Hadoop을 이용한 인터넷 측정 데이터 분석

2012-07-26

충남대학교
이영석
lee@cnu.ac.kr
http://networks.cnu.ac.kr

내용

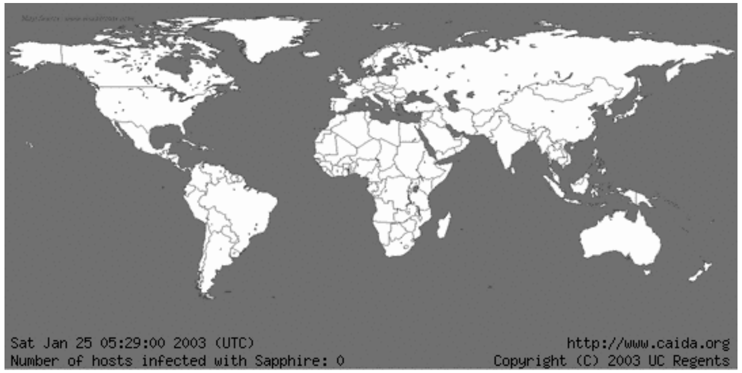
- 인터넷 측정 (Internet Measurement) ?
- 왜 Hadoop ?
- Hadoop 기반 인터넷 측정 및 분석시스템
- 결론

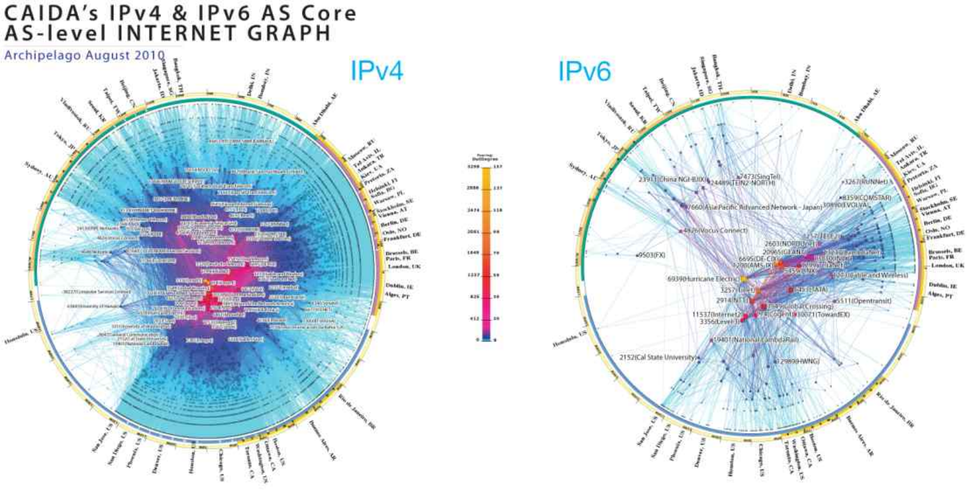
인터넷 측정

- 주요 측정 대상
  - 링크
    - 패킷, 플로우(Cisco NetFlow)
  - 서버
    - 로그, SNMP MIB
  - 라우터
    - 라우팅 테이블(RIB, updates), SNMP MIB
  - 단말/사용자
    - 공개/동의 사용자 데이터(트위터, 위치, 웹, 블로그, 검색, 쇼핑, 이메일)
  - 종단간 측정
    - Ping, traceroute, skitter
- 응용
  - 보안
    - firewall (snort, bro), IDS/IPS, DDoS
  - 네트워크 관리
    - Trouble shouting, performance management, capacity planning/traffic engineering
  - 학문적 연구
    - 트래픽 모델링(poisson vs. self-similarity)
    - 인터넷 네트워크 토폴로지 분석
  - 인터넷 센서쉽
    - 이집트, 리비아

인터넷 측정연구를 통한 발견

- 웜 전파
  - CAIDA
- 인터넷 토폴로지
- 새로운 인터넷 구조
- 인터넷 검열

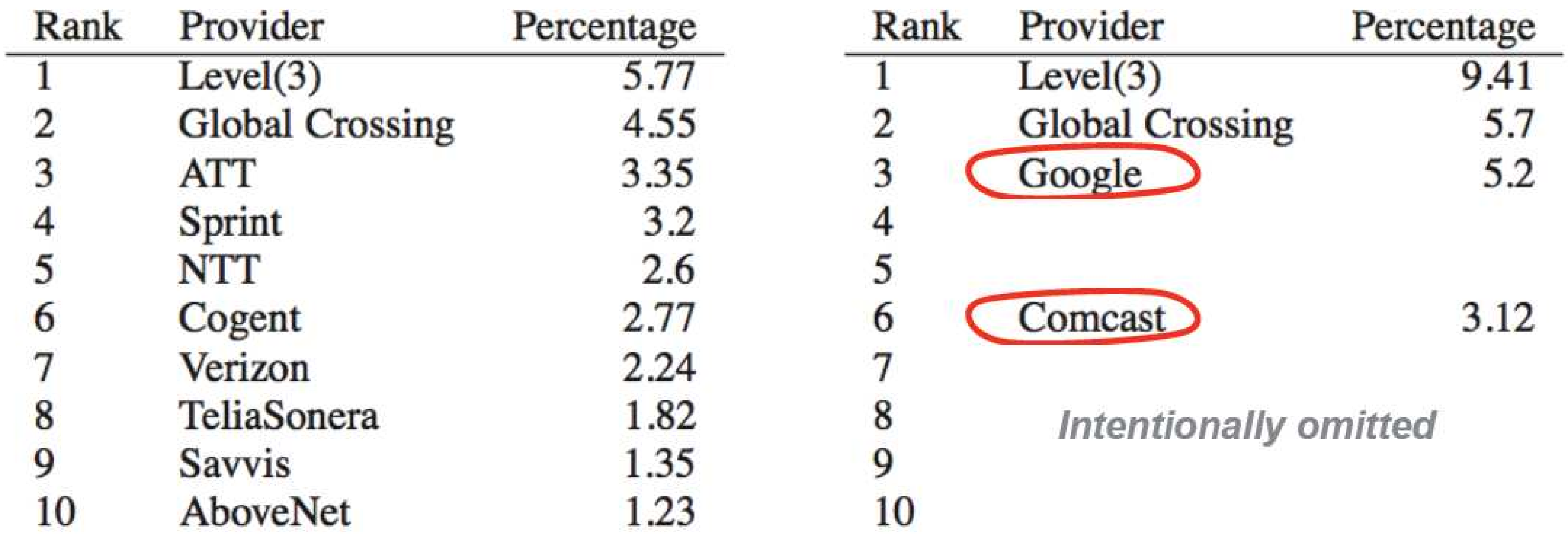

Sat Jan 25 05:29:00 2003 (UTC)
Number of hosts infected with Sapphire: 0
http://www.caida.org
Copyright (C) 2003 UC Regents



copyright © 2010 UC Regents. all rights reserved.

2009 Internet Observatory Report

- By Arbor Networks
  - C. Labovitz et al., "Internet Inter-Domain Traffic," ACM SIGCOMM2010
  - 110+ ISPs / Content Providers



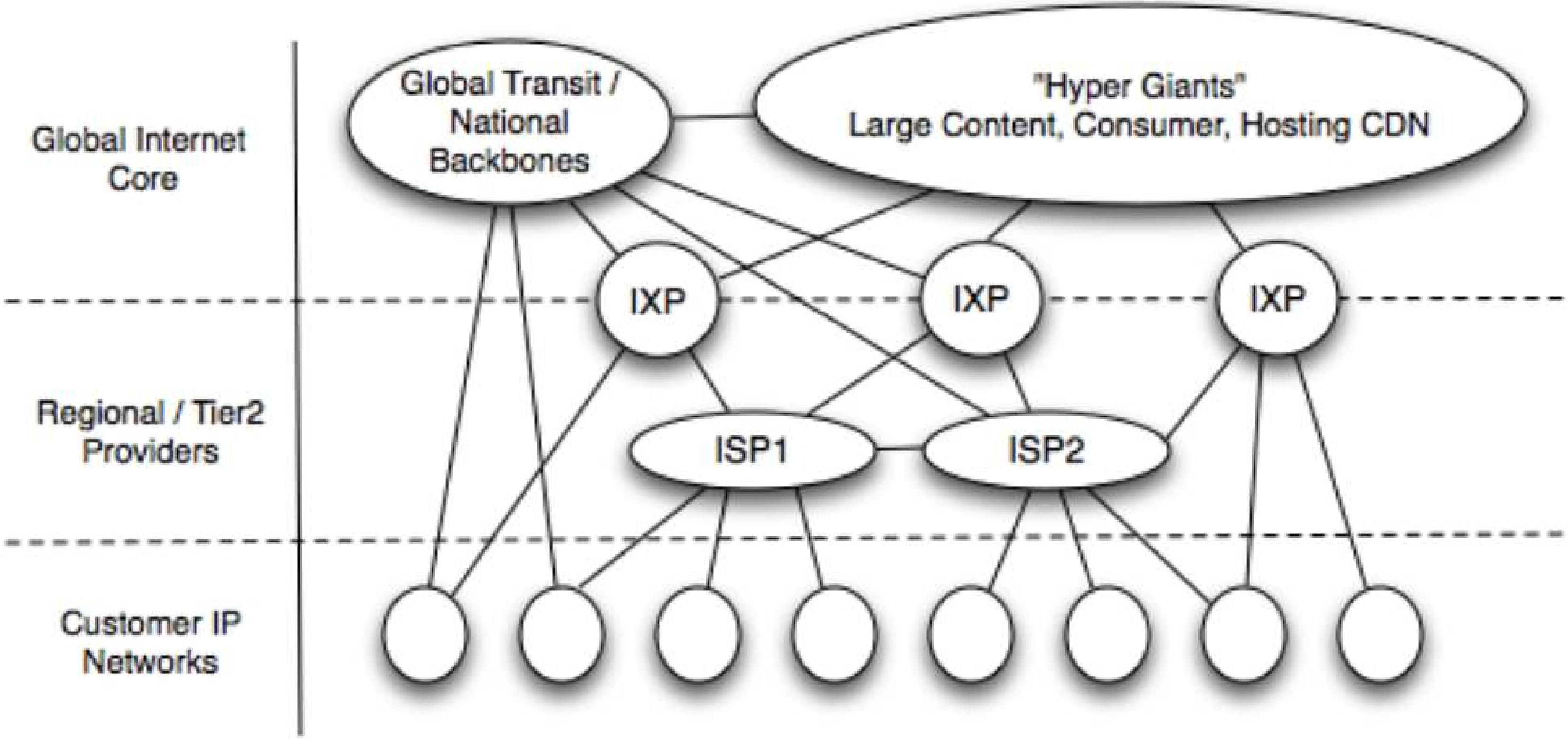
Rank	Provider	Percentage
1	Level(3)	5.77
2	Global Crossing	4.55
3	ATT	3.35
4	Sprint	3.2
5	NTT	2.6
6	Cogent	2.77
7	Verizon	2.24
8	TeliaSonera	1.82
9	Savvis	1.35
10	AboveNet	1.23

