BIG DATA FOR SECURITY: A CYBERSTRATEGIC APPROACH

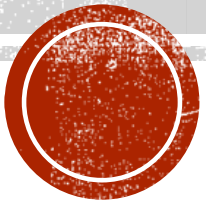
Kavé Salamatian

Professor of computer Science

University of Savoie, France

Castex Chair on CyberStrategy

Labex Persyval, Grenoble Alps University



CYBERSTRATEGY ?

- Strategy
  - The art of positioning, commanding, directing, coordinating forces in space in order to attain the objectives
- Space ?
  - What is the space of cyberspace
- What are cyber-objectives ?
  - How do define cyber-objectives? What is the constraint space ?
- What are cyber-forces?
  - Economics, Hackers, military, startups, research
- How to position, command, direct, coordinate
  - Governance
- **Is there anything new under the sun ?**

UNIVERSITÉ SAVOIE MONT BLANC

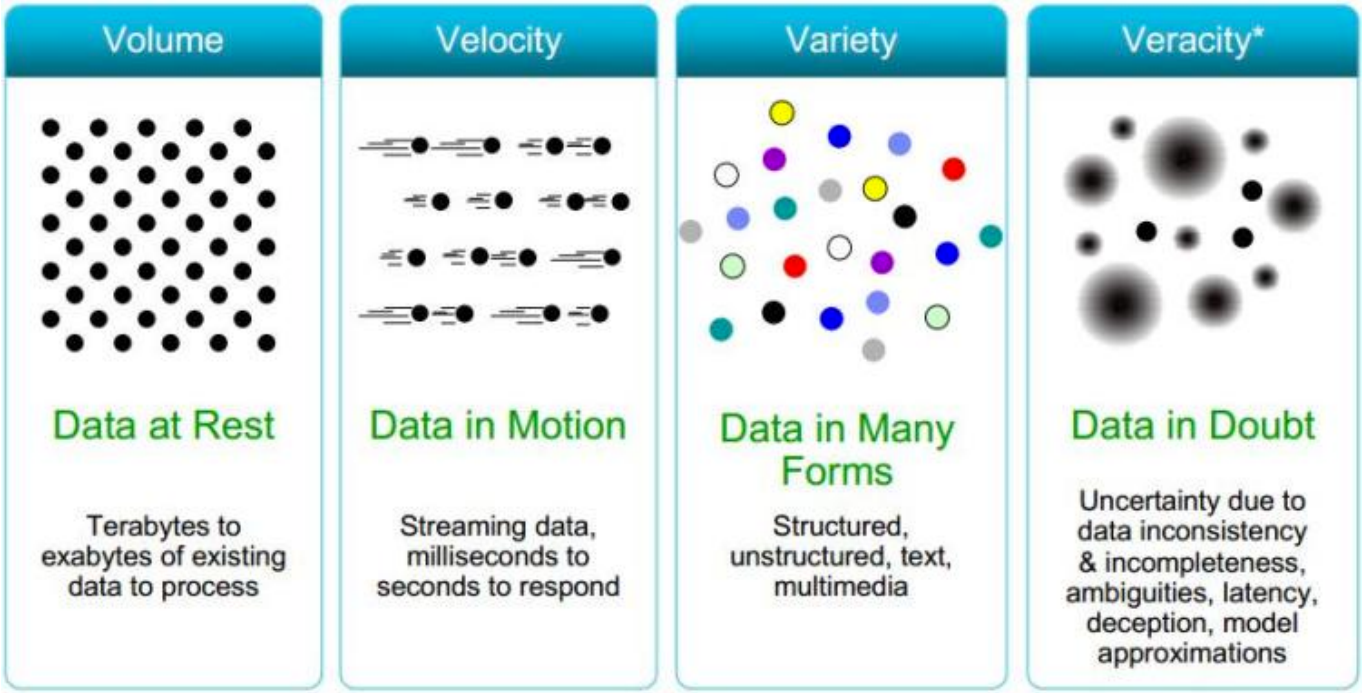
CYBERSTRATEGY : RISKS AND OPPORTUNITIES

- Risks
  - All networks have been used as vector of attack
    - Gengis Khan and Silk Road, railroads and armored trains
  - A network *per se* is a target for attack
    - First cyber-attack goes back to american civil war
    - The more important the information the more likely the attack
  - Opportunities
    - Economics
      - Google, Facebook
    - Strategic
      - Lay cables to get strategic position

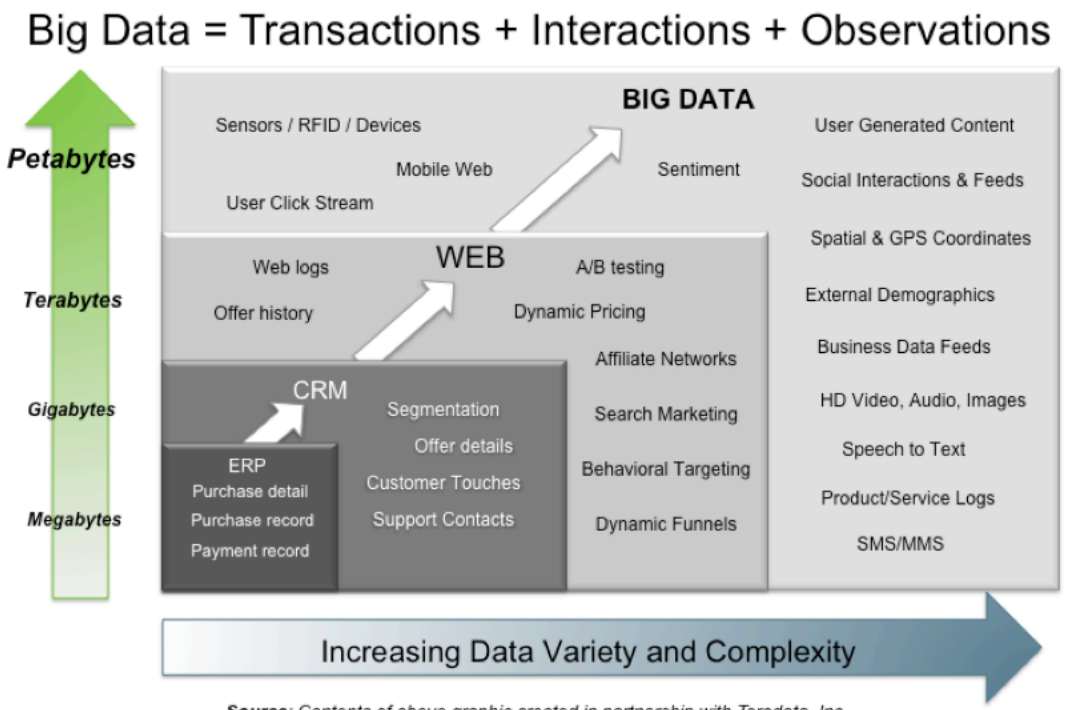
UNIVERSITÉ SAVOIE MONT BLANC

BIG DATA ?

- Four V



EVOLUTION



***Source**: Contents of above graphic created in partnership with Teradata, Inc.*

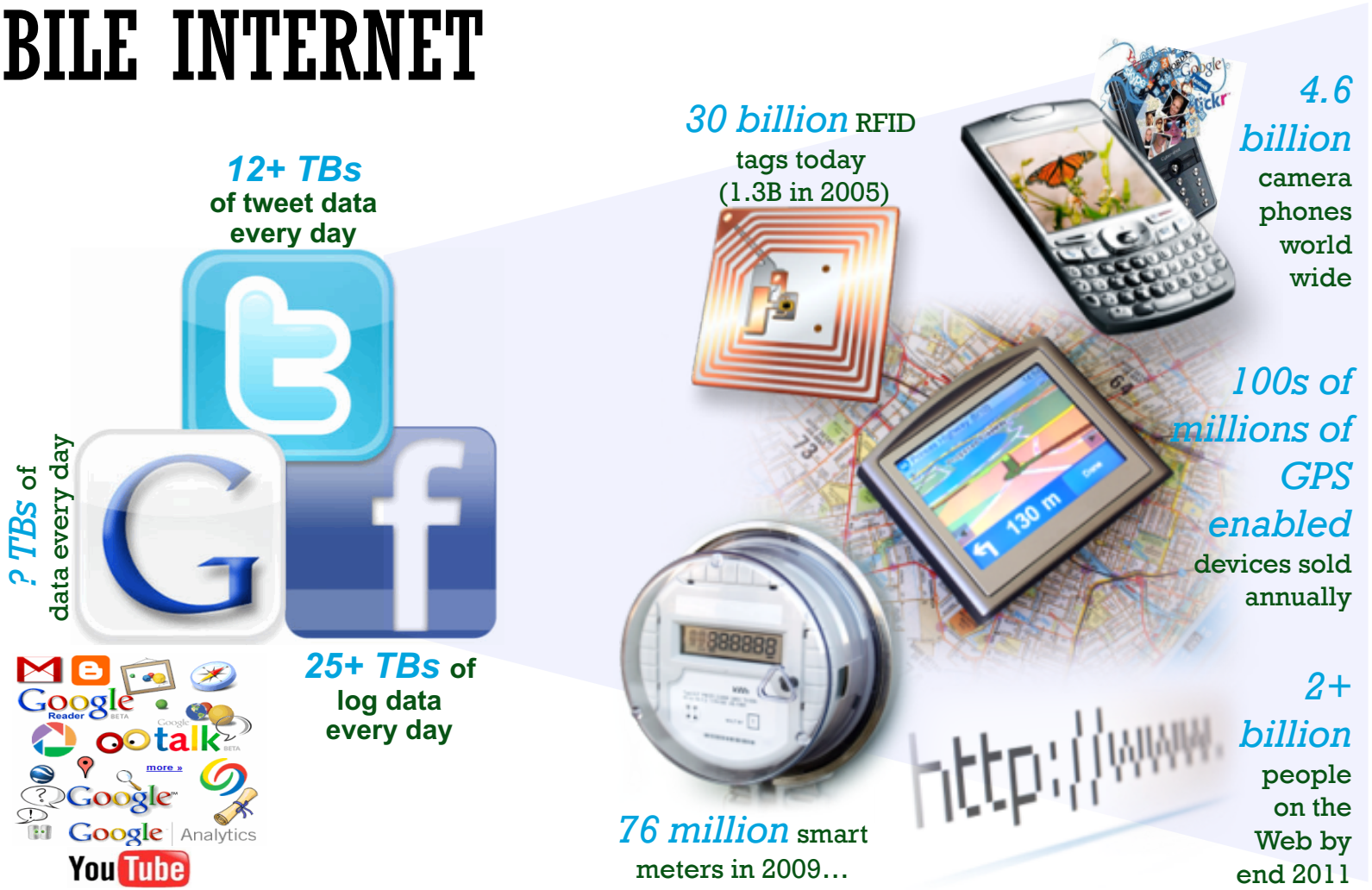
UNIVERSITÉ SAVOIE MONT BLANC

INTERACTIONS AND CO-EVOLUTION

- Networks as a source of Big data
- Big Data for networks
- Network systems for Big Data

UNIVERSITÉ SAVOIE MONT BLANC

MOBILE INTERNET



UNIVERSITÉ SAVOIE MONT BLANC

A CARTOGRAPHY OF CYBERSPACE

- 3 possible frameworks
  - Cyberspace embedded in geography
  - Geography embedded in cyberspace
  - Cyberspace as a space on its own

UNIVERSITÉ SAVOIE MONT BLANC

CYBERSPACE AS EMBEDDED IN GEOGRAPHY

- Cyberspace cannot exist without physical equipment and concrete infrastructures deployed in the geographical space
  - cables, servers, datacenters, exchange points, NOCs (Network Operating Centers), *etc.*
- These artefact depend on their geographical environment
  - Physical, political and economic constraints
- Example of issues
  - Risk analysis of the Internet connections between Europe and Asia
    - Classical geopolitics
  - Territory planning and of spatial inequalities.
    - Classical Geography

