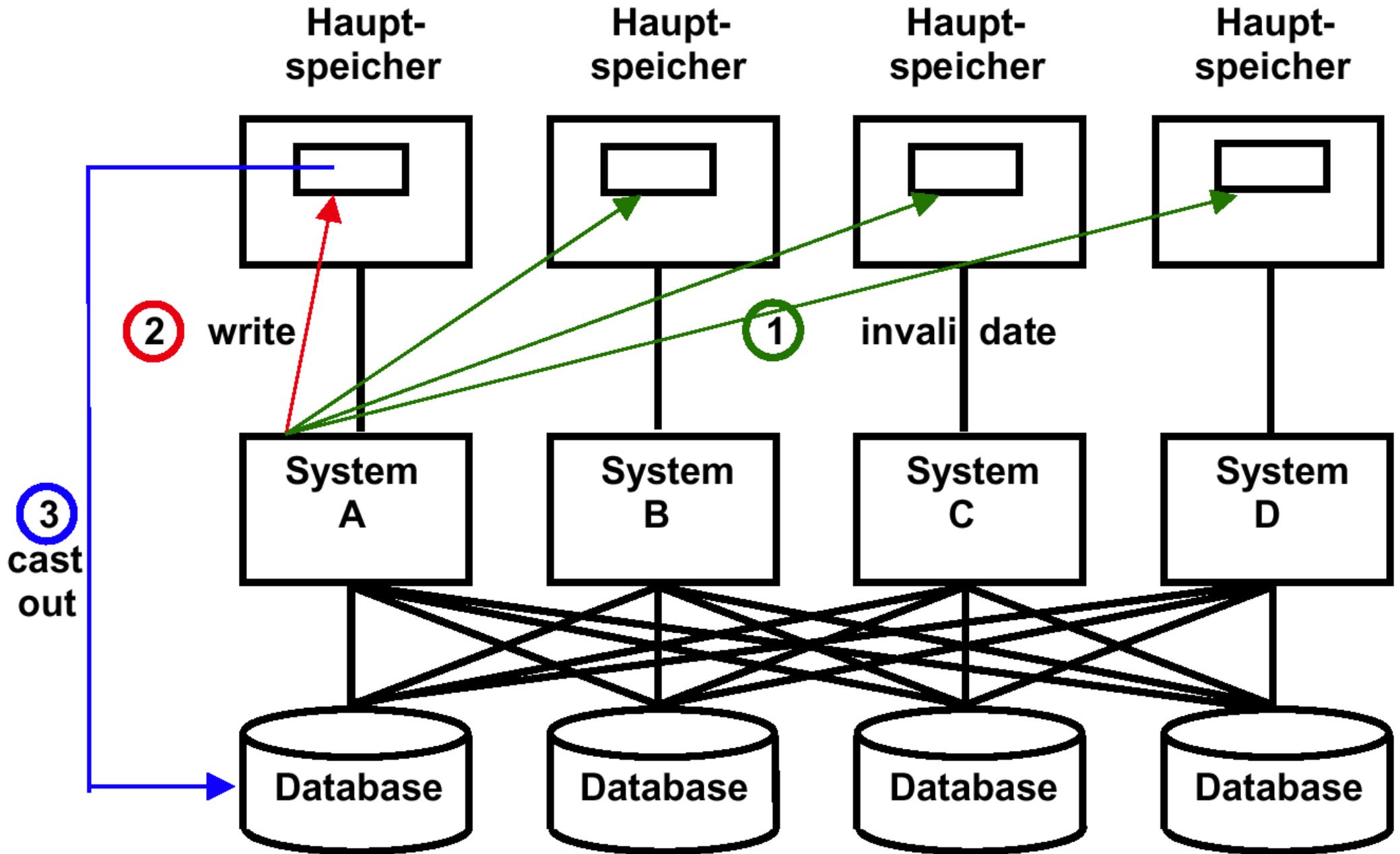


Invalidate-Broadcast Kohärenzsteuerung

Nur System A besitzt Write Lock. B, C und D besitzen nur Read Lock. Ein Invalidate Broadcast an alle Systeme des Clusters benachrichtigt B, C und D dass die Kopie des Kocks nicht mehr gültig ist.

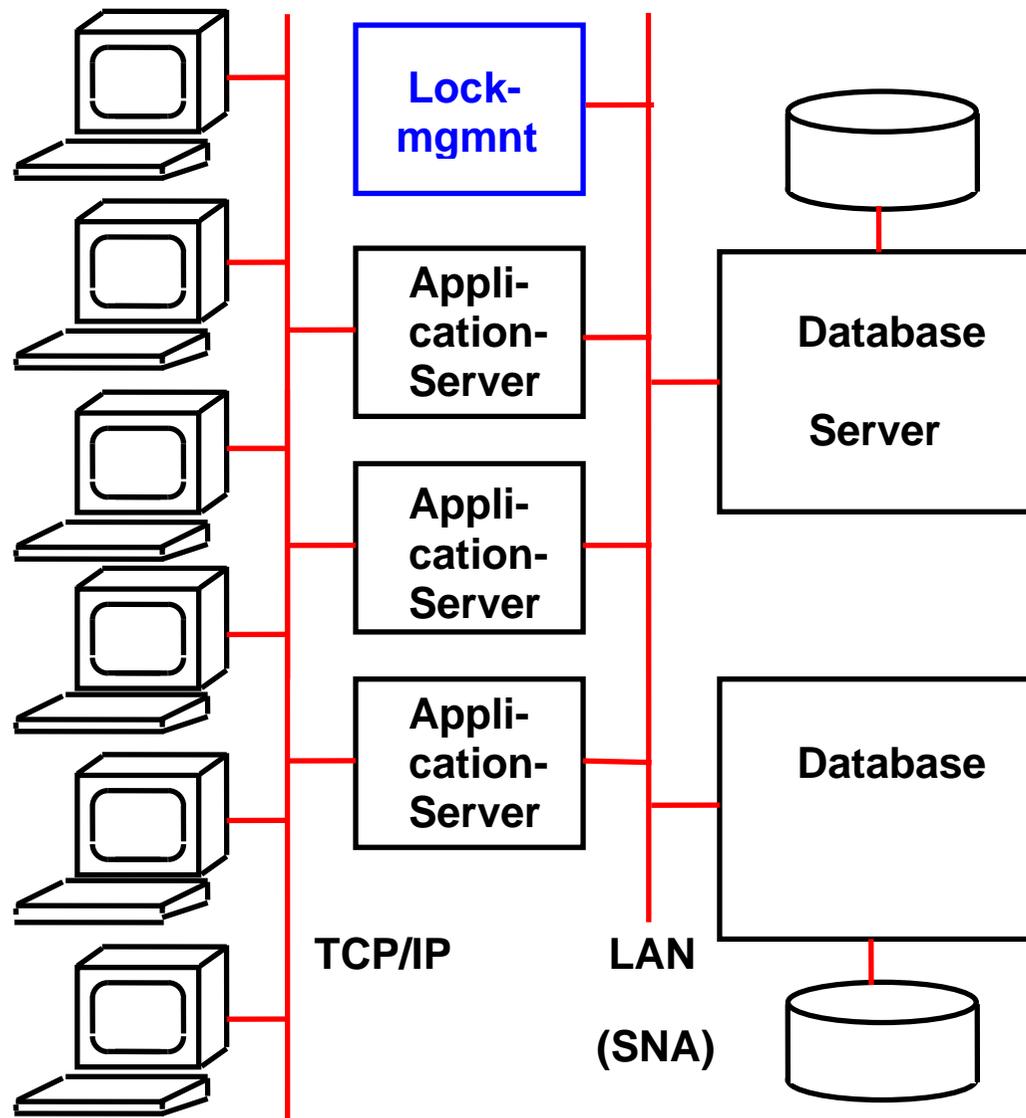
Beispiele für diese Lösung sind die VAX DBMS und VAX Rdb/VMS Datenbanksysteme.

Eine Faustregel besagt: **You cannot build a cluster that scales, if you do not solve the locking problem .**
 Jim Gray, Andreas Reuter, 1993



Invalidate-Broadcast Kohärenzsteuerung

Presentation



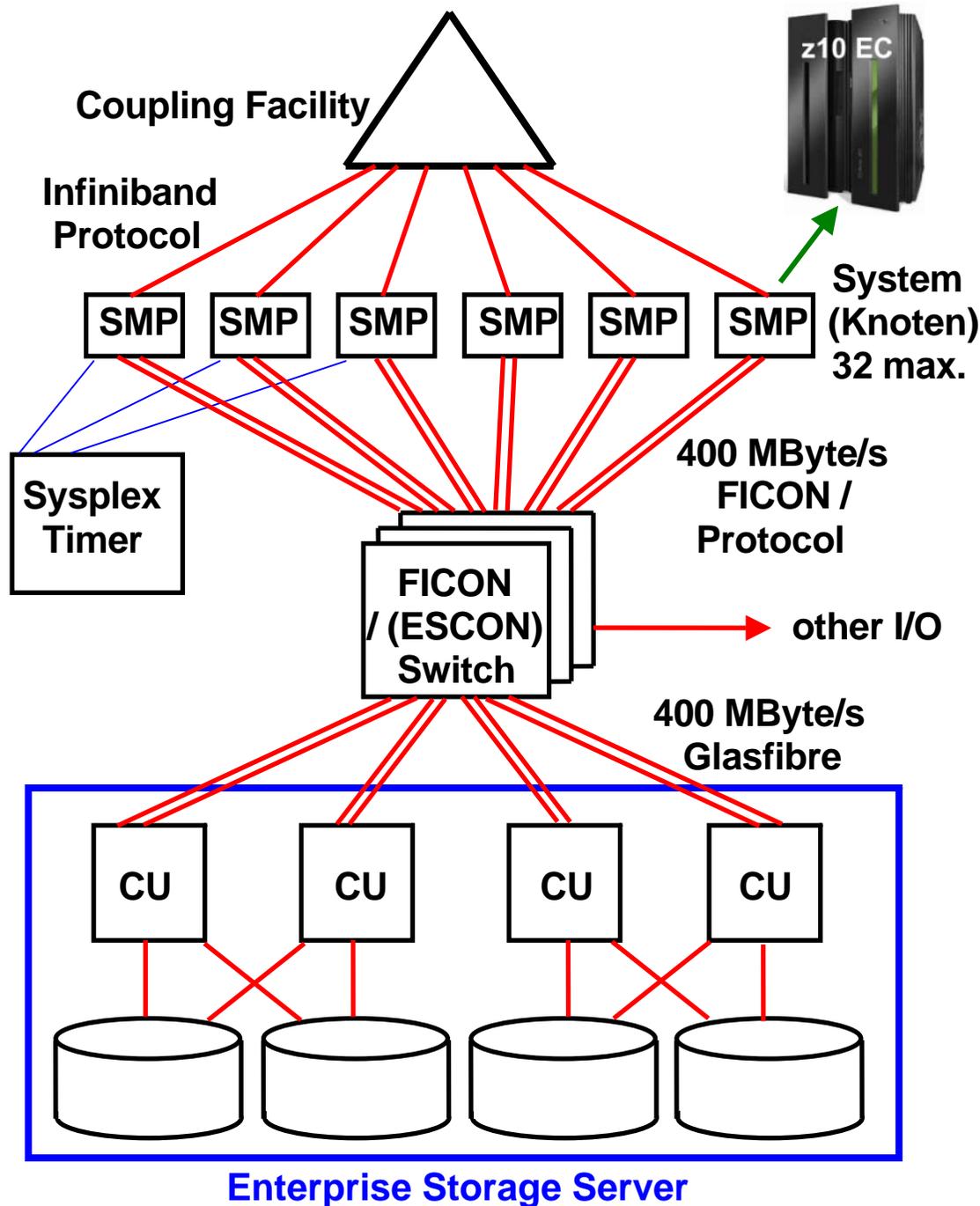
Dezentrale Locks arbeiten mit einer Invalidate-Broadcast Kohärenzsteuerung, die bei großen Transaktionsraten schlecht skaliert. Dieses Problem kann mit Hilfe eines zentralisierten Lockverwaltung und einem zentralen Sperrverwaltungsserver adressiert werden.

Eine typische SAP/R3 Konfiguration verwendet einen getrennten Server für das Lock management (Sperr-Server). Wenn ein Knoten ein S Lock erwerben möchte, wendet er sich an den Sperr-Server um zu erfragen, ob ein anderer Knoten ein Interesse an diesem Kock hat. Wenn ja, wird nur dieser benachrichtigt.

