GATEWAY SPECIAL INTEREST GROUP MEETING NOTES

STATUS OF THIS MEMO

   This memo is a report on a meeting. No conclusions, decisions, or
   policy statements are documented in this note.

INTRODUCTION

   This memo is a report on the Gateway Special Interest Group Meeting
   that was held at ISI in Marina del Rey, California on 28 and 29
   February 1984. Robert Hinden of BBNCC chaired, and Jon Postel of ISI
   hosted the conference. Approximately 35 gateway designers and
   implementors attended. These notes are based on the recollections of
   Jon Postel and Mike Muuss. Under each topic area are Jon Postel's
   brief notes, and additional details from Mike Muuss.

   The rest of this memo has three sections: the agenda, notes on the
   talks, and the attendees list.

MEETING AGENDA

   Tuesday, February 28

      9:00 Opening Remarks -- BBN - Hinden
      9:15 Opening Remarks -- ISI - Postel
      9:30 The MIT C Gateway -- MIT - Martin
      10:00 The Butterfly Gateway -- BBN - Hinden
      10:30 Break
      11:00 The EGP C Gateway -- ISI - Kirton
      11:20 The BRL Gateway -- BRL - Natalie
      11:40 The CMU Gateway -- CMU - Accetta
      12:00 Lunch
      1:30 The Wisconsin BITNET/CSNET Gateway -- UWisc - Solomon
      2:00 LAN to X.25 Gateway -- Computer Gateways Inc. - Buhr
      2:20 ISI-UCI Gateway -- UCI - Rose
      2:40 FACC Gateway -- FACC - Holkenbrink
      3:00 Break
      3:30 Lincoln IP/ST Gateway -- LL - Forgie/Kantrowitz
      3:50 Minimal Stub Gateways -- MITRE - Nabielsky
      4:10 Discussion

    Wednesday, February 29

        9:00 Opening Remarks -- BBN - Hinden
        9:10 SPF routing -- BBN - Seamonson
        9:35 Multiple Constraint Routing -- SRI - Shacham
        10:00 FACC Multinet Gateway Routing -- FACC - Cook
        10:30 Break
        11:00 Metanet Gateway -- SRI - Denny
        11:20 Address Mapping and Translation -- UCL - Crowcroft
        11:40 Design of the FACC Multinet Gateway -- FACC - Cook
        12:00 Lunch
        1:30 SAC Gateway -- SRI - Su/Lewis
        2:00 EGP -- Linkabit - Mills
        2:30 Congestion Control -- FACC - Nagle
        3:00 Break
        3:30 A Gateway Congestion Control Policy--NW Systems - Niznik
        4:00 Discussion

NOTES ON THE MEETING

    The MIT C Gateway -- MIT - Martin

        Postel: A description of the gateway implemented at MIT. The
        gateway was first developed by Noel Chiappa. It is written in C.
        The MIT environment has 32 internal networks which are treated as
        subnets of the MITNET on the Internet. The MIT gateways then do
        subnet routing in their interior protocol. The subnet routing
        scheme is similar to GGP. Liza has added an EGP implementation to
        this gateway.

        Muuss:

        Campus network/project Athena
        Dynamic routing
        Congestion control - grad student
                    +---------------+---+
          Class A net : | 18|subnet|res|host|
                    +---------------+---+

        "Bridges" forward between subnets.

        Campus Network and Project Athena 65 VAX 750s, 200 IBM PCs.

        Hosts: Now = 400, 1986 = 3,000, 1990 = 10,000

        Subnets: Now = 42, 1985 = 60, 1990 = 200, (4 subnets/building)

        Protocols: Internet, DECnet, Chaosnet

     FiberOptic spine between campus buildings.

     MIT gateways:

        11/03s and 11/23s
        68000 on Abus
        6800 on Multibus (Bridge communications)

        MIT C gateway -
        Runs under MOS, bridge OS, homegrown OS. Multiple protocols,
        multiple interfaces.

        11/03 - 100 packets/sec.
        11/23 - 180 packets/sec.

