Delay-Tolerant Networking: *Issues & Recent Results*

Kevin Fall

Intel Research, Berkeley

kfall@intel.com

http://WWW.DTNRG.ORG

19th IEEE Workshop on Computer Communications

October 19, 2004 / Bonita Springs, FL

Intel Research Berkeley

Outline

– *Why the Internet Architecture is not a ‘one-size-fits-all’ solution*
– *The DTN Routing Problem*
– *Recent Implementation Results*

Unstated Internet Assumptions

- End-to-end RTT is not terribly large
  - A few seconds at the very most [typ < 500ms]
  - (window-based flow/congestion control works)
- Some path exists between endpoints
  - Routing usually finds single "best" existing route
    - [ECMP is an exception]
- E2E Reliability using ARQ works well
  - True for low loss rates (under 2% or so)
- Packet switching is the right abstraction
  - Internet/IP makes packet switching interoperable

Intel Research Berkeley

Non-Internet-Like Networks

- Stochastic mobility
  - Military/tactical networks
  - Mobile routers w/disconnection (e.g. ZebraNet)
- Periodic/predictable mobility
  - Spacecraft communications (LEO sats)
  - Busses, mail trucks, delivery trucks, etc. (InfoStations)
- "Exotic" links
  - Deep space [Mars: 40 min RTT; episodic connectivity]
  - Underwater [acoustics: low capacity, high error rates & latencies]
  - Sensor networks, mules

DTN challenges…

- Intermittent/Scheduled/Opportunistic Links
  - Scheduled transfers can save power and help congestion; scheduling common for esoteric links
- High Link Error Rates / Low Capacity
  - RF noise, light or acoustic interference, LPI/LPD concerns
- Very Large Delays
  - Natural prop delay could be seconds to minutes
  - If disconnected, may be (effectively) much longer
- Different Network Architectures
  - Many specialized networks won’t/can’t ever run IP

What to Do?

- Some problems surmountable using Internet/IP
  - ‘cover up’ the link problems using PEPs
  - Mostly used at “edges,” not so much for transit
- Performance Enhancing Proxies (PEPs):
  - Do “something” in the data stream causing endpoint (TCP/IP) systems to not notice there are problems
  - Lots of issues with transparency– security, operation with asymmetric routing, etc.
- Some environments *never* have an e2e path
  - Consider an approach tolerating disconnection, etc...

Delay-Tolerant Networking Architecture

- Goals
  - Support interoperability across ‘radically heterogeneous’ networks
  - Acceptable performance in high loss/delay/error/disconnected environments
  - Decent performance for low loss/delay/errors
- Components
  - Flexible naming scheme with *late binding*
  - Message overlay abstraction and API
  - Routing and link/contact scheduling w/CoS
  - Per-(overlay)-hop reliability and authentication

Message Overlay Abstraction

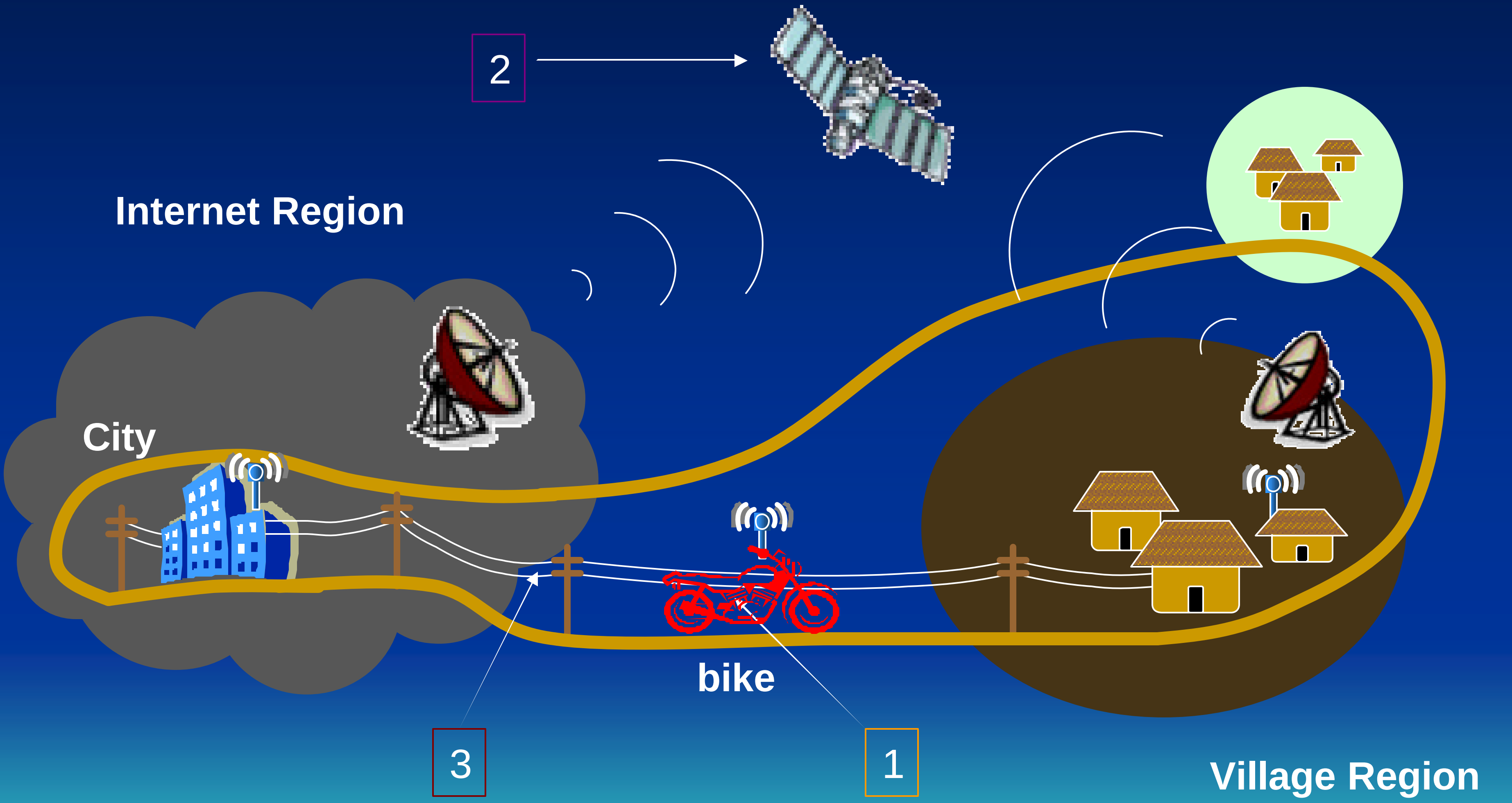
- E2E Async Message Service: “Bundles”
  - “postal-like” message delivery over regional transports with coarse-grained CoS [4 classes]
  - *Options*: return receipt, “traceroute”-like function, alternative reply-to field, custody transfer
  - Supportable on nearly any type of network
- Applications send/receive <u>messages</u>
  - “Application data units” of possibly-large size
  - May require framing above some transport protocols
  - API supports response processing long after request was sent (application *re-animation*)