UNIVERSITÉ SAVOIE MONT BLANC

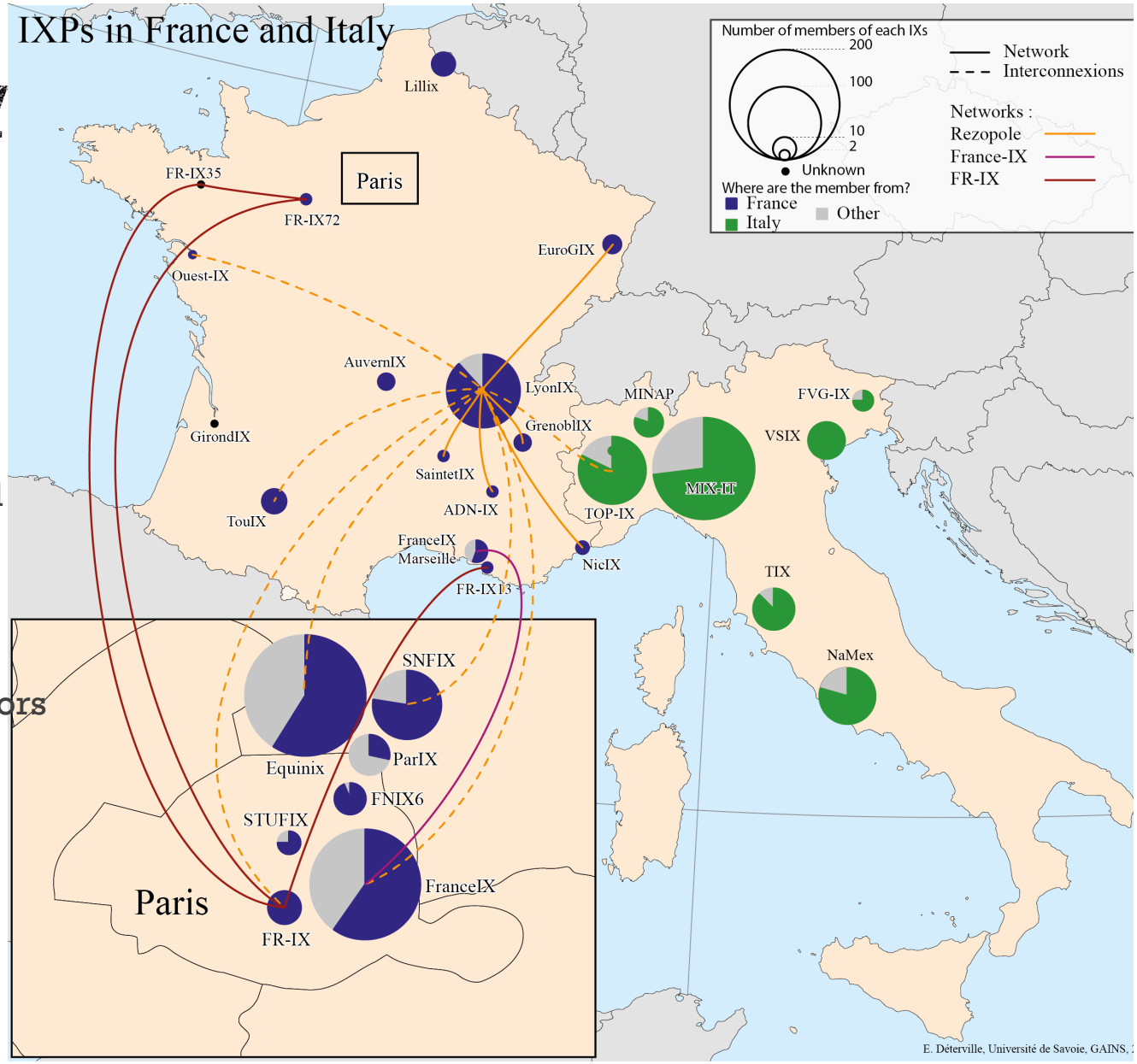
COMPARATIVE ANALYSIS OF IXPS IN FRANCE AND ITALY

France

- No central government planning
- Strong historical operator
- Emergence of a large number of local IXPs

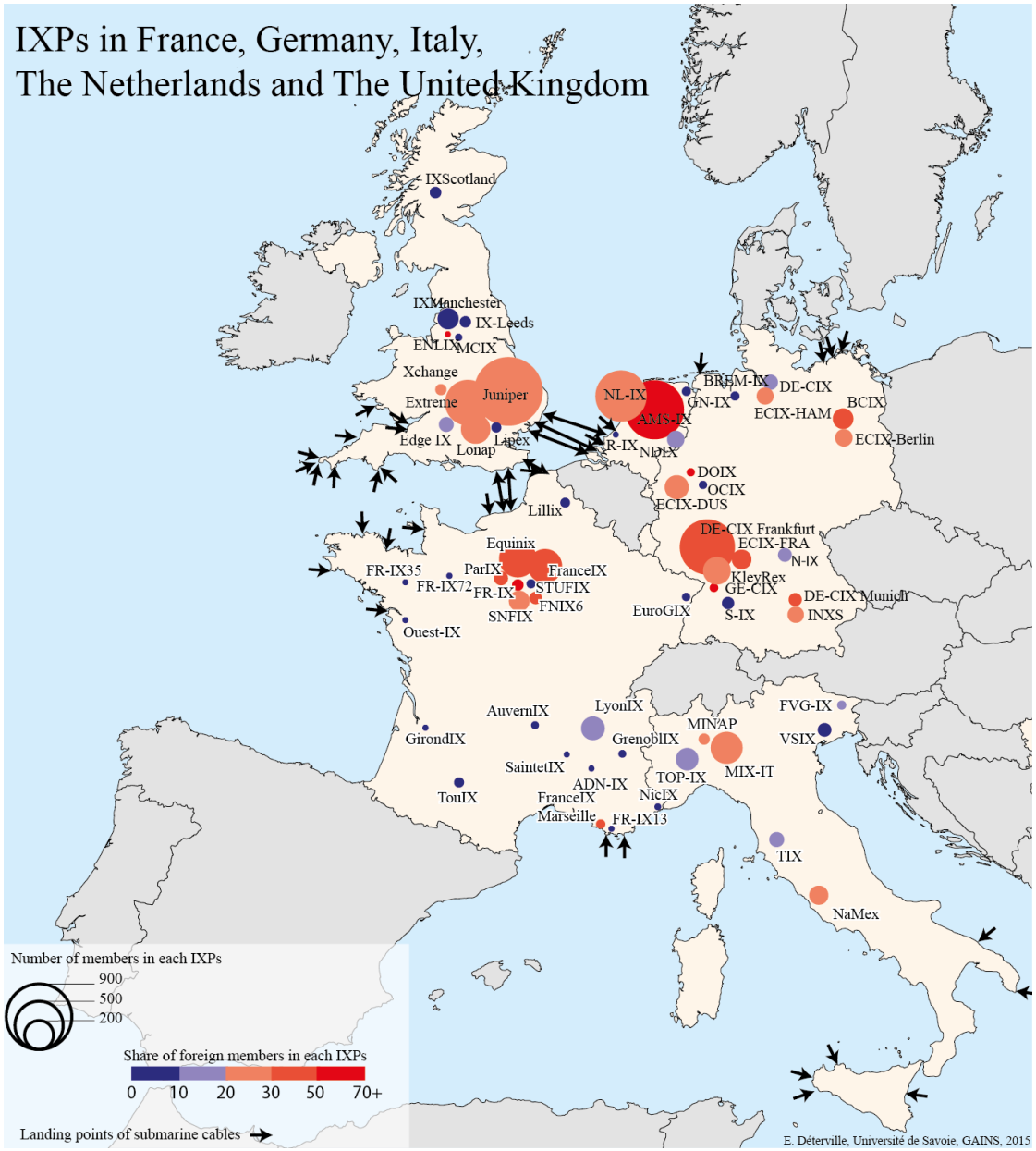
Italie

- No central government planning
- Concurrence between several operators
- Strong North vs. South divide



UNIVERSITÉ SAVOIE MONT BLANC

MAJOR IXPS IN EUROPE



UNIVERSITÉ SAVOIE MONT BLANC

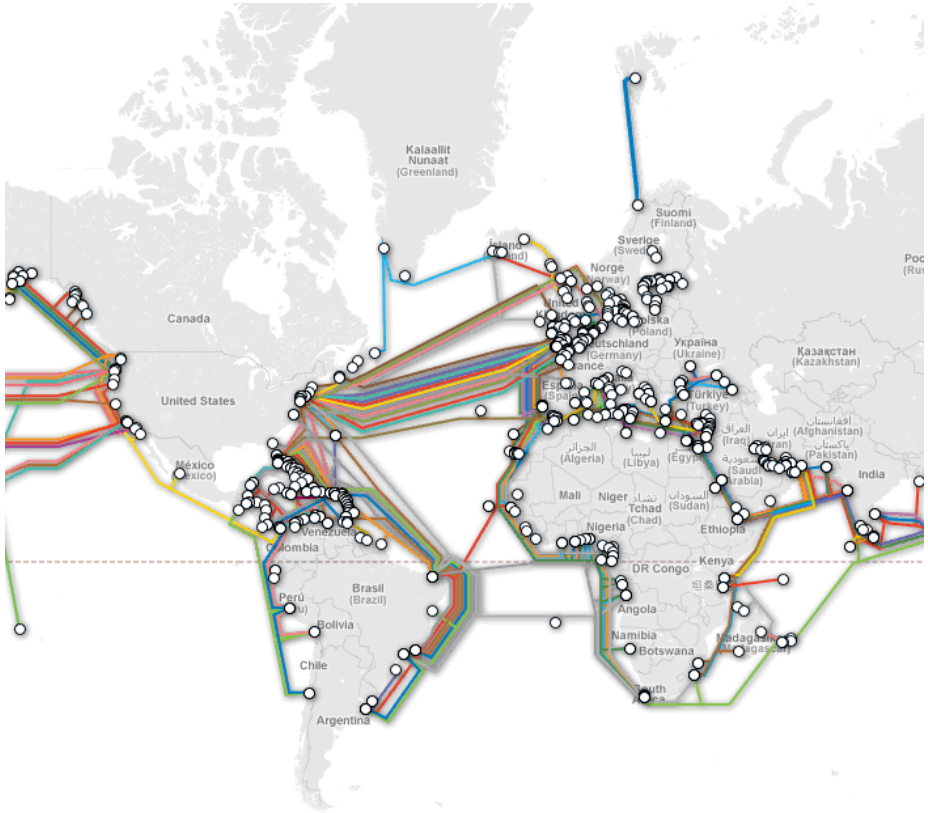
EURO-ASIA INTERNET AUTOBAHN : A STRATEGIC PATH

- Risk ?
  - Reliability
    - We need high speed and reliable paths
      - Low delay is important (because of Fast Trading)
    - Physical reliability
      - Cable cut
    - Network reliability
      - BGP Flaps
  - Surveillance
    - Several major surveillance actors on path
      - Europe (France, Germany, UK,Italy)
      - Israel, US/UK, Middle East actors, Singapour, China
  - Security
    - Data center mafia, Rogue ASes

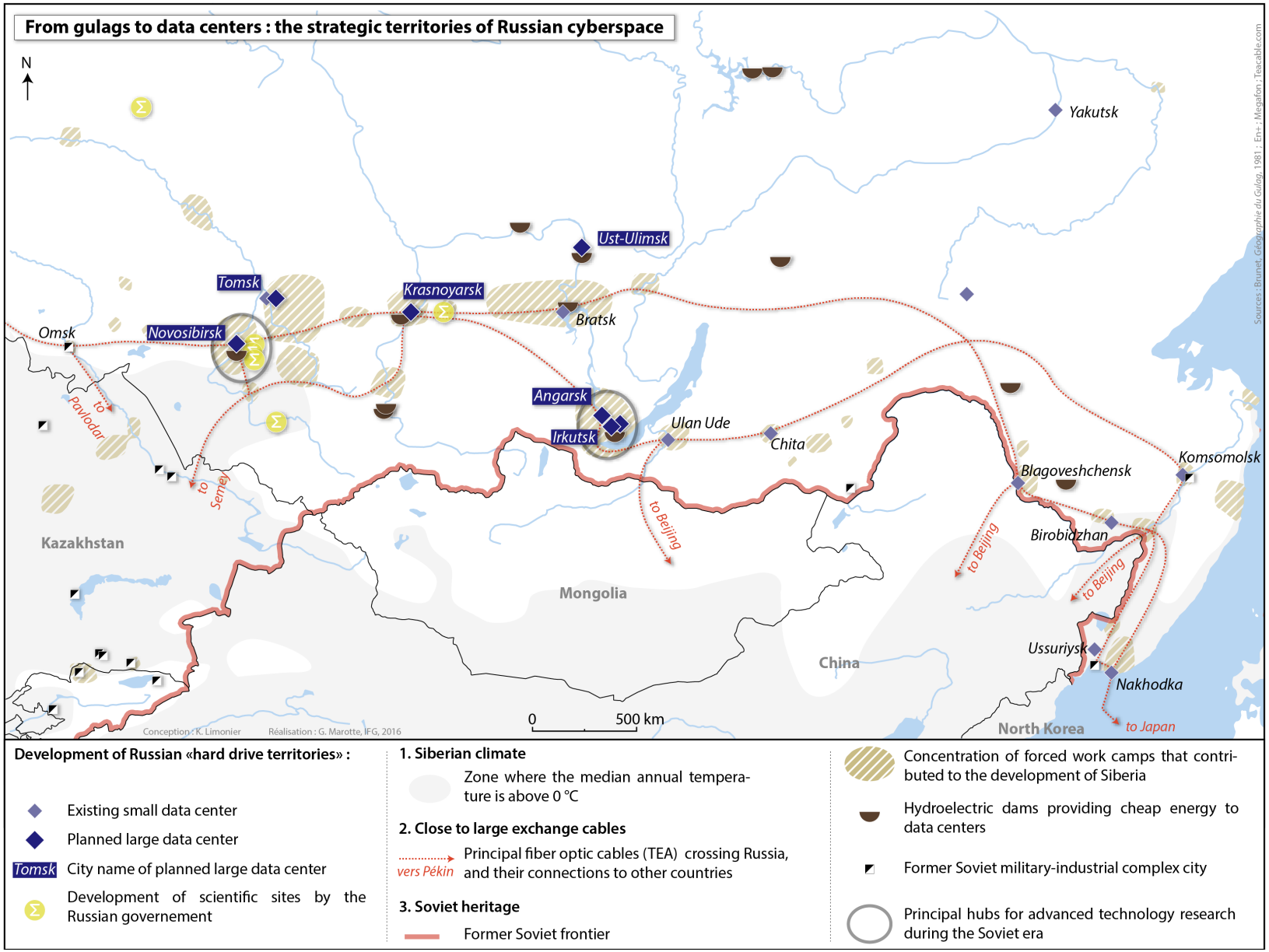
UNIVERSITÉ SAVOIE MONT BLANC

EURO-ASIA CONNECTIVITY

- 3 paths
  - Maritime via Suez
    - Traditional path with 90% of traffic
  - Land via russia
    - Expensive, low bandwidth
    - New silk road ?
  - Land-Maritime via russia-caucasus-Indian sea
    - New path
  - A fourth via artic sea
    - Not sure if it will be finalized

