Datacenter Design for High Capacity Web Cache and Streaming

Vijay Gill <vijaygill9@aol.com>

Chris Wright <cjwright@aol.com>

Overview

- Problem Space

- Streaming

- Caching

A Moment of Reflection

The irony is that, we subtly contribute to this by helping
to build network and organization infrastructure which
supports a highly networked economy. We think we're
doing better because our salaries go up and up and
up, when in fact, we spend more and more of our
income at... amazon.com. translation: less money
now than when I was flat broke.

-Faisal Jawdat

The Internets

BB
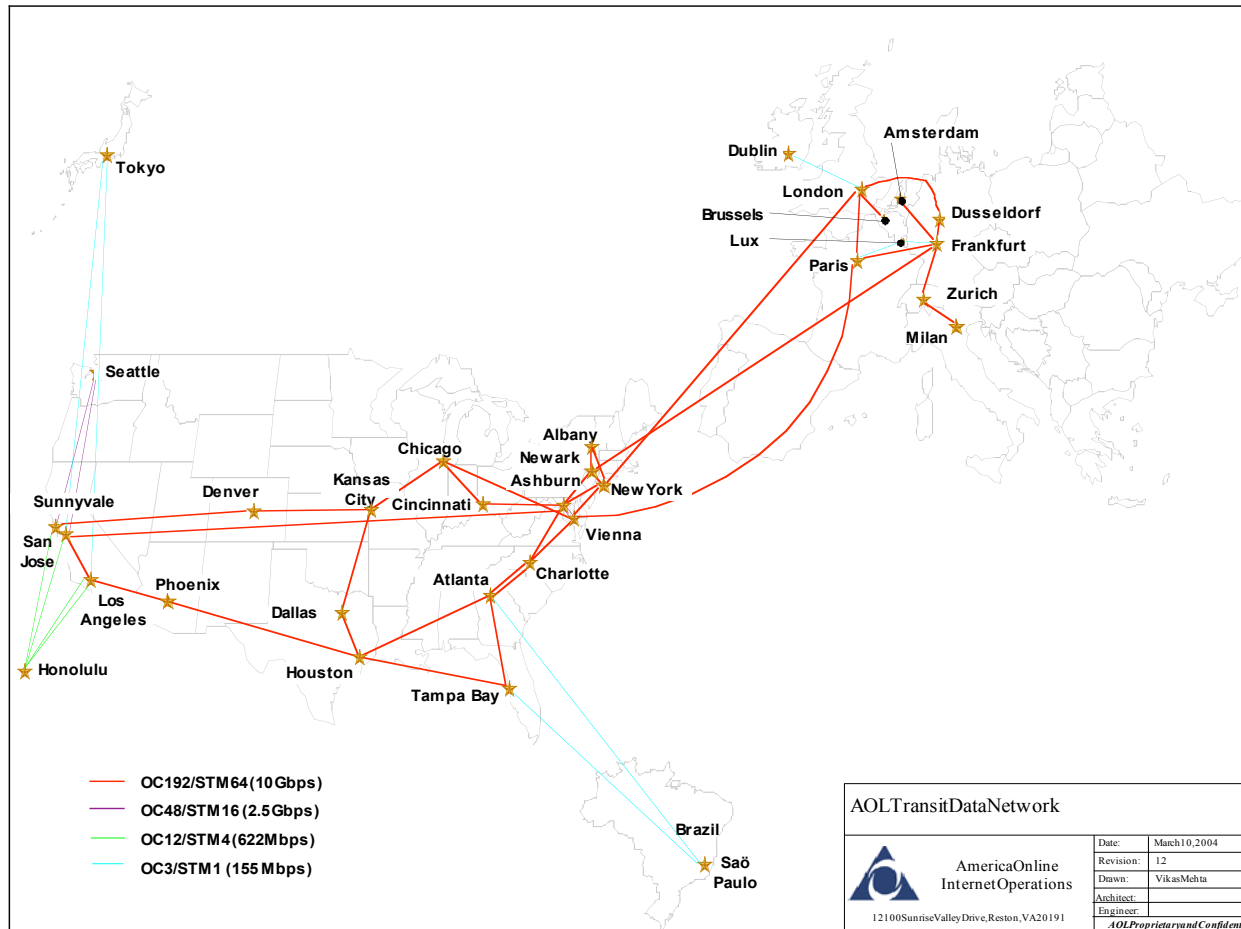
BB

POP

POP
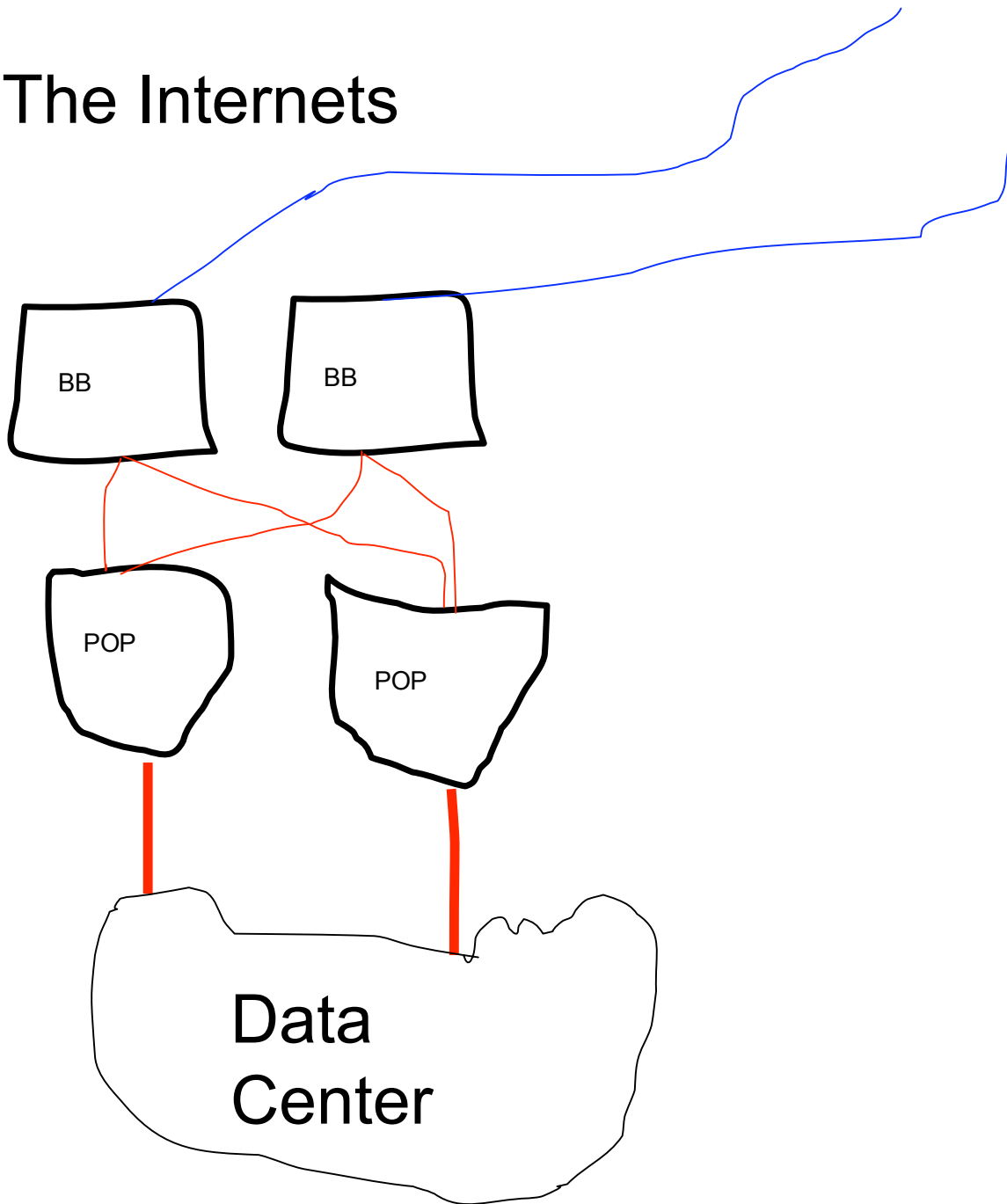
Data
Center

Data Center Design Philosophy

Complete life cycle in NetOps
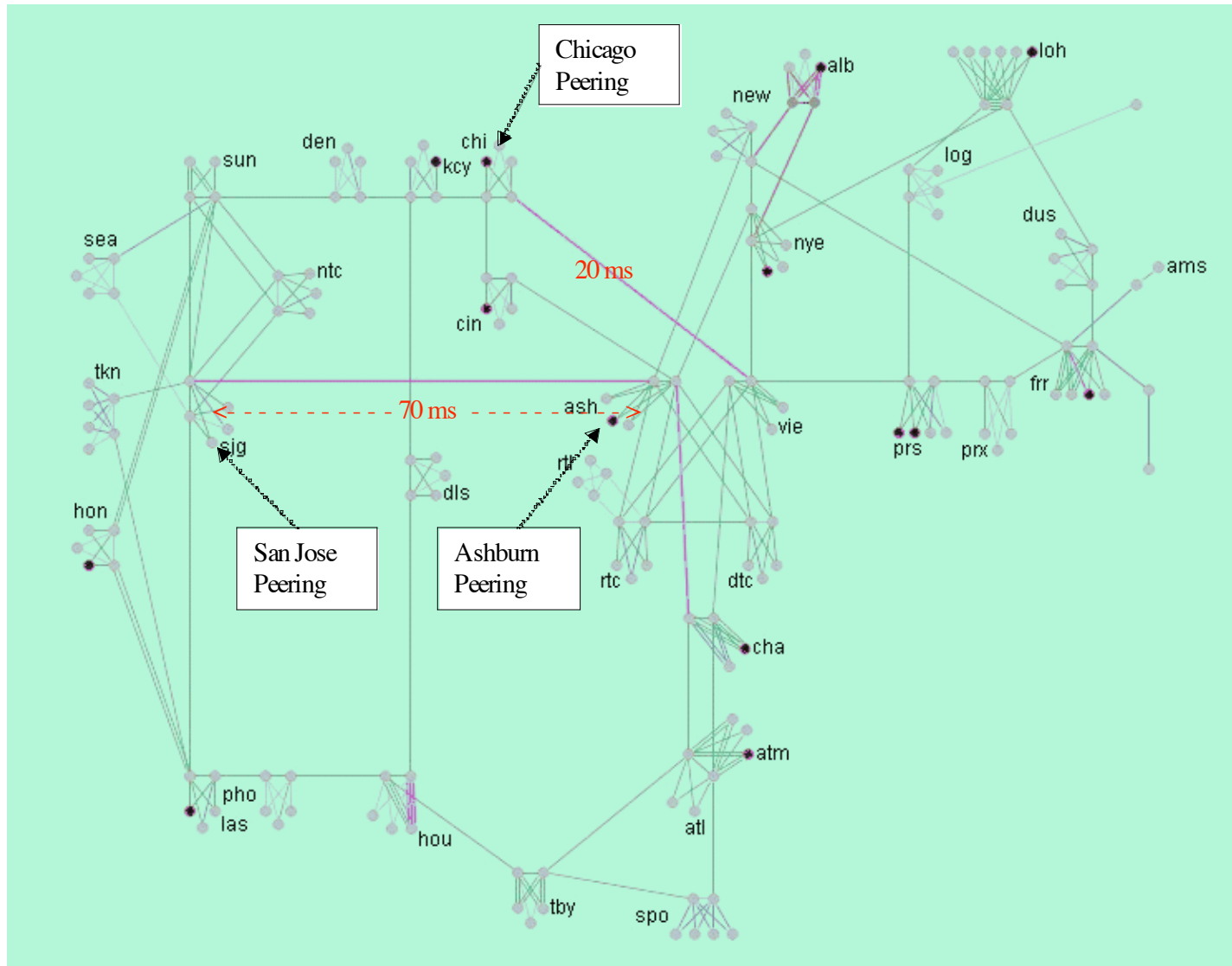
- Architect, design, implement, and operate

Design Criteria

- Diversity of component, paths, logical units
- No Single Point of Failure (SPOF)
    - ➢Redundant capacity to support peak load
- Routed topology
- System review to ensure performance & survivability goals

ATDN Map

Tokyo

Dublin

Amsterdam

London

Brussels

Dusseldorf

Lux

Frankfurt

Paris

Zurich

Milan

Seattle

Chicago

Albany

Newark

Kansas City

Ashburn

New York

Sunnyvale

Denver

Cincinnati

Vienna

San Jose

Phoenix

Atlanta

Charlotte

Los Angeles

Dallas

Honolulu

Houston

Tampa Bay

Brazil

Saö Paulo

OC192/STM64 (10Gbps)
OC48/STM16 (2.5Gbps)
OC12/STM4 (622Mbps)
OC3/STM1 (155Mbps)

AOLTransitDataNetwork

Date:	March10,2004
Revision:	12
Drawn:	VikasMehta
Architect:	
Engineer:	

