



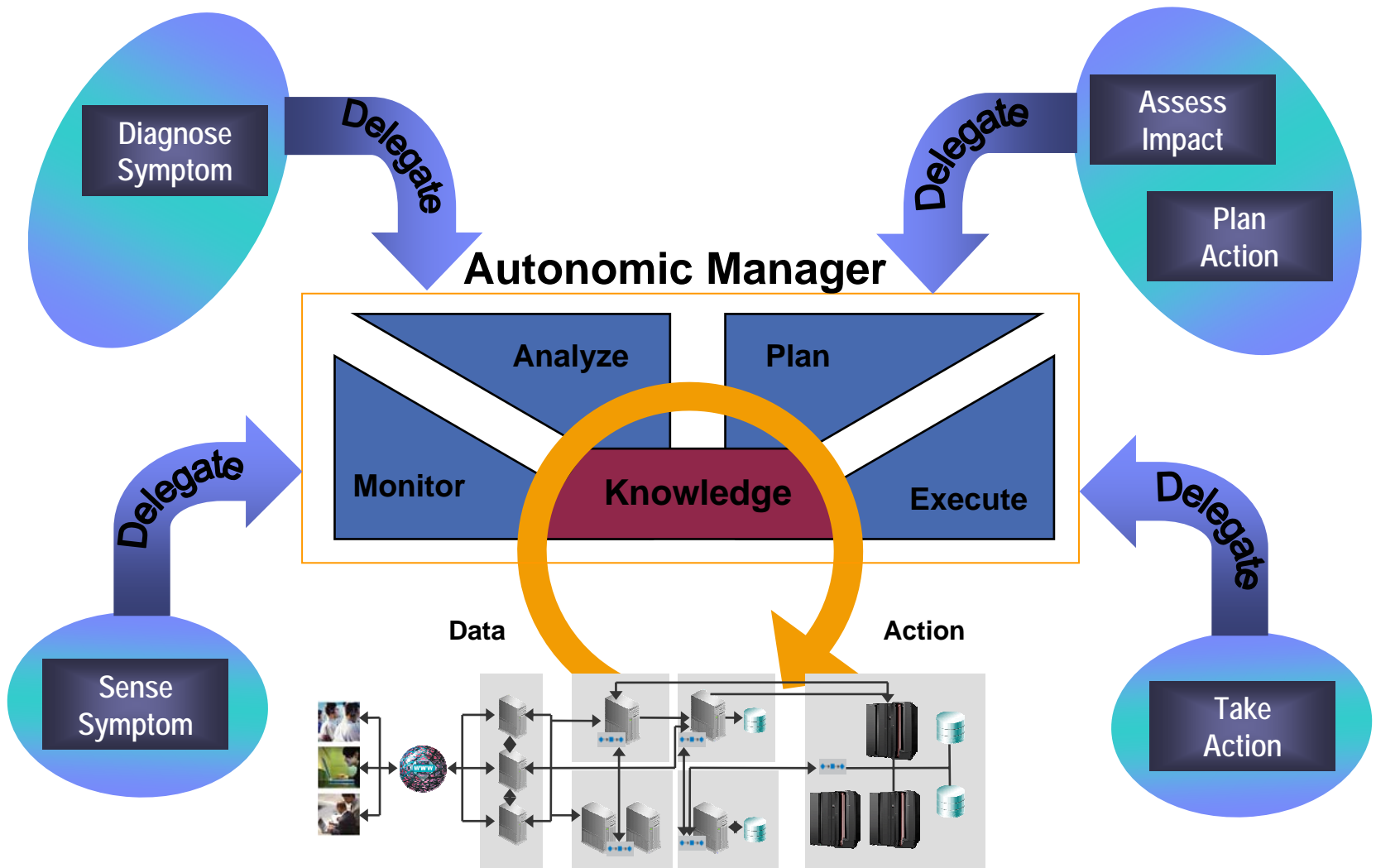
IBM T.J. Watson Research Center

Towards Autonomic Networking Middleware

Alexander Keller, alexk@us.ibm.com



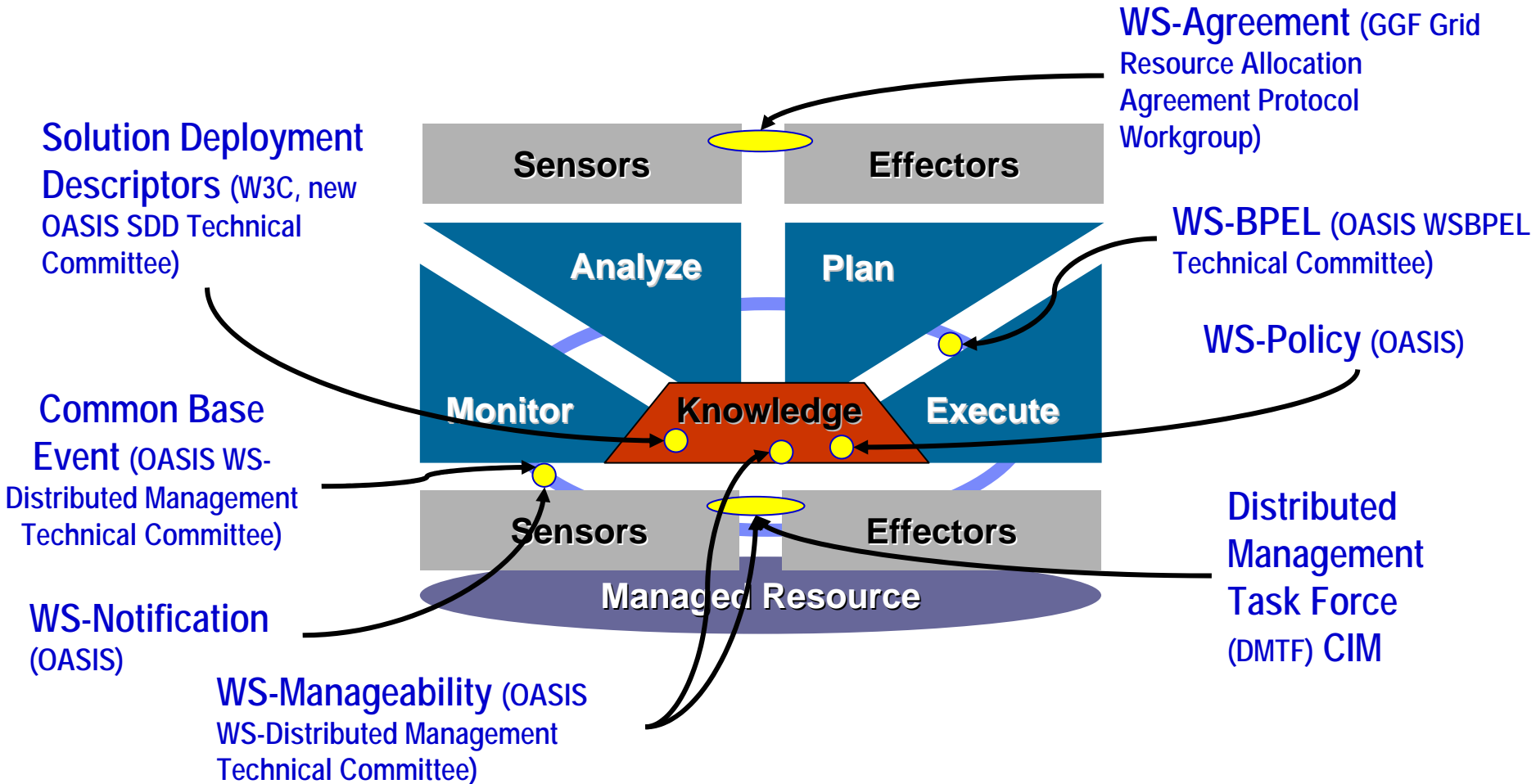
The Autonomic Control Loop (1)



The Autonomic Control Loop (2)

- **Concept exists for a long time**
 - Borrowed from control theory, manufacturing
- **Primary focus on management logic**
 - Addresses what an Autonomic Manager does, and how
 - Standardization: Knowledge representation
 - Differentiation through M-A-P-E component logic
- **Secondary focus on manageability**
 - How a resource exposes its sensors and effectors
 - Objective: create open standards

Standards needed for Autonomic Computing



What does “autonomic” actually mean?