Nachteilig: Die Anfrage an den Sperr-Server geht über ein LAN mit dem entsprechenden TCP/IP Stack Overhead.

“You cannot scale a transaction processing system, if you do not solve the locking problem”
Jim Grav. Andreas Reuter. 1993

Typische SAP/R3 Konfiguration



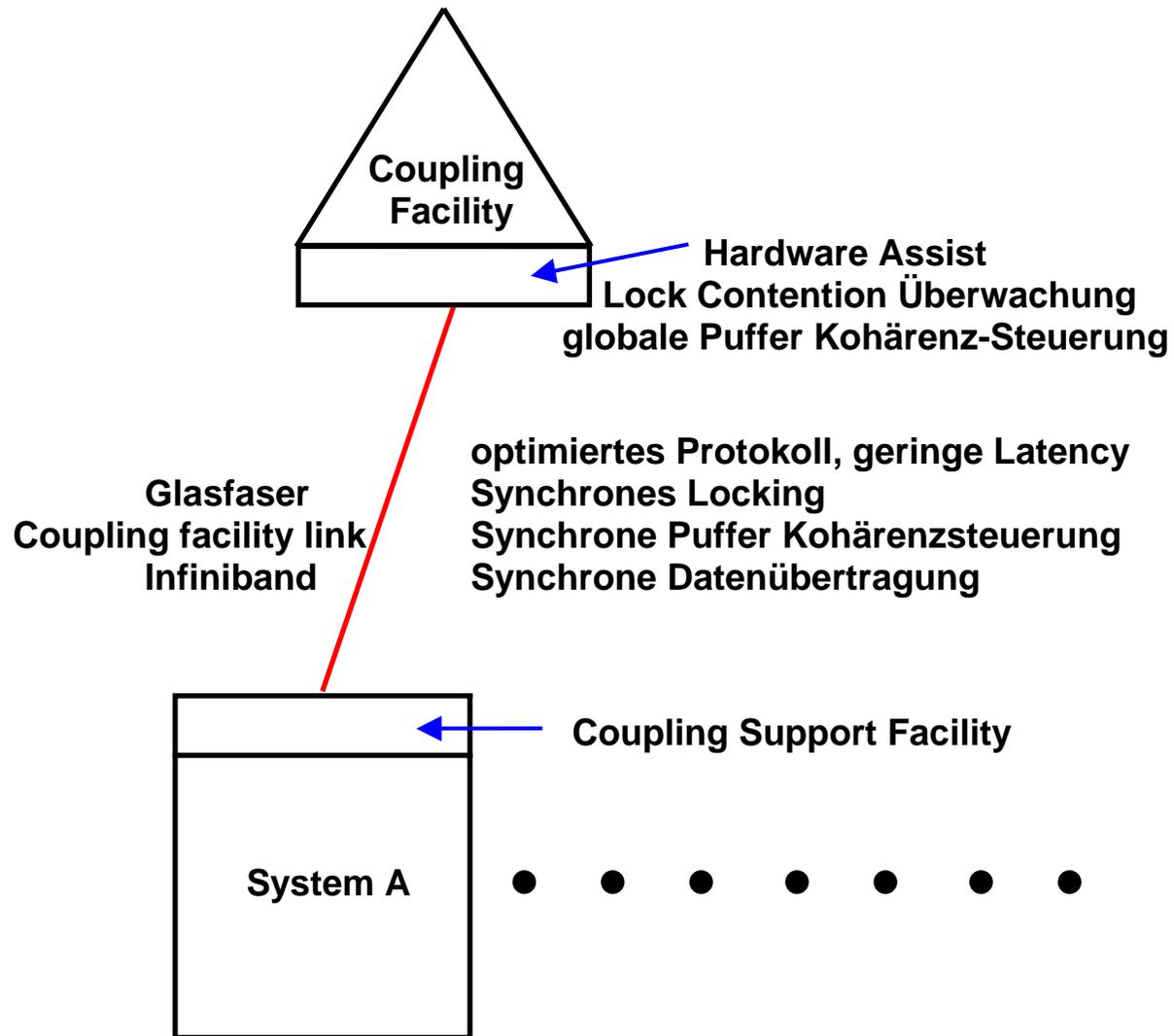
Sysplex with Coupling Facility

Ein Sysplex besteht aus bis zu 32 System z Rechnern, die über FICON Glasfasern und FICON Switches mit Plattenspeichern in der Form von einem oder mehreren Enterprise Storage Servern verbunden sind. Ein Enterprise Storage Server emuliert mehrere Control Units (CU).

Weiterhin sind 1 oder in der Regel 2 Coupling Facilities vorhanden.

Eine Coupling Facility ist ein regulärer System z Rechner, auf dem „Coupling Facility Code“ an Stelle eines Betriebssystems läuft.

Weiterhin sind 1 oder 2 Zeitgeber (Sysplex Timer) vorhanden

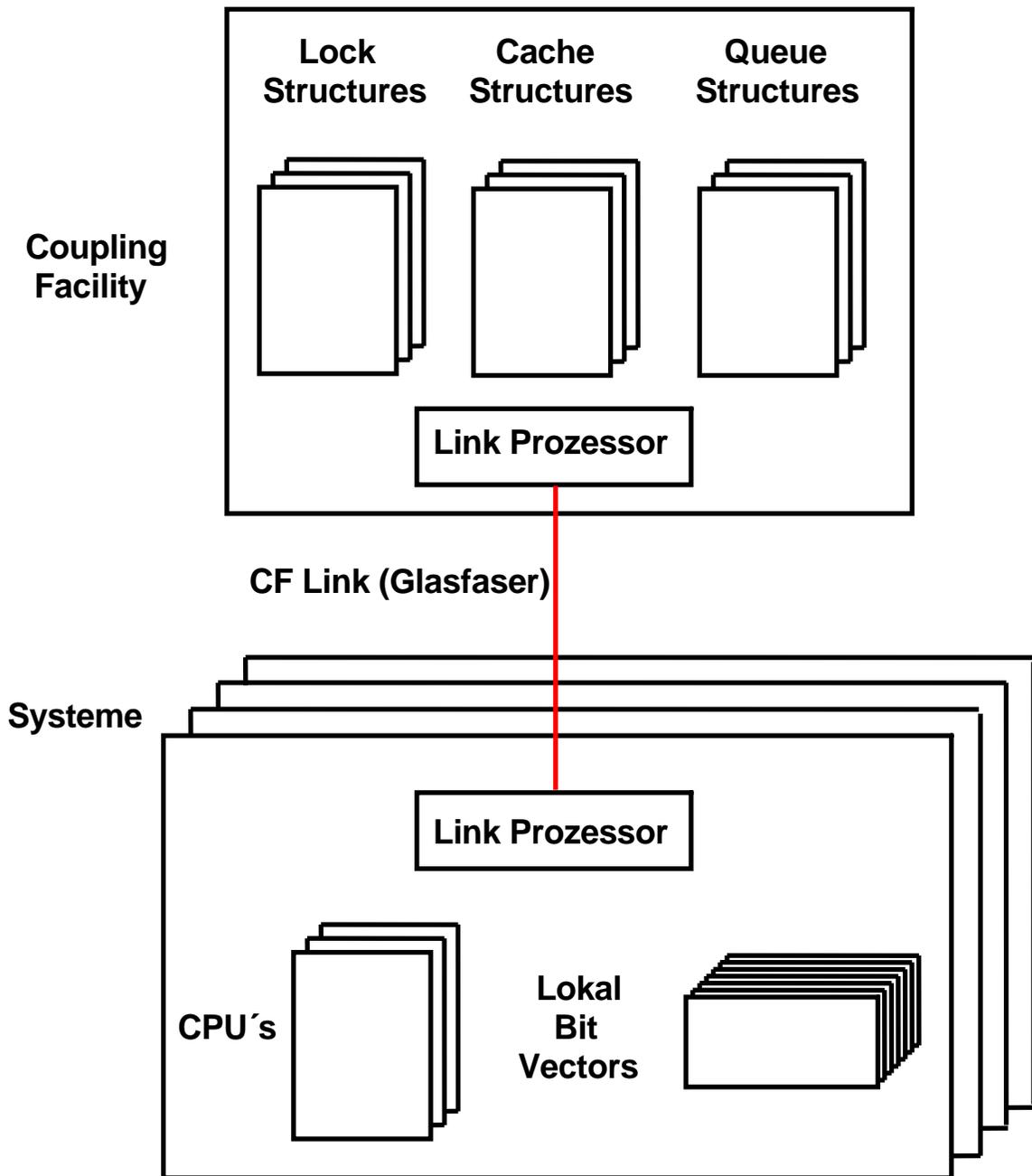


Die Coupling Facility ist durch eine Punkt-zu-Punkt Glasfaserleitung mit jedem System (Knoten) des Sysplex verbunden (Coupling facility link).

Es wird ein spezielles Verbindungsprotokoll mit besonders geringer Latency eingesetzt .

Die CF Glasfaser Verbindung wird durch spezielle Hardware Einrichtungen und durch zusätzliche Maschinenbefehle in jedem angeschlossenen System unterstützt. Die Coupling Facility kann in einem angeschlossenen Rechner ohne Unterbrechung des laufenden Prozesses Daten in spezielle Speicherbereiche (Bit Vektoren) abändern.

Anbindung eines Systems an die Coupling Facility



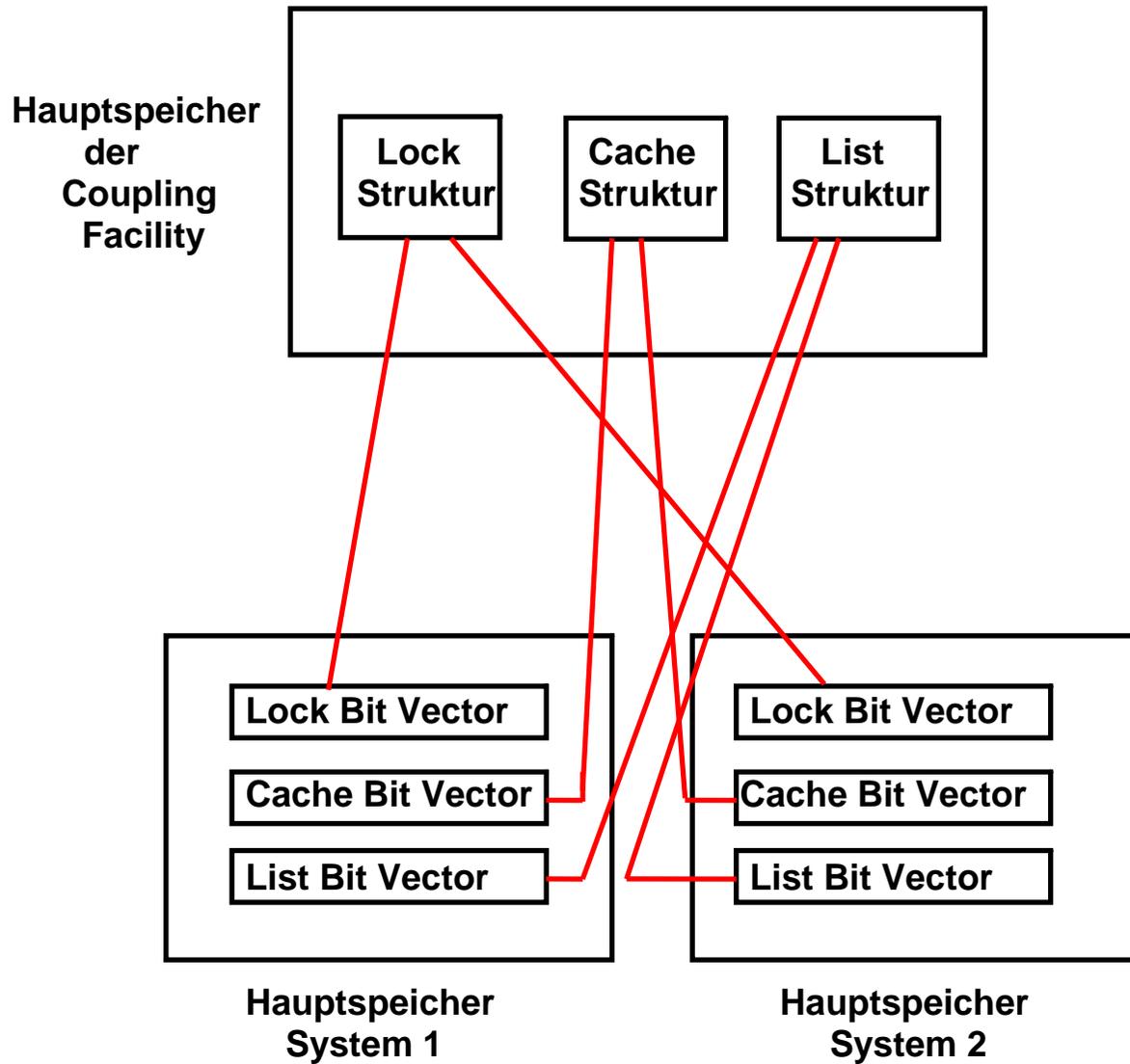
Die Coupling Facility unterhält in ihrem Hauptspeicher getrennte Strukturen für die Verwaltung von

- Locks (Sperrern)
- Sysplex weiter Daten Cache
- Listen für Sysplex-weite Aufgaben.

