To BPEL or not to BPEL?

Wednesday, April 07, 2010
5:56 PM

To BPEL or not to BPEL?

- BPEL disadvantage: more time to product…
- BPEL advantage: better reusability, analysis…

A tale of two strategies…

Wednesday, April 07, 2010
5:57 PM

A tale of two strategies…

- The older strategy: salesforce.com:
  - Write business logic in a custom language.
  - Provide a large variety of services.
  - No BPEL.
- The newer strategy: intalioWorks.com:
  - Write overall strategy in BPMN;
  - Refine into BPEL;
  - Plumb BPEL to services graphically without writing code.

The big picture of BPEL

Wednesday, April 07, 2010
10:47 AM

The big picture

Define your business processes via BPMN,
and (**using design patterns** in WS-CDL), **convert them to executable** BPEL,
which you **execute on a BPEL engine**, e.g., Apache ODE or IBM Websphere,
which **contacts web services** (e.g., generated by apache AXIS or salesforce APEX),
which **utilize cloud services** via java APIs.

So, why are we this high up?

BPMN is the Cobol of the new millenium!

Easily readable by a manager, and

Clearly expresses **observable intent**, but

Darn near **impossible for a manager to write!**

BPEL is not understandable to managers, but need not be, because it only **implements the BPMN observables.**

It might be argued that BPEL **caused clouds as we know them,** by justifying their use.

The real effect of BPMN

Allows one to judge just how important reliability is,
by answering "what-if" questions about service performance,
and analyzing the cost of service delays,
which justifies high-reliability services,
which led to clouds!

Wednesday, April 07, 2010
11:01 AM

The hidden use for BPMN models

- Can use the model, without code, to predict business performance and value.
- Can model the effects of performance and reliability.
- Can predict bottom-line behaviors.

Wednesday, April 07, 2010
11:03 AM

How to analyze BPMN

BPMN analysis is a simple generalization of PERT (Program Evaluation and Review Technique), which is a generalization of Critical Path Method analysis (CPM).

Wednesday, December 02, 2009
11:08 AM

Critical Path Method

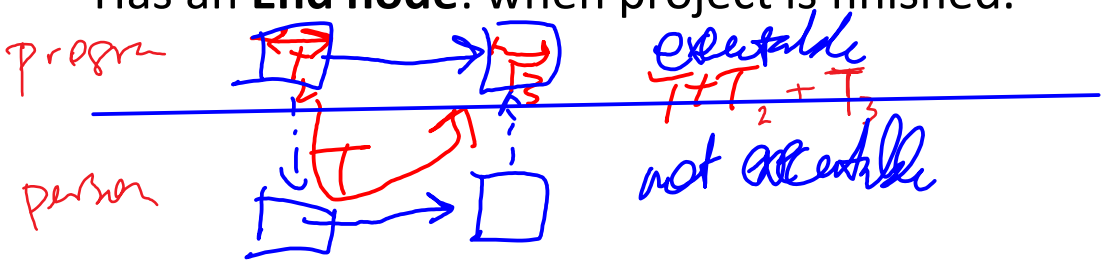
A basic tool for project management

Express task precedences as a **task graph**.

Task A -> B means A should precede B.

Has a **Start node**: when project starts.

Has an **End node**: when project is finished.