- **“Smart” management system executes the MAPE-K loop**
 - Interacts with “dumb” managed resources
 - If sub-optimal, configures them so that global optimum is achieved
 - E.g., OSI Network Management: Manager/Agent paradigm
 - How systems management is done today
- **Every individual system runs its own MAPE-K loop**
 - and, somehow, optimal behavior of the overall system emerges
 - E.g., Routing algorithms, as implemented in OSPF, BGP
 - Smart resources, no management system needed

These are 2 fundamentally different viewpoints

...and these 2 Viewpoints have been articulated long ago

- ***“What we see are merely reflections of perfect archetypes” (Plato)***
 - Basically, the world is in equilibrium
 - If a system gets pushed outside of the equilibrium, push it back again
- ***“The world is in a constant state of flux” (Heraclitus)***
 - The world is in a process of flow and change
 - Systems are interacting in a variety of combinations
- **Variations of these viewpoints reflected in middleware paradigms**
 - Synchronous vs. asynchronous communication (RPC vs. messaging)
 - Manager/Agent model vs. peer-to-peer interactions
 - Workflows vs. business objects

Solving the right problems

- **Scientific community's main focus areas:**
 - Optimization techniques
 - (Mathematical) system modeling and design
 - Rigorous evaluation of results through experiments
 - Achieving automation
- **Major headaches in practical IT service management:**
 - Availability and problem determination
 - Dealing with human error
 - More than 50% of service outages due to misconfiguration
 - Repeatable procedures for changing systems
 - More than 50% of service outages happen during maintenance windows
 - Labor Costs
 - “Good enough” solutions
 - Reluctance to immediately implement automation, based on past negative experience

Plenty of opportunities for the scientific community!

Data Center Activities
Resolve Problems
Configure/Install
Collect change request and gain approval
Support SW (e.g. OS files, processes)
Perform customer-driven changes
Define Requirements / Plan
Upgrade
Apply patches
Analyze and maintain user requirements
Manage system availability, capacity
Test
Maintain Policy, processes, procedures
Support Hardware (e.g. server, disks)
Perform internal-driven changes
Audit Compliance
Perform User Mgmt and housekeeping
Maintain Configuration info
Perform health check
Handle special projects/requests

Priority

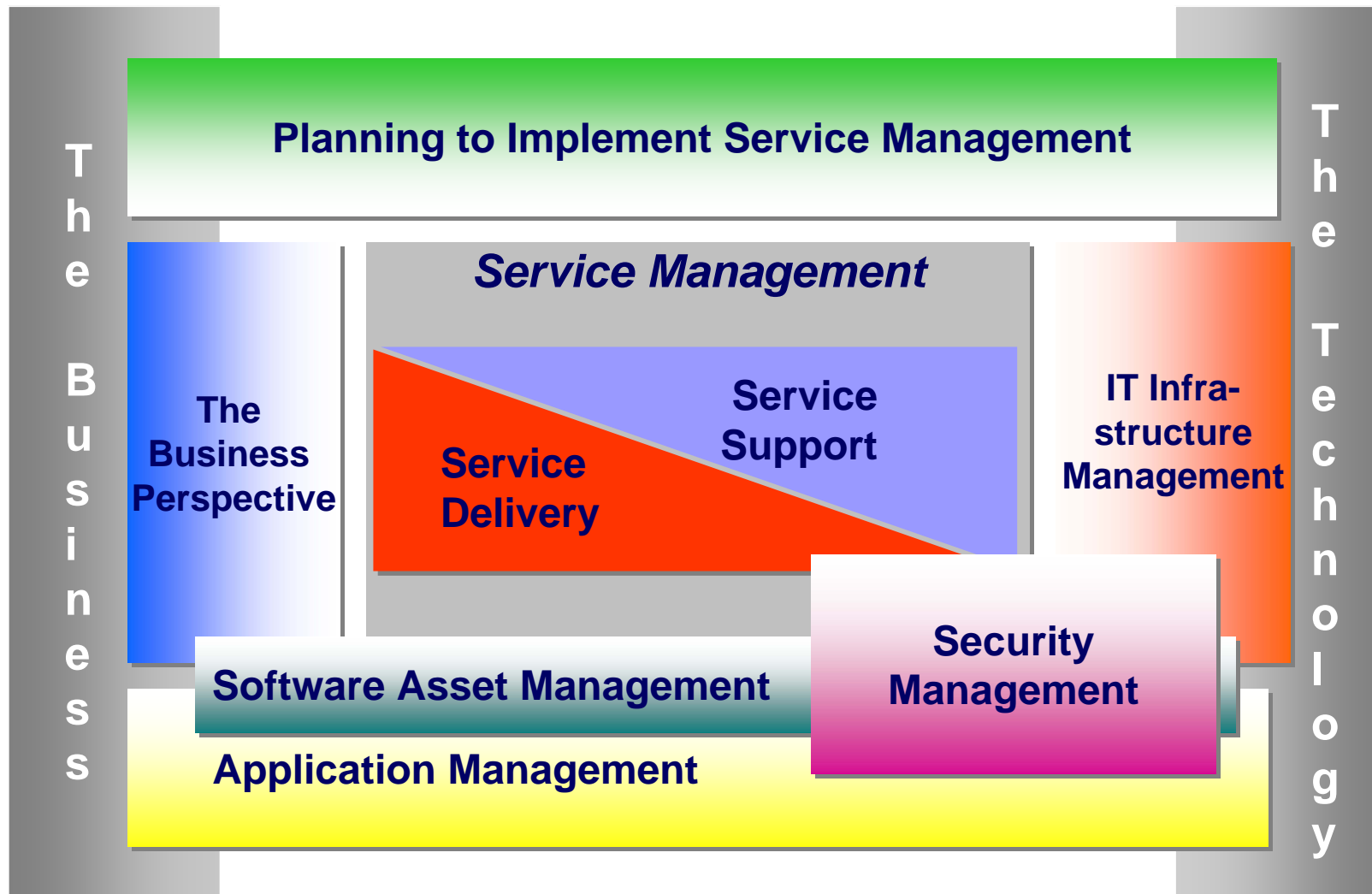
Research Challenges in Networking Middleware

