# IBM System z10 Business Class (z10 BC) Reference Guide

*The New Face of Enterprise Computing*

October 2009

## Table of Contents

## *IBM System z10 Business Class (z10 BC) Overview*

As the pace of business continues to accelerate and the planet becomes smarter, the physical and digital foundations on which progress depends are straining to keep up. They're too complex, too inefficient, and too inflexible. With our dynamic infrastructure strategy, IBM is positioned to help clients succeed in this new world, addressing today's challenges to improve service, reduce cost, and manage risk, while also laying the foundation for what is to come. We are helping clients address the increasing cost and complexity of infrastructure, link and manage all their IT and business assets, and make their business and IT infrastructure as dynamic as the business demands.

The System z10™ family of servers is well positioned to participate in this new dynamic model. It delivers many innovative technologies for flexible enterprise computing and includes proven leadership capabilities for security, availability, scalability, virtualization, and management. As environmental concerns raise the focus on energy consumption, the System z10 is designed to reduce energy usage and save floor space when consolidating workloads from distributed servers. The System z10 specialty engines continue to help users expand the use of the mainframe for a broad set of applications, while helping to lower the cost of ownership.

Protection of the IT infrastructure and data continues to be of key importance. This announcement strengthens the System z10 position in security with the next generation of cryptographic feature. Crypto Express3 is a state-of-the-art, tamper-sensing and tamper-responding programmable cryptographic feature available for the System z10. This feature can be configured as a secure key coprocessor or an accelerator.  This new generation raises the bar in error checking by using a two lock-step cross-checking CPUs process for enhanced error detection and fault isolation of cryptographic operations performed by this coprocessor.

New usability enhancements in this announcement enable the grouping of domains across multiple coprocessor features, helping to simplify the management and migration of coprocessor configuration with a new TKE Migration wizard.

At the heart of a dynamic infrasructure is the need to get to data fast and to scale with business demands and growth. The new generation of FICON® - FICON Express8 – is designed to provide faster access to data with a link rate of 8 gigabits per second (Gbps) and the ability to auto-negotiate to 2, 4  or 8 Gbps, offering investment protection for existing infrastructures - servers, switches, directors, storage devices, printers. With its performance improvements, FICON Express8 may also provide opportunities for growth of your FICON infrastructure when you are limited by availability of I/O slots. You may be able to "do more with less."

The IBM System z10 Business Class™ (z10 BC™), the technology that could change the way you think about Enterprise solutions. The technology that delivers the scalability, flexibility, virtualization, and breakthrough performance you need—at the lower capacity entry point you want. This is the technology that fights old myths and perceptions—that's not just for banks and insurance companies. This is the technology for any business that wants to ramp up innovation, boost efficiencies and lower costs—pretty much any enterprise, any size, any location. This is a mainframe technology for a new kind of data center—resilient, responsive, energy efficient—this is z10™ BC. And it's about to rewrite the rules and deliver new freedoms for your business. Whether you want to deploy new applications quickly, grow your business without growing IT costs or consolidate your infrastructure for reduced complexity, look no further—

z Can Do IT

## Think Big, Virtually Limitless

The Information Technology industry has recognized the business value of exploiting virtualization technologies on any and all server platforms. The leading edge virtualization capabilities of IBM System z®, backed by over 40 years of technology innovation, are the most advanced in the industry. With utilization rates of up to 100%, it's the perfect platform for workload consolidation, both traditional and new.

- *Want to deploy dozens—or hundreds—of applications on a single server for lower total cost of ownership? Want a more simplified, responsive infrastructure?*
- *Want investment protection where new generation technology typically allows application growth at no extra cost?*

The virtualization technology found in z/VM® with the System z platform may help clients achieve all of these operational goals while also helping to maximize the financial return on their System z investments.

The z10 BC can have big advantages over traditional server farms. The z10 BC is designed to reduce energy usage and save floor space when used to consolidate x86 servers. With increased capacity the z10 BC virtualization capabilities can help to support hundreds of virtual servers in a single 1.42 square meters footprint. When consolidating on System z you can create virtual servers on demand; achieve network savings through HiperSockets™ (internal LAN); improve systems management of virtual servers; and most importantly, consolidate software from many distributed servers to a single consolidated server.

So why run hundreds of standalone servers when z10 BC could do the work more efficiently, in a smaller space, at a lower cost, virtually? Less power. Less space. Less impact on the environment.

## More Solutions, More Affordable

Today's businesses with extensive investments in hardware assets and core applications are demanding more from IT—more value, more transactions, more for the money. Above all, they are looking for business solutions that can help enable business growth while driving costs out of the business. System z has an ever growing set of solutions that are being enhanced to help you lower IT costs. From enterprise wide applications such as SAP or Cognos® BI to the consolidation of infrastructure workloads, z10 BC has low cost solutions that also help you save more as your demand grows. So, consider consolidating your IT workloads on the z10 BC server if you want the right solutions on a premier platform at a price you can afford.

The convergence of Service-Oriented Architecture (SOA) and mainframe technologies can also help liberate these core business assets by making it easier to enrich, modernize, extend and reuse them well beyond their original scope of design. The ultimate implementation of flexibility for today's On Demand Business is a Service Oriented Architecture—an IT architectural style that allows you to design your applications to solve real business problems. The z10 BC, along with the inherent strengths and capabilities of multiple operating system choices and innovative System z software solutions from WebSphere®, CICS®, Rational® and Lotus® strengthen the flexibility of doing SOA and strengthen System z as an enterprise hub.

**Special workloads, Specialty engines, affordable technology**

The z10 BC continues the long history of providing integrated technologies to optimize a variety of workloads. The use of specialty engines can help users expand the use of the mainframe for new workloads, while helping to lower the cost of ownership. The IBM System z specialty engines can run independently or complement each other. For example, the zAAP and zIIP processors enable you to purchase additional processing capacity exclusively for specific workloads, without affecting the MSU rating of the IBM System z model designation. This means that adding a specialty engine will not cause increased charges for IBM System z software running on general purpose processors in the server.

**In order of introduction:**

The Internal Coupling Facility (ICF) processor was introduced to help cut the cost of Coupling Facility functions by reducing the need for an external Coupling Facility. IBM System z Parallel Sysplex® technology allows for greater scalability and availability by coupling mainframes together. Using Parallel Sysplex clustering, System z servers are designed for up to 99.999% availability.

The Integrated Facility for Linux® (IFL) processor offers support for Linux and brings a wealth of available applications that can be run in a real or virtual environment on the z10 BC. For customers interested to use a z10 BC only for Linux workload, the z10 BC can be configured as a server with IFLs only.

The System z10 Application Assist Processor (zAAP) is designed to help enable strategic integration of new application technologies such as Java™ technology-based Web applications and XML-based data interchange services with core business database environments.

The System z10 Integrated Information Processor (zIIP) is designed to support select data and transaction processing and network workloads and thereby make the consolidation of these workloads on to the System z platform more cost effective.

Announced August 2009, z/OS is enhanced with a new capability that can enable zAAP eligible workloads to run on zIIPs. This capability can allow you to run zIIP and zAAP eligible workloads together on one type of specialty engine – the zIIP. For additional information, see IBM System z Specialty Engine Update, Including zAAP on zIIP Capability Frequently Asked Questions XSQ0036-USEN-01.

**The New Face Of System z**

IBM's mainframe capabilities are legendary. Customers deploy systems that remain available for years because they are expected to, and continue to, work above expectations. However, these systems have seen significant innovative improvements for running new applications and consolidating workloads in the last few years, and customers can see real gains in price/performance by taking advantage of this new technology.

IBM provides affordable world-class technology to help today's enterprises respond to business conditions quickly and with flexibility. From automation to advanced virtualization technologies to new applications supported by open industry standards such as SOA, IBM servers teamed with IBM's Storage Systems, Global Technology Services and IBM Global Financing help deliver competitive advantages for a dynamic infrastructure.

z Can Do IT. The future runs on IBM System z and the future begins today!

## z/Architecture

The z10 BC continues the line of upward compatible mainframe processors and retains application compatibility since 1964. The z10 BC supports all z/Architecture®-compliant Operating Systems. The heart of the processor unit is the IBM z10 Enterprise Quad Core processor chip running at 3.5 GHz, designed to help improve CPU intensive workloads.

The z10 BC, like its predecessors, supports 24, 31, and 64-bit addressing, as well as multiple arithmetic formats. High-performance logical partitioning via Processor Resource/Systems Manager™ (PR/SM™) is achieved by industry-leading virtualization support provided by z/VM.

A change to the z/Architecture on z10 BC is designed to allow memory to be extended to support large (1 megabyte (MB)) pages. Use of large pages can improve CPU utilization for exploiting applications.

Large page support is primarily of benefit for long-running applications that are memory-access-intensive. Large page is not recommended for general use. Short-lived processes with small working sets are normally not good candidates for large pages.

Large page support is exclusive to System z10 running either z/OS or Linux on System z.

### z10 BC Architecture

Rich CISC Instruction Set Architecture (ISA)

- *894 instructions (668 implemented entirely in hardware)*
- *Multiple address spaces robust inter-process security*
- *Multiple arithmetic formats*

Architectural extensions for z10 BC

- *50+ instructions added to z10 BC to improve compiled code efficiency*
- *Enablement for software/hardware cache optimization*
- *Support for 1 MB page frames*
- *Full hardware support for Hardware Decimal Floating-point Unit (HDFU)*

### z/Architecture operating system support

Delivering the technologies required to address today's IT challenges also takes much more than just a server; it requires all of the system elements to be working together. IBM System z10 operating systems and servers are designed with a collaborative approach to exploit each other's strengths.

The z10 BC is also able to exploit numerous operating systems concurrently on a single server, these include z/OS, z/VM, z/VSE, z/TPF, TPF and Linux for System z. These operating systems are designed to support existing application investments without anticipated change and help you realize the benefits of the z10 BC. z10 BC—the new business equation.

### z/OS

Pressures on the global economy create business risks requiring increased focus on improving time-to-market, reducing costs, and optimizing profit. You need a dynamic infrastructure that can respond quickly to change while reducing cost and driving profit. This latest release of the operating system delivers technology

and virtualization to help deliver improved performance, availability, security, and efficiency, simplified administration, reduced risk, improved service, and enhanced flexibility – all of which can ultimately help reduce costs and drive profit.

Improvements for z/OS® V1.11 include:

- *Failure avoidance: Predictive failure analysis is designed to help provide early warning about system trends that can cause system or application impacts, in many cases before they impact your business. The output from predictive failure analyses can be used for reporting and logging or to drive responsive automation intended to prevent serious problems.*

- Simplified management: A new z/OS Management Facility, updates to z/OS health checks, new migration health checks, new autonomics for defining and managing I/O, simplified application of network security, as well as other operational improvements help simplify systems management, improve administrator, operator, and developer productivity, and ultimately provide less opportunity for error.

- Responsive networking: New z/OS Communications Server designs are intended to improve networking in a Parallel Sysplex®, enable more efficient workload distribution, and help improve the quality of the load balancing in multitiered z/OS server and application environments.

- Trusted system: z/OS V1.11 continues to enhance its security features and to expand them to larger enterprise-wide roles. The ability to implement centralized authentication, create a comprehensive audit and risk management plan, configure secure networks, and centrally manage digital certificate lifecycle can not only help reduce the risk from fraud and security breaches, but also help meet industry compliance guidelines.

- Accountability: Superior measurement and data collection and reporting capabilities are updated and can be used for comprehensive risk management, auditing, and compliance plans. For example, a new identity propagation function can allow z/OS subsystems (like CICS TS V4.1) to associate distributed identities to RACF® for improved cross-platform interoperability and accounting capabilities.

- Storage scalability: An additional data set type is now supported on extended address volumes (EAVs), which can be as large as 223 GB. This helps to relieve storage constraints as well as simplify storage management by providing the ability to manage fewer, large volumes as opposed to many small volumes.

- Improved resource optimization and economics: z/OS has expanded specialty engine usage. z/OS is updated to enable z/OS CIM (Common Information Model) server processing to run on the System z Integrated Information Processor (zIIP). In addition, z/OS V1.11, together with IBM DB2® for z/OS Version 8 or DB2 9 DB2 utilities, is updated to enable part of sort utility processing to run on a zIIP. Also, z/OS V1.11 is enhanced with a new function that can enable System z Application Assist Processor (zAAP) eligible workloads to run on System z Integrated Information Processors (zIIPS). This can enable you to run zIIP- and zAAP-eligible workloads on the zIIP (additional restrictions apply).

- Platform synergies: Support for IBM System StorageTM DS8000® solid state drives, DS8000 High Performance FICON multitrack function, FICON Dynamic Channel Path Management (DCM), updated IBM System z10 HiperDispatch, and IBM System z10 instruction prefetch heuristics can help improve end-to-end application, transaction, and batch performance.

- z/OS release-to-release performance improvement: For the past several releases z/OS has provided a release-to-release performance improvement by systematically reducing or eliminating constraints and inefficiencies within the base operating system. Imagine going to a new release of the z/OS operating system and not needing more system resource! Performance improvements are measured under IBM laboratory conditions and use IBM Large System Performance Reference (LSPR) workloads.

August 5, 2008, IBM announced z/OS V1.10. This release of the z/OS operating system builds on leadership capabilities, enhances time-tested technologies, and leverages deep synergies with the IBM System z10 and IBM System Storage™ family of products. z/OS V1.10 supports new capabilities designed to provide:

- *Storage scalability. Extended Address Volumes (EAVs) enable you to define volumes as large as 223 GB to relieve storage constraints and help you simplify storage management by providing the ability to manage fewer, large volumes as opposed to many small volumes.*

- *Application and data serving scalability. Up to 64 engines, up to 1.5 TB per server with up to 1.0 TB of real memory per LPAR, and support for large (1 MB) pages on the System z10 can help provide scale and performance for your critical workloads.*

- *Intelligent and optimized dispatching of workloads. HiperDispatch can help provide increased scalability and performance of higher n-way System z10 systems by improving the way workload is dispatched within the server.*

- *Low-cost, high-availability disk solution. The Basic HyperSwap™ capability (enabled by IBM TotalStorage® Productivity Center for Replication Basic Edition for System z) provides a low-cost, single-site, high-availability disk solution which allows the configuration of disk replication services using an intuitive browser-based graphical user interface (GUI) served from z/OS.*

- *Improved total cost of ownership. zIIP-Assisted Hiper-Sockets for Large Messages, IBM Scalable Architecture for Financial Reporting™ enabled for zIIP (a service offering of IBM Global Business Services), zIIP-Assisted z/OS Global Mirror (XRC), and additional z/OS XML System Services exploitation of zIIP and zAAP help make these workloads more attractive on System z.*

- *Improved management of temporary processor capacity. Capacity Provisioning Manager which is available on z/OS V1.10, and on z/OS V1.9 with PTFs, can monitor z/OS systems on System z10 servers. Activation and deactivation of temporary capacity can be suggested or performed automatically based on user-defined schedules and workload criteria. RMF™ or equivalent function is required to use the Capacity Provisioning Manager.*

- *Improved network security. z/OS Communications Server introduces new defensive filtering capability. Defensive filters are evaluated ahead of configured IP filters, and can be created dynamically, which can provide added protection and minimal disruption of services in the event of an attack.*

- *z/OS V1.10 also supports RSA key, ISO Format-3 PIN block, 13-Digit through 19-Digit PAN data, secure key AES, and SHA algorithms.*

- *Improved productivity. z/OS V1.10 provides improvements in or new capabilities for: simplifying diagnosis and problem determination; expanded Health Check Services; network and security management; automatic dump and re-IPL capability; as well as overall z/OS, I/O configuration, sysplex, and storage operations*

**Commitment to system integrity**

First issued in 1973, IBM's MVS™ System Integrity Statement and subsequent statements for OS/390® and z/OS stand as a symbol of IBM's confidence and commitment to the z/OS operating system. Today, IBM reaffirms its commitment to z/OS system integrity.

IBM's commitment includes designs and development practices intended to prevent unauthorized application programs, subsystems, and users from bypassing z/OS security—that is, to prevent them from gaining access, circumventing, disabling, altering, or obtaining control of key z/OS system processes and resources unless allowed by the installation. Specifically, z/OS "System Integrity" is defined as the inability of any program not authorized by a mechanism under the installation's control to circumvent or disable store or fetch protection, access a resource protected by the z/OS Security Server (RACF), or obtain control in an authorized state; that is, in supervisor state, with a protection key less than eight (8), or Authorized Program Facility (APF) authorized. In the event that an IBM System Integrity problem is reported, IBM will always take action to resolve it.

IBM's long-term commitment to System Integrity is unique in the industry, and forms the basis of the z/OS industry leadership in system security. z/OS is designed to help you protect your system, data, transactions, and applications from accidental or malicious modification. This is one of the many reasons System z remains the industry's premier data server for mission-critical workloads.

### z/VM

Version 6 Release 1 (V6.1) is the newest version of z/VM and is intended to be the base for all future z/VM enhancements. This release implements a new Architecture Level Set (ALS) available only on the IBM System z10 Enterprise Class server and System z10 Business Class server and future generations of System z servers. System z10 technology together with z/VM V6.1:

- *Acknowledges the highly attractive economics of workload consolidation on the more environmentally friendly, highly secure, and reliable System z10 servers*
- *Allows z/VM to take advantage of newer hardware technology for future exploitation*

Much of the success of Linux on System z can be attributed to the advanced virtualization technologies available in z/VM. Many IBM System z clients have found that the combination of Linux and z/VM not only offers server consolidation savings, but also enables the flexible deployment of business-critical enterprise solutions. Clients can run Linux on System z as a guest of z/VM on a System z server also running z/OS, z/VSEtm, and/or z/TPF systems. IBM also offers System z servers configured with Integrated Facility for Linux (IFL) specialty engines for running Linux on z/VM.

The z/VM hypervisor is designed to help clients extend the business value of mainframe technology across the enterprise by integrating applications and data while providing exceptional levels of availability, security, and operational ease. IBM System z has a rich heritage of innovation in the area of virtualization and has refined the infrastructure with a coordinated investment in hardware, firmware, hypervisors, and operating systems to enable exceptional qualities of service in the support of virtualization. z/VM virtualization technology is designed to provide the capability for clients to run hundreds or thousands of Linux servers in a single mainframe, running with other System z operating systems such as z/OS and z/VSE, or as a large-scale Linux enterprise-server solution. z/VM is designed to help improve productivity by hosting non-Linux workloads such as z/OS, z/VSE, and z/TPF.

z/VM® V6.1 is designed to offer:

- *Exploitation of the System z10 server cache management instructions to help improve the performance of virtual networking*
- *Closer integration with IBM Systems Director by shipping the Manageability Access Point Agent for z/VM*
- *Inclusion of post-z/VM V5.4 enhancements delivered in the IBM service stream including:*
  - *A CMS-based z/VM SSL server*
  - *Linux guests using dynamic storage reconfiguration*
  - *IBM FlashCopy® SE*
  - *Worldwide port name (WWPN) prediction tool*
  - *OSA-Express QDIO data connection isolation*
  - *Additional tape encryption*

z/VM V5.4 is designed to extend its System z virtualization technology leadership by exploiting more capabilities of System z servers including:

- *Greater flexibility, with support for the new z/VM-mode logical partitions, allowing all System z processor-types (CPs, IFLs, zIIPs, zAAPs, and ICFs) to be defined in the same z/VM LPAR for use by various guest operating systems*
- *Capability to install Linux on System z as well as z/VM from the HMC on a System z10 that eliminates the need for any external network setup or a physical connection between an LPAR and the HMC*
- *Enhanced physical connectivity by exploiting all OSA-Express3 ports, helping service the network and reducing the number of required resources*
- *Dynamic memory upgrade support that allows real memory to be added to a running z/VM system. With z/VM V5.4, memory can be added nondisruptively to individual guests that support the dynamic memory reconfiguration architecture. Systems can now be configured to reduce the need to re-IPL z/VM. Processors, channels, OSA adapters, and now memory can be dynamically added to both the z/VM system itself and to individual guests.*

The operation and management of virtual machines has been enhanced with new systems management APIs, improvements to the algorithm for distributing a guest's CPU share among virtual processors, and usability enhancements for managing a virtual network.

Security capabilities of z/VM V5.4 provide an upgraded LDAP server at the functional level of the z/OS V1.10 IBM Tivoli® Directory Server for z/OS and enhancements to the RACF Security Server to create LDAP change log entries in response to updates to RACF group and user profiles, including user passwords and password phrases. The z/VM SSL server now operates in a CMS environment, instead of requiring a Linux distribution, thus allowing encryption services to be deployed more quickly and helping to simplify installation, service, and release-to-release migration.

The z/VM hypervisor is designed to help clients extend the business value of mainframe technology across the enterprise by integrating applications and data while providing exceptional levels of availability, security, and operational ease. z/VM virtualization technology is designed to provide the capability for clients to run hundreds to thousands of Linux servers in a single mainframe, together with other System z operating systems such as z/OS, or as a large-scale Linux-only enterprise-server solution. z/VM V5.4 can also help to improve productivity by hosting non-Linux workloads such as z/OS, z/VSE, and z/TPF.

August 5, 2008, IBM announced z/VM 5.4. Enhancements in z/VM 5.4 include:

- *Increased flexibility with support for new z/VM-mode logical partitions*
- *Dynamic addition of memory to an active z/VM LPAR by exploiting System z dynamic storage-reconfiguration capabilities*
- *Enhanced physical connectivity by exploiting all OSA-Express3 ports*
- *Capability to install Linux on System z from the HMC without requiring an external network connection*
- *Enhancements for scalability and constraint relief*
- *Operation of the SSL server in a CMS environment*
- *Systems management enhancements for Linux and other virtual images*

For the most current information on z/VM, refer to the z/VM Web site at: http://www.vm.**ibm.com**.

**z/VSE**

z/VSE Version 4 Release 2 (z/VSE V4.2) is built on a heritage of ongoing refinement and innovation that spans more than four decades. It is designed to help protect and leverage existing investments in VSE information assets. z/VSE V4.2 is designed to address the needs of clients with growing core VSE workloads and/or those who wish to exploit Linux on System z for new, Web-based business solutions and infrastructure simplification.

z/VSE V4.2 exploits selected features of innovative IBM System z and IBM System Storage technology. It executes in z/Architecture mode only and supports 64-bit real addressing for selected system functions. z/VSE V4.2 is not designed to support 64-bit virtual addressing or allow 64-bit addressing for user applications.

z/VSE V4.2 continues the z/VSE focus on scalability, security, service-oriented architecture (SOA), and hybrid solutions that exploit z/VSE and Linux on System z.

z/VSE V4.2 is designed to:

- *Support up to 512 VSE tasks. That is two times the 255 task limit in z/VSE V4.1 and prior versions. Removing the task constraint may help VSE clients grow their CICS workloads and consolidate their VSE systems, and may ease the migration from CICS/VSE® to CICS Transaction Server for VSE/ESA.*

- *Exploit up to 32 GB of processor storage. That is four times the 8 GB limit in z/VSE V4.1 and 16 times the limit in z/VSE V3.1 and VSE/ESA V2.*

- *Support the Parallel Access Volume (PAV) feature of DS8000, DS6000, and ESS series. PAV allows accessing an ECKD-type volume with multiple concurrent requests. PAV may result in a performance improvement compared to traditional I/O.*

- *Enable LDAP (Lightweight Directory Access Protocol) sign-on for z/VSE V4.2. A user ID and password can be authenticated using an LDAP server that is connected to z/VSE V4.2 via a TCP/IP network. z/VSE V4.2 provides an LDAP client only. An LDAP server on another platform is required. The use of "company-wide" user IDs connects z/VSE with the centralized management of user IDs. As a result, password rules and password renewal can be enforced via the LDAP server. For example, LDAP sign-on can overcome the previous limitation that z/VSE user IDs must be between 4 and 8 characters, and passwords are limited to 8 characters. Using LDAP sign-on, user IDs and passwords can both be up to 64 characters. LDAP authorization is designed to integrate z/VSE into "Identity Management Systems," such as IBM Tivoli products, or systems based on the Microsoft® Active Directory.*

- *Support the SAN Volume Controller (SVC). In addition to IBM System Storage qualified for use with z/VSE FCP-SCSI support, VSE clients can now access the IBM System Storage disk subsystems and OEM SCSI disk devices supported by the SVC. SVC support offers more storage choices to clients and may help lower Total Cost of Ownership (TCO).*

z/VSE V4.2 is the preferred follow-on VSE product for clients with z/VSE V4.1, z/VSE V3, or VSE/ESA installed. You may choose to use the Fast Service Upgrade (FSU) process when migrating from z/VSE V4.1 or z/VSE V3.1.