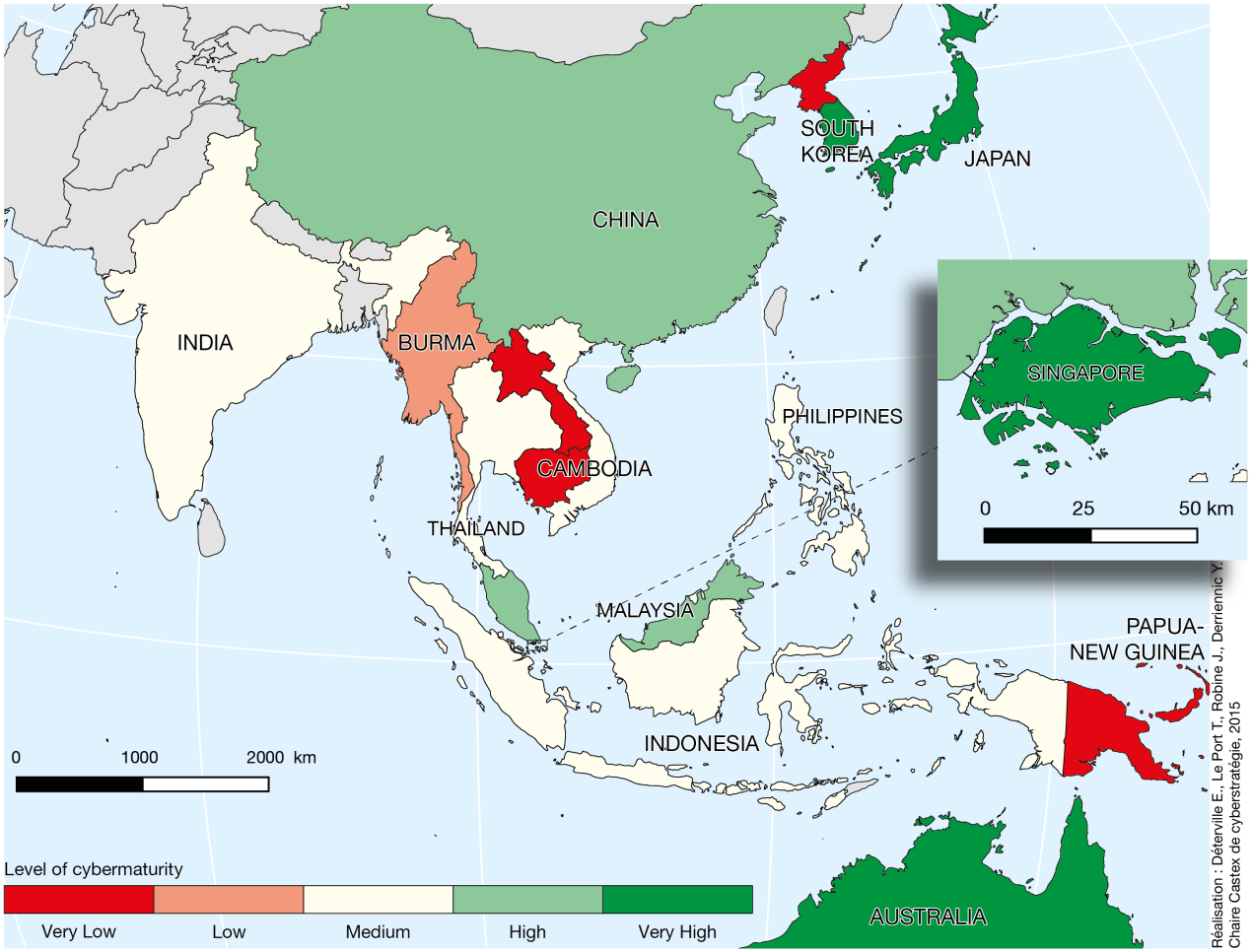


UNIVERSITÉ SAVOIE MONT BLANC



UNIVERSITÉ SAVOIE MONT BLANC

Cybermaturity of Asian Nations in 2015, According to ASPI
— Global Score —



Source : Cybermaturity in Asia Pacific, ASPI, 2015

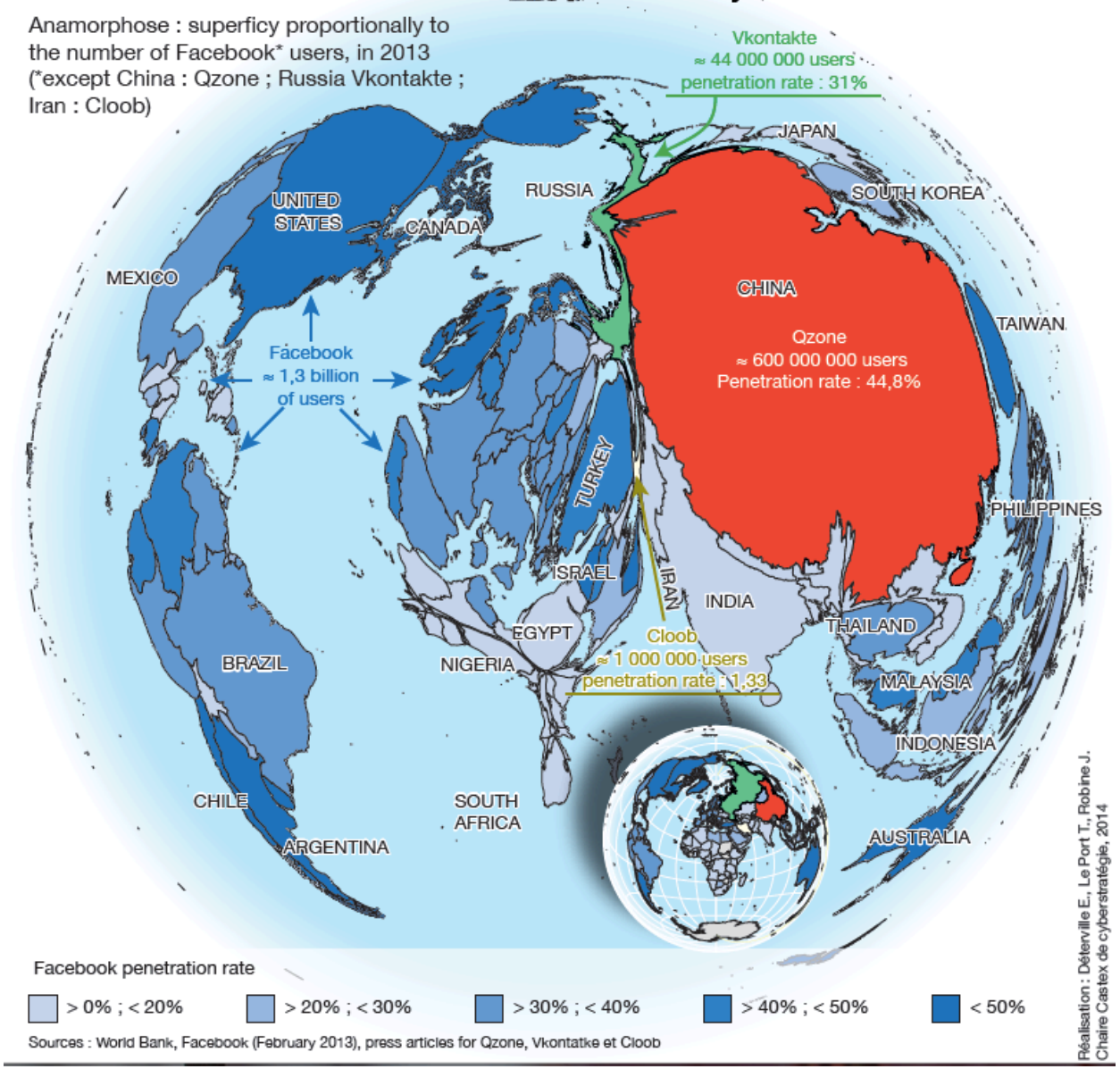
UNIVERSITÉ SAVOIE MONT BLANC

GEOGRAPHY EMBEDDED IN CYBERSPACE

- The cyberspace projects itself into a new geography
  - Traces of activities in cyberspace tells something about geography
- Intermediation
  - Projecting the real world into cyberspace information and leveraging on it to act on real world

UNIVERSITÉ SAVOIE MONT BLANC

2. - The social network sector dominated by Facebook



UNIVERSITÉ SAVOIE MONT BLANC

DATASET

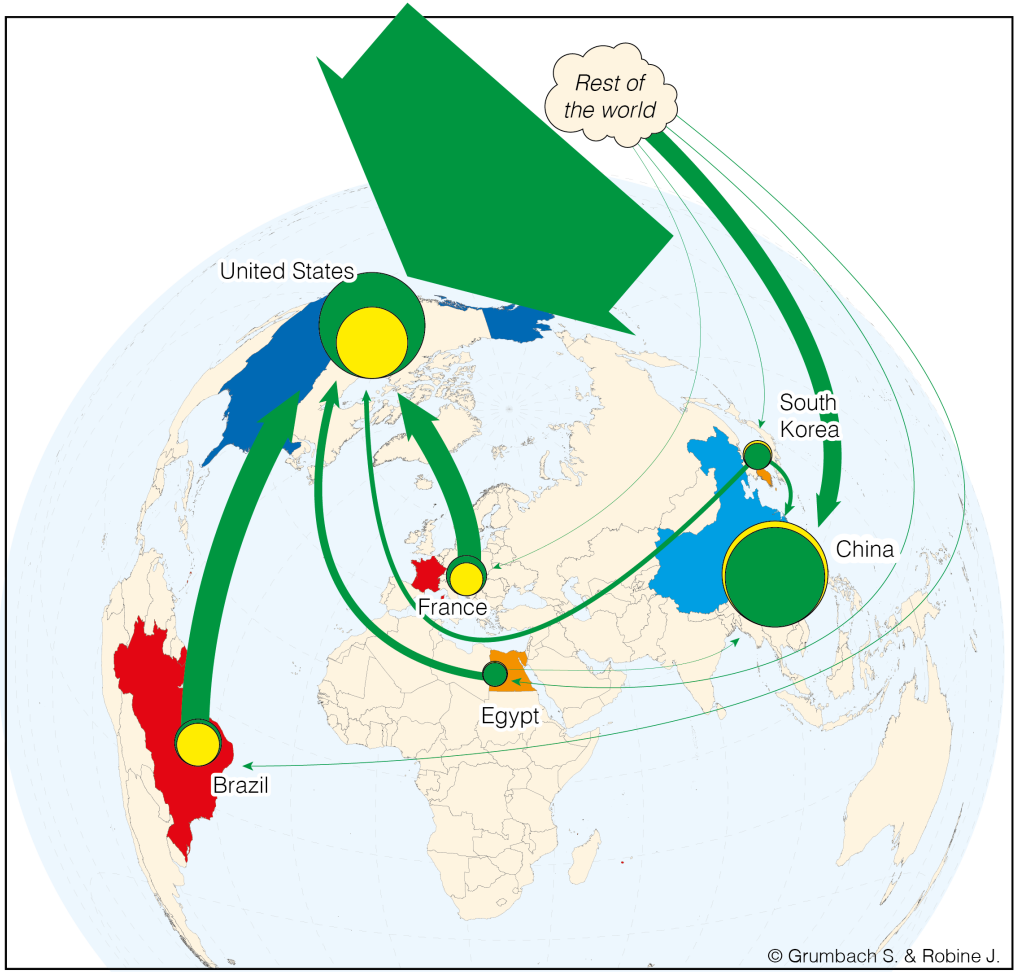
- All DNS requests forwarded to the DNS servers of a major mobile and ADSL Internet Service Provider, ISP, in China during two days, in July 2015.
  - Data gathered from servers located in 27 Chinese regions
  - Each record has five fields: a timestamp (rounded at 1 second timescale), the source IP sending the request, the domain name queried, resolution results, and the list of resolved IP addresses
- Advertisers and trackers identification
  - we merged three lists: EasyList, SimpleAd, and Simple Malvertising. Each
  - The EasyList contains 50 thousands URLs, the Simple Ad list contains 2703 sites, Simple Malvertising contains 5643 site

