

Voice Is Not a Network

- Voice is an **application**
- Voice traffic engineering is used to plan for a certain level of service
- It is applicable to both circuit switched and packet-based networks



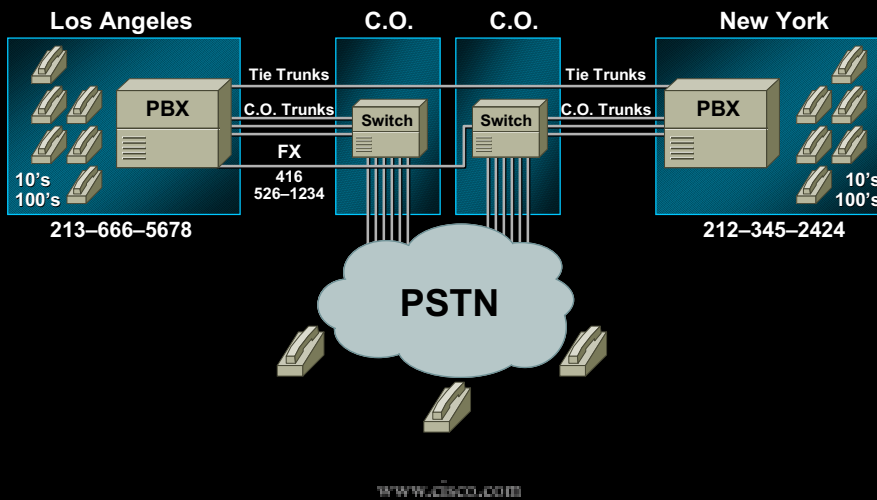
www.cisco.com

Agenda

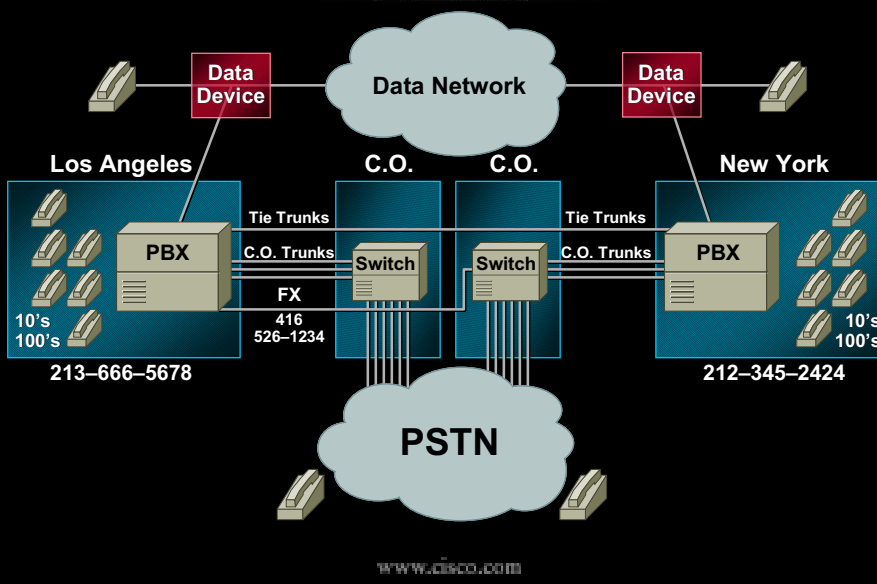
- Basic Voice Traffic Engineering
- How Costs Impact Trunk Groups
- How to apply Voice Traffic Engineering to a Packet-Based Data Network

www.cisco.com

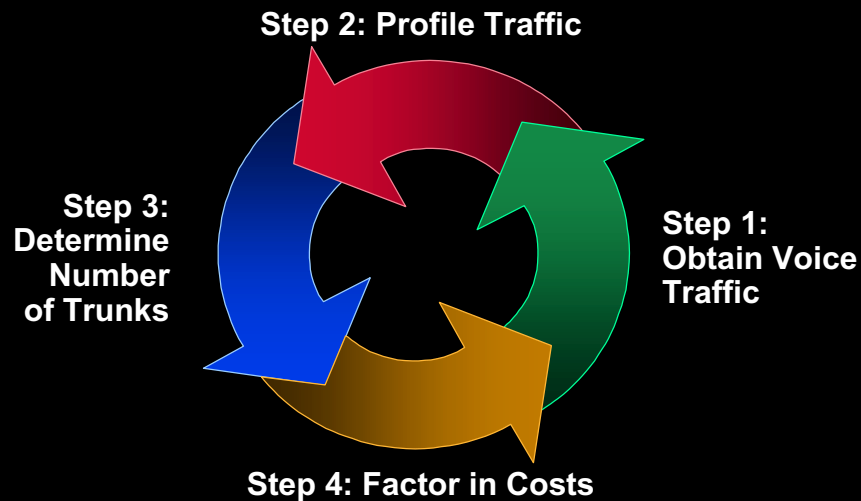
Today's Enterprise Network



Voice Networking in Transition



Traffic Engineering: 4-Step Process



Step 1: Obtain Voice Traffic

- Carrier bills
- Traffic reports from PBX
 - CDR (Call Detail Report)
 - Reports are specific to PBX manufacturer
- Carrier design studies or traffic reports
- Third party software and hardware

www.cisco.com

Defining the Traffic Flow

$$A = C \times T$$

A, The Traffic Flow Is the Product of:

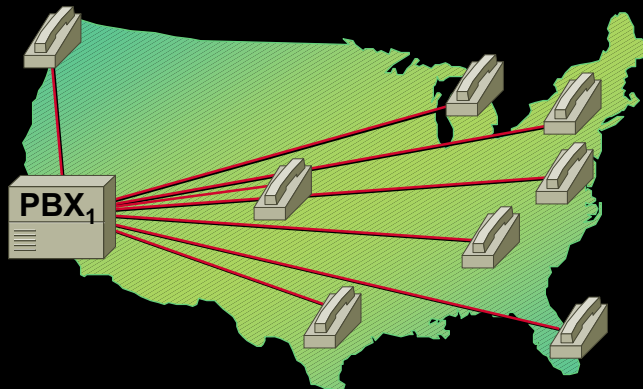
C: The Number of Calls Originated
During a Period of 1 Hour

T: The Average Holding Time of the Call

www.cisco.com

For Example

- 200 calls of an average duration of two minutes are generated during a period of one hour
- The traffic flow is equal to 400 call-minutes per hour



www.cisco.com

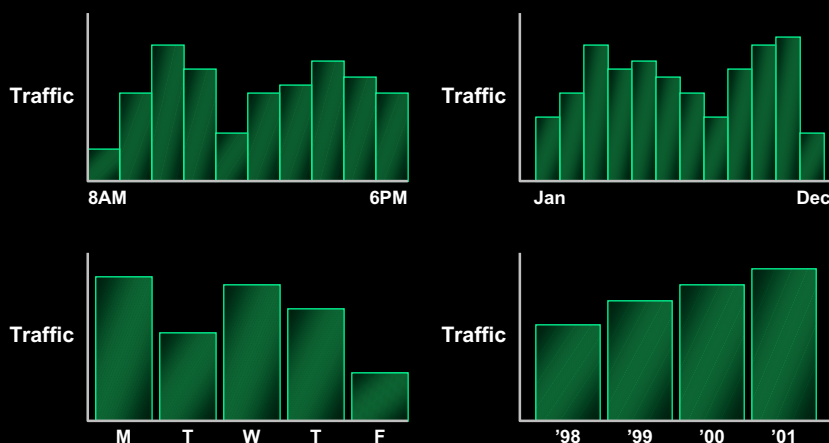
Converting to a Common Measurement

- Converting call-minutes to call-hours, divide by 60
- In our example: $400/60 = 6.67$ call-hours
- Typically we use erlangs, which is defined as the continuous use of a circuit for one hour
- Another common measurement is CCS (Centum call seconds or 100 call seconds)

1 erlang = 36 CCS

www.cisco.com

Traffic Variation



Traffic Varies by Hour, Day, Month and Year

www.cisco.com

Busy Hour

- **Busy hour =**
Total traffic in a month x% in busy day
x% in busy hour
- **BH is always used to determine the required number of trunks**

www.cisco.com

Factors Affecting the Holding Time

- **Holding time = total time trunk in use (for inband signaling)**
Dialing + call setup + ringing + conversation + release
- **Remember that telephone bills and PBX cdrs only have carried calls, i.e. no incomplete calls**
- **Other sources of trunk use: ring-no-answer, busy signal, etc.**
- **Call processing may or may not be included**

www.cisco.com

Adjusting the Holding Time

- Add 10% to 16% to length of all calls—to account for overhead
- Call billing: 6 second vs. 1 minute increments
- PBXs that use 1 minute increments round upwards; calls on average will have 30 seconds of extra hold time per call

www.cisco.com

For Example

A Bill Shows 404 Calls Totaled 1834 Minutes
Billing Is in One Minute Increments
What Is the Adjusted Traffic?

$$404 \times (0.5) = 202$$

$$1834 - 202 = 1632 \text{ "Real Traffic"}$$

www.cisco.com

Now We Have Numbers

What Do They Mean?