Relationship between CPM and BPMN

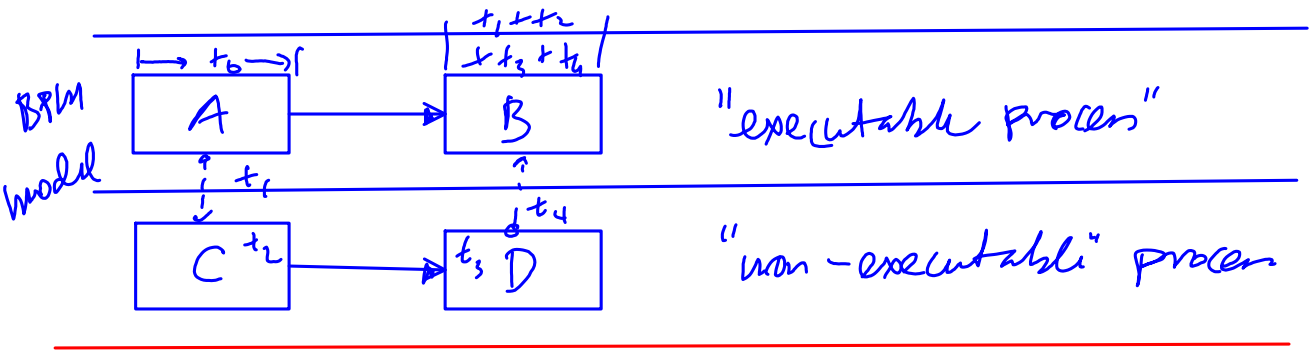
Wednesday, April 07, 2010
11:16 AM

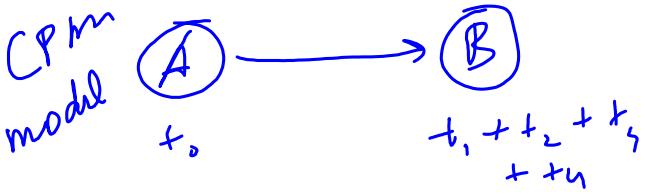
Relationship between CPM and BPMN:

- BPMN processes take time.
- Time delays are measurable and predictable.
- Effects of delays are predictable.
- Can use statistics on delays, as well as traces of business load, to predict overall performance, and even profit.
- Can determine which processes are critical, which determines which cloud functions are critical.
- Can ask "what-if" questions about what will happen if critical processes change. E.g., via a different SLA.

Relating CPM and BPMN

Wednesday, April 07, 2010
11:27 AM





Two BPMN processes (request and response)
One CPM task (get data).

Trace-driven simulation

A common business decision support technique

Input:

statistics on subprocess time requirements, and an actual trace of customer requests for a given day.

Output:

predictions of how that day would have gone if processes had been more (or less) efficient.

Key is cost/value analysis:

Cost: what you pay for service, or what it costs you to provide it yourself.

Value: what customers do, and how much money you make.

The big deal

Wednesday, April 07, 2010
6:24 PM

The big deal:

BPEL can be tuned based upon observation.
It is not static, but its performance changes when you reassign parts of it to different services.

CPM Definitions

Duration: how long it takes to do a task.

Earliest start time: earliest time a task can be started = max of earliest completion times for prerequisites.

Latest start time: latest time a task can be started and still finish with minimum completion time for project.

Earliest completion time: earliest start time + task duration.

Latest completion time: latest start time + task duration.

Slack time: latest completion time - earliest completion time = latest start time - earliest start time.

A task is critical if its earliest start time is its latest start time, i.e., slack time = 0!

Changing the time of a critical task changes overall behavior.

Key idea: if a service is critical, improving it means something.

The critical path theorem

Critical tasks do not occur in isolation, but instead lie on a critical path from start to finish.

There may be multiple critical paths.

If any task on a critical path changes in duration, the whole project duration can change.

The critical path algorithm

Objective: **compute slack time** for every task.

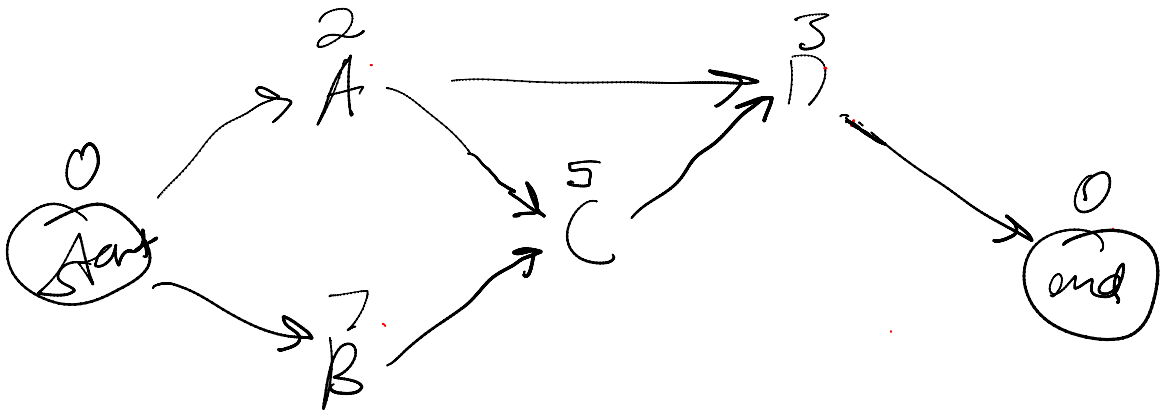
Two phases: **forward scan** and **backward scan**.

Forward scan: compute **earliest** completion time.

Backward scan: compute **latest** completion time.

A task graph

Wednesday, December 02, 2009
11:16 AM



A task graph has a start, end, and intermediate tasks.

Forward algorithm

Wednesday, December 02, 2009
11:24 AM

Forward scan:

Label each task with its time.

Mark start's earliest start time and earliest completion time as 0.

