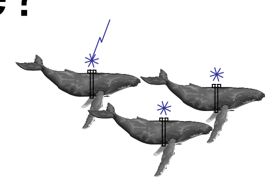
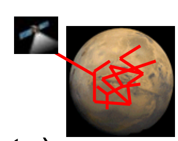
Delay-Tolerant Networking for Extreme Environments

Kevin Fall
Intel Research
Berkeley, CA
kfall@intel-research.net
http://www.intel-research.net

What is Extreme?



- Deep Space Communications
  - Beyond near-earth
  - Landers, Orbiters, Deep Space Probes
- Sensor Networks
  - Terrestrial: Ocean or Land Based
  - Extra-terrestrial objects (on planets, etc)



- High-Stress Physical Environments
  - Battlefield, Civil Emergency, Submarines

RFC1149

- "…encapsulation of IP datagrams in avian carriers" (i.e. birds, esp carrier pigeons)
- Delivery of datagram:
  – Printed on scroll of paper in hexadecimal
  – Paper affixed to AC by duct tape
  – On receipt, process is reversed, paper is scanned in via OCR

Implementation of RFC1149

- See http://www.blug.linux.no/rfc1149/

Ping Results

```
Script started on Sat Apr 28 11:24:09 2001
vegard@gyversalen:~$ /sbin/ifconfig tun0
tun0      Link encap:Point-to-Point Protocol
          inet addr:10.0.3.2  P-t-P:10.0.3.1  Mask:255.255.255.255
          UP POINTOPOINT RUNNING NOARP MULTICAST  MTU:150  Metric:1
          RX packets:1 errors:0 dropped:0 overruns:0 frame:0
          TX packets:2 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0
          RX bytes:88 (88.0 b)  TX bytes:168 (168.0 b)

vegard@gyversalen:~$ ping -i 900 10.0.3.1
PING 10.0.3.1 (10.0.3.1): 56 data bytes
64 bytes from 10.0.3.1: icmp_seq=0 ttl=255 time=6165731.1 ms
64 bytes from 10.0.3.1: icmp_seq=4 ttl=255 time=3211900.8 ms
64 bytes from 10.0.3.1: icmp_seq=2 ttl=255 time=5124922.8 ms
64 bytes from 10.0.3.1: icmp_seq=1 ttl=255 time=6388671.9 ms

--- 10.0.3.1 ping statistics ---
9 packets transmitted, 4 packets received, 55% packet loss
round-trip min/avg/max = 3211900.8/5222806.6/6388671.9 ms
vegard@gyversalen:~$ exit

Script done on Sat Apr 28 14:14:28 2001
```

About 1.5 Hrs

Comms Challenges

- Large Delays
- Intermittent and Scheduled Links
- Limited Power Nodes
- Bandwidth Asymmetry
- Limited Emission Requirements (LPI/LPD)
- Heterogeneous Network Architectures
- Link Security Needs

Extreme Delay Links

- The Problem:
  - Delivery may be prop or tx-time dominant
  - Both can be extreme:
    - Very long propagation → problems w/RTX
    - Very slow links → longer tx time, more storage
- Long propagation delay especially difficult
  - Can't buy less latency
  - Adversely affects conventional reliable transports

TCP Dependence on RTT

- Slow-start ramp: *time to window W*:

$$T = (RTT \bullet b)\log_2 \frac{W}{W_0} \text{ for initial window } W_0$$

- Steady-State Throughput limited by

$$B(p) \approx \min\left(\frac{W_{max}}{RTT}, \frac{1}{RTT\sqrt{\frac{2bp}{3}} + T_0 \min(1, 3\sqrt{\frac{3bp}{8}})\, p(1+32p^2)}\right)$$

p = loss probability, b = pkts ACKd per ACK, T_0 = initial RTO

Asymmetric Links

- Significant differences in each direction
  - Latency (MAC behavior, path)
  - Bandwidth (cost/engineering, technology)
  - Loss characteristics (power, path, buffering)
- Problems in one direction affect the other
  - ACK congestion → lost ACKs
  - Burst ACK arrivals → burst sending
- Some cases have *no* reverse channel
  - Possibly applicable to erasure coding…

