Network Performance: An MPE/iX Overview

Jeff Bandle
HP MPE/iX Networking Architect

CONTENTS

- General Networking
  - Common networking terms
  - Networking concepts independent of MPE/iX
- MPE/iX Specific Networking
  - Overview of MPE/iX networking stacks
  - Ideas for performance changes on MPE/iX
- System Performance
  - How MPE/iX networking performance affects the system

INTRODUCTION

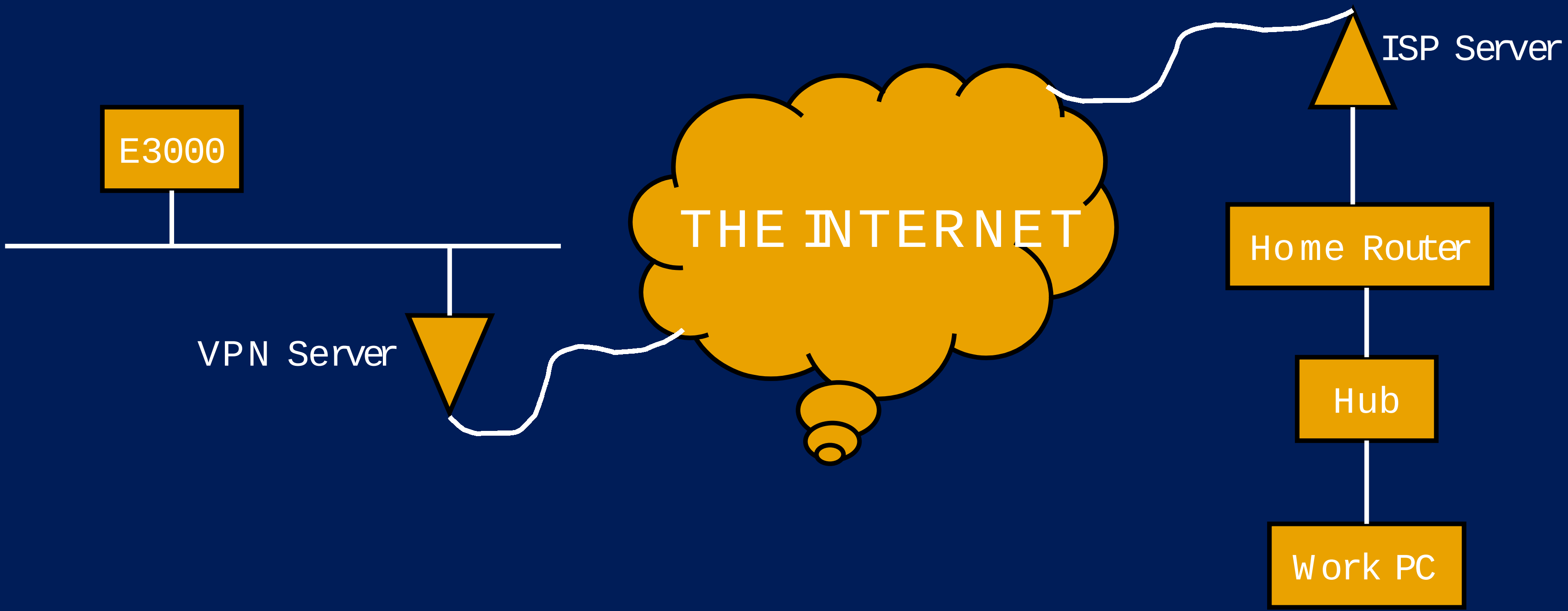
- What is performance?
  - Bandwidth
  - Response time
  - System
- General Networking vs. System Specific Networking

GENERAL NETWORKING

- Network setup complexity is a factor
  - Simple Network – Fewer layers to propagate data

GENERAL NETWORKING

- Network setup complexity
  - Complex network – More layers/hardware delays data propagation
  - Study of "pings" to 3000 international sites – 150 ms avg.



