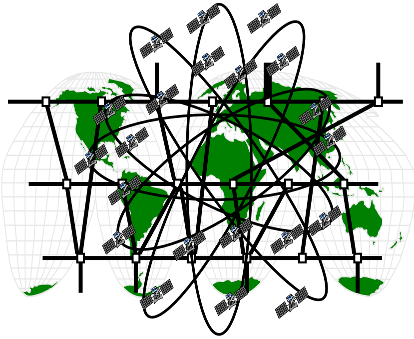


UNIVERSITÀ DEGLI STUDI DI TRENTO

Dipartimento di Ingegneria e Scienza dell'Informazione

The Future Internet: Challenges and Solutions

Yoram Ofek

University of Trento – Italy

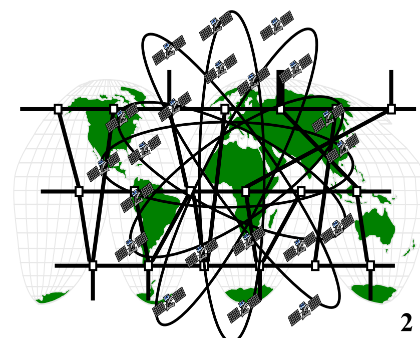
ofek@dit.unitn.it; http://dit.unitn.it/ofek/ +39 0461 88 3954 (office)

Keynote Presentation – NAEC2008

September 25-28, 2008 – Riva del Garda, Italia

Outline

- Vision and Challenges
- Adding the global time (UTC) dimension
- Consequences of networking with time
- Deployment strategy – as the Internet underlay
- Solutions to the various challenges
- Summary



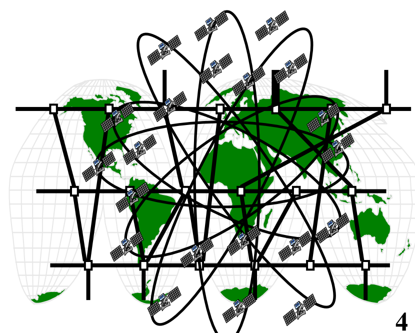
An Internet Vision

- Two primary Internet (telecom) transitions:
  - 1. From **business** to **home/mobile (wireless) users**
    - The capacity per home will equal campus capacity
      - **20 Mb/s campus pays 1000 Euro/month vs. home 30 Euro/month**
  - 2. From (scheduled) **broadcast** to (all-IP) **on-demand:**
    - Triple-any:
      - **Anyone (any skills)** **[2 billion users or more in the future]**
      - **Anything (any-service)**
      - **Anytime (no rigid schedule)**
    - Global scale: **from anywhere to anywhere**

An Internet Vision (Cont.)

- Continuous exponential traffic growth
  - 50-100 (perhaps more) folds in 5-15 years
    - – Faster than "Moore's Law" (in switching)
  - It's envisioned
    >90% of traffic will be to home/mobile users

Mostly for entertainment - non-business!

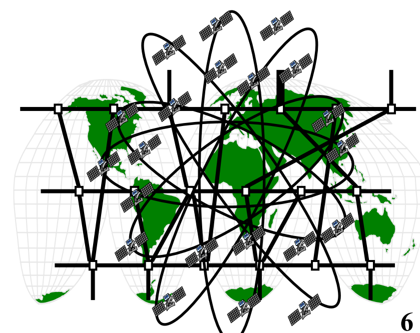


Why (IP) Packet Switching?

- It is the most flexible method to match:
  - User's diverse desires and capabilities (dynamic range up to 1000):
    - Display quality
    - Processing speed
    - Encoding
    - Access capacity
    - …
  - For combined wired (optical) and wireless networks
- While liberating users from the rigid TV broadcasting schedules and quality!

The Cost/Revenue Challenges

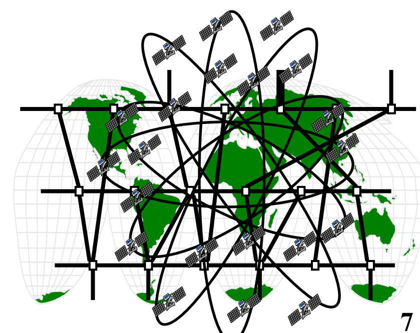
- **Who will pay?** **[for 50-100 (or more) larger Internet]**
  - **Given that service providers will NOT lose money again …**
- **What is roughly required**
  - **Factor of 20 lower cost (better scalability index)**
    PLUS:
  - **Revenue from selected premium services, which are guaranteed so people will be willing to pay extra**
    - **STREAMING HIGH QUALITY LIVE EVENTS**



The Switch/Router Scalability Challenge

- Cisco CRS-1 [carrier routing system]: 92/(2*72) ≈ 0.64 Tb/s per chassis (72 Chassis)
  - Factor of 2 improvement in 5 years/$500 millions
- **Internet traffic is doubling, say, every 18 mo.**
  - **50-100 folds in 5-15 years!**

- **Switching scalability to multi-terabit per second (10-100 Tb/s) per chassis is required**



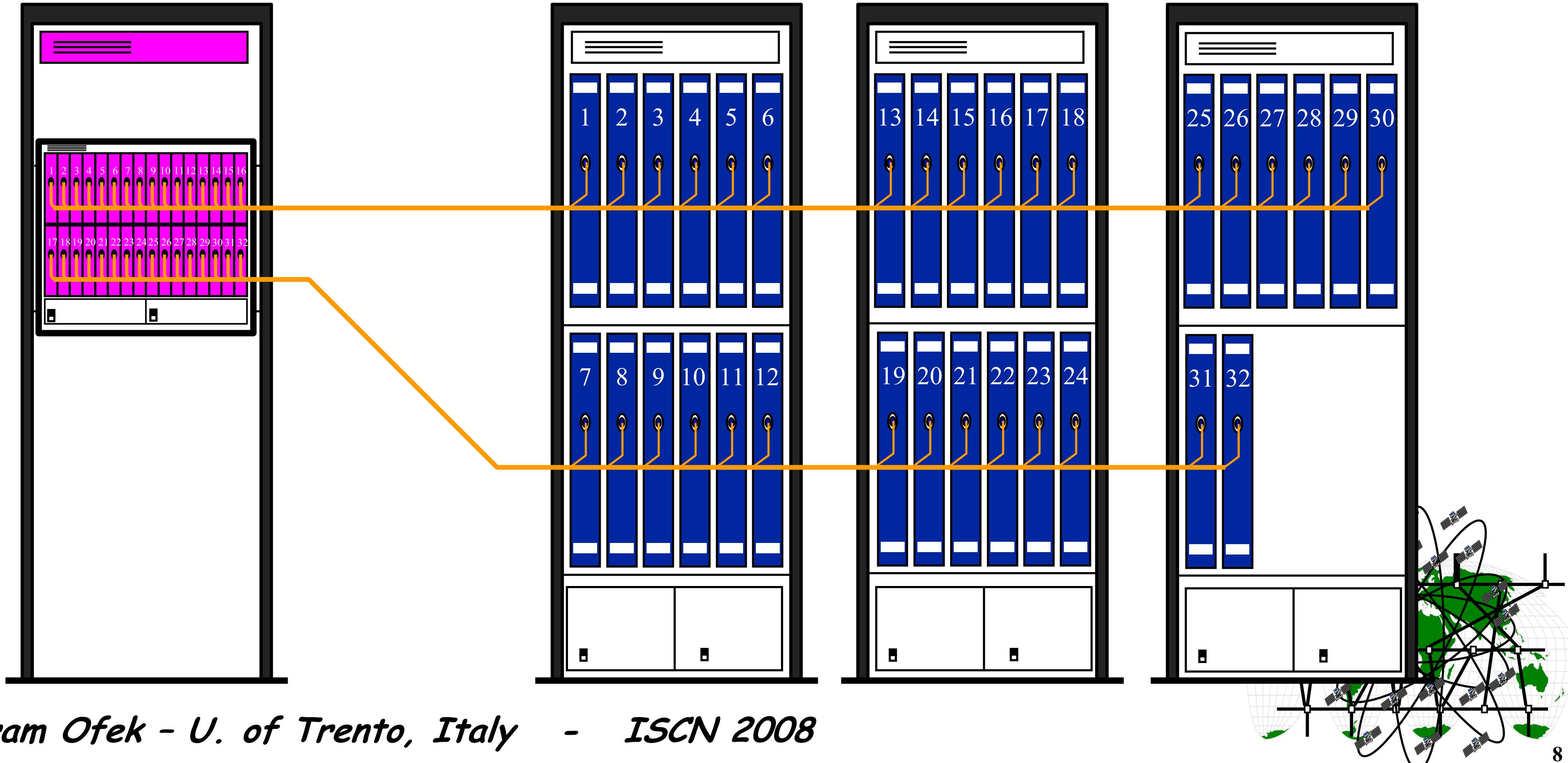
The Switch/Router Scalability Challenge

Fourth-Generation Switches/Routers

Clustering and Multistage – 46 Tb/s

Cisco CRS-1 --- 72 Chassis

(0.64 Tb/s per chassis)



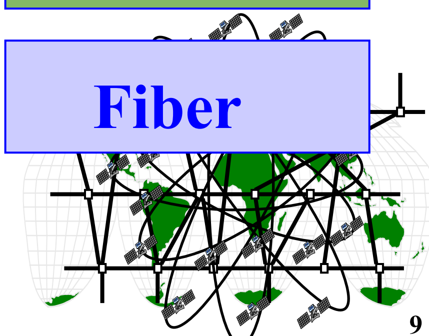
Protocol Challenge:

Confusing, Costly and No Single Network …

Away from ATM →

"Zoo" of Protocols – WHY?

