Université de Liège

SDN++ : Beyond programmable plumbing

Laurent Mathy
University of Liège, Belgium

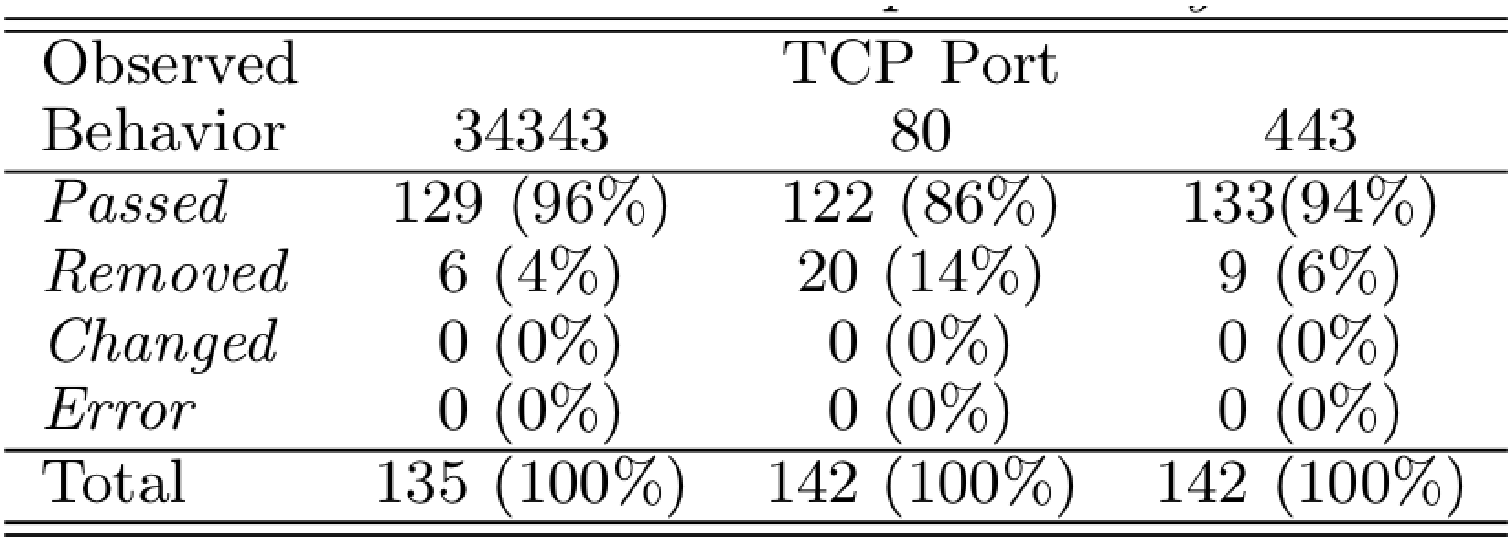
Global Future Internet Summit 2013, Seoul, Korea
- Beyond the Horizon -

What's really wrong with the current Internet

What actually happens to TCP in the wild?

- IMC 2011 paper
- 142 access networks in 24 countries.
- Tests to measure what actually happened to TCP.
  - Are new options actually permitted?
  - Does re-segmentation occur in the network?
  - Are sequence numbers modified?
  - Do middleboxes proactively ack?

Middleboxes and new TCP Options in SYN



Observed	TCP Port		
Behavior	34343	80	443
Passed	129 (96%)	122 (86%)	133(94%)
Removed	6 (4%)	20 (14%)	9 (6%)
Changed	0 (0%)	0 (0%)	0 (0%)
Error	0 (0%)	0 (0%)	0 (0%)
Total	135 (100%)	142 (100%)	142 (100%)

- Middleboxes that remove unknown options are not so rare, especially on port 80

What actually happens to TCP in the wild?

- Rewrote sequence numbers:
  - 10% of paths (18% on port 80)
- Resegmented data:
  - 3% of paths (13% on port 80)
- Proxy Ack:
  - 3% of paths (7% on port 80)
- Ack data not sent:
  - 26% of paths (33% on port 80) do strange things if you send an ack for data not yet sent.

Not to mention…

- NAT
  - Pretty nearly ubiquitous, but comparatively benign
- DPI-driven rate limiters
- Lawful intercept equipment
- Application optimizers
- Anything at the server end:
  - Firewalls
  - Reverse proxies
  - Server load balancers
  - Traffic scrubbers
  - Normalizers, etc

TCP option work will not detect most of these.