        GGP - Gw/Gw
        EGP - Exterior Gw
        IGP - Interior Gw

        EGP: Autonomous systems

        EGP:
          Neighbor acquisition
          Hello/I heard you
          Net reachability poll
          Net reachability message

     MIT IGP:

        IP header on EGP protocol
        Dest: net number, subnet number, 0, 0377 (broadcast address)

     IGP header:

        Autonomous system number
        Sequence number
        Tasks:
            Propagate exterior and subnet routing.

     Packets
        Ext route request, and update Routing server
            Default gateway
            Exceptional gateways
            Nets reached

     MIT - Gw broadcasts initial routings when it comes up, and again
     on each change, net is flooded on each change several times. Each
     bridge can ask for help.

    Future: Wideband net gateway from BBN will also sit on net 18,
    and an MIT routing server to acquire routing information. Trick -
    BBN-Gw will be on an Ethernet, and a modified ARP will be used by
    the bridges to "fool" the BBN gateway into acquiring the routes.

    Subnet Routing - inspired by PUP and CHAOS
       Neighbor Bridge
         Net I/F
         Bridge address
         Latest seq number
         Aging value
       Route to subnet
         Distance

       Packets
         Request
         I'm up

          Route update
            Distance vector (256 bytes)
                   0 - Direct
                   1 -127 - hop count
                   128-255 - "Interface used for next hop" to subnet
                             and hop count
                   255 - Unreachable

    Problem -
       Many neighbors --> too much time and traffic needed for
    processing.

       3 level addressing and routing strategy
       Ext Gw:
         Routing server
         Default Gw
       Subnet routing
         Small but rich subnet routing updates.

  The Butterfly Gateway -- BBN - Hinden

    Postel: A description of the butterfly hardware and a discussion
    of the plans for the new gateway software to be implemented on it.
    The butterfly machine is a multiprocessor (MC68000's)
    interconnected with a funny switch. The new software will
    incorporate the so called "Shortest Path First" or SPF routing
    algorithm.

        Muuss:

        Replacement for existing 30 PDP-11 "core" gateways.
        Problems to be solved.

             o Replace GGP
                  - Routing updates filling up
                  - Neighbor probes (N**2)
                  - Few buffers

             o Present GGP updates only hold 70 net numbers, repacking
                data will increase that to approximately 100 nets, but
                this is just short term.

        Features of Butterfly -
             o 1000's of nets
             o Partitioned nets
             o Type of service routing, access control
             o Flow control
             o Large and small gateway configurations

        New functions -
             o Routing
             o Neighbor discovery
             o Reduce neighbor pinging
             o Access/departure model
             o Connect gateways with point-to-point lines

        Routing -
             o SPF - shortest path first
             o Gateway based routing (opposed to network routing)
             o Routing updates
                   Gw ID
                  <nets directly connected>
                  <neighbor, distance>
             o Updates flooded to other gateways

        Next-door - Neighbors
             o Neighbor gateways closest to gateway
             o Ping next-door-neighbors only
             o For up/down acquisition, partition into rings. Reduces
                pinging.

        Access/departure model

          First Gw (entrance) picks exit gateway

        First Gw adds Gw - Gw header

     Butterfly gateway

        Processor nodes and switch nodes

        4-legged switch nodes, decision is simply UP or DOWN. 2
        inputs
        and 2 outputs.

        Processor: MC 68000
        Memory management Unit
        Processor node controller - 2901 bit slice
        PVC is the memory controller.

        Butterfly -
        32 M bps/path
        Bandwith: approximately N - speed
        Size: approximately N/2   log N 2

        Butterfly will support multibus interface; 1822, HDLC,
        Ethernet, Ring

     Terminal and load device will be a personal computer

     Small Gw for ARPA is approximately $20K

     New Gw processor structure

     Buffer Management
        o   Scatter/gather buffers minimum size and extensions
        o   Buffer pool on processors with I/O
        o   Primary and secondary collections per device
            ==> guaranteed minimum service per device
                 (implemented w/counts)

  The EGP C Gateway -- ISI - Kirton

     Postel: A user process was installed in Berkeley 4.2 Unix to do
     EGP protocol functions leaving the normal router kernel function
     in charge of forwarding datagrams. The EGP user process may do
     system calls to update the kernel routing data. Based on the work
     of Liza Martin.

