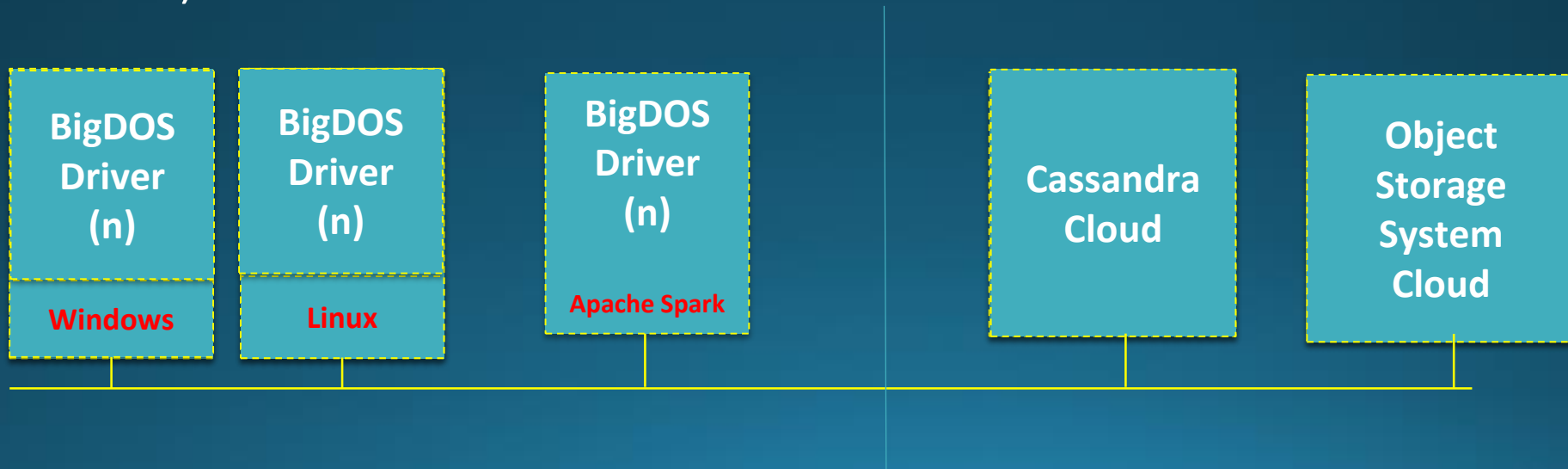# BigDOS

Decentralized File System for Unstructured Big Data

Private Cloud Architecture (vs Web, vs Client/Server)

- BigDOS filesystem driver (client side)
  - For Windows/Linux (Standard Filesystem API Access)
  - For Apache Spark (Big Data Analytics Platform)
- BigDOS Cloud of unstructured big data
  - Cassandra cloud for storing metadata (as file attributes)
  - Object Storage (such as Ceph and WOS) cloud for storing data (as file contents)

Global Distributed (vs LAN)

- ELASTIC
  - Read and write throughput both increase linearly as new machines are added, with no downtime or interruption to applications.
- SCALABLE
  - Complicated geometrical topology is supported
  - Thousands of servers can be added.
  - Unlimited files are supported.

High Availability (vs most likely)

- DECENTRALIZED
  - There are no single points of failure.
  - There are no network bottlenecks.

- FAULT TOLERANT
  - Meta data is automatically replicated to multiple nodes for fault-tolerance.
  - Replication across multiple data centers is supported.
  - Failed nodes can be replaced with no downtime.

- LOCATION AWARE
  - The same data (meta data + unstructured data itself) can have multi copies in cloud to boost performance and improve availability.
  - User will be led to the nearest server to get meta data and file contents.
  - System will automatically fall back to another if one copy is not available.

Standard Filesystem API (vs proprietary)

- Fully featured
  - All kinds of file operations are supported
  - Linux/Windows and Apache Spark are supported
- Fully compliant
  - All kinds of existing application (include MS Office) can run without modification
  - All big data analytics (Apache Spark) applications can run without modification
- Fully extended
  - The industry standard extended file attributes are supported.

Customizable

- SECURITY
  - All kinds of security model can be customized at server side.
- LOCATION AWARE
  - All geometrical topology and access policy can be customized at server side.

BigDOS: Standard Access to Global Distributed Data

- Part public information can be found at https://github.com/MRYingLEE/BigDOS

- MR.YING.LEE@gmail.com
- WeChat: ying713560
- Mobile/Whatsapp: +852-98178373