“

**There Are Lies,
Damn Lies
and Statistics**

”

Mark Twain

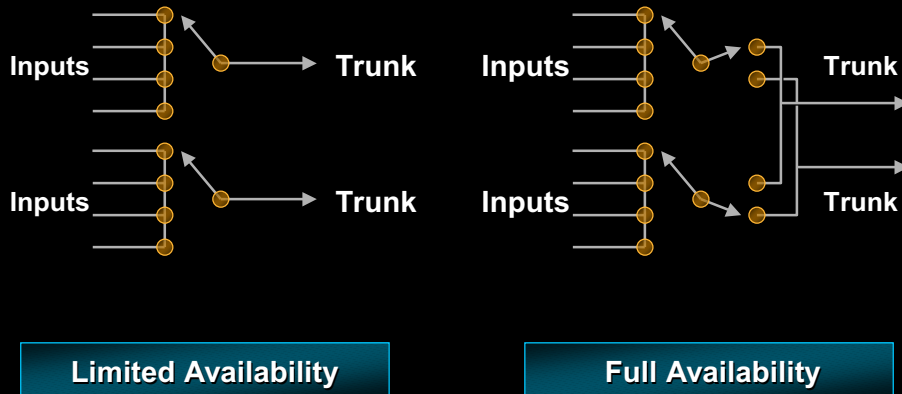
www.cisco.com

Traffic Probability

- **How the switch handles trunk allocation**
- **Traffic source characteristics**
- **How lost calls are handled**

www.cisco.com

Switch Trunk Availability



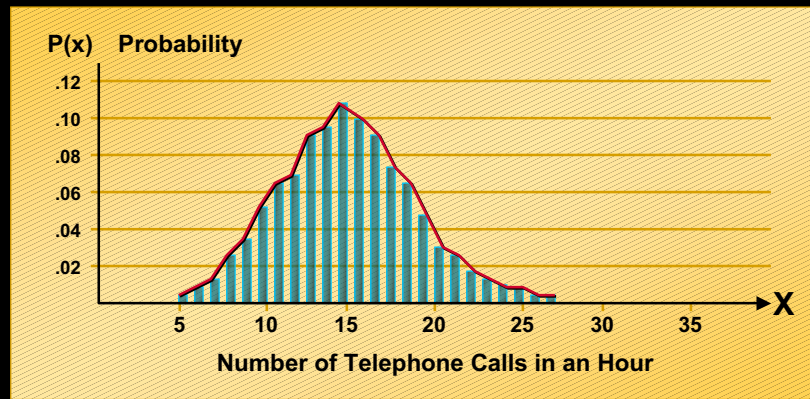
www.cisco.com

Traffic Probability

- How the switch handles trunk allocation
- **Traffic source characteristics**
- How lost calls are handled

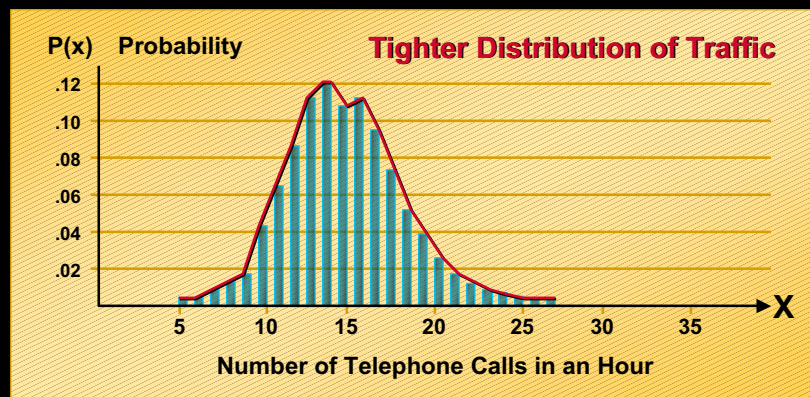
www.cisco.com

Random Traffic



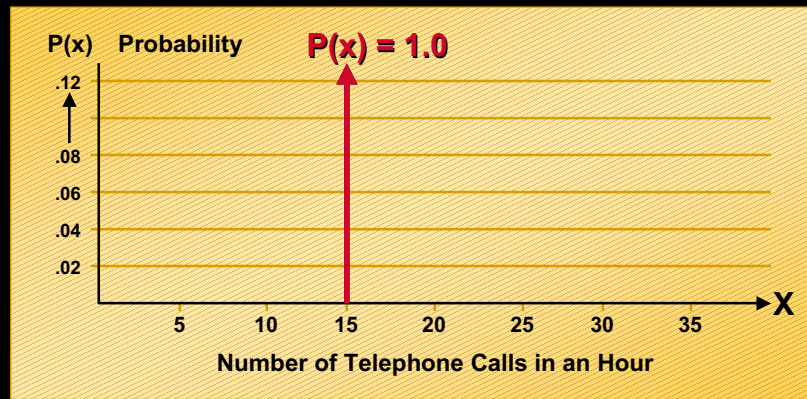
www.cisco.com

Smooth Traffic



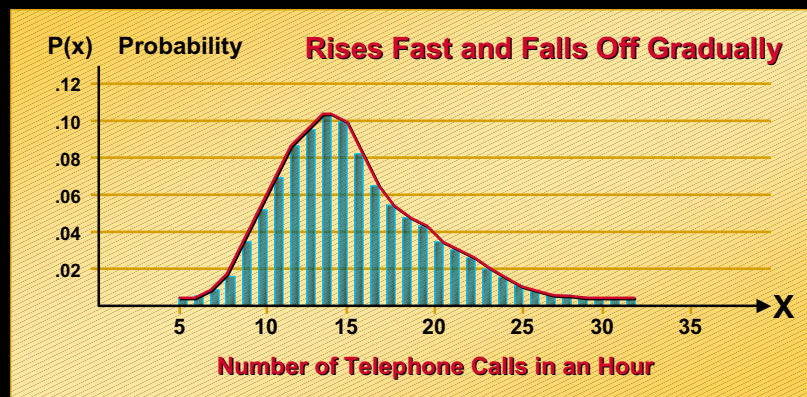
www.cisco.com

Exaggerated Case of Smooth Traffic



www.cisco.com

Rough or Peaked Traffic



www.cisco.com

Number of Originating Sources

10 Trunks and a Probability of .01 Poisson Distribution

- Infinite—probability of call arrival is constant
- Finite—probability of call arrival varies with the number of sources already connected

Number of Sources	Traffic Capacity (Erlangs)
Infinite	4.13
100	4.26
75	4.35
50	4.51
25	4.84
20	5.08
15	5.64
13	6.03
11	6.95
10	10

www.cisco.com

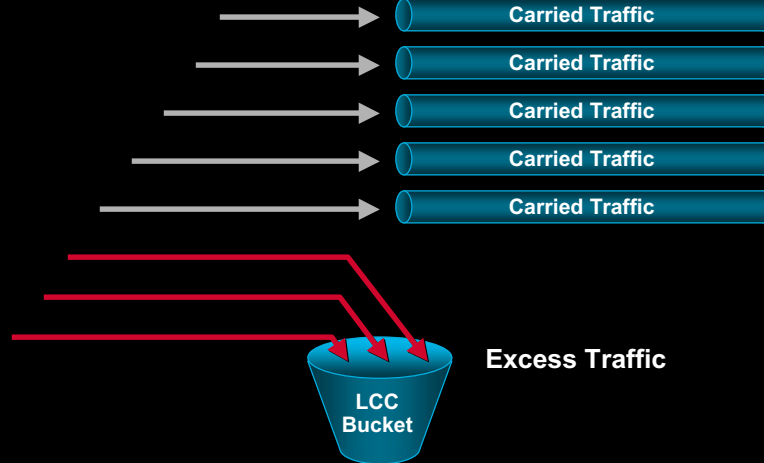
Traffic Probability

- How the switch handles trunk allocation
- Traffic source characteristics
- **How lost calls are handled**

www.cisco.com

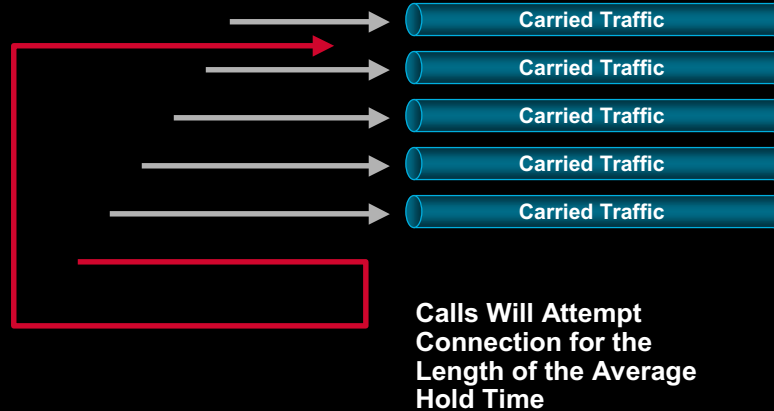
Handling Lost Calls: Lost Calls Cleared (LCC)