2000	Alternative 1	Alternative 2	Alternative 3	Alternative 4
IP	IP	IP	IP	IP
ATM			RPR	RPR
Sonet	NG Sonet		NG Sonet	
Lambda	Lambda	G.709 DWDM	Lambda	G.709 DWDM
Fiber	Fiber	Fiber	Fiber	Fiber

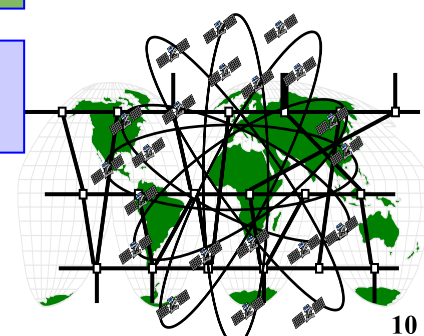


Protocol Challenge:
Confusing, Costly and No Single Network … …

"Zoo" of Protocols - WHY?

Alternative 5	Alternative 6	Alternative 7	Alternative 8
IP			
GMPLS	IP		
Ethernet	GMPLS	IP	IP
NG Sonet	Ethernet	GMPLS	GMPLS
Lambda	G.709 DWDM	Ethernet	G.709 DWDM
Fiber	Fiber	Fiber	Fiber

And more …

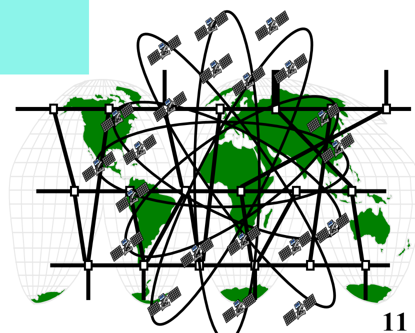


Live Streaming & Download Challenges

Video, Video, …

Streaming / Content Playing

- **On-demand:** TV programs, live sport events (e.g., foot-ball!), movies, HD, 3D, …
  - Guaranteed capacity low delay & loss
  - Easy access via the WWW
    - “TV guide-like” on-line
    - Movie guide on-line
    - …



Live Streaming & Download Challenges

From

1-Click = 50 Packets (50 KB)

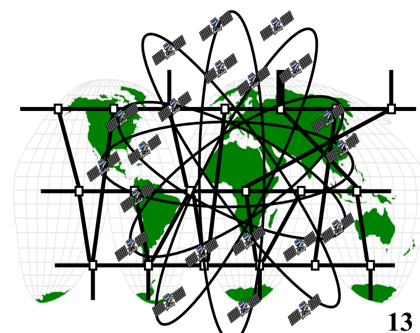
To

1-Click = 1,000,000-10,000,000 Packets (1-10 GB)

Few "clicks" and the network is "flooded"

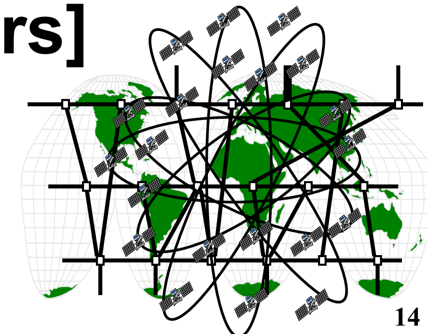
Live Streaming & Download Challenges
the “Good News”

- Traffic characterization is predictable
- Not all (media content) bits should arrive correctly
    - Reliable transmission is not a requirement (in most cases)
- In 5-10 year advanced allocation will be applicable to >90% of the IP traffic



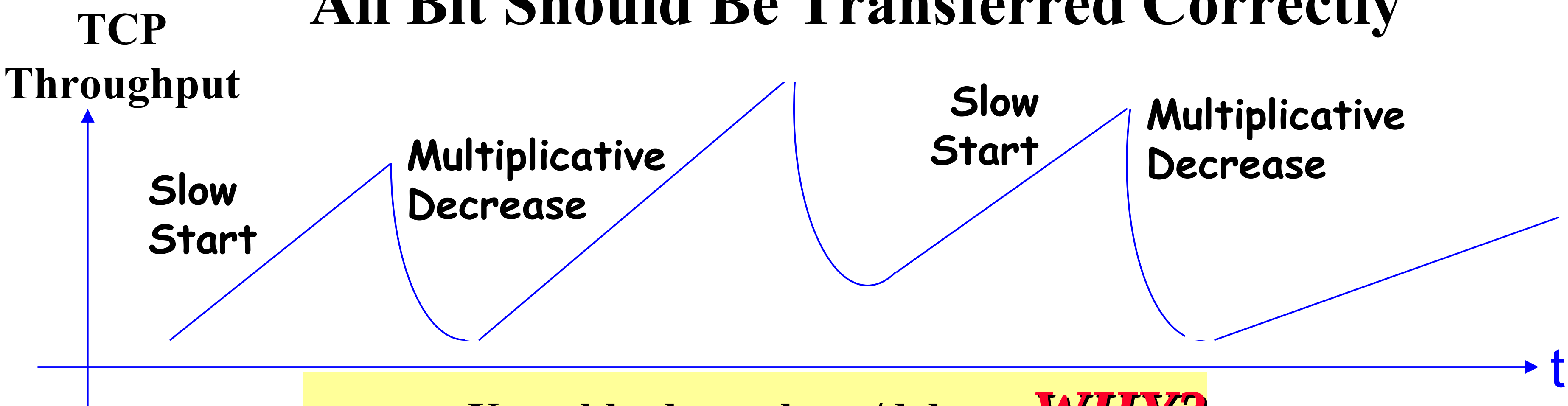
Live Streaming & Download Challenges
the "Unfortunate News"

- Today, 70%-80% of the Internet traffic is peer-to-peer downloads Over TCP!?
  - TCP provides reliable transmission [all bits will arrive correctly]
    - But without predicable throughput and delay
  - Streaming media, download, interactive, VoIP may tolerate defined packet loss but require predictable throughput and delay [rather than adaptive – sub-optimal for users]



TCP the "Unfortunate News"

Reliable Transmission:
All Bit Should Be Transferred Correctly

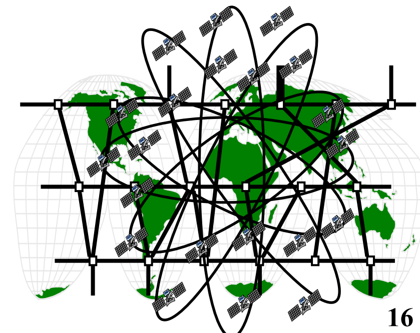


Unstable throughput/delay – *WHY?*

- **The issues:**
  - **"Neurotic" mechanism (with "ups" and "downs")**
    - **Unpredictable throughput – low utilization**
      BUT as a result the INTERNET doesn't COLLAPSE
  - **Assuming that all users are FAIR?**

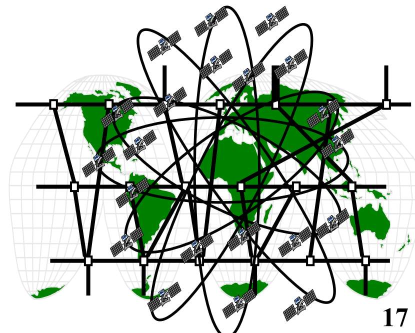
Interactive Video-based App. Challenges TCP cannot be used

- Interactive apps are even more challenging
- Users expect:
  - Minimal delay
  - Continuous play
  - Constant video quality
  - Lips synchronization



The Wireless and Mobility Challenges

- Shared medium
- Collisions/contentions/interference/SINR
  - Random back-off
    - Unpredictability for streaming (premium service)
      - Much like TCP
    - The random back-off is critical to avoid collapse



Green Internet: Power Usage

- **World population: 6,676,120,288**
  - **Number of Internet users 1,407,724,920**
  - **Penetration 21.1 %**
    - **Source: www.internetworldstats.com**
- **Switching and Routing 34%**
- **Regeneration 27%**
- **Processing 22%**
- **Storage 10%**
- **Transport 7%**
    - **"Data Centers Network Power Density Challenges" By Alex Vukovic, ASHRAE Journal, (Vol. 47, No. 4, April 2005).**

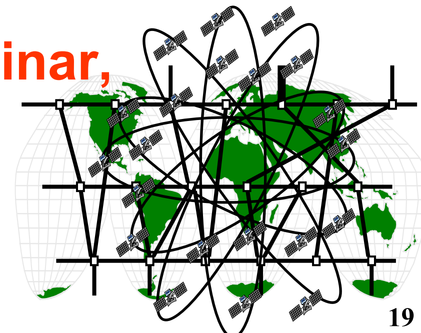
Green Internet: Power Usage

Today Internet
uses ~1% of total world electricity usage
(excluding PCs, customers equipment, mobile etc.)