        Muuss:

        EGP under 4.2

        Elimination of nonrouting gateways

        Design -
            Forwarding done in kernel
            Kernel does not send redirects
            EGP user process for route updates
            Written in C
            EGP based on Liza Martin's code

        Routing Tables
          o   Kernel
          o   EGP Process

        EGP Process Table -
          o   External updates
          o   Internal information

        Facilities -

          Configuration file-
              o   Trusted neighbors
              o   Internal non - routing gateways

          Acquisition -
            o   Predetermined number of core gateways are EGP'd to
            o   Only accept from trusted neighbors
            o   Cannot acquire neighbors indirectly, for now

          Unix Interfaces -
            Reuse IP socket (problem with protocol number)
            Listening to ICMP for redirects
            System calls for -
              o   Route updates
              o   I/F config reading
              o   I/F status check

          Performance -
              o   60 ms/packet pair (CPU time)
              o   Typically 1% of CPU for 1 minute polling

          Protocol function going
          Routing updates being implemented

          Should be all going in April.

    The BRL Gateway -- BRL - Natalie

        Postel: This was a description of the BRL dumb gateway. More
        interesting was the description of the BRL complex and the
        inteconnections between machines. The gateway is written in C
        (and derived from the MIT C-Gateway) and based on a simple
        multiprocess operating system called LOS.

        Muuss:

        BRL history

        LOS design
          Message passing
          Memory Management
          No copying of data, buffer size

    The CMU Gateway -- CMU - Accetta

        Postel: This was a description of the CMU dumb gateway.

        Muuss:

        History -
          o   "Logical-Host" multiplexor (March 81)
          o   Gateway (Oct 82) remote debugger and monitor
          o   Router (Oct 83)
                 - Modular device and protocol support
                 - Stub IP dynamic routing
                 - Local inter-network cable routing.
          o   Written in "C"

        Uses low memory for buffers (maximum 32K)!
           (autoboot of 3M bps Ethernet)
        Auto-configuration of devices
        Individual stack contents
        Round-robin scheduler
        Dynamic memory allocation

        Device driver
          Network interfaces
          Auxiliary support devices

        Does IP, ICMP, UDP

           Splicing through of PUP and CHAOS on chaos net, uses ARP.

           Configuration testing protocol (as in Ethernet Spec).

        IP Processing-

            o   Consistency checks
            o   Redirects does not forward misrouted packets
            o   Fragmentation - ICMP dest unreach If DF Set
            o   Access list for who can pass through

        No GGP, no EGP, Uses known gateways

        Ordinary devices and PDP-10 and PDP-20

    The Wisconsin BITNET/CSNET Gateway -- UWisc - Solomon

        Postel: This was a discussion of a mail relay between the
        Internet and BITNET to be installed at Wisconsin.

        Muuss:

        WISC-IBM (192.5.2.24) will connect to BITNET

        Mail gateway, BITNET uses RFC 822 headers!

    LAN to X.25 Gateway -- Computer Gateways Inc. - Buhr

        Postel: This was a description of a protocol translation device
        between an X.25 world and the DATAPOINT ARCNET world.

        Muuss:

        ARCNET to X.25 Bridge

        ARCNET - from Datapoint,
           Baseband coax, 2.5 mbps
           Token passing
           Reserve/send/wait/ack protocol
           RIM chip implements this

        "The OSI models seem less clear than the Internet models, perhaps
        because they are less well developed."

        Wraps the subnetwork in an enhanced subnetwork layer.

        Every pair of subnetworks must be connected in this design - hence
        a bridge not a gateway.

        Bridge is a network layer RELAY.

        ARCNET address is sent as X.25 data

    ISI-UCI Gateway -- UCI - Rose

        Postel: This was a description of the UCI dumb gateway. This one
        is made up of two hosts (VAX 750s) 50 miles apart. The VAXs are
        connected via a 9.6 Kbs leased line. One is interfaced to the
        ISI-NET (an Ethernet) and the other to UCIICS net (also an
        Ethernet). The VAXs run Berkeley Unix 4.1. These VAXs run as
        regular hosts too.

        Muuss:

        MTU is 512. Effective bandwidth of approximately 6000 baud over
        9600 baud line.

