

Implementing TCP/IP on the CMCC

This chapter describes considerations for designing a TCP/IP network using CMCCs, such as transport alternatives, scalability, and availability. The key issue covered in this guide is where to place the features for optimal network performance and scalability. The choices evaluated in this chapter include IP Datagram (using either CLAW or CMPC+) and TCP/IP Offload. The TN3270 Server design issues and options are discussed in a separate document, *TN3270 Design and Implementation Guide* at www.cisco.com/warp/public/cc/pd/ibsw/tn3270sr/tech/tndg_toc.htm.

Basic Design Considerations

When you are designing a TCP/IP network using the CMCC, you must decide the following:

- The number of channel-attached routers and CMCCs required to support scalability (throughput) requirements
- How to design your TCP/IP network for maximum availability
- How to easily migrate from an IBM 3172 Interconnect Controller to a CMCC

When you answer the design questions, you can design a network with optimal performance, high availability, and minimal cost.

Data Center Scalability

In an SNA environment, the key to determining scalability of the CMCC or other channel controller is the number of active sessions. SNA is characterized by relatively small frames of interactive, screen data.

In contrast, TCP/IP traffic is mixed. Some traffic is interactive traffic with small data frames. However, much of the traffic in a typical TCP/IP environment is from operations such as file transfers that require very high bandwidth and consume vast network resources. Therefore, the primary key to determining scalability of the CMCC in a TCP/IP environment is the overall throughput required. Secondary considerations include function placement and the number of TCP/IP connections.

You can determine the number of CMCCs required based on measuring the amount of traffic and thus the throughput required. This section provides some guidelines; however, you should consult your local systems engineer as well.

Cisco has performed a wide variety of throughput tests in its labs. The results of these tests, which have been confirmed in numerous real customer networks, indicate that the CMCC (running IP Datagram) does not inhibit throughput in any of the configurations tested. (A complete list of the configurations tested is available from Cisco.) In other words, the CMCC is capable of supporting as much IP Datagram traffic as the mainframe is able to generate and the channel connection is able to accommodate.

For example, in one test the CIP was able to sustain 116.8 Mbps throughput.¹ The results, as shown in Table 5-1, indicate a 95 percent utilization of each ESCON channel on the dual-port CIP when using a test application with the limited windowing characteristics observed. The results do not represent the absolute throughput capacity and capability for either the CIP or the RSP4. The CIP CPU utilization is 17 percent, which means that the CIP CPU still has a large amount of bandwidth available for use while transmitting data over two channels.

Table 5-1 shows the actual CIP and RSP4 metrics.²

Table 5-1 CIP and RSP4 Metrics

Metric	Utilization
CMCC CPU	17%
ECA0 (Actual Channel Utilization CHPID 1)	95%
ECA1 (Actual Channel Utilization CHPID 2)	95%
RSP4 CPU	1%

Extrapolating from these results, one can conclude that a single Cisco 7513 can support six fully loaded CIP interfaces and five additional LAN/WAN interfaces. The TCP/IP throughput data for a single port varies depending on the channel protocol you use.

Throughput as a Function of Channel Protocols and CMCC Performance

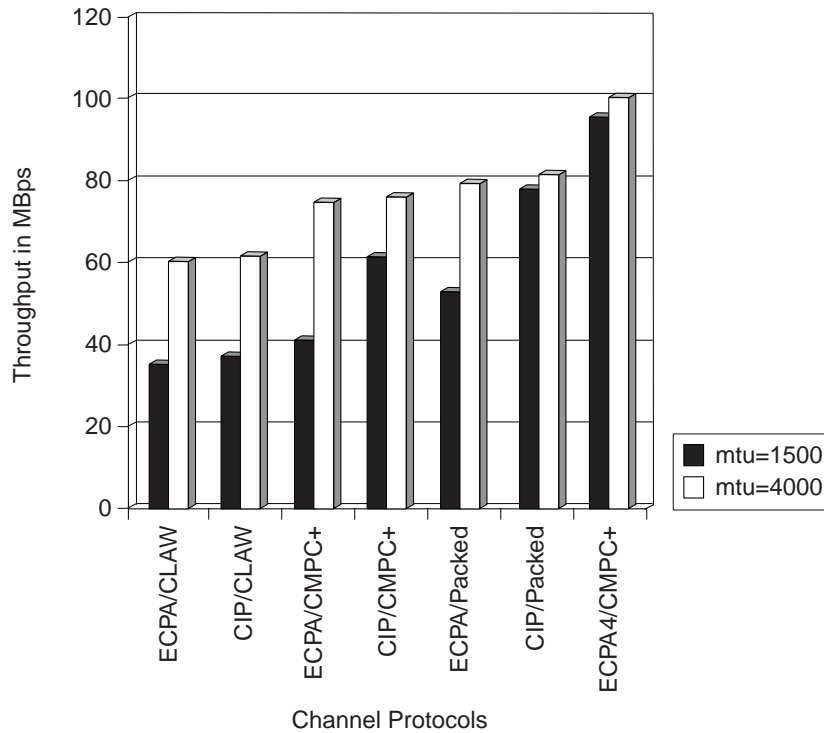
The channel protocol you select affects throughput. In general, using the CMPC+ protocol results in a higher throughput than the use of the CLAW protocol. OS/390 Version 2, Release 8 supports a feature called CLAW packing, which allows more data to be sent in a CLAW frame. IBM recommends using MPC+ in OS/390 Version 2, Release 5 or later.