Cascaded Intermittent Links

- Prob success (iid fail prob p_f) over *k* links:

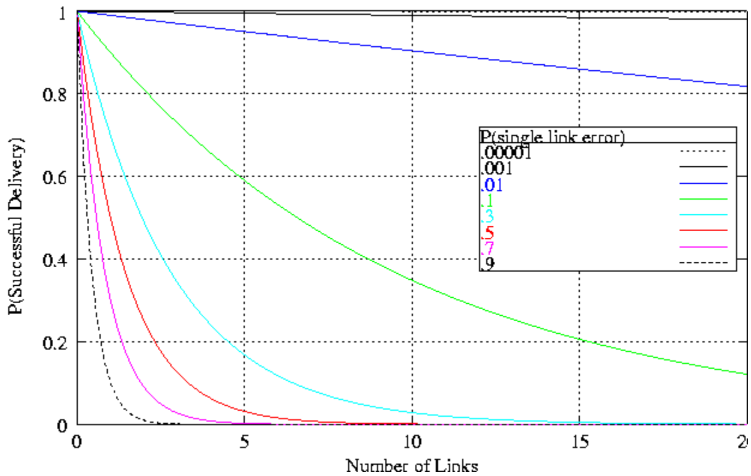
$$p_s = 1 - p_f \,;\, p_s(k) = (1 - p_f)^k$$

- For E2E delivery must have all links up

- But, expected # of failed links is

$$kp_f$$

Probability of Delivery



Routing Issues

- End-to-end path may not exist
  - Lack of many redundant links [there are exceptions]
  - Traditional routing assumes at least one path exists, fails otherwise
- Routing Algorithms More Complex
  - Scheduled links and contact opportunities
  - Need to match between pending messages, send opportunities and message priority
  - Available power may affect link selection decision
  - Typical routing algorithms optimize 1 metric. Those that don't [e.g. BGP] can become unstable.

Limited Power

- Extreme devices tend to use batteries

Battery Type	Cell (V)	Storage Density (Wh/kg)	Temp Range (celcius)	Rapid Charge (hrs)	Memory Effect
NiCD	1.2	40-60	-10 to +50	.5 to 1	Yes
NiMH	1.25	60-80	-10 to +50	2 to 3	No
Li-Ion	3.6	100	-20 to +60	3 to 6	No
Li-Poly	3	140	-30 to +55	8 to 15	No

- Power Requirements

	Intel 2011	RFMTR1000	LQUWM1k	LQUWM7k	Sojourner
Rx Cur	300mA	4.6mA	45mA	43mA	35mA
Tx Cur	500mA	12mA	60mA	1.09A	170mA
Sleep Cur	25mA	5uA	480uA	348uA	28mA
Tx Power	63mW	0.75 mW	?	?	100mW

Expected Lifetime

- An example with the RFM Radio:
  - 2 AA cells (2900 mAH each), 3v
  - 100% Duty Cycle Xmit: 242 hrs (10 days)
  - 100% Duty Cycle Rcv: 630 hrs (26 days)
  - 100% Duty Cycle Sleep: 580k hrs (66 yrs)
- Clearly:

Power management is fundamental

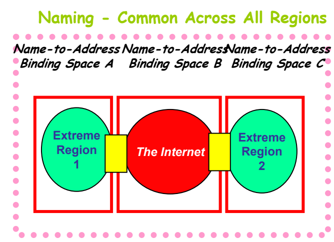
Heterogeneous Architectures

- Most extreme systems do not (won't) run IP
  - Oceans: just beginning to investigate routing
  - Space: very limited routing [e.g. rover to lander]
  - Sensors: novel or simple routing, low power
  - Too much overhead, no need for global routing
- Most have domain-specific naming scheme
  - Typically, a flat node ID or name
- But we don't want to scrap existing (Internet) software

