Network Algorithmics: Making the Internet Faster

George Varghese,
UCLA

Algorithmics to speed abstractions in all systems (not just networks)

- Example 1: Virtual Memory:
  - *Abstraction*: Illusion of infinite memory
  - *Algorithmics:* Paging Algorithms
- Example 2: Relational Databases
  - *Abstraction:* Operations on Logical tables
  - *Algorithmics:* Query Planning

How did algorithmics start

- Context: Web exploding, traffic doubling, address doubling in 1990s
- Problem: TCP (connected queues) and IP (datagram) slow, as were routers & servers
- Network Algorithmics: techniques to restore speed of abstractions to that of fiber.
- This lecture: trying to give you a sense of what this field is/was about

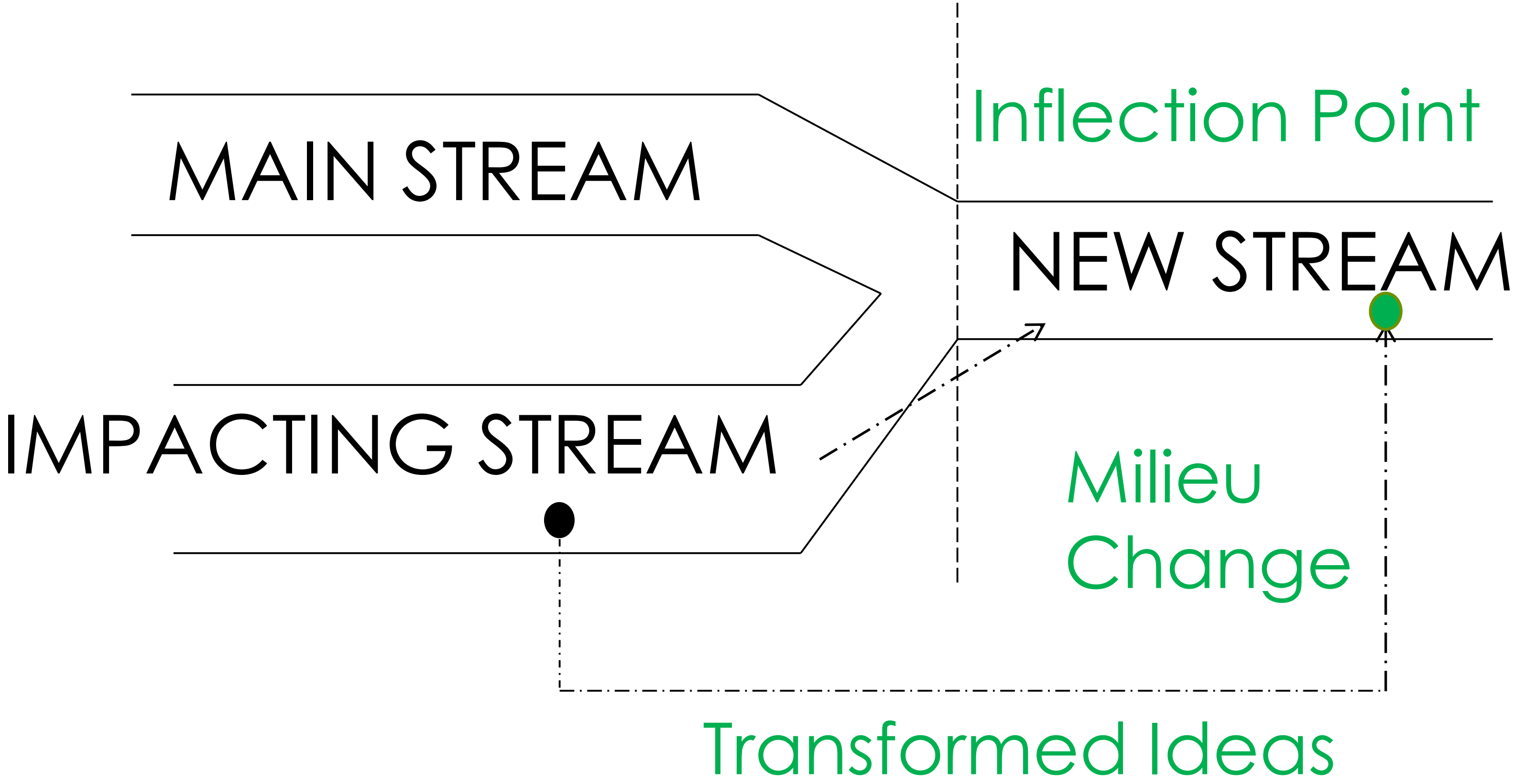
Outline

- Confluences as a way of thinking of inter-disciplinary research.
- Network Algorithmics viewed as confluences from lookups to logging
- Ideas for creating yourself