(a) Top Ten 2007

Rank	Provider	Percentage
1	Level(3)	9.41
2	Global Crossing	5.7
3	Google	5.2
4		
5		
6	Comcast	3.12
7		
8	*Intentionally omitted*	
9		
10		

(b) Top Ten 2009

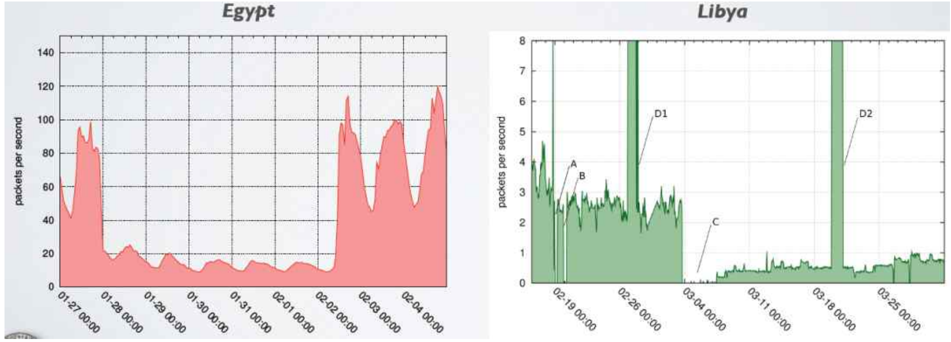
The New Internet


Global Internet Core
Regional / Tier2 Providers
Customer IP Networks
Global Transit / National Backbones
"Hyper Giants" Large Content, Consumer, Hosting CDN
IXP
IXP
IXP
ISP1
ISP2

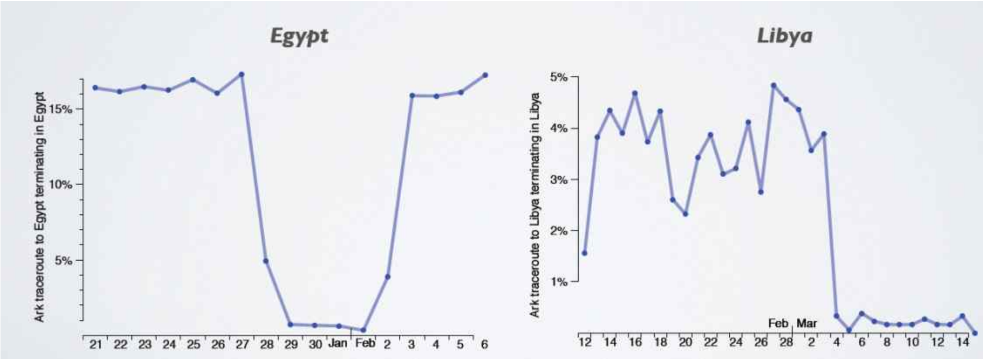
C. Labovitz et al., "Internet Inter-Domain Traffic," ACM SIGCOMM2010

인터넷 검열

- CAIDA malware monitoring

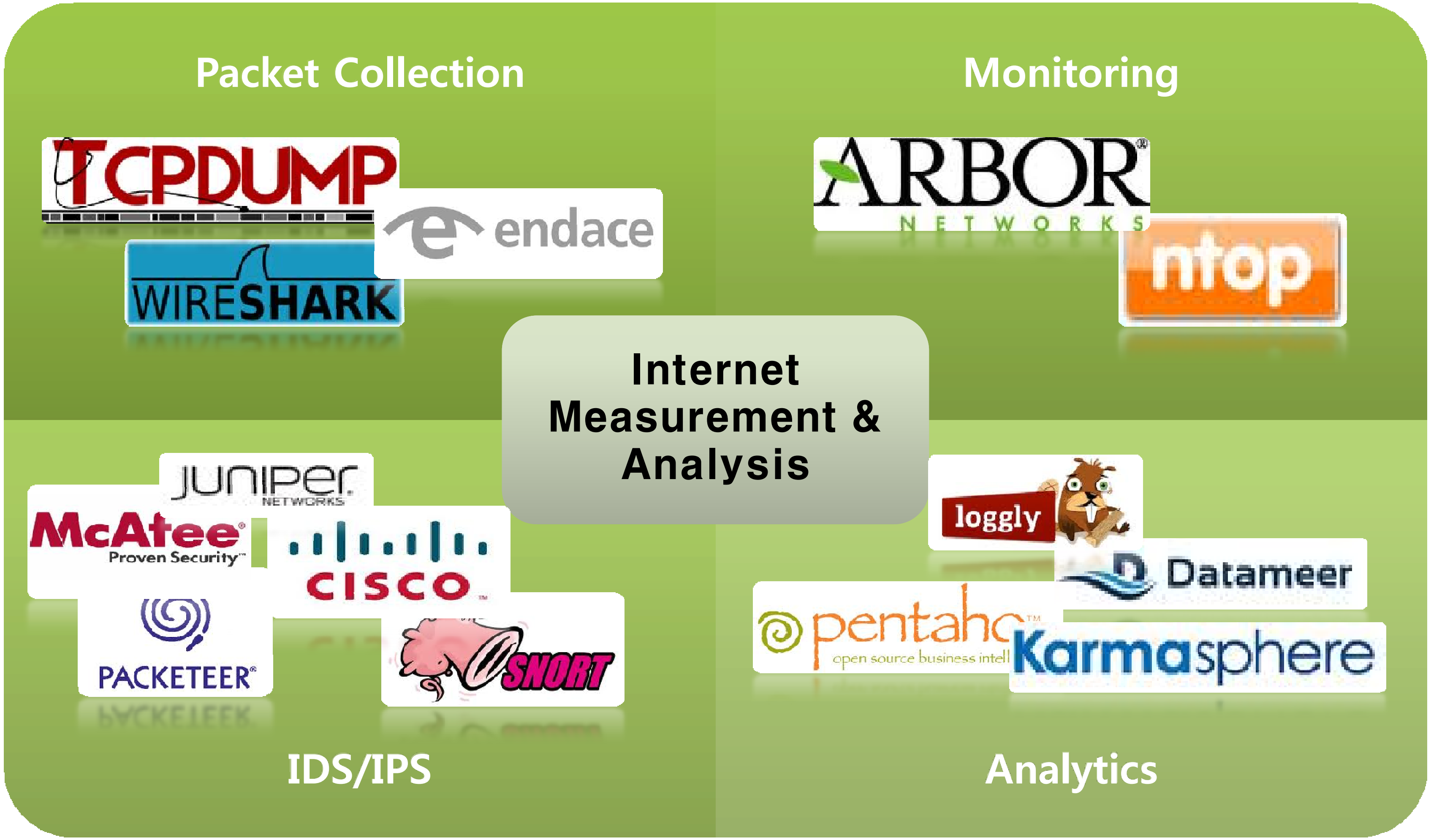


- CAIDA Ark



A. Dainotti et al., "Analysis of country-wide Internet outages caused by censorship, " ACM IMC2011

인터넷 측정 관련 분야



Packet Collection

Monitoring

Internet Measurement & Analysis

IDS/IPS

Analytics

관련도구

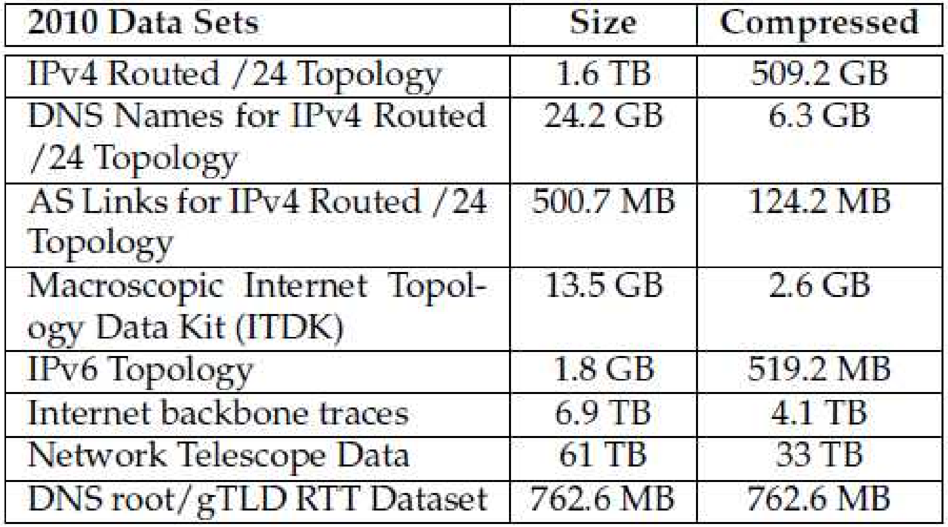
- Open source vs. commercial
- CAIDA
  - Internet telescope
- Active probing tools
  - Ping, traceroute, skitter, scamper
  - Archipelago
- Passive tools
  - CoralReef, flow-tools, tcpdump/wireshark
  - Bro, snort