    FACC Gateway -- FACC - Holkenbrink

        Postel: A description of a gateway designed by Ford. The gateway
        is based on a MC68000 multiprocessor and a VME bus. An
        interesting question that came up during this presentation was
        "What is the least information a host (or gateway) must have when
        it comes up, and how can it acquire the rest of what it needs to
        go into full operation from the environment?"

        Muuss:

        Inter-segment Processor. M68000 CPU with various co-processors.
        68000 IOPS, 1822, IOP Ethernet IOP. 1 cpu does IP, routing.
        Multi-cpu version of MOS

    Lincoln IP/ST Gateway -- LL - Forgie/Kantrowitz

        Postel: This was a discussion of the design of the Lincoln
        gateways used primarily in the WBCNET for speech transmission
        research. This gateway uses special I/O interfaces to promote a
        high packet processing rate. The gateway implements both the
        regular IP, and the ST protocol which permits resource
        reservations to minimize the variation in transmission delay.
        These gateways can, of course, act as regular internet gateways,
        and have achieved very good performance in terms of datagrams per
        second.

        Muuss:

        Packet voice experiments, wideband SATNET. Concentrate traffic
        from local nets to trunk net. Needed enough performance to load
        WBSATNET. 11/44 and ACC IF11 (Z-80). T1 trunk protocol converter.
        (voice T1 <--> datagram)

        IP problems -
          o   Congestion
          o   High packet header overhead
          o   No support for conference call

        ST -
          o   Virtual circuit
          o   Know capacity in advance, schedule channel
          o   Abbreviated header

     11/44 - 900 to 1000 pkts/sec.

     Port processor:
        Sync low speed: 600K bits/sec.
        Packet processing: 500 pkts/sec. average
          20-talker LPC voice loop, 28 data
             bytes/pkt, 50% duty cycle
        Data handling
          4 pcm voice stream loop 64K bps
          184 data bytes/pkt, 100% duty cycle

     Dispatcher Requirements
        o Timely do ST
        o Utilize rest of circuit for IP
        o Performance measurement

     Reservations on the SATNET: Each host makes a reservation for
     Nbytes of M messages every INTERVAL. Reservations are absolute.

     ST and IP for each distant run = MPP multipurpose packets.

     12,000 lines of C code in 11/44 portion.

  Minimal Stub Gateways -- MITRE - Nabielsky

     Postel: This was a more abstract discussion of how stub gateways
     could interact and acquire information about the topology of the
     Internet.

     Muuss:

     Ethernet stub to Internet
     Inexpensive, single-band ISBC 186/51 Intel @ $3000
     High performance. EGP?

     128K bytes/board

     The Internet forest

         Alternative to ARP using Multicast

      SPF routing -- BBN - Seamonson

      Postel: This was a fine presentation of the principles of the
      "Shortest Path First" (SPF) routing procedures with some remarks
      on how it is tailored to the Internet gateway situation. One
      point that was impressed on me was that when using SPF in a set of
      gateways (say, the core autonomous system) the procedure will do
      routing to an "exit" gateway. Somehow I had not thought about it
      in those terms before, but (obviously) just as there is a source
      and a destination IMP in the ARPANET there will be an entrance and
      an exit gateway in an SPF autonomous system.

      Muuss:

      Features -
        Metric, update procedures, path calculation, forwarding

      Current GGP problems -
        o   Counting to infinity
        o   Not enough topology information in each Gw
        o   Updates potentially very large

      SPF in ARPANET
        o   Single path (not optimal) - no split of flow
        o   Delay based, to minimize delay
        o   Global knowledge of connection topology and delays

      Metric used -
        o   Delay, delay of each packet averaged
               (queueing plus transmission plus propagation)
               arrival-to-arrival time.
        o   Average delay on each trunk computed every 9.6 seconds.
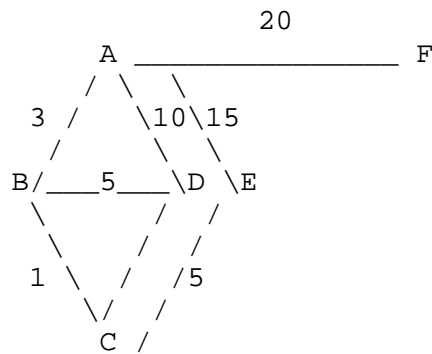   Report large changes in delay, fast

