IBM z Systems

# TPFUG – z/TPF Communications and Security

Jamie Farmer, TPF Development Lab

March 23, 2015

IBM

# Disclaimer

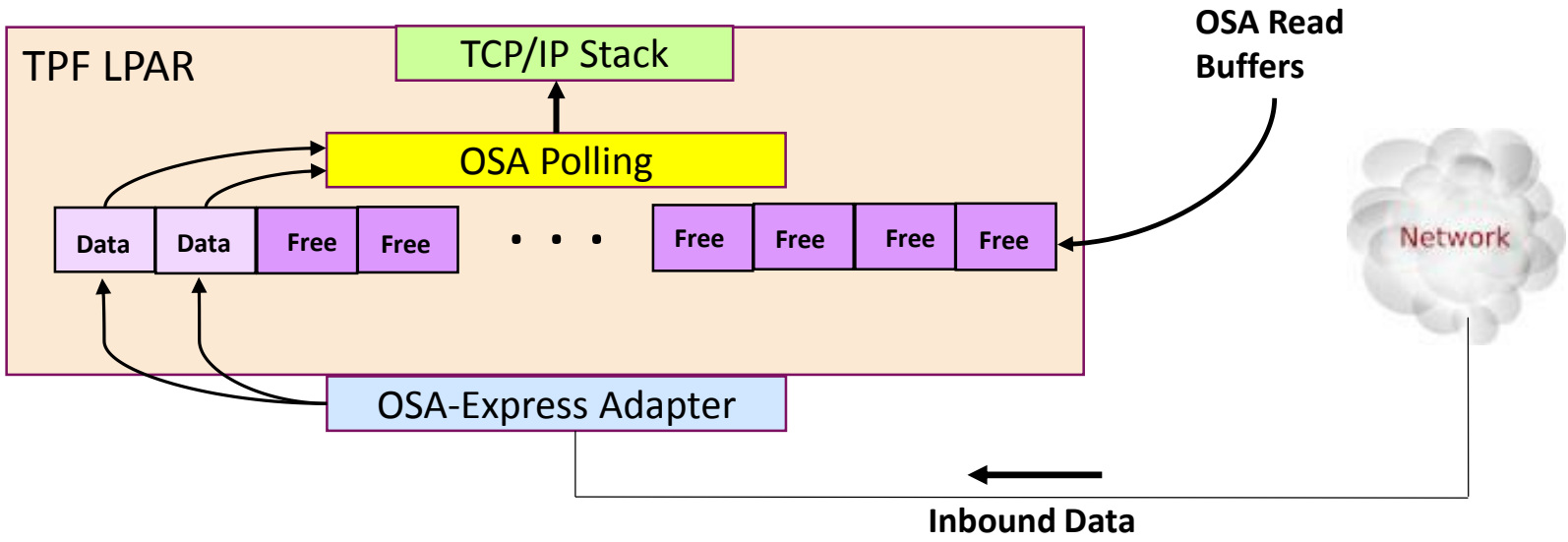
- Any reference to future plans are for planning purposes only. IBM reserves the right to change those plans at its discretion. Any reliance on such a disclosure is solely at your own risk. IBM makes no commitment to provide additional information in the future.

# Agenda

- OSA-Express and TCP/IP Enhancements
- z/TPF HTTP Server Enhancements
- Cryptography and Security Enhancements
- Demonstrating the Benefits of Running z/TPF and zLinux as LPARs in a “z System”