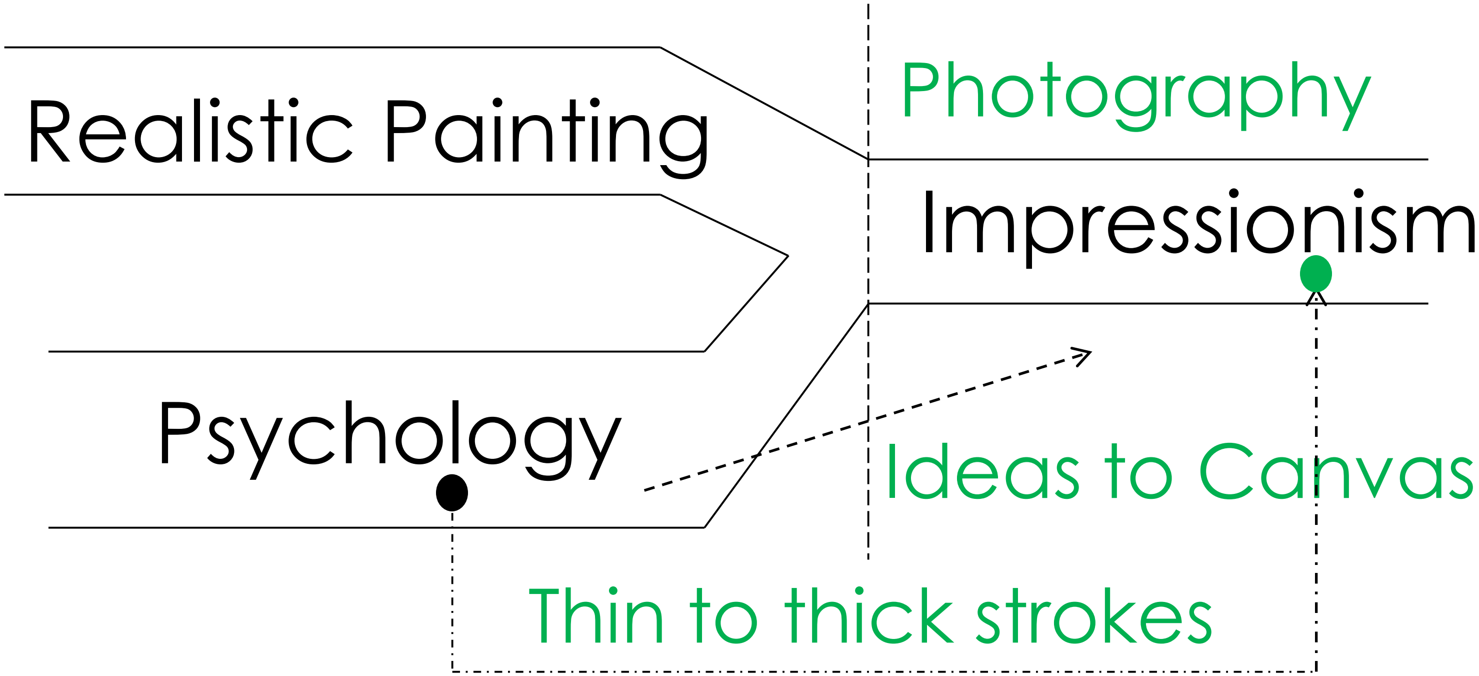
What is a Confluence?

