Kavita Chavda - DE & CTO Systems Services SPL – Global Technology Services
Tom Bish - IBM Senior Technical Staff Member Scale-out Storage Systems - Storage Cloud Architecture

Nov 5, 2009

Storage Cloud Architecture Blueprint & Technology Roadmap

Agenda

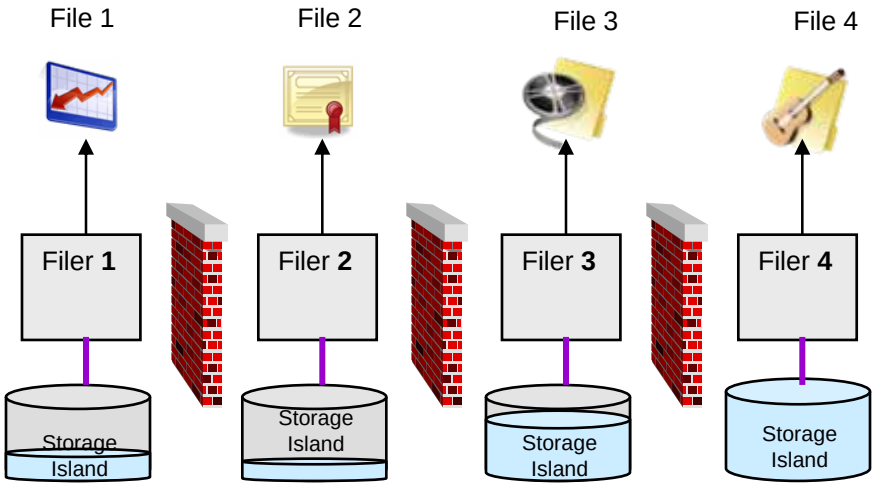
- Storage Components – Scale out NAS
- Storage Cloud offerings Overview
- Public Storage Cloud Architecture Blueprint
- Questions

Agenda

- Storage Components – Scale out NAS
- Storage Cloud offerings Overview
- Public Storage Cloud Architecture Blueprint
- Questions

What is ‘Scale Out NAS’?

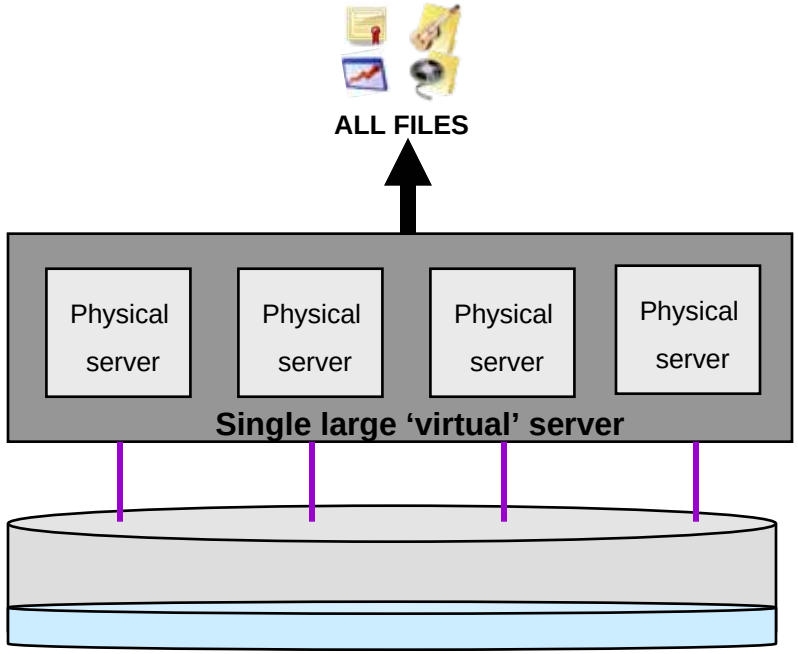
“Traditional NAS”



A few “traditional NAS” challenges:

- NAS appliances and simple file servers have many limitations - # of objects, bandwidth, capacity.
- Manageability, backup, tiering become impossible at scale.
- Simply adding NAS appliances leads to fragmented data
  - “find the file” becomes “find the server that has the file and *then* find the file”
  - Workflow, regulatory compliance, ILM all hampered as a result
- Isolated direct-attached or SAN-based disk leads to underutilization
  - Disks allocated to one server or NAS appliance cannot be used by others
  - Total utilization can be as low as 15%
- Adding more and more NAS appliances leads to management nightmare
  - File systems have to be stitched together “as if” they were a unified whole
- Collections of NAS appliances leads to hot-spots and poor performance
  - NAS appliance with most in-demand files is pegged while others are idle

“Scale Out NAS”

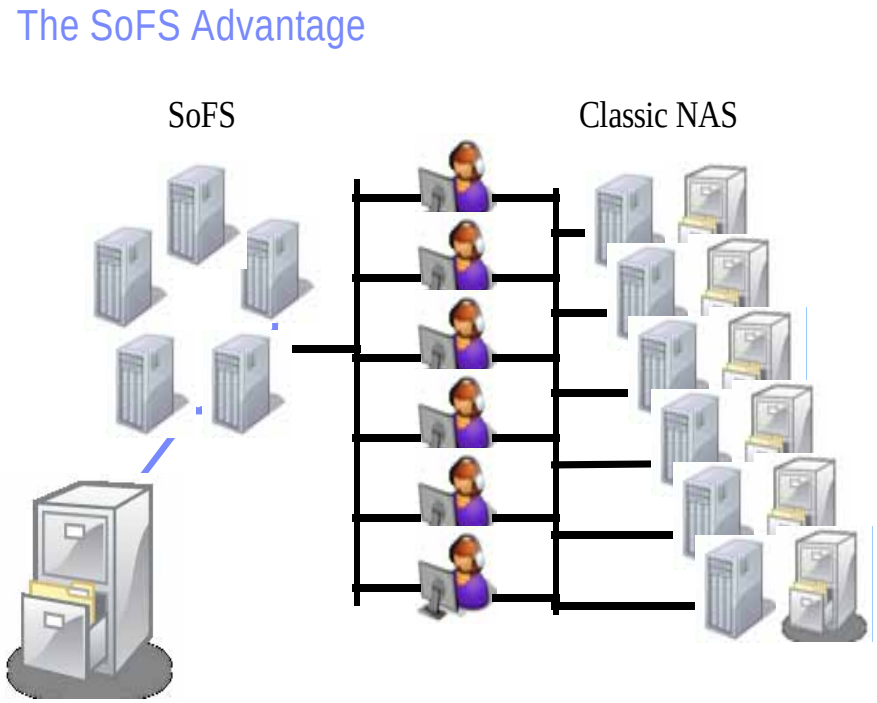


Different implementations of “Scale Out NAS” offer different benefits:

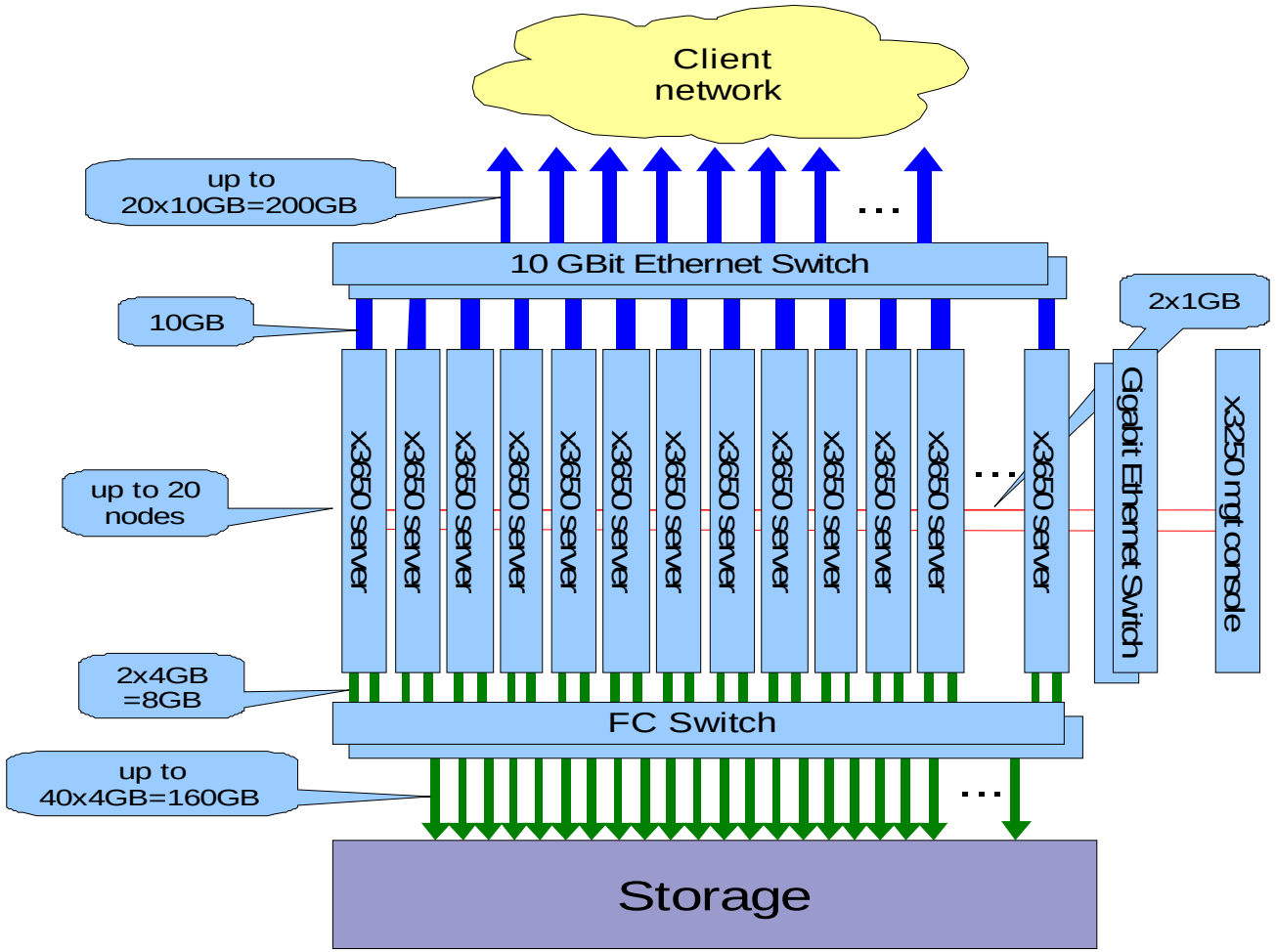
- Very high aggregate performance and scalability
- Very high availability
- Some have much lower cost
- Greatly simplified management, backups, load balancing, etc.
- Some integrate disk to disk to tape tiering.
- Some can scale servers (I/O) and storage (PB) independently.
- Some integrate scalable backups, de-dupe, other features.

What is SoFS or Smart Business Storage Cloud?