### z/TPF

z/TPF is a 64-bit operating system that allows you to move legacy applications into an open development environment, leveraging large scale memory spaces for increased speed, diagnostics and functionality. The open development environment allows access to commodity skills and enhanced access to open code libraries, both of which can be used to lower development costs. Large memory spaces can be used to increase both system and application efficiency as I/Os or memory management can be eliminated.

z/TPF is designed to support:

- *64-bit mode*
- *Linux development environment (GCC and HLASM for Linux)*
- *32 processors/cluster*
- *Up to 84* engines/processor*
- *40,000 modules*
- *Workload License Charge*

### Linux on System z

The System z10 BC supports the following Linux on System z distributions (most recent service levels):

- *Novell SUSE SLES 9*
- *Novell SUSE SLES 10*
- *Red Hat RHEL 4*
- *Red Hat RHEL 5*

The IBM System z10 Business Class (z10 BC) delivers innovative technologies for small and medium enterprises that give you a whole new world of capabilities to run modern applications. Ideally suited in a Dynamic Infratructure, this competitively priced server delivers unparalleled qualities of service to help manage growth and reduce cost and risk in your business.

The z10 BC further extends the leadership of System z by delivering expanded granularity and optimized scalability for growth, enriched virtualization technology for consolidation of distributed workloads, improved availability and security to help increase business resiliency, and just-in-time management of resources. The z10 BC is at the core of the enhanced System z platform and is the new face of System z.

The z10 BC has the machine type of 2098, with one model (E10) offering between one to ten configurable Processor Units (PUs). This model design offers increased flexibility over the two model IBM System z9® Business Class (z9® BC) by delivering seamless growth within a single model, both temporary and permanent.

The z10 BC delivers improvements in both the granular increments and total scalability compared to previous System z midrange servers, achieved by both increasing the performance of the individual PU as well as increasing the number of PUs per server. The z10 BC Model E10 is designed to provide up to 1.5 times the total system capacity for general purpose processing, and over 40% more configurable processors than the z9 BC Model S07.

The z10 BC advances the innovation of the System z10 platform and brings value to a wider audience. It is built using a redesigned air cooled drawer package which replaces the prior "book" concept in order to reduce cost and increase flexibility. A redesigned I/O drawer offers higher availability and can be concurrently added or replaced when at least two drawers are installed. Reduced

capacity and priced I/O features will continue to be offered on the z10 BC to help lower your total cost of acquisition. The quad core design z10 processor chip delivers higher frequency and will be introduced at 3.5 GHz which can help improve the execution of CPU intensive workloads on the z10 BC. These design approaches facilitate the high-availability, dynamic capabilities and lower cost that differentiate this z10 BC from other servers.

The z10 BC supports from 4 GB up to 248 GB of real customer memory. This is almost four times the maximum memory available on the z9 BC. The increased available memory on the server can help to benefit workloads that perform better with larger memory configurations, such as DB2, WebSphere and Linux. In addition to the customer purchased memory, an additional 8 GB of memory is included for the Hardware System Area (HSA). The HSA holds the I/O configuration data for the server and is entirely fenced from customer memory.

High speed connectivity and high bandwidth out to the data and the network are critical in achieving high levels of transaction throughput and enabling resources inside and outside the server to maximize application requirements. The z10 BC has a host bus interface with a link data rate of 6 GB using the industry standard InfiniBand protocol to help satisfy requirements for coupling (ICF and server-to-server connectivity), cryptography (Crypto Express3 and Crypto Express2 with secure coprocessors and SSL transactions), I/O (ESCON, FICON or FCP) and LAN (OSA-Express3 Gigabit, 10 Gigabit and 1000BASE-T Ethernet features). High Performance FICON for System z (zHPF) also brings new levels of performance when accessing data on enabled storage devices such as the IBM System Storage DS8000.

PUs defined as Internal Coupling Facilities (ICFs), Integrated Facility for Linux (IFLs), System z10 Application Assist Processor (zAAPs) and System z10 Integrated Information Processor (zIIPs) are no longer grouped together in one pool as on the IBM eServer™ zSeries® 890 (z890), but are grouped together in their own pool, where they can be managed separately. The separation significantly simplifies capacity planning and management for LPAR and can have an effect on weight management since CP weights and zAAP and zIIP weights can now be managed separately. Capacity BackUp (CBU) features are available for IFLs, ICFs, zAAPs and zIIPs.

LAN connectivity has been enhanced with the introduction of the third generation of Open Systems Adapter-Express (OSA-Express3). This new family of LAN adapters have been introduced to reduce latency and overhead, deliver double the port density of OSA-Express2 and provide increased throughput. The z10 BC continues to support OSA-Express2 1000BASE-T and GbE Ethernet features, and supports IP version 6 (IPv6) on HiperSockets. While OSA-Express2 OSN (OSA for NCP) is still available on System z10 BC to support the Channel Data Link Control (CDLC) protocol, the OSA-Express3 will also provide this function.

Additional channel and networking improvements include support for Layer 2 and Layer 3 traffic, FCP management facility for z/VM and Linux for System z, FCP security improvements, and Linux support for HiperSockets IPv6. STP enhancements include the additional support for NTP clients and STP over InfiniBand links.

The Crypto Express3 utilizes PCI-E adapters, and like the Crypto Express2 feature, with PCI-X adapters, can be individually configured as a secure coprocessor or an accelerator for SSL. Both The Crypto Express3 and Crypto Express2 support the attachment of a TKE workstation with optional Smart Card Reader for additional security when loading master and operational keys.

The CP Assist for Cryptographic Function (CPACF) provides:

- *DES, TDES, AES-128, AES-192, AES-256*
- *SHA-1, SHA-224, SHA-256, SHA-384, SHA-512*
- *Pseudo Random Number Generation (PRNG)*

z10 BC is designed to deliver the industry leading Reliability, Availability and Serviceability (RAS) customers expect from System z servers. RAS is designed to reduce all sources of outages by reducing unscheduled, scheduled and planned outages. Planned outages are further designed to be reduced by reducing preplanning requirements.

z10 BC preplanning improvements are designed to avoid planned outages and include:

- *Flexible Customer Initiated Upgrades*
- *Enhanced Driver Maintenance*
  - *Multiple "from" sync point support*
- *Reduce Pre-planning to avoid Power-On-Reset*
  - *8 GB for HAS*
  - *Dynamic I/O enabled by default*
  - *Add Logical Channel Subsystems (LCSS)*
  - *Change LCSS Subchannel Sets*
  - *Add/delete Logical partitions*

- *Designed to eliminate a logical partition deactivate/ activate/IP*
  - *Dynamic Change to Logical Processor Definition*
  - *Dynamic Change to Logical Cryptographic Coprocessor Definition*
- Concurrent I/O drawer add/repair

Additionally, several service enhancements have also been designed to avoid unscheduled outages and include continued focus on firmware quality, reduced chip count on Single Chip Module (SCM) and memory subsystem improvements. In the area of scheduled outage enhancements include redundant 100 Mb Ethernet service network with VLAN, rebalance of PSIFB and I/O fanouts, and single processor core sparing and checkstop. Exclusive to the System z10 is the ability to hot swap ICB-4 and InfiniBand hub cards.

Enterprises with IBM System z9 BC and IBM eServer zSeries 890 (z890) may upgrade to any z10 BC model. Model upgrades within the z10 BC are concurrent. If you desire a consolidation platform for your mainframe and Linux capable applications, you can add capacity and even expand your current application workloads in a cost-effective manner. If your traditional and new applications are growing, you may find the z10 BC a good fit with its base qualities of service and its specialty processors designed for assisting with new workloads. Value is leveraged with improved hardware price/performance and System z10 BC software pricing strategies.

The z10 BC is specifically designed and optimized for full z/Architecture compatibility. New features enhance enterprise data serving performance, industry leading

z10 BC Design and Technology

virtualization capabilities, energy efficiency at system and data center levels. The z10 BC is designed to further extend and integrate key platform characteristics such as dynamic flexible partitioning and resource management in mixed and unpredictable workload environments, providing scalability, high availability and Qualities of Service (QoS) to emerging applications such as WebSphere, Java and Linux.

With the logical partition (LPAR) group capacity limit on z10 BC, z10 EC, z9 EC and z9 BC, you can now specify LPAR group capacity limits allowing you to define each LPAR with its own capacity and one or more groups of LPARs on a server. This is designed to allow z/OS to manage the groups in such a way that the sum of the LPARs' CPU utilization within a group will not exceed the group's defined capacity. Each LPAR in a group can still optionally continue to define an individual LPAR capacity limit.

The z10 BC has one model with a total of 130 capacity settings available as new build systems and as upgrades from the z9 BC and z890.

The z10 BC model is designed with a Central Processor Complex (CPC) drawer with Single Chip Modules (SCM) that provides up to 10 Processor Units (PUs) that can be characterized as either Central Processors (CPs), IFLs, ICFs, zAAPs or zIIPs.

Some of the significant enhancements in the z10 BC that help bring improved performance, availability and function to the platform have been identified. The following sections highlight the functions and features of the z10 BC.

The System z10 BC is designed to provide balanced system performance. From processor storage to the system's I/O and network channels, end-to-end bandwidth is provided and designed to deliver data where and when it is needed.

The processor subsystem is comprised of one CPC, which houses the processor units (PUs), Storage Controllers (SCs), memory, Self-Time-Interconnects (STI)/InfiniBand® (IFB) and Oscillator/External Time Reference (ETR). The z10 BC design provides growth paths up to a 10 engine system where each of the 10 PUs has full access to all system resources, specifically memory and I/O.

The z10 BC uses the same processor chip as the z10 EC, relying only on 3 out of 4 functional cores per chip. Each chip is individually packaged in an SCM. Four SCMs will be plugged in the processor board providing the 12 PUs for the design. Clock frequency will be 3.5 GHz.

There are three active cores per PU, an L1 cache divided into a 64 KB cache for instructions and a 128 KB cache for data. Each PU also has an L1.5 cache. This cache is 3 MB in size. Each L1 cache has a Translation Look-aside Buffer (TLB) of 512 entries associated with it. The PU, which uses a high-frequency z/Architecture microprocessor core, is built on CMOS 11S chip technology and has a cycle time of approximately 0.286 nanoseconds.

The PU chip includes data compression and cryptographic functions. Hardware data compression can play a significant role in improving performance and saving costs over doing compression in software. Standard clear key cryptographic processors right on the processor translate to high-speed cryptography for protecting data in storage, integrated as part of the PU.

Speed and precision in numerical computing are important for all our customers. The z10 BC offers improvements for decimal floating point instructions, because each z10 processor chip has its own hardware decimal floating point unit, designed to improve performance over that provided by the System z9. Decimal calculations are often used in financial applications and those done using other floating point facilities have typically been performed by software through the use of libraries. With a hardware decimal floating point unit some of these calculations may be done directly and accelerated.

The design of the z10 BC provides the flexibility to configure the PUs for different uses; There are 12 PUs per system, two are designated as System Assist Processors (SAPs) standard per system. The remaining 10 PUs are available to be characterized as either CPs, ICF processors for Coupling Facility applications, or IFLs for Linux applications and z/VM hosting Linux as a guest, System z10 Application Assist Processors (zAAPs), System z10 Integrated Information Processors (zIIPs) or as optional SAPs and provide you with tremendous flexibility in establishing the best system for running applications.

The z10 BC can support from the 4 GB minimum memory up to 248 GB of available real memory per server for growing application needs. A new 8 GB fixed HSA which is managed separately from customer memory. This fixed HSA is designed to improve availability by avoiding outages that were necessary on prior models to increase its size. There are up to 12 I/O interconnects per system at 6 GBps each.

The z10 BC supports a combination of Memory Bus Adapter (MBA) and Host Channel Adapter (HCA) fanout cards. New MBA fanout cards are used exclusively for ICB-4. New ICB-4 cables are needed for z10 BC. The InfiniBand Multiplexer (IFB-MP) card replaces the Self-Timed Interconnect Multiplexer (STI-MP) card. There are two types of HCA fanout cards: HCA2-C is copper and is always used to connect to I/O (IFB-MP card) and the HCA2-O which is optical and used for customer InfiniBand coupling.

The z10 BC has been designed to offer high performance and efficient I/O structure. The z10 BC ships with a single frame: the A-Frame which supports the installation of up to four I/O drawers. Each drawer supports up to eight I/O cards, four in front and four in the rear, providing support for up to 480 channels (32 I/O features).

To increase the I/O device addressing capability, the I/O subsystem has been enhanced by introducing support for multiple subchannels sets (MSS), which are designed to allow improved device connectivity for Parallel Access Volumes (PAVs). To support the highly scalable system design, the z9 BC I/O subsystem uses the Logical Channel SubSystem (LCSS) which provides the capability to install up to 512 CHPIDs across the I/O drawers (256 per operating system image). The Parallel Sysplex Coupling Link architecture and technology continues to support high speed links providing efficient transmission between the Coupling Facility and z/OS systems. HiperSockets provides high speed capability to communicate among virtual servers and logical partitions. HiperSockets is now improved with the IP version 6 (IPv6) support; this is based on high speed TCP/IP memory speed transfers and provides value in allowing applications running in one partition to communicate with applications running in another without dependency on an external network. Industry standard and openness are design objectives for I/O in z9 BC.

## z10 BC Model

The z10 BC has one model, the E10, (Machine Type 2098) offering between 1 to 10 processor units (PUs), which can be configured to provide a highly scalable solution designed to meet the needs of both high transaction processing applications and On Demand business. The PUs can be characterized as either CPs, IFLs, ICFs, zAAPs, zIIPs or option SAPs. An easy-to-enable ability to "turn off" CPs or IFLs is available on z10 BC, allowing you to purchase capacity for future use with minimal or no impact on software billing. An MES feature will enable the "turned off" CPs or IFLs for use where you require the increased capacity. There are a wide range of upgrade options available in getting to and within the z10 BC.

The z10 BC hardware model number (E10) on its own does not indicate the number of PUs which are being used as CPs. For software billing purposes only, there will be a Capacity Indicator associated with the number PUs that are characterized as CPs. This number will be reported by the Store System Information (STSI) instruction for software billing purposes only. There is no affinity between the hardware model and the number of CPs.

### z10 BC capacity identifiers
nxx, where n = subcapacity engine size and xx = number of CPs

- *Total 130 Capacity Indicators for "software settings"*
- *A00 for systems with IFL(s) or ICF(s) only.*

### Memory DIMM sizes: 2 GB and 4 GB
- *Maximum physical memory: 256 GB per system*
  - *Minimum physical installed = 16 GB of which 8 GB is for Fixed HSA*
- *For 8 to 32, 4 GB increments, from 32 to 248, 8 GB increments*

### z10 BC model upgrades
The z10 BC provides for the dynamic and flexible capacity growth for mainframe servers. There are full upgrades within the z10 BC and upgrades from any z9 BC or z890 to any z10 BC. Temporary capacity upgrades are available through On/Off Capacity on Demand (CoD).
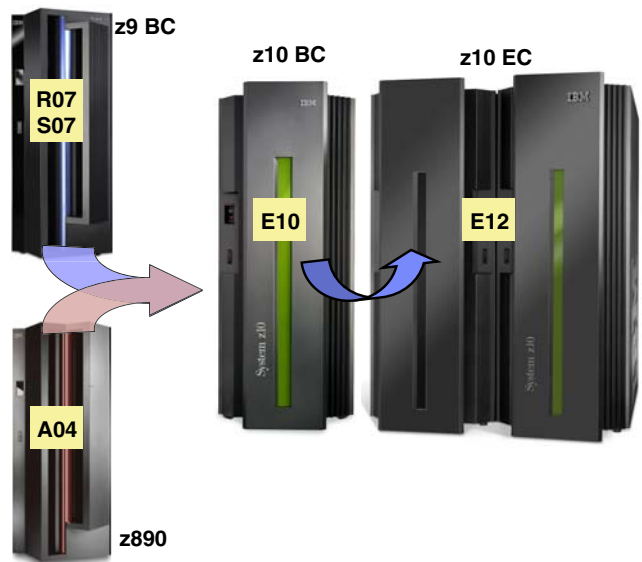
For the z10 BC models, there are twenty-six capacity settings per engine for central processors (CPs). Sub-capacity processors have availability of z10 BC features/functions and any-to-any upgradeability is available within the sub-capacity matrix. All CPs must be the same capacity setting size within one z10 BC. All specialty engines run at full speed.

The one for one entitlement to purchase one zAAP and/or one zIIP for each CP purchased is the same for CPs of any speed.

**z10 BC Model Capacity IDs:**
- *A00, A01 to Z01, A01 to Z02, A03 to Z03, A04 to Z04, and A05 to Z05*
- *Capacity setting A00 does not have any CP engines*
- *Nxx, where n = the capacity setting of the engine, and xx = the number of PU characterized as CPs in the CPC*

| | | | | |
|---|---|---|---|---|
| Z01 | Z02 | Z03 | Z04 | Z05 |
| Y01 | Y02 | Y03 | Y04 | Y05 |
| X01 | X02 | X03 | X04 | X05 |
| W01 | W02 | W03 | W04 | W05 |
| V01 | V02 | V03 | V04 | V05 |
| U01 | U02 | U03 | U04 | U05 |
| T01 | T02 | T03 | T04 | T05 |
| S01 | S02 | S03 | S04 | S05 |
| R01 | R02 | R03 | R04 | R05 |
| Q01 | Q02 | Q03 | Q04 | Q05 |
| P01 | P02 | P03 | P04 | P05 |
| O01 | O02 | O03 | O04 | O05 |
| N01 | N02 | N03 | N04 | N05 |
| M01 | M02 | M03 | M04 | M05 |
| L01 | L02 | L03 | L04 | L05 |
| K01 | K02 | K03 | K04 | K05 |
| J01 | J02 | J03 | J04 | J05 |
| I01 | I02 | I03 | I04 | I05 |
| H01 | H02 | H03 | H04 | H05 |
| G01 | G02 | G03 | G04 | G05 |
| F01 | F02 | F03 | F04 | F05 |
| E01 | E02 | E03 | E04 | E05 |
| D01 | D02 | D03 | D04 | D05 |
| C01 | C02 | C03 | C04 | C05 |
| B01 | B02 | B03 | B04 | B05 |
| A01 | A02 | A03 | A04 | A05 |
| 1-way | 2-way | 3-way | 4-way | 5-way |
| Specialty Engine | Specialty Engine | Specialty Engine | Specialty Engine | Specialty Engine | Specialty Engine | Specialty Engine | Specialty Engine | Specialty Engine | Specialty Engine |

TThe performance design of the z/Architecture can enable the server to support a new standard of performance for applications through expanding upon a balanced system approach. As CMOS technology has been enhanced to support not only additional processing power, but also more PUs, the entire server is modified to support the increase in processing power. The I/O subsystem supports a greater amount of bandwidth than previous generations through internal changes, providing for larger and faster volume of data movement into and out of the server. Support of larger amounts of data within the server required improved management of storage configurations, made available through integration of the operating system and hardware support of 64-bit addressing. The combined balanced system design allows for increases in performance across a broad spectrum of work.

**Large System Performance Reference**
IBM's Large Systems Performance Reference (LSPR) method is designed to provide comprehensive z/Architecture processor capacity ratios for different configurations of Central Processors (CPs) across a wide variety of system control programs and workload environments. For z10 BC, z/Architecture processor capacity identifier is defined with a (A0x-Z0x) notation, where x is the number of installed CPs, from one to five. There are a total of 26 sub-capacity levels, designated by the letters A through Z.

In addition to the general information provided for z/OS V1.9, the LSPR also contains performance relationships for z/VM and Linux operating environments.

Based on using an LSPR mixed workload, the performance of the z10 BC (2098) Z01 is expected to be:

• *Up to 1.4 times that of the z9 BC (2096) Z01.*

Moving from a System z9 partition to an equivalently sized System z10 BC partition, a z/VM workload will experience an ITR ratio that is somewhat related to the workload's instruction mix, MP factor, and level of storage over commitment. Workloads with higher levels of storage over commitment or higher MP factors are likely to experience lower than average z10 BC to z9 ITR scaling ratios. The range of likely ITR ratios is wider than the range has been for previous processor migrations.

The LSPR contains the Internal Throughput Rate Ratios (ITRRs) for the z10 BC and the previous-generation zSeries processor families based upon measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user may experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, and the workload processed. Therefore no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated. For more detailed performance information, consult the Large Systems Performance Reference (LSPR) available at: http://www.**ibm.com**/servers/eserver/zseries/lspr/.

**CPU Measurement Facility**

The CPU Measurement Facility is a hardware facility which consists of counters and samples. The facility provides a means to collect run-time data for software performance tuning. The detailed architecture information for this facility can be found in the System z10 Library in Resource Link™.

A new host bus interface using InfiniBand with a link data rate of 6 GBps, was introduced on the z10 BC. It provides enough throughput to support the full capacity and processing power of the CPC. The z10 BC contains an I/O subsystem infrastructure which uses up to four I/O drawers that provides eight I/O slots in each drawer. There are two I/O domains per drawer, and four I/O cards per domain. I/O cards are horizontal and may be added concurrently. Concurrent replacement and/or repair is available with systems containing more than one I/O drawer. Drawers may be added concurrently should the need for more connectivity arise.

ESCON, FICON Express8, FICON Express4, FICON Express2, FICON Express, OSA-Express3, OSA-Express2, Crypto Express3 and Crypto Express2 features plug into the z10 BC I/O drawer along with any ISC-3s and InfiniBand Multiplexer (IFB-MP) cards. All I/O features and their support cards can be hot-plugged in the I/O drawer. Each model ships with one I/O drawer as standard in the A-Frame (the A-Frame also contains the Central Processor Complex [CPC]), where the I/O drawers are installed. Each IFB-MP has a bandwidth up to 6 GigaBytes per second (GB/sec) for I/O domains and MBA fanout cards provide 2.0 GB/sec for ICB-4s.

The z10 BC continues to support all of the features announced with the System z9 BC such as:

• *Logical Channel Subsystems (LCSSs) and support for up to 30 logical partitions*

• *Increased number of Subchannels (63.75k)*

• *Multiple Subchannel Sets (MSS)*

• *Redundant I/O Interconnect*

• *Physical Channel IDs (PCHIDs)*

• *System Initiated CHPID Reconfiguration*

• *Logical Channel SubSystem (LCSS) Spanning*

**System I/O Configuration Analyzer**

Today the information needed to manage a system's I/O configuration has to be obtained from many separate applications. The System's I/O Configuration Analyzer (SIOA) tool is a SE/HMC-based tool that will allow the system hardware administrator access to the information from these many sources in one place. This will make it much easier to manage I/O configurations, particularly across multiple CPCs. The SIOA is a "view-only" tool. It does not offer any options other than viewing options.