- **Codify best practices for running IT systems**
 - IT Process Management, Service Flows
- **Service Level Agreements**
 - Express cost and profit functions
 - Multiple layers of SLAs
 - Automated reconciliation of external with internal goals (aka Policies)
- **Determining the location of a managed resource**
 - Often, servers in a data center are in 192.168 subnets
 - Legal implications on where (corporate) data is stored
- **(Security and Trust)**
 - Discrepancy between attributed importance and business opportunity

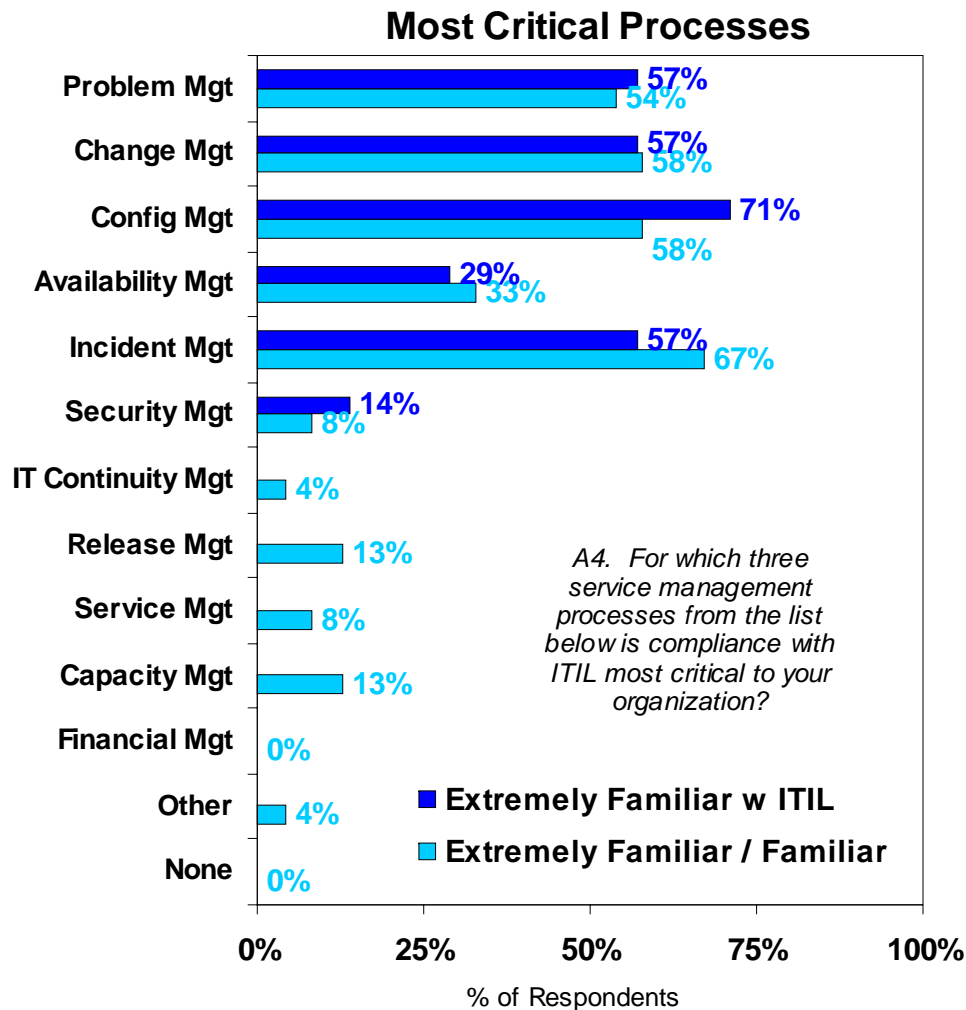
The need for best practices

- **Development organizations:**
 - management costs viewed as support, not development costs