HotNets 2011 paper reports 600+ middleboxes for 900 routers in a typical enterprise net

Middleboxes

- NATs are ubiquitous
  - We've become pretty good at working around them.

- Firewalls are ubiquitous
  - Ability to communicate using one port does not imply that communication is possible on any other port.

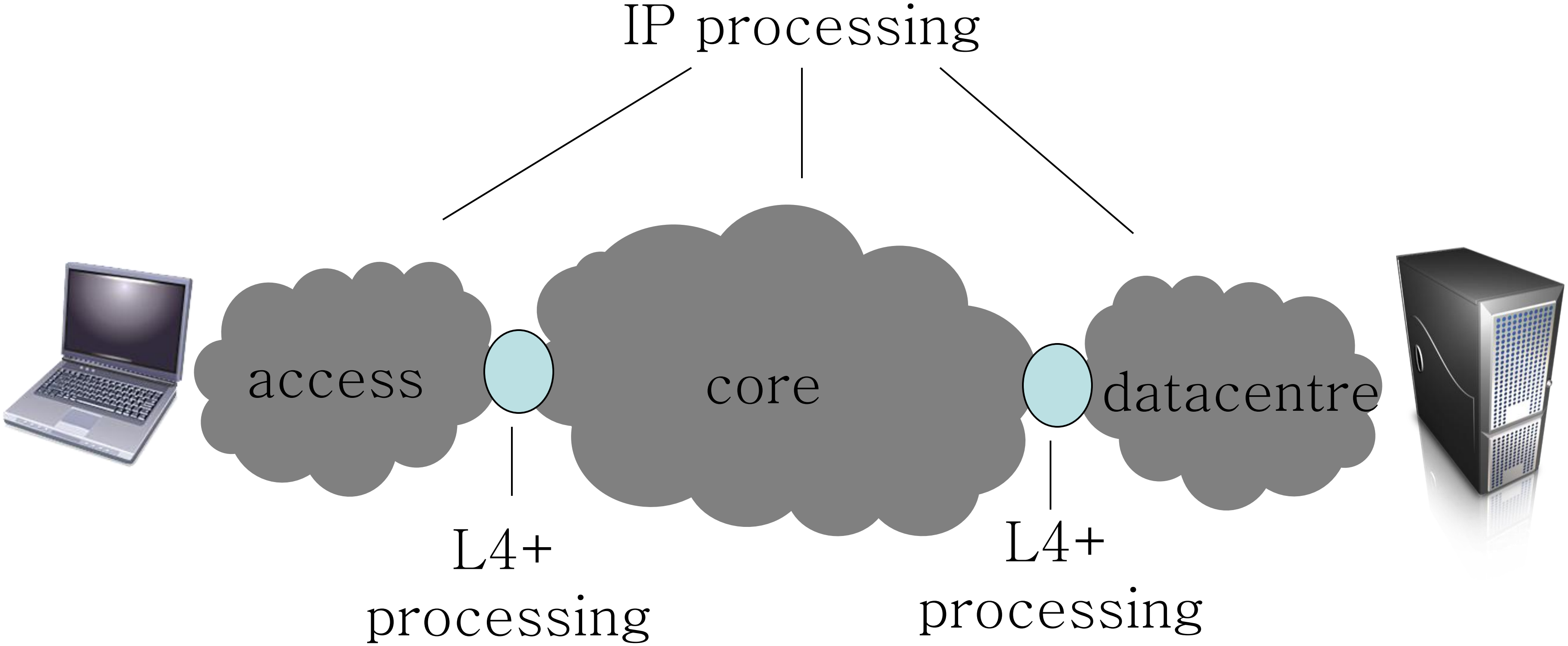
Middleboxes everywhere

- Plenty of box vendors will sell you a solution.
  - Whatever you think your problem is.

- Current apps get optimized and set in silicon.
  - Application Entrenchment
- Future apps tunnelled over HTTP
  - (but what do all those port 80 specialized middleboxes do?)

- Impossible to reason about the concatenation of middleboxes.

A segmented Internet



Observation: The Internet is becoming a concatenation of IP networks interconnected by L4+ functionality.

Why are middleboxes everywhere?