      Update procedure -
        o   Updates report delay to each neighbor
        o   Update triggered by topology change, significant delay
            change, or 1 time/minute.
            Decay of threshold to direct to send update
        o   Sequence numbers
        o   Flooding on all trunks sent out on all lines
        o   Receipt of echo is acknowledgement
        o   Retransmission
        o   Aging of information
        o   Updates are 2*n*l packet growth. n = number imps,
            l = number lines

        - When lines goes up, rather than dumping routing
          table,just waits one minute until all updates have
          been heard.

     Path calculation
        o   Dijkstras Algorithm

```
                      20
               A _____ F
              / \   \
          3 /    \10\15
            /      \   \
         B/___5___\D \E
          \      /   /
           \    /   /
         1 \  /   /5
            \/   /
            C /
```

  1.   A   B(A, 3), D(A, 10), E(A, 15). F(A, 20)

  2.   A   C(B, 4), D(B, 8), E(A, 15), F(A, 20)
            |
            B

  4.   A   E(C, 9), F(A,20)
            |
            B
          / \
        C   D

  5.   A
            |
            B
            |
            C
           /
          E

     Then tree is inverted into a "go here to get to this destination."

     For Internet -

        Similar algorithm, needs special packet header to
        indicate "exit" gateway to get to destination network.

      Update procedure -
         Neighbor interface, neighbors, and delay to neighbor.

          "Next door neighbors" for minimizing traffic.
          Ability to package multiple updates in one average
          explicit Acks.

       Path calculation -
          o   Possible to build different trees based on type of
               service.

       Forwarding -
          o   Exit Gw
          o   Consistent databases are important.

   Multiple Constraint Routing -- SRI - Shacham

      Postel: This was a clear presentation of some of the consequences
      of the idea of type of service routing. The level of complexity
      of the routing procedure is determined to depend on how many
      catagories of service there are and how many selections there are
      in each catagory. A few examples were discussed including the
      current type of service parameters of IP.

      Muuss:

      Both current and proposed ARPANET algorithms provide "best" path
      under single constraint (number of hops, delay).
      Internet will have diverse characteristics, it would be nice to
      consider more than one constraint.

         o   Determine a set of measures.
         o   Represent each measure as a single number.
         o   Determine range of values. (complexity $0(c^{**}n)$ range of n)
         o   Define path measure as a function of measure of length.
                sum (delay, cost)
                min/capacity, length, security)

      If just one cost is used, then SPF (or whatever) can be used for
      each cost. However, under multiple constraints there is a more
      difficult problem. e.g.: minimum delay with packet size of at
      least 1000 bytes.

      RUMC has been shown to be in the NP complete family.

      RUMC needs bigger tables, more processing and routing overhead.
      Its not awful for 2-choice TOS, like in IP.

      Table size is random, we have to be prepared for the worst case.

      Possible strategies: flood a "search packet," dropped when

constraints are not met, see if it makes it though. Good only for
virtual circuit. Weighted sum (VC only) works only with some
probability.

TOS is needed for Internet, but the algorithms are costly.
Complexity for providing TOS IP style is not too high.

FACC Multinet Gateway Routing -- FACC - Cook

Postel: This approach considered hop count to be an inadequate
metric for routing decsions in a system of different types of
networks (e.g., Ethernets, ARPANETs, 2.4Kb lines). Delay was
selected as the metric to use. There are some interesting issues
in the measurement of delay for some types of networks. Also, the
design considers the use of multiple paths when they are avaiable,
and routing to provide connectivty between the parts of
partitioned networks.

Muuss:

Routing with a single constraint.
A network of gateways Access, Transport, or Dual networks.
Some networks are used as backbones between gateways only.

