z/OS® IP Network Security:

Capacity Planning for zIIP Assisted IPSec

Bruce Armstrong
Patrick Brown
Tom Moore
Jerry Stevens

July 26 2007
z/OS Communications Server Design

Acknowledgments

Many thanks to the following, for their contributions to this project.

Rick Armstrong - *IBM z/OS CommServer Performance Test*

Walt Caprice Jr - *IBM Washington Systems Center*

Chuck Gardiner - *IBM ENTS Business Development and Sales Enablement*

Linwood Overby - *IBM Enterprise Network Solutions*

Bernard Pierce - *IBM z/OS Performance*

Don Schmidt - *IBM Systems and Technology Group*

Overview

In IBM announcement 107-190 on April 18, 2007, IBM previewed that beginning with z/OS V1R8, the IBM System z9™ Integrated Information Processor (zIIP) can be used to handle much of the CPU-intensive processing involved in the IPSec AH (Authentication Header) and ESP (Encapsulating Security Payload) protocols. This further positions IBM System z™ as a cost-effective server in environments requiring end-to-end security for IP network traffic. For additional information about the zIIP refer to http://www-03.ibm.com/systems/z/ziip.

The objective of this document is to assist z/OS customers with capacity planning related to IPSec workloads executing on zIIP. This document does not provide installation and configuration information related to deploying IPSec on z/OS. For information about deploying IPSec on z/OS reference the z/OS Communication Server library and Communications Server for z/OS V1R8 TCP/IP Implementation Volume 4: Policy-Based Network Security SG24-7342 from the IBM Redbooks library.

The document is divided into 3 major topics as follows:

Part 1 of this document expands upon the announcement, provides a summary of the results of enabling zIIP Assisted IPSEC, and describes the configuration requirements for zIIP Assisted IPSec.

Part 2 of this document provides capacity-planning information for:

- Customers running on z/OS V1R8 who have already deployed IPSec on general purpose processors[1] and are considering migrating IPSec workload to zIIP. This information will help you estimate the zIIP capacity required to handle your current IPSec workload, as well as help you estimate the reduction in general purpose processor utilization once you've enabled the zIIP Assisted IPSec feature.

- Customers who have already deployed IPSec on z/OS, but are running on a version older than V1R8. Some traffic modeling may be necessary in order to estimate zIIP and CP utilization since you are running on a version of z/OS older than V1R8. This modeling is fairly straightforward, since you are already running IPSec.

- Customers not already running IPSec on z/OS, but who are considering future IPSec deployment on a System z9 server with zIIP. Again, traffic modeling will be necessary in order to do the projection analysis, but you'll have to make some assumptions about the percentage of your existing traffic to be secured when you do deploy IPSec. This paper contains some tips for such modeling.

---

[1] This paper will use the terms "general purpose processors" and "CPs" interchangeably. zIIP engines will always be referred to as 'zIIPs'.

In Part 3, early zIIP IPSec performance data, collected in the IBM Research Triangle Park Lab, is presented for three common IP Network traffic patterns. These data may be of general interest to network and system-capacity planners.

Future updates to this paper

As part of our development work, we regularly evaluate the performance of IPSec. If we identify and implement changes which improve IPSec, we will update this paper with new results.

Copyrights

Part 1: Functional Description

The new zIIP Assisted IPSec function allows Communication Server to interact with z/OS Workload Manager to have its enclave Service Request Block (SRB) work directed to zIIP. Within z/OS V1R8 Communications Server, much of the processing related to security routines (Encryption/Authentication algorithms and AH|ESP protocol overhead) runs in enclave SRBs, and this enclave SRB workload can be directed to available zIIPs.

Summary of results

If you are running IPSec, you may be able to achieve significant reduction in general purpose CPU consumption using the new zIIP Assisted IPSec function. As we discuss in subsequent sections, the amount of CPU reduction depends on the type of traffic being run, combined with other factors.

In a controlled lab test environment, we have measured zIIP Assisted IPSec relative to IPSec with no zIIP:

- Interactive workload CPU utilization reduction of up to 24%
- Bulk data transfer workload CPU reduction of up to 93%
- Throughput Improvement for Bulk Data workload when using zIIP:
    - TCP flow control optimizations become enabled in zIIP environments, enabling better overlap of security algorithms between sender and receiver sides of the connection (these flow control optimizations are discussed in "**IPSec Bulk-Data Throughput Discussion"** within Part 3 of this document);
    - The amount of IPSec throughput improvement on a zIIP server varies, related to TCP buffer sizes in use at both ends of the connection, natural latency on the communications link, and the performance characteristics of the IPSec node at the other side of the connection. With zIIPs at both ends of the connection (for z/OS to z/OS communication), we achieved approximately twice the throughput achievable in a non-zIIP configuration.
    - Approximately Linear scaling of multi-session IPSec throughput was observed, as more zIIPs are added to configuration

Performance data contained in this document was obtained in a specific operating environment, under the conditions described, and is presented as an illustration only. Actual performance in other operating environments may vary.

Enabling the Support

Software Requirements:

In order to enable the zIIP Assisted IPSec Communication Server function there are two z/OS V1R8 software changes related to this support:

1. Required - z/OS Communications Server TCP/IP APAR PK40178, which provides the support for zIIP Assisted IPSec.

2. Optional - z/OS APAR OW20045. Refer to the IIPHONORPRIORITY topic below

Configuration Requirements:

Assuming you already have IPSec enabled, you need to make only two additional configuration changes to enable zIIP Assisted IPSec.

z/OS Communications Server Configuration

The configuration statement:

GLOBALCONFIG ZIIP IPSECURITY

within the TCP/IP profile triggers CommServer to request z/OS to direct this IPSec enclave SRB processing to available zIIPs. The default for this configuration statement is GLOBALCONFIG ZIIP NOIPSECURITY.

WLM Policy Customization

Although this is an optional customization task, it is strongly recommended that you also complete this step. The enclave created for IPSec traffic is an *independent* enclave, meaning it can be classified and managed (within z/OS Workload Manager) differently than its owning address space (i.e., it can be classified/managed differently than the TCP/IP address space). The reason for providing the ability to treat IPSec traffic workload differently than other TCP/IP workloads is that the execution times of IPSec SRBs may be much longer than any other work directed to zIIP(s). Running these SRBs at very high priority could lead to significant processor delays for other work on the zIIP. In particular:

Bulk data workload (such as FTP or Tivoli Storage Manager (TSM) Backups) may generate SRBs which carry many IP datagrams, and when secured, each datagram needs to be driven through the authentication and/or encryption algorithms. This can result in SRBs with extremely long execution times.

If you're using zIIP to also handle other subsystems' workloads with demanding response time requirements (such as DB2/DRDA), we recommend you classify the IPSec enclave as less important, and with an execution velocity goal that allows WLM to achieve the goals for the other latency-sensitive zIIP workload. Even if you're using zIIP exclusively for IPSec workload, we recommend you classify the IPSec enclave, to keep the IPSec traffic workload from falling into the SYSOTHER service class (the catch-all service class for unclassified work).