For TCP/IP throughput, the CIP generally outperforms the CPA. However, the difference is relatively small, as shown in Figure 5-1. The new ECPA4 adapter using the CMPC+ protocol offers the best throughput for a CMCC.

1. For more information, refer to the white paper "CIP Performance Test: Specific Protocol Throughput Tests" at www.cisco.com/warp/public/cc/pd/ifa/ifpz/chifpz/tech/cipps_wp.htm.

2. RMFMON and the Systems Activity Display (SAD) were used to confirm the channel path identifier (CHPID) utilization.

Figure 5-1 TCP/IP Throughput as a Function of the Selected Channel Protocol



The performance figures in this document are quoted in Mbps. The quotation of throughput figures is not restricted to a particular metric. The use of Mbps (millions of bits per second), or MBps (millions of bytes per second), or KBps (thousands of bytes per second) depends largely on the environment in which the test was performed.

For example, the common terminology used to describe data throughput across IBM channels is MBps. This metric is used partly because the ESCON channel's bandwidth is designated as 10 or 17 MBps. LAN throughput traditionally has been measured in Mbps.

Performance benchmark tests are commonly used to demonstrate the absolute throughput of a given piece of internetworking equipment as described in Request for Comments (RFC) 1944. This document, *Benchmarking Methodology for Network Interconnect Devices Status*, discusses and defines the tests you can use to describe the performance characteristics of a network-interconnecting device. In addition, RFC 1944 describes specific formats for reporting the test results.

In the tests described in this chapter, the absolute throughput of the internetworking device (the Cisco 7507) has not been tested. The throughput of a single CMCC card and ESCON channels using a single given channel protocol and specific application were tested. (The CMCC can support multiple channel protocols—CSNA, CLAW, CMPC—simultaneously). Furthermore, tests were conducted when sending and receiving frames containing single protocol packets from multiple Token Rings. Consequently, these tests are not considered absolute performance or capability benchmarking tests.