So, is this just e-mail?

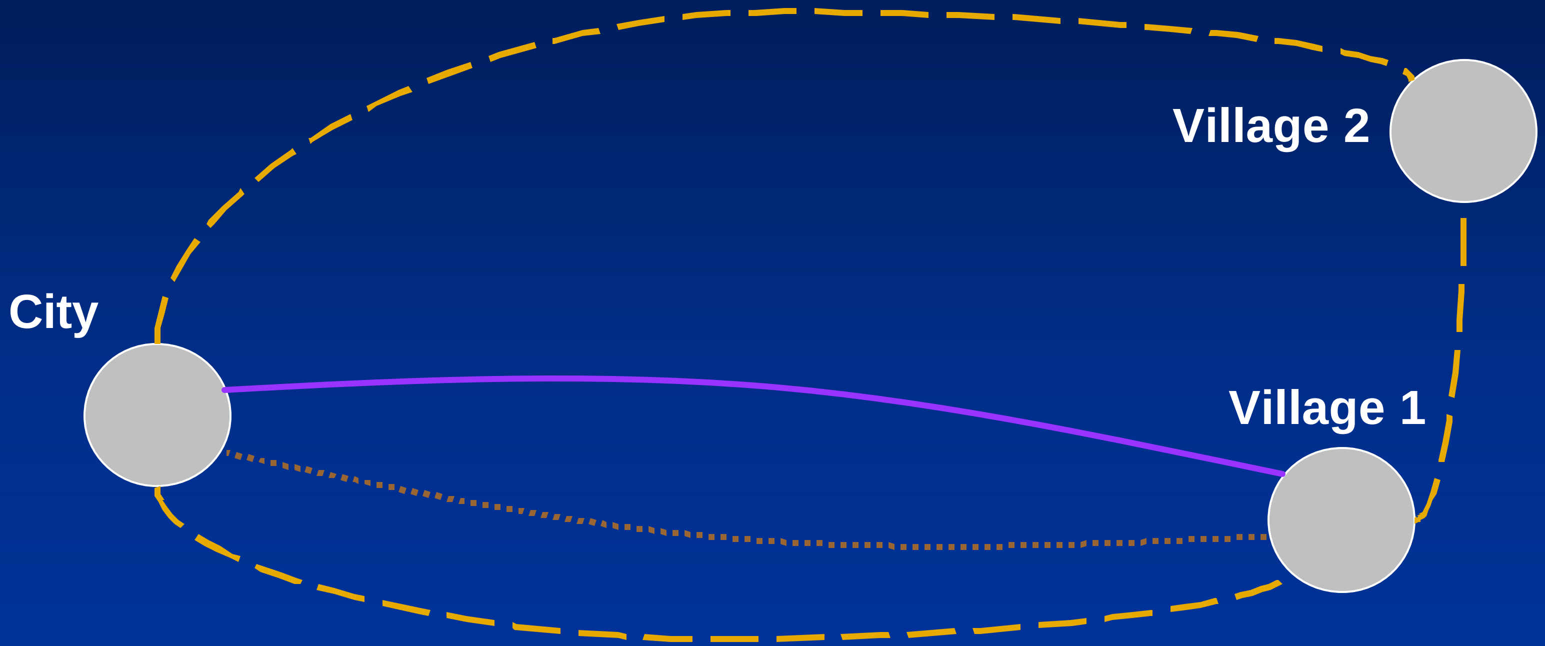
	naming/ late binding	routing	flow contrl	multi- app	security	reliable delivery	priority
e-mail	Y	N	N(Y)	N(Y)	opt	Y	N(Y)
DTN	Y	Y	Y	Y	opt	opt	Y