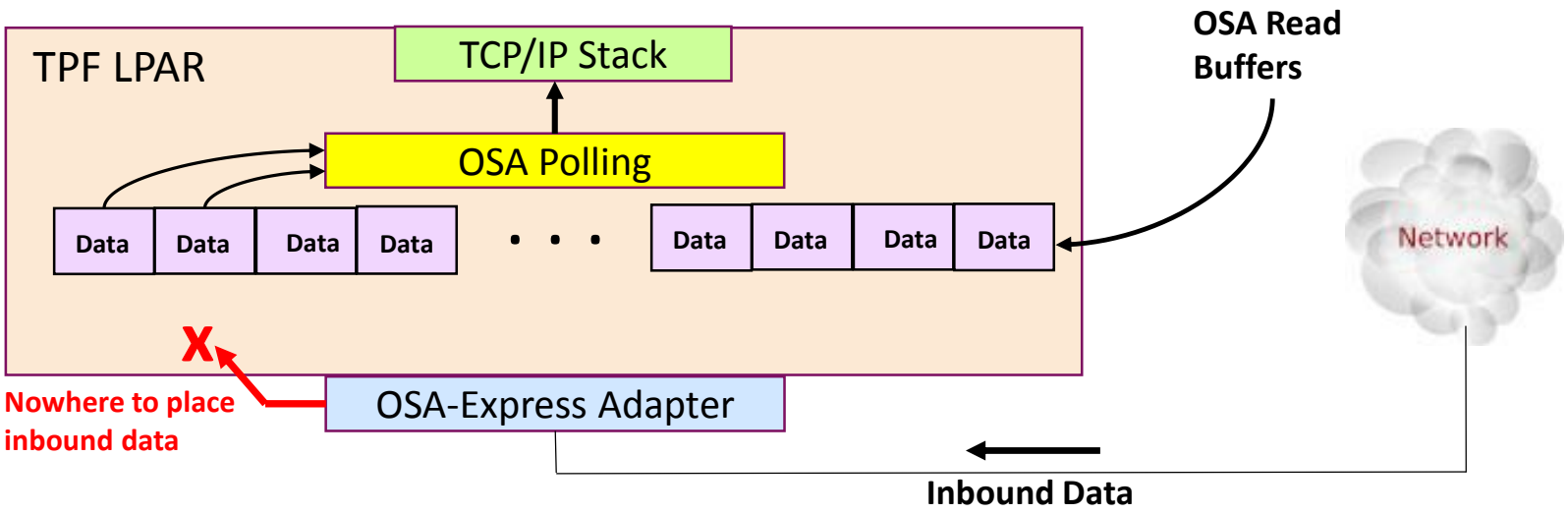
# OSA-Express Enhancements

# OSA Read Buffer Usage




- Shared memory is used to exchange inbound data from OSA to TPF.
- OSA places data received from the network into a given read buffer
- TPF's OSA polling processing pulls the data from the OSA read buffer and introduces the data into the system
  - OSA polling is called each time through the CPU loop and during dump processing
- Number of OSA Read buffers is configurable – 16, 32, 64
  - Wraparound set of buffers

# All OSA Read Buffers Full Condition




- OSA card will begin to queue data waiting for an OSA read buffer to become free.
  - If queue becomes too large, OSA will drop the inbound data
- Occurs when OSA polling is not being called frequently enough for the inbound message rate, for example during system error processing
  - Delivered APAR PJ42029 to poll during system error processing more frequently
- Rare occurrence, but currently there is no indication that this condition occurred.

