Cloudline Autonomous Driving Solutions

Accelerating insights through a new generation of Data and Analytics

October, 2018
HPE big data analytics solutions power the data-driven enterprise

Secure, workload-optimized platforms that power:

- Intelligent edge analytics
- Streaming analytics
- Core analytics/cloud
- AI edge to core

Unique, optimized reference architectures
Partner ecosystem
Flex consumption
Advice & expertise

Outcomes as a Service

“The industry is expected to see a compound annual growth rate (CAGR) of 43.6% through 2022, according to the report Machine Learning: Global Market to 2022.” Source
The 4G / Edge Autonomous Vehicle Computing Stack
Current Generation: 4G and Cloud Data Centers

A car can travel 4 feet in the amount of time it takes for data to travel round trip at 80ms latency.

Autonomous Vehicles
A future with IoT at its center requires fast computing solutions that current infrastructure does not support. An autonomous car is said to require approximately 4,000 gigabytes per day. *Mashable*

The 4G Network
This is the current generation of cellular networks that exist in the US today.

Cloud Data Centers
Not suitable for the low latency real-time predictive analytics autonomous vehicles will require. These could be used to store the large amounts of unstructured data for future data-mining and analytics. A car would have traveled 4 feet with 80ms of data latency.
The 5G / Edge Autonomous Vehicle Computing Stack
Next Generation: 5G and Micro Data Centers

Autonomous Vehicles
A future with IoT at its center requires fast computing solutions that current infrastructure does not support. An autonomous car is said to require approximately 4,000 gigabytes per day. [Mashable](https://www.mashable.com)

The 5G Network
The speed and latency requirements of autonomous vehicles will require 5G technology. The carriers are spending $100’s billions in order to upgrade their infrastructure to handle the 5G demand.

Micro Data Centers
With autonomous vehicles constantly on the move a requirement for data centers on the edge will evolve. Companies like Vapor IO are designing self-enclosed micro data centers that will be installed under cell towers.

A car can travel 4 inches in the amount of time it takes for data to travel round trip at 80ms latency.

5 Miles / ~5ms Latency

![Image of a car driving with 5G network and edge computing stacks highlighted.](image-url)
HPE Cloudline Building Blocks for Big Data Streaming

Purpose-built with Open Standards Based Hardware

Different requirements along the data pipeline stages demand different node geometries

“IoT Event Producers”
Edge Processing of data in motion
HPE Cloudline 2200 or 2800
HPE Edgeline Products

“Fast Data”
Core Processing of data in motion
HPE Cloudline 3100 with NVME

“Big Data”
Analysis of data at rest
HPE Cloudline 3100 with LFF HDDs

“Data Lake”
Object Storage- Warm/Cold
HPE Cloudline 5200 or 5800

Streaming Analytics Storage

HPE Cloudline 3100 with NVME

HPE Cloudline 5200 or 5800

Streaming Analytics Storage

HPE Cloudline 3100 with LFF HDDs
HPE Cloudline for Autonomous Driving Solutions

IoT Events Producers

Kafka Cluster
- Messaging system for real time analysis of streaming data
- A distributed configuration & synchronization service

Zookeeper

Kafka S3 Connector
- Data is serialized/converted to S3

Schema Registry
- Avro-Schema-Connector

Kafka Data Stored in Avro format

S3 Cloud Storage via Scality Ring

HPE Cloudline 2100/2600 (S3 Metadata)
- 107 operations/sec @1-1.5ms latency

HPE Cloudline 5200/5800 (S3 Data)

HPE Cloudline 2200/2800

HPE Cloudline 3100

HPE Cloudline 3100

Data Ingest

Analytics

Object Storage

Edge to Cloud

Processing 1.3M messages/sec

12B+ messages processed/serialized in 24 hours
The SMACK Stack for Big Data / AI / ML

<table>
<thead>
<tr>
<th>Apache Zookeeper™</th>
</tr>
</thead>
<tbody>
<tr>
<td>An open source Apache project that provides centralized infrastructure and services that enable synchronization across a Hadoop cluster. ZooKeeper maintains common objects needed in large cluster environments. Examples of these objects include configuration information, hierarchical naming space, and so on. Applications leverage these services to coordinate distributed processing across large clusters.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spark</th>
</tr>
</thead>
<tbody>
<tr>
<td>A fast, in-memory data processing engine with elegant and expressive development APIs to allow data workers to efficiently execute streaming, machine learning or SQL workloads that require fast iterative access to datasets.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mesos</th>
</tr>
</thead>
<tbody>
<tr>
<td>A cluster manager that provides efficient resource isolation and sharing across distributed applications or frameworks.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Akka</th>
</tr>
</thead>
<tbody>
<tr>
<td>A set of open-source libraries for designing scalable, resilient systems that span processor cores and networks.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cassandra</th>
</tr>
</thead>
<tbody>
<tr>
<td>A distributed database for managing large amounts of structured data across many commodity servers, while providing highly available service and no single point of failure.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kafka</th>
</tr>
</thead>
<tbody>
<tr>
<td>A fast, scalable, durable, and fault-tolerant publish-subscribe messaging system. Kafka works in combination with Apache Storm, Apache HBase and Apache Spark for real-time analysis and rendering of streaming data.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flink</th>
</tr>
</thead>
<tbody>
<tr>
<td>An open-source stream processing framework for distributed, high-performing, always-available, and accurate data streaming applications.</td>
</tr>
</tbody>
</table>
Autonomous Driving Software Eco-System (Example)

High Availability Configuration

Fronting Kafka

Cluster 1

Cluster N

Router

Event Acquisitions Pipeline

HPE Cloudline
Scality RING

Hewlett Packard Enterprise
Thank you!