CONFLUENCE: Where Two Rivers meet

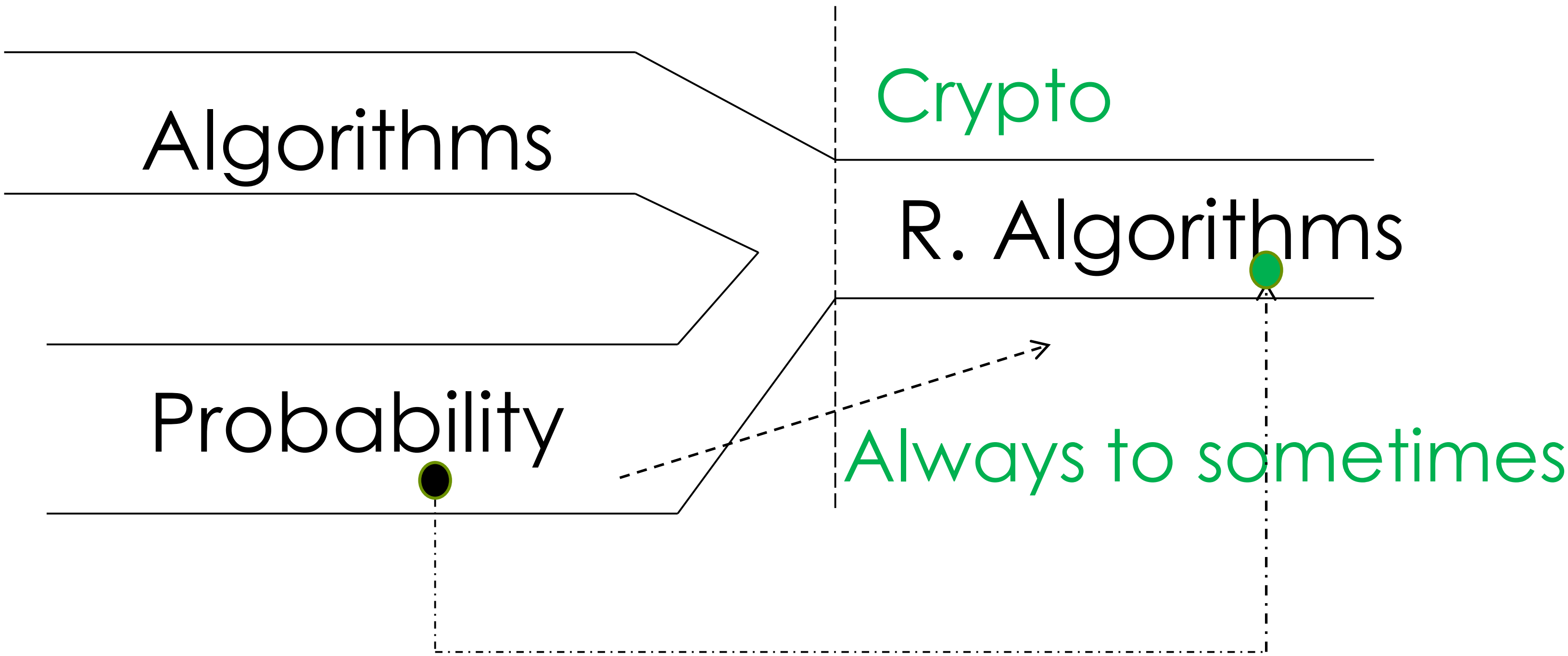
Confluence Definition



Example 1: Impressionism



Example 2: Randomized Algorithms



Sieve of Eratosthenes to Miller-Rabin

Algorithm	TIme on $10^{100} + 267$
Miller-Rabin (100 trials)	0.3 seconds
Best Deterministic (AKS)	37 weeks