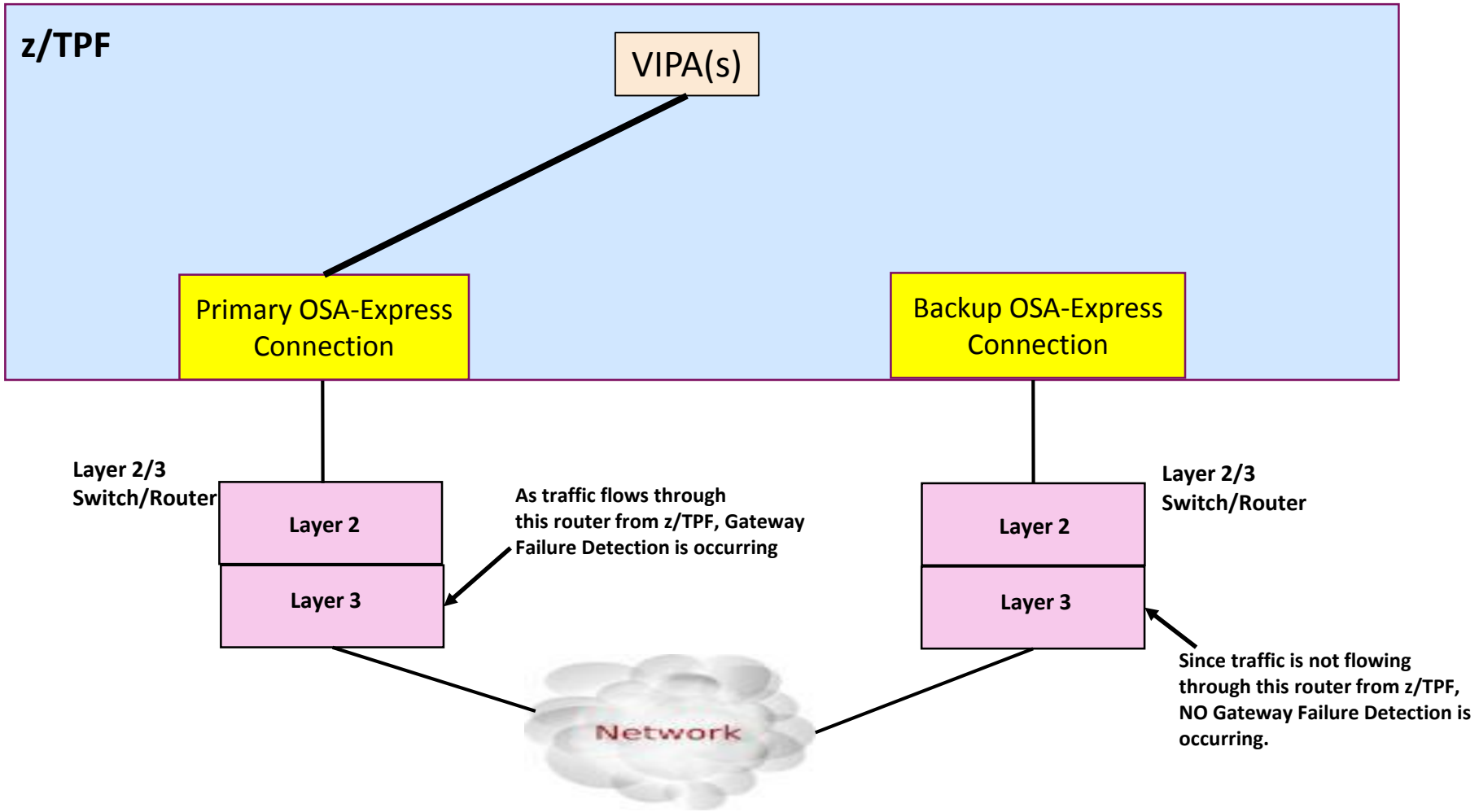
# OSA Read Buffers Full Enhancement – PJ42415

- New warning message displayed to the console when OSA read buffers full condition occurs.
  - Driven from OSA polling when invoked and all read buffers contain data
  - Message displayed at most once every 30 seconds to prevent flooding the console
- Using this message a user can more easily determine if read buffer full conditions were the cause of some network disruption.
- If this condition is hit frequently, could increase the number of OSA read buffers.
  - OSABUFF parameter in CTK2