AmericaOnline InternetOperations

12100SunriseValleyDrive,Reston,VA20191

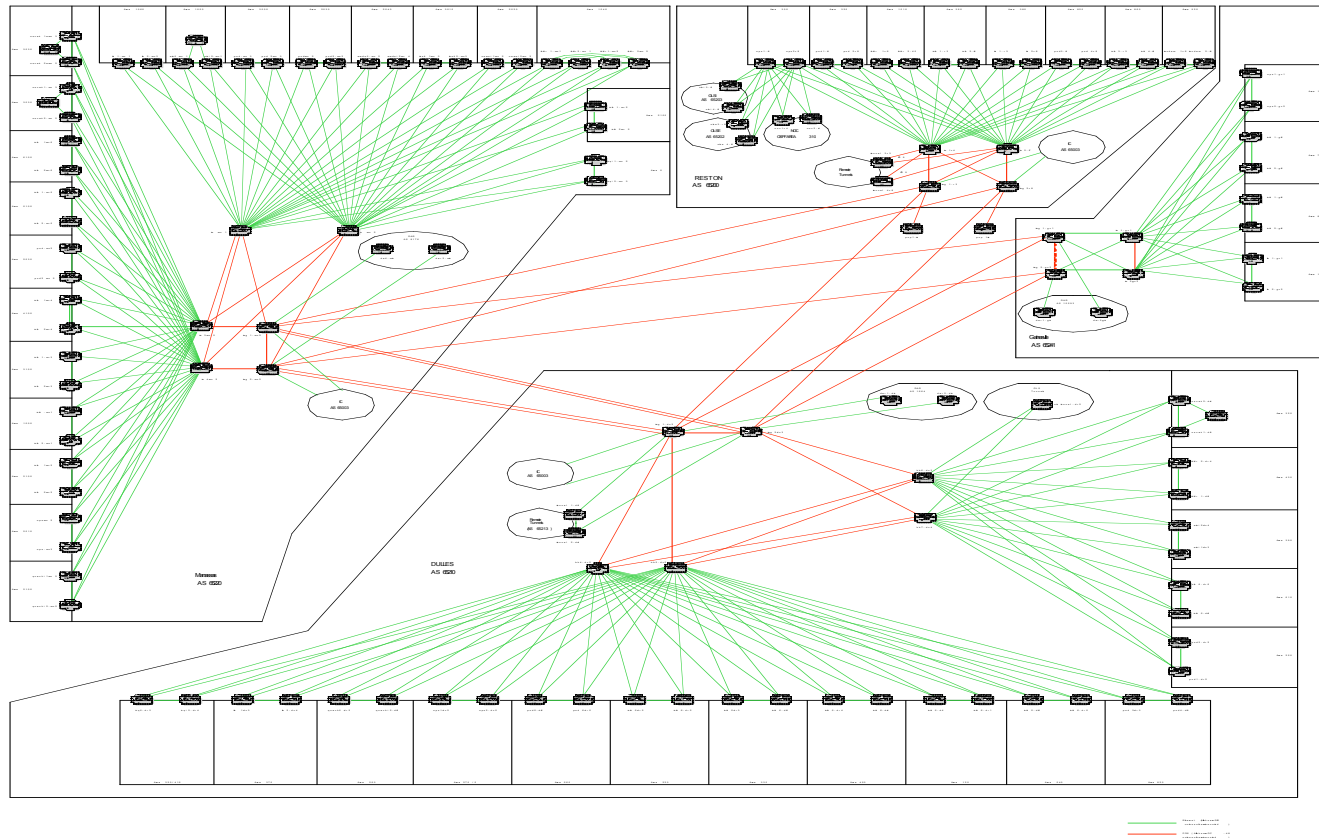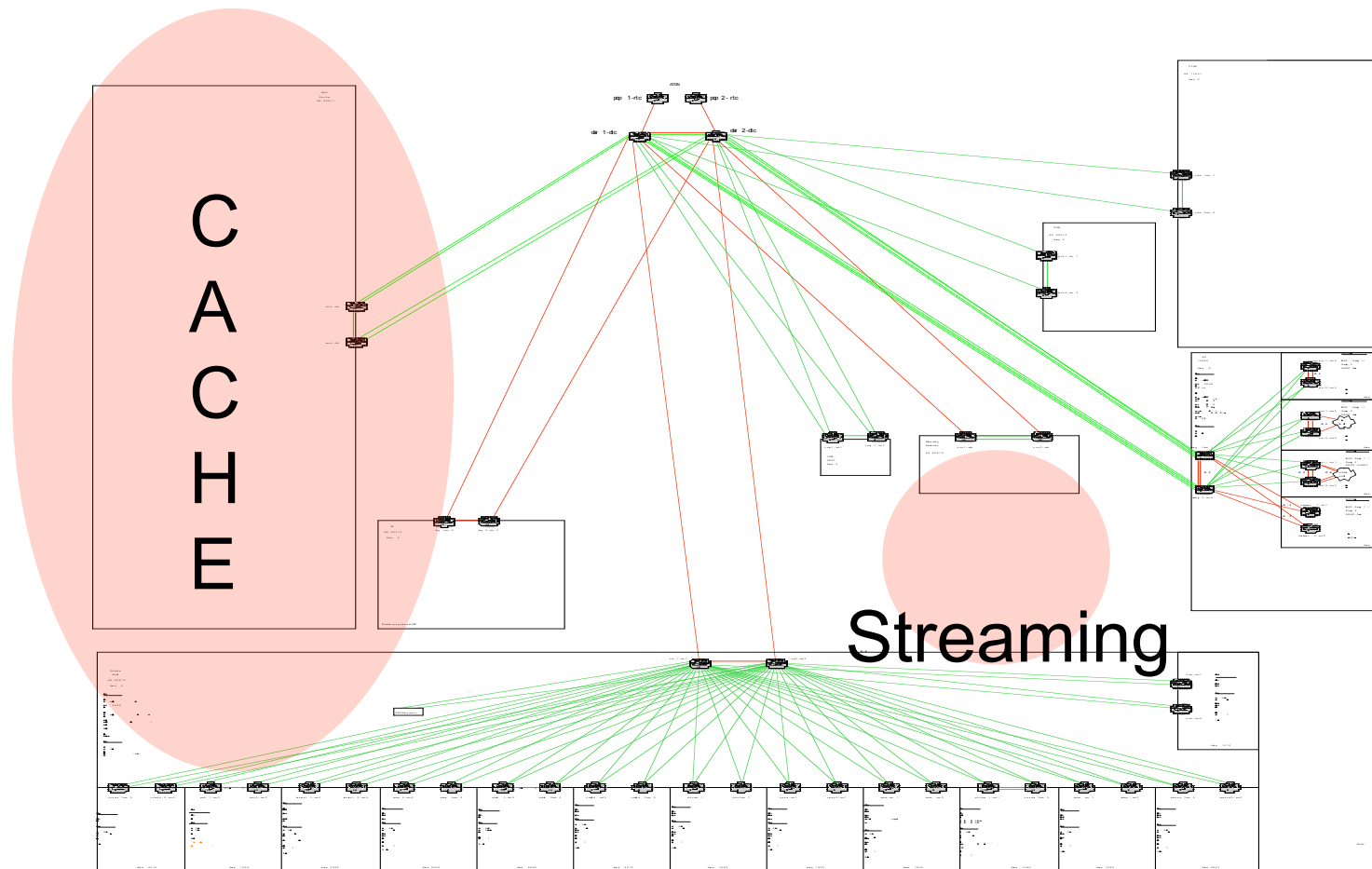*AOLProprietaryandConfidential*