To classify the independent enclave used for IPSec workload perform the following WLM Service definitions using the WLM ISPF panels:

1. Create a workload for the IPSec traffic that will be operating on the independent enclave. From the primary WLM ISPF panel select option 2 "Workloads".
2. Create a service class that will contain an appropriate performance goal for the IPSec independent enclave. From the primary WLM ISPF panel, select option 4 "Service Classes". From this panel you will define your new service class and associate it with the workload you previously defined. When you define BASE GOAL information for your single defined period, choose a goal type of "Execution velocity". After you select this, then you will need to define a Velocity and Importance for the service class being defined. It is important to set an appropriate value depending on other traffic that may be competing for zIIP or General CPU resources (General CPs become a factor when you have defined the IIPHONORPRIORITY parameter located in the IEAOPTxx member of SYS1.PARMLIB to a value of YES).
3. Create a WLM "subsystem type" for TCP/IP . You must specify the subsystem type name **TCP** and you can define it by using the WLM ISPF application. From the primary WLM ISPF panel select option 6 "Classification Rules". From this panel "Subsystem Type Selection List for Rules" you will move your cursor to the field "Subsystem-Type" and press the enter key. You will then be prompted for the type of operation you wish to perform. Since you want to create an new subsystem type, select option 1 "Create". From this new screen "Create Rules for the Subsystem Type" specify the "Subsystem Type" of **TCP** and a desired description of this new subsystem type.
4. Create a classification rule for the created "subsystem type" of **TCP** by using the WLM ISPF application screen "Create Rules for the Subsystem Type". (Reach this screen using option 6 from the primary screen or you may already be in this screen after the creation of the new subsystem type). At this point, define a classification rule for the subsystem type. This rule determines what work is associated with a service class for this subsystem type.
   You may use the following work qualifiers for the new independent enclave for IPSec work:

   •Subsystem Instance (SI) will be set to the TCP/IP stack's jobname.
   •Transaction Name will be set to a value of **TCPENC01**

Other z/OS MVS Tuning Controls

Three IEAOPTxx statements in SYS1.PARMLIB are relevant for zIIP Assisted IPSec.

1. **IIPHONORPRIORITY:** (added by APAR OA20045):

Specifying IIPHONORPRIORITY=YES allows the zIIP eligible workload to run on standard CPs, if zIIP work is not completed in a reasonable time period (see ZIIPAWMT below). This is the default and recommended value.

Specifying IIPHONORPRIORITY=NO disallows any zIIP eligible work from running on CPs (unless no zIIPs are online, or zIIP work is holding system locks or other resources impeding non zIIP work). When the NO value is set and zIIPs are present in the configuration, zIIP eligible work will be contained on the zIIPs. During periods of very high zIIP utilization, throughput and response

time may suffer. It may be reasonable to tradeoff throughput/response time in some environments, where minimizing utilization of the standard CPs is paramount.

2. **ZIIPAWMT**:

ZIIPAWMT controls how aggressive z/OS will be in requesting help from other zIIPs or CPs, (when IIPHONORPRIORITY=YES and all zIIPs are busy). We recommend using the default value for this setting (12 milliseconds).

3. **PROJECTCPU:** PROJECTCPU is used for projection purposes, and discussed in detail in the following sections.

Part 2: Capacity Planning for zIIP Assisted IPSec

In this section, we assume zIIP(s) are not on your z/OS image, but you are interested in *projecting the effectiveness* of zIIP for your existing (or future) IPSec workload. *Projecting the effectiveness* here means

- Determining the percentage of your current (or future) Communications Server-related CPU consumption eligible to be moved to zIIP (off of the general purpose processors)

- Determining how many zIIPs would be required to handle the amount of workload you calculated above, and

- Determining the percentage of reduction in CPU utilization on general purpose CPs once you have added zIIPs to the configuration.

It is also possible IPSec workload performance (response time and/or aggregate throughput) will be improved when zIIPs are added to the configuration. Such response time/throughput improvement is likely if your network performance is currently being constrained by high CPU utilization, and addition of zIIP(s) relieves this constraint. This paper does not discuss throughput projection modeling; the reader may assume response time and throughput will not be degraded with zIIP, as long as sufficient zIIP capacity is available.

Three methods for projecting zIIP effectiveness are described:

1. **PROJECTCPU Method:** very accurate and very simple, but you have to be on z/OS V1R8 AND, in order to use this method, you need to be running your representative IPSec workload already.

2. **Projection method based upon current TCP/IP SRB-mode CPU Consumption:** This method is also very accurate and simple, and you can be running on a z/OS older than V1R8. But you need to already be running your representative IPSec workload AND the accuracy of this projection method will be questionable if a significant portion of your total IP network workload is running non-secured. This method assumes all current SRB CPU consumption within the TCP/IP address space is related to IPSec operation. zIIP utilization (and general purpose CP utilization reduction) could be seriously overestimated if a significant portion of your traffic is running non-secured.

3. **IPSec Traffic Model Method :** If neither of the two methods above is usable in your environment, you will need to use the IPSec Traffic Model method. This method projects zIIP effectiveness based on several modeling variables. IBM provides a tool (CP3KIPSec) to observe traffic patterns and system behavior, and will collect data to be fed into the model. Once you are finished collecting data, send the output to IBM for processing. We analyze the data and provide you with a report documenting the results. We'll likely need some additional information from you (see "IPSec Modeling Questionnaire" in Part 2) before we can fully construct the model.

Method 1: Simplest zIIP Performance Projection method. Use PROJECTCPU Service

Target Audience: This method is the recommended projection option for all customers who are already on z/OS V1R8 and who are already running IPSec workload.

Function exists within z/OS allowing customers already running IPSec (but not currently using zIIPs) to accurately project the amount of their existing workload eligible to move the zIIPs. This function builds upon the PROJECTCPU service present in z/OS. PROJECTCPU gives a very precise accounting of zIIP eligible work. Using PROJECTCPU for zIIP capacity planning purposes is therefore very accurate and simple, since no extra analysis of network traffic is required. z/OS V1R8 is required to use PROJECTCPU for such IPSec zIIP performance analysis. (Although PROJECTCPU is present in z/OS 1.6 and 1.7, CommServer's IPSec enclave support does not exist in these older z/OS versions - so PROJECTCPU is not usable for IPSec projection modeling in z/OS 1.6 and 1.7.)

In order to use PROJECTCPU to project zIIP and general CP utilization for IPSec traffic, perform the following:

- Set PROJECTCPU=YES in parmlib member IEAOPTxx
- Set GLOBALCONFIG ZIIP IPSECURITY in your TCPIP profile dataset[2]
- Run your IPSec workload, and generate an RMF Workload Activity Report for representative interval(s)

Example Analysis 1:

Here's an example on interpretation of the RMF Workload Activity Report, based on the PROJECTCPU function. The characteristics of this network benchmark:

- TCP bulk data inbound to z/OS, similar to large FTP PUTs (into z/OS)
- IP MTU = 1500 bytes; connectivity via OSA-Express2 Gigabit Ethernet
- 5 concurrent TCP connections running (here we're using the IBM Application Workload Modeler, simulating the traffic pattern of 5 concurrent FTP PUTs)
- Traffic secured by IPSec tunnel using Triple DES Encryption; AH HMAC_MD5 Authentication
- Five System z9 (2094) CPs configured, zIIP NOT configured; using CPACF facility for crypto operations via ICSF.

[2] If you don't yet have zIIPs and you're configuring GLOBALCONFIG ZIIP IPSECURITY solely to do your performance projection work for future zIIPs, please note: TCP/IP will consume slightly more CPU if no zIIPs are online and you've coded GLOBALCONFIG ZIIP IPSECURITY, so you should remove GLOBALCONFIG ZIIP IPSECURITY from your TCPIP profile after you've completed your zIIP performance projection runs. Once you have zIIPs ONLINE, GLOBALCONFIG ZIIP IPSECURITY will have no negative effect on CPU consumption.

Our WLM Policy has defined a workload named IPSECWK, with an associated Service Class named IPSECCL, and we've created a classification rule to tie the IPSec independent enclave to this Service Class. (See the 'WLM Policy Customization' section in Part 1 for a detailed discussion on customizing WLM Policy for zIIP Assisted IPSec.)