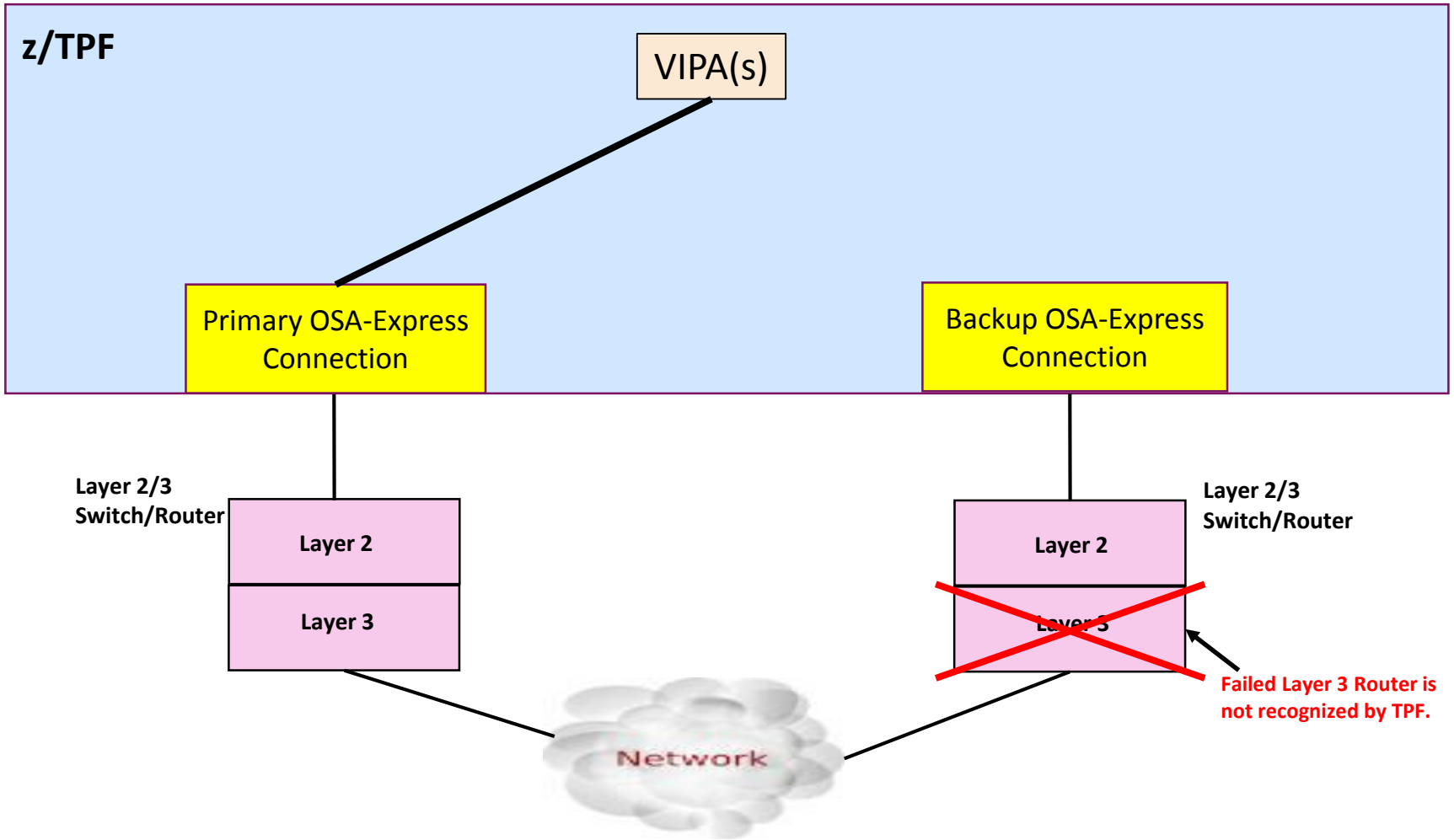
# Current OSA-Express Gateway Failure Detection

- An OSA-Express connection can have up to two default gateways defined to it.
  - Used as the first-hop for traffic destined outside of the subnet for this OSA-Express connection.
- The detection of failed gateways for an OSA-Express connection only occurs when data is sent through that gateway.
  - Backup OSA-Express connections do not have the gateways monitored.
- Failures of gateways defined to a backup OSA-Express connection are only detected when traffic is swung to that OSA-Express connection.