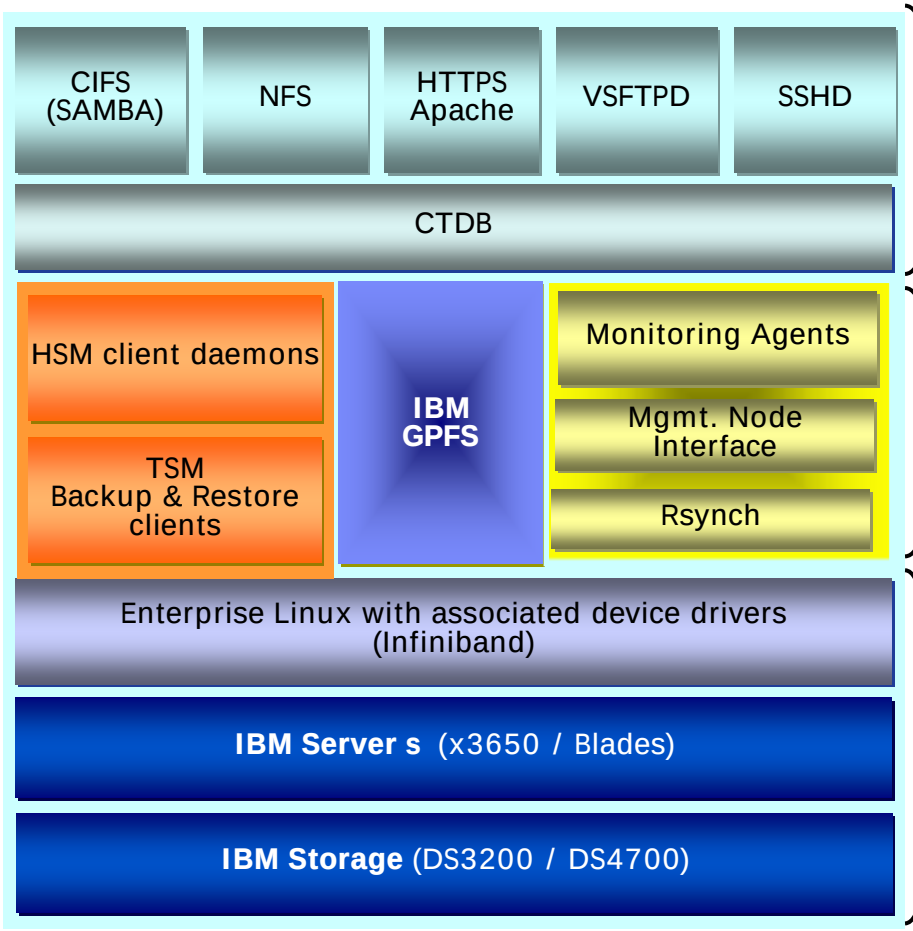
- GTS offering to deliver a scalable NAS solution
  - Utilizes x3650 servers, block storage like DSxK or XiV and SAN Fabric
  - Nodes loaded with RHEL, GPFS, Samba, CTDB and IBM developed software to form solution
    - Customer acquires hardware and open source software, GTS assembles solution
    - Small variances in configuration causes support issues
- All nodes are equal - Interface and connection to storage
- All nodes must connect to all block storage
  - Requires SAN fabric between nodes and Storage as it
  - SAN must scale with nodes and storage



SoFS Topology



SoFS / SoNAS Software Component Stacks

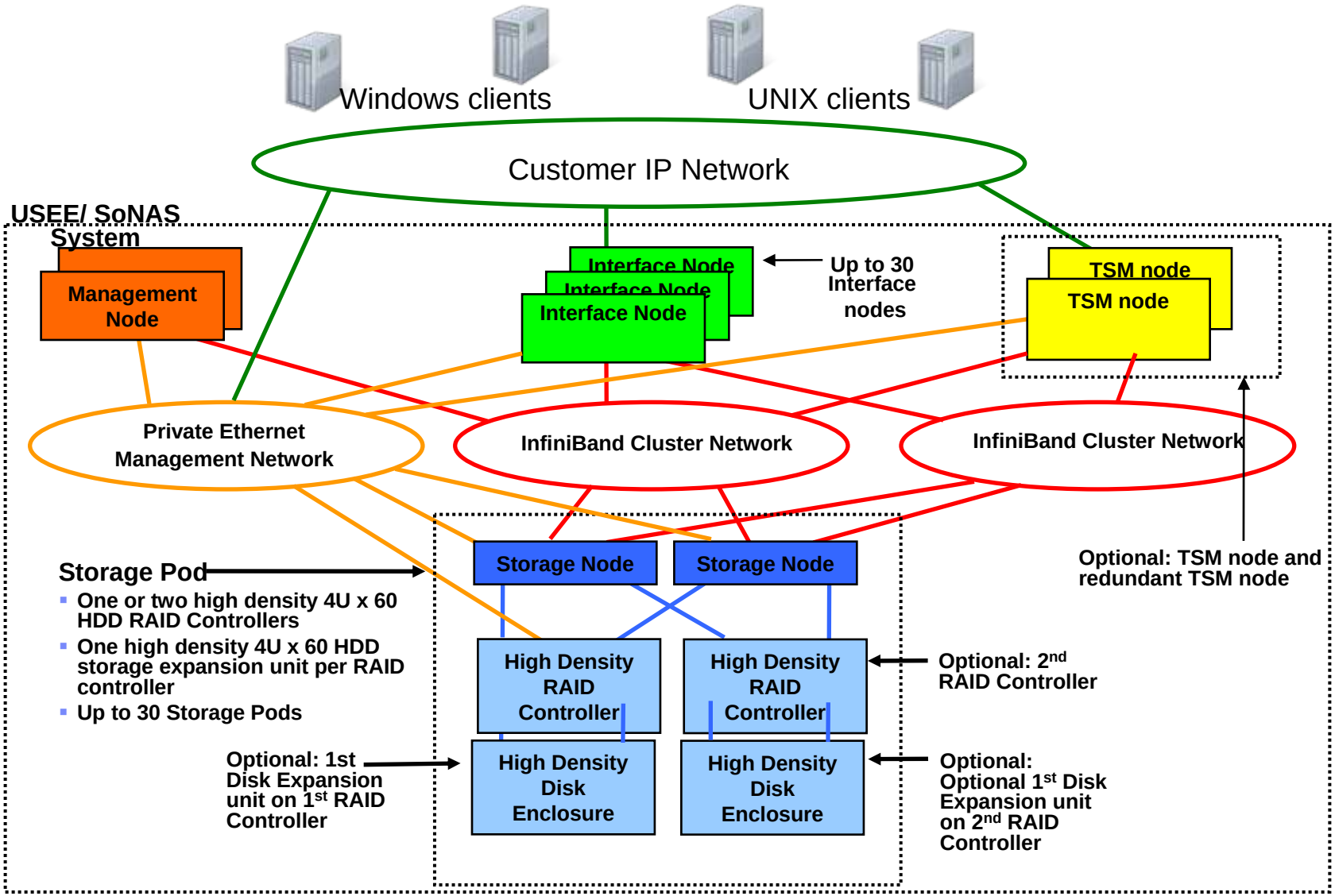


- CTDB – Clustered Trivial Database Daemon, Controls the cluster and the file service daemons
- Enhanced CIFS Server with NTFS Semantics to Support Active Directory Integration
- General Parallel Filesystem (GPFS) - IBM's High end clustered file system
- SoFS Package – Provides Management GUI, Apache file server module, acceleration tools, etc.
- TSM integration for optional tape access (HSM, backup, restore)
- SLES 11 + Device Drivers
- IBM COTS Hardware
  - x3650 Servers / Blades
  - DS3200/4700 storage

Unified Scalable Storage – Enterprise Edition – USSEE System SoNAS is "Appliance" version of SoFS - Fixed hardware with new machine type/model

- Appliance is pre-integrated set of components
  - x3650 M2 Servers with DDN 6620 RAID storage
  - Defined feature code/configuration offerings
  - Built by manufacturing
  - Service packages for hardware and software included in hardware base
- Provides flexible, scalable NAS solution
- Two tier architecture divides Interface Nodes with Storage Nodes
  - Can scale upper interface independently of lower storage nodes
  - Block storage is scaled with Storage Nodes - no SAN fabric
- Infini-band infrastructure connects Interface Nodes and storage nodes with high speed, low latency fabric
- Adds CLI for external management and configuration - provides hooks for automation
- Provides hardware SNMP alerting for internal hardware failures
  - Unified alerting as a single product, not collection of components

USSEE System - Topology



Unified Scalable Storage – Enterprise Edition – USSEE System SoNAS is "Appliance" version of SoFS - Fixed hardware with new machine type/model

- Appliance is pre-integrated set of components
  - x3650 M2 Servers with DDN 6620 RAID storage
  - Defined feature code/configuration offerings
  - Built by manufacturing
  - Service packages for hardware and software included in hardware base
- Provides flexible, scalable NAS solution
- Two tier architecture divides Interface Nodes with Storage Nodes
  - Can scale upper interface independently of lower storage nodes
  - Block storage is scaled with Storage Nodes - no SAN fabric
- Infini-band infrastructure connects Interface Nodes and storage nodes with high speed, low latency fabric
- Adds CLI for external management and configuration - provides hooks for automation
- Provides hardware SNMP alerting for internal hardware failures
  - Unified alerting as a single product, not collection of components

Agenda

- Storage Components – Scale out NAS
- Storage Cloud offerings Overview
- Public Storage Cloud Architecture Blueprint
- Questions

GTS has leveraged GPFS, SoFS to create a private & public Storage cloud Service Product

Smart Business Storage Cloud

Privately owned and managed.
Access limited to client and its partner network.
Drives efficiency, standardization and best practices while retaining greater customization and control

Smart Business Storage on the IBM Cloud *(coming soon)*

Service provider owned and managed.
Access by subscription.
Delivers select set of standardized business process, application and/or infrastructure services on a pay-per-usage basis.

….Security, privacy customization & control

Hybrid …
The combination of public and private models for the greatest efficiencies and broadest workload support.

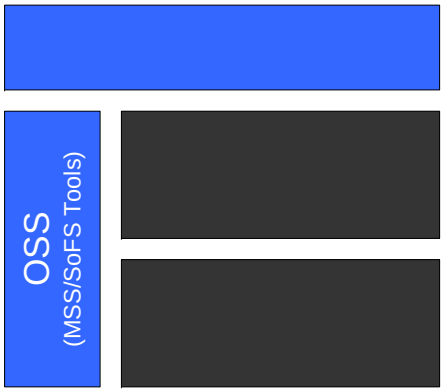
….Standardization, capital preservation, flexibility and time to deploy

Smart Business Storage Cloud (private) can be implemented behind the client firewall in managed or un-managed configurations and as an IBM hosted offering via SO.

Storage Cloud – Private & Public

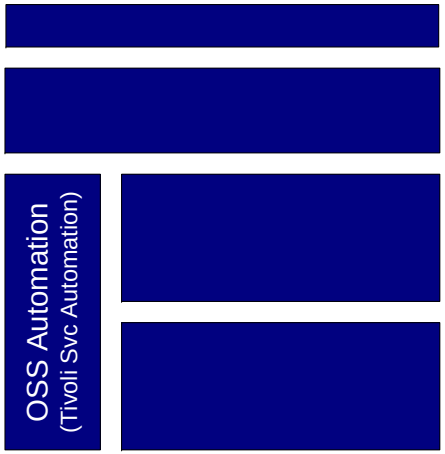
Previous Service:

- Client Owns Assets
- Not SO Enabled
- Stand-alone Tools
- No Multitenancy
- ITD Support Structure



Private Cloud:

- Client/SO Owns Assets
- Sale of Implementation plus Basic or Premium Support
- Custom SO Enabled
- No Multitenancy
- Bolstered Support



Public Cloud:

- IBM Owns Assets
- Pay-per-GB Usage
- Provisioning Automation
- Multitenant/Shared
- Bolstered Support

Scale Out File Services (retired)

Which offering will suit your customer?

Present storage cloud to your clients as one offering – Smart Business Storage - but understand the differences between the two…

	Smart Business Storage Cloud (private)	**Smart Business Storage on the IBM Cloud (public)**
Geography?	Global	**U.S. Only in 2009. High speed services in limited major metro areas.**
Asset Ownership?	Client or SO owns assets	**IBM owns assets**
Pay-per-Usage?	No	**Yes**
Automated Provisioning?	No	**Yes**
Data Security?	Highest	**High**
Performance?	Customizable to Client Needs	**Varies by network connection type**
Managed Solution?	Optional	**Yes**
Hosted Solution?	Supported with SO	**Yes**
SLA?	No	**99.99% Availability**
Cost?	Varies Widely based on Client Need	**Low**

Sales, Delivery & Steady State Processes are required to manage end to end Client satisfaction for Public Storage Cloud

Standard EP

Understand & Explore Phases

- Quick Reference Guide (CE)
- Solution Overview Presentation
- Education Materials
- ibm.com webpage (CE)
- All-in-One Slide (CE)
- Data Sheet (CE)
- White Paper *(Not in plan for initial release)*
- 2 tier data collection checklist
- Discussion Guide
- VSC Client Networking Guide
- TCO ROI analysis – estimating cost benefit

Develop Phase

- Opportunity Management process
- Capability Management process
- 2 tier data collection checklist
- eLearning on eConfig tool
- Telco pre-Qualification form
- Storage Cloud Qualifier Worksheet
- Statement of Work (SOW)
- ROI Report Template *(Not in plan for initial release)*
- Proposal Presentation *(Not in plan for initial release)*

Implement & Confirm Phases

- **Capacity Planning/Management**
- **PoD Deployment / Expansion**
- **Service Initiation / On Boarding**
- **Technical On Boarding**
- **Transitioning**
- **Steady State/Run Offering Team**
- **Steady State/Run Business Project Office**
- **Steady State/Run Integration Center**
- **Steady State/Run Production**