Die einzelnen Systeme des Sysplex sind mit jeder dieser Strukturen über das CF Link und die Link Prozessoren direkt verbunden.

Jedes System unterhält je 1 lokalen „Bit Vector“ für eine logische Verbindung zu jeder Struktur in der CF.

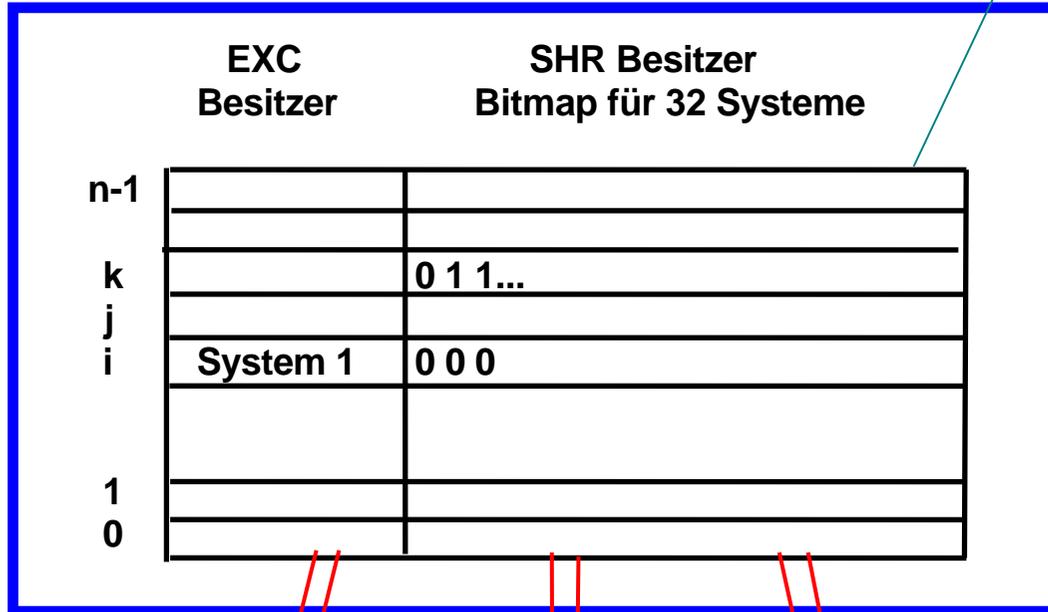
Für die Kommunikation CPU – CF sind spezifische Maschinenbefehle und zusätzliche Link Prozessor vorhanden.



Zuordnung von Bit Vektoren zu CF Strukturen

Coupling Facility

Lock Table (Lock Struktur)



lokaler
State

	i	j	k	
	E			

lokale
Queue

A, i, P1, E

System 1

	i	j	k	
			S	

C, k, P2, S

System 2

	i	j	k	
			S	

D, k, P4, S

System 3

1

Nutzung der CF Lock Struktur

Der Lock - Zustand eines Data Items kann 3 Werte haben:

Frei	0
Shared	S = SHR
Exclusive	E = EXC

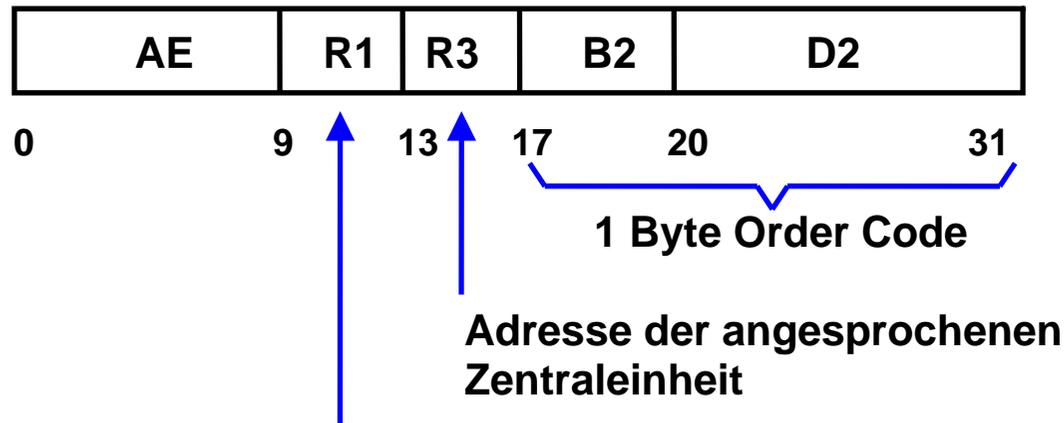
A Symbolischer Lock Name

I Lock Klasse

P1 Prozess 1

E exklusiver Besitz des Locks

Signal Processor Maschinenbefehl



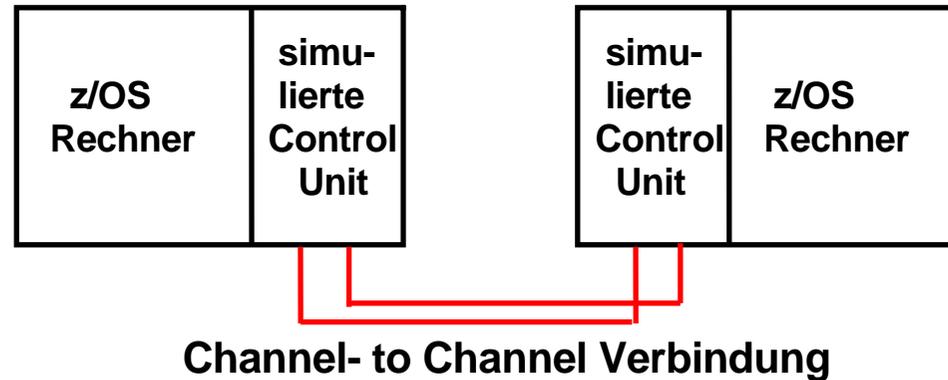
Nimmt Status Information auf

An eight-bit order code and, if called for, a 32-bit parameter are transmitted to the CPU designated by the CPU address contained in the third operand. The result is indicated by the condition code and may be detailed by status assembled in bit positions 32-63 of the first-operand location.

The second-operand address is not used to address data; instead, bits 56-63 of the address contain the eight-bit order code. Bits 0-55 of the second-operand address are ignored. The order code specifies the function to be performed by the addressed CPU.

The 16-bit binary number contained in bit positions 48-63 of general register R3 forms the CPU address. Bits 0-47 of the register are ignored.

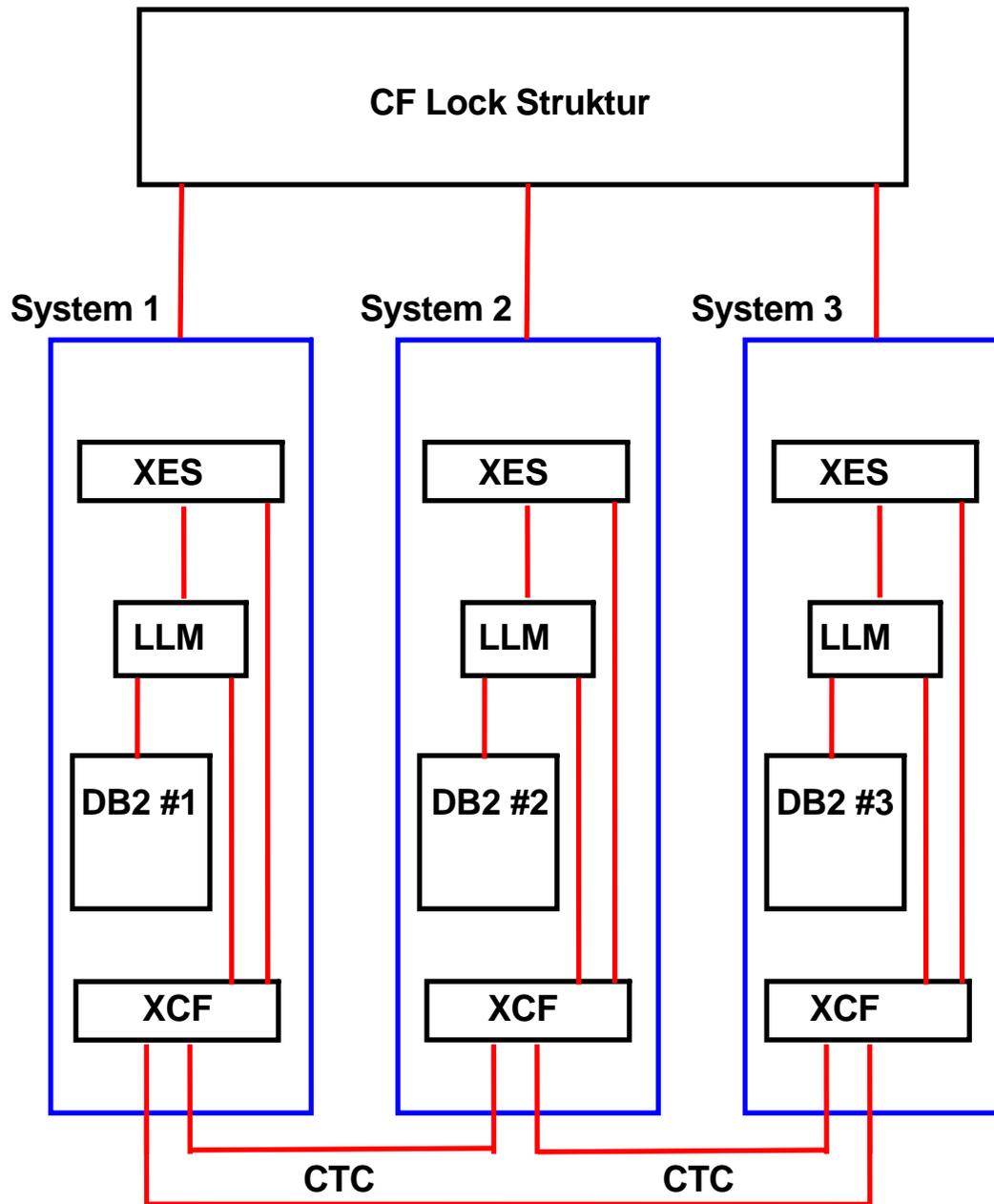
Channel- to Channel Verbindung (CTC) Cross-System Coupling Facility (XCF)



Die Channel- to Channel Verbindung wird durch eine Hardware Einrichtung eines zSeries Systems verwirklicht, die dieses System gegenüber einem anderen zSeries System wie eine E/A Einheit erscheinen lässt.

Für eine Full Duplex Verbindung werden normalerweise zwei CTC Verbindungen eingesetzt.

Die Cross-System Coupling Facility (XCF) ist eine Komponente des z/OS Betriebssystems. Sie verwendet das CTC Protokoll und die CTC Hardware. Sie stellt die Coupling Services bereit, mit denen z/OS Systeme innerhalb eines Sysplex miteinander kommunizieren.