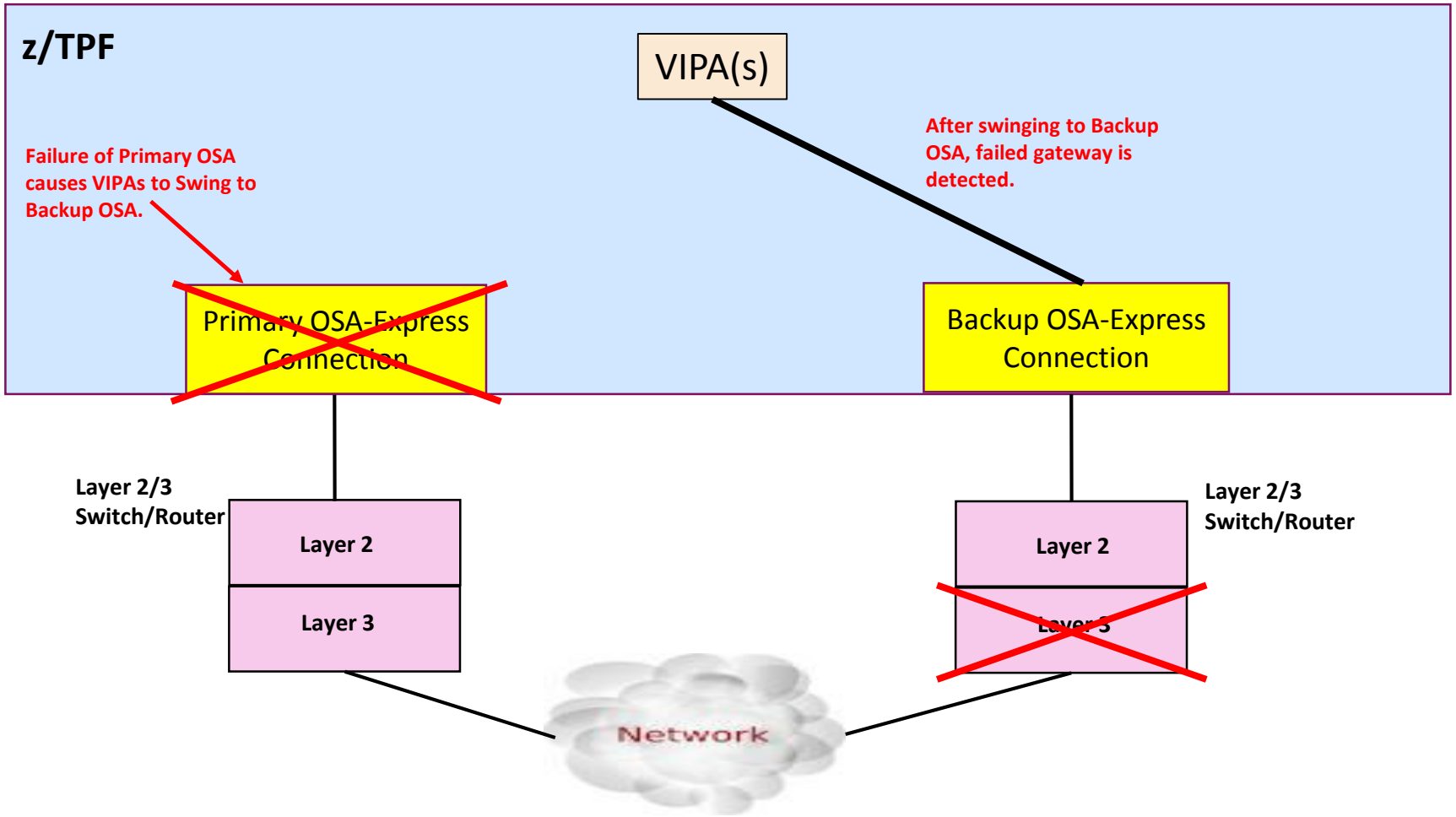
# Example of a Typical OSA-Express Setup



# Example of a Failed Gateway on Backup OSA-Express Connection



# Primary OSA-Express Connection Failure




**Lost connectivity to the External Network!!!**

# Detection of Gateway Failure on Backup OSA-Express Connections – PJ42322

- z/TPF now performs gateway failure detection on backup OSA-Express connections.
  - Allowing users to detect problems with gateways and correct them before they are needed.
- New messages have been created to notify operator of gateway failures on backup OSA-Express connections.
  - Message indicating an individual gateway failure

```
OSA00028E  GATEWAY-1.1.1.1 FOR BACKUP OSA-OSABACK FAILED
```

  - Message indicating when all gateways on a backup OSA have failed

```
OSA00029A  NO DEFAULT GATEWAYS ACTIVE FOR BACKUP OSA-OSABACK
```

# Expansion of z/TPF TCP/IP Client Port Range

# Client Port Range for z/TPF

- Every IP address on any system contains a pool of port numbers that are used for most client sockets
  - Used when the local port is not relevant
- Currently, z/TPF uses port numbers 1024-5000 as the client port range
  - Conflicts with many well known server ports (MQ uses port 1414)
  - Number of client sockets is limited to approximately 4000 per IP address.
  - Limited port range could cause port number reuse issues on some remote systems.

