Performance Factors in Cloud Computing
Share Session 12778

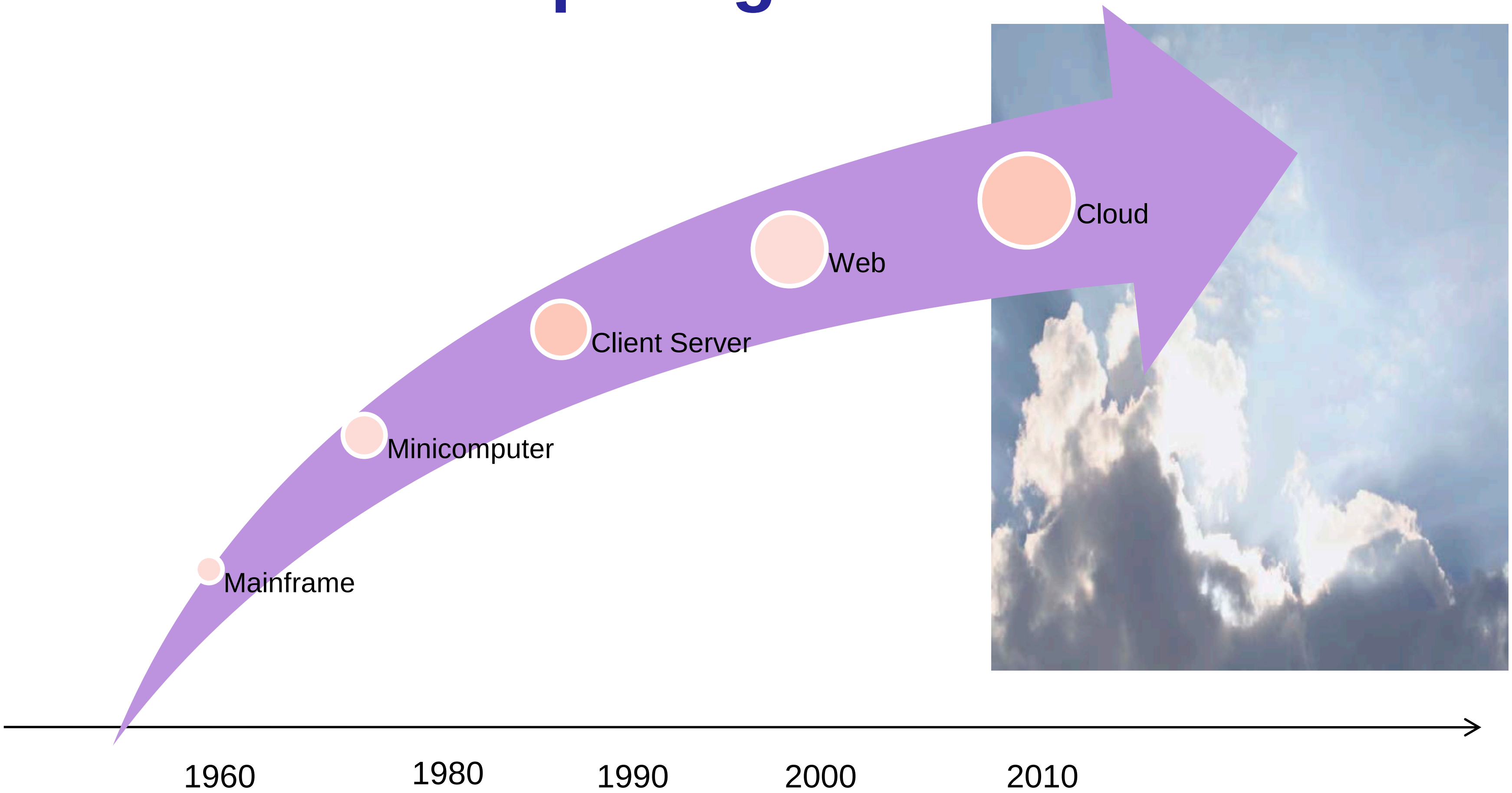
Laura Knapp
WW Business Consultant
Laurak@aesclever.com

Background

Performance Challenges

Maintaining Mainframe Network Performance in a Cloud

Cloud is a Computing Model



Why Cloud

BENEFIT	COMMENT
Cost Savings	Organizations can reduce or eliminate IT capital expenditures and reduce ongoing operating expenditures by paying only for the services they use and, potentially, by reducing the size of their IT staffs.
Flexibility	Cloud computing offers more flexibility (often called “elasticity”) in matching IT resources to business functions than past computing methods.
Scalability	Organizations using cloud computing need not scramble to secure additional hardware and software when user loads increase, but can instead add and subtract capacity as the network load dictates.
Access to Top-End IT Capabilities	Particularly for smaller organizations, cloud computing can allow access to hardware, software, and IT staff of a caliber far beyond that which they can attract and/or afford for themselves.
Redeployment of IT Staff	By reducing or doing away with constant server updates and other computing issues, and eliminating expenditures of time and money on application development, organizations may be able to concentrate at least some of their IT staff on higher-value tasks.
Focusing on Core Competencies	Arguably, the ability to run data centers and to develop and manage software applications is not necessarily a core competency of most organizations.
Sustainability	The poor energy efficiency of most existing data centers, due to substandard design or inefficient asset utilization, is now understood to be environmentally and economically unsustainable.

Cloud Types

PRIVATE CLOUD

Operated solely for an organization

COMMUNITY CLOUD

Shared by several organizations and supports a specific community that has shared concerns

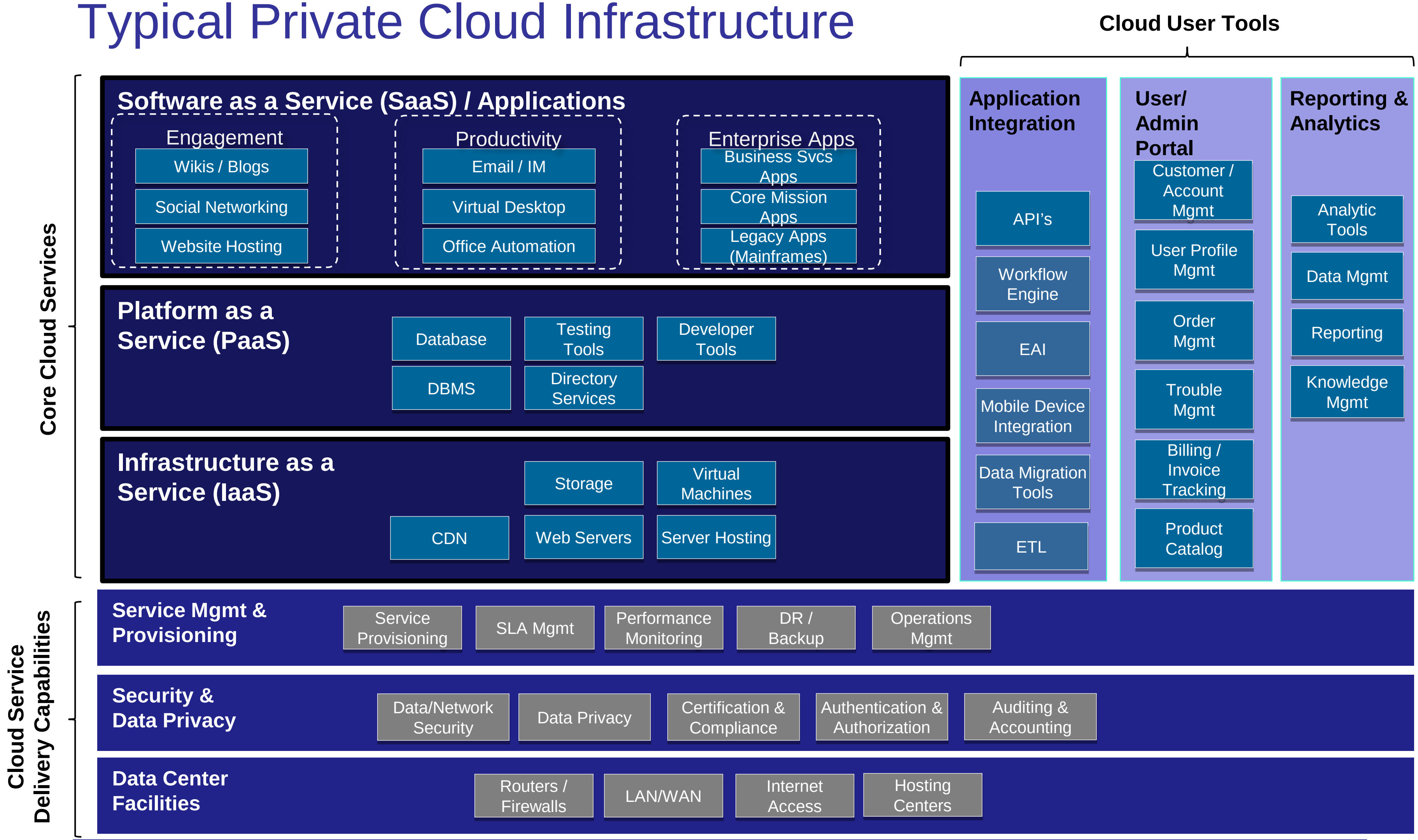
PUBLIC CLOUD

Made available to the general public or a large industry group and is owned by an organization selling cloud services