First the SIOA tool analyzes the current active IOCDS on the SE. It extracts information about the defined channel, partitions, link addresses and control units. Next the SIOA tool asks the channels for their node ID information. The FICON channels support remote node ID information, so that is also collected from them. The data is then formatted and displayed on five screens:

1) *PCHID Control Unit Screen – Shows PCHIDs, CSS. CHPIDs and their control units*

2) *PCHID Partition Screen – Shows PCHIDS, CSS. CHPIDs and what partitions they are in*

3) *Control Unit Screen – Shows the control units, their PCHIDs and their link addresses in each of the CSS's*

4) *Link Load Screen – Shows the Link address and the PCHIDs that use it*

5) *Node ID Screen – Shows the Node ID data under the PCHIDs*

The SIOA tool allows the user to sort on various columns and export the data to a USB flash drive for later viewing.

**ESCON**

The z10 BC supports up to 480 ESCON channels. The high density ESCON feature has 16 ports, 15 of which can be activated for customer use. One port is always reserved as a spare which is activated in the event of a failure of one of the other ports. For high availability the initial order of ESCON features will deliver two 16-port ESCON features and the active ports will be distributed across those features.

**FICON**

With introduction of FICON Express8 on the System z10 BC, you now have additional growth opportunities for your storage area network (SAN). FICON Express 8 provides synergy with existing switches, directors, and storage devices. With support for native FICON, High Performance FICON for System z (zHPF), and Fibre Channel Protocol (FCP), the System z10 servers enable you to position your SAN for even higher performance - helping you to pre-pare for an end-to-end 8 Gbps infrastructure to meet the increased bandwidth demands of your applications.

The z10 BC continues to support your current fiber optic cabling environments with FICON Express8.

New with the System z10, all FICON channel error informa-tion is forwarded to the HMC where it is analyzed to help detect and report the trends and thresholds for all FICON channels on System z10, whether CHPID type FC or FCP. This report shows an aggregate view of the data and can span multiple systems.

**FICON Express8**

The z10 BC supports up to 128 FICON Express8 channels, each one operating at 2, 4, or 8 Gb/sec auto-negotiated. The FICON Express8 features are available in long wave-length (LX) and short wavelength (SX with four channels per feature.

FICON Express8 10KM LX designed to support unrepeated distances up to 10 kilometers (6.2 miles) over 9 micron single mode fiber optic cabling without performance degradation. To avoid performance degradation at extended distance, FICON switches or directors (for buffer credit provision) or Dense Wavelength Division Multiplexers (for buffer credit simulation) may be required.

FICON Express8 SX is designed to support 50 or 62.5 micron multimode fiber optic cabling.

For details regarding the unrepeated distances for FICON Express8 10KM LX and  FICON Express8 SX refer to System z Planning for Fiber Optic Links (GA23-0367) available on System z10 servers at planned availability in the Library section of Re-source Link.
www.**ibm.com**/servers/resourcelink


All channels on a single FICON Express8 feature are of the same type - 10KM LX or SX. Both features support small form factor pluggable optics (SFPs) with LC Duplex connectors. The optics continue to permit each channel to be individually serviced in the event of a fiber optic module failure.

The FICON Express8 features are exclusive to z10 EC and z10 BC servers.

### FICON Express4
The z10 BC supports up to 128 FICON Express4 channels, each one operating at 1, 2 or 4 Gb/sec auto-negotiated. The FICON Express4 features are available in long wavelength (LX) and short wavelength (SX). For customers exploiting LX, there are two options available for unre-

peated distances of up to 4 kilometers (2.5 miles) or up to 10 kilometers (6.2 miles). Both LX features use 9 micron single mode fiber optic cables. The SX feature uses 50 or 62.5 micron multimode fiber optic cables. Each FICON Express4 feature has four independent channels (ports) and can be configured to carry native FICON traffic or Fibre Channel (SCSI) traffic. LX and SX cannot be intermixed on a single feature. The receiving devices must correspond to the appropriate LX or SX feature. The maximum number of FICON Express4 features is 32 using four I/O drawers.

Exclusive to the z10 BC and z9 BC is the availability of a, lower cost FICON Express4 2-port feature, the FICON Express4-2C 4KM LX and FICON Express4-2C SC. These features support two FICON 4 Gbps LX and SX channels respectively. The FICON Express4-2-port cards are designed to operate like the 4 port card but with the flexibility of having fewer ports per card.


### FICON Express2
The z10 BC supports carrying forward FICON Express2 channels, each one operating at 1 or 2 Gb/sec auto-negotiated.


### FICON Express Channels
The z10 BC also supports carrying forward FICON Express LX and SX channels from z9 BC and z990 each channel operating at 1 or 2 Gb/sec auto-negotiated. Each FICON Express feature has two independent channels (ports).

The System z10 BC Model E10 is limited to 32 features

– any combination of FICON Express8, FICON Express4, FICON Express2 and FICON Express LX and SX features.

The FICON Express8, FICON Express4, FICON Express2 and FICON Express feature conforms to the Fibre Connection (FICON) architecture and the Fibre Channel (FC) architecture, providing connectivity between any combination of servers, directors, switches, and devices in a Storage Area Network (SAN).

For more information about FICON, visit the IBM Redbooks® Web site at: http://www.redbooks.**ibm.com**/ search for SG24-5444. There are also various FICON I/O Connectivity information at: www-03.**ibm.com**/systems/z/connectivity/.

### Concurrent Update
The FICON Express8 and FICON Express4 SX and LX features may be added to an existing z10 BC concurrently. This concurrent update capability allows you to continue to run workloads through other channels while the new features are being added. This applies to CHPID types FC and FCP.

### Modes of Operation
There are two modes of operation for supported FICON Express features. These modes are configured on a channel-by-channel basis – each of the channels can be configured in either of two supported modes.

- *Fibre Channel (CHPID type FC), which is native FICON or FICON Channel-to-Channel (server-to-server)*
- *Fibre Channel Protocol (CHPID type FCP), which supports attachment to SCSI devices via Fibre Channel switches or directors in z/VM, z/VSE, and Linux on System z10 environments*
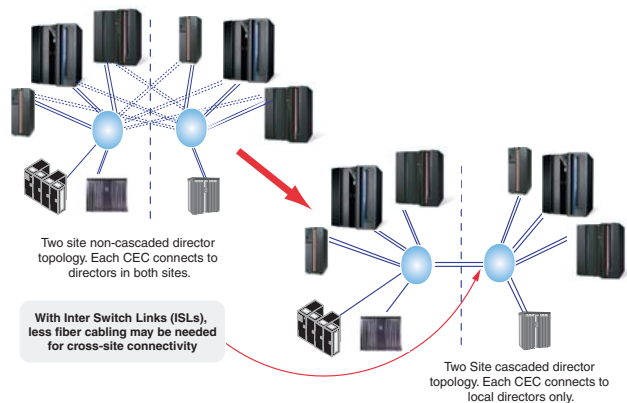
### Native FICON Channels
Native FICON channels and devices can help to reduce bandwidth constraints and channel contention to enable easier server consolidation, new application growth, large business intelligence queries and exploitation of On Demand Business.

Because the FICON CTC function is included as part of the native FICON (FC) mode of operation, FICON CTC is not limited to intersystem connectivity (as is the case with ESCON), but will support multiple device definitions.

### FICON Support for Cascaded Directors
Native FICON (FC) channels support cascaded directors. This support is for a single hop configuration only. Two-director cascading requires a single vendor high integrity fabric. Directors must be from the same vendor since cascaded architecture implementations can be unique. This type of cascaded support is important for disaster recovery and business continuity solutions because it can help provide high availability, extended distance connectivity, and (particularly with the implementation of 2 Gb/sec Inter Switch Links) has the potential for fiber infrastructure cost savings by reducing the number of channels for interconnecting the two sites.



Two site non-cascaded director topology. Each CEC connects to directors in both sites.

With Inter Switch Links (ISLs), less fiber cabling may be needed for cross-site connectivity

Two Site cascaded director topology. Each CEC connects to local directors only.

**FCP Channels**

The same FICON features used for native FICON channels can be defined to be used for Fibre Channel Protocol (FCP) channels. FCP channels are defined as CHPID type FCP. The 8 Gb/sec capability on the FICON Express8 channel means that 8 Gb/sec link data rates are available for FCP channels as well.

**FICON Express8 performance improvements for FCP on the System z10 servers**

The FICON Express8 FCP channel, when operating at 8 Gbps, is designed to achieve a maximum throughput of up to 800 MBps when processing large sequential read I/O operations and up to 730 MBps when processing large sequential write I/O operations.  This represents an 80 to 100% increase compared to a FICON Express4 FCP channel operating at 4 Gbps on System z10.

The FICON Express8 FCP channel is designed to achieve a maximum of 84,000 IO/sec when processing read or write small data transfer I/O operations. This represents approximately a 40% increase compared to a FICON Express4 FCP channel when operating at 4 Gbps on a System z10 server.

These measurements for FCP (CHPID type FCP supporting attachment to SCSI devices) are examples of the maximum MB/sec and IO/sec that can be achieved in a laboratory environment, using one FICON Express8 channel on a System z10 server with z/VM V5.4 or Linux on System z distribution Novell SUSE SLES 10 with no other processing occurring, and do not represent actual field measurements

**SCSI IPL now a base function**

The SCSI Initial Program Load (IPL) enablement feature, first introduced on z990 in October of 2003, is no longer required. The function is now delivered as a part of the server Licensed Internal Code. SCSI IPL allows an IPL of an operating system from an FCP-attached SCSI disk.

**FCP Full fabric connectivity**

FCP full fabric support means that any number of (single vendor) FCP directors/ switches can be placed between the server and an FCP/SCSI device, thereby allowing many "hops" through a Storage Area Network (SAN) for I/O connectivity. FCP full fabric connectivity enables multiple FCP switches/directors on a fabric to share links and therefore provides improved utilization of inter-site connected resources and infrastructure.

High performance FICON for System z - improving upon the native FICON protocol: The FICON Express8 features support High Performance FICON for System z (zHPF) which was introduced in October 2008 on the System z10 servers. zHPF provides optimizations for online transaction processing (OLTP) workloads. zHPF is an extension to the FICON architecture and is designed to improve the execution of small block I/O requests. zHPF streamlines the FICON architecture and reduces the overhead on the channel processors, control unit ports, switch ports, and links by improving the way channel programs are written and processed. zHPF-capable channels and devices support both native FICON and zHPF protocols simultaneously (CHPID type FC).

High Performance FICON for System z now supports multitrack operations: zHPF support of multitrack operations can help increase system performance and improve FICON channel efficiency when attached to the IBM System Storage DS8000 series.  zFS, HFS, PDSE, and other applications that use large data transfers with Media Manager are expected to benefit.

In laboratory measurements, multitrack operations (e.g. reading 16x4k bytes/IO) converted to the zHPF protocol on a FICON Express8 channel, achieved a maximum of up to 40% more MB/sec than multitrack operations using the native FICON protocol.

zHPF and support for multitrack operations is exclusive to the System z10 servers and applies to all FICON Express8, FICON Express4,  and  FICON Express2 features (CHPID type FC). Exploitation is required by z/OS and the control unit.

zHPF with multitrack operations is available in the DS8000 series Licensed Machine Code (LMC) level 5.4.3.xx (bundle version 64.3.xx.xx) or later with the purchase of the DS8000 High Performance FICON for System z feature.

Previously, zHPF was limited to read or write sequential I/O's transferring less than a single track size (for example, 12 4k byte records or 12x4k bytes/IO).

FICON Express8 performance improvements for zHPF and native FICON on the System z10 servers: A FICON Express8 channel exploiting the High Performance FICON for System z (zHPF) protocol, when operating at 8 Gbps, is designed to achieve a maximum throughput of up to 800 MBps when processing large sequential read I/O operations and up to 730 MBps when processing large sequential write I/O operations.  This represents an 80 to 100% increase in performance compared to a FICON Express4 channel operating at 4 Gbps on a System z10 server. For those large sequential read or write I/O operations that use the native FICON protocol, the FICON Express8 channel, when operating at 8 Gbps, is designed to achieve up to 510 MBps. This represents 1 a 45 to 55% increase compared to a FICON Express4 channel operating at 4 Gbps on a System z10 server.

The FICON Express8 channel, when operating at 8 Gbps, is also designed to achieve a maximum of 52,000 IO/sec for small data transfer I/O operations that can exploit the zHPF protocol. This represents approximately a 70% increase compared to a FICON Express4 channel operating at 4 Gbps and executing zHPF I/O operations on

System a z10 server. For those small data transfer I/O operations that use the native FICON protocol, the FICON Express8 channel, when operating at 8 Gbps, is designed to achieve a maximum of 20,000 IO/sec, which represents approximately a 40% increase compared to a FICON Express4 channel operating at 4 Gbps on a System z10 server. The FICON Express8 features support both the native FICON protocol and the zHPF protocol concurrently in the server Licensed Internal Code.

These measurements for FICON (CHPID type FC) using both the native FICON and zHPF protocols are examples of the maximum MB/sec and IO/sec that can be achieved in a laboratory environment using one FICON Express8 channel on a System z10 server with z/OS V1.10 and no other processing occurring and do not represent actual field measurements.

**Platform and name server registration in FICON channel**
The FICON channel now provides the same information to the fabric as is commonly provided by open systems, registering with the name server in the attached FICON directors. With this information, your storage area network (SAN) can be more easily and efficiently managed, enhancing your ability to perform problem determination and analysis.

Registration allows other nodes and/or SAN managers to query the name server to determine what is connected to the fabric, what protocols are supported (FICON, FCP) and to gain information about the System z10 using the attributes that are registered. The FICON channel is now designed to perform registration with the fibre channel's Management Service and Directory Service.

It will register:

- *Platform's:*
  - *Worldwide node name (node name for the platform – same for all channels)*
  - *Platform type (host computer)*
  - *Platform name (includes vendor ID, product ID, and vendor specific data from the node descriptor)*
- *Channel's:*
  - *Worldwide port name (WWPN)*
  - *Node port identification (N_PORT ID)*
  - *FC-4 types supported (always 0x1B and additionally 0x1C if any Channel-to-Channel (CTC) control units are defined on that channel)*
  - *Classes of service support by the channel*

Platform registration is a service defined in the Fibre Channel – Generic Services 4 (FC-GS-4) standard (INCITS (ANSI) T11 group).

Platform and name server registration applies to all of the FICON Express8, FICON Express4, FICON Express2, and FICON Express features (CHPID type FC). This support is exclusive to System z10 and is transparent to operating systems.

**Preplanning and setup of SAN for a System z10 environment**
The worldwide port name (WWPN) prediction tool is now available to assist you with preplanning of your Storage Area Network (SAN) environment prior to the installation of your System z10 server.

This standalone tool is designed to allow you to setup your SAN in advance, so that you can be up and running much faster once the server is installed. The tool assigns WWPNs to each virtual Fibre Channel Protocol (FCP) channel/port using the same WWPN assignment algorithms a system uses when assigning WWPNs for channels utilizing N_Port Identifier Virtualization (NPIV).

The tool needs to know the FCP-specific I/O device definitions in the form of a .csv file. This file can either be created manually, or exported from Hardware Configuration Definition/Hardware Configuration Manager (HCD/HCM). The tool will then create the WWPN assignments, which are required to set up your SAN. The tool will also create a binary configuration file that can later on be imported by your system.

The WWPN prediction tool can be downloaded from Resource Link and is applicable to all FICON channels defined as CHPID type FCP (for communication with SCSI devices). Check Preventive Service Planning (PSP) buckets for required maintenance.
http://www.**ibm.com**/servers/resourcelink/

**Extended distance FICON – improved performance at extended distance**
An enhancement to the industry standard FICON architecture (FC-SB-3) helps avoid degradation of performance at extended distances by implementing a new protocol for "persistent" Information Unit (IU) pacing. Control units that exploit the enhancement to the architecture can increase the pacing count (the number of IUs allowed to be in flight from channel to control unit). Extended distance FICON also allows the channel to "remember" the last pacing update for use on subsequent operations to help avoid degradation of performance at the start of each new operation.

Improved IU pacing can help to optimize the utilization of the link, for example help keep a 4 Gbps link fully utilized at 50 km, and allows channel extenders to work at any distance, with performance results similar to that experienced when using emulation.

The requirements for channel extension equipment are simplified with the increased number of commands in flight. This may benefit z/OS Global Mirror (Extended Remote Copy – XRC) applications as the channel extension kit is no longer required to simulate specific channel commands. Simplifying the channel extension requirements may help reduce the total cost of ownership of end-to-end solutions.

Extended distance FICON is transparent to operating systems and applies to all the FICON Express 8, FICON Express4 and FICON Express2 features carrying native FICON traffic (CHPID type FC). For exploitation, the control unit must support the new IU pacing protocol. The channel will default to current pacing values when operating with control units that cannot exploit extended distance FICON.

Exploitation of extended distance FICON is supported by IBM System Storage DS8000 series Licensed Machine Code (LMC) level 5.3.1xx.xx (bundle version 63.1.xx.xx), or later.

To support extended distance without performance degradation, the buffer credits in the FICON director must be set appropriately. The number of buffer credits required is dependent upon the link data rate (1 Gbps, 2 Gbps, or 4 Gbps), the maximum number of buffer credits supported by the FICON director or control unit, as well as application and workload characteristics. High bandwidth at extended distances is achievable only if enough buffer credits exist to support the link data rate.

**FICON Express enhancements for Storage Area Networks**

**N_Port ID Virtualization**
N_Port ID Virtualization is designed to allow for sharing of a single physical FCP channel among multiple operating system images. Virtualization function is currently available for ESCON and FICON channels, and is now available for FCP channels. This function offers improved FCP channel utilization due to fewer hardware requirements, and can reduce the complexity of physical FCP I/O connectivity.

**Program Directed re-IPL**
Program Directed re-IPL is designed to enable an operating system to determine how and from where it had been loaded. Further, Program Directed re-IPL may then request that it be reloaded again from the same load device using the same load parameters. In this way, Program Directed re-IPL allows a program running natively in a partition to trigger a re-IPL. This re-IPL is supported for both SCSI and ECKD™ devices. z/VM 5.3 provides support for guest exploitation.

**FICON Link Incident Reporting**
FICON Link Incident Reporting is designed to allow an operating system image (without operating intervention) to register for link incident reports, which can improve the ability to capture data for link error analysis. The information can be displayed and is saved in the system log.

**Serviceability Enhancements**
Requests Node Identification Data (RNID) is designed to facilitate the resolution of fiber optic cabling problems. You can now request RNID data for a device attached to a native FICON channel.

**Local Area Network (LAN) connectivity**

**OSA-Express3 – the newest family of LAN adapters**
The third generation of Open Systems Adapter-Express (OSA-Express3) features have been introduced to help reduce latency and overhead, deliver double the port density of OSA-Express2, and provide increased throughput.

Choose the OSA-Express3 features that best meet your business requirements.

To meet the demands of your applications, provide granularity, facilitate redundant paths, and satisfy your infrastructure requirements, there are seven features from which to choose. In the 10 GbE environment, Short Reach (SR) is being offered for the first time.

| Feature | Infrastructure | Ports per Feature |
|---|---|---|
| OSA-Express3 GbE LX | Single mode fiber | 4 |
| OSA-Express3 10 GbE LR | Single mode fiber | 2 |
| OSA-Express3 GbE SX | Multimode fiber | 4 |
| OSA-Express3 10 GbE SR | Multimode fiber | 2 |
| OSA-Express3-2P GbE SX | Multimode fiber | 2 |
| OSA-Express3 1000BASE-T | Copper | 4 |
| OSA-Express3-2P 1000BASE-T | Copper | 2 |

Note that software PTFs or a new release may be required (depending on CHPID type) to support all ports.

**OSA-Express3 for reduced latency and improved throughput**
To help reduce latency, the OSA-Express3 features now have an Ethernet hardware data router; what was previously done in firmware (packet construction, inspection, and routing) is now performed in hardware. With direct memory access, packets flow directly from host memory to the LAN without firmware intervention. OSA-Express3 is also designed to help reduce the round-trip networking time between systems. Up to a 45% reduction in latency at the TCP/IP application layer has been measured.

The OSA-Express3 features are also designed to improve throughput for standard frames (1492 byte) and jumbo frames (8992 byte) to help satisfy the bandwidth requirements of your applications. Up to a 4x improvement has been measured (compared to OSA-Express2).

The above statements are based on OSA-Express3 performance measurements performed in a laboratory environment on a System z10 and do not represent actual field measurements. Results may vary.

**Response time improvements with OSA-Express3 optimized latency mode:**
886 Optimized Latency Mode (OLM) can help improve performance for z/OS workloads with demanding low latency requirements. This includes interactive workloads such as SAP using DB2 Connect. OLM can help improve performance for applications that have a critical requirement to minimize response times for inbound and outbound data when servicing remote clients. This enhancement applies exclusively to OSA-Express3 QDIO mode (CHPID type OSD).

**Port density or granularity**
The OSA-Express3 features have Peripheral Component Interconnect Express (PCI-E) adapters. The previous table identifies whether the feature has 2 or 4 ports for LAN connectivity. Select the density that best meets your business requirements. Doubling the port density on a single feature helps to reduce the number of I/O slots required for high-speed connectivity to the Local Area Network.

The OSA-Express3 10 GbE features support Long Reach (LR) using 9 micron single mode fiber optic cabling and Short Reach (SR) using 50 or 62.5 micron multimode fiber optic cabling. The connector is new; it is now the small form factor, LC Duplex connector. Previously the SC Duplex connector was supported for LR. The LC Duplex connector is common with FICON, ISC-3, and OSA-Express2 Gigabit Ethernet LX and SX.

**The OSA-Express3 features are exclusive to System z10**
There are operating system dependencies for exploitation of two ports in OSD mode per PCI-E adapter. Whether it is a 2-port or a 4-port feature, only one of the ports will be visible on a PCI-E adapter if operating system exploitation updates are not installed.

**OSA-Express3 Ethernet features – Summary of benefits**
OSA-Express3 10 GbE LR (single mode fiber), 10 GbE SR (multimode fiber), GbE LX (single mode fiber), GbE SX (multimode fiber), and 1000BASE-T (copper) are designed for use in high-speed enterprise backbones, for local area network connectivity between campuses, to connect server farms to System z10, and to consolidate file servers onto System z10. With reduced latency, improved through-put, and up to 96 ports of LAN connectivity, (when all are 4-port features, 24 features per server), you can "do more with less."

**The key benefits of OSA-Express3 compared to OSA-Express2 are:**
- *Reduced latency (up to 45% reduction) and increased throughput (up to 4x) for applications*
- *More physical connectivity to service the network and fewer required resources:*

 – *Fewer CHPIDs to define and manage*
 – *Reduction in the number of required I/O slots*
 – *Possible reduction in the number of I/O drawers*
 – *Double the port density of OSA-Express2*
 – *A solution to the requirement for more than 48 LAN ports (now up to 96 ports)*

The OSA-Express3 features are exclusive to System z10.