Routing updates
  Variable length
  Broadcast routing updates

Unitary ends - A - Gw - B - Rest
   Routing for A is really just routing to B
   Neighbor Gws, nets
   Lots and lots of tables

Metanet Gateway -- SRI - Denny

Postel: This is a project to invent several new addressing
features for gateways. In particular, there is a scheme to use an
option much like the source route option to do multi-addressing of
IP datagrams. It seems as if the gateways that implement this
option will have to know which other gateways do and don't
implement it. Also, there was discussion of a gateway to a
network that is in radio silence, and how to keep TCP connections
going with hosts that can't talk. This project is also concerned
about network reconstitution, security, survivability, congestion
control, and supporting multimedia data (voice, bitmaps, etc.) in
applications. A gateway is being developed in ADA for a MC68000
machine (SUN), and the initial version of the gateway is to be up
in May 84.

Muuss:

Navy internet
  Multimedia mail and conf.
    Radio silence (EMCON)
    Security and Survivability.

EMCON - Causes special problems for EGP and IGP one way nonTCP
mail delivery. No Acks. Uses name screen to redirect mail to
special one-way mail catcher, who then forwards using ordinary
methods.

Security and survivability
Access control - "capability" - 32/64 bit key which changes
frequently (every hour or so)

Reconstitution - Partitioning, coalescing, mobile host
Test and monitoring - HMP

Gateway target - 68000 in ADA. Telesoft compiler

Address Mapping and Translation -- UCL - Crowcroft

Postel: This was a discussion of some of the issues in
interconnecting networks of different types including the Internet
and networks in England such as the Universe network. The
Universe network is made up of Cambridge Rings at several sites
linked via a satellite channel.

Muuss:

ARPA - SATNET - NULLNET - UCLNET UNIVERSE Satellite, 3 UCL rings

SAM -
  o   IP switch to several 1822 hosts
  o   IP/universe mapper, overlays UCLNET on universe
  o   Mask and match
         128. 11. code. host

Three types:

   1. Direct: code --> subnet
      2. Redirect: 2nd lookup (for multihoming)
      3. Logical: Logical address into a table of universe
   names.
                    Name lookups give addresses and routes.

IP tunnels through X.25

      BBN Van gateway PSS - IPSS -Telenet - for hosts that can't use
      SATNET.

      SAM does access control and multihoming. Clever Multihoming gives
      host a second address and sends an ICMP/Redirect to force TCP
      connection to go through a different route, but wind up at same
      place!!!

      Wrote EGP in ADA. It didn't help at all.

   Design of the FACC Multinet Gateway -- FACC - Cook

      Postel: This is a distributed multiprocessor machine using a
      special bus network for the interprocessor communication. The
      softaware is written in C. The gateways is in an early test
      phase.

      Muuss:

      RADC program

      Started with AUTODIN II, switched to DDN.
      Small to large switching devices.
      DoD uses of PDNs, and partitioned network problems.

      Distributed processing architecture -
        Parallel contention, 90M bps bus, 22 wires. Each node has cpu,
        memory, optimal comm line. Wire - OR presentation of address,
        contention happens each time bus becomes free, all requestors
        put out type of msg, pri, and address.   Reads back wire - OR of
        result, and highest gwy wins, sorted by (pri, type, higher
      addr).
        Bus was originally designed for our FAA fail-soft application
        Z-800l w/MMU. Not binary addressing, but unitary (base1)
      One element resolved per bus transaction.
      Boards may be plugged in while running.
      Inherent parallelism in layered protocols.

      Interface connector clues board to modem levels and date rate. Up
      to 100K bps now, soon up to T1 rate.

      Multiprocessor approach allows routing calculation to take place
      out-of-band from the measurement of delay and traffic, and allows
      use of more compute power for routing.

      Mostly written in C, with some assembler. Multiprocessor
      operating system, designed from scratch.

     SAC Gateway -- SRI - Su/Lewis

        Postel: This was a presentation of the design for the gateways to
        be used in the advanced SAC demo experiments on network
        partitioning and reconstitution, and communication between
        intermingiling mobile networks. Much of these demonstrations will
        be done with packet radio units and networks. Some of the ideas
        are to use a gateway-centered type of addressing and double
        encapsulation (i.e., an extra IP header) to route datagrams.

