

Highly Scalable Cognitive Storage Management Platform Using Cloud Native Services

Ramakrishna Vadla, IBM Maneesh Rapelly, IBM

Acknowledgement : Sumant Padbidri, Anbazhagan Mani



- Server Downtimes and Consequences
- Storage Management Evolution
- Next Generation Storage Management Platform
- Cognitive (AI) Storage Management Platform
- Predictive Analytics
- Scalability

Server Downtimes and Consequences

Average cost per hour of enterprise server downtime worldwide in 2017 and 2018





Storage Administrator

Storage Management Evolution



- Each storage device type has it own management module
- ✓ Challenges
 - No consolidated view of the storage
 - Management complexity Login to multiple consoles to monitor the devices
 - Difficult to debug the problems those are part of the other devices



- Consolidated view of all the storage devices including third party devices
- On-premise deployment on dedicated server
- ✓ Challenges
 - Dedicated resources for deployment
 - Support issues turn around time is more to debug the issues
 - No information about the other deployments
 - Scalability is the challenge
 - High TCO
 - Running predictive analytics

Next Generation Storage Management Platform

- ✓ Deploy thin Meta Data Collection Service in client data center that connects to storage devices
- \checkmark Run all the data processing micro services on the cloud
- \checkmark Supports thousands of tenants with less resources
- ✓ Highly Scalable and reliable using cloud auto scale feature
 - Horizontally
 - Vertically
- \checkmark Processing of billions of metrics per minute
- ✓ Recover from site disasters (DR)
- ✓ Secure data in motion, data at rest, RBAC
- ✓ Data lake based on NoSQL such as Cassandra deployed on the cloud.
- ✓ Predictive analytics
- \checkmark Proactive support faster time to resolution
- ✓ Different roles of the organization can view the same details



Cloud Native Services Based Architecture

✓ Microservice based architecture

- ✓ Data Services Kubernetes and Containers
 - Highly scalable using advanced auto scale features
 - High Availability and reliability
- Lambda/Cloud functions Used for small repetitive tasks that can be processed in less time
- ✓ Data Lake No SQL such as Cassandara database, AWS DynamoDB, Azure datalake
- ✓ Meta Data Storage Object storages such as IBM Cloud Object storage, AWS S3, Azure object store
- ✓ Messaging Service Kafka-as-a-service platform from IBM Cloud, AWS Streaming service
- ✓ LogAnalytics Elasticsearch (ELK) service from Cloud -Elastic, IBM Cloud, AWS
- ✓ ML/DL service using IBM Watson/ Amazon Sagemaker /MS Azure ML



Cognitive (AI) Storage Management Platform

- Predict Data Traffic issues high response times/declines in throughput
 - Noisy neighbor (Correlation Analysis)
 - Slow responding Hosts (Correlation Analysis)
- Analyze patterns and correlate with other customer datasets
 - Performance prediction in heterogeneous environments
 - Tracking of known issues Learn from other customer issues -(Classification)
 - Classify the workload types based on the performance data patterns
- Predictive Analytics
 - Capacity Forecasting (Regression)
 - Power consumption in data centers (Regression)
- > Performance anomaly detection
 - Performance metrics analysis (Time-series data analysis)
 - Automated Triaging and Root Cause Analysis (Classification)
 - Log analysis (Clustering)
- Configuration best practices recommendations
 - Manual upgrades/Automated upgrades
 - Configuration validation to avoid interruptions in service
- > Intelligent Performance Tuning
- Monitoring and improving SLAs

Failures

- Device Failures
- Network Failures
- Protocol Failures



IBM Watson

AI Services

AI Based Predictive Analytics

Predict Data Traffic Issues – High response times/Declines in throughput

Goal – Find a host that causes data traffic issues - high response times/decline in throughput



Al Based Predictive Analytics

Predict Data Traffic Issues – High response times/Declines in throughput

Goal – Find a host that causes data traffic issues - high response times/decline in throughput

Slow responding hosts

- FC port buffer is not utilized properly by hosts
- Difficult to find the host.
- Host with highest correlation is the culprit

Noisy neighbor

• Extremely busy volumes create problems for other volumes in the cluster



Configuration & Log Analytics

✓ Configuration Analytics

- ✓ Different versions of storage devices deployment report
- ✓ Total amount of storage (PB) deployed across the customers
- ✓ Different type of storage devices deployed
- ✓ No. of devices deployed across geo
- ✓ Customers require upgrades
- ✓ Log Analytics
 - ✓ Errors
 - ✓ Warning

Elastic - Opensource distributed real time data search and analytics index based database engine with schema free JSON documents

Logstash – Ship logs from any source, parse them, get the right timestamp, index them, and search them.

Kibana – Data visualization engine allows to natively interact with data via custom dashboards



Highly Scalable Platform

Scalability based on Cloud native micro services – Kubernetes & containers

- Application expected to be available 24/7, Frequent Deployment of new versions
- ✓ Containers help avoiding downtime
- Kubernetes does container orchestration by managing pods
- Pods can control one or more containers
- ✓ Replica sets responsible for specified number of active pods during scale out or in
- Deployment controller changes the actual state to desired state
- ✓ Service is an abstraction which defines logical set of pods and policy to access them



One deployment with single pod

Deploy with multiple pods

Highly Scalable Platform

Scalability based on Cloud native micro services – Kubernetes & containers

- Scale a deployment to fixed number of replicas: --replicas=10
- Horizontal pod autoscaling
 - --min –max –cpu-percent
- Proportional scaling:
 - support running multiple versions of an application at the same time
 - When rolling update is in progress, balances the additional replicas in the existing active ReplicaSets.
- Exposing the service:
 - NodePort
 - Load Balancer
 - Kubernetes Ingress



Q&A Thank You