 - Core business: adding new functionality to products
 - Not: Making products easier to administer
 - Fortunately, this begins to change
- **IT organizations:**
 - Estimating the impact of a change is extremely difficult
 - Results in:
 - Setup of dedicated staging environment
 - Manually try out what works best (or at all)
 - Create *IT run books* detailing the procedures
 - Huge costs (equipment, personnel)
 - Takes up to 90 days for complex changes
 - Still no guarantee that the procedure will really work in production
 - Every IT organization reinvents the wheel

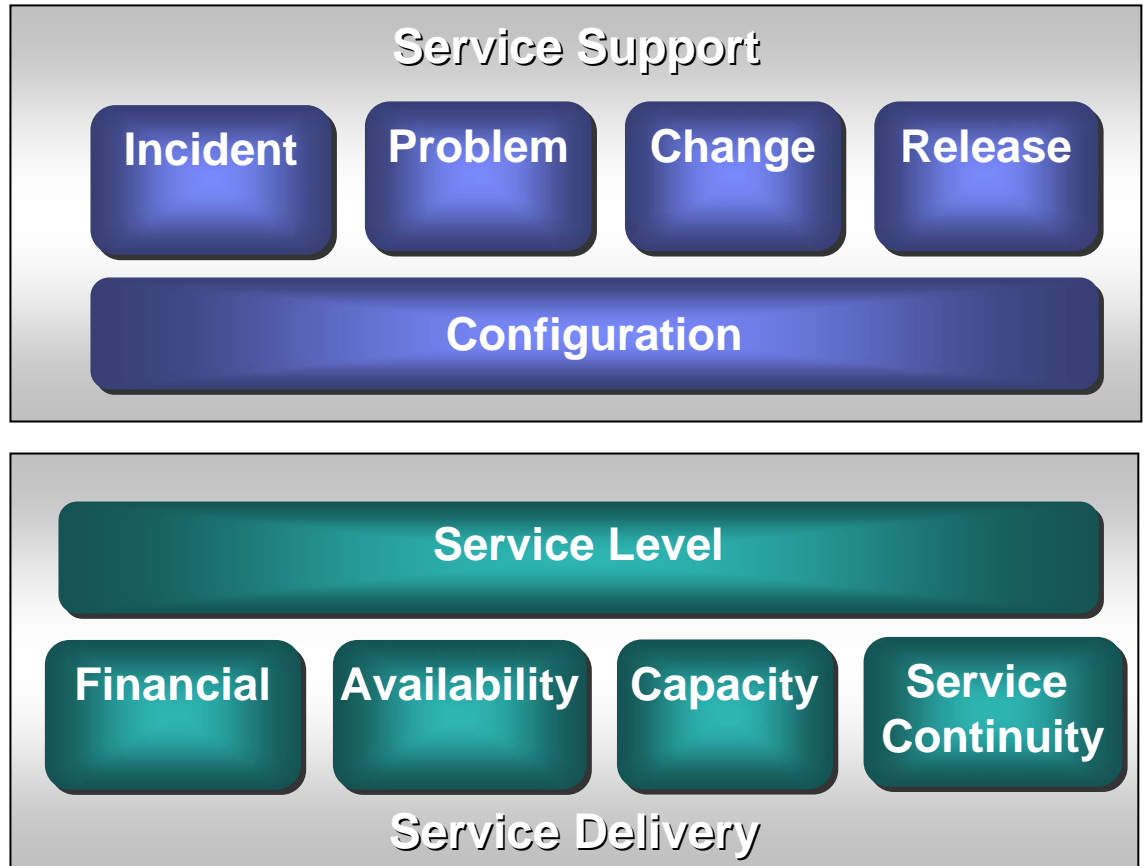
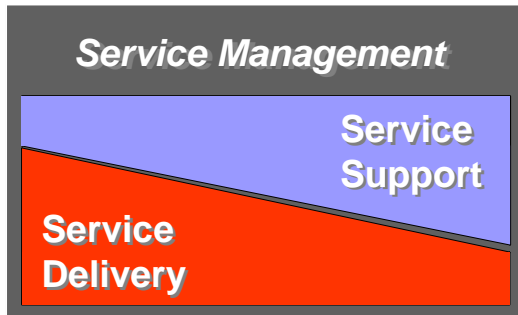
Codifying Best Practices: IT Infrastructure Library (ITIL)