        Muuss:

        Network dynamics due to component mobility or failure.
        Mobile host, reconstitution, partitioning.
           H/W: 11/23
           S/W: Some "C" gateway
           OS: VMOS (SRI)

        Gateway-centered addressing, rather than network.
           Gw host instead of net.host.
        Double encapsulation: additional IP header.
           TCP uses addr as an ID, IP uses it as an ADDRESS (-> route)
           Need to separate these dual uses of this address field.
        Incremental Routing (next-hop indication)

     EGP -- Linkabit - Mills

        Postel: A presentation of the EGP design. EGP has three major
        aspects, neighbor acquisition, neighbor reachability, and network
        reachability. The autonomous system concept was discussed.

        Muuss:

        Background, Implementation, Experience, Disparaging Remarks

        Design goals -
           o   Established demarcations
           o   Decouple implementations
           o   Confine routing loops
           o   Exchange reachability information
           o   Provide flow control for connectivity information
           o   Medium-term lifetime

        Non goals   Not trying to do these!
           o   Flexibility of topology
           o   Rapid response   Very slow update

        o   Adaptive routing
        o   Common routing metric   No agreement at all
        o   Load sharing or splitting

    "Good news travels fast and bad news travels forever."
    Not for routing, but only provides reachability

    RFC827 initial mode, RFC888 stub protocol

    Neighbor acquisition protocol
        o   2-way shake
        o   Flow - rates
        o   Explicit acquisition/cause

    Neighbor reachability protocol
        o   Periodic polling
        o   Parasitic information
        o   Reachability algorithm Network reachability
             protocol
        o   Periodic pulling
        o   Remote information
        o   Direct and indirect neighbors
        o   Indirect internal and indirect external
             neighbors
        o   Distance information

    EGP neighbors do not need to peer with more than one
    CORE gateway, but you may peer with anybody you wish.

    Shortcomings -
        o   Slow reaction due polling
        o   Tree-structured routing constraint
          - Rigid topology
          - Administrative resistance to odering
          - Lack of adaptive connectivity
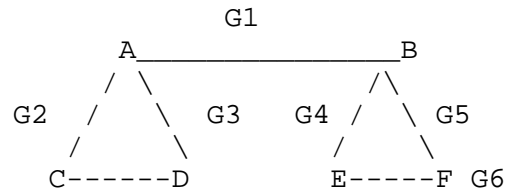   o   Neighbor acquisition incomplete.

    Loops between autonomous systems will last a long
    time, and are a real no-no.

    System models -
        o   "Appropriate first hop" criterion
          - Not useful for implementation
          - Requires global information
          - Inadequate for verification
        o   Graph models
          - N-graph shows net connectivity
          - T-graph shows system connectivity

                 - T-acycloc criterion insures loop-free
          o   Derived features
                 - Induces spanning tree

       N-graph

```
                                 G1
                       A_____B
                      / \              /\
                 G2  /   \ G3    G4 /  \ G5
                    /     \         /    \
                  C------D        E-----F G6
```

       AS1 = G2, G3, G6   A   B
       AS2 = G1
       AS3 = G4, G5   AS1 ----- AS2 ----- AS3

                                     T-graph

       Test: to ensure that there are no cycles

       Spanning subtree

       Specification effort - Status report State machine designed

       Remaining issues -
          o   Remove extra hop in core system
          o   Expand tables
          o   Test backdoor "GGP"
          o   Resolve specification issues
          o   Resolve full gateway configuration
                 - Back door connectivity guidance
                 - can only advertise 1 path at a time.
                 - APF rule guidancee
                 - Self organization issues
          o   Implement and distribute for operational systems.

    Congestion Control -- FACC - Nagle

       Postel: This was a discussion of the situation leading to the
       ideas presented in RFC 896, and how the policies described there
       improved overall performance.

Muuss:

First principle of congestion control:

   DON'T DROP PACKETS (unless absolutely necessary)

Second principle:

   Hosts must behave themselves (or else)

   Enemies list -

      1. TOPS-20 TCP from DEC
      2. VAX/UNIX 4.2 from Berkeley

Third principle:

   Memory won't help (beyond a certain point).

   The small packet problem: Big packets are good, small are bad
   (big = 576).

Suggested fix: Rule: When the user writes to TCP, initiate a send
only if there are NO outstanding packets on the connection. [good
for TELNET, at least] (or if you fill a segment). No change when
Acks come back. Assumption is that there is a pipe-like buffer
between the user and the TCP.