**OSA-Express3 10 GbE**

**OSA-Express3 10 Gigabit Ethernet LR**
The OSA-Express3 10 Gigabit Ethernet (GbE) long reach (LR) feature has two ports. Each port resides on a PCIe adapter and has its own channel path identifier (CHPID). There are two PCIe adapters per feature. OSA-Express3 10 GbE LR is designed to support attachment to a 10 Gigabits per second (Gbps) Ethernet Local Area Network (LAN) or Ethernet switch capable of 10 Gbps. OSA-Express3 10 GbE LR supports CHPID type OSD exclu-sively. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

**OSA-Express3 10 Gigabit Ethernet SR**
The OSA-Express3 10 Gigabit Ethernet (GbE) short reach (LR) feature has two ports. Each port resides on a PCIe adapter and has its own channel path identifier (CHPID). There are two PCIe adapters per feature. OSA-Express3 10 GbE SR is designed to support attachment to a 10 Gigabits per second (Gbps) Ethernet Local Area Network (LAN) or Ethernet switch capable of 10 Gbps. OSA-Express3 10 GbE SR supports CHPID type OSD exclu-sively. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

### OSA-Express3 Gigabit Ethernet LX

The OSA-Express3 Gigabit Ethernet (GbE) long wavelength (LX) feature has four ports. Two ports reside on a PCIe adapter and share a channel path identifier (CHPID). There are two PCIe adapters per feature. Each port supports attachment to a one Gigabit per second (Gbps) Ethernet Local Area Network (LAN). OSA-Express3 GbE LX supports CHPID types OSD and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

### OSA-Express3 Gigabit Ethernet SX

The OSA-Express3 Gigabit Ethernet (GbE) short wavelength (SX) feature has four ports. Two ports reside on a PCIe adapter and share a channel path identifier (CHPID). There are two PCIe adapters per feature. Each port supports attachment to a one Gigabit per second (Gbps) Ethernet Local Area Network (LAN). OSA-Express3 GbE SX supports CHPID types OSD and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

### OSA-Express3-2P Gigabit Ethernet SX

The OSA-Express3-2P Gigabit Ethernet (GbE) short wavelength (SX) feature has two ports which reside on a single PCIe adapter and share one channel path identifier (CHPID). Each port supports attachment to a one Gigabit per second (Gbps) Ethernet Local Area Network (LAN). OSA-Express3 GbE SX supports CHPID types OSD and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

### Four-port exploitation on OSA-Express3 GbE SX and LX

For the operating system to recognize all four ports on an OSA-Express3 Gigabit Ethernet feature, a new release and/or PTF is required. If software updates are not applied, only two of the four ports will be "visible" to the operating system.

Activating all four ports on an OSA-Express3 feature provides you with more physical connectivity to service the network and reduces the number of required resources (I/O slots, I/O cages, fewer CHPIDs to define and manage).

Four-port exploitation is supported by z/OS, z/VM, z/VSE, z/TPF, and Linux on System z.

### OSA-Express3 1000BASE-T Ethernet

The OSA-Express3 1000BASE-T Ethernet feature has four ports. Two ports reside on a PCIe adapter and share a channel path identifier (CHPID). There are two PCIe adapters per feature. Each port supports attachment to either a 10BASE-T (10 Mbps), 100BASE-TX (100 Mbps), or 1000BASE-T (1000 Mbps or 1 Gbps) Ethernet Local Area Network (LAN). The feature supports auto-negotiation and automatically adjusts to 10, 100, or 1000 Mbps, depending upon the LAN. When the feature is set to autonegotiate, the target device must also be set to autonegotiate. The feature supports the following settings: 10 Mbps half or full duplex, 100 Mbps half or full duplex, 1000 Mbps (1 Gbps) full duplex. OSA-Express3 1000BASE-T Ethernet supports CHPID types OSC, OSD, OSE, and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

When configured at 1 Gbps, the 1000BASE-T Ethernet feature operates in full duplex mode only and supports jumbo frames when in QDIO mode (CHPID type OSD).

### OSA-Express3-2P 1000BASE-T Ethernet

The OSA-Express3-2P 1000BASE-T Ethernet feature has two ports which reside on a single PCIe adapter and share one channel path identifier (CHPID). Each port supports attachment to either a 10BASE-T (10 Mbps), 100BASE-TX (100 Mbps), or 1000BASE-T (1000 Mbps or 1 Gbps) Ethernet Local Area Network (LAN). The feature supports auto-negotiation and automatically adjusts to 10, 100, or 1000

Mbps, depending upon the LAN. When the feature is set to autonegotiate, the target device must also be set to auto-negotiate. The feature supports the following settings: 10 Mbps half or full duplex, 100 Mbps half or full duplex, 1000 Mbps (1 Gbps) full duplex. OSA-Express3 1000BASE-T Ethernet supports CHPID types OSC, OSD, OSE, and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs. Software updates are required to exploit both ports.

When configured at 1 Gbps, the 1000BASE-T Ethernet feature operates in full duplex mode only and supports jumbo frames when in QDIO mode (CHPID type OSD).

**OSA-Express2 availability**
OSA-Express2 Gigabit Ethernet and 1000BASE-T Ethernet continue to be available for ordering, for a limited time, if you are not yet in a position to migrate to the latest release of the operating system for exploitation of two ports per PCI-E adapter and if you are not resource-constrained.

Historical summary: Functions that continue to be supported by OSA-Express3 and OSA-Express2:

- *Queued Direct Input/Output (QDIO) – uses memory queues and a signaling protocol to directly exchange data between the OSA microprocessor and the network software for high-speed communication.*
  - *QDIO Layer 2 (Link layer) – for IP (IPv4, IPv6) or non-IP (AppleTalk, DECnet, IPX, NetBIOS, or SNA) work-loads. Using this mode the Open Systems Adapter (OSA) is protocol-independent and Layer-3 indepen-dent. Packet forwarding decisions are based upon the Medium Access Control (MAC) address.*
  - *QDIO Layer 3 (Network or IP layer) – for IP workloads. Packet forwarding decisions are based upon the IP address. All guests share OSA's MAC address.*

- *Jumbo frames in QDIO mode (8992 byte frame size) when operating at 1 Gbps (fiber or copper) and 10 Gbps (fiber).*
- *640 TCP/IP stacks per CHPID – for hosting more images*
- *Large send for IPv4 packets – for TCP/IP traffic and CPU efficiency, offloading the TCP segmentation processing from the host TCP/IP stack to the OSA-Express feature*
- *Multiple Image Facility (MIF) and spanned channels – for sharing OSA among logical channel subsystems*

The OSA-Express3 and OSA-Express2 Ethernet features support the following CHPID types:

| CHPID Type | OSA-Express3, OSA-Express2 Features | Purpose/Traffic |
|---|---|---|
| OSC | 1000BASE-T | OSA-Integrated Console Controller (OSA-ICC) TN3270E, non-SNA DFT, IPL to CPC and LPARs Operating system console operations |
| OSD | 1000BASE-T GbE 10 GbE | Queued Direct Input/Output (QDIO) TCP/IP traffic when Layer 3 Protocol-independent when Layer 2 |
| OSE | 1000BASE-T passthru (LCS) | Non-QDIO, SNA/APPN®/HPR and/or TCP/IP |
| OSN | 1000BASE-T GbE | OSA for NCP Supports channel data link control (CDLC) |

**OSA-Express QDIO data connection isolation for the z/VM environment**
Multi-tier security zones are fast becoming the network configuration standard for new workloads. Therefore, it is essential for workloads (servers and clients) hosted in a virtualized environment (shared resources) to be protected from intrusion or exposure of data and processes from other workloads.

With Queued Direct Input/Output (QDIO) data connection isolation you:

- *Have the ability to adhere to security and HIPAA-security guidelines and regulations for network isolation between the operating system instances sharing physical network connectivity*
- *Can establish security zone boundaries that have been defined by your network administrators*

Internal "routing" can be disabled on a per QDIO connection basis. This support does not affect the ability to share an OSA-Express port. Sharing occurs as it does today, but the ability to communicate between sharing QDIO data connections may be restricted through the use of this support. You decide whether an operating system's or z/VM's Virtual Switch OSA-Express QDIO connection is to be non-isolated (default) or isolated.

QDIO data connection isolation applies to the device statement defined at the operating system level. While an OSA-Express CHPID may be shared by an operating system, the data device is not shared.

QDIO data connection isolation is supported by z/VM and applies to OSA-Express3 and OSA-Express2 features (CHPID type OSD) on System z10 and to the OSA-Express2 features on System z9.

**Network Traffic Analyzer**

With the large volume and complexity of today's network traffic, the z10 BC offers systems programmers and network administrators the ability to more easily solve network problems. With the introduction of the OSA-Express Network Traffic Analyzer and QDIO Diagnostic Synchronization on the System z and available on the z10 BC, customers will have the ability to capture trace/trap data and forward it to z/OS 1.8 tools for easier problem determination and resolution.

This function is designed to allow the operating system to control the sniffer trace for the LAN and capture the records into host memory and storage (file systems), using existing host operating system tools to format, edit, and process the sniffer records.

OSA-Express Network Traffic Analyzer is exclusive to the z10 BC, z9 BC, z10 EC, and z9 EC, and is applicable to the OSA-Express3 and OSA-Express2 features when configured as CHPID type OSD (QDIO), and is supported by z/OS.

**Dynamic LAN idle for z/OS**

Dynamic LAN idle is designed to reduce latency and improve network performance by dynamically adjusting the inbound blocking algorithm. When enabled, the z/OS TCP/IP stack is designed to adjust the inbound blocking algorithm to best match the application requirements.

For latency sensitive applications, the blocking algorithm is modified to be "latency sensitive." For streaming (throughput sensitive) applications, the blocking algorithm is adjusted to maximize throughput. The z/OS TCP/IP stack can dynamically detect the application requirements, making the necessary adjustments to the blocking algorithm. The monitoring of the application and the blocking algorithm adjustments are made in real-time, dynamically adjusting the application's LAN performance.

System administrators can authorize the z/OS TCP/IP stack to enable a dynamic setting, which was previously a static setting. The z/OS TCP/IP stack is able to help determine the best setting for the current running application, based on system configuration, inbound workload volume, CPU utilization, and traffic patterns.

## Link aggregation for z/VM in Layer 2 mode

z/VM Virtual Switch-controlled (VSWITCH-controlled) link aggregation (IEEE 802.3ad) allows you to dedicate an OSA-Express2 (or OSA-Express3) port to the z/VM operating system when the port is participating in an aggregated group when configured in Layer 2 mode. Link aggregation (trunking) is designed to allow you to combine multiple physical OSA-Express3 and OSA-Express2 ports (of the same type for example 1GbE or 10GbE) into a single logical link for increased throughput and for nondisruptive failover in the event that a port becomes unavailable.

- *Aggregated link viewed as one logical trunk and containing all of the Virtual LANs (VLANs) required by the LAN segment*
- *Load balance communications across several links in a trunk to prevent a single link from being overrun*
- *Link aggregation between a VSWITCH and the physical network switch*
- *Point-to-point connections*
- *Up to eight OSA-Express3 or OSA-Express2 ports in one aggregated link*
- *Ability to dynamically add/remove OSA ports for "on demand" bandwidth*
- *Full-duplex mode (send and receive)*
- *Target links for aggregation must be of the same type (for example, Gigabit Ethernet to Gigabit Ethernet)*

The Open Systems Adapter/Support Facility (OSA/SF) will provide status information on an OSA port – its "shared" or "exclusive use" state. OSA/SF is an integrated component of z/VM.

Link aggregation is exclusive to System z10 and System z9, is applicable to the OSA-Express3 and OSA-Express2 features in Layer 2 mode when configured as CHPID type OSD (QDIO), and is supported by z/VM 5.3 and later.

## Layer 2 transport mode: When would it be used?

If you have an environment with an abundance of Linux images in a guest LAN environment, or you need to define router guests to provide the connection between these guest LANs and the OSA-Express3 features, then using the Layer 2 transport mode may be the solution. If you have Internetwork Packet Exchange (IPX), NetBIOS, and SNA protocols, in addition to Internet Protocol Version 4 (IPv4) and IPv6, use of Layer 2 could provide "protocol independence."

The OSA-Express3 features have the capability to perform like Layer 2 type devices, providing the capability of being protocol- or Layer-3-independent (that is, not IP-only). With the Layer 2 interface, packet forwarding decisions are based upon Link Layer (Layer 2) information, instead of Network Layer (Layer 3) information. Each operating system attached to the Layer 2 interface uses its own MAC address. This means the traffic can be IPX, NetBIOS, SNA, IPv4, or IPv6.

An OSA-Express3 feature can filter inbound datagrams by Virtual Local Area Network identification (VLAN ID, IEEE 802.1q), and/or the Ethernet destination MAC address. Filtering can reduce the amount of inbound traffic being processed by the operating system, reducing CPU utilization.

Layer 2 transport mode is supported by z/VM and Linux on System z.

## OSA Layer 3 Virtual MAC for z/OS

To simplify the infrastructure and to facilitate load balancing when an LPAR is sharing the same OSA Media Access Control (MAC) address with another LPAR, each operating system instance can now have its own unique "logical" or "virtual" MAC (VMAC) address. All IP addresses associated with a TCP/IP stack are accessible using their own

VMAC address, instead of sharing the MAC address of an OSA port. This applies to Layer 3 mode and to an OSA port shared among Logical Channel Subsystems.

This support is designed to:

- *Improve IP workload balancing*
- *Dedicate a Layer 3 VMAC to a single TCP/IP stack*
- *Remove the dependency on Generic Routing Encapsulation (GRE) tunnels*
- *Improve outbound routing*
- *Simplify configuration setup*
- *Allow WebSphere Application Server content-based routing to work with z/OS in an IPv6 network*
- *Allow z/OS to use a "standard" interface ID for IPv6 addresses*
- *Remove the need for PRIROUTER/SECROUTER function in z/OS*

OSA Layer 3 VMAC for z/OS is exclusive to System z, and is applicable to OSA-Express3 and OSA-Express2 features when configured as CHPID type OSD (QDIO).

### Direct Memory Access (DMA)

OSA-Express3 and the operating systems share a common storage area for memory-to-memory communication, reducing system overhead and improving performance. There are no read or write channel programs for data exchange. For write processing, no I/O interrupts have to be handled. For read processing, the number of I/O interrupts is minimized.

### Hardware data router

With OSA-Express3, much of what was previously done in firmware (packet construction, inspection, and routing) is now performed in hardware. This allows packets to flow directly from host memory to the LAN without firmware intervention.

With the hardware data router, the "store and forward" technique is no longer used, which enables true direct memory access, a direct host memory-to-LAN flow, returning CPU cycles for application use.

This avoids a "hop" and is designed to reduce latency and to increase throughput for standard frames (1492 byte) and jumbo frames (8992 byte).

### IBM Communication Controller for Linux (CCL)

CCL is designed to help eliminate hardware dependencies, such as 3745/3746 Communication Controllers, ESCON channels, and Token Ring LANs, by providing a software solution that allows the Network Control Program (NCP) to be run in Linux on System z freeing up valuable data center floor space.

CCL helps preserve mission critical SNA functions, such as SNI, and z/OS applications workloads which depend upon these functions, allowing you to collapse SNA inside a z10 BC while exploiting and leveraging IP.

The OSA-Express3 and OSA-Express2 GbE and 1000BASE-T Ethernet features provide support for CCL. This support is designed to require no changes to operating systems (does require a PTF to support CHPID type OSN) and also allows TPF to exploit CCL. Supported by z/VM for Linux and z/TPF guest environments.

**OSA-Express3 and OSA-Express2 OSN (OSA for NCP)**

OSA-Express for Network Control Program (NCP), Channel path identifier (CHPID) type OSN, is now available for use with the OSA-Express3 GbE features as well as the OSA-Express3 1000BASE-T Ethernet features.

OSA-Express for NCP, supporting the channel data link control (CDLC) protocol, provides connectivity between System z operating systems and IBM Communication Controller for Linux (CCL). CCL allows you to keep your business data and applications on the mainframe operating systems while moving NCP functions to Linux on System z.

CCL provides a foundation to help enterprises simplify their network infrastructure while supporting traditional Systems Network Architecture (SNA) functions such as SNA Network Interconnect (SNI).

Communication Controller for Linux on System z (Program Number 5724-J38) is the solution for companies that want to help improve network availability by replacing Token-Ring networks and ESCON channels with an Ethernet network and integrated LAN adapters on System z10, OSA-Express3 or OSA-Express2 GbE or 1000BASE-T.

OSA-Express for NCP is supported in the z/OS, z/VM, z/VSE, TPF, z/TPF, and Linux on System z environments.

**OSA Integrated Console Controller**

The OSA-Express Integrated Console Controller (OSA-ICC) support is a no-charge function included in Licensed Internal Code (LIC) on z10 BC, z10 EC, z9 EC, z9 BC, z990, and z890 servers. It is available via the OSA Express3, OSA-Express2 and OSA-Express 1000BASE-T Ethernet features, and supports Ethernet-attached TN3270E consoles.

The OSA-ICC provides a system console function at IPL time and operating systems support for multiple logical partitions. Console support can be used by z/OS, z/OS.e, z/VM, z/VSE, z/TPF, and TPF. The OSA-ICC also supports local non-SNA DFT 3270 and 328x printer emulation for TSO/E, CICS, IMS™, or any other 3270 application that communicates through VTAM®.

With the OSA-Express3 and OSA-Express2 1000BASE-T Ethernet features, the OSA-ICC is configured on a port by port basis, using the Channel Path Identifier (CHPID) type OSC. Each port can support up to 120 console session connections, can be shared among logical partitions using Multiple Image Facility (MIF), and can be spanned across multiple Channel Subsystems (CSSs).

Four port exploitation for OSA-Express3 1000BASE-T (#3367) and two port exploitation for OSA-Express3-2P 1000BASE-T (#3369) for OSA-ICC will be available in the first quarter of 2010.

**Remove L2/L3 LPAR-to-LPAR Restriction**

OSA port sharing between virtual switches can communicate whether the transport mode is the same (Layer 2 to Layer 2) or different (Layer 2 to Layer 3). This enhancement is designed to allow seamless mixing of Layer 2 and Layer 3 traffic, helping to reduce the total cost of networking. Previously, Layer 2 and Layer 3 TCP/IP connections through the same OSA port (CHPID) were unable to communicate with each other LPAR-to-LPAR using the Multiple Image Facility (MIF).

This enhancement is designed to facilitate a migration from Layer 3 to Layer 2 and to continue to allow LAN administrators to configure and manage their mainframe network topology using the same techniques as their non-mainframe topology.

**OSA/SF Virtual MAC and VLAN id Display Capability**

The Open Systems Adapter/Support Facility (OSA/SF) has the capability to support virtual Medium Access Control (MAC) and Virtual Local Area Network (VLAN) identifications (IDs) associated with OSA-Express2 feature configured as a Layer 2 interface. This information will now be displayed as a part of an OSA Address Table (OAT) entry. This information is independent of IPv4 and IPv6 formats. There can be multiple Layer 2 VLAN Ids associated to a single unit address. One group MAC can be associated to multiple unit addresses.

For additional information, view IBM Redbooks, IBM System z Connectivity Handbook (SG24-5444) at: www.redbooks.**ibm.com**/.

The HiperSockets function, also known as internal Queued Direct Input/Output (iDQIO) or internal QDIO, is an integrated function of the z10 BC server that provides users with attachments to up to sixteen high-speed "virtual" Local Area Networks (LANs) with minimal system and network overhead. HiperSockets eliminates the need to utilize I/O subsystem operations and the need to traverse an external network connection to communicate between logical partitions in the same z10 BC server.

Now, the HiperSockets internal networks on z10 BC can support two transport modes: Layer 2 (Link Layer) as well as the current Layer 3 (Network or IP Layer). Traffic can be Internet Protocol (IP) version 4 or version 6 (IPv4, IPv6) or non-IP (AppleTalk, DECnet, IPX, NetBIOS, or SNA). HiperSockets devices are now protocol-independent and Layer 3 independent. Each HiperSockets device has its own Layer 2 Media Access Control (MAC) address, which is designed to allow the use of applications that depend on the existence of Layer 2 addresses such as DHCP servers and firewalls.

Layer 2 support can help facilitate server consolidation. Complexity can be reduced, network configuration is simplified and intuitive, and LAN administrators can configure and maintain the mainframe environment the same as they do a non-mainframe environment. With support of the new Layer 2 interface by HiperSockets, packet forwarding decisions are now based upon Layer 2 information, instead of Layer 3 information. The HiperSockets device performs automatic MAC address generation and assignment to allow uniqueness within and across logical partitions (LPs) and servers. MAC addresses can also be

locally administered. The use of Group MAC addresses for multicast is supported as well as broadcasts to all other Layer 2 devices on the same HiperSockets network. Datagrams are only delivered between HiperSockets devices that are using the same transport mode (Layer 2 with Layer 2 and Layer 3 with Layer 3). A Layer 2 device cannot communicate directly with a Layer 3 device in another LPAR.

A HiperSockets device can filter inbound datagrams by Virtual Local Area Network identification (VLAN ID, IEEE 802.1q), the Ethernet destination MAC address, or both. Filtering can help reduce the amount of inbound traffic being processed by the operating system, helping to reduce CPU utilization.

Analogous to the respective Layer 3 functions, HiperSockets Layer 2 devices can be configured as primary or secondary connectors or multicast routers. This is designed to enable the creation of high performance and high availability Link Layer switches between the internal HiperSockets network and an external Ethernet or to connect the HiperSockets Layer 2 networks of different servers. The HiperSockets Multiple Write Facility for z10 BC is also supported for Layer 2 HiperSockets devices, thus allowing performance improvements for large Layer 2 datastreams.

HiperSockets Layer 2 support is exclusive to System z10 and is supported by z/OS, Linux on System z environments, and z/VM for Linux guest exploitation.

.

**HiperSockets Multiple Write Facility for increased performance**

Though HiperSockets provides high-speed internal TCP/IP connectivity between logical partitions within a System z server – the problem is that HiperSockets draws excessive CPU utilization for large outbound messages. This may lead to increased software licensing cost – HiperSockets large outbound messages are charged to a general CPU which can incur high general purpose CPU costs. This may also lead to some performance issues due to synchronous application blocking – HiperSockets large outbound messages will block a sending application while synchronously moving data.

A solution is HiperSockets Multiple Write Facility. HiperSockets performance has been enhanced to allow for the streaming of bulk data over a HiperSockets link between logical partitions (LPARs). The receiving LPAR can now process a much larger amount of data per I/O interrupt. This enhancement is transparent to the operating system in the receiving LPAR. HiperSockets Multiple Write Facility, with fewer I/O interrupts, is designed to reduce CPU utilization of the sending and receiving LPAR.

The HiperSockets Multiple Write solution moves multiple output data buffers in one write operation.

If the function is disabled then one output data buffer is moved in one write operation. This is also how HiperSockets functioned in the past.

If the function is enabled then multiple output data buffers are moved in one write operation. This reduces CPU utilization related to large outbound messages. When enabled, HiperSockets Multiple Write will be used anytime a message spans an IQD frame requiring multiple output data buffers (SBALs) to transfer the message. Spanning multiple output data buffers can be affected by a number of factors including:

- *IQD frame size*
- *Application socket send size*
- *TCP send size*
- *MTU size*

The HiperSockets Multiple Write Facility is supported in the z/OS environment.  For a complete description of the System z10 connectivity capabilities refer to IBM System z Connectivity Handbook, SG24-5444.

**HiperSockets Enhancement for zIIP Exploitation**
In z/OS V1.10, specifically, the z/OS Communications Server allows the HiperSockets Multiple Write Facility processing for outbound large messages originating from z/OS to be performed on a zIIP. The combination of HiperSockets Multiple Write Facility and zIIP enablement is described as "zIIP-Assisted HiperSockets for large messages." zIIP-Assisted HiperSockets can help make highly secure, available, virtual HiperSockets networking a more attractive option. z/OS application workloads based on XML, HTTP, SOAP, Java, etc., as well as traditional file transfer, can benefit from zIIP enablement by helping to lower general purpose processor utilization for such TCP/IP traffic.

Only outbound z/OS TCP/IP large messages which originate within a z/OS host are eligible for HiperSockets zIIP-Assisted processing. Other types of network traffic such as IP forwarding, Sysplex Distributor, inbound processing, small messages, or other non TCP/IP network protocols are not eligible for zIIP-Assisted HiperSockets. When the workload is eligible, then the TCP/IP HiperSockets device driver layer (write) processing is redirected to a zIIP, which will unblock the sending application. zIIP Assisted Hiper-Sockets for large messages is available with z/OS V1.10 with PTF and System z10 only. This feature is unsupported if z/OS is running as a guest in a z/VM environment and is supported for large outbound messages only.

To estimate potential offload, use PROJECTCPU for current and existing workloads. This is accurate and very simple, but you have to be on z/OS 1.10 with the enabling PTFs AND System z10 server AND you need to be performing HiperSockets Multiple Write workload already on z/OS.