Enterprise-centric Pricing Structure

- Pricing model is built to be **all-inclusive** with no additional transfer charges associated with the network or support.
- Primary part is a *full variable* model with no capacity commitment on a monthly basis.
- High speed services require minimum commits:
  - 20TB for 1Gbps
  - 50TB for 10Gbps
- Additionally, we will offer lower price points based on:
  - Longer terms
  - Greater minimum commits
- Remote copies of the data and extra point-to-point VPN circuits are addt'l and provided via separate services contract

Part Numbers (Draft):

1GB @ 12Month Int Srvc Full Variable

1GB @ 12Month Int Srvc 20TB Minimum

1GB @ 36Month Int Srvc 20TB Minimum

1GB @ 12Month 1Gbps Cir 20TB Minimum

1GB @ 36Month 1Gbps Cir 20TB Minimum

1GB @ 12Month 1Gbps Cir 200TB Minimum

1GB @ 36Month 1Gbps Cir 200TB Minimum

1GB @ 12Month 10Gbps Cir 50TB Minimum

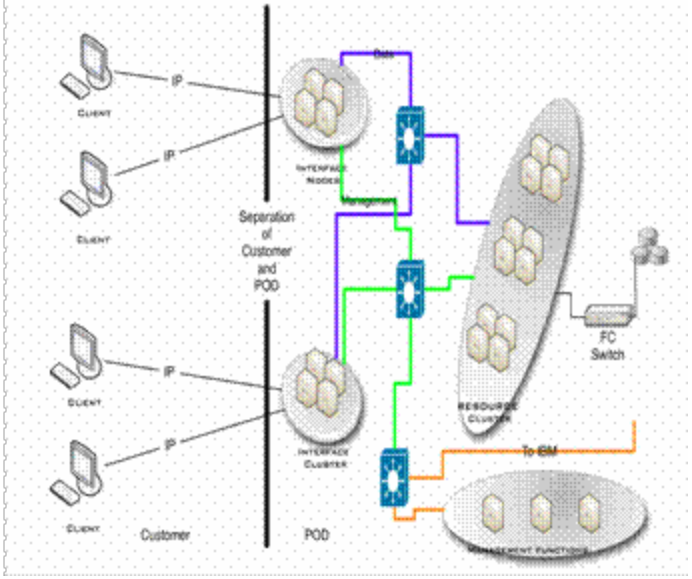
1GB @ 36Month 10Gbps Cir 50TB Minimum

1GB @ 12Month 10Gbps Cir 200TB Minimum

1GB @ 36Month 10Gbps Cir 200TB Minimum

IBM Virtual Storage Services – Landscape & Differentiation

- Key requirements of an *enterprise* storage cloud are **performance, availability, security** and **cost**.
- No competitive *high-performance* Cloud solutions exist for storage that provide a *strong SLA* to the client.
- Competitive offerings have embraced the commodity model relying on commodity storage devices and standard internet connectivity. This provides an advantage in cost *at the expense of performance and availability*.
- IBM differentiates through it's global datacenter infrastructure and strong network provider partnerships to *bring the cloud to the client.*



- IBM's Virtual Storage Cloud concept uses globally distributed infrastructure and dedicated circuits to enable:
  - Storage performance drastically greater than the competition
  - SLAs that an enterprise client can count on
  - An On-Ramp to the cloud through seamless migration of existing high-performance applications
  - An entry into the fastest growing part of the market – mid-tier, file-based storage
  - A financial model that allows customers to cut down on up-front Capital costs
  - Simple, secure and easy way to extend customer's file services to the cloud

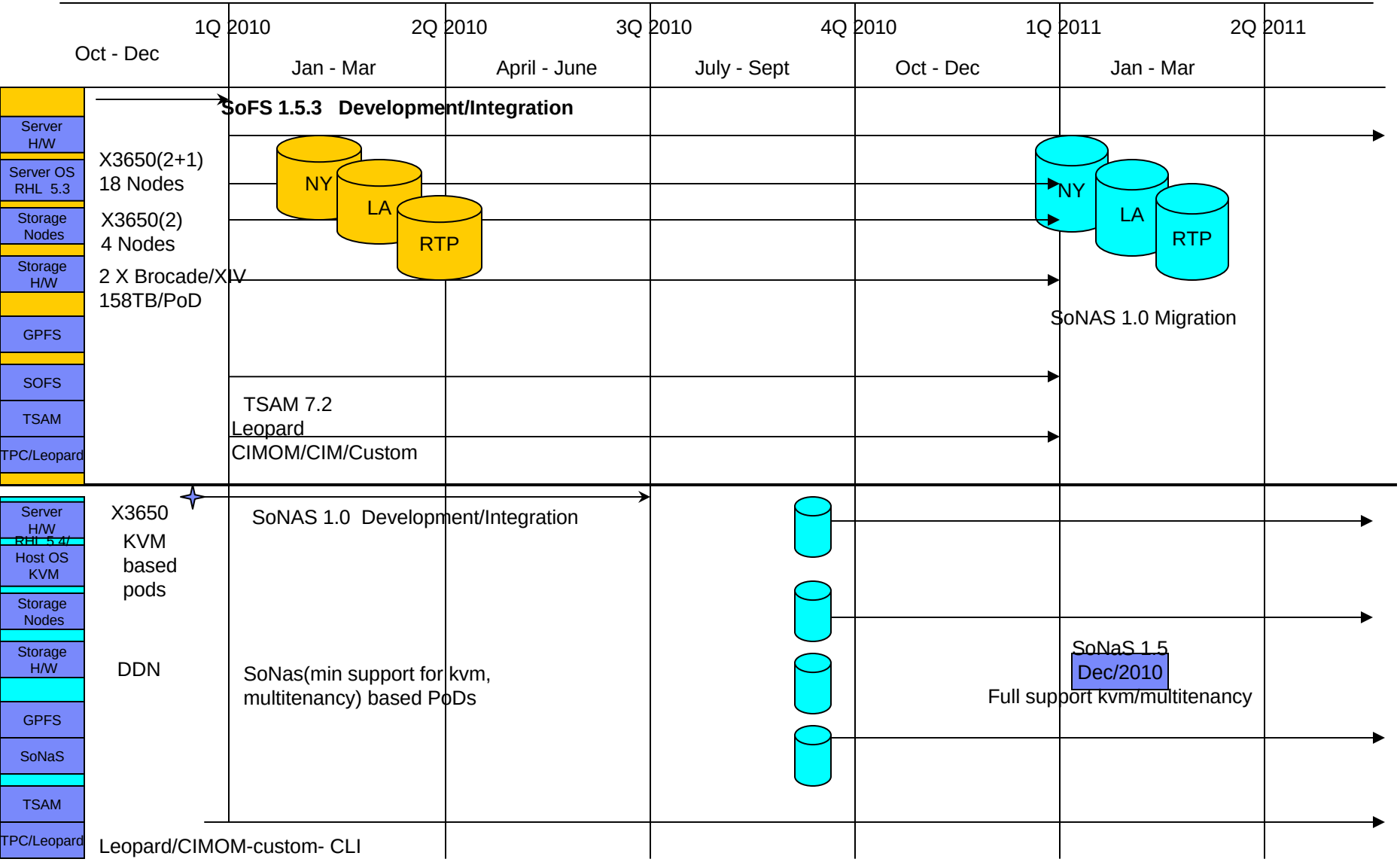
Agenda

- Storage Components – Scale out NAS
- Storage Cloud offerings Overview
- Public Storage Cloud Architecture Blueprint

Architecture Blueprint is a collection of documents that describe the overall architecture, its OSS & BSS components and hardware & software required

- Requirements
- Non functional requirements
- Architecture Overview Diagram
- Architecture Decisions
- Component Model & Build Plan
- Operational Model
- Security
- Use Case
- Standards
- Technology Roadmap

Technology Roadmap for Public Storage Cloud PoDs



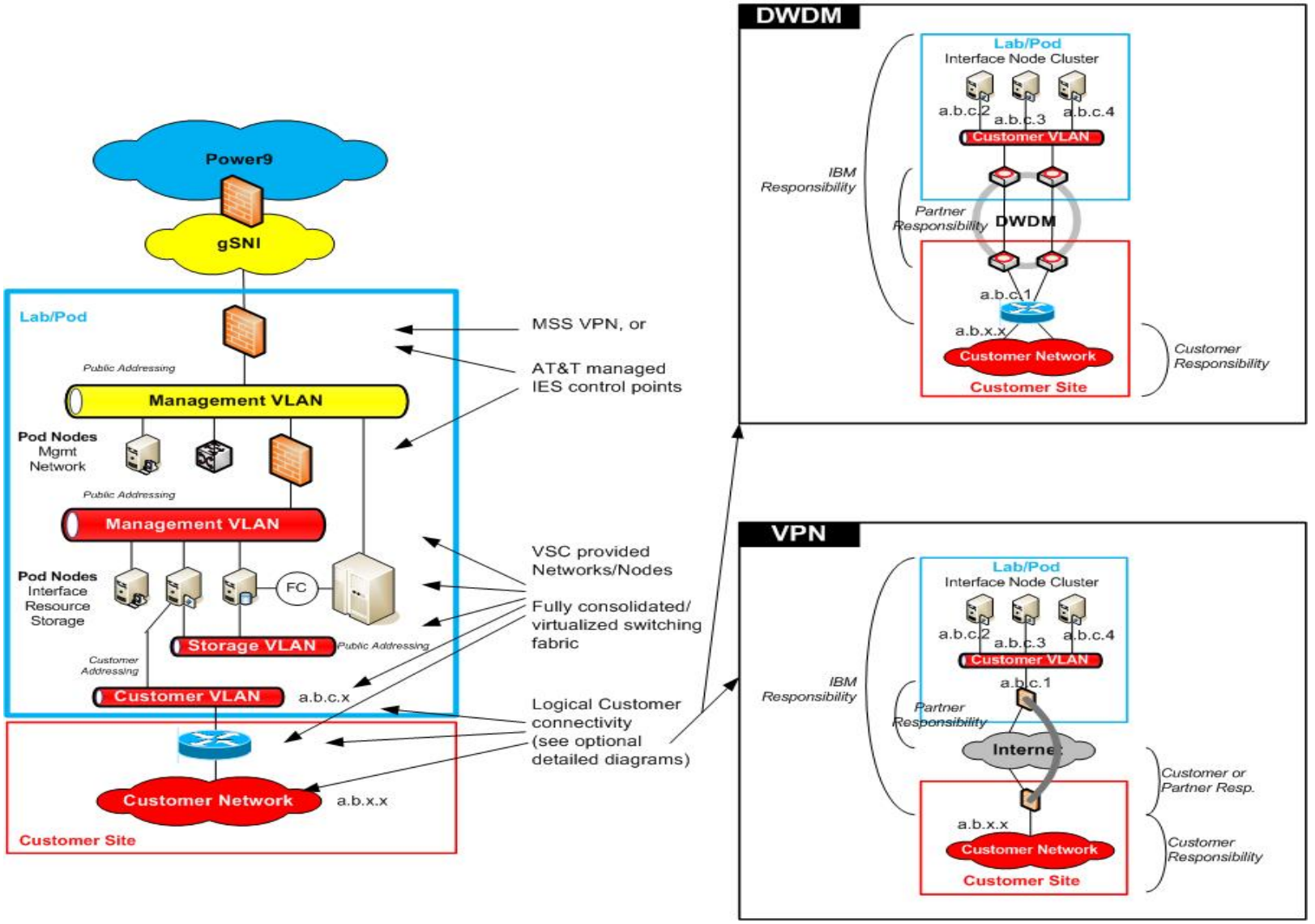
Use Cases Overview

- CONSULTING (SALES) – F2F – NO INTERNET
- DELIVERY
  - PROVISION SERVICE WORKFLOWS
    - INITIATE SERVICE
    - PROVISION CUSTOMER SERVICE
    - CONFIGURE SERVICE
    - CREATE FILE SYSTEM
    - ESTABLISH NEW POD
    - ADD NEW CUSTOMER TO THE PORTAL
  - MANAGE SERVICE
    - ADD / MODIFY QUOTA OR FILE SETS
    - MODIFY CAPACITY PLAN
    - MANAGE CAPACITY
    - DEACTIVATE AND DECOMMISSION SERVICE
    - DEACTIVATE SERVICE
    - DECOMMISSION SERVICE
    - CHANGE CUSTOMER LOCATION
    - DISCOVER GROUPS
  - USE SERVICE
    - ACCESS STORAGE
    - VIEW REPORTS INCLUDING TRENDING & UTILIZATION
    - REQUEST PERFORMANCE REPORT
    - BILLING

