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

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

Director
Metron

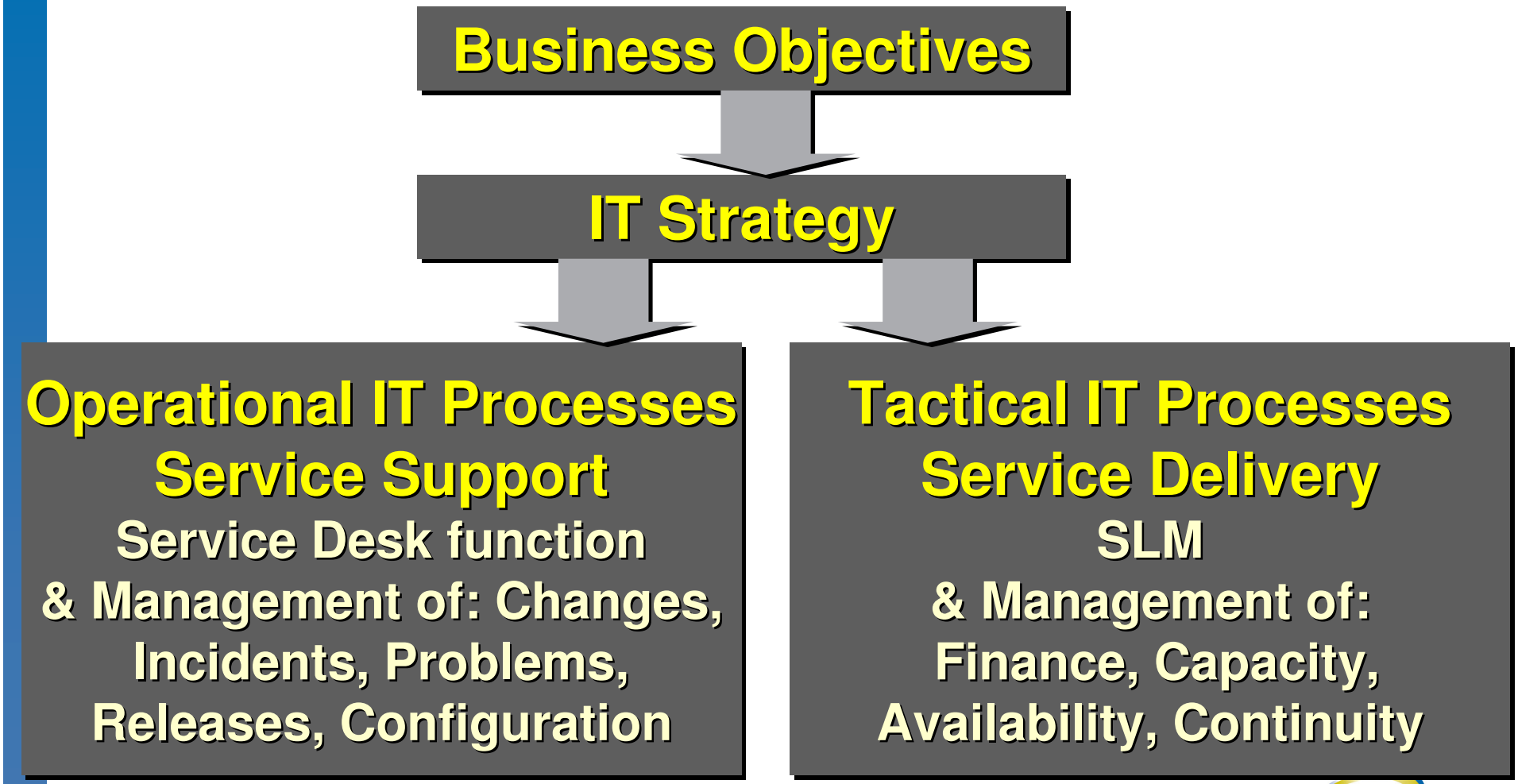
Abstract: PA by CM

- 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

