HP-UX Performance Assured by Capacity Management & refined through ITIL practice

Adam Grummitt – adam.grummitt@metron.co.uk
Director
Metron
Abstract: PA by CM

- ITIL, itSMF and ITSM processes
- Capacity Management objectives
- Performance Assurance targets
- Metrics and data sources
- Data model for repositories
- Dataflows for processes
- Processes and interfaces
- Modelling of servers
- Case study – sample reports
ITIL

- The IT Infrastructure Library - books & definitions
  - Service Support & Service delivery
  - Business, Infrastructure, Development, Service
- Good practice for managing IT
- Basis of BS15000, 7799 & ISO 17799 standards
- Developed by UK’s OGC in the 90’s
- Metron key contributor to initial Demonstrator
- itSMF
  - The IT Service Management Forum for ITIL users
  - Promotes exchange of info & experience
  - GB, NL, B, AUS, ZA, CDN, F, CH/A/D, USA
ITIL overview

Business Objectives

IT Strategy

Operational IT Processes
Service Support
Service Desk function
& Management of: Changes, Incidents, Problems, Releases, Configuration

Tactical IT Processes
Service Delivery
SLM
& Management of: Finance, Capacity, Availability, Continuity
ITIL Service Delivery Processes

SLM = Service Level Management
Service Catalogue

Availability M
Availability DB

Capacity M
Config-DB CMDB

IT Service Continuity M
ITSCM Plan

Security M
Security Plan

Operational Processes

IT Financial M
Financial System
IT Infrastructure Planning (ITIP)

Management

Operations

ITI Planning
- Performance Assurance
- Quality Assurance
- Performance Engineering

Results

Network (NW)
- NW Planning
- NW Management
- End-User Support

Targets

Service
- SLM
- Availability M
- Continuity M
- Finance M

Support

Development
ITIP - Performance Assurance Processes

Capacity Management
- Performance Management
  - Performance Monitoring
  - Performance Analysis
  - Resource Accounting
- Capacity Planning
  - Workload Characterisation
  - Workload Prediction
  - Hardware Planning
- Capacity Forecasting

Performance Engineering
- Performance Prediction
- Performance Testing

Service Planning
- SLA & SLM
- ITIP Audit
ITIP objectives

• Ensure the right level of ITI investment
• Identify and resolve bottlenecks
• Evaluate tuning strategies
• Improve and report/publish performance
• “Right-size” or “consolidate servers”
• Ensure accurate and timely procurements
• Ensure effective service level management
• Plan for workload growth, new apps / sites
• Avoid performance disasters
Data Sources

- Business volumes - planners, reports
- Workload volumes - users, logs, trails
- Service Levels - SLAs, users, ops, monitors, logs
- Resource Usage - Monitors, Accounting systems
- New Systems - Developers, users
- All metrics v key metrics - overhead v clarity
- Metrics vary: snapshot, gas meter, average, peak
- Cockpit dynamic v control panel management
- Invasive instrumentation v MIBs/APIs/utilities
Corporate Objectives
Underpinning IT applications

IT Procurement Policy
Hardware, software, development

Service Level Objectives
TP (response, vol) & Batch (thruput)

Inventory

Applications

Data Centre

Workload

Config’n
Topology, OS

Network
Node, link, protocol

Environment
Language, RDBMS

Key Tx
Vol, trend, pattern

Existing apps

Databases
Structure, map, items

Traffic for CP
Resource usage, arrivals

Traffic for PE
I/O per file, arrivals

New apps
## ITIL Objectives

### Effective and timely Performance Assurance
- Alerting of performance problems across all nodes
- Effective reporting of status of all target machines
- Effective advice for all target machines
- Model library for all pools of machines

### Formal and effective liaison with other teams
- Product & Design process provide CM input
- QA, Network Planning, SLA, DBA ditto

### Configurations matched to workloads
- Application views
## ITIL CM Processes and current status

<table>
<thead>
<tr>
<th>Process</th>
<th>1 2 3 4 5 6 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Management</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td><strong>Performance Measurement</strong></td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td><strong>Performance Monitoring</strong></td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td><strong>Performance Analysis</strong></td>
<td>1 2 3 4</td>
</tr>
<tr>
<td><strong>Performance Reporting</strong></td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td><strong>Performance Alerting</strong></td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td><strong>Capacity Planning</strong></td>
<td>1 2 3 4</td>
</tr>
<tr>
<td><strong>Workload Characterisation</strong></td>
<td>1 2 3 4</td>
</tr>
<tr>
<td><strong>Workload Prediction</strong></td>
<td>1 2 3 4</td>
</tr>
<tr>
<td><strong>Response Forecasting etc</strong></td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>
Why model?