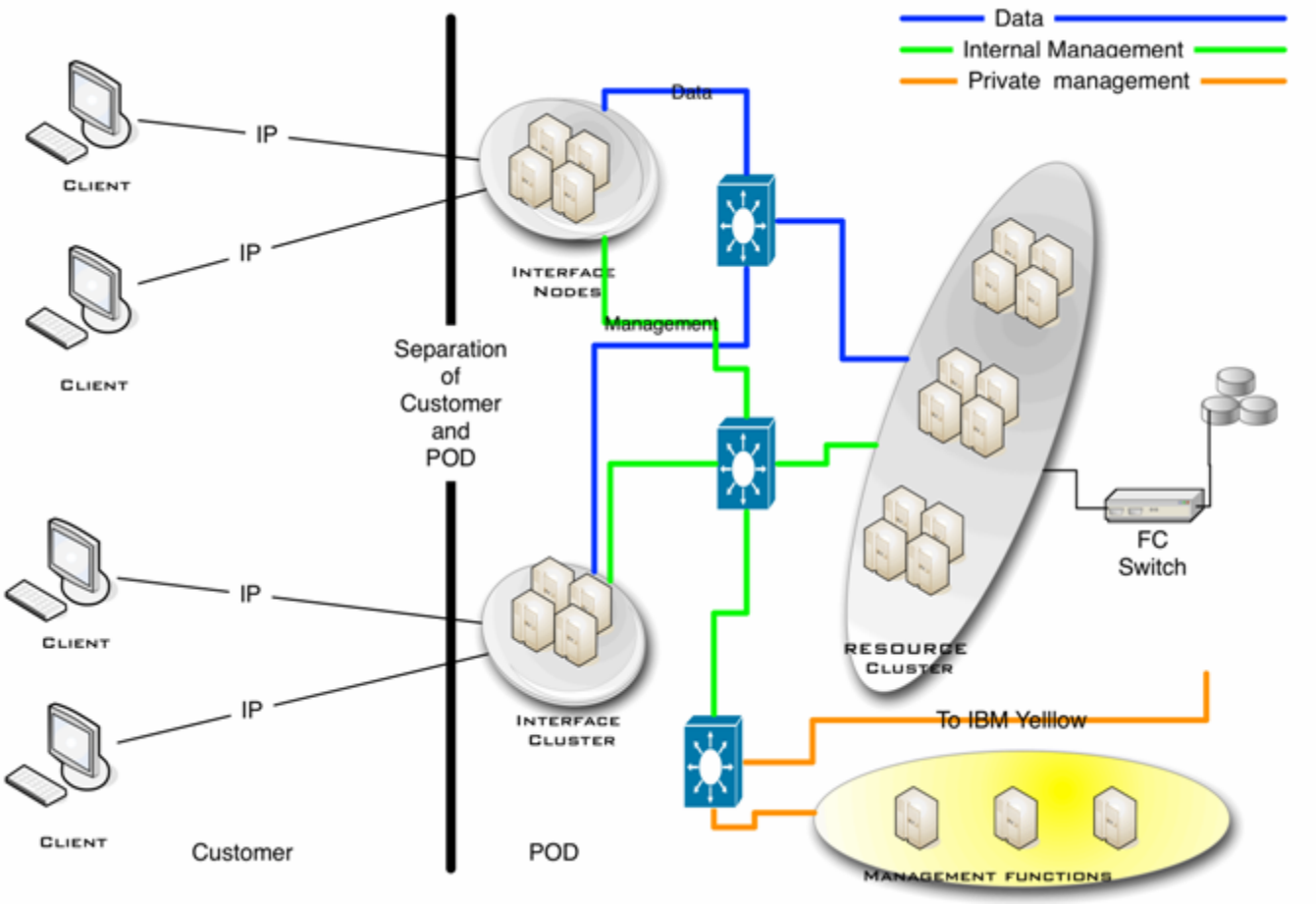
Design Points for the Storage Cloud : A design that segregates customers from each others. Interface clusters are dedicated to each customer & establishes the connectivity to the customer

- Public Storage Cloud PoD dedicated to the Customer will sit in their IP address space & leverage MS Active Directory
  - Protocols used by customer to access the storage pod is HTTPS, SFTP, SCP, CIFS, NFS, AD
  - Certificates are on the Interface nodes
  - SoFS has a self signed certificate
  - Interface Nodes communicates using NSD protocol to Storage Nodes
  - SSH is used for intra cluster operations
- Security
  - Interface clusters are trusted by Storage cluster. The Block storage is shared by all customers.
  - Customer devices are managed through IES ACLs/local firewalls to protect
  - All the devices are being managed to ITCS policies, controls
  - Green Network is RED ZONE NETWORK
  - Orange Network is YELLOW Zone
  - Each Interface node will be hardened with additional ports lock down
  - Layer 3 ACL'S implemented to control flow between the green and orange networks within the RED zone
  - Authentication of customers/users will be controlled by the customer through AD systems on their sites. Customers accessing the Storage PoD (Interface nodes through the AAA system) will authenticated. IBM does not control authentication as Active Directory on customer premises is used.
  - Node logging component will capture audit logs from the operating systems on the nodes. The logging server will take a copy from the nodes and move it to the Blue logging server
  - Software provisioning stack maintains unique transaction ID, tracks requester and establishes trust relationships between components via mutually-authenticated SSL
- The hardware components are based on 3 x3650 dedicated to each customer with shared Storage Nodes, SAN Switches & XIV Storage

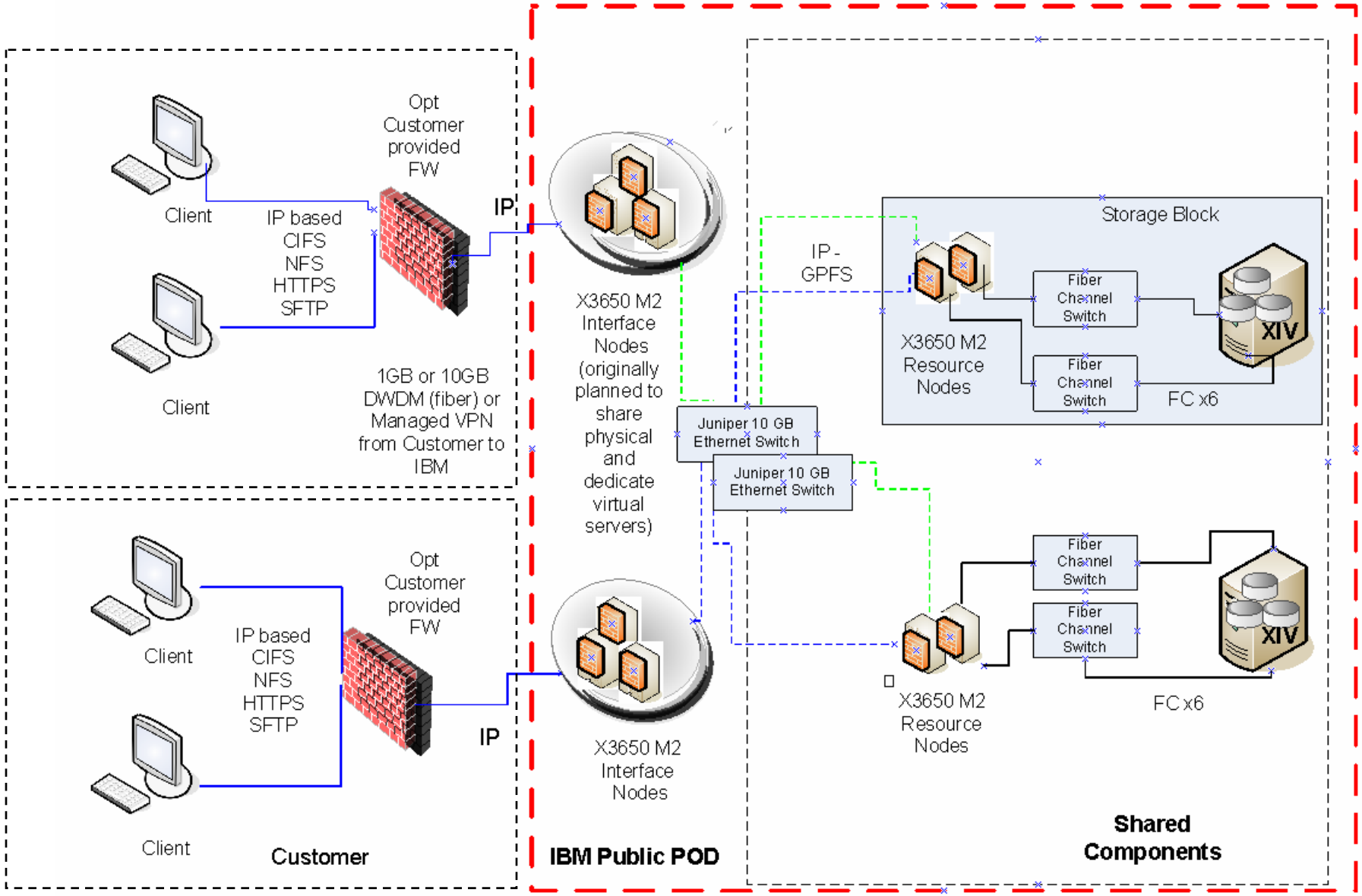
Public Storage Cloud High level Core & Customer Connectivity Options
The network is NOT auto-provisioned nor metered and will be Monitored by ATT



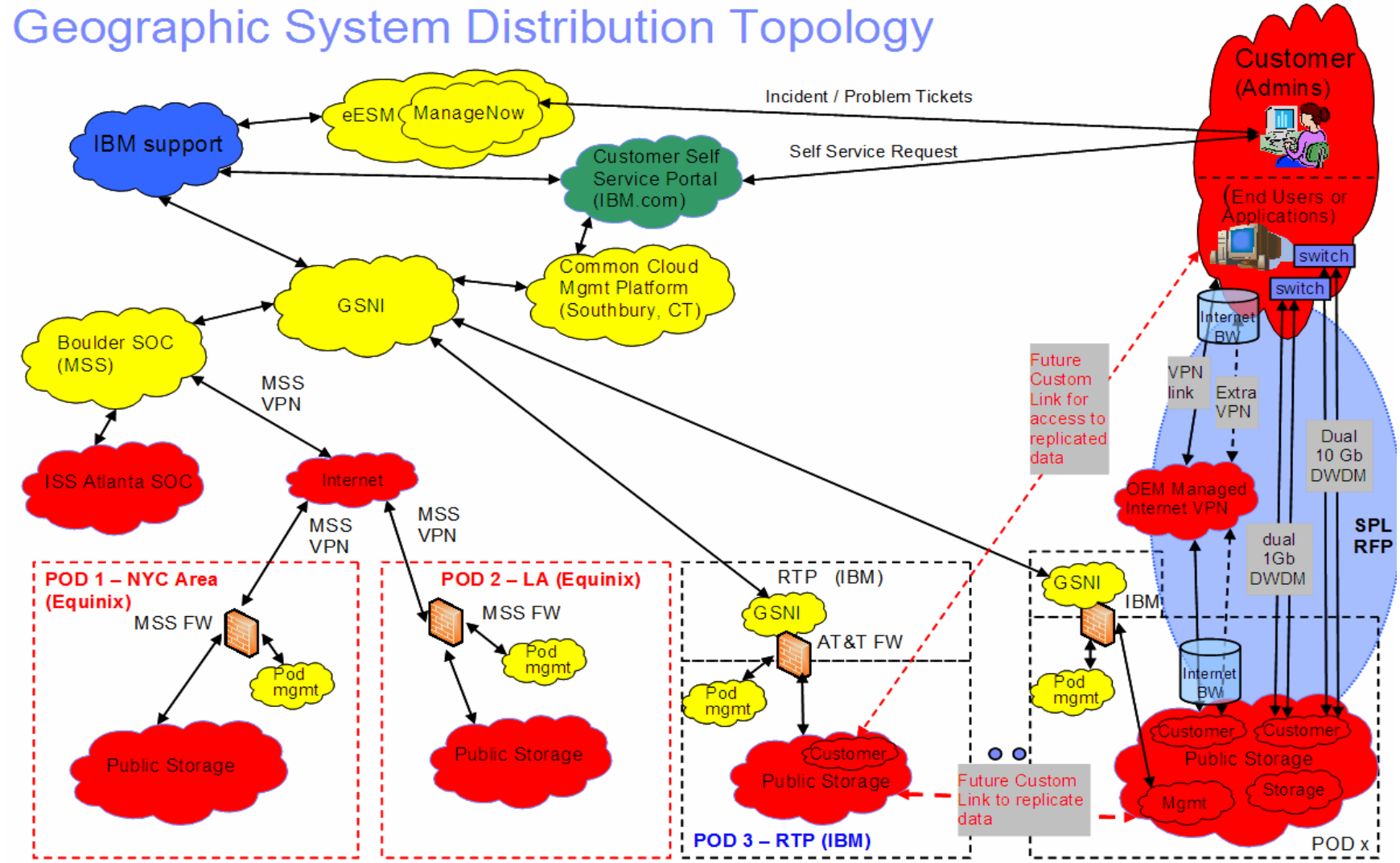
Public Storage Cloud PoD High level Concept View