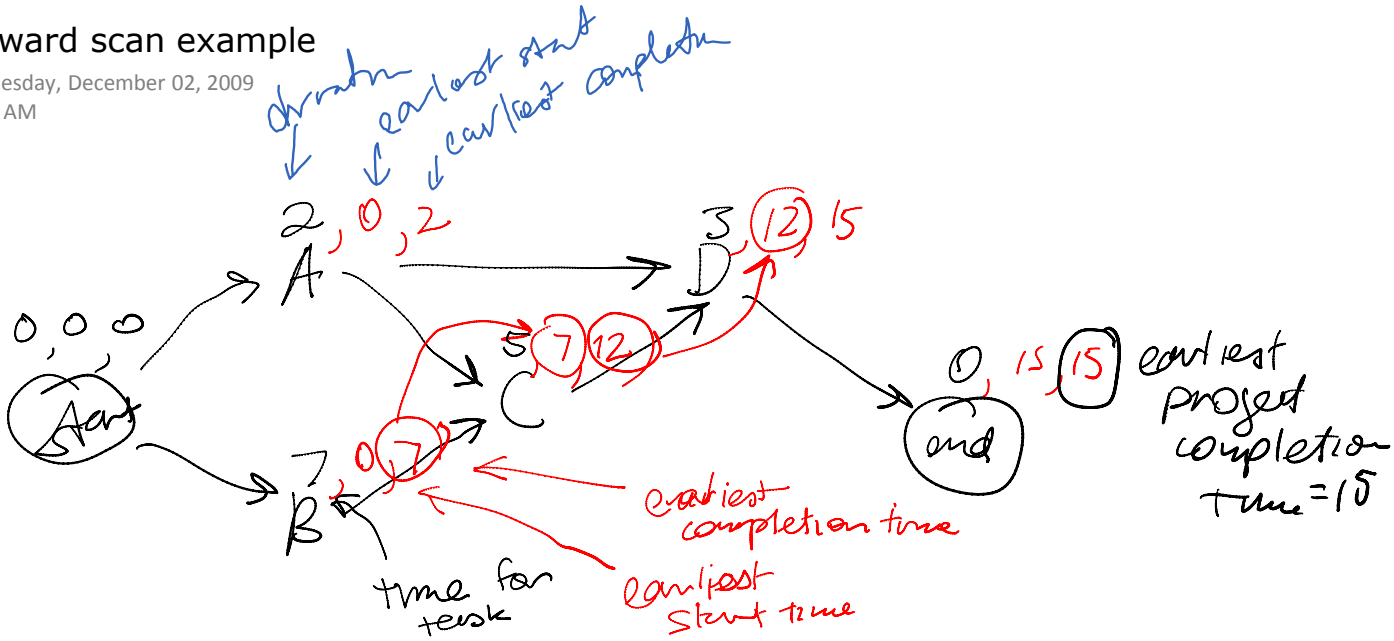
From left to right (start to finish),

- If a task's predecessors are marked with their earliest completion times,
  - its earliest start time is the **maximum** of **predecessor earliest completion times** and
  - its earliest completion time is its earliest start time plus its duration.

At the end, every node is labeled with its earliest start and completion times.

Forward scan example

Wednesday, December 02, 2009
11:16 AM



Start at beginning

Compute earliest start time and earliest completion time.

Backward algorithm

Wednesday, December 02, 2009
11:24 AM

Backward scan:

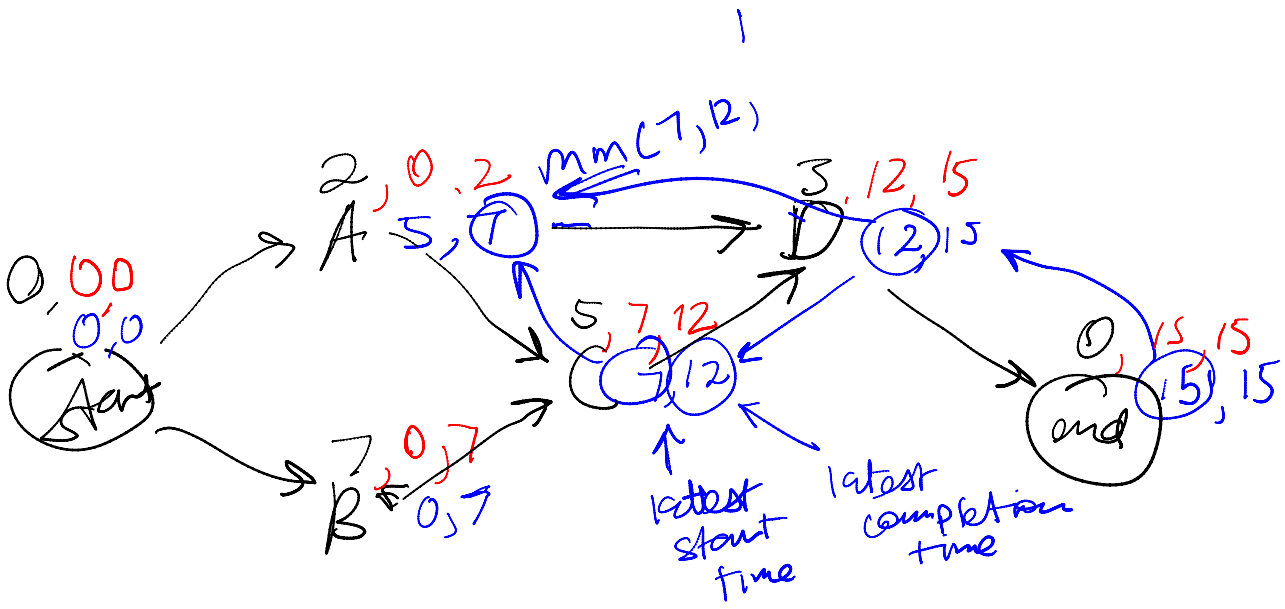
Mark end's latest start time and latest completion time as its earliest start time and completion time.
From right to left (finish to start),

- If a task's **successors** are marked with latest start time,
  - its **latest completion time** is the **minimum** of **successor latest start times** and
  - its latest start time is its latest completion time minus its duration.

At the end, every node is labeled with its latest start and completion times.

Backward scan example

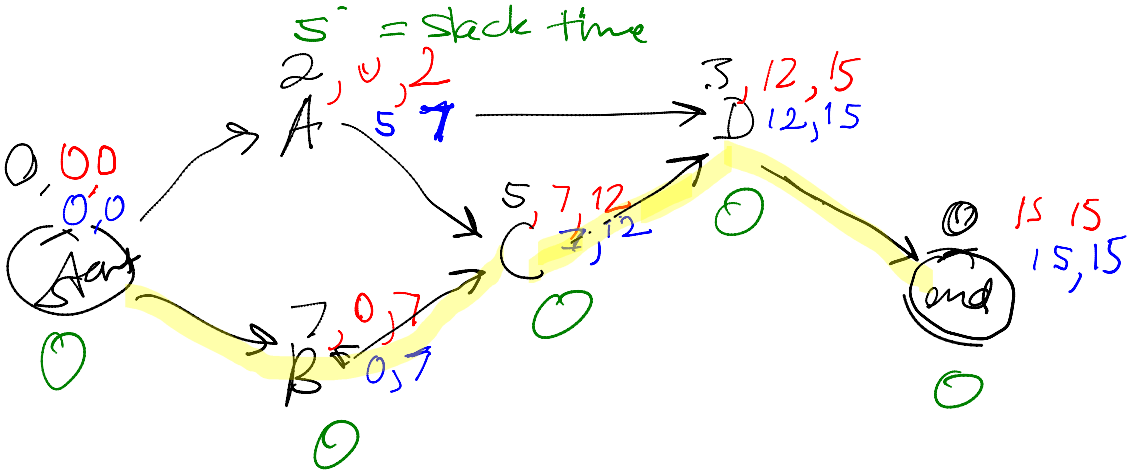
Wednesday, December 02, 2009
11:16 AM



Start at end
Compute latest completion and start times from end to beginning

The critical path

Wednesday, December 02, 2009
11:16 AM



A critical path has earliest start time = latest start time.

CPM and PERT

- CPM: Critical Path Method
  - input: task times and precedences
  - output: minimum completion time, slack time, etc.
- PERT: Program Evaluation and Review Technique
  - input: task time **probability distributions** and precedences
  - output: probability distribution of minimum completion time, slack time, etc.

PERT

Wednesday, December 02, 2009
12:14 PM

How (simple) PERT works

Assign a normally-distributed probability distribution $n(\mu,\sigma)$ to each task:

μ=mean completion time

σ=standard deviation of completion time

$n(\mu,\sigma)(t) = \exp(-(t-\mu)^2/\sigma^2)$

Compute the PDF of the earliest start time and latest start time from the PDF of the task completion times!

If we have two tasks in sequence, and they're independent, then the probability that the two take time t is the integral of the product of PDF(t1)*PDF(t2), for all t=t1+t2 (convolution).