## Security

Today's world mandates that your systems are secure and available 24/7. The z10 BC employs some of the most advanced security technologies in the industry—helping you to meet rigid regulatory requirements that include encryption solutions, access control management, and extensive auditing features. It also provides disaster recovery configurations and is designed to deliver 99.999% application availability to help avoid the downside of planned downtime, equipment failure, or the complete loss of a data center.

The z10 BC received Common Criteria Evaluation Assurance Level 5 (EAL5) certification May 4, 2009. EAL5 certification provides assurance that many disparate applications running in different operating environments in separate logical partitions on one z10 BC will be secure and distinct from each other. The System z10 servers, and previous IBM mainframes, are the world's only servers with the highest level of hardware security certification, Common Criteria Evaluation Assurance Level 5.

The System z10 servers represent the continuing commitment to the evolution of the mainframe while laying the foundation for future growth

System z security is one of the many reasons why the world's top banks and retailers rely on the IBM mainframe to help secure sensitive business transactions.

**z Can Do IT** securely.

## Cryptography

The z10 processor chip has on board cryptographic functions. Standard clear key integrated cryptographic coprocessors provide high speed cryptography for protecting data in storage. CP Assist for Cryptographic Function (CPACF) supports DES, TDES, Secure Hash Algorithms (SHA) for up to 512 bits, Advanced Encryption Standard (AES) for up to 256 bits and Pseudo Random Number Generation (PRNG).

The z10 BC includes both standard cryptographic hardware and optional cryptographic features for flexibility and growth capability. IBM has a long history of providing hardware cryptographic solutions, from the development of Data Encryption Standard (DES) in the 1970s to delivering integrated cryptographic hardware in a server to achieve the US Government's highest FIPS 140-2 Level 4 rating for secure cryptographic hardware.

The IBM System z10 BC cryptographic functions include the full range of cryptographic operations needed for e-business, e-commerce, and financial institution applications. In addition, custom cryptographic functions can be added to the set of functions that the z10 BC offers.

System z is investing in technology that provide improved performance for specialized functions. The Crypto Express2 and Crypto Express3 features for cryptography are examples. The Crypto Express2 and Crypto Express3 features can be configured as a secure key coprocessor or as accelerators, providing Secure Sockets Layer (SSL) acceleration. As a coprocessor, the feature includes support for Personal Account Numbers from 13 to 19 digits in length for stronger protection of data.

New integrated clear key encryption security features on z10 EC include support for longer keys with the Advanced Encryption Standard and more secure hashing algorithms.

Performing these functions in hardware is designed to contribute to improved performance. The newest support includes Protected Keys which provide stronger security for keys used in the CP Assist for Cryptographic Function (CPACF).

Enhancements to eliminate preplanning in the cryptography area include the System z10 function to dynamically add Crypto to a logical partition. Changes to image profiles, to support Crypto Express2 and Crypto Express3 features, are available without an outage to the logical partition. Crypto Express2 and Crypto Express3 features can also be dynamically deleted or moved.

Audit logging has been added to the new TKE workstation to enable better problem tracking and it also provides a new domain grouping function.

**CP Assist for Cryptographic Function (CPACF)**

The CP Assist for Cryptographic Function is available to every processor unit defined as a CP. It provides a set of symmetric cryptographic functions that enhance the encryption/decryption performance of clear-key operations for SSL, Virtual Private Network (VPN), and data storing applications not requiring a high level of security such as FIPS 140-2 level 4.

Cryptographic keys must be protected by the application system or operational procedures, as these keys are provided in the clear-key form to the CPACF. CPACF must be explicitly enabled, using a no-charge enablement feature. SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512 are shipped enabled on all servers with processor units (PUs) defined as CPs, IFLs, zIIPs, or zAAPs.

The CP Assist for Cryptographic Function offers:

For data privacy and confidentiality

- *Data Encryption Standard (DES)*
- *Triple Data Encryption Standard (TDES)*
- *Advanced Encryption Standard (AES) for 128-bit, 192-bit, and 256-bit keys*

For data integrity

- *Secure Hash Algorithms*
- *SHA-1: 160 bit*
- *SHA-2: 224, 256, 384, and 512 bit*

For message authentication codes (MAC)

- *Single-key MAC*
- *Double-key MAC*

**Protected key CPACF - blending clear-key and secure-key cryptography:**

The security of encryption relies upon keeping the value of the key a secret. A secure key should NEVER exist in the clear outside of the secure boundary of the card. If and when a secure key needs to exist outside of the tamper-resistant hardware it should be encrypted under another key, usually a master key.

An enhancement to Central Processor Assist to Cryptographic Function (CPACF) is designed to help facilitate the continued privacy of cryptographic key material when used by the CPACF for high-performance data encryption. Leveraging the unique z/Architecture, protected key CPACF helps to ensure that key material is not visible to applications or operating systems when used for encryption operations.

Protected key CPACF is designed to provide additional security for keys used on the CPACF without a significant performance impact. An enhancement to the information management tool, IBM Encryption Tool for IMS and DB2 Databases is expected to take advantage of this new level of key security.

**The newest-generation cryptographic feature - Crypto Express3:**
Crypto Express3 represents the newest-generation cryptographic feature designed to complement the cryptographic functions of CPACF. The Crypto Express3 resides in the I/O cage of the z10 EC and the I/O drawer of the z10 BC and continues to support all of the cryptographic functions available on Crypto Express2.

Crypto Express3 is a state-of-the-art, tamper-sensing and tamper-responding, programmable cryptographic feature. The cryptographic electronics and microprocessor provide a secure cryptographic environment using two PCI-Express (PCI-E) adapters. Dual processors operate in parallel to support the IBM Common Cryptographic Architecture (CCA) with high reliability.

**Crypto Express3 applications**
The Crypto Express3 feature is suited to applications requiring high-speed, security-sensitive, RSA acceleration, cryptographic operations for data encryption and digital signing, secure management, and use of cryptographic keys, or custom cryptographic applications. These can include financial applications such as PIN generation and verification in automated teller and point-of-sale (POS) transaction servers, remote key loading of ATMs and POS terminals, Web-serving applications, Public Key Infrastructure applications, smart card applications, and custom proprietary solutions. Applications can benefit from the strong security characteristics of the coprocessor and the opportunity to offload computationally intensive cryptographic processing.

**Crypto Express3-1P**
An option of one PCI-Express adapter per feature, in addition to the current two PCI-Express adapters per feature, is being offered for the z10 BC to help address small and midrange security requirements while maintaining high performance.

The Crypto Express3-1P feature with one PCI-Express adapter can continue to be defined as either a Coprocessor or an Accelerator. A minimum of two features must be ordered.

**Crypto Express3 key features**
Key features of Crypto Express3 include:
- *Dynamic power management to maximize RSA performance while keeping within temperature limits of the tamper-responding package*
- *For virtualization, the ability of 787 all logical partitions (LPARs) in all Logical Channel Subsystems (LCSSs) to access the the Crypto Express3 feature, up to 32 LPARs per feature*
- *Improved reliability, availability, and serviceability (RAS); even better than the excellent RAS offered for the Crypto Express2 feature*
- *Secure code loading that enables the updating of functionality while installed in application systems*
- *Lock-step checking of dual CPUs for enhanced error detection and fault isolation of cryptographic operations performed by a coprocessor when a PCI-E adapter is defined as a coprocessor*

- *Dynamic addition and configuration of cryptographic features to logical partitions without an outage*
- *Updated cryptographic algorithms used in loading the Licensed Internal Code (LIC) with the TKE workstation to keep in step with current recommendations for cryptographic strength*
- *Support for smart card applications using Europay, MasterCard Visa specifications*

Crypto Express3 is designed to provide Improved performance for symmetric and asymmetric operations.

**Crypto Express3 continues to support the following features:**
- *Cryptographic key generation*
  - *Pseudo Random Number Generation (PRNG)*
  - *Random Number Generation Long (RNGL) - 8 bytes to 8096 bytes*
- *Personal identification number (PIN) processing*
  - *PIN generation, verification, and translation functions*
  - Each Crypto Express3 feature can be configured as:
  - *Two PCI-E cryptographic coprocessors (default mode)*
  - *One PCI-E cryptographic coprocessor and one PCI-E cryptographic accelerator*
  - *Two PCI-E cryptographic accelerators*

**Crypto Express3 PCI-E adapter defined as a coprocessor:**

When one or both of the two PCI-E cryptographic adapters are configured as a coprocessor, the adapter can be used to:

- *Encrypt and decrypt data by utilizing secret-key algorithms. Algorithms supported for data confidentiality are:*
  - *Single-length key DES*
  - *Double-length key DES*
  - *Triple-length key DES*
  - *AES algorithms that use 128, 192, or 256-bit data-encrypting keys*
- *Generate, install, and distribute cryptographic keys securely using both public and secret key cryptographic methods*

- *Generate, verify, and translate personal identification numbers (PINs)*
- *Generate, verify, and translate 13- through 19-digit personal account numbers (PANs)*
- *Ensure the integrity of data by using message authentication codes (MACs), hashing algorithms, and Rivest-Shamir-Adelman (RSA) public key algorithm (PKA) digital signatures*
- *Perform financial PIN processing and other specialized banking functions*
- *Manage DES, TDES, AES, and RSA keys*
- *Offer highly secure encryption processing, use of secure encrypted key values, and User Defined Extensions (UDX) to CCA*
- *Provide secure remote key loading of encryption keys to ATMs, point of sale terminals (POS), and PIN entry devices*
- *Exchange cryptographic keys between IBM CCA and non-CCA servers*
- *Generate high-quality random numbers for keys and other cryptographic applications*

**Crypto Express3 Accelerator:**

The Crypto Express3 accelerator is configured from the HMC so that it uses only a subset of the coprocessor functions at a higher speed. When one or both of the two PCI-E cryptographic adapters are configured as an accelerator, the Crypto Express3 feature can be used for:

- *High-performance clear-key RSA functions*
- *Acceleration of modular arithmetic operations, the RSA cryptographic operations used with the SSL/TLS protocol*
- *Offloading of compute-intensive RSA public-key and private-key cryptographic operations employed in the SSL protocol*

Supported functions include:

- *PKA Decrypt (CSNDPKD), with PKCS-1.2 formatting*
- *PKA Encrypt (CSNDPKE), with zero-pad formatting*
- *Digital Signature Verify (CSNDDSV)*

The RSA encryption and decryption functions support key lengths of 512 bits to 4,096 bits, in the Modulus Exponent (ME) and Chinese Remainder Theorem (CRT) formats.

More information on IBM System z10 cryptographic performance can be found in the IBM System z Security Web site at www.**ibm.com**/systems/z/advantages/security/z10cryptography.html

**Configurable Crypto Express2**

The Crypto Express2 feature has two PCI-X adapters and the Crypto Express2-1P has a single PCI-X adapter. Each of the PCI-X adapters can be defined as either a Coprocessor or an Accelerator.

Crypto Express2 Coprocessor – for secure-key encrypted transactions (default) is:

- *Designed to support security-rich cryptographic functions, use of secure-encrypted-key values, and User Defined Extensions (UDX)*
- *Designed to support secure and clear-key RSA operations*
- *The tamper-responding hardware and lower-level firmware layers are validated to U.S. Government FIPS 140-2 standard: Security Requirements for Cryptographic Modules at Level 4*

Crypto Express2 Accelerator – for Secure Sockets Layer (SSL) acceleration:

- *Offloads compute-intensive RSA public-key and private-key cryptographic operations employed in the SSL protocol Crypto Express2 features can be carried forward on an upgrade to the System z10 BC, so users may continue to take advantage of the SSL performance and the configuration capability*

The configurable Crypto Express2 feature is supported by z/OS, z/VM, z/VSE, and Linux on System z. z/VSE offers support for clear-key operations only. Current versions of z/OS, z/VM, and Linux on System z offer support for both clear-key and secure-key operations.

**Crypto Express2-1P**

An option of one PCI-X adapter per feature, in addition to the current two PCI-X adapters per feature, is being offered for the z10 BC to help satisfy small and midrange security requirements while maintaining high performance.

The Crypto Express2-1P feature, with one PCI-X adapter, can continue to be defined as either a Coprocessor or an Accelerator. A minimum of two features must be ordered.

Additional cryptographic functions and features with Crypto Express2 and Crypto Express2-1P.

Key management – Added key management for remote loading of ATM and Point of Sale (POS) keys. The elimination of manual key entry is designed to reduce downtime due to key entry errors, service calls, and key management costs.

Improved key exchange – Added Improved key exchange with non-CCA cryptographic systems. New features added to IBM Common Cryptographic Architecture (CCA) are designed to enhance the ability to exchange keys between CCA systems, and systems that do not use control vectors by allowing the CCA system owner to define permitted types of key import and export while preventing uncontrolled key exchange that can open the system to an increased threat of attack.

These are supported by z/OS and by z/VM for guest exploitation.

**Support for ISO 16609**

Support for ISO 16609 CBC Mode T-DES Message Authentication (MAC) requirements ISO 16609 CBC Mode T-DES MAC is accessible through ICSF function calls made in the PCI-X Cryptographic Adapter segment 3 Common Cryptographic Architecture (CCA) code.

**Support for RSA keys up to 4096 bits**

The RSA services in the CCA API are extended to support RSA keys with modulus lengths up to 4096 bits. The services affected include key generation, RSA-based key management, digital signatures, and other functions related to these.

Refer to the ICSF Application Programmers Guide, SA22-7522, for additional details.

**Cryptographic enhancements to Crypto Express2 and Crypto Express2-1P**

**Dynamically add crypto to a logical partition.**

Today, users can preplan the addition of Crypto Express2 features to a logical partition (LP) by using the Crypto page in the image profile to define the Cryptographic Candidate List, Cryptographic Online List, and Usage and Control Domain Indexes in advance of crypto hardware installation.

With the change to dynamically add crypto to a logical partition, changes to image profiles, to support Crypto Express2 features, are available without outage to the logical partition. Users can also dynamically delete or move Crypto Express2 features. Preplanning is no longer required.

This enhancement is supported by z/OS, z/VM, z/VSE, and Linux on System z.

**Secure Key AES**

The Advanced Encryption Standard (AES) is a National Institute of Standards and Technology specification for the encryption of electronic data. It is expected to become the accepted means of encrypting digital information, including financial, telecommunications, and government data.

AES is the symmetric algorithm of choice for new applications, instead of Data Encryption Standard (DES) or Triple-DES, for the encryption and decryption of data. The AES encryption algorithm will be supported with secure (encrypted) keys of 128, 192, and 256 bits. The secure key approach, similar to what is supported today for DES and TDES, provides the ability to keep the encryption keys protected at all times, including the ability to import and export AES keys, using RSA public key technology.

Support for AES encryption algorithm includes the master key management functions required to load or generate AES master keys, update those keys, and re-encipher key tokens under a new master key.

**Support for 13- through 19-digit Personal Account Numbers**

Credit card companies sometimes perform card security code computations based on Personal Account Number (PAN) data. Currently, ICSF callable services CSNBCSV (VISA CVV Service Verify) and CSNBCSG (VISA CVV Service Generate) are used to verify and to generate a VISA Card Verification Value (CVV) or a MasterCard Card Verification Code (CVC). The ICSF callable services currently support 13-, 16-, and 19-digit PAN data. To provide additional flexibility, new keywords PAN-14, PAN-15, PAN-17, and PAN-18 are implemented in the rule array for both CSNBCSG and CSNBCSV to indicate that the PAN data is comprised of 14, 15, 17, or 18 PAN digits, respectively.

Support for 13- through 19-digit PANs is exclusive to System z10 and z9 and is offered by z/OS and z/VM for guest exploitation.

**Stronger cryptography encryption for TKE protocols inbound /outbound authentication**

TKE uses cryptographic algorithms and protocols in communication with the target cryptographic adapters in the host systems it administers. Cryptography is first used to verify that each target adapter is a valid IBM cryptographic coprocessor. It then ensures there are secure messages between the TKE workstation and the target Crypto Express2 and Crypto Express3 feature.

The cryptography has been updated to keep pace with industry developments and with recommendations from experts and standards organizations.

The following enhancements have been made:

- *TKE Certificate Authorities (CAs) initialized on a TKE workstation with TKE 6.0 LIC can issue certificates with 2048-bit keys. Previous versions of TKE used 1024-bit keys.*
- *The transport key used to encrypt sensitive data sent between the TKE workstation and a Crypto Express3 coprocessor has been strengthened from a 192-bit TDES key to a 256-bit AES key.*
- *The signature key used by the TKE workstation and the Crypto Express3 coprocessor has been strengthened from 1024-bit to a maximum of 4096-bit strength.*
- *Replies sent by a Crypto Express3 coprocessor on the host are signed with a 4096-bit key.*

**TKE smart card support:**

TKE 6.0 contains support to increase the key strength for TKE Certificate Authority (CA) smart cards, TKE smart cards, and signature keys stored on smart cards from 1024-bit to 2048-bit strength.

Only feature number 0884 smart cards with the feature number 0885 smart card reader support the creation of TKE Certificate Authority (CA) smart cards, TKE smart cards, or signature keys with the new 2048-bit key

strength. Existing feature #0888 smart cards and #0887 smart card readers are limited to 1024-bit key strengths.

**Simplified usability with Crypto Express3 migration wizard:**

A wizard is now available to allow users to collect configuration data from a Crypto Express2 and Crypto Express3 coprocessor and migrate the data to a different Crypto Express coprocessor. The target Crypto Express coprocessor must have the same or greater capabilities.

Benefits of using this wizard include:

- *Reduces migration steps, thereby minimizing user errors*
- *Minimizes the number of user "clicks"*
- *Significantly reduces migration task duration*

**Usability enhancements with TKE 6.0:**

Trusted Key Entry (TKE) 6.0 Licensed Internal Code (LIC) includes domain grouping, a significant usability enhancement. The TKE 6.0 LIC provides capabilities for:

- *Grouping of up to 16 domains across one or more cryptographic adapters. These adapters may be installed in one or more servers or LPARs. Grouping of domains applies to Crypto Express3 and Crypto Express2 features.*
- *Greater flexibility and efficiency by executing domain-scoped commands on every domain in the group. For example, a TKE user can load master key parts to all domains with one command.*
- *Efficiency by executing Crypto Express2 and Crypto Express3 scoped commands on every coprocessor in the group. This allows a substantial reduction of the time required for loading new master keys from a TKE workstation into a Crypto Express3 or Crypto Express2 feature.*

TKE 6.0 can be attached to a z9 or a z10 processor and can manage cryptographic domains across either machine type, however a CEX3 is required to take advantage of the strongest algorithms.

## TKE 5.3

TKE 5.3 workstation can continue to be used with the z9 and z10, however TKE 5.3 cannot take advantage of the stronger algorithms available on the Crypto Express3. The Trusted Key Entry (TKE) workstation and the TKE 5.3 level of Licensed Internal Code are optional features on the System z10 BC. The TKE 5.3 Licensed Internal Code (LIC) is loaded on the TKE workstation prior to shipment. The TKE workstation offers security-rich local and remote key management, providing authorized persons a method of operational and master key entry, identification, exchange, separation, and update. The TKE workstation supports connectivity to an Ethernet Local Area Network (LAN) operating at 10 or 100 Mbps. Up to ten TKE workstations can be ordered.

The TKE 5.3 level of LIC includes support for the AES encryption algorithm, adds 256-bit master keys, and includes the master key management functions required to load or generate AES master keys to cryptographic coprocessors in the host.

Also included is an imbedded screen capture utility to permit users to create and to transfer TKE master key entry instructions to diskette or DVD. Under 'Service Management' a "Manage Print Screen Files" utility will be available to all users.

The TKE workstation and TKE 5.3 LIC  are available on the z10 EC, z10 BC, z9 EC, and z9 BC.

TKE 5.3 LIC has added the capability to store key parts on DVD-RAMs and continues to support the ability to store key parts on paper, or optionally on a smart card. TKE 5.3 LIC has limited the use of floppy diskettes to read-only. The TKE 5.3 LIC can remotely control host cryptographic coprocessors using a password-protected authority signature key pair either in a binary file or on a smart card.

## System z10 BC cryptographic migration

Clients using a User Defined Extension (UDX) of the Common Cryptographic Architecture should contact their UDX provider for an application upgrade before ordering a new System z10 BC machine; or before planning to migrate or activate a UDX application to firmware driver level 73 and higher.

- *The Crypto Express2 feature is supported on the z9 BC and can be carried forward on an upgrade to the System z10 BC*
- *You may continue to use TKE workstations with 5.3 licensed internal code to control the System z10 BC*
- *TKE 5.0 and 5.1 workstations may be used to control z9 EC, z9 BC, z890 and z990 servers*

## Remote Loading of Initial ATM Keys

Typically, a new ATM has none of the financial institution's keys installed. Remote Key Loading refers to the process of loading Data Encryption Standard (DES) keys to Automated Teller Machines (ATMs) from a central administrative site without the need for personnel to visit each machine to manually load DES keys. This has been done by manually loading each of the two clear text key parts individually and separately into ATMs. Manual entry of keys is one of the most error-prone and labor-intensive activities that occur during an installation, making it expensive for the banks and financial institutions.

## Remote Key Loading Benefits
- *Provides a mechanism to load initial ATM keys without the need to send technical staff to ATMs*
- *Reduces downtime due to key entry errors*
- *Reduces service call and key management costs*
- *Improves the ability to manage ATM conversions and upgrades*

Integrated Cryptographic Service Facility (ICSF), together with Crypto Express2, support the basic mechanisms in Remote Key Loading. The implementation offers a secure bridge between the highly secure Common Cryptographic Architecture (CCA) environment and the various formats and encryption schemes offered by the ATM vendors. The following ICSF services are offered for Remote Key loading:

- *Trusted Block Create (CSNDTBC): This callable service is used to create a trusted block containing a public key and some processing rules*
- *Remote Key Export (CSNDRKX): This callable service uses the trusted block to generate or export DES keys for local use and for distribution to an ATM or other remote device*

Refer to Application Programmers Guide, SA22-7522, for additional details.

**Improved Key Exchange With Non-CCA Cryptographic Systems**
IBM Common Cryptographic Architecture (CCA) employs Control Vectors to control usage of cryptographic keys. Non-CCA systems use other mechanisms, or may use keys that have no associated control information. This enhancement provides the ability to exchange keys between CCA systems, and systems that do not use Control Vectors. Additionally, it allows the CCA system owner to define permitted types of key import and export which can help to prevent uncontrolled key exchange that can open the system to an increased threat of attack.

These enhancements are exclusive to System z10, and System z9 and are supported by z/OS and z/VM for z/OS guest exploitation.

It may sound revolutionary, but it's really quite simple. In the highly unpredictable world of On Demand business, you should get what you need, when you need it. And you should pay for only what you use. Radical? Not to IBM. It's the basic principle underlying IBM capacity on demand for the IBM System z10.

The z10 BC also introduces a architectural approach for temporary offerings that can change the thinking about on demand capacity. One or more flexible configuration definitions can be used to solve multiple temporary situations and multiple capacity configurations can be active at once (for example, activation of just two CBUs out of a definition that has four CBUs is acceptable). This means that On/Off CoD can be active and up to seven other offerings can be active simultaneously. Tokens can be purchased for On/Off CoD so hardware activations can be prepaid.

All activations can be done without having to interact with IBM—when it is determined that capacity is required, no passwords or phone connections are necessary. As long as the total z10 BC can support the maximums that are defined, then they can be made available. With the z10 BC it is now possible to add permanent capacity while a temporary capacity is currently activated, without having to return first to the original configuration.

Capacity on Demand – Temporary Capacity

The set of contract documents which support the various Capacity on Demand offerings available for z10 BC has been completely refreshed. While customers with existing contracts for Capacity Back Up (CBU) and Customer Initiated Upgrade (CIU) – On/Off Capacity on Demand (On/Off CoD) may carry those contracts forward to z10 BC machines, new CoD capability and offerings for z10 BC is only supported by this new contract set.