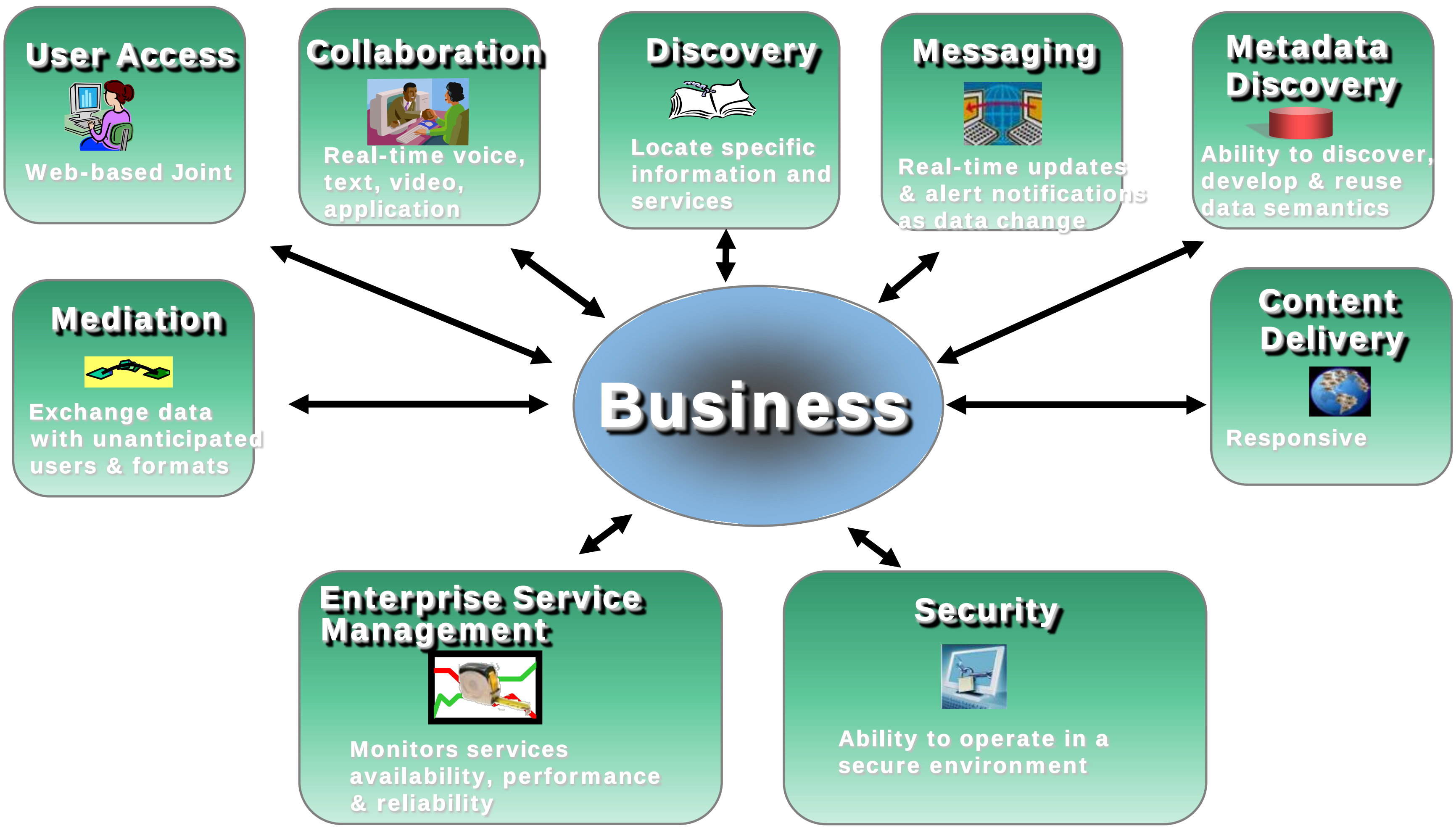
HYBRID CLOUD

Composition of two or more clouds (private, community, or public) that remain unique entities but are bound together

Typical Private Cloud Infrastructure

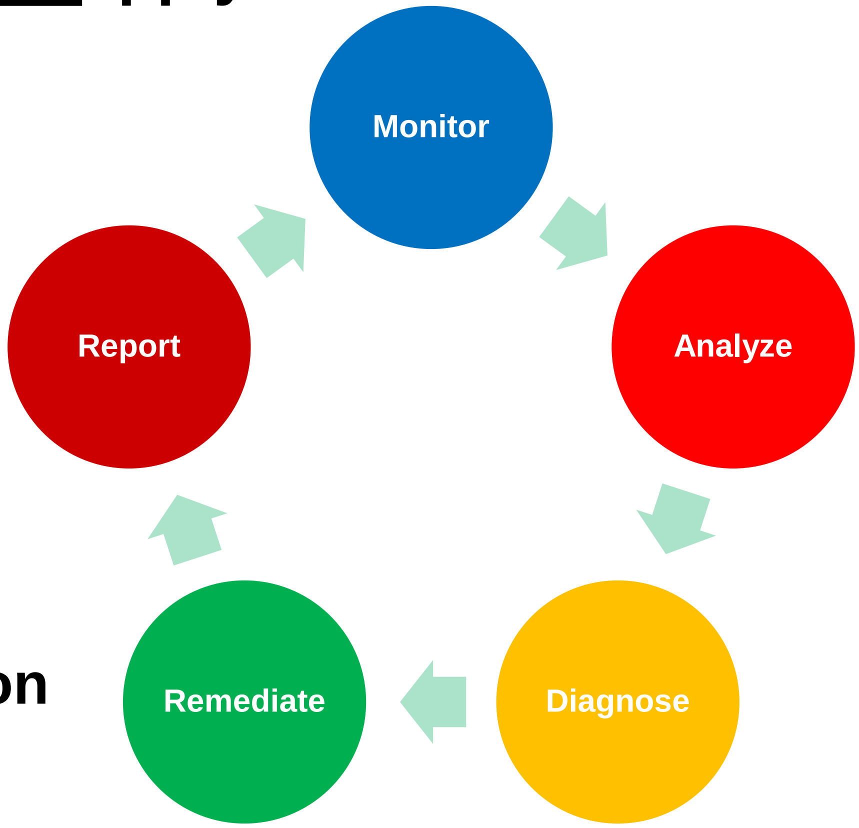


Core Business Services



Managing Cloud Data Center

- Fundamentals of management apply FCAPS
  - Fault
  - Configuration
  - Availability
  - Performance
  - Security



- Leading to
  - Service level achievement
  - Optimum resource utilization
  - Highly available systems
  - High performing systems