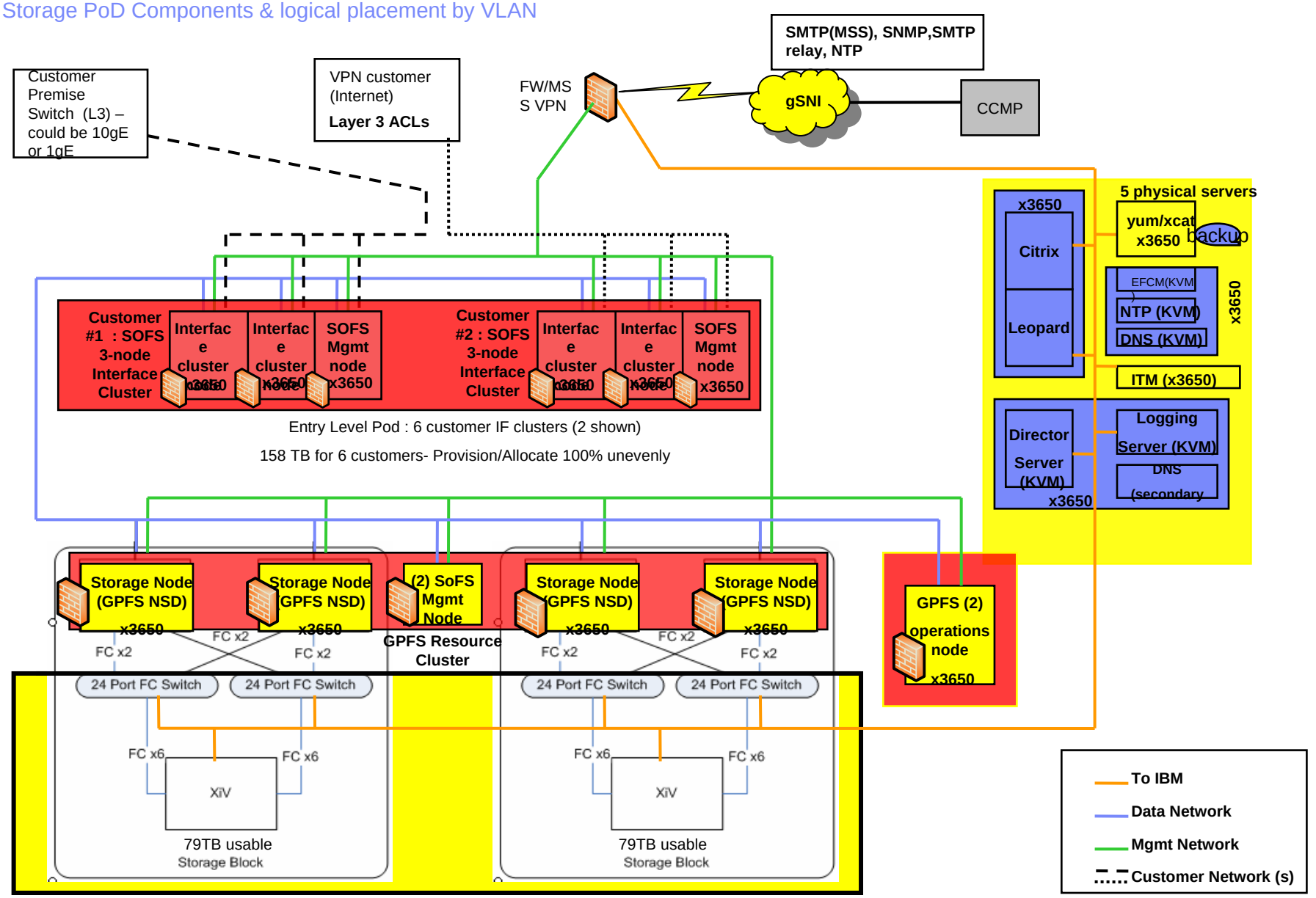
High Level Public Storage Cloud Pod Component view



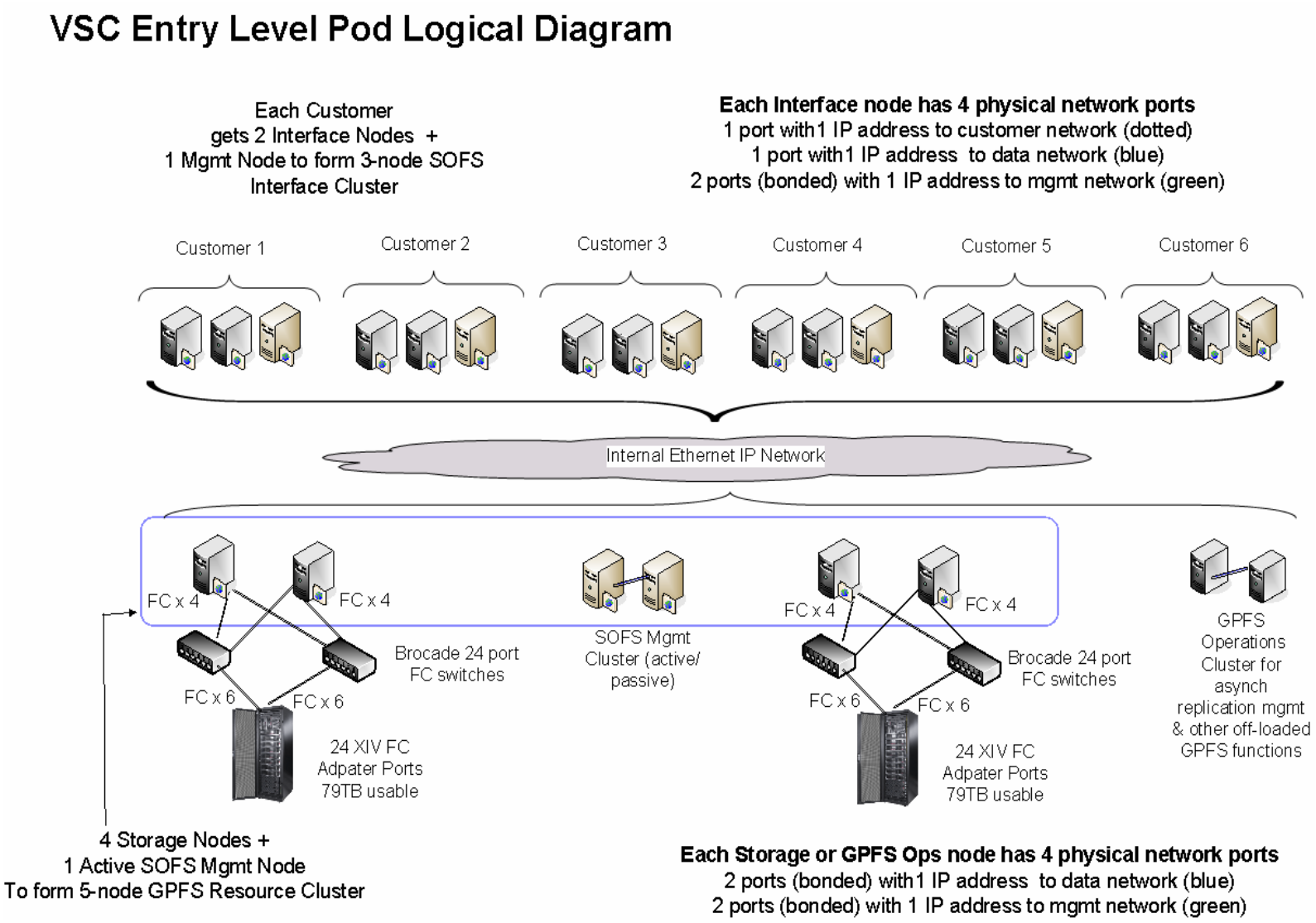
Geographic System Distribution Topology

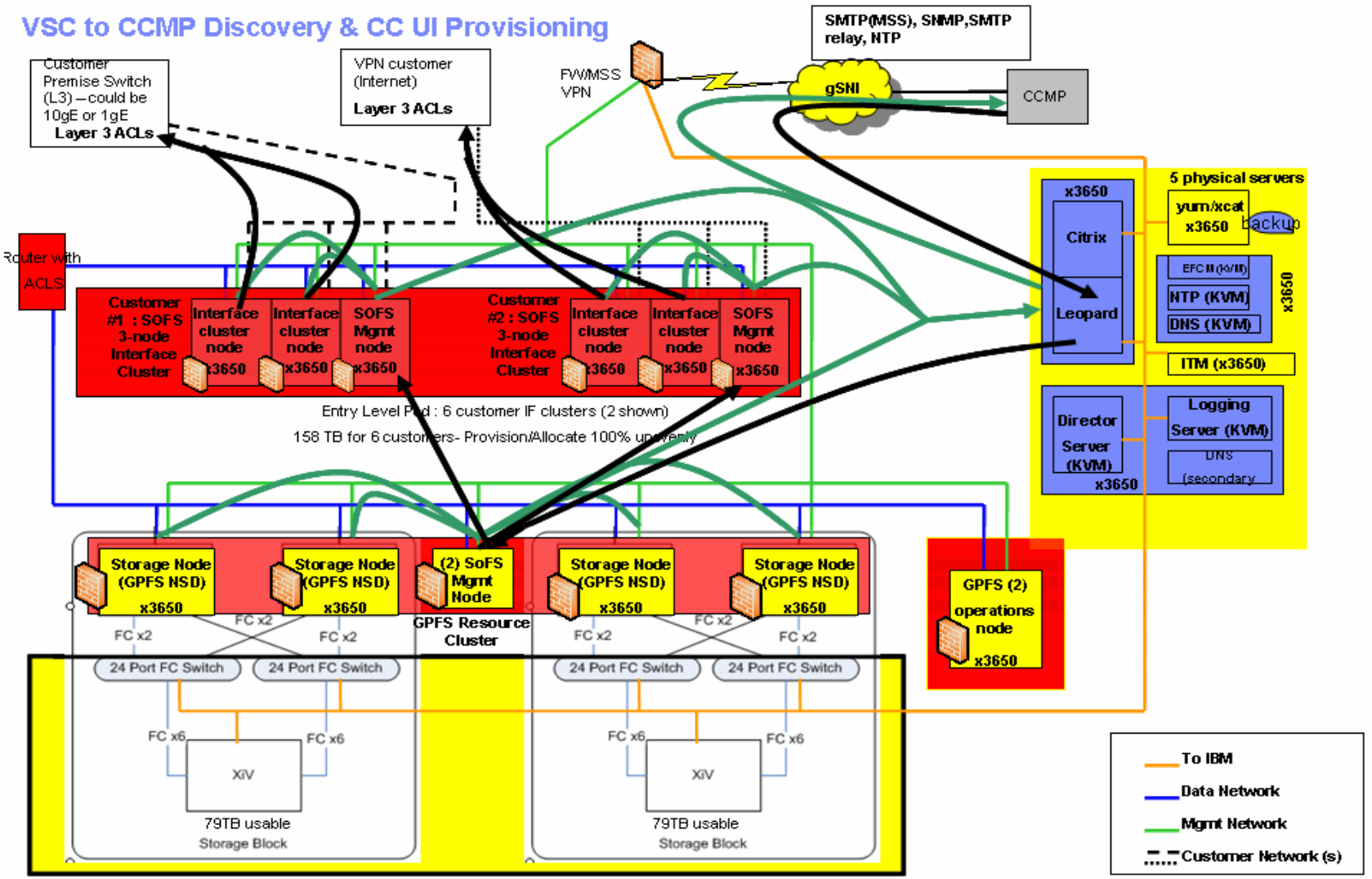


Storage PoD Components & logical placement by VLAN



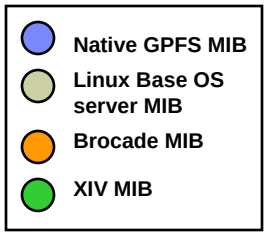
VSC Entry Level Pod Logical Diagram





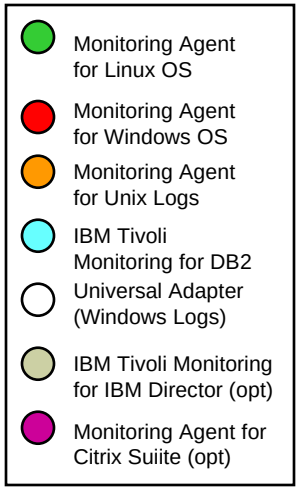
Cloud Hardware & Software SNMP Event Mgmt