Num Records	Num IP	Num destinations
149,619,580,908	18,507,392	711,660,375

UNIVERSITÉ SAVOIE MONT BLANC

The USA, a hegemonic actor of data flows

Traffic in top 10 sites in selected countries



Internet users
(million people)

500
250
100
50
0

Traffic of top 10 sites
(million monthly visits)

5.000
1.000
500
250
0

Traffic of top 10 sites
(million monthly visits)

10
100
1.000
10.000

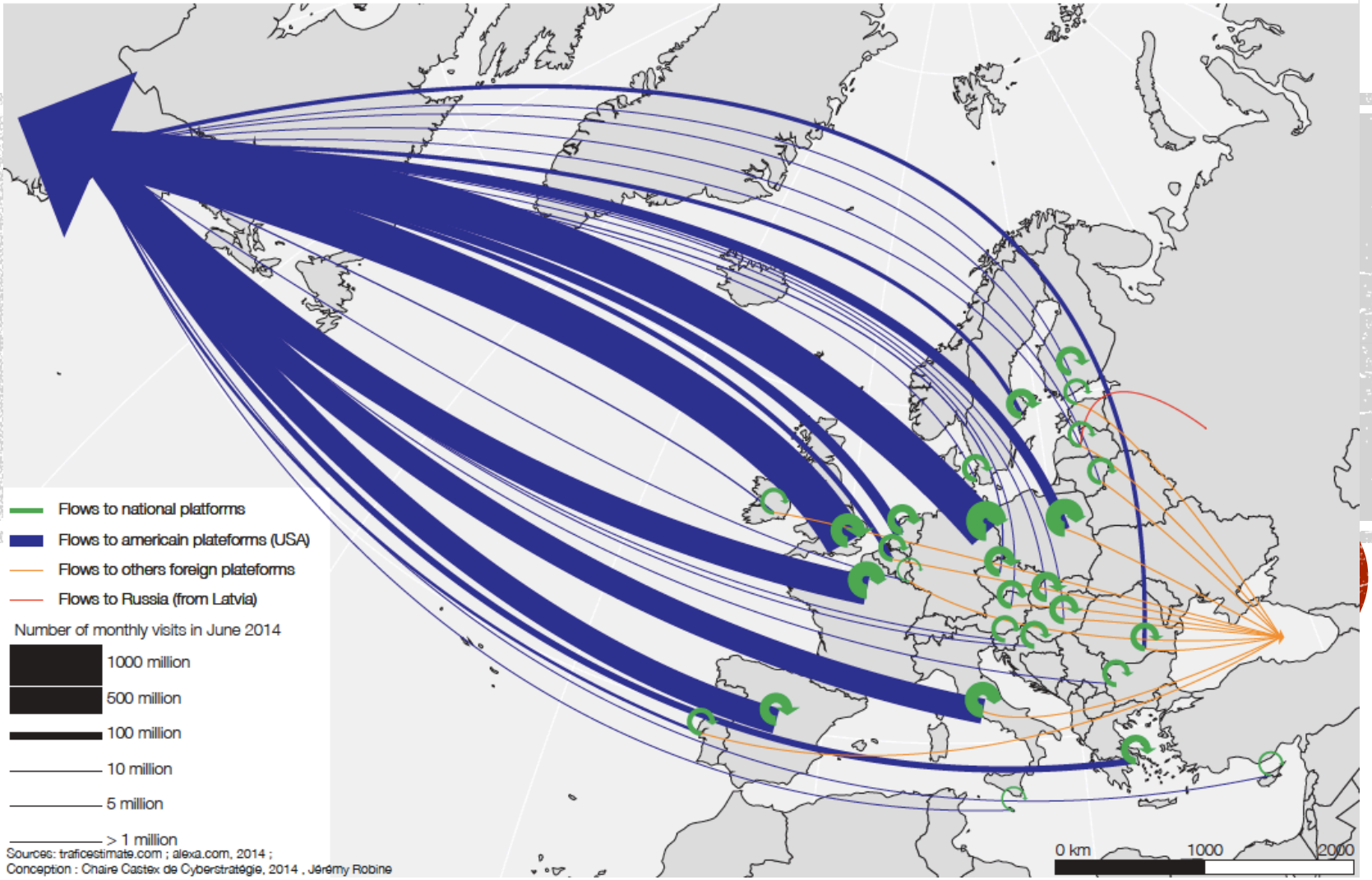
Balance inflow and outflow
(million monthly visits)

-∞ -500 0 +500 +∞

UNIVERSITÉ SAVOIE MONT BLANC

Data Flows in Europe

Estimation of the monthly visits of the 25 most visited websites from each of the EU 28 countries and distribution of their visitors by the websites' country of origin

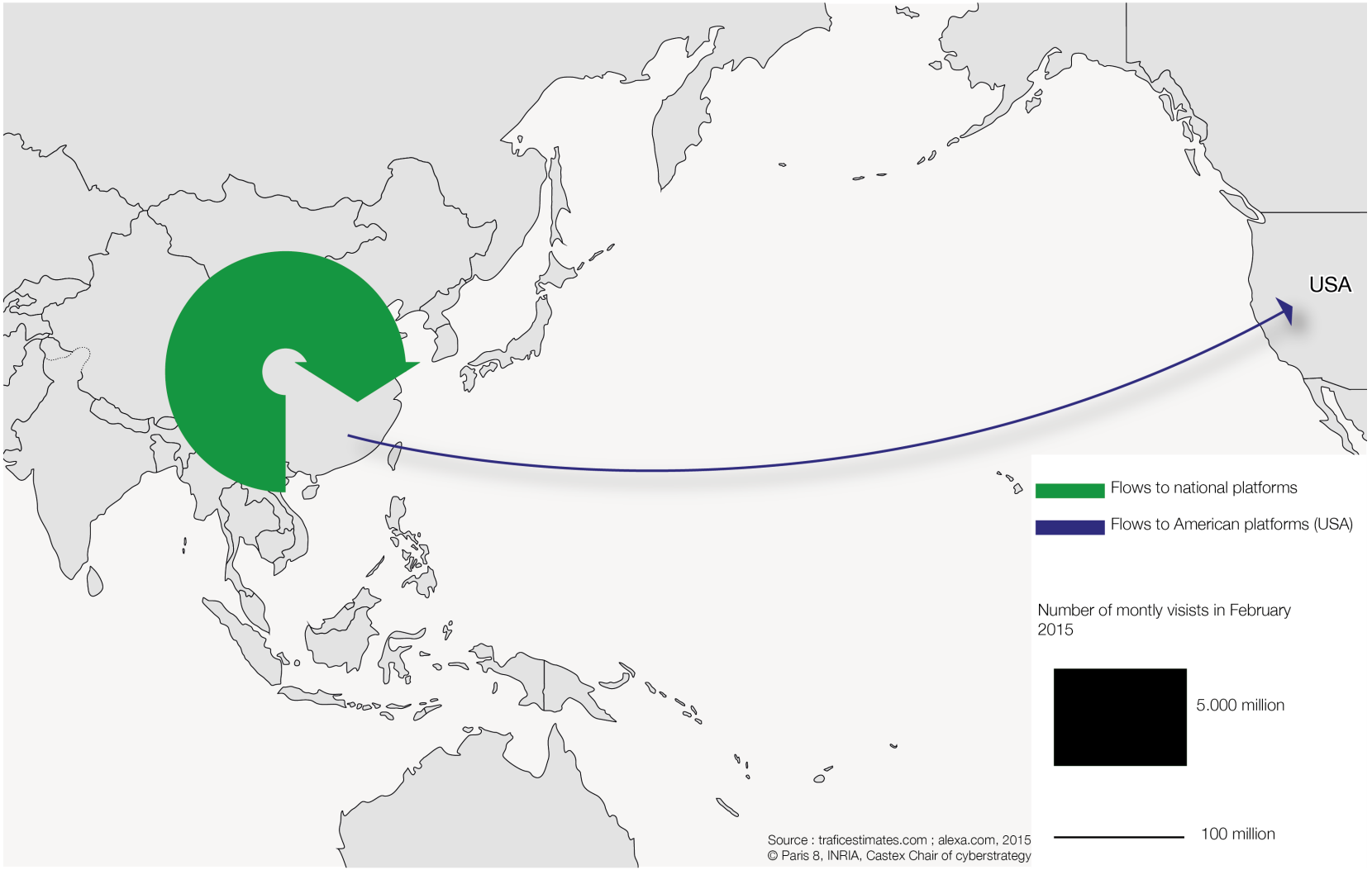


Sources: traficestimate.com ; alexa.com, 2014 ;
Conception : Chaire Castex de Cyberstratégie, 2014 , Jérémy Robine

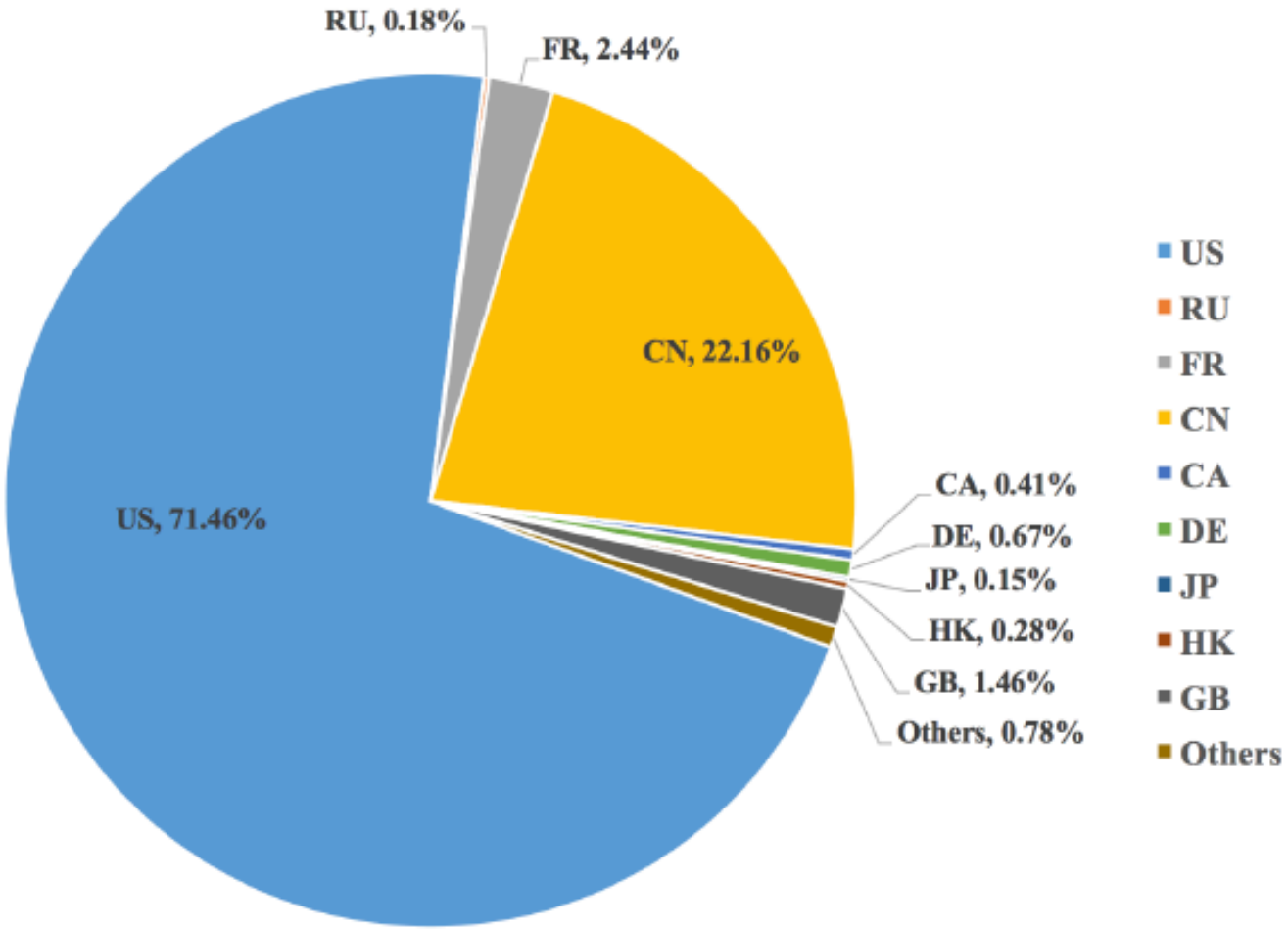
UNIVERSITÉ SAVOIE MONT BLANC

Data Flows in China, massively stay in China

Estimation of the monthly visits of the 25 most visited websites from China and distribution of their visitors by the websites' country of origin