Example of PERT analysis

Wednesday, April 07, 2010
6:36 PM

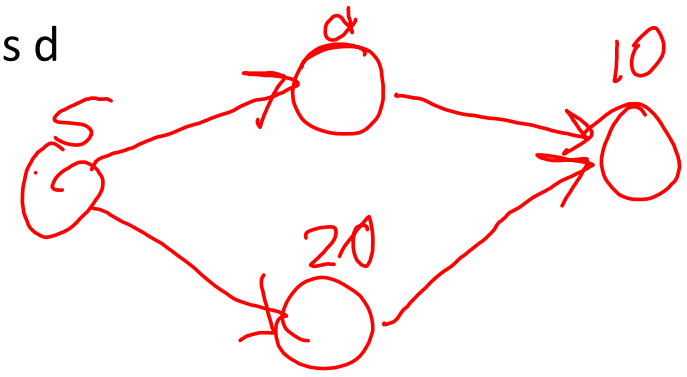
Suppose we have an empirical distribution as follows:

0-10	50 instances p=50/100
10-20	40 instances p=40/100
20-30	10 instances p=10/100

Caveat:

distributions change over time.
sensitive to time of day and external load.
sensitive to external events.

Call this d



If 0<=d<10, d is not critical, delay is 35 with p=1/2
If 10<=d<20, d is not critical, delay is 35, p=4/10
if 20<=d<30, d is critical, delay is 35-45, p=1/10
SO
We can compute the expected value, which is the sum of value * Prob(value), where the value is an interval
= (35, 35)*9/10 + (35,45)*1/10

Two probability distributions

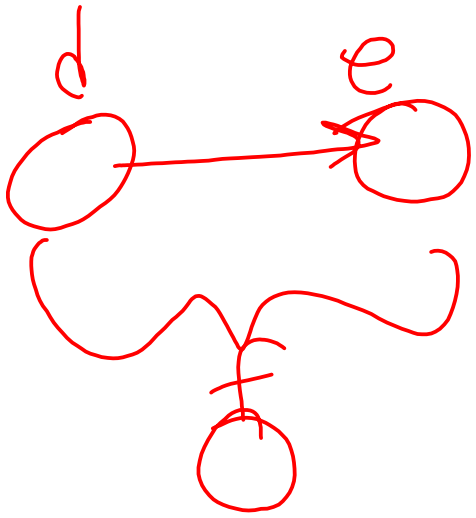
Wednesday, April 07, 2010
6:44 PM

Suppose that d is

10	p=.2
20	p=.5
30	p=.3

and e is

10	p=.1
20	p=.3
30	p=.6

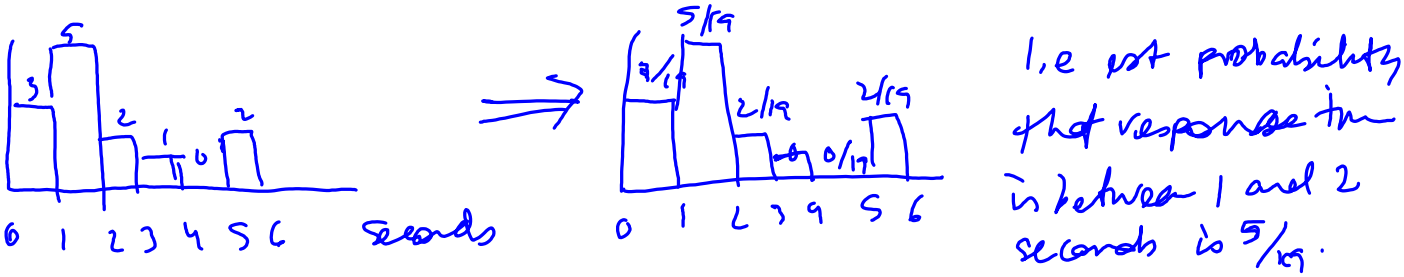


10+10 at probability .2*.1
10+20 at probability .2*.3
10+30 at probability .2*.6
….

Wednesday, April 07, 2010
11:43 AM

Practical modeling

- Completion times expressed in terms of frequency tables for observations:
  - Completion time range on X
  - Frequency on Y
  - Empirical probability: frequency/total trials.



Limitations of BPMN modeling

Wednesday, April 07, 2010
11:48 AM

Limitations of modeling

- Response times for external services vary with (unpredictable) external load (and time of day).
- **Cannot model unforeseen situations.**

Amusement: IBM has tried to beat this, by doing modeling that injects unforeseen situations!

- **E.g., gnomes running around causing hardware failures!**

Alas,…

Wednesday, April 07, 2010
10:59 AM

Alas,…

I haven't been able to locate a freeware BPMN analyzer,
that:

Assigns probabilities to failures,
Simulates resulting operational behavior, and predicts effects upon profits.

But these do exist; they just cost money!