- Responses are Non-linear
- Traffic related queuing
- Lists, cache, freeslots
- Constraints of O/S
- Constraints of RDBMS etc
- Feedback loops
- Non-intuitive
Model

“as is”

& “what if”

Baseline Model “as is”

Modified Model “what if”

CONFIG’N
cpu 1, 2, ...
Disk 1, 2, ...

WORKLOAD
Oracle
Test, System

Oracle
Test, System
User related e.g.: 1,000 orders per hour @ 1 CPU sec & 10 disk I/O per order or SLA secs response
Primary Results

System
AppOne
AppTwo
Test
Dev
MIS
Projections

Response Time

- mandatory
- desirable

Planning horizons

Now 1 2 3 4 5 6
**Definition**
Determine objectives & Identify constraints
Define project plan, gain commitment, create infrastructure

**Baseline**
Establish current workload, characterise relevant components
Identify resources; monitor usage; build & calibrate model

**Forecast**
Define future workload and resources
Modification analysis “what if” & Sensitivity analysis “So what”

**Action**
Establish effective reporting & exceptions handling
Report and make recommendations/procurements

**Track**
Monitor outcomes and report accordingly
Modify models and revise forecasts
Pyramid of Needs

- Disorder / Lack of control
- Acquisition of relevant metrics / Context related knowledge
- Monitoring & Basic Control
- Basic pre-emption of Problems
- Prediction
- Optimal usage of available resources
- Full Control
- Advice Planning Modelling
- Trending Analysis Reporting
- Data Capture Management Alerting

22
Pyramid of Tools

- Real time performance monitor
- Threshold Alarms
- Automatic Reporting
- Aggregate & Correlate & Trend
- Advice and alerts
- Models
- Meta-models
# Athene Functions & SPIR

<table>
<thead>
<tr>
<th>Feature</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capture and Collection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CustomDataBase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Explorer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Automatic Reporting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Advisor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Planner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Client-Server</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>SCOPE– PROTOTYPE –</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IMPLEMENT – REVIEW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Automatic reports and advice

### Status Report

**Bulletin Name:** Head Office servers - Daily Performance

**Installation:** Wshop.3

**Analysis Period:** 16/05/2001 to 16/05/2001

**Date and Time of Report:** 31/10/2001 08:54

#### Status Report

<table>
<thead>
<tr>
<th>System</th>
<th>Oracle Database(s)</th>
<th>CPU</th>
<th>Memory</th>
<th>I/O</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>daffodil</td>
<td>N/A</td>
<td>🌷</td>
<td>🌷</td>
<td>🌷</td>
<td>🌷</td>
</tr>
<tr>
<td>rose</td>
<td>N/A</td>
<td>🌺</td>
<td>🌺</td>
<td>🌺</td>
<td>🌺</td>
</tr>
<tr>
<td>snowdrop</td>
<td>N/A</td>
<td>🌹</td>
<td>🌹</td>
<td>🌹</td>
<td>🌹</td>
</tr>
<tr>
<td>tulip</td>
<td>N/A</td>
<td>🌸</td>
<td>🌸</td>
<td>🌸</td>
<td>🌸</td>
</tr>
<tr>
<td>crocus</td>
<td>N/A</td>
<td>🌺</td>
<td>🌺</td>
<td>🌺</td>
<td>🌺</td>
</tr>
</tbody>
</table>

#### Interpretation of UNIX Performance Report

**Average value compared with thresholds**

The total CPU utilisation, averaged over the analysis period, is high, because it is above the lower (warning) and upper (alarm) threshold of 85.9%. There is some, but not enough, slack in the workload to handle peaks in the workload without encountering resource contention.

**Proportion of Time compared with thresholds**

The amount of time when the total CPU utilisation was above 90% was 59.4% of the total time. The amount of time when the total CPU utilisation was above 80% was 6.3% of the total time.
Multi-tier models
“What-if” scenarios
That’s what it’s all about

- Right kit in the right place at the right time
- Predict when it will all fall apart
- Take action to avoid that in time
- Consolidate servers effectively
- Don’t waste money on redundant kit
- Or on un-necessary interim upgrades
HP-UX Performance Assured by Capacity Management & refined through ITIL practice

Adam Grummitt – adam.grummitt@metron.co.uk
Director
Metron

Any Questions???