Overview

• A quick skate across the top of an entire suite of technology-based issues that exist within the IP architecture:
  – IP Carriage
  – IP, TCP and UDP
  – IP Addresses
  – IP V6
  – DNS
  – IP Routing
  – Network Management
  – VPNs
  – MPLS
  – VOIP
  – Wireless
IP Carriage Architectures

Issues in designing an efficient high speed IP backbone network
Carriage Networks and IP packets

- Each speed shift places greater functionality into the IP packet header and requires fewer services from the carriage system.
- IP networks need to get faster, not smarter.

<table>
<thead>
<tr>
<th>NETWORK</th>
<th>PACKET</th>
</tr>
</thead>
<tbody>
<tr>
<td>real time bit streams</td>
<td>asynchronous data packet flows</td>
</tr>
<tr>
<td>network data clock</td>
<td>per-packet preamble data clock</td>
</tr>
<tr>
<td>end-to-end circuits</td>
<td>address headers and destination routing</td>
</tr>
<tr>
<td>fixed resource segmentation</td>
<td>variable resource segmentation</td>
</tr>
<tr>
<td>network capacity management</td>
<td>adaptive dynamic utilization</td>
</tr>
<tr>
<td>single service platform</td>
<td>multi-service payloads</td>
</tr>
</tbody>
</table>
The Evolution of the IP Transport Stack

Multiplexing, protection and management at every layer

Higher Speed, Lower cost, complexity and overhead
Engineering Internet Backbone Networks

- Data Networks were originally designed as overlays on the PSTN network
- As the Internet evolved its demands for carriage capacity have increased more than one million-fold
  - This massive increase in volume requires rethinking how to efficiently build data networks
- This has lead to engineering data networks without an underlying PSTN
  - Such IP trunk networks are very recent developments to the carrier engineering domain

- Current High Speed IP platform architectures consist of:
  - DWDM fibre systems
  - 10G optical channels
  - 10GiGE Ethernet framing
  - Multi-router POPs
  - Load distribution through topology design and ISIS link metrics
Faster Core IP Networks

• From Silicon to Photons
  – Reduce the number of optical / electrical conversions in order to increase network speed

• The optical switched backbone
  – Gigabit to Terabit network systems using multi-wavelength optical systems
  – Single hop routing to multi-hop optical Traffic-Engineering control planes
A whole new Terminology Set: Gigabit Networking Technology Elements

- **Ethernet packet frames**
  - Faster Ethernet: 100mFE, GigE, 10GigE
  - VLANs: 802.1Q
  - Rings (802.17) and T-Bit Fast Switches

- **Optical Transports**
  - CWDM / DWDM
  - Wavelength-Agile Optical Cross-Connect control systems with GMPLS controls

- **Traffic Engineering**
  - Rapid Response, Rapid Convergence IP Routing Systems
  - MPLS to maintain path vector sets
Current High Speed IP Network Architectures

Access Network

DWDM 10G links

GigE VLAN Edge
IP Giga Network Architecture
IP Architecture

- IP is a simple end-to-end overlay level 3 datagram protocol
  - End-to-end header semantics
  - No signalled connection between link level conditions and transport services
  - Universal abstraction of a common simple packet transmission service that has been adapted to operate efficiently over wires, modems, Frame Relay, ATM, Ethernet, broadcast radio, packet radio, satellite circuits, SDH, fibre, pigeons
Yes, Pigeons!

- RFC 1149 “Standard for the transmission of IP datagrams on avian carriers”
- RFC 2549 “IP over Avian Carriers with Quality of Service”
- Implemented in 2001 in Norway
  - [http://www.blug.linux.no/rfc1149/](http://www.blug.linux.no/rfc1149/)
Implementation of BBI2 (RFC 1149 Firewalling)

Bert pulled an all-nighter to implement the ICPM_TRIGGLATE message (top left), resizing the firewall equipment (top right) and presenting the somewhat shocking result of filtering 63 packets to the working group (bottom left).

Bert regrets dropping a wrong packet (bottom right).