If 2 Billion people have broadband access (1Mb/s) then ~5%?

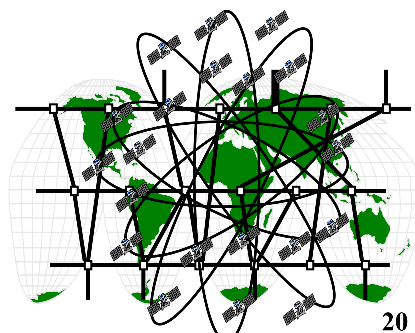
If 2 Billion people have broadband access (10 Mb/s) then ~50%?

Source: R.S Tucker, "A Green Internet" - May 2007, CUBIN Seminar, The University of Melbourne



Outline

- Vision and Challenges
- Adding the global time (UTC) dimension
- Consequences of networking with time
- Deployment strategy – as the Internet underlay
- Solutions to the various challenges
- Summary

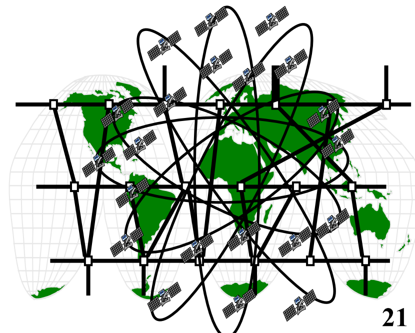


Adding Global Time (UTC) Dimension

http://www.youtube.com/watch?v=MU0UEPdcGE8&feature=related

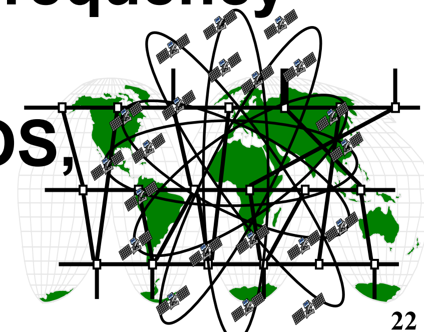
http://www.youtube.com/watch?v=Ezv8U_Flukk&feature=related

- An optional feature – since IP remains as is
  - Used to solve the various challenges when needed:
    - Cost & revenue
      - **Lower** complexity=cost of IP infrastructure – switch scalability
      - **Increase** revenue from premium services
    - Live streaming and interactive applications – without TCP
    - "Zoo" of protocols
    - Wireless
    - Green Internet



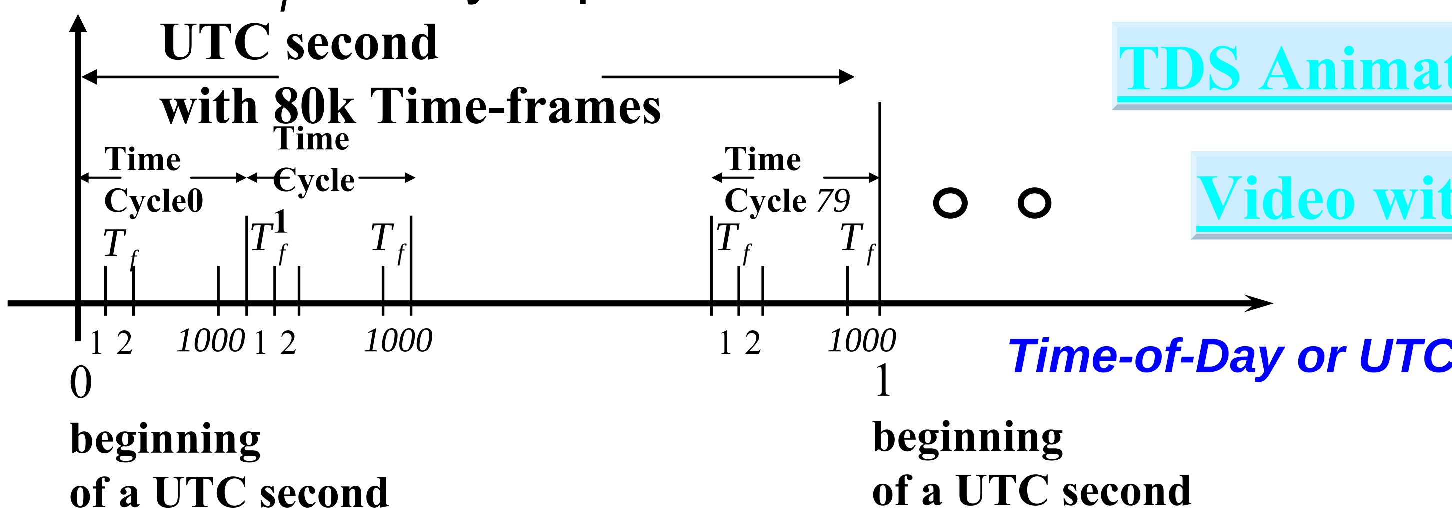
Global Time Distribution (Free) Solutions

- <u>In space 4 satellite constellations:</u>
  - GPS – US – 31 satellites,
  - Glonass – Russia –24 satellites,
  - Galileo – EU – 1 experimental satellite
  - Compass/Beidou – China – 1 experimental satellite.
- <u>In-band time distribution:</u>
  - With/without standard protocols: IEEE 1558, IETF NTPv2…
- <u>On earth:</u>
  - E-LORAN - Enhanced Loran (LOng RAnge Navigation) is a terrestrial radio navigation system using low frequency radio transmitters
  - other time distribution, such as: WAAS, EGNOS, MSAS, DGPS (differential GPS), CDMA, …



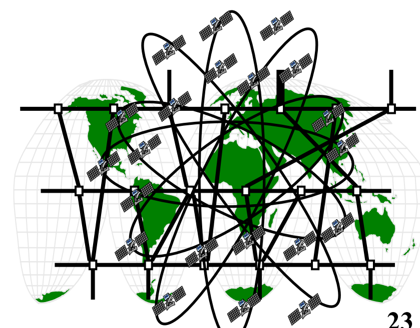
Pipeline Forwarding with UTC
Factor of 20 Lower Cost / Premium Services

- **Pipelines are deployed to increase efficiency:**
  - **Optimal method** - independent of a specific realization
  - **Factory (automotive) / computers (CPU)**
- **Internet Pipeline thanks:** **GPS/Galileo/multitude of other sources**
  - **Time frames as virtual containers for IP packets**
    - **Thus, no header processing**
    - T_f **accuracy of 1µs is sufficient**



[TDS Animation]

[Video with TDS]



Switch Scalability with Global Time

[Optimal Complexity → Lower Cost/Power → Green Internet]

Scheduling Controller (per time frame 10-100μs)

Computation complexity: *O(NlgN)*

Input port

BW(link)

$q_{in} = 1$ (optimal)
$s = 1$ (optimal)

0 *0*

Switching Fabric
***N* Ports**
Complexity:
NlgN
(optimal)
$s = 1$
(optimal)

N-1 *N-1*

BW(link)

BW(link)

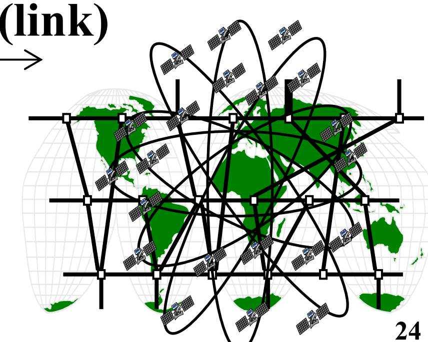
$q_{in} = 1$ (optimal)
$s = 1$ (optimal)

BW(link)

Input port

s – speedup
q_{in} – buffer size

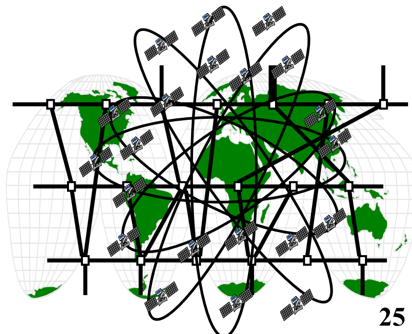
Pipeline



Networking with Global Time (UTC)

[Lower Cost/Power → Green Internet]

- No "stopping" of the serial bit stream
  - No header processing / no segmentation / no reassembly
- Delay per hop: constant
  - Jitter per hop: zero
- Loss: none due to congestion
- "Bonus": QoS for streaming media / interactive
  - – [Sort of a "negative option"]



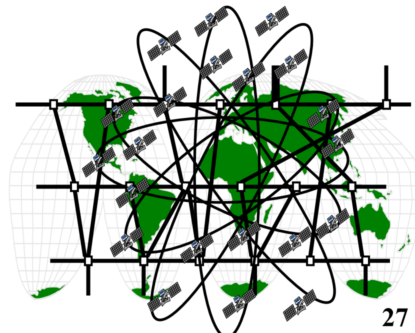
Outline

- Vision and Challenges
- Adding the global time (UTC) dimension
- Consequences of networking with time
- Deployment strategy – as the Internet underlay
- Solutions to the various challenges
- Summary