We begin our projection analysis with the RMF Workload Activity Report for the IPSECCL Service Class:

```
REPORT BY: POLICY=SDPOL      WORKLOAD=IPSECWK      SERVICE CLASS=IPSECCL    RESOURCE GROUP=*NONE
                                                   CRITICAL     =NONE
                                                   DESCRIPTION  =zIIP assisted IPSec service cls

  TRANSACTIONS   TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE----  SERVICE TIMES  ---APPL %---  PAGE-IN RATES   ---STORAGE---
  AVG      1.00  ACTUAL                 0  SSCHRT   0.0  IOC          0  CPU    237.5   CP    395.82  SINGLE     0.0  AVG      0.00
  MPL      1.00  EXECUTION              0  RESP     0.0  CPU      62088K  SRB      0.0   AAPCP   0.00  BLOCK      0.0  TOT      0.00
  ENDED       0  QUEUED                 0  CONN     0.0  MSO          0  RCT      0.0   IIPCP 395.82  SHARED     0.0  CEN      0.00
  END/S    0.00  R/S AFFIN              0  DISC     0.0  SRB          0  IIT      0.0                 HSP        0.0  EXP      0.00
  #SWAPS      0  INELIGIBLE             0  Q+PEND   0.0  TOT      62088K  HST      0.0   AAP       N/A  HSP MISS   0.0
  EXCTD       0  CONVERSION             0  IOSQ     0.0  /SEC      1035K  AAP      N/A   IIP      0.00  EXP SNGL   0.0  SHR      0.00
  AVG ENC  1.00  STD DEV                0                                 IIP      0.0                 EXP BLK    0.0
  REM ENC  0.00                                        ABSRPTN  1035K                                  EXP SHR    0.0
  MS ENC   0.00                                        TRX SERV 1035K

  PER IMPORTANCE  PERF     --TRANSACTIONS--     -------------RESPONSE TIME--------------   -EX VEL%-   TOTAL   -EXE--
                  INDX     -NUMBER-     -%-     ------GOAL------  ---ACTUAL---     TOTAL   GOAL  ACT   USING%  DELAY%
  1   3            0.4            0       0                                          40 99.1    99.1     0.9
```

Analysis of the zIIP-specifics in the Workload Activity report for this benchmark:

Under the APPL% Column:

IIP N/A - zIIP is not configured.

IIPCP 395.82 - This is the percentage of CPU time used by zIIP-eligible work (in the IPSECCL Service Class) running on standard CPs. This statistic is normalized to the capacity of a single standard CP, and does not include certain z/OS activity such as the dispatcher and interrupt handlers, which are essential to the IPSec function being tested. The interpretation is: *This IPSec workload would more than fully saturate four zIIPs. The additional time for z/OS functions above would typically require something like 10% of the RMF times, resulting in about 407% total or about 7% of a fourth zIIP (or general purpose CP if spillover is allowed via the IIPHONORPRIORITY setting discussed earlier).* If less than five zIIPs are added to the configuration and IIPHONORPRIORITY=NO is specified (disallowing spillover to standard CPs), the zIIPs will run at full utilization, and we likely will not be able to achieve the full throughput achieved in this benchmark (because we'll be constrained by available processing power). If five zIIPs are added to the configuration, we'd predict each to run at approx 407/5 = 81.4% utilization, and we *would* be able to achieve the full throughput achieved in this benchmark.

CP 395.82 - Since there is no other workload (other than IPSec) running in the IPSECCL Service Class, all of the standard CP utilization in this Service Class is attributable to zIIP-eligible IPSec workload. So CP and IIPCP will be identical in a non-zIIP configuration.

By summing up the CPU busy for each of our service classes (or by reviewing the RMF CPU Activity Report for this measurement interval), we note that the z/OS image is running at 431.05% busy (each of

the 5 CPs is averaging 86.21% busy). So in this benchmark, the zIIP eligible IPSec workload accounts for 395.82/431.05=91.83% of the entire system activity in this benchmark. It might then be useful to generate a "rule of thumb" to estimate Network CPU Consumption Per Megabyte for this type of workload, when using the above-specified IPSec Ciphersuite (TDES+AH HMAC_MD5). This benchmark achieved an aggregate throughput of 105.09 MB/Sec (as reported by the IBM Application Workload Modeler), therefore:

CPU Consumption per MB (for this workload and configuration) is

[(4.3105 CPUSec/Sec)/(105.09 MB/Sec)] = 41.02 CPU milliseconds per Megabyte

If five zIIPs are added to the configuration to handle the zIIP eligible workload, the CPU consumption per MB will be approximately distributed as 3.36 CPU ms/MB on CPs and 37.66 CPU ms/MB on zIIPs. (91.8%*41.02 ms/MB should move to the zIIPs, with the remaining 8.2%*41.02 ms/MB staying on CPs.)

Four zIIP Measurement:

Below, four zIIPs have been added to the configuration - So how closely do these 'actual' results match up with the above projections?

```
REPORT BY: POLICY=SDPOL      WORKLOAD=IPSECWK    SERVICE CLASS=IPSECCL    RESOURCE GROUP=*NONE
                                                 CRITICAL     =NONE
                                                 DESCRIPTION  =zIIP assisted IPSec service cls

TRANSACTIONS   TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE----  SERVICE TIMES  ---APPL %---  PAGE-IN RATES   ---STORAGE---
AVG      1.00  ACTUAL                 0  SSCHRT   0.0  IOC          0  CPU    239.3   CP     14.62  SINGLE     0.0  AVG      0.00
MPL      1.00  EXECUTION              0  RESP     0.0  CPU      67878K  SRB      0.0   AAPCP   0.00  BLOCK      0.0  TOT      0.00
ENDED       0  QUEUED                 0  CONN     0.0  MSO          0   RCT      0.0   IIPCP  14.62  SHARED     0.0  CEN      0.00
END/S    0.00  R/S AFFIN              0  DISC     0.0  SRB          0   IIT      0.0                HSP        0.0  EXP      0.00
#SWAPS      0  INELIGIBLE             0  Q+PEND   0.0  TOT      67878K  HST      0.0   AAP       N/A  HSP MISS   0.0
EXCTD       0  CONVERSION             0  IOSQ     0.0  /SEC      1131K  AAP      N/A   IIP    384.17  EXP SNGL   0.0  SHR      0.00
AVG ENC  1.00  STD DEV                0                                 IIP    230.5                EXP BLK    0.0
REM ENC  0.00                                         ABSRPTN  1131K                                 EXP SHR    0.0
MS ENC   0.00                                         TRX SERV 1131K

PER IMPORTANCE  PERF     --TRANSACTIONS--     -------------RESPONSE TIME--------------   -EX VEL%-   TOTAL   -EXE--
                INDX     -NUMBER-     -%-     ------GOAL------  ---ACTUAL---     TOTAL   GOAL  ACT   USING%  DELAY%
1   3            0.4            0       0                                              40 89.6    89.6    10.4
```

IIP 384.17 - 3.84 zIIPs are fully consumed handling this IP workload. (IPSec accounts for 96.04% busy on each of the zIIPs.)

IIPCP 14.62 - This benchmark achieved 100 MB/Sec; which is slightly lower than what was observed for five zIIPs. In the earlier analysis (Workload Activity Report with NO zIIP), we said at least 7% of the zIIP-eligible workload will spill over to standard CPs, if we added four zIIPs to the configuration. In this case we're seeing 14.62% spilling over. The combined IIP eligible is 96.04+14.62% = 110.66%, which is within 4% of capacity projected using general purpose processors above.

CP 14.62 - In the earlier analysis (Workload Activity Report with NO zIIP), we noted that the IPSECCL Service Class accounted for 79.16% busy on each of the five standard CPs. With the zIIP now configured, IPSECCL-related CP utilization (averaged on each CP) has dropped to 14.62/2 = 7.31% busy; *a utilization reduction of over 71 percentage points on each of the standard CPs.* The IPSECCL service class work remaining on the standard CPs here is work that "spilled over" from the four zIIPs.

Total (CP+zIIP) CPU consumption per MB in the zIIP run is 43.9 CPU ms/MB. This is very close to the result seen in the non-zIIP run, but obviously the standard CPs are much less busy in the zIIP run.