# Expansion of Client Port Range – PJ42623

- The z/TPF client port range now uses port numbers 49152-65535.
  - Compliance with RFC 6335 regarding port number assignments.
- Allows for over 16,000 client port numbers per IP address.
- Further reduces conflicts with server port numbers
  - Well defined servers are generally not in this range
- Reduces port number reuse issues on remote platforms

- With PJ42623 applied or without, it is recommended servers be defined in the Network Services Database (NSD)
  - Servers defined in the new client port number range (49152-65535) are strongly recommended to be defined in the NSD
    - Prevents clients on z/TPF from using ports defined in the NSD

# z/TPF HTTP Server Enhancements

# z/TPF HTTP Server - Background

- z/TPF supports the Apache Server
  - Full functioning HTTP Server
- Working through the TPFUG, customers asked for a lightweight HTTP server for message transport.
  - Including support for asynchronous message processing
    - ECB receiving the request does not need to be the ECB sending the reply
- APAR PJ39252 delivered support for the z/TPF HTTP Server
  - PJ41171 added SSL support to the z/TPF HTTP Server
- Initial usage of the z/TPF HTTP Server identified some new requirements
  - Required for adoption and usability
  - Will continue to be the lightweight HTTP Server solution for z/TPF

# PUT 11 - z/TPF HTTP Server Enhancements (PJ42624)

**Version HTTP/1.0 Support**

- Original z/TPF HTTP Server support allowed for clients using HTTP/1.1 version of the protocol
  - Clients still exist using the HTTP/1.0 version of the protocol.
- PJ42624 adds support for clients that want to use HTTP/1.0 version.
  - Allow the receipt of HTTP/1.0 version requests and the z/TPF HTTP Server will build HTTP/1.0 responses
  - Support for persistent connections using HTTP/1.0
    - Requires support for Connection: keep-alive and Keep-Alive headers

# More PUT 11 z/TPF HTTP Server Enhancements (PJ42624)

## Support of Additional HTTP Methods

- Original z/TPF HTTP Server only allowed requests with either the GET or POST methods
  - PJ42624 adds support for the HEAD, PUT, DELETE and OPTIONS methods on HTTP requests

## Support for Associating User Data With an HTTP Connection

- PJ42624 added support to associate user data with an HTTP connection
  - Supplying an 8-byte user data area
  - Useful in asynchronous message processing to pass data from request ECB to the response ECB.
- New tpf_httpUserData() API to set or retrieve the user data associated with the HTTP connection

# PUT 12 - z/TPF HTTP Server Enhancements (PJ42695)

**Chunked Transport Encoding Support**

- Chunked Transport Encoding allows an HTTP node to send a single logical HTTP message as multiple HTTP requests.
  - Generally used by intermediate nodes for performance to start sending an HTTP message without having to wait for the entire message to arrive.
- PJ42695 added support for chunked inbound HTTP requests
  - The z/TPF HTTP Server doesn't ever send chunked encoded messages

# More PUT 12 z/TPF HTTP Server Enhancements (PJ42695)

**Support of Expect 100-continue Header**

- The Expect 100-continue header is a way for clients to ask a server if it can accept an HTTP request
  - Usually sent by client as it is building the rest of the message body
  - Client waits a period of time for a 100 status code response from the server – indicating the server can accept the message
    - If 100 status code response is not received, client sends entire message anyway
  - Before PJ42695, z/TPF HTTP Server did not reply to Expect 100-continue header
- With PJ42695, the z/TPF HTTP Server responds to an Expect 100-continue header
  - Recommended to disable this function in clients when possible for performance.