Reference:

[BBI2] RFC1149 Firewalling
IP Architecture

- TCP and UDP are DIFFERENT end-to-end transport services
  - UDP is an unreliable datagram service
  - TCP is a flow-controlled reliable stream service
  - Most IP payload is TCP (95% by volume)
  - Real-time services use a UDP base

- UDP and TCP have a widely different operating model
  - TCP attempts to saturate network resources using a cooperative model of congestion limit probing (network-clocking of data transfer)
  - UDP uses an external clocking model that is normally impervious to network conditions
  - The fit is often not entirely comfortable
    - hence the QoS effort to attempt to impose some level of network-based arbitration
IP Architecture Pressures

• Now under some pressure
  – QoS signalling between application and network
  – NATS, ALGs, intercepting caches break end-to-end semantic with middleware
  – IPSEC, SIP, HTTPS tunnels, IPV6 tunnelling (…) now being used to 2\textsuperscript{nd} guess middleware in order to recreate end-to-end associations
  – Transport services under pressure to be more aggressive in recovery vs making UDP more ‘reliable’
  – Identity semantics all confused with application, end-to-end and network level identity assertions

• This new architecture no longer simple, scaleable or efficient
Addresses -- How to get here from there

- Addresses provide information on how to locate something, e.g., what route to take from here to there.
- Internet addresses combine
  - a routing portion, known as the network part
  - a name portion known as the host part
IP Addresses

- IP uses overloaded semantics of an “address”
  - AN IP address is used as an IDENTITY, a LOCATOR and a ROUTING ELEMENT
  - These are separable concepts:
    - What is the best PATH to reach YOUR current LOCATION?
  - IP makes no distinction at present between these three roles
  - Consequent serious issues with Mobility, NATs, SIP, URLs, Security
  - This is common to both V4 and V6
IP V4 Addresses

- V4 remains the overwhelmingly dominant protocol choice
  - 32 bit (4G) address space
    - 50% allocated
    - 25% deployed
    - 5%- 10% utilization density achieved
    - Consumption at a rate of 32M p.a.
    - Anticipated lifespan of a further 10 – 15 years in native mode
    - Indefinite lifespan in NAT mode
IPV6

• “IP with larger addresses”
• Address space requirements are no longer being easily met by IPv4
• This is an issue for high volume deployments including:
  – GPRS mobile
  – 3G Mobile
  – WebTV
  – Pocket IP devices
• IPV6 appears to offer reasonable technology solutions that preserve IP integrity, reduce middleware dependencies and allow full end-to-end IP functionality
• Issues are concerned with co-existence with the IPv4 base and allowing full inter-working between the two protocol domains
IPv6 Strengths

- Larger addresses to match
  - consumer electronics
  - disposable passive devices (labels and tags)
  - automated conversation and distributed control functions
IPv6 Weaknesses

• Not sufficiently “different” from IPv4
  – No ‘value add” to fuel investment in transition
  – Reuses large amounts of V4 infrastructure to there’s an expectation of identical outcomes
    • http://www.kame.net

• Not sufficiently “similar” to IPv4
  – The coupling of address and identity functions in the IP architecture makes transparent address translation a challenge
  – Referential integrity issues – is the DNS protocol independent or loosely/tightly coupled between V6 and V4

• Still working on the technology
  – Address architecture
  – Site-Local addressing
  – Multi-homing
  – Mobility
  – Transition mechanisms
V4 and V6 – direction?

- No change and no widespread adoption of V6 - yet
- Most growth in IP is being absorbed by NATs and DHCP
- Likely deployment model is in vendor-push walled garden deployments with application-specific gateway portals into and out of the V6 domain
- The next 2 years appear to be a critical period for V6 deployment
- The hype surrounding V6 is unhelpful
  - V6 is IP with larger addresses – nothing more
- The lack of production high speed routing code from vendors is frustrating
  - Noone wants to deploy ‘experimental’ code!
Domain Names

• Hierarchical name space with an associated distributed caching database (the “DNS”)

• The DNS:
  – Maps names to IP addresses
  – Maps IP addresses to names
  – Maps service names to other names
  – Maps E.164 numbers to service addresses
  – Can contain unstructured text elements
    • Key signatures
    • Identity
Domain Name Issues

- Single root of the hierarchy
- Control of root by USG
- Short-cut name spaces
- Multi-lingual DNS
- Security and resilience
- Alternative Identity name space (DNSSEC + Dynamic Update)
- Trademarks and IPR issues
- Generic TLDs
Routing

