NetApp

Go further, faster®

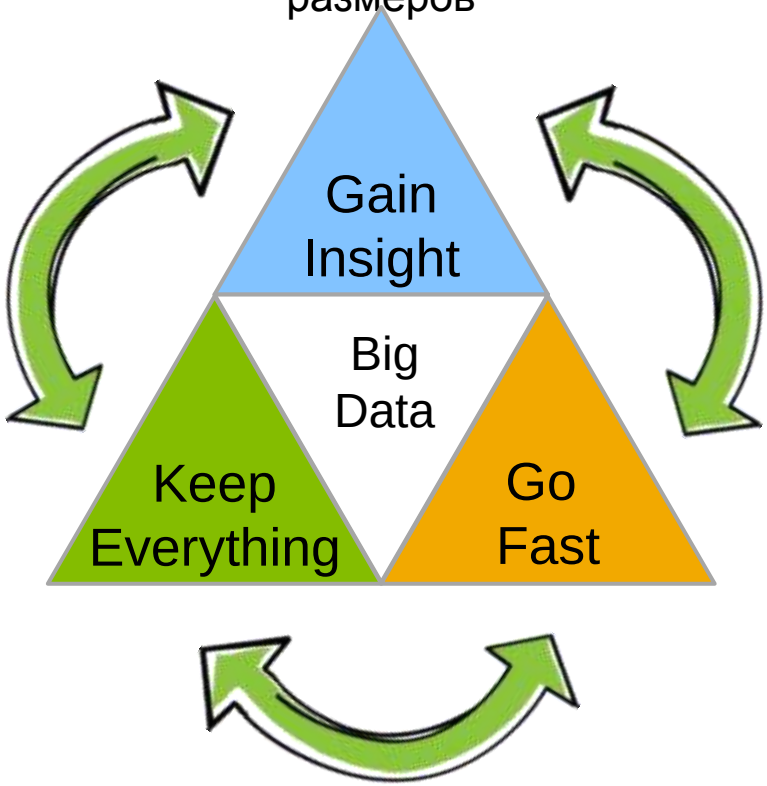
Big Data-Archiving using a Storage-Grid

Gunther.Thiel@netapp.com

Big Data ABCs

Analytics

Аналитика для наборов данных очень больших размеров



Content

Безонасное хранение данных без ограничений

Bandwidth

Производительность для высоких рабочих нагрузок

Introducing: NetApp StorageGRID

Large content repository for big, unstructured data

- Hundreds of millions of data sets, Petabytes

Create, manage and consume content globally

- Predictable access to data independent of location
- Policy-controlled data stores at each site

Intelligent Data classification and access

- Metadata-based management

What is Object Storage?