- Why?
  - Packets are an artifact of the network
  - As soon as you reason about applications, you think in terms of flows

Going with the flow…

- Currently flow processing in middleboxes serves to inhibit new applications.
  - Optimization of the present
  - Inextensible inflexible network security

- But middleboxes are there for a purpose
  - They are not going away any time soon

- Key question: is it possible to re-claim the middlebox as a force for enabling end-to-end innovation?

What can we do about it?

- Those L4+ platforms need to be more general than today's middleboxes.
  - More open and explicit
  - More upgradable, as new apps arrive.
    - Programmable
  - Aggregate functionality, so it is manageable.
  - Identifiable, so we can reason about them
  - Cheap and scalable.
- This is the essence of Software-defined networking
  - OpenFlow is SDN at layer 2

The End-to-Middle-to-End Principle

- The End-to-End Principle
  - Application specific functions should reside in the end- hosts of a network rather than the intermediary nodes, provided they can be implemented “completely and correctly” in the end hosts.
  - Essentially this is a recipe for enabling application innovation.
    - » But it only works if the network operator really doesn’t care about which applications are running.
    - » Security, performance, legal requirements, the NSA are some reasons they do in practice care.
- The End-to-Middle-to-End Principle
  - When application-specific functions are placed in the intermediary nodes, it must be possible to reason about the emergent behaviour.

Programmable Data path

- Winning the Future Internet battle is not primarily about building a better middlebox.
  - Though much of the effort must go on this.
- Programmability is key to decoupling infrastructure and functionality
- Virtualization is key to decoupling infrastructure ownership and functionality ownership
- But programmability/virtualization and performance usually don't go together
  - What should a high performance programmable network box look like?

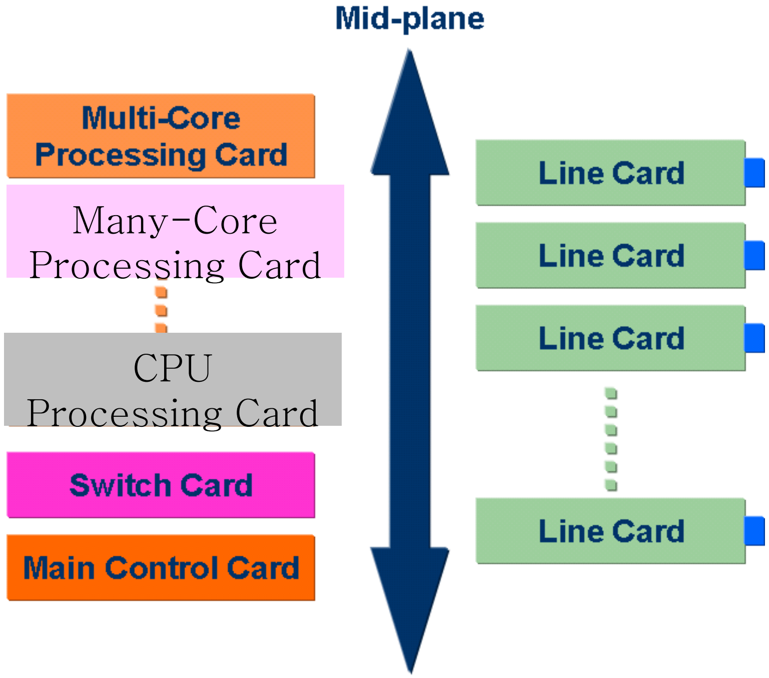
Programmable Data Path (2)

- But programmability/virtualization and performance usually don't go together
  - Network Function Virtualization (NFV) currently runs on commodity servers
    - Fairly poor traffic aggregate rates
    - Low port density
    - Wrong system assumptions (context switching, etc)

➔ What should a high performance programmable network box look like?