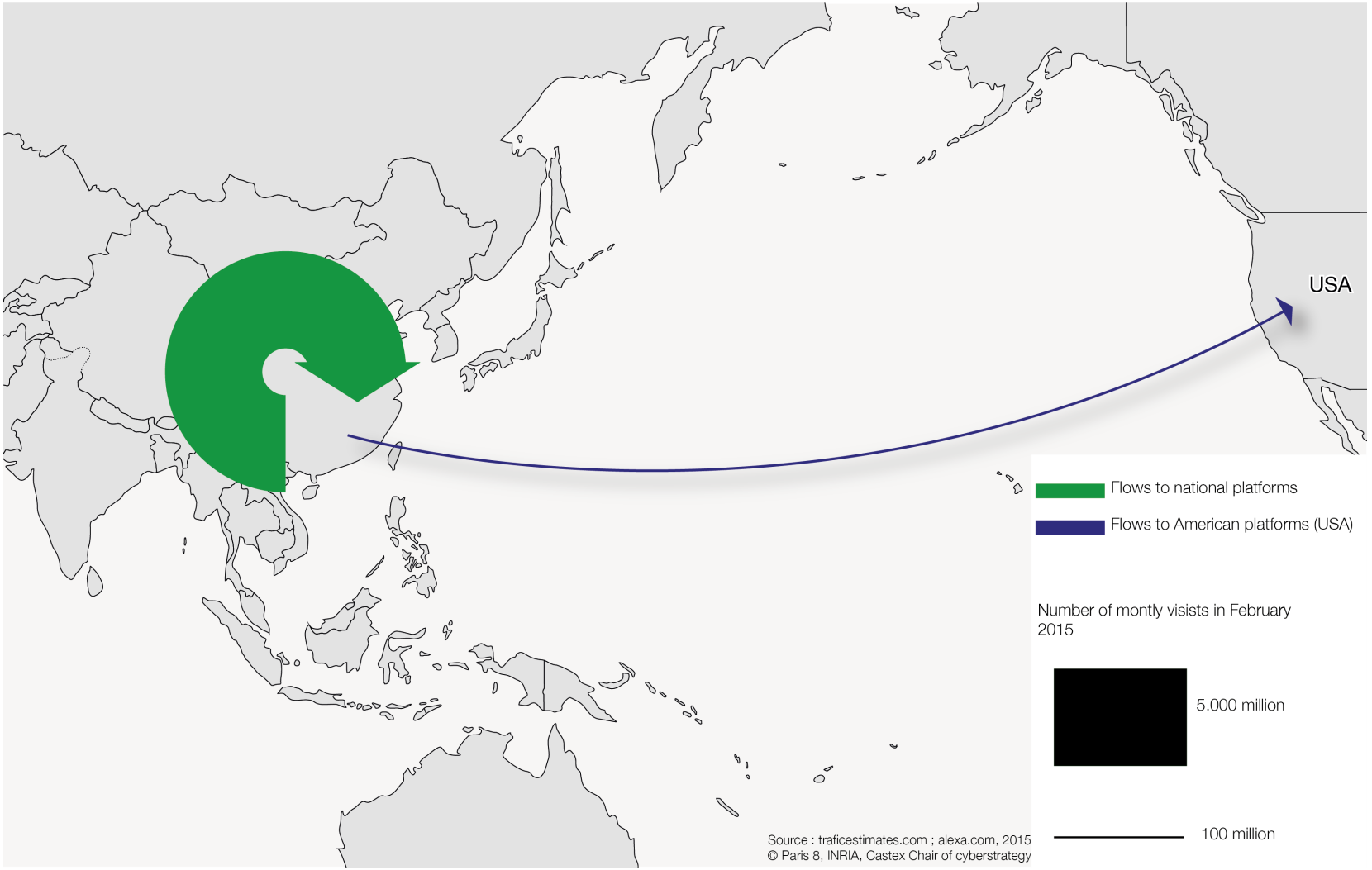


ADVS SHARE BY COUNTRY IN THE ADVS LIST

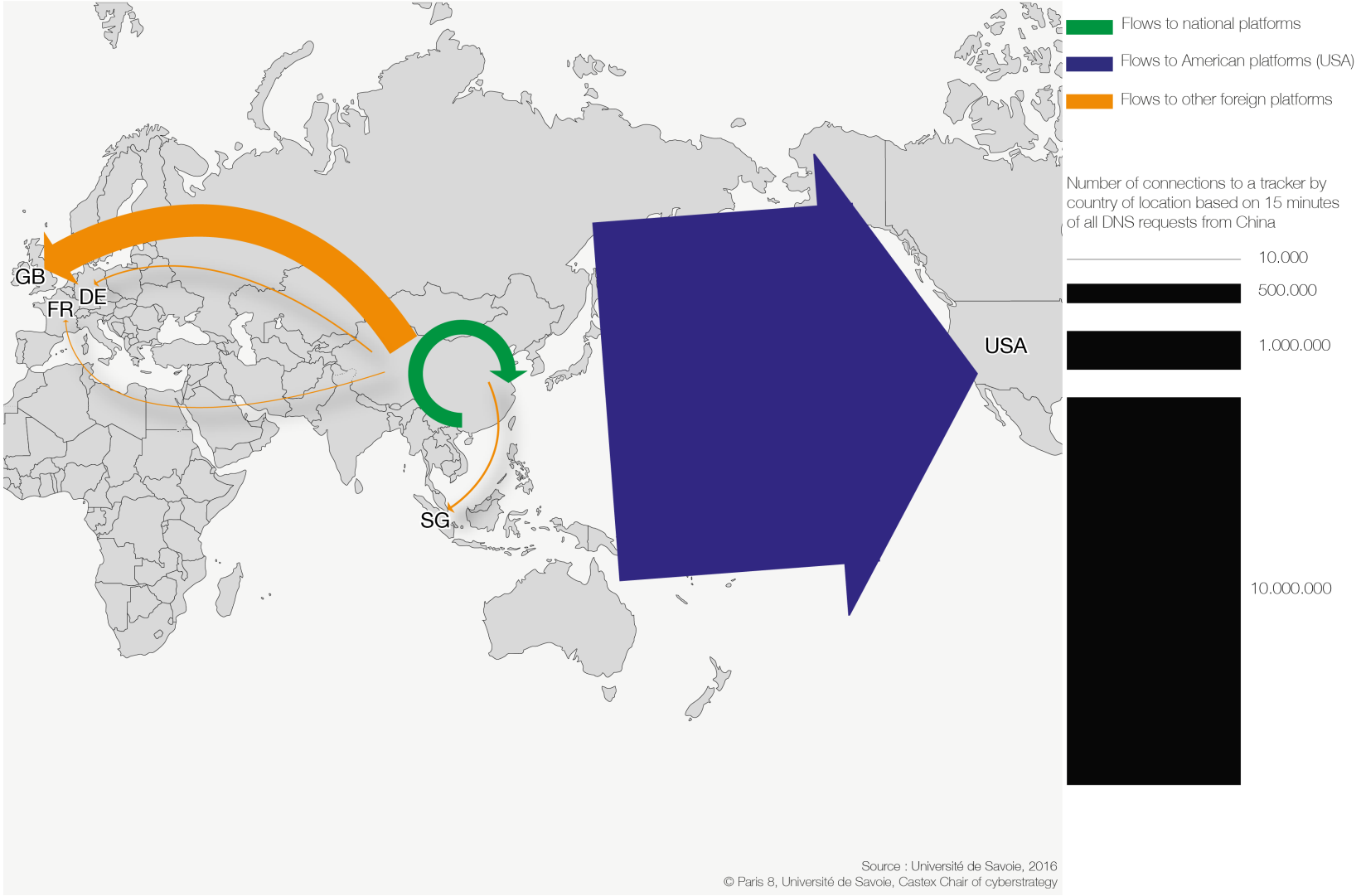


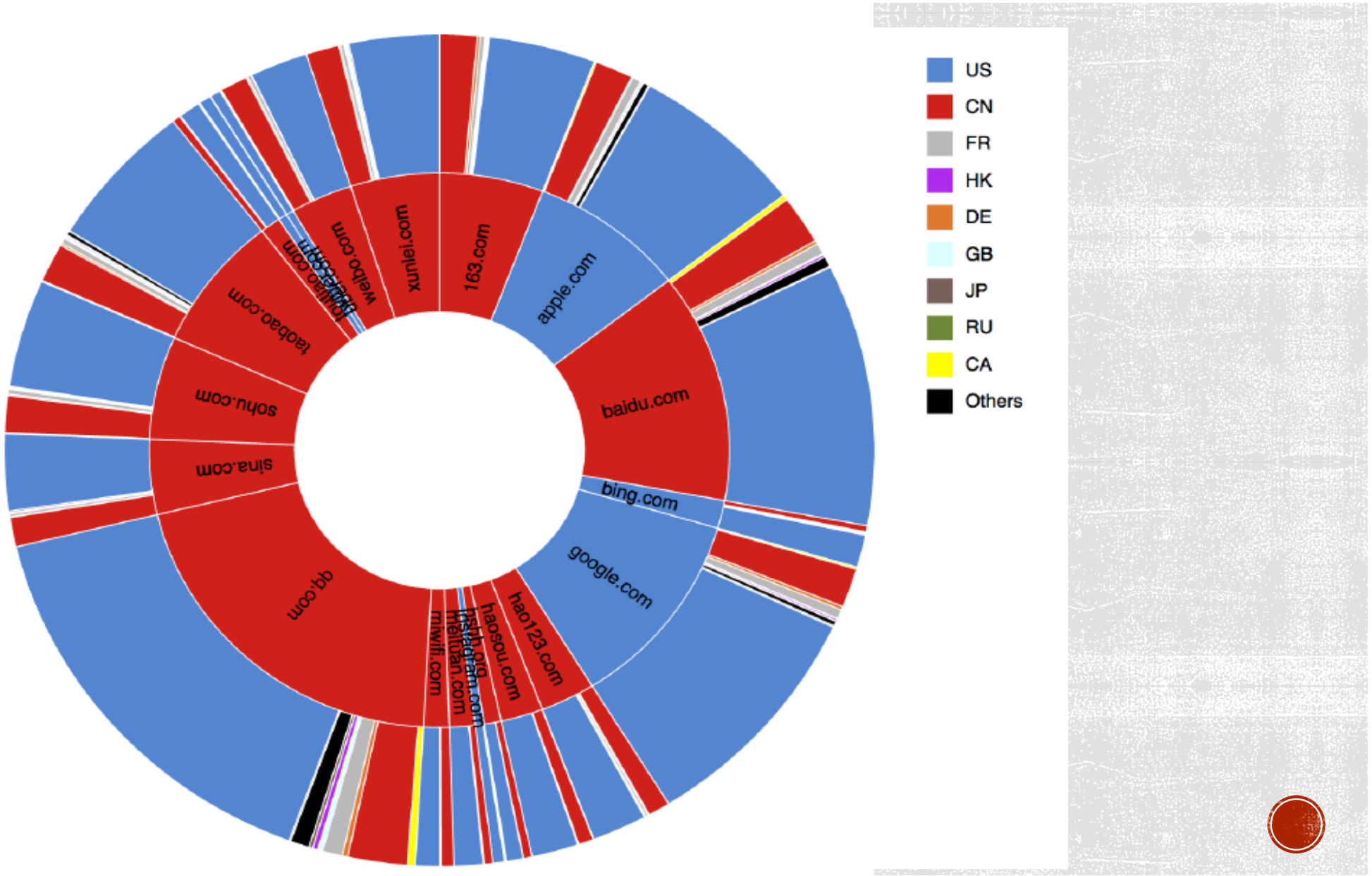
Data Flows in China, massively stay in China

Estimation of the monthly visits of the 25 most visited websites from China and distribution of their visitors by the websites' country of origin