인터넷 측정 연구의 이슈

- 대규모 데이터 처리
  - 캡춰, 저장, 분석
    - 고속 링크: 10 Gbps 이더넷
    - 실시간 vs 비실시간
- 데이터 마이닝
  - 보안, 성능, 장애, 비즈니스 인텔리전스
- 툴
  - 오픈소스 vs 상용
  - 파일시스템 vs DB

인터넷 측정 데이터

- CAIDA
  - 토폴로지 데이터
    - IPv4/IPv6 by Archipelago
  - 트래픽
    - Internet core backbone links
    - UCSD Network Telescope

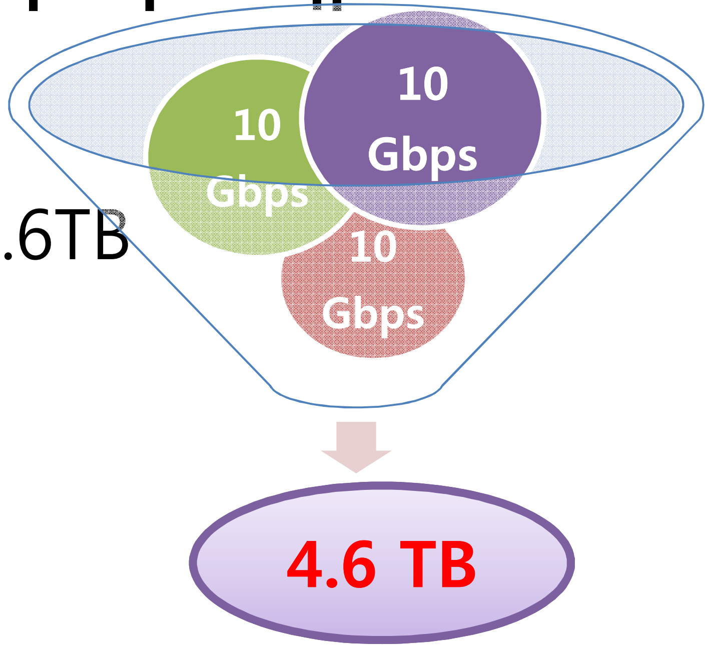


2010 Data Sets	Size	Compressed
IPv4 Routed /24 Topology	1.6 TB	509.2 GB
DNS Names for IPv4 Routed /24 Topology	24.2 GB	6.3 GB
AS Links for IPv4 Routed /24 Topology	500.7 MB	124.2 MB
Macroscopic Internet Topology Data Kit (ITDK)	13.5 GB	2.6 GB
IPv6 Topology	1.8 GB	519.2 MB
Internet backbone traces	6.9 TB	4.1 TB
Network Telescope Data	61 TB	33 TB
DNS root/gTLD RTT Dataset	762.6 MB	762.6 MB

Table 1: Data CAIDA Regularly Collected in 2010

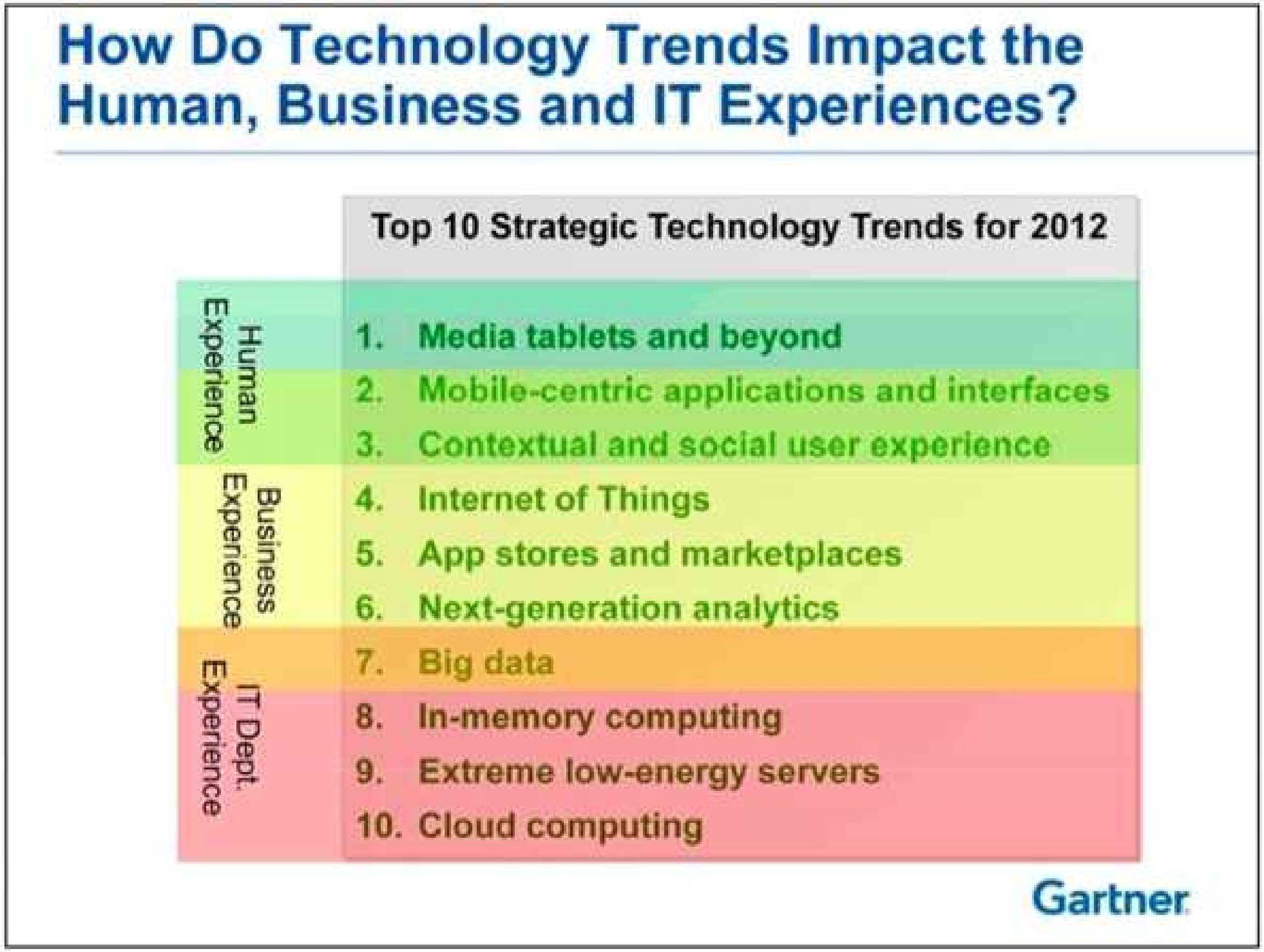
M. Fomenkov and k. claffy, "Internet measurement data management challenges'', in Workshop on Research Data Lifecycle Management. Jul 2011, Workshop on Research Data Lifecycle Management.

인터넷 측정 데이터: 예



- 패킷
  - 10GE링크에서 1시간 측정시 4.6TB
- Cisco NetFlow
  - CNU 1일 1.2GB
- 라우팅 테이블
  - Routeviews in bzip2
    - RIBS: 45MB every 2 hr -> 540MB for 1 day ->197GB for 1 year
    - Updates: 1MB for every 15 minutes

Gartner: 10 Key IT Trends for 2012



How Do Technology Trends Impact the Human, Business and IT Experiences?

	Top 10 Strategic Technology Trends for 2012
Human Experience	1. Media tablets and beyond
	2. Mobile-centric applications and interfaces
	3. Contextual and social user experience
Business Experience	4. Internet of Things
	5. App stores and marketplaces
	6. Next-generation analytics
IT Dept. Experience	7. Big data
	8. In-memory computing
	9. Extreme low-energy servers
	10. Cloud computing

Gartner

http://www.gartner.com/it/page.jsp?id=1826214

분산/병렬처리 기술의 발전

- 분산컴퓨팅
  - Google MapReduce, 2004
  - 1 PB sorting by Google
    - 2008: 6 hours and 2 minutes on 4,000 computers
    - 2011: 33 minutes on 8000 computers

      http://googleresearch.blogspot.kr/2011/09/sorting-petabytes-with-mapreduce-next.html