GENERAL NETWORKING

- Use or Routers, Switches and Hubs
  - Hub - A **hub** is a small, simple, inexpensive device that joins multiple computers together at a low-level network protocol layer.
  - Switch - A **switch** is a small device that joins multiple computers together at a low-level network protocol layer. Technically, switches operate at layer two (Data Link Layer) of the OSI model.
  - Router - A **router** is a physical device that joins multiple networks together. Technically, a router is a "layer 3 gateway," meaning that it connects networks (as gateways do), and that it operates at the network layer of the OSI model.

GENERAL NETWORKING

- Common tools to check complex networking
  - Ping

```
ping [-oprv] [-i address] [-t ttl] host packet-size [[-n] count]

: ping nack.cup.hp.com
PING nack.cup.hp.com: 64 byte packets
64 bytes from 15.13.195.50: icmp_seq=0. time=1. ms
64 bytes from 15.13.195.50: icmp_seq=1. time=1. ms
64 bytes from 15.13.195.50: icmp_seq=2. time=1. ms
64 bytes from 15.13.195.50: icmp_seq=3. time=1. ms
64 bytes from 15.13.195.50: icmp_seq=4. time=1. ms
64 bytes from 15.13.195.50: icmp_seq=5. time=1. ms
64 bytes from 15.13.195.50: icmp_seq=6. time=1. ms
64 bytes from 15.13.195.50: icmp_seq=7. time=1. ms
```

GENERAL NETWORKING

- Common tools to check complex networking
  - Traceroute

```
traceroute [-dnrv] [-w wait] [-m max_ttl] [-p port#] [-q nqueries] [-s src_addr] host [data size]

traceroute to cup.hp.com (15.75.208.53), 30 hops max, 20 byte packets
 1 cup47amethyst-oae-gw2.cup.hp.com (15.244.72.1)       1 ms   1 ms   1 ms
 2 hpda.cup.hp.com (15.75.208.53)                       1 ms   1 ms   1 ms
```

GENERAL NETWORKING

- Traceroute (cont)

```
traceroute to atl.hp.com (15.45.88.30), 30 hops max, 20 byte packets
 1 cup47amethyst-oae-gw2.cup.hp.com (15.244.72.1)  1 ms  1 ms  1 ms
 2 cup44-gw.cup.hp.com (15.13.177.65)              1 ms  1 ms  1 ms
 3 cupgwb01-legs1.cup.hp.com (15.61.211.71)        1 ms  1 ms  1 ms
 4 palgwb02-p7-4.americas.hp.net (15.243.170.45)   2 ms  1 ms  1 ms
 5 atlgwb02-p6-1.americas.hp.net (15.235.138.17)   60 ms  60 ms  60 ms
 6 atlgwb03-vbb102.americas.hp.net (15.227.140.7)  60 ms  60 ms  60 ms
 7 atldcrfc5.tio.atl.hp.com (15.41.16.205)    61 ms  60 ms  60 ms
 8 i3107at1.atl.hp.com (15.45.88.34)          60 ms  60 ms  60 ms
```

GENERAL NETWORKING

- Traceroute (cont)

```
traceroute to www-dev.bri.hp.com (15.144.120.100), 30 hops max, 20 byte packets
 1 cup47amethyst-oae-gw2.cup.hp.com (15.244.72.1) 1 ms  1 ms  1 ms
 2 cup44-gw.cup.hp.com (15.13.177.65)  1 ms  1 ms  1 ms
 3 cupgwb01-legs1.cup.hp.com (15.61.211.71)  1 ms  1 ms  1 ms
 4 palgwb02-p7-4.americas.hp.net (15.243.170.45)  2 ms  2 ms  1 ms
 5 atlgwb02-p6-1.americas.hp.net (15.235.138.17)  60 ms  60 ms  61 ms
 6 15.227.138.42 (15.227.138.42)  183 ms  204 ms  183 ms
 7 bragwb02.europe.hp.net (15.203.204.2)  183 ms  184 ms  184 ms
 8 15.203.202.18 (15.203.202.18)  188 ms  227 ms  188 ms
 9 15.144.16.4 (15.144.16.4)  189 ms  188 ms  189 ms
10 www-dev.bri.hp.com (15.144.120.100)  189 ms  188 ms  188 ms
```