First Attempt Offered Traffic

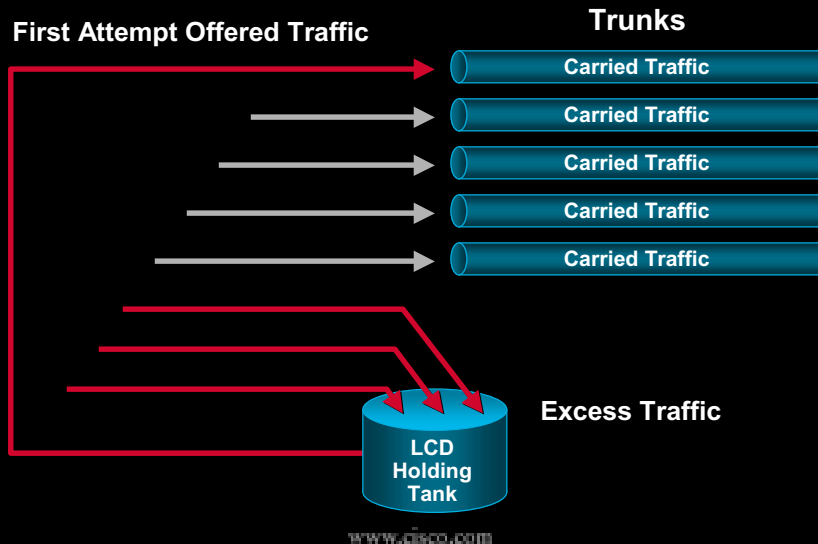


Handling Lost Calls: Lost Calls Held (LCH)

First Attempt Offered Traffic



Handling Lost Calls: Lost Calls Delayed (LCD)



Step 3: Determining the Number of Trunks

- Decide how many trunk groups we want based on how we profiled the traffic
- Determine grade of service for each group
- Apply probability tables or programs to calculate the number of trunks

www.cisco.com

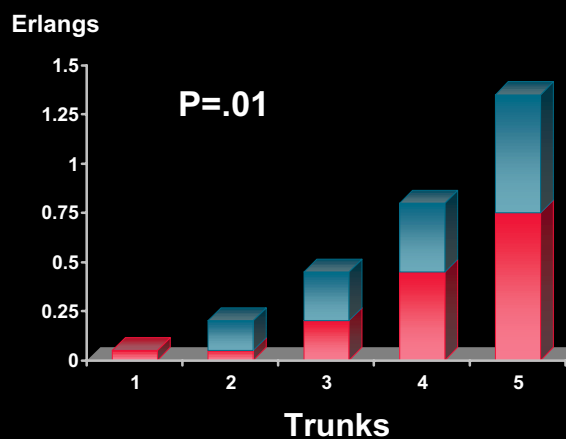
Large or Small Trunk Groups?

- Larger trunk groups are more efficient
- Volume discount benefits
- Iterative design approach

www.cisco.com

Theory of Large Trunk Groups

- Total capacity increases in a non-linear and not a linear fashion
- The incremental addition of one trunk approaches one erlang



www.cisco.com

Grade of Service

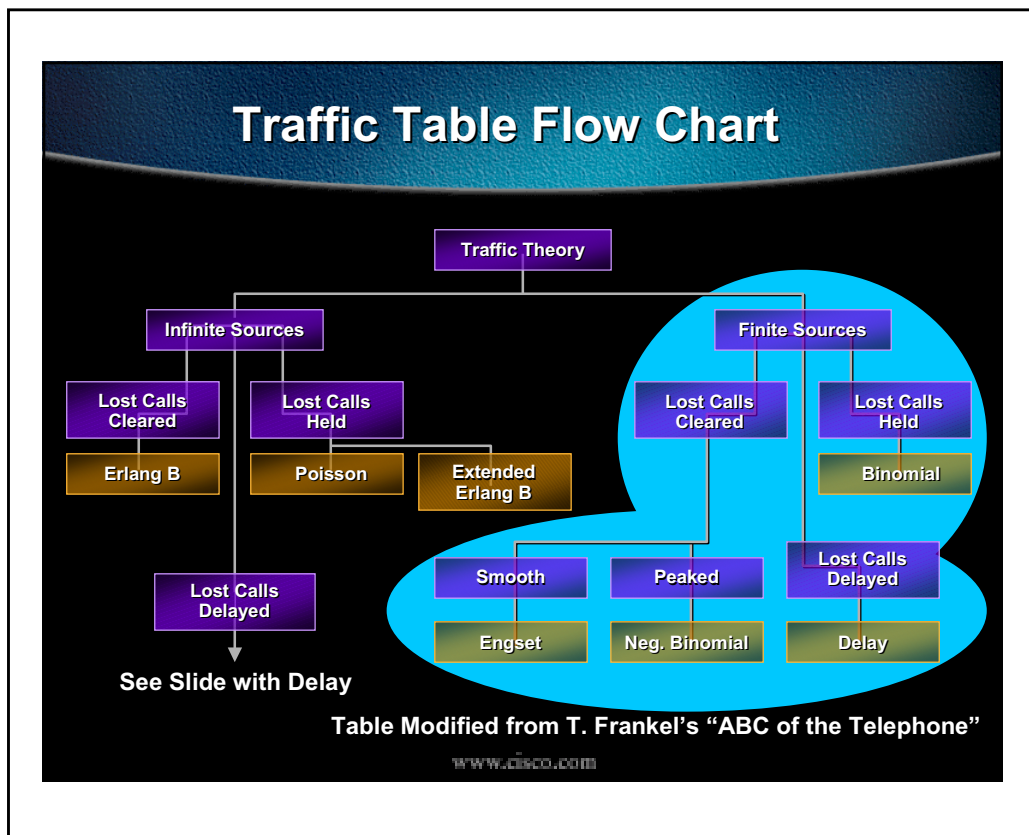
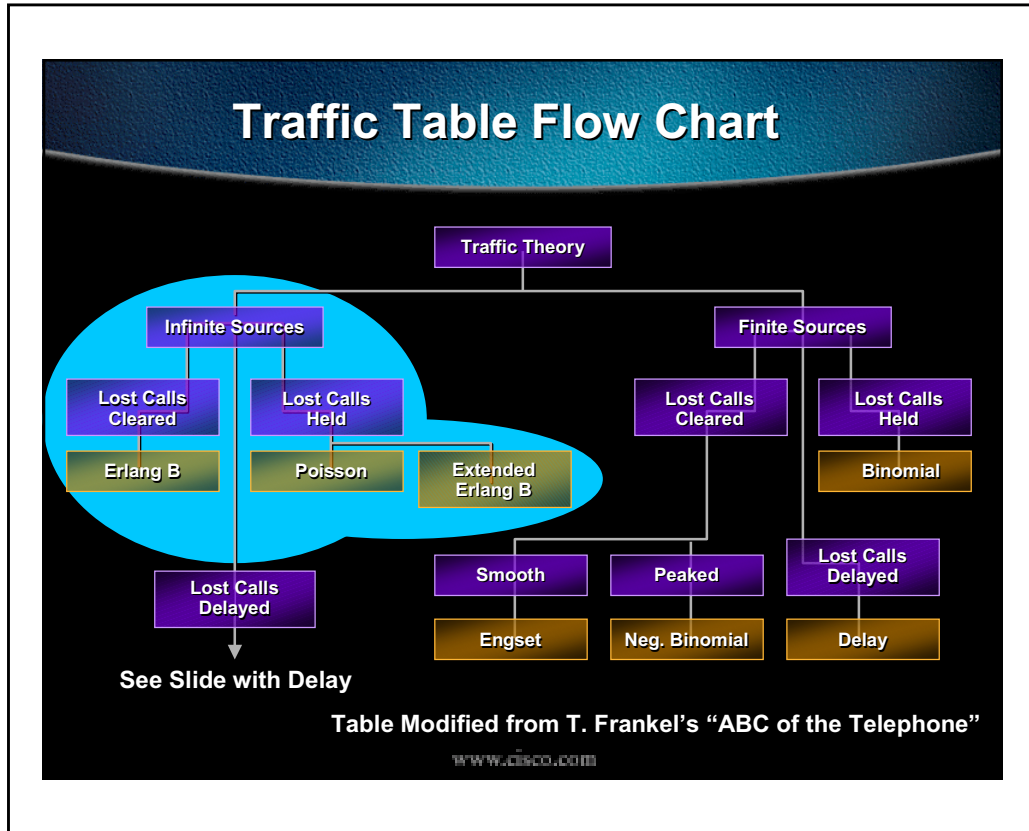
- **Grade of service based on the probability of blockage**
- **Based on the busy hour**
- **Grade of service differ between:
PBXs
Automatic call distribution with queuing**

www.cisco.com

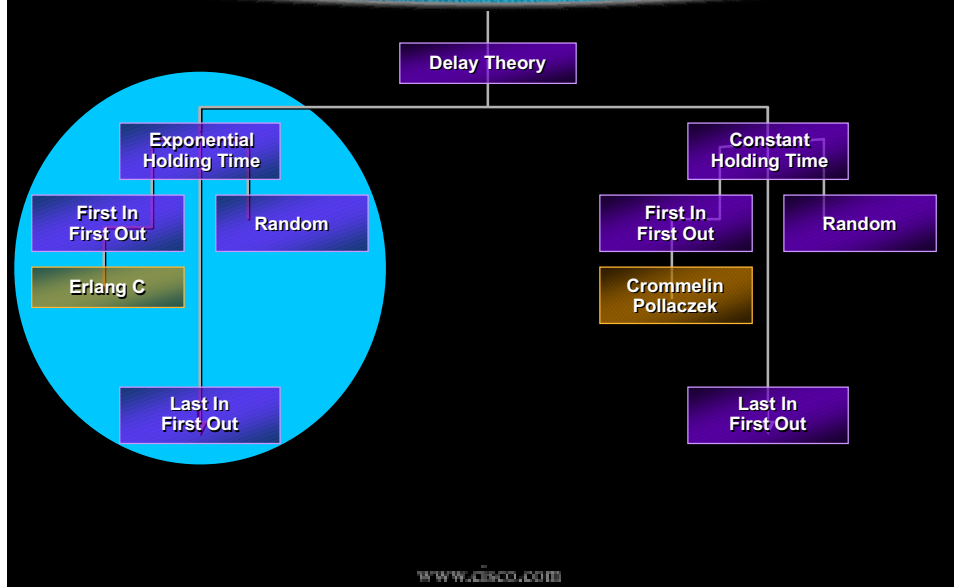
Pick the Traffic Table or Program

- **Number of sources: Finite and infinite**
- **Traffic characteristics: Random, smooth or rough**
- **How blocked calls are handled**
- **Switch availability**

www.cisco.com



Traffic Table Flow Chart (Cont.)