Consequences: Networking with Global Time

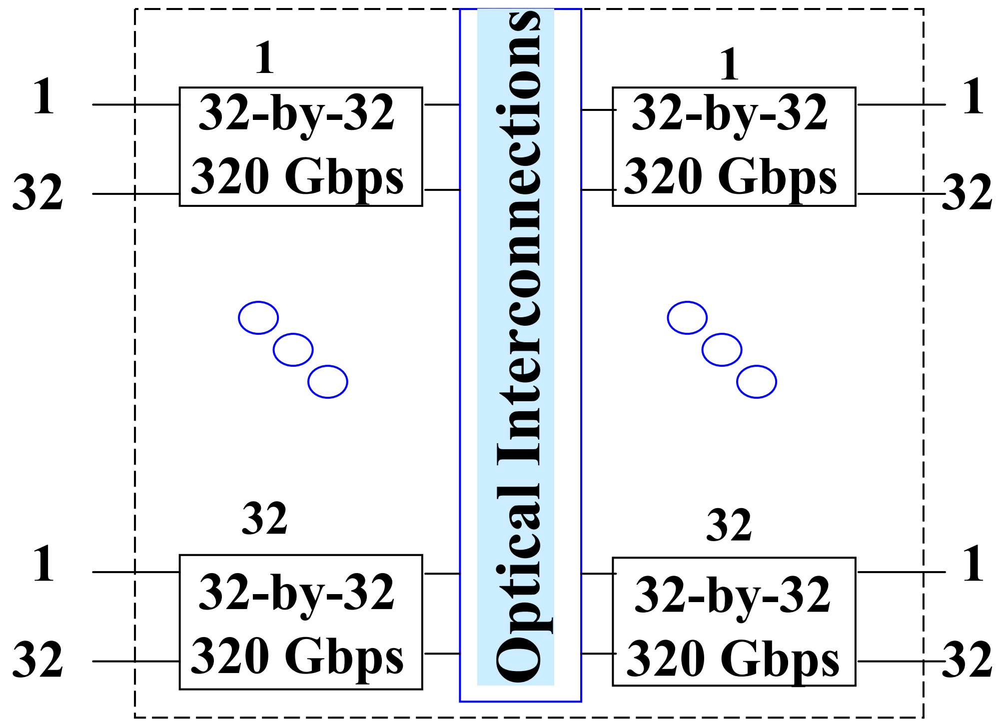
- Easy to implement

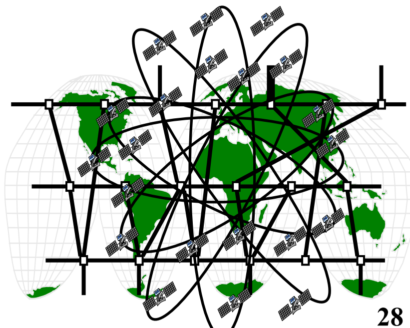
- Simple to analyze – time blocking



Current testbed: 10 Terabit/s per Chassis (Banyan-based)

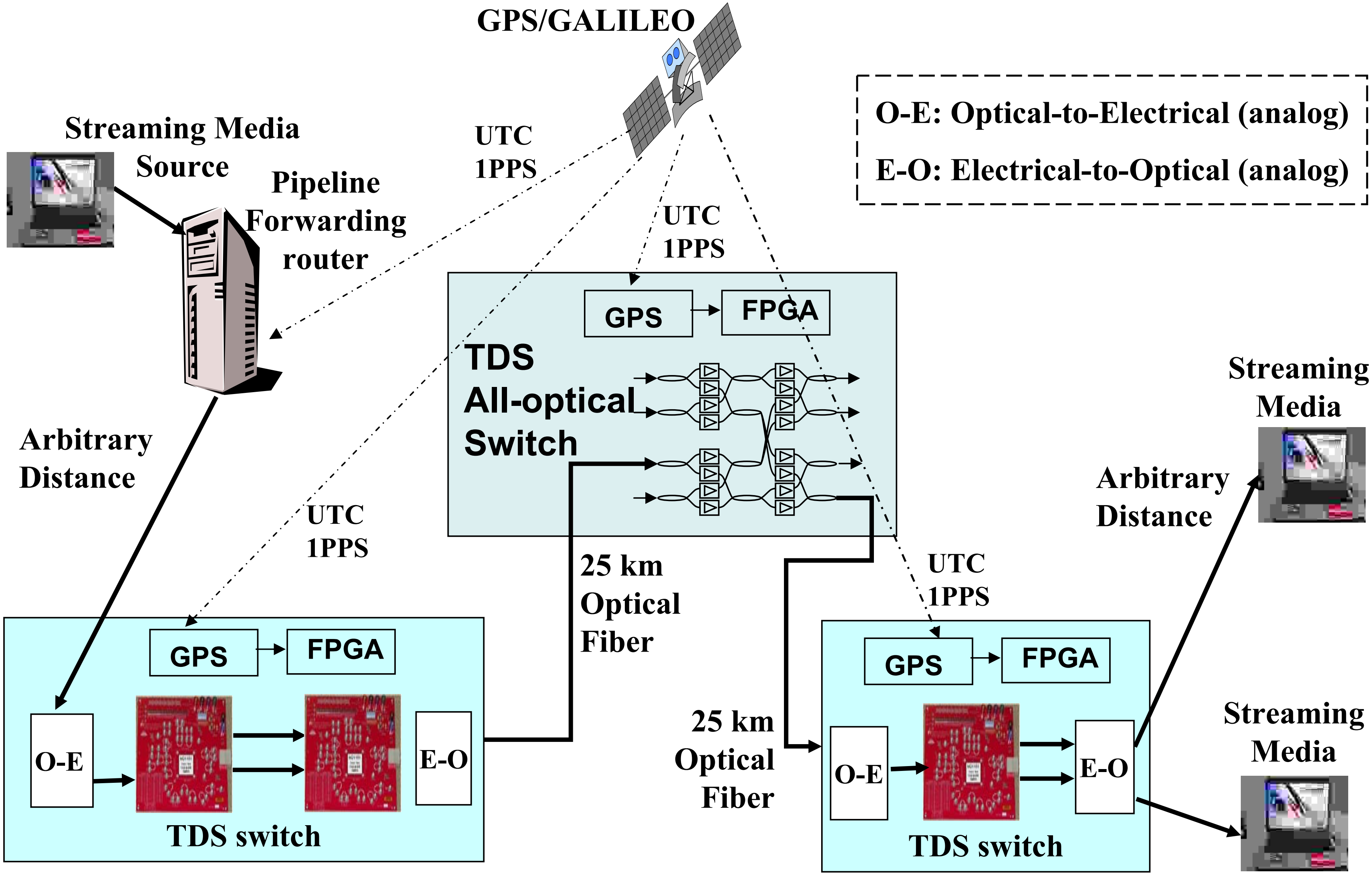
- **Simple implementation (9-month) with all off-the-shelf components**
  - Using existing (7 years old technology) Mindspeed M21151 cross-point switches