Heterogeneous Hardware Systems

- Get the best of all worlds (PEARL2@ICT/CAS)



Mid-plane

Multi-Core Processing Card

Many-Core Processing Card

CPU Processing Card

Switch Card

Main Control Card

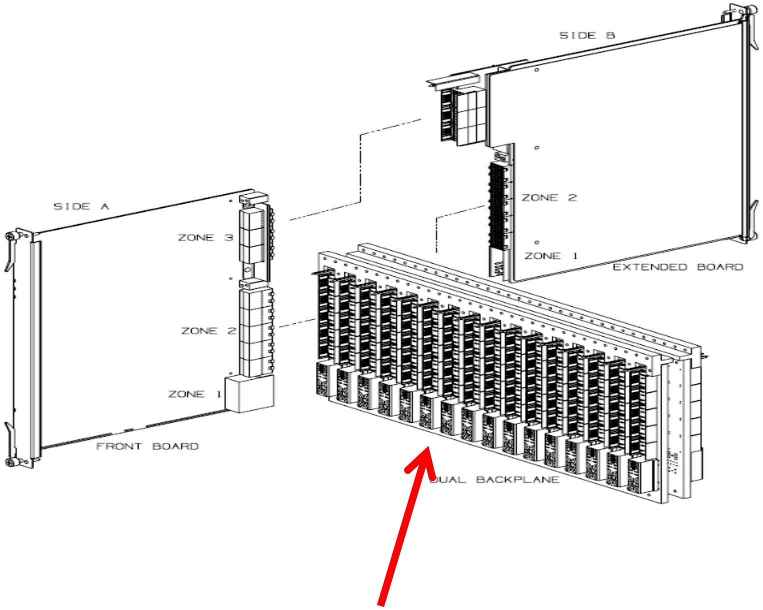
Line Card

Line Card

Line Card

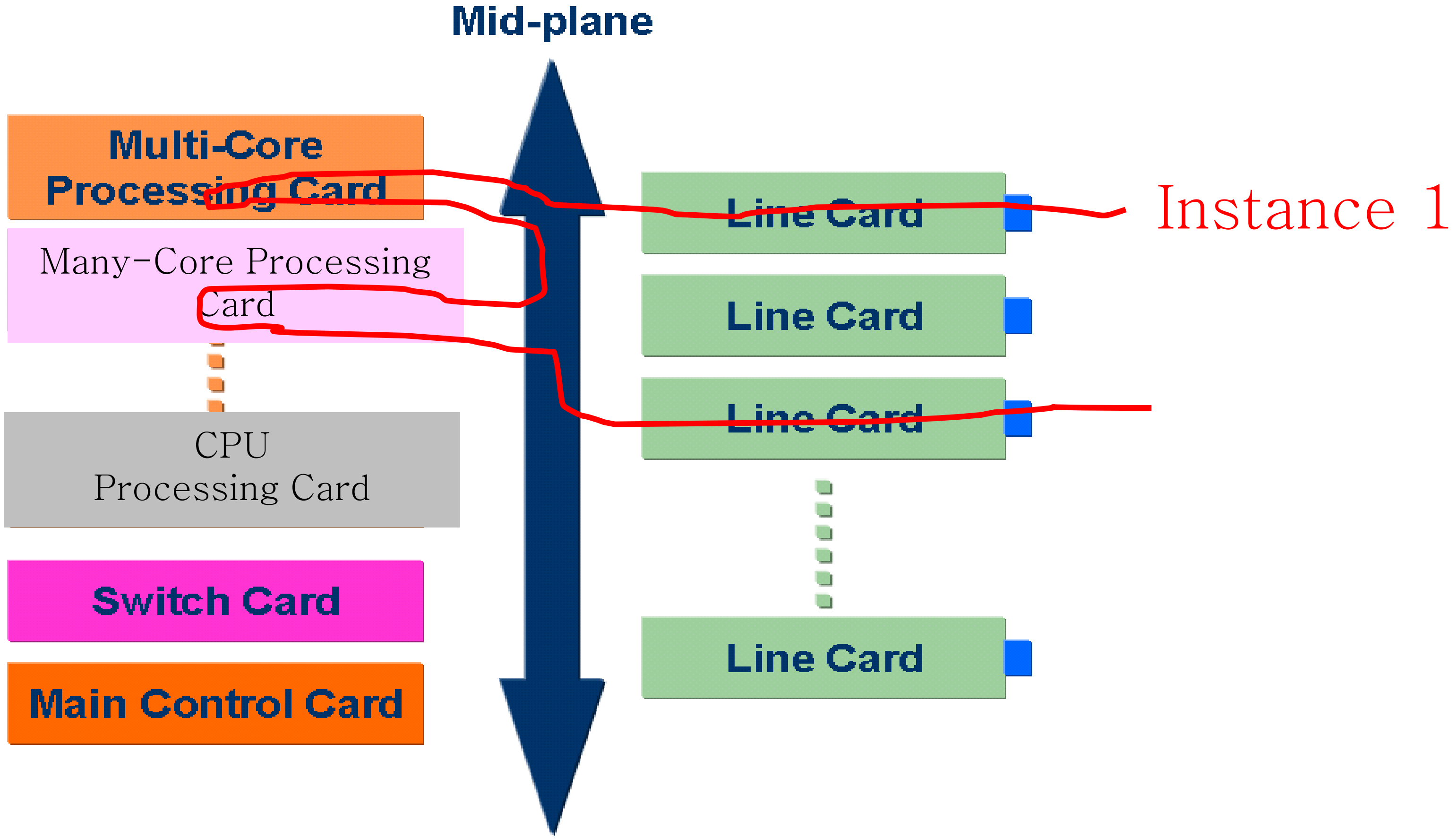
Line Card

TCAMs, FPGAs, GPUs, NPs, storage, etc

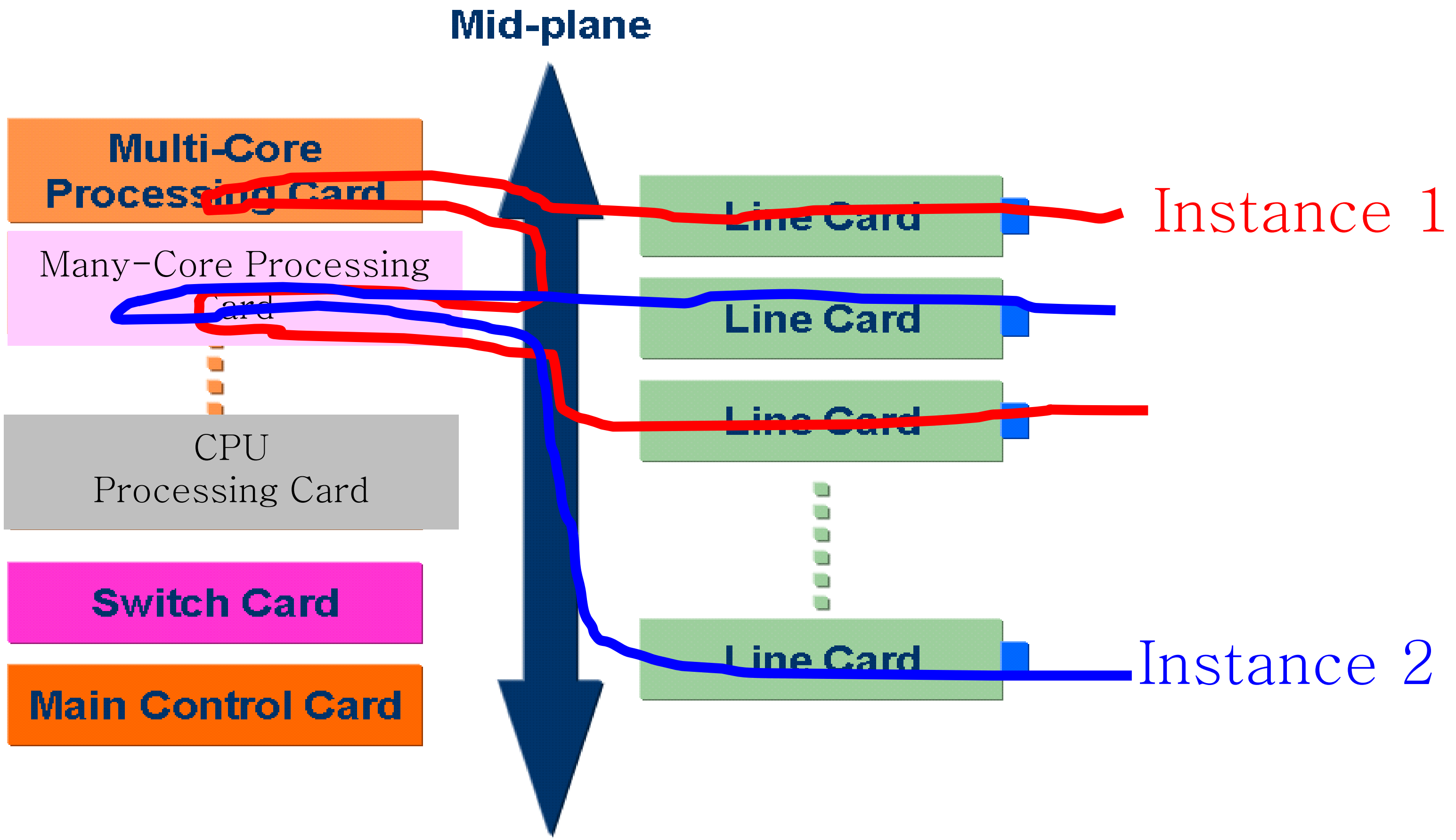