Determining the Number of CMCCs

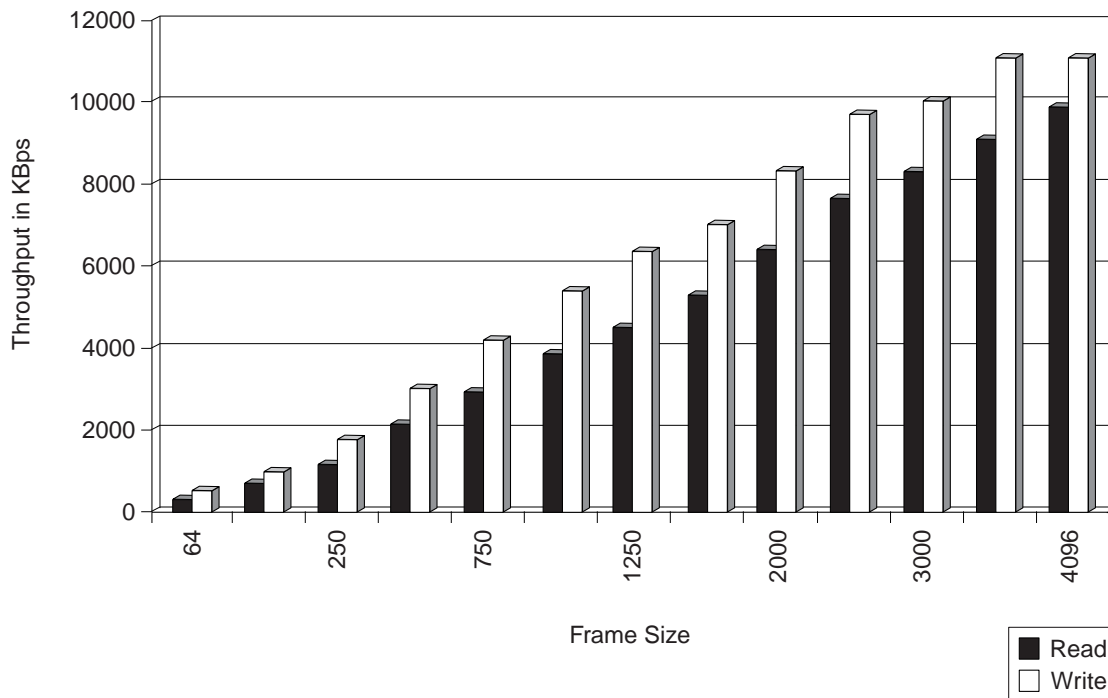
Determining how many CMCCs are required depends on the transaction rate, the transaction size, and the number of TCP/IP connections. It also depends on the number of hosts you attach to the CMCC and the type of channel you have, either ESCON or bus and tag. If you run multiple functions in a single CMCC (such as IP Datagram, TCP/IP Offload, or TN3270 Server), consult your systems engineer for assistance. The following numbers assume that only the Cisco IP Datagram feature is running on the CMCC:³

- A CMCC with a single ESCON can attach to between one and 32 hosts using either the ESCON Director or EMIF.
- A CIP with two daughter cards can attach to between one and 64 hosts by using either the ESCON Director or EMIF.


Other factors that may increase the number of CMCCs required are availability and redundancy requirements and the number of other features running in the channel-attached router.

Figure 5-2 shows CLAW throughput data based on the frame size and fiber length. CLAW has more overhead in the read direction, so Figure 5-2 shows both read and write directions. As the frame size increases, the CLAW channel throughput increases in both read and write directions. The throughput data for CMPC+ follows a similar trend as CLAW. As the frame size increases, so does the throughput.

Figure 5-2 CLAW Throughput as a Function of Frame Size



3. VTAM may limit the number of control units available for XCA devices.



Determining the Number of Channel-Attached Routers

If you run only IP Datagram functions in the channel-attached router, you can place four to five CMCCs in a single Cisco 7500 Series router. Performance is not the determining factor in this case, but rather availability, redundancy, and risk.

Other limiting factors in the router processor can be OSPF, Frame Relay, X.25, or any of the SNA features (for example, DLSw+). These features limit the capacity of the router, not the CMCC; however, if the features run in CMCC-attached routers, they may limit the capacity of the combined solution. This document does not describe how to determine the route processor limitations of these features.

Scalability in a TCP/IP Offload Environment

Before OS/390 Version 2, Release 5, one drawback to using TCP/IP on the mainframe was the amount of mainframe resources required to handle the TCP/IP header, checksumming, and lost packet retransmission. The TCP/IP Offload CMCC feature alleviates this problem by running the TCP/IP stack on the CMCC. The TCP/IP applications remain on the mainframe.

TCP/IP Offload with a Cisco CMCC significantly reduces the number of cycles needed on the mainframe while still providing high throughput capability. Between 30 and 50 percent of the cycles used by the mainframe stack can be saved by using TCP/IP Offload. Reduced cycle utilization on the mainframe means that the mainframe has more capacity to handle other traffic. The tradeoff for the reduced mainframe utilization is increased utilization on the CMCC. Compared to IP Datagram, the CMCC has a lot more work to do with TCP/IP Offload. Therefore, at some point the CMCC does become a limiting factor for scalability.

TCP/IP Offload can act as an offload host for several ESCON Director-attached mainframes or EMIF LPARs, as well as acting as a single offload host for directly connected ESCON- or bus and tag-connected hosts.