> Utilize vendor SNMP MIBs for base monitoring
- GPFS MIB
- Linux Server Base OS MIB
- XIV MIB
- Brocade MIB



SNMP traps caught by ITD TEC/Netview in SOC

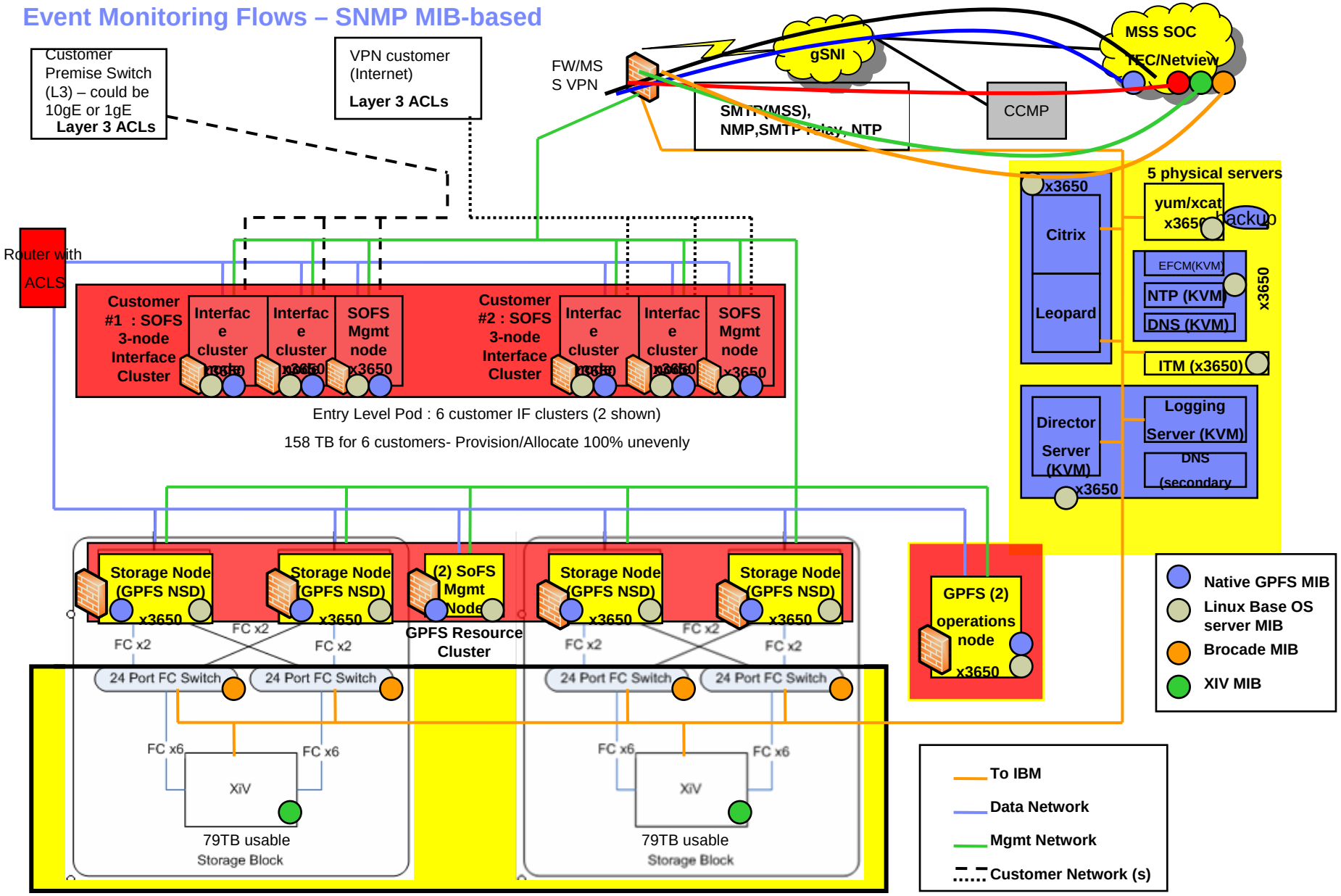
> Utilize Tivoli ITM 6.2.1 for extended monitoring to trigger SNMP
- OSS server Up/Down
- Monitoring Agent for Linux OS
- Monitoring Agent for Windows OS
- DB2 Monitoring Agent
- Monitoring Agent for Linux Logs
  - To catch SOFS-specific CTBD & Samba errors
- Other Optional agents available to extend monitoring scope



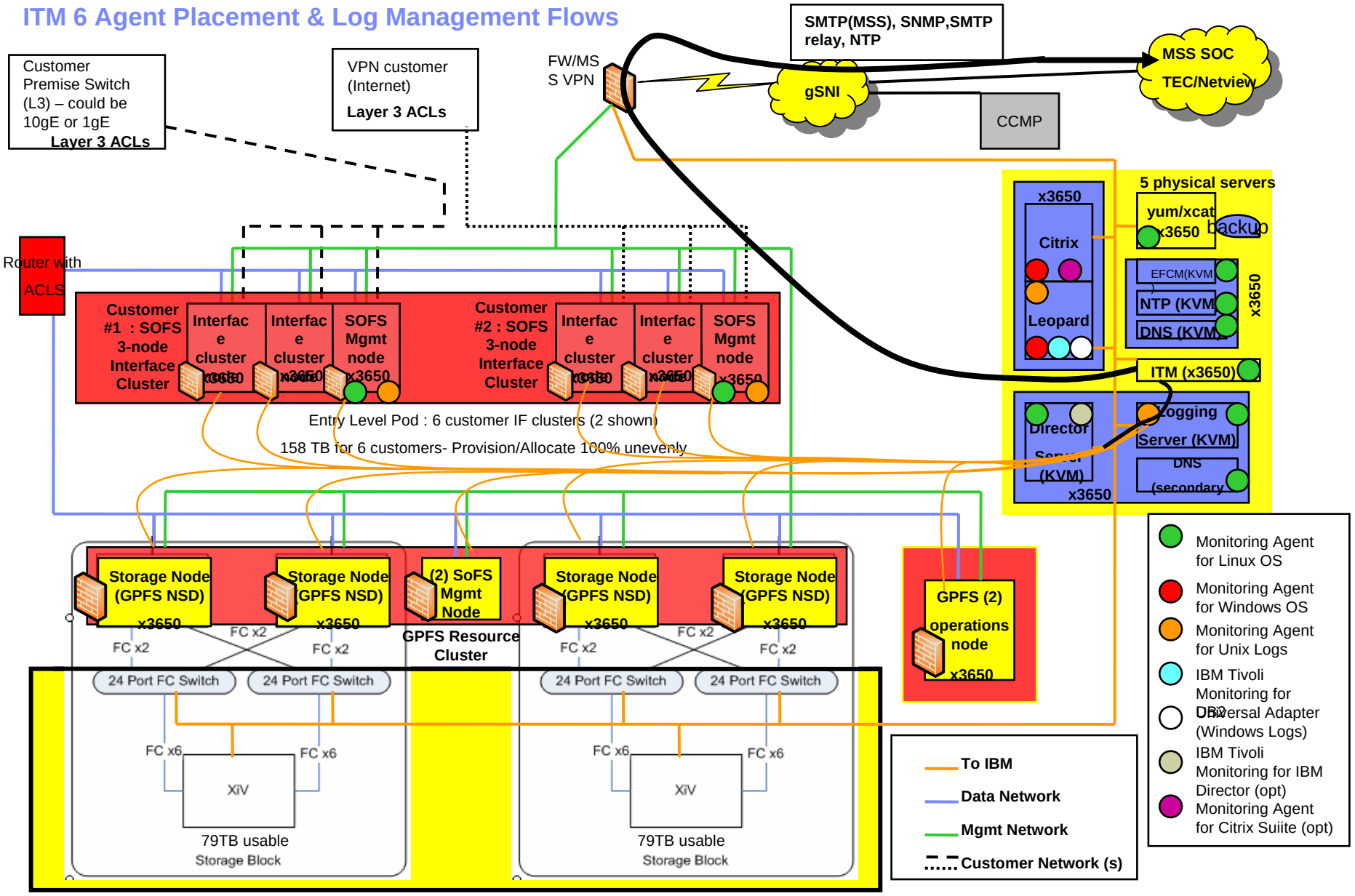
➢New Availability SLA Probe to enable real-time SNMP monitoring of each customer file system from customer premise switch

> Utilize Traditional SMTP for emails to customer service admins on quota notifications