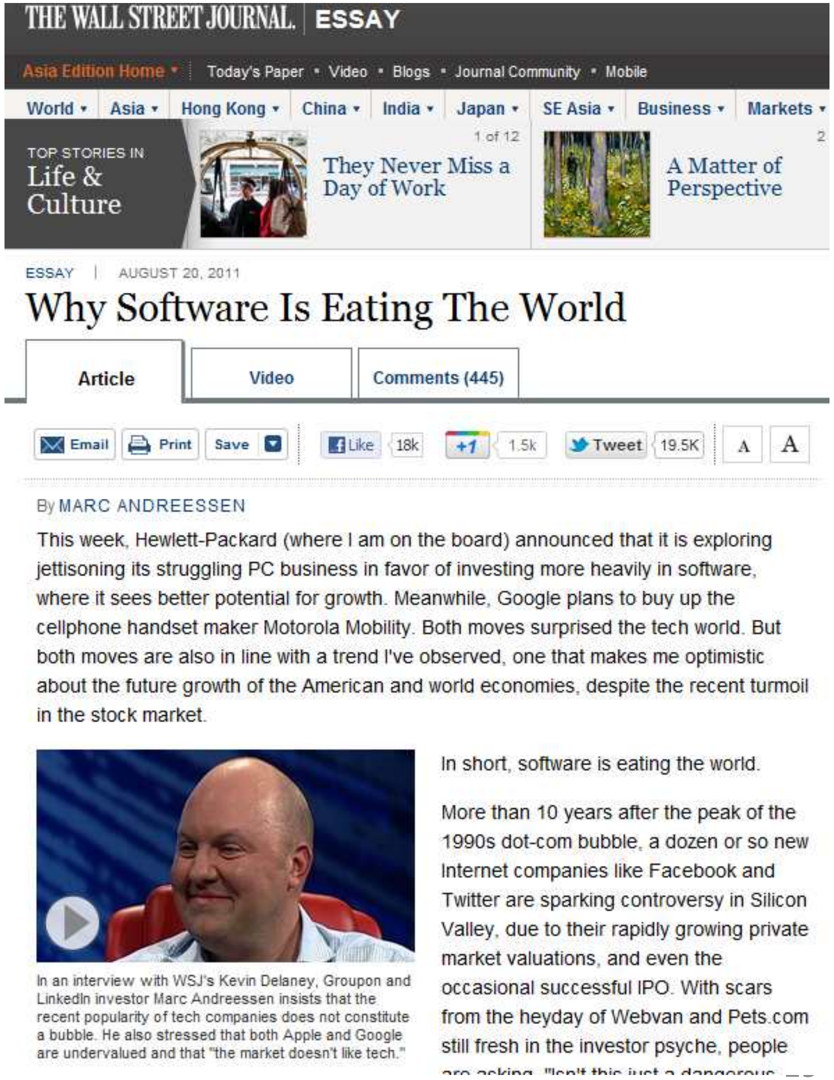
- 병렬/멀티코어
  - multi-core CPU, GPU, FPGA
  - 40Gbps IP forwarding capability
    - S. Han et al., "PacketShader: A GPU-Accelerated Software Router" ACM SIGCOMM, 2010

Why Software Is Eating the World

by Marc Andreessen, Aug. 20, 2011

http://online.wsj.com/article/SB10001424053111903480904576512250915629460.html

- Hewlett-Packard bought
  - Autonomy for $10 billion
- IBM invest
  - $100 million for big data analysis research
- EMC bought
  - Greenplum for $300 million
- Oracle bought
  - Endeca Technologies



ESSAY | AUGUST 20, 2011

Why Software Is Eating The World

By MARC ANDREESSEN

This week, Hewlett-Packard (where I am on the board) announced that it is exploring jettisoning its struggling PC business in favor of investing more heavily in software, where it sees better potential for growth. Meanwhile, Google plans to buy up the cellphone handset maker Motorola Mobility. Both moves surprised the tech world. But both moves are also in line with a trend I've observed, one that makes me optimistic about the future growth of the American and world economies, despite the recent turmoil in the stock market.

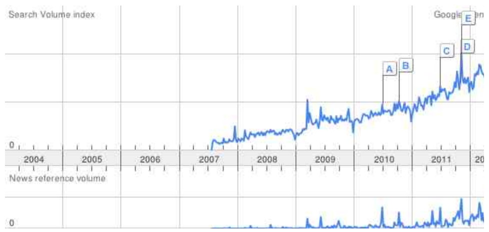
In an interview with WSJ's Kevin Delaney, Groupon and LinkedIn investor Marc Andreessen insists that the recent popularity of tech companies does not constitute a bubble. He also stressed that both Apple and Google are undervalued and that "the market doesn't like tech."

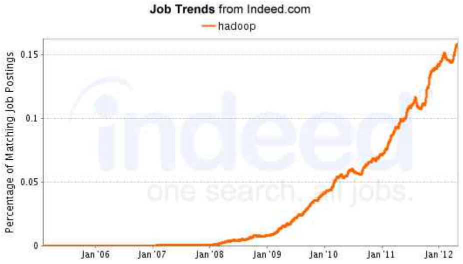
In short, software is eating the world.

More than 10 years after the peak of the 1990s dot-com bubble, a dozen or so new Internet companies like Facebook and Twitter are sparking controversy in Silicon Valley, due to their rapidly growing private market valuations, and even the occasional successful IPO. With scars from the heyday of Webvan and Pets.com still fresh in the investor psyche, people are asking, "Isn't this just a dangerous hadoop

http://www.google.com/trends/?q=hadoop

- Open-source framework for running applications on large clusters built of commodity hardware
- Implementation of MapReduce and HDFS
  - MapReduce : computational paradigm
  - HDFS : distributed file system





http://www.indeed.com/jobtrends?q=hadoop&l=

twitter facebook YAHOO! cloudera YAHOO! hortonworks

IBM LinkedIn Datameer JASPERSOFT

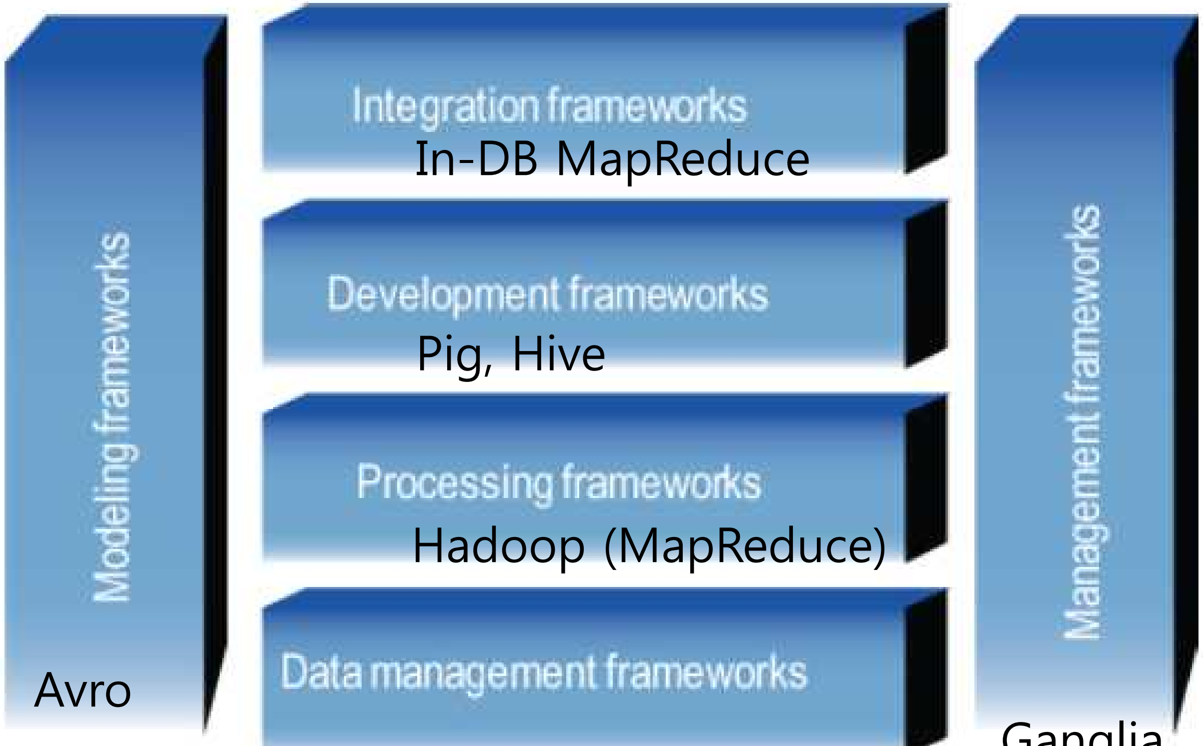
NETFLIX pentaho open source business intelligence loggly

CBS Interactive DataStax MAPR TECHNOLOGIES

ORBITZ NEXR TOWARD OPEN PLATFORM innogrid

ebay NTT Communications

Hadoop Technology



In-DB MapReduce

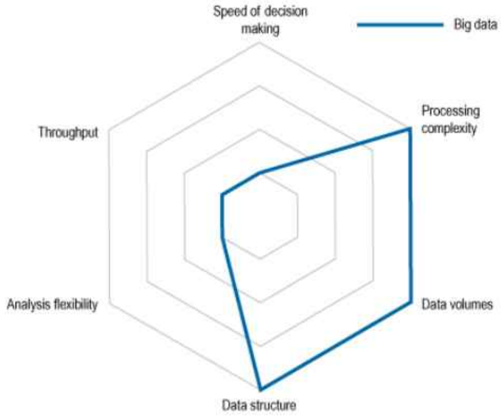
Pig, Hive

Hadoop (MapReduce)