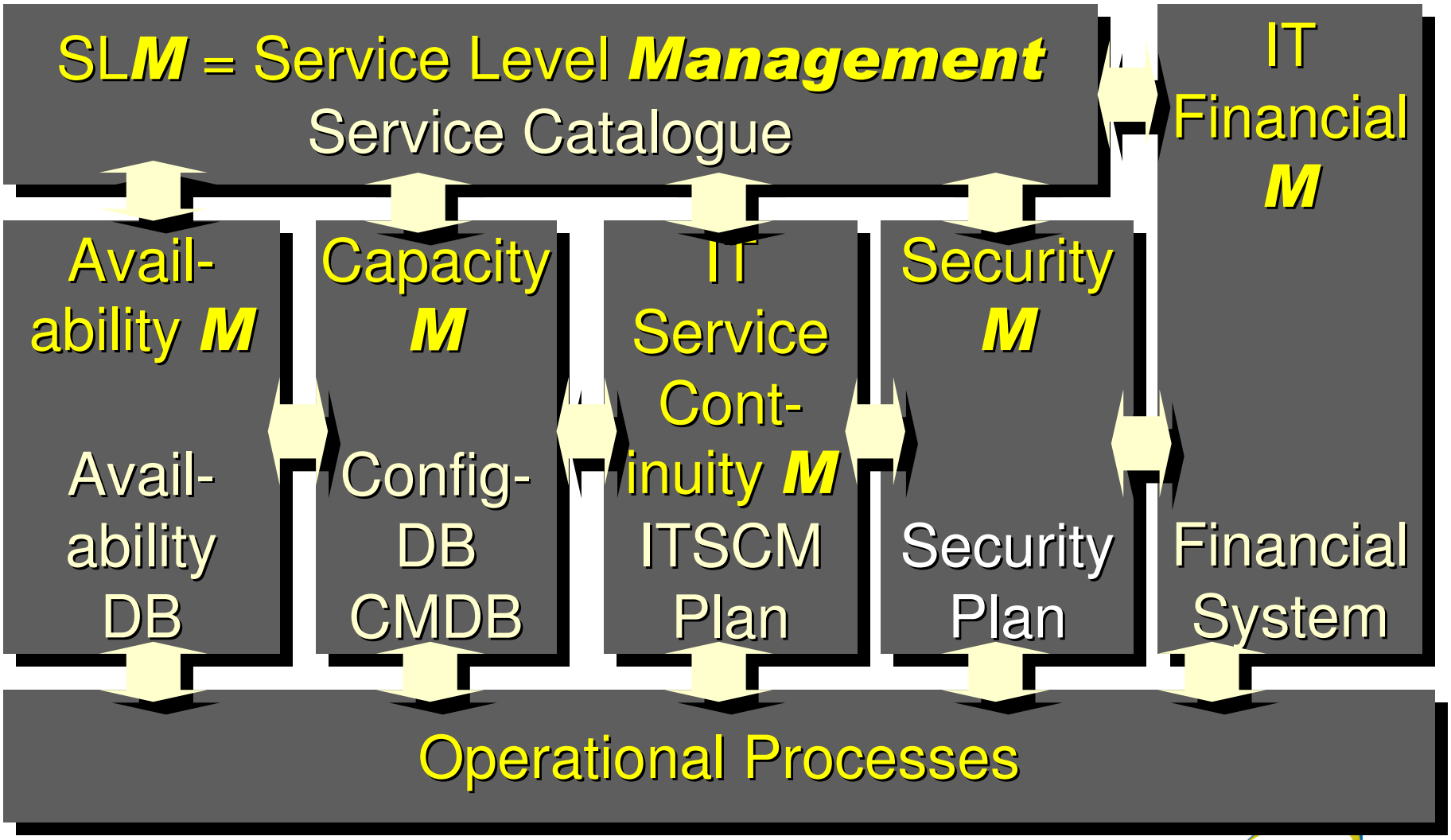
ITIL

- 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

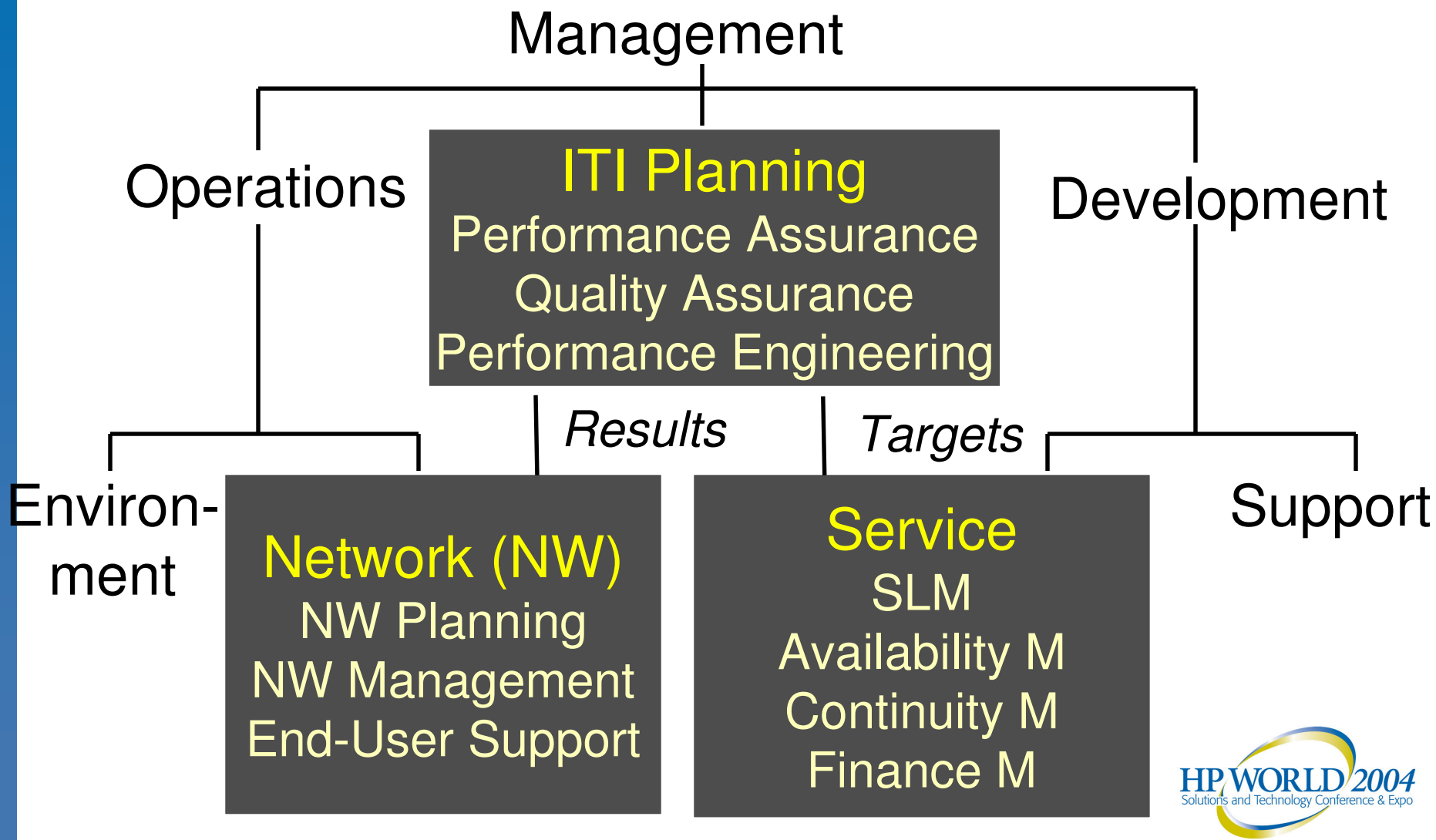
ITIL overview



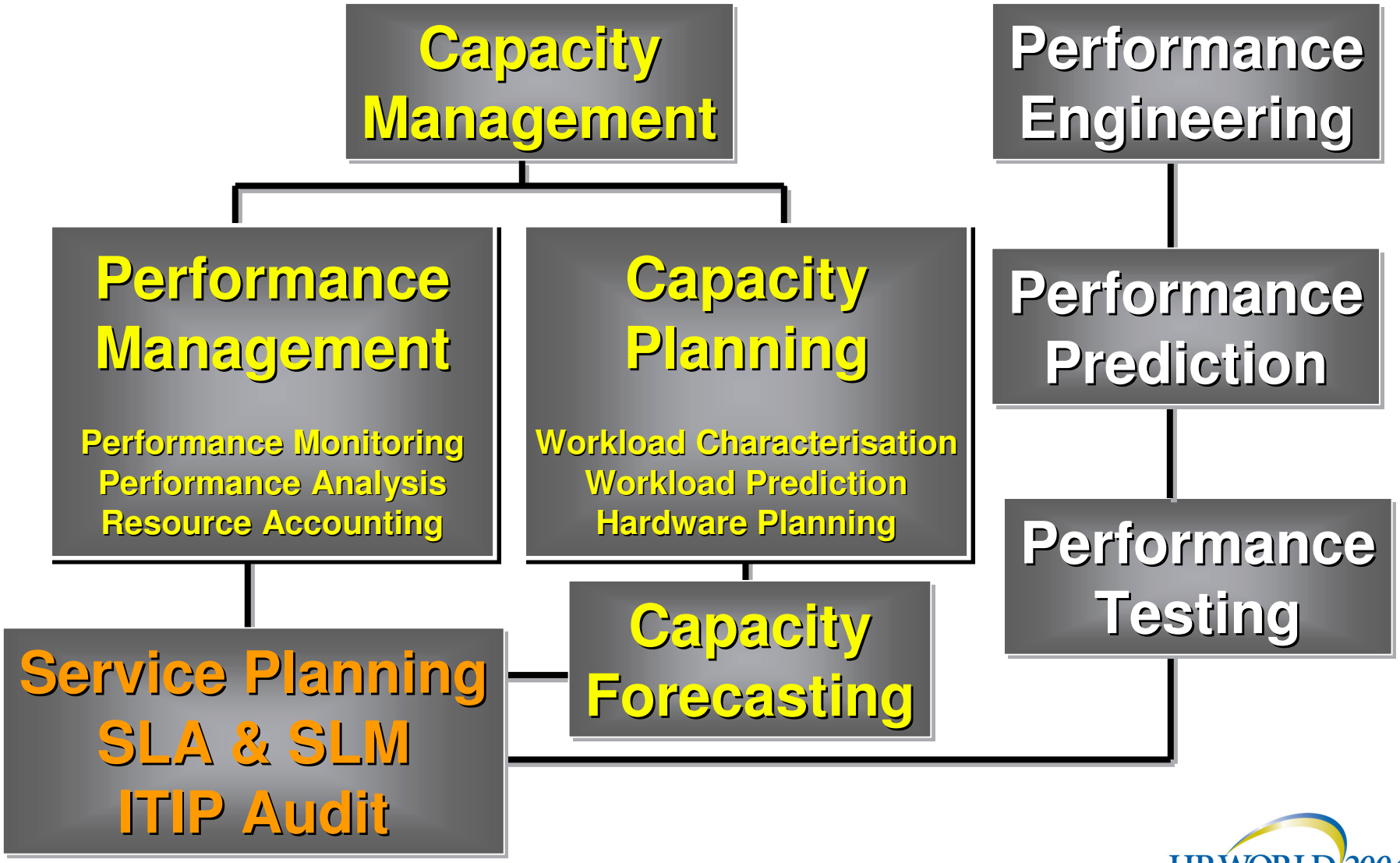
ITIL Service Delivery Processes



IT Infrastructure Planning (ITIP)