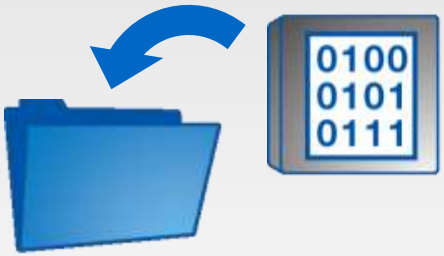
Block

Specific location on disks / memory

Tracks

Sectors

File



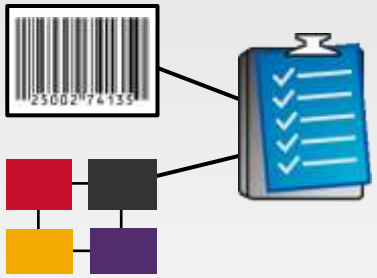
Specific folder in fixed logical order

File path

File name

Date

Object



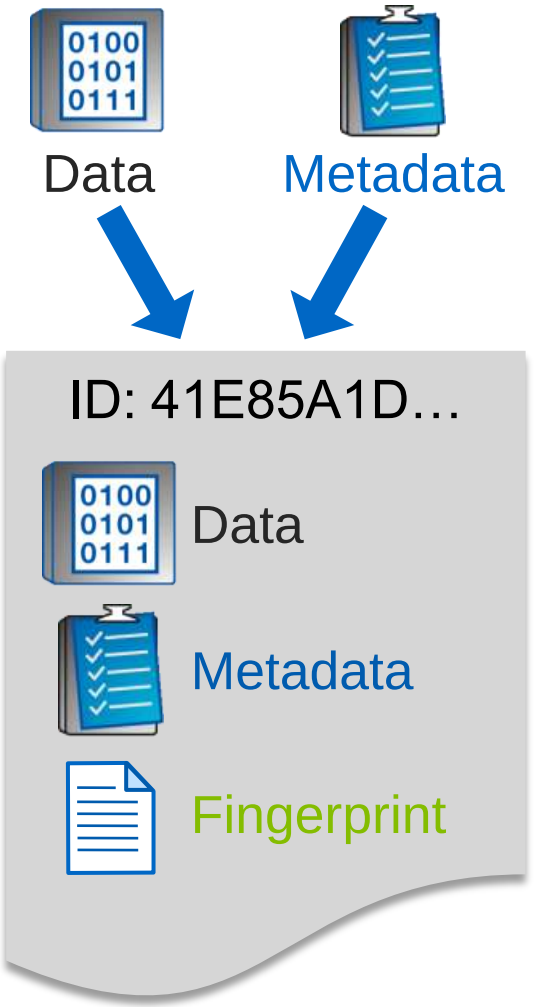
Flexible container size

Data and Metadata

Unique ID

Object Fundamentals

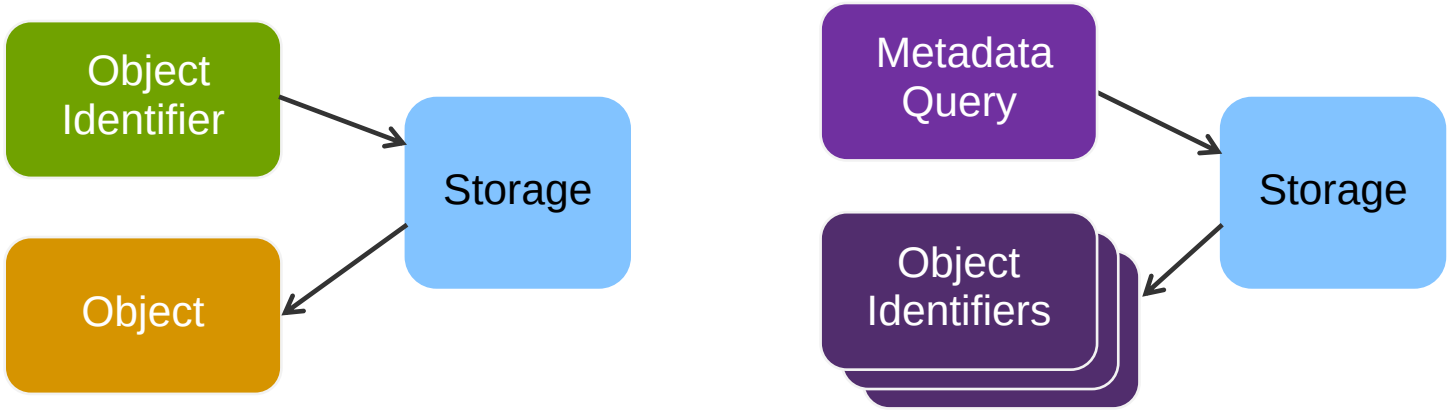
- Based on Object Storage
- Each object can include:
  - User-provide data
  - User-provided metadata
  - User-provided identifiers
  - System-generated metadata
  - System-generated identifiers
  - System-generated fingerprint
- Objects are immutable once committed



