Kentik turns network traffic into operational and business value.

**Executive Summary**

- **FOUNDED**: 2014
- **HQ**: San Francisco
- **CUSTOMERS**: 200+
- **TEAM MEMBERS**: 70+
- **RUN BY**: Network and measurement nerds
- **GROWTH**: 20x since January 2016
- **FOCUS**: SPs and Enterprise
- **TECHNOLOGY**: In-house bg data platform
  Delivered as a service
Kentik Data Engine
Millions of Flow Records / Second

Network
Systems
Applications

Context Enrichment
User, Location, Source

Kentik Platform

Automate
ML-Driven Insights and Alerts

Investigate
Real-Time Data Instant Queries

Integrate
Network Observability and Intelligence Use Cases

**App-Aware Infrastructure Operations**
- Traffic Engineering
- Performance Management

**Application Operations**
- App vs. Network Performance
- Application Governance

**Attack Detection, Mitigation & Investigation**
- Threat Detection
- DDoS Defense
- Digital Forensics

**Business Operations**
- Service Creation
- Cost Management
Adding context by enriching traffic data
What do we want?

Enrichment:

Going beyond the basics, and contextualizing network data.
How do we do it?

Tags and tables:

Dynamic routing tables of context - what does the traffic *mean*?
Making Data Useful: Basic Flow Enrichment

### SNMP
- Interface names/desc
- Interface capacity
- IP addressing

### (MP)BGP
- AS Path
- Next Hop
- VRF/VXLAN MPLS

### GeoIP
- IP → ASN
- ASN → AS Name

### Customs
- Ports/Protocols
- Address Families
- Curated data-feeds

Enrichment at Ingest

NetFlow v*, IPFIX, sFlow, Flow logs, LB logs

Enriched, Context bearing Flow record
Next Up: Infrastructure Context

**1st Class Citizen / Built-In**

<table>
<thead>
<tr>
<th>Interface Classification</th>
<th>Network Classification</th>
<th>Custom AS Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Customer/Provider tagging</td>
<td>Custom Geo</td>
</tr>
</tbody>
</table>

**Overlay Service, Threat Data**

<table>
<thead>
<tr>
<th>Clouds</th>
<th>Threat feeds</th>
<th>Applications tagging</th>
</tr>
</thead>
</table>

**Application**

| k8s, Orchestration | CDN Logs | Istio, nginx, Load Balancer |
## Next Up: Infrastructure Context

### 1st Class Citizen / Built-In

<table>
<thead>
<tr>
<th>Interface Classification</th>
<th>Network Classification</th>
<th>Custom AS Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inside/outside directionality</td>
<td>e2e directionality</td>
<td>Networks w/ multiple ASNs</td>
</tr>
<tr>
<td>Connectivity type</td>
<td>Customer/Provider tagging</td>
<td>Private ASNs</td>
</tr>
<tr>
<td>Provider vs Customer</td>
<td>CRM meets flows</td>
<td></td>
</tr>
</tbody>
</table>

### Custom Geo

- Country groups/Markets
- Sub country groups

### Clouds

- ISP Embedded + Self-hosted CDNs
- Cloud providers

### Threat Data

- Threat feeds
  - Botnets
  - Infected hosts
- Applications tagging
  - OTT services
  - Well known Apps

### Application

- k8s, Orchestration
- CDN Logs
- Istio, nginx, Load Balancer
Data Fusion
The How

- LOTS of routing tables!
- Can be IP, MAC, VLAN, device, interface, BGP, or other traffic field-based
- Up to dozens of tables with millions of entries, per customer
- Live updating in real time through ingest system
- With persistence
- Tables must be synced with load balancing
- Includes global tables of Geo, threats
- In production - millions of tags * millions of FPS, ~20ms avg update
Enrichment requires cross-disciplinary skills

- Distributed Systems Software Engineers
- Network Protocol Developers
- UX Designers
- Frontend Software Engineers
- Site Reliability Engineers
- Network Engineers
How to make more unicorns?

- Distributed Systems Software Engineers
- Network Protocol Developers
- Network Engineers
- UX Designers Frontend Software Engineers
- Site Reliability Engineers

UNICORN
What does enrichment enable?
Kentik CDN / OTT Tracking
Match DNS request with flows in real time

1. What is the IP of download.updates.com

2. Live DNS requests
   - Requester: subscriber IP
   - Hostname: download.updates.com
   - IP: 1.2.3.4

3. Flow ingest
   - dst_ip: subscriber IP
   - src_ip: 1.2.3.4
   - src_OTT_service: OTT Download

4. Kentik curated DB hostnames → OTT Provider
   - download.updates.com
   - live.music.com
   - IP: 1.2.3.4

ISP Network

Commercial CDN #1

OTT Download
download.updates.com
IP: 1.2.3.4

OTT Music
live.music.com
IP: 1.2.3.4

DNS tap

OTT Music
live.music.com
IP: 1.2.3.4

Kentik CDN / OTT Tracking
Match DNS request with flows in real time
BGP Ultimate Exit
Examining traffic exit points on your network

All flow tools allow you to view where traffic enters your network, can your tool show you where it leaves?