ITIP - Performance Assurance Processes

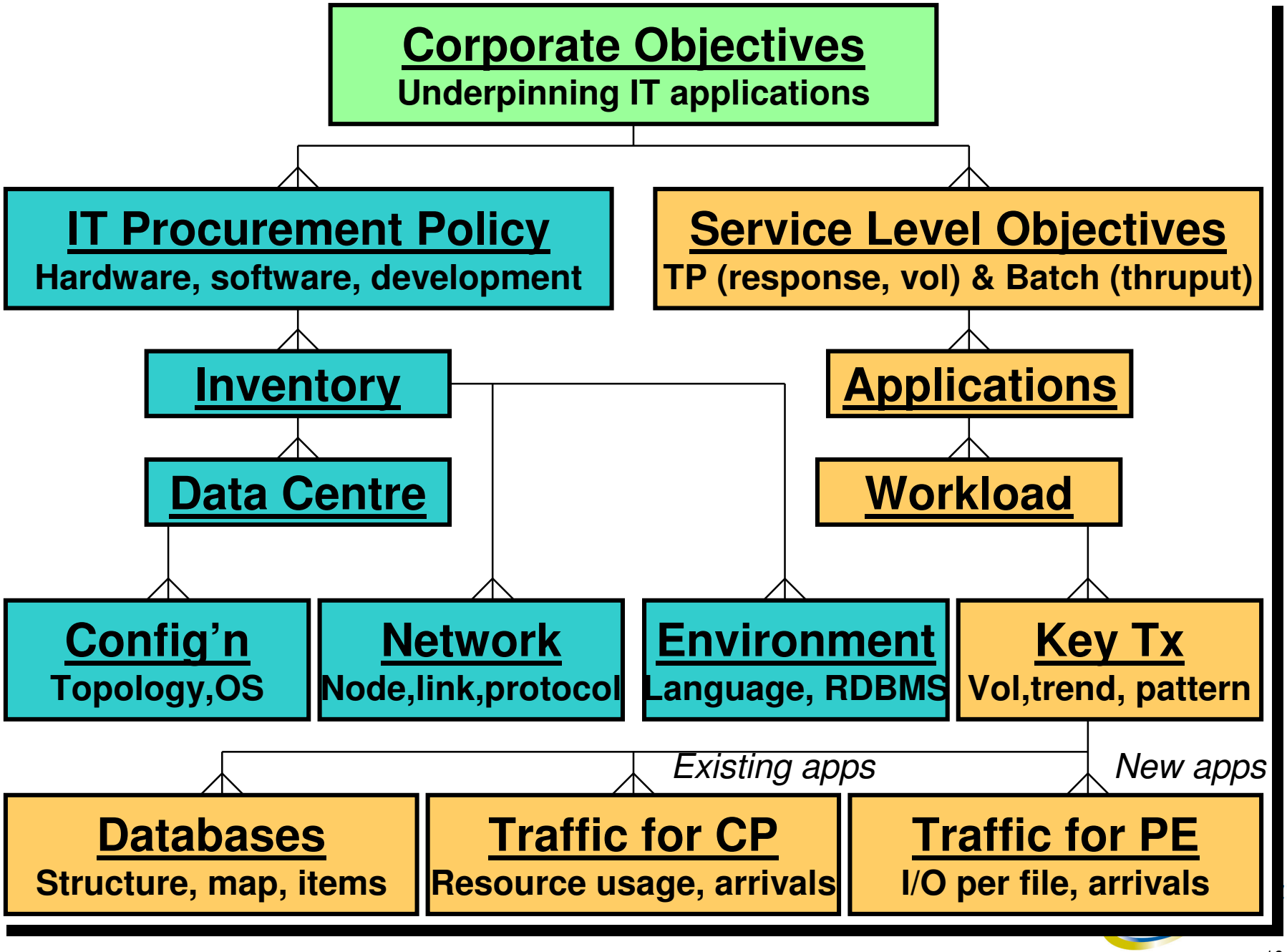


ITIP objectives

- 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
New apps
Databases
Structure, map, items
Traffic for CP
Resource usage, arrivals
Traffic for PE
I/O per file, arrivals

ITIL Objectives

Effective and timely Performance Assurance

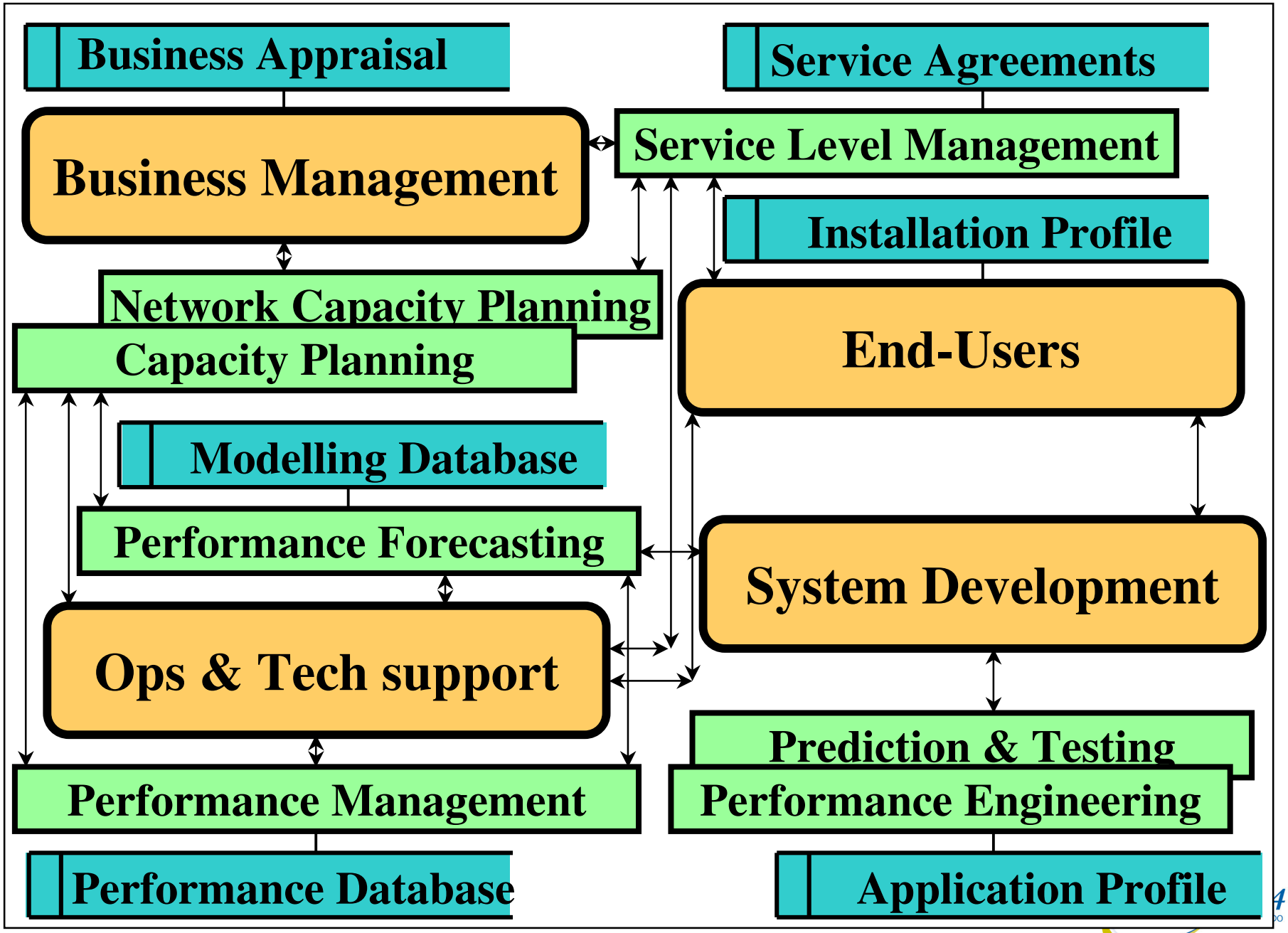
- Alerting of performance problems across all nodes
- Effective reporting of status of all target machines
- Effective advice for all target machines
- Model library for all pools of machines

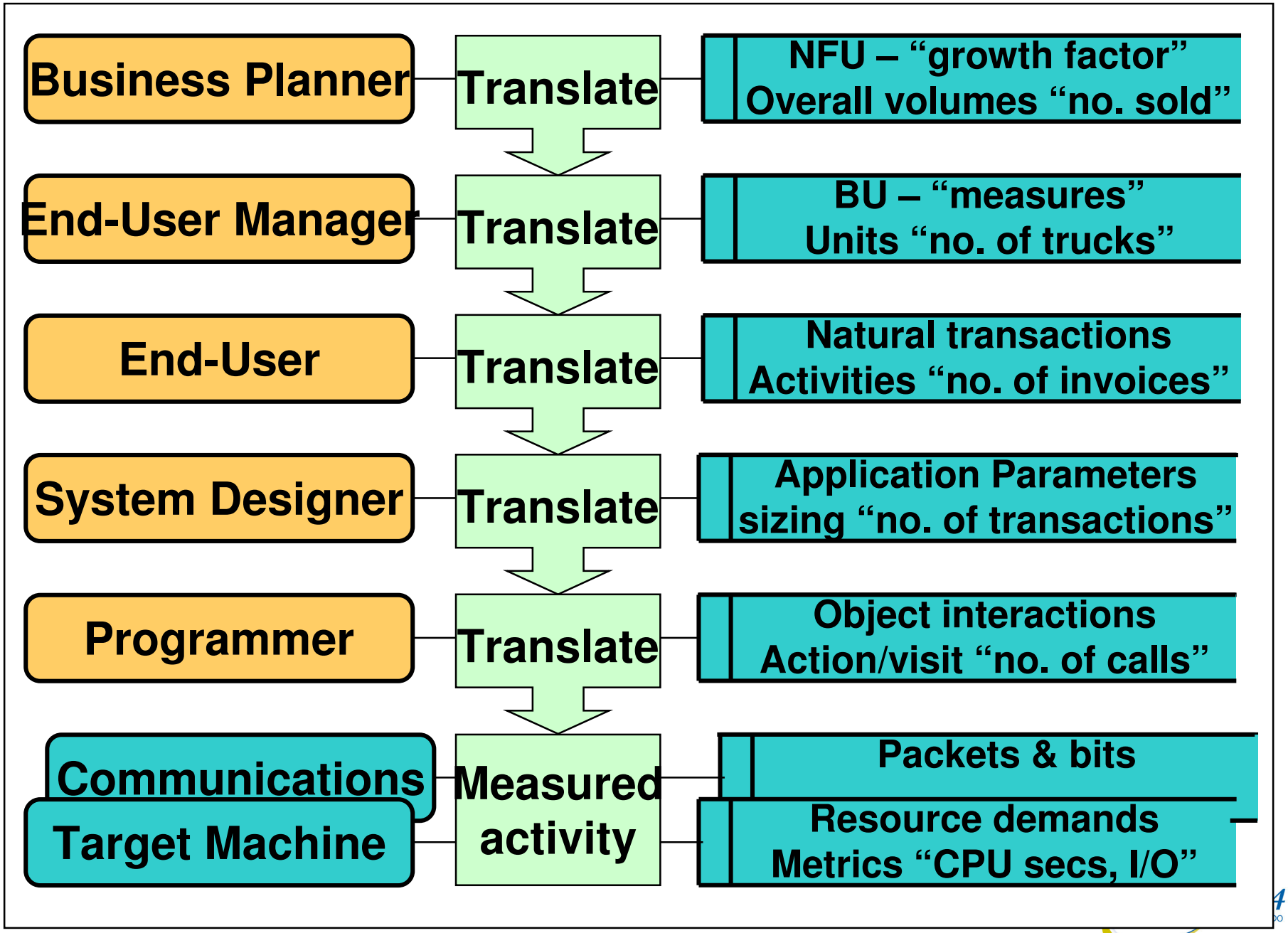
Formal and effective liaison with other teams