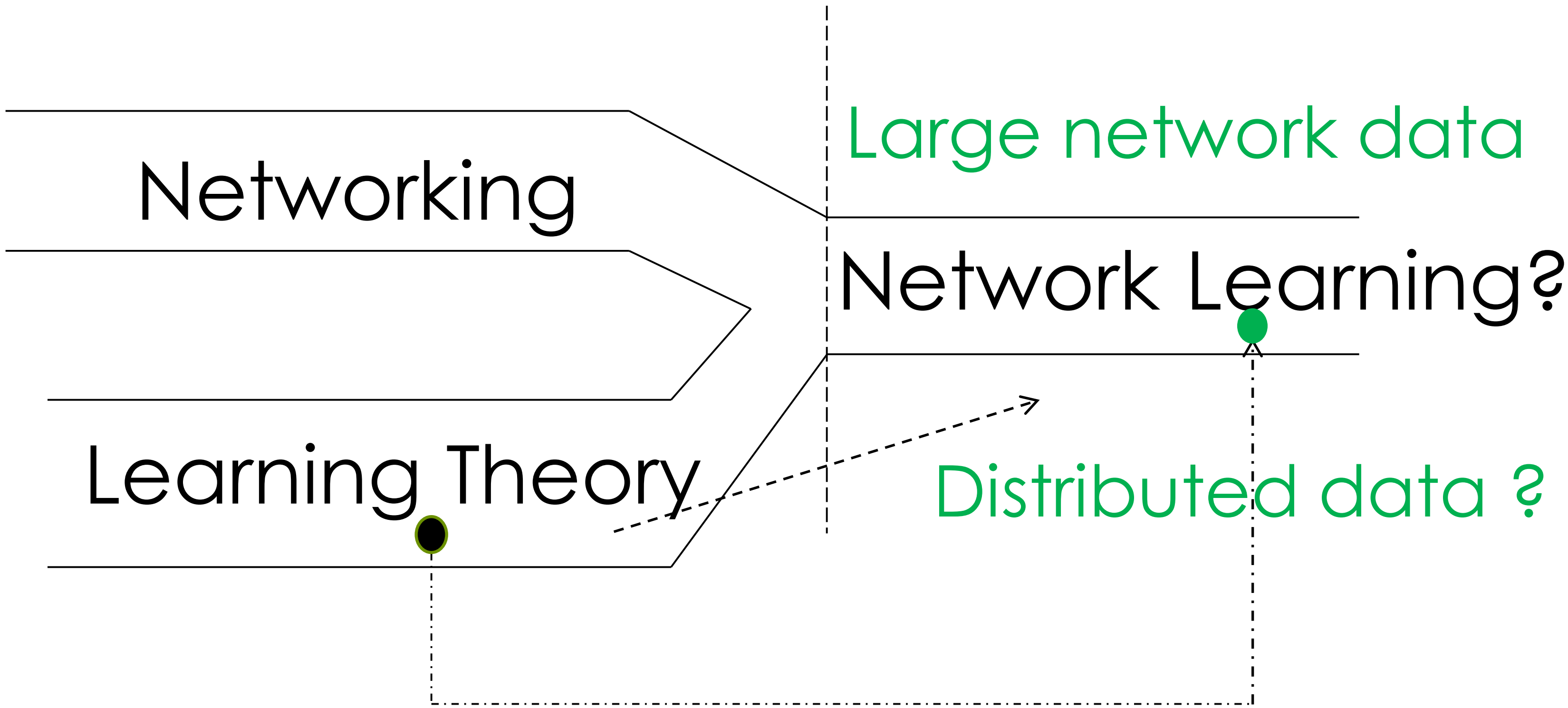
More Computer Science examples

- Distributed Algorithms
  - *Streams:* Algorithms, Networks
  - *Inflection Point:* Popularity of Internet
  - *Mileu Change*: Asynchrony, partial failure
- Computational Economics
  - *Streams:* Economics, Computer Science
  - *Inflection Point:* Internet Auctions
  - *Milieu Change*: *L*arge scale, small latency

Why Confluences?

- Separate trends from fads
- Provide a research theme
- Balance desire for beauty *and* impact
- Suggest a new field in making, especially when the original field has matured

All Interdisciplinary work ≠ confluence



What concept has changed?