Current Networking Test-bed Setup

http://dit.unitn.it/ip-flow/

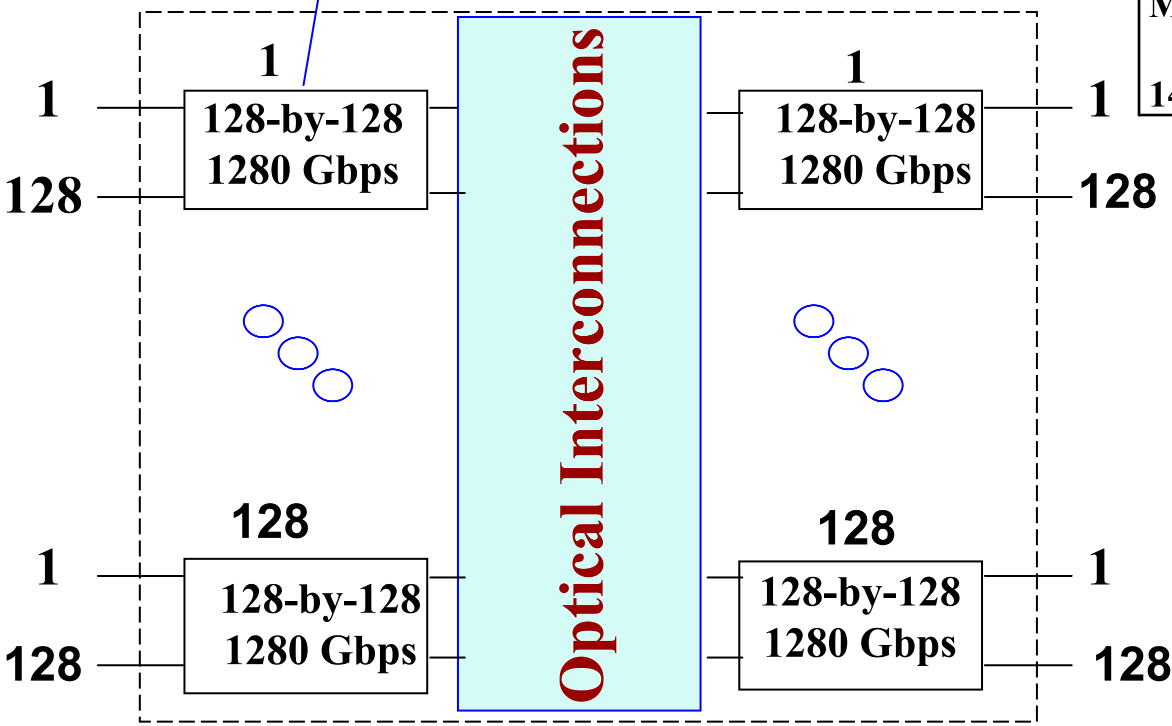


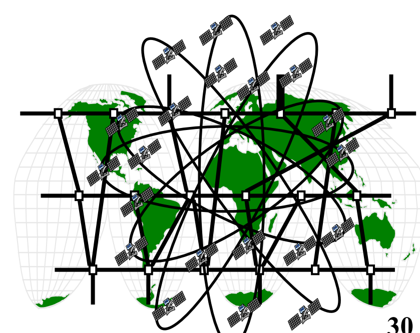
Example: 160 Terabit/s Fabric (Banyan-based)

[Lower Cost/Power → Green Internet]

Switching Element	Maximum Bit Rate	Power per Input-Output	Power per 1 Gb/s
Vitesse VSC3040 144 channel	11 Gb/s	0.11 W	0.01 W
Mindspeed M21151 144 channels	3.2 Gb/s	0.15 W	0.05 W

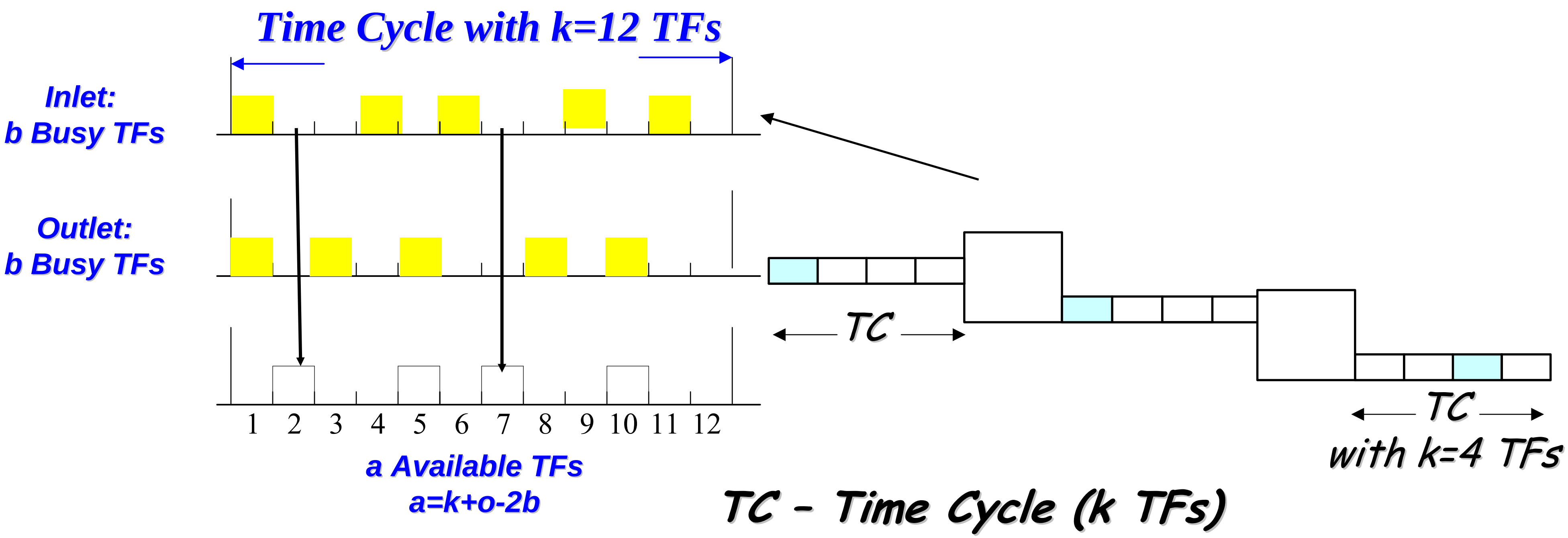
Crosspoint Switches:
Vitesse – VSC3040
[144-by-144: 11 Gb/s]



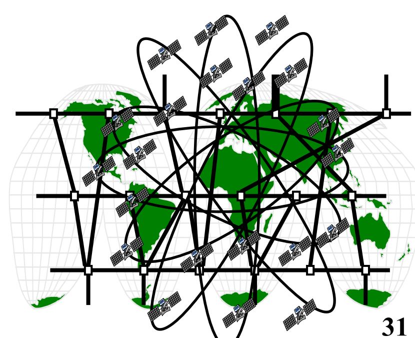


Time Blocking Problem

Finding available TF (time frame): inlet and outlet

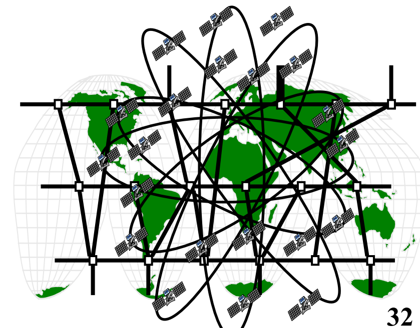


No time domain blocking up to 50% utilization

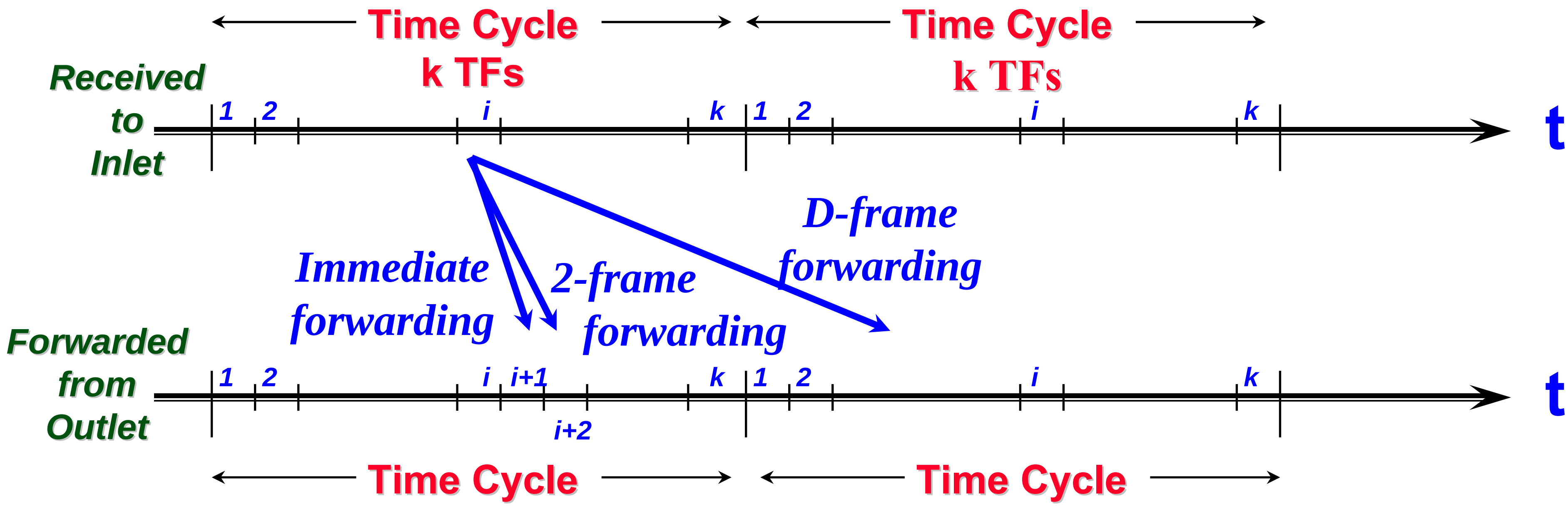


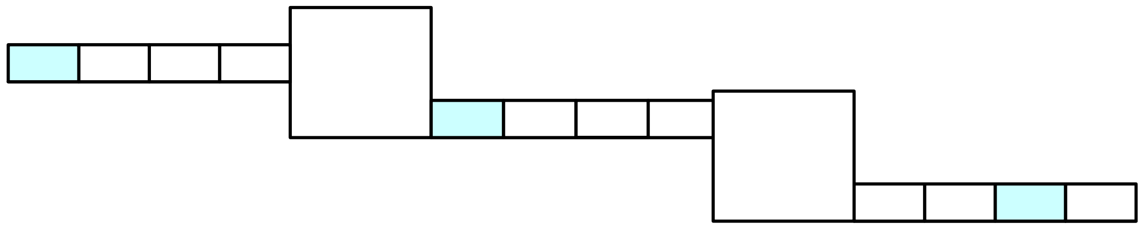
Time Domain Blocking Analysis for Immediate Forwarding (IF)