- Product & Design process provide CM input
- QA, Network Planning, SLA, DBA ditto

Configurations matched to workloads

- Application views


Business Appraisal
Service Agreements
Business Management
Service Level Management
Installation Profile
Network Capacity Planning
Capacity Planning
End-Users
Modelling Database
Performance Forecasting
System Development
Ops & Tech support
Prediction & Testing
Performance Management
Performance Engineering
Performance Database
Application Profile

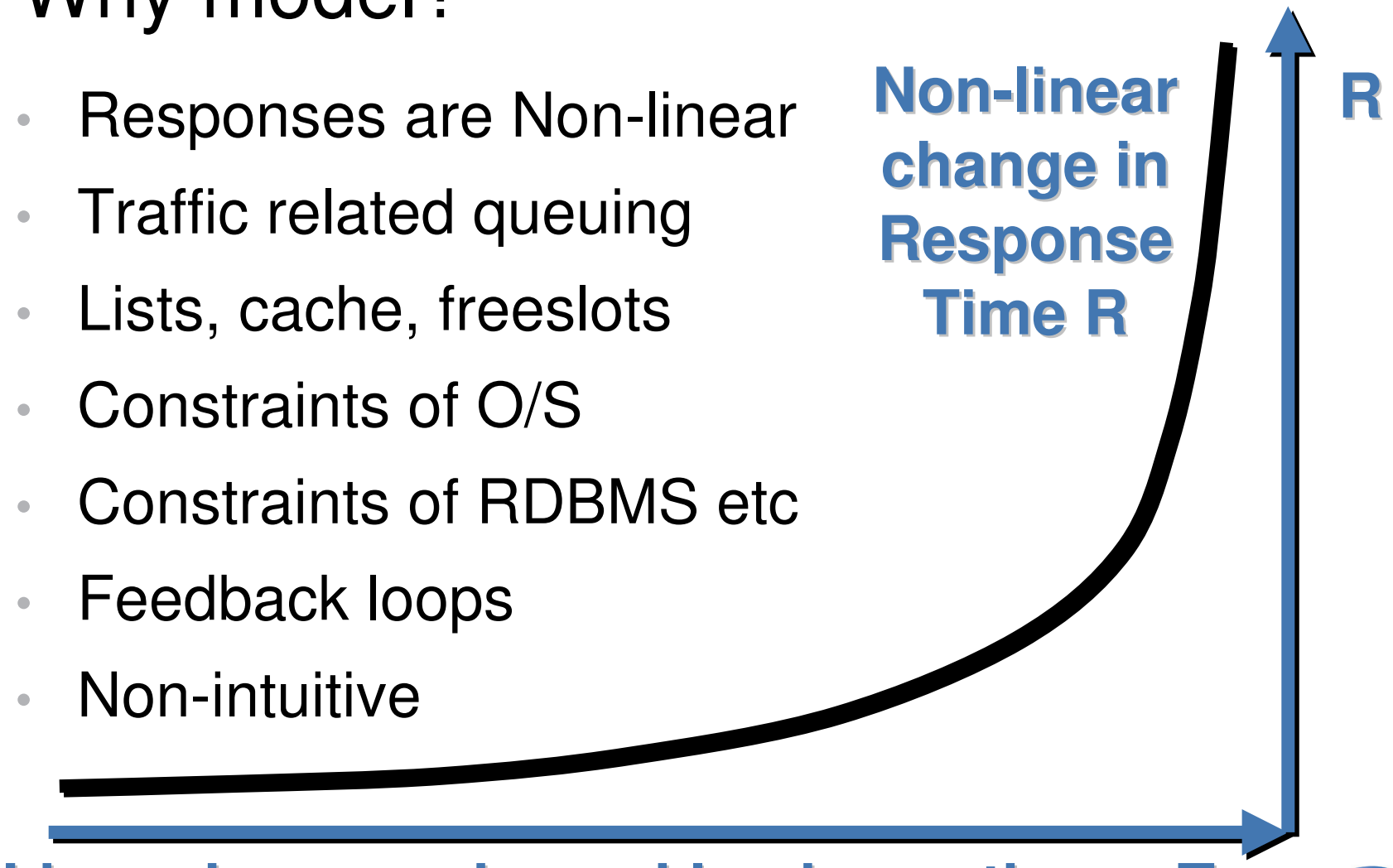

Business Planner
Translate
NFU – "growth factor"
Overall volumes "no. sold"
End-User Manager
Translate
BU – "measures"
Units "no. of trucks"
End-User
Translate
Natural transactions
Activities "no. of invoices"
System Designer
Translate
Application Parameters
sizing "no. of transactions"
Programmer
Translate
Object interactions
Action/visit "no. of calls"
Communications
Target Machine
Measured activity
Packets & bits
Resource demands
Metrics "CPU secs, I/O"

ITIL CM Processes and current status

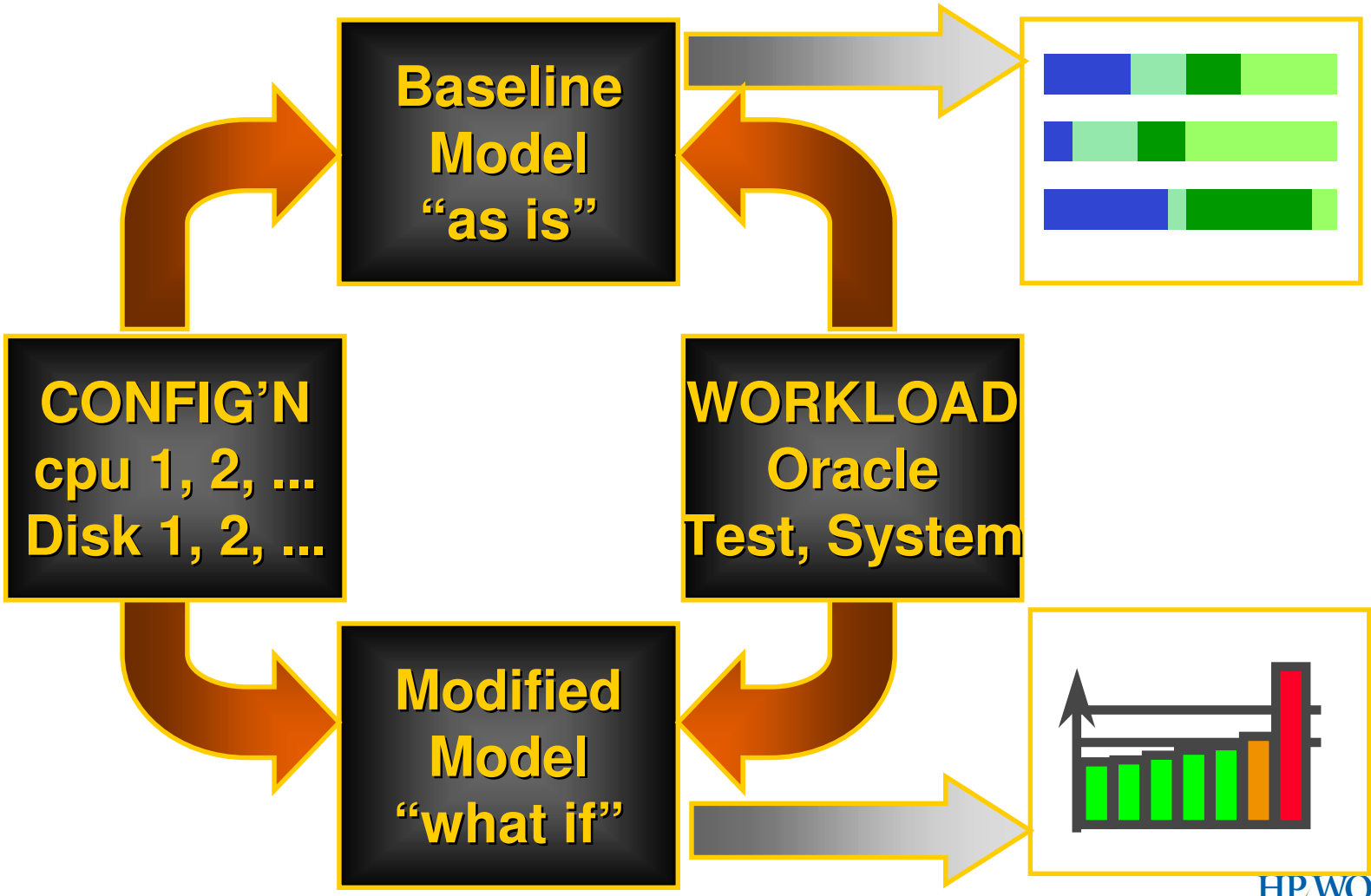
Performance Management	❶❷❸❹❺⑥⑦
Performance Measurement	❶❷❸④⑤⑥⑦
Performance Monitoring	❶❷❸④⑤⑥⑦
Performance Analysis	❶②③④
Performance Reporting	❶❷③④⑤⑥⑦
Performance Alerting	①②③④⑤⑥⑦
Capacity Planning	①②③④
Workload Characterisation	①②③④
Workload Prediction	①②③④
Response Forecasting etc	

Why model?