# IBM z13 Crypto Express5S Support

# The z13 Processor Introduces Crypto Express5S

- March 10, 2015 - IBM System z13 became generally available (GA)
- IBM introduces the new Crypto Exress5S hardware accelerator
- Crypto Express5S is the only supported Cryptographic Adapter for IBM System z13
  - Older Cryptographic Adapters like Crypto Express4S are not supported on IBM System z13

# z/TPF Support of Crypto Express5S – PJ42625

- Provides support for z/TPF use of Crypto Express5S in accelerator mode
- z/TPF uses Crypto Express5S to accelerate RSA operations in hardware
  - SSL Session Establishment and RSA application APIs
- If you are using cryptographic accelerators on z/TPF
  - PJ42625 is required before moving to IBM System z13

# OpenSSL Update

# OpenSSL Vulnerabilities

- Over the past year, vulnerabilities have been identified against the SSL protocol.
  - Some of these apply to the SSL support on z/TPF.
  - APARs were delivered to address these vulnerabilities
- **PJ42658 & PJ43014**
  - CVE-0214-3566 (POODLE Exposure)
    - A Man in the Middle Vulnerability
  - CVE-0214-3568
    - Enforcement of no-ssl3 build option exposure
- **PJ42340**
  - CVE-2014-0224
    - A Man in the Middle Vulnerability with ChangeCipherSpec

# Upgrading OpenSSL to Latest Version

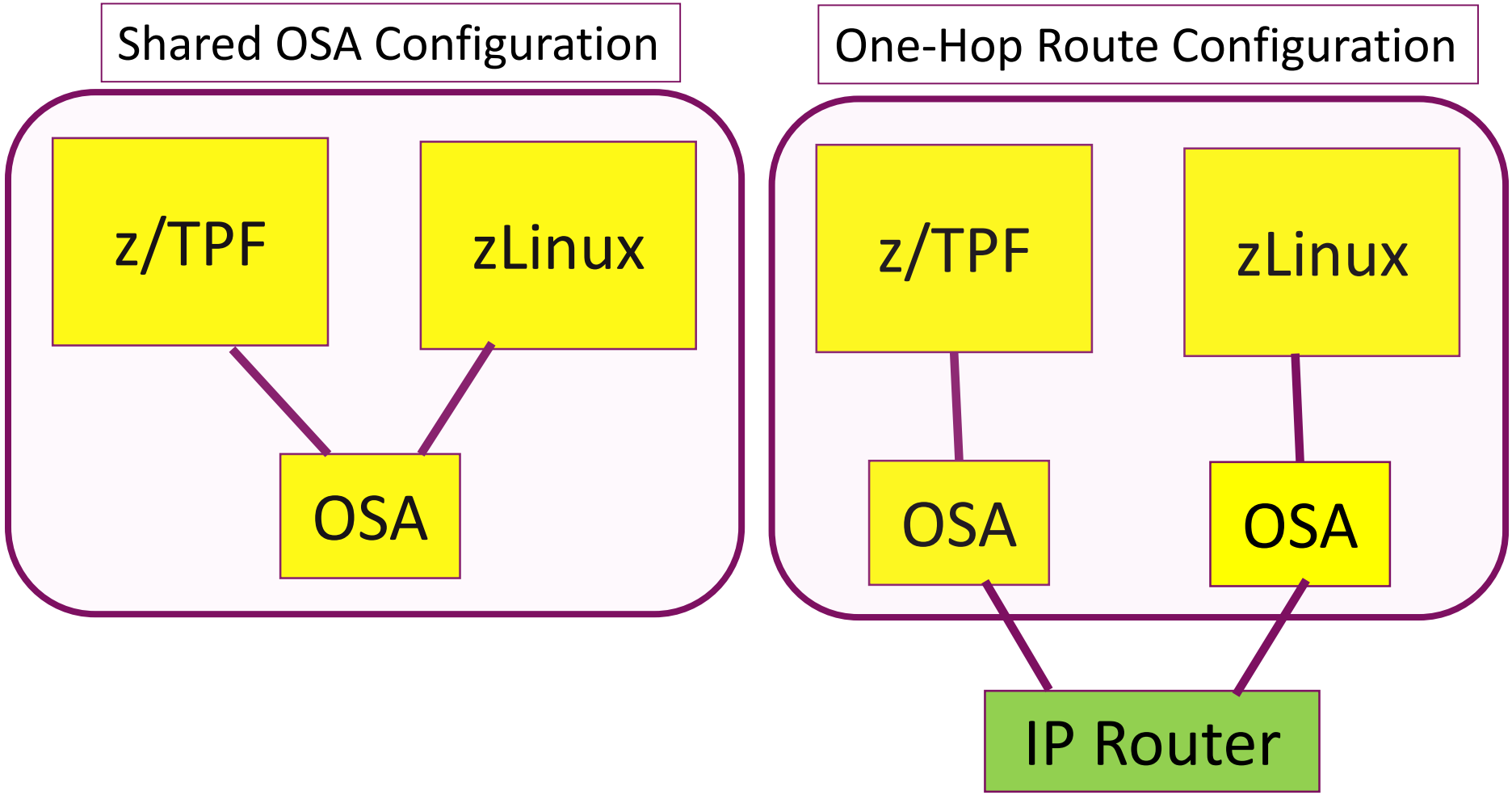
- Effort underway to upgrade to the latest OpenSSL level on z/TPF
  - OpenSSL 1.0.2 - GA date January 22, 2015
- z/TPF has made modifications to OpenSSL
  - Support for shared SSL
  - Asynchronous I/O – SSL_aor()
  - Hardware acceleration support
- Part of the re-port is to redesign how these modifications are applied
  - Making it easier to port newer versions in the future
- What we support today in OpenSSL will be supported in new version
- New version of OpenSSL supports additional features
  - Looking for sponsored users to provide feedback on what additional features are needed.