Background

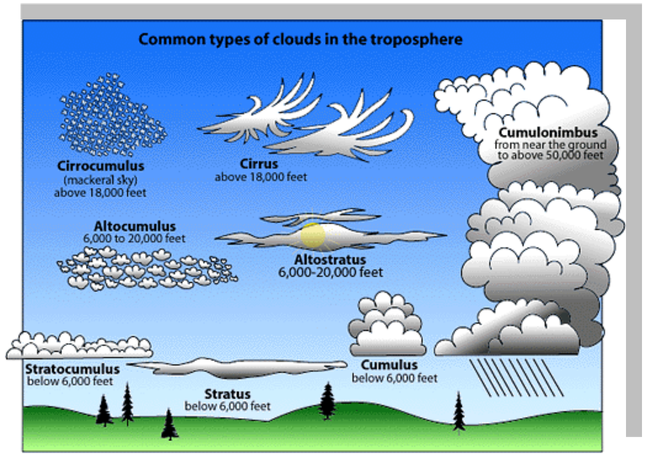
Performance Challenges

Maintaining Mainframe Network Performance in a Cloud

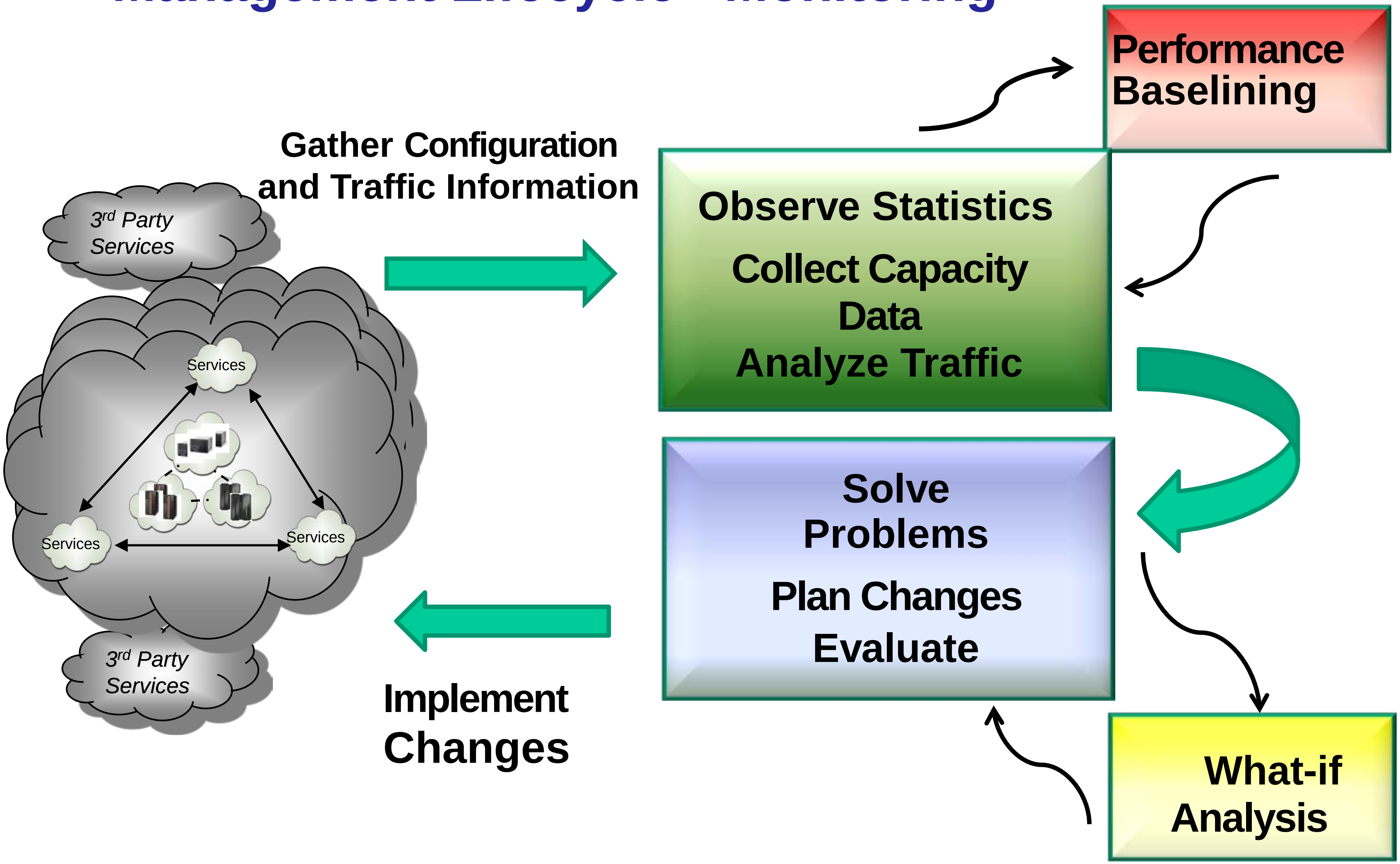
Approaches

Top Down or bottom Up doesn't matter
Consistency does

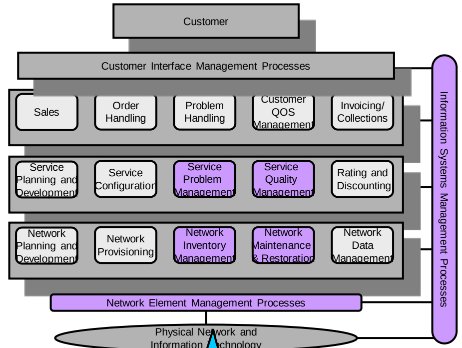
- **Applications**
- **Middleware**
- **Guest OS**
- **VM**
- **Network**



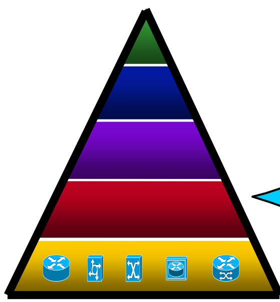
Management Lifecycle - Monitoring



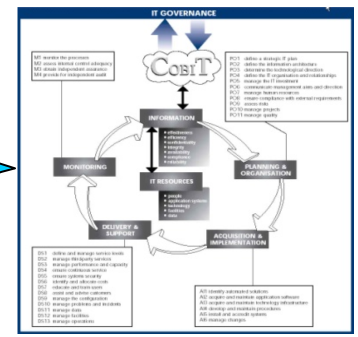
eTOM
- Extends M.3xxx
- Process & Functional Architecture
- Defines processes for providing services



Integrated Service Management



ITU – M.3xxx
- Physical Focus
- Defines interfaces & functions
- M.3400 focuses on functions
- Recommended architecture for TMN
- Recommended interfaces Q_x CMIP

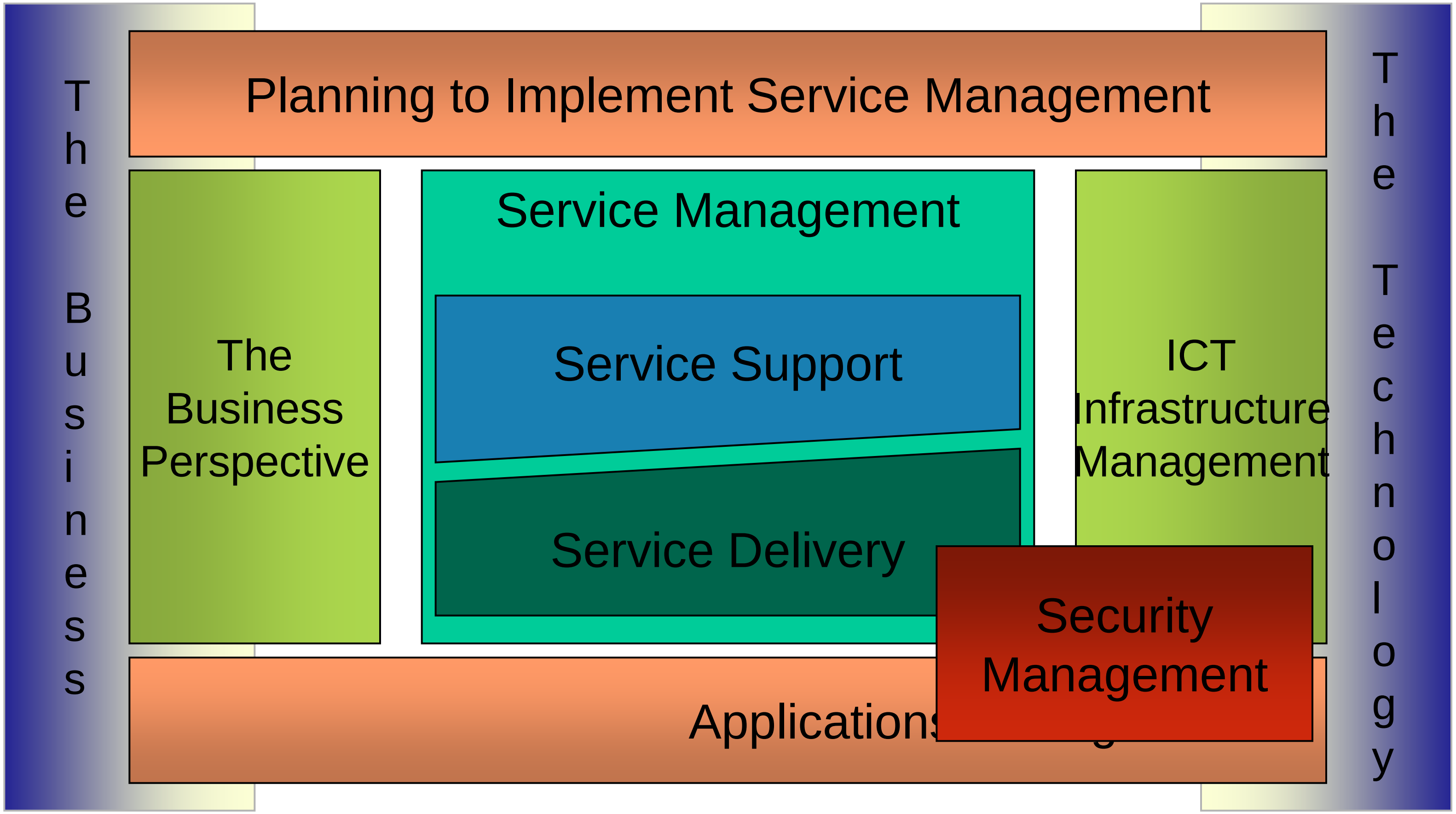


COBIT
- IT Infrastructure management focus
- IT Governance
  - Planning
  - Investment
  - Projects
  - Quality
  - Delivery
  - Support

www.ITIL.org

ITIL
- Process Focus
- IT Service management
- Service level
- Equates to COBIT Dxxx processes

ITIL


The Business
Planning to Implement Service Management
The Business Perspective
Service Management
Service Support
Service Delivery
ICT Infrastructure Management
Applications
Security Management
The Technology