Five-zIIP Measurement

Since our analysis suggested the above workload could not be fully contained with four zIIPs, we've added a fifth zIIP to the configuration:

```
REPORT BY: POLICY=SDPOL      WORKLOAD=IPSECWK    SERVICE CLASS=IPSECCL    RESOURCE GROUP=*NONE
                                                 CRITICAL     =NONE
                                                 DESCRIPTION  =zIIP assisted IPSec service cls

TRANSACTIONS   TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE----  SERVICE TIMES  ---APPL %---  PAGE-IN RATES   ---STORAGE---
AVG      1.00  ACTUAL                0  SSCHRT   0.0  IOC          0   CPU    250.4   CP      0.11  SINGLE     0.0  AVG      0.00
MPL      1.00  EXECUTION             0  RESP     0.0  CPU      71034K  SRB      0.0   AAPCP   0.00  BLOCK      0.0  TOT      0.00
ENDED       0  QUEUED                0  CONN     0.0  MSO          0   RCT      0.0   IIPCP   0.11  SHARED     0.0  CEN      0.00
END/S    0.00  R/S AFFIN             0  DISC     0.0  SRB          0   IIT      0.0                 HSP        0.0  EXP      0.00
#SWAPS      0  INELIGIBLE            0  Q+PEND   0.0  TOT      71034K  HST      0.0   AAP       N/A  HSP MISS   0.0
EXCTD       0  CONVERSION            0  IOSQ     0.0  /SEC      1184K  AAP      N/A   IIP    417.22  EXP SNGL   0.0  SHR      0.00
AVG ENC  1.00  STD DEV               0                                  IIP    250.3                 EXP BLK    0.0
REM ENC  0.00                                        ABSRPTN  1184K                                  EXP SHR    0.0
MS ENC   0.00                                        TRX SERV 1184K

PER IMPORTANCE  PERF     --TRANSACTIONS--     -------------RESPONSE TIME--------------   -EX VEL%-   TOTAL   -EXE--
                INDX     -NUMBER-     -%-     ------GOAL------  ---ACTUAL---     TOTAL   GOAL  ACT   USING%  DELAY%

1   3            0.4           0       0                                                  40  100     100     0.4
```

IIP 417.22 - 4.17 zIIPs are fully consumed handling this IPSec workload. (IPSec accounts for 83.4% busy on each of the zIIPs.)

IIPCP 0.11 - Practically no zIIP-eligible work has spilled over to the CPs.

This benchmark achieved 105.68 MB/Sec; which is slightly higher than what was observed for four zIIPs and no-zIIP runs.

Averaged over the entire measurement interval, we noted that the total CP+zIIP utilization was 452.82%, with each of the five zIIPs averaging 79.86% busy and each of the two general CPs averaging 26.76% busy. Normalizing our CP and zIIP consumption per unit of throughput we have:

CPU Consumption per MB (for this workload and configuration) is

[(4.5282 CPUSec/Sec)/(105.68 MB/Sec)] = 42.8 CPU milliseconds per Megabyte

These 42.8 ms/MB are distributed as 39.4 ms/MB on zIIPs, with the remaining 3.4 ms/MB on the CPs. These measured results are within five percent of the projections we stated earlier, using PROJECTCPU with no zIIPs.

Summary of this analysis: If this were the representative workload in your installation, you would conclude:

- Four zIIPs would be completely consumed by this workload, with more than 14% of the zIIP-eligible work spilling over to the CPs, or

- Five zIIPs would be required to fully handle the workload. Since it’s difficult to accurately predict these throughput improvements when adding zIIPs, capacity planning should be performed on a normalized “per-Megabyte-transferred” basis, rather than on a “raw throughput” basis.

When doing zIIP Capacity Planning for IPSec, don't forget to include zIIP capacity consumed by other zIIP exploiters (such as DB2/DRDA).

Method 2: zIIP Performance Projection based upon current SRB-mode CPU consumption in TCP/IP Address Space

Target Audience: Customers not yet on z/OS V1R8 who cannot use the PROJECTCPU function.

If you are already running IPSec on z/OS, and you are configured such that ALL IP traffic is secured via IPSec on z/OS, then zIIP Performance Projection is a very simple task. The amount of zIIP-eligible IPSec work present on the z/OS image is approximately equal to the amount of SRB-mode activity occurring within the TCP/IP address space.

While it is unlikely that you actually secure *all* your traffic via IPSec, this projection method should be fairly accurate even if you have a very high ratio of secured-to-nonsecured traffic. If it is not valid for you to assume all (or almost all) of your IP traffic is being secured via IPSec on z/OS, then you will have to use the more complex modeling technique described below, in Method 3.

The amount of SRB-mode activity within the TCP/IP address space can be determined from the RMF Monitor II Address Space Resource Data (ARD) report[3] . Subtract the "TCB TIME" value from the "CPU TIME" value, and average this statistic across your measurement interval.

Analysis based upon current SRB-mode CPU Consumption

```
00:16:18 DEV    FF   FF PRIV LSQA X C SRM  TCB   CPU  EXCP SWAP LPA CSA NVI V&H
JOBNAME  CONN  16M   2G   FF  CSF M R ABS  TIME  TIME RATE RATE  RT  RT  RT  RT
TCPIP    0.000   0  104    6  128 X    0M  0.04 35.21 0.00 0.00 0.0 0.0 0.0 0.0

SOF      0.000   3   57    5   74     ---  0.00  0.00 ---- ---- --- --- --- ---
RACF     0.000   0   66    1   91 X   1.5  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
TNF      0.000   0   33    0   48 X   0.0  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
VMCF     0.000   0   33    0   48 X   0.2  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
CATALOG  0.000   0  179    0  217 X    31  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
JES2MON  0.000   0   55    1   72     200  0.01  0.02 0.00 0.00 0.0 0.0 0.0 0.0
BPXOINIT 0.000   0   75    0   99     ---  0.00  0.00 ---- ---- --- --- --- ---
```

Here, the configuration and workload are identical to that described earlier (2 CPs, no zIIPs, TCP inbound Bulk Data with IPSec TDES+AH HMAC_MD5).

This is a DELTA-mode Monitor II ARD report with a measurement interval of 30 seconds. In this 30 second measurement interval, TCPIP has consumed 35.21-.04 = 35.17 SRB mode CPU Seconds. Our starting assumption for Method 2 is that ALL traffic is secured by IPSec, and this being the case, almost all SRB-mode activity within TCPIP is eligible to move to zIIP (for this model, it's reasonable to assume ALL SRB mode activity is zIIP eligible) . So 35.17 CPU seconds of processing is zIIP eligible during

[3] Once GLOBALCONFIG ZIIP IPSEC is configured in your TCPIP profile, you no longer can use the RMF Monitor II ARD report to observe IPSec-related CPU time (because IPSec processing will move to an enclave SRB). You'll need to use the Workload Activity Report to view the CPU consumption incurred by TCPIP enclave SRBs.

every 30 second interval, meaning it will take approximately 35.17/30 = 1.17 zIIPs to fully handle this workload.

We also noted the RMF CPU Activity Report for this run showed the two standard CPs each averaging 61% busy (or the system is consuming 1.22 CPU Seconds per Second, or 36.6 CPU Seconds each 30 second interval). If we provide enough zIIPs to fully handle this workload, the standard CPs will consume only 36.6-35.17 = 1.43 CPU seconds each 30 second interval; so CP utilization (on each CP) will drop to approximately .5*(1.43/30)=2.4%.