BGP-UE uses iBGP data from your devices to determine where flows exit, allowing you to get a deeper understanding of your traffic.

Is my CDN customer keeping traffic local?
Is my transit customer using expensive subsea capacity?
Find out with Ultimate Exit.
Enrichment on Traditional Networks: Interfaces on Devices

Enriched flow record

src_int:  {INTERNAL, BACKBONE, n/a}
dst_int:  {EXTERNAL, TRANSIT, LEVEL42}
### Useful Enrichment: Interface Classification

#### INTERFACE DESCRIPTION (SNMP)

- [TRANSIT][AS33356][Level42][more description text]

#### DESCRIPTION MATCH REGEX (Enrichment engine)

- `^\[TRANSIT\] [.\*] [.\*] \.$`

#### INTERFACE CLASSIFIERS

- **SET INTERFACE NETWORK BOUNDARY:** EXTERNAL
- **SET INTERFACE CONNECTIVITY TYPE:** TRANSIT
- **SET INTERFACE PROVIDER:** (LEVEL42)

---

**Enriched flow record**

- **src_int:** {INTERNAL, BACKBONE, n/a}
- **dst_int:** {EXTERNAL, TRANSIT, LEVEL42}
Useful Enrichment: Interface Classification
FLOW RECORD:
- Ultimate Exit (country, site, device, interface): {country, site, MyRouter#2, customer#3}
- Ultimate Exit Connectivity Type: customer
- Ultimate Exit Connectivity Provider/Customer: CUSTOMER#3
Ultimate Exit Discrimination

SOURCE PROVIDER | INGRESS SITE | ULTIMATE EXIT SITE | ULTIMATE EXIT CONNECTIVITY TYPE | ULTIMATE EXIT PROVIDER
---|---|---|---|---
customer #1 | Seattle | | | customer_1
| Trans-Atlantic | Amsterdam | | customer
| Domestic Long Haul | Los Angeles | | customer_2
| Trans-Atlantic | Paris | | customer_3
| | Denver | | customer_4
| | Vancouver | | customer_5
| | New York | | ix_peering
| | France IX | | customer_6

Revenue Generating

Hot Potato

Moderate Cost
Subscriber Tracking: Implementation

ISP NETWORK

Large IP-to-subscriber in-memory DB
N x Million IPs, churning

Ingest API

+ IP1 → Sub1
- IP2 → Sub2

INGEST + ENRICH

Flows inbound

to Flow Datastore
OTT Service Tracking
Mark flows with Src/Dst OTT services being used
CDN Attribution
Discovering sources of traffic in neighboring networks

Is all the traffic from my peer actually from their customers or from embedded CDNs?

Using DNS data extracted from your network, we can determine if your neighboring ASNs have embedded CDN caches and what percentage of the overall traffic they represent.

Is my CDN customer delivering traffic locally?

Find out with CDN Attribution.
## OTT Tracking in Action

<table>
<thead>
<tr>
<th>Gaming → Microsoft → Level3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaming → Microsoft → UNKNOWN_CDN</td>
</tr>
<tr>
<td>→ gs2 ww prod dl playstation.net → UNKNOWN_CDN</td>
</tr>
<tr>
<td>Gaming → Microsoft → EdgeCast Verizon</td>
</tr>
<tr>
<td>→ gs2 ww prod dl playstation.net → Amazon + AWS</td>
</tr>
<tr>
<td>→ epicgames-download1 akamaized.net → UNKNOWN_CDN</td>
</tr>
<tr>
<td>Gaming → Microsoft → Level3</td>
</tr>
<tr>
<td>→ gs2 ww prod dl playstation.net → Level3</td>
</tr>
<tr>
<td>Gaming → Microsoft → Akamai</td>
</tr>
<tr>
<td>→ epicgames-download1 akamaized.net → Akamai</td>
</tr>
<tr>
<td>→ gs2 ww prod dl playstation.net → Akamai</td>
</tr>
</tbody>
</table>

**Timeline:**

<table>
<thead>
<tr>
<th>AM</th>
<th>05 AM</th>
<th>06 AM</th>
<th>07 AM</th>
<th>08 AM</th>
<th>09 AM</th>
<th>10 AM</th>
<th>11 AM</th>
<th>12 PM</th>
<th>01 PM</th>
<th>02 PM</th>
<th>03 PM</th>
<th>04</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Use Case: Subscriber Tracking

Simple tools for high-tier customer support
“UserID 1234 peaks at XX Mbps, let’s compare to their plan”

Fair-Use quota monitoring
“User 4567 complains about excess usage charges, which apps / services are responsible?”