Cloud Performance Risks

Challenge	Consideration	Impact
Provisioning response delay or failure	Multitude of reasons that could impede provisioning process especially faulty provisioning policies	Application or service slow down or failure
Performance unpredictability or service unavailability	Incomplete understanding of topology at time of performance issue	Significant increase in time to remediate
Arrival rates of workloads	Seasonal and time dependent aspects of workloads are not always considered	Inability of provisioning systems to respond resulting in application or service slow down or failure

Public Cloud Performance Risks

Challenge	Consideration	Impact
Isolation and visibility of component parts for performance testing	Existing performance tools are blocked from accessing component parts due to security issues	Reliance on vendor for data points and no guarantees you are getting what you pay for.
Distance and skinny straw problems	Greater distances results in higher latency and a smaller network pipe results in bottlenecks	Degradation of response time, application timeouts, application failures

Public Cloud Performance Risks

Challenge	Consideration	Impact
Application workflow characteristics	Applications with the same number of tasks, data transfer quantities have different resource use characteristics	Definitive trade-off between performance and cost. Weigh the resource performance versus the overall cost
I/O, memory, CPU, and VM usage	Pricing is often based on usage of these resources	Does cloud provisioning software take cost issue into decision process
I/O bound applications	Require high performance infrastructure	Use of commodity infrastructure components may impede response time

Steps to Effective Performance Management

Monitor

Establish Baselines

Monitor over a long period of time to develop baselines and trends

Data analysis with no preconceived bias for capacity and performance trend development and any time dependencies

Report on trends, changes, and exceptions

Baseline re-evaluation and resetting

Review and Remediate

Analyze

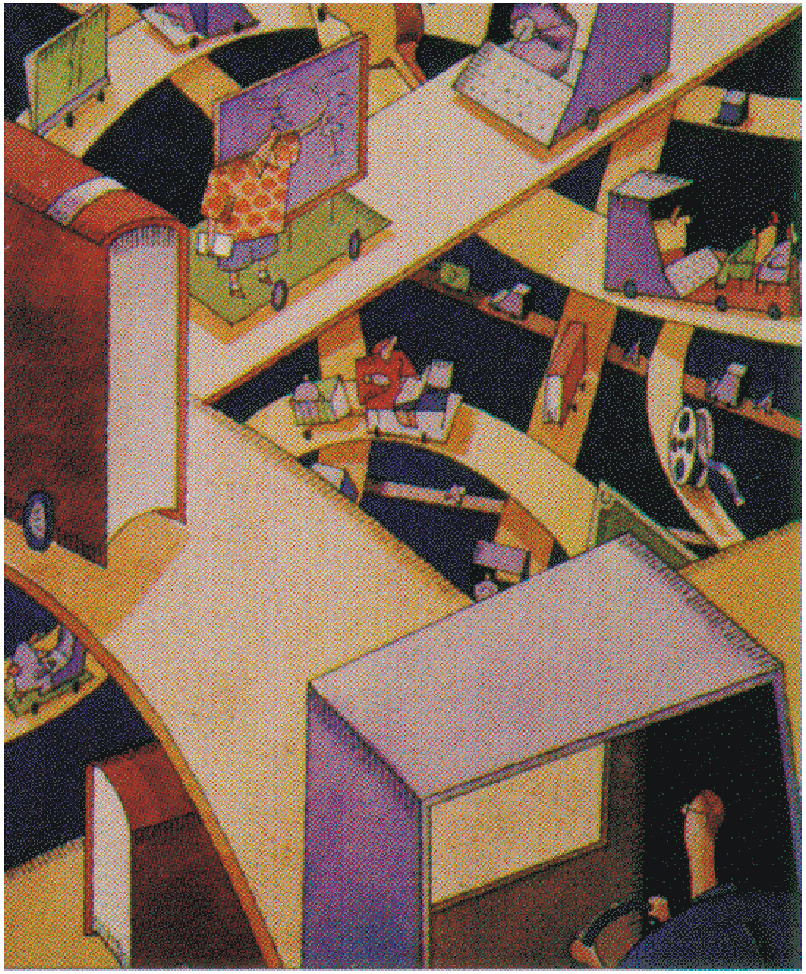
Background

Performance Challenges

Maintaining Mainframe Network Performance in a Cloud

IP Resource Bottlenecks

CPU
Memory
Buffering, queuing, and latency
Interface and pipe sizes
Network capacity
Speed and Distance
Application Characteristics

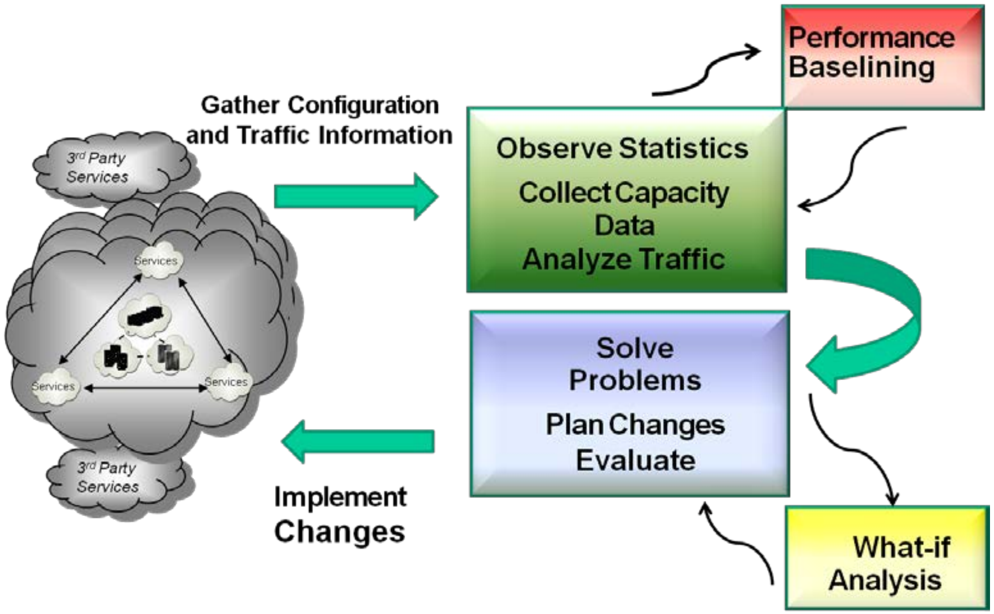


Results in:

Network capacity problems
Utilization overload
Application slowdown or failure

Information to Collect

- Link/segment utilization
- CPU Utilization
- Memory utilization
- Response Time
- Round Trip Time
- Queue/buffer drops
- Broadcast volumes
- Traffic shaping parameters
- RMON statistics
- Packet/frame drop/loss
- Environment specific



CPU Utilization

In Virtualized systems CPU utilization can be misleading

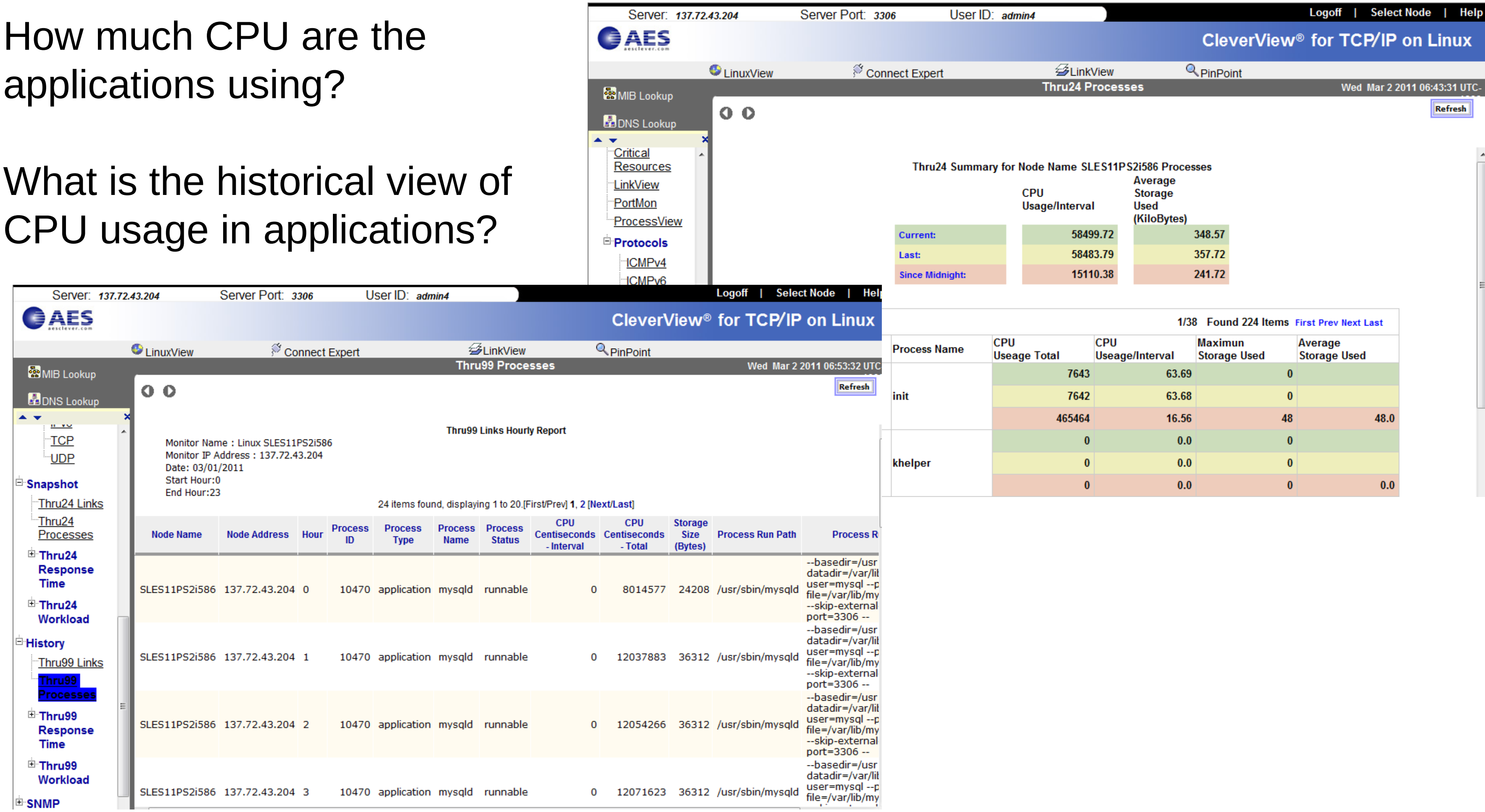
Running low on CPU any system can cause
- immediate application failure
- system slowdown impacting all applications
- need to restart system