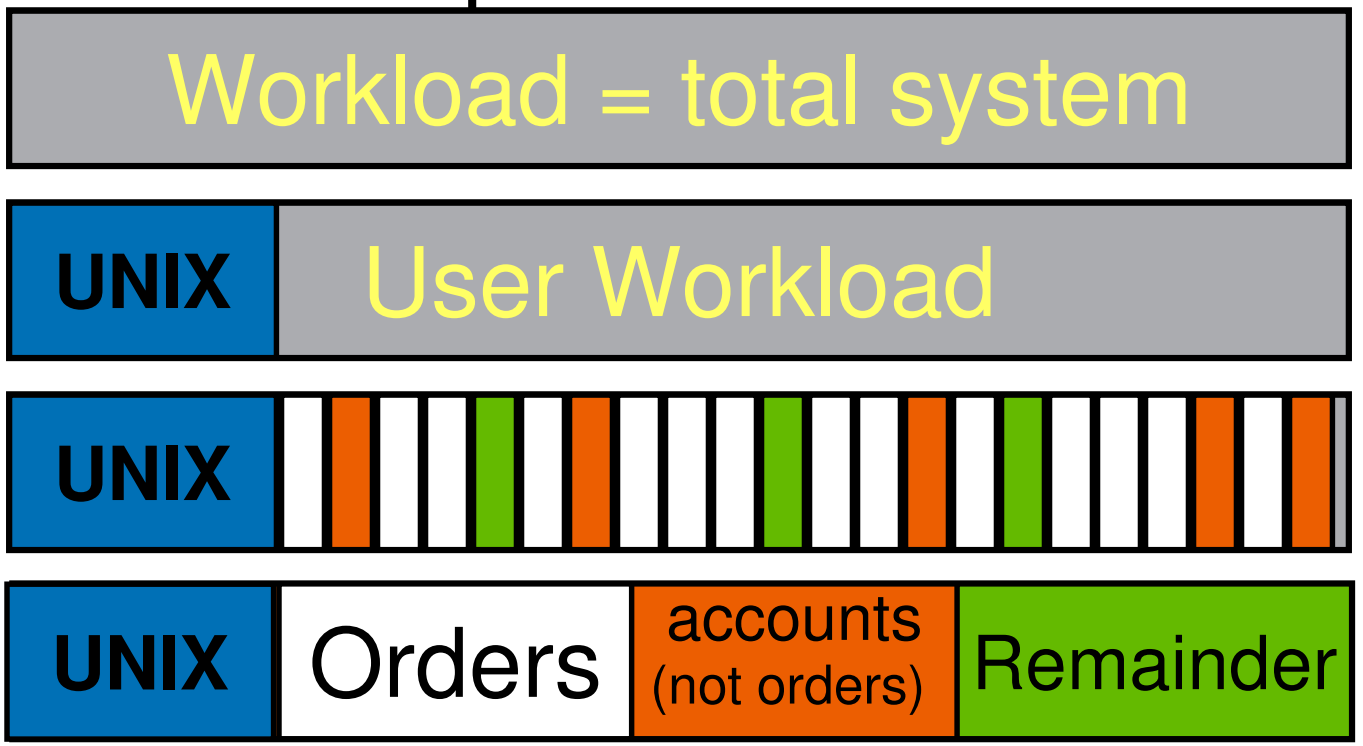
- Responses are Non-linear
- Traffic related queuing
- Lists, cache, freeslots
- Constraints of O/S
- Constraints of RDBMS etc
- Feedback loops
- Non-intuitive



Model “as is” & “what if”