# Demonstrating the Benefits of Running z/TPF and zLinux as LPARs in a “z System”

# LPAR-LPAR Shared OSA Latency Testing

- Several customers are looking at exploiting Linux systems interfacing with z/TPF
  - Question being should zLinux or distributed x86 Linux be used?
- Measure the response time from an application perspective between two LPARs in the same z System sharing an OSA card.
  - Compare with communicating to a node that is 1-hop away
- Environment for Testing
  - All tests performed on a zEC12 machine
    - z/TPF LPAR : 1 dedicated CP
    - zLinux LPAR : 1 dedicated IFL
- Driver calculates round trip time from an application perspective
- Only difference between shared OSA and 1-hop tests was the network path used

# OSA Latency Test - Environment Comparison



# OSA Latency Test – Summary

- Tests were performed using hundreds -> 10's of thousands messages per second
- Round Trip Time is an application RTT – time application sent a request to time it received a response

| Message Size (Bytes) | Shared OSA Average RTT (microseconds) | 1-Hop Route Average RTT (microseconds) | Average RTT Ratio | Average Extra Time Per Message (microseconds) |
|---|---|---|---|---|
| 100 | 204 | 474 | 2.3 | 270 |
| 500 | 167 | 446 | 2.7 | 279 |
| 1400 | 154 | 924 | 6.0 | 770 |
| 5000 | 551 | 2191 | 4.0 | 1640 |
| 10,000 | 412 | 2549 | 6.2 | 2137 |
| 20,000 | 365 | 4350 | 11.9 | 3985 |

# OSA Latency Test – Analysis of results

- Latency of a LPAR-LPAR sharing an OSA is better than the latency of going through 1-hop
- Other factors to consider include:
  - Processing time to service a request
  - Size of messages exchanged for a given transaction
  - Number of messages exchanged per transaction
- If distributed Linux is more than 1-hop away the latency increases

# Communications / Security Enhancement Summary

- OSA-Express Enhancements
  - Detecting when all OSA read buffers are full – **PJ42415**
  - Detecting OSA gateway failures on backup OSA-Express connections – **PJ42322**
- Expansion of TCP/IP Client Port Range – **PJ42623**
- z/TPF HTTP Server Enhancements - **PJ42624(PUT 11) & PJ42695 (PUT 12)**
- Crypto Express5S Support for z13 – **PJ42625**
- OpenSSL Update
  - SSL Vulnerabilities – **PJ42340 & PJ42658**
  - OpenSSL Version 1.0.2 – ***In Progress***
- Demonstrating the Benefits of Running z/TPF and zLinux as LPARs in a "z System"

# Trademarks

- IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.
- *(Include any special attribution statements as required – see Trademark guidelines on https://w3-03.ibm.com/chq/legal/lis.nsf/lawdoc/5A84050DEC58FE31852576850074BB32?OpenDocument#Developing%20the%20Special%20Non-IBM%20Tr)*

**Notes**

- Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage 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 here.
- All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.
- This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
- All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
- Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.
- Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.
- This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.