- Many similarities to (abstract) e-mail service
- Primary difference involves routing/restart and API
- E-mail depends on an underlying layer's routing:
  - Cannot generally move messages closer to their destinations in a partitioned network
  - In the Internet (SMTP) case, not disconnection-tolerant or efficient for long RTTs due to "chattiness"
- E-mail security authenticates only user-to-user

Example Routing Problem



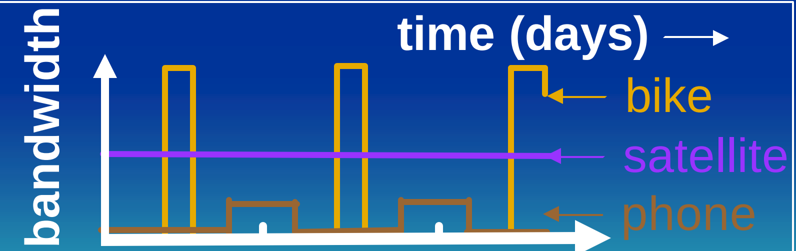
Example Graph Abstraction



bike (data mule)
intermittent high capacity

Geo satellite
medium/low capacity

dial-up link
low capacity



Connectivity: Village 1 – City

Routing on Dynamic Graphs

- DTN routing takes place on a time-varying topology
  - Links come and go, sometimes predictably
  - Use any/all links that can possibly help
- Scheduled, Predicted, or Unscheduled Links
  - May be direction specific [e.g. ISP dialup]
  - May learn from history to predict schedule
- Messages fragmented based on dynamics
  - <u>Proactive fragmentation</u>: optimize contact volume
  - <u>Reactive fragmentation</u>: resume where you failed
  - Both are important for high utilization of precious link resources