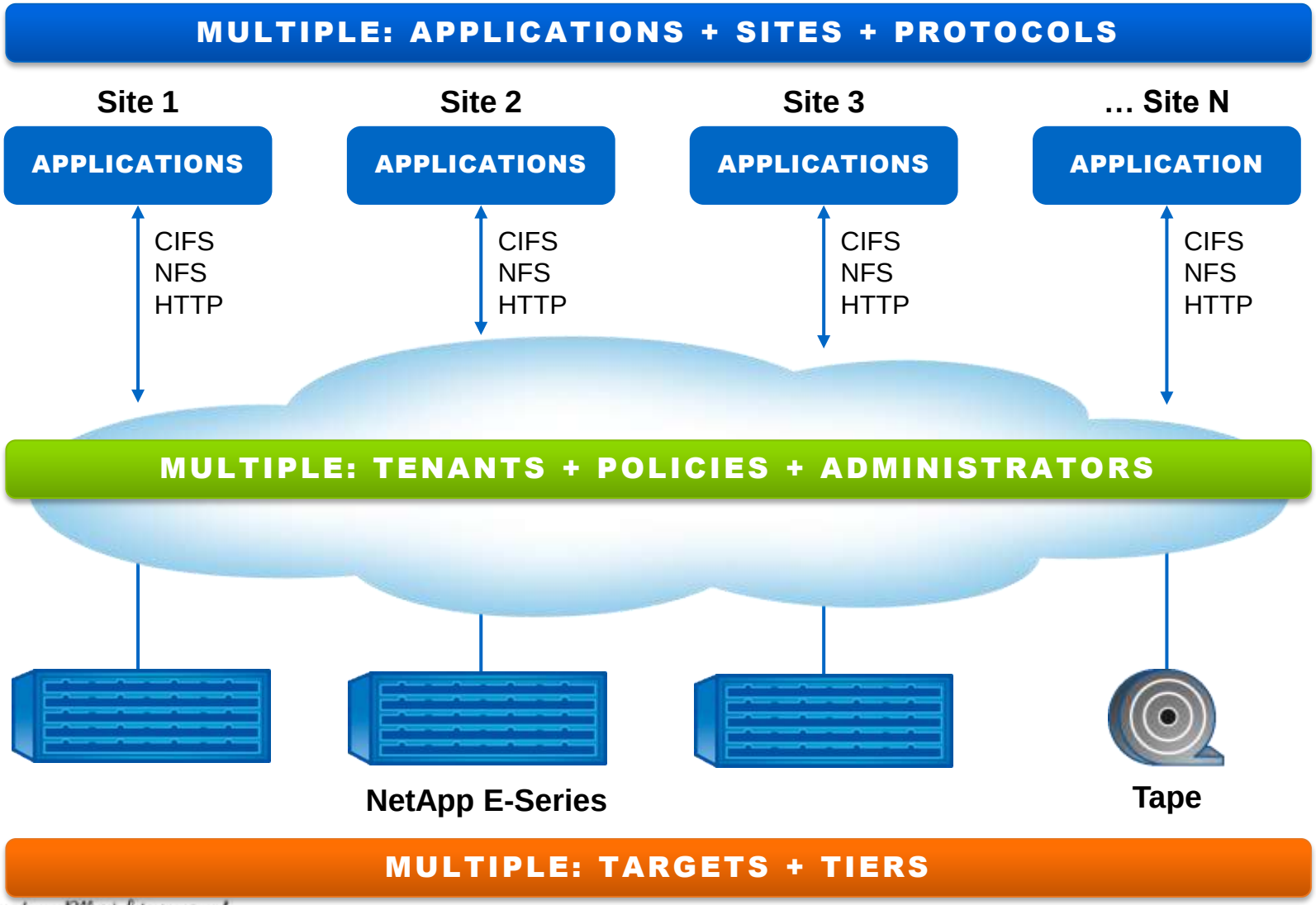
Accessing Objects

Objects are accessed in two ways:

- Accessing an object using object identifier assigned when the object was created, or
- Querying for a list of identifiers for all objects that have specific metadata values.