Traffic Topology

Data Centers

This thing severely violated the Rule of Complexity. The Rule of Complexity states that if an answer seems too complicated to be the right answer, it is the wrong answer.
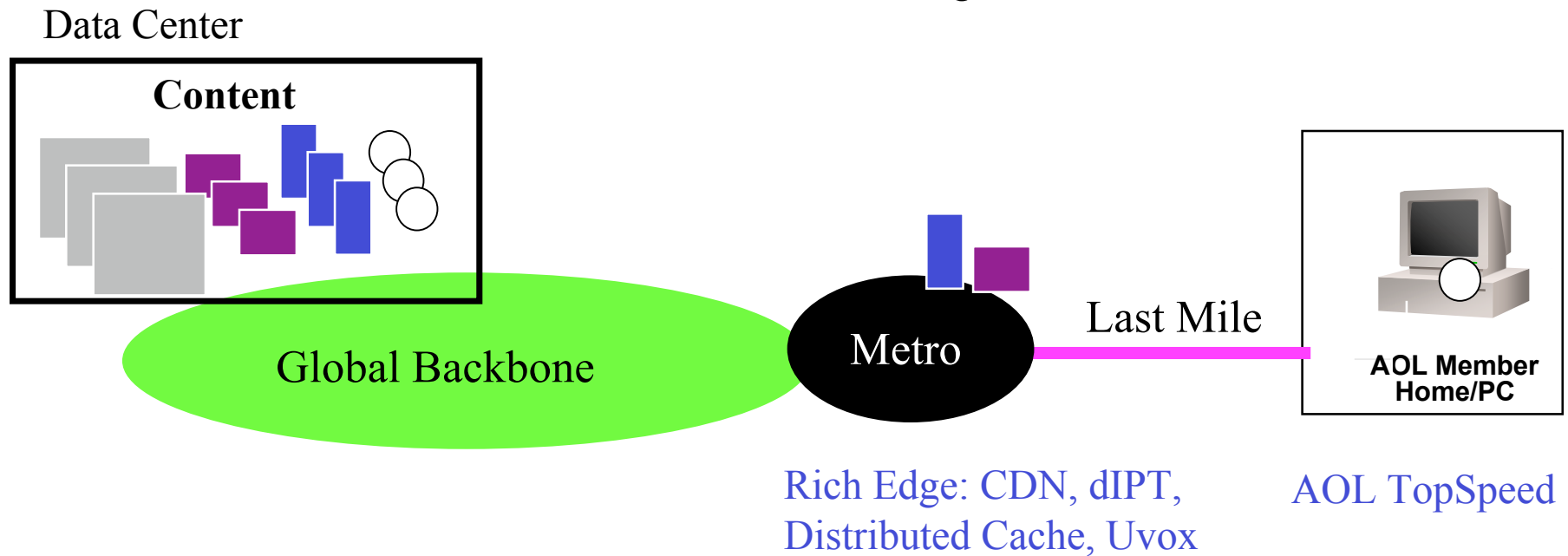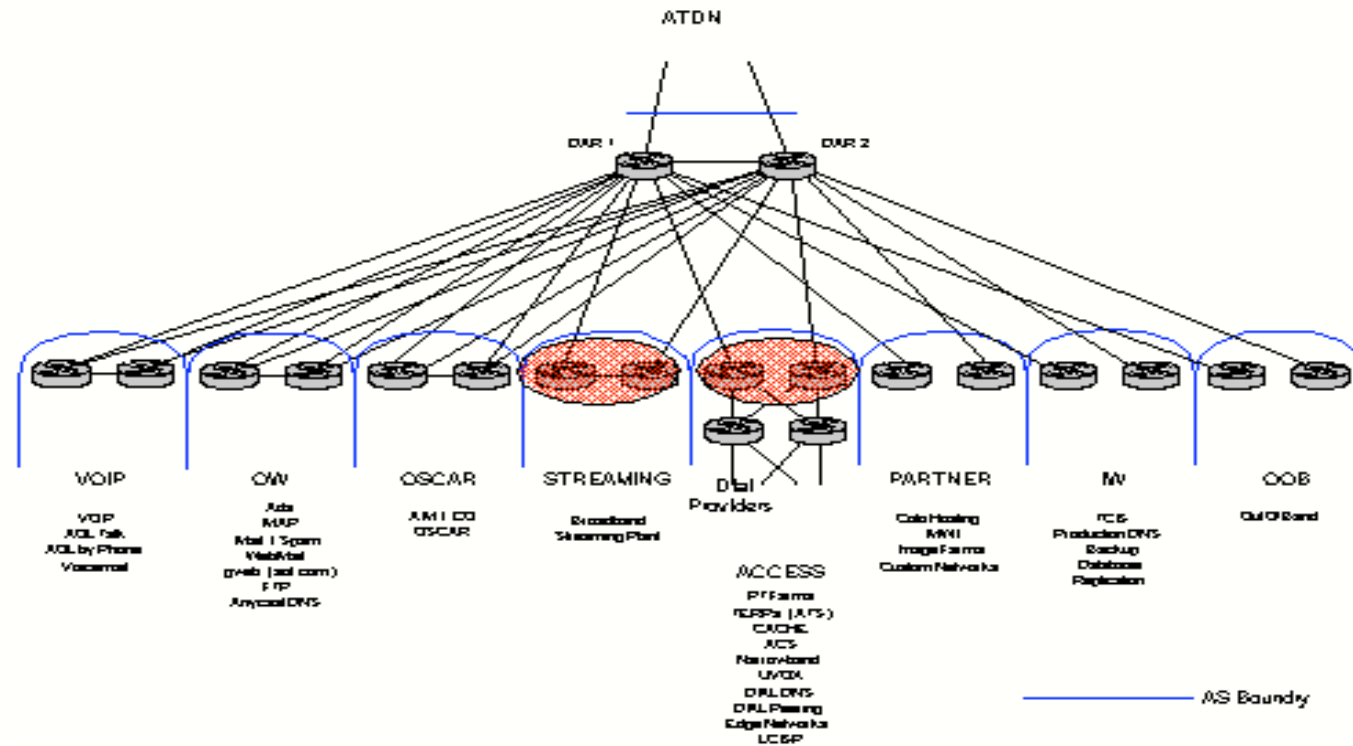 -Steve Cutchen

Data Centers

C
A
C
H
E

Streaming

Content Distribution Evolution

- Distribute the "Right Content"...

  - To the "Right Place"...