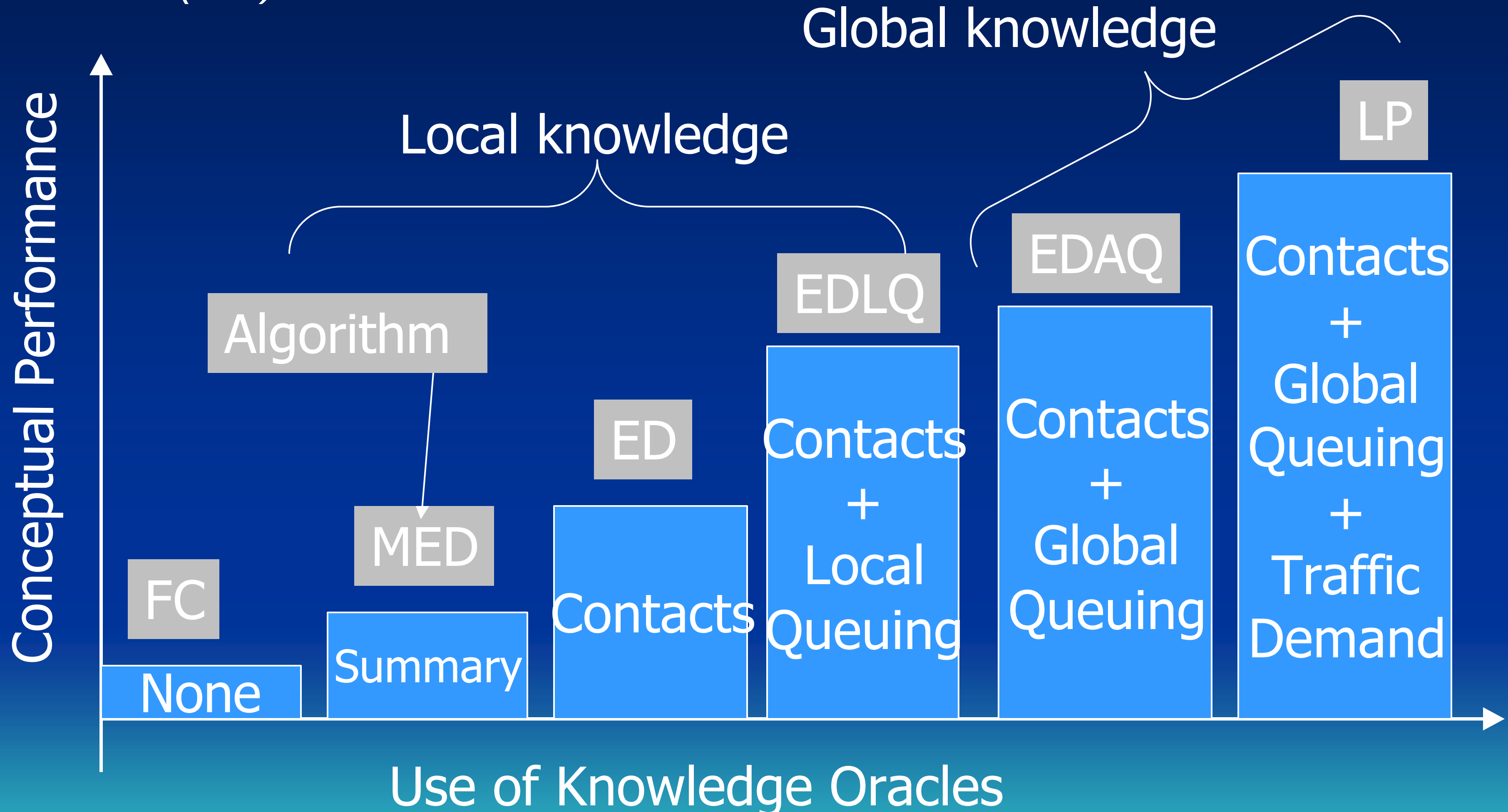
Intel Research Berkeley

The DTN Routing Problem

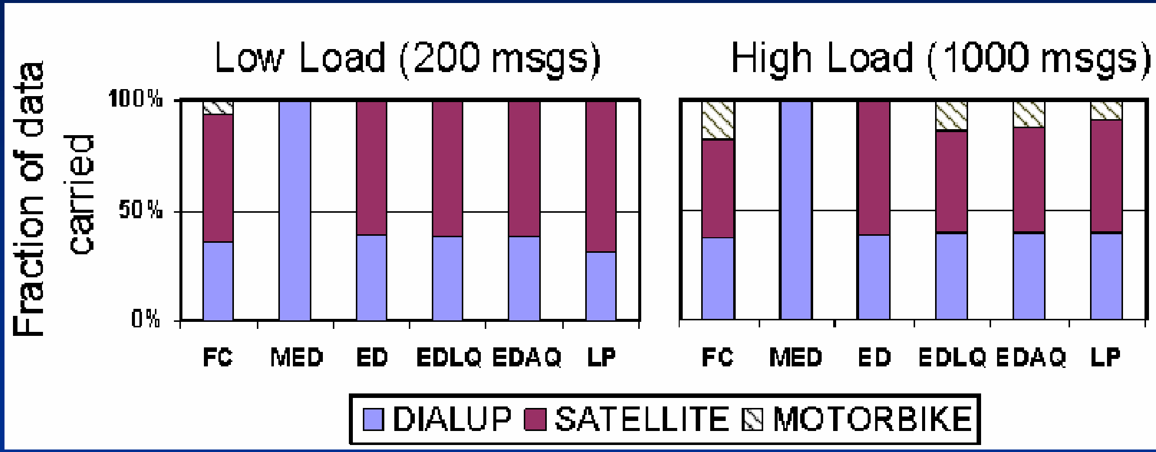
- *<u>Inputs</u>*: topology (multi)graph, vertex buffer limits, contact set, message demand matrix (w/priorities)
- An ***edge*** is a possible opportunity to communicate:
  - One-way: (S, D, c(t), d(t))
  - (S, D): source/destination ordered pair of contact
  - c(t): capacity (rate); d(t): delay
  - A ***Contact*** is when c(t) > 0 for some period $[i_k, i_{k+1}]$
- Vertices have buffer limits; edges in graph if ever in any contact, multigraph for multiple physical connections
- *Problem*: optimize some metric of delivery on this structure
  - Sub-question: what metric to optimize?

Knowledge vs Performance

S. Jain (UW): SIGCOMM 2004

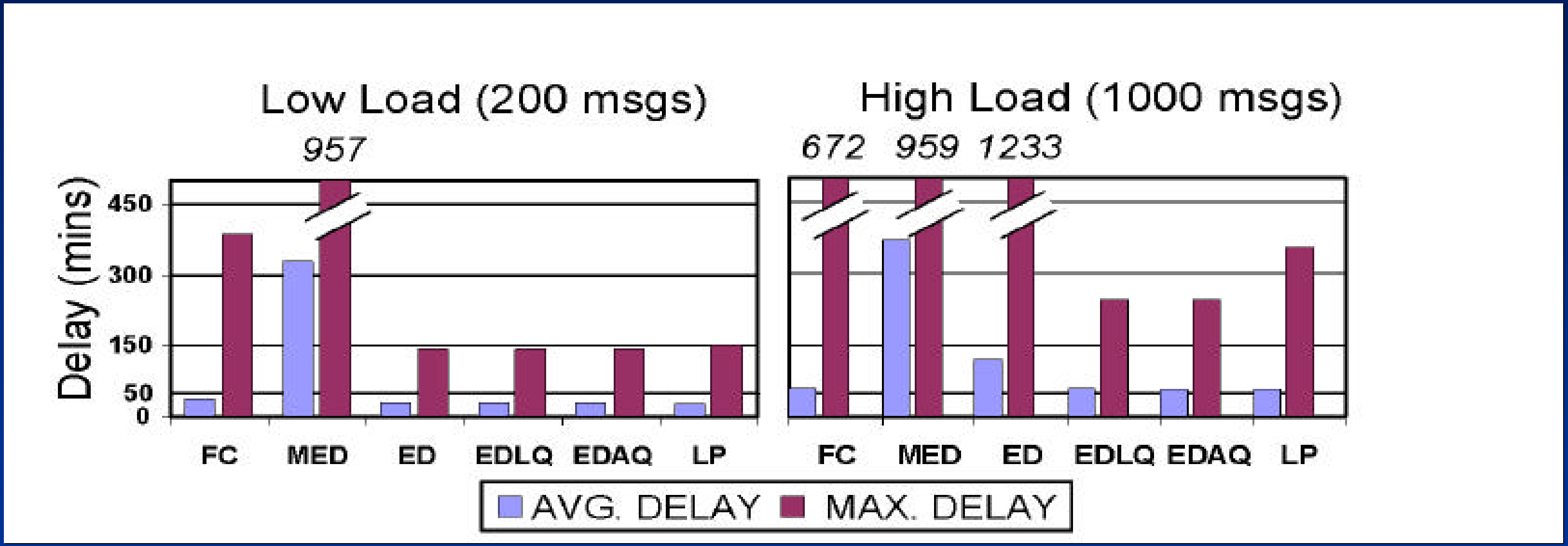


Data Allocations by Algorithm



Min Expected Delay (MED): All data is carried by dialup
Earliest Delivery (ED): Same for low and high load.
{Split between dialup and satellite}
ED, EDLQ, EDAQ make same choices for low load
EDLQ, EDAQ start to use bike also