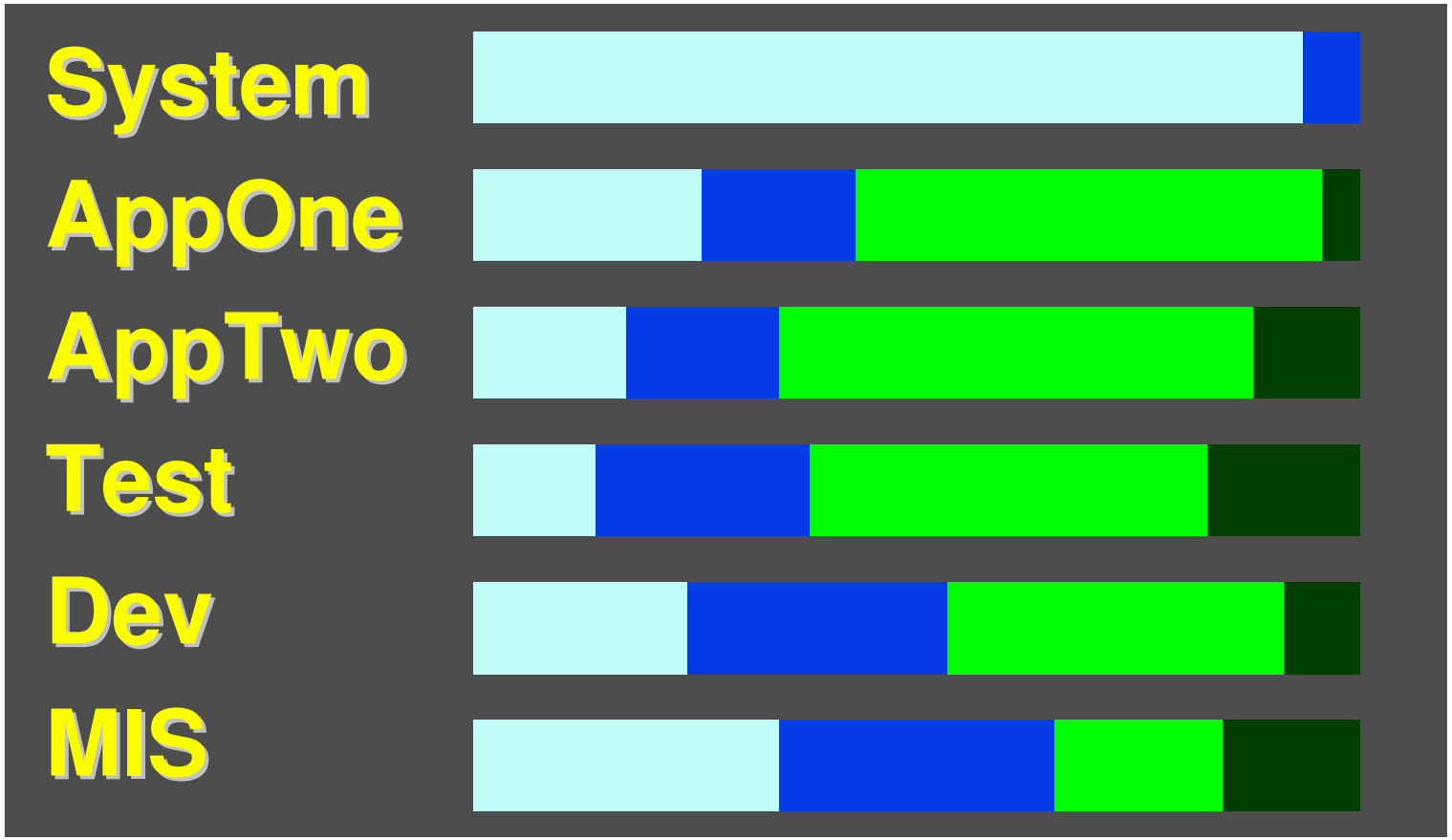


Workload Components

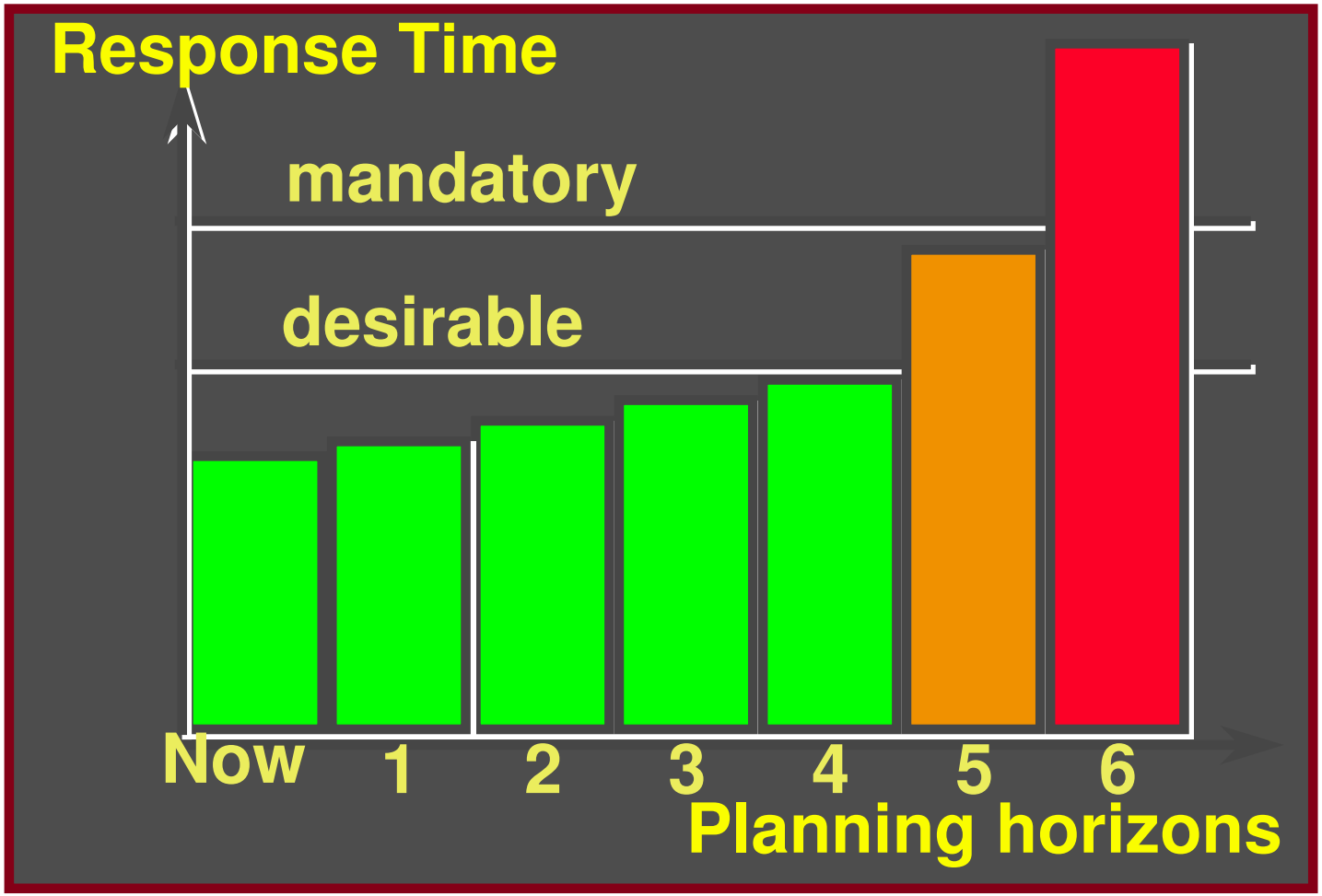


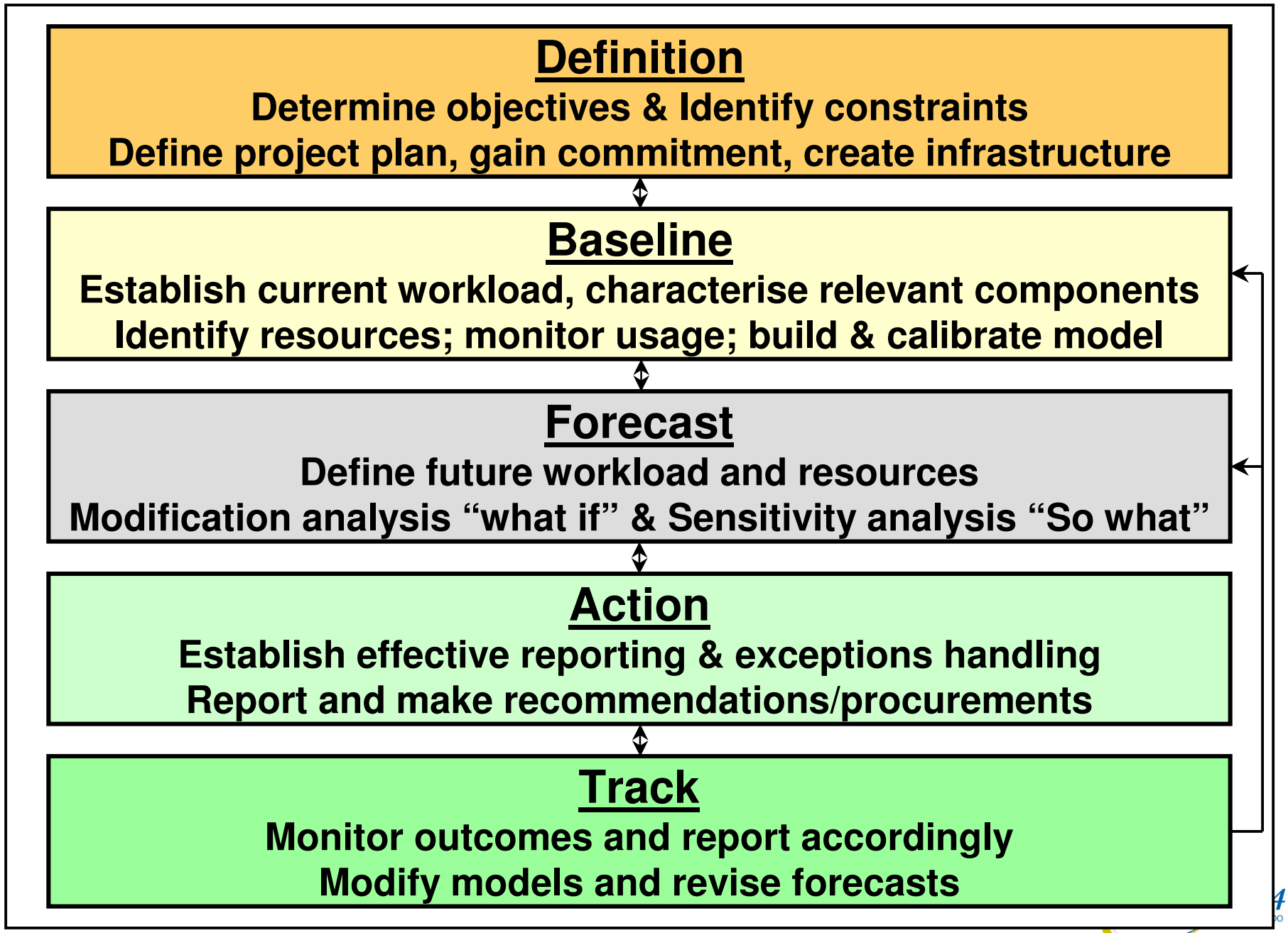
User related e.g.: 1,000 orders per hour
@ 1 CPU sec & 10 disk I/O per order
or SLA secs response