Avro

HDFS, HBase

Ganglia,
Nagios



http://info.cloudera.com/GartnerReportHadoopJanuary2011.html

충남대 데이터네트워크 연구실

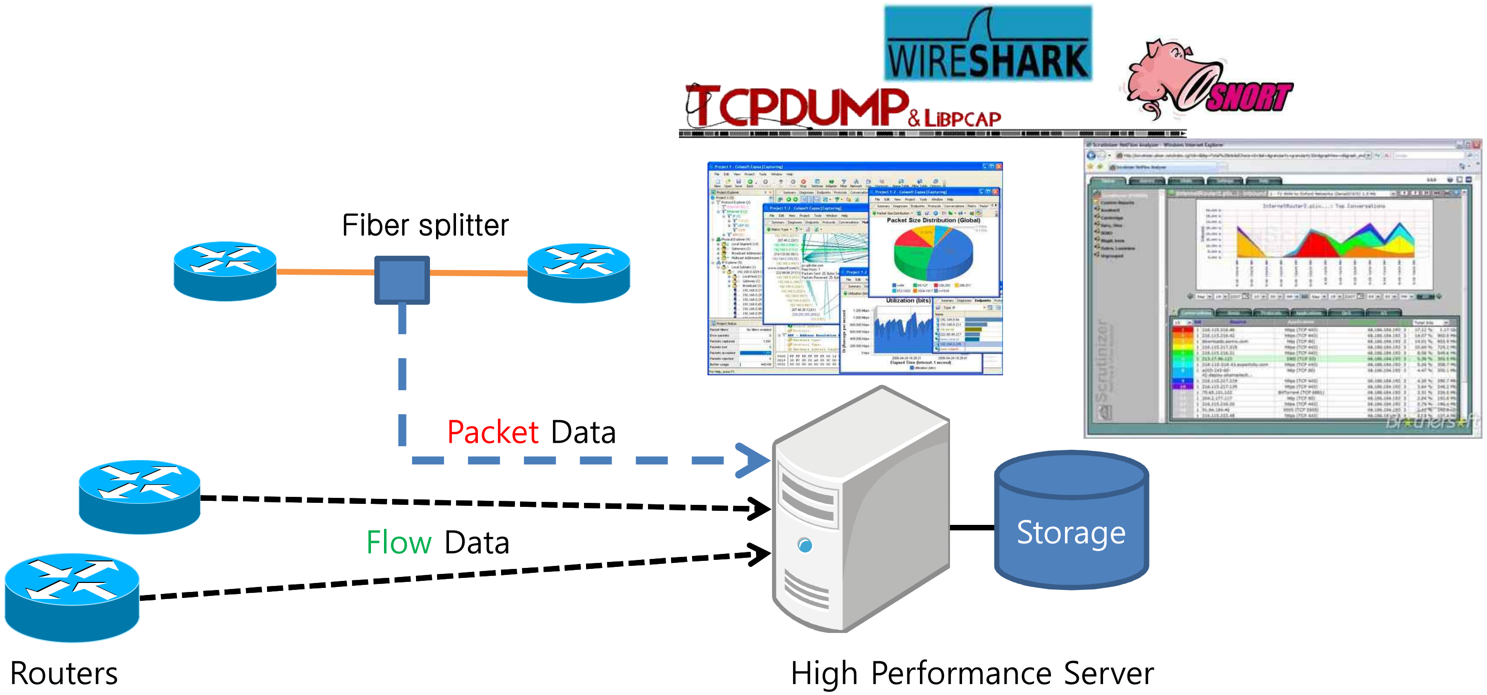
http://networks.cnu.ac.kr

- 인터넷 데이터분석 프레임워크 연구
  - Input: packet/flow/BGP
  - Output: statistics/analytics results
    - Ex) top 10 heavy users, popular applications, favorite websites, anomaly detection
- 접근방법
  - 오픈소스 분산 컴퓨팅 플랫폼 활용
    - Google's programming model, MapReduce
    - Open-source system, Hadoop
- 이유
  1. 대규모 파일 저장을 위한 분산파일시스템
  2. 대규모 데이터 처리 및 분석을 위한 분산컴퓨팅
     - More data usually beats better algorithm
  3. 시스템 장애에 견고한 구조

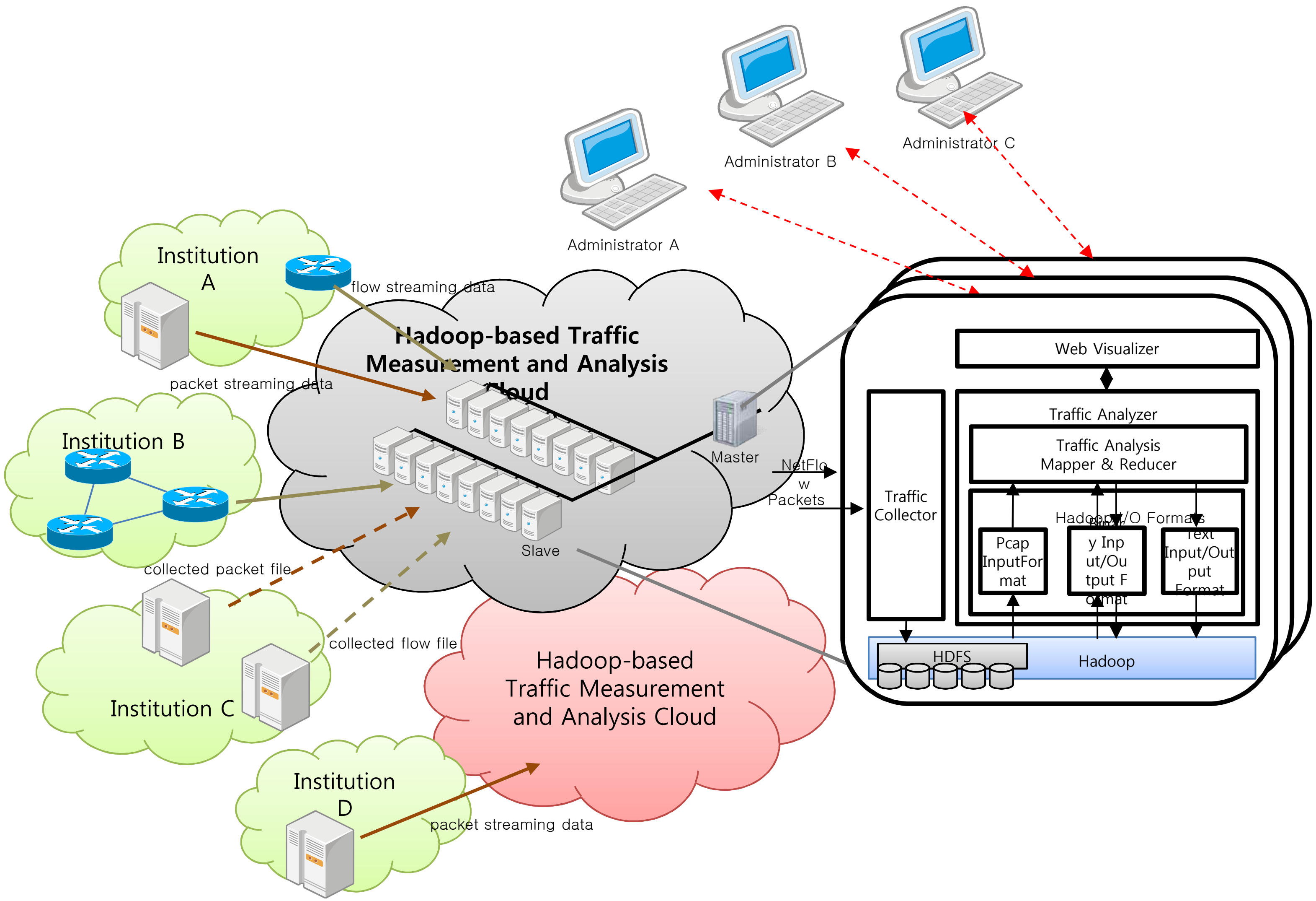
- 인력양성
  - Hadoop 관련 교육(2009 ~)
    - Cloudera, Hadoop summit, Hadoop world, Strata conference
  - Platformday, JCO 발표

기존 인터넷 트래픽 측정 및 분석

- Packet
  - HTTP request packet, DNS packet
- Abstract of a set of packets: flow (Cisco router)
  - (168.188.1.10, 61.10.1.1, 3000, 80) 1MB, 1000 packets, 30 secs



Hadoop 기반 인터넷 트래픽 측정 및 분석 시스템 구조



연구 결과

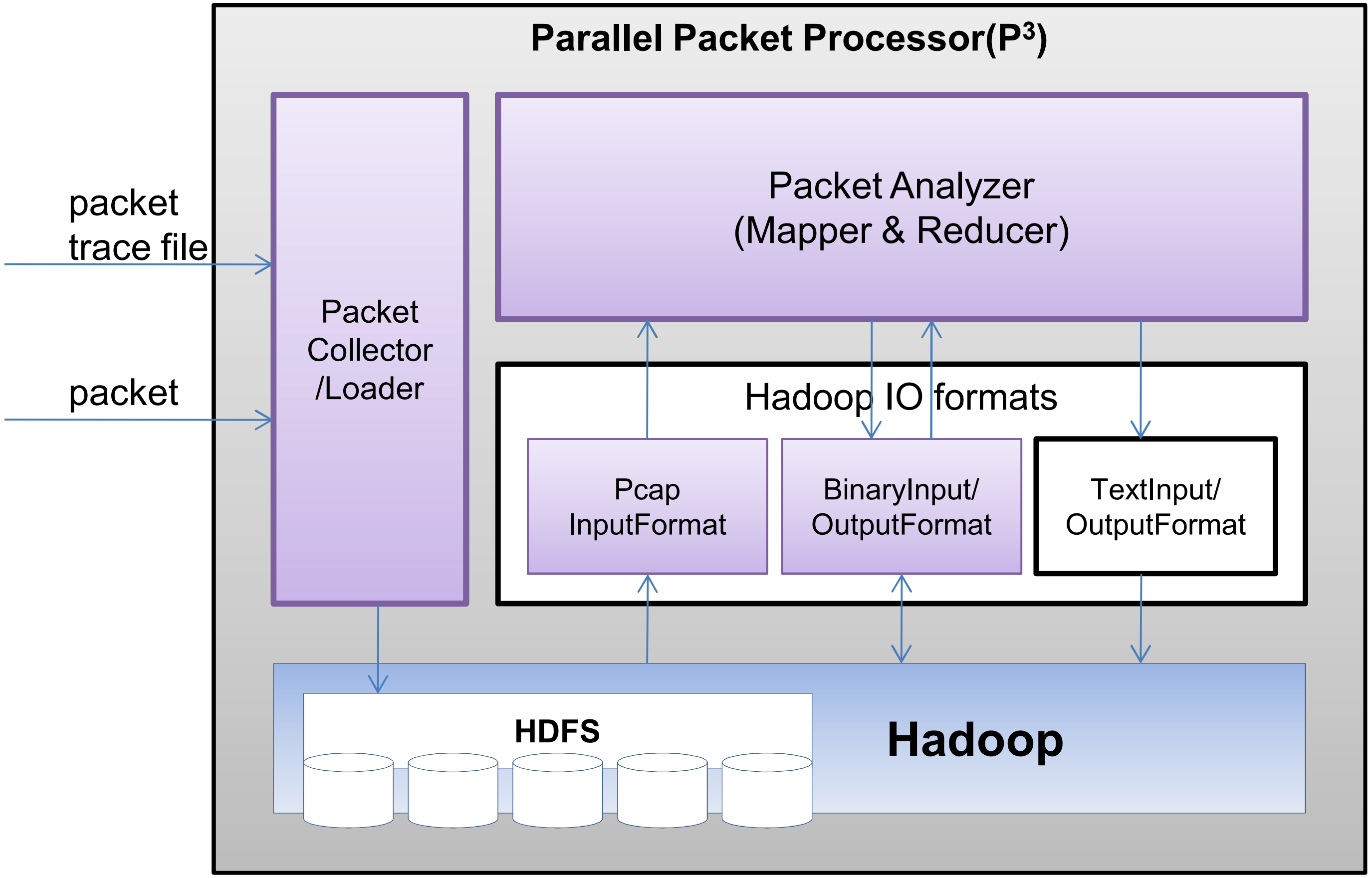
https://sites.google.com/a/networks.cnu.ac.kr/dnlab/publication