Network Algorithmics via Confluences

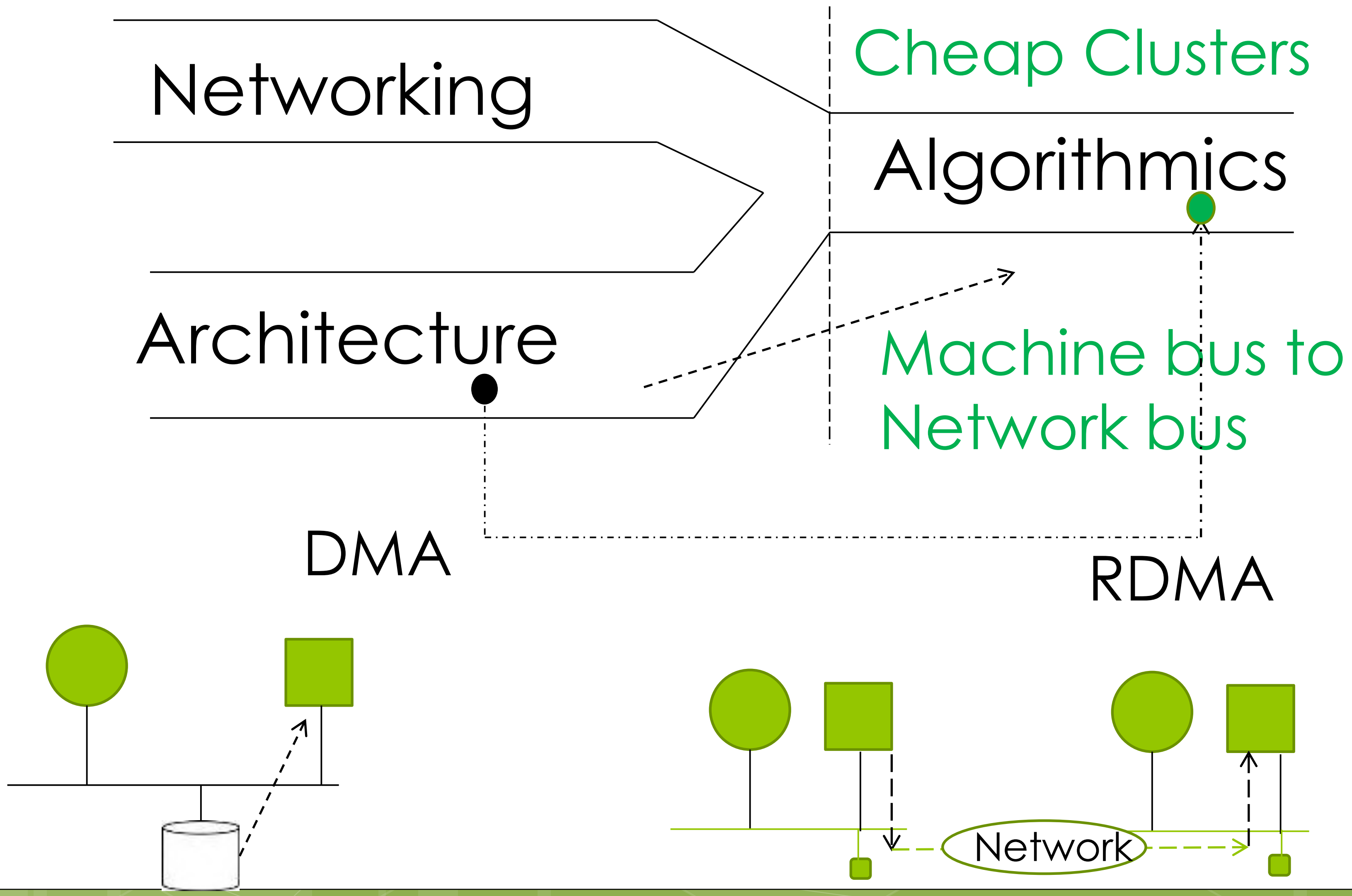
Network Algorithmics

- *Definition:* A set of algorithms to make network devices (servers, routers) faster
- *Streams:* A confluence between computer architecture, algorithms & networking
- *Inflection Points*: cheap clusters (servers), explosion of web, IPv6 (for routers
- *Two examples of transformed ideas*: RDMA and Binary Search on prefix lengths. Many more

1. RDMA Example (Servers)

- Servers getting slow as they used remote disks. READs from disk to memory slow
- Inspiration from computer architecture: DMA from disk to memory, bypass CPU
- Do same thing across network (RDMA). Needs lots of new ideas
- RDMA a major force in data centers of Google, Microsoft to make servers fast.

Example 1: RDMA [KSL 86]



DESIGN

Is RDMA the Future of Data Center Storage Fabrics?

Inside the CERN data center in Meyrin, Switzerland, 2017

Big data analytics, ever larger databases, and dense workload consolidation are driving the rise of RDMA.

Mary Branscombe | Oct 04, 2017

DataCenter Knowledge

https://www.datacenterknowledge.com/design/rdma-future-data-center-storage-fabrics

From RDMA to Fast Servers: Other transformed ideas

Inflection point: Internet heating up (90s)

- Fast Buffers: Avoid packet copies without changing protocol → 0 copy interfaces
- Application Device Channels → Bypass OS and go direct from App to Network
- Header Prediction → Predict that next packet will be in order to make common case of TCP processing fast