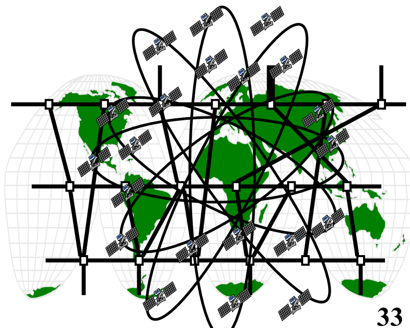
- **Probability of at least one available *TF***
- $P_{available} = C_{available\text{-}immediate} / C_{total} = \sum_{o=2b-k+1}^{o=b} \binom{b}{o}\binom{k-b}{b-o} \Big/ \binom{k}{b}$
- **And, thus, the blocking probability for a given *k*,*b* is:**
- $P_{blocking} = 1 - P_{available} = 1 - \sum_{o=2b-k+1}^{o=b} \binom{b}{o}\binom{k-b}{b-o} \Big/ \binom{k}{b}$



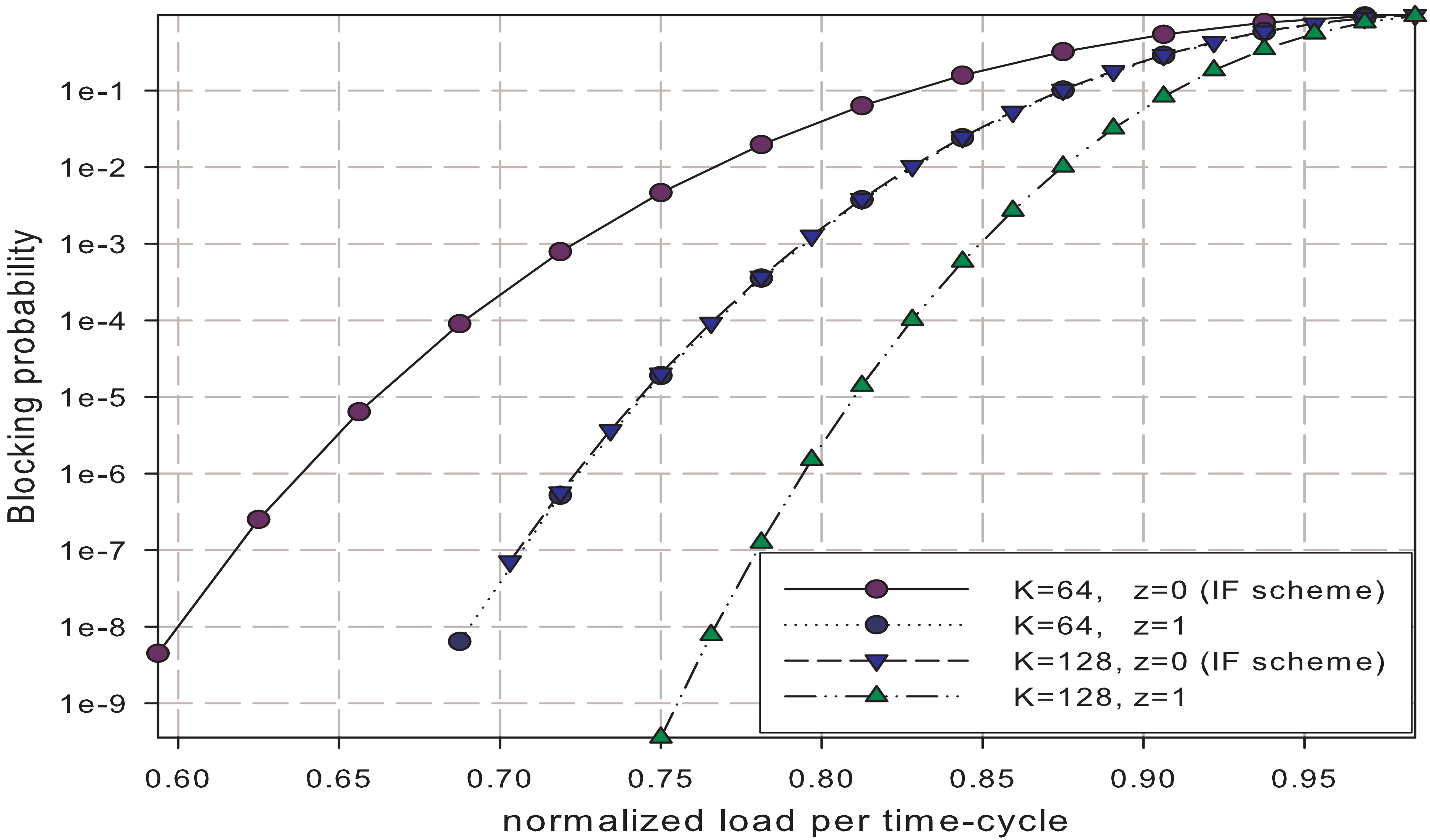
Reducing Blocking with Non-immediate Pipeline Forwarding – Additional Delay Buffers



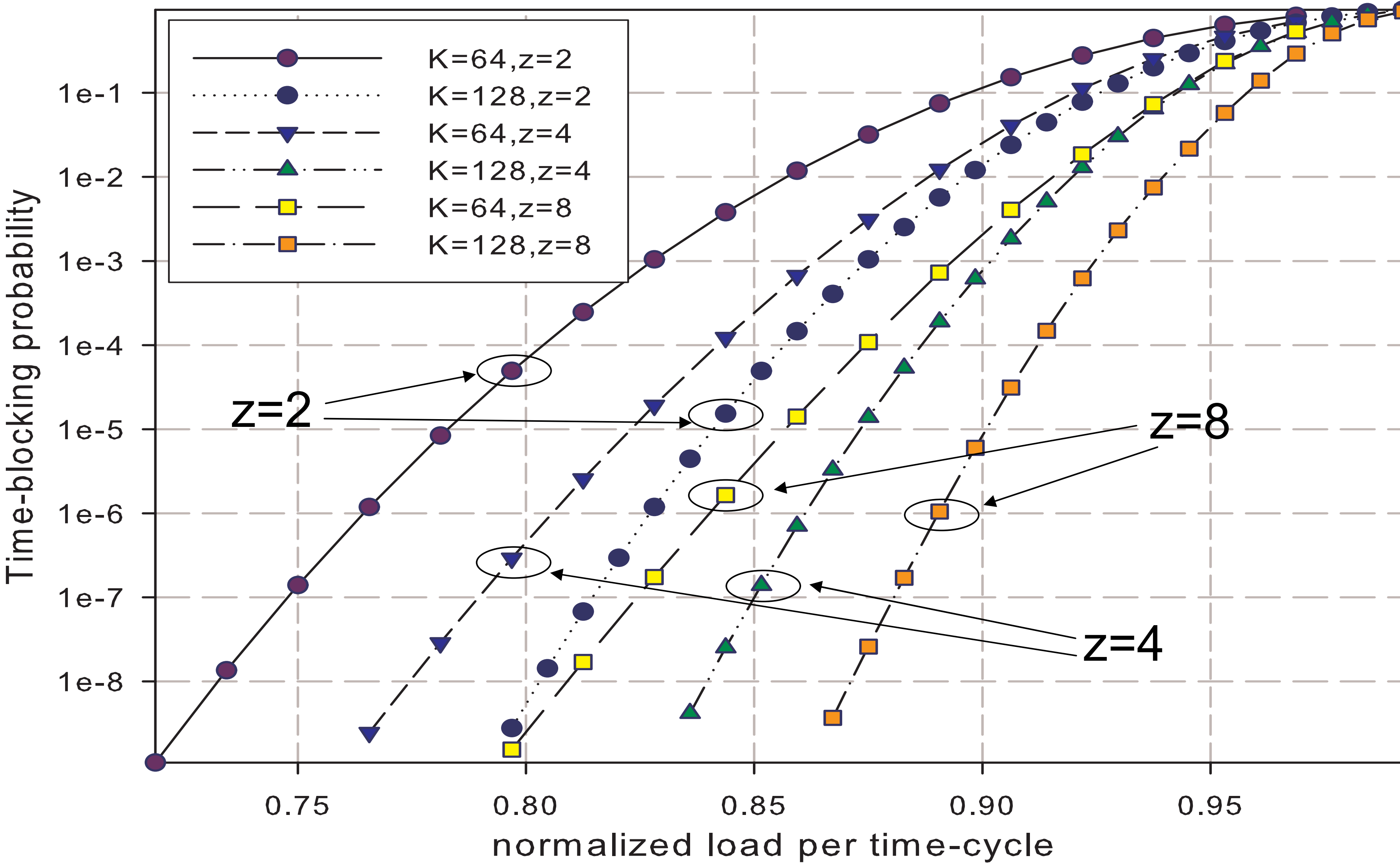




Results for z=0,1
Non-immediate with 1 buffer delay



Results for z>1
Non-immediate with 1 buffer delay



Closed-form Time Blocking Probability Analysis Using Combinatorial (Counting) Approach:

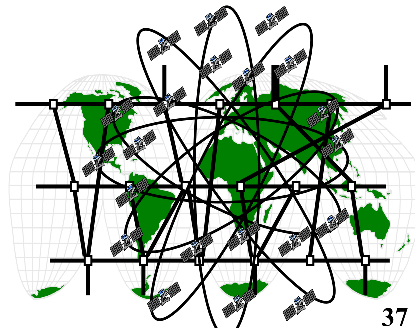
Throughput and delay are known deterministically

(1) Immediate forwarding (IF) analysis –
IEEE T. on Communications 2008
(2) Non-immediate forwarding (NIF) & multi-hop
IEEE INFOCOM 2008

(1) More TFs in each time cycle lower time blocking
(2) More optical channels lower time blocking

Outline

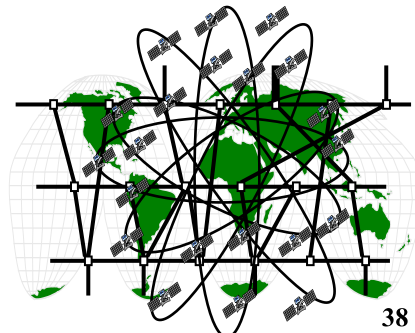
- Vision and Challenges
- Adding the global time (UTC) dimension
- Consequences of networking with time
- Deployment strategy –
  as the Internet underlay
- Solutions to the various challenges
- Summary



Deployment Solution Strategy

- **KEEP THE INTERNET AS IS!**
  - **It is operating beautifully for many applications**
  - **Scaling 50-100 folds much too expensive, and therefore, not realistic**

- **So how to proceed?**



The Underlay Deployment Solution

- **Future Internet:**
  - **Current Internet evolving slowly**
  - **Underlay – ultra scalable >90% of the traffic**
    - **Video-based and interactive (VoIP) applications**
    - **Over the same fiber infrastructure (with WDM)**

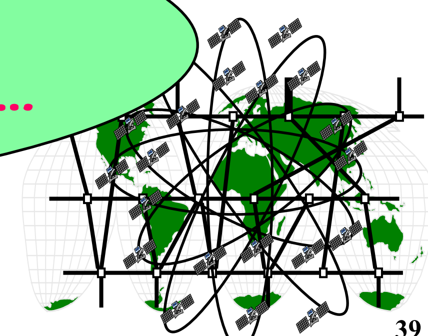
Internet As Is
Control/Manage/Monitor
the Underlay

10%

"Green Internet"

90%

Time-based Pipeline Forwarding
Underlay – "Super-highway" for:
- Streaming media (video) / IPTV / HD / 3D / VoIP …
- Large (content) file transfer / on-demand



Underlay over the Same Infrastructure

- Using WDM (wavelength division multiplexing)
- Existing protocols: DiffServ, GMPLS, …

DiffServ

[The only place where packet header is processed]

Animation

Internet As Is 10%

Control/Manage/Monitor the Underlay

UTC for "Total Event Ordering" in order to simplify network management

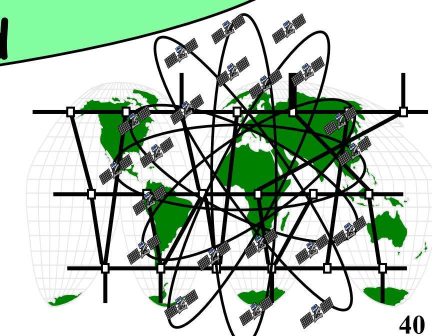
Pipeline Forwarding

Underlay – "Super-highway" for:

- Streaming media (video) / IPTV / HD / 3D / …
- Large (content) file transfer / on-demand

90%

"Green Internet"



The Internet Underlay
Simplified Protocol Structure

Internet *As Is*
Control/Manage/Monitor the Underlay

IP

10%

GMPLS

UTC for "Total Event Ordering" in order to simplify management, routing and GMPLS control plane

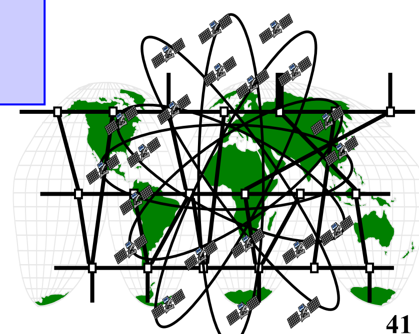
"Green Internet"

TDS - Time-driven Switching
In All-optical: FλS – Fractional Lambda Switching

90%

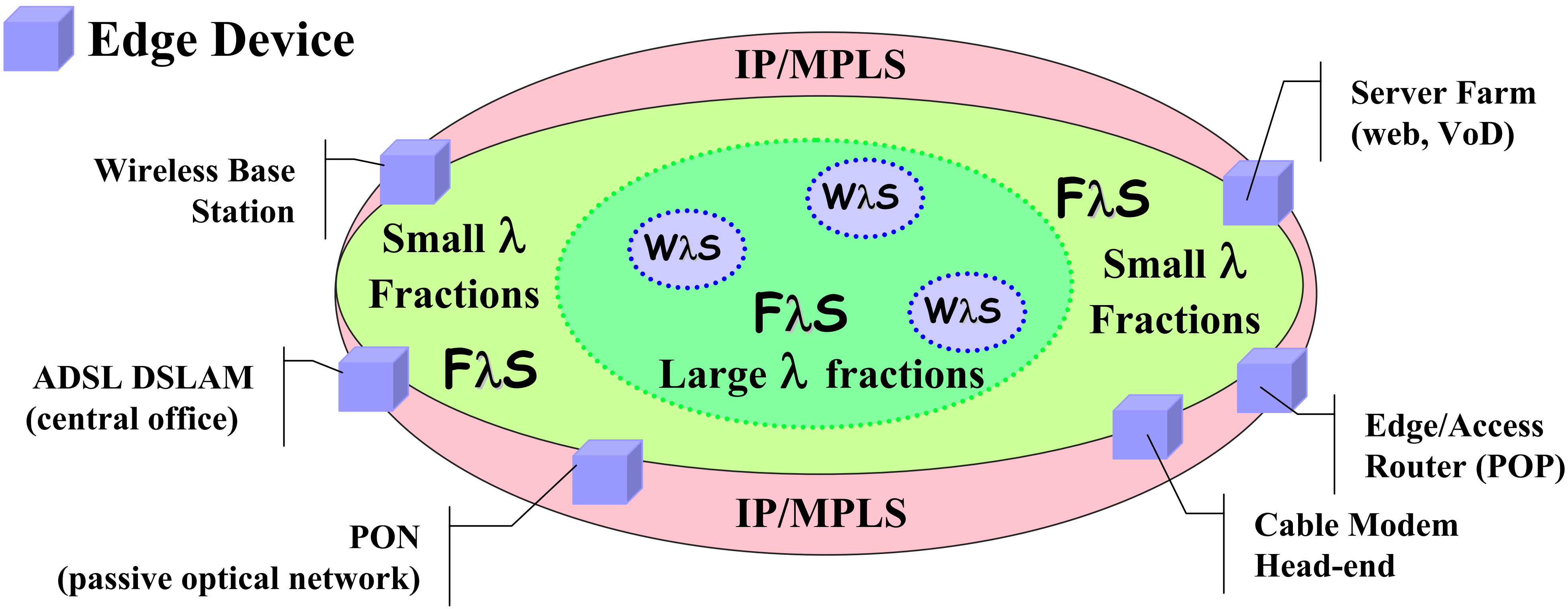
Fiber / WDM

Pipeline Forwarding Underlay – "Super-highway"



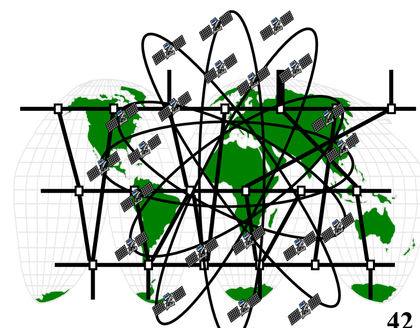
The Underlay Deployment Solution

Header Processing & Buffering Only at the Edges



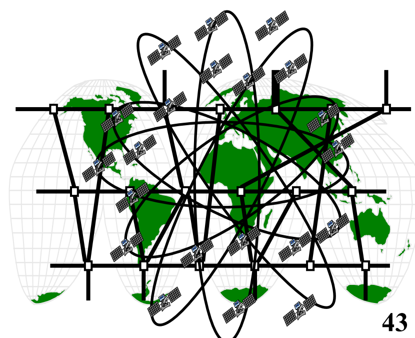
WλS - Whole λ Switching or λ Routing

FλS - Fractional λ Switching



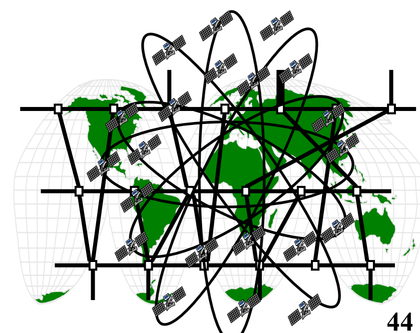
The Underlay Deployment Solution

- **First stage:**
  - World wide live streaming & download to distribution centers in metro area
  - Where every major metro area will have it's own "super-highway" underlay
- **Second stage:**
  - Integration with rural areas
  - Integration with wireless: 3G, 4G, …