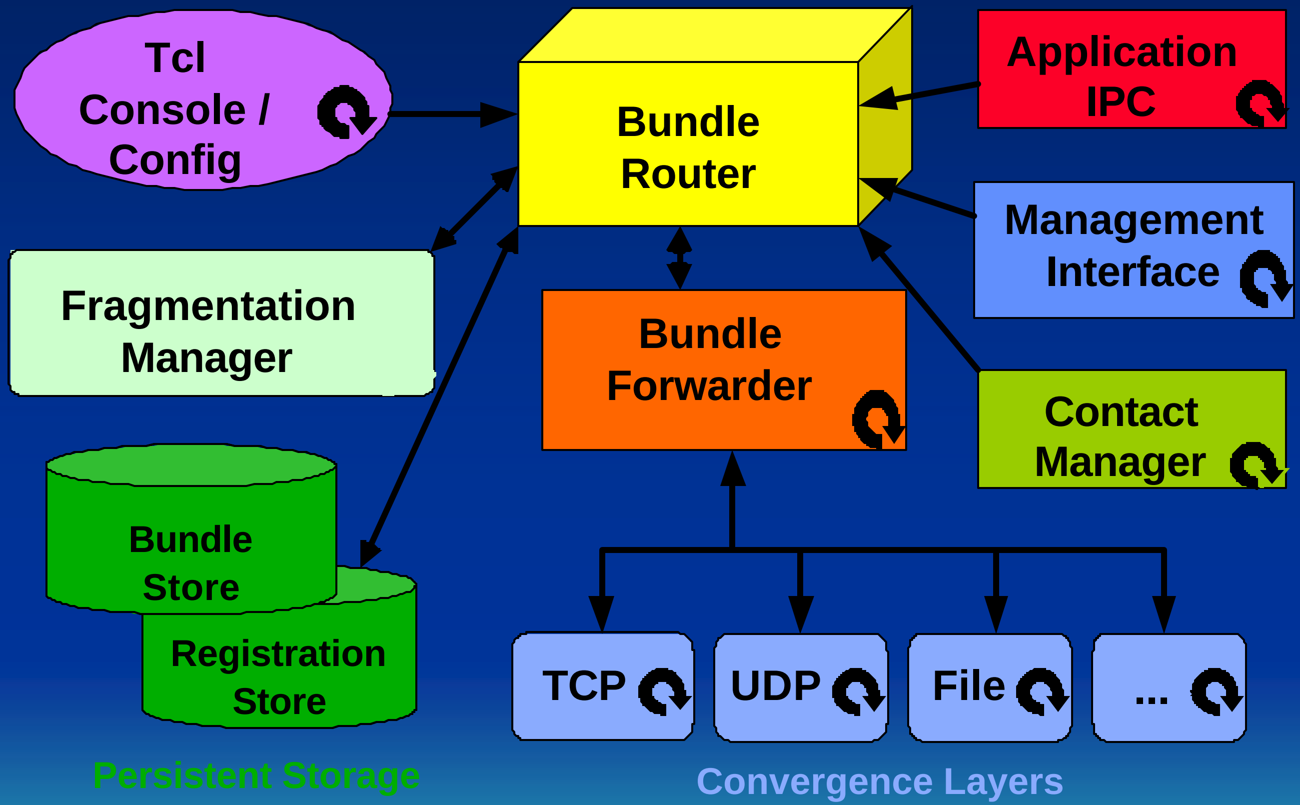
Delivery Delay Comparison



Low load: ED, EDLQ, EDAQ approx. same performance
High load: EDLQ, EDAQ are optimal. ED is much worse
MED has high delay in both cases
FC performs well on average delay
but has much worse max delay