GENERAL NETWORKING

- Hardware Potential Performance Changes
  - Routers –
    - Use router tools to analyze networking traffic
    - Readjust traffic loads to balance across different connections (if possible)
    - Use tools to verify memory usage is not being compromised for connections

GENERAL NETWORKING

- Hardware Potential Performance Changes
  - Routers –
    - Since routers have intelligence inside of them, data is stored in buffers
    - Common performance problems related to buffer allocation

      Middle buffers, 600 bytes (total 150, permanent 25): 147 in free list (10 min, 150 max allowed)
      61351931 hits, 137912 misses, 51605 trims, 51730 created 91652 failures (0 no memory)

      - **permanent**: take the number of total buffers in a pool and add about 20%.
      - **min-free**: set min-free to about 20-30% of the permanent number of allocated buffers in the pool.
      - **max-free**: set max-free to something greater than the sum of permanents and minimums

        buffer middle permanent 180

        buffer middle min-free 50

        buffer middle max-free 235
      - Adjust for traffic burst
        - Slow traffic – Min free goes up
        - Fast traffic – Permanent goes up

GENERAL NETWORKING

- Hardware Potential Performance Changes
  - Switches
    - Dependant on type and brand, changeable parameters vary
      - Change speed (10/100/1000 kbps) to match other devices
      - Change Duplex level (Half/Full to relieve conflicts)
      - Autonegotiation isnt' full foolproof (If possible nail port paramters)
      - Link multiple ports together in a trunk (not all switches)
        - Limited to direct connections with peer switch

GENERAL NETWORKING

- Hardware Potential Performance Changes
  - Hubs
    - Hubs usually don't have parameters that can be changed for performance
    - If they are bundled with a switch, use switch information to make changes
    - Most hubs, by default, are half-duplex in operation
      - Need to validate that connections into the hub are half duplex

GENERAL NETWORKING

- Other Potential Issues
  - Difference in software standards
    - HTTP 1.0 vs. HTTP 1.1 – Persistent connections
    - Large data frames – Not standard in all hardware
  - Systems need to work to keep “pipes” full
    - Introduction of Fiberchannel starting to push 100BT
- Other places for tips and tricks
  - www.web100.org - Pointers to tools for performance analysis
  - www.compnetworking.about.com – High level info on networks
  - www.racticallynetwork.com - SOHO networking information

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stacks Made of Multiple Layers

F Intrinsics	Sockets/Net IPC APIs
ADCP	Telnet
AFCP	TCP/IP/UDP
Network Links	

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – Links
  - 100BT/100VG – Full Duplex vs. Half Duplex
    - Full Duplex allows for send and receive traffic at the same time
    - 100 VG had some advantages but lost out on marketing side VHS vs. Beta
    - Full Duplex can be affected by connections
    - Full Duplex can be affected by application design
    - MP systems also affect Full Duplex behavior

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – Links

  - ACC – WAN Link
    - Speeds limited by connection medium
      - Phone speeds and satellite technologies – 2 mbps possible
    - Best used as an access point into a network, not as interconnect between systems.

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – Transports
  - AFCP – Used to communicate with DTC device – HP Proprietary
    - Configuration within NMMGR to change parameters
    - After selecting DTC to configure, select TUNE DTC option

      Set 1: Normal timer mode

      Set 2: Short retransmission timer mode

      Set 3: Long retransmission timer mode

      Set 4: Variable timer mode

      Set 5: MPE XL Release 1.2 timer mode

      Set 6: MPE XL Release 2.1 timer mode

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – Transports
  - TCP/IP – Used to communicate with open standards based devices
    - Configuration with NMMGR
    - Within the NS->UNGUIDED CONFIG->NETXPORT->GPROT->TC

[1024] Maximum Number of Connections

[2] Retransmission Interval Lower Bound (Secs)

[180] Maximum Time to Wait For Remote Response (Sec)

[4] Initial Retransmission Interval (Secs)

[4] Maximum Retransmissions per Packet

[600] Connection Assurance Interval(Secs)