    - And at the "Right Time"…

Data Center

Content

Global Backbone

Metro

Last Mile

AOL Member Home/PC

Rich Edge: CDN, dIPT, Distributed Cache, Uvox
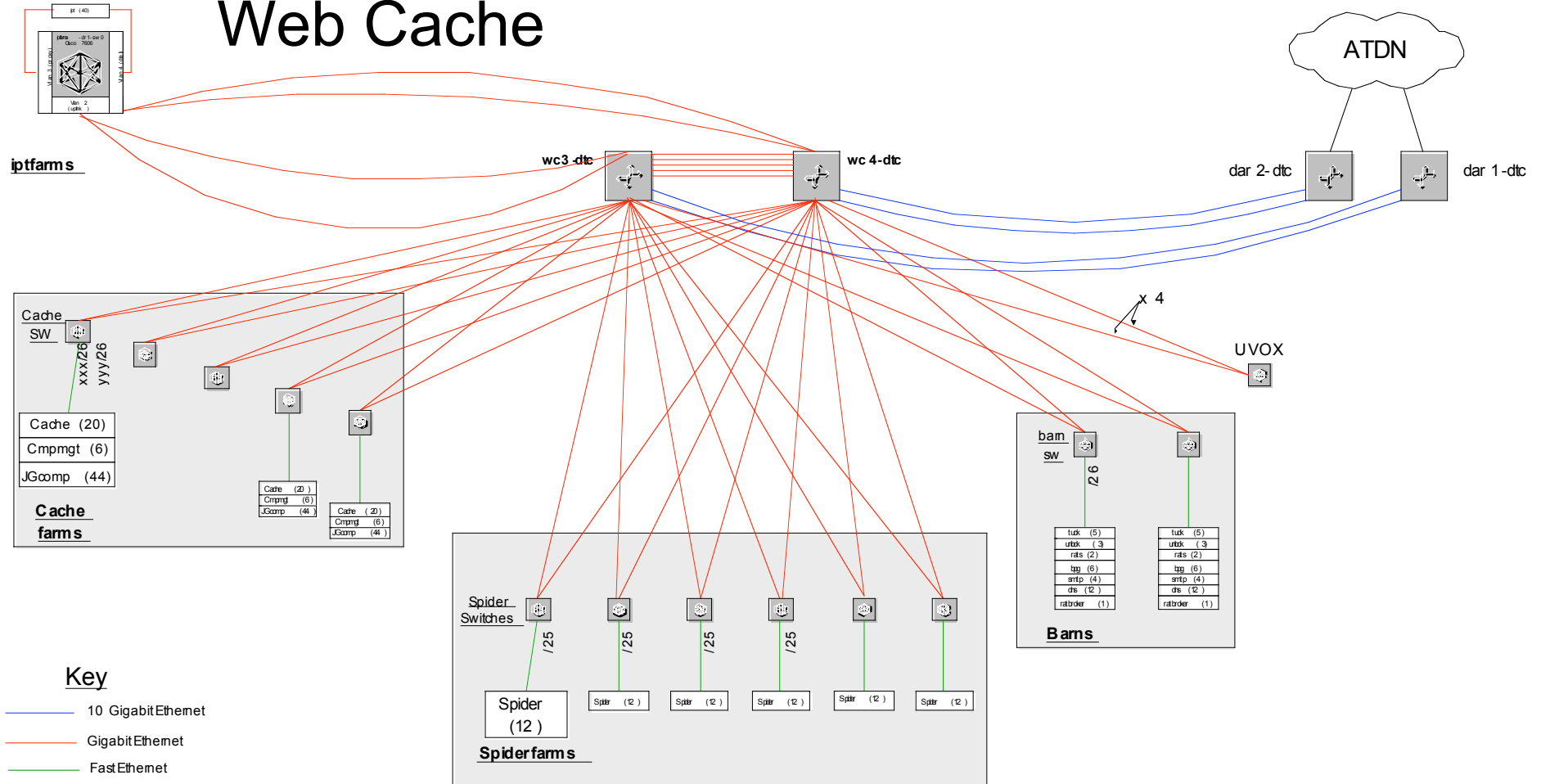
AOL TopSpeed