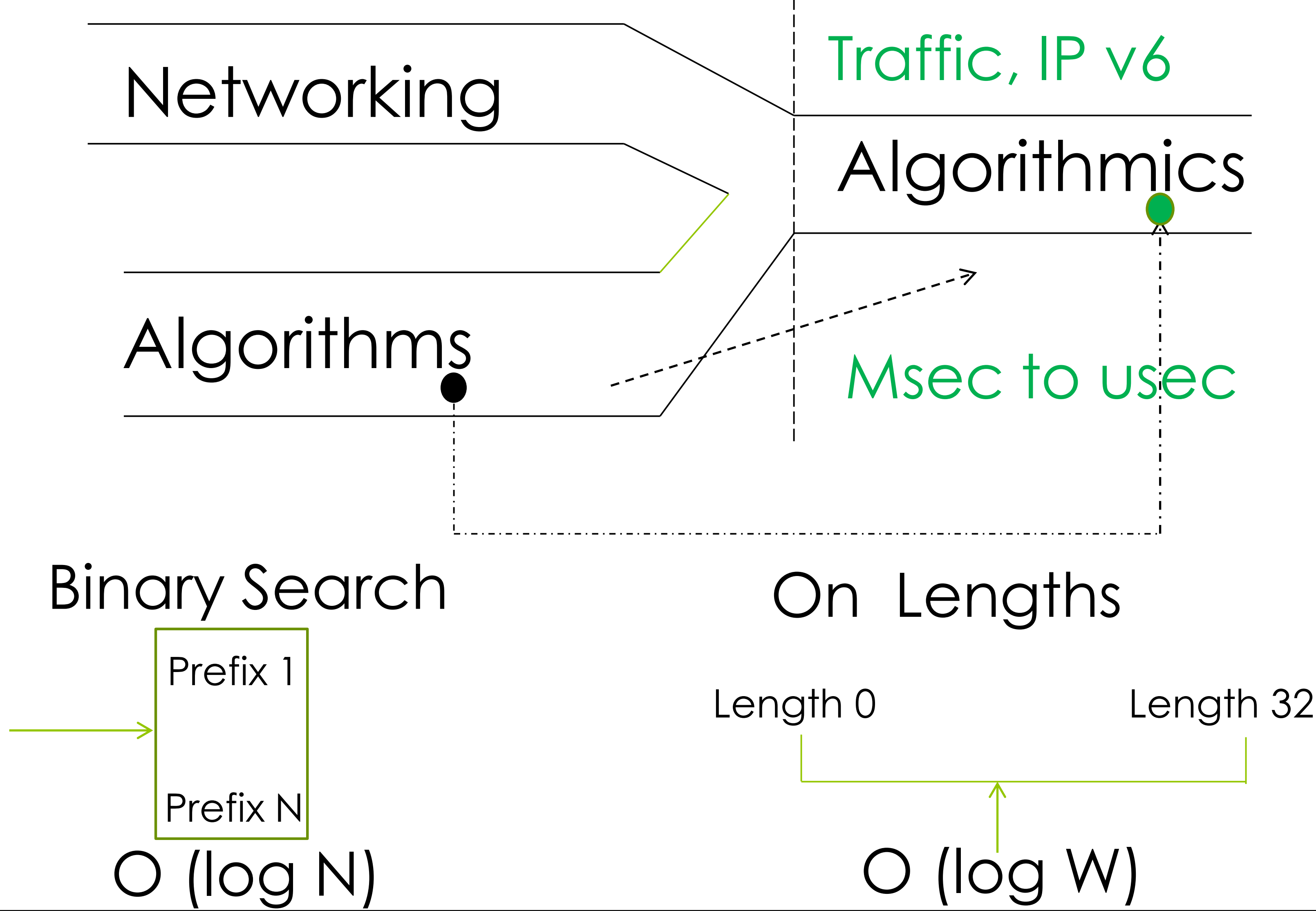
2. Prefix Lookup (Routers)

- Your Facebook feed is a set of Internet messages carrying a 32 bit IP address
- Routers send your message onwards by looking up a forwarding table
- To avoid a table of a billion addresses in each router, routers store prefixes
- This reduces router memory but routers now have to solve longest prefix match for every packet billions of times per second! Even harder for *IPv6 with 128 bit* addresses. How to cope?