[4] Maximum Connection Assurance Retransmissions

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – APIs
  - Sockets – Standards based networking connectivity interface
    - Sending data requires use of data buffers
    - Tradeoff between efficiency in application and efficiency in networking
    - Studies seem to point to 1k byte buffers being optimal balance
    - Only works if application can package data.
    - Connection startup/teardown is expensive – AVOID IF POSSIBLE

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – APIs

  – NetIPC – HP Propriety networking connectivity interface
    - Similar to open standards sockets
    - 1k byte buffers are optimal if application allows
    - Fix length data blocks remove need to negotiate buffer length
      – Eliminates an extra IPCRECEIVE call for get length of data
    - Connection startup/teardown is expensive – AVOID IF POSSIBLE

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – Services
  - Telnet – Open standards terminal connectivity
    - Based on very inefficient 1-character transfer mode
    - Most common complaint is character echo response
    - Block mode response is comparable to VT/DTC
    - Character echo improved with Advanced Telnet functionality
      - Requires terminal emulator that supports it.
      - QCTERM as an example

MPE/iX SPECIFIC NETWORKING

- MPE/iX Networking Stack – Services
  - DTC TIO/ADCP – HP Proprietary terminal connectivity
    - Efficient block mode data transfer
    - Higher cost due to needing DTCs and special applications
    - DTSTUNEB can be used to adjust buffer parameters
      - WARNING – Due so at your own risk.
      - Change total number of data buffers created - # per ldev
      - Change maximum number of data buffers useable per ldev – 24 is default

MPE/iX SYSTEM PERFORMANCE to NETWORKING

- Networking connections use resources
  - Data structures for each socket/NetIPC connection
  - Data buffers for each DTC/Telnet connection
  - Timer structures used by all layers
  - Busy connections on small systems can exhaust resources
    - "Fake" system by creating more "dummy" devices

MPE/iX SYSTEM PERFORMANCE to NETWORKING

- System is very busy servicing interrupts

  – Tradeoff between "smart" cards and "dumb" cards
    - Network adapters could do more work
    - Newer cards are cheaper, but system needs to do processing
    - High LAN traffic situations see this more often as problem
    - Solution is to get more CPU
    - Efficiencies have been introduced into MPE/iX stacks

MPE/iX SYSTEM PERFORMANCE to NETWORKING

- Connectivity mix can affect system performance
  - VT vs. DTC vs. Telnet
    - DTC is most efficient
      - Handles data away from the system
      - Very few data transfers per I/O request
    - VT is efficient also because of HP prorietary
      - Has limits because of sitting on TCP/IP stack
      - Requires driver applications on sending and receiving systems

MPE/iX SYSTEM PERFORMANCE to NETWORKING

- Connectivity mix can affect system performance (cont)
  - Telnet is least efficient because of need to support open standards
    - Block mode applications (VPLUS) comparable to VT
      - Telnet is 90% as efficient as VT in block mode
    - CI commands most overhead for Telnet – 1 character at a time response
      - Telnet is 70% as efficient as VT in character mode

MPE/iX SYSTEM PERFORMANCE to NETWORKING

- Check for application type with regards to I/O
  - Block mode access vs. character mode access
  - Internal studies show that frame size is either:
    - Very small - < 140 bytes
    - Max value – 1500 bytes
    - Nothing in between
  - If many character mode applications being used, system network will bog down
  - Move to block mode alternative, higher CPU speeds or offload to other systems

MPE/iX SYSTEM PERFORMANCE to NETWORKING

- Check for application type with regards to I/O (cont)

  - Check to see how networking connections are being made
  - Multiple starts/shutdowns for connection are EXPENSIVE
  - On small 918 class system, 15 user test FAFFed system
    - Higher CPU
    - Different connectivity methods
    - More memory

WRAPUP

- If you suspect networking performance problems, what can you do?
  - Characterize problem – can't connect, lost packets, system is bogging down
  - Understand where heaviest use is coming from
    - Single application use – Can application/parms be tweaked to ease performance pressure?
    - Multiple users rapidly connecting to system – Can users be directed to connect by differing methods
    - Network is experiencing problems – Isolate segment that is causing problem
      - Check router, switch for potential problems

invent