Sysplex Communication

The Sysplex component of z/OS uses these facilities:

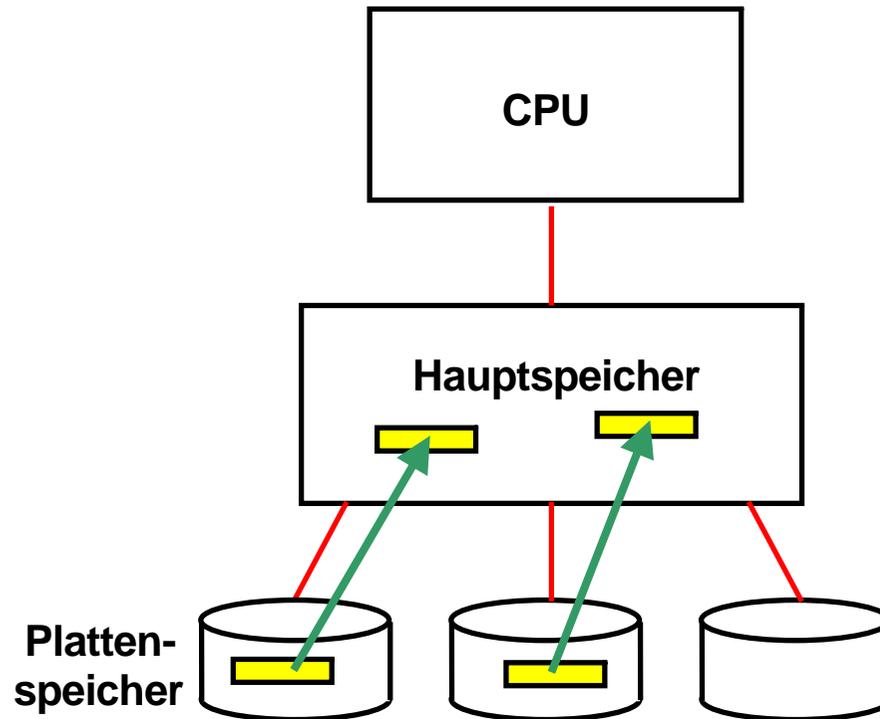
- CTC** Channel to Channel Verbindung
- XCF** Cross System Coupling Facility
- XES** Corss System extended Services
- LLM** Local Lock Manager, (Inter Resource Lock Manager, IRLM)

Cross-system extended services (XES) is the z/OS component that enables communications with the coupling facility. XES is an extension to the XCF component.

Coupling Facility Locking

Coupling Facility Caching

Coupling Facility Control/List Strukturen



Ein/Ausgabepuffer (I/O Buffer)

In der Vergangenheit hat ein Anwendungsprogramm jeweils einzelne Datensätze (Records) vom Plattenspeicher in einen Ein/Ausgabepuffer (I/O Buffer) im Hauptspeicher gelesen und dort verarbeitet. Zur Leistungssteigerung hat man bald mehrere logische Records zu einem physischen Record zusammengefasst. Mit etwas Glück findet das Anwendungsprogramm beim nächsten Zugriff die Daten bereits im Ein/Ausgabepuffer und ein Plattenspeicherzugriff erübrigt sich.

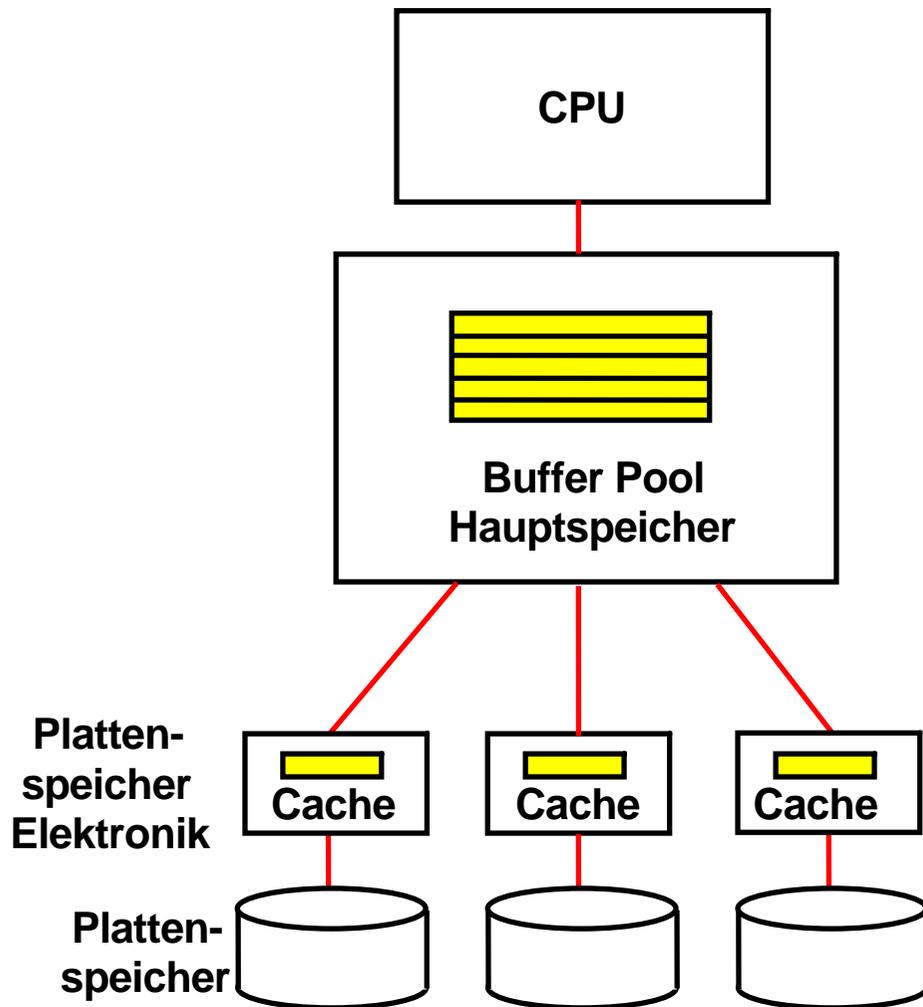
In der Regel arbeitet ein Anwendungsprogramm mit mehreren Dateien oder Datenbanken gleichzeitig, die alle ihren eigenen Ein/Ausgabepuffer benutzen. Die Menge der Ein/Ausgabepuffer wird als „Buffer Pool“ bezeichnet und vom Betriebssystem optimal verwaltet.

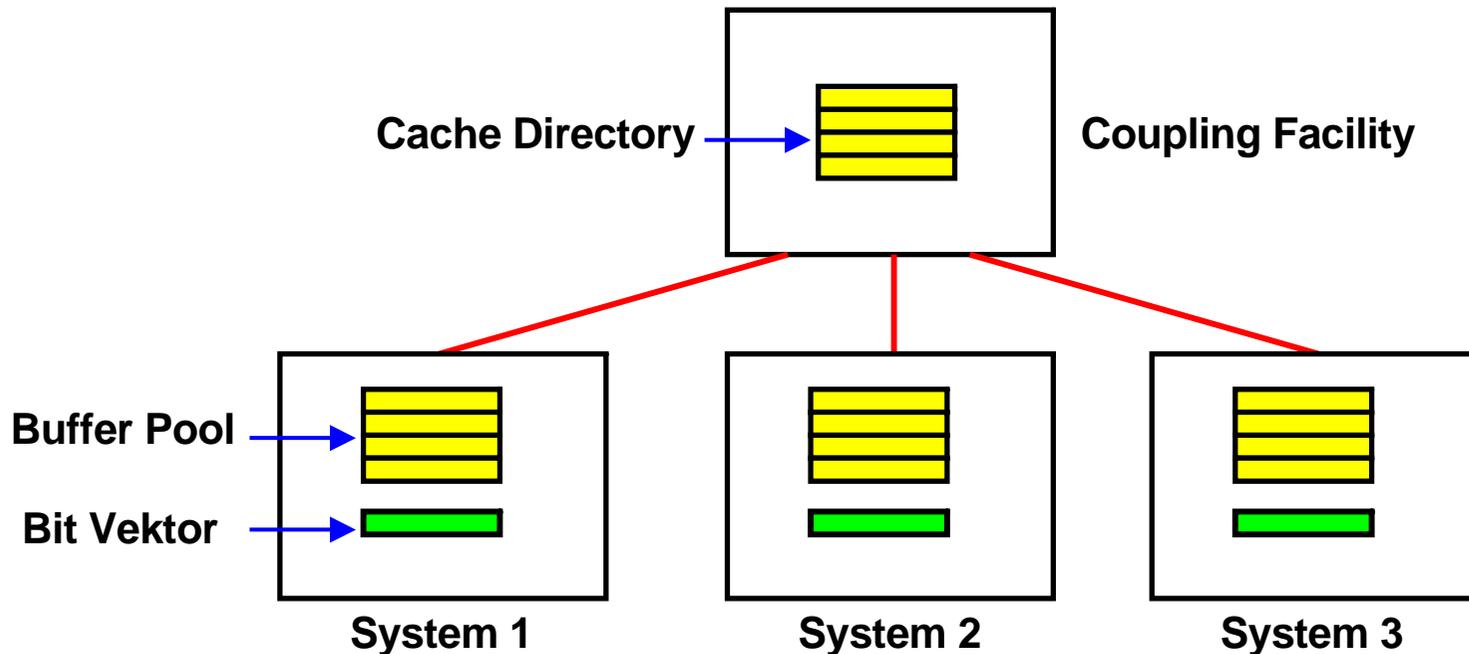
Plattenspeicher Cache und Hauptspeicher Buffer Pool

Der Buffer Pool stellt eine Art Plattenspeicher Cache im Hauptspeicher dar. Zusätzlich werden Daten in einem Plattenspeicher Cache gespeichert, der Bestandteil der Plattenspeicher Elektronik ist. Das Betriebssystem bemüht sich, den Speicherplatz im Buffer Pool und im Plattenspeicher Cache optimal zu verwalten.

Der Buffer Pool besteht aus einzelnen Puffern (Buffers), die Datenbankobjekte oder Teile einer Datei aufnehmen.

In einem Cluster ist nicht auszuschließen, dass Datensätze oder Datenbankrecords gleichzeitig in den Buffer Pools mehrerer Knoten (Systeme) abgespeichert werden.





Cache Directory in der Coupling Facility

Der Buffer Pool in jedem System enthält Blöcke (Buffer), die möglicherweise gerade bearbeitet werden. Es kann sein, dass sich in zwei unterschiedlichen Systemen Buffer mit den gleichen Datenbankrecords befinden.

Die Coupling Facility unterhält ein "Cache Directory", in dem sich jeweils ein Eintrag für jeden Buffer in den angeschlossenen Systemen befinden. Analog zur Lock Verwaltung befinden sich außerdem in jedem System Bit Vektoren, die den Inhalt des Cache Directories teilweise replizieren.

Bei einer Änderung eines Eintrags im Cache Directory erfolgt ein automatisches Update der Bit Vektoren in allen angeschlossenen Systemen.

Coupling Facility Cache Directory

Der lokale Buffer Pool im System 1 enthält Puffer (Blöcke) mit Records, die gerade bearbeitet werden. Solange die Transaktion nicht abgeschlossen ist, verhindert der Lock Manager einen Zugriff durch ein anderes System (z.B. System 2).