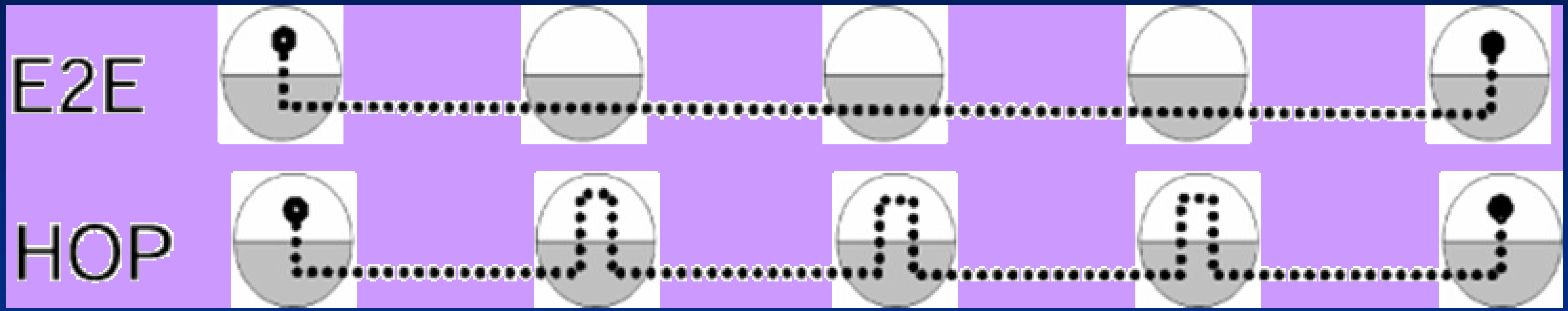
Intel Research Berkeley

‘DTN2’ Implementation



Intel Research

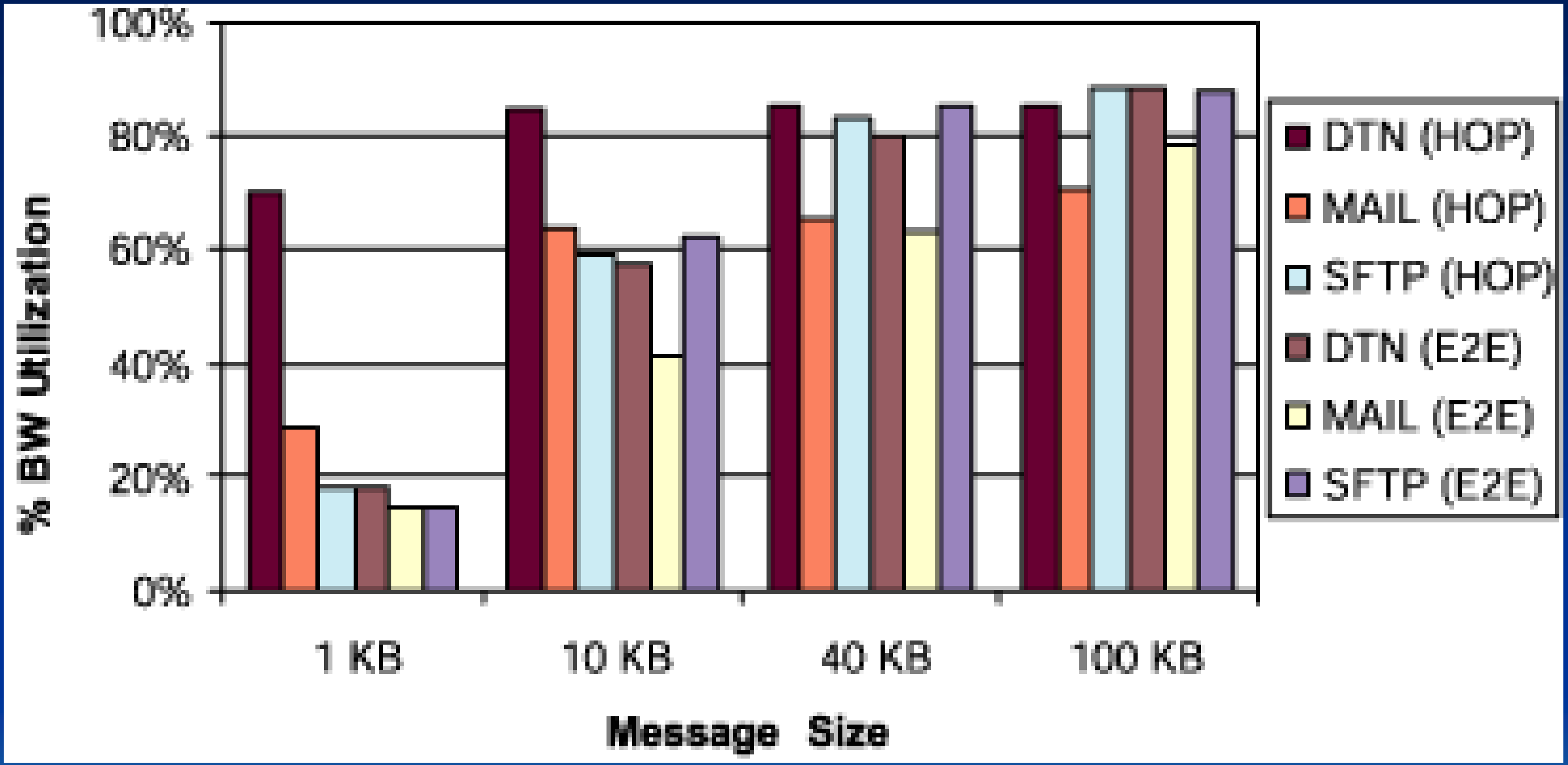
Experiment Setup



- Compare robustness to interruption / link errors
- Approaches compared
  - End-to-end TCP (kernel routing)
  - Proxied (TCP ‘plug proxies’)
  - Store-and-forward (Sendmail, no ckpoint/restart)
  - DTN (store-and-forward with restart)
- Link up/down patterns: aligned, shifted, sequential, random