The offload TCP/IP stack can support full ESCON channel throughput of 9 MBps. In addition to offering high throughput, the CMCC, with its 64 MB of memory, maintains 7500 Telnet sessions with a 2-KB window or 500 FTP sessions with a 32-KB window. It simultaneously communicates to 32 host connections through the same ESCON adapter.

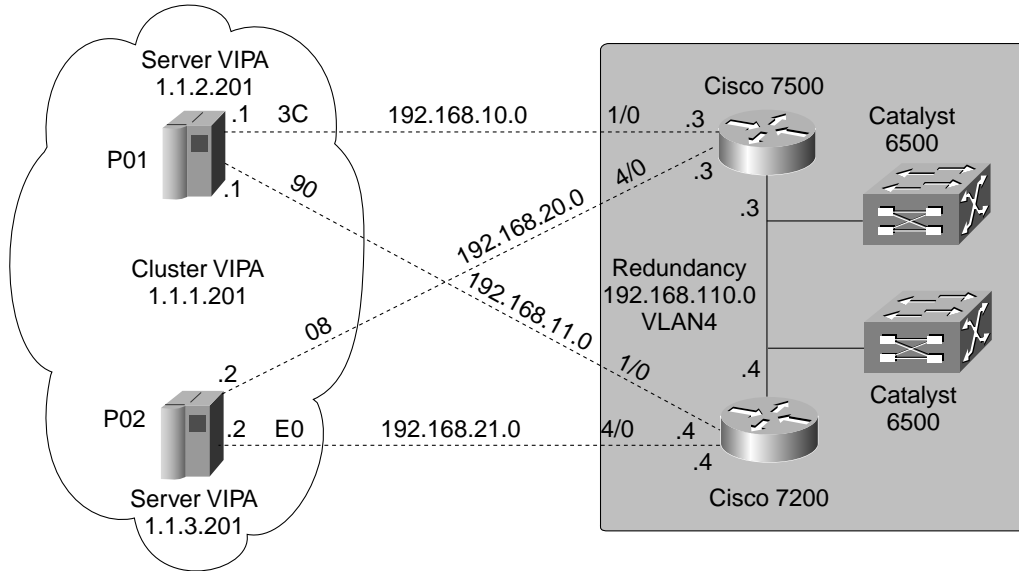
Data Center Availability

This section describes the channel-attached routers and CMCCs required to design your data center for maximum availability. Availability is the degree to which your data center resources are available to process transactions. This section provides some guidelines; however, you should consult your local systems engineer as well. Choices include using virtual IP addressing for redundancy and using MultiNode Load Balancing (MNLB). Another important availability issue is the prevention of the loss of traffic. Using Cisco IOS quality of service (QoS) features, the potential for the loss of mainframe traffic can be minimized.

Using the Virtual IP Address Feature for TCP/IP Redundancy

Virtual IP Address (VIPA) is a feature of OS/390 that allows you to configure a single virtual IP address on the mainframe. The network routing protocols Routing Information Protocol (RIP), called ROUTED on the mainframe for an IBM TCP/IP stack, and OSPF dynamically locate an alternate route to the VIPA (configured on the host TCP/IP stack definitions) through routing updates, as shown in Figure 5-3.

Figure 5-3 Redundancy Using the Virtual IP Address



Without the VIPA feature, if the CMCC adapter fails, alternate route information is often ignored because as long as the router's interface on the subnet is active, it sends the packet directly over the subnet.

VIPA solves this problem by creating a virtual subnet within the mainframe TCP/IP address space. When a remote host sends a packet to the virtual IP address, it is forced to use IP routing regardless of whether the host is on a directly attached subnet or not. The remote host must use its routing tables to decide on the best path to the virtual IP address.

Implementing Cisco TCP/IP Sysplex Solution: MNLB

With the rapid migration from SNA- to IP-based networks, Web access to mission-critical mainframe applications residing on host server platforms is essential. The host server cluster must scale to meet future growth requirements and be easy to maintain and support without disrupting application availability to the end user. In addition, the host server components should provide feedback on the network to be used as the basis of load-balancing decisions. MNLB is a Cisco IOS Software feature that provides easy access to IBM TCP/IP application servers residing on multiple hosts. With this feature, you do not need to keep track of which TCP/IP stack is running a particular application or to which host it is best to connect.