Delay-Tolerant Architecture

- Goals
  - Interoperability across network architectures
  - Reliability robust to link and node failure
- Components
  - Flexible Naming Scheme
  - Reliable Message Overlay with Routing
  - Per-hop Authentication with CoS
  - Interoperability Gateways

Naming



- Naming *region*:
  - Local region naming
  - Inter-region prefix
- Tuples of the form (A, T)
  - A: administrative [valid inside a region]
  - T: topological [valid globally]
- Inter-region next hop determinable by local *f(T)* directly
- A is externally opaque and region-specific
  - May encompass esoteric routing [e.g. diffusion]
  - Could be object names, addresses, etc.

Reliable Message Overlay

- End-to-End Message Service: “Bundles”
  - “postal-like” delivery over regional transports
  - Optional class of service/notification
- Key Idea: *Custody Transfer*
  - *Custodian* owns reliable-delivery guarantee
  - Bundles transferred between custodians toward destination
  - Sender may free resources upon successful custody transfer

Bundles

- Bundles
  - Arbitrarily long messages delivered end-to-end between DTN capable nodes over distinct (but possibly identical) transport layers.
  - May have associated delivery characteristics. Thus, CoS is always at bundle granularity.
  - Bundles may be fragmentary and require reassembly to be complete.
  - Authenticated/verified during delivery.

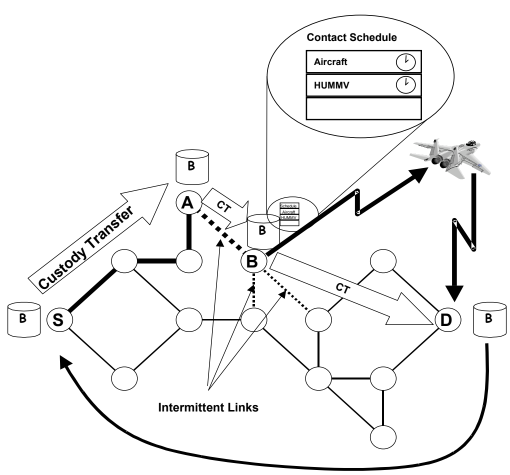
Routing, Forwarding and Custody Transfer

- "Classic" Concepts (Internet):
  - *Routing*: selecting best next hop for every possible destination
  - *Forwarding*: sending packet to best next hop
    - Typically, "on demand" [statistical multiplexing]
    - Forwarders know *a-priori* next hop for every destination
- DTN Concepts:
  - *Routing*: selecting best DTN next hop for destination
  - *Forwarding*: sending a bundle p2p when possible
  - *Custody Transfer*: reliable intra-DTN delivery (<u>with storage</u>)

DTN Node Types

- Non Persistent Node [NP node]
  - no stable storage
  - Build/consume bundles, forwards bundles, participates in time synchronization
  - May forward or cache bundle or bundle parts
  - Never assumes custody
- Persistent Node [P node]
  - stable storage
  - Does everything an NP node does
  - Always accepts custody of a bundle on success
  - Notifies prior custodian of custody transfer
- Exception: SRC/DST always accept custody

Bundle Routing Example



End-to-end Acknowledgement
(Sent using bundles, path omitted for clarity)

Comparison

- Trials until success [end-to-end]:

$$E(X) = \frac{1-(1-p_f)^k}{(1-p_f)^k}$$

- Trials until success [link-by-link]:

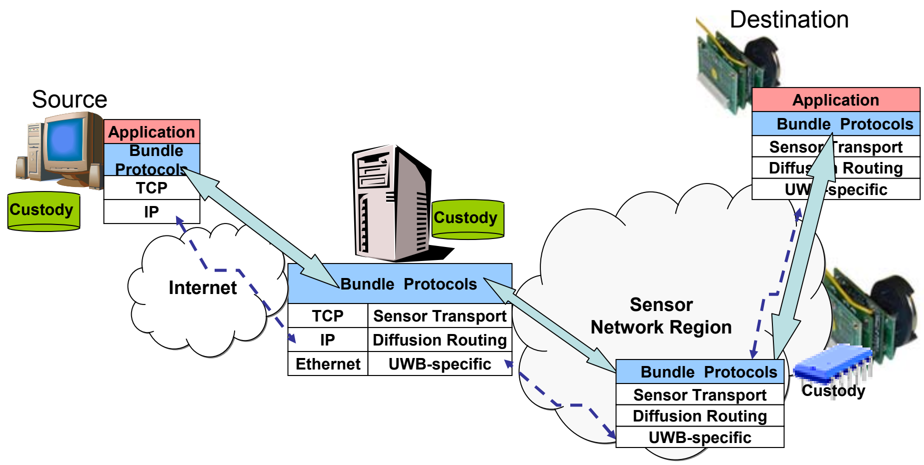
$$E(X) = \sum_k \frac{q}{p} = k\,\frac{p_f}{(1-p_f)}$$

(assuming p_f < 1)

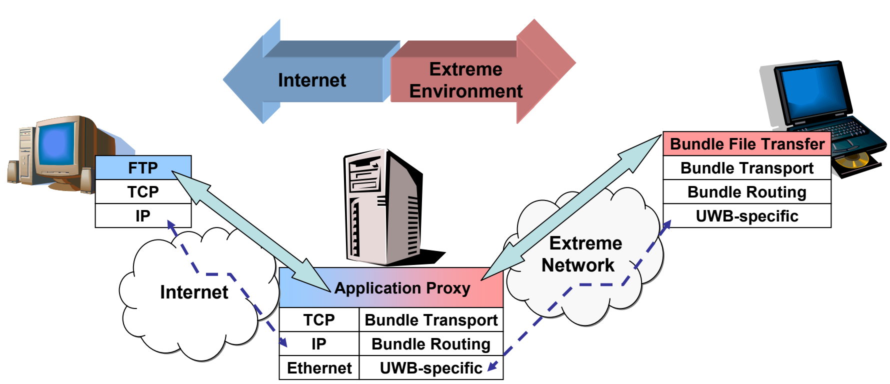
Expected Retries

Pf:	0.1	0.3	0.5	0.7	0.9		0.1	0.3	0.5	0.7	0.9
links	LINK-LINK						E2E				
	0.1	0.3	0.5	0.7	0.9		0.1	0.3	0.5	0.7	0.9
1	0.11	0.43	1.00	2.33	9.00		0.11	0.43	1.00	2.33	9.00
2	0.22	0.86	2.00	4.67	18.00		0.23	1.04	3.00	10.11	99.00
3	0.33	1.29	3.00	7.00	27.00		0.37	1.92	7.00	36.04	999.00
4	0.44	1.71	4.00	9.33	36.00		0.52	3.16	15.00	122.46	9999.00
5	0.56	2.14	5.00	11.67	45.00		0.69	4.95	31.00	410.52	99999.00
6	0.67	2.57	6.00	14.00	54.00		0.88	7.50	63.00	1370.74	999999.00
7	0.78	3.00	7.00	16.33	63.00		1.09	11.14	127.00	4571.47	9999999.00
8	0.89	3.43	8.00	18.67	72.00		1.32	16.35	255.00	15240.58	99999999.00
9	1.00	3.86	9.00	21.00	81.00		1.58	23.78	511.00	50804.26	999999999.00
10	1.11	4.29	10.00	23.33	90.00		1.87	34.40	1023.00	169349.88	9999999999.00
11	1.22	4.71	11.00	25.67	99.00		2.19	49.57	2047.00	564501.93	99999999999.00
12	1.33	5.14	12.00	28.00	108.00		2.54	71.25	4095.00	1881675.42	999999999999.00
13	1.44	5.57	13.00	30.33	117.00		2.93	102.21	8191.00	6272253.74	9999999999999.03
14	1.56	6.00	14.00	32.67	126.00		3.37	146.44	16383.00	20907514.81	99999999999999.40
15	1.67	6.43	15.00	35.00	135.00		3.86	209.63	32767.00	69691718.38	1000000000000000.00
16	1.78	6.86	16.00	37.33	144.00		4.40	299.91	65535.00	232305730.25	10000000000000000.00
17	1.89	7.29	17.00	39.67	153.00		5.00	428.87	131071.00	774352436.51	100000000000000000.00
18	2.00	7.71	18.00	42.00	162.00		5.66	613.09	262143.00	2581174790.71	1000000000000000000.00
19	2.11	8.14	19.00	44.33	171.00		6.40	876.28	524287.00	8603915971.38	10000000000000000000.00
20	2.22	8.57	20.00	46.67	180.00		7.23	1252.25	1048575.00	28679719906.92	100000000000001000000.00
LINK-LINK formula: kp_f/(1-p_f)											
E2E formula: [1-(1-p_f)^k]/(1-p_f)^k											