Intel Research Berkeley

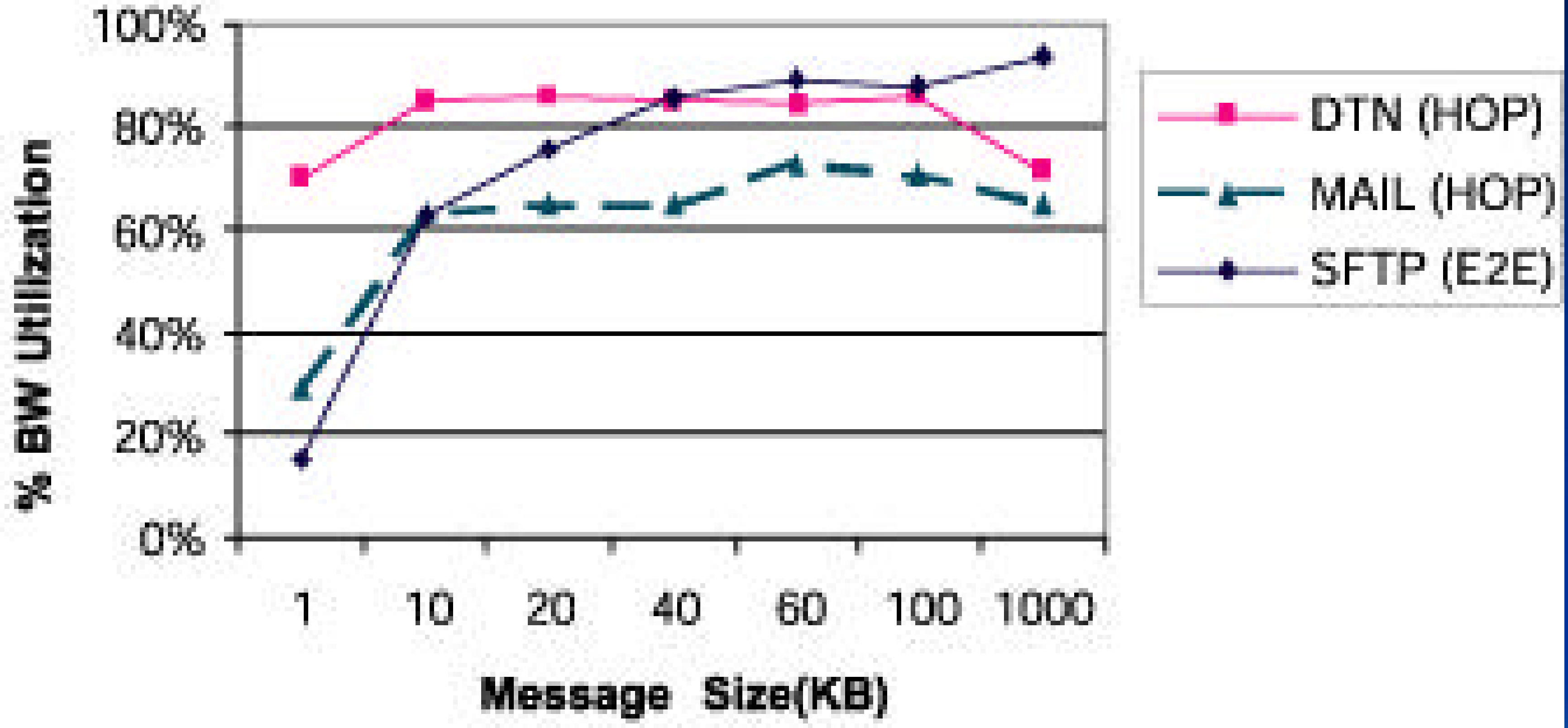
BW Efficiency



No disruptions: DTN does well for small msgs, little overhead overall

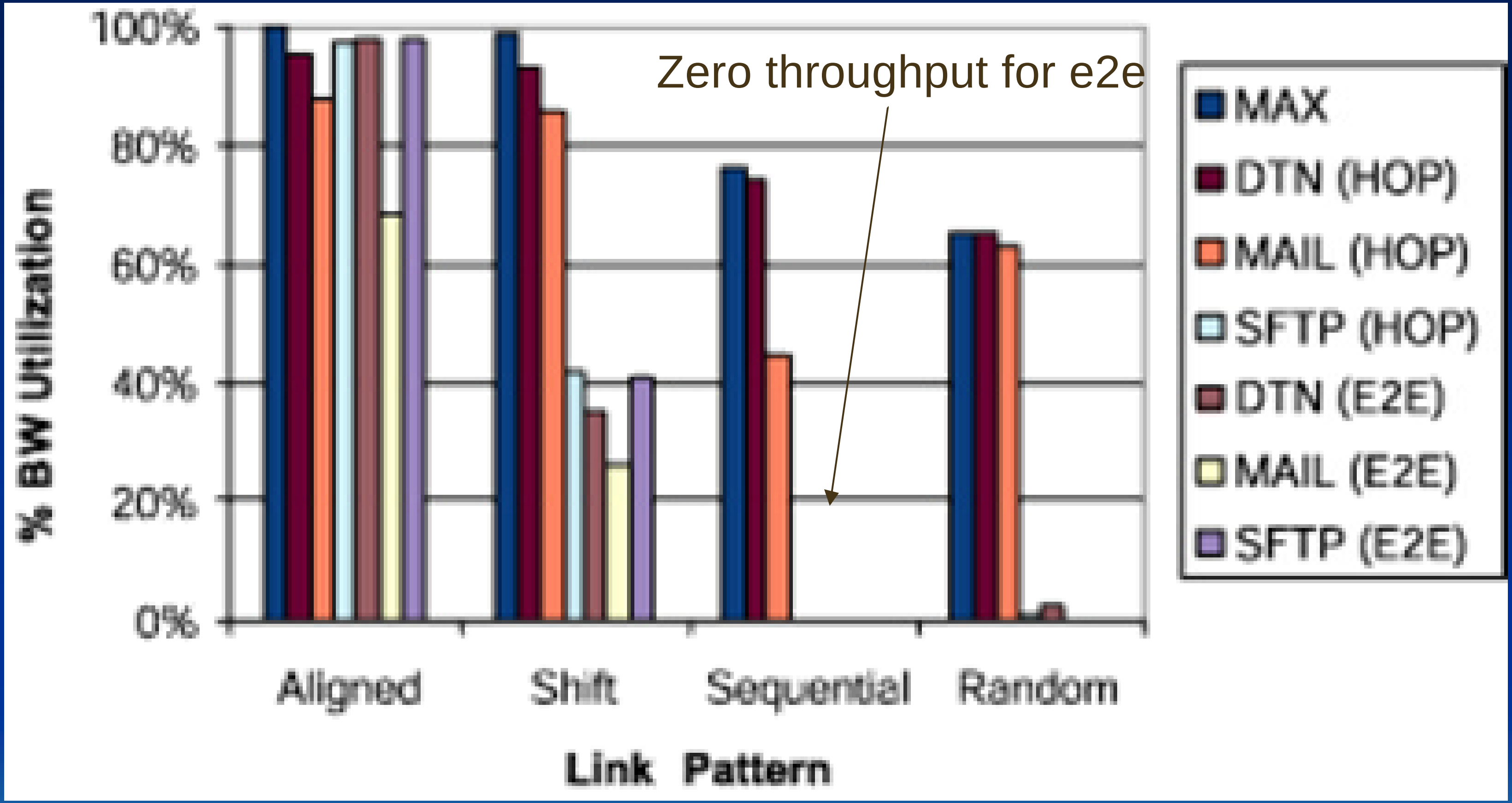
Intel Research Berkeley

Efficiency Trend



Store-and-forward delays increase w/msg size

Interruption Tolerance



Up/down 1m/3min; 40kb messages; shift 10s

Conclusions

- DTN foundational concepts appear to have wide applicability
- DTN Routing is a rich and challenging problem
- Reference implementation can be tricky
- Early performance results suggest our approach to disruption tolerance is effective

Status

- IETF/IRTF DTNRG formed end of 2002
  - See http://www.dtnrg.org
- DTN1 Agent Source code released 3/2003
- SIGCOMM Papers: 2003 [arch], 2004 [routing]
- Several other documents (currently ID’s):
  - DTNRG Architecture document
  - Bundle specification
  - Application of DTN in the IPN
- Basis for new DARPA DTN program
- Part of NSF ‘ICT4B’ Project (with UCB)

Acknowledgements

- DTN/ICT4B People:
  - Eric Brewer, Mike Demmer, Rabin Patra (UCB)
  - Bob Durst, Keith Scott (MITRE)
  - Kevin Fall, Melissa Ho (Intel Research Berkeley)
  - Sushant Jain (UW Intern)
  - S. Keshav (U Waterloo)
  - Ting Liu (Princeton)
- Other DTN (IPN) People:
  - Vint Cerf (MCI)
  - Scott Burleigh, Adrian Hooke (NASA/JPL)
  - Forrest Warthman (Warthman)
  - Stephen Farrell (Trinity College, Ireland)
  - The *dtn-interest* mailing list

http://www.dtnrg.org

Thank You!

http://tier.cs.berkeley.edu

Intel Research Berkeley

On to an application…

ICT for Billions (ICT4B)

- Information and Communication Technologies for Developing Regions of the World
- Networking focus: *intermittent networking*
  - *Mission-specific architecture and API*