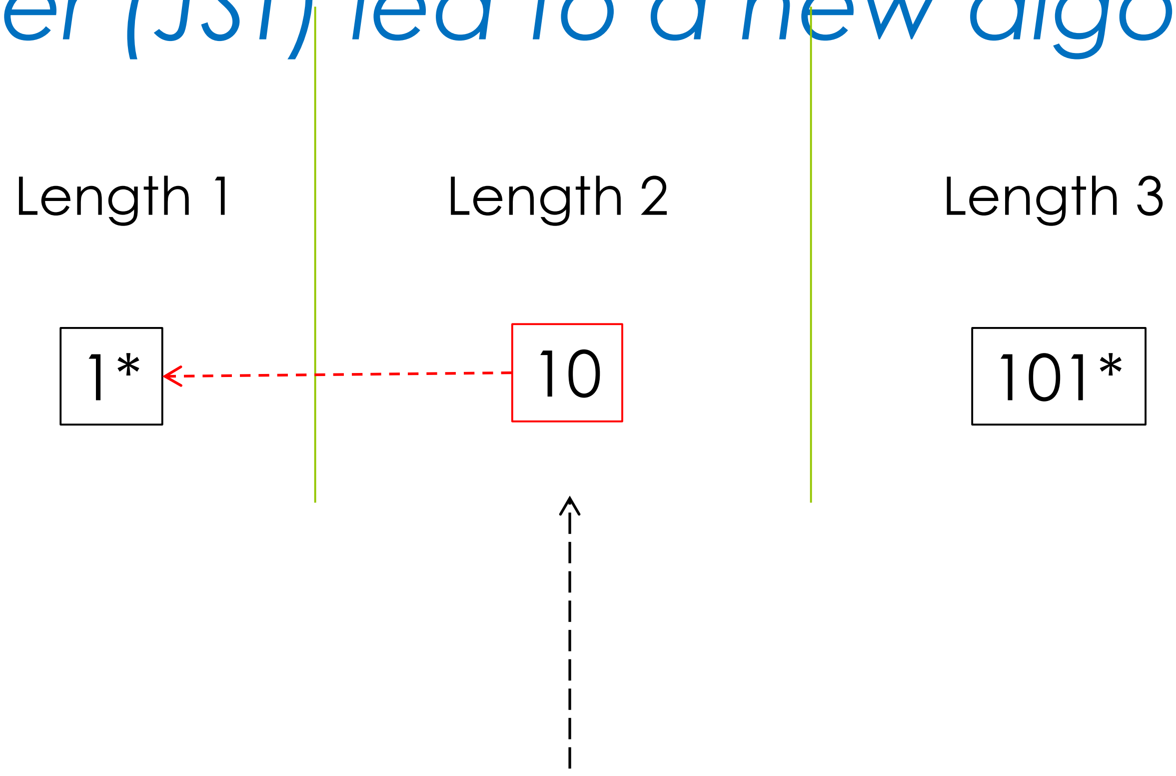
Prefix Lookup Details

- A prefix is the start bits common to lots of computers in an organization like UCLA
- Core routers don't store all addresses in UCLA: just a few prefixes (see kb.ucla.edu)
- One UCLA prefix: 128.97.0.0/16
- Write each number between dots in binary and take first 16 bits →
- 10000000 01100001 *, where * is don't care

Example 2: IP Lookup [WVTP 97]



Binary Search on prefix *lengths: how a* ***collision*** *with an outsider Jon Turner (JST) led to a new algorithm*