Erlang B

- **Infinite sources**
- **Lost calls cleared**
- **Constant or exponential holding time**
- **Random traffic**
- **Application**
 - Outbound trunks with overflow, i.e. alternate routes are used
- **Commonly used throughout the world**

www.cisco.com

Design Example #1

How Many Trunks to Handle this Traffic?

- 352 hours of first-attempt traffic in the month
- 22 business days/month
- 10% call processing
- 15% of traffic in busy hour
- Chance of blocking = 1%

www.cisco.com

Design Example #1

First, What Is the Busy Hour?

www.cisco.com

Design Example #1

First, What Is the Busy Hour?

$$\text{BH} = 352 \text{ hr} \div 22 \text{ d/m} \times 0.15\% \text{ Busy Hour} \times 1.10 \text{ Call Processing} \\ = 2.64 \text{ E}$$

www.cisco.com

Design Example #1

Second, What Table Would You Use?

- Infinite sources
- Overflow excess traffic
- No queuing

www.cisco.com

Design Example #1

Second, What Table Would You Use?

Erlang B

www.cisco.com

Design Example #1

Finally, Calculate the Number of Trunks Required?

N	.003	.005	P	.02	.03	.05
1	.003	.005	.011	.021	.031	.053
2	.081	.106	.153	.224	.282	.382
3	.289	.349	.456	.603	.716	.9
4	.602	.702	.87	1.093	1.259	1.525
5	.995	1.132	1.361	1.658	1.876	2.219
6	1.447	1.622	1.909	2.276	2.543	2.961
7	1.947	2.158	2.501	2.936	3.25	3.738
8	2.484	2.73	3.128	3.627	3.987	4.543
9	3.053	3.333	3.783	4.345	4.748	5.371
10	3.648	3.961	4.462	5.084	5.53	6.216
11	4.267	4.611	5.16	5.842	6.328	7.077
12	4.904	5.279	5.876	6.615	7.141	7.95
13	5.559	5.964	6.608	7.402	7.967	8.835
14	6.229	6.664	7.352	8.201	8.804	9.73
15	6.913	7.376	8.108	9.01	9.65	10.63

Table Extracted from T. Frankel's "ABC of the Telephone"

www.cisco.com

Design Example #1

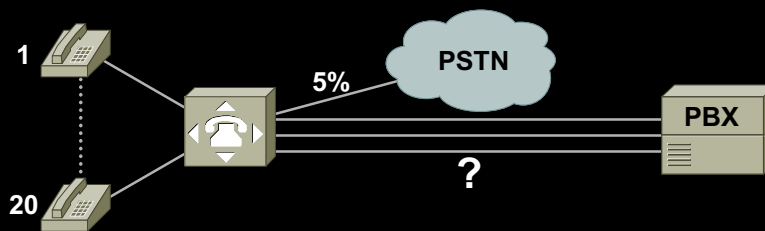
Finally, Calculate the Number of Trunks Required?

N	P					
	.003	.005	.01	.02	.03	.05
1	.003	.005	.011	.021	.031	.053
2	.081	.106	.153	.224	.282	.382
3	.289	.349	.456	.603	.716	.9
4	.602	.702	.87	1.093	1.259	1.525
5	.995	1.132	1.361	1.658	1.876	2.219
6	1.447	1.622	1.909	2.276	2.543	2.961
7	1.947	2.158	2.501	2.936	3.25	3.738
8	2.484	2.73	3.128	3.627	3.987	4.543
9	3.053	3.333	3.783	4.345	4.748	5.371
10	3.648	3.961	4.462	5.084	5.53	6.216
11	4.267	4.611	5.16	5.842	6.328	7.077
12	4.904	5.279	5.876	6.615	7.141	7.95
13	5.559	5.964	6.608	7.402	7.967	8.835
14	6.229	6.664	7.352	8.201	8.804	9.73
15	6.913	7.376	8.108	9.01	9.65	10.63

Table Extracted from T. Frankel's "ABC of the Telephone"

www.cisco.com

Design Example #2



How Many Tie Lines Are Required to Support a Remote Office Key System if the following Conditions Exist:

- There are 20 subscribers
- Average five calls per hour per user
- Average duration of 100 seconds per call
- 5% of the traffic will overflow to PSTN

www.cisco.com

Design Example #2

- $A = C * T$
- $A = \frac{20 * 5 * 100}{3600} = 2.78$ Erlang
- $L = 20$ users
- $P = .05$

A = Traffic Flow

C = Number of Calls Originated in 1 Hour

T = Average Holding Timer of One Call

L = Number of Callers

P = Probability of a Call being Blocked

www.cisco.com

Design Example #2

Using Engset

N	P							
	0.001	0.002	0.003	0.005	0.01	0.02	0.03	0.05
	L = 20							
5	0.88	1.03	1.13	1.28	1.53	1.84	2.07	2.43
6	1.34	1.54	1.68	1.87	2.17	2.56	2.83	3.26
7	1.88	2.13	2.29	2.52	2.88	3.33	3.65	4.14
8	2.49	2.77	2.96	3.23	3.65	4.16	4.51	5.06
9	3.15	3.48	3.7	3.99	4.46	5.02	5.42	6.02

Six Trunks Satisfy the Requirements

Table Extracted from T. Frankel's "ABC of the Telephone"

www.cisco.com

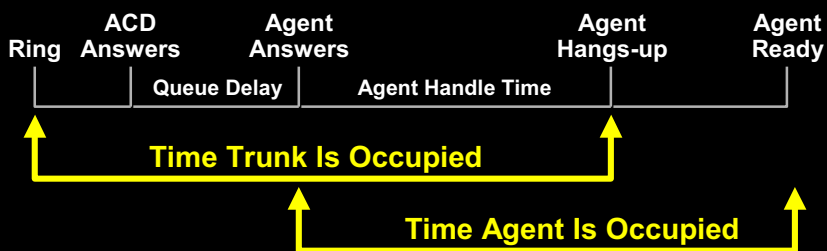
Design Example #3: Mixed-Model

- **Call center design**
 - Uses 1-800 number to sales, service, etc.
 - Some call centers provide function to several companies
 - Every minute can equate to \$10,000s
- **Real-time issues**
 - Peaked traffic
 - Agent breaks, sickness, etc.

www.cisco.com

Call Center Design

- Call centers use a device called an Automatic Call Distributor (ACD) to queue and distribute calls to agents
 - May be part of a PBX or a separate device
- Holding time of a trunk



www.cisco.com

Call Center Traffic Models

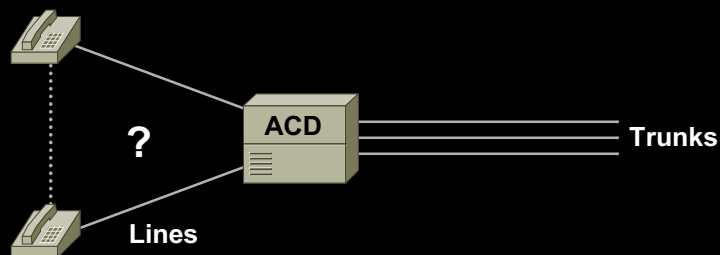
- Erlang C is used to design number of agent positions

Queuing

No overflow

www.cisco.com

Design Example #3



How Many Lines Are Required to Support a Call Center that has:

- 230 calls per hour with an average delay not longer than 30 sec
- No more than 20 percent of the calls delayed over one minute
- Average hold time is 120 seconds
- Queuing is first in first out

www.cisco.com

Design Example #3

- $A = C * T$
 - $A = \frac{230 * 120}{3600} = 7.67$ Erlangs
 - $D1 = \frac{30 \text{ sec}}{120 \text{ sec}} = .25$ holding time
 - $t = \frac{60 \text{ sec}}{120 \text{ sec}} = .5$ holding time
 - $P(t)$ (for $t = .5$) must be less than 20%
- A** = Traffic Flow
C = Number of Calls Originated in 1 Hour
T = Average Holding Timer of One Call
D1 = Average Delay on all Calls as a Percentage of the Holding Time
t = Time of no Greater than Delay as a Percentage of the Holding Time
P(t) = Probability of Delay Greater than t

www.cisco.com

Design Example #3

Using Erlang C

A	N	P(0)	D1	D2	P(t) for t=						
					0.1	0.2	0.3	0.4	0.5	0.75	1
7.7	8	0.889	2.94	3.33	0.856	0.831	0.806	0.782	0.759	0.704	0.653
	9	0.564	0.434	0.769	0.496	0.435	0.382	0.336	0.295	0.213	0.154
	10	0.346	0.15	0.435	0.274	0.218	0.173	0.138	0.109	0.062	0.035
	11	0.202	0.061	0.303	0.145	0.104	0.075	0.054	0.039	0.017	0.007

1. Found the Entry for Erlangs = 7.7
2. Found the Entry for $d1 < .25$
3. Looking at $t = 0.5$ Find $p(t) < 0.20$
4. 10 Agent Lines Will Meet My Requirements

Table Extracted from T. Frankel's "ABC of the Telephone"

www.cisco.com

Agenda

- **Basic Voice Traffic Engineering**
- **How Costs Impact Trunk Groups**
- **How to Apply Voice Traffic Engineering to a Packet-Based Data Network**

www.cisco.com

Step 4: Least-Cost Trunking

- **Calculate the number of trunks required to carry the busy hour traffic**
- **Account for traffic overflowing to alternate paths, i.e. lost calls cleared**
- **Primary and alternate routes (overflow) are dictated by economics, e.g. carrier pricing**
- **Packet based networks can serve as either primary or alternate routes, given we understand the cost associated with their use**

www.cisco.com

Cost of Service

- **Cost per minute often-quoted parameter**
 - Unique to each organization
 - Varies with each organization
- **In reality, costs should include transmission, equipment, administration and maintenance**

www.cisco.com

Cost—Optimization Rule

- **Use average usage figures instead of busy hour calculations**
- **Use the least costly circuits until the cost per hour becomes more expensive than the next best route**

www.cisco.com

Back to Our First Example

- 2.64 Erlangs in the busy hour requires eight trunks for a p of .01
- 352 hours of first-attempt traffic in the month
- 22 business days per month
- 10% call processing
- Average hourly usage is equal to:
352 hours/month—22 days/month—8 hours/day *
1.10 processing = 2.2 Erlangs per hour

www.cisco.com

Example

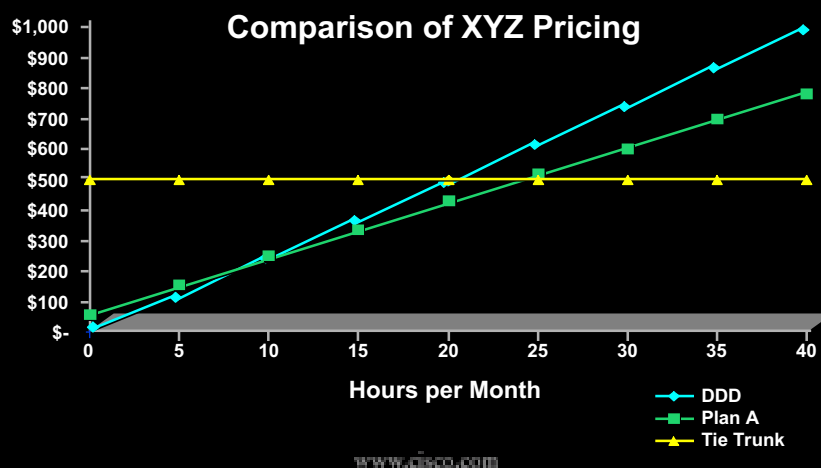
- XYZ carrier has offered our company three choices for carrying voice traffic between its head office, and its sales office
 1. DDD \$25 per hour
 2. Plan A \$60 fixed charge plus \$18 per hour
 3. Tie trunk \$500

Which Option Is Best?

www.cisco.com

Graph of Cost vs. Usage

...Depends on Traffic Volume



Economic Cross over Points

- 8.57 hours between DDD and plan A
- 24.4 hours between plan A and tie trunks

www.cisco.com

Traffic Distribution Based on Average Usage

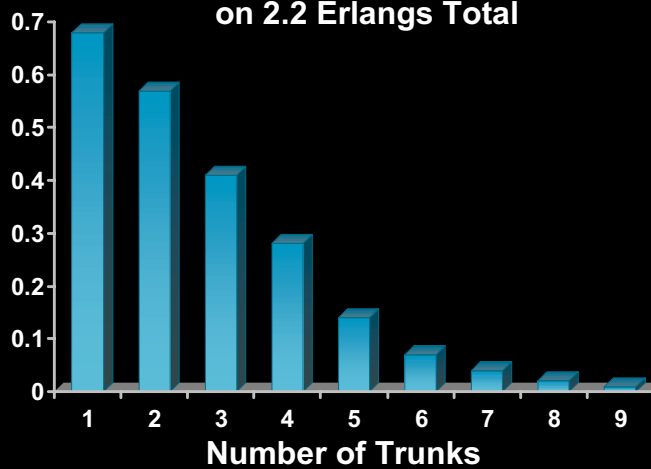
Number Trunks	Offered Hours	Carried per Trunk	Cumulative Carried	Grade of Service
1	2.2	0.688	0.688	0.688
2	1.513	0.565	1.253	0.431
3	0.947	0.419	1.672	0.24
4	0.528	0.271	1.943	0.117
5	0.257	0.149	2.093	0.049
6	0.107	0.069	2.161	0.018
7	0.039	0.027	2.188	0.005
8	0.012	0.009	2.197	0.002
9	0.003	0.003	2.199	0

Traffic on Each Trunk Based on 2.2 Erlangs

www.cisco.com

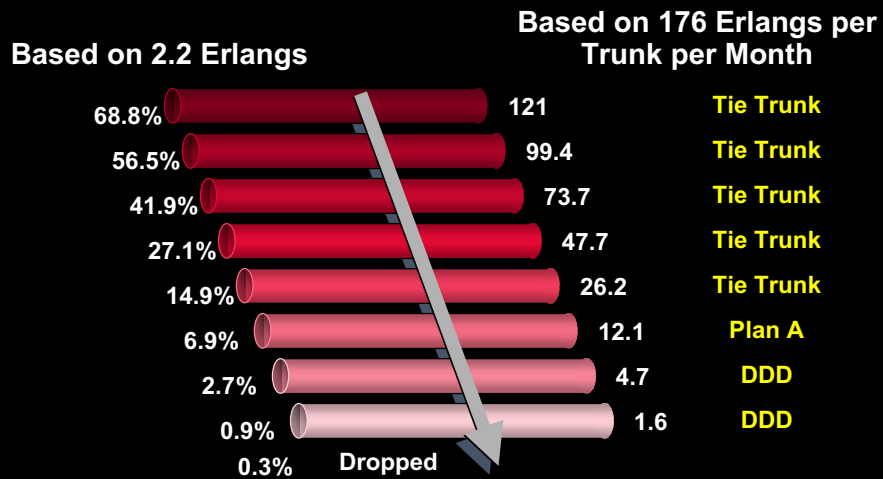
Traffic Distribution Based on Average Usage (Cont.)

Erlangs Carried on Each Trunk Based on 2.2 Erlangs Total



www.cisco.com

Average Trunk Usage on a Monthly Basis

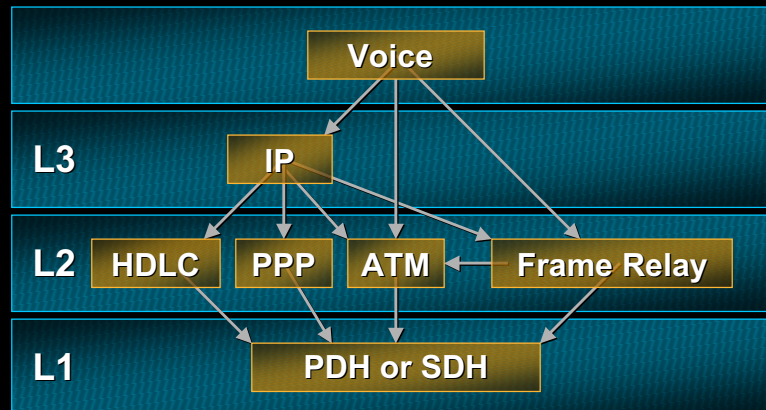


Agenda

- **Basic Voice Traffic Engineering**
- **How Costs Impact Trunk Groups**
- **How to Apply Voice Traffic Engineering to a Packet-Based Data Network**

www.cisco.com

Voice Transport Choices



Each Layer Adds a Bandwidth Requirement which Must Be Accounted for in Proper Network Traffic Engineering

www.cisco.com

Erlang to Packet Conversion

- 1 Erlang = 60 minutes = 64Kbps * 3600 seconds / 8bits/byte = 28.8 Mbytes
- IP
1 Erlang = 450K packets (64 byte packets) or **125 pps**
- Frame Relay
1 Erlang = 1.44M frames (20 byte frames) or **400 fps**
- ATM
1 Erlang = (AAL1) 655K cells or **182 cells/second**
= (AAL5) 600K cells or **167 cells/second**