Running low on CPU can
- cause immediate application failure
- domino effect on related resources and applications
- intermittent application oddities

Questions to Answer on CPU Utilization

How much CPU are the applications using?

What is the historical view of CPU usage in applications?



Scenario 1 – Linux CPU Usage High

Situation

A client had a very successful beta with Linux on system z. As they added additional workloads onto the Linux systems overall CPU was increasing much higher then when the application was running on a standalone server.

Trouble Shooting

Using a Linux TCP/IP Monitor check the overall flow of information through both the IP and TCP layers. The CPU utilization was viewed over time. Verify that listeners are available for the applications. View alerts and determine if any would suggest the problem being seen. Check the buffer count. In this system the buffer count had never been raised and was still set at 16.

Solution

Increasing the buffer to 50 reduced the CPU utilization for this linux server as we added more applications.

As you increase the buffer additional memory will be used

SUSE SLES11: in /etc/udev/rules.d/51-qeth-0.0.f200.rules add ACTION=="add", SUBSYSTEM=="ccwgroup", KERNEL=="0.0.f200", ATTR{buffer_count}="128"

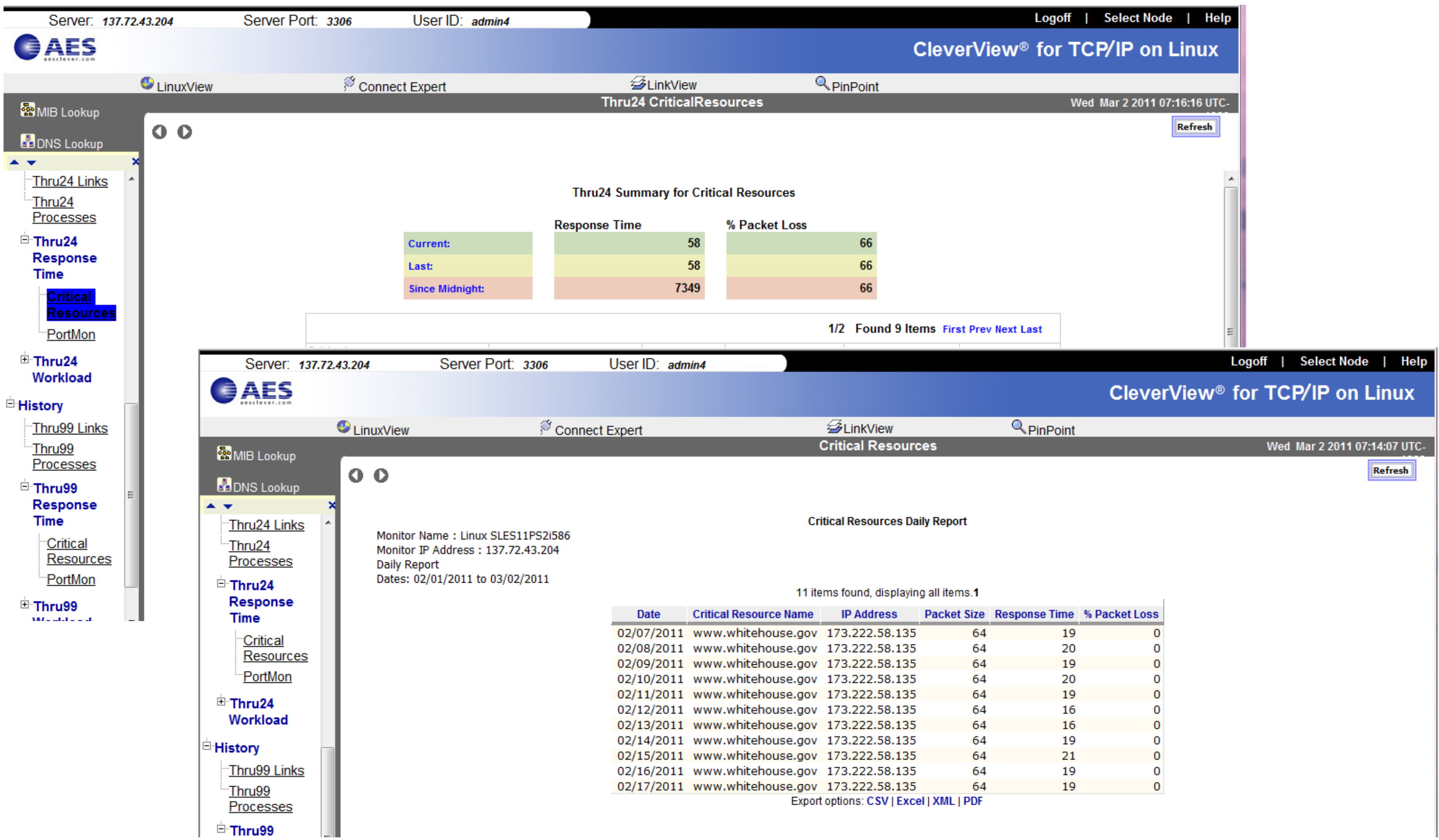
Response Time

No one is ever happy with what they get

External customers may go elsewhere

Where is the problem?
- Network?
  - Router have long queus?
  - Is the LAN to slow?
  - Is the route long?
- Operating system?
  - Too long to queue for transmit?
- Application?
- Protocol?
  - Window size improperly set?
  - MTU size improperly set?

Now and Historical Response Time



	Response Time	% Packet Loss
Current:	58	66
Last:	58	66
Since Midnight:	7349	66

Date	Critical Resource Name	IP Address	Packet Size	Response Time	% Packet Loss
02/07/2011	www.whitehouse.gov	173.222.58.135	64	19	0
02/08/2011	www.whitehouse.gov	173.222.58.135	64	20	0
02/09/2011	www.whitehouse.gov	173.222.58.135	64	19	0
02/10/2011	www.whitehouse.gov	173.222.58.135	64	20	0
02/11/2011	www.whitehouse.gov	173.222.58.135	64	19	0
02/12/2011	www.whitehouse.gov	173.222.58.135	64	16	0
02/13/2011	www.whitehouse.gov	173.222.58.135	64	16	0
02/14/2011	www.whitehouse.gov	173.222.58.135	64	19	0
02/15/2011	www.whitehouse.gov	173.222.58.135	64	21	0
02/16/2011	www.whitehouse.gov	173.222.58.135	64	19	0
02/17/2011	www.whitehouse.gov	173.222.58.135	64	19	0

Scenario 2– Slow Application Response

Situation

A client had a Linux on system environment and they were about ready to grow the production use of Linux. One of the applications accessed an outside website which was critical to the service the application provided. As they moved the application to a virtualized system they noticed a decline in response time. What was causing the added time?

Trouble Shooting

Using a Linux TCP/IP Monitor check the overall flow of information through both the IP and TCP layers. Since outside resources were required they were set up as critical resources and monitored for packet loss and response time. The response times were measured before the move and after the move.

Solution

It was determined that after the move the firewall in front of the virtualized server needed to be reconfigured in order to return the overall response time to normal.