Next generation ATCA architec

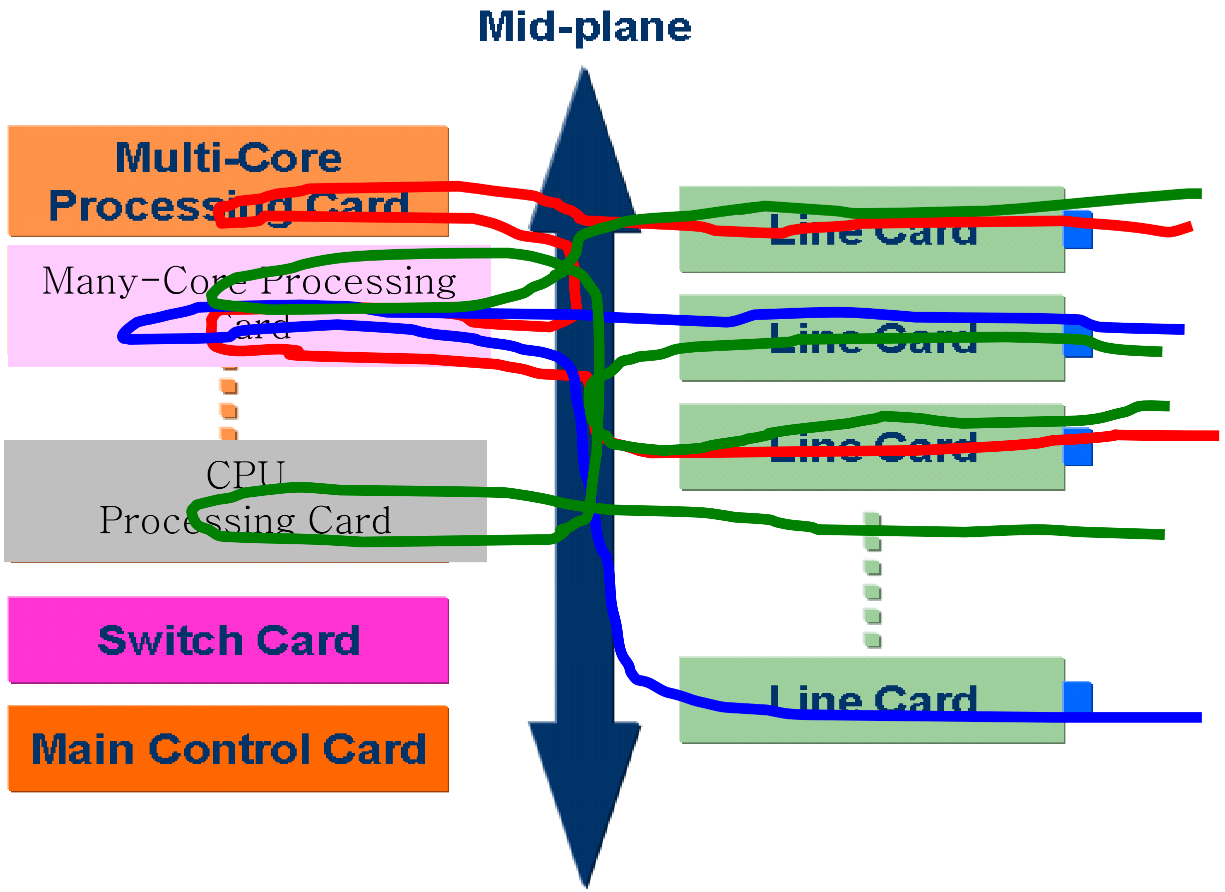
Middlebox instances



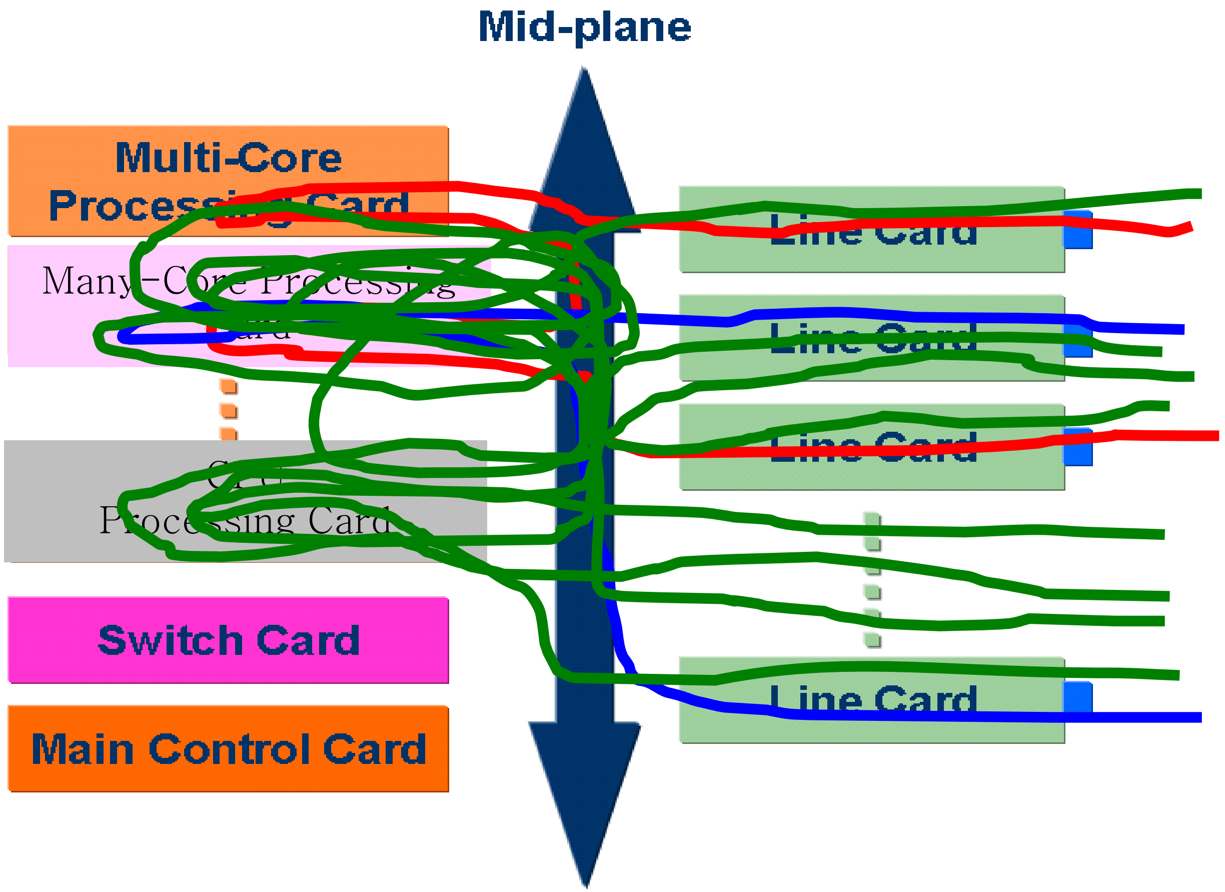
Middlebox instances



Middlebox instances



Middlebox instances



Hardware is not enough

- Need programming abstraction
  - Hide nitty-gritty details
    - » Distributed memory modules with independent address spaces
    - » No cache coherence
    - » Multiple instruction sets
  - Write once, run everywhere
    - » JIT
  - Crucial for adoption

Hardware is not enough

- Run-time/OS
  - Instantiate high-level network processing applications
    - » Allocates network processing elements to HW components
    - » Small change in configuration can result in big performance swings
  - Performance optimization
  - Virtualization
  - Element migration

Conclusions

- This is active/programmable networks
  - Kind of, but from a systems perspective
  - I prefer "programmable network infrastructure"
- Asking what a "general purpose" programmable network processing environment should look like
- Heterogeneous systems are a nightmare, but hope as focusing on networking only
- A bunch of those boxes and tunnels (OF interconnect?)
  - incremental deployment of new Internet