Note: The Method-2 projection technique just discussed assumes IPSec traffic rate will not differ in zIIP vs non-zIIP configurations. But as we describe later (in the "IPSec Bulk-Data Throughput Discussion" section), throughput for your bulk data applications (such as FTP) might be improved simply by enabling the zIIP Assisted IPSec function (an obvious contradiction to the "throughput will remain flat" assumption). Customers with a high degree of bulk-data traffic may therefore find the zIIP Capacity estimate produced by Method-2 is somewhat low; Method-2 will project *just enough* zIIP capacity to handle the non-zIIP traffic rate, but higher throughput rates may be achievable with additional zIIP capacity.

Method 3: Projection based upon IPSec Traffic Modeling

Target Audience: Customers who can't use either Projection Methods (1) or (2) above.

If the above projection methods can't be used in your environment (because either you're not yet running IPSec, or you're on a version of z/OS older than V1R8), some modeling will be necessary to derive zIIP projections for your (current or future) IPSec workload. IBM will assist with this modeling.

Overview of modeling methodology:

We'll need to develop an understanding of your future IPSec configuration (even if you're running IPSec on an older z/OS version, we cannot translate behavior directly from your current release to z/OS V1R8, so modeling is required). The general modeling approach we'll take:

As a starting point, IBM will provide you with a tool to gather various statistics describing your current z/OS Communication Server activity. The data collected by this tool will be spaced at regular intervals (e.g., every 15 minutes) so we can model your system on a discrete-interval basis. The statistics we'll need to feed into the zIIP projection model are the following:

- **IR:** Inbound Datagram Rate
- **OR:** Outbound Datagram Rate
- **IS:** Avg Inbound Datagram Size
- **OS:** Avg Outbound Datagram Size
- **PPI:** Avg Inbound Packets per read-side interrupt
- **TS, TR:** Asymmetry Factors due to TCB mode activity [4]

Once the tool has gathered these per-interval statistics, you will need to help shape the collected data to reflect what your eventual IPSec configuration will look like. Specifically, you need to consider your network, and answer three questions, composed in the IPSec Modeling Questionnaire.

---

[4] An asymmetry regarding zIIP utilization may exist due to outbound Socket calls (e.g., SEND) being executed in TCB mode. In many cases, the ***outbound*** TCP/IP (and IPSec) data path is executed on the application's thread, and if this thread is operating in TCB mode, it will remain on the general purpose processor (it will not be zIIP eligible). By contrast, the entire ***inbound*** TCPIP (and IPSec) data path runs in SRB mode, and for IPSec traffic, this SRB will always be zIIP eligible (if GLOBALCONFIG ZIIP IPSECURITY is specified). Hence we have a potential inbound|outbound asymmetry. Our zIIP Projection modeling needs to take this asymmetry into account.

IPSec Modeling Questionnaire:

- Should IBM assume the same traffic load you're currently handling? Or a somewhat lighter or heavier traffic load than is currently present on the z/OS server?

- What percentage of total traffic do you think will eventually be secured by IPSec? You can pick a range, and IBM will model across the range.

- What percentage of total IPSec-secured traffic will be of an interactive nature (e.g., TN3270, CICS, IMS, etc)? In the modeling, IBM will then assume the remaining traffic is of a bulk nature (e.g., FTP, Tivoli Storage Manager (TSM), NFS, etc). These two workload types behave quite differently, especially regarding the TCB-mode asymmetry issue on the outbound data path. So we'll need to take a stab at differentiating these two general classes of workload. (And since we'll be modeling over discrete intervals, the relative percentage of interactive vs bulk can differ per interval.)

Once the above questions are answered and data is collected, IBM will run the zIIP Projection model and report back on projected zIIP utilization ranges (per interval), number of zIIPs required to handle peak periods, projected reduction in standard CP utilization (per interval), etc..

Part 3: Early zIIP Assisted IPSec Performance Data

In this section, we present performance data obtained for three common traffic patterns. Note: we used IBM Application Workload Modeler http://www.ibm.com/software/network/awm (Product Number 5655-J62) to generate these traffic patterns; the actual applications cited were *not* used.

We show Throughput and Network CPU Consumption[5] for these workloads:

- Bulk Data Inbound to z/OS - similar to FTP PUT or TSM Backup
- Bulk Data Outbound from z/OS - similar to FTP GET or TSM Restore
- Interactive (Request/Response) Traffic - Similar to TN3270, CICS, DB2, IMS

The above workloads were run with IPSec enabled, and we compare configurations with: NO zIIP, ONE zIIP, TWO zIIPs, Three zIIPs and Five zIIPs. All configurations have the server defined with TWO standard CPs. The server under test is a 2094-S38.

For completeness, we also show the performance metrics for the same workloads running nonsecured (IPSec not enabled). This should give the reader a feeling for the magnitude of the additional processing overhead involved in the IPSec protocols.

IPSec Bulk-Data Throughput Discussion

During development of the zIIP Assisted IPSec function, we discovered certain detrimental (to bulk-data throughput) behavior in TCP's flow control, when operating over an IPSec tunnel. When operating over an IPSec tunnel, bulk-data TCP connections were falling into a "start/stop" mode (as opposed to a continuous "streaming" mode), with the communication channel repeatedly going idle. This behavior was leading to additional latency at both ends of each TCP connection, which was severely limiting attainable throughput.

In order to sustain a high degree of throughput for bulk workload, TCP **must** avoid having the communication channel ever go idle, so with the zIIP Assisted IPSec function, we've incorporated TCP flow control changes to keep the communication pipe "full", thereby avoiding the latency issues visible with earlier code. On a normalized (per-MB transferred) basis, these flow control changes **do** result in slightly higher CPU consumption than earlier code, so we've designed these throughput improvement changes such that they become enabled only if zIIP Assisted IPSec is enabled. *The result: significantly higher bulk data throughput is achievable in zIIP environments than in non-zIIP environments.*

---

[5] *Network CPU Consumption* is a measure of all TCP/IP communications-related CPU resource consumed by an application on a per-transaction (or per-Megabyte) basis. Applications consume an additional amount of CPU resource (per transaction) which is unrelated to communications, and this amount varies from application to application. In this study, we've ignored application-related CPU consumption, and instead focus on IPSec's impact to *Network CPU Consumption* for several workload types.

Single-Session vs Multi-Session Throughput Considerations: In the lab, we've demonstrated that a single TCP connection's throughput will become capped at the point of full (100%) utilization of a single zIIP. That is, a single TCP connection cannot make use of more than 100% of one zIIP. On the other hand, when multiple bulk-data TCP connections are operating concurrently, we observe approximately linear gains in aggregate throughput as more zIIPs are added to the configuration (until we approach the point of Gigabit Ethernet saturation). All of the bulk-data workload throughput data below is for five concurrent TCP connections.

Bulk-Data Throughput Disclaimer: In order to achieve high throughput rates, both the transmit (encrypting) and receive (decrypting) sides of the connection need to perform well. At this time, we have little data on the IPSec performance capabilities of other (non-z/OS) platforms. Therefore, all of the data presented in this paper is for z/OS to z/OS communication, and we make no IPSec performance claims for any other platform communicating with z/OS.

Bulk Data Inbound to z/OS

Inbound bulk data traffic (e.g., FTP PUT or TSM Backup) is a workload type that will benefit greatly from zIIP Assisted IPSec. Figure 1 demonstrates that enabling IPSec *without* zIIPs can result in over an 11X increase in CP utilization (normalized on a per-megabyte basis). However, once sufficient zIIP capacity is added to the configuration, the vast majority of IPSec processing moves to zIIP, with the CPs handling only TCB-mode activity and any SRB-mode IPSec activity that spills over from zIIP.