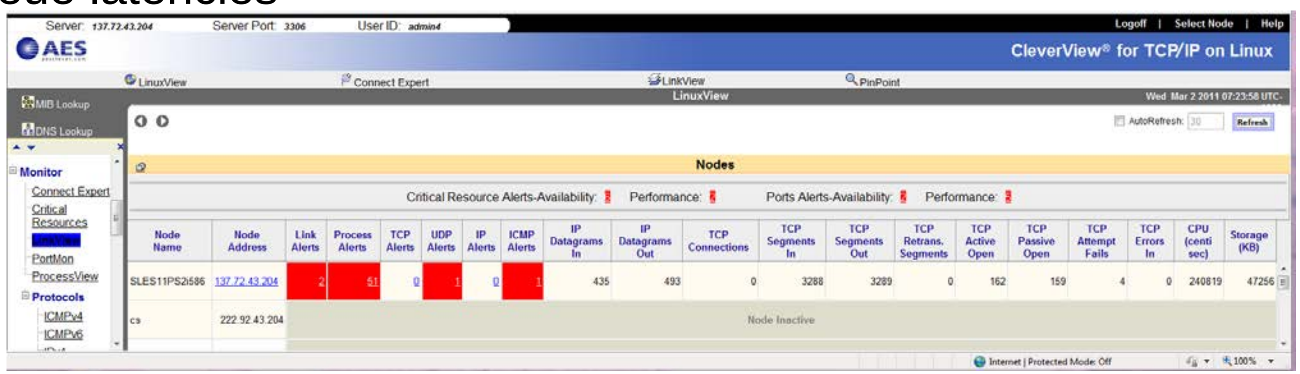
System Utilization

Since you cannot over-provision your system (add as much memory as you want, as much DASD, etc) you need to optimize

Determining what is currently being used on the system will assist in determining how much you can grow the system

An application behaving poorly may be due to improper design, improper setting of system resources to use, or application configuration

Sluggishness of a system may be due to not enough CPU, I/O overloads, or queue latencies



Scenario 3– Can I Add more Applications

Situation

A task force was recommending adding additional applications to the virtualized mainframe. The initial move went well and they wanted to increase the usage of Linux and decrease their distributed servers. The task force approved the move without looking at any data to see if the system could handle the workload.

Trouble Shooting

Due to the environment OSA was inspected to see if it could handle the traffic. CPU utilization was investigated on both the VM and Linux partitions. On the Linux system the ethernet interface was checked to see how loaded it was. While the task force made a broad and quick decision a lot of worked followed to ensure a tuned system.

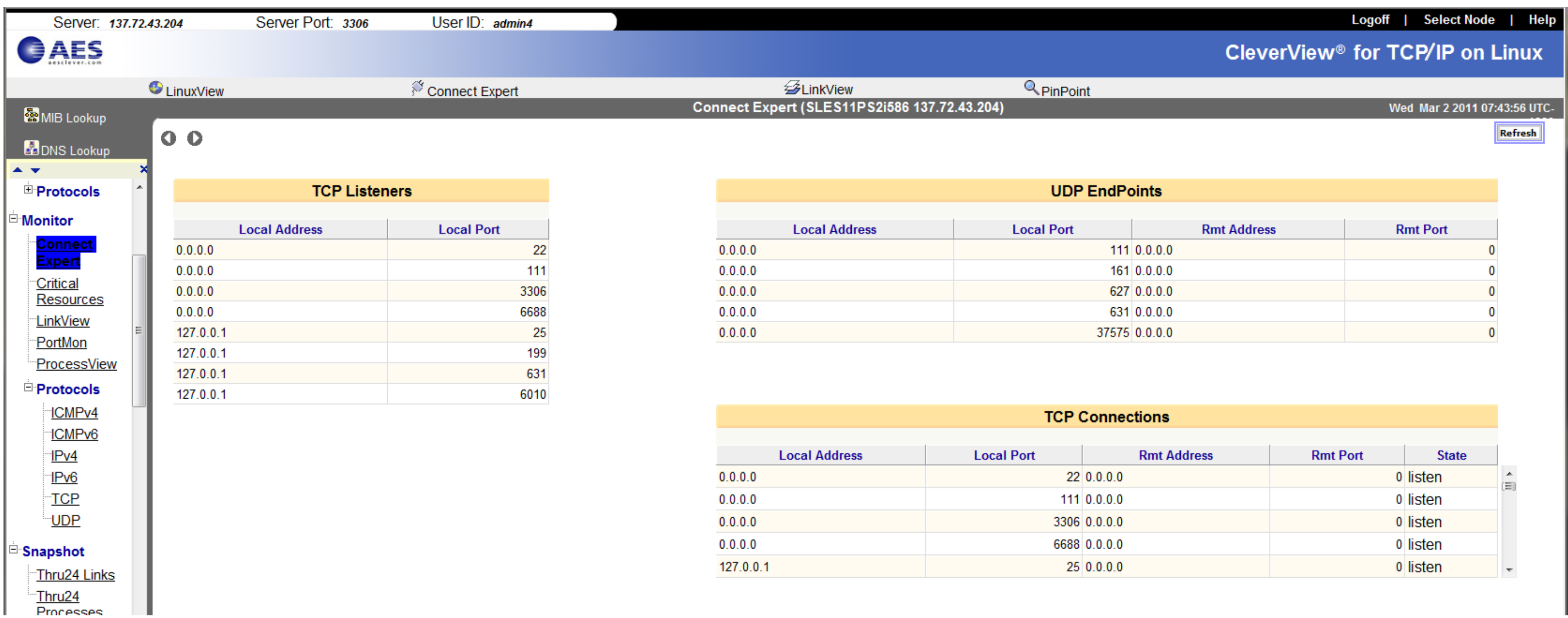
Solution

In order to prevent future fragmentation issues we reset the MTU size to 1492 and defined that as the standard for their linux systems. While this didn't cause an issue when the workload on Linux was small over time it could be a major problem.

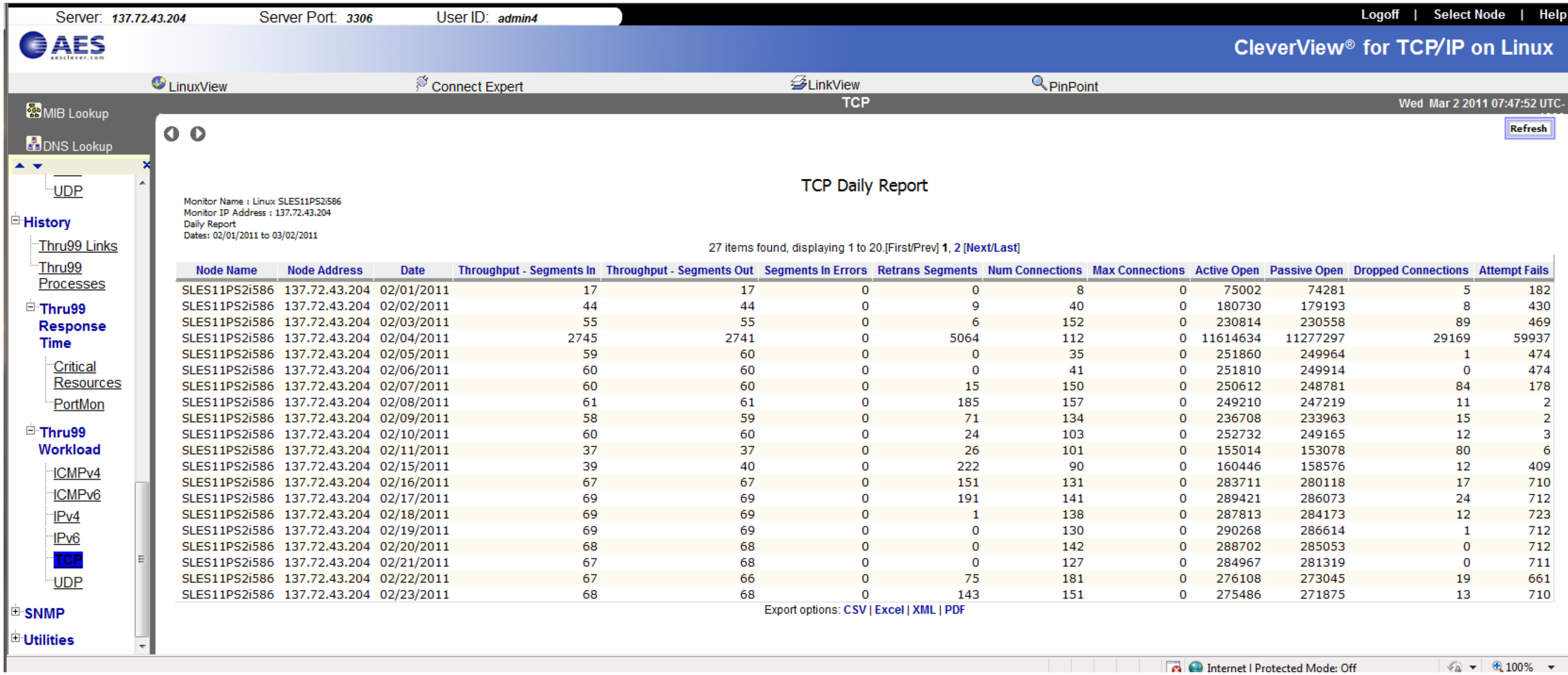
Overall Connections

Most Resources, applications, network components connect with either TCP or UDP

If a TCP listen is not available then a service will not be able to function



Connections



TCP Daily Report

Monitor Name : Linux SLES11PS2i586
Monitor IP Address : 137.72.43.204
Daily Report
Dates: 02/01/2011 to 03/02/2011