www.cisco.com

Voice Compression Bandwidth Requirements

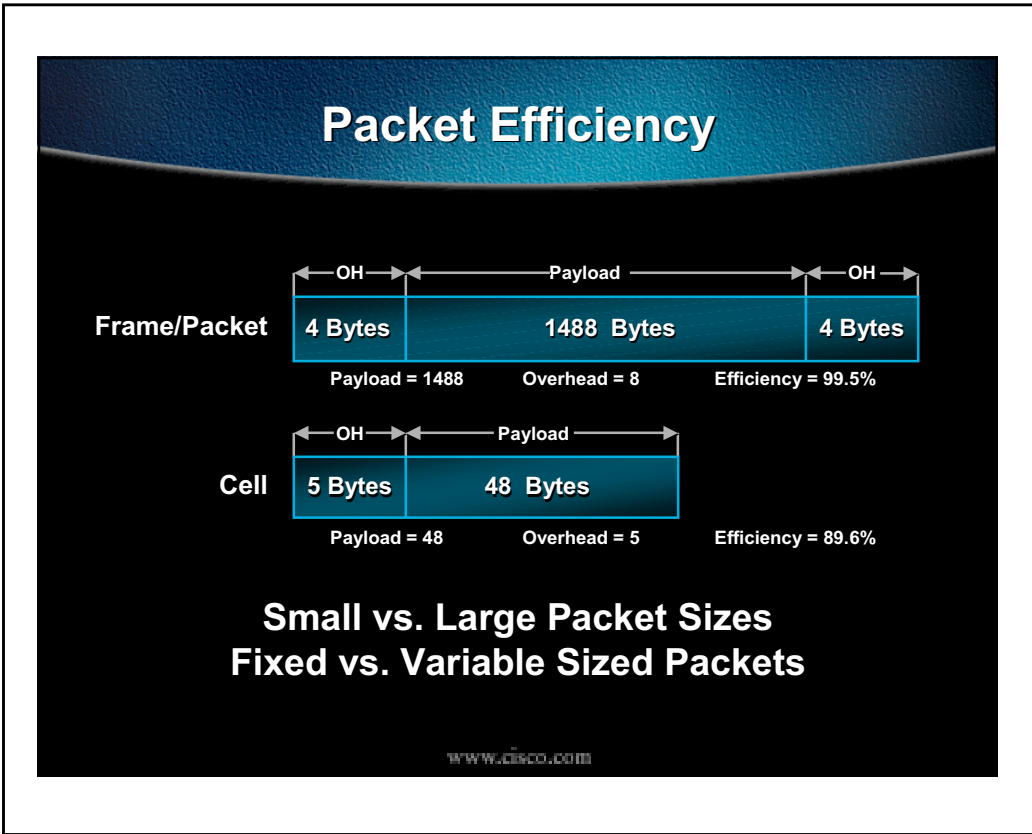
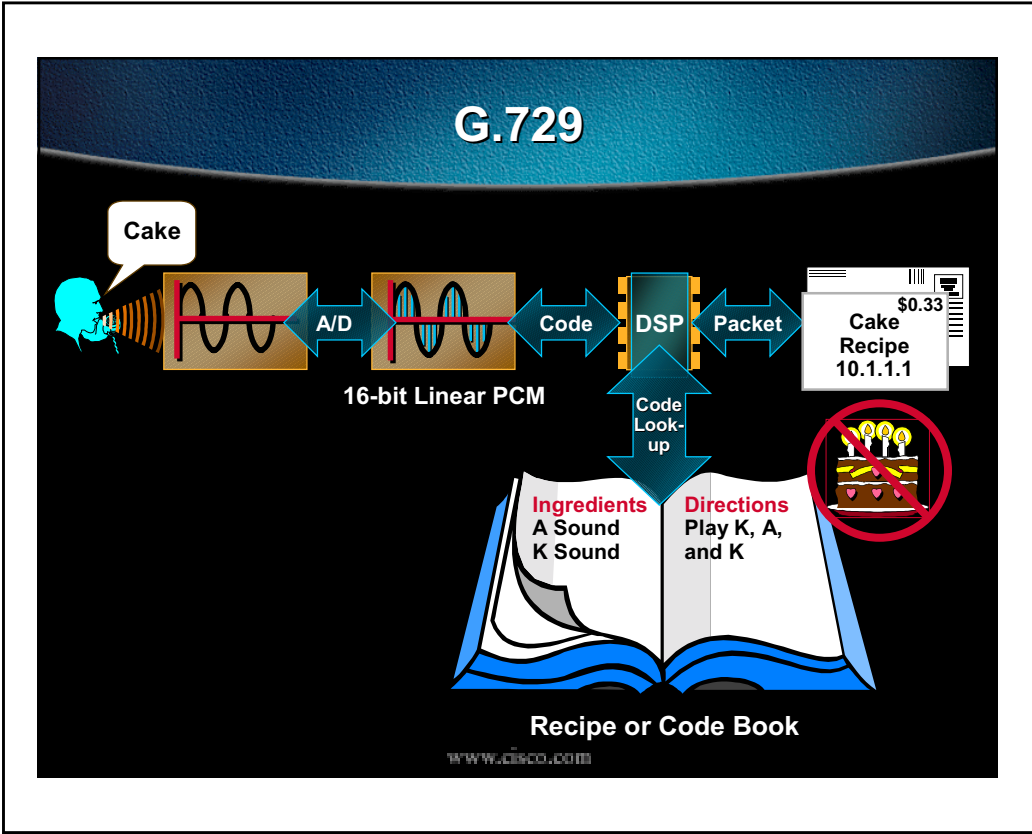
Encoding/Compression	Result Bit Rate
G.711 PCM	64 kbps (DS0)
G.726 ADPCM	16, 24, 32, 40 kbps
G.727 E-ADPCM	16, 24, 32, 40 kbps
G.729 CA-CELP	8 kbps
G.728 LD-CELP	16 kbps
G.723.1	6.3/5.3 kbps Variable

www.cisco.com

G.729 Voice Compression

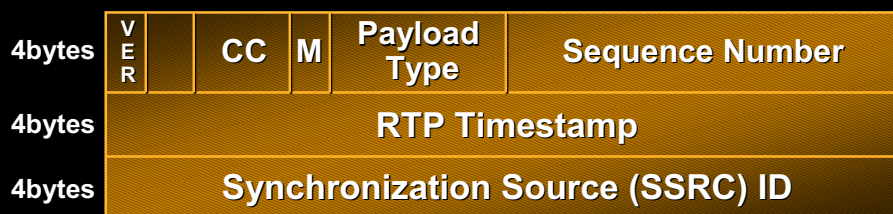
- Will reduce the bandwidth required to 1/8 th, 1 Erlang = 3.6Mbytes
- G.729 coder will output voice frames in 10 msec segments; 1–10msec frame = 10 bytes
- Bandwidth may or may not fit perfectly into a packet; for example, a 48 byte ATM cell will have at least 8 bytes of padding

www.cisco.com



RTP/RTCP—RFCs 1889/1890

- End-to-end network transport function
 - Payload type identification—voice, video, compression type
 - Sequence numbering
 - Time stamping
 - Delivery monitoring
- RTCP (Real-Time Control Protocol)

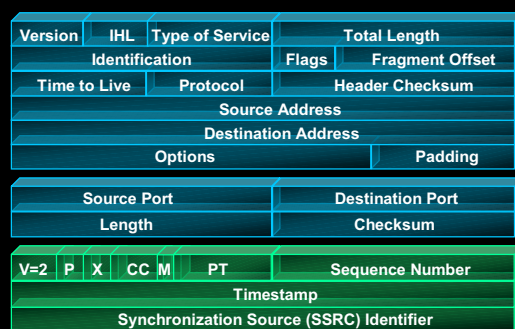


www.cisco.com

VoIP Bandwidth Solution

RTP Header Compression

- 20ms@8kb/s yields 20 byte payload
- IP header 20; UDP header 8; RTP header 12
2X payload!!!!!!!
- Header compression 40 bytes to 2 or 4 bytes
- **Hop-by-hop** on slow links <512kbps
- CRTP—Compressed Real-time Protocol (CRTP)



www.cisco.com

RSVP: Resource Reservation Protocol

- IETF signaling protocol
Reservation of bandwidth and delay
- Flow can be signaled by end station or by router (static reservation)
- Basically reserves queue space



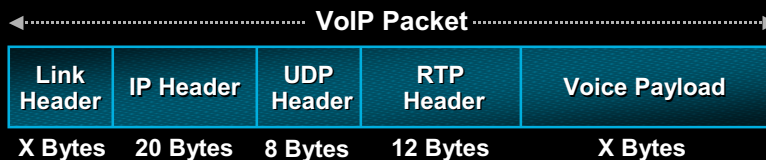
Non RSVP Enabled Routers Pass the VoIP Flow as Best Effort