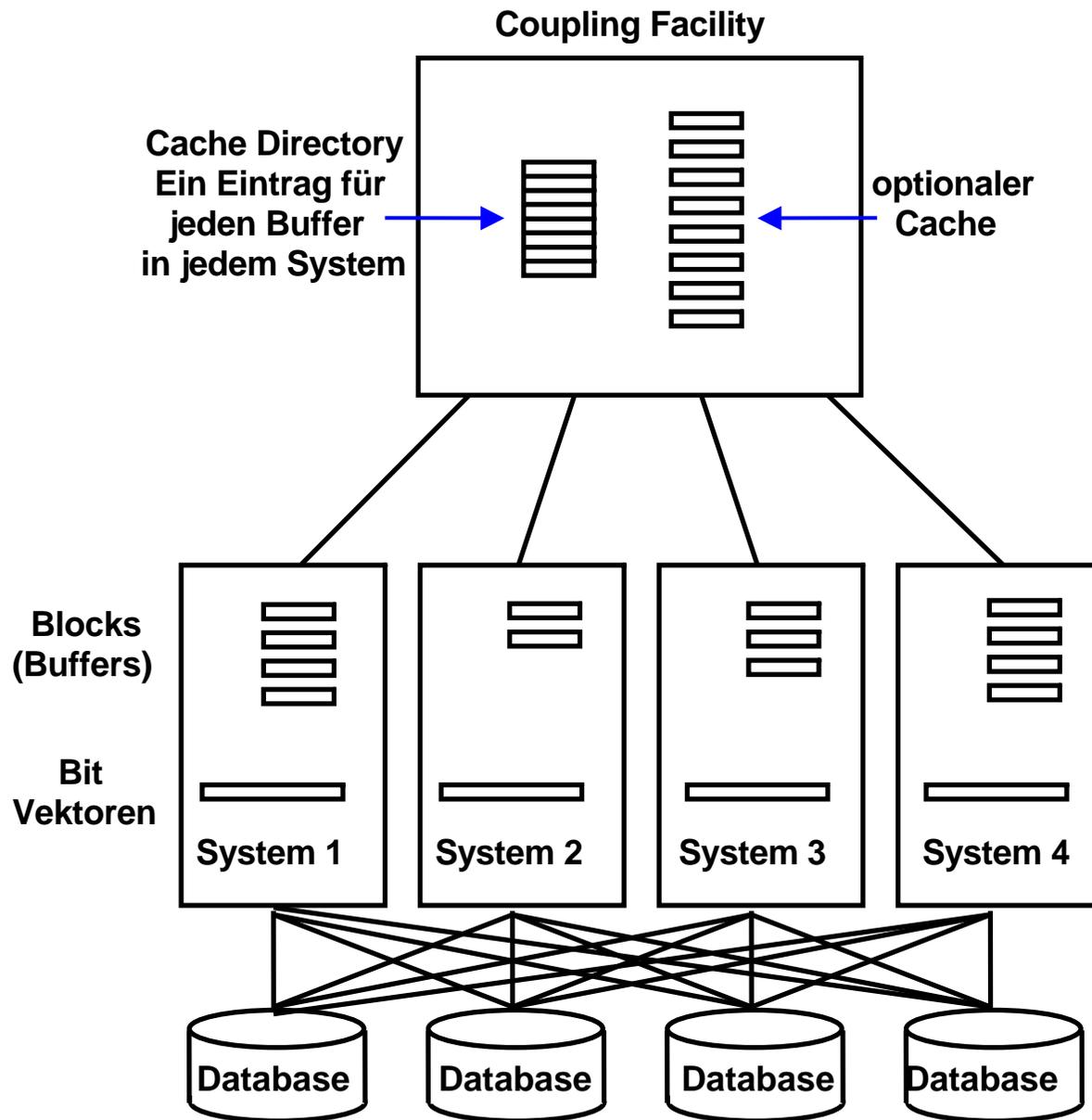
Wenn die Transaktion abgeschlossen ist (commit), werden die Locks freigegeben. Die Puffer bleiben in System 1 - evtl. werden sie demnächst wieder gebraucht.

Greift System 2 jetzt auf einen Buffer mit dem gleichen Datenbank Record zu, entsteht ein Kohärenzproblem. Die beiden Buffer in den Systemen 1 und 2 haben nicht den gleichen Inhalt.

Lösung: „Force-at-Commit“ . Bei Transaktionsabschluss erfolgt ein update des Cache Directories durch System 1.

Die CF sendet eine „Cross-Invalidate (CI) Nachricht an alle anderen Systeme.

Die CI Nachricht ändert den lokalen State Vector innerhalb des Hauptspeichers eines jeden Systems ab. Dies geschieht durch den Link Prozessor und verursacht keine CPU Unterbrechung !



Aller Datentransfer in 4 KByte Blöcken.

Das Cache Directory in der Coupling Facility enthält einen Eintrag für jeden Block (Buffer), der Teil eines Buffer Pools in einem der beteiligten Systeme ist.

a) System 1 Read from Disk

1. Load Block from Disk
2. Register with CF Directory
3. add Bit in Bit Vector

b) System 2 Read from Disk

1. Load Block from Disk
2. Register with CF Directory
3. add Bit in Bit Vector

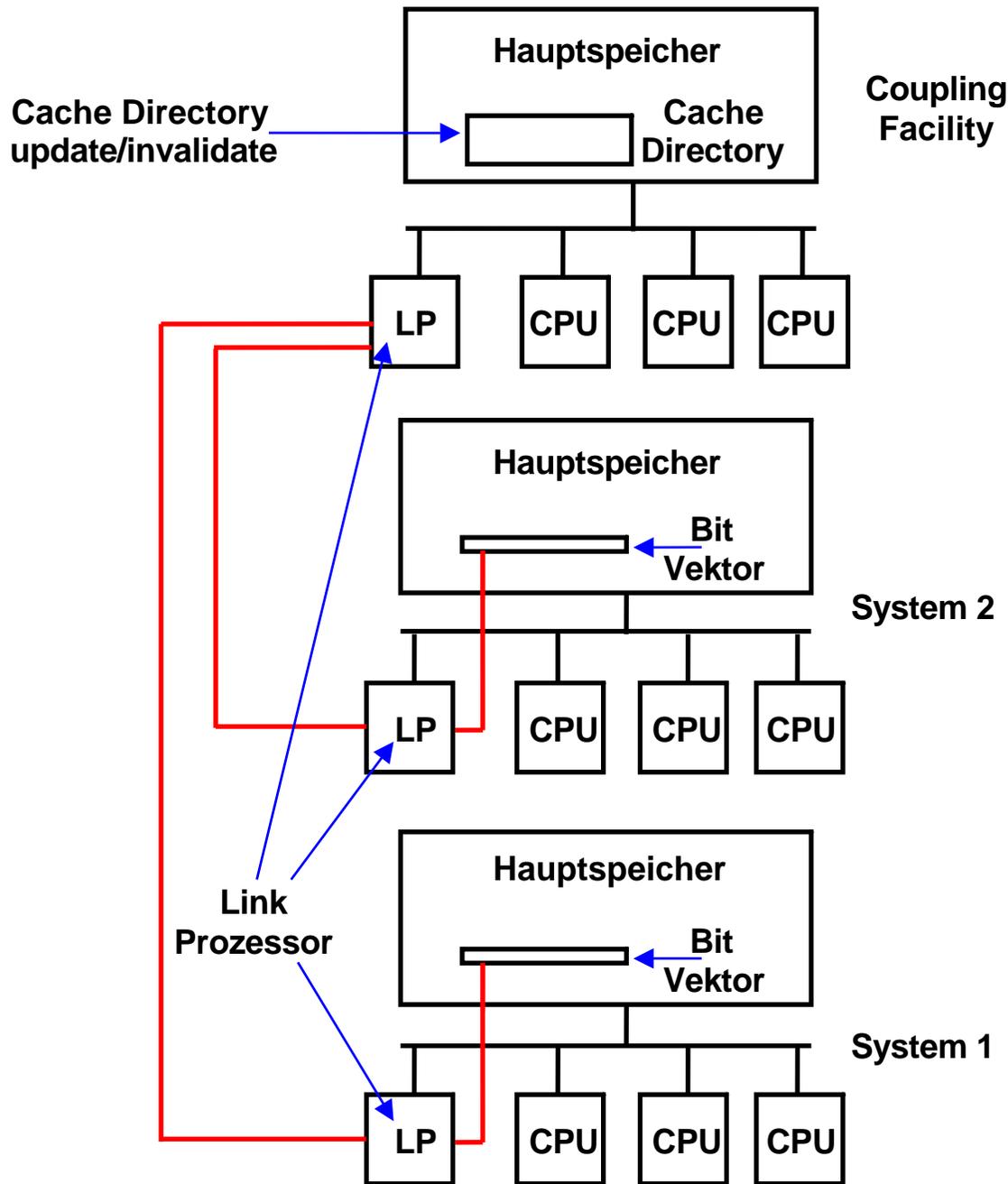
einige Zeit später

c) System 1 Intend to write (to local Buffer)

1. Register with CF
2. CF invalidates all Bit Vectors
3. Write to local Buffer

d) System 2 Read from Buffer

1. Read
2. detect invalid Bit in local Bit Vector
3.



Das Cache Directory in der Coupling Facility enthält einen Eintrag für jeden Block (Buffer), der Teil eines Buffer Pools in einem der beteiligten Systeme ist.

Bei Transaktionsabschluss erfolgt ein update des Cache Directories durch System 1 (Force-at-Commit).

Hierzu sendet die CF eine „Cross-Invalidate (CI) Nachricht an alle anderen Systeme.

Link Prozessoren haben einen Direct Memory Access auf den Hauptspeicher. Über die Link Prozessoren der Coupling Facility und die der Systeme können Bit Vektoren im Hauptspeicher abgeändert werden, ohne daß der normale Programmablauf dadurch beeinflußt wird (kein Prozesswechsel). Dies verbessert das Leistungsverhalten, da jeder Prozesswechsel eine Pfadlänge von mehreren Tausend Maschinenbefehlen erfordert.

Coupling Facility Cache.

Neben dem Cache Directory kann die Coupling Facility auch als Plattenspeicher-Cache genutzt werden. Hiervon machen einige, aber nicht alle Datenbanksysteme Gebrauch. DB2 und Adabas nutzen die Coupling Facility auch als Plattenspeicher-Cache.

DB2, IMS und Adabas sind die wichtigsten unter z/OS eingesetzten Datenbanksysteme.

Coupling Facility locking

Coupling Facility Caching

Coupling Facility Control/List Strukturen

CF Control/List Strukturen

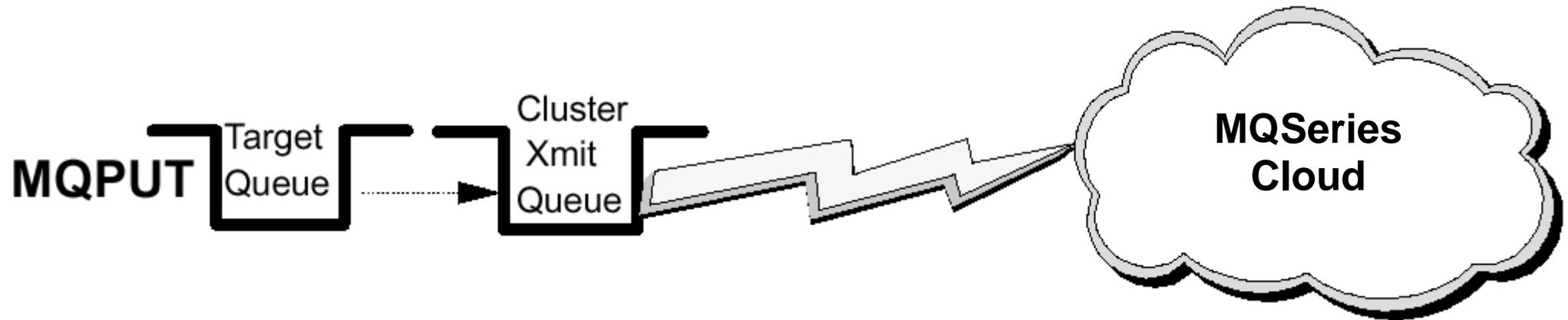
Neben dem Lock und dem Cache Management enthält die Coupling Facility Listen/Queue Strukturen, die vor allem für eine zentrale Verwaltung aller angeschlossenen Systeme eingesetzt werden.

Beispiele hierfür sind:

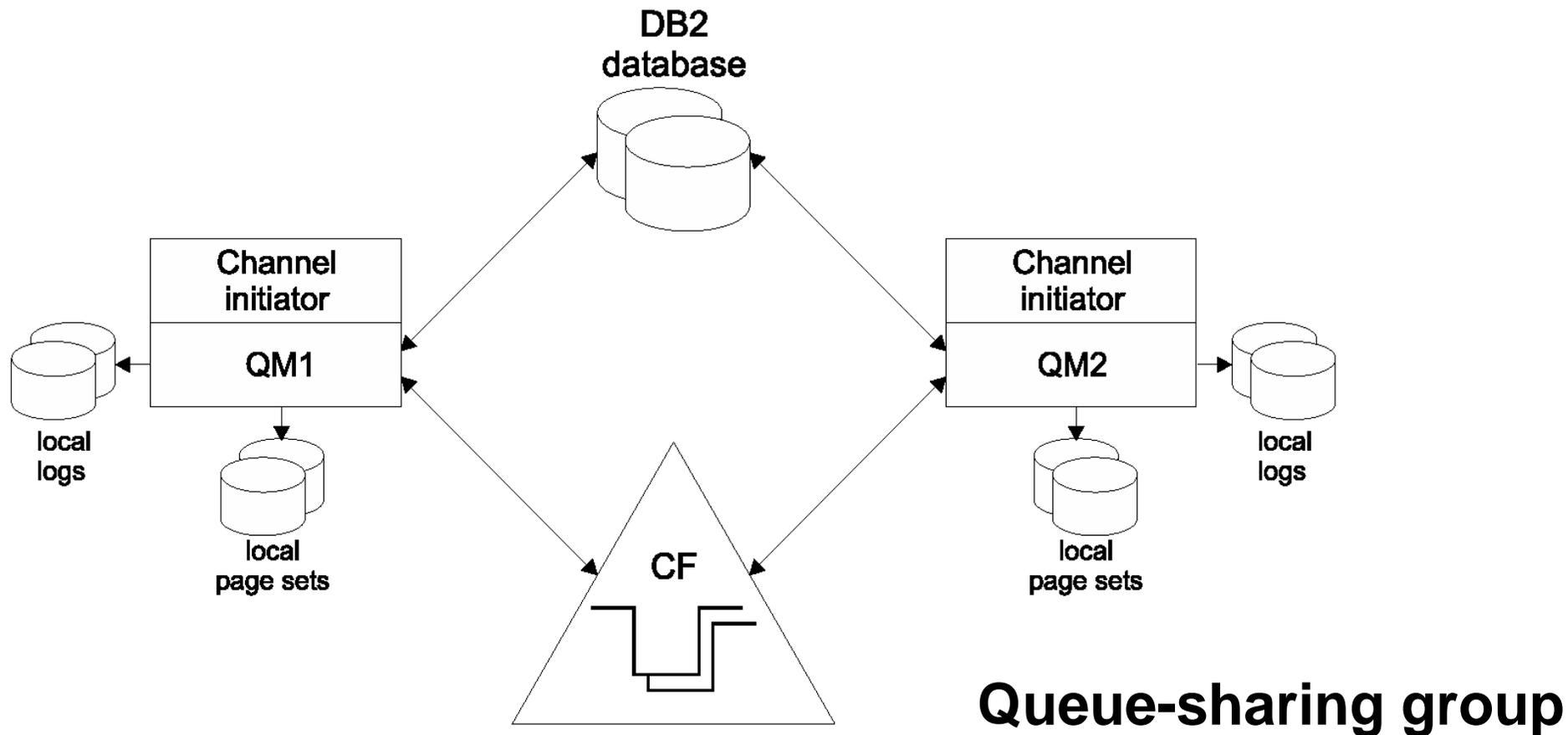
- **Clusterweite RACF Steuerung.** Ein Sysplex Cluster besteht aus mehreren z/OS Instanzen. Im einfachsten Fall müsste sich ein Benutzer mit getrennten Passworten in jedes System einzeln einloggen. „**Single Sign On**“ ist eine Einrichtung, mit der der Benutzer mit einem einzigen Sign-On Zugriffsrechte auf alle Ressourcen eines Sysplex erhält. Die entsprechenden RACF Benutzerprofile werden von der Coupling Facility zentral in einer Control/List Struktur verwaltet.
- **Work Load Management (WLM) Instanzen tauschen periodisch Status Information aus um Transaktionen dynamisch an unterbelastete Systeme weiter zu reichen**

Für einen Zugriff auf die Control/List Strukturen bestehen drei Möglichkeiten:

- LIFO Queue
- FIFO Queue
- Key Sequenced



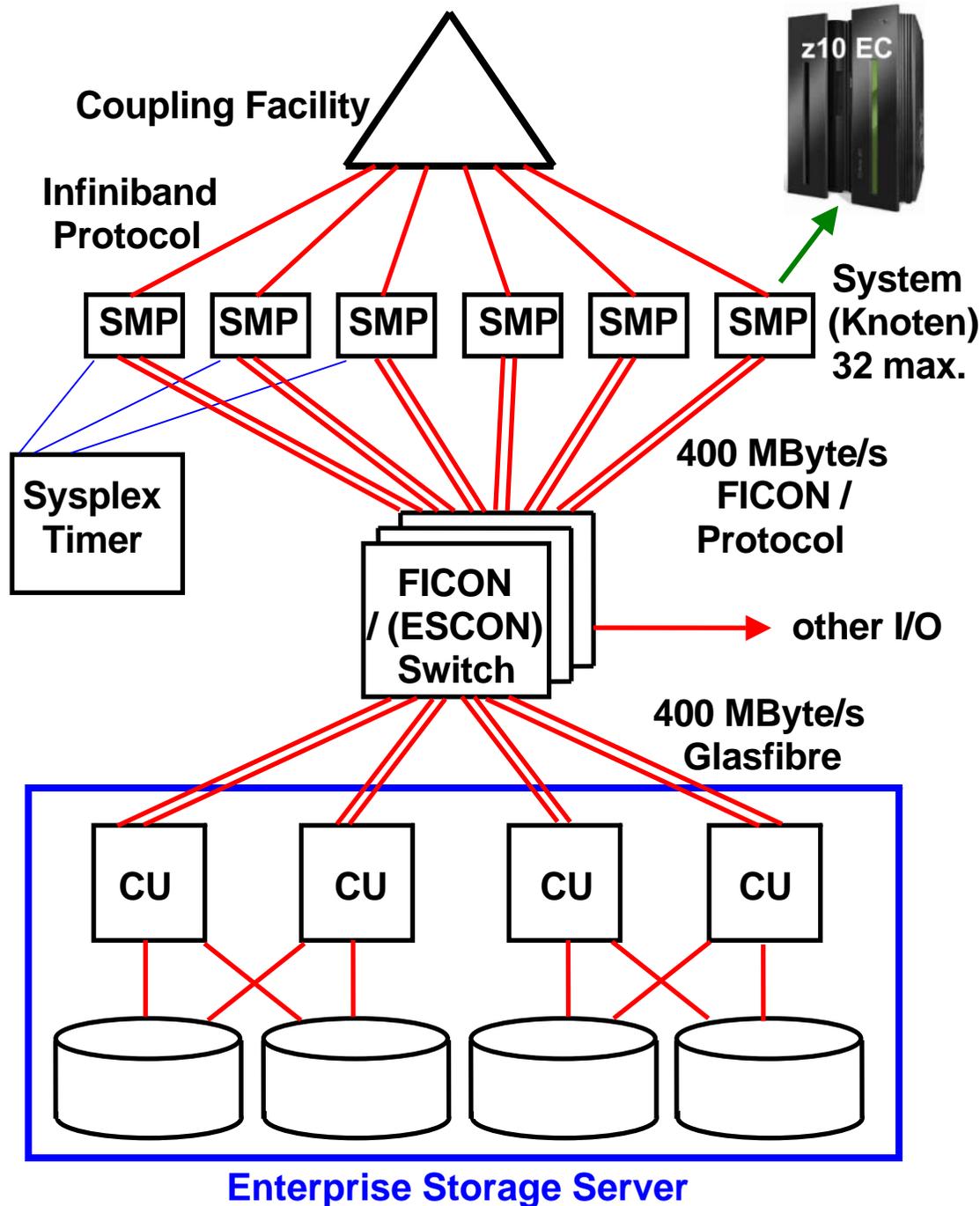
One example of using the Coupling Facility Control/List structure is MQSeries Cloud processing, leaving it up to the z/OS Work Load Manager to access one of several Channel Initiators in an MQSeries Cloud.



The group of queue managers that can access the same shared queues is called a queue-sharing group. Each member of the queue-sharing group has access to the same set of shared queues.

Queue-sharing groups have a name of up to four characters. The name must be unique in your network, and must be different from any queue manager names.

The figure above illustrates a queue-sharing group that contains two queue managers. Each queue manager has a channel initiator and its own local page sets and log data sets. A coupling facility list structure manages the queues



Sysplex with Coupling Facility

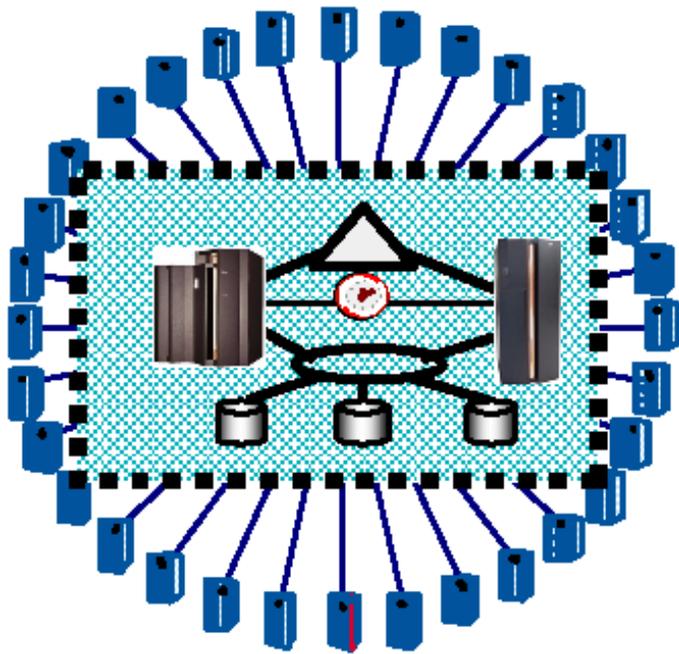
Ein Sysplex besteht aus bis zu 32 System z Rechnern, die über FICON Glasfasern und FICON Switche mit Plattenspeichern in der Form von einem oder mehreren Enterprise Storage Servern verbunden sind. Ein Enterprise Storage Server emuliert mehrere Control Units (CU).

Weiterhin sind 1 oder in der Regel 2 Coupling Facilities vorhanden.

Eine Coupling Facility ist ein regulärer System z Rechner, auf dem „Coupling Facility Code“ an Stelle eines Betriebssystems läuft.

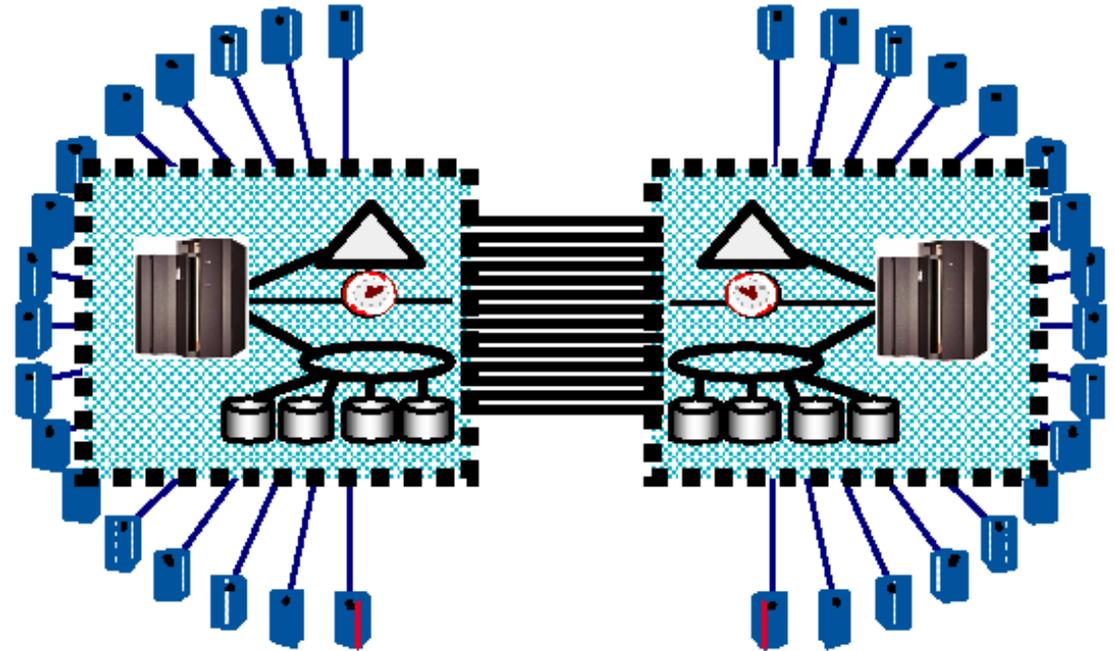
Weiterhin sind 1 oder 2 Zeitgeber (Sysplex Timer) vorhanden