27 items found, displaying 1 to 20.[First/Prev] 1, 2 [Next/Last]

Node Name	Node Address	Date	Throughput - Segments In	Throughput - Segments Out	Segments In Errors	Retrans Segments	Num Connections	Max Connections	Active Open	Passive Open	Dropped Connections	Attempt Fails
SLES11PS2i586	137.72.43.204	02/01/2011	17	17	0	0	8	0	75002	74281	5	182
SLES11PS2i586	137.72.43.204	02/02/2011	44	44	0	9	40	0	180730	179193	8	430
SLES11PS2i586	137.72.43.204	02/03/2011	55	55	0	6	152	0	230814	230558	89	469
SLES11PS2i586	137.72.43.204	02/04/2011	2745	2741	0	5064	112	0	11614634	11277297	29169	59937
SLES11PS2i586	137.72.43.204	02/05/2011	59	60	0	0	35	0	251860	249964	1	474
SLES11PS2i586	137.72.43.204	02/06/2011	60	60	0	0	41	0	251810	249914	0	474
SLES11PS2i586	137.72.43.204	02/07/2011	60	60	0	15	150	0	250612	248781	84	178
SLES11PS2i586	137.72.43.204	02/08/2011	61	61	0	185	157	0	249210	247219	11	2
SLES11PS2i586	137.72.43.204	02/09/2011	58	59	0	71	134	0	236708	233963	15	2
SLES11PS2i586	137.72.43.204	02/10/2011	60	60	0	24	103	0	252732	249165	12	3
SLES11PS2i586	137.72.43.204	02/11/2011	37	37	0	26	101	0	155014	153078	80	6
SLES11PS2i586	137.72.43.204	02/15/2011	39	40	0	222	90	0	160446	158576	12	409
SLES11PS2i586	137.72.43.204	02/16/2011	67	67	0	151	131	0	283711	280118	17	710
SLES11PS2i586	137.72.43.204	02/17/2011	69	69	0	191	141	0	289421	286073	24	712
SLES11PS2i586	137.72.43.204	02/18/2011	69	69	0	1	138	0	287813	284173	12	723
SLES11PS2i586	137.72.43.204	02/19/2011	69	69	0	0	130	0	290268	286614	1	712
SLES11PS2i586	137.72.43.204	02/20/2011	68	68	0	0	142	0	288702	285053	0	712
SLES11PS2i586	137.72.43.204	02/21/2011	67	68	0	0	127	0	284967	281319	0	711
SLES11PS2i586	137.72.43.204	02/22/2011	67	66	0	75	181	0	276108	273045	19	661
SLES11PS2i586	137.72.43.204	02/23/2011	68	68	0	143	151	0	275486	271875	13	710

Export options: CSV | Excel | XML | PDF

Scenario 4– Excessive Segmentation

Situation

As you can see on the previous chart on 2/4/2011 there were a significant number of segmented TCP packets, dropped connections, and failed attempts. What was going on?

Trouble Shooting

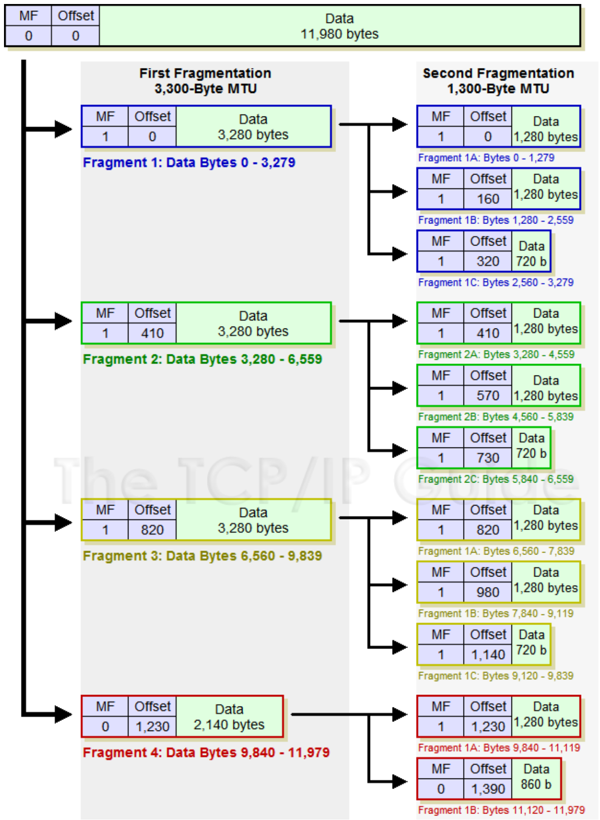
Using a Linux TCP/IP Monitor check the overall flow of information through both the IP and TCP layers. The OSA adapter was inspected and traffic was moving through it smoothly. Look at the MTU settings on your links and the fragmentation on the IP stack. While there was not significant fragmentation, the MTU size was set at 1500. This wasn't a good value for IP fragments, but this would not impact TCP Segmentation.

Solution

It was clear that this Linux system was not using 'Large-Send' The default for Linux is no. We changed this to TSO which now had segmentation done by the OSA adapter freeing up resources in the Linux system.

MTU Size

- Optimizing MTU size can provide optimum performance improvements
- Set the maximum size supported by all hops between the source and destination
- Traceroute can provide details on the MTU size but some router administrators block traceroute
- If you application sends
- frames <= 1400 bytes use an MTU size of 1492
- Jumbo frames use and MTU size of 8992
- TCP uses MTU size for window size calculation
- For VSWITCH an MTU of 8992 is recommended



Scenario 6– Excessive Fragmentation

Situation

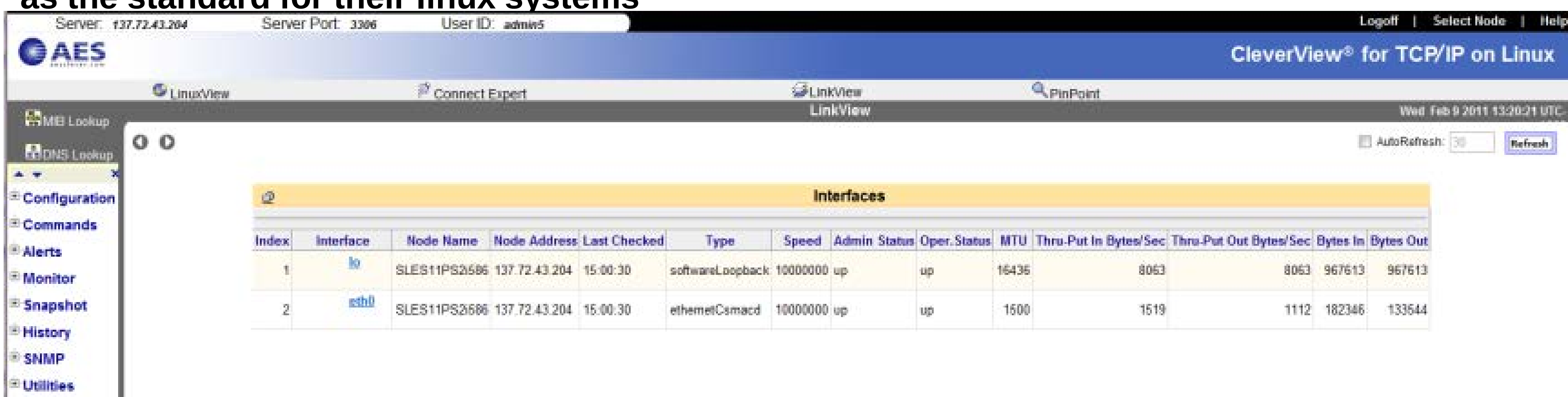
A client had a Linux on system environment and they were about ready to grow the production use of Linux. While they did not have any major problems they new of they asked for an overall health check.

Trouble Shooting

Using a Linux TCP/IP Monitor check the overall flow of information through both the IP and TCP layers. Look at the MTU settings on your links and the fragmentation on the IP stack. While there was not significant fragmentation, the MTU size was set at 1500.