Customer Rankings of ITIL Processes



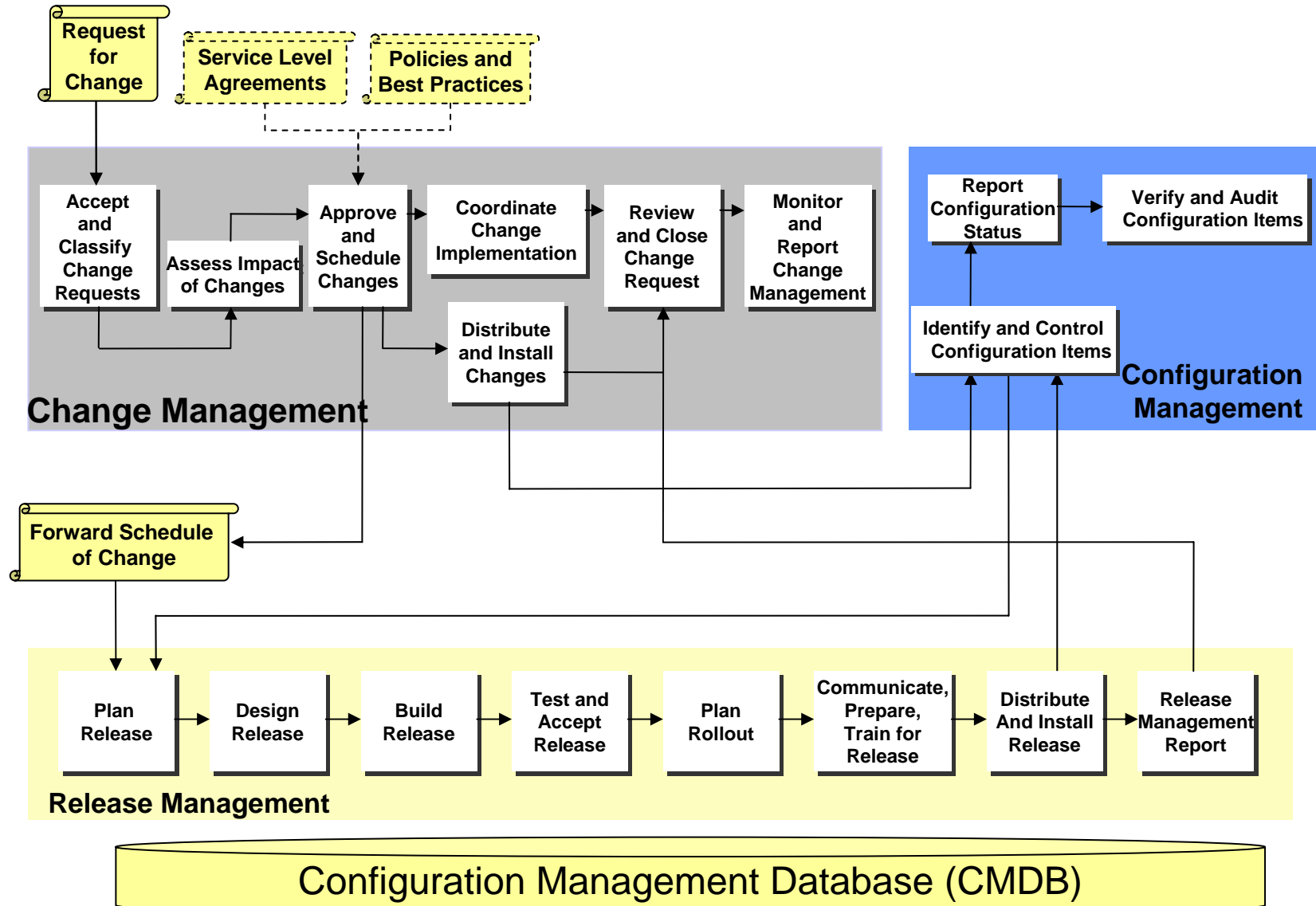
ITIL Service Management Processes