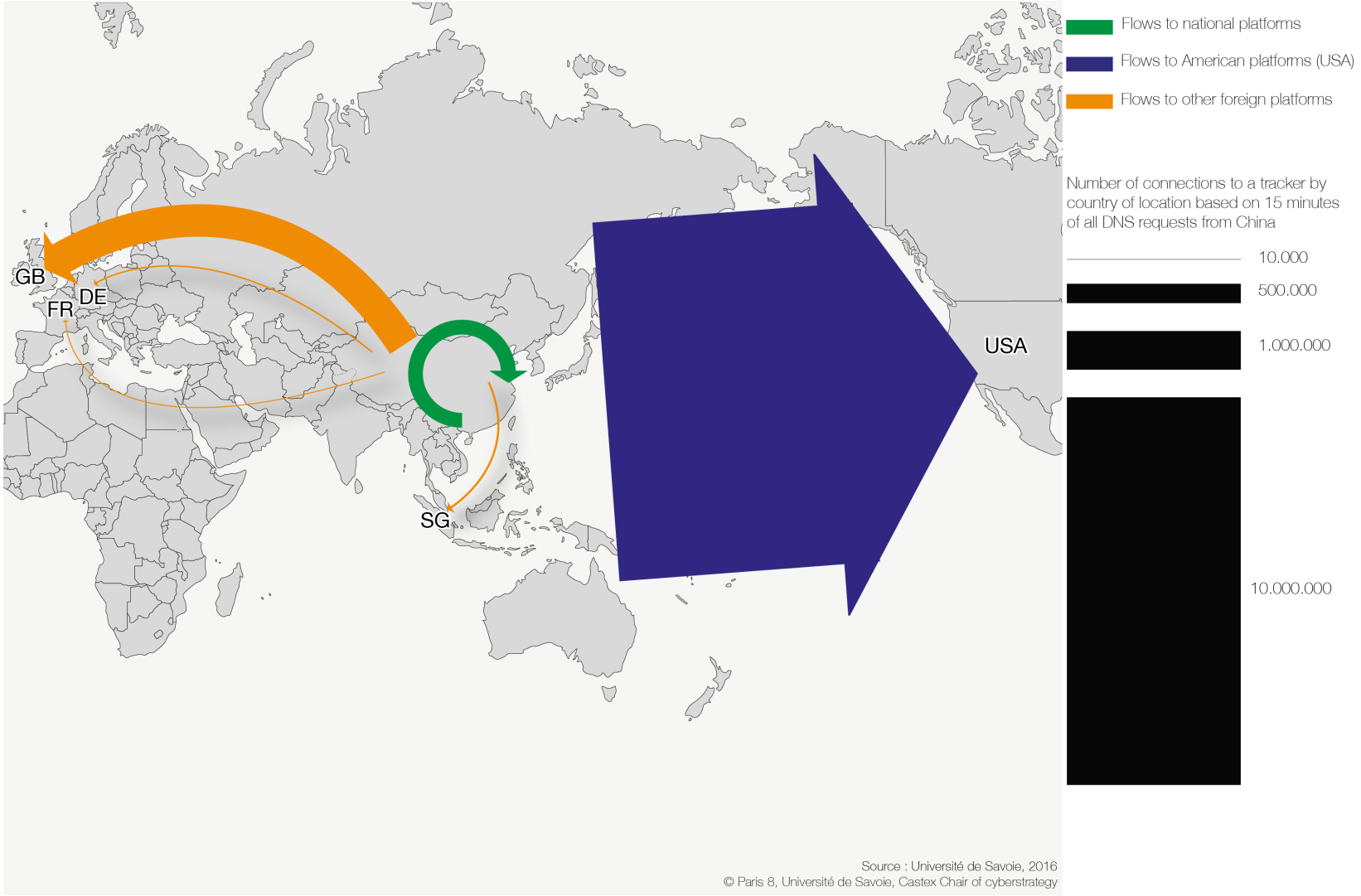


Trackers accessed from China, massively located in the US

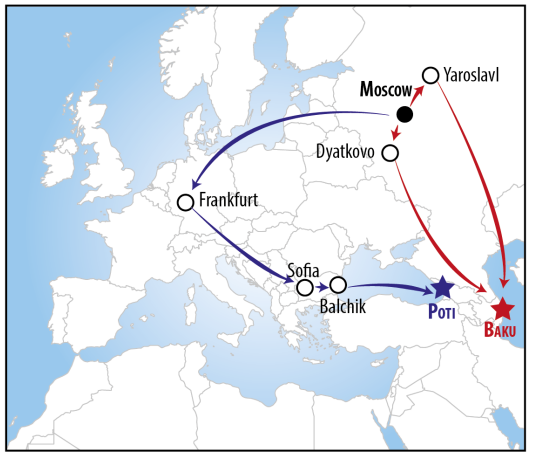


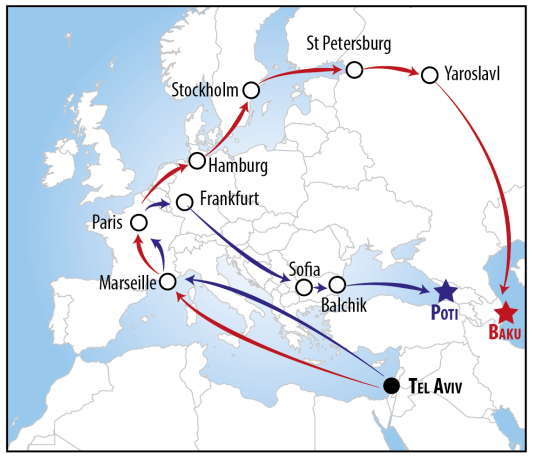

US
CN
FR
HK
DE
GB
JP
RU
CA
Others
xunlei.com
163.com
apple.com
baidu.com
bing.com
google.com
hao123.com
haosou.com
mwifi.com
taobao.com
sohu.com
sina.com
qq.com

Trackers accessed from China, massively located in the US



How data from various origins «travels» to Georgia and Azerbaijan



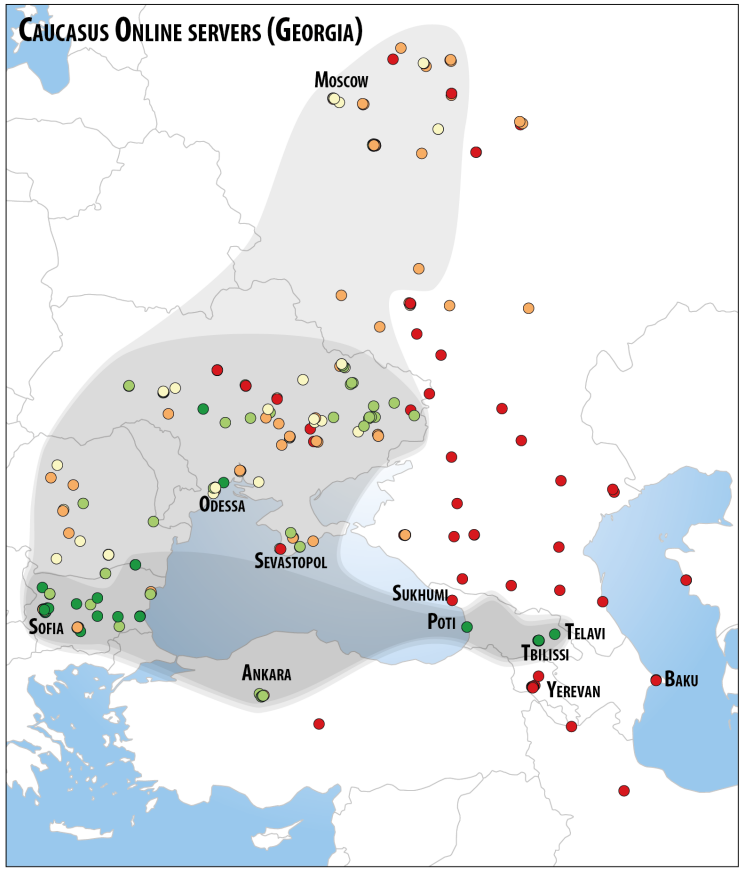


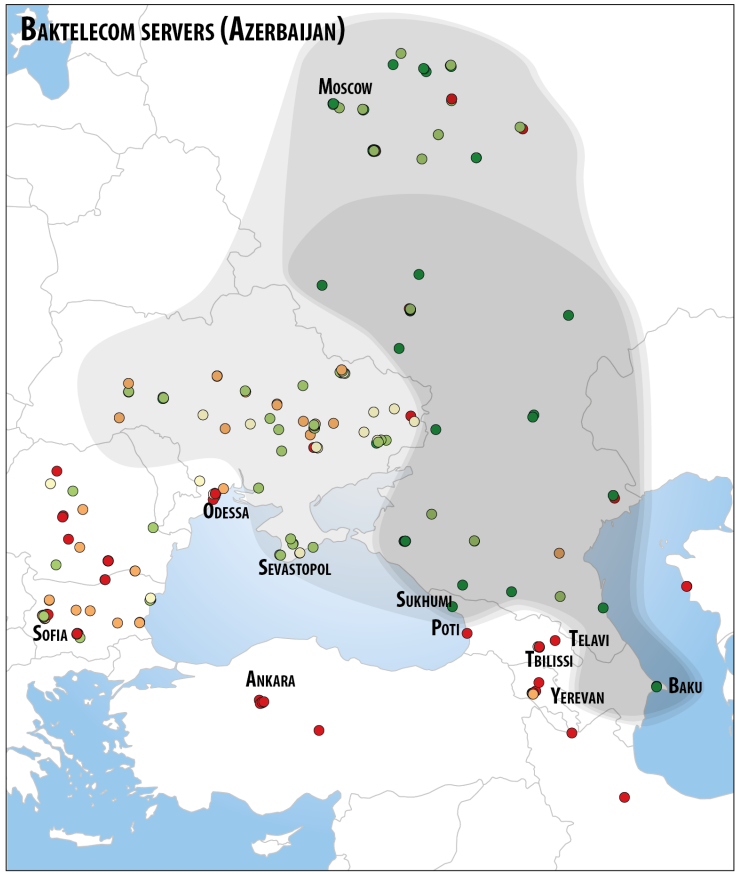
- Signal origin
- Intermediary connection points
- «Hops» heading to Azerbaijan
- «Hops» heading to Georgia

Sources : up to 10 000 traceroute requests using Nmap and RIPE Atlas probes on IPs from major Georgian and Azeri Autonomous Systems (AS). Maps and Datas : Kevin Limonier

UNIVERSITÉ SAVOIE MONT BLANC

TIME FOR DATA TO TRAVEL TO GEORGIA AND AZERBAIJAN





Latency (in milliseconds)

- 1 - 63
- 63 - 87
- 87 - 97
- 97 - 109
- 109 - 788

Average Latency Areas

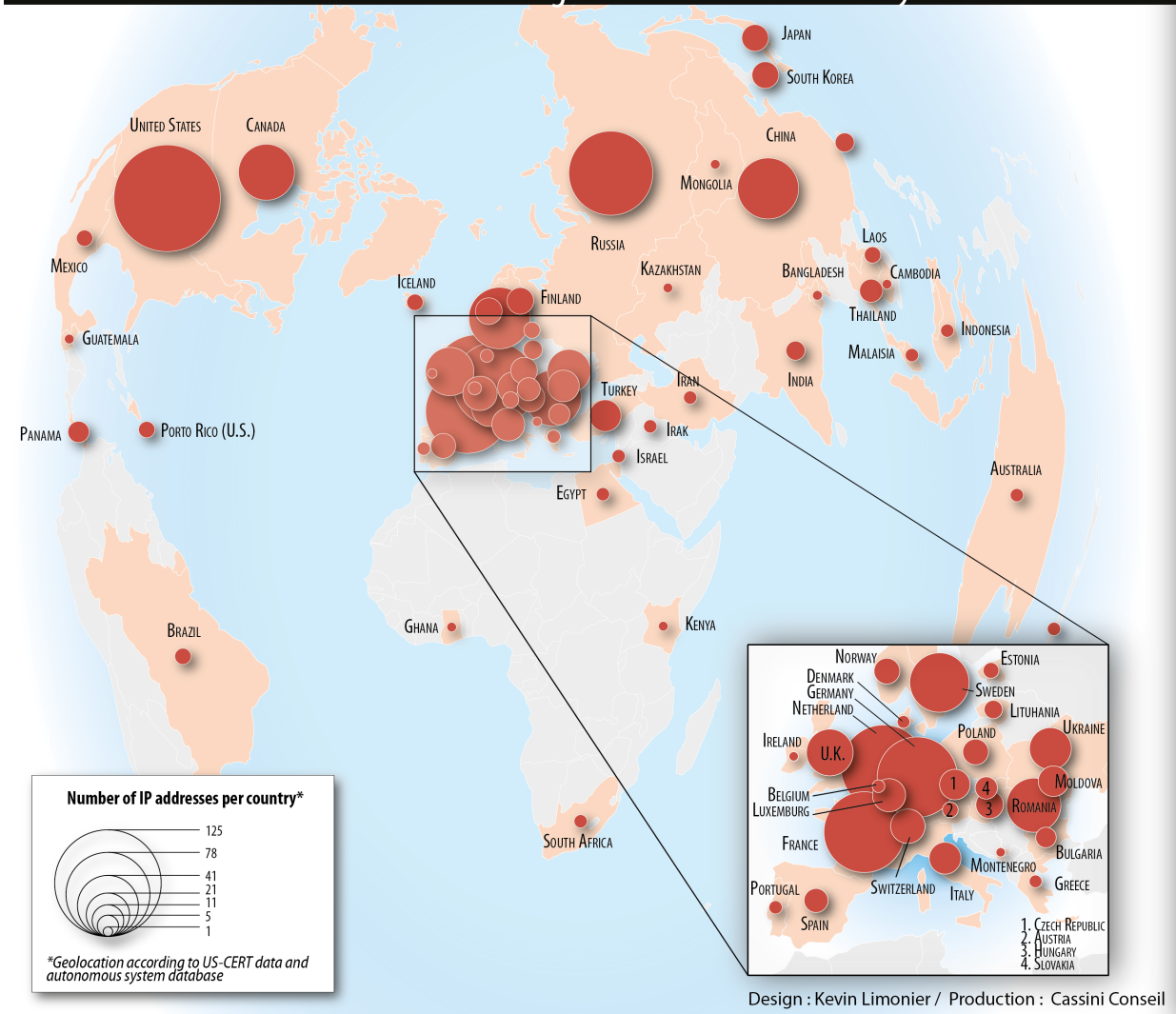
- Low speed
- Medium speed
- High speed

Sources : RIPE Atlas Network

UNIVERSITÉ SAVOIE MONT BLANC

Where does Russian Malicious Cyber Activity come from

according to US Homeland Security declassified files



Design : Kevin Limonier / Production : Cassini Conseil

On december 29, 2016, the United States Computer Emergency Readiness Team (US-CERT), an organization within the Department of Homeland Security (DHS), published declassified documents about a malicious cyber-activity attributed to Russia and known as GRIZZLY STEPPE. Among other data, US-CERT made public a list of approximately 890 IP addresses supposed to be used by Russian civilian and military intelligence services against American interests. This list, made for network administrators, contains information about the geographical origin of each suspicious IP address.