Solution

In order to prevent future fragmentation issues we reset the MTU size to 1492 and defined that as the standard for their linux systems



Linux: OSA LAN Timer or Blocking Timer

OSA inbound blocking function

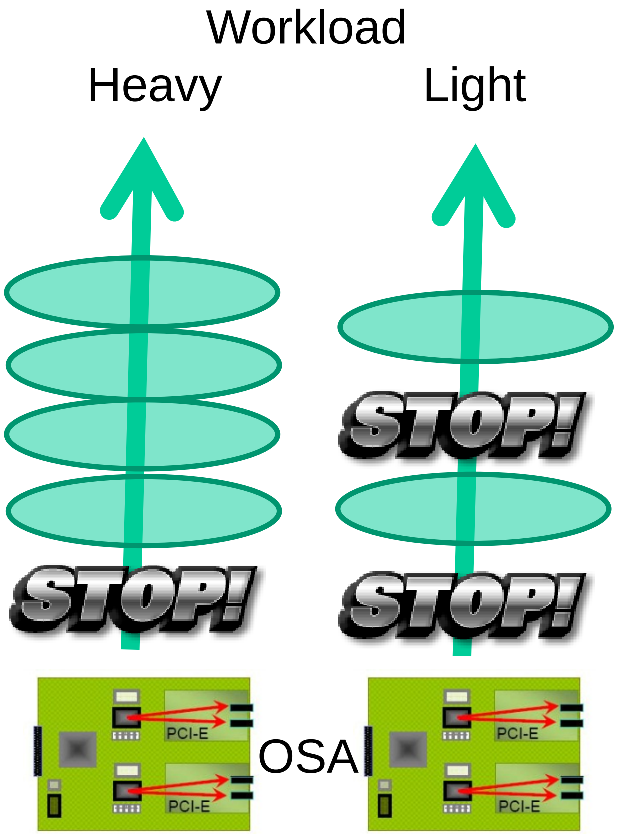
- Determines how long OSA will hold packets
- Indirectly affects
  - Frequency of host interrupt
  - Payload per interrupt

Linux has 3 potential values for OSA2

- For frames under 1536:Time between 2 incoming packets
- For Jumbo frames: Duration of inter-packet gap
- Total duration that OSA holds a single inbound buffer
- Default mode is NO LAN idle which is a good compromise for both transactional and streaming workloads

Linux behaves differently with OSAExpress3

- Using the default for OSA2 results in short latency but high CPU utilization



Scenario 7 – High CPU Utilization after move to OSA3

Situation

A system with an even mix of transactional and streaming workloads had a hardware upgrade and was now running with an OSA3 adapter. The Linux CPU became excessively high for no clearly visible reason.

Trouble Shooting

Historical data was viewed to ensure that the spike in CPU activity did occur when the OSA3 adapter was activated. In viewing the bytes in/out and other workload data no glaring inconsistencies were seen.

Solution

When the change was made the original OSA2 values for BLKT were used (inter=0, inter_jumbo=0, total=0). Due to the difference in OSA2 and OSA3 behavior these numbers were changed (inter=5, inter_jumbo=15, total=250). CPU utilization returned to normal

OSA2 default value on OSA3 results in shortest latency and highest CPU utilization

Best to use MTU size of 1492 for OSA3

Supported in
SLES10SP3+kernel update
SLES 11
RHEL 5.5

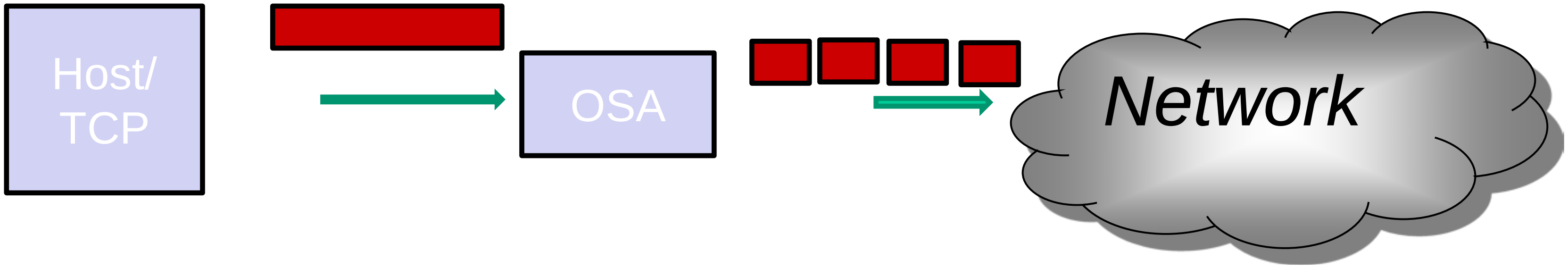
Red Hat: /etc/sysconfig/network-scripts/ifcfg-eth0 add OPTIONS="blkt/inter=5 blkt/inter_jumbo=15 blkt/total=250"

TCP Segmentation Offload (Large Send)

Segmentation consumes large amount of CPU

Allows most IPv4 TCP segmentation processing to be handled by OSA

Increases data transfer efficiency for IPv4 packets



Scenario 8 – 2 Tiered Database System

Situation

Client had a 2 tiered Database system and OSA 3 adapters. The front end database servers created many TCP/IP connections with high transactional volumes. The responses resulted in large TCP segments and the CPU utilization was unbearable.

Trouble Shooting

Look at the detailed TCP connections and transfer information. Use a packet trace tool. Is there a correlation between large segments sizes resulted in excessive CPU utilization. If so, go in an looking at the OSA adapter TCP Offload was not turned on.

Solution

TCP Segmentation was turned on for the OSA adapter.

On average anywhere from 25% to 45% CPU improvement was observed.

Use the large_send parameter

- no: no large send
- TSO: OSA adapter does segmentation
- EDDP: the qeth driver performs segmentation

TCPIP will still do segmentation for:

- LPAR-LPAR packets
- IPSec encapsulated packets
-
- When Multipath is in effect (unless all interfaces support segmentation offload)

Scenario 9 – Linux Hipersocket Performance Slow

Situation

A client had a very successful beta with Linux on system z. As they added additional workloads onto the Linux systems overall network performance declined. Hipersockets was used and the expectation was that performance should have been better.

Trouble Shooting

Using a Linux TCP/IP Monitor check the overall flow of information through both the IP and TCP layers. Verify that listeners are available for the applications. View alerts and determine if any would suggest the problem being seen. Check the buffer count. In this system the buffer count had never been raised and was still set at 16.

Solution

Tests by IBM have shown that using the default of 16 limits throughput as the number of parallel sessions increases with HiperSockets. The buffers were increased to 50 with acceptable results

As you increase the buffer additional memory will be used

SUSE SLES11: in /etc/udev/rules.d/51-qeth-0.0.f200.rules add ACTION=="add", SUBSYSTEM=="ccwgroup", KERNEL=="0.0.f200", ATTR{buffer_count}="128"

Cloud Computing…. It's a Journey

A Simple Idea

- User:
  - Builds a web application,
  - Using a standard platform
  - Using a standard database
  - Upload this application to a cloud provider
  - Only pays for what is used
  - Everything else is an implementation detail.
- Cloud provider automatically
  - Provisions the services
  - Scales the application and the database together

Clear Tenets

- Application Flexibility
  - Standardized
  - Increasing "click to run" services
  - Live in remote Internet data centers
  - Scalable to millions
- Procurement
  - Efficient
  - Rapid
  - Commoditized
  - "Pay by the sip"
- Security
  - Simplified
  - Streamlined

Multi-faceted Enablement

- Infrastructure
  - Consolidation
  - Global Information Grid
  - Capacity Services
  - Virtualization
  - Rapid Provisioning
  - Facility Analysis
- Software
  - Network-centric Services
  - Software-as-a-Service (Saas)
- Processes
  - ITIL
  - Security (Certification & Accreditation)

Murphy's Law

If anything can go wrong, it will

If anything just cannot go wrong it will

Left to themselves, things tend to go from bad to worse

If everything seems to be going well, you have obviously overlooked something

AES Sessions

Session	Title	Day	Time	Room
12152	IPv6 Basics	Tuesday February 5	1:30 PM	Golden Gate 4
12777	Network Problem Diagnosis with Packet Traces	Wednesday February 6	9:30 AM	Golden Gate 3
12778	Performance Factors in Cloud Computing	Wednesday February 6	11:00 AM	Golden Gate 4
12150	I'm Running IPv6 How Do I Access?	Wednesday February 6	3:00 PM	Golden Gate 4
12158	Managing an IPv6 Network	Thursday February 7	11:00 AM	Golden Gate 4
12149	Kick Start your IPv6 Skills using your home network	Friday February 8	8:00 AM	Golden Gate 4
12153	IPv6 Deep Dive	Friday February 8	9:30 AM	Golden Gate 4

Vielen Dank

ありがとうございました

Köszönettel

Obi

Спасибо

ขอบคุณ

شكراً

Bedankt

Gracias

شكرا

Ευχαριστώ

धन्यवाद

THANK YOU

Merci

Díky

Grazie

Danke

Hvala

תודה

Merci

ขอบคุณ

Teşekkürler

धन्यवाद

Hindi

Gracias

laurak@aesclever.com

www.aesclever.com

650-617-2400

:

감사합니다

நன்றி

Tamil

Obrigado