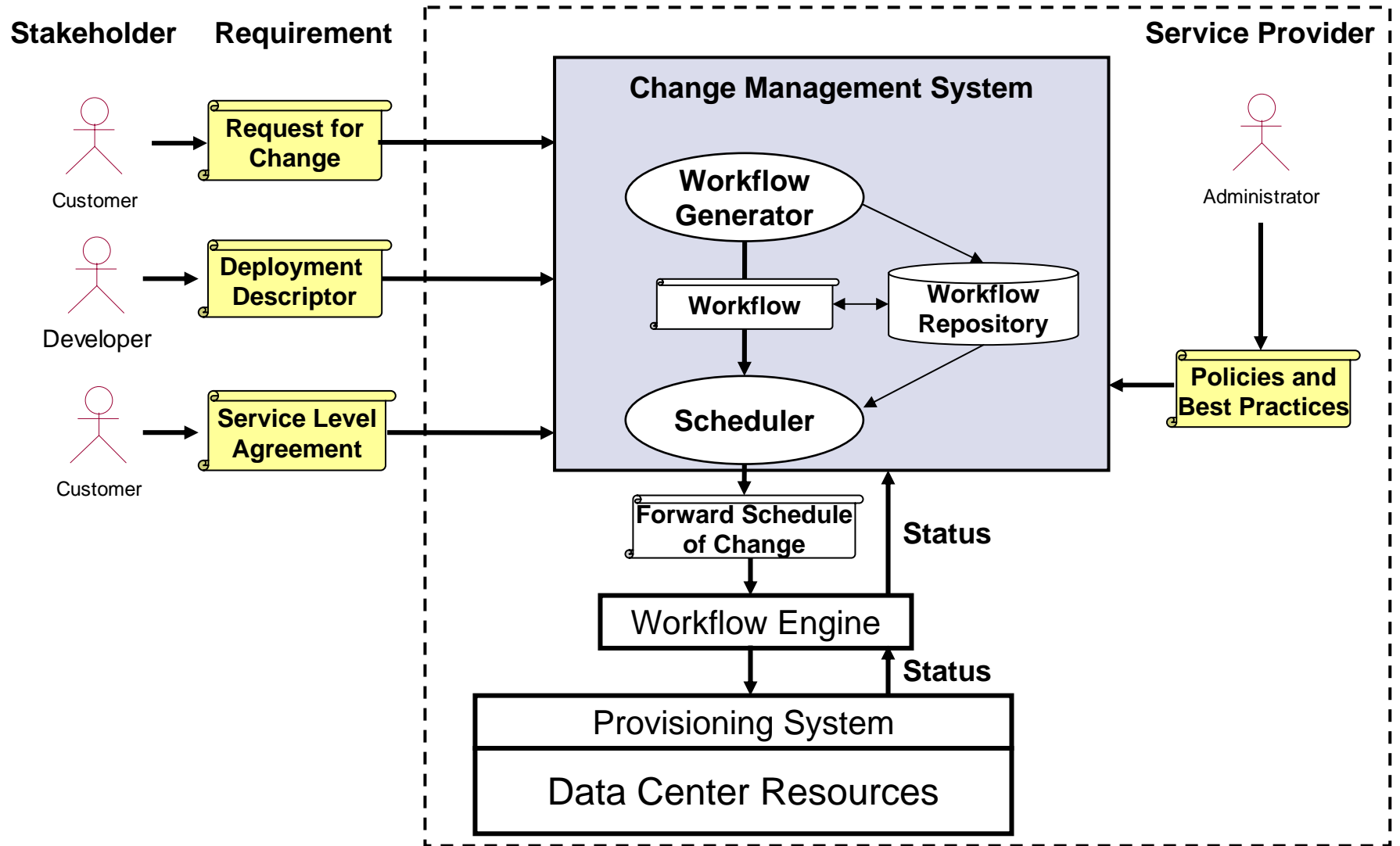
Problem: These Processes are described in textual form