- IP uses a de-coupled routing architecture
  - Routing architectures can (and do) change without disrupting the service platform
- Two level hierarchy
  - Interior routing to undertake topology maintenance and best path identification
  - Exterior routing to undertake connectivity maintenance and conformance to external policies
Routing – Interior Routing

- Predominant use of SPF algorithms for topology maintenance
  - OSPF
  - IS-IS
- Overlay external routes with iBGP
- Little evidence of takeup of MPLS-based approaches
Routing – Exterior Routing

• BGP is the protocol of choice for exterior routing
  – Operator base highly familiar with BGP characteristics and capabilities
  – Easily disrupted
    • Poor security model with massive levels of distributed trust and no coupled authentication mechanisms
  – Poor scaling performance
  – Highly unstable (oscillation and damping)
  – Unresponsive to dynamic changes
  – No TE / QoS Support
    • And none likely!
  – No alternative to field!
Network Management

• SNMP-based architecture
  – In-band management model
  – Query-response polling architecture using a structured set of query variables
  – Problems:
    • Insecure
    • Vulnerable implementations
    • Too simple?
  – Efforts underway to create a successor architecture to SNMP to incorporate better security, lock and confirm actions (mutex plus confirm), shared management state
IP VPNs

• Sharing of a common base packet switching platform by a collection of IP networks
• Issues of integrity of the platform and integrity of the offered IP service to the VPN client
• Critical areas of technology development include
  – MPLS – Multi-Protocol Label Switching
  – MPR – Multi-Protocol Routing
  – VLANS – Virtual LAN Packet Frame formats
  – IPSEC – end-to-end IP authentication and encryption services
  – QoS – various forms of Quality of Service network mechanisms
  – PPP / MPLS / VLAN / VC inter-working – the enterprise-wide VPN service model
  – Dynamic VPN technologies – secure edge-based discovery tools
MPLS

- Where ATM collides with IP
- MPLS is an encapsulation technology that adds a network-specific egress label of a packet, and then uses this for each hop-by-hop switching decision
- Originally thought of as a faster switching technology than IP-level switching. This is not the case
- Now thought of as a more robust mechanism of network-specific encap than <IP in IP>, or <IP in L2TP in IP>
- Has much of the characteristics of a solution looking for a problem:
  - IP-VPNs? IP-TE? IP-QoS? Multi-protocol variants of these?
VOIP

- In theory voice is just another IP application
- In practice it’s a lot harder than that
- Issues of Quality and Signalling
- Quality
  - Voice is a low jitter, low loss, low latency, constant load application
  - TCP is a high jitter, medium loss, variable load transport
  - The problem is to get VOIP into the network without it being unduly impaired by TCP flows
  - Either overprovision the network and minimize the impacts or
  - differentiate the traffic to the network and allow the network elements to treat VOIP packets differently from TCP packets
VOIP

• How can you map the E.164 telephone number space into the Internet environment?
  – Allow VOIP gateways to operate autonomously as an agent of the caller rather than the receiver
  – ENUM technology to use the DNS to map an E.164 number to a URL service location
  – Use the DNS to map the URL service location to an IP address of the service point
  – What happens with NATs?
Wireless

• *In theory*
  – IP makes minimal assumptions about the nature of the transmission medium. IP over wireless works well.

• *In practice*
  – high speed TCP over wireless solutions only works in environments of low radius of coverage and high power
  – TCP performance is highly sensitive to packet loss and extended packet transmission latency
Wireless

- 3G IP-based wireless deployments will not efficiently interoperate with the wired IP Internet
- Likely 3G deployment scenario of wireless gateway systems acting as transport-level bridges, allowing the wireless domain to use a modified TCP stack that should operate efficiently in a wireless environment
- 802.11 is different
- Bluetooth is yet to happen (or not)
IP Extensions & Refinements

- **IP Multicast technologies**
  - Extension of IP into support of common broadcast / conferencing models
  - Large-scale multicast
  - Small-scale multicast – conferencing
  - No widescale deployment as yet

- **IP Mobility**
  - IP support of mobility functions for mobile hosts and mobile subnets
  - Difference between nomadic operation and roaming operation

- **IP QoS**
  - IP support of distinguished service responses from the network
  - Per-flow responses or per-traffic class response models exist
  - No real uptake of either approach so far