RSVP Enabled Router See the PATH and RESERVE Messages and Allocate the Appropriate Queue Space for the Given Flow

www.cisco.com

VoIP Packet Format



Note: Link Layer Sizes Vary per Media

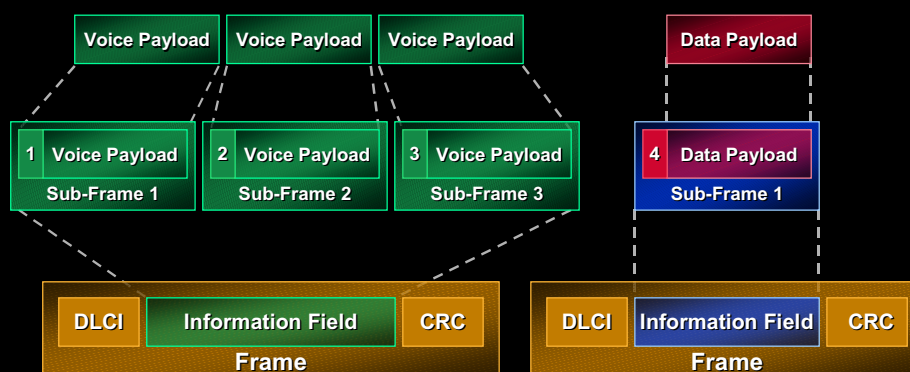
www.cisco.com

VoIP Bandwidth

- **Voice payload calculation**
20 msec voice sample * 8 kbps (for G.729) /
8 bits/byte = 20 bytes
- **Packet size calculations**
20 byte payload, uncompressed header (40 bytes) or
RTP/CRTP (5 bytes)
- **Bandwidth calculations**
Router using G.711 = 160 byte voice payload at 50 pps (**80 kbps**)
Router using G.729 = 20 byte payload at 50 pps (**24 kbps**)
IP Phone using G.711 = 240 byte payload at 33 pps (**74.6 kbps**)
IP Phone using G.723.1 = 24 byte payload at 33 pps (**17 kbps**)
(Not including link layer header or CRTP)

www.cisco.com

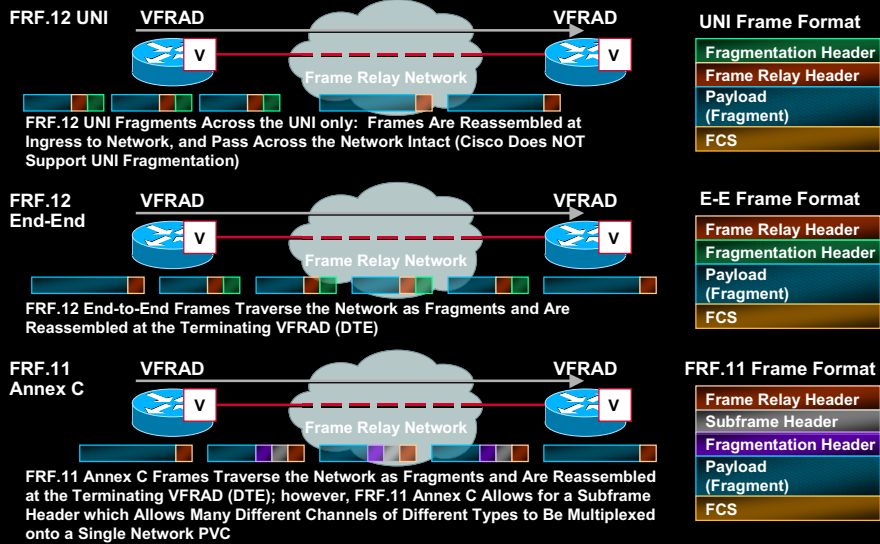
Multiple Sub-Channel Payloads in an FRF.11 Frame



Source: Frame Relay Forum

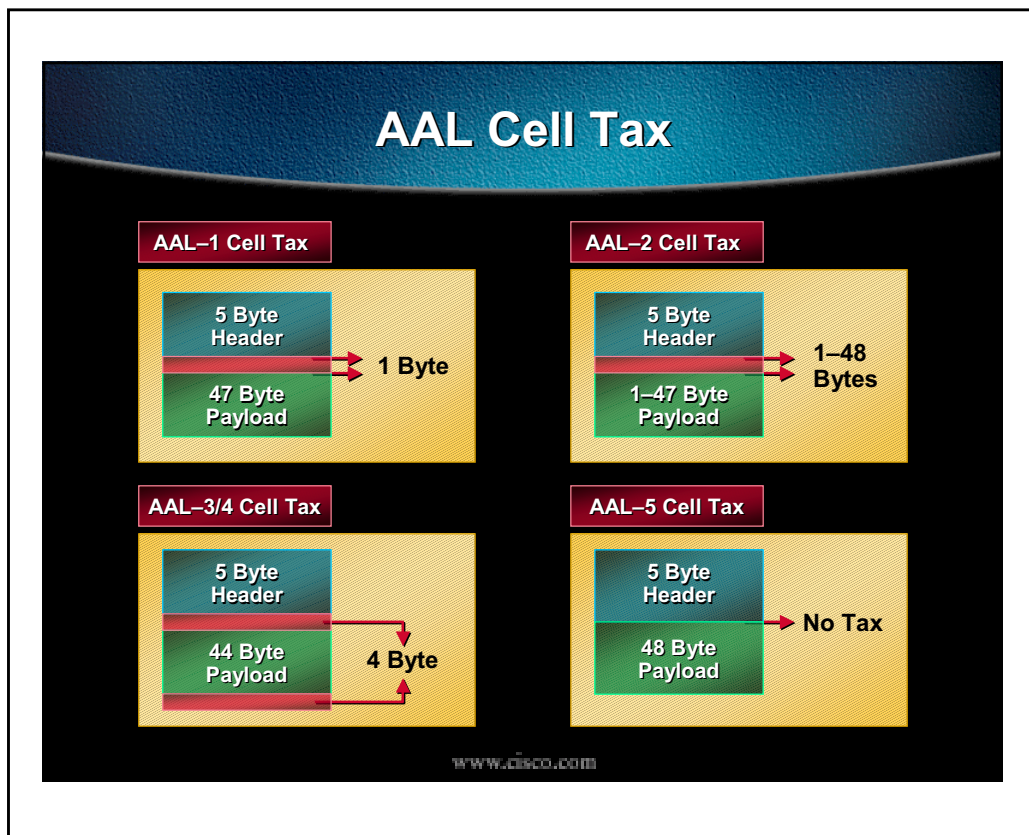
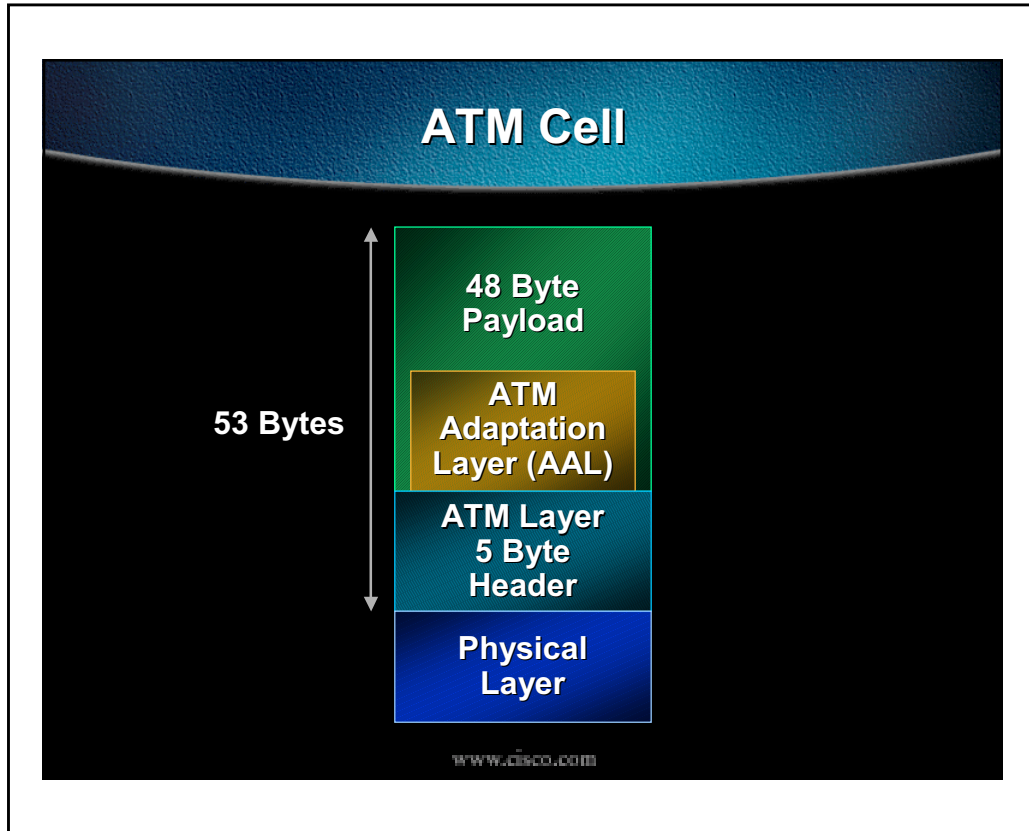
www.cisco.com