MNLB is designed for loosely coupling individual computing servers that balance the workload across the systems. This balancing is transparent to the requesting clients. The architecture employs an IP-based feedback mechanism enabling continuous adjustment of load-balancing decisions. The Cisco Appliance Services architecture (CASA) customizes routing in neighboring IP routing engines (or forwarding agents) under the direction of an MNLB Services Manager.

The MNLB architecture does not require all inbound traffic for a server cluster to pass through a single load-balancing engine. It enables a combination of fast forwarding agents (routers or IP-capable switches) and load-balancing and backup managers to synchronize and control traffic.



The MNLB design allows multiple MNLB routers to distribute load across multiple hosts in a cluster. A global virtual IP address is assigned to the cluster, which must be configured to all hosts as well as the MNLB routers. It allows the hosts to recognize the packets addressed to them and allows the MNLB routers to identify the packets that must be intercepted and rerouted.

The MNLB Services Manager provides a synchronization point for the assignment of affinities and maintenance of the affinity cache for clients and servers. An optional forward IP address identifies the new destination for the packet if it is not sent to the manager. The Services Manager assigns affinities using criteria specific to the network function. For example, the Services Manager uses load management information either sent by the host or statically configured to balance workloads.

A load-sensing agent interacts with the IBM Workload Manager to obtain host load information. This interaction allows the MNLB Services Manager to connect to each host in the cluster and retrieve load metrics that are used in balancing affinities across the cluster. Beginning with zOS Version 1, Release 2, the IBM Sysplex Distributor performs the functions of the MNLB Services Manager. For earlier releases, the MNLB Services Manager function is performed by the Cisco LocalDirector.

Forwarding agents can intercept packets that match local cache affinity and process them as instructed. If a matching affinity is not found, the packet is compared against the wildcard affinities to find managers that are interested in this type of packet. If no appropriate wildcard affinity is found, normal IP routing prevails. Generally, a manager uses the wildcard affinity to be informed of flows. When a manager has determined how a flow should be handled, it sets a full cached affinity so that subsequent packets for that same flow can be offloaded to the forwarding agent.

In the case of load balancing, full affinity is used to identify the server that is to receive the data. However, a wildcard affinity is used to define packet criteria and to identify the manager that makes the balancing decision for the IP packets that match wildcard criteria.

Using the Cisco IOS QoS Features

Cisco enables system administrators to prioritize mainframe traffic over other IP traffic. When you use IP precedence with a Cisco infrastructure, it improves the end-to-end network response time by leveraging Cisco IOS QoS⁴ features, such as Weighted Fair Queuing (WFQ) and Weighted Random Early Detection (WRED).

These features ensure that mainframe IP traffic takes precedence over other traffic and minimizes the chance of mainframe traffic being dropped during periods of congestion. This is especially important for SNA sessions that are being supported by TN3270 or HPR/IP. The SNA protocol is designed so that sessions can be dropped if an expected response is not received within a given amount of time. By judiciously utilizing QoS features, you can minimize the possibility of lost SNA sessions due to dropped traffic.

Migrating from an IBM 3172 Interconnect Controller

The IBM 3172 Interconnect Controller, introduced in the 1980s as a specialized PC server for connecting TCP/IP LANs to the mainframe, enjoyed considerable success in the late 1980s and early 1990s. IBM added SNA support to the original TCP/IP support, and it became a common LAN-to-mainframe connectivity device. As a result, many enterprises continue to have an installed base of IBM 3172 Interconnect Controllers in the data center.

4. For more information about QoS, refer to the white paper "Delivering Predictable Host Integration Services" at www.cisco.com/warp/public/cc/so/neso/ibso/ibm/s390/phost_wp.htm.

However, the IBM 3172 is based on dated technology and can no longer provide the throughput and availability characteristics required by the modern enterprise data center. The CMCC is a complete functional replacement for the IBM 3172 that provides vastly improved scalability and availability. Because of this, enterprises with multiple IBM 3172s can simplify their environments by migrating to a CMCC solution.

The CMCC implements the CLAW protocol for both IP Datagram and TCP/IP Offload features. This is the protocol used for the TCP/IP Offload feature offered on the IBM 3172. Therefore, enterprises that have implemented this feature on the IBM 3172 can very easily migrate to the CMCC by simply ensuring that the CMCC definitions match those of the existing interconnect controllers. Host definitions for IBM 3172s that are operating in an equivalent manner to the IP Datagram support on the CMCC need to change, but the magnitude of the change is very small.