  - *Multiple layers of network intermittency*

Intel Research Berkeley

ICT4B Application Areas

- E-Government
  - Forms, status updates, certifications
- Health
  - Early screening
- Trade
  - Price dissemination, market making
- Education
  - Various topics: health, agriculture, microfinance, etc.
- Alerts / News / Weather
- General communication

Intel Research Berkeley

ICT4B Technology Areas

- Task-Specific Devices
  - Hand-held with speech recognition
  - Local wireless
  - Sensors
  - Uses: Medical, data entry, information, etc.
- Intermittent Networking
  - DTN forms the underlying networking technology
  - Capable of supporting async messaging over most any comms technology
- Distributed System Architecture
  - Back-end services in data center (databases, trading system, etc.)
  - Village-level kiosks (cache, I/O capability with devices, printer)
- Speech Recognition
  - Speaker-independent small-vocabulary approach
  - (currently taking samples in Tamil)
- Very Low Cost Displays
  - Using ink-jet printing approach

Some of The Team…[7/2004]

Intel Research Berkeley

MSSRF (Villianur)…[7/2004]

MSSRF (Kizhur?)...[7/2004]

MSSRF (Veerampattinam)…[7/2004]

Intel Research Berkeley

ICT4B Project Status

- ICT4B NSF ITR funded 10/2003 (5yr)
- DTN forwarding layer and early apps being tested (code released 3/2003)
- Joint UCB/Intel attendance at ‘ICT for Sustainable Development’ conference Jan 2004/Bangalore; ‘Bridging the Divide’ conference Mar 2004/Berkeley; ‘Digital Rally’ Apr 2004/San Jose; PolicyMaker’s Workshop July 2004/Delhi
- Fellow travelers: HP Labs India, IIT Bombay/Kanpur/Madras, Univ. of Washington, MITRE, DARPA, NSF, CMU, UCLA, JPL, U Waterloo, MCI

Intel Research Berkeley