- 소프트웨어
  - Hadoop의 바이너리 InputFormat 모듈
  - Hadoop에서 pcaplib 패킷 처리 모듈
  - Hadoop에서 Cisco NetFlow 처리 모듈
  - MapReduce기반 IP/TCP/HTTP/BGP 분석 모듈
  - https://sites.google.com/a/networks.cnu.ac.kr/dnlab/research/hadoop
  - https://github.com/ssallys/pcap-on-Hadoop
- 논문

  [1] "An Internet Traffic Analysis Method with MapReduce", Cloudman workshop, April 2010

  [2] "A Hadoop-based Packet Trace Processing Tool," TMA 2011, April, 2011
- 특허
- 관련연구
  - NHN

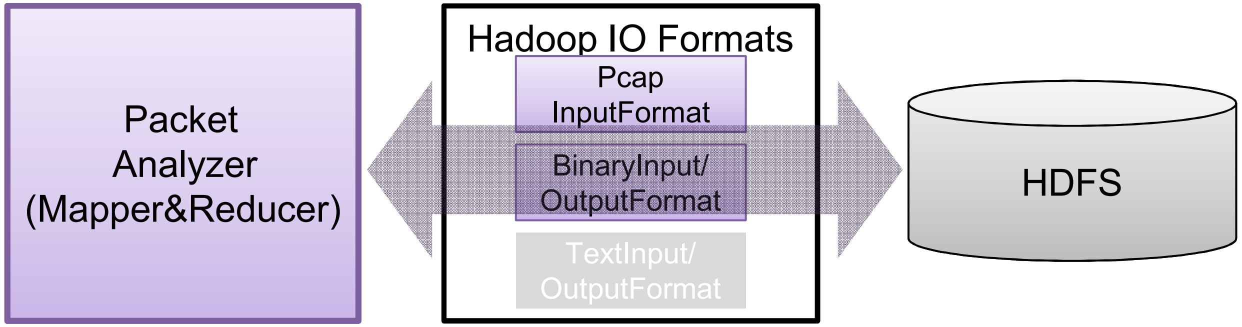
Hadoop에서 IP 패킷 분석 방법



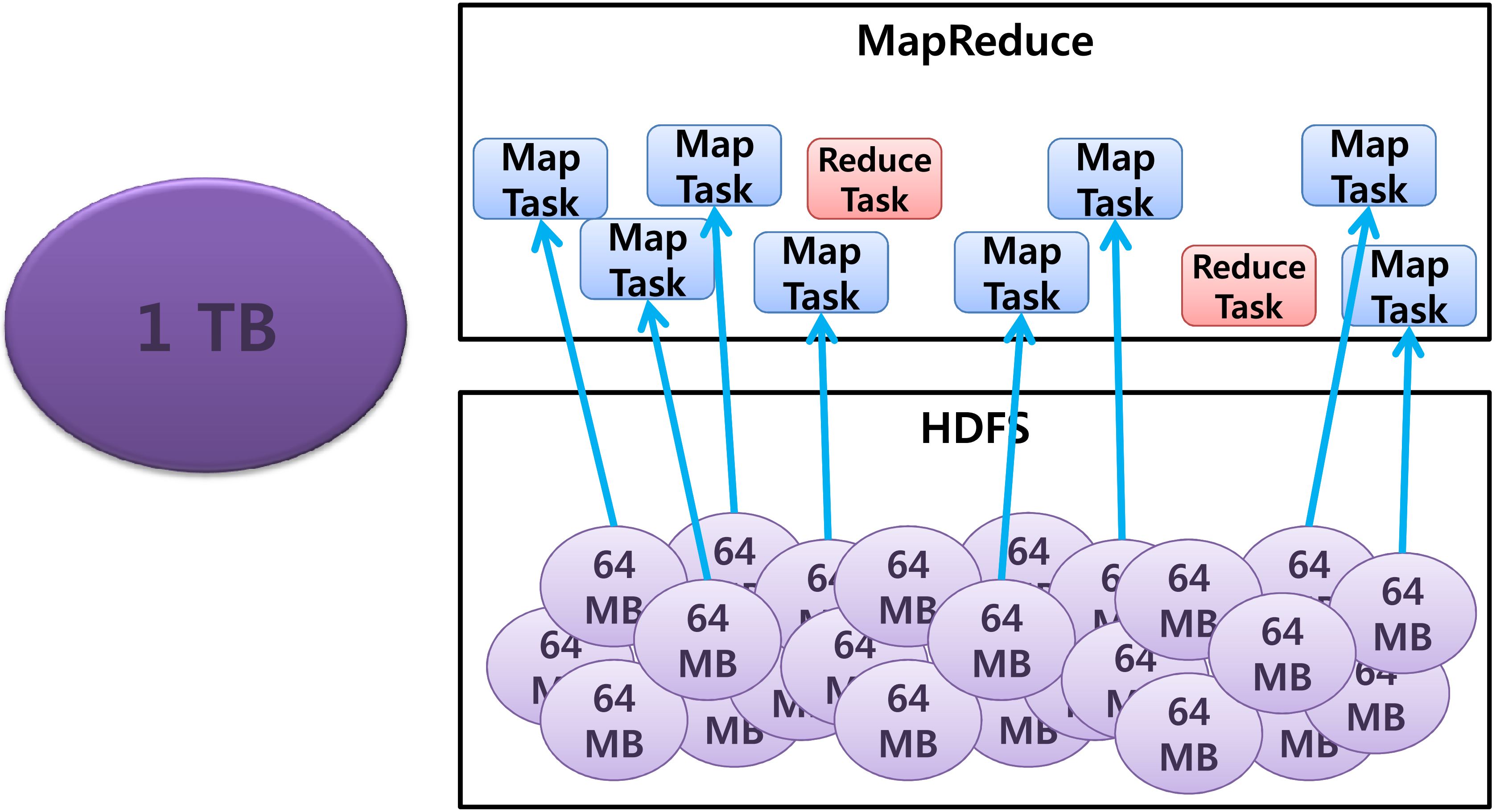
패킷 처리를 위한 Hadoop InputFormat

- PcapInputFormat
  - InputFormat for pcap packet trace file
  - uses pcap header fields to detect Split boundary
  - can parse variable-length of binary packet records

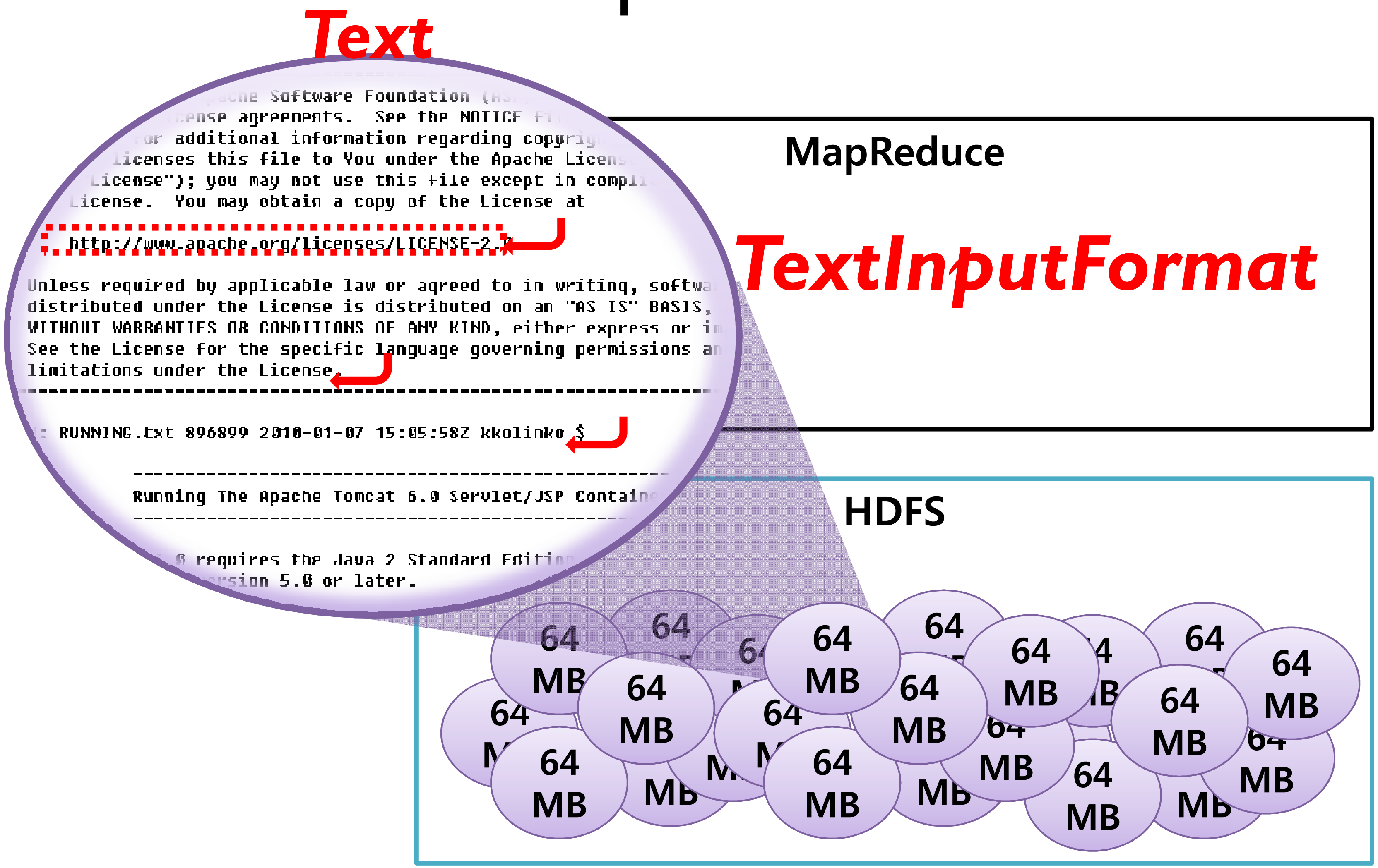
- BinaryInputFormat
  - InputFormat for binary data files which contains fixed-lengh of binary records
  - uses record's length to detect Split boundary
  - can parse fixed-length of binary records