The new contract set is structured in a modular, hierarchical approach. This new approach will eliminate redundant terms between contract documents, simplifying the contracts for our customers and IBM.

Just-in-time deployment of System z10 BC Capacity on Demand (CoD) is a radical departure from previous System z and zSeries servers. This new architecture allows:

- Up to eight temporary records to be installed on the CPC and active at any given time
- Up to 200 temporary records to be staged on the SE
- Variability in the amount of resources that can be activated per record
- The ability to control and update records independent of each other
- Improved query functions to monitor the state of each record
- The ability to add capabilities to individual records concurrently, eliminating the need for constant ordering of new temporary records for different user scenarios
- Permanent LIC-CC upgrades to be performed while temporary resources are active

These capabilities allow you to access and manage processing capacity on a temporary basis, providing increased flexibility for on demand environments. The CoD offerings are built from a common Licensed Internal Code – Configuration Code (LIC-CC) record structure. These Temporary Entitlement Records (TERs) contain the information necessary to control which type of resource can be accessed and to what extent, how many times and for how long, and under what condition – test or real workload. Use of this information gives the different offerings their personality.

**Capacity Back Up (CBU):** Temporary access to dormant processing units (PUs), intended to replace capacity lost within the enterprise due to a disaster. CP capacity or any and all specialty engine types (zIIP, zAAP, SAP, IFL, ICF) can be added up to what the physical hardware model can contain for up to 10 days for a test activation or 90 days for a true disaster recovery.

On system z10 the CBU entitlement records contain an expiration date that is established at the time of order and is dependent upon the quantity of CBU years. You will now have the capability to extend your CBU entitlements through the purchase of additional CBU years. The number of CBU years per instance of CBU entitlement remains limited to five and fractional years are rounded up to the near whole integer when calculating this limit. For instance, if there are two years and eight months to the expiration date at the time of order, the expiration date can be extended by no more than two additional years. One test activation is provided for each additional CBU year added to the CBU entitlement record.

**CBU Tests:** The allocation of the default number of test activations changed.  Rather than a fixed default number of five test activations for each CBU entitlement record, the number of test activations per instance of the CBU entitlement record will coincide with the number of CBU years, the number of years assigned to the CBU record. This equates to one test activation per year for each CBU entitlement purchased.Additional test activations are now available in quantities of one and the number of test activations remains limited at 15 per CBU entitlement record.

These changes apply only to System z10 and to CBU entitlements purchased through the IBM sales channel or directly from Resource Link.

There are terms governing System z Capacity Back Up (CBU) now available which allow customers to execute production workload on a CBU Upgrade during a CBU Test.

While all new CBU contract documents contain the new CBU Test terms, existing CBU customers will need to execute a contract to expand their authorization for CBU Test upgrades if they want to have the right to execute production workload on the CBU Upgrade during a CBU Test.

**Amendment for CBU Tests**

The modification of CBU Test terms is available for existing CBU customers via the IBM Customer Agreement Amendment for IBM System z Capacity Backup Upgrade Tests (in the US this is form number Z125-8145). This amendment can be executed at any time, and separate from any particular order.

**Capacity for Planned Event (CPE):** Temporary access to dormant PUs, intended to replace capacity lost within the enterprise due to a planned event such as a facility upgrade or system relocation. CPE is similar to CBU in that it is intended to replace lost capacity; however, it differs in its scope and intent. Where CBU addresses disaster recovery scenarios that can take up to three months to remedy, CPE is intended for short-duration events lasting up to three days, maximum. Each CPE record, once activated, gives you access to all dormant PUs on the machine that can be configured in any combination of CP capacity or specialty engine types (zIIP, zAAP, SAP, IFL, ICF).

CPE is changing with this announcement. Effective December 31, 2009 CPE will allow you to select the capacity to meet your business needs rather than providing temporary access to all dormant capacity.

**On/Off Capacity on Demand (On/Off CoD):** Temporary access to dormant PUs, intended to augment the existing capacity of a given system. On/Off CoD helps you contain workload spikes that may exceed permanent capacity such that Service Level Agreements cannot be met and business conditions do not justify a permanent upgrade. An On/Off CoD record allows you to temporarily add CP capacity or any and all specialty engine types (zIIP, zAAP, SAP, IFL, ICF) up to the following limits:

- *The quantity of temporary CP capacity ordered is limited by the quantity of purchased CP capacity (permanently active plus unassigned)*
- *The quantity of temporary IFLs ordered is limited by quantity of purchased IFLs (permanently active plus unassigned)*
- *Temporary use of unassigned CP capacity or unassigned IFLs will not incur a hardware charge*
- *The quantity of permanent zIIPs plus temporary zIIPs can not exceed the quantity of purchased (permanent plus unassigned) CPs plus temporary CPs and the quantity of temporary zIIPs can not exceed the quantity of permanent zIIPs*
- *The quantity of permanent zAAPs plus temporary zAAPs can not exceed the quantity of purchased (permanent plus unassigned) CPs plus temporary CPs and the quantity of temporary zAAPs can not exceed the quantity of permanent zAAPs*
- *The quantity of temporary ICFs ordered is limited by the quantity of permanent ICFs as long as the sum of permanent and temporary ICFs is less than or equal to 16*
- *The quantity of temporary SAPs ordered is limited by the quantity of permanent SAPs as long as the sum of permanent and temporary SAPs is less than or equal to 32*

Although the System z10 BC will allow up to eight temporary records of any type to be installed, only one temporary On/Off CoD record may be active at any given time. An On/Off CoD record may be active while other temporary records are active.

Management of temporary capacity through On/Off CoD is further enhanced through the introduction of resource tokens. For CP capacity, a resource token represents an amount of processing capacity that will result in one MSU of SW cost for one day – an MSU-day. For specialty engines, a resource token represents activation of one engine of that type for one day – an IFL-day, a zIIP-day or a zAAP-day. The different resource tokens are contained in separate pools within the On/Off CoD record. The customer, via the Resource Link ordering process, determines how many tokens go into each pool. Once On/Off CoD resources are activated, tokens will be decremented from their pools every 24 hours. The amount decremented is based on the highest activation level for that engine type during the previous 24 hours.

Resource tokens are intended to help customers bound the hardware costs associated with using On/Off CoD. The use of resource tokens is optional and they are available on either a prepaid or post-paid basis. When prepaid, the customer is billed for the total amount of resource tokens contained within the On/Off CoD record. When post-paid, the total billing against the On/Off Cod record is limited by the total amount of resource tokens contained within the record. Resource Link will provide the customer an ordering wizard to help determine how many tokens they need to purchase for different activation scenarios. Resource tokens within an On/Off CoD record may also be replenished.

Resource Link offers an ordering wizard to help determine how many tokens you need to purchase for different activation scenarios. Resource tokens within an On/Off CoD record may also be replenished. For more information on the use and ordering of resource tokens, refer to the Capacity on Demand Users Guide, SC28-6871.

**Capacity Provisioning**

Hardware working with software is critical. The activation of On/Off CoD on z10 EC can be simplified or automated by using z/OS Capacity Provisioning (available with z/OS V1.10 and z/OS V1.9). This capability enables the monitoring of multiple systems based on Capacity Provisioning and Workload Manager (WLM) definitions. When the defined conditions are met, z/OS can suggest capacity changes for manual activation from a z/OS console or the system can add or remove temporary capacity automatically and without operator intervention. z10 BC Can Do IT better.

z/OS Capacity provisioning allows you to set up rules defining the circumstances under which additional capacity should be provisioned in order to fulfill a specific business need. The rules are based on criteria, such as: a specific application, the maximum additional capacity that should be activated, time and workload conditions. This support provides a fast response to capacity changes and ensures sufficient processing power will be available with the least possible delay even if workloads fluctuate.

An installed On/Off CoD record is a necessary prerequisite for automated control of temporary capacity through z/OS Capacity Provisioning.

See z/OS MVS Capacity Provisioning User's Guide (SA33-8299) for more information.

**On/Off CoD Test:** On/Off CoD allows for a no-charge test. No IBM charges are assessed for the test, including IBM charges associated with temporary hardware capacity, IBM software, or IBM maintenance. This test can be used to validate the processes to download, stage, install, activate, and deactivate On/Off CoD capacity non-disruptively. Each On/Off CoD-enabled server is entitled to only one no-charge test. This test may last up to a maximum duration of 24 hours commencing upon the activation of any capacity resources contained in the On/Off CoD record. Activation levels of capacity may change during the 24 hour test period. The On/Off CoD test automatically terminates at the end of the 24 hours period. In addition to validating the On/Off CoD function within your environment, you may choose to use this test as a training session for your personnel who are authorized to activate On/Off CoD.

SNMP API (Simple Network Management Protocol Application Programming Interface) enhancements have also been made for the new Capacity On Demand features. More information can be found in the System z10 Capacity On Demand User's Guide, SC28-6871.

**Capacity on Demand – Permanent Capacity**

**Customer Initiated Upgrade (CIU) facility:** When your business needs additional capacity quickly, Customer Initiated Upgrade (CIU) is designed to deliver it. CIU is designed to allow you to respond to sudden increased capacity requirements by requesting a System z10 BC PU and/or memory upgrade via the Web, using IBM Resource Link, and downloading and applying it to your System z10 BC server using your system's Remote Support connection. Further, with the Express option on CIU, an upgrade may be made available for installation as fast as within a few hours after order submission.

**Permanent upgrades:** Orders (MESs) of all PU types and memory for System z10 BC servers that can be delivered by Licensed Internal Code, Control Code (LIC-CC) are eligible for CIU delivery. CIU upgrades may be performed up to the maximum available processor and memory resources on the installed server, as configured. While capacity upgrades to the server itself are concurrent, your software may not be able to take advantage of the increased capacity without performing an Initial Programming Load (IPL).

| | System z9 | System z10 |
|---|---|---|
| Resources | CP, zIIP, zAAP, IFL, ICF | CP, zIIP, zAAP, IFL, ICF, SAP |
| Offerings | Requires access to IBM/ RETAIN® to activate | No password required or access to IBM/RETAIN to activate |
| | CBU, On/Off CoD One offering at a time | CBU, On/Off CoD, CPE Multiple offerings active |
| Permanent upgrades | Requires de-provisioning of temporary capacity first | Concurrent with temporary offerings |
| Replenishment | No | Yes w/ CBU & On/Off CoD |
| CBU Tests | 5 tests per record | Up to 15 per record |
| CBU Expiration | No expiration | Specific term length |
| Capacity Provisioning Manager Support | No | Yes |

## Reliability, Availability, and Serviceability (RAS)

In today's on demand environment, downtime is not only unwelcome—it's costly. If your applications aren't consistently available, your business suffers. The damage can extend well beyond the financial realm into key areas of customer loyalty, market competitiveness and regulatory compliance. High on the list of critical business requirements today is the need to keep applications up and running in the event of planned or unplanned disruptions to your systems.

While some servers are thought of offering weeks or even months of up time, System z thinks of this in terms of achieving years. The z10 BC continues our commitment to deliver improvements in hardware Reliability, Availability and Serviceability (RAS) with every new System z server. They include microcode driver enhancements, dynamic segment sparing for memory and fixed HSA, as well as a new I/O drawer design. The z10 BC is a server that can help keep applications up and running in the event of planned or unplanned disruptions to the system.

The System z10 BC is designed to deliver industry leading reliability, availability and security our customers have come to expect from System z servers. System z10 BC RAS is designed to reduce all sources of outages by reducing unscheduled, scheduled and planned outages. Planned outages are further designed to be reduced with the introduction of concurrent I/O drawer add and eliminating pre-planning requirements. These features are designed to reduce the need for a Power-on-Reset (POR) and help eliminate the need to deactivate/activate/IPL a logical partition.

**RAS Design Focus**

**High Availability (HA)** – The attribute of a system designed to provide service during defined periods, at acceptable or agreed upon levels and masks UNPLANNED OUTAGES from end users. It employs fault tolerance, automated failure detection, recovery, bypass reconfiguration, testing, problem and change management.

**Continuous Operations (CO)** – The attribute of a system designed to continuously operate and mask PLANNED OUTAGES from end users. It employs non-disruptive hardware and software changes, non-disruptive configuration and software coexistence.

**Continuous Availability (CA)** – The attribute of a system designed to deliver non-disruptive service to the end user 7 days a week, 24 HOURS A DAY (there are no planned or unplanned outages). It includes the ability to recover from a site disaster by switching computing to a second site.

## *Availability Functions*

With the z10 BC, significant steps have been taken in the area of server availability with a focus on reducing pre-planning requirements. Pre-planning requirements are minimized by delivering and reserving 8 GB for HSA so the maximum configuration capabilities can be exploited. Features that carry forward from previous generation processors include the ability to dynamically enable I/O, and the dynamic swapping of processor types.

### Hardware System Area (HSA)

Fixed HSA of 8 GB is provided as standard with the z10 BC. The HSA has been designed to eliminate planning for HSA and makes all the memory purchased by customers available for customer use. Preplanning for HSA expansion for configurations will be eliminated as HCD/IOCP will, via the IOCDS process, always reserve:

- *2 Logical Channel Subsystems (LCSS), pre-defined*
- *30 Logical Partitions (LPARs), pre-defined*
- *Subchannel set 0 with 63.75k devices*
- *Subchannel set 1 with 64K-1 devices*
- *Dynamic I/O Reconfiguration – always enabled by default*
- *Add/Change the number of logical CP, IFL, ICF, zAAP, zIIP, processors per partition and add SAPs to the configuration*
- *Dynamic LPAR PU assignment optimization CPs, ICFs, IFLs, zAAPs, zIIPs, SAPs*
- *Dynamically Add/Remove Crypto (no LPAR deactivation required)*

### Redundant I/O Interconnect

In the event of a failure or customer initiated action such as the replacement of an HCA/STI fanout card, the z10 BC is designed to provide access to your I/O devices through another HCA/STI to the affected I/O domains. This is exclusive to System z10 and System z9.

### Enhanced Driver Maintenance

One of the greatest contributors to downtime during planned outages is Licensed Internal Code (LIC) updates. When properly configured, z10 BC is designed to permit select planned LIC updates.

A new query function has been added to validate LIC EDM requirements in advance. Enhanced programmatic internal controls have been added to help eliminate manual analysis by the service team of certain exception conditions.

With the z10 BC, PR/SM code has been enhanced to allow multiple EDM 'From' sync points. Automatic apply of EDM licensed internal change requirements is now limited to EDM and the licensed internal code changes update process.

There are several reliability, availability, and serviceability (RAS) enhancements that have been made to the HMC/SE based on the feedback from the System z9 Enhanced Driver Maintenance field experience.

- *Change to better handle intermittent customer network issues*
- *EDM performance improvements*
- *New EDM user interface features to allow for customer and service personnel to better plan for the EDM*
- *A new option to check all licensed internal code which can be executed in advance of the EDM preload or activate.*

**Dynamic Oscillator Switchover**

The z10 BC has two oscillator cards, a primary and a backup. For most cases, should a failure occur on the primary oscillator card, the backup can detect it, switch over, and provide the clock signal to the system transparently, with no system outage. Previously, in the event of a failure of the active oscillator, a system outage would occur, the subsequent system Power On Reset (POR) would select the backup, and the system would resume operation. Dynamic Oscillator Switchover is exclusive to System z10 and System z9.

**Transparent Sparing**

The z10 BC offers 12 PUs, two are designated as System Assist Processors (SAPs). In the event of processor failure, if there are spare processor units available (undefined), these PUs are used for transparent sparing.

**Concurrent Memory Upgrade**

Memory can be upgraded concurrently using LIC-CC if physical memory is available on the machine either through the Plan Ahead Memory feature or by having more physical memory installed in the machine that has not been activated.

**Plan Ahead Memory**

Future memory upgrades can now be preplanned to be nondisruptive. The preplanned memory feature will add the necessary physical memory required to support target memory sizes. The granularity of physical memory in the System z10 design is more closely associated with the granularity of logical, entitled memory, leaving little room for growth. If you anticipate an increase in memory requirements, a "target" logical memory size can now be specified in the configuration tool along with a "starting" logical memory size. The configuration tool will then calculate the physical memory required to satisfy this target memory. Should additional physical memory be required, it will be fulfilled with the preplanned memory features.

The preplanned memory feature is offered in 4 gigabyte (GB) increments. The quantity assigned by the configuration tool is the number of 4 GB blocks necessary to increase the physical memory from that required for the "starting" logical memory to the physical memory required for the "target" logical configuration. Activation of any preplanned memory requires the purchase of preplanned memory activation features. One preplanned memory activation feature is required for each preplanned memory feature. You now have the flexibility to activate memory to any logical size offered between the starting and target size.

**Service Enhancements**

z10 BC service enhancements designed to avoid scheduled outages include:

- *Concurrent firmware fixes*
- *Concurrent parts replacement*
- *Concurrent hardware upgrades*
- *DIMM FRU indicators*
- *Single processor core checkstop*
- *Single processor core sparing*
- *Rebalance PSIFB and I/O Fanouts*
- *Redundant 100 Mb Ethernet service network with VLAN*

## *Environmental Enhancements*

Power and cooling discussions have entered the budget planning of every IT environment. As energy prices have risen and utilities have restricted the amount of power usage, it is important to review the role of the server in balancing IT spending.

### Power Monitoring
The "mainframe gas gauge" feature introduced on the System z9 servers, provides power and thermal information via the System Activity Display (SAD) on the Hardware Management Console and will be available on the z10 BC giving a point in time reference of the information. The current total power consumption in watts and BTU/hour as well as the air input temperature will be displayed.

### Power Estimation Tool
To assist in energy planning, Resource Link provides tools to estimate server energy requirements before a new server purchase. A user will input the machine model, memory, and I/O configuration and the tool will output an estimate of the system total heat load and utility input power. A customized planning aid is also available on Resource Link which provides physical characteristics of the machine along with cooling recommendations, environmental specifications, system power rating, power plugs/receptacles, line cord wire specifications and the machine configuration.

**IBM Systems Director Active Energy Manager**
IBM Systems Director Active Energy Manager™ (AEM) is a building block which enables customers to manage actual power consumption and resulting thermal loads IBM servers place in the data center. The z10 BC provides support for IBM Systems Director Active Energy Manager (AEM) for Linux on System z for a single view of actual energy usage across multiple heterogeneous IBM platforms within the infrastructure. AEM for Linux on System z will allow tracking of trends for both the z10 BC as well as multiple server platforms. With this trend analysis, a data center administrator will have the data to help properly estimate power inputs and more accurately plan data center consolidation or modification projects.

On System z10, the HMC will now provide support for the Active Energy Manager (AEM) which will display power consumption/air input temperature as well as exhaust temperature. AEM will also provide some limited status/ configuration information which might assist in explaining changes to the power consumption. AEM is exclusive to System z10

## *Parallel Sysplex Cluster Technology*

IBM System z servers stand alone against competition and have stood the test of time with our business resiliency solutions. Our coupling solutions with Parallel Sysplex technology allow for greater scalability and availability.

Parallel Sysplex clustering is designed to bring the power of parallel processing to business-critical System z10, System z9, z990 or z890 applications. A Parallel Sysplex cluster consists of up to 32 z/OS images coupled to one or more Coupling Facilities (CFs or ICFs) using high-speed specialized links for communication. The Coupling Facilities, at the heart of the Parallel Sysplex cluster, enable high speed, read/write data sharing and resource sharing among all the z/OS images in a cluster. All images are also connected to a Sysplex Timer® or by implementing the Server Time Protocol (STP), so that all events can be properly sequenced in time.
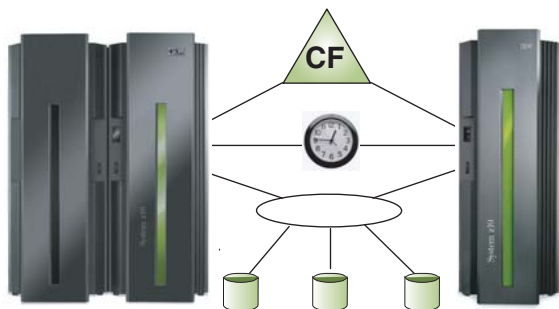


Parallel Sysplex Resource Sharing enables multiple system resources to be managed as a single logical resource shared among all of the images. Some examples of resource sharing include JES2 Checkpoint, GRS "star," and Enhanced Catalog Sharing; all of which provide simplified systems management, increased performance and/or scalability.

Although there is significant value in a single footprint and multi-footprint environment with resource sharing, those customers looking for high availability must move on to a database data sharing configuration. With the Parallel Sysplex environment, combined with the Workload Manager and CICS TS, DB2 or IMS, incoming work can be dynamically routed to the z/OS image most capable of handling the work. This dynamic workload balancing, along with the capability to have read/write access data from anywhere in the Parallel Sysplex cluster, provides scalability and availability. When configured properly, a Parallel Sysplex cluster is designed with no single point of failure and can provide customers with near continuous application availability over planned and unplanned outages.

With the introduction of the z10 EC, we have the concept of n-2 on the hardware as well as the software. The z10 BC participates in a Sysplex with System z10 EC, System z9, z990 and z890 only and currently supports z/OS 1.8 and higher and z/VM 5.2 for a guest virtualization coupling facility test environment.

For detailed information on IBM's Parallel Sysplex technology, visit our Parallel Sysplex home page at http://www.**ibm.com**/systems/z/pso/

**Coupling Facility Control Code (CFCC) Level 16**
CFCC Level 16 is available on the IBM System z10 BC.

**Improved service time with Coupling Facility Duplexing enhancements:** Prior to Coupling Facility Control Code (CFCC) Level 16, System-Managed Coupling Facility (CF) Structure Duplexing required two duplexing protocol exchanges to occur synchronously during processing of

each duplexed structure request. CFCC Level 16 allows one of these protocol exchanges to complete asynchronously. This allows faster duplexed request service time, with more benefits when the Coupling Facilities are further apart, such as in a multi-site Parallel Sysplex environment.

**List notification improvements:** Prior to CFCC Level 16, when a shared queue (subsidiary list) changed state from empty to non-empty, the CF would notify ALL active connectors. The first one to respond would process the new message, but when the others tried to do the same, they would find nothing, incurring additional overhead.

CFCC Level 16 can help improve the efficiency of coupling communications for IMS Shared Queue and WebSphere MQ Shared Queue environments. The Coupling Facility notifies only one connector in a sequential fashion. If the shared queue is processed within a fixed period of time, the other connectors do not need to be notified, saving the cost of the false scheduling. If a shared queue is not read within the time limit, then the other connectors are notified as they were prior to CFCC Level 16.

When migrating CF levels, lock, list and cache structure sizes might need to be increased to support new function. For example, when you upgrade from CFCC Level 15 to Level 16 the required size of the structure might increase. This adjustment can have an impact when the system allocates structures or copies structures from one coupling facility to another at different CF levels.

The coupling facility structure sizer tool can size structures for you and takes into account the amount of space needed for the current CFCC levels.

Access the tool at:
http://www.**ibm.com**/servers/eserver/zseries/cfsizer/.

CFCC Level 16 is exclusive to System z10 and is supported by z/OS and z/VM for guest exploitation.