Comparison of FRF.11/12 Fragmentation Schemes

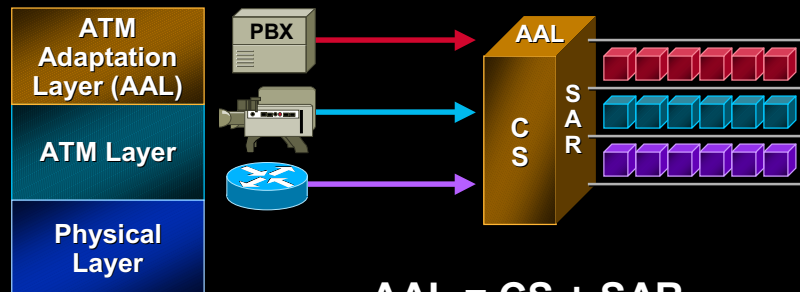


VoFR Bandwidth

- **Voice payload calculation**
 $20 \text{ msec voice sample} * 8 \text{ kbps (for G.729)} / 8 \text{ bits/byte} = 20 \text{ bytes}$
 Note: to derive the payload for G.711, substitute 64 kbps = 160 bytes
 - **Packet size calculations**
 $20 \text{ byte payload} + 7 \text{ byte Header} = 27 \text{ bytes}$
 (Header = DLCI/FRF.11/seqn/CRC)
 - **Bandwidth calculations**
 $27 \text{ b/voice packet} * 8 \text{ bits/byte} * 50 \text{ pps} = 10.8 \text{ kbps per call}$
- www.cisco.com



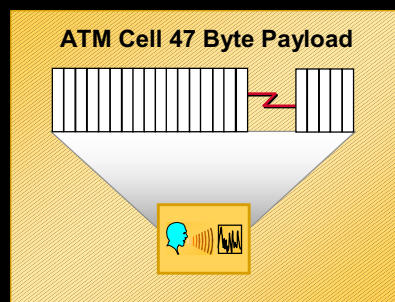
ATM Adaptation Layer—AAL



- **AAL = CS + SAR**
- **CS—cell tax**
- **SAR—cell <-> packet**

www.cisco.com

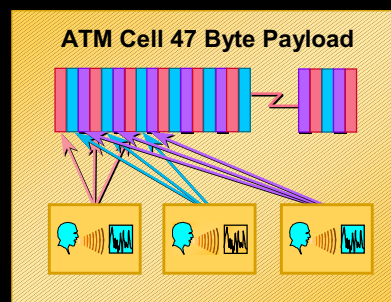
ATM Voice Transport Choices



Unstructured circuit emulation service

- Entire payload filled with one voice channel
- Or a continuous stream of bytes from an E1 circuit is placed into cells

Note: AAL1 Uses One Byte for Overhead

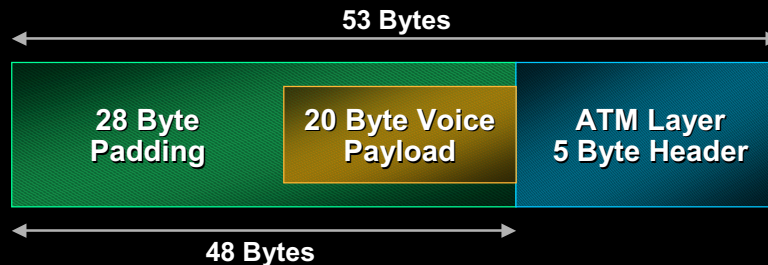


Structured circuit emulation service

- Entire payload filled with many voice channels
- DS0 structure information is maintained

www.cisco.com

ATM AAL5 Voice Cells



- G.729 compression with 20 ms voice sample
- No AAL5 CS “Cell Tax”
- 28 Bytes “Overhead” due to padding

www.cisco.com

VoATM Bandwidth

- **Voice payload calculation**
 $20 \text{ msec voice sample} * 8 \text{ kbps (for G.729)} / 8 \text{ bits/byte} = 20 \text{ bytes}$
Note: to derive the payload for G.711, substitute 64 kbps = 160 bytes
- **Packet size calculations**
 $20 \text{ byte payload} + 28 \text{ byte pad} + 5 \text{ byte header} = 53 \text{ bytes}$
- **Bandwidth calculations**
 $53 \text{ b/voice packet} * 8 \text{ bits/byte} * 50 \text{ pps} = 21.2 \text{ kbps per call}$

www.cisco.com

Various Link Layer Header Sizes “Varying Bit Rates per Media”

Example

G.729 with 60 byte packet (Voice and IP Header) at 50pps
(No RTP Header Compression)

Media	Link Layer Header Size	Bit Rate
Ethernet	14 Bytes	29.6 kbps
PPP	6 Bytes	26.4 kbps
Frame Relay	4 Bytes	25.6 kbps
ATM	5 Bytes per Cell	42.4 kbps

Note: For ATM a Single 60 Byte Packet Requires Two 53 Byte ATM Cells

www.cisco.com

Traffic Distribution: How Do Voice Trunks Equate to Bandwidth?

- Still use 8 trunks to give us the required grade of service
- At any point in time we could be using $8 * 64$ kbps (PCM) or **512 kbps** of bandwidth
- During the busy hour we will be using 2.64 trunks x 64 kbps or **169 kbps** (or 33% of 512 kbps)
- During the average hour we will be using 2.2 trunks x 64 kbps or **141 kbps** (or 27.5% of 512 kbps)
- Over a period of a month we will be using 352 Erlangs out of a possible 1408 (8 trks * 8 hrs/day * 22 days) or 25%

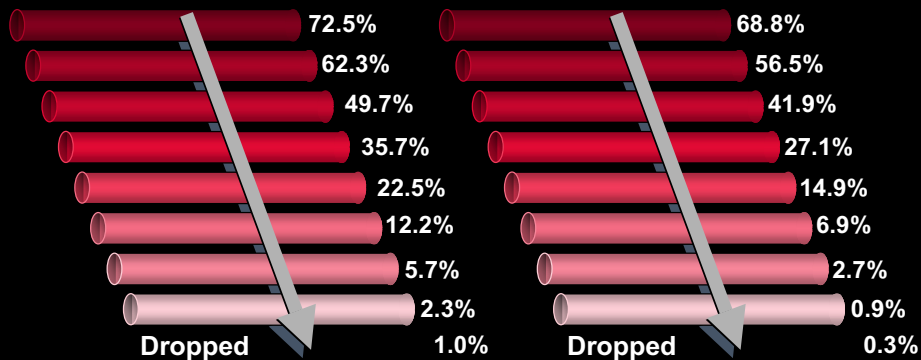
www.cisco.com

Traffic Distribution: How Do Voice Trunks Equate to Bandwidth?

Using Our Earlier Example

2.64 Erlangs during BH

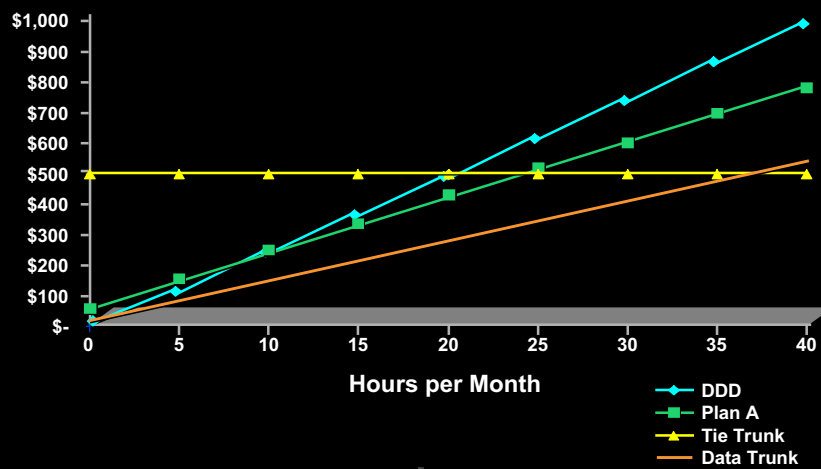
2.2 Erlangs during AH



www.cisco.com

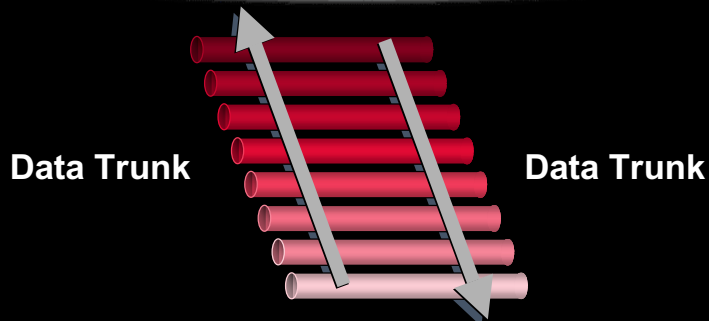
Cost Analysis Including Packet Pricing

Comparison of XYZ Pricing



www.cisco.com

Other Trunk Selection Criteria



- Do we take the most busy trunks and move them on to the data network, or do we take the least busy trunks?
- The decision will be based on desired voice quality and costs, but will have an impact on the traffic load delivered to the data network

www.cisco.com

Agenda

- **Basic Voice Traffic Engineering**
- **How Costs Impact Trunk Groups**
- **How to Apply Voice Traffic Engineering to a Packet-Based Data Network**

www.cisco.com