Parallel Sysplex



1 to 32 Systems

GDPS



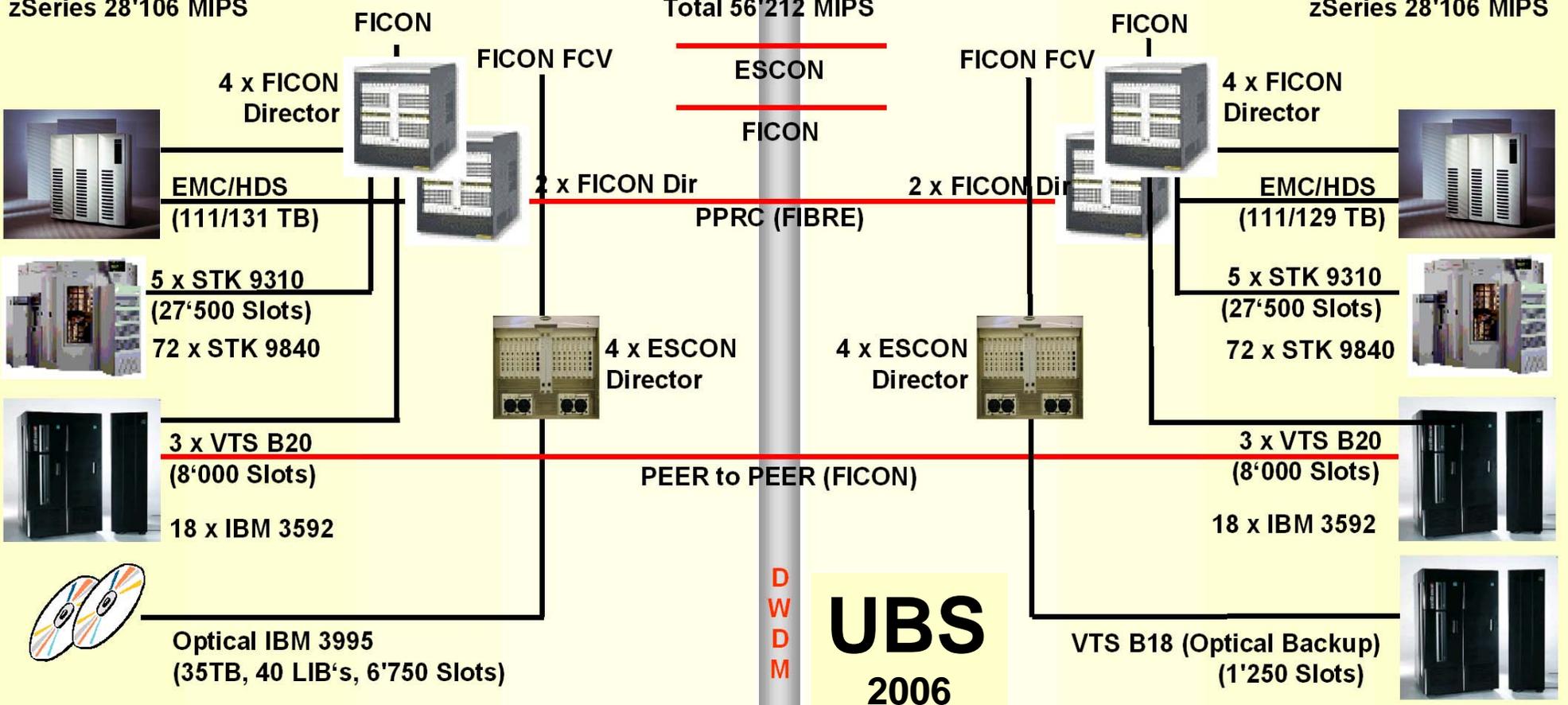
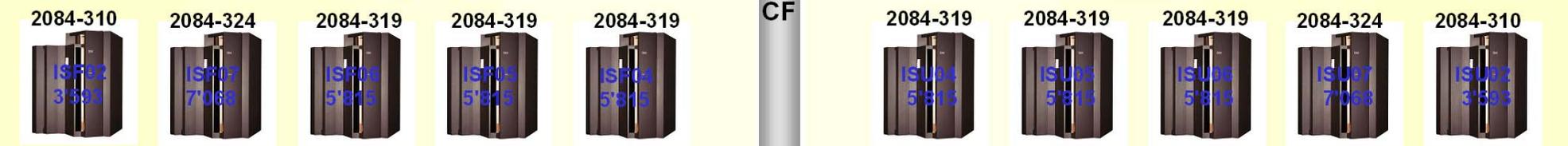
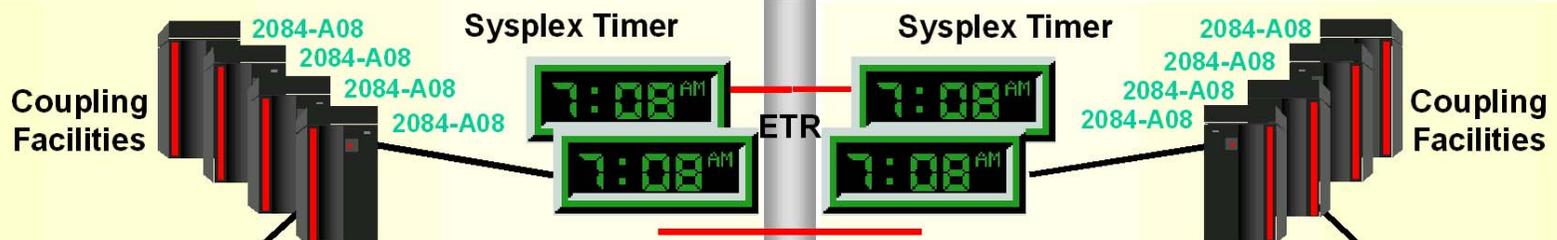
Site 1

Site 2

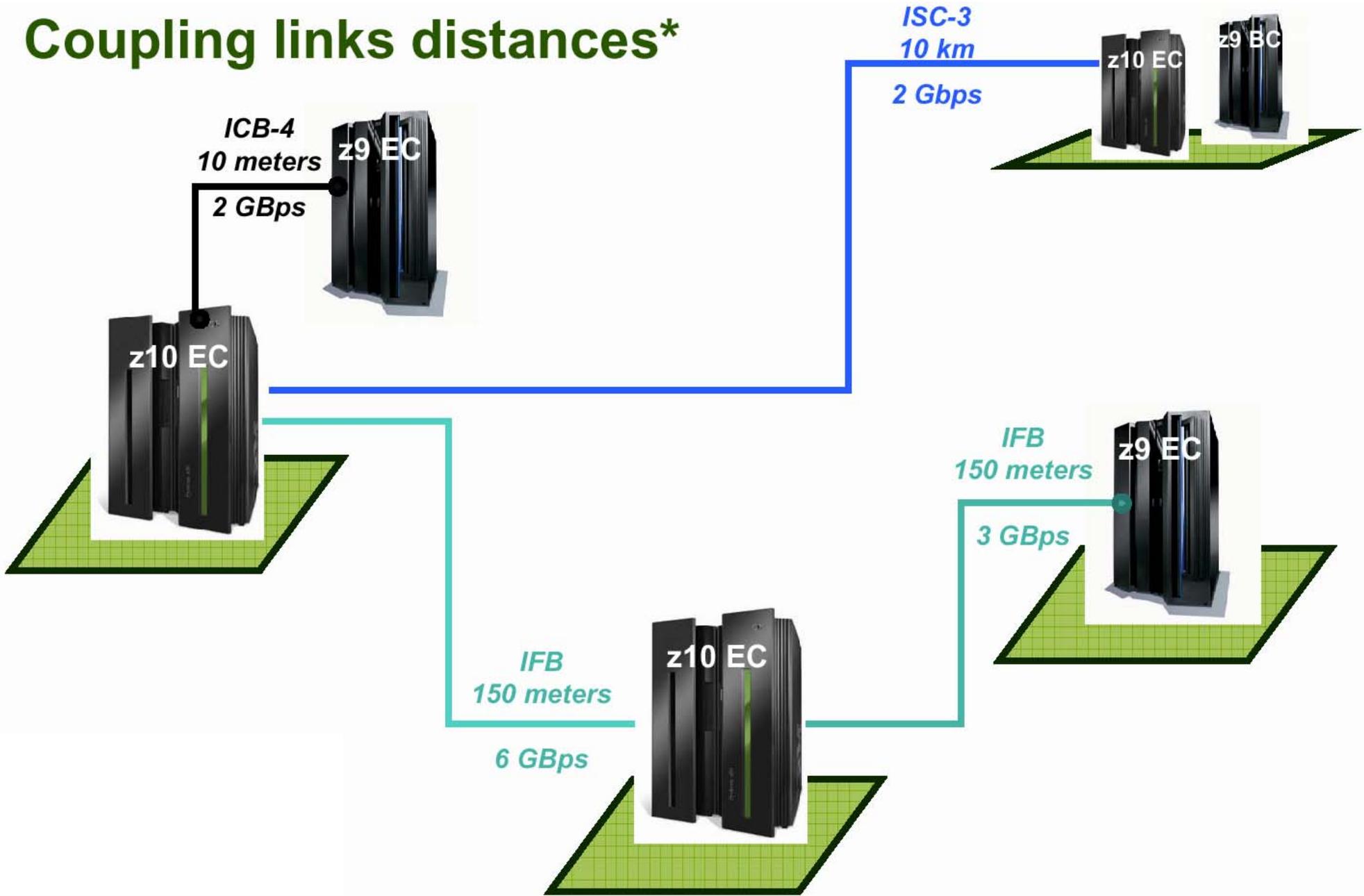
A **Geographically Dispersed Parallel Sysplex (GDPS)** is the ultimate Disaster Recovery and Continuous Availability solution for a System z multi-site enterprise

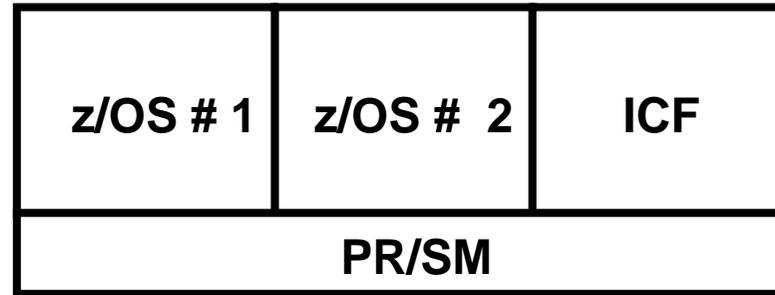
**Site 1
FLUR**

**Site 2
Urdorf**



Coupling links distances*





Integrierte Coupling Facility

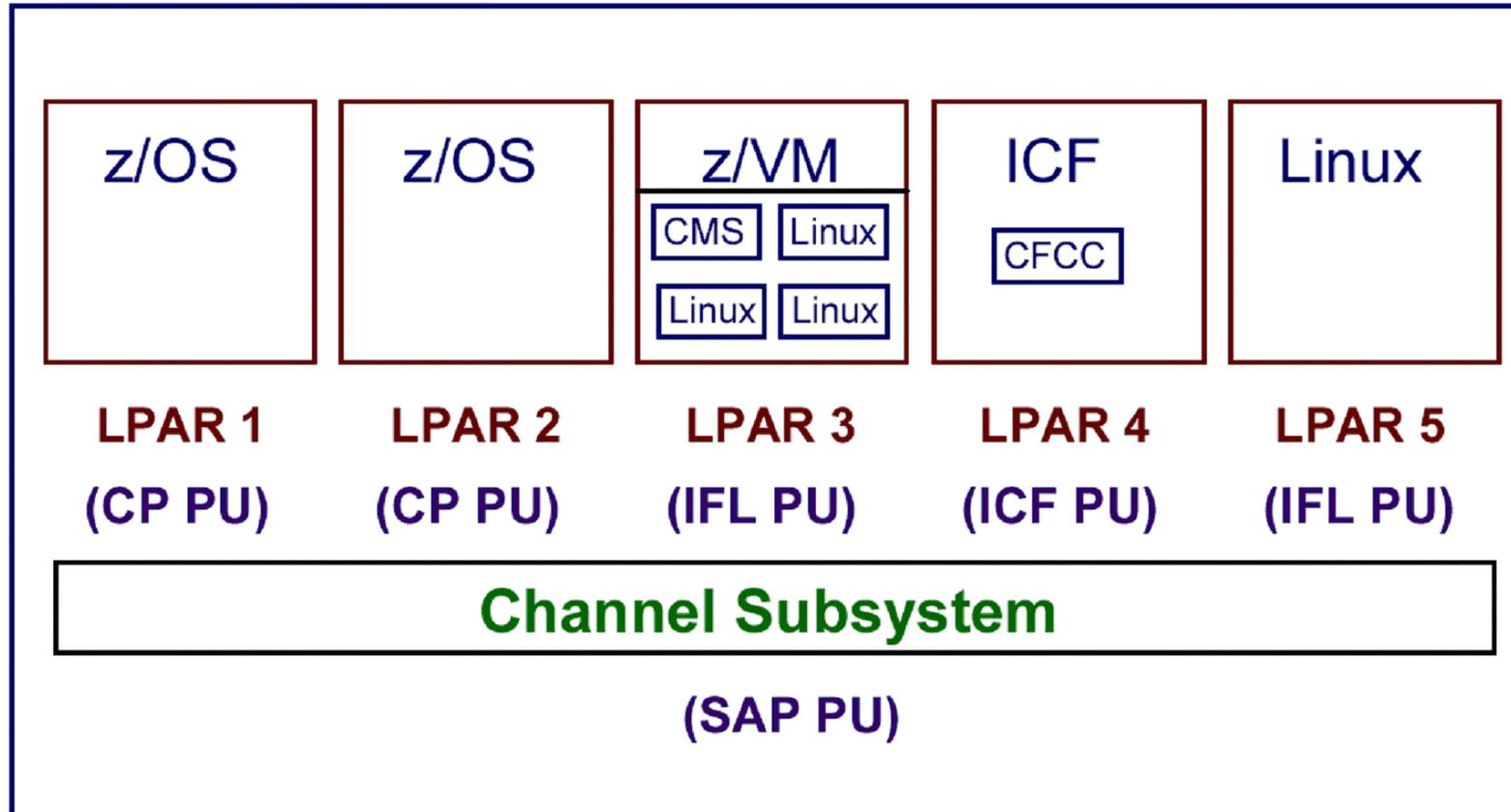
ICF

Die Coupling Facility (CF) kann statt in einem getrennten Rechner auch in einer Logischen Partition (LPAR) eines zSeries Knotens (System) untergebracht werden, wobei in anderen Partitionen zwei oder mehr z/OS Images laufen. Diese Art der CF wird als Integrated Coupling Facility (ICF) bezeichnet. Die ICF läuft auf einer oder mehreren hierfür dedizierten CPUs eines SMP. ICFs haben in den letzten Jahren zunehmend an Popularität gewonnen.