**Coupling Facility Configuration Alternatives**
IBM offers multiple options for configuring a functioning Coupling Facility:

• *Standalone Coupling Facility: The standalone CF provides the most "robust" CF capability, as the CPC is wholly dedicated to running the CFCC microcode — all of the processors, links and memory are for CF use only. A natural benefit of this characteristic is that the standalone CF is always failure-isolated from exploiting z/OS software and the server that z/OS is running on for environments without System-Managed CF Structure Duplexing. The z10 BC with capacity indicator A00 is used for systems with ICF(s) only. There are no software charges associated with such a configuration.*

• *Internal Coupling Facility (ICF): Customers considering clustering technology can get started with Parallel Sysplex technology at a lower cost by using an ICF instead of purchasing a standalone Coupling Facility. An ICF feature is a processor that can only run Coupling Facility Control Code (CFCC) in a partition. Since CF LPARs on ICFs are restricted to running only CFCC, there are no IBM software charges associated with ICFs. ICFs are ideal for Intelligent Resource Director and resource sharing environments as well as for data sharing environments where System-Managed CF Structure Duplexing is exploited.*

**System-Managed CF Structure Duplexing**
System-Managed Coupling Facility (CF) Structure Duplexing provides a general purpose, hardware-assisted, easy-to-exploit mechanism for duplexing CF structure data. This provides a robust recovery mechanism for failures such as loss of a single structure or CF or loss of connectivity to a single CF, through rapid failover to the backup instance of the duplexed structure pair. CFCC Level 16 provides CF Duplexing enhancements described previously in the section titled "Coupling Facility Control Code (CFCC) Level 16".

*A robust failure recovery capability*
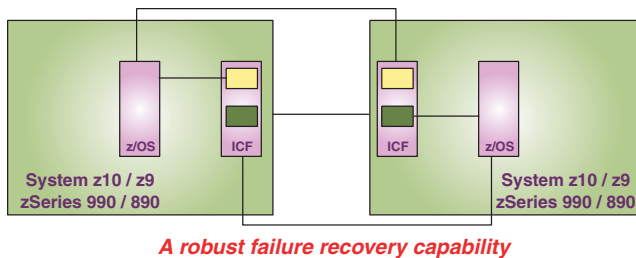
### Parallel Sysplex Coupling Connectivity

The Coupling Facilities communicate with z/OS images in the Parallel Sysplex environment over specialized high-speed links. As processor performance increases, it is important to also use faster links so that link performance does not become constrained. The performance, availability and distance requirements of a Parallel Sysplex environment are the key factors that will identify the appropriate connectivity option for a given configuration.

When connecting between System z10, System z9 and z990/z890 servers the links must be configured to operate in Peer Mode. This allows for higher data transfer rates to and from the Coupling Facilities. The peer link acts simultaneously as both a CF Sender and CF Receiver link, reducing the number of links required. Larger and more data buffers and improved protocols may also improve long distance performance.
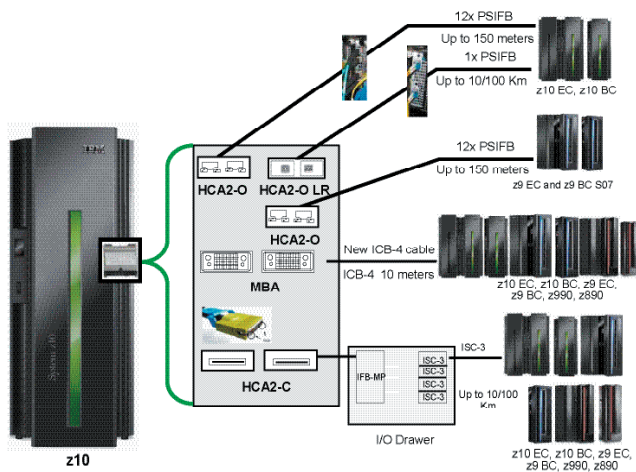


### Long reach InfiniBand coupling links

Now, InfiniBand can be used for Parallel Sysplex coupling and STP communication at unrepeated distances up to 10 km (6.2 miles) and even greater distances when attached to a qualified optical networking solution. InfiniBand coupling links supporting extended distance are referred to as 1x (one pair of fiber) IB-SDR or 1x IB-DDR.

- *Long reach 1x InfiniBand coupling links support single data rate (SDR) at 2.5 gigabits per second (Gbps) when connected to a DWDM capable of SDR*
- *Long reach 1x InfiniBand coupling links support double data rate (DDR) at 5 Gbps when connected to a DWDM capable of DDR.*

Depending on the capability of the attached DWDM, the link data rate will automatically be set to either SDR or DDR.

InfiniBand coupling link technology is designed to provide a high-speed solution and increased distance (150 meters) compared to ICB-4 (10 meters).

InfiniBand coupling links also provide the ability to define up to 16 CHPIDs on a single PSIFB port, allowing physical coupling links to be shared by multiple sysplexes. This also provides additional subchannels for Coupling Facility communication, improving scalability, and reducing contention in heavily utilized system configurations. It also allows for one CHPID to be directed to one CF, and another CHPID directed to another CF on the same target server, using the same port.

Like other coupling links, external InfiniBand coupling links are also valid to pass time synchronization signals for Server Time Protocol (STP). Therefore the same coupling links can be used to exchange timekeeping information and Coupling Facility messages in a Parallel Sysplex environment.

InfiniBand coupling links are CHPID type CIB.

**Coupling Connectivity for Parallel Sysplex**

Five coupling link options: The z10 BC supports Internal Coupling channels (ICs), Integrated Cluster Bus-4 (ICB-4), InterSystem Channel-3 (ISC-3) (peer mode), and 12x and 1x InfiniBand (IFB) links for communication in a Parallel Sysplex environment.

1) *Internal Coupling Channels (ICs) can be used for internal communication between Coupling Facilities (CFs) defined in LPARs and z/OS images on the same server.*

2) *Integrated Cluster Bus-4 (ICB-4) links are for short distances. ICB-4 links use 10 meter (33 feet) copper cables, of which 3 meters (10 feet) is used for internal routing and strain relief. ICB-4 is used to connect z10 BC-to-z10 BC, z10 EC, z9 EC, z9 BC, z990, and z890. Note: If connecting to a z9 BC or a z10 BC with ICB-4, those servers cannot be installed with the non-raised floor feature. Also, if the z10 BC is ordered with the non-raised floor feature, ICB-4 cannot be ordered.*

3) *InterSystem Channel-3 (ISC-3) supports communication over unrepeated distances of up to 10 km (6.2 miles) using 9 micron single mode fiber optic cables and even greater distances with System z qualified optical networking solutions. ISC-3s are supported exclusively in peer mode (CHPID type CFP).*

4) *12x InfiniBand coupling links (12x IB-SDR or 12x IB-DDR) offer an alternative to ISC-3 in the data center and facilitate coupling link consolidation; physical links can be shared by multiple systems or CF images on a single system. The 12x IB links support distances up to 150 meters (492 feet) using industry-standard OM3 50 micron fiber optic cables.*

*System z now supports 12x InfiniBand single data rate (12x IB-SDR) coupling link attachment between System z10 and System z9 general purpose (no longer limited to standalone coupling facility)*

5) *Long Reach 1x InfiniBand coupling links (1x IB-SDR or 1x IB-DDR) are an alternative to ISC-3 and offer greater distances with support for point-to-point unrepeated connections of up to 10 km (6.2 miles) using 9 micron single mode fiber optic cables. Greater distances can be supported with System z qualified optical networking solutions. Long reach 1x InfiniBand coupling links support the same sharing capability as the 12x InfiniBand version allowing one physical link to be shared across multiple CF images on a system.*

Note: The InfiniBand link data rates do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. Specifically, with 12x InfiniBand coupling links, while the link data rate can be higher than that of ICB, the service times of coupling operations are greater, and the actual throughput is less.

Refer to the Coupling Facility Configuration Options whitepaper for a more specific explanation of when to continue using the current ICB or ISC-3 technology versus migrating to InfiniBand coupling links.

The whitepaper is available at: http://www.**ibm.com**/systems/z/advantages/pso/whitepaper.html.

**z10 Coupling Link Options**

| Type | Description | Use | Link data rate | Distance | z10 BC z10 EC Max | z10 Max |
|------|-------------|-----|----------------|----------|-------------------|---------|
| PSIFB | 1x IB-DDR LR | z10 to z10 | 5 Gbps | 10 km unrepeated (6.2 miles) 100 km repeated | 12*/32* | |
| PSIFB | 12x IB-DDR | z10 to z10 z10 to z9 | 6 Gbps 3 GBps** | 150 meters (492 ft)*** | 12*/32* | |
| IC | Internal Coupling Channel | Internal Communi-cation | Internal Speeds | N/A | 32/32 | 64 CHPIDS |
| ICB-4 | Copper connection between OS and CF | z10, z9 z990, z890 | 2 GBps | 10 meters*** (33 ft) | 12/16 | |
| ISC-3 | Fiber connection between OS and CF | z10, z9 z990, z890 | 2 Gbps | 10 km unrepeated (6.2 miles) 100 km repeated | 48/48 | |

- The maximum number of Coupling Links combined cannot exceed 64 per server (PSIFB, ICB-4, ISC-3). There is a maximum of 64 Coupling CHPIDs (CIB, ICP, CBP, CFP) per server.

- For each MBA fanout installed for ICB-4s, the number of possible customer HCA fanouts is reduced by one

\* Each link supports definition of multiple CIB CHPIDs, up to 16 per fanout

\*\* z10 negotiates to 3 GBps (12x IB-SDR) when connected to a System z9

\*\*\* 3 meters (10 feet) reserved for internal routing and strain relief

Note: The InfiniBand link data rates of 6 GBps, 3 GBps, 2.5 Gbps, or 5 Gbps do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the link data rate may be higher than that of ICB (12x IB-SDR or 12x IB-DDR) or ISC-3 (1x IB-SDR or 1x IB-DDR), the service times of coupling operations are greater, and the actual throughput may be less than with ICB links or ISC-3 links.

**Time synchronization and time accuracy on z10 BC**

If you require time synchronization across multiple servers (for example you have a Parallel Sysplex environment) or you require time accuracy either for one or more System z servers or you require the same time across hetero-geneous platforms (System z, UNIX, AIX®, etc.) you can meet these requirements by either installing a Sysplex Timer Model 2 (9037-002) or by implementing Server Time Protocol (STP).

The z10 BC server requires the External Time Reference (ETR) feature to attach to a Sysplex Timer. The ETR fea-ture is standard on the z10 BC and supports attachment at an unrepeated distance of up to three kilometers (1.86 miles) and a link data rate of 8 Megabits per second. The distance from the Sysplex Timer to the server can be extended to 100 km using qualified Dense Wavelength Division Multiplexers (DWDMs). However, the maximum repeated distance between Sysplex Timers is limited to 40 km.

**Message Time Ordering (Sysplex Timer Connectivity to Coupling Facilities)**

As processor and Coupling Facility link technologies have improved, the requirement for time synchronization toler-ance between systems in a Parallel Sysplex environment has become ever more rigorous. In order to enable any exchange of timestamped information between systems in a sysplex involving the Coupling Facility to observe the correct time ordering, time stamps are now included in the message-transfer protocol between the systems and the Coupling Facility. Therefore, when a Coupling Facility is configured on any System z10 or System z9, the Coupling Facility will require connectivity to the same 9037 Sysplex Timer or Server Time Protocol (STP) configured Coordi-nated Timing Network (CTN) that the systems in its Parallel Sysplex cluster are using for time synchronization. If the ICF is on the same server as a member of its Parallel Sys-plex environment, no additional connectivity is required, since the server already has connectivity to the Sysplex Timer.

However, when an ICF is configured on any z10 which does not host any systems in the same Parallel Sysplex cluster, it is necessary to attach the server to the 9037 Sysplex Timer or implement STP.

**Server Time Protocol (STP)**

STP is a message-based protocol in which timekeeping information is transmitted between servers over externally defined coupling links. ICB-4, ISC-3, and InfiniBand coupling links can be used to transport STP messages.

Recent enhancements to STP Licensed Internal Code (LIC) delivers system management improvements by saving the STP configuration across PORs and power failures for a single server and a two server STP-only CTN. The attached documents outline the prerequisites, set up required, and steps needed to enable this functionality. Please make sure you refer to the appropriate document that pertains to your configuration, since the considerations are different between a single and two server CTN.

**Improved STP system management with new z/OS messaging:**

This new function is designed to generate z/OS messages when various hardware events that affect the External Time Sources (ETSs) configured for an STP-only Coordinated Timing Network (CTN) occur. This may improve problem determination and correction times. Previously, the messages were generated only on the Hardware Management Console (HMC).

The ability to generate z/OS messages is supported on IBM System z10 and System z9 servers with z/OS V1.11 with enabling support rolled back to z/OS V1.10 and V1.9.

NTP client support: This enhancement addresses the requirements of customers who need to provide the same accurate time across heterogeneous platforms in an enterprise.

The STP design has been enhanced to include support for a Simple Network Time Protocol (SNTP) client on the Support Element. By configuring an NTP server as the STP External Time Source (ETS), the time of an STP-only Coordinated Timing Network (CTN) can track to the time provided by the NTP server, and maintain a time accuracy of 100 milliseconds.

Note: NTP client support has been available since October 2007.

Enhanced accuracy to an External Time Source: The time accuracy of an STP-only CTN has been improved by adding the capability to configure an NTP server that has a pulse per second (PPS) output signal as the ETS device. This type of ETS device is available worldwide from several vendors that provide network timing solutions.

STP has been designed to track to the highly stable, accurate PPS signal from the NTP server, and maintain an accuracy of 10 microseconds as measured at the PPS input of the System z server. A number of variables such as accuracy of the NTP server to its time source (GPS, radio signals for example), and cable used to connect the PPS signal will determine the ultimate accuracy of STP to Coordinated Universal Time (UTC).

In comparison, the IBM Sysplex Timer is designed to maintain an accuracy of 100 microseconds when attached to an ETS with a PPS output. If STP is configured to use a dial-out time service or an NTP server without PPS, it is designed to provide a time accuracy of 100 milliseconds to the ETS device.

For this enhancement, the NTP output of the NTP server has to be connected to the Support Element (SE) LAN, and the PPS output of the same NTP server has to be connected to the PPS input provided on the External Time Reference (ETR) card of the System z10 or System z9 server.

Continuous Availability of NTP servers used as External Time Source: Improved External Time Source (ETS) availability can now be provided if you configure different NTP servers for the Preferred Time Server (PTS) and the Backup Time Server (BTS). Only the PTS or the BTS can

be the Current Time Server (CTS) in an STP-only CTN. Prior to this enhancement, only the CTS calculated the time adjustments necessary to maintain time accuracy. With this enhancement, if the PTS/CTS cannot access the NTP Server or the pulse per second (PPS) signal from the NTP server, the BTS, if configured to a different NTP server, may be able to calculate the adjustment required and propagate it to the PTS/CTS. The PTS/CTS in turn will perform the necessary time adjustment steering.

This avoids a manual reconfiguration of the BTS to be the CTS, if the PTS/CTS is not able to access its ETS. In an ETR network when the primary Sysplex Timer is not able to access the ETS device, the secondary Sysplex Timer takes over the role of the primary – a recovery action not always accepted by some customers. The STP design provides continuous availability of ETS while maintaining the special roles of PTS and BTS as – signed by the customer.

The availability improvement is available when the ETS is configured as an NTP server or an NTP server using PPS.

NTP Server on Hardware Management Console: Improved security can be obtained by providing NTP server support on the HMC. If an NTP server (with or without PPS) is configured as the ETS device for STP, it needs to be attached directly to the Support Element (SE) LAN. The SE LAN is considered by many users to be a private dedicated LAN to be kept as isolated as possible from the intranet or Internet.

Since the HMC is normally attached to the SE LAN, providing an NTP server capability on the HMC addresses the potential security concerns most users may have for attaching NTP servers to the SE LAN. The HMC, via a separate LAN connection, can access an NTP server available either on the intranet or Internet for its time source. Note that when using the HMC as the NTP server, there is

no pulse per second capability available. Therefore, you should not configure the ETS to be an NTP server using PPS.

Enhanced STP recovery when Internal Battery Feature is in use: Improved availability can be obtained when power has failed for a single server (PTS/CTS), or when there is a site power outage in a multi site configuration where the PTS/CTS is installed (the site with the BTS is a different site not affected by the power outage).

If an Internal Battery Feature (IBF) is installed on your System z server, STP now has the capability of receiving notification that customer power has failed and that the IBF is engaged. When STP receives this notification from a server that has the role of the PTS/CTS, STP can automatically reassign the role of the CTS to the BTS, thus automating the recovery action and improving availability.

Recent enhancements to STP Licensed Internal Code (LIC) delivers system management improvements by saving the STP configuration across PORs and power failures for a single server and a two server STP-only CTN. See Techdoc TD105103, at http://www-03.**ibm.com**/support/techdocs/atsmastr.nsf/WebIndex/TD105103 which outlines the prerequisites, set up required, and steps needed to enable this functionality.

Please make sure you refer to the appropriate document that pertains to your configuration, since the considerations are different between a single and two server CTN.

Application Programming Interface (API) to automate STP CTN reconfiguration: The concept of "a pair and a spare" has been around since the original Sysplex Couple Data Sets (CDSs). If the primary CDS becomes unavailable, the backup CDS would take over. Many sites have had automation routines bring a new backup CDS online to avoid a single point of failure. This idea is being extended to STP.

With this enhancement, if the PTS fails and the BTS takes over as CTS, an API is now available on the HMC so you can automate the reassignment of the PTS, BTS, and Arbiter roles. This can improve availability by avoiding a single point of failure after the BTS has taken over as the CTS.

Prior to this enhancement, the PTS, BTS, and Arbiter roles had to be reassigned manually using the System (Sysplex) Time task on the HMC.

For additional details on the API, please refer to System z Application Programming Interfaces, SB10-7030-11.

Additional information is available on the STP Web page: http://www.**ibm.com**/systems/z/pso/stp.html.

The following Redbooks are available on the Redbooks Web site: http://www.redbooks.**ibm.com**/.

- *Server Time Protocol Planning Guide, SG24-7280*
- *Server Time Protocol Implementation Guide, SG24-7281*

**Internal Battery Feature Recommendation**

**Single data center**
- *CTN with 2 servers, install IBF on at least the PTS/CTS*
  - *Also recommend IBF on BTS to provide recovery protection when BTS is the CTS*
  - *CTN with 3 or more servers IBF not required for STP recovery, if Arbiter configured*

**Two data centers**
- *CTN with 2 servers (one in each data center) install IBF on at least the PTS/CTS*
  - *Also recommend IBF on BTS to provide recovery protection when BTS is the CTS*
- *CTN with 3 or more servers, install IBF on at least the PTS/CTS*
  - *Also recommend IBF on BTS to provide recovery protection when BTS is the CTS*

The new functions available on the Hardware Management Console (HMC) version 2.10.2 as described apply exclusively to System z10.

The 2.10.2 HMC will continue to support up to two 10/100 Mbps Ethernet LANs. Token Ring LANs are not supported. The 2.10.2 HMC applications have been updated to support HMC hardware without a diskette drive. DVD-RAM, CD-ROM, and/or USB flash memory drive media will be used.

| Family | Machine Type | Firmware Driver | SE Version |
|--------|--------------|-----------------|------------|
| z10 BC | 2098 | 76 | 2.10.2 |
| z10 EC | 2097 | 76 | 2.10.2 |
| z9 BC | 2096 | 67 | 2.9.2 |
| z9 EC | 2094 | 67 | 2.9.2 |
| z890 | 2086 | 55 | 1.8.2 |
| z990 | 2084 | 55 | 1.8.2 |
| z800 | 2066 | 3G | 1.7.3 |
| z900 | 2064 | 3G | 1.7.3 |
| 9672 G6 | 9672/9674 | 26 | 1.6.2 |
| 9672 G5 | 9672/9674 | 26 | 1.6.2 |

**New HMC security features:**
The Hardware Management Console (HMC) and Support Element (SE) versions 2.10.2 provide a new feature called Digitally Signed Firmware (Licensed Internal Code). This new feature provides the following benefits.

- *It helps ensure no malware can be installed on System z products during LICC updates.*
- *It enables, with other existing security functions, System z10 CPACF functions to comply to Federal Information Processing Standard (FIPS) 140-2 Level 1 for Cryptographic Licensed Internal Code (LIC) changes.*
- *This enables use of the High Performance 897 Secure Key Bulk Encryption (Secure CPACF) feature.*

This new Digitally Signed Firmware follows the System z focus of security for the Hardware Management Console and Support Element. More details of the security aspects of the Hardware Management Console and Support Element are described in the z10 publication System z Hardware Management Console Security which can be found in the z10 Technical Notes area on IBM Resource Link.

**Serviceability enhancement for FICON channels:**
Problem determination can now be simplified by using the Hardware Management Console (HMC) to more quickly pinpoint fiber optic cabling issues in your Storage Area Network (SAN) fabric without IBM service personnel involvement.

All FICON channel error information is forwarded to the HMC where it is analyzed to help detect and report the trends and thresholds for all FICON channels on System z10. This report shows an aggregate view of the data and can span multiple systems.

**Internet Protocol, Version 6 (IPv6)**
HMC version 2.10.1 and Support Element (SE) version 2.10.1 can now communicate using IP Version 4 (IPv4), IP Version 6 (IPv6), or both. It is no longer necessary to assign a static IP address to an SE if it only needs to communicate with HMCs on the same subnet. An HMC and SE can use IPv6 link-local addresses to communicate with each other.

**HMC/SE support is addressing the following requirements:**
- *The availability of addresses in the IPv4 address space is becoming increasingly scarce.*
- *The demand for IPv6 support is high in Asia/Pacific countries since many companies are deploying IPv6.*
- *The U.S. Department of Defense and other U.S. government agencies are requiring IPv6 support for any products purchased after June 2008.*

More information on the U.S. government requirements can be found at: http://www.whitehouse.gov/omb/memoranda/fy2005/m05-22.pdf and

http://www.whitehouse.gov/omb/egov/documents/IPv6_FAQs.pdf.

**HMC/SE Console Messenger**
On systems prior to System z9, the remote browser capability was limited to Platform Independent Remote Console (PIRC), with a very small subset of functionality. Full functionality via Desktop On-Call (DTOC) was limited to one user at a time; it was slow, and was rarely used.

With System z9, full functionality to multiple users was delivered with a fast Web browser solution. You liked this, but requested the ability to communicate to other remote users.

There is now a new Console Manager task that offers basic messaging capabilities to allow system operators or administrators to coordinate their activities. The new task may be invoked directly, or via a new option in Users and Tasks. This capability is available for HMC and SE local and remote users permitting interactive plain-text communication between two users and also allowing a user to broadcast a plain-text message to all users. This feature is a limited instant messenger application and does not interact with other instant messengers.

**HMC z/VM Tower System Management Enhancements**
Building upon the previous z/VM Systems Management support from the Hardware Management Console (HMC), which offered management support for already defined virtual resources, HMC capabilities are available allowing selected virtual resources to be defined. In addition, further enhancements have been made for managing defined virtual resources.

Enhancements are designed to deliver out-of-the-box integrated graphical user interface-based (GUI-based) management of selected parts of z/VM. This is especially targeted to deliver ease-of-use for enterprises new to System z. This helps to avoid the purchase and installation of additional hardware or software, which may include complicated setup procedures. You can more seamlessly perform hardware and selected operating system management using the HMC Web browser-based user interface.

Support for HMC z/VM tower systems management enhancements is exclusive to z/VM 5.4 and the System z10.

**Enhanced installation support for z/VM using the HMC**
HMC version 2.10.1 along with Support Element (SE) version 2.10.1 on z10 BC and corresponding z/VM 5.4 support, will now give you the ability to install Linux on System z in a z/VM virtual machine using the HMC DVD drive. This new function does not require an external network connection between z/VM and the HMC, but instead, uses the existing communication path between the HMC and SE.

This support is intended for customers who have no alternative, such as a LAN-based server, for serving the DVD contents for Linux installations. The elapsed time for installation using the HMC DVD drive can be an order of magnitude, or more, longer than the elapsed time for LAN-based alternatives.

Using the legacy support and the z/VM 5.4 support, z/VM can be installed in an LPAR and both z/VM and Linux on System z can be installed in a virtual machine from the HMC DVD drive without requiring any external network setup or a connection between an LPAR and the HMC.