This map represents the worldwide repartition of these IPs, providing a visual representation of how US DHS may see the threat attributed to Russia. Of course, a large part of these IPs are likely to be proxies, that is to say intermediary infrastructures between the target and the source of the attack. These addresses could unintendedly be used as proxies for malicious activities, via infected machines.

Nevertheless, the map clearly shows the importance of Western European countries, as they host a major part of the suspicious IPs US-CERT revealed. Netherlands, Germany, France and Sweden host almost 30% of all these IPs, while only 78 (8%) are officially located in Russia.

Chaire Castex de Cyberstrategie
www.cyberstrategie.org

CASSINI
www.cassini-conseil.com

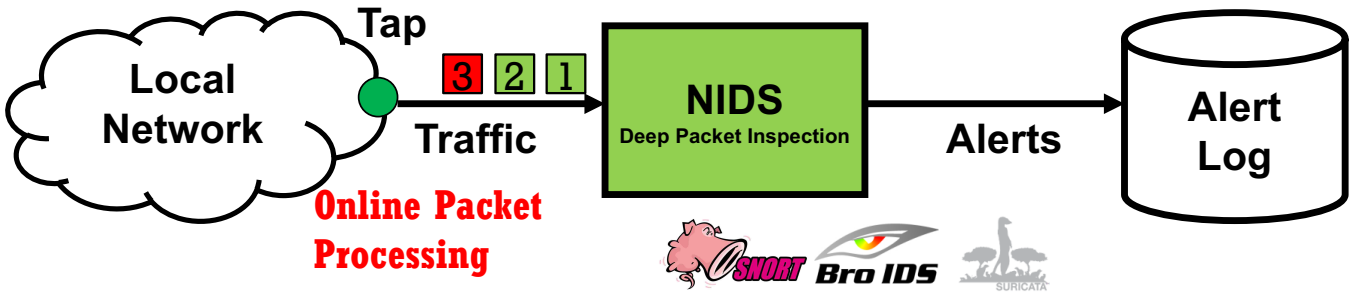
CYBERSPACE AS A SPACE ON ITS OWN

- Two examples
  - BGP
  - Net Neutrality

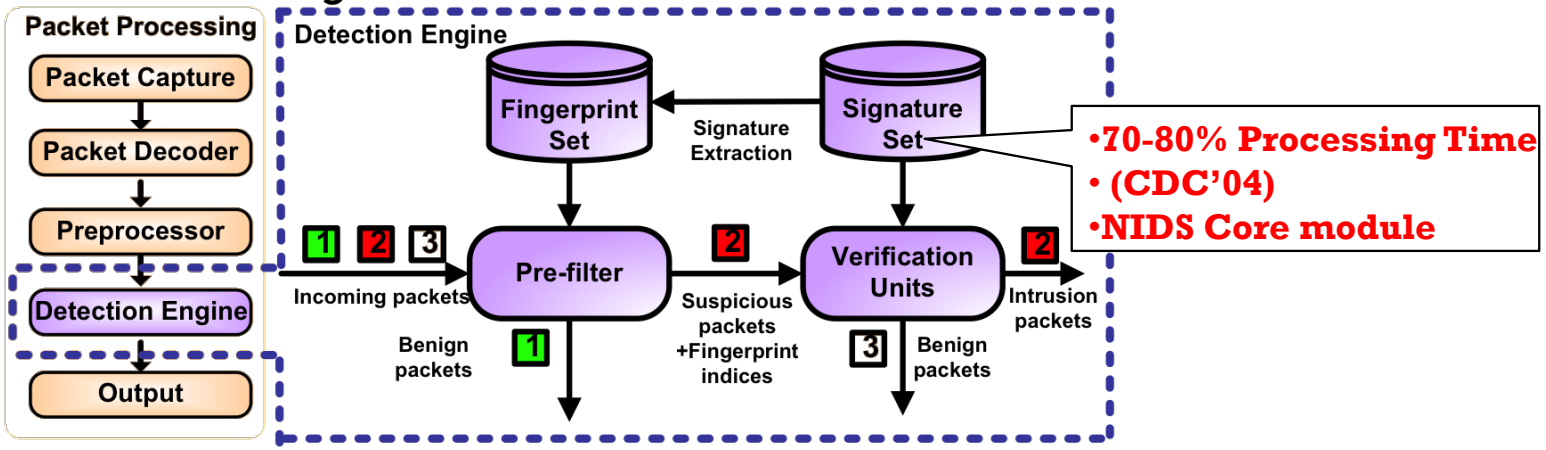
UNIVERSITÉ
SAVOIE
MONT BLANC

BACKGROUND-NIDS

- NIDS Deployment



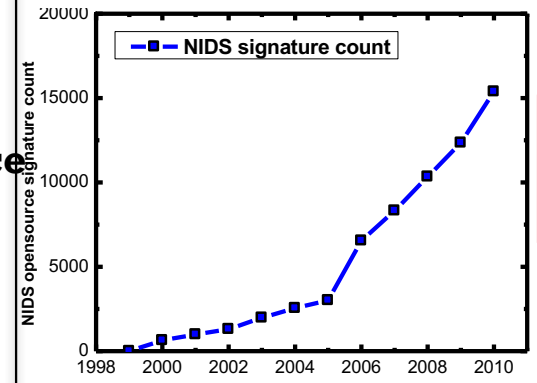
- NIDS Packet Processing



UNIVERSITÉ SAVOIE MONT BLANC

BACKGROUND-NIDS CHALLENGE

Performance Issue

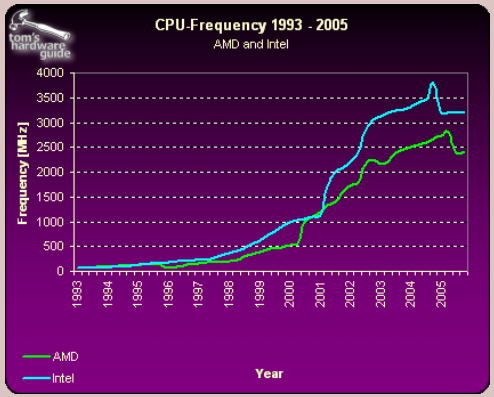


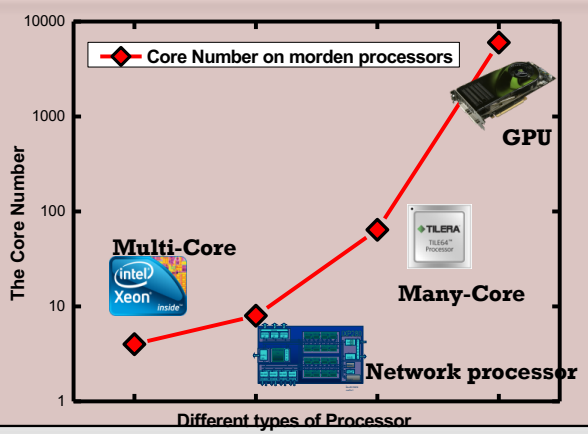
Grown Ruleset

CPU 2.5GHZ	Cycle/Byte
1Gbps	20 Cycle
10Gbps	2 Cycle
40Gbps	0.5 Cycles

Grown Traffic

Underlying Processors

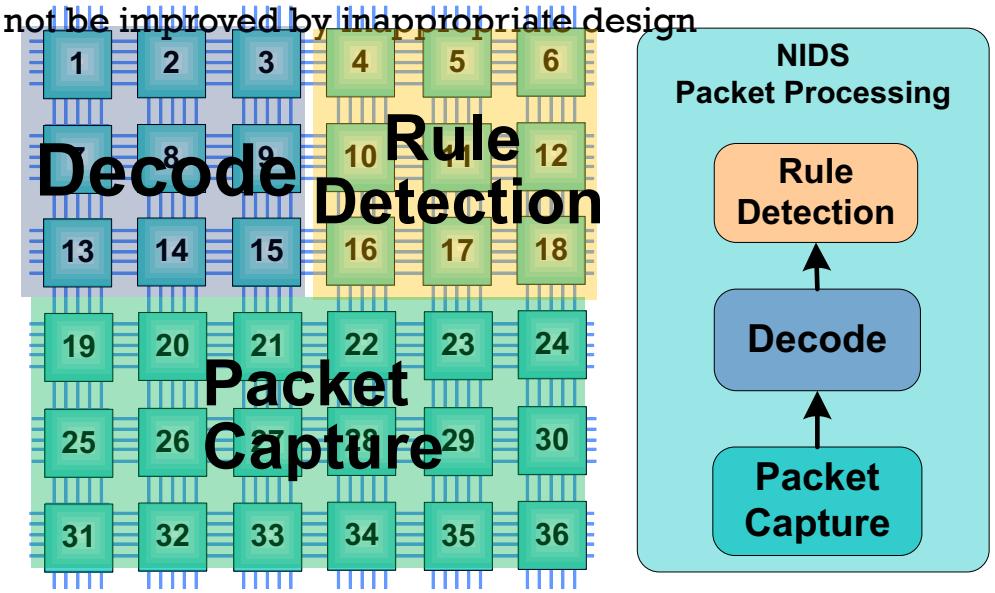




UNIVERSITÉ SAVOIE MONT BLANC

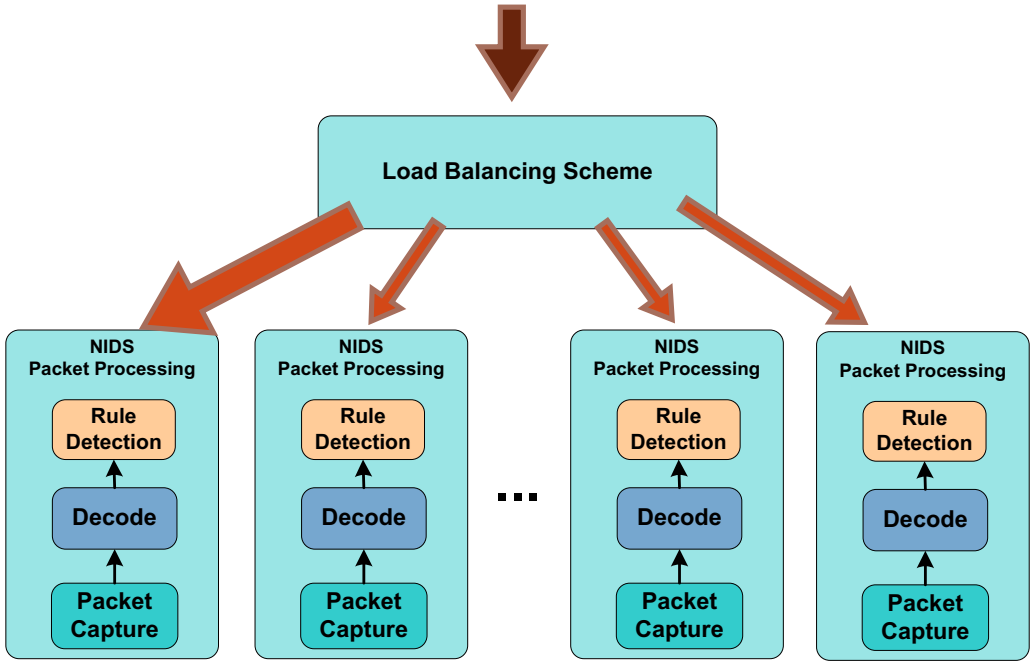
SOFTWARE BASED NIDS(1)

- Issue 1 : Parallel Design
  - Divide NIDS packet processing into steps and map these steps onto available computing resource
    - Performance can not be improved by inappropriate design



SOFTWARE BASED NIDS(2)