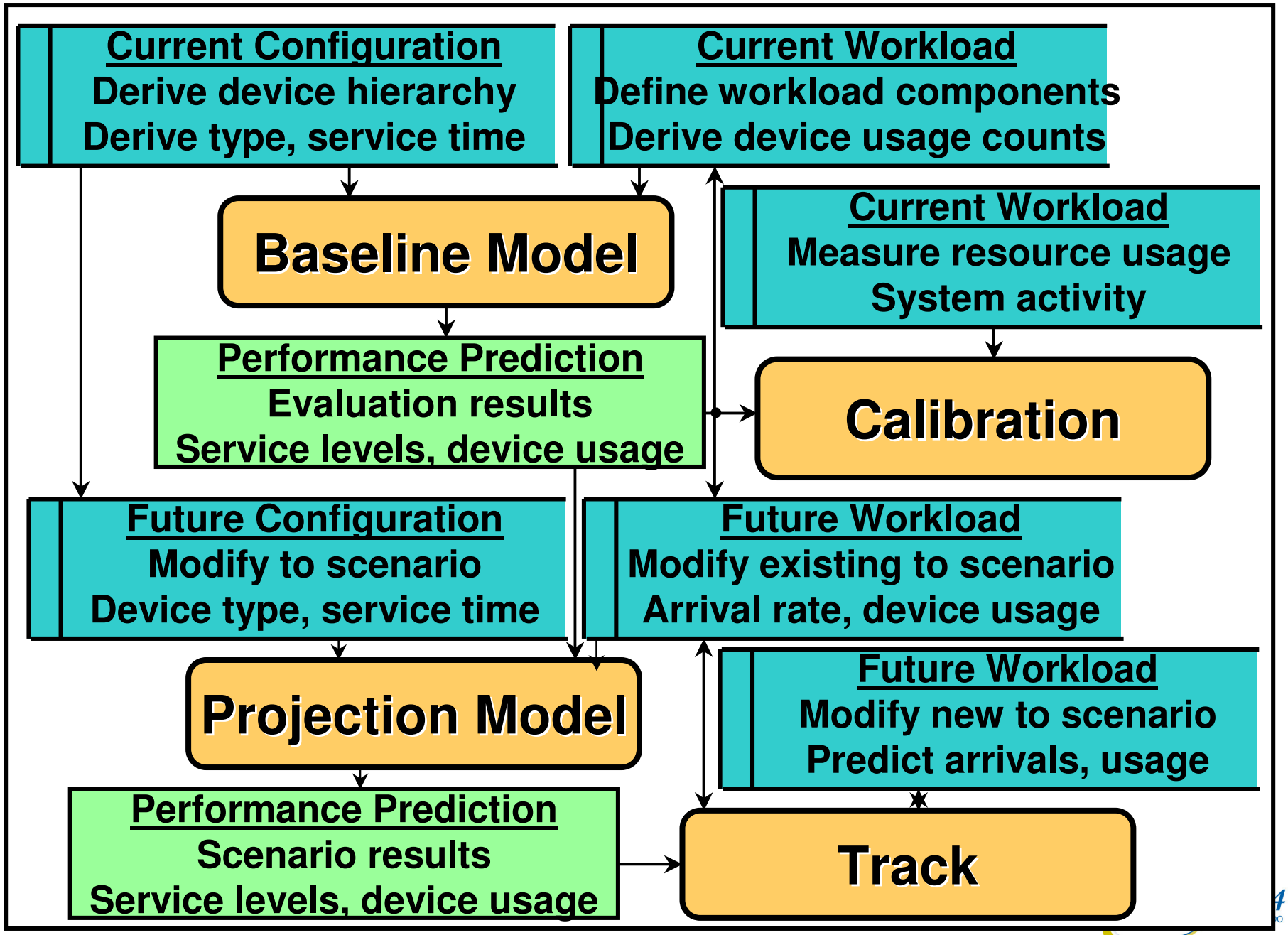
Primary Results



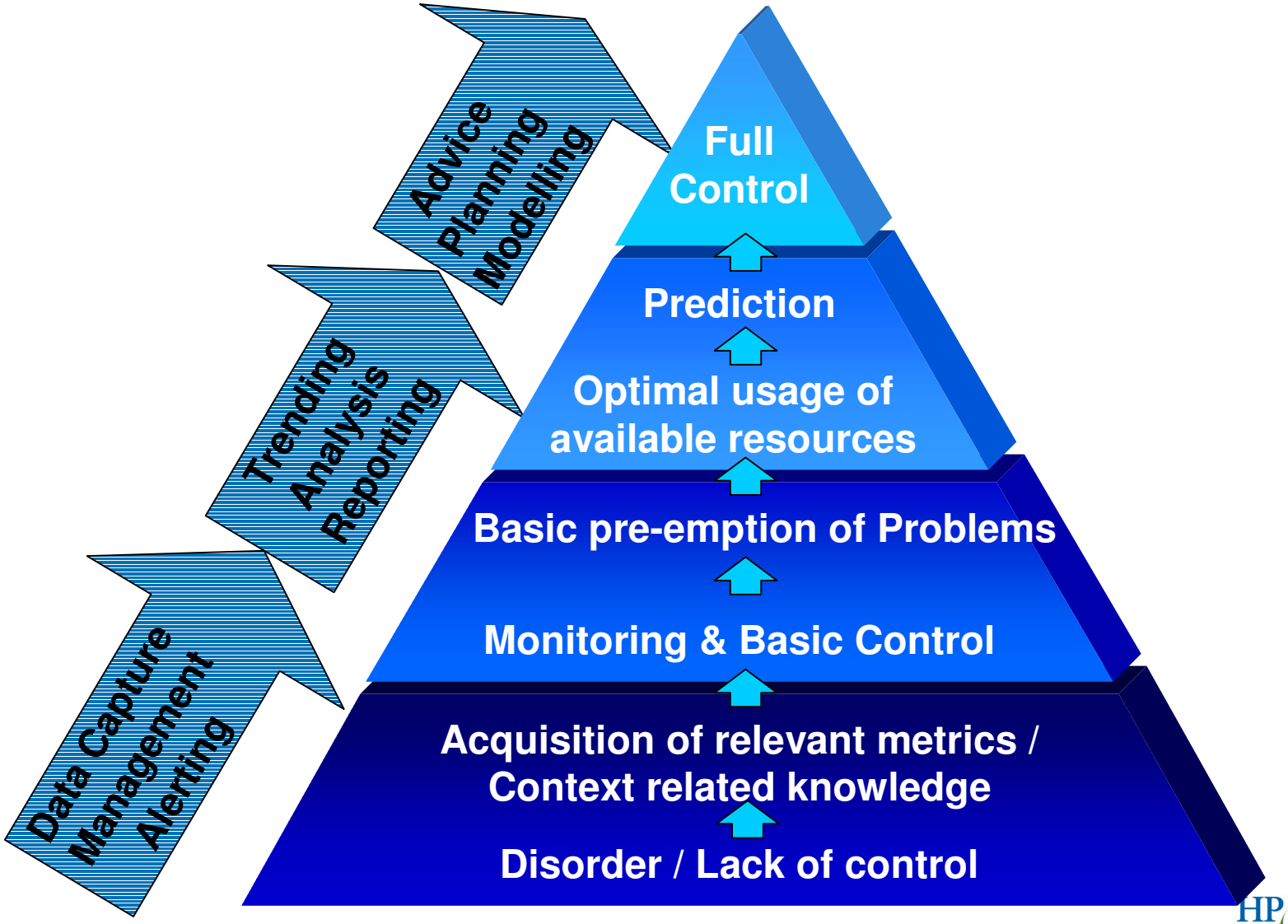
Projections



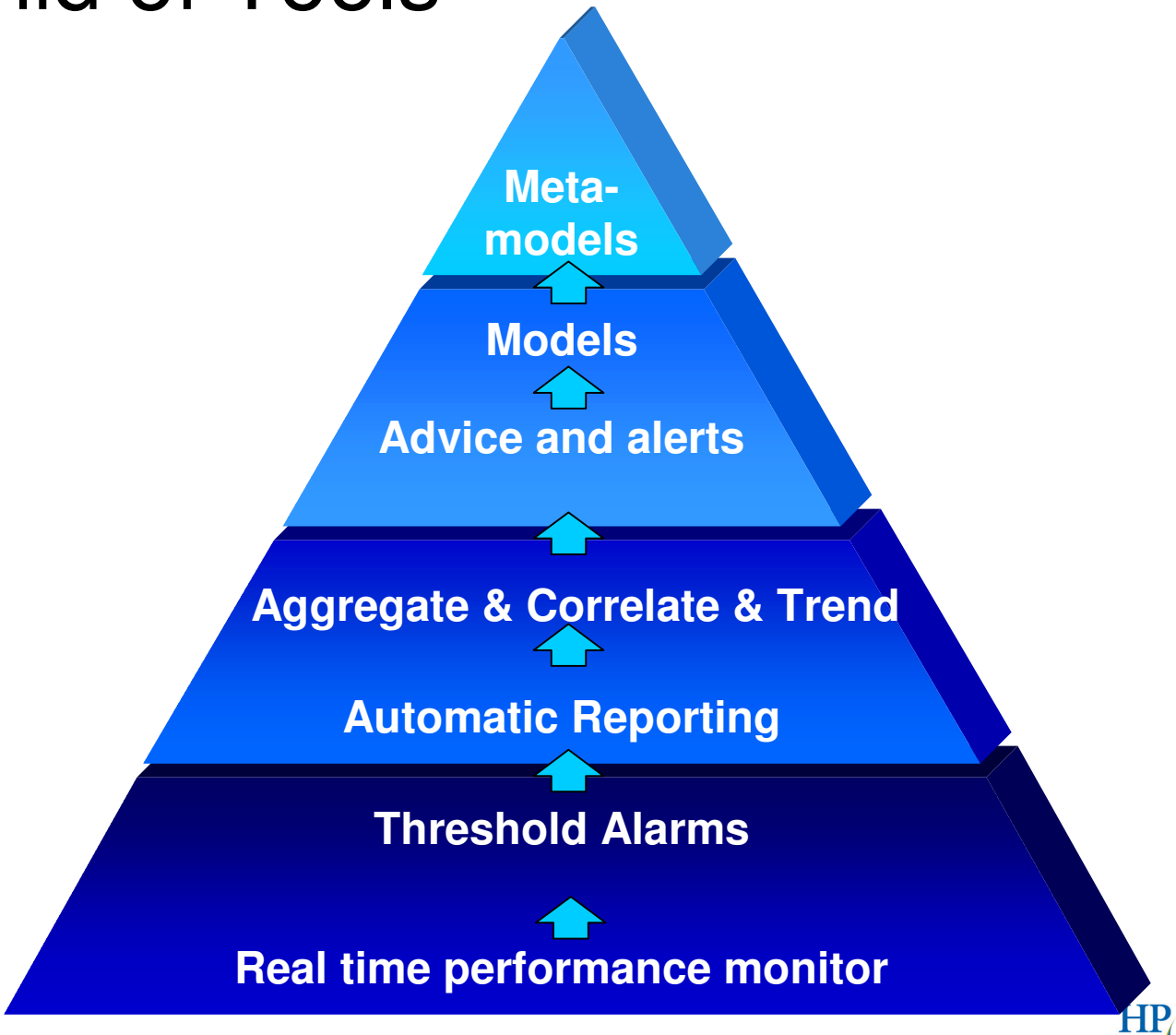

Definition
Determine objectives & Identify constraints
Define project plan, gain commitment, create infrastructure
Baseline
Establish current workload, characterise relevant components
Identify resources; monitor usage; build & calibrate model
Forecast
Define future workload and resources
Modification 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