Die ICF stellt CF Funktionalität ohne Coupling Links zu Verfügung. Letztere werden durch PR/SM emuliert.

Aus Zuverlässigkeitsgründen sollten immer mindestens 2 CFs vorhanden sein.

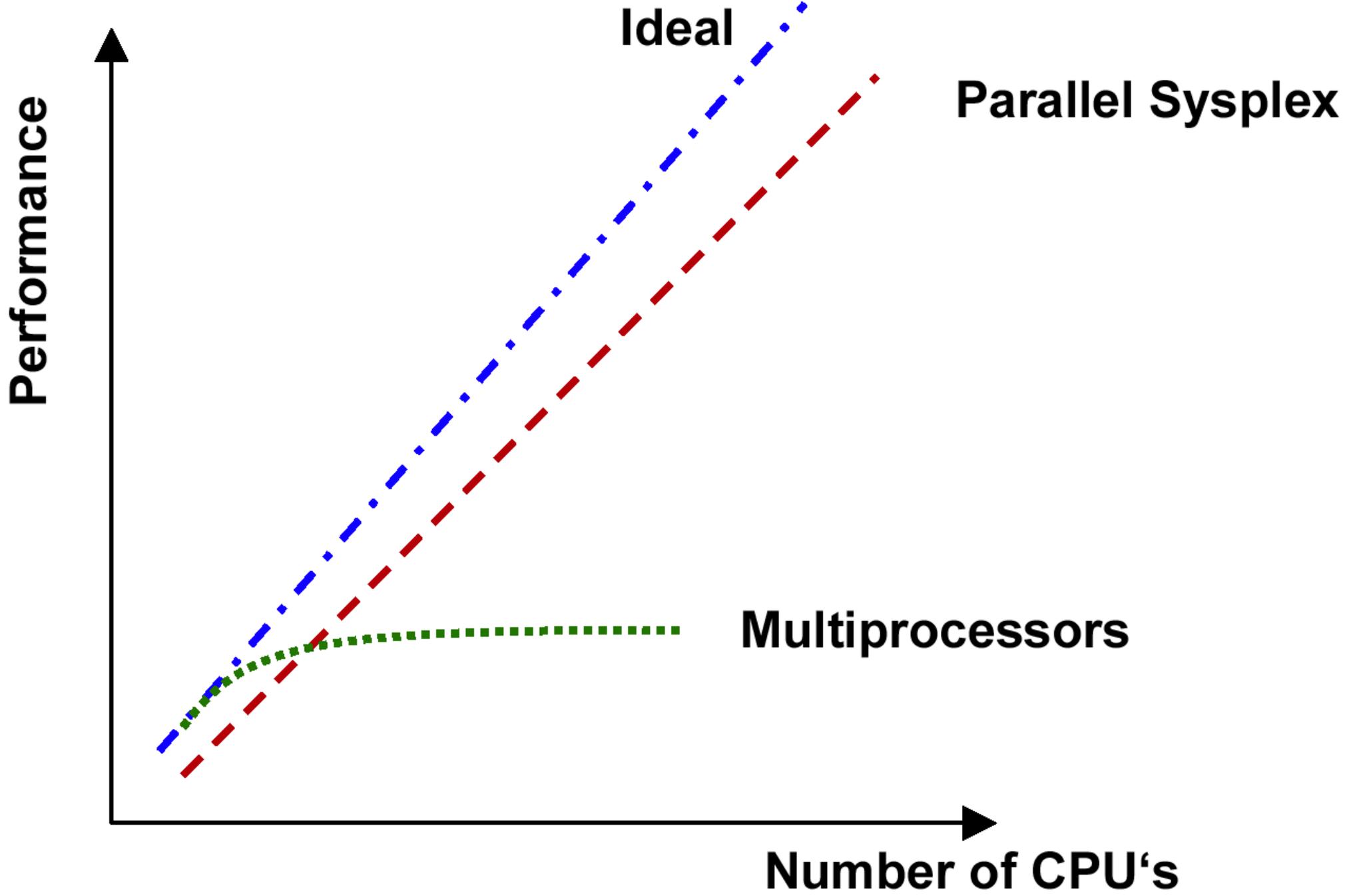
zSeries Server



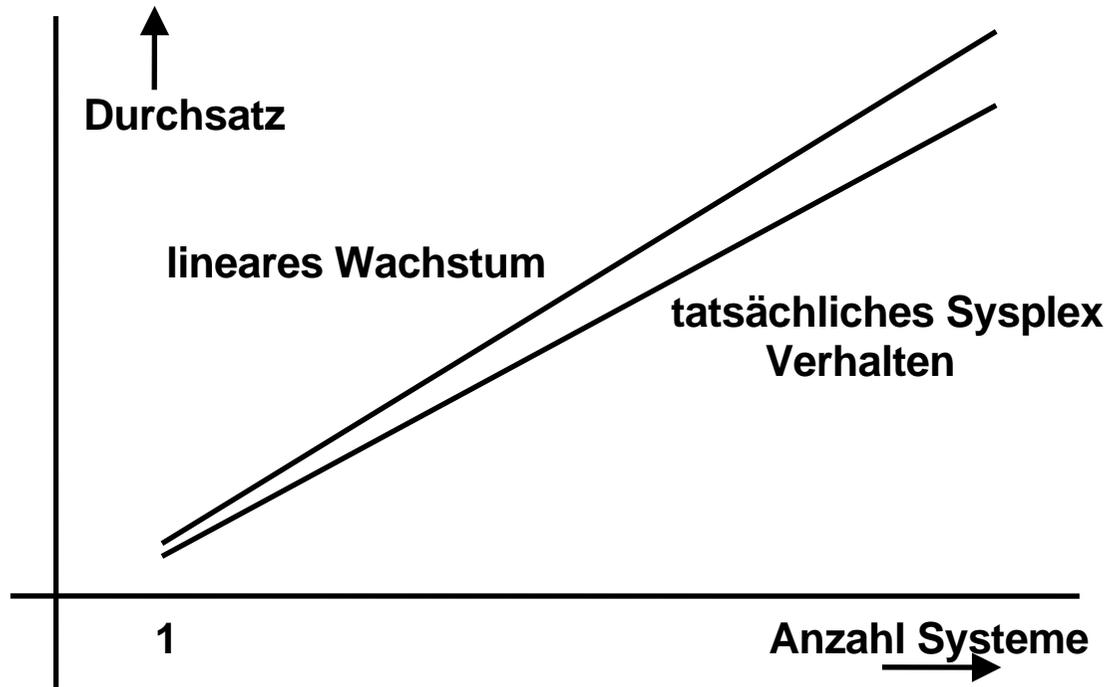
Nutzung der Coupling Facility

Die Coupling Facility ist eng mit dem z/OS Betriebssystem verknüpft. Andere System z Betriebssysteme werden nicht unterstützt.

Gezeigt ist eine Konfiguration mit 5 LPARs: 2 x z/OS, je 1 x ICF, zVM und zLinux. Nur die beiden z/OS Betriebssysteme nutzen die integrierte Coupling Facility; zVM und zLinux bleiben außen vor.

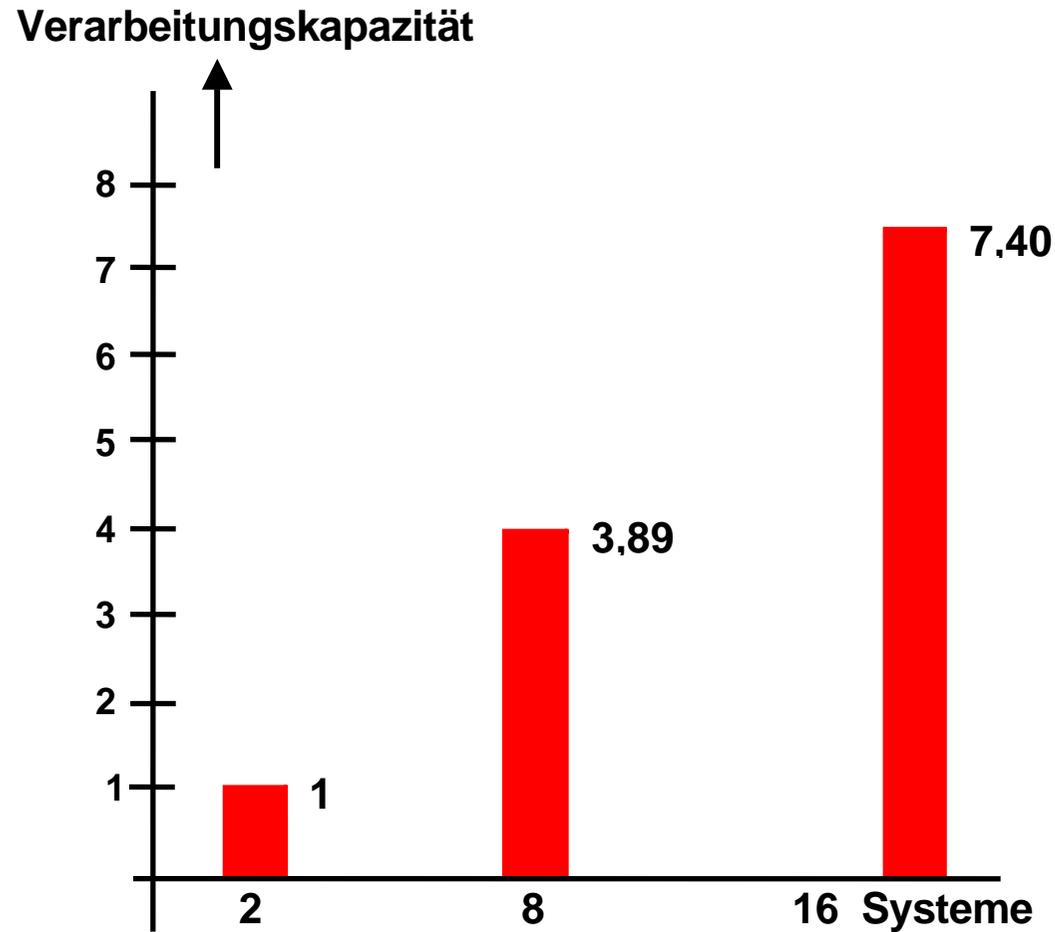


Sysplex Overhead



Installation	Anzahl Systeme	% Sysplex Overhead
A	4	11 %
B	3	10
C	8	9
D	2	7
E	11	10
Relational Warehouse Workload	2	13

Die Sysplex Software (wenn installiert) erzeugt in jedem System zusätzlichen Overhead, selbst wenn der Sysplex nur aus einem einzigen System besteht. In jedem System wird zusätzliche CPU Kapazität benötigt um den gleichen Durchsatz zu erreichen.



Parallel Sysplex Leistungsverhalten

CICS Transaktionsmanager, CICSplex System Manager, IMS Datenbank.
Mischung von OLTP, Reservierung, Data Warehouse und Bankanwendungen

Literatur: Coupling Facility Performance: A Real World Perspective, IBM Redbo