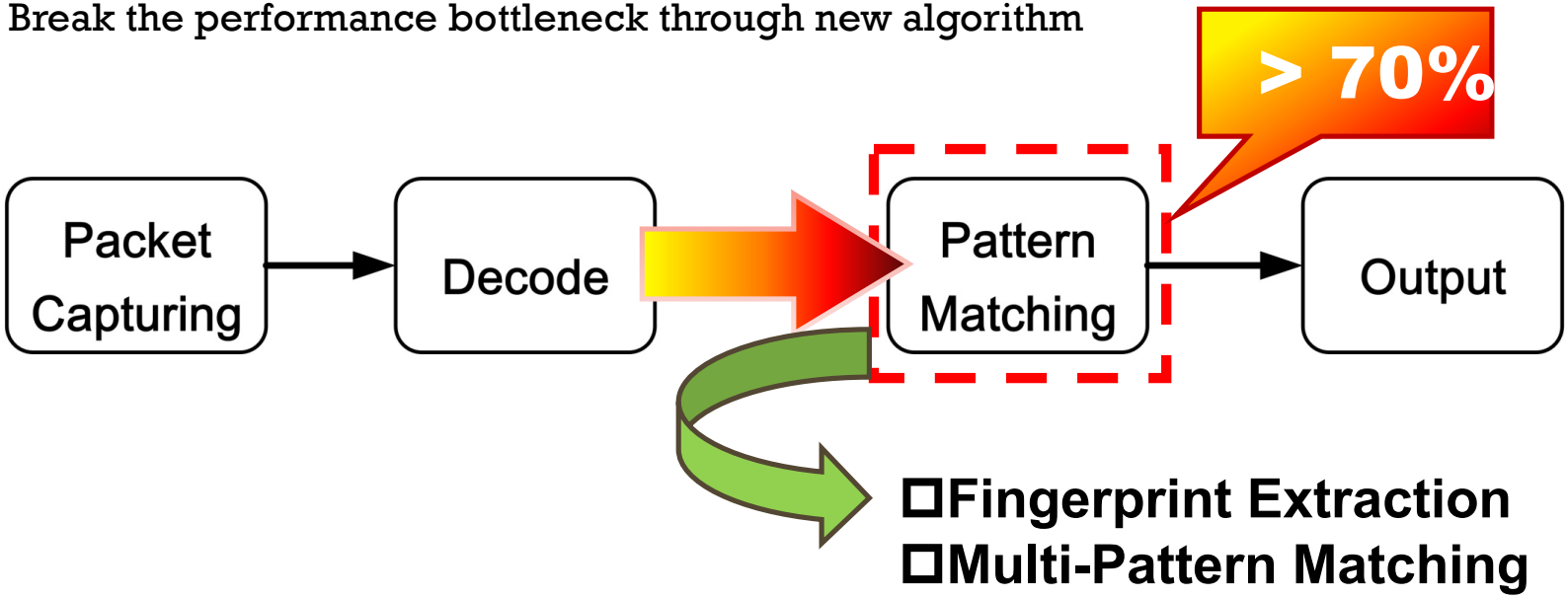
- Issue 2 : Load Balancing
  - Overload partition among the parallel NIDS components



UNIVERSITÉ SAVOIE MONT BLANC

SOFTWARE BASED NIDS(3)

- Issue3: Efficient pattern matching component
  - Break the performance bottleneck through new algorithm

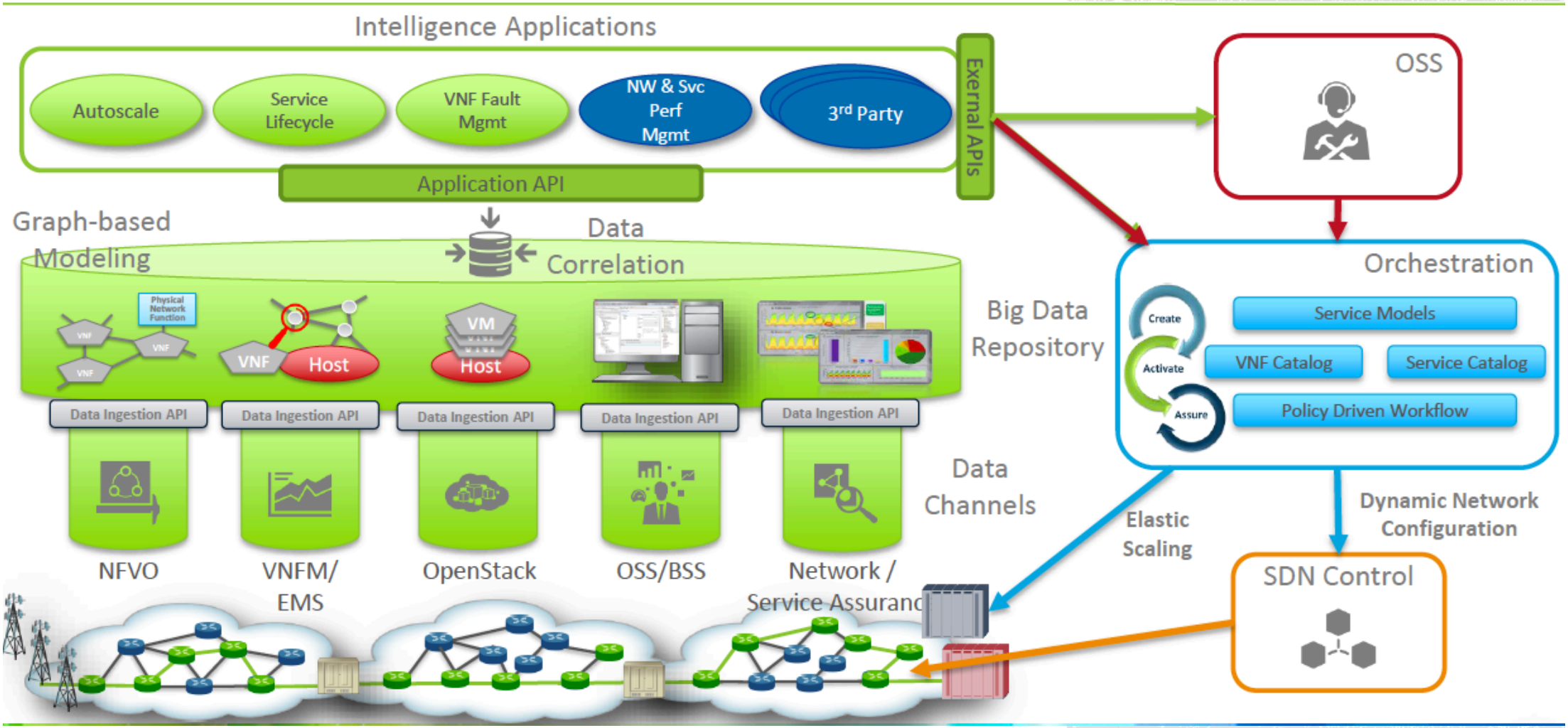


UNIVERSITÉ SAVOIE MONT BLANC

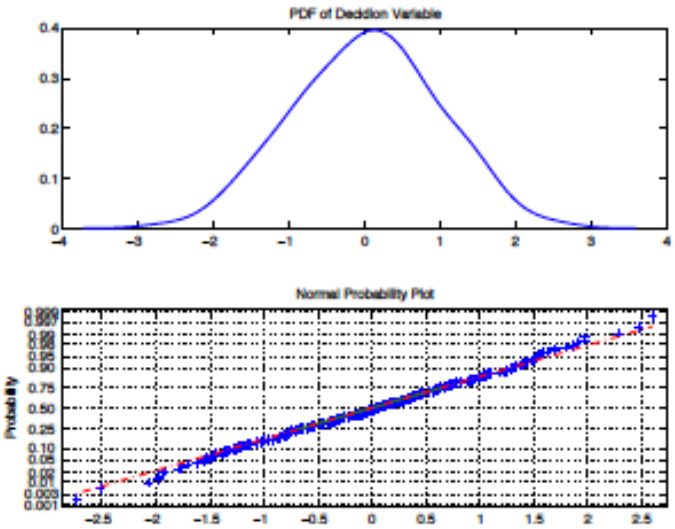
PERFORMANCE RESULTS

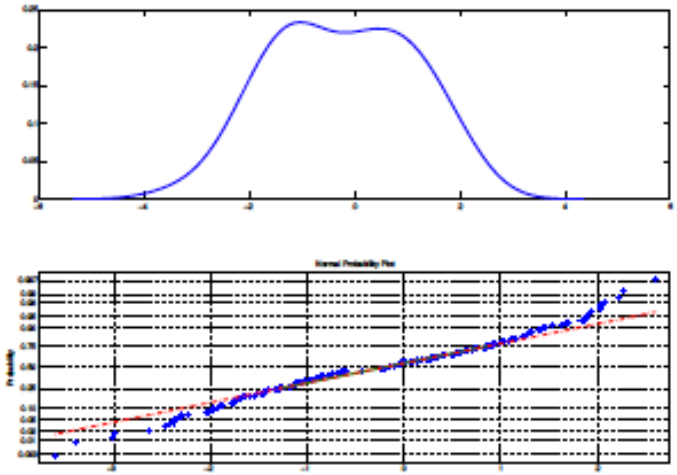
- NIDS Hybrid parallel design
  - Linear Speedup rate
  - 8.4Gbps processing ability when processing 100 bytes length packet
- RPB scheme
  - Resolve the load balancing issue due to the Internet traffic unbalancing
  - 42% performance upgrade
- Efficient fingerprint extraction scheme
  - 69% performance upgrade

UNIVERSITÉ SAVOIE MONT BLANC


Intelligence Applications
Autoscale
Service Lifecycle
VNF Fault Mgmt
NW & Svc Perf Mgmt
3rd Party
Application API
Exernal APIs
OSS
Graph-based Modeling
Data Correlation
Physical Network Function
VNF
VNF
VNF
Host
VM
Host
Big Data Repository
Orchestration
Create
Activate
Assure
Service Models
VNF Catalog
Service Catalog
Policy Driven Workflow
Data Ingestion API
Data Ingestion API
Data Ingestion API
Data Ingestion API
Data Ingestion API
Data Channels
Elastic Scaling
Dynamic Network Configuration
NFVO
VNFM/ EMS
OpenStack
OSS/BSS
Network / Service Assuranc
SDN Control

ANOMALY DETECTION 1

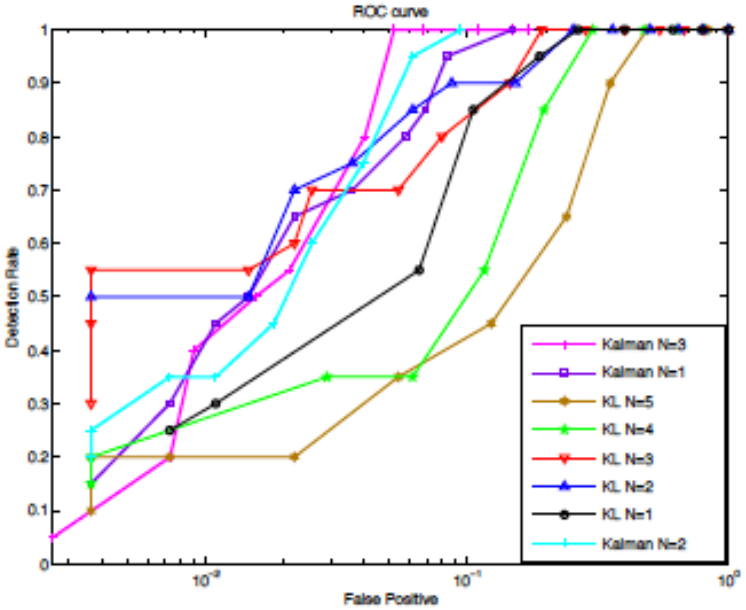


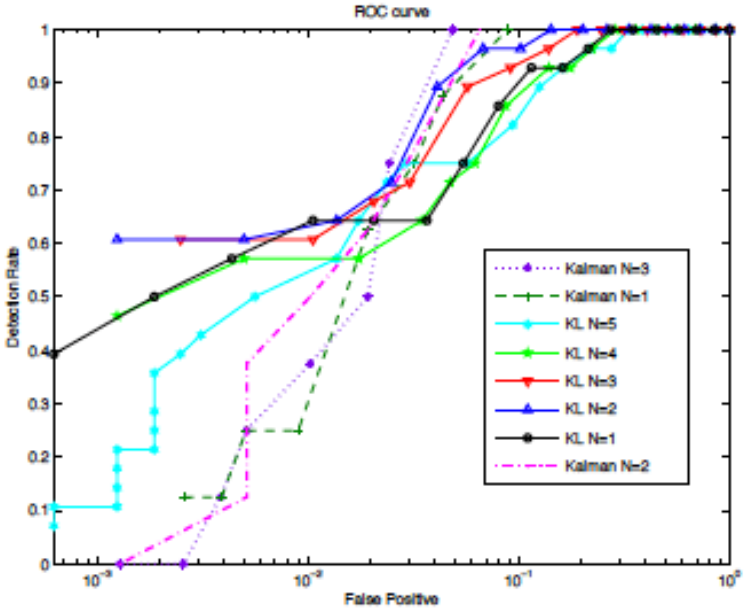


UNIVERSITÉ
SAVOIE
MONT BLANC

ANOMALY DETECTION 2

- Mix of methods work much better





UNIVERSITÉ SAVOIE MONT BLANC

SYSTEMS FOR BIG DATA

- Parallelisms
  - Granularity
    - Many cores, clusters, grids
- Hardware
  - TCAM, FPGA
  - Systems for heterogeneous architecture
- Mapping data to architecture
  - Data management
  - Algorithmics
  - Mapping resources to computation needs

UNIVERSITÉ SAVOIE MONT BLANC

CONCLUSION

- International collaboration is needed on cybersecurity
  - In particular information exchange is fundamental
- Security is not only a technical issue it is a strategic issue
  - We need to interact much more with other very important stakeholder
  - Interdisciplinarity is tough
- There is

- Many Thanks to Labex Persyval, Grenoble that funded my travel

UNIVERSITÉ SAVOIE MONT BLANC