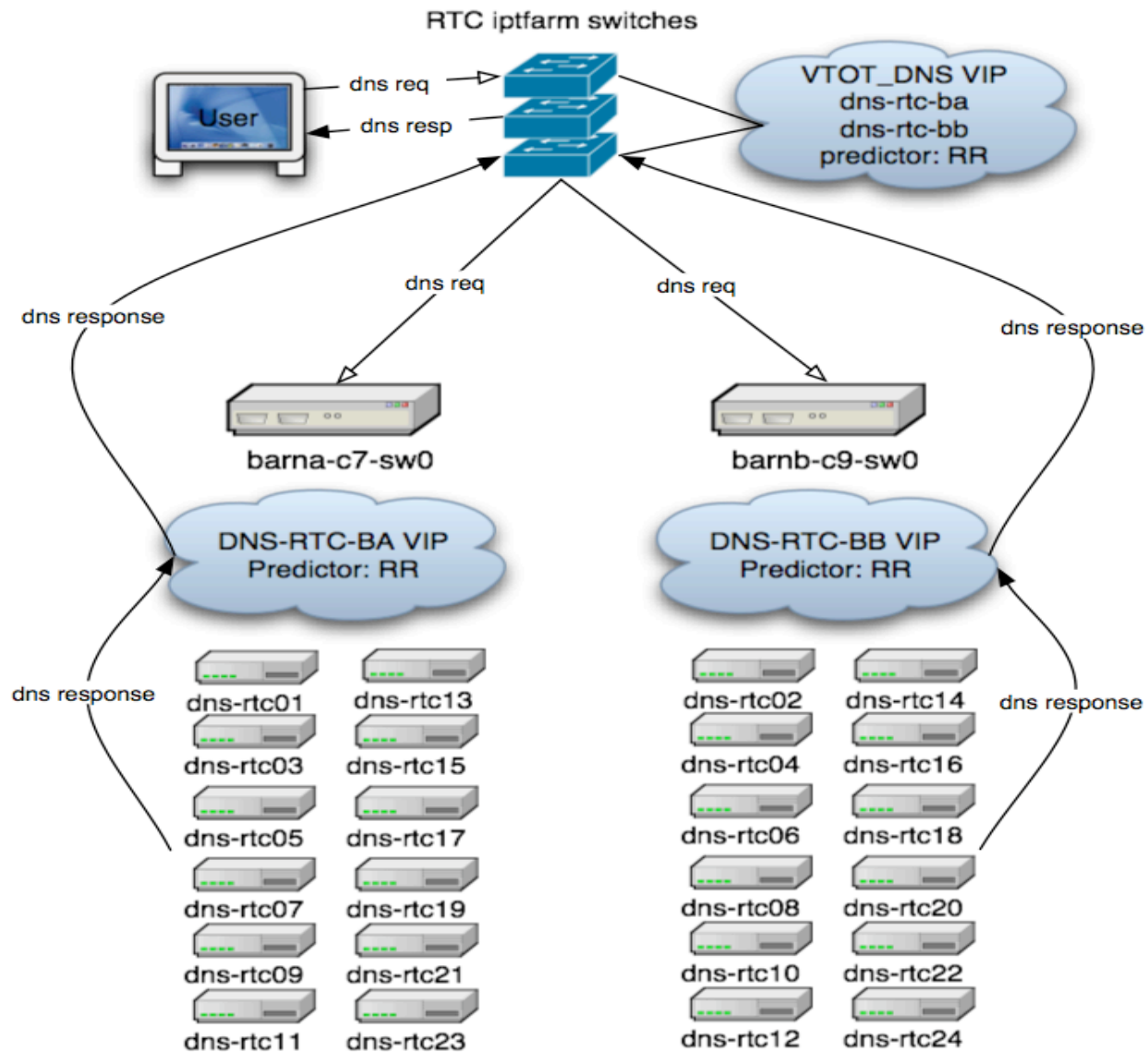# 10K Foot View

Farmed Design - Dial
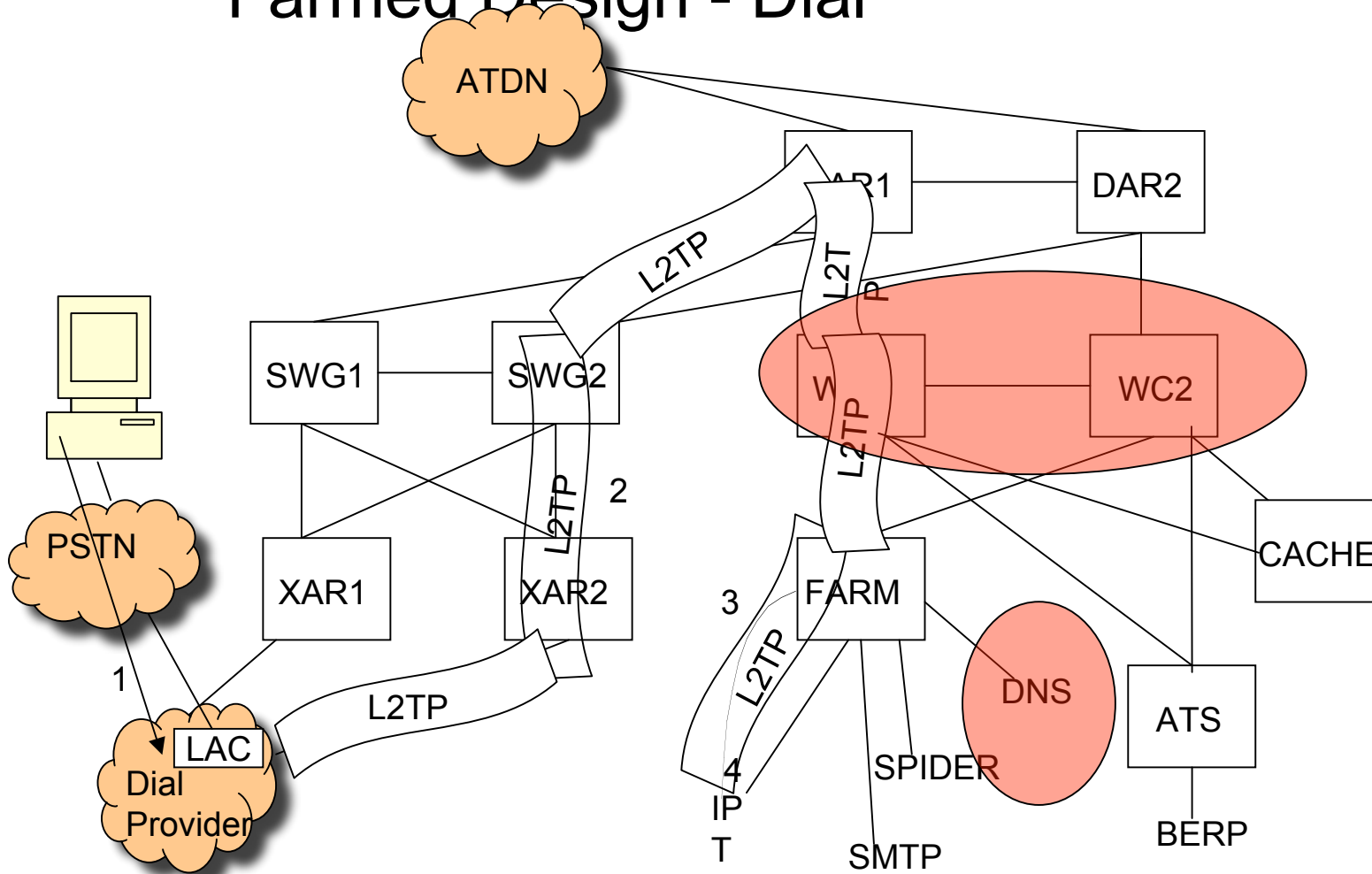
The day a vendor starts handing out leather jackets with high end products is the day he starts getting a lot of my business. -Jonathan Hartford

Web Cache

ATDN

iptfarms

ipt (40)

pfarms   - dt1 -sw 0
Cisco   7606

Vlan 2
(uplink)

wc3 -dtc

wc 4-dtc

dar 2- dtc

dar 1 -dtc

Cache
SW

xxx /26
yyy /26

Cache   (20)
Cmpmgt   (6)
JGcomp   (44)

**Cache
farms**

Cache   (20)
Cmpmgt   (6)
JGcomp   (44)

Cache   (20)
Cmpmgt   (6)
JGcomp   (44)

x 4

UVOX

barn
SW

/26

tuck	(5)
urbrk	(3)
rats	(2)
bng	(6)
smtp	(4)
dns	(12)
ratbroker	(1)

tuck	(5)
urbrk	(3)
rats	(2)
bng	(6)
smtp	(4)
dns	(12)
ratbroker	(1)

Barns

Spider
Switches

/25   /25   /25   /25

Spider
(12)

Spider (12)   Spider (12)   Spider (12)   Spider (12)   Spider (12)

Spider farms

Key

— 10 Gigabit Ethernet

— Gigabit Ethernet

— Fast Ethernet

Web Requests Served Daily: Over 13 Billion

DNS Load Balancing Architecture
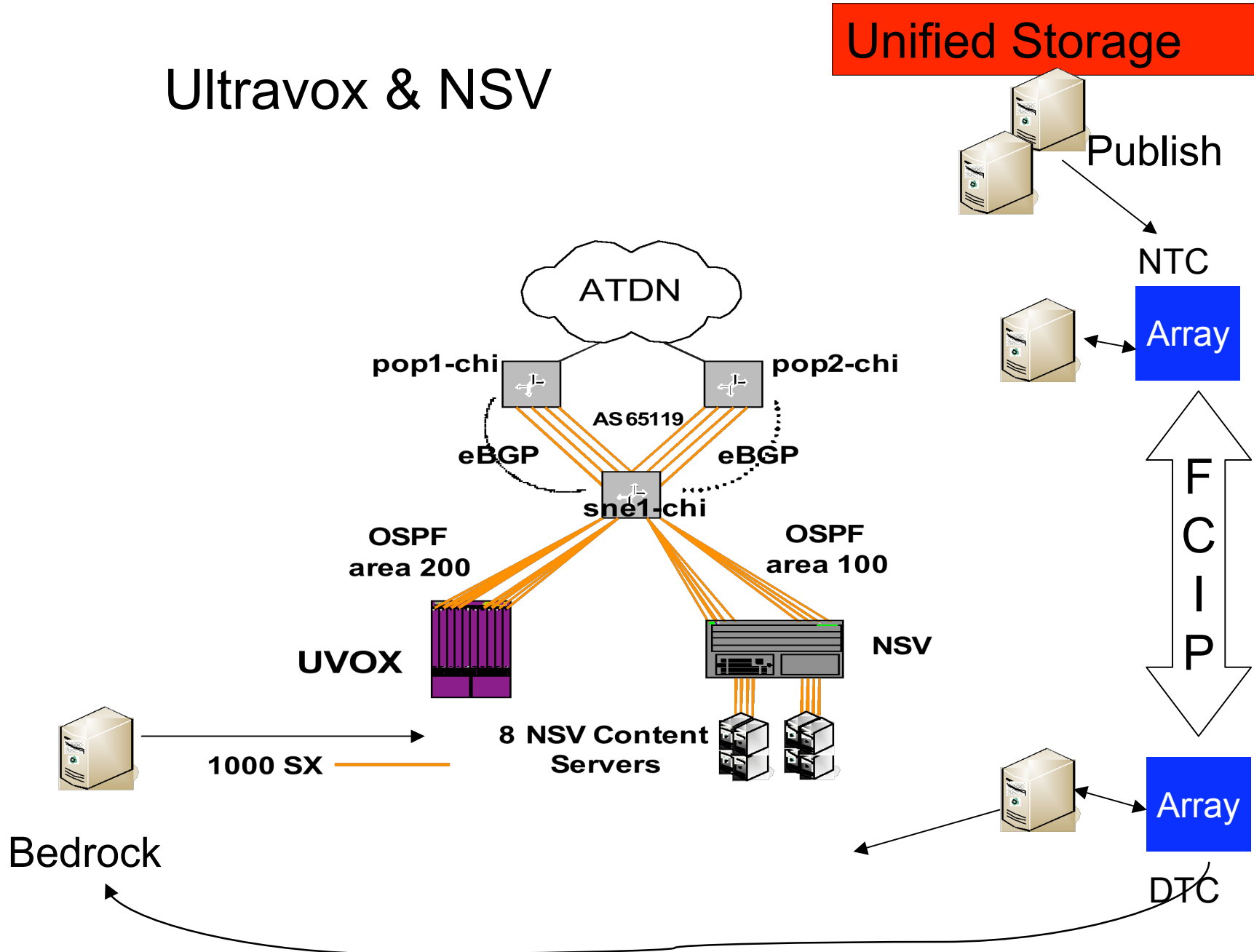
Streaming - Blocks