Current Configuration
Derive device hierarchy
Derive type, service time
Current Workload
Define workload components
Derive device usage counts
Baseline Model
Current Workload
Measure resource usage
System activity
Performance Prediction
Evaluation results
Service levels, device usage
Calibration
Future Configuration
Modify to scenario
Device type, service time
Future Workload
Modify existing to scenario
Arrival rate, device usage
Projection Model
Future Workload
Modify new to scenario
Predict arrivals, usage
Performance Prediction
Scenario results
Service levels, device usage
Track

Pyramid of Needs



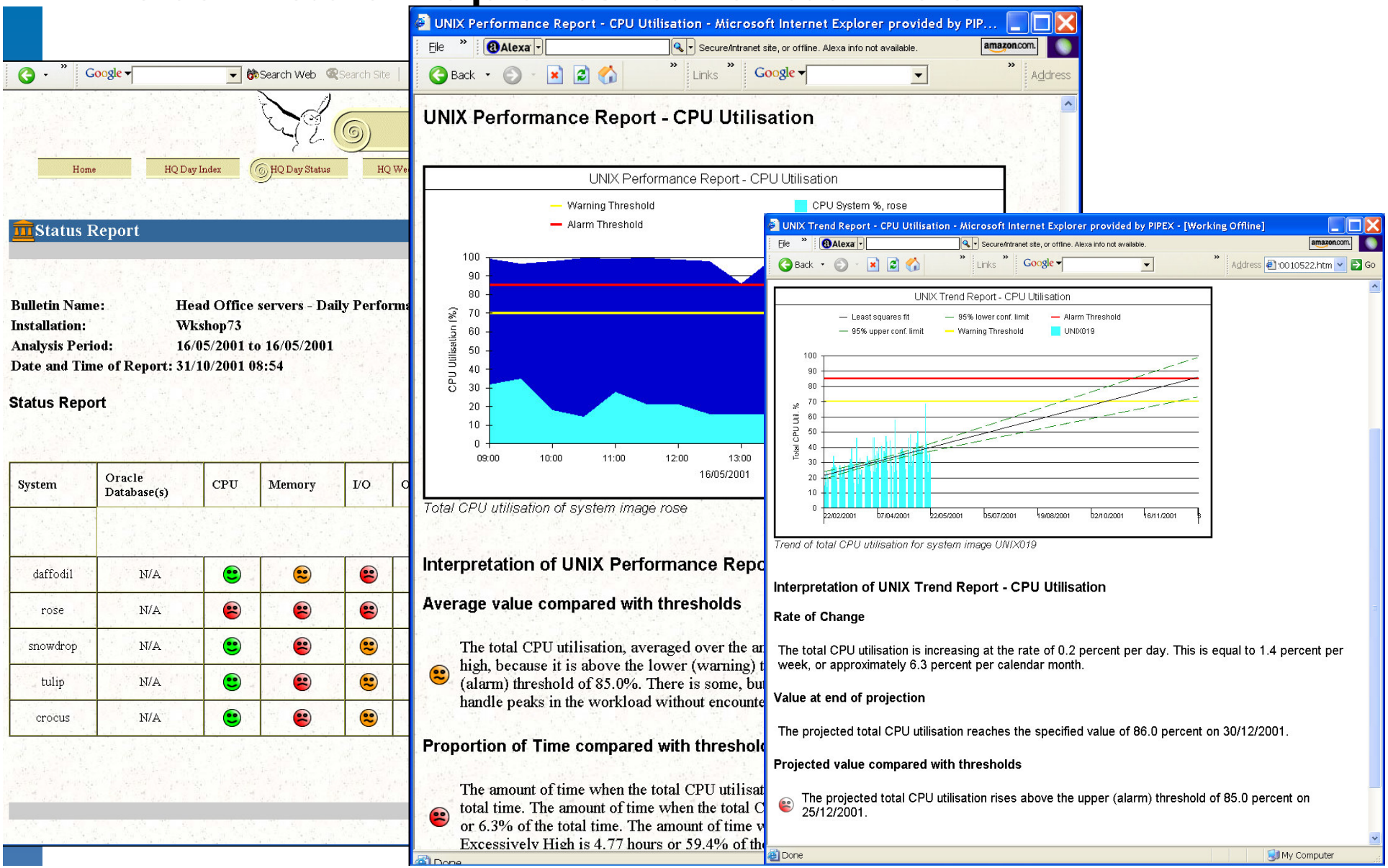
Pyramid of Tools



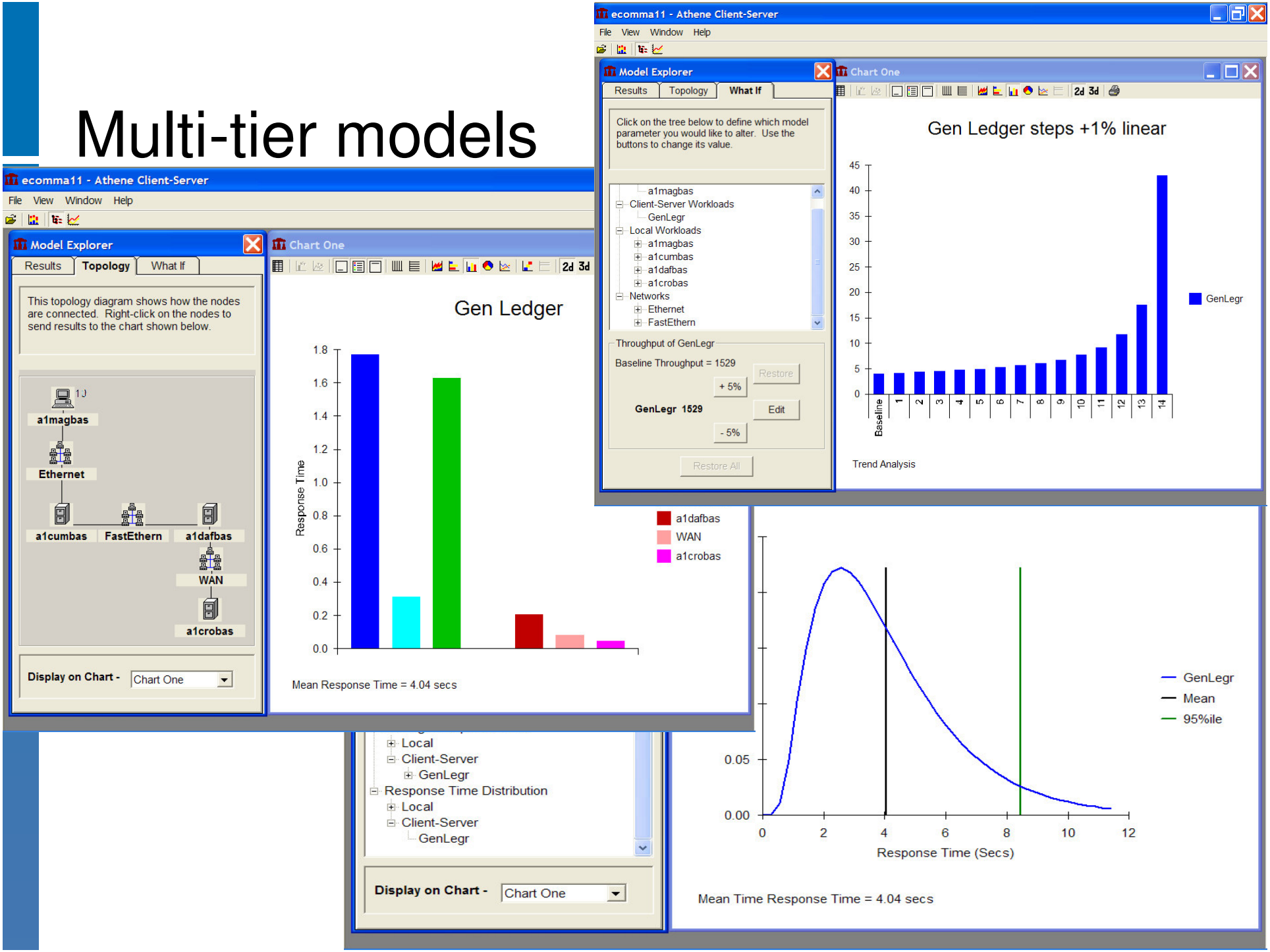
Athene Functions & SPIR

Capture and Collection	❶❷❸④⑤⑥⑦
Data Management	①②
CustomDataBase	❶②③④⑤
Analysis	❶②③④
Explorer	①②③
Automatic Reporting	❶❷③④⑤⑥⑦
Advisor	❶②③④⑤⑥
Planner	①②③④
Client-Server	①
SCOPE– PROTOTYPE –	IMPLEMENT - REVIEW

Automatic reports and advice



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 ???

HP WORLD 2004

Solutions and Technology Conference & Expo

Co-produced by:

interex

shared knowledge • shared power

encompass

AN HP USER GROUP