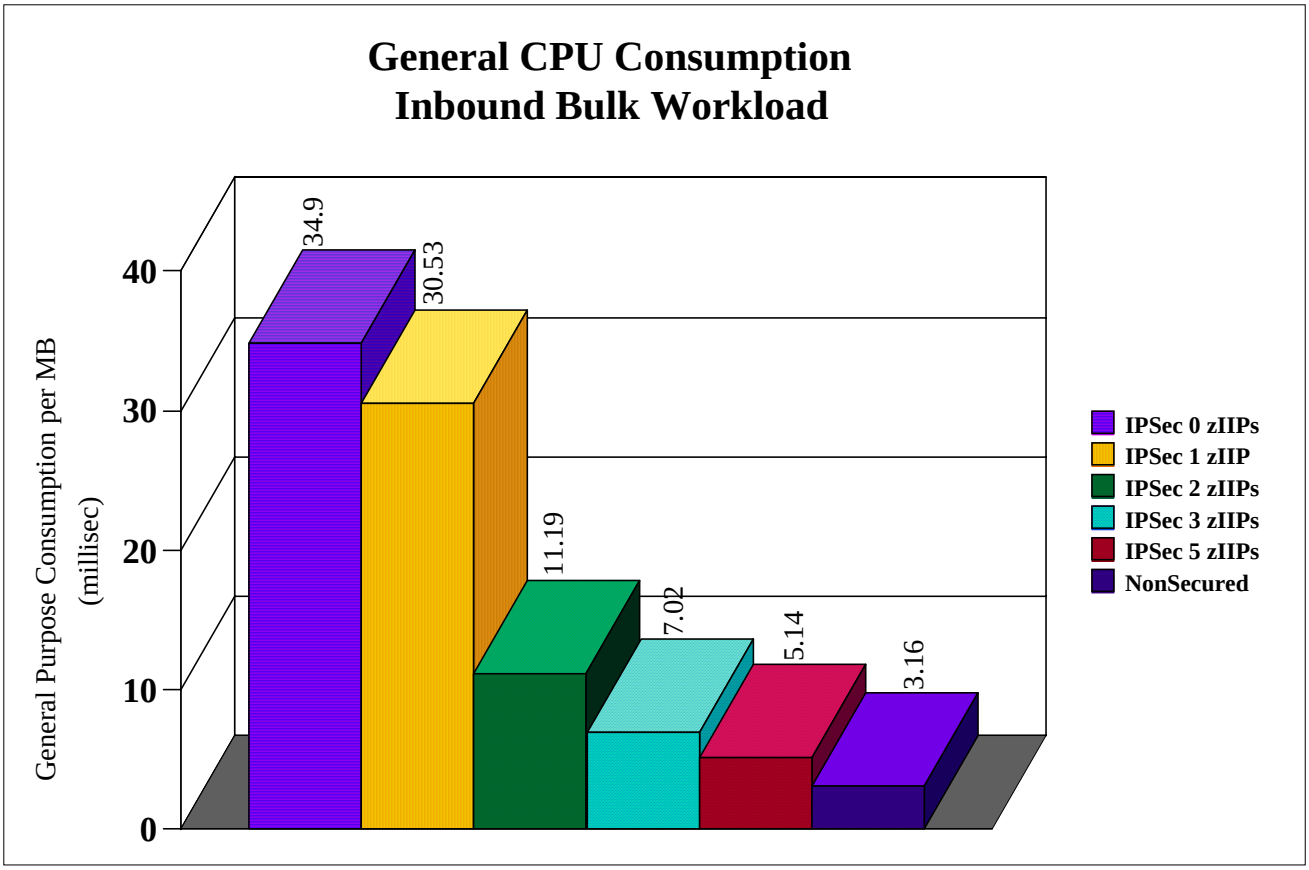

Figure 1: Network CPU Consumption Per MB - Inbound Bulk Data

Figure 2 demonstrates the impact of additional zIIP capacity on multi-session throughput. Where a 2 CP configuration achieved a max of 31.84 MB/Sec, addition of 5 zIIPs has raised aggregate throughput approximately to the point of full Gigabit Ethernet saturation (104.53 MB/Sec). By contrast (but not shown in the chart), a non-zIIP configuration could not achieve more than 55 MB/Sec, regardless of number of configured general purpose CPs.

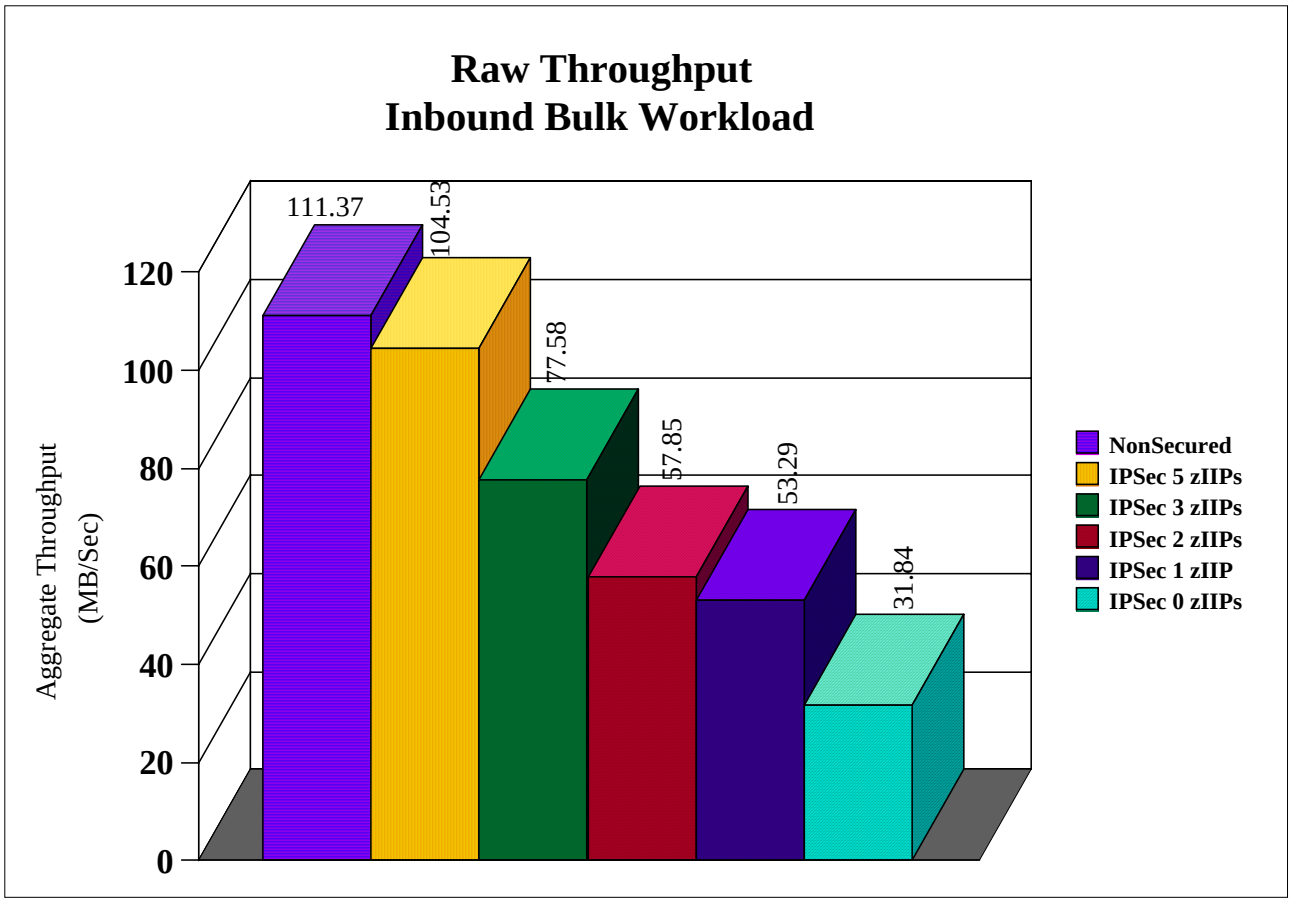

Figure 2: Raw Throughput - Inbound Bulk Data

Bulk Data Outbound from z/OS

Outbound bulk data traffic (e.g., FTP GET or TSM RESTORE) is a workload type that will benefit greatly from zIIP Assisted IPSec.