- Unified Storage – Replicated content storage and distribution

- UVOX – Bedrock servers take content from Unified Storage Complex

  - Already in correct codec

  - Place it in UVOX format

  - Send it to UVOX

- NSV

  - Store up 6 hours of cached content in Bedrock Servers
  - Send to Deathstar

spinner.com is the greatest thing to happen to the internet since aol's irc server got klined.
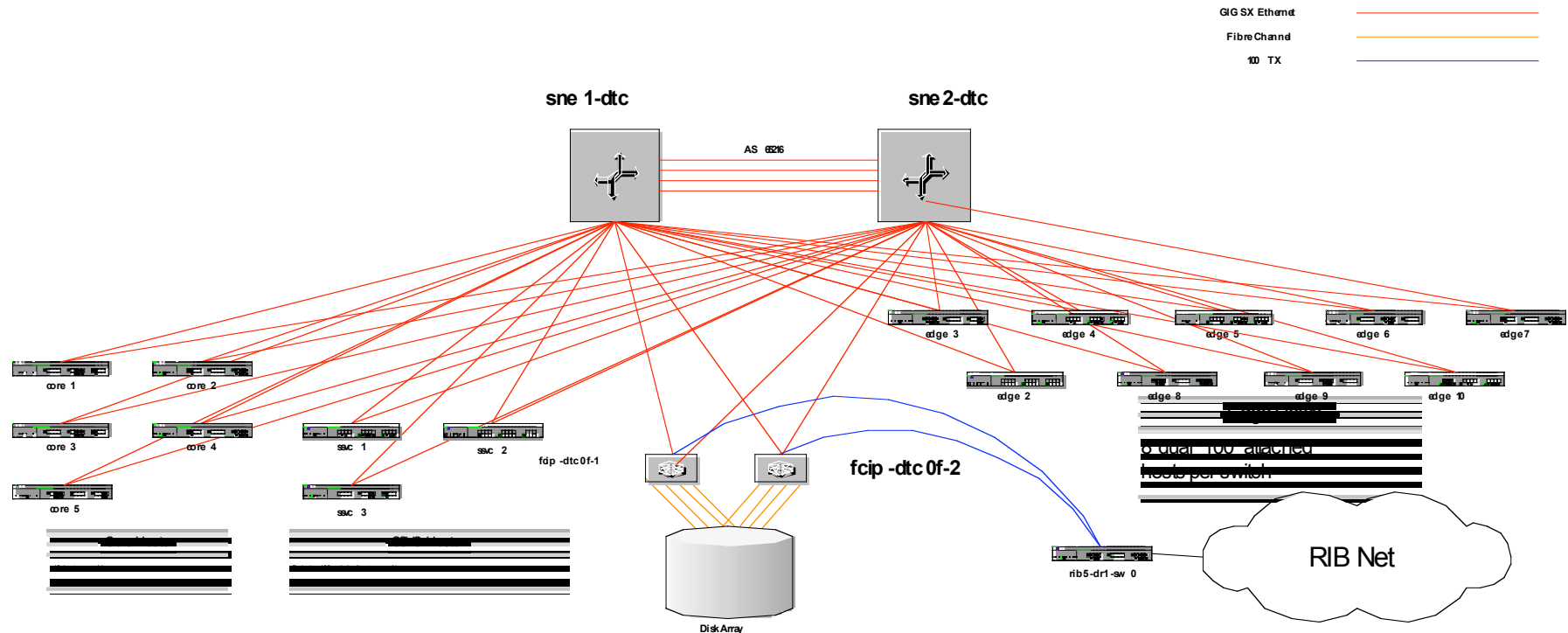
-Faisal Jawdat

Ultravox & NSV

Unified Storage

Publish

NTC

Array

ATDN

pop1-chi   pop2-chi

AS 65119

eBGP   eBGP

sne1-chi

OSPF
area 200

OSPF
area 100

UVOX

NSV

FCIP

8 NSV Content
Servers

1000 SX

Bedrock

Array

DTC

Bryan Blank: overclocking leads to instability, instability leads to hate
Bill Fumerola: hate leads to filing PRs and using the phrase "it worked in linux"

Unified Storage Complex

GIG SX Ethernet

FibreChannel

100 TX

sne 1-dtc

sne 2-dtc

AS 65216

edge 3   edge 4   edge 5   edge 6   edge 7

core 1   core 2

edge 2

edge 8   edge 9   edge 10

core 3   core 4   ssvc 1   ssvc 2

fdip -dtc0f-1   fcip -dtc0f-2

core 5   ssvc 3

rib5-dr1-sw 0

RIB Net

Disk Array

I lost my CD case. You know how I am when I lose my music. Sucked to be anything in view that night. -Jeremy Hartman

Live Content

- Some live content is encoded in the BOC

- BOC sends content to Unified Storage Complexes
  - Multiple levels of reflection and splitting happens

- Partners (Large News Corp)
  - Each encoder at LNC transmits both a low and a high bit video feed to the live transmitters in DTC. From there the feeds are then forwarded to the transmitters in each data center. Then via multicast it is pushed to the edge. (see. USC NTC DTC MTC)