Need to create models that can be interpreted by workflow systems

Relationships between ITIL Service Support Processes



Architecture and Inputs of a Change Management System



Service Level Agreements

- **What are the key performance indicators (KPIs)?**
 - Availability, response time, throughput, bandwidth...
- **How are they defined?**
 - Need to break down KPIs into measurable parameters
- **Who monitors compliance?**
 - Service Provider
 - 3rd party Measurement Services, e.g. Keynote
- **What are the maintenance intervals?**
 - Different time periods have different cost/profit functions
- **How is all this information put together?**
 - SLA templates are too inflexible
 - SLA languages, e.g. WS-Agreement
- **Is automated SLA negotiation a priority?**



Are SLAs hard to specify? (e.g., WS-Agreement)

wsag:ServiceProperties	wsag:ServiceName=ServerResourceInCustomer.Application	wsag:GuaranteeTerm (4)	
wsag:Name	wsag:ServiceScope	wsag:Service...	wsag:BusinessValueList
1 ResourcesAvailableFrom_900am_To_500pm_probed_every_Minute	wsag:ServiceScope	09:00:00-17:00:00	wsag:BusinessValueList <ul style="list-style-type: none"> wsag:Penalty <ul style="list-style-type: none"> wsag:AssessmentInterval <ul style="list-style-type: none"> wsag:TimeInter... PT1M wsag:ValueUnit USD wsag:ValueExpr... 5.00 wsag:AssessmentInterval <ul style="list-style-type: none"> wsag:TimeInter... PT10M wsag:ValueUnit USD wsag:ValueExpr... 50.00 wsag:AssessmentInterval <ul style="list-style-type: none"> wsag:TimeInter... PT30M wsag:ValueUnit USD wsag:ValueExpr... 500.00 wsag:AssessmentInterval <ul style="list-style-type: none"> wsag:TimeInter... PT1H wsag:ValueUnit USD wsag:ValueExpr... 5000.00 wsag:Reward <ul style="list-style-type: none"> wsag:AssessmentInterval <ul style="list-style-type: none"> wsag:TimeInter... PT1M wsag:ValueUnit USD wsag:ValueExpr... 10 wsag:AssessmentInterval <ul style="list-style-type: none"> wsag:TimeInter... P1D wsag:ValueUnit USD wsag:ValueExpr... 100
2 ResourcesAvailableFrom_5pm_to_2am_probed_every_Minute	wsag:ServiceScope	17:00:00-02:00:00	wsag:BusinessValueList
3 ResourcesInMaintenanceFrom_2am_to_3am_probed_every_Minute	wsag:ServiceScope	02:00:00-03:00:00	wsag:BusinessValueList
4 ResourcesAvailableFrom_3am_to_9am_probed_every_Minute	wsag:ServiceScope	03:00:00-09:00:00	wsag:BusinessValueList

Time intervals

Costs and profits are specified the same way

5 \$ per Minute

Additional penalties if not resolved within a certain time period

Additional bonus for continuous uptime

Crafting SLAs is an Art!

- **If Servers become overloaded, which Customer will be starved out?**
 - Classify customers according to revenue
 - SLA violation may not be due to technical failure, but result of business decision
- **Customer A signs an SLA for 99% availability over a year**
 - Service provider is able to keep the site up throughout the year
 - Busy Christmas shopping season stretches provider's resources
 - What will the service provider do?
 - Leave Customer A's Internet storefront untouched
 - Shut down Customer A's storefront for 3.5 days in a row?
 - Without paying any penalty as he is still within the allowable range of downtime
 - Ask customer for additional money to keep his site up and running
- **What will an automated Service Provisioning System do?**

Autonomic Networking Middleware: Takeaway Points

- **Major progress in core technologies for Networking Middleware**
 - Self-* technologies, context and location awareness, P2P systems
 - However, these core technologies seem isolated
- **Need to put the parts together, think in a bigger context**
 - Address the major pain points of *real-life* environments
 - Study the problems that yield the most “bang for the buck” first
- **Some focus areas:**
 - IT Processes as a way to codify best practices
 - Service Level Agreements are prerequisite for automation
- **Automation will only happen gradually over time**
 - Deal with human error, e.g., in system configuration
- **Some of this may sound like “soft topics”**
 - However, under the covers, there is plenty of “rigorous science”



IBM T.J. Watson Research Center

Q & A