This addresses security concerns and additional configuration efforts using the only other previous solution of the external network connection from the HMC to the z/VM image.

Support for the enhanced installation support for z/VM using the HMC is exclusive to z/VM 5.4 and the System z10.

# Implementation Services for Parallel Sysplex

**IBM Implementation Services for Parallel Sysplex CICS and WAS Enablement**
IBM Implementation Services for Parallel Sysplex Middleware – CICS enablement consists of five fixed-price and fixed-scope selectable modules:

*1) CICS application review*

*2) z/OS CICS infrastructure review (module 1 is a prerequisite for this module)*

*3) CICS implementation (module 2 is a prerequisite for this module)*

*4) CICS application migration*

*5) CICS health check*

IBM Implementation Services for Parallel Sysplex Middleware – WebSphere Application Server enablement consists of three fixed-price and fixed-scope selectable modules:

*1) WebSphere Application Server network deployment planning and design*

*2) WebSphere Application Server network deployment implementation (module 1 is a prerequisite for this module)*

*3) WebSphere Application Server health check*

For a detailed description of this service, refer to Services Announcement 608-041, (RFA47367) dated June 24, 2008.

**Implementation Services for Parallel Sysplex DB2 Data Sharing**
To assist with the assessment, planning, implementation, testing, and backup and recovery of a System z DB2 data sharing environment, IBM Global Technology Services announced and made available the IBM Implementation Services for Parallel Sysplex Middleware – DB2 data sharing on February 26, 2008.

This DB2 data sharing service is designed for clients who want to:

1) Enhance the availability of data
2) Enable applications to take full utilization of all servers' resources
3) Share application system resources to meet business goals
4) Manage multiple systems as a single system from a single point of control
5) Respond to unpredicted growth by quickly adding computing power to match business requirements without disruption
6) Build on the current investments in hardware, software, applications, and skills while potentially reducing computing costs

The offering consists of six selectable modules; each is a stand-alone module that can be individually acquired. The first module is an infrastructure assessment module, followed by five modules which address the following DB2 data sharing disciplines:

1) DB2 data sharing planning
2) DB2 data sharing implementation
3) Adding additional data sharing members
4) DB2 data sharing testing
5) DB2 data sharing backup and recovery

For more information on these services contact your IBM representative or refer to: www.**ibm.com**/services/server.

### GDPS

Geographically Dispersed Parallel Sysplex (GDPS) is designed to provide a comprehensive end-to-end continuous availability and/or disaster recovery solution for System z servers, Geographically Dispersed Open Clusters (GDOC) is designed to address this need for open systems. When available, GDPS 3.5 will support GDOC for coordinated disaster recovery across System z and non-System z servers if Veritas Cluster Server is already installed. GDPS and the new Basic HyperSwap (available with z/OS V1.9) solutions help to ensure system failures are invisible to employees, partners and customers with dynamic disk-swapping capabilities that ensure applications and data are available. z10 BC—big on service, low on cost.

GDPS is a multi-site or single-site end-to-end application availability solution that provides the capability to manage remote copy configuration and storage subsystems (including IBM TotalStorage), to automate Parallel Sysplex operation tasks and perform failure recovery from a single point of control.

GDPS helps automate recovery procedures for planned and unplanned outages to provide near-continuous availability and disaster recovery capability.

For additional information on GDPS, visit: http://www-03.**ibm.com**/systems/z/gdps/.

### IBM Implementation Services for Parallel Sysplex CICS and WAS Enablement

IBM Implementation Services for Parallel Sysplex Middleware – CICS enablement consists of five fixed-price and fixed-scope selectable modules:

1) CICS application review
2) z/OS CICS infrastructure review (module 1 is a prerequisite for this module)
3) CICS implementation (module 2 is a prerequisite for this module)
4) CICS application migration
5) CICS health check

IBM Implementation Services for Parallel Sysplex Middleware – WebSphere Application Server enablement consists of three fixed-price and fixed-scope selectable modules:

*1) WebSphere Application Server network deployment planning and design*

*2) WebSphere Application Server network deployment implementation (module 1 is a prerequisite for this module)*

*3) WebSphere Application Server health check*

For a detailed description of this service, refer to Services Announcement 608-041, (RFA47367) dated June 24, 2008.

**Implementation Services for Parallel Sysplex DB2 Data Sharing**

To assist with the assessment, planning, implementation, testing, and backup and recovery of a System z DB2 data sharing environment, IBM Global Technology Services announced and made available the IBM Implementation Services for Parallel Sysplex Middleware – DB2 data sharing on February 26, 2008.

This DB2 data sharing service is designed for clients who want to:

*1) Enhance the availability of data*

*2) Enable applications to take full utilization of all servers' resources*

*3) Share application system resources to meet business goals*

*4) Manage multiple systems as a single system from a single point of control*

*5) Respond to unpredicted growth by quickly adding computing power to match business requirements without disruption*

*6) Build on the current investments in hardware, software, applications, and skills while potentially reducing computing costs*

The offering consists of six selectable modules; each is a stand-alone module that can be individually acquired. The first module is an infrastructure assessment module, followed by five modules which address the following DB2 data sharing disciplines:

*1) DB2 data sharing planning*

*2) DB2 data sharing implementation*

*3) Adding additional data sharing members*

*4) DB2 data sharing testing*

*5) DB2 data sharing backup and recovery*

For more information on these services contact your IBM representative or refer to: www.**ibm.com**/services/server.

**GDPS**

Geographically Dispersed Parallel Sysplex (GDPS) is designed to provide a comprehensive end-to-end continuous availability and/or disaster recovery solution for System z servers, Geographically Dispersed Open Clusters (GDOC) is designed to address this need for open systems. When available, GDPS 3.5 will support GDOC for coordinated disaster recovery across System z and non-System z servers if Veritas Cluster Server is already installed. GDPS and the new Basic HyperSwap (available with z/OS V1.9) solutions help to ensure system failures are invisible to employees, partners and customers with dynamic disk-swapping capabilities that ensure applications and data are available. z10 BC—big on service, low on cost.

GDPS is a multi-site or single-site end-to-end application availability solution that provides the capability to manage remote copy configuration and storage subsystems (including IBM TotalStorage), to automate Parallel Sysplex operation tasks and perform failure recovery from a single point of control.

GDPS helps automate recovery procedures for planned and unplanned outages to provide near-continuous availability and disaster recovery capability.

For additional information on GDPS, visit: http://www-03.**ibm.com**/systems/z/gdps/.

## Fiber Quick Connect for FICON LX Environments

Fiber Quick Connect (FQC), an optional feature on z10 BC, is offered for all FICON LX (single-mode fiber) channels, in addition to the current support for ESCON (62.5 micron multimode fiber) channels. FQC is designed to significantly reduce the amount of time required for on-site installation and setup of fiber optic cabling. FQC facilitates adds, moves, and changes of ESCON and FICON LX fiber optic cables in the data center, and may reduce fiber connection time by up to 80%.

FQC is for factory installation of Fiber Transport System (FTS) fiber harnesses for connection to channels in the I/O drawer. FTS fiber harnesses enable connection to FTS direct-attach fiber trunk cables from IBM Global Technology Services.

FQC, coupled with FTS, is a solution designed to help minimize disruptions and to isolate fiber cabling activities away from the active system as much as possible.

IBM provides the direct-attach trunk cables, patch panels, and Central Patching Location (CPL) hardware, as well as the planning and installation required to complete the total structured connectivity solution. An ESCON example: Four trunks, each with 72 fiber pairs, can displace up to 240 fiber optic jumper cables, the maximum quantity of ESCON channels in one I/O drawer. This significantly reduces fiber optic jumper cable bulk.

At CPL panels you can select the connector to best meet your data center requirements. Small form factor connectors are available to help reduce the floor space required for patch panels.

CPL planning and layout is done prior to arrival of the server on-site using the default CHannel Path IDdentifier (CHPID) placement report, and documentation is provided showing the CHPID layout and how the direct-attach harnesses are plugged.

FQC supports all of the ESCON channels and all of the FICON LX channels in the I/O drawer of the server. On an upgrade from a z890 or z9 BC, ESCON channels that are NOT using FQC cannot be used on the z10 BC FQC feature.

## z10 BC Physical Characteristics

### Physical Planning

A System z10 BC feature may be ordered to allow use of the z10 BC in a non-raised floor environment. This capability may help ease the cost of entry into the z10 BC; a raised floor may not be necessary for some infrastructures.

The non-raised floor z10 BC implementation is designed to meet all electromagnetic compatibility standards. Feature #7998 must be ordered if the z10 BC is to be used in a non-raised floor environment. A Bolt-down kit (#7992) is also available for use with a non-raised floor z10 BC, providing frame stabilization and bolt-down hardware to help secure a frame to a non-raised floor. Bolt-down kit (#7992) may be ordered for initial box or MES starting January 28, 2009.

The Installation Manual for Physical Planning (GC28-6875) is available on Resource Link and should always be referred to for detailed planning information.
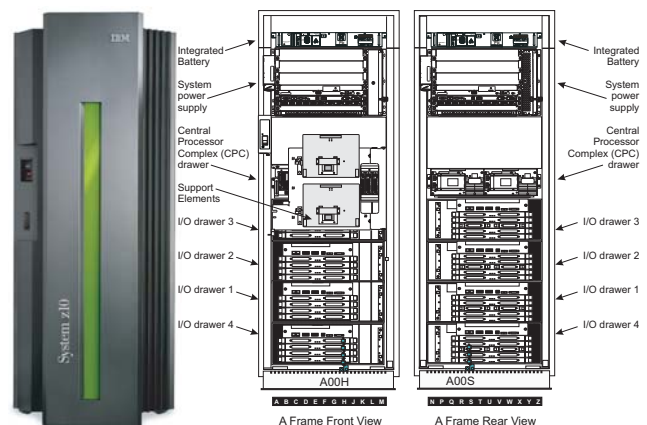
## z10 BC Configuration Detail

### z10 BC System Power

| | 1 I/O Drawer | 2 I/O Drawers | 3 I/O Drawers | 4 I/O Drawers |
|---|---|---|---|---|
| normal room (<28 degC) | 3.686 kW | 4.542 kW | 5.308 kW | 6.253 kW |
| warm room (>=28 degC) | 4.339 kW | 5.315 kW | 6.291 kW | 7.266 kW |

### z10 BC Highlights and Physical Dimensions

| | z10 BC | z9 BC |
|---|---|---|
| Number of Frames | 1 Frame | 1 Frame |
| Height (with covers) Width (with covers) Depth (with covers) | 201.5 cm/79.3 in (42 EIA) 77.0 cm /30.3 in 180.6 cm /71.1 in | 194.1 cm/76.4 in (40 EIA) 78.5 cm /30.9 in 157.7 cm /62.1 in |
| Height Reduction Width Reduction | 180.9 cm / 71.2 in (EIA) None | 178.5 cm / 70.3 in (EIA) None |
| Machine Area Service Clearance | 1.42 sq. m. /15.22 sq. ft. 3.50 sq. m. /37.62 sq. ft. (IBF Contained w/in Frame) | 1.24 sq. m. /13.31 sq. ft. 3.03 sq. m. /32.61 sq. ft. (IBF Contained w/in Frame) |

Maximum of 480 CHPIDs, four I/O drawers, 32 I/O slots (8 I/O slots per I/O drawer):

| Features | Min # Features | Max # Features | Max Connections | Increments per Feature | Purchase Increments |
|---|---|---|---|---|---|
| 16-port ESCON | 0 [1] | 32 | 480 channels | 16 channels 1reserved as as a spare | 4 channels |
| FICON Express8 | 0 [1] | 32 | 128* | 4* | 4* |
| FICON Express4* | 0 [1] | 32 | 64/128* channels | 2/4* channels | 2/4* channels |
| FICON Express2** | 0 [1] | 20 | 80 channels | 4 channels | 4 channels |
| FICON Express** | 0 [1] | 20 | 40 channels | 2 channels | 2 channels |
| ICB-4 | 0 [1] | 6 | 12 links [2] [3] | 2 links | 1 link |
| ISC-3 | 0 [1] | 12 | 48 links [2] | 4 links | 1 link |
| 1x PSIFB | 0 [1] | 6 | 12 links [2] | 2 links | 2 links |
| 12x PSIFB | 0 [1] | 6 | 12 links [2] [3] | 2 links | 2 links |
| OSA-Express3* | 0 | 24 | 48/96* ports | 2 or 4 | 2 ports/ 4 ports |
| OSA-Express2** | 0 | 24 | 24/48 ports | 1 or 2 | 2 ports/ 1 port |
| Crypto Express2* | 0 | 8 | 8/16 PCI-X adapters | 1/2* PCI-X adapters | 2* PCI-X adapters [4] |
| Crypto Express3* | 0 | 8 | 8/16 PCI-e adapters | 1/2* PCI-e adapters | 2* PCI-e adapters |

1) Minimum of one I/O feature (ESCON, FICON) or Coupling Link (PSIFB, ICB-4, ISC-3) required.

2) The maximum number of external Coupling Links combined cannot exceed 56 per server. There is a maximum of 64 coupling link CHPIDs per server (ICs, ICB-4s, active ISC-3 links, and IFBs)

3) ICB-4 and 12x IB-DDR are not included in the maximum feature count for I/O slots but are included in the CHPID count.

4) Initial order of Crypto Express2 is 2/4 PCI-X adapters (two features). Each PCI-X adapter can be configured as a coprocessor or an accelerator.

\* FICON Express4-2C 4KM LX has two channels per feature, OSA-Express3 GbE and 1000BASE-T have 2 and 4 port options. Crypto Express3-1P and Crypto Express2-1P has 1 coprocessor

\** Available only when carried forward on an upgrade from z890 or or z9 BC. Limited availability for OSA-Express2 GbE features.

## z10 BC Concurrent PU Conversions

- *Must order (characterize one PU as) a CP, an ICF or an IFL*
- *Concurrent model upgrade is supported*
- *Concurrent processor upgrade is supported if PUs are available*
  - *Add CP, IFL, unassigned IFL, ICF, zAAP, zIIP or optional SAP*
- *PU Conversions*
  - *Standard SAP cannot be converted to other PU types*

| From \ To | CP | IFL | Unassigned IFL | ICF | zAAP | zIIP | Optional SAP |
|---|---|---|---|---|---|---|---|
| CP | X | Yes | Yes | Yes | Yes | Yes | Yes |
| IFL | Yes | X | Yes | Yes | Yes | Yes | Yes |
| Unassigned IFL | Yes | Yes | X | Yes | Yes | Yes | Yes |
| ICF | Yes | Yes | Yes | X | Yes | Yes | Yes |
| zAAP | Yes | Yes | Yes | Yes | X | Yes | Yes |
| zIIP | Yes | Yes | Yes | Yes | Yes | X | Yes |
| Optional SAP | Yes | Yes | Yes | Yes | Yes | Yes | X |

Exceptions: Disruptive if ALL current PUs are converted to different types may require individual LPAR disruption if dedicated PUs are converted.

## z10 BC Model Structure

| Model | PU | PUs for Customer CPs | Max Avail Subcapacity | Standard SAPs | Standard Spares | CP/IFL/ICF/zAAP/zIIP** | Max Customer Memory | Max Chan. |
|---|---|---|---|---|---|---|---|---|
| E10 | 4 | 10 | 5 | 2 | 0 | 5/10/10/5/5 | 248 GB | 480* |

*\* Max is for ESCON channels.*
*\*\* For each zAAP and/or zIIP installed there must be a corresponding CP. The CP may satisfy the requirement for both the zAAP and/or zIIP. The combined number of zAAPs and/or zIIPs can not be more than 2x the number of general purpose processors (CPs).*

| z10 BC | Minimum | Maximum |
|---|---|---|
| E10 | 4 GB | 248 GB |

Memory DIMM sizes: 2 GB and 4 GB. (Fixed HSA not included, up to 248 GB for customer use June 30, 2009)

## System z CF Link Connectivity – Peer Mode only

| Connectivity Options | z10 ISC-3 | z10 ICB-4 | z10 1x PSIFB | z10 12x PSIFB |
|---|---|---|---|---|
| z10/z9/z990/z890 ISC-3 | 2 Gbps | N/A | N/A | N/A |
| z10/z9/z990/z890 ICB-4 | N/A | 2 Gbps | N/A | N/A |
| z9 with PSIFB | N/A | N/A | N/A | 3 GBps* |
| z10 1x PSIFB (>150m) | N/A | N/A | 5 Gbps* | N/A |
| z10 12x PSIFB | N/A | N/A | N/A | 6 GBps* |

- *N-2 Server generation connections allowed*
- *Theoretical maximum rates shown*
- *1x PSIFBs support single data rate (SDR) at 2.5 Gbps when connected to a DWDM capable of SDR speed and double data rate (DDR) at 5 Gbps when connected to a DWDM capable of DDR speed*
- *System z9 does NOT support 1x IB-DDR or SDR InfiniBand Coupling Links*

*Note: The InfiniBand link data rate of 6 GBps, 3 GBps or 5 Gbps does not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the link data rate may be higher than that of ICB, the service times of coupling operations are greater, and the actual throughput may be less than with ICB links.

## z10 BC System weight and IBF hold-up times

| z10 Model E10 – Single Frame | |
|---|---|
| w/o IBF | w/ IBF |
| 1890 lbs. | 2100 lbs. |

## z10 BC IBF hold uptime

| | 1 I/O Drawer | 2 I/O Drawers | 3 I/O Drawers | 4 I/O Drawers |
|---|---|---|---|---|
| 1 CPC Drawer | 13 min | 11 min | 9 min | 7 min |

## Coupling Facility – CF Level of Support

| CF Level | Function | z10 EC z10 BC | z9 EC z9 BC | z990 z890 |
|----------|----------|---------------|-------------|-----------|
| 16 | CF Duplexing Enhancements<br>List Notification Improvements<br>Structure Size increment increase from 512 MB –> 1 MB | X | | |
| 15 | Increasing the allowable tasks in the CF from 48 to 112 | X | X | |
| 14 | CFCC Dispatcher Enhancements | | X | X |
| 13 | DB2 Castout Performance | | X | X |
| 12 | z990 Compatibility 64-bit CFCC<br>Addressability Message Time Ordering<br>DB2 Performance SM Duplexing Support for zSeries | | X | X |
| 11 | z990 Compatibility SM Duplexing Support for 9672 G5/G6/R06 | | X | X |
| 10 | z900 GA2 Level | | X | X |
| 9 | Intelligent Resource Director IC3 / ICB3 / ISC3 Peer Mode<br>MQSeries® Shared Queues<br>WLM Multi-System Enclaves | | X | X |

Note: zSeries 900/800 and prior generation servers are not supported with System z10 for Coupling Facility or Parallel Sysplex levels.

## *Statement of Direction*

IBM intends to support optional water cooling on future high end System z servers. This cooling technology will tap into building chilled water that already exists within the datacenter for computer room air conditioning systems. External chillers or special water conditioning will not be required. Water cooling technology for high end System z servers will be designed to deliver improved energy efficiencies.

IBM intends to support the ability to operate from High Voltage DC power on future System z servers.   This will be in addition to the wide range of AC power already supported. A direct HV DC datacenter power design can improve data center energy efficiency by removing the need for an additional DC to AC inversion step.

The System z10 will be the last server to support Dynamic ICF expansion. This is consistent with the System z9 hardware announcement 107-190 dated April 18, 2007,  IBM System z9 Enterprise Class (z9 EC) and System z9 Business Class (z9 BC) – Delivering greater value for everyone, in which the following Statement of Direction was made: IBM intends to remove the Dynamic ICF expansion function from future System z servers.

The System z10 will be the last server to support connections to the Sysplex Timer (9037). Servers that require time synchronization, such as to support a base or Parallel Sysplex, will require Server Time Protocol (STP). STP has been available since January 2007 and is offered on the System z10, System z9, and zSeries 990 and 890 servers.

ESCON channels to be phased out: It is IBM's intent for ESCON channels to be phased out.  System z10 EC and System z10 BC will be the last servers to support greater than 240 ESCON channels.

ICB-4 links to be phased out: Restatement of SOD) from RFA46507) IBM intends to not offer Integrated Cluster Bus-4 (ICB-4) links on future servers.  IBM intends for System z10 to be the last server to support ICB-4 links.

Power Sequence Controller (PSC) feature quantities: The optional PSC feature provides the ability to turn on and off specific control units from the central processor complex (CPC). IBM intends to make three changes in the area of PSC support:

- *IBM intends for System z10 to be the last platform to support greater than two Power Sequence Controller (PSC) features*
- *Systems with water-cooling will further limit the maximum quantity of PSC features to one.*
- *IBM intends for System z10 to be the last platform to allow the PSC feature to be ordered individually when not part of a new-build server or when not part of a box MES order.*

Support for optional overhead cabling:  On future System z servers, IBM intends to support optional overhead cabling. This would be applicable to some data center environments and would apply to cabling for I/O (fiber optic and 1000BASE-T Ethernet).

Removal of specific smart card features:  The IBM System z10 EC and System z10 BC will be the last platforms to support smart card feature number #0888 and the #0887 smart card reader. The #0888 smart card has been replaced by the #0884 smart card. The #0887 smart card reader has been replaced by the #0885 smart card reader. The #0885 smart card reader and the #0884 smart card were made available on October 28, 2008. Refer to "IBM System z10 Enterprise Class - The future runs on System z10, the future begins today," Hardware Announcement 108-794, (RFA48381) dated October 21, 2008.

Customers should begin to migrate information from the #0888 smart card to the #0884 smart card to prepare for the change. Refer to the "Trusted Key Entry PCIX Workstation User's Guide" for instructions on how to make backups of TKE Certificate Authority (CA) smart cards and how to move key material from one TKE smart card to another.

Removal of Crypto Express2 feature:  The IBM System z10 EC and z10 BC will be the last servers to offer Crypto Express2 (#0863) as a feature, either as part of a new-build order, or carried forward on an upgrade.

All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

## *Publications*

**The following publications are available at announce in the Library section of Resource Link**

| Title | Order Number |
|---|---|
| z10 BC System Overview | SA22-1085 |
| z10 BC Installation Manual for Physical Planning (IMPP) | GC28-6875 |
| System z Functional Matrix | ZSW0-1335 |
| z10 PR/SM Planning Guide | SB10-7153 |

**The following publications are shipped with the product and available in the Library section of Resource Link**

| Title | Order Number |
|---|---|
| System z Service Guide for TKE Workstations | GC28-6862 |
| z10 BC Installation Manual | GC28-6874 |
| z10 BC Safety Inspection | GC28-6877 |
| z10 BC Service Guide | GC28-6878 |
| Systems Safety Notices | G229-9054 |
| System z Statement of Limited Warranty | GC28-6883 |

**The following publications will be available at planned availability in the Library section of Resource Link**

| Title | Order Number |
|---|---|
| System z API for Java | API-JAVA |
| System z Application Programming Interfaces | SB10-7030 |
| System z HMC Operations Guide (Version 2.10.2) | C28-6881 |
| System z CIM Management Interface | SB10-7154 |

| | |
|---|---|
| System z CHPID Mapping Tool User's Guide | GC28-6825 |
| System z Service Guide for HMCs and SEs | GC28-6861 |
| z10 Capacity on Demand User's Guide | SC28-6871 |
| z10 SE Operations Guide (Version 2.10.2) | SC28-6882 |

Publications for System z10 can be obtained at Resource Link by accessing the following Web site: www.**ibm.com**/servers/resourcelink

Using the instructions on the Resource Link panels, obtain a user ID and password. Resource Link has been designed for easy access and navigation.

**The following IBM Redbooks have been updated:**

| Title | Order Number |
|---|---|
| IBM System z Connectivity Handbook | SG24-5444 |
| IBM System z10 Business Class Technical Overview | SG24-7632 |

ZSO03021-USEN-04