Our measurements suggest enabling IPSec for outbound bulk workload will result in a 29X increase in CPU processing. With zIIP in the configuration, approximately 94% of this additional IPSec processing moves to zIIP.

In figure 3 below the outbound IPSec bulk workload fully utilized three zIIPs and required the assistance of a fourth zIIP to handle all of the IPSec traffic that could be directed to the specialty engines.

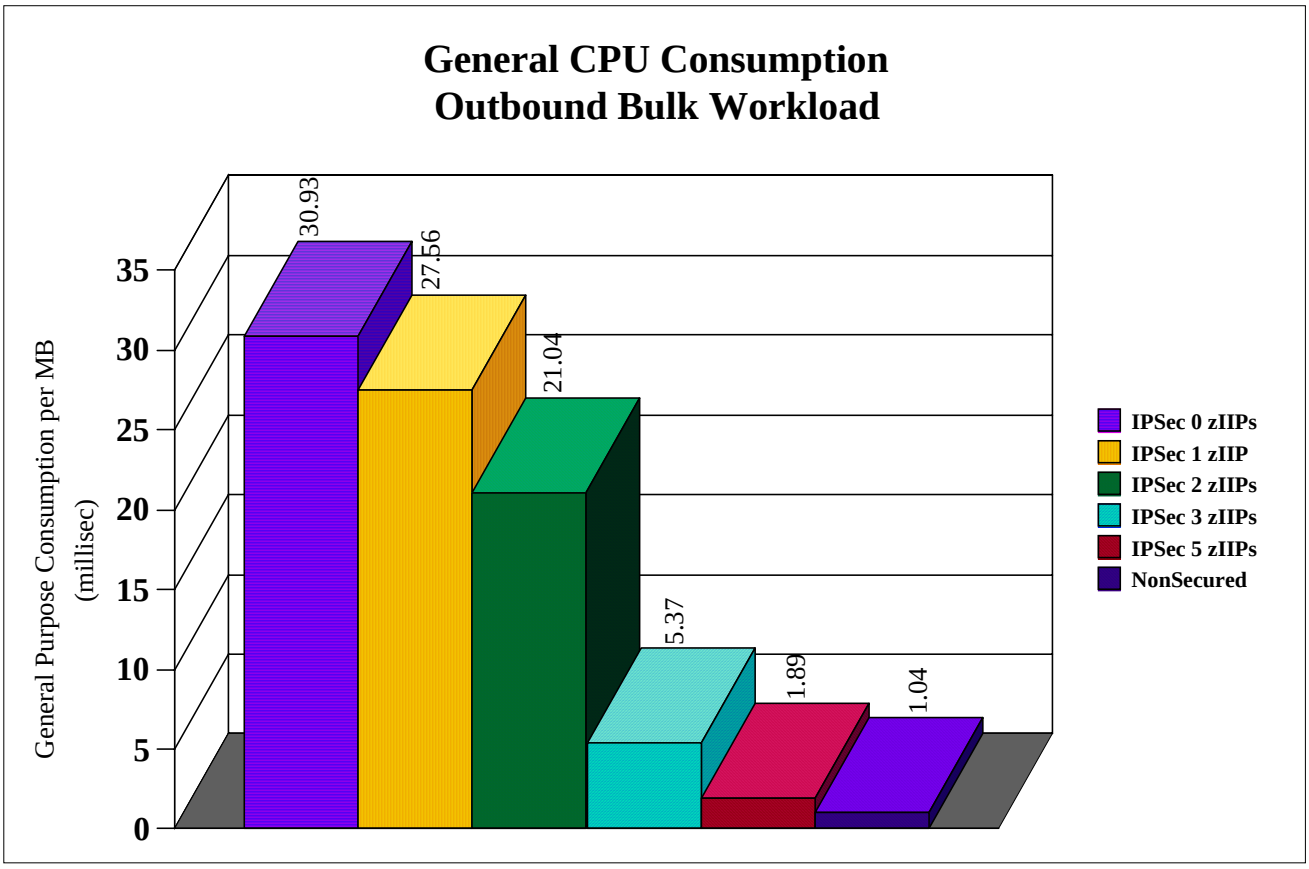

Figure 3: Network CPU Consumption per MB- Outbound Bulk Data

Outbound bulk workload does not appear to be significantly affected by the TCB-mode asymmetry problem discussed in Part 2. This is because with bulk traffic workload, most data is driven outbound in SRB-mode (as a result of inbound TCP ACK segments). Since the majority of outbound data is processed in SRB mode, it is zIIP eligible.

Figure 4 is the raw outbound IPSec bulk data throughput achieved over a Gigabit Ethernet link and this throughput data was collected from the same benchmark run described in figure 3 (CPU consumption data for Outbound Bulk IPSec traffic).

As illustrated in figure 4, when you have sufficient CPs (General or zIIP specialty engines) to handle the IPSec traffic then you can nearly achieve throughput rates that were obtained for non-secured data transmissions.