Validate engineering assumptions on user bandwidth consumption
“Users on this CMTS/DSLAM/PoP consume in 95%ile YY Mbps at peak time”
Thanks!

kentik.com/nfd19
OTT Tracking Challenges

CDNs vs Owned vs Embedded delivery infrastructure. CDNs host multiple OTT providers.
Subscriber and Service Utilization Analytics

• Real-time enrichment of network traffic data
• Adds context for business insight
  • Subscriber IDs (usernames, MAC, tier)
  • OTT service names
  • Originating CDN
• Understand network utilization per subscriber
• Or across subscriber segments
• Reduce customer service caseload
  • “Why am I over plan / quota?”
  • “Why is my connection slow?”
• Optimize traffic delivery from CDNs for cost and performance
• Create more profitable product pricing and packaging
Interface Classifications
Filter your traffic based on interface or device roles, and the providers/customers behind.

Accurately classify your flow data by identifying roles for your device interfaces. Is this an internal or an external interface? If this is external, is it for Peering, or Transit or for a Customer?

Interface Classification gives you tight controls for classification that help you filter to the exact view of the data you need.
Over-the-Top (OTT) traffic enrichment

- **Hard, but feasible**
  - OTT providers rely on owned infrastructure and CDNs
  - Combine Flows + DNS query data + Curated host patterns
  - Still done near real time at ingest.
  - A high cardinality / frequency flow tagging backend is required

- **Business impact**
  - **Identify** traffic or cache embedding opportunities
  - Additional, end-to-edd end-user support tool
Kentik Ops

- Containerized microservice architecture
- Hybrid of private cloud in Equinix, + cloud proxies in AWS, Google
- In Equinix
- Our own provisioning system
- Linux + Docker + ZFS
- All nodes PXE
- MX + QFX network stack
- In-house backend for ingest, fusion, column store, streaming
- Go, Rust, C, C++
- Gigabits inbound via Internet, interconnection
- Unencrypted UDP and encrypted TCP
Once in Use, Kentik Spreads

SecOps
- ebay
- tierpoint
- IBM Cloud
- GitHub
- DreamHost
- UBER
- neustar
- DigitalOcean
- Bank of America

BizOps
- TATA
- gtt
- Limelight
- CenturyLink
- Bell

DevOps
- credit karma
- GitHub
- Expedia
- Cisco
- AppNexus
- MediaMath
- DigitalOcean
- Twitch
- IMPERVA
- GoDaddy
- CBS Interactive
What’s New
New Functionality

**Cloud Infrastructure**
- AWS, GCP Flow Logs
- Auto-provisioned host monitoring
- CDN logs

**Service Provider**
- Subscriber Intelligence
- CDN Tracking
- OTT Tracking
- My Kentik Portal

**Automation**
- ServiceNow
- FlowSpec
- OpsGenie
New Functionality: Behind the Scenes

**Enrichment - Scale**
- Tens of millions of tags
- Updated in < second
- Millions at a time

**Enrichment - Integration**
- Kubernetes pod/service
- AWS tagging
- NSX tagging

**User Interface**
- Content library
- Universal search
- Guided dashboards
- Linked dashboards
- Expanded query
- Geo visualizations
- Performance bracketing
- Raw flow viewer
What’s Ahead

kentik.com/nfd19
Negative Road Map

- Up / down pingthings
- Complete config parsing / “what if”
- App stack instrumentation (APM)
- Generic logging platform
- Generic BI
- IGP analytics/forensics
- Storage monitoring
- Wireless monitoring
- DB analytics
- SIEM
- Endpoint patch / version management
- Help desk / ticketing
What’s Next: Q4 ‘18, 1H ‘19

- Azure logs, tap
- Metrics, Streaming Telemetry
- Segmentation, policy intent
- Wider CDN, load balancer log integration
- Turnkey orchestration integration
- White label cloud visibility platform
- Edge compute/deployed Kentik

What’s Next: 2H 2019

- Traffic-based synthetic measurement
- INT / p4 integration