Internetwork Operation



Interoperability Gateways



CoS the USPS Way…

Option Name	Mailing Receipt	Delivery Record	Air Delivery	Recipient Pays	Moves Money	Delivery Confirm	Return Receipt	Careful Handling	Insurance	Restricted Delivery	Signature Confirm
Cert. Of Mailing-RM	Y		(w/PAL)					(w/SH)			
ParcelAirLift (PAL)			Y								
Special Handling SH			(w/PAL)	(w/COD)		(w/DC)	(w/RR)	Y	(w/IM)		(w/SC)
Certified Mail CM	Y	Y					(w/RR)			(w/RD)	
COD	(w/RM)	Y		Y		(w/DC)	(w/RR)	(w/SH)	(w/RM)	(w/RD)	(w/SC)
Delivery Confirm DC				(w/COD)		Y	(w/RM)	(w/SH)	(w/IM or RM)		
Insured Mail IM			(w/PAL)			(w/DC)		(w/SH)	Y		(w/SC)
Money Order					Y						
Return Receipt RR	Y	Y	(w/PAL)			(w/DC)	Y	(w/SH)		(w/RD)	(w/SC)
Registered Mail RM	Y	Y		(w/COD)		(w/DC)	(w/RR)		Y	(w/RD)	(w/SC)
Restricted Delivery RD			(w/PAL)			(w/DC)	(w/RR)	(w/SH)		Y	(w/SC)
Sig. Confirm		Y				Y					Y

DTN CoS

- Classes of Service for a Bundle:
  - Types: Expedited, Regular, Bulk
  - Options: send notification, keep delivery record, inform on delivery
- Stamps encode CoS, are not forgeable, and are obtained by sender from trusted service
- DT routers can verify CoS in stamp using network “forwarding service” key

Postage Stamps

- Each bundle contains a cryptographically-signed "postage stamp"
  - Similar to Kerberos tickets
- Provides authorization to use the network at a particular class of service for a particular message
- Postage stamps are verified at each P node
  - NP nodes may not store any complete bundle
  - Endpoint P nodes are special (later)

Related Work

- Protocol Architecture
  - ARPANET design, NewArch, IPN
- Naming, Addressing and Routing
  - Intentional naming, CHORD, CAM
  - IPNL, TRIAD, RON
  - Diffusion Routing, MPLS (sort of)
- Extreme Links
  - Many (WHOI, JPL, UCB, UW, MIT, …)

Futures

- DTN work based on earlier IPN Architecture
  - Interplanetary Internet www.ipnsig.org
  - Mitre, JPL, MCI and others
  - DTN generalizes to non-space environments
- Investigations
  - Army TI and Special Forces Ops
  - Heterogeneous UCB/Intel/JPL Sensor Nets
  - UWB Developments
    - BWRC, Intel, UCSD, Rutgers, USC

Thank you for listening…