Outline

- Vision and Challenges
- Adding the global time (UTC) dimension
- Consequences of networking with time
- Deployment strategy – as the Internet underlay
- Solutions to the various challenges
- Summary



Extra Revenue from Live Streaming

- Live events: sport (foot-ball, …), news, …
  - ON DEMAND and EVERYWHERE
- An example: assume that there are:
  - 1000 prime events per year in which
  - 10 million users **really would like to watch**
    - **but unable do it during the rigid TV broadcasting schedules and capacity (IP may provide any capacity)** and perhaps are:
  - **(1) willing to pay 10 Euro/event (or targeted advertising) – provided they receive it with (2) SLA guarantee/assurance**
- Consequently, the potential to provide the telecom business with **additional 100 Billion Euro in revenue** (1,000*10,000,000*10) through **new Internet services**
- Of course, all current Internet services are unchanged

Wireless Solutions with Time

- Major capacity increase for predictable traffic utilizing **TIME** to optimize:
  - Spectrum: OFDMA (orthogonal frequency division multiple access), …
    - Together with:
  - Space: SDMA (space division multiple access), smart antennas, …
- Optimize multi-hop wireless networks:
  - Mesh
  - Ad-hoc

Wireless Already Relies on Time (Phase Sync.)

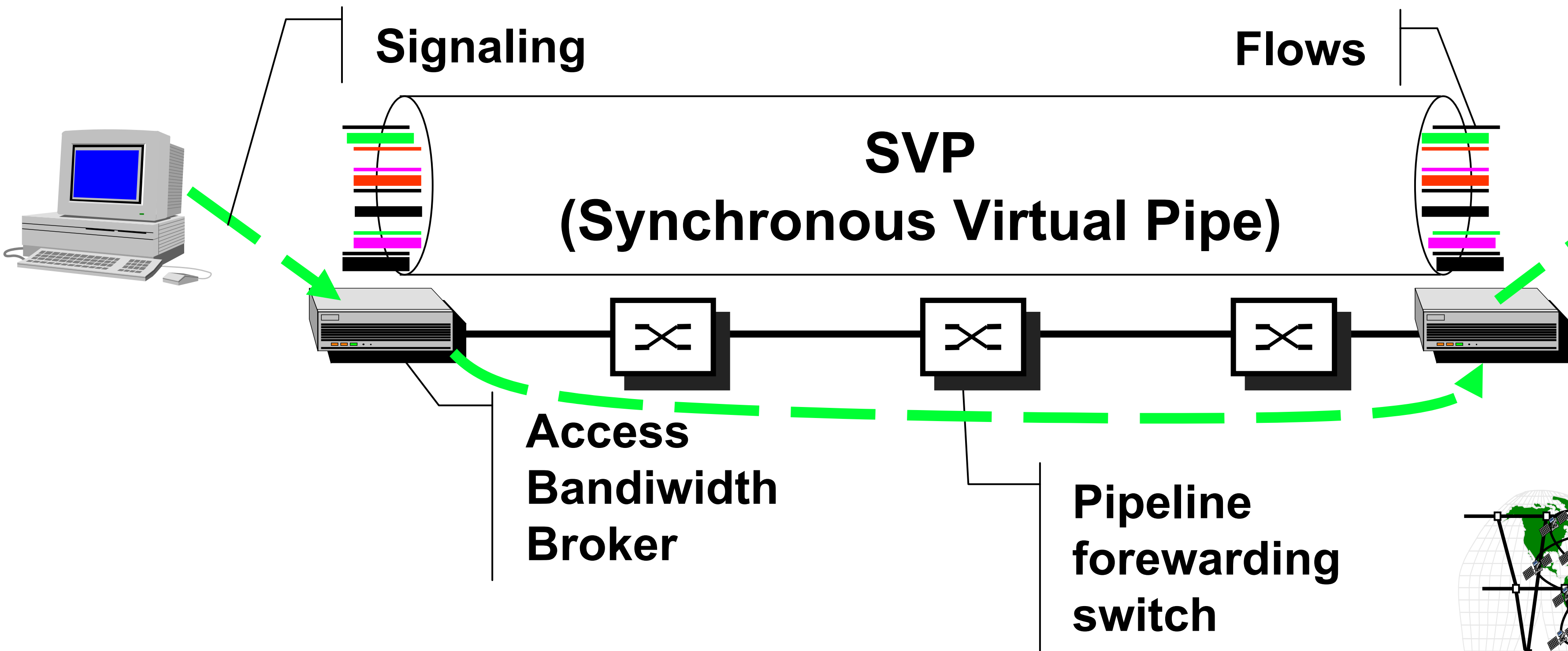
Synchronization service	Application	Expected quality
Global time (Time-of-Day) (Phase sync.)	**WiMAX**	**TDD – Time-division duplex - 1μs – MOBILE WL**
	Femto-Cell	**Solution for Fixed Mobile Convergence – home cellular base station via DSL - 1μs UTC to the home**
	3GPP MBMS/LTE	MBMS Content synchronization - TBD
	IP SLA One-way delay	Target precision in few orders of magnitude below average delay (i.e. ~ 10-100μs)
System specific time (Phase sync.)	802.16(D/)E	Depends on: mode, modulation, application, implementation Strongest needs for optimized radio frequency utilization, mobility and HO/Fast BS switching and MBS options
	3GPP2 CDMA Base Stations	Frequency assignment *shall* be less than ± 5×10^{-8} (like 3GPP) Time alignment error *should* be less than 3 μs
	3GPP MBMS/LTE	Cell synchronization accuracy better than to 3μs for SFN support Different options under study; one is to get precise time in-band

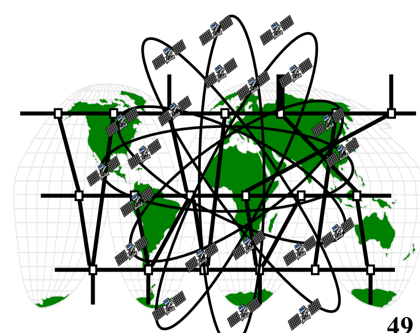
Protocol Simplifications Solutions over a Single IP Network

- Global time enables TEO - total event ordering for management, routing and control plane
- Reduction in the number of protocols as a result the "super-highway" underlay
  - Reduced need for:
    - ATM, SONET/SDH – all variants
    - VPN (virtual private network)
      - – simply use synchronous virtual pipe (SVP) see next
  - Simplified SLAs and security (e.g., firewall)
    - Using secure/authenticated global time stamps
  - Well-defined network interface

SVP (Synchronous Virtual Pipes) for Plurality of Flows over the "Super Highway" Underlay

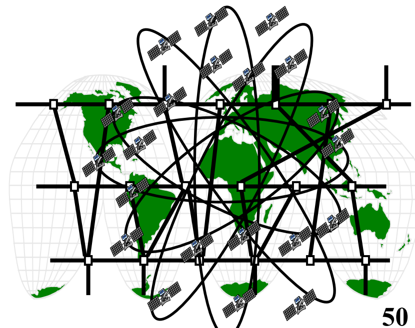
Like FEC (forwarding equivalent class) in MPLS
Setup with GMPLS (+ global time semantic)
[or Like VP (virtual pipe) in ATM]





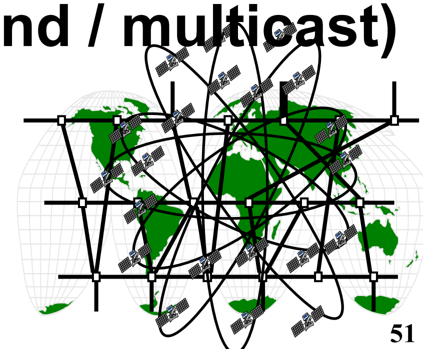
Outline

- Vision and Challenges
- Adding the global time (UTC) dimension
- Consequences of networking with time
- Deployment strategy – as the Internet underlay
- Solutions to the various challenges
- Summary



Summary

- Adding the time dimension (UTC) solves:
  - Revenue from premium events on-demand
    - Live streaming with deterministic performance guarantees
    - Interactive applications and services
  - Switching scalability and cost problem
    - Namely, much larger network for much lower cost
    - All-optical networking solution
  - Green Internet
  - Single integrated IP network
  - Reduction in protocol complexity while using existing protocols: GMPLS, DiffServ
  - Wireless/mobile with live streaming (on-demand / multicast)
- Underlay as a realistic deployment strategy for >100 larger Internet



Thank you!

***"Every great scientific truth goes through three stages.
First, people say it conflicts with the Bible.
Next, they say it had been discovered before.
Lastly, they say they always believed it."***

**Louis Agassiz, 1807-1873
(Swiss-born, Harvard professor,
discovered the Ice age)**