Hadoop에서 패킷 분석 구조 개요

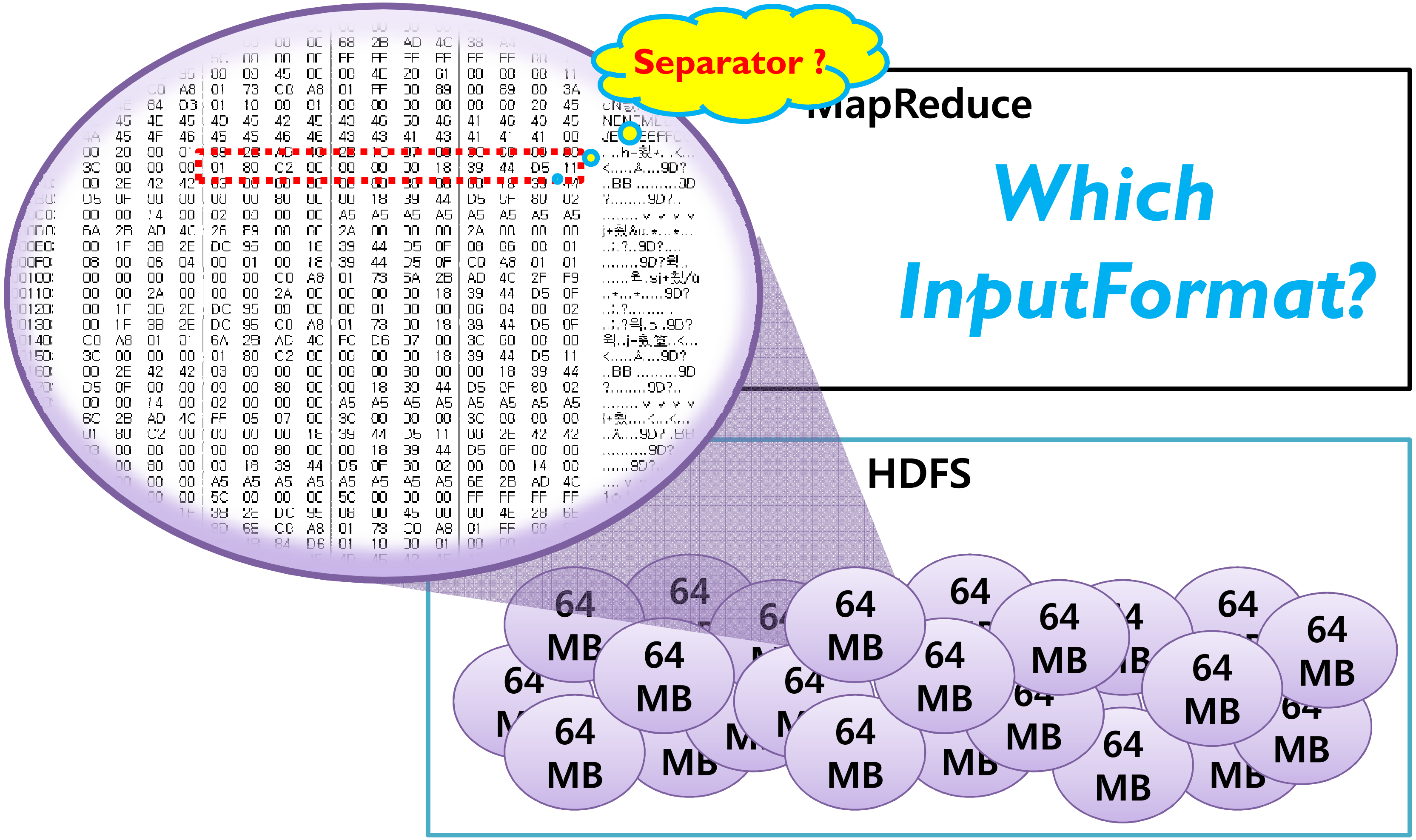


TextInputFormat



```
pache Software Foundation (AS
cense agreements.  See the NOTICE fi
For additional information regarding copyrig
licenses this file to You under the Apache Licens
License"); you may not use this file except in compli
License.  You may obtain a copy of the License at

 http://www.apache.org/licenses/LICENSE-2.

Unless required by applicable law or agreed to in writing, softwa
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or im
See the License for the specific language governing permissions an
limitations under the License.
=================================================================

: RUNNING.txt 896899 2010-01-07 15:05:58Z kkolinko $

          --------------------------------------------------
          Running The Apache Tomcat 6.0 Servlet/JSP Containe
          ==================================================

              0 requires the Java 2 Standard Edition
            ersion 5.0 or later.
```

Packet Record in HDFS File ?



Separator ?

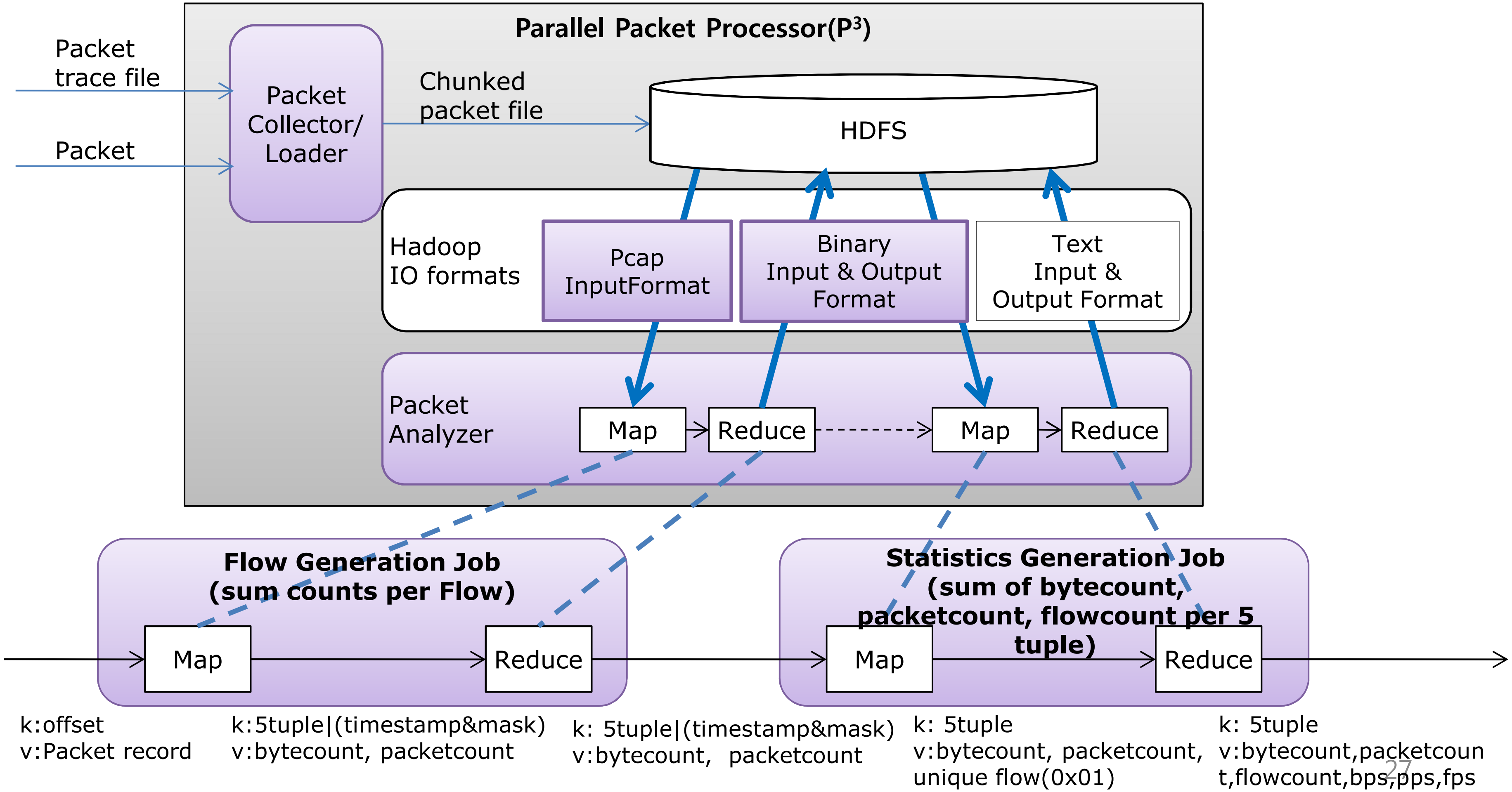
MapReduce

Which InputFormat?