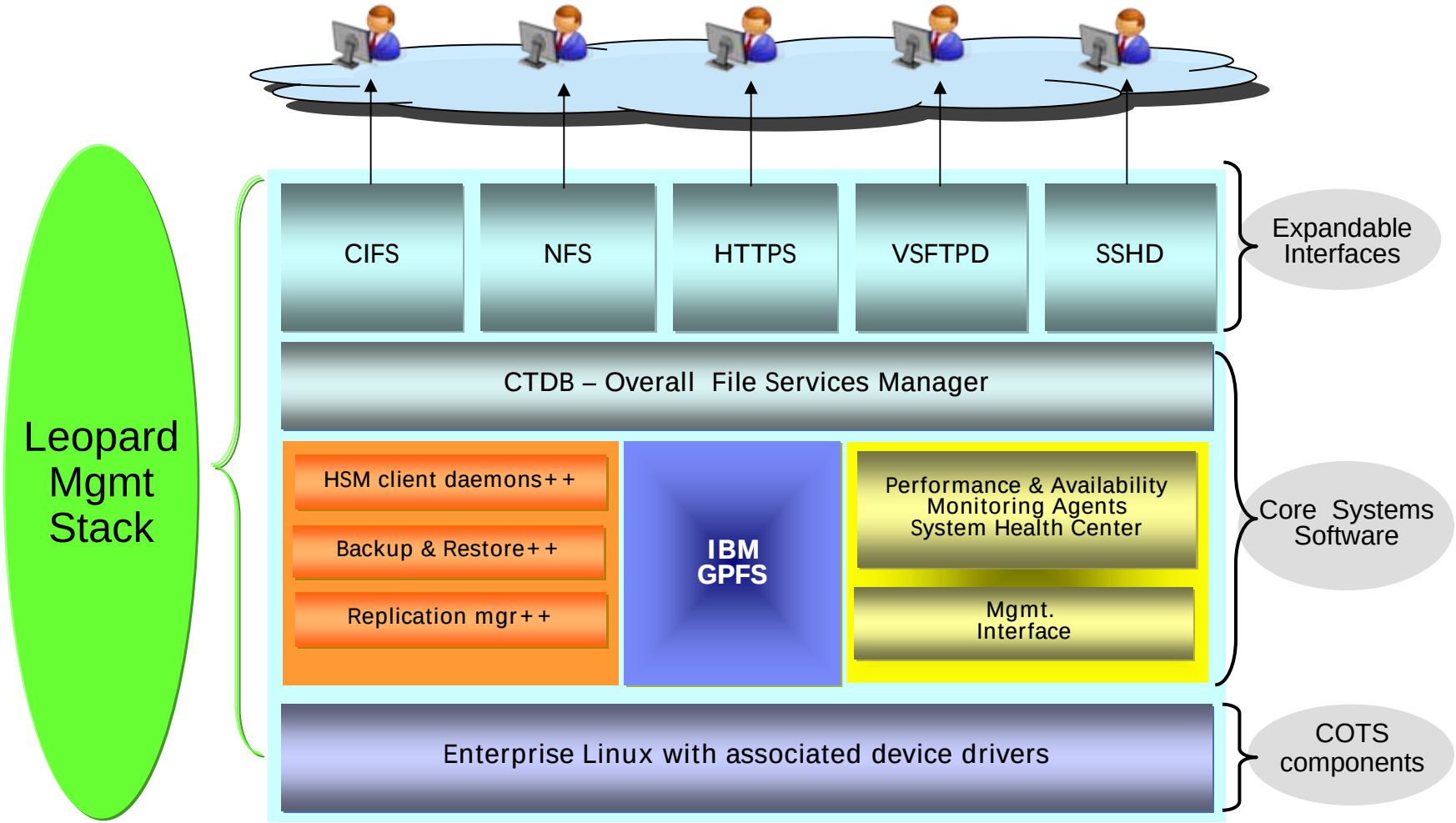
Event Monitoring Flows – SNMP MIB-based



ITM 6 Agent Placement & Log Management Flows



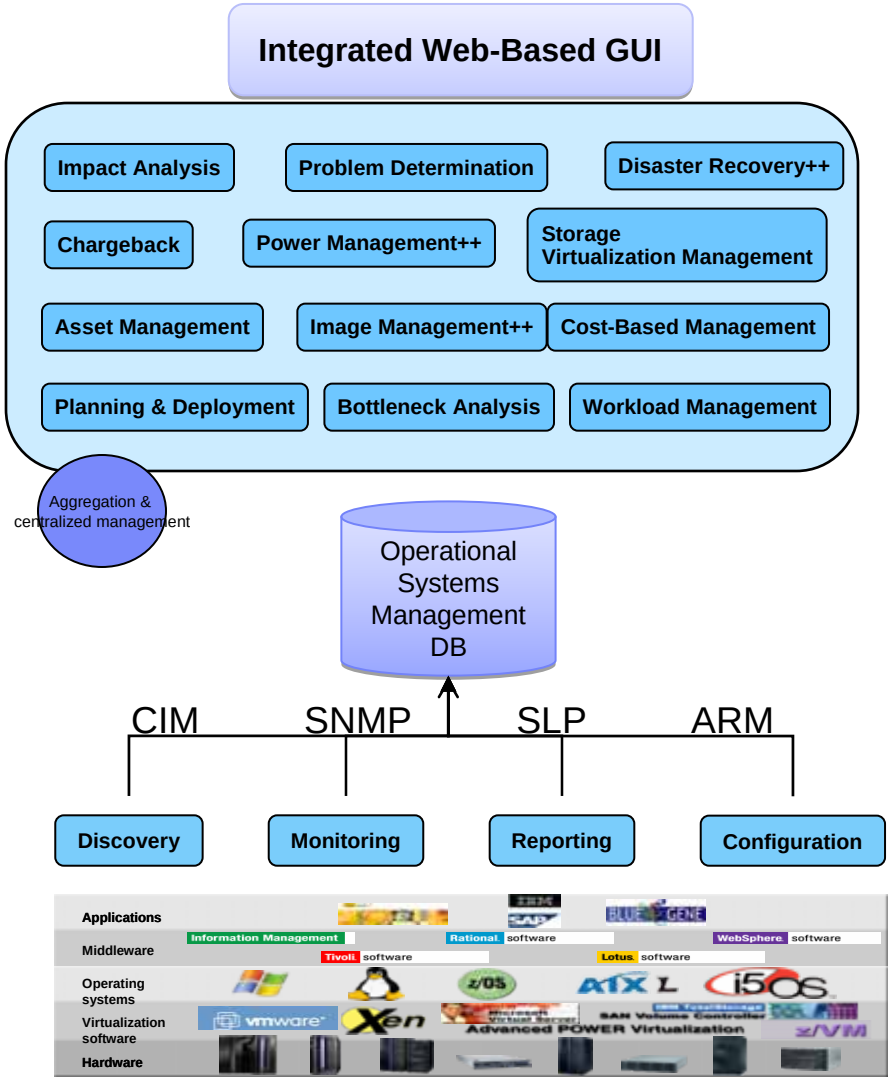
Cloud Storage software stack



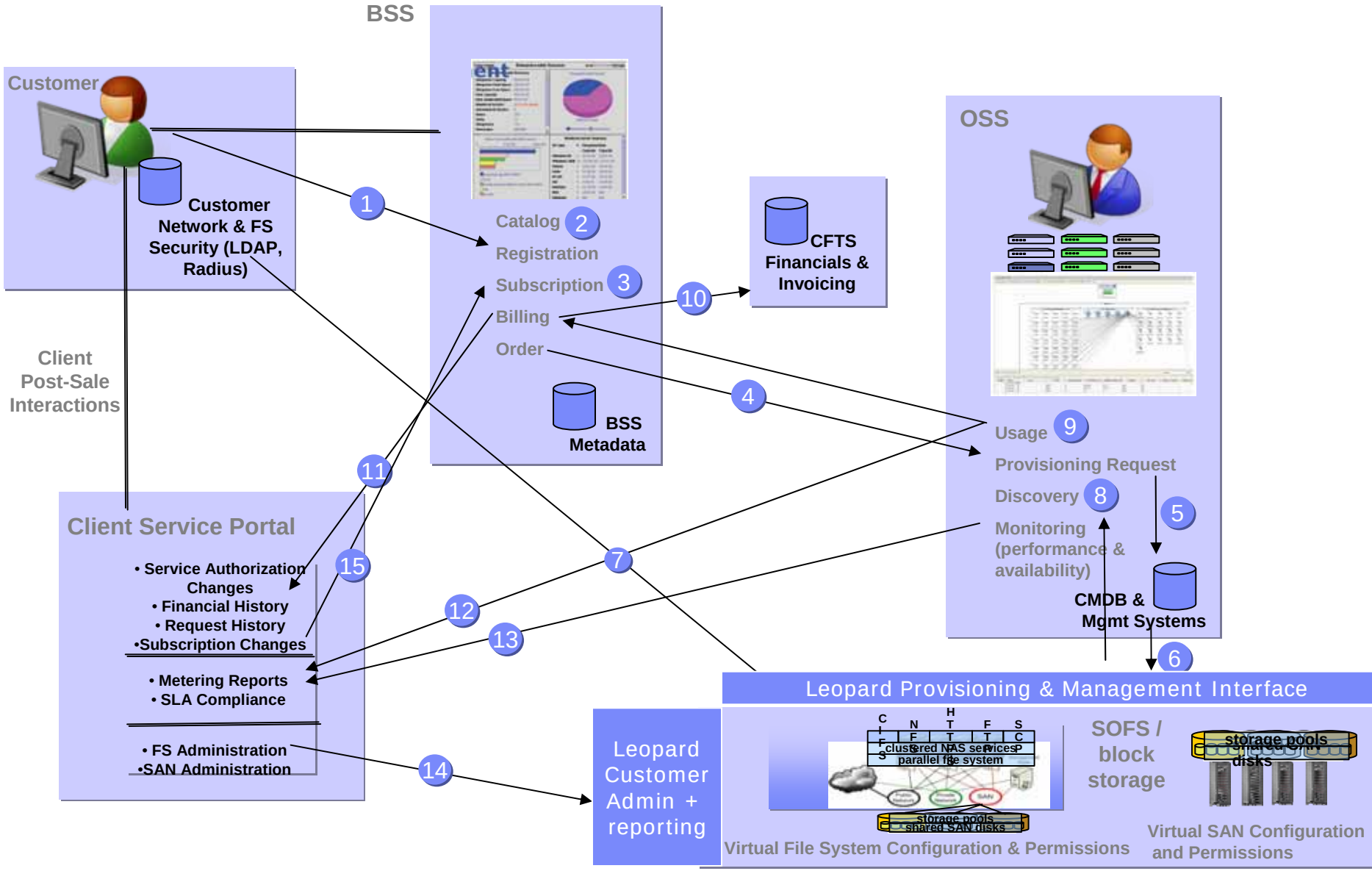
++ Not in scope for R1

TPC/Leopard Systems Management

- Improve the way in which server, storage and network resources are used in a data center
  - Centralize management
  - Greater synergy
- Automated analytics for cost-effective management
  - Virtualization Management
  - Resiliency Management
  - Problem Determination and what-if analysis
  - Scalable GPFS & NAS Management Solution
  - Energy Management
- Chargeback
- A simple and easy-to-use Web 2.0 GUI for cloud management and customer access
  - Single topology view showing servers, storage, and networks
  - Correlation of various info sources in one view (topology, alerts, charts)
  - Support for collaboration & user annotation



IBM Smart Business Storage Cloud uses TPC/Leopard for automation



VSC: CCMP + TSAM + TPC/Leopard

- End to End Monitoring
  - Health
  - Performance
  - Utilization
- End to End Provisioning Automation
  - FileSet
  - ACL
  - Quota
  - Share

Salsa: Integrated File System and Storage Provisioning and Analytics++

Problem

•Current file systems and storage provisioning mechanisms are not integrated
→ No end-to-end optimization
→ No end-to-end service differentiation or performance isolation
→ May result in wastage of resources or SLO violation or higher cost
•Manual provisioning is inefficient, costly, error-prone, time consuming

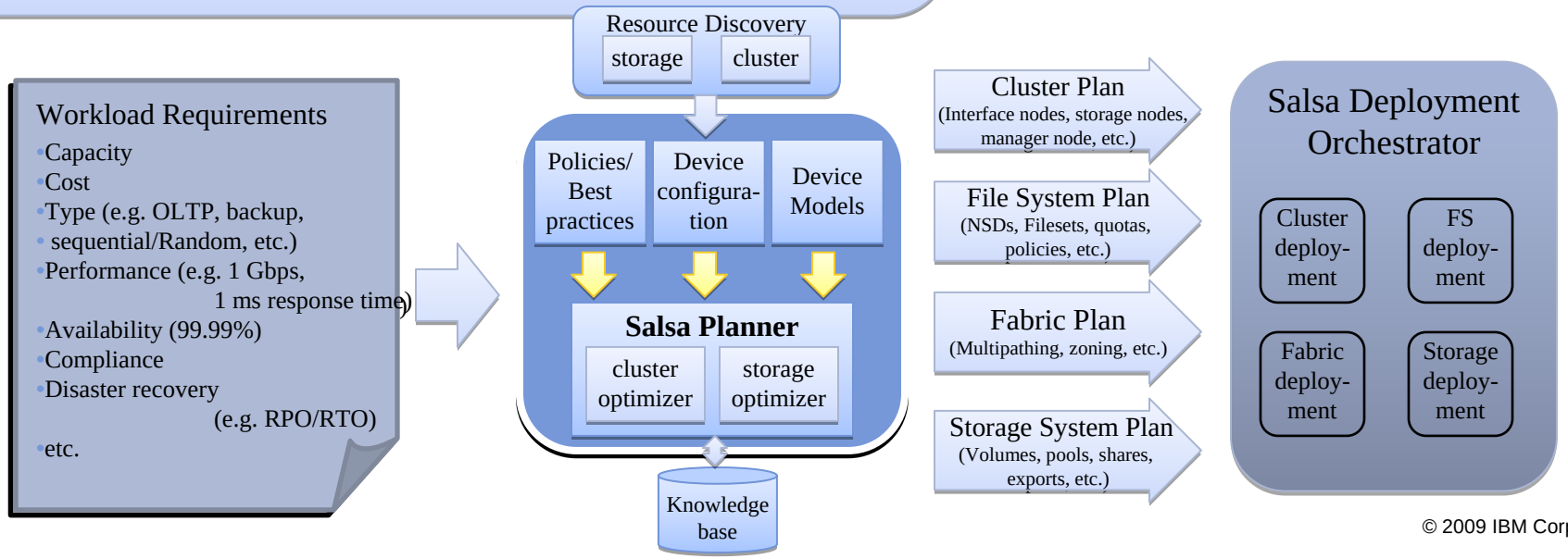
Challenges

•Resource heterogeneity, varied workload requirements, shared resources
•Multiple layer of file system and storage system virtualization
•Multi-objective multi-constraint problem → Millions of possible configurations
•(Cost, Performance, Resiliency) = Complex function of hardware and software configuration