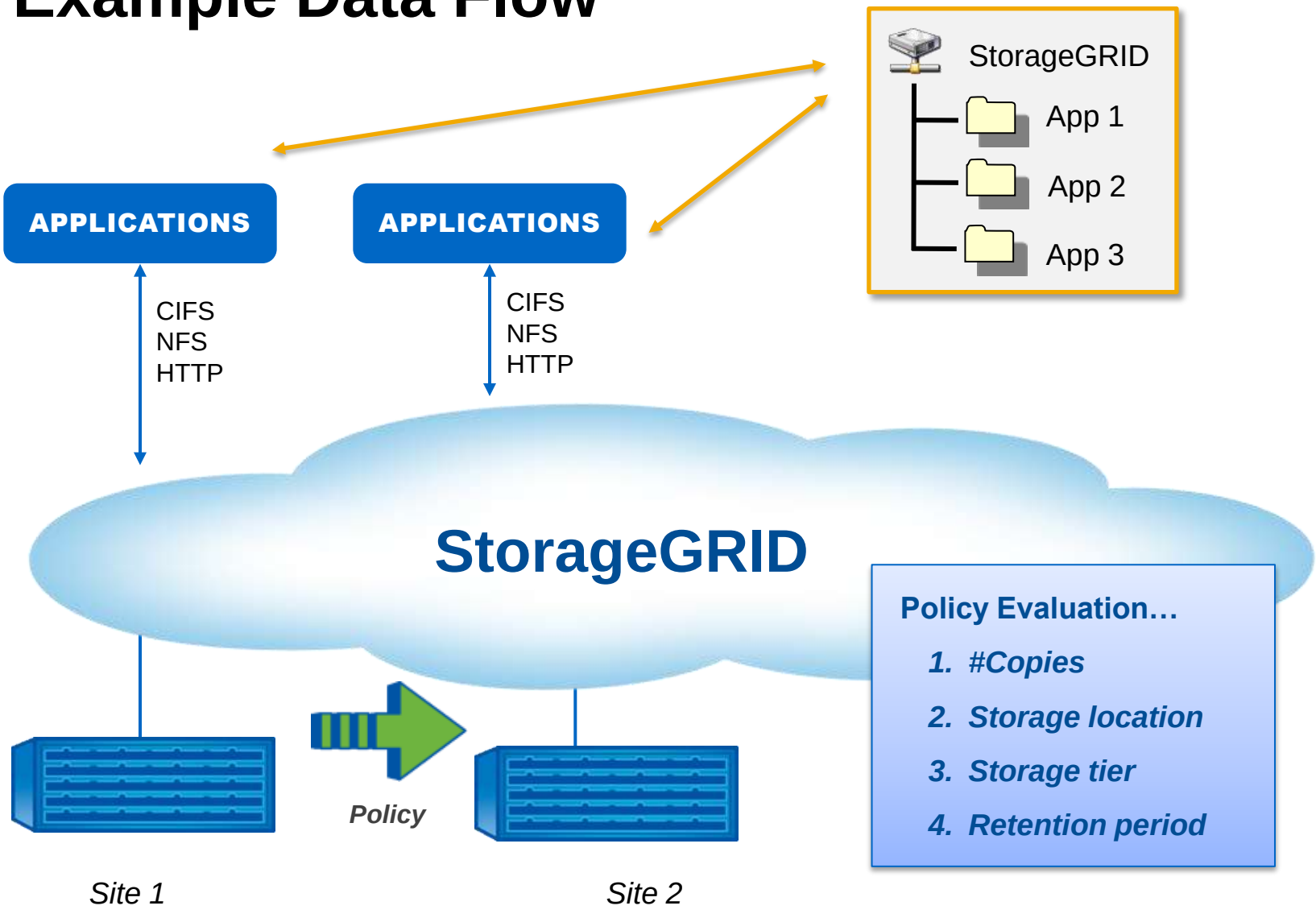


NetApp StorageGRID Solution

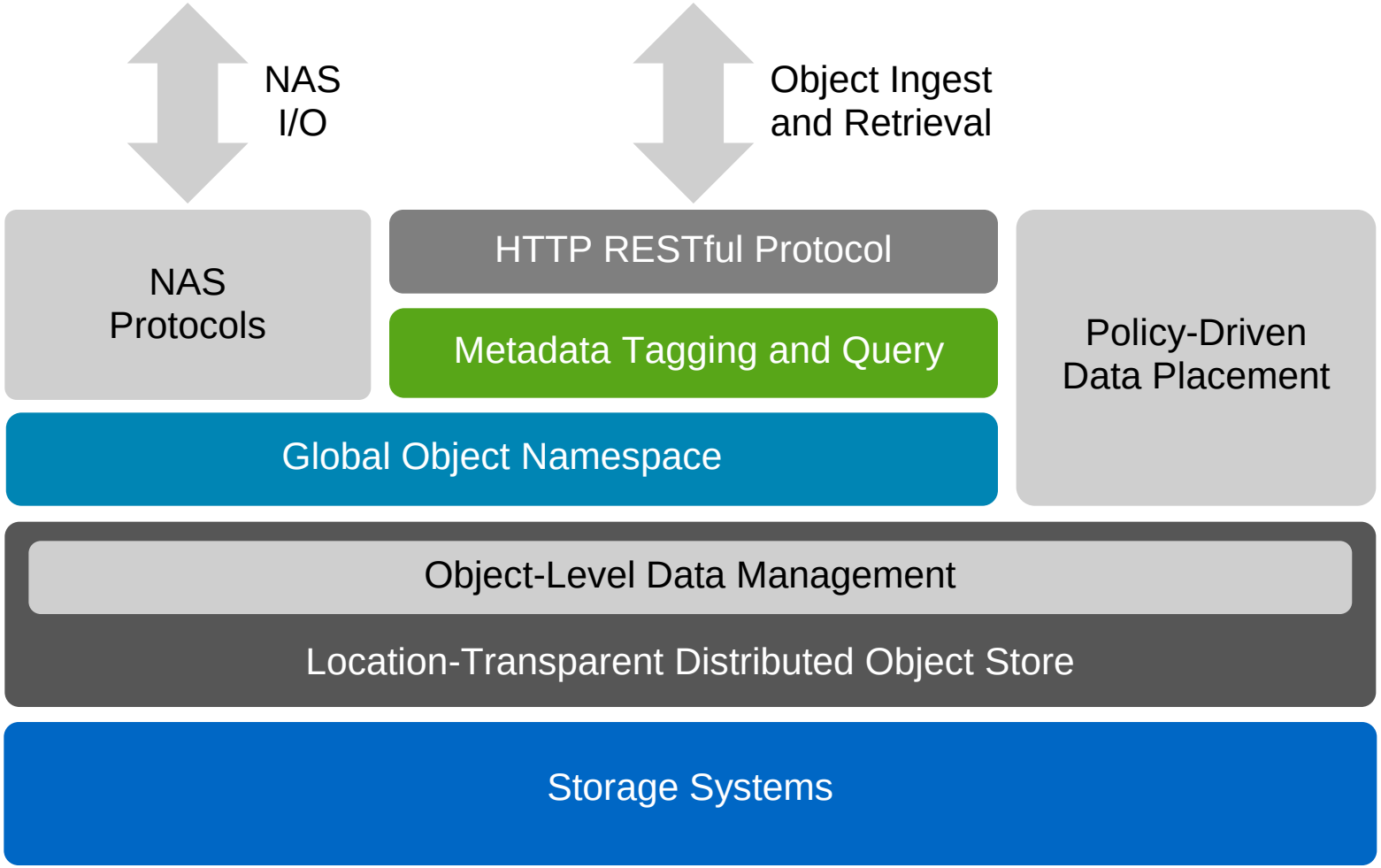


NetApp Confidential – Limited Use

Example Data Flow



StorageGRID Functional Diagram



Big Data Requires Multiple Architectures

FAS Family
with
Data ONTAP®

E-Series
With
Hadoop
Lustre
StorNext
StorageGRID

- Analytics
  - Combines FAS and E-Series to deliver enterprise Hadoop
- Bandwidth
  - Solutions based on E-Series provide extremely high performance and storage density to deliver high performance for intensive workloads
- Content
  - FAS with Data ONTAP Cluster Mode provides scale for enterprise file services and content repositories
  - E-Series with StorageGRID provides scale for distributed multi-site content repositories