Large News Network Live Streaming
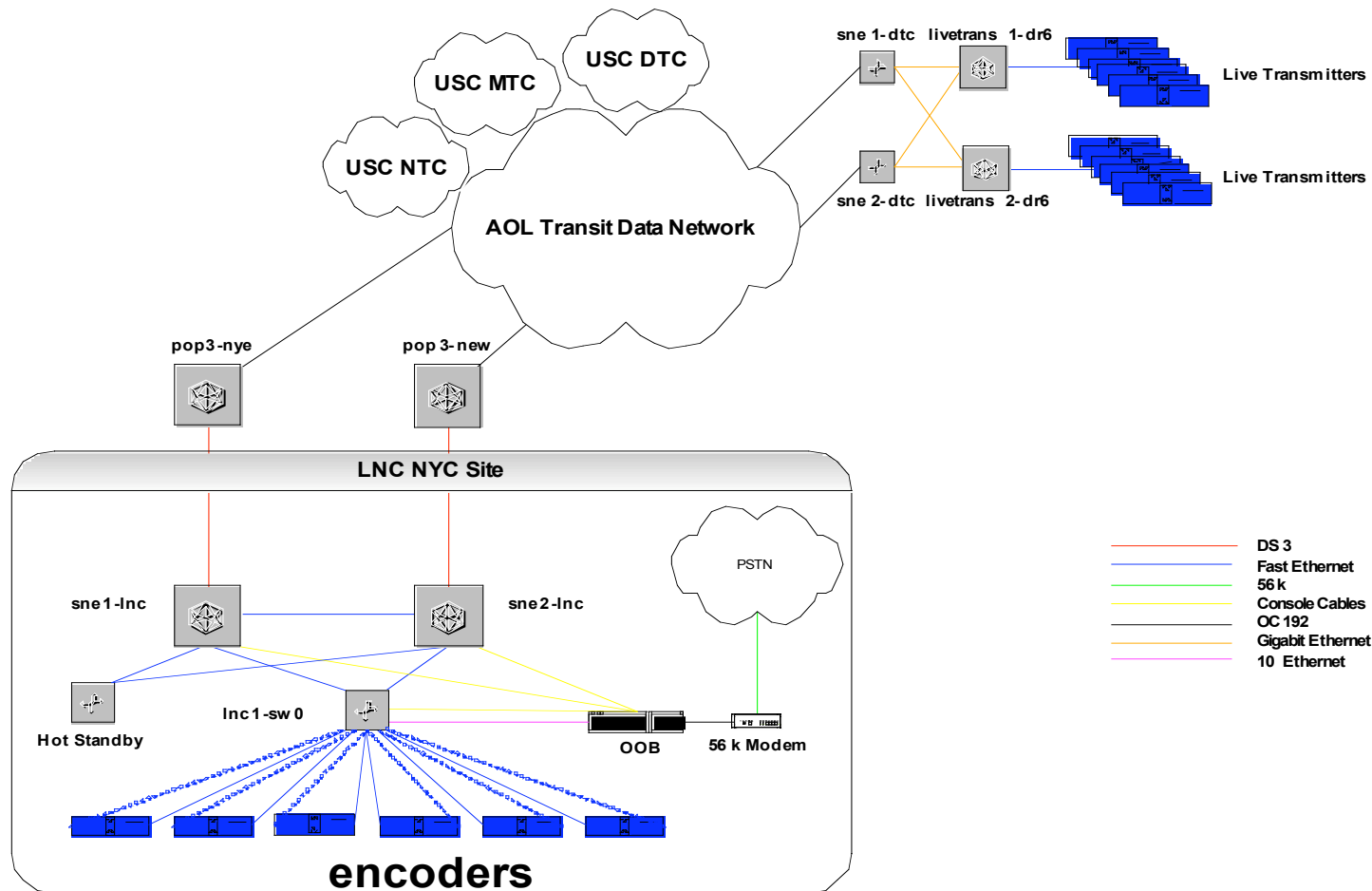
BOC
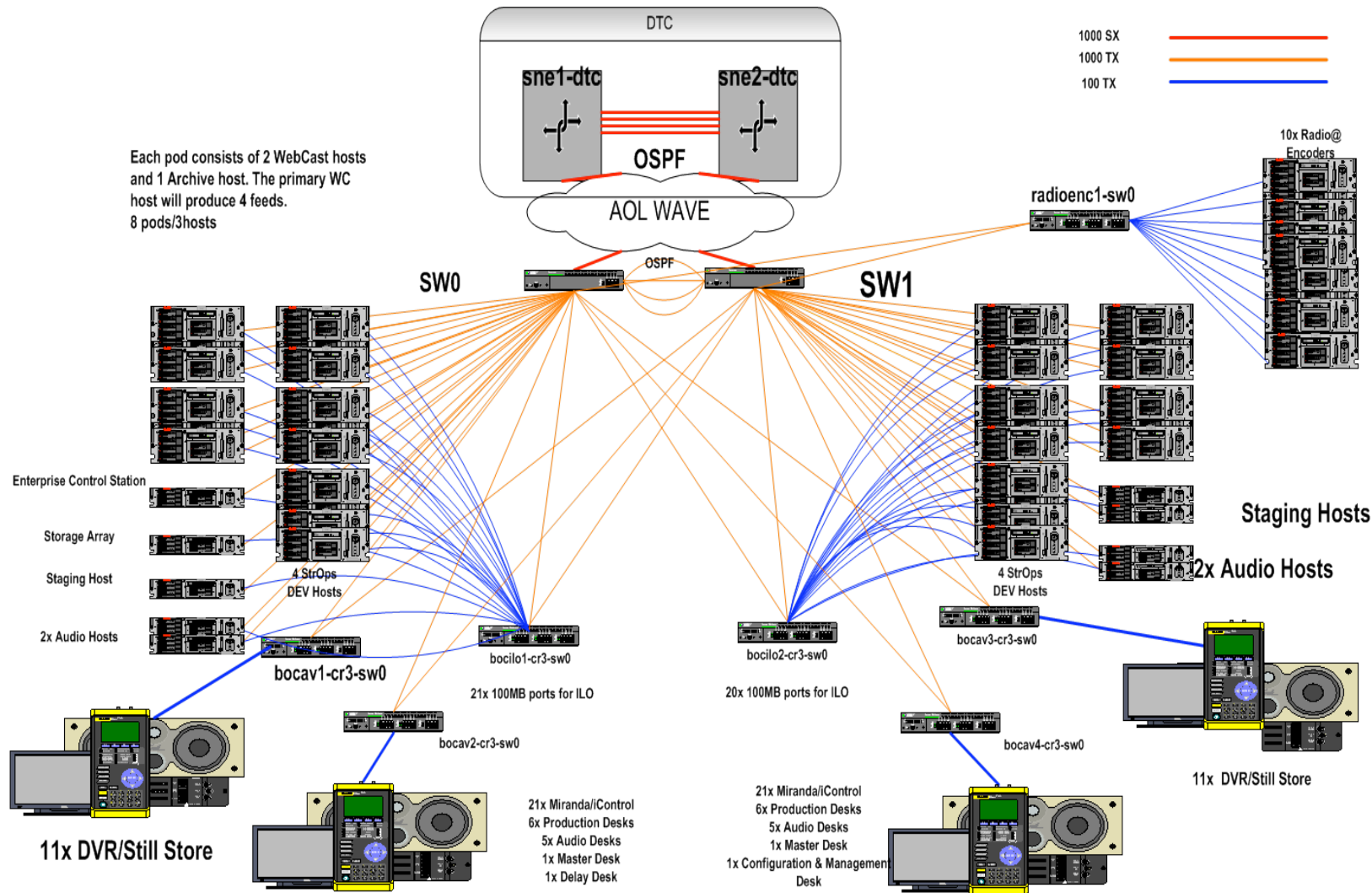
There comes a time in one's life, when you must take a risk, or live forever with your dreams. - Late Extreme Skier Trevor Petersen

Questions

- Many thanks to Matt Schechtman, John Watson, James Tarry, Daniel Shames, Brenda Rapp, Pat Gunn, Mark Muehl, Teresa Dietrich-Gurney