The source quench problem Rule: When a TCP gets an ICMP Source
Quench, it must reduce the number of outstanding datagrams on
relevant TCP connections.

Rule: When a gateway nears overload, before starting to drop
packets, send a Source Quench.

Node capacity: Each node ought to have one buffer for each TCP
connection, plus some for overload.

Both fixes really need to be done together, although the first one
is often helpful by itself. Side effect: FTPs start off "slowly,"
until the first Ack comes back Dave Mills thinks this will
increase the mean delay for medium-size interactions. This
probably will not work so well for SATNET.

Problems about propagation time of links biasing the validity of
this result!!

   A Gateway Congestion Control Policy--NW Systems - Niznik

      Postel: This talk was (for Postel) hard to follow. There were a
      number of references to well known results in queuing theory etc,
      but I could not follow how they were being used.

      Muuss:

      Replacements for IMP SPF
      Topological observations
      Nodal congestion control policy
         GMD - control application [from German network]
         RPN - relational Petri net
         DCT - dynamic congestion table
      NCCP performance evaluation
      Planned GCCP: Gateway congestion control policy

      Lots of diagrams and figures.

      Better throughput than SPF, but somewhat higher delay.

      Cubic structure of table.

   DISCUSSION (Postel's personal comments)

      There was very little organized discussion during the meeting and
      not really very much question and answer interaction during the
      presentation. There was a lot of discussion during the breaks,
      and at lunch time, and at the end of each day.

      Some things that occured to me during the meeting that may have
      been triggered by something someone said (or maybe by the view out
      the window):

         Don't design a protocol where you expect to get a lot of
         messages from a lot of sources at the same time. For example,
         don't ask all the hosts on an Ethernet to send you an ack to a
         broadcast packet.

         Has anyone worked out in detail the routing traffic costs for
         the GGP vs the SPF procedures for the actual case of the
         Internet?

         How will the fact that thinking of the routing in the core
         autonomous system is cast in terms of an entry and an exit
         gateway effect other things? Will there be special

arrangements between the entry and exit gateway? Will an
autonomous system become a circuit switch connecting pairs of
entry/exit gateways?

Is TOS routing worth the cost?

Should we allow (as a new type of ICMP message) redirects to
Gateways?

Does making memory larger ever hurt? If a gateway's memory is
full of inappropriately retransmitted TCP segments would it be
better if there were less memory?

Is there something reasonable to do with source quench at the
TCP? Re: RFC-896.

If there are links (or networks) of vastly differing delay and
thruput characteristics what impact would an IP level load
splitting (say by gateways) have on TCP connections (some of
the segments of the connection go one path and others go a
different path)?

Are any problems avoided (either way) by using double IP
headers vs a "source route like" IP option to separate the IP
level addressing and routing function from the TCP level
end-point naming function of the IP addresses.

What bad things could happen from the proposed IP
multidestination routing option?

MEETING ATTENDEES

    Mike Accetta - CMU
    R. Buhr - Canada
    J. Noel Chiappa - MIT
    Paul Cook - Ford
    Jon Crowcroft - UCL
    Barbara Denny - SRI
    Jim Forgie - LL
    Steve Groff - BBN
    Phill Gross - Linkabit
    Kjell Hermansen - NTA
    Robert Hinden - BBN
    Patrick Holkenbrink - FACC
    Ruth Hough - AIRINC
    Willie Kantrowitz - LL
    Paul Kirton -ISI
    Mark Lewis -SRI
    Liza Martin - MIT
    Doug Miller - MITRE
    Dave Mills - Linkabit
    Mike Muuss - BRL
    Jose Nabielsky - MITRE
    Ron Natalie - BRL
    John Nagle - Ford
    Carol Niznick NW Systems
    Jon Postel - ISI
    Joyce Reynolds -ISI
    Marshall Rose - UCI
    Joe Sciortino - AIRINC
    Linda Seamonson - BBN
    Nachum Shacham - SRI
    Alan Sheltzer - UCLA
    Marvin Solomon - WISC
    Zaw-Sing Su - SRI
    Mitch Tasman - BBN