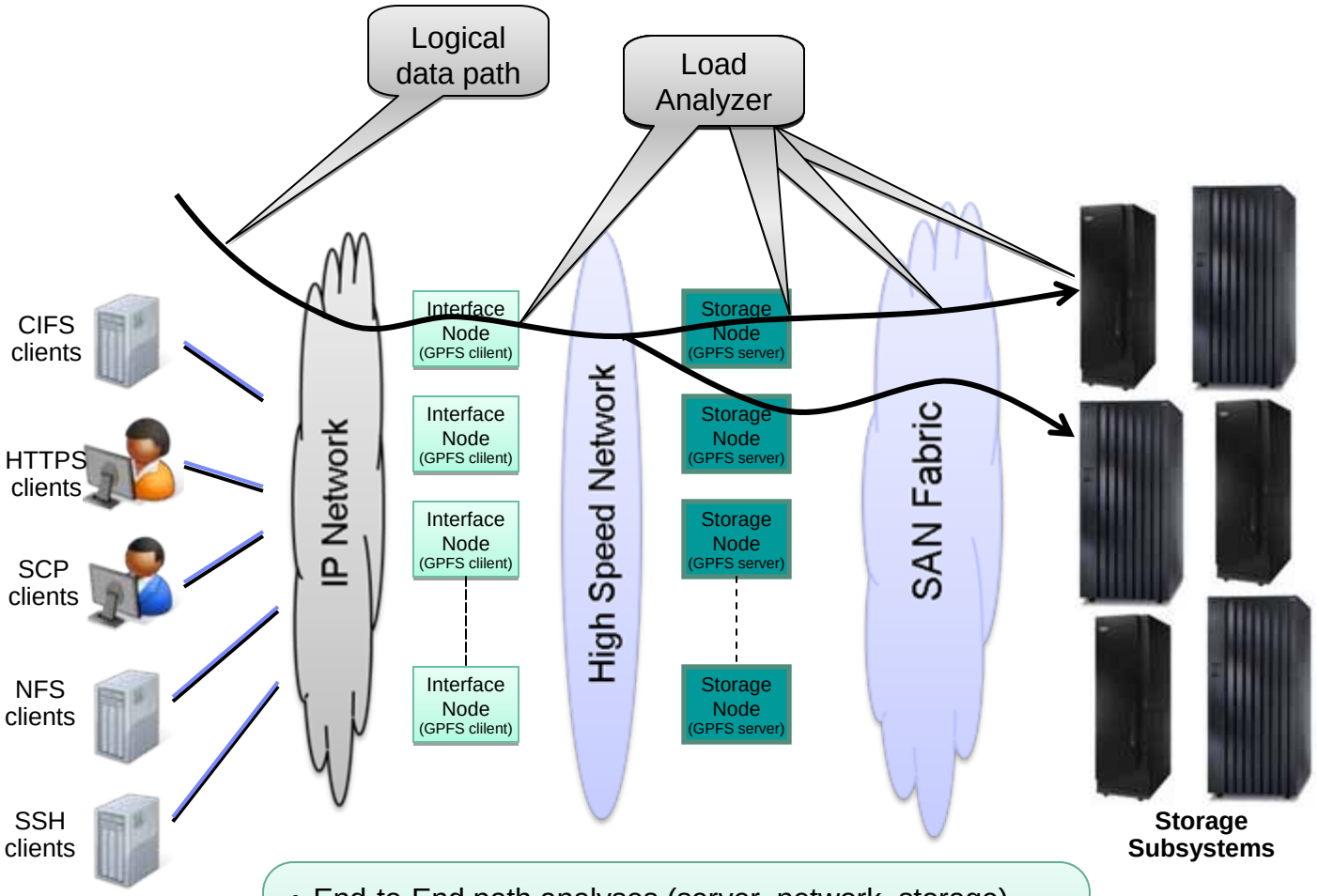
Our solution: Salsa

•Integrated tool to simplify and automate file system and storage provisioning
•Coordinated and automated deployment
•Policies and best practices automation
•End-to-End performance and configuration analyses and optimization
- Problem determination, What-if analyses

•Reduce cost, Increase efficiency
•Administrators specify high level workload requirements (cost, performance, availability, etc.)
- Not the underlying mechanisms (subsystem configuration, NSD configuration, etc.)



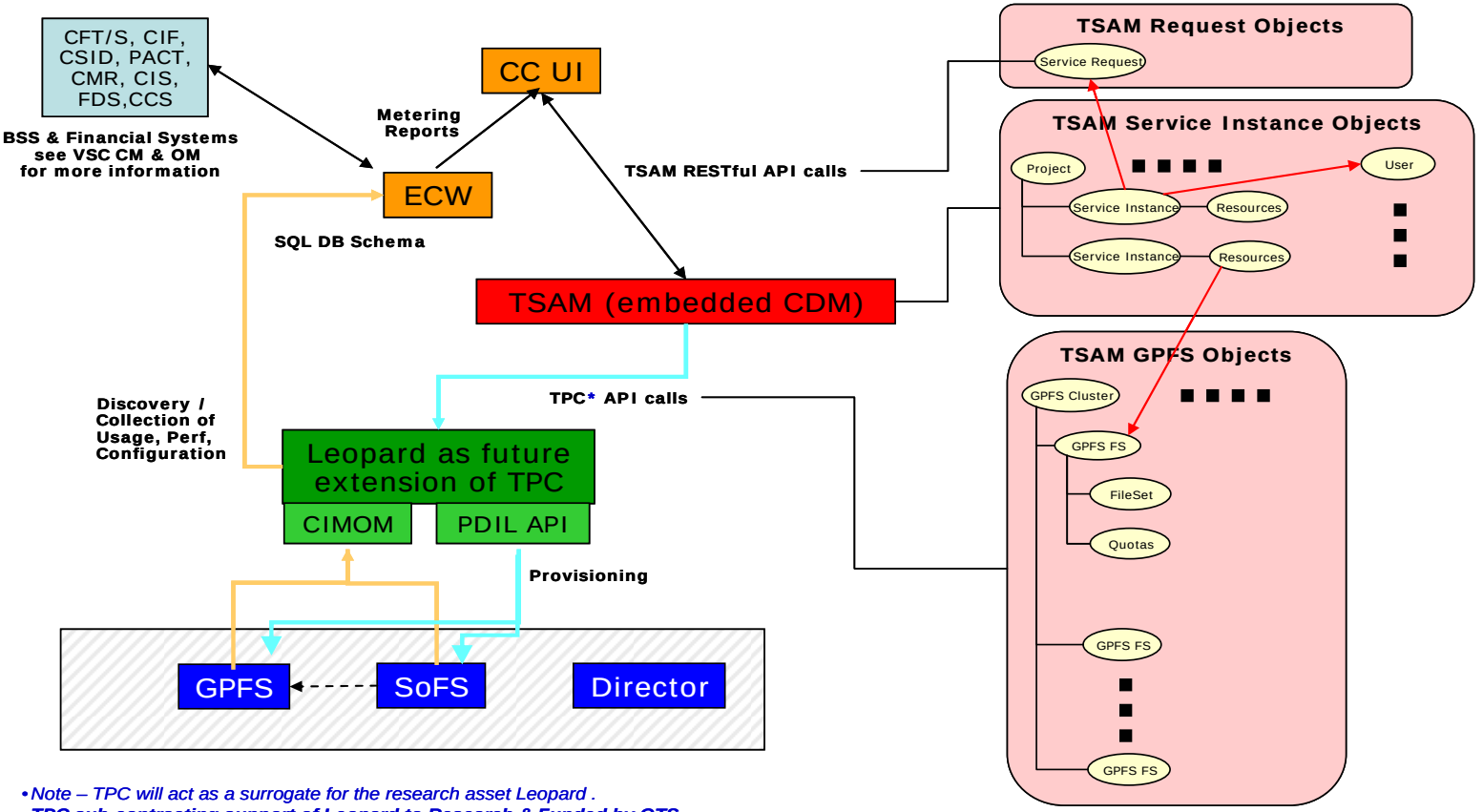
End-to-End Analytics: Problem Determination & What-If Analyses



- End-to-End path analyses (server, network, storage)
- Automated bottleneck detection
- What-If analyses
- Fast reaction to problematic symptoms

CCMP & VSC Flows

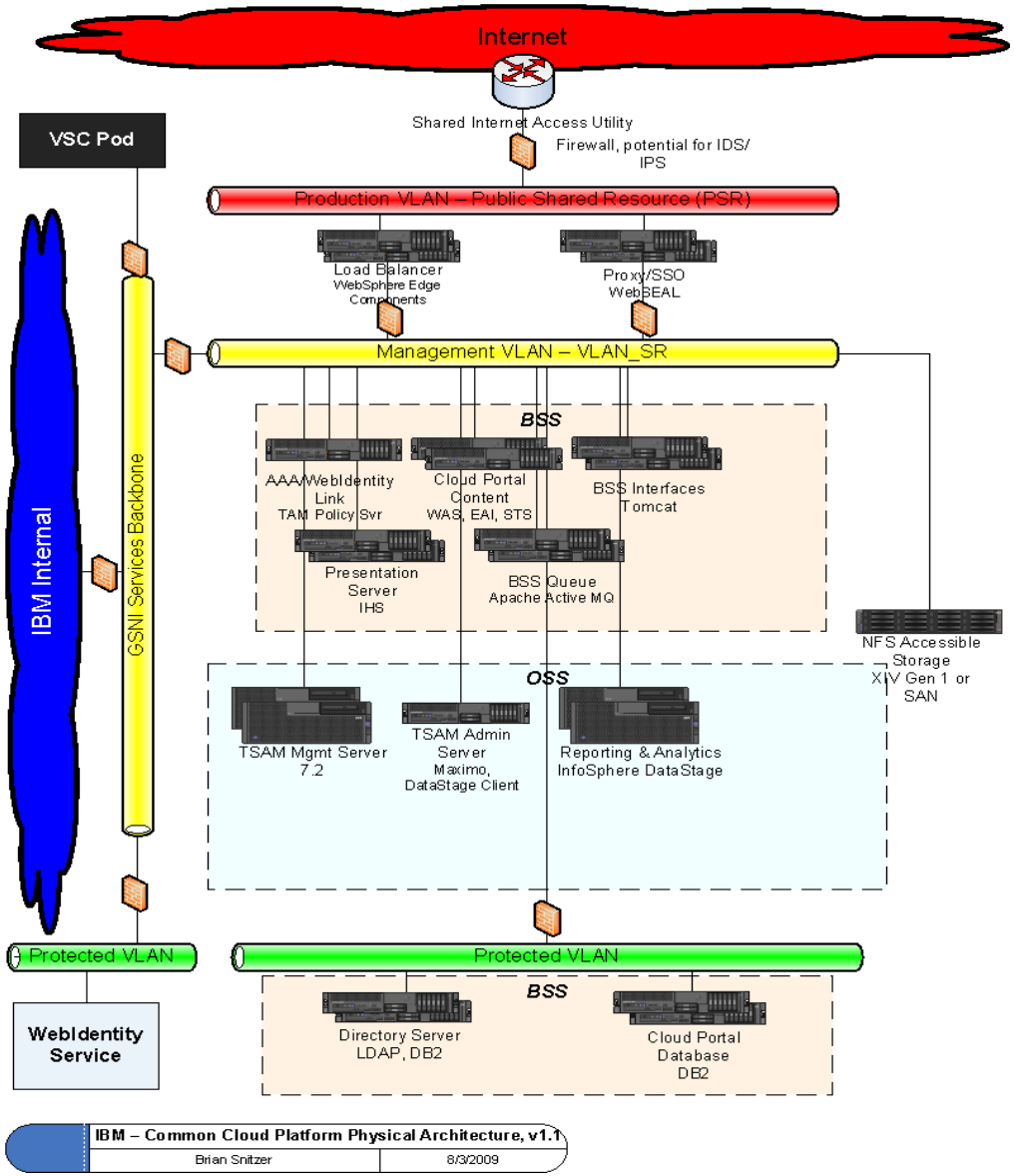
VSC Release 1 System Data Flows



- *Note – TPC will act as a surrogate for the research asset Leopard .*
- ***TPC sub-contracting support of Leopard to Research & Funded by GTS***

Common Cloud Mgmt Platform

Note: CCMP Will be doing a separate IESC Certification



Source: Common Cloud Operational Model

BSS System interaction with WDP, Web Identity, ECW, CC UI

- Flexible data load to onboard customer contract
- Monthly billing cycle coincides with customer-facing utilization reports
- 30-minute utilization snapshots with detailed billing reports
- Ability to adjust Peak Usage and enter Incidental one-time charges
- Generation of output file to CFTs

Thank you!

Questions