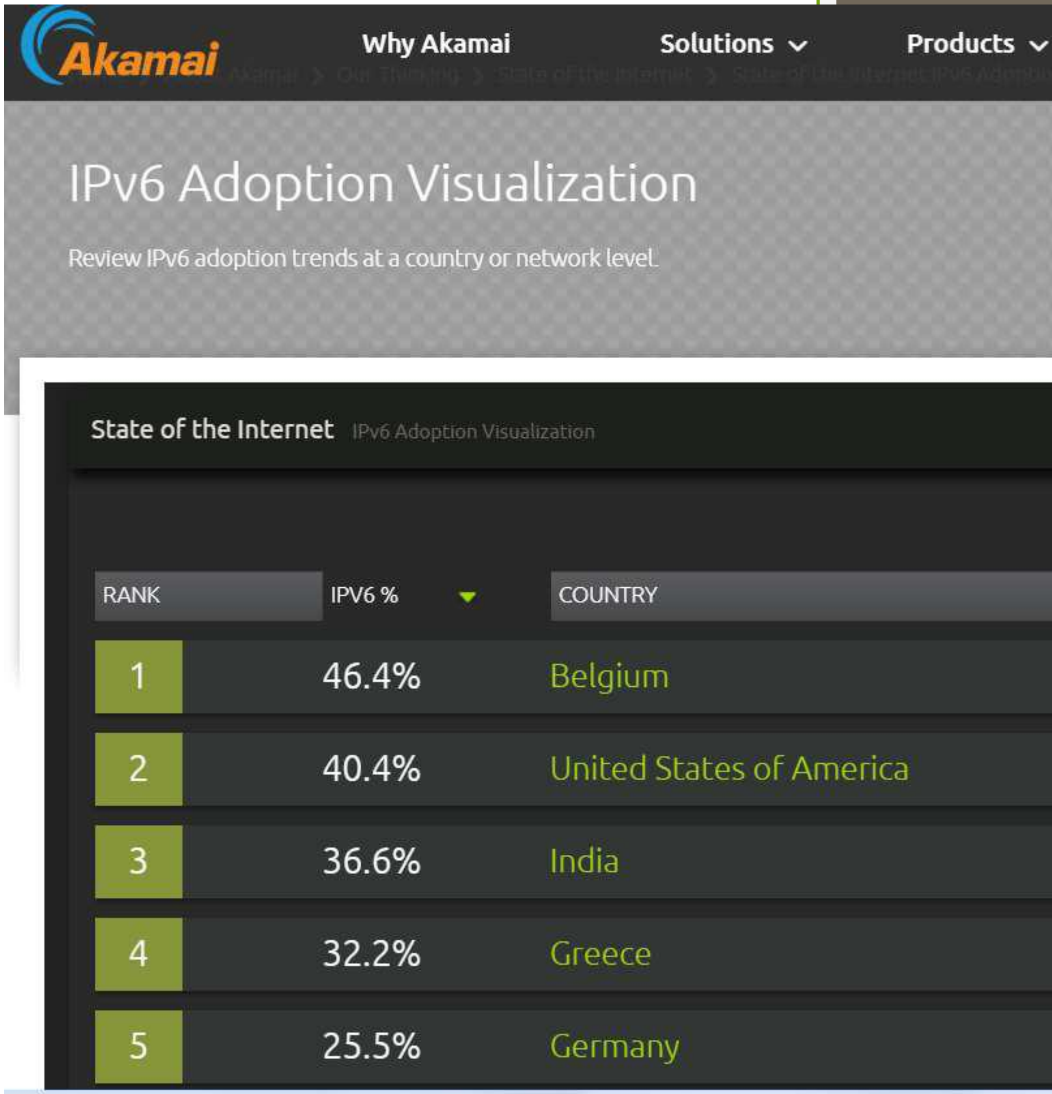


D. D. & GV, JST: Bug: For binary search start in middle marker

Day 2: JST, Doesn't work, Oh, just add markers

Algorithm Summary

- Arrange prefixes by lengths. Start in middle length table. If you get a match go to right half else go to left half
- To make it work, have to add dummy prefixes as *markers*. A 32 bit prefix leaves markers in the Length 16 table, length 24, length 28 table etc.
- Markers can cause us to go on a wild goose chase to right half when answer is in left, So we *precompute* longest match of every marker and remember that when we go to right.
- **Bottom Line**: only log (128) = 7 accesses to tables even for 128 bit 1P addresses. Not classical binary search



Akamai Why Akamai Solutions Products

IPv6 Adoption Visualization

Review IPv6 adoption trends at a country or network level.

State of the Internet IPv6 Adoption Visualization

RANK	IPV6 %	COUNTRY
1	46.4%	Belgium
2	40.4%	United States of America
3	36.6%	India
4	32.2%	Greece
5	25.5%	Germany

https://www.akamai.com/us/en/about/our-thinking/state-of-the-internet-report/state-of-the-internet-ipv6-adoption-visualization.jsp

Fast Routers common by 2000s

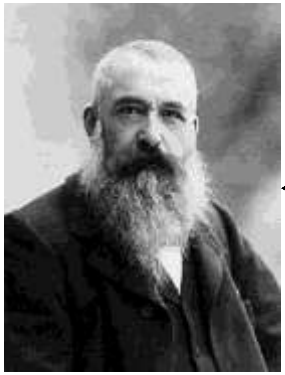
- Cisco Cat 6K, GSR, Juniper M40
- All the problems (switching, lookups, ACLs, scheduling) had reasonable hardware
- Solutions scaled as link speeds scaled
- New Algorithmics problems after 2000 in measurement & security (e.g., catching worm attacks) & *flexible* routers (P4)

Create yourself

What should CS at UCLA give you?

- You will certainly learn to code in CS 31. 32. You will learn algorithms
- Just these will get you a job at Google or Facebook if you study for the interview
- You will also learn about internals (OS, networks, architecture, databases)
- But as part of an amazing UCLA freshman batch, I hope you will learn to create something new, an idea or a product. How?

1. Embrace Collisions with people with new perspectives



Monet

Paris
1860

Renoir

Impressionism

Dyson

Princeton
1973

Montgomery

Confluence:
Number Theory
with Physics

Rabin

MIT, 1975

Miller

Randomized
Algorithms

2. Pick your problem carefully

- Watch for Trends (Top Down)
  - Read Trade Rags
  - Listen to Grapevine
  - Talk to others (teenagers, kids)
- Know your Strengths (Bottom up)
  - Collaborators
  - Personal skill set
  - Access to Data (secret weapons)

3. Keep thinking via an Idle Loop of Old Important Problems

Keep thinking of older problems in background as one learns new techniques

- *Synchronize routers after partition heals*
  → Set Difference algorithm 10 years later
- *Bridge Learning via sending SYSIDs*
  → Carousel solution 10 years later

4. Be *contrarian* in picking problems

Examples I have worked on:

- Industry felt route lookups too slow

  Fast IP Lookups common today
- Academics felt packets had to be sorted by deadlines to be fair across conversations

  Cheap modification of Round Robin works (DRR)
- Humans must produce attack signatures

  Automated signature extraction.

But balance risk . . .

- *Analogy from Football:* Don't just throw long balls, run the football occasionally.
- *Analogy from Finance*: Balance your portfolio. Buttress your stocks with bonds.
- *Similarly*: keep at least one risky problem but add safer ideas. Have to make a living!

Conclusions

- *Confluences:* Look for the inflection point and the new milieu before jumping in
- *Network Algorithmics:* Not same as algorithms, different measures & models
- *Differences:* Sometimes (RDMA) we change OS structure; in longest match we sacrifice insertion time (msec) for fast search (nsec)
- *Create Yourself:* by embracing collisions, being contrarian, and having an idle loop

5. Finally, avoid extremes

Influenza, commonly known as "the **flu**", is an infectious disease common among mammals. The most common symptoms are chills and fever.

Confluenza, commonly known as "the **conflu**", is an infectious disease unique to researchers. The most common symptoms is excessive preoccupation with finding confluences in every aspect of life.

Get your conflu shot today Thank you!