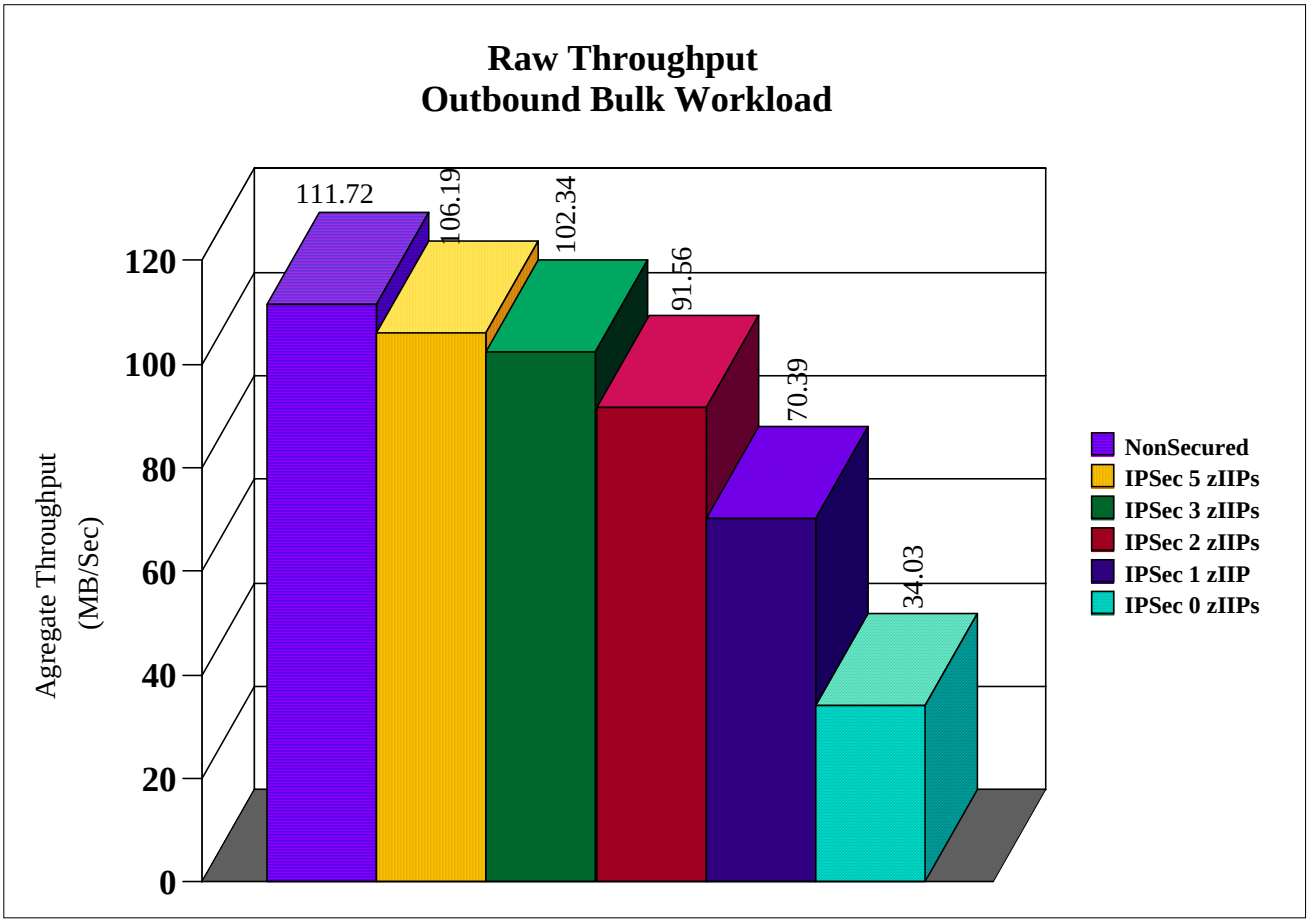

Figure 4: Raw Throughput - Outbound Bulk Data

Interactive Traffic Pattern

Interactive workloads will be less affected by IPSec than will the bulk workloads. There are two reasons for this:

- The interactive workloads we've studied tend to move relatively small amounts of data. (Less than 1500 bytes per network round trip is very common.) Since IPSec processing cost is approximately linear with the amount of user data carried per transaction, there's naturally less IPSec overhead for small-data applications than for large-data applications.

- The interactive workloads we've studied also have a "thicker" application layer. The network pathlength represents a smaller portion of total system pathlength for these applications than we see with the common bulk data applications, so an increase in the communications pathlength (for IPSec) is somewhat diluted by all the pathlength incurred within the application.

Since interactive workloads are less affected by the presence of IPSec, it should be expected the effectiveness of zIIP Assisted IPSec for interactive applications will be less pronounced than we saw with the bulk workloads. One additional factor comes into play with most interactive workloads -- *the outbound data path may be executed in TCB mode (on the application's thread), meaning those IPSec cycles are not zIIP eligible.*

Even given the above factors, our measurements below (10 concurrent interactive TCP connections sending and receiving 100 bytes) show that zIIP Assisted IPSec does provide a significant reduction in general purpose CP busy for secured interactive workloads. Without zIIP, enabling IPSec results in approximately 3X the general purpose CP consumption vs nonsecured. With a zIIP in the configuration, the increase in general purpose CP utilization has dropped to 2X (vs nonsecure). Note - this workload was relatively light, and we really didn't need a second zIIP.

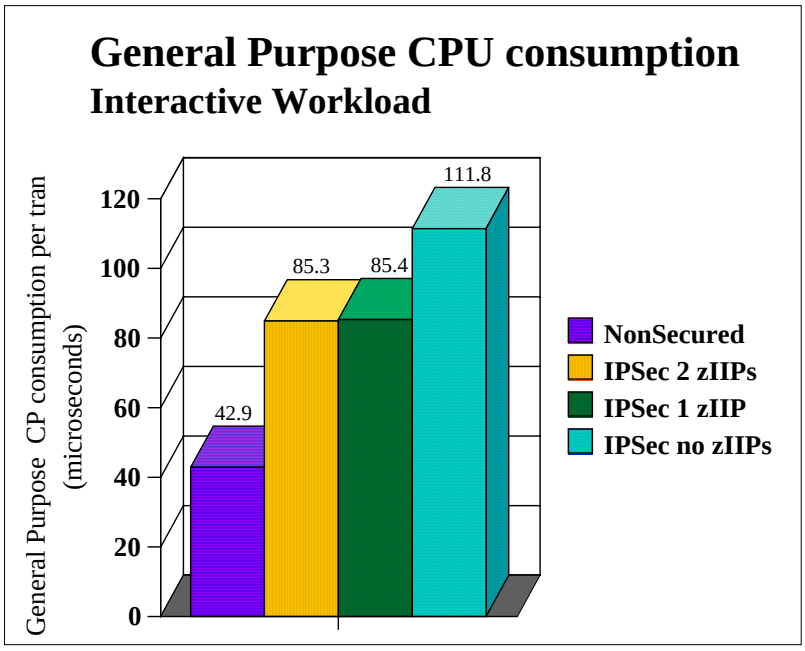

Figure 5: Network CPU Consumption per Transaction - Interactive Workload

Application vs Network CPU Consumption for Interactive Workloads

When operating over a non-secure (non-IPSec) network, some common interactive applications have an application-to-network CPU consumption ratio of approximately 12:1 (i.e., for every 12 cycles of pure application processing, there is 1 cycle of TCP/IP communications processing). Applying this 12:1 ratio to the figure 5 IPSec 1-zIIP data, the additional 42.4 cpu microseconds consumed per transaction (on the general purpose CPs, related to IPSec processing) will translate to a 7.5% increase in general purpose CP busy. By contrast, without zIIPs in the configuration, the additional 68.9 cpu microseconds consumed per transaction will translate to a 12.5% increase in general purpose CP busy.

Throughput for our IPSec interactive workload is relatively flat, regardless of number of zIIPs (as displayed below in figure 6). This is to be expected, since our benchmark did not raise CPU consumption very high in any of these configurations. So we have no unusual processing constraint in any of the configurations. The extra overhead of IPSec processing results in added latency which appears to be flat across our various configurations, resulting in approximately 24% lower throughput for IPSec vs nonsecured.

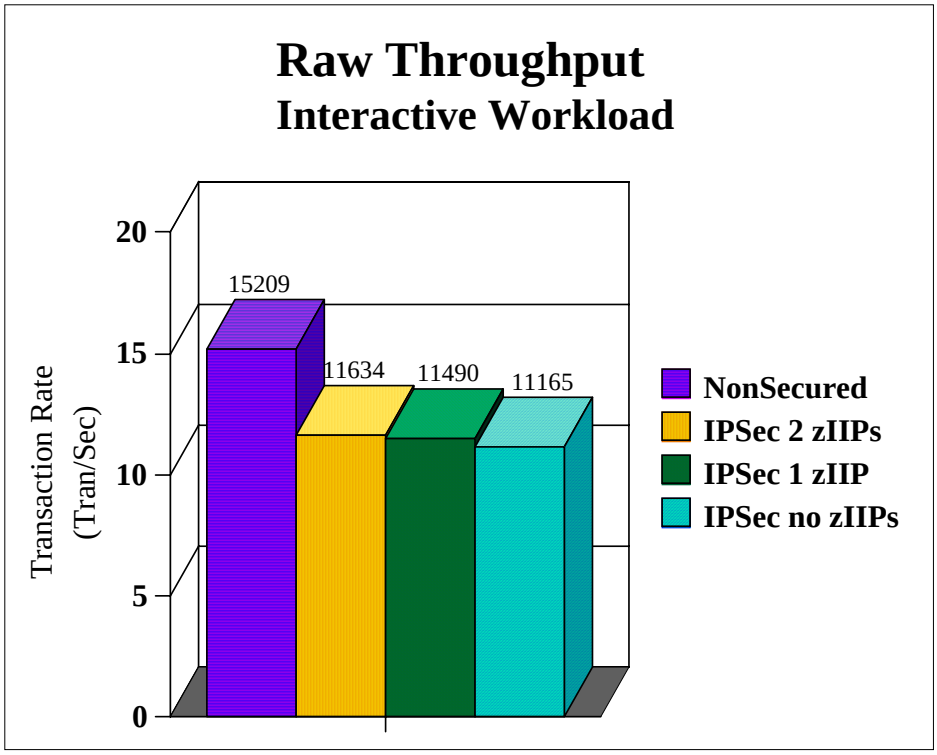

Figure 6: Raw Throughput - Interactive Workload