HDFS

InputFormat 예제

- Periodic flow statistics



Parallel Packet Processor(P³)
Packet trace file
Packet
Packet Collector/ Loader
Chunked packet file
HDFS
Hadoop IO formats
Pcap InputFormat
Binary Input & Output Format
Text Input & Output Format
Packet Analyzer
Map
Reduce
Map
Reduce
Flow Generation Job (sum counts per Flow)
Map
Reduce
Statistics Generation Job (sum of bytecount, packetcount, flowcount per 5 tuple)
Map
Reduce
k:offset
v:Packet record
k:5tuple|(timestamp&mask)
v:bytecount, packetcount
k: 5tuple|(timestamp&mask)
v:bytecount, packetcount
k: 5tuple
v:bytecount, packetcount, unique flow(0x01)
k: 5tuple
v:bytecount,packetcount,flowcount,bps,pps,fps

패킷 분석 도구

Traffic Analysis	P^3 Command	MapReduce Jobs	CoralReef
Total traffic and host/port count statistics	PcapTotalStats	▪ count up bytes/packets ▪ emit unique IP/port and flow ▪ summarize # of unique IP/Port and flows	**crl_stats** –Csource=[pcap file]
Periodic flow statistics	PcapTotalFlowStats	▪ generate flows and count up bytes/packets ▪ summarize periodical flow statistics	**crl_flow** –I –b –Ci=[interval] Csource=[pcap file] **/t2_rate** -s
Periodic simple traffic statistics	PcapRate	▪ compute periodic bytes/packets regarding IPv4/v6/non-IP per interval	**crl_rate** –Ci=[interval] Csource=[pcap file]
Top N	PcapTopN	▪ sort records and emit top N record	[command line] **/t2_top** –Sb –n[n]
Total count grouping by key	PcapCountUp	▪ count up bytes/packets per key	-

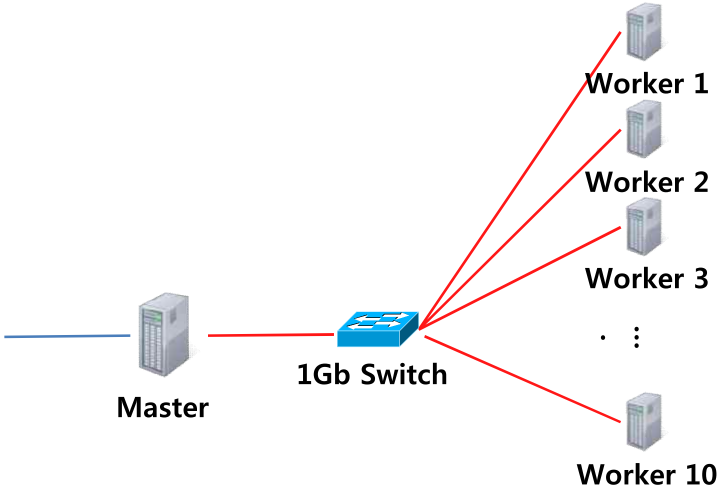
실험

Testbed

	Type	Nodes	CPU	Memory	HardDisk
CoralReef	Single	1	2.83GHz (Quad-core)	4 GB	1.5 TB
Hadoop	Standard	5	2.83GHz (Quad-core)	4 GB	1.5 TB
	High-performance	10	2.93GHz (Octo-core)	16 GB	1 TB

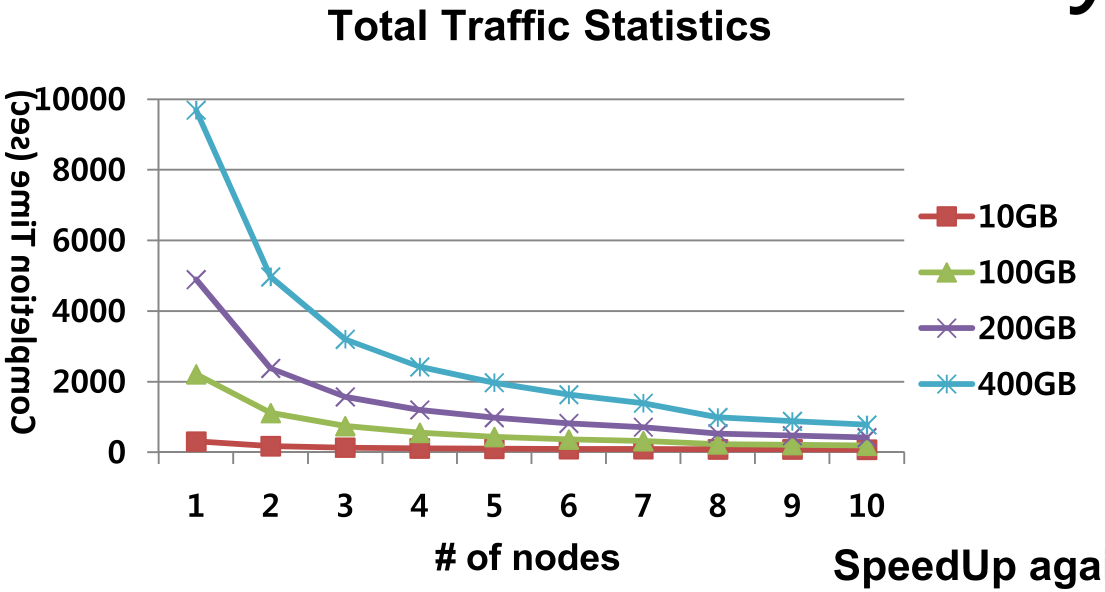
Packet Trace files

Type	# of Packet files	# of Packets
10 GB	1	9.4 M
100 GB	1	92.7 M
200 GB	2	185.4 M
400 GB	7	441.1 M



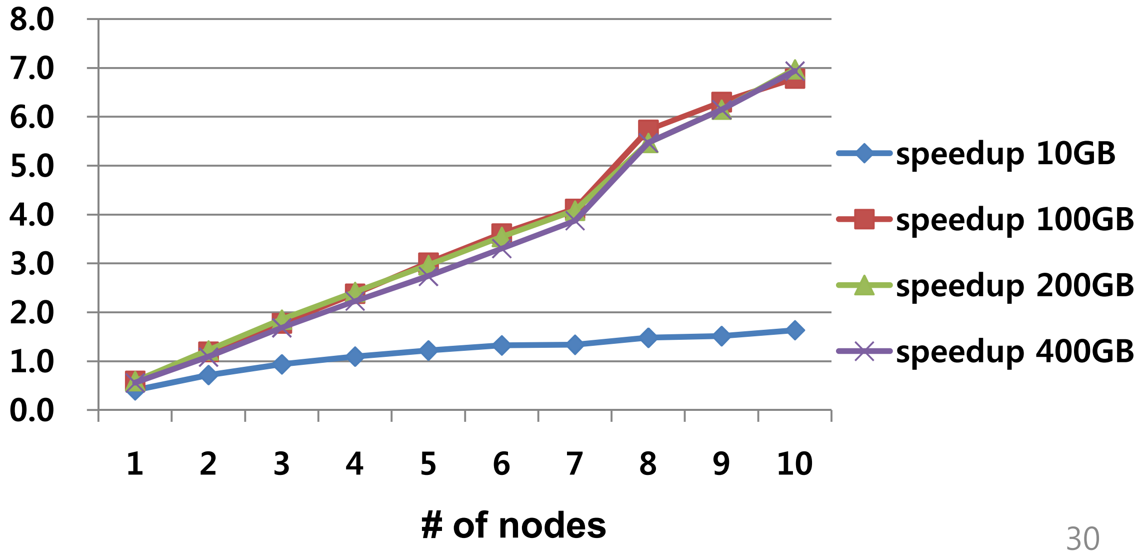
Scalability

Total Traffic Statistics



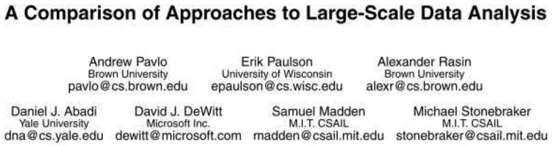
Resource-proportional Computing !

SpeedUp against CoralReef



결론

- Hadoop fits well into big Internet data analysis
- Hadoop 장단점
  - One size does not fit all !
  - MapReduce 알고리즘의 복잡성
    - Aggregation 문제에서는 좋은 장점
    - 복잡한 문제처리를 위한 MapReduce 알고리즘의 비효율성: TCP
  - Hadoop 안정성/버전/성능 이슈



A Comparison of Approaches to Large-Scale Data Analysis

Andrew Pavlo
Brown University
pavlo@cs.brown.edu

Erik Paulson
University of Wisconsin
epaulson@cs.wisc.edu

Alexander Rasin
Brown University
alexr@cs.brown.edu

Daniel J. Abadi
Yale University
dna@cs.yale.edu

David J. DeWitt
Microsoft Inc.
dewitt@microsoft.com

Samuel Madden
M.I.T. CSAIL
madden@csail.mit.edu

Michael Stonebraker
M.I.T. CSAIL
stonebraker@csail.mit.edu

- 현재 연구
  - Hive 연동 interactive 분석 방법 및 시각화
  - 실시간처리: 인메모리, SSD
  - 데이터마이닝기법 연동: Mahout, R